Agricultural and Forest Meteorology 301-302 (2021) 108340

Contents lists available at ScienceDirect

Agricultural and Forest Meteorology

journal homepage: www.elsevier.com/locate/agrformet

Research paper

Understanding the non-stationary relationships between corn yields and meteorology via a spatiotemporally varying coefficient model

Hao Jiang [a], Hao Hu [b], Bo Li [c], Zhe Zhang [d], Shaowen Wang [b,*], Tao Lin [a,*]

[a] College of Biosystems Engineering and Food Science, Zhejiang University, Hangzhou, Zhejiang, China
[b] Department of Geography and Geographic Information Science, University of Illinois at Urbana-Champaign, Urbana, Illinois, United States
[c] Department of Statistics, University of Illinois at Urbana-Champaign, Urbana, Illinois, United States
[d] Department of Geography, Texas A&M University, College Station, Texas, United States

ARTICLE INFO

Keywords:
Bayesian hierarchical model
Crop yields
Meteorology
Spatiotemporal analysis
Spatiotemporally varying coefficient

ABSTRACT

The relationships between crop yields and meteorology are naturally non-stationary because of spatiotemporal heterogeneity. Many studies have examined spatial heterogeneity in the regression model, but only limited research has attempted to account for both spatial autocorrelation and temporal variation. In this article, we develop a novel spatiotemporally varying coefficient (STVC) model to understand non-stationary relationships between crop yields and meteorological variables. We compare the proposed model with variant models specialized for time or spatial, namely spatial varying coefficient (SVC) model and temporal varying coefficient (TVC) model. This study was conducted using the county-level corn yield and meteorological data, including seasonal Growing Degree Days (GDD), Killing Degree Days (KDD), Vapor Pressure Deficit (VPD), and precipitation (PCPN), from 1981 to 2018 in three Corn Belt states, including Illinois, Indiana, and Iowa. Allowing model coefficients varying in both temporal and spatial dimensions gives the best performance of STVC in simulating the corn yield responses toward various meteorological conditions. The STVC reduced the root-mean-square error to 10.64 Bu/Ac (0.72 Mg/ha) from 15.68 Bu/Ac (1.06 Mg/ha) for TVC and 16.48 Bu/Ac (1.11 Mg/ha) for SVC. Meanwhile, the STVC resulted in a higher R^2 of 0.81 compared to 0.56 for SVC and 0.64 for TVC. The STVC showed better performance in handling spatial dependence of corn production, which tends to cluster estimation residuals when counties are close, with the lowest Moran's I of 0.10. Considering the spatiotemporal non-stationarity, the proposed model significantly improves the power of the meteorological data in explaining the variations of corn yields.

1. Introduction

Understanding the effects of meteorology variability on crop yields is central to yield risk measurement, farm management, and even food security (Lobell and Burke, 2010; Olesen et al., 2011; Peng et al., 2020). Two main modeling approaches have been extensively studied to deepen understanding: the process-based models and statistical models, where process-based models are also referred to as crop simulation models. In addition to meteorology variables, process-based models, e.g. DSSAT (Jones et al., 2003), often require various input data such as cultivar, soil conditions, and farm management, which are difficult to obtain for large scale studies. The other approach is statistical models, which have been widely used to quantify empirical relationships between crop yields and meteorology from historical records (Bornn and Zidek, 2012; Lobell and Burke, 2010; McGrath et al., 2015; Ray et al., 2015; Schlenker and Roberts, 2009).

In statistical modeling, regression methods are often used to quantify the relationships between the interested outcome variable and a set of covariates. Two major issues are often discussed in existing regression methods for crop yield estimation. First, spatially correlated error terms violate the assumption of independent and normally distributed residuals in linear regression models. This spatial dependence of residuals indicates that the model is inadequate to explain the data, thus resulting in poor model fitting and less accurate predictions (Bornn and Zidek, 2012; Hoeting, 2009; Jiang et al., 2009). Second, the impact of meteorology variability on crop yields might follow a spatiotemporally non-stationary process, i.e., regression coefficients do not necessarily remain fixed from location to location or time to time, primarily when

* Corresponding authors.
E-mail addresses: shaowen@illinois.edu (S. Wang), lintao1@zju.edu.cn (T. Lin).

https://doi.org/10.1016/j.agrformet.2021.108340
Received 11 July 2020; Received in revised form 9 December 2020; Accepted 17 January 2021
Available online 4 February 2021
0168-1923/© 2021 Elsevier B.V. All rights reserved.

the study area covers a variety of spatially heterogeneous landscapes and temporally varied agronomic practices (Choi et al., 2012; Sharma et al., 2011). Therefore, there is a need for space-time statistical models to take both spatial autocorrelation and spatiotemporally non-stationarity of the crop production system into account (An et al., 2015).

The spatial autocorrelation and heterogeneity of meteorological and yield data are the direct cause of spatially correlated errors. The tendency for areas or sites that are geographically close together to have similar values of spatial attributes leads to spatial autocorrelation. Characteristics vary over a continuous surface, such as soil property, could respond to precipitation similarly in soil moisture and cause similar effects on crop yield. The disparity in geographic conditions and management factors heavily affects the spatial heterogeneity of crop systems and the level of climate stress tolerance (Butler and Huybers, 2015; Lobell and Azzari, 2017). For example, the impact of climate extremes on crop yield is much more substantial in rain-fed regions than irrigated regions (Troy et al., 2015). Besides, per unit of increase of meteorological factors, such as precipitation, influences the corn yield differently in different conditions. Li et al. have identified that the negative sensitivity towards excessive precipitation and the positive sensitivity of precipitation under drought is of a similar magnitude (Li et al., 2019). Therefore, the stationary coefficients (i.e., intercept and slopes) of the classical multiple linear regression model present a challenge in describing the relationships between crop yields and meteorological factors on large spatial scales.

Panel analysis and spatial varying coefficients (SVC) model are commonly used approaches for modeling non-stationary spatial processes (Banerjee et al., 2014; Gelfand et al., 2003; Mahalingam and Orman, 2018; Shand et al., 2018). Spatially non-stationary processes can be explained by the fact that relationships between the outcomes and covariates are intrinsically different across space (Fortheringham et al. 1998). With varied managing practices and environmental conditions, responses of corn production toward meteorological variability naturally conform to non-stationary spatial processes. Both panel analysis and SVC are adapted to varying spatial associations, but their implementations are much different. In panel analysis, the linear regressions are conducted by data collected over time and over the same individuals. The assumptions of different distributions of error terms across the spatial dimension brought unique attributes of individuals. SVC has been applied to areas like reproducing house prices (Gelfand et al., 2003) and violent rates (Waller et al., 2007). In SVC, the overall mean of coefficients is estimated first. Then local deviations from the mean are estimated by applying a spatial random effect such as conditional autoregressive (CAR) models (Waller et al., 2007). Since prior information such as the model for spatial random effects can be included, SVC models fit nicely into the Bayesian hierarchical spatial modeling framework. In terms of handling the spatial autocorrelation or dependency between neighbors inherent in data, panel analysis assumes errors to be independent over spatial sectors. In contrast, SVC directly models the autocorrelation by decomposing the residuals into structured random effects and white noise (Waller et al., 2007).

Temporal non-stationary processes, including the improved agronomy practices (e. g. stress-tolerant cultivars) and extreme climate events, are still not reliably reproduced in Panel and SVC models. For both Panel regressions and SVC models, the local relationships between outcomes and covariates are time-invariant (Choi et al., 2012; Gelfand et al., 2003). However, variations of crop systems and external factors over time can substantially influence crop yield sensitivity towards meteorology variability. For instance, the "Great Midwest Flood of 1993″ was one of the greatest and damaging disasters ever to occur in the Midwest and severely affected the corn yield (Junker et al., 1999). Record-breaking precipitation spanned from June to August, which led to widespread crop failures. The improvements in modern agronomy practice and genetic engineering tend to increase the vulnerability of corn production to drought (Lobell et al., 2014). In consequence, the quantitative performance of statistical models is subject to which period they are applied, and the models could underestimate the effect of environmental stress. Flexible non-stationary temporal processes should be considered to reduce the uncertainty of modeling.

To make an analytical tool for yield response, an essential improvement is to develop models that accommodate both spatial and temporal non-stationary processes. This information is also critical for crop models that provide an essential benchmark for calibrating spatiotemporal varied crop yield responses (Folberth et al., 2019). SVC has presented a great ability to capture correlated spatial datasets in prior work (Gelfand et al., 2003; Waller et al., 2007). However, considering the heterogeneity of the crop production system and significant interannual meteorological variability, the coefficients of meteorological variables are expected to vary in space and time. As a further step toward spatiotemporal modeling, we aim to develop a novel spatiotemporal model based on SVC to account for spatiotemporally non-stationary responses of crop yields to meteorological variables. Therefore, a Spatiotemporally Varying Coefficient (STVC) model, is proposed to provide a broader and flexible inferential basis for spatiotemporal yield variability analysis.

In this study, we examine the spatiotemporal varying yield-meteorology relationships with the STVC using county-level data in the Midwestern U.S. from 1981 to 2018. To demonstrate the advantage of this model, we further compare it with ordinary least squares (OLS), spatial panel regression analysis on model fitting performance. The rest of the paper is organized as follows. First, we introduce methodological foundations for the STVC model in detail. Then, we briefly describe the basic OLS model and a panel regression model for spatiotemporal problems. Next, we utilize the STVC model to study the variability of corn yields in response to meteorological covariates, and the results of the model comparison are discussed accordingly. Finally, we briefly summarize our research findings and future work. The purposes of this article are to (1) implement the STVC model to analyze spatiotemporal non-stationary processes; (2) reproduce spatiotemporally varying meteorological impacts on corn yields using the STVC model; (3) compare the model performance with competitive methods.

2. Materials and methodology

2.1. Spatiotemporally varying coefficients (STVC) model

Gelfand et al. (2003) developed a spatially varying coefficient model as follows:

$$y_{(s)} = \mathbf{x}_{(\mathbf{s})}^{\mathbf{T}}\boldsymbol{\beta} + \mathbf{x}_{(\mathbf{s})}^{\mathbf{T}}\boldsymbol{\beta}_{(\mathbf{s})} + \varepsilon_{(s)} \tag{1}$$

where $\boldsymbol{\beta}_{(s)}$ is a second-order stationary mean zero Gaussian process independent of the white noise error process. The formulation in Eq. (1) allows decomposing the total effects of covariates into an overall mean effect at the global level, denoted by $\boldsymbol{\beta}$, and a local deviation at location s to the overall mean effect, denoted by $\boldsymbol{\beta}_{(s)}$. The intrinsic conditional autoregressive (ICAR) model is a popular choice for modeling the random spatial process $\boldsymbol{\beta}_{(s)}$ (Dong et al., 2016; Ozaki et al., 2008; Waller et al., 2007). Specifically, the model for $\boldsymbol{\beta}_{(s)}$ governed by ICAR is given as:

$$\beta_{i(s)}|\beta_{i(s^*)} \sim N\left(\frac{1}{m_s}\sum\beta_{i(s^*)}, \frac{\sigma^2_{\eta_s}}{m_s}\right), i = 1, 2, \ldots, p \tag{2}$$

where $\beta_{i(s)}$ is the i^{th} element in β_s, S^* is the set of neighbors of region s, m_s is the cardinality of the set S^* or the number of neighbors, and $\sigma^2_{\eta_s}$ controls the magnitude of spatial variation.

We further extend the SVC model in Gelfand et al. (2003) to spatiotemporal data. The innovation of this extension is to allow the spatially varying coefficient to change over time. First, we construct the Temporally Varying Coefficients (TVC) Model by random walk process

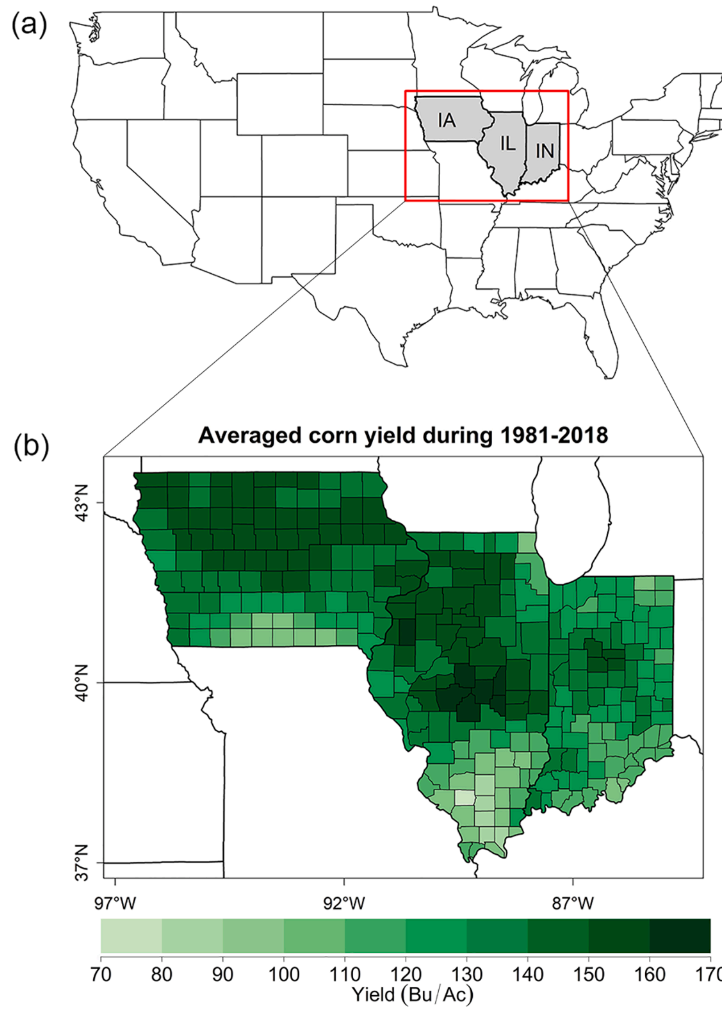

Fig. 1. Map of the study area. (a) '3I' states include Illinois (IL), Indiana (IN), and Iowa (IA) in the U.S. Midwest; (b) averaged corn yield at the county-level.

as:

$$y_t = \mathbf{x}_{(t)}^T\boldsymbol{\beta} + \mathbf{x}_{(t)}^T\boldsymbol{\beta}_{(t)} + \varepsilon_{(t)} \tag{3}$$

where $\boldsymbol{\beta}_{(t)}$ represents the temporal random effects that can be modeled by a random walk or an autoregressive process. For example, elements in term $\boldsymbol{\beta}_{(t)}$ can be defined as the random walk process:

$$\beta_{(t)} = \beta_{(t-1)} + \varepsilon_{(t)},\ i\ =\ 1,\ 2,\ 3,\ \ldots.,p \tag{4}$$

where $\varepsilon_{(t)}$ is white noise, and $\beta_{(0)}\ = 0$. The explanation behind a random walk process is that the current value of the random variable is determined by the past value plus an independent error term.

Accordingly, our model is constructed based on SVC and TVC as:

$$y_{(s,t)} = \mathbf{x}_{(s,t)}^T\boldsymbol{\beta} + \mathbf{x}_{(s,t)}^T\boldsymbol{\beta}_{(s,t)} + \varepsilon_{(s,t)} \tag{5}$$

where each element in $\boldsymbol{\beta}_{(s,t)}$ is further decomposed as:

$$\boldsymbol{\beta}_{(s,t)} = \boldsymbol{\beta}_{(s)} + \boldsymbol{\beta}_{(t)} \tag{6}$$

where $\boldsymbol{\beta}_{(s)}$ represents the spatial random effects that can take the form defined in Eq. (2), and $\boldsymbol{\beta}_{(t)}$ represents the temporal random effects defined in Eq. (4).

By combining Eq. (1)-(6), the relationship between corn yields and meteorological variables, including growing degree days (GDD), killing degree days (KDD), vapor pressure deficit (VPD), and precipitation (PCPN) are represented as follows:

$$\begin{aligned} Yield_{s,t} = {} & \beta^{Itercept} + \beta^{GDD}GDD_{s,t} + \beta^{KDD}KDD_{s,t} + \beta^{VPD}VPD_{s,t} + \beta^{PCPN}PCPN_{s,t} + \beta_s^{Itercept} \\ & + \beta_{s,t}^{GDD}GDD_{s,t} + \beta_{s,t}^{KDD}KDD_{s,t} + \beta_{s,t}^{VPD}VPD_{s,t} + \beta_{s,t}^{PCPN}PCPN_{s,t} + \varepsilon_{s,t} \end{aligned} \tag{7}$$

where s is the index for the county, t is the index for year, $\beta^{Intercept}$, β^{GDD}, β^{KDD},β^{VPD} and β^{PCPN} control the overall mean process of coefficients at the global level, among which $\beta_s^{Intercept}$ represents the county-specific intercept to account for time-invariant spatial heterogeneity, and $\varepsilon_{s,t}$ is the white noise. To consider temporal variation in the spatially non-stationary process, $\beta_{s,t}^{GDD}$, $\beta_{s,t}^{KDD}$, $\beta_{s,t}^{VPD}$, and $\beta_{s,t}^{PCPN}$are defined as spatio-temporally varying coefficients for GDD, KDD, VPD, and precipitation. Those county-specific coefficients can be further decomposed as:

$$\beta_{s,t}^{factor} = \beta_s^{factor} + \alpha_t^{factor},\ factor = \{GDD, KDD, VPD, PCPN\} \tag{8}$$

where β_s^{GDD}, β_s^{KDD}, β_s^{VPD}, and β_s^{PCPN} represent the time-invariant spatial random effects and α_t^{GDD}, α_t^{KDD}, α_t^{VPD}, and α_t^{PCPN} represent the location-invariant temporal random effects. All county-specific terms, $\beta_s^{Intercept}$, β_s^{GDD}, β_s^{KDD}, β_s^{VPD}, and β_s^{PCPN}, are considered to follow the ICAR prior defined in Eq. (2). The temporal changes are modeled as random walk processes as follows:

$$\alpha_t^{factor} = \alpha_{t-1}^{factor} + \varepsilon_t^{factor},\ factor = \{GDD, KDD, VPD, PCPN\} \tag{9}$$

where ε_t^{factor} is the white noise term. By choosing the random walk process, we assume that the local coefficient within a specific year is composed of its coefficient from last year plus a random error.

2.1.1. Parameter estimation

Following the estimation of SVC, we use a Bayesian method to estimate our STVC model. The likelihood of the crop yields in Eq. (7) is expressed as:

$$\begin{aligned} f(Yield|\Theta) = \prod_s \prod_t N(&Yield_{s,t}|factor_{s,t}, \beta^{Intercept}, \beta^{factor}, \beta_s^{Intercept}, \\ &\beta_s^{factor}, \alpha_t^{factor}, \sigma_{s,t}^2),\ factor = \{GDD, KDD, VPD, PCPN\} \end{aligned} \tag{10}$$

where Θ is a set of all the hyper-parameters included in the model. The prior specification for our model is as follows. $\beta^{Intercept}$, β^{GDD}, β^{KDD},β^{VPD}, and β^{PCPN} have ICAR prior. α_t^{GDD}, α_t^{KDD}, α_t^{VPD}, and α_t^{PCPN} follow the random walk process. A vague normal prior, $N(0,\ 10^6)$, is used for $\beta^{Intercept}$, β^{GDD}, β^{KDD}, β^{VPD}, and β^{PCPN}. A vague inverse gamma, $IG(1{,}0.01)$, is set as the prior distribution for all white noise parameters. The coefficients of meteorological factors can be interpreted as the changes of corn yield (Bu/Ac) induced by per unit of accumulated GDD or KDD change (°C), or per unit of accumulated precipitation change (mm), or per unit of the seasonal mean of VPD (Pa) over the growing season for a specific county. The inferred coefficient for each meteorological factor is a multivariate conditional probabilistic distribution based on the meteorological conditions in the specific year and county. Therefore, the estimated coefficient for one specific factor could be influenced by others. In this study, although we calculate the mean values of posterior distributions as estimated coefficients, the yield response is considered the result of combined meteorological impacts.

