Essays in yield modeling and basis price forecasting

by

Noah James Miller

B.A., University of Oklahoma, 2007
M.Sc., London School of Economics and Political Science, 2011
M.S. Kansas State University, 2015

AN ABSTRACT OF A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Agricultural Economics
College of Agriculture

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2020

Abstract

Concerns over the sustainability of small-holder commercial agriculture in the face of extended periods of extreme weather have increased across much of the country in the past two decades. A growing body of research has linked interannual variation in growing season weather to regional and national yields and prices, however relatively less work exists delineating the impacts of weather on disaggregated farm-level yield and elevator-level price outcomes. Here I make use of fine-scaled weather datasets tied to historic farm yields and elevator harvest basis levels to explicitly model these weather impacts using panel fixed-effects regressions. Results suggest that sorghum yields are quite sensitive to warming temperatures - as moderate increases of 2°C in growing season temperatures lead to an average 24% reduction in yields – thereby raising doubts about its potential for offsetting climate change impacts relative to other crops. We also consider whether warming impacts can be lessened through adjustment of the growing season (i.e. shifting up, shortening, or extending the growing season) and find very little support for this form of adaptation. Inclusion of soil moisture data into base price forecasts indicate that improvement in forecasts can be made by inclusion of within-season weather data – as inclusion of weekly cubic precipitation improves naïve harvest basis forecasts for Kansas grain elevators by 14%. This result improves to upward of 36% when the same model is applied to the other states in the Southern Great Plains.

Essays in yield modeling and basis price forecasting

by

Noah James Miller

B.A., University of Oklahoma, 2007
M.Sc., London School of Economics and Political Science, 2011
M.S. Kansas State University, 2015

A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Agricultural Economics
College of Agriculture

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2020

Approved by:

Co-Major Professor
Dr. Jason Bergtold

Approved by:

Co-Major Professor
Dr. Jesse Tack

Copyright

Abstract

Concerns over the sustainability of small-holder commercial agriculture in the face of extended periods of extreme weather have increased across much of the country in the past two decades. A growing body of research has linked interannual variation in growing season weather to regional and national yields and prices, however relatively less work exists delineating the impacts of weather on disaggregated farm-level yield and elevator-level price outcomes. Here I make use of fine-scaled weather datasets tied to historic farm yields and elevator harvest basis levels to explicitly model these weather impacts using panel fixed-effects regressions. Results suggest that sorghum yields are quite sensitive to warming temperatures - as moderate increases of 2°C in growing season temperatures lead to an average 24% reduction in yields – thereby raising doubts about its potential for offsetting climate change impacts relative to other crops. We also consider whether warming impacts can be lessened through adjustment of the growing season (i.e. shifting up, shortening, or extending the growing season) and find very little support for this form of adaptation. Inclusion of soil moisture data into base price forecasts indicate that improvement in forecasts can be made by inclusion of within-season weather data – as inclusion of weekly cubic precipitation improves naïve harvest basis forecasts for Kansas grain elevators by 14%. This result improves to upward of 36% when the same model is applied to the other states in the Southern Great Plains.

Table of Contents

List of Figures

List of Tables

Acknowledgements

This dissertation represents the culmination of my work as a PhD student. The path to completing it was not always straightforward or convenient and as I walked along the path I encountered a number of obstacles (several of them self-created) and setbacks. Fortunately, the chairs of my committee, Dr. Jason Bergtold and Dr. Jesse Tack, served as excellent mentors and guides. Whatever is of merit in this dissertation is largely a product of their thoughtful and thorough criticism of my work. I am deeply indebted to them. I am also grateful to the other members of my PhD committee - Dr. Mykel Taylor, Dr. Ignacio Ciampitti, and Dr. Allen Featherstone - who were extremely generous with their time and always willing to share their thoughts on the research discussed herein. I also thank my parents, Forrest and Helen, and my sister, Elizabeth who were instrumental in providing encouraging words and stability in what seemed to be a very turbulent period of my life - but now in the age of coronavirus seems idyllic.

Written below are the words I turned to for strength from time to time throughout the PhD, whenever life seemed inhospitable and bleak:

"Let this be our endeavor, then: to master self, to grow stronger each day, to advance in virtue."

– Thomas à Kempis

Preface

This work is supported by AFRI Education and Workforce Development grant no. 2019-67011-29508 from the USDA National Institute of Food and Agriculture. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the author and do not necessarily reflect the view of the U.S. Department of Agriculture.

Chapter 1 - The Impacts of Warming Temperatures on US Sorghum Yields and the Potential for Adaptation

Introduction

The findings of the 2014 Intergovernmental Panel on Climate Change (IPCC) state that even with implementation of stringent mitigation policies, average global surface temperatures are projected to increase beyond 1.5°C by the end of the century (Pachauri et al. 2014). Though region-specific anomalies are expected, increased mean temperatures will very likely be accompanied by more frequent occurrences of extreme heat and heat waves of longer intensity and duration (Pachauri et al. 2014). Research linking a large portion of historic variation in crop yields to changes in temperature suggests that the impacts of projected climate outcomes on agricultural production, in the absence of adaptation, will be profound (Lobell and Field 2007; Lobell, Schlenker, and Costa-Roberts 2011). Individual and meta-analysis studies of corn, soybean, wheat, and rice production show that warming will produce overall aggregate negative effects on yields for all four crops, and that these estimates are robust to alterations in model specification (Miao et al. 2016; Tack et al. 2015; Challinor et al. 2014; Lobell and Asseng 2017).

The effects of warming on sorghum however, are less well understood. Sorghum is known for its tolerance to hot, arid climates, which has allowed it to be grown on marginal lands otherwise ill-suited for crop production. This along with a protein content comparable to corn has made it the leading grain produced for human consumption and animal feed in the Sahel region of Africa, and is central to many food security programs throughout the developing world (Belton and Taylor, 2004). US sorghum production is mainly used as an input for biofuel and livestock production much like corn, making these crops logical on-farm substitutes. It is also gluten-free and is not genetically modified, which makes it an attractive alternative to wheat and corn for many food products. The majority of US production is located in the Southern Plains (United Sorghum Checkoff Program 2016), a region known for its harsh weather conditions, and it has been speculated that a potential adaptation strategy for corn farmers facing climate change is to switch from corn to sorghum production (Burke and Lobell 2010; Brown and Funk 2008). This argument assumes, among other things, that sorghum will not suffer similar negative yield impacts as corn; however, recent evidence based on field-trial data suggests that sorghum also faces yield losses due to extreme heat (Tack et al. 2017). There is therefore a need to quantify the effects of projected climate change on sorghum yields at the farm level.

To address this gap in the literature, we harness a unique farm-level panel dataset (Kansas Farm Management Association (KFMA) Dataset) on sorghum yields to delineate the relationship of weather on yields. The dataset contains 45,971 observations and is spread spatially across Kansas and temporally over 38 years. We estimate change in yield as a potentