Center for Climate, Regional, Environmental and Trade Economics
www.create.illinois.edu

Revisiting the impact of climate change on agriculture through spatially-varying and place-tailored Ricardian estimates

Noé J Nava[a], Sandy Dall'erba[b], Chang Cai[c], and A. Stewart Fotheringham[d]

[a]Economic Research Service, U.S. Department of Agriculture, Kansas City, MO 64105, USA, noe.nava@usda.gov

[b]University of Illinois at Urbana-Champaign, Center for Regional, Environmental, and Trade Economics, dallerba@illinois.edu

[c]University of Illinois at Urbana-Champaign, Center for Regional, Environmental, and Trade Economics, ccai5@illinois.edu

[d]School of Geographical Sciences and Urban Planning, Arizona State University, sfotheri@asu.edu

CREATE Discussion Paper ##-22-02
February, 2022

Abstract: The Ricardian framework has been widely used to study the impact of climate change on agriculture across U.S. counties over the past few decades. While spatial heterogeneity of climate change is well-accepted, the literature struggle to reach an agreement on how to model it, hence leading to a wide range of forecasted impacts. This paper employs Multiscale Geographically-Weighted-Regression (MGWR) to avoid setting an a priori definition of heterogeneity and to generate county-specific marginal effects of climate change impacts. This manuscript tests the predictive power of our MGWR application with other functional forms found in the literature on a homogenized dataset of historical climate, demographic and soil quality controls. Our cross-validation exercise indicates that our MGWR approach has higher predictive power than studies that cluster spatial units, and that other approaches have a downward bias. We attribute the divergence in results to unspecified heterogeneity. Our place-specific marginal effects will help guide the development of place-tailored mitigation and adoption strategies to climate change.

Climate, Regional, Environmental and Trade Economics
65-67 Mumford Hall
1301 West Gregory Drive
Urbana, IL, 61801

1 Introduction

The estimation of damages from global warming, particularly in the agricultural sector, is receiving an increasing amount of attention. The Ricardian approach, a reduced-form hedonic analysis developed by Mendelsohn et al. (1994) to calibrate such an impact over a single cross-section of data, has been extensively employed to predict climate change induced damages on agriculture across the spatial units (farms, municipalities, regions) of the country under study, including the U.S. (Schlenker et al., 2005; Schlenker and Roberts, 2006; Dall'erba and Domínguez, 2016), China (Liu et al., 2004; Wang et al., 2009) and South-Africa (Gbetibouo and Hassan, 2005).[1] For a single country, however, the consequences of climate change can diverge widely across studies. Cai and Dall'Erba (2021) investigate the role of group-membership i.e., clustering spatial units that are thought to respond similarly to climate change, in the U.S. Their findings suggest that a priori definitions of group-membership are the single determinant that explains the divergence in results across Ricardian studies for U.S. counties.

The global nature of the marginal effect of climate on spatial units has motivated the development of such grouping strategies. Early contributions of the Ricardian literature (see the work of Mendelsohn *et al.*, 1994, and Mendelsohn and Dinar, 2003) have been criticized for not fully accounting for region specific features. Subsequent studies, such as Schlenker et al. (2005), Deschênes and Greenstone (2007), and Dall'erba and Domínguez (2016), have remedied this shortcoming by grouping counties based on irrigation status, state boundaries and elevation. In addition, studies add a quadratic term to the climate covariates such that the marginal effect is dependent on space; but as we demonstrate in the next section, a quadratic term does not fully incorporate the role of geographic determinants into the estimated impact of climate. These approaches allow the generation of between two and fifty different marginal effects for each covariate. However, Cai and Dall'Erba (2021) indicates that among the county groupings chosen in the literature focusing on the US agriculture, there is little evidence that one grouping outperforms the other in terms of predictive power.

[1] For an exhaustive review of these Ricardian applications see Mendelsohn and Massetti (2017).

Ex-ante grouping definitions raise endogeneity concerns, so a potential solution to this confusion is to generate spatially varying estimates without any a priori clustering of counties via local regression techniques. Then, clear county-specific marginal effects can be calculated and place-tailored mitigation strategies can be derived from them. Several empirical applications of local regression techniques have outperformed global approaches in terms of their predictive power (see, among others Paez *et al.*, 2008, Zhang *et al.*, 2011, and Soler and Gemar, 2018). However, few applications of local regression models focus on agriculture and, when they do, their outcome variable is crop yield.[2] Consequently, spatial heterogeneity in the marginal effects of climate change has been underexplored. We focus on the United States, a country that has been subject to a large number of previous county-level studies with which we can compare our approach using the newly-developed Multiscale Geographically Weighted Regression (MGWR) framework (Fotheringham et al., 2017).

Our primary contribution, therefore, is to demonstrate the predictive gains from using the MGWR estiamator compared to the global approach, with or without spatial heterogeneity, used in the works of Mendelsohn et al. (1994), Dall'erba and Domínguez (2016) and Schlenker et al. (2005). In all four cases, we rely on the same data, measured in 2012, and the same functional form and covariate choices as defined in Mendelsohn et al. (1994). We are aware that the more recent contributions of Dall'erba and Domínguez (2016) and Schlenker et al. (2005) use more and different covariates, and that other contributions, such as Deschênes and Greenstone (2007), have extended the Ricardian approach to a panel context. Nevertheless, it is only by comparing models for which the only difference is the definition of spatial heterogeneity that we can assess the relative advantage of our MGWR framework compared to global approaches. We do so by comparing the capacity of each model to accurately predict the 2017 observed values when relying on marginal effects calibrated in 2012. We find that our MGWR estimator outperforms previous handlings of the spatial heterogeneity in terms of predictive accuracy. For example, previous approaches predict losses when we observe gains, and only our MGWR predicts gains.

[2] See for example Cai et al. (2014), and Shiu and Chuang (2019).

We attribute this difference to a miss-specification of clustering in previous studies that can be corrected with the use of MGWR.

In the next section, we develop the conceptual framework of our spatial non-stationary Ricardian approach that extends the original idea of Mendelsohn et al. (1994) to a local, multiscale structure. We describe in section 3 the data employed in our estimations and in the cross-validation exercise. Contrary to the a priori clustering approaches adopted by previous studies, MGWR allows us to be agnostic about any possible clustering of counties in terms of their response to climate change. In order to show the appeal of this approach to our case study, we employ in section 4 a cross-validation approach on the data observed in 2017. Results provide evidence of the relative greater predictive power of our local approach compared to any previous specification.

2 Spatial non-stationary Ricardian approach

The idea at the core of the Ricardian approach is that landowners adapt their production choices to new climate conditions. The decisions made vis-à-vis the optimal use of the land are reflected in the land price which is commonly modeled as the discounted sum of future net returns (Mendelsohn et al., 1994; Plantinga et al., 2002):

$$V_i = \int \pi_i^*(t) e^{-rt} dt \tag{1}$$

where V_i is farmland value, π_i^* is the optimal net returns to land at location i, and r is the discount rate. Net returns are a function of the product of the land's output market price P_k and a vector of input prices w. They depend also on a set of exogenous variables, on different land uses denoted as k, and on $\mathbf{z}_i = (\mathbf{C}_i, \mathbf{G}_i, \mathbf{S}_i)$ that is composed of climate attributes $\mathbf{C}_i$, socio-economic characteristics $\mathbf{G}_i$, and soil quality indicators $\mathbf{S}_i$ (Hsiang, 2016):

$$\pi_i^* = max_{k \epsilon \{1,\ldots,K\}} \pi_{ik}(P_k, \mathbf{w}, \mathbf{z_i}) \tag{2}$$

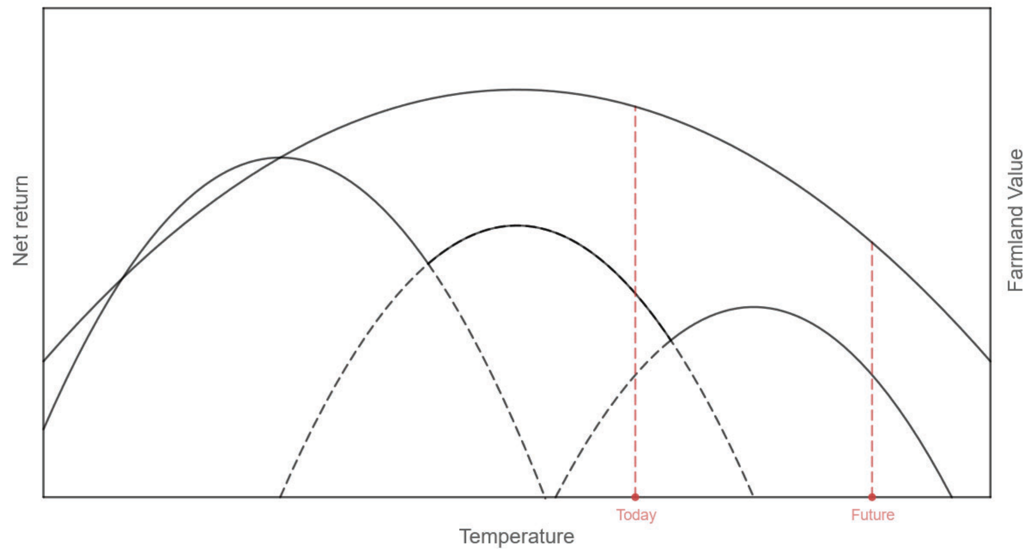
In this framework, the effect of climate change on farmland values can be conceptualized in panel (a) of Figure 1. Note that the x-axis focuses on changes in temperature, but a similar approach could be used for changes in precipitation. The three overlapping parabolas represent the envelope formed by three different types of land use reflecting that, as temperature increases, farmers would choose to switch their inputs and outputs to maximize the land's net return (y-axis left) for a given temperature level. The land value this process leads to (y-axis right) is reflected in the outer parabola for the entire range of temperature. Traditionally, Equation (1) is parametrized through the econometric model depicted in Equation (3) where $(\boldsymbol{\beta_1'}, \boldsymbol{\beta_2'}, \boldsymbol{G}, \boldsymbol{\theta})$ is a set of exogenous coefficients (Fezzi and Bateman, 2015).[3]

$$V_i = \alpha_0 + \boldsymbol{\beta_1}'\boldsymbol{C_i} + \boldsymbol{\beta_2}'\boldsymbol{C_i^2} + \boldsymbol{\gamma}'\boldsymbol{G_i} + \boldsymbol{\theta}'\boldsymbol{S_i} + \epsilon_i \tag{3}$$

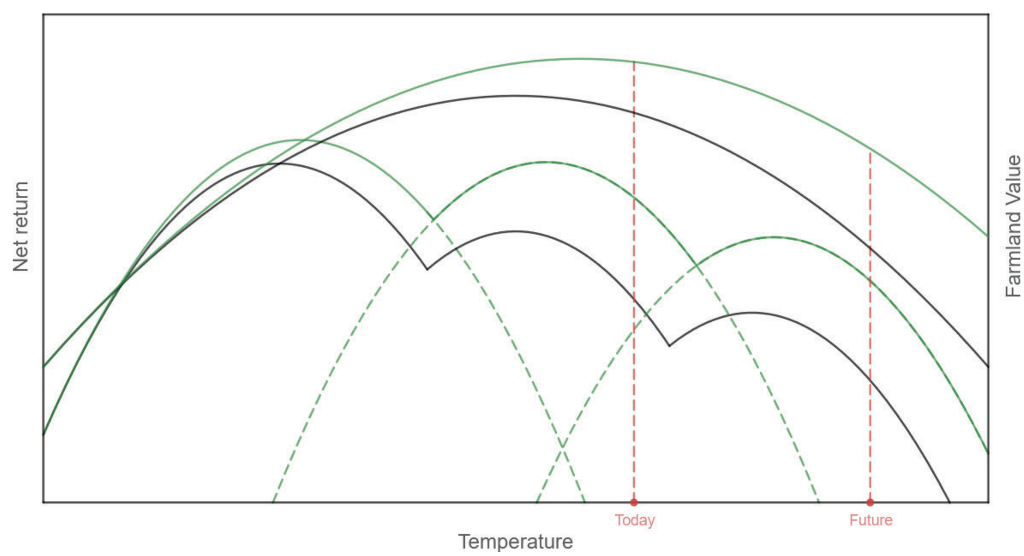
The average marginal effect of climate on farmland value, $\frac{\partial V_i}{\partial \boldsymbol{C_i}} = \boldsymbol{\beta_1}' + 2\boldsymbol{\beta_2}'\boldsymbol{C_i}$, is a function of $\mathbf{C}_i$. However, in this expression, the two beta coefficients are a-spatial. Their magnitude is assumed to be valid for any county in the sample, from Florida to Minnesota, so that the only source of spatially varying marginal effects lies in the value of the covariate $\mathbf{C}_i$. Several authors have challenged the assumption that the marginal effect of climate change is constant for all locations (Mendelsohn and Dinar, 2003; Schlenker et al., 2005; Timmins, 2006). Panel (b) of Figure 1 illustrates their rationale for the simplest case of two different responses. If the econometric model omits the different responses to climate change, future farmland predictions will fall between the green and the black outer parabolas.

The slope heterogeneity illustrated in panel (b) of Figure 1 can be treated by disaggregating the estimations to a finer spatial scale. The Ricardian literature offers different grouping strategies that consist of clustering counties that are thought to be subject to similar effects from climate change. Some of the earliest investigations in this direction find that climate change has little or no effect on irrigated farmlands, suggesting that in earlier studies the gains associated with climate change are underestimated (Mendelsohn

[3]Based on their approach, Mendelsohn et al. (1994) forecast that global warming may bring average gains to U.S. agriculture. Since then, their pioneering work has led to a surge of studies using a variation of the Ricardian approach, either in the U.S. or abroad (Mendelsohn and Massetti, 2017).



(a) Spatial homogeneous responses



(b) Spatial heterogeneous responses

Figure 1: Hedonic approach under different spatial assumptions.

Note: For each color, the three overlapping paraboles represent three different land uses, so the upper envelop maximizes net returns with respect to the different activities and temperature levels. The outer parabole represents the land value over the different temperature and illustrates the farmers' net return expectations. Panel (a) illustrates the case for homogenous responses to climate change, while panel (b) represents the case for two different responses to climate change. The assumption of homogenous responses when the data generating process exhibits heterogenous responses as in panel (b) would lead to future predictions falling between the two outer paraboles.

and Dinar, 2003). Once irrigation is accounted for, Schlenker et al. (2005) find that the projected climate impact estimates for rain-fed areas converge to a national annual loss of $9.62 billion. In their panel data estimation, Deschênes and Greenstone (2007) allow the climate coefficients to fully interact with irrigation and state fixed effects. Their results indicate that climate change will benefit farmers by $3.24 billion.[4] More recently, Dall'erba and Domínguez (2016) split the U.S. Southwestern counties into two groups based on the median value of elevation. Their results show that the probability of a negative future impact of climate change is larger in highland counties which are forecasted to experience heat waves more frequently than in the lowland counties.

When Cai and Dall'Erba (2021) investigate the effect of the above grouping specifications on the marginal impact of climate change on farmlands, they conclude that the predicted change in farmland values varies only slightly when either irrigation or elevation is used as the criterion for grouping spatial units. Contradictory findings about the direction and/or magnitude of the impact of climate change in U.S. agriculture arising from different grouping strategies of spatial units suggest that the clustering of spatial units in such studies has an effect on the analytical results – a problem often referred to as the modifiable areal unit problem (MAUP) (Fotheringham and Wong, 1991). This effect has been recognized in the context of studies on the impact of climate change on agricultural productivity by Deschênes and Greenstone (2007, 2012).

Given the unsuitability of a priori groupings of spatial units with which to demonstrate spatially heterogeneous effects of climate change on agricultural production, we take a different approach and rely on local regressions to relax the assumption that the data generating process is constant over space without having to specify any a priori regionalization of units. Several types of statistical models exist that generate locally varying responses to the same stimuli, such as the Bayesian spatially varying coefficients model (Gelfand et al., 2003; Fotheringham et al., 2017, 2021) and the frequentist multiscale geographically weighted regression approach (Yu et al., 2020a,b), although comparative

[4] Fisher et al. (2012) find and correct several irregularities in the paper of Deschênes and Greenstone (2007), hence concluding with a negative impact of climate change on U.S. agriculture across different model specifications.

studies suggest they produce similar results (Wolf et al., 2018).

We selected the GWR framework because of its ease-of-use, familiarity, scalability and the fact that it has never been applied to Ricardian studies.[5] Equation (4) depicts a spatially varying model where y_i is the dependent variable, and x_{ik} is the k^{th} explanatory variable. This model is known as the basic-GWR. The local coefficients are denoted as $\beta_0(u_i, v_i)$ for the intercept and $\beta_k(u_i, v_i)$ for the coefficients associated with the k^{th} explanatory variable. In turn, coefficients are a function of the i^{th} observation's latitude u_i and longitude v_i. The stochastic error is denoted by ϵ_i. To retrieve the local coefficients associated with Equation (4), GWR employs a variation of the Generalized Least Squares: $\boldsymbol{\beta} = [\boldsymbol{X'W}(u_i, v_i)\boldsymbol{X'}]^{-1}\boldsymbol{X'W}(u_i, v_i)\boldsymbol{Y}$, where the diagonal weighted matrix is denoted by $\mathbf{W}(u_i, v_i)$. Because its elements depend on the location of the i^{th} observation, each parameter is location-specific.

$$y_i = \beta_0(u_i, v_i) + \Sigma_k \beta_k(u_i, v_i) x_{ik} + \epsilon_i \tag{4}$$

One problem with basic-GWR is that it imposes the same bandwidth (i.e. the same degree of spatial heterogeneity) on each set of local parameter estimates. To remove this problem, Fotheringham et al. (2017) derive a generalization of the basic-GWR, known as Multiscale Geographically Weighted Regression (MGWR), that permits multiple local bandwidths. This model is depicted in Equation (5) where the superscript *bw* indicates that each set of local parameter may have a different bandwidth, distinguishing it from Equation (4).

$$y_i = \beta_0^{bw}(u_i, v_i) + \Sigma_k \beta_k^{bw}(u_i, v_i) x_{ik} + \epsilon_i \tag{5}$$

MGWR is estimated with a back-fitting algorithm that maximizes the expected log-likelihood of each term in an additive model: $\boldsymbol{y} = \boldsymbol{\Sigma_k f_k}$, where each $\boldsymbol{f_k} = \beta_k^{bw}(u_i, v_i) x_{ik}$

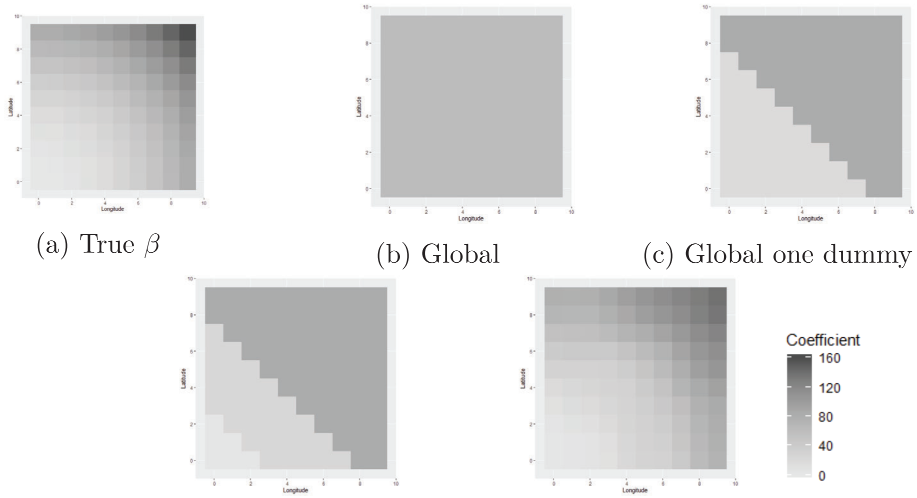
[5] In contrast to other GWR methods, the MGWR is more flexible in how spatial heterogeneity is modelled since it allows the calibration of coefficients to have different bandwidths depending on the variable i.e., each covariate is regressed across different subsets of observations. For example, if the nature of one variable is global and the rest are not, MGWR will assign a bandwith equal to the total of observations to the former covariate and a lower bandwidth to the latter one.

in Equation (5) and $\boldsymbol{y}$ is its dependent variable (Fotheringham et al., 2017). In each iteration, the back-fitting algorithm calculates the residual $\boldsymbol{\epsilon} = \boldsymbol{y} - \boldsymbol{\Sigma_k f_k}$ and regresses each term of the form $\boldsymbol{\epsilon} + \boldsymbol{f_k}$ against its corresponding data vector x_{ik} using the basic-GWR estimator described above until a pre-specified convergence criterion is reached.[6] Estimating the Ricardian model of Equation (3) with the capabilities of the MGWR estimator allows the estimation of the localized marginal effects of the climate conditions on agricultural production without pre-specifying any grouping of spatial units and without imposing any restriction on optimized covariate-specific bandwidths. In this framework, the global model in Equation (3) relating agricultural land value to a series of covariates is depicted in Equation (6).

$$V_i = \alpha_0^{bw}(u_i, v_i) + \boldsymbol{\beta_1^{bw}}(u_i, v_i)'\boldsymbol{C} + \boldsymbol{\beta_2^{bw}}(u_i, v_i)'\boldsymbol{C^2} + \boldsymbol{\gamma^{bw}}(u_i, v_i)'\boldsymbol{G} + \boldsymbol{\theta^{bw}}(u_i, v_i)'\boldsymbol{S} + \epsilon_i \quad (6)$$

To illustrate how local models improve upon Equation (3), panel (a) in Figure 2 shows the extreme case in which the coefficient β varies across all spatial units. Panels (b) – (d) employ different grouping techniques such that β is allowed to vary across clusters, and panel (e) is the basic-GWR estimator. As more interactions are included global specifications improve but they still present severe biases. The set of β values recovered with the GWR specification captures closely the true specification of β without imposing any dummy interaction. Panels (b) – (d) reflect the bias that ad-hoc specifications would have on local estimates while panel (e) illustrates our proposed correction. The key feature of GWR methods is therefore to allow the researcher to uncover the presence of local clusters and to include them in the estimation process without prior knowledge of the number of existing groups, group location and group membership.

[6]MGWR inference is discussed in the appendix A.



(a) True β (b) Global (c) Global one dummy

(d) Global Two dummies (e) GWR

Figure 2: Spatially varying data generating process across different estimators.

Note: True $\beta = 1 + (u_i + v_i)^2$. Panels (a) is the simulated spatial data set. Panels (b) - (d) employ different grouping techniques such that β is allowed to vary across clusters, and panel (e) is the basic-GWR estimator. As more interactions are included, global specifications improve, but they still present severe biases. The set of β values recovered with the GWR specification captures closely the true specification of β without imposing any dummy interaction.

3 Data

The Ricardian literature has three distinct traditions in terms of their identification strategy with each new development raising upon the alleged shortcomings of the other. First, issues have been raised about the cross sectional application of farmland prices in reference to endogeneity, and instead they exploit a panel of random variations in weather to identify their key parameters (Deschênes and Greenstone, 2007, 2012). A recent development in the literature considers the quality of estimated coefficients to account for long-run adaptation and instead exploits county-level climate trends via long-difference estimators (Hsiang, 2016; Burke and Emerick, 2016). We follow the canonical work of Mendelsohn et al. (1994) because our manuscript is interested in showcasing how local-regressions, our MGWR estimator, can offer empirical applications limited to a cross-section an alternative to account for the spatial non-stationarity of the impact of climate change. Following closely the canonical work of Mendelsohn et al. (1994), our model relates average farmland value per acre with seasonal climate data while controlling for demographic and soil quality characteristics as shown in Equation (6).

We initially considered calibrating our model with data from all 3,008 non-metropolitan contiguous U.S. counties, but we followed Schlenker et al. (2005) and Dall'erba and Domínguez (2016) by removing urban counties, 207 of them, defined as counties with more than 1,600 inhabitants per square mile. A further 34 counties were removed because of missing data for at least one of our variables. Our final sample is therefore composed of 2,687 U.S. counties. Table 1 contains basic summary statistics related to the dependent and the main explanatory variables employed in our study. Our variable of interest is the average county farmland value per acre. The U.S. Department of Agriculture prepares a census every five years in which it collects farmers' land prices. The two most recent censuses were in 2012 and 2017, so we will use the former for calibration purposes and the later to test the predictive accuracy of each model compared to observed values. Our historical climate variable are precipitation and temperature. Precipitation is measured in cubic inches and temperature is measured in degree Celsius. The climate data are collected from the Parameter-elevation Regression on Independent Slopes Model (PRISM) database (PRISM Climate Group, 2021). We process PRISM data to estimate a historical average for precipitation and temperature for the U.S. counties over the 1993-2012 and 1993-2017 periods. The Ricardian literature raises concerns about the non-concavity in the farmland value's responses to seasonal climate. Indeed, one should expect farmland value to increase with respect to climate at a decreasing rate as in Figure 1. This limitation prevents researchers from interpreting extremum as maximum and is the result of aggregating monthly climate into seasonal climate (Darwin, 1999). Disaggregating climate into monthly values, however, would cause serious multicollinearity issues, and dropping some arbitrary variable may cause an omitted variable bias (Mendelsohn and Massetti, 2017). Including all seasonal variables for precipitation and temperature is paramount for our analysis since we aim to capture the heterogeneity of growing seasons in the U.S.

Table 1 contains basic summary statistics related to our control variables. The county's population density per square mile is obtained from the 2010 U.S. Census Bureau of Statistics, and the county's per-capita income is obtained from the U.S. Bureau

Table 1: Descriptive statistics: Dependent variable and historical climate variables

Variable	Mean	s.d.	Max.	Min.
2012				
Farmland value per acre	$3,216	$2,029	$ 21,801	$ 192
Temperature				
Winter (Dec. - Feb.)	1.94	5.94	19.56	-11.99
Spring (Mar. - May.)	12.50	4.93	24.42	0.87
Summer (Jun. - Aug.)	23.50	3.14	32.61	13.45
Autumn (Sep. - Nov.)	13.33	4.31	25.15	2.56
Precipitation				
Winter (Dec. - Feb.)	2.79	1.87	17.29	0.06
Spring (Mar. - May.)	3.34	1.27	10.55	0.04
Summer (Jun. - Aug.)	3.649	1.44	9.83	0.01
Autumn (Sep. - Nov.)	2.86	1.28	10.77	0.08
2017				
Farmland value per acre	$3,165	$2,095	$ 28,569	$ 162
Temperature				
Winter (Dec. - Feb.)	2.23	6.03	19.65	-12.03
Spring (Mar. - May.)	12.88	4.90	24.74	1.21
Summer (Jun. - Aug.)	23.44	3.13	32.77	13.53
Autumn (Sep. - Nov.)	13.11	4.33	24.99	2.28
Precipitation				
Winter (Dec. - Feb.)	2.79	1.90	17.65	0.03
Spring (Mar. - May.)	3.47	1.29	10.96	0.00
Summer (Jun. - Aug.)	3.61	1.56	9.82	0.00
Autumn (Sep. - Nov.)	2.76	1.22	12.15	0.02
Observations				2,689

Note: Temperature is reported in degrees Celsius. Precipitation is reported in cubic millimeters per day.

of Economic Analysis of the same year. Controlling for socio-economic characteristics allows us to remove unwanted variation that can bias our forecast. Soil quality is a measure of a land's productivity, so we include information about salinity, flood frequency ratio, wetland, slope steepness, erosion index, sand contents, clay contents, permeability, and moisture capacity. Soil indicators are obtained from the U.S. Geological Survey. Because we do not have reliable data for our control variables in 2017 and those for 2012 are proxied with data from 2010, we follow Deschênes and Greenstone (2007, 2012) and assume they are constant. Thus, we can forecast changes in the value of farmland prices as a function of climate change (first-degree difference in our climate variables).

In the subsequent sections, we test the robustness of our model relative to previous specifications forecasting the impacts of climate change on agricultural land values,

Table 2: Descriptive statistics for the value of land and control variables

Variable	Mean	s.d.	Max.	Min.
Socioeconomic Variables				
Population density per square mile in 2010	73.22	99.79	1582.74	0.13
Per capita income in 2010	\$39,209	\$10,531	\$52,380	\$18,404
Soil Quality Variables				
Salinity	0.20	0.56	7.19	0.00
Flood frequency ratio	0.08	0.12	0.73	0.00
Wetland ratio	0.55	0.38	1.00	0.00
Slope steepness	11.46	11.75	68.41	0.54
Erosion index	0.27	0.08	0.49	0.00
Sand contents	38.37	19.68	96.33	0.98
Clay content	18.54	8.15	55.12	1.12
Permeability	20.46	18.68	113.63	0.91
Moisture capacity	0.16	0.04	0.38	0.01
Observations				2,687

namely: Mendelsohn et al. (1994) (no grouping of spatial units); Schlenker et al. (2005) (two-way division of spatial units based on irrigated ratio ≥ 0.2); and Dall'erba and Domínguez (2016) (two-way division based on median elevation). For each approach we employ the same 2012 dataset to predict the 2017 farmland values. This allows us to assess the predictive power of the local regression approach described here relative to other approaches for a year for which the data are observed. We then employ our updated data in 2017 to estimate the marginal effect of yearly and seasonal climate variables on farmland values to further study the capabilities of our estimator choice.

4 Results

4.1 Cross-Validation Results

Cai and Dall'Erba (2021) recognize the shortcomings associated with the wrongful grouping of U.S. counties in the Ricardian approach, so the authors perform cross-validation exercises across several groupings for the same year. Our manuscript expands their approach and calibrates all the models on 2012 data to predict the changes in farmland values by 2017. To our knowledge, no previous Ricardian study has used a cross-validation approach across years to demonstrate the better predictive power of one model over another. The results of our cross-validation exercise are depicted in Figure 3 where the observed 2012-2017 changes are displayed on a map and in a histogram and the 2012-2017 predicted changes based on the various models are also displayed.

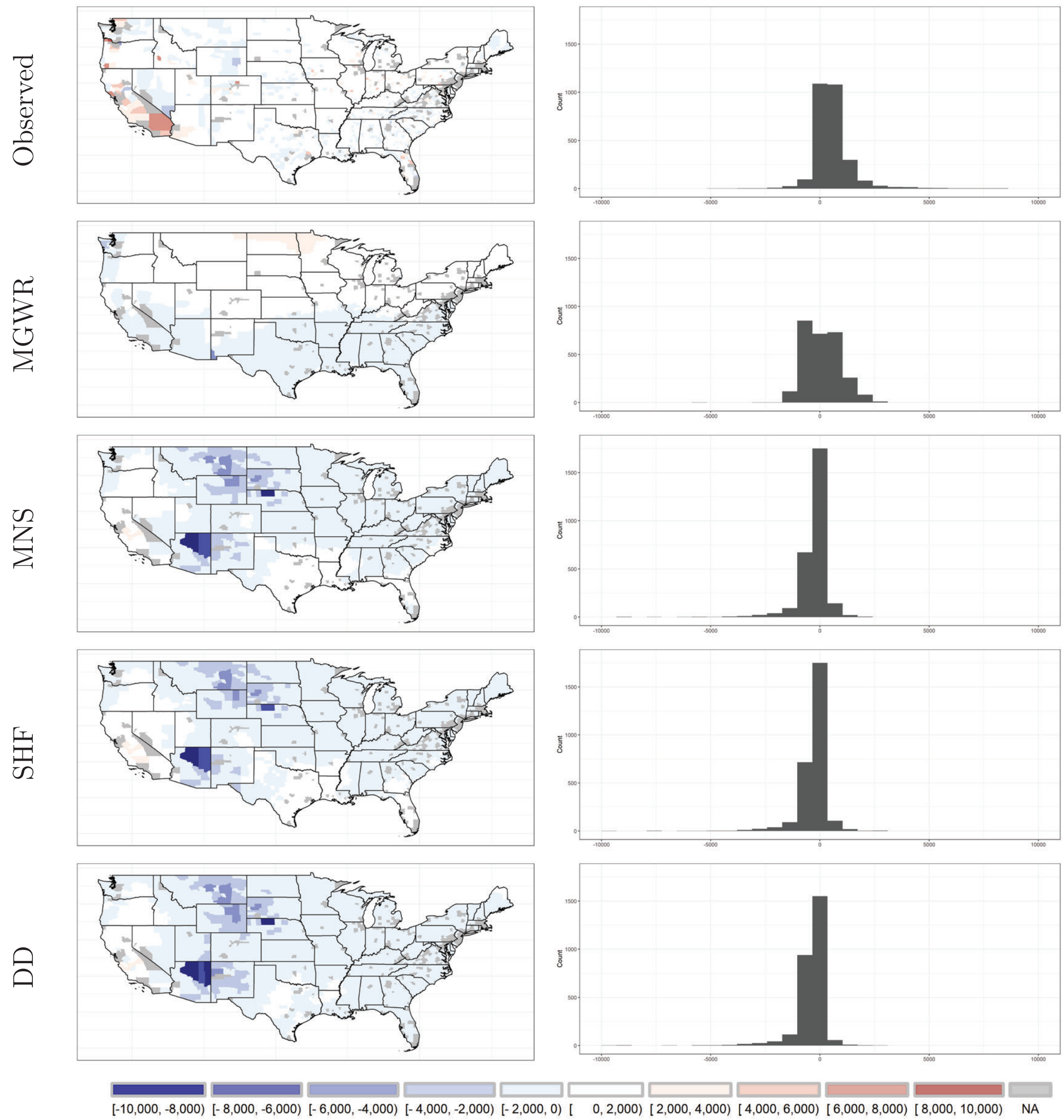
The table at the bottom of the figure reports certain prediction diagnostics. All the global regression models significantly underestimate the changes in farmland values. Many of the predictions from the global models are centered around zero, but with some high-valued cluster of losses, which is reflected in the average estimated county loss of -\$412.08 to -\$280.79 per acre. Only MGWR predicts gains between 2012-2017 with an average of \$109.86 per acre, which is still lower than the observed values in 2017 at \$573.92 per acre. Despite this difference, our local estimator significantly outperforms any of the global specifications as they each predict losses by 2017. The better predictive performance of MGWR is also evidenced by its low Mean Squared Error (MSE)[7] value compared to other models.[8] Further examination of Figure 3 indicates that the results from the local model reproduce a pattern similar to the isotherms formed when temperature is mapped across latitudes in the U.S. In contrast, none of the global specifications shows this pattern, but instead they highlight ill-defined clusters of high losses.[9]

Two additional elements can be drawn from these results. First, the local capabilities of the non-stationary Ricardian approach do not impose any kurtosis in the distribution of gains and losses. Figure 3 shows that the tails from the global predictions are thinner and longer than those from the local predictions. As a result, most global predictions fall around the mean, hence creating high peaks in their distribution. Because our MGWR Ricardian approach does not restrict the kurtosis of our predictions, it is able to uncover important unspecified heterogeneity. The second feature we note is that local specifications ignore outliers, such as the land appreciation in the Californian counties which is likely to be a result of the California Drought of 2012. This suggests that the local model

[7] $MSE = \frac{1}{n} \sum_N (V_i - \hat{V}_i)^2$, where $\hat{V}_i$ is the predicted value.

[8] Our approach also ignores some outliers along the West Coast which every global estimator predict would increase slightly in 2017, suggesting that our estimator exchanges complexity in predictions for generality in the model.

[9] The MGWR predictions to the middle of the century do not follow the same temperature/latitude approach, demonstrating that our MGWR approach incorporates unspecified clusters beyond weather.



	Observed	MGWR	MNS	SHF	DD
Average gain/loss	573.92	109.86	-280.79	-285.44	-412.08
Total gains/losses	1,588,041.72	303,994	-766,938	-789,816	-1,140,231
MSE	-	2,033,204	2,241,155	2,181,644	2,490,719
Min	-5,074	-5,848.89	-9,127.98	-9,812.79	-45,441.29
Max	22,353	2,547.22	3,087.78	3,087.11	75,197.74

Figure 3: Spatial and density distribution of observed and predicted values for 2017 across specification.

Note: MGWR is our local specification. MNS employs no grouping strategy following Mendelsohn et al. (1994). SHF splits counties by irrigation ratio (0.2) following Schlenker et al. (2005). DD splits the counties by median elevation following Dall'erba and Domínguez (2016).

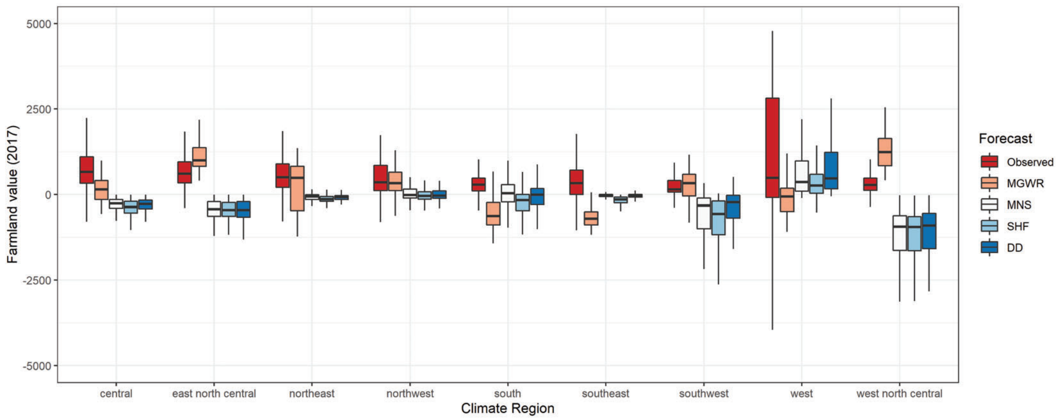


Figure 4: Observed and forecasted farmland values in 2017.
Note: Figure investigates the regional differences across estimators by dividing forecast by climate regions (Karl and Koss 1984). Outliers falling above $5,000 and below -$5,000 are trimmed for exposition.

underestimates the dependent variable in the locations that are impacted by extreme weather events but still does a much better job in capturing the impacts of these events than any global model.

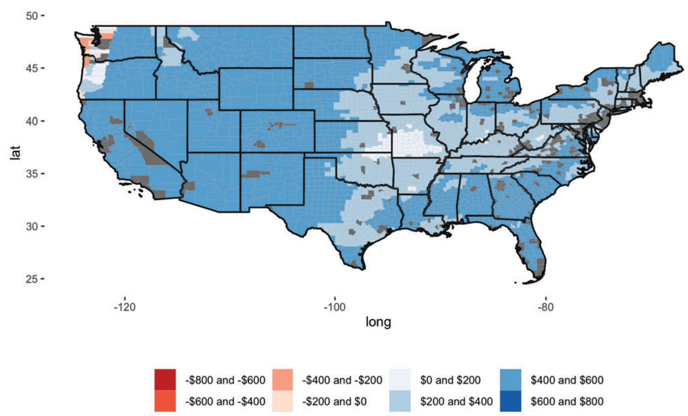
We investigate further the differences between our MGWR estimator and the global estimators in Figure 4 by dividing forecasts by climate regions (Karl and Koss, 1984). This figure shows that our local estimator outperforms all global estimators across the different climate regions except for the Southern and Southeastern regions. The underperformance of global estimators is clearly observed in the Southwest and in the Northwest and Central regions where all global estimators predict significantly lower values than the observed 2017 farmland values. MGWR predictions in the Northwest and Central regions surpassed the observed 2017 farmland values. The absolute value of the difference between our local and the global regressions favors our MGWR application. For the Southern and Southeastern regions, the median predictions from the global estimators are close to the median values of those regions with values above the average predictions, suggesting that the local estimator outperforms the global estimators.

We acknowledge that there are two caveats about our MGWR estimation that also affect the studies this manuscript examines. The functional form is assumed to be constant through time. While this is a necessary assumption to make predictions, adaptation strategies to mitigate the impact of climate change can take place, affecting the model's parameters. The second caveat is the role of omitted variable bias in our estimations. Studies of farmland prices are constrained by cross-sectional data, so practioners are unable to control for time-invariant influences. Despite these limitations, the nature of our study is not to develop the *best* functional form, but to propose an estimator to treat the spatial heterogeneity in the study of farmland prices. As this section demonstrates, our MGWR Ricardian approach incorporates the spatial heterogeneity present in our dataset. This having been said, the next section is concerned with the estimation of county-specific marginal effects of climate that can be employed to develop place-tailored mitigation stragies to climate change.

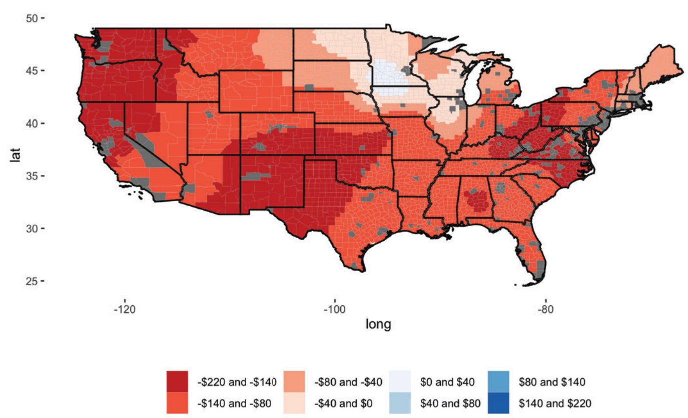
4.2 Place-tailored Weather Impacts on Farmland Prices

Figure 5 reports the MGWR parameter estimates of yearly temperature and precipitation obtained based on 2017 data and Equation (6): $\sum_{k\in\Phi}\frac{\partial V_i}{\partial \boldsymbol{C_i}} = \sum_{k\in\Phi}\{\boldsymbol{\beta_{1k}^{bw}}(u_i, v_i)' + 2\boldsymbol{\beta_{2k}^{bw}}(u_i, v_i)'\boldsymbol{C_{ik}}\}$, where $k \in \Phi$ corresponds to the seasonal weather variables (either temperature or precipitation). Panel (a) indicates the marginal effect of yearly precipitation on farmland values. Not surprisingly, most counties benefit from an increase in precipitation. In contrast, most counties will experience losses from rising temperature as shown in panel (b). Our findings indicate significant heterogeneity in the impact of climate change in the U.S. In both maps, the overall trend with respect to climate change seems to be either positive (for precipitation) or negative (for temperature), but the direction of the impact changes for a small subset of counties. For example, the impact of more precipitation is positive for most of the states except in the counties close to the coast in Oregon and Washington. Similarly, an increase in temperature damages most farmers in the U.S. except for some counties in the upper part of Midwest (South Minnesota and North Iowa). While all our estimated seasonal effects are statistically different from zero, we are unable to test if the yearly summation are statistically different from zero i.e., no GWR method exists to test the linear combination of parameters.