The Markov chain Monte Carlo (MCMC) algorithms embedded in WinBUGS software are used to perform parameter estimation and draw a statistical inference. Bayesian formulations of ICAR were discussed with details by (Besag et al., 1991) and (Besag and Kooperberg, 1995). A nice feature of the BHM is that we can easily estimate the uncertainty along with the point estimates. Given the complexity of posterior distributions in hierarchical models, both Gibbs sampler and Metropolis Hasting algorithms are used in MCMC to generate posterior samples (Gilks, 1995). MCMC sampling methods provide a general approach to fitting complex hierarchical models in a Bayesian framework (Cressie and Wikle 2015). MCMC generates samples of model parameters iteratively from a set of Markov chains, which usually take a number of steps to converge. The samples generated after the convergence are then considered to form the posterior distribution of all the unknown parameters. In our study, two chains are initialized, and posterior distributions for each model parameter were estimated using 100,000 iterations with the first 97,000 iterations discarded as the burn-in period. Regarding the significance test for parameters in STVC, we use the 2.5% and 97.5% quantiles of the posterior distribution to evaluate if zero is covered by this interval. If yes, the parameters are considered not significantly different from zero.

2.2. Yield data and meteorological variables

The study area is located in the Midwestern U.S., including the states of Illinois, Indiana, and Iowa (Fig. 1). The corn production is mainly rain-fed in these states. County-average corn yield data for the period of 1981 to 2018 are collected from the U.S. Department of Agriculture's National Agricultural Statistics Service (USDA, 2019). The 38-year average corn yields range from 99.41 to 165.63 Bu/Ac (6.69 Mg/ha to 11.15 Mg/ha) for each county (*SI Appendix*, Figure S1). When aggregated at the state level, all three states show increasing trends of corn yield over the past 38 years except for a significant hit by severe drought in the year 2012 (*SI Appendix*, Figure S2). Here we detrend corn yields before running the regression analysis, as suggested by Quiring and Papakryiakou (2003). The temporal linear trend induced by the combined effects of changes in governmental policies, technological improvements, and climate change was therefore removed. Thus, we can focus on the variation of corn yields explained by meteorology variabilities. From the available data, 239 yield records are missing for multiple reasons, such as only aggregated records available for several counties or skipped surveys in multiple years for some counties. To address this issue, we replace the missing data with the mean yield value of adjacent counties, which results in 11,134 county-year records in total.

Meteorology data are derived from the PRISM Climate Group (PRISM Climate Group, 2004). We acquire this data from the Applied Climate Information System API (2017). In particular, maximum daily temperature and minimum daily temperature, daily precipitation information, and daily VPD are used. Among these variables, precipitation and VPD are used to measure the availability of water supply and severity of drought stress for crop growth, respectively. Instead of using both maximum and minimum temperature information as covariates, we calculate the indicators of GDD and KDD, which often serve as simple single-dimensional measurements for describing crops' exposure to heat. GDD serves as necessary thermal resources for crop growth, while KDD directly indicates the heat stress. Therefore, four meteorological factors, including PCPN, VPD, GDD, and KDD, are selected as STVC covariates to cover various water and heat conditions. GDD is calculated by taking an average of the daily minimum and maximum and subtracting a base temperature value. KDD is calculated as the measure of exposure to extreme heat. KDD is measured by daily maximum temperature above a threshold T_{UT}:

$$GDD = \frac{T_{MAX} + T_{MIN}}{2} - T_{BASE} \tag{11}$$

$$KDD = T_{MAX} - T_{UT} \tag{12}$$

There are different methods for calculating GDD. In this study, the most commonly used method in calculating GDD for corn is employed McMaster and Wilhelm, 1997). Constraints on maximum and minimum temperatures are applied for the purpose of eliminating the effect of low or high temperatures that prevent or retard the growth of corn. More specifically, before entering temperature data into Eqs. (11) and ((12), T_{MAX} and T_{MIN} are set equal to T_{BASE} if less than T_{BASE}, and set equal to T_{UT} when greater than T_{UT}. T_{BASE} and T_{UT} are set equal to 8 °C and 29 °C according to previous studies (Butler and Huybers, 2013; Schlenker and Roberts, 2009).

Since the PRISM meteorology data are served as continuous surfaces (spatial raster layers) with a spatial resolution at 4 km, zonal statistics are used by overlapping the data with county boundaries to estimate the average values of GDD, KDD, VPD, and precipitation at the county scale (*SI Appendix*, Figure S3). GDD, KDD, and precipitation data are aggregated by summing over the major growing season, and the VPD is aggregated using average value over the growing season. The growing season is considered from the start of May to the end of August, which is the typical growing period for corn in the Midwest.

2.3. Model evaluation

The STVC model is employed to study the variability of corn yields in response to meteorological covariates, including seasonal GDD, KDD,

Table 1
Summary of parameters and performance for OLS, Panel Regression with Spatial Fixed Effects, Panel Regression with Spatial and Time Fixed Effects, SVC, TVC, and STVC models.

Parameters	OLS	Panel regression (Spatial fixed effect)	Panel regression (Spatial and time fixed effect)	SVC	TVC	STVC
Coefficient for GDD (Bu/Ac/ °C)	0.06	0.07	0.08	0.07 (0.06,0.09)	0.03 (−0.09,0.21)	0.03 (−0.09,0.15)
Coefficient for KDD (Bu/Ac/ °C)	−0.28	−0.29	−0.19	−0.29 (−0.35,−0.24)	−0.21 (−0.44,0.01)	−0.18 (−0.44,0.10)
Coefficient for VPD (Bu/Ac/Pa)	−0.05 [NS]	−0.01	−0.05	−0.03 (−0.01, 0.01)	−0.02 (−0.23,0.12)	−0.02 (−0.16,0.11)
Coefficient for PCPN (Bu/Ac/mm)	−0.04	−0.04	−0.04	−0.03 (−0.08,0.00)	−0.02 (−0.16,0.19)	−0.00 (−0.18,0.15)
Moran's *I* of residuals	0.16 (0.08,0.43)	0.17 (0.08,0.49)	0.17 (0.08,0.48)	0.18 (0.06,0.47)	0.11 (0.07,0.22)	**0.10** (0.05,0.19)
R-Squared	0.40	0.45	0.64	0.60	0.64	**0.84**
Adj. R-Squared	0.40	0.43	0.63	0.56	0.64	**0.81**
RMSE (Bu/Ac)	20.32	16.94	13.62	16.48	15.68	**10.64**

Notes: All the parameters for the STVC model are posterior means. Adjusted R-squared for STVC is also calculated based on the posterior mean of fitted values. Coefficients denoted with N.S. represent they are not significant at the 0.05 confidence level. The coefficients can be interpreted as the changes of corn yield (Bu/Ac) induced by per unit change in meteorological factirs.

and precipitation cumulates and seasonal VPD mean at the county scale in three Corn Belt states in the Midwestern United States. Based on the historical corn yield and meteorology information, we comprehensively investigate the impacts of temperature and precipitation on the variation of corn yield over space and time. We seek to understand whether such impacts follow any spatially as well as temporally non-stationary process and compared them with commonly used statistical approaches.

2.3.1. OLS model

In an ordinary least squares (OLS) regression model for spatial data, the dependent variable is modeled as a linear function of a set of independent variables plus errors,

$$y_{(s)} = \mathbf{x}_{(s)}^{\mathbf{T}}\boldsymbol{\beta} + \varepsilon_{(s)} \tag{13}$$

where $y_{(s)}$ represents the observed yield at location s,$\mathbf{x}_s$ is a set of p covariates including a column of 1′s for the intercept, $\boldsymbol{\beta}$ is a $p \times 1$ vector of coefficients which can be estimated by OLS method, and $\varepsilon_{(s)}$ is white noise error.

2.3.2. Panel regression model

When spatial data are temporally referenced, we have spatiotemporal data to be considered in regression analysis. Another term "panel data" is often used to describe such data in econometrics (Mahalingam and Orman, 2018). Regression models developed for panel data are referred to as panel regression models. Compared to OLS, a panel regression model has an advantage in capturing the uniqueness of spatial effects. For example, a basic panel regression that incorporates spatial heterogeneity can be represented as:

$$y_{(s,t)} = \mathbf{x}_{(s,t)}^{T}\boldsymbol{\beta} + a_{(s)} + \varepsilon_{(s,t)} \tag{14}$$

where $y_{(s,t)}$ represents the observed yield at location s in time t, $\mathbf{x}_{\mathbf{s,t}}$ is a set of covariates specific to location and time, $a_{(s)}$ is a site-specific term for controlling time-invariant spatial heterogeneity. There are two distinct approaches to modeling this site-specific term (Hsiao, 2014). One is to treat $a_{(s)}$ as a fixed but unknown parameter to estimate. In this case, Eq. (14) is known as a fixed effects model. In our case, spatial fixed effects models are selected to control the omitted bias caused by time-invariant unobserved heterogeneity and result in a site-specific regression intercept for each county. Similarly, the panel regression with both spatial and time fixed effects leads to site- and time-variant regression intercepts.

2.3.3. Model comparison

For the comparison purpose, we run the following alternative models against the data: the OLS model in Eq. (13), the panel regression model in Eq. (14), SVC and TVC models defined as above. As suggested by Lobell and Burke (2010), the panel regression model includes a fixed spatial effects term to capture time-invariant heterogeneity, such as soil quality.

We evaluate our STVC model with regard to model fitting performance. The goodness of model fitting is measured through the coefficient of determination adjusted by the number of covariates (adjusted R-squared). Since adjusted R-squared is not reported by Bayesian methods, we use the posterior mean of estimated values to calculate the equivalent adjusted R-squared for the STVC model. Besides, Moran's *I* (Moran, 1950) is used to measure spatial autocorrelation of estimation residuals by Eq. (15).

$$I = \frac{n}{S_0}\frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}(z_i - \bar{z})(z_j - \bar{z})}{\sum_{i=1}^{n}(z_i - \bar{z})} \tag{15}$$

where n is the number of samples; z_i is the value of the residual yield at county i; z_j is the residual yield at other counties (where $j \neq i$); $\bar{z}$ is the mean value of residual yield across study area; $S_0 = \sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}$ and ***wij*** is the spatial weighting between county *i* and county *j*. Moran's *I* ranges from −1 to +1, with +(-)1 indicating perfect dispersion and correlation, respectively. A positive value of Moran's *I* indicates clustered estimation residuals, and a near-zero Moran's *I* indicates less spatially correlated estimation residuals. In our study, the spatial weighting was defined based on an inverse distance between counties.

To evaluate the model's performance, the root mean squared error (RMSE) defined in Eq. (16) is then calculated over all 293 counties as:

$$RMSE = \sqrt{\sum_{s}\sum_{t}\left(Yield_{s,t} - \widehat{Y}ield_{s,t}\right) \Big/ (N_s \times N_t)} \tag{16}$$

where $\widehat{Y}ield_{s,t}$ represents the predicted yield value at county *s* and year *t*; N_s and N_t represent the total number of counties and years.

3. Results and discussion

3.1. Models' performance in simulating yield responses

STVC model, based on spatiotemporally non-stationary processes, can be used to fit reliable yield-meteorology relationships. The STVC outperforms alternative models, including OLS, spatial panel regression, SVC, and TVC, in fitting detrended yield with the lowest RMSE at 10.64 Bu/Ac (0.72 Mg/ha, Table 1). The accuracy of the panel regression with fixed spatial and time effect is comparable to STVC with RMSE of 13.62 Bu/Ac (0.92 Mg/ha). The other four models with either spatial or

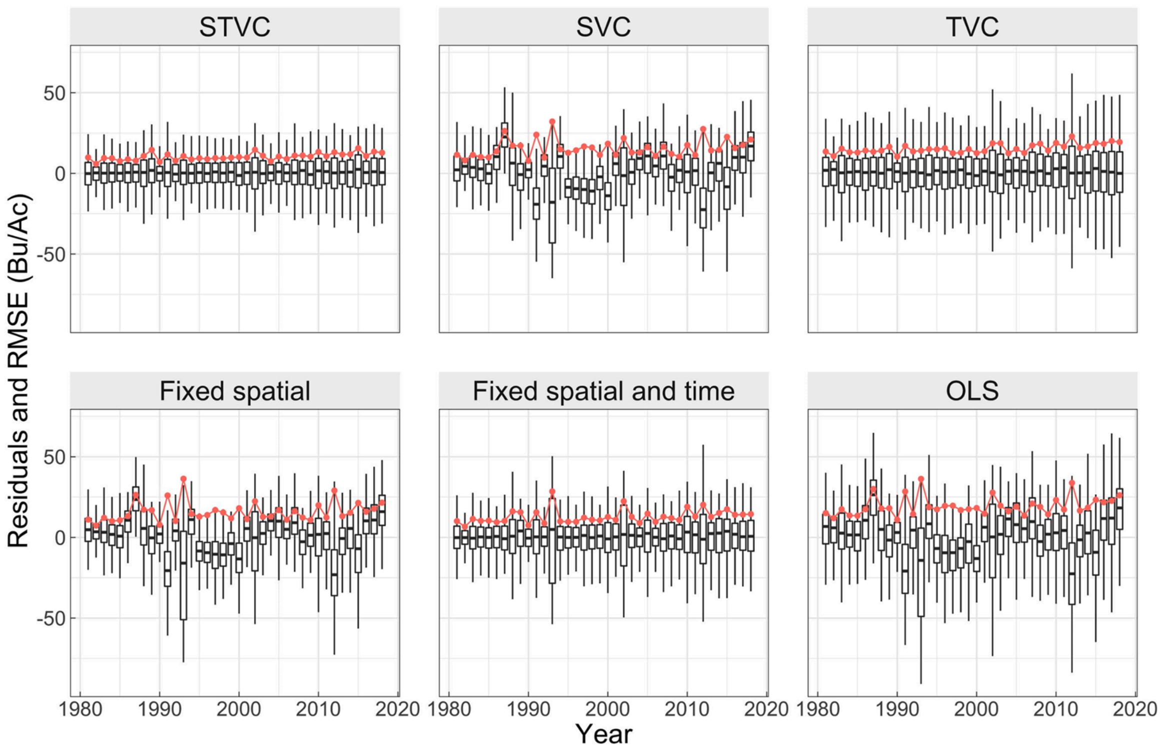

Fig. 2. Boxplot of the estimation residuals and RMSE (red lines) of different methods at each year during 1981–2018. Vertical bars show interquartile range (25th to 75th percentile), and the whiskers indicate the 5th and 95th percentiles. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

temporal non-stationary processes result in higher estimation errors, whose RMSE values range between 15.68 and 20.32 Bu/Ac (1.06 and 1.37 Mg/ha). According to R^2, the STVC achieves the highest value of 0.84 among all models, compared to 0.60 for STV and 0.64 for TVC. The spatiotemporally non-stationary processes have improved the performance for linear regressions. With the expense of additional parameters of spatially varying intercepts in its mean structure, the spatial panel model results in a higher R^2 than OLS (0.45 vs. 0.40). The spatiotemporally varying intercepts structure made the panel model gain considerably, with a R^2 value of 0.64. The improved performance indicates that non-stationary spatiotemporal processes more appropriately describe the relationship between meteorology and yields.

3.1.1. Spatial autocorrelation of residuals

STVC model accounts for spatial autocorrelation and results in less spatially clustered residuals. Interestingly, temporal non-stationarity becomes a key factor in resolving the spatial autocorrelation. For all competing models, the values of Moran's *I* are positive. The results demonstrate that there is a significant ($p < 0.05$) positive spatial autocorrelation of estimation residues, meaning that adjacent counties tend to have a similar level of estimation bias. However, by allowing the meteorological coefficients to change over time, STVC and TVC significantly reduce mean value of Moran's *I* of residuals to 0.10 and 0.11, compared to 0.18 of SVC ($p < 0.01$ for all years and models, Table 1). They also demonstrate a higher estimation accuracy with RMSE of 10.74 and 15.68 Bu/Ac (0.72 and 1.06 Mg/ha) for STVC and TVC, comparing to 16.48 Bu/Ac (1.11 Mg/ha) for SVC, despite the direct links between non-stationary spatial structure of SVC and spatial associations (Table 1).

Although the SVC model, which allows different coefficients in space, provides a better estimation accuracy than OLS, it results in a more evident spatial clustering of residuals. For OLS, a strong positive spatial autocorrelation of residuals is detected by Moran's *I* index with an average value of 0.16 ($p < 0.01$ for all years). In the OLS model, the residuals are assumed to be independent and normally distributed and represent vertical distances between the actual values of the dependent variable and their mean values. The estimates for the coefficients are chosen to minimize the sum of squared residuals (SSR). When OLS is applied to spatial data, the residuals are often correlated rather than being independent (Lobell and Burke, 2010). However, Moran's *I* of residuals are higher in panel regression and SVC models, 0.17 for two types of Panel regression, and 0.18 for the SVC model ($p < 0.01$ for all years and models), indicating a severe spatial autocorrelation of errors (Table 1). Therefore, these results indicate spatially dependent estimated residuals even for the panel with site and year varied intercept terms and SVC with site-varied coefficients.

3.1.2. Inter-annual variations in models' performance

The STVC model provides reduced interannual variations in model performance than competitive models. Considering the mean of estimation residuals, a designed temporal structure is important for unbiased results. STVC, TVC, and Panel regression with fixed spatial and time effects have zero means across all 38 years. However, the temporal inconsistency of residuals' distribution can be found in models without a dedicated temporal structure. OLS, SVC, and Panel regression models, which are temporal-invariant models, result in significant nonzero means ($p < 0.05$) in 31, 34, and 35 years (Fig. 2). For OLS, Panel, and SVC models, their peak values of Moran's *I* for residuals exceed 0.40 in 1993 ($p < 0.01$), indicating highly spatial correlated errors (*SI Appendix*, Figure S4). Meanwhile, the STVC has a lower lvel of spatially correlated errors with a peak value of 0.19 in 1981 ($p < 0.01$; *SI Appendix*, Figure S4). Besides, the STVC model provides a higher level of accuracy in capturing year-to-year differences. As measured in each year, the RMSE values of the STVC model range from 5.89 Bu/Ac (0.40 Mg/ha) in 1982 to 15.60 Bu/Ac (1.05 Mg/ha) in 2015 (Fig. 2). In contrast, the highest RMSE is 36.37, 28.50, and 36.27 Bu/Ac (2.45, 1.92, and 2.44 Mg/ha) in 1993 for panel with fixed spatial effect, the panel with spatial time effect, and OLS models, respectively (Fig. 2).