In Figure 6 and Figure 7, we dissagregate yearly impacts into seasonal impacts: $\frac{\partial V_i}{\partial \boldsymbol{C_i}} =$



(a) Marginal Effect: Yearly Precipitation



(b) Marginal Effect: Yearly Temperature

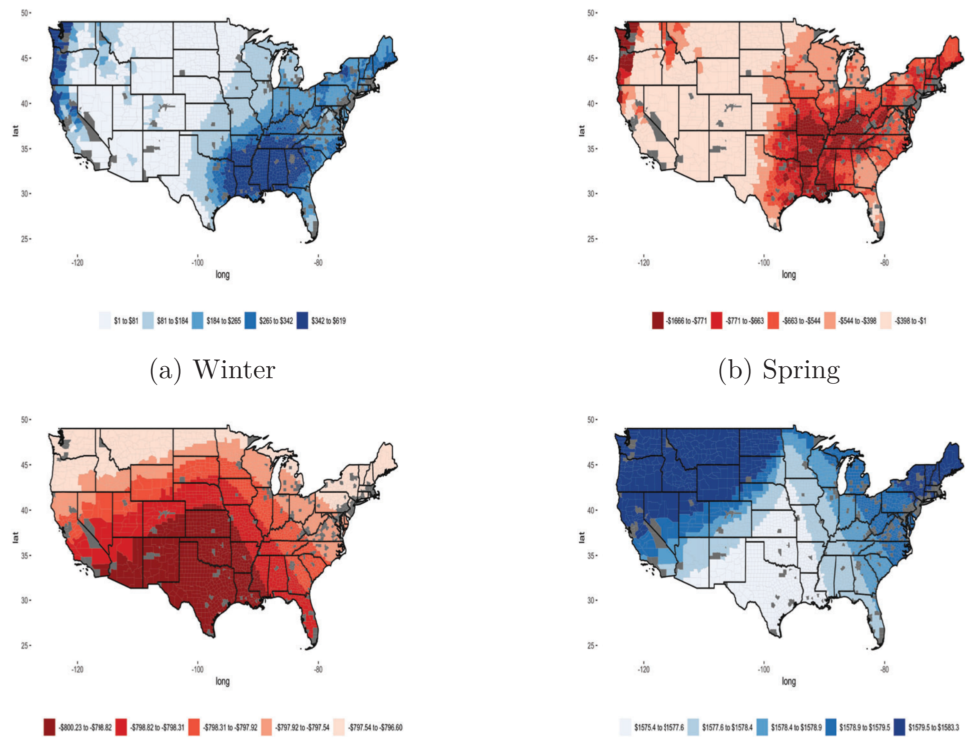
Figure 5: Marginal effect of yearly precipitation and temperature.

Note: Figure shows the yearly marginal effect of precipitation and temperature, obtained with equation 3: $\sum_{k\in\Phi}\frac{\delta V_i}{\delta C_i} = \sum_{k\in\Phi}\boldsymbol{\beta}^{bw}_{1,k}(u_i,v_i)' + 2\boldsymbol{\beta}^{bw}_{2,k}(u_i,v_i)'C,k$, where $k \in \Phi$ indicates the variables associated with the weather variable i.e., either temperature or precipitation.

$\boldsymbol{\beta_{1k}^{bw}}(u_i, v_i)' + 2\boldsymbol{\beta_{2k}^{bw}}(u_i, v_i)'\boldsymbol{C_{ik}}$.[10]

In reference to the previous heterogeneity observed in Figure 5, Figure 6 suggests that precipitation in Spring severly affects counties in the coastal areas of Washington and Oregon, forming an isolated cluster in the affected area. No seasonal temperature, however, seems to dominate the direction of the sign for the exceptional cluster of counties perceived in Figure 7. Additionally, seasonal impacts observed in Figure 6 and Figure 7 have different levels of variance. A strip of counties between Northern California and Washington benefit the most from Winter precipitation, highlighting their status of irrigated counties. Precipitation in the Spring, however, damages these counties. Notice too that a great deal of spatial heterogeneity is found within the group of the 100^{th} meridian, which some studies that split regions by such longitude ignore. All these marginal effects have been calculated and are available upon request.

[10] In a yet unpublished paper by Gammans et al. (2020), the authors estimate local marginal effects following the work of Deschênes and Greenstone (2007) i.e., exploiting variations in weather. Our work differs from theirs in which our identification strategy relies on cross-sectional variations in climate.

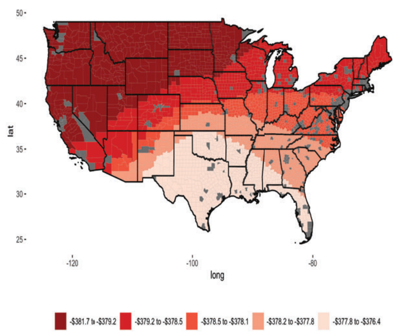


(a) Winter (b) Spring

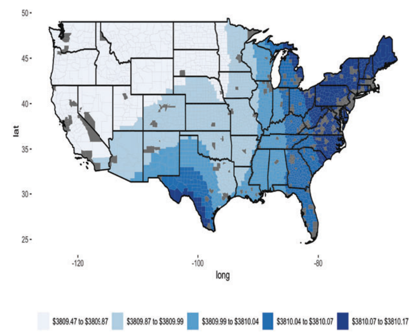
(c) Summer (d) Autumn

Figure 6: Seasonal precipitation marginal effect: $\frac{\delta V_i}{\delta C_i} = \boldsymbol{\beta}_{1,k}^{bw}(u_i, v_i)' + 2\boldsymbol{\beta}_{2,k}^{bw}(u_i, v_i)'C_{i,k}$. In contrast to yearly precipitation, the seasonal effect of weather does not show the same level of variation for some seasons (except for Winter and Spring). Thus, we describe their values by quantiles.

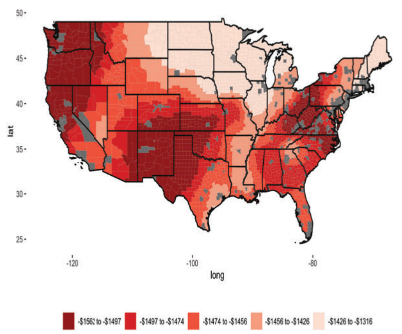
The Ricardian application of Dall'erba and Domínguez (2016) uncovers that the marginal impact of climate change is conditional on altitude. The results in Figure 7 show a cluster of counties displaying a large seasonal marginal effect in the Appalachian mountains, but only for Summer temperature, providing evidence that our MGWR Ricardian approach supports the conclusions derived in Dall'erba and Domínguez (2016). The cluster does not appear for precipitation, suggesting that the impact of precipitation is not altitude-dependent. In two separate studies, Schlenker et al. (2005) and Schlenker and Roberts (2006) claim that the 100^{th} meridian should serve as an appropriate split to cluster counties. The rationale is that counties East of the 100^{th} meridian are rainfed and as a result, they are more susceptible to precipitation variations. Panel (a) in Figure 6 and panel (a) in Figure 5 provide evidence to support a split by the 100^{th} meridian in terms of the impact of precipitation on farmland values. However, the marginal impact of temperature on farmland value does not display the same pattern. Indeed, panel (b) indicates the presence of substantial spatial heterogeneity in the marginal impact of temperature on both sides of the U.S. Furthermore, irrigated counties West of the 100^{th} meridian show more heterogeneity than their Eastern counterparts. Schlenker et al. (2007) consider this phenomenon in their own Ricardian study of agriculture. Their study suggests that water availability is capitalized into land prices and concludes that farmers benefit from rain after the growing season (Fall and Winter). Our results support their conclusions. Panel (a) in Figure 5 shows that most counties West of the 100th meridian benefit from precipitation, which replenishes aquifers. On the other hand, rain-fed counties east of the 100^{th} meridian show more heterogeneity, and even losses in some cases, suggesting that too much precipitation may lead to floods. As we previously indicated, researchers often split counties a priori by the 100^{th} meridian and include a dummy variable to generate two sets of slope coefficients. The results shown in Figure 5 support their rationale.



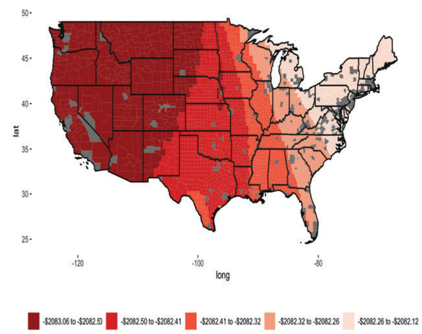
(a) Winter



(b) Spring



(c) Summer



(d) Autumn

Figure 7: Seasonal temperature marginal effect: $\frac{\delta V_i}{\delta C_i} = \boldsymbol{\beta}_{1,k}^{bw}(u_i, v_i)' + 2\boldsymbol{\beta}_{2,k}^{bw}(u_i, v_i)'C_{i,k}$. In contrast to yearly temperature, the seasonal effect of weather does not show the same level of variation for some seasons (except for Summer). Thus, we describe their values by quantiles.