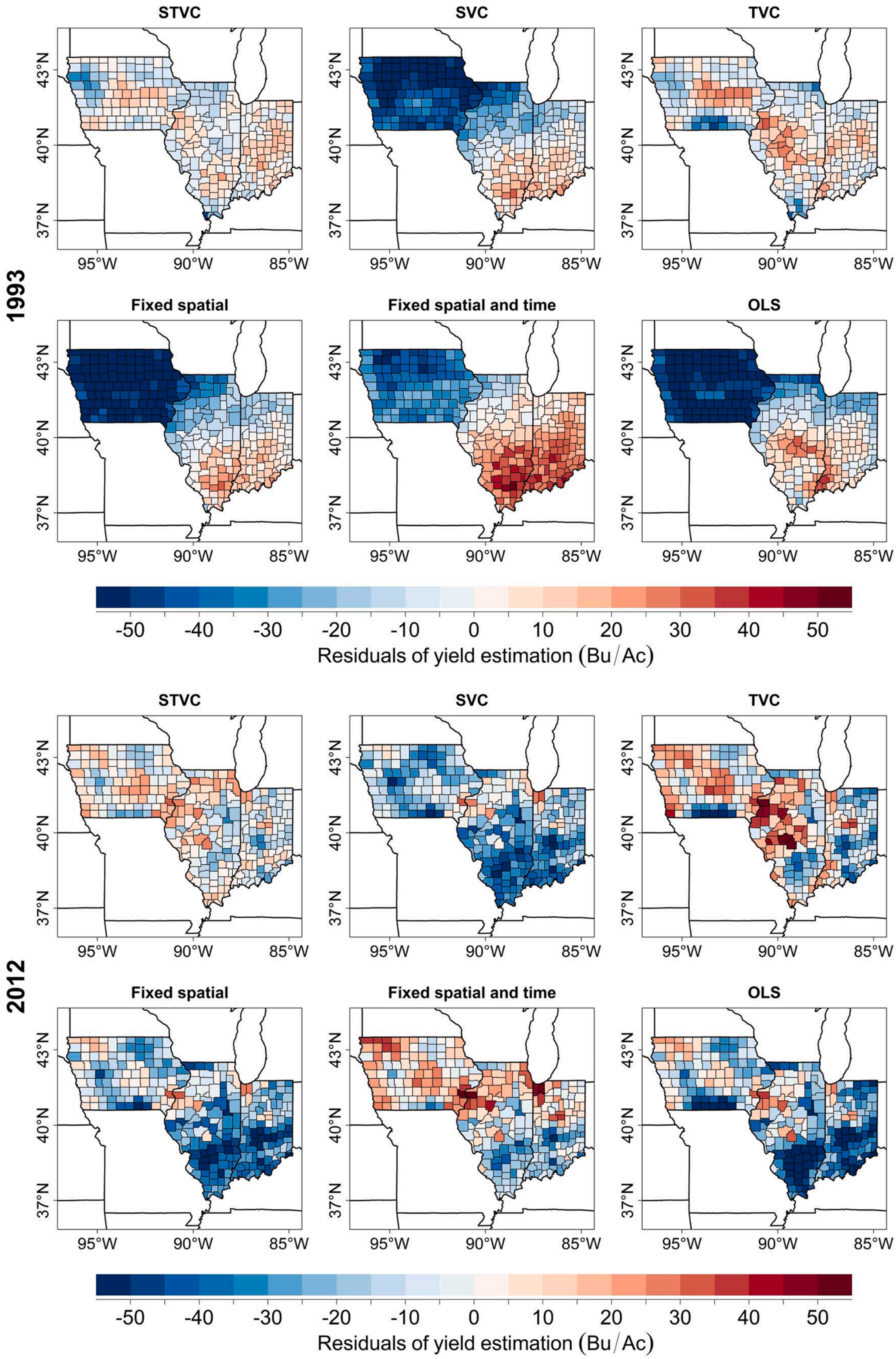

Fig. 3. Spatial distribution of yield estimation residuals for different methods in 1993 and 2012.

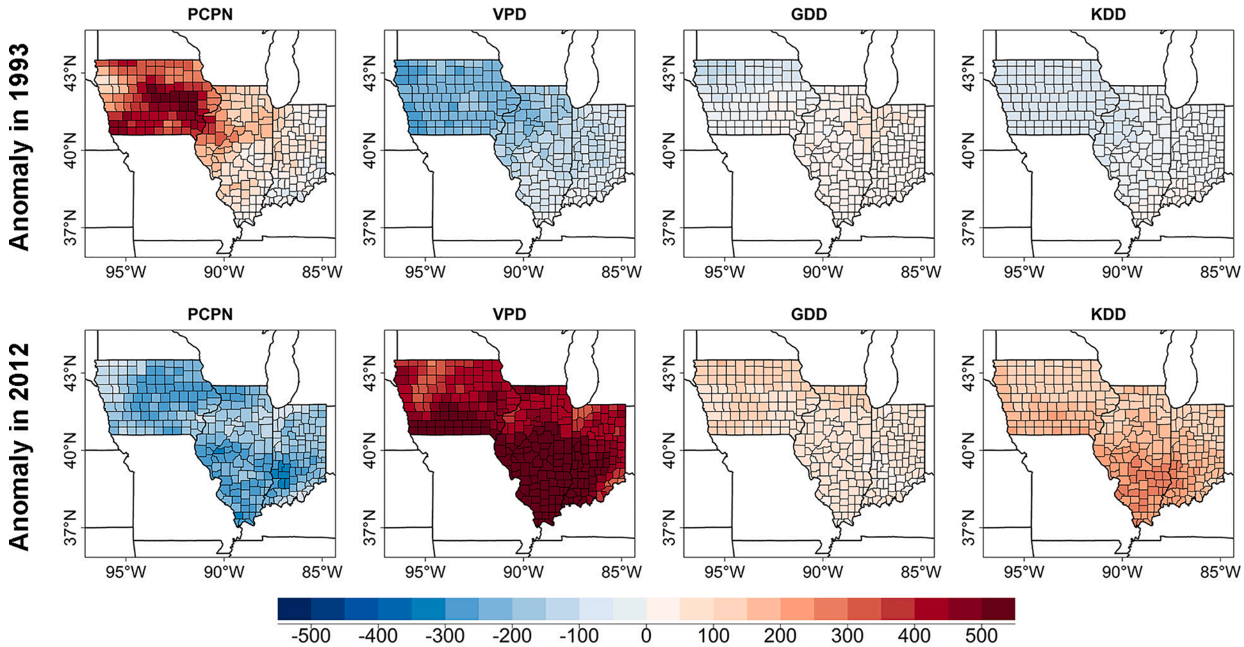

Fig. 4. Meteorology anomaly of four meteorological factors in 1993 and 2012, calculated by the departure in 1993 and 2012 from the mean of 1981–2018.

3.2. Simulating yield anomalies under extreme climate events

STVC is a robust estimator under severe flooding events. When extreme flooding occurred in 1993, a severe corn yield loss can be found in Iowa than Illinois and Indiana (*SI Appendix*, Figure S2). Under such a circumstance, OLS, Panel regressions, and SVC models significantly underestimate the yield loss caused by excessive precipitation (Fig. 3). The underestimated yields are also clustered in the south area for panel regression with fixed spatial and time effects (Fig. 3). In contrast, the STVC and TVC models with embedded non-stationary temporal processes provide nearly randomly distributed residuals across the counties.

STVC also provides a reliable tool for reproducing yield response under extreme drought and heat events. In 2012, the extreme heat and drought profoundly threatened corn production, especially in Indiana (Fig. 4). However, all models, except for STVC and TVC, overestimate the yield in most of the southern counties during this year. Differences between temporal-variant and temporal-invariant models include: (1) STVC estimates a negative impact of GDD in 2012 (−0.01 Bu/Ac, equals −0.67 Kg/ha, in average) while other temporal-invariant models keep a

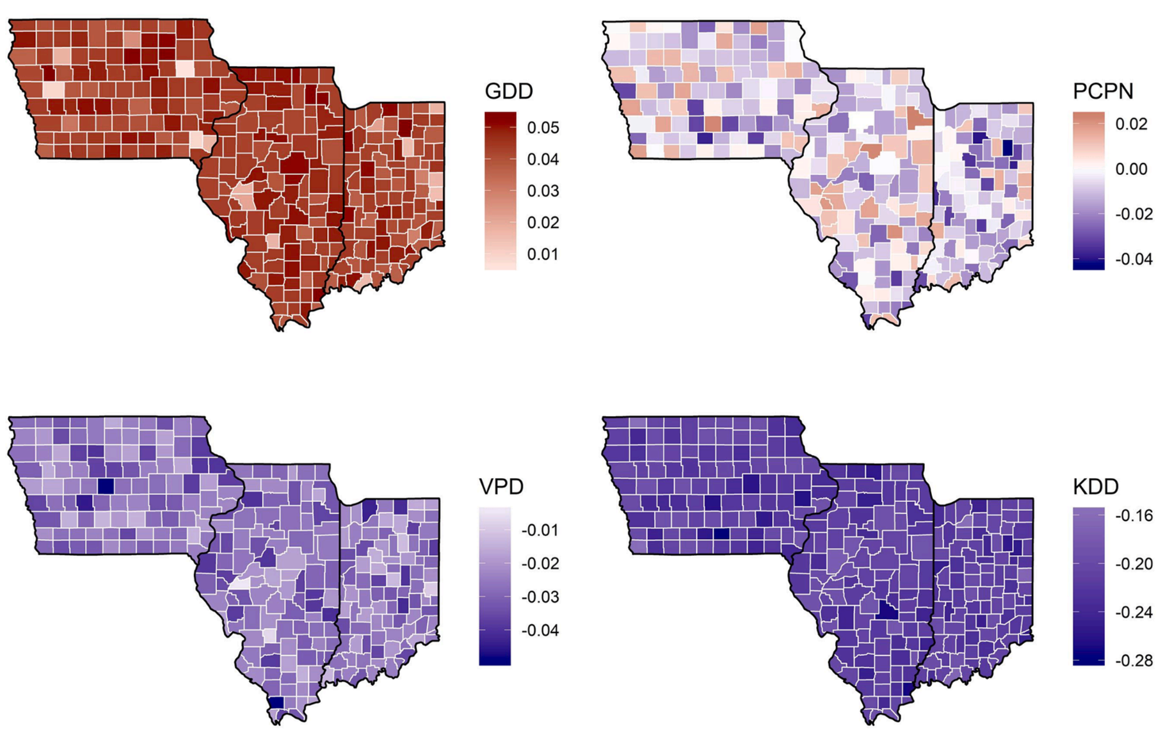

Fig. 5. STVC estimated county-level coefficients for the GDD, KDD, VPD, and precipitation averaged from 1981–2018 (only include the coefficients significant at 0.05 level).