5 Conclusion

Economic estimates of the agricultural damages from climate change have largely relied on the Ricardian approach and on econometric techniques that either disregard or reduce spatial heterogeneity to a small number of groups, often two. Given the size of the U.S. territory and the large variety of local climate conditions it offers, we believe that a much greater degree of spatial heterogeneity is present in the impacts climate change can have on the U.S. agriculture. This paper explores this question by relying on the recently developed MGWR estimator. Compared to global frameworks, this modeling approach allows us to uncover spatial heterogeneity in the effects of climate change on agricultural production without having to set a priori the number of clusters, their location and their membership; hence allowing us to generate place-specific marginal effects relying on a purely data-driven approach. Compared with global models, the MGWR forecasts exhibit a better capacity for predicting the large spatial variations in climate impact on farmland values and for accurately predicting future farmland values, as examplified in our cross-validation exercise for the year 2017. By generating place-specific estimates, our approach offers policy makers and stakeholders a tool to address the effects of climate change on a case-by-case basis in each county and suggest mitigation and adaptation strategies that are place-tailored.

Despite the novelty and strengths of our approach, we acknowledge that our approach does not include any form of interregional dependence that may play a large role in climate change mitigation. For instance, a recent contribution by Dall'Erba et al. (2021) estimates that domestic trade is able to mitigate the negative impact of climate change on crop profit by a magnitude of $14.5 billion. Fan et al. (2018) arrive to similar conclusions but argue that it is not only trade but also migration that are mitigation tools against climate change. As such, further research shall not only offer county-level predictions, but it shall also account for domestic trade which has been largely ignored in the literature, interregional linkages based on domestic trade, supply-chain and migration such as in the case of the Dust Bowl period.

References

Burke, M. and Emerick, K. (2016). Adaptation to climate change: Evidence from us agriculture. *American Economic Journal: Economic Policy*, 8(3):106–40.

Cai, C. and Dall'Erba, S. (2021). On the evaluation of heterogeneous climate change impacts on us agriculture: does group membership matter? *Climatic Change*, 167(14).

Cai, R., Yu, D., and Oppenheimer, M. (2014). Estimating the spatially varying responses of corn yields to weather variations using geographically weighted panel regression. *Journal of Agricultural and Resource Economics*, pages 230–252.

Dall'Erba, S., Chen, Z., and Nava, N. J. (2021). U.S. interstate trade will mitigate the negative impact of climate change on crop profit. *American Journal of Agricultural Economics*, 103(5):1720–1741.

Dall'erba, S. and Domínguez, F. (2016). The impact of climate change on agriculture in the Southwestern United States: The Ricardian approach revisited. *Spatial Economic Analysis*, 11(1):46–66.

Darwin, R. (1999). The impact of global warming on agriculture: A Ricardian analysis: Comment. *American Economic Review*, 89(4):1049–1052.

Deschênes, O. and Greenstone, M. (2007). The economic impacts of climate change: Evidence from agricultural output and random fluctuations in weather. *American Economic Review*, 97(1):354–385.

Deschênes, O. and Greenstone, M. (2012). The economic impacts of climate change: Evidence from agricultural output and random fluctuations in weather: Reply. *American Economic Review*, 102(7):3761–73.

Fan, Q., Fisher-Vanden, K., and Klaiber, H. A. (2018). Climate change, migration, and regional economic impacts in the United States. *Journal of the Association of Environmental and Resource Economists*, 5(3):643–671.

Fezzi, C. and Bateman, I. (2015). The impact of climate change on agriculture: Nonlinear effects and aggregation bias in Ricardian models of farmland values. *Journal of the Association of Environmental and Resource Economists*, 2(1):57–92.

Fisher, A. C., Hanemann, W. M., Roberts, M. J., and Schlenker, W. (2012). The economic impacts of climate change: Evidence from agricultural output and random fluctuations in weather: Comment. *American Economic Review*, 102(7):3749–60.

Fotheringham, A. S., Li, Z., and Wolf, L. J. (2021). Scale, context, and heterogeneity: A spatial analytical perpective on the 2016 US presidential election. *Annals of the American Association of Geographers*, pages 1–20.

Fotheringham, A. S. and Wong, D. W. (1991). The modifiable areal unit problem in multivariate statistical analysis. *Environment and Planning A*, 23(7):1025–1044.

Fotheringham, A. S., Yang, W., and Kang, W. (2017). Multiscale geographically weighted regression (mgwr). *Annals of the American Association of Geographers*, 107(6):1247–1265.

Gammans, M., Mérel, P., and Paroissien, E. (2020). Reckoning climate change damages along an envelope.

Gbetibouo, G. A. and Hassan, R. M. (2005). Measuring the economic impact of climate change on major South African field crops: A Ricardian approach. *Global and Planetary Change*, 47(2-4):143–152.

Gelfand, A. E., Kim, H.-J., Sirmans, C., and Banerjee, S. (2003). Spatial modeling with spatially varying coefficient processes. *Journal of the American Statistical Association*, 98(462):387–396.

Hsiang, S. (2016). Climate Econometrics. *Annual Review of Resource Economics*, 8:43–75.

Karl, T. and Koss, W. J. (1984). Regional and national monthly, seasonal, and annual temperature weighted by area, 1895-1983.

Liu, H., Li, X., Fischer, G., and Sun, L. (2004). Study on the impacts of climate change on China's agriculture. *Climatic Change*, 65(1):125–148.

Mendelsohn, R. and Dinar, A. (2003). Climate, water, and agriculture. *Land Economics*, 79(3):328–341.

Mendelsohn, R., Nordhaus, W. D., and Shaw, D. (1994). The impact of global warming on agriculture: A Ricardian analysis. *American Economic Review*, pages 753–771.

Mendelsohn, R. O. and Massetti, E. (2017). The use of cross-sectional analysis to measure climate impacts on agriculture: Theory and evidence. *Review of Environmental Economics and Policy*, 11(2):280–298.

Plantinga, A. J., Lubowski, R. N., and Stavins, R. N. (2002). The effects of potential land development on agricultural land prices. *Journal of Urban Economics*, 52(3):561–581.

PRISM Climate Group, O. S. U. (2021). Prism Gridded Climate Data. data retrieved from PRISM website, http://prism.oregonstate.edu,.

Schlenker, W., Hanemann, W. M., and Fisher, A. C. (2005). Will U.S. agriculture really benefit from global warming? Accounting for irrigation in the hedonic approach. *American Economic Review*, 95(1):395–406.

Schlenker, W., Hanemann, W. M., and Fisher, A. C. (2007). Water availability, degree days, and the potential impact of climate change on irrigated agriculture in California. *Climatic Change*, 81(1):19–38.

Schlenker, W. and Roberts, M. J. (2006). Nonlinear effects of weather on corn yields. *Review of Agricultural Economics*, 28(3):391–398.

Shiu, Y.-S. and Chuang, Y.-C. (2019). Yield estimation of paddy rice based on satellite imagery: Comparison of global and local regression models. *Remote Sensing*, 11(2):111.

Timmins, C. (2006). Endogenous land use and the Ricardian valuation of climate change. *Environmental and Resource Economics*, 33(1):119–142.

Wang, J., Mendelsohn, R., Dinar, A., Huang, J., Rozelle, S., and Zhang, L. (2009). The impact of climate change on China's agriculture. *Agricultural Economics*, 40(3):323–337.

Wolf, L. J., Oshan, T. M., and Fotheringham, A. S. (2018). Single and multiscale models of process spatial heterogeneity. *Geographical Analysis*, 50(3):223–246.

Yu, H., Fotheringham, A. S., Li, Z., Oshan, T., Kang, W., and Wolf, L. J. (2020a). Inference in multiscale geographically weighted regression. *Geographical Analysis*, 52(1):87–106.

Yu, H., Fotheringham, A. S., Li, Z., Oshan, T., and Wolf, L. J. (2020b). On the measurement of bias in geographically weighted regression models. *Spatial Statistics*, 38:100453.

Appendices

A Inference in the MGWR

Contrary to the basic-GWR, MGWR is defined as a Generalized Additive Model (GAM) whose values are estimated by a back-fitting algorithm (Fotheringham et al., 2017). In this case inference is a limitation since there is no single hat matrix mapping the estimated values $\hat{y}_i$ onto y_i. Yu et al. (2020a) reframe the basic-GWR as a GAM, and demonstrate that the hat matrix for a MGWR model can be derived within the back-fitting estimation algorithm. We display their derivation, by first showing that the hat matrix for a basic-GWR can be written as an additive term. Then, we demonstrate how a hat matrix can be estimated and used for the computation of local standard errors.