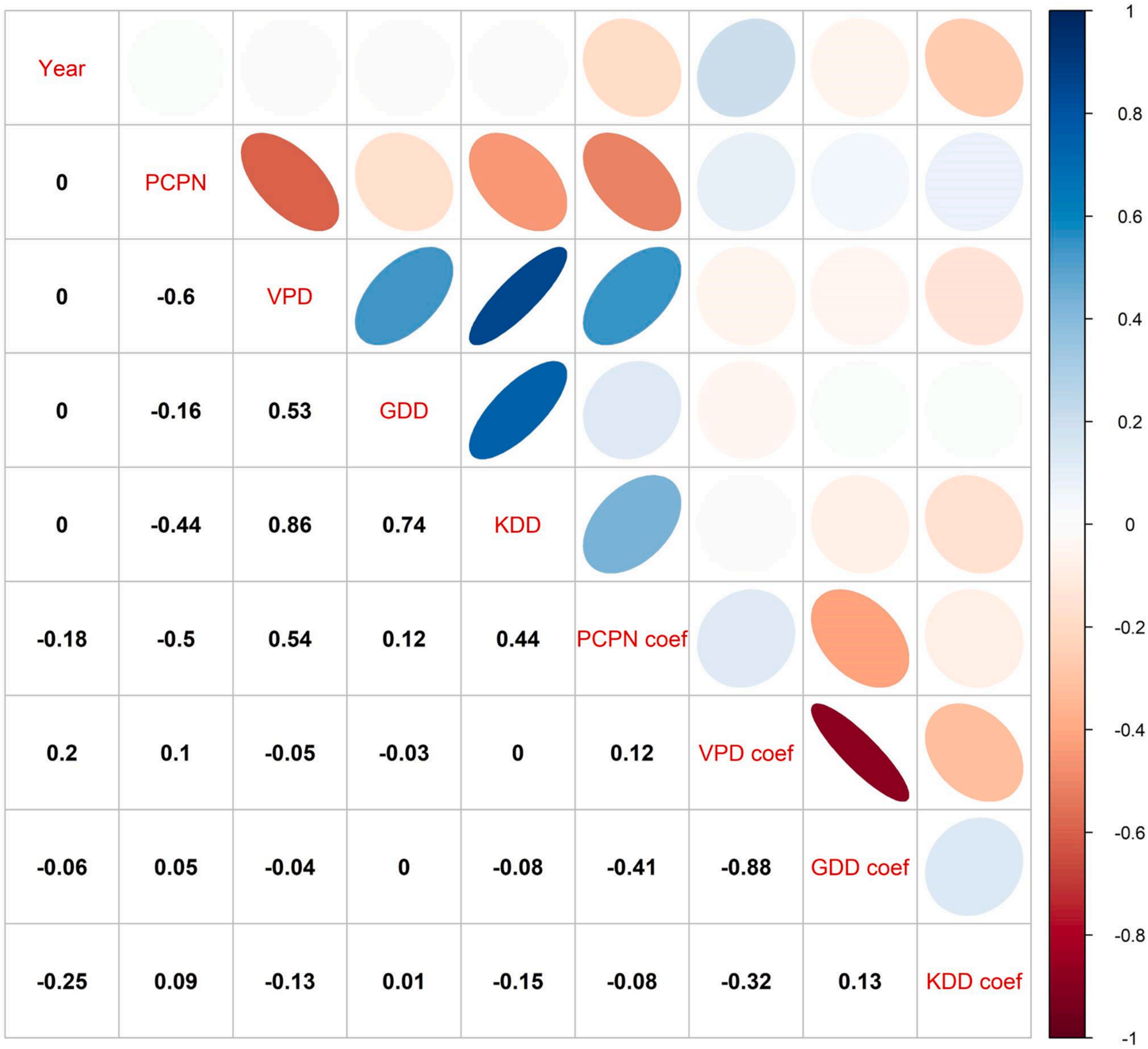

Fig. 6. Correlation matrix between STVC estimated county-level yield sensitivity and meteorological factors.

positive impact (Table 1); (2) STVC suggests a severer yield loss due to one degree increase of KDD compare to OLS (−0.30 vs. −0.28 Bu/Ac, equal 20.19 and 18.84 Kg/ha, in average; *SI Appendix*, Table S1). These evidences indicate the importance of modeling temporal varying crop responses to resolve crop production's spatial autocorrelationof.

The success of STVC is partially attributed to its unique structure. STVC is markedly different from other approaches in its structure, where time-space dynamics are considered in a hierarchical order. Disentangling the complex form of spatiotemporally correlated data is critical for understanding yield responses. The OLS tends to ignore spatial and temporal patterns because commonly used regression methods fail to account for either spatial or temporal correlation. On the fitting side, spatially correlated observations do not satisfy the independence assumption, which is required for unbiased coefficients estimation for OLS. In the SVC, the ICAR in WinBUGS software is used to take nearby counties into account. In the TVC, the random walk simulates the year-to-year association. However, isolating the spatial- or temporal- correlation is insufficient for explicit yield dependencies. By integrating ICAR and temporal random walk, STVC has the potential to account for both spatial and temporal correlations with flexibility in defining the hierarchical structure.

The STVC provides insights into meteorological yield response on a regional scale for process-based models. Comparing to process-based models, quantifying meteorological impacts using STVC requires relatively less effort on a regional scale. Therefore, STVC has the potential to be a robust benchmark of regional yield response for crop process-based models to scale up from field to large scale. Crop processed-based models could adjust the parameters to prompt its yield estimation robustness. Previous studies found underestimated yield under extreme heat events, which could partly due to inappropriate temperature response functions (Wang et al., 2018). For example, the CERES-Maize model explicitly simulates the heat stress impact on yield using temperature relevant functions for kernel number and kernel growth (Jones and Kiniry,1986). When the temperature exceeds the optimal, the CERES-Maize model assumes the kernel growth would stop. However, the occurrence of extreme heat could do harm to the plant and slow down the kernel growth even when the temperature returns to optimal. With the estimated yield loss and KDD coefficient using STVC, a negative impact on final yields of accumulated extreme heat can be better described.

3.3. Role of spatiotemporally non-stationary processes in the yield response analysis

The STVC model computes non-stationary coefficients at the spatial dimension. Coefficients estimated from all models considered in this study are summarized in Table 1. As the heterogeneity of maize yield rose in the Midwest U.S., the fixed spatial effect could cause considerable uncertainty in regions with diversified landscape and agronomy conditions (Lobell and Azzari, 2017). OLS and panel regression analysis

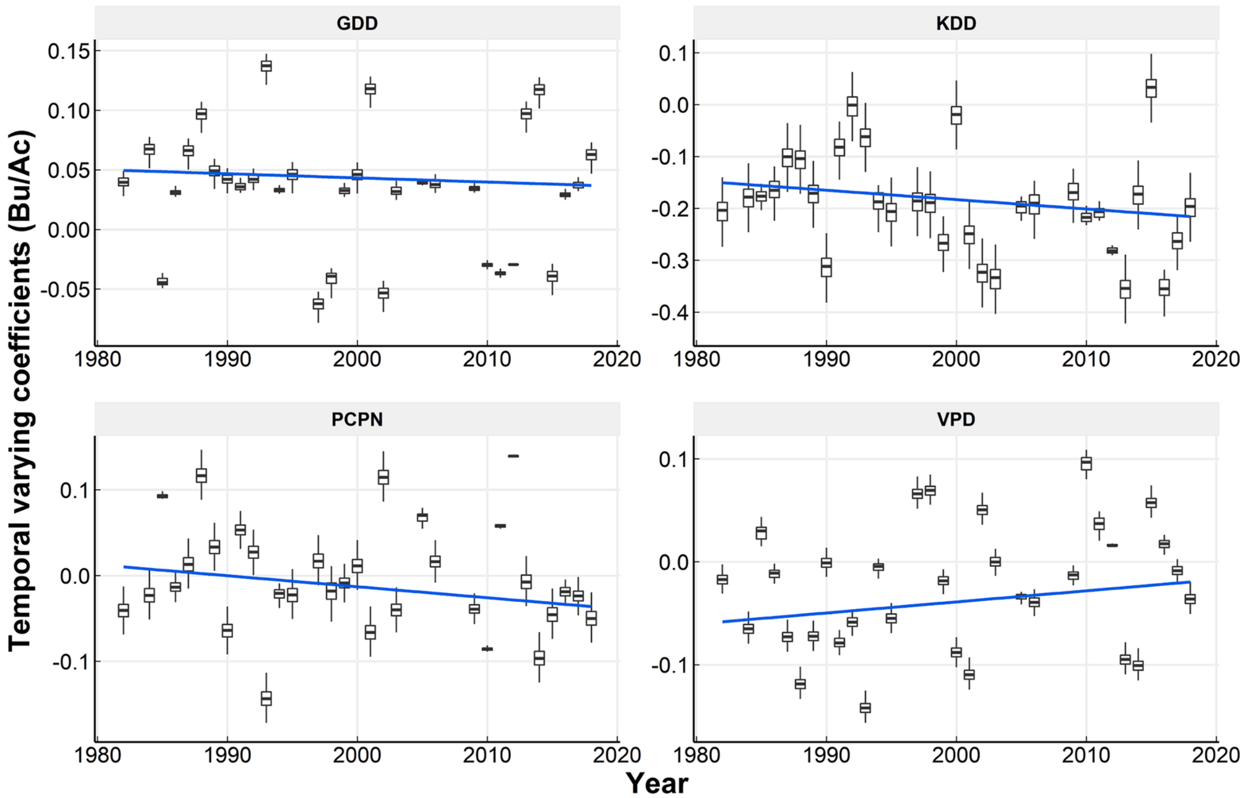

Fig. 7. Boxplot of temporal varying coefficients for various meteorological factors (only include the coefficients significant at 0.05 level). Vertical bars show interquartile range (25th to 75th percentile), and the solid lines indicate the regression line over the years.

can only result in temporal-invariant coefficients. The results from the OLS model demonstrate that corn yields are negatively associated with precipitation and KDD but have a slightly positive relationship to GDD, and the coefficient for VPD is not significant at 0.05 level. Compared to OLS, the panel regression models with spatial fixed effect and spatial-time fixed effect suggest a more significant relationship between VPD and crop yields. In contrast, for the STVC model, the coefficients for each meteorological factor are not fixed in spatial and temporal dimensions. The spatial non-stationarity of coefficients estimated by the STVC model is first examined by averaging out their changes over time (Fig. 5). When temporally varying coefficients are also included in testing the significance of parameters in STVC, it generates 38 parameter surfaces to represent the significance of parameters in different years (Fig. 5).

STVC model accounts for the temporal variation of coefficients by introducing a random walk process to allow spatially varying coefficients to change over time. Identifying temporally inconsistent impact on crop yields is critical for understanding long-term yield anomalies (Ceglar et al., 2020). We further investigate the correlation between the temporal variation of coefficients and meteorological factors. We find that the yield response towards precipitation is highly correlated with the amount of seasonal precipitation (Fig. 6). For example, in the years 1988 and 2012, the coefficients of precipitation increase to 0.12 Bu/Ac (9.08 Kg/ha; *SI Appendix*, Table S1). Considering the nearly zero value of historical average (Table 1), it's a dramatic increase. It is found that in 1988 and 2012, the average precipitation attains the lowest level among the years we included. That means the marginal increase of corn yield is significant under conditions with insufficient rains. These results are consistent with the common sense that excessive or inadequate precipitation could put pressure on rain-fed corn.

The STVC model with spatiotemporally non-stationary processes can contribute to large scale parameterization for crop models. Gridded crop models are useful tools for yield and risk assessment (Folberth et al., 2019). Because the crop process-based models are usually calibrated at the field level, proper spatial parameterization is necessary for large scale applications in the form of gridded models. However, as climatic and agronomic conditions varied in space and time, inappropriate parameterization of crop models may result in huge uncertainties because of complex crop and environment interactions (Bassu et al., 2014). Various types of information, e. g. management, are needed to articulate spatiotemporally non-stationary crop responses but unavailable at the large scale for process-based models. The STVC could bridge the gap by providing a quantitive measurement of spatiotemporally varied yield responses using limited meteorology information and historical observations.

3.4. Temporal trends of yield response

Coefficients estimated by STVC suggest that the corn production become more resilient to drought but more sensitive to high temperature. The STVC model demonstrates the changes in yield response to various meteorological indicators over time. By grouping all '3I' states, the corn yield response to KDD, GDD, and precipitation becomes increasingly negative over time with correlation coefficients of -0.25, -0.06, and -0.18, while the response to VPD is positively correlated with time, with a correlation of 0.2 (Fig. 6). Among these factors, the KDD has the steepest time slope, while the GDD has the most moderate slope (Fig. 7). Importantly, all predictors included in our analysis show insignificant time trends regardless of evident peaks or valleys (*SI Appendix*, Figure S3). Therefore, despite the heavier climate pressures for some years (drought and heat in 1988 and 2012, and flooding in 1993), changes in corn yield coefficients are partially caused by time-dependent factors other than climate change.

One possible explanation is that the continuously improved cultivar is relevant to the time-variant relationships between corn yields and meteorology. Considering the dominant role of precipitation for rain-fed corn, hybrids with the drought-tolerant trait have been bred to increase yield (Cooper et al., 2014). The estimated increasing yield loss due to one unit of KDD is consistent with the previous study that high temperature has caused more severe yield decline in Indiana (Libecap and Steckel, 2011). An important factor that influences the yield is the length of the grain filling phase, which is the key stage of dry matter accumulation and is sensitive to heatwaves. Recent research provides evidence for a lengthened grain filling phase of modern cultivars (Zhu et al., 2018). Therefore, the renewal cultivar is likely to suffer more heat stress in the key growth stage. In addition to improved cultivar, other factors (such as sowing density and CO_2 concentration) may also change the response to KDD (Lobell et al., 2013). Furthermore, many relevant covariables, such as management and soil property, are not considered in this research. Therefore, our yield response analysis still contains various uncertainties. To better understand the spatiotemporal variations of crop development, extra factors and spatial effect should be further investigated.

One limitation of this work is that we have not fully explored possible temporal structures or more complicated spatiotemporal nested structures. As pointed out by (Gelfand et al., 2003), Eq. (6) is not the only way to decompose the spatiotemporally varying coefficients and may cause extra estimation uncertainties. More sophisticated forms of spatiotemporal relationships could be specified. One possible extension is to model the spatiotemporal coefficients as spatiotemporal processes governed by non-stationary space-time dependency structure (Shand and Li, 2017). Moreover, it will be interesting to compare different temporal processes, e.g., autoregressive process versus random walk, to better understand temporal structures.

4. Conclusions

In this study, we develop a novel spatiotemporally varying coefficient (STVC) model to investigate the non-stationary relationship between meteorology and corn yields in the Midwestern U.S. counties from 1981 to 2018. The model treats regression effects as spatially and temporally correlated processes within a Bayesian framework to enable statistical inference on the regression associations. The results indicate heterogeneous spatial effects of GDD, KDD, VPD, and precipitation on crop yields. Meanwhile, the changes in impacts over time are revealed. Compared to alternative models, including OLS and spatial panel regression, the STVC model significantly improves the amount of variation in corn yields that can be explained by meteorology variabilities after taking spatiotemporal non-stationarity into account. More specifically, changes in GDD, KDD, VPD, and precipitation account for 84% of the variations in corn yields. The marginal increase of corn yields is significant under conditions with insufficient rains, while high temperatures accelerate the rate of decrease in corn yields.

The STVC model accommodates the non-stationarity of regression coefficients in both spatial and temporal dimensions. The model directly captures non-stationary processes in coefficients by its design, rather than allowing them to be reflected through the error terms. Therefore, the issue of spatial autocorrelation in residuals is mitigated. From an application perspective, the contribution of our study helps better understand the relationship between corn yields and meteorological factors. Although previous studies have been conducted to examine the spatial heterogeneity of the relationship, this study represents the first effort that incorporates the temporal variation of the relationship. By revealing this variation, the data can be further utilized through more confirmatory analysis. For instance, given the fact that the impact of meteorology on crop yields is not constant over space and time, further analysis can be conducted to examine the potential factors, such as meteorology and non-meteorology related, which contribute to the variation.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was partially funded by the National Key Research and Development Program of China under Grant Number 2017YFD0700605 and National Natural Science Foundation of China under Grant Number 32071894 and 31701316.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.agrformet.2021.108340.

References

An, L., Tsou, M.-H., Crook, S.E.S., Chun, Y., Spitzberg, B., Gawron, J.M., Gupta, D.K., 2015. Space–time analysis: concepts, quantitative methods, and future directions. Ann. Assoc. Am. Geogr. 105, 891–914. https://doi.org/10.1080/00045608.2015.1064510.

Banerjee, S., Carlin, B.P., Gelfand, A.E., Carlin, B.P., Gelfand, A.E., 2014. Hierarchical Modeling and Analysis for Spatial Data. Chapman and Hall/CRC. https://doi.org/10.1201/b17115.

Bassu, S., et al., 2014. How do various maize crop models vary in their responses to climate change factors? Glob. Change Biol. 20, 2301–2320. https://doi.org/10.1111/gcb.12520.

Besag, J., Kooperberg, C., 1995. On conditional and intrinsic autoregressions. Biometrika 82, 733–746. https://doi.org/10.1093/biomet/82.4.733.

Besag, J., York, J., Molli, A., 1991. Bayesian image restoration, with two applications in spatial statistics. Ann. Inst. Stat. Math. 43, 1–20. https://doi.org/10.1007/BF00116466.

Bornn, L., Zidek, J.V., 2012. Efficient stabilization of crop yield prediction in the Canadian Prairies. Agric. For. Meteorol. 152, 223–232. https://doi.org/10.1016/j.agrformet.2011.09.013.

Butler, E.E., Huybers, P., 2013. Adaptation of U.S. maize to temperature variations. Nat. Clim. Chang. 3, 68–72. https://doi.org/10.1038/nclimate1585.

Choi, J., Lawson, A.B., Cai, B., Hossain, M.M., Kirby, R.S., Liu, J., 2012. A Bayesian latent model with spatio-temporally varying coefficients in low birth weight incidence data. Stat. Methods Med. Res. 21, 445–456. https://doi.org/10.1177/0962280212446318.

Butler, E.E., Huybers, P., 2015. Variations in the sensitivity of US maize yield to extreme temperatures by region and growth phase. Environ. Res. Lett. 10, 034009 https://doi.org/10.1088/1748-9326/10/3/034009.