The hat matrix for the basic-GWR $\boldsymbol{S}$, is defined in Equation (7) with each row given by $\boldsymbol{s_i} = \boldsymbol{X_i}(\boldsymbol{X'W}(u_i, v_i)\boldsymbol{X})^{-1}\boldsymbol{X'W}(u_i, v_i)$. Here, $\boldsymbol{X_i}$ is the i^{th} row in the matrix of all predictors $\boldsymbol{X}$.

$$\boldsymbol{S} = \begin{pmatrix} \boldsymbol{X_1}(\boldsymbol{X'W}(u_1, v_1)\boldsymbol{X})^{-1}\boldsymbol{X'W}(u_1, v_1) \\ \dots \\ \boldsymbol{X_n}(\boldsymbol{X'W}(u_n, v_n)\boldsymbol{X})^{-1}\boldsymbol{X'W}(u_n, v_n) \end{pmatrix} \tag{7}$$

Let us define the additive hat matrices by $\boldsymbol{R_k}$ for the set of local parameters $\beta_{\mathbf{k}}$ associated with the kth explanatory variable, such that it has the following two properties regarding the dependent variable $\boldsymbol{y}$:

$$\hat{\boldsymbol{f}}_{\boldsymbol{k}} = \hat{\boldsymbol{R}}_{\boldsymbol{k}} y \tag{8}$$

$$\hat{\boldsymbol{y}} = \boldsymbol{S}\boldsymbol{y} = \Sigma_k \boldsymbol{R_k}\boldsymbol{y} \tag{9}$$

Equation (8) states that $\boldsymbol{R_k}$ projects $\boldsymbol{y}$ onto each fitted additive term $\hat{\boldsymbol{f}}_{\boldsymbol{k}}$, and Equation (9) implies $\Sigma_k \boldsymbol{R_k} = \boldsymbol{S}$. Next, we express each fitted additive term. Each comes from within each iteration of the back-fitting algorithm described in the main text, as a column vector:

$$\hat{\boldsymbol{f}}_{\boldsymbol{k}} = \begin{pmatrix} x_{1k}\hat{\beta}_{1k} \\ \boldsymbol{...} \\ x_{nk}\hat{\beta}_{nk} \end{pmatrix} \tag{10}$$

Each estimated local parameter can be written as $\hat{\beta}_{ik} = \boldsymbol{e_k}\hat{\boldsymbol{\beta}}_{\boldsymbol{i}}$, where $\boldsymbol{e_k}$ is the k^{th} row of an identity matrix whose dimension is the number of regressors plus 1. Thus, we can re-write $\hat{\beta}_{ik}$ as $\hat{\beta}_{ik} = \boldsymbol{e_k(X'W}(u_i, v_i)\boldsymbol{X})^{\boldsymbol{-1}}\boldsymbol{X'W}(u_i, v_i)\boldsymbol{y}$, and Equation (10) as:

$$\hat{\boldsymbol{f}}_{\boldsymbol{k}} = \begin{pmatrix} x_{1k}\boldsymbol{e_k(X'W}(u_i, v_i)\boldsymbol{X})^{\boldsymbol{-1}}\boldsymbol{X'W}(u_i, v_i) \\ \boldsymbol{...} \\ x_{nk}\boldsymbol{e_k(X'W}(u_i, v_i)\boldsymbol{X})^{\boldsymbol{-1}}\boldsymbol{X'W}(u_i, v_i) \end{pmatrix}\boldsymbol{y} \tag{11}$$

Therefore, if the $\boldsymbol{R_k}$ is a hat matrix that projects $\boldsymbol{y}$ onto each fitted additive term $\hat{\boldsymbol{f}}_{\boldsymbol{k}}$ (Equation (8)), Equation (11) allows us to write each additive hat matrix as:

$$\hat{\boldsymbol{R}}_{\boldsymbol{k}} = \begin{pmatrix} x_{1k}\boldsymbol{e_k(X'W}(u_i, v_i)\boldsymbol{X})^{\boldsymbol{-1}}\boldsymbol{X'W}(u_i, v_i) \\ \boldsymbol{...} \\ x_{nk}\boldsymbol{e_k(X'W}(u_i, v_i)\boldsymbol{X})^{\boldsymbol{-1}}\boldsymbol{X'W}(u_i, v_i) \end{pmatrix} \tag{12}$$

Now, we show how the additive hat matrix, and the MGWR hat matrix, can be estimated within the back-fitting algorithm. First, let $\boldsymbol{A_j}$ be the hat matrix of the partial model such that $\hat{\boldsymbol{f}}^*_{\boldsymbol{k}} = \boldsymbol{A_k}(\hat{\boldsymbol{f}}_{\boldsymbol{k}} + \hat{\boldsymbol{\epsilon}})$. Within the back-fitting algorithm, $\hat{\boldsymbol{f}}^*_{\boldsymbol{k}}$ is the updated fitted additive term from the previous iteration fitted additive term $\hat{\boldsymbol{f}}_{\boldsymbol{k}}$. In each iteration, the back-fitting algorithm calculates the fitted residual $\hat{\boldsymbol{\epsilon}} \boldsymbol{= y - \Sigma_k}\hat{\boldsymbol{f}}_{\boldsymbol{k}}$; but Equation (9) implies that $\boldsymbol{S} = \Sigma_k\boldsymbol{R_k}$, so we can write the following equality:

$$\hat{\boldsymbol{f}}^*_{\boldsymbol{k}} = \boldsymbol{A_k}(\hat{\boldsymbol{f}}_{\boldsymbol{k}} + \hat{\boldsymbol{\epsilon}}) = \boldsymbol{A_k}(\boldsymbol{R_k y} + \boldsymbol{y} - \boldsymbol{S y}) = (\boldsymbol{A_k R_k} + \boldsymbol{A_k} - \boldsymbol{A_k S})\boldsymbol{y} \tag{13}$$

The updated additive hat matrix is $\boldsymbol{R}^*_{\boldsymbol{k}} = \boldsymbol{A R_k} + \boldsymbol{A_k} - \boldsymbol{A_k S}$ and the updated MGWR hat matrix is $\boldsymbol{S}^* = \boldsymbol{S} - \boldsymbol{R_k} + \boldsymbol{R}^*_{\boldsymbol{k}}$. With $\boldsymbol{R_k}$ and $\boldsymbol{S}$, local standard errors can be computed. First, let us re-write Equation (10) as $\hat{\boldsymbol{f}}_{\boldsymbol{k}} = diag(\hat{X}_k)\hat{\boldsymbol{\beta}}_{\boldsymbol{k}}$, where $diag(\hat{\boldsymbol{X}}_{\boldsymbol{k}})$

is shown in Equation (14). Then, we can re-write $\hat{\boldsymbol{\beta}}_{\boldsymbol{k}} = [diag(\boldsymbol{X}_{\boldsymbol{k}})]^{-1}\boldsymbol{R}_{\boldsymbol{k}}\boldsymbol{y} = \boldsymbol{C}\boldsymbol{y}$, where $\boldsymbol{C} = [diag(\boldsymbol{X}_{\boldsymbol{k}})]^{-1}\boldsymbol{R}_{\boldsymbol{k}}$.

$$diag(\boldsymbol{X}_{\boldsymbol{k}}) = \begin{pmatrix} x_{1k} & 0 & 0 & 0 \\ 0 & x_{2k} & 0 & 0 \\ 0 & 0 & ... & 0 \\ 0 & 0 & 0 & x_{nk} \end{pmatrix} \tag{14}$$

Finally, we can write the variances of the local parameters $\hat{\boldsymbol{\beta}}_{\boldsymbol{k}}$ as in Equation (15), where the normalized residual sum of squares from MGWR is defined by Equation (16) and the trace of $\boldsymbol{S}$ is defined by $\tau_1 = trace(\boldsymbol{S})$. Therefore, we can compute local standard errors for each parameter with: $SE(\hat{\boldsymbol{\beta}}_{\boldsymbol{k}}) = \sqrt{var(\hat{\boldsymbol{\beta}}_{\boldsymbol{k}})}$.

$$var(\hat{\boldsymbol{\beta}}_{\boldsymbol{k}}) = diag(\boldsymbol{C}\boldsymbol{C'}\hat{\boldsymbol{\sigma}}^{\boldsymbol{2}}) \tag{15}$$

$$\hat{\sigma}^2 = \frac{\Sigma_N (y_i - \hat{y}_i)^2}{n - \tau_1} \tag{16}$$

The local standard errors for all local parameters estimated with the MGWR are defined by $SE(\hat{\boldsymbol{\beta}}) = [SE(\hat{\boldsymbol{\beta}}_{\boldsymbol{1}}, SE(\hat{\boldsymbol{\beta}}_{\boldsymbol{2}}), ..., SE(\hat{\boldsymbol{\beta}}_{\boldsymbol{k}},)]$, and the pseudo-t test for each local parameter is given by:

$$\frac{\hat{\beta_{ik}}}{SE(\hat{\beta_{ik}})} \sim t_{n-\tau_1} \tag{17}$$