Ceglar, A., Zampieri, M., Gonzalez-Reviriego, N., Ciais, P., Schauberger, B., Velde, M.V., der, 2020. Time-varying impact of climate on maize and wheat yields in France since 1900. Environ. Res. Lett. 15, 094039 https://doi.org/10.1088/1748-9326/aba1be.

Cooper, M., Gho, C., Leafgren, R., Tang, T., Messina, C., 2014. Breeding drought-tolerant maize hybrids for the U.S. corn-belt: discovery to product. J. Exp. Bot. 65, 6191–6204. https://doi.org/10.1093/jxb/eru064.

Dong, G., Ma, J., Harris, R., Pryce, G., 2016. Spatial Random slope multilevel modeling using multivariate conditional autoregressive models: a case study of subjective travel satisfaction in Beijing. Ann. Am. Assoc. Geogr. 106, 19–35. https://doi.org/10.1080/00045608.2015.1094388.

Folberth, C., Baklanov, A., Balkovič, J., Skalský, R., Khabarov, N., Obersteiner, M., 2019. Spatio-temporal downscaling of gridded crop model yield estimates based on machine learning. Agric. For. Meteorol. 264, 1–15. https://doi.org/10.1016/j.agrformet.2018.09.021.

Gelfand, A.E., Kim, H.-J., Sirmans, C.F., Banerjee, S., 2003. Spatial modeling with spatially varying coefficient processes. J. Am. Stat. Assoc. 98, 387–396. https://doi.org/10.1198/016214503000170.

Gilks, W.R., 1995. Markov Chain Monte Carlo in Practice, 1st ed. Chapman and Hall/CRC. https://doi.org/10.1201/b14835.

Hoeting, J.A., 2009. The importance of accounting for spatial and temporal correlation in analyses of ecological data. Ecol. Appl. 19, 574–577. https://doi.org/10.1890/08-0836.1.

Hsiao, C., 2014. Analysis of Panel Data, 3rd ed. Cambridge University Press, Cambridge. https://doi.org/10.1017/CBO9781139839327.

Jiang, P., He, Z., Kitchen, N.R., Sudduth, K.A., 2009. Bayesian analysis of within-field variability of corn yield using a spatial hierarchical model. Precis. Agric. 10, 111–127. https://doi.org/10.1007/s11119-008-9070-4.

Jones, C.A., Kiniry, J.R., 1986. CERES-Maize; a Simulation Model of Maize Growth and Development. Texas A &M Univ. Press, College Station.

Jones, J.W., Hoogenboom, G., Porter, C.H., Boote, K.J., Batchelor, W.D., Hunt, L.A., Wilkens, P.W., Singh, U., Gijsman, A.J., Ritchie, J.T., 2003. The DSSAT cropping system model. Eur. J. Agron. 18, 235–265. https://doi.org/10.1016/S1161-0301(02)00107-7.

Junker, N.W., Schneider, R.S., Fauver, S.L., 1999. A Study of heavy rainfall events during the Great Midwest flood of 1993. Wea. Forecast. 14, 701–712. https://doi.org/10.1175/1520-0434(1999)014<0701:ASOHRE>2.0.CO;2.

Li, Y., Guan, K., Schnitkey, G.D., DeLucia, E., Peng, B., 2019. Excessive rainfall leads to maize yield loss of a comparable magnitude to extreme drought in the United States. Glob. Change Biol. 25, 2325–2337. doi:10.1111/gcb.14628.

Libecap, G.D., Steckel, R.H., 2011. The Economics of Climate Change: Adaptations Past and Present. University of Chicago Press.

Lobell, D.B., Burke, M.B., 2010. On the use of statistical models to predict crop yield responses to climate change. Agric. For. Meteorol. 150, 1443–1452. https://doi.org/10.1016/j.agrformet.2010.07.008.

Lobell, D.B., Hammer, G.L., McLean, G., Messina, C., Roberts, M.J., Schlenker, W., 2013. The critical role of extreme heat for maize production in the United States. Nat. Clim. Change 3, 497–501. https://doi.org/10.1038/nclimate1832.

Lobell, D.B., Roberts, M.J., Schlenker, W., Braun, N., Little, B.B., Rejesus, R.M., Hammer, G.L., 2014. Greater Sensitivity to Drought Accompanies Maize Yield Increase in the U.S. Midwest. Science 344, 516–519. https://doi.org/10.1126/science.1251423.

Lobell, D.B., Azzari, G., 2017. Satellite detection of rising maize yield heterogeneity in the U.S. Midwest. Environ. Res. Lett. 12, 014014 https://doi.org/10.1088/1748-9326/aa5371.

Mahalingam, B., Orman, W.H., 2018. GDP and energy consumption: a panel analysis of the U.S. Appl. Energy 213, 208–218. https://doi.org/10.1016/j.apenergy.2018.01.036.

McGrath, J.M., Betzelberger, A.M., Wang, S., Shook, E., Zhu, X.-.G., Long, S.P., Ainsworth, E.A., 2015. An analysis of ozone damage to historical maize and soybean yields in the United States. Proc. Natl. Acad. Sci. 112, 14390–14395. https://doi.org/10.1073/pnas.1509777112.

McMaster, G.S., Wilhelm, W.W., 1997. Growing degree-days: one equation, two interpretations. Agric. For. Meteorol. 87, 291–300. https://doi.org/10.1016/S0168-1923(97)00027-0.

Moran, P.a.P., 1950. Notes on continuous stochastic phenomena. Biometrika 37, 17–23. https://doi.org/10.1093/biomet/37.1-2.17.

Olesen, J.E., Trnka, M., Kersebaum, K.C., Skjelvåg, A.O., Seguin, B., Peltonen-Sainio, P., Rossi, F., Kozyra, J., Micale, F., 2011. Impacts and adaptation of European crop production systems to climate change. Eur. J. Agron. 34, 96–112. https://doi.org/10.1016/j.eja.2010.11.003.

Ozaki, V.A., Ghosh, S.K., Goodwin, B.K., Shirota, R., 2008. Spatio-temporal modeling of agricultural yield data with an application to pricing crop insurance contracts. Am. J. Agric. Econ. 90, 951–961. https://doi.org/10.1111/j.1467-8276.2008.01153.x.

Peng, B., Guan, K., Tang, J., Ainsworth, E.A., Asseng, S., Bernacchi, C.J., Cooper, M., Delucia, E.H., Elliott, J.W., Ewert, F., Grant, R.F., Gustafson, D.I., Hammer, G.L., Jin, Z., Jones, J.W., Kimm, H., Lawrence, D.M., Li, Y., Lombardozzi, D.L., Marshall-Colon, A., Messina, C.D., Ort, D.R., Schnable, J.C., Vallejos, C.E., Wu, A., Yin, X., Zhou, W., 2020. Towards a multiscale crop modelling framework for climate change adaptation assessment. Nat Plants 6, 338–348. https://doi.org/10.1038/s41477-020-0625-3.

Quiring, S.M., Papakryiakou, T.N., 2003. An evaluation of agricultural drought indices for the Canadian prairies. Agric. For. Meteorol. 118, 49–62. https://doi.org/10.1016/S0168-1923(03)00072-8.

Ray, D.K., Gerber, J.S., MacDonald, G.K., West, P.C., 2015. Climate variation explains a third of global crop yield variability. Nat. Commun. 6, 5989. https://doi.org/10.1038/ncomms6989.

Schlenker, W., Roberts, M.J., 2009. Nonlinear temperature effects indicate severe damages to U.S. crop yields under climate change. Proc. Natl. Acad. Sci. 106, 15594–15598. https://doi.org/10.1073/pnas.0906865106.

Shand, L., Li, B., 2017. Modeling non-stationarity in space and time. Biometrics 73, 759–768. https://doi.org/10.1111/biom.12656.

Shand, L., Li, B., Park, T., Albraccin, D., 2018. Spatially varying auto-regressive models for prediction of new human immunodeficiency virus diagnoses. J. Roy. Stat. Soc. C 67, 1003–1022. https://doi.org/10.1111/rssc.12269.

Sharma, A.I., Kabenge, I., Irmak, S., 2011. Application of GIS and geographically weighted regression to evaluate the spatial non-stationarity relationships between precipitation vs. irrigated and rainfed maize and soybean yields. Trans. ASABE 54, 953–972. https://doi.org/10.13031/2013.41227.

Troy, T.J., Kipgen, C., Pal, I., 2015. The impact of climate extremes and irrigation on U.S. crop yields. Environ. Res. Lett. 10, 054013 https://doi.org/10.1088/1748-9326/10/5/054013.

Waller, L.A., Zhu, L., Gotway, C.A., Gorman, D.M., Gruenewald, P.J., 2007. Quantifying geographic variations in associations between alcohol distribution and violence: a comparison of geographically weighted regression and spatially varying coefficient models. Stoch. Environ. Res. Risk Assess. 21, 573–588. https://doi.org/10.1007/s00477-007-0139-9.

Wang, N., Wang, E., Wang, J., Zhang, J., Zheng, B., Huang, Y., Tan, M., 2018. Modelling maize phenology, biomass growth and yield under contrasting temperature conditions. Agric. For. Meteorol. 250–251, 319–329. https://doi.org/10.1016/j.agrformet.2018.01.005.

Zhu, P., Jin, Z., Zhuang, Q., Ciais, P., Bernacchi, C., Wang, X., Makowski, D., Lobell, D., 2018. The important but weakening maize yield benefit of grain filling prolongation in the U.S. Midwest. Glob. Change Biol. 24, 4718–4730. https://doi.org/10.1111/gcb.14356.

PRISM Climate Group, 2004, PRISM Gridded Climate Data, http://prism.oregonstate.edu. (Accessed 20 December, 2019).

USDA, 2019. United States Department of Agriculture – National Agricultural Statistics Service Database Quick Stats, https://www.nass.usda.gov/Quick_Stats/. (Accessed 22 December 2019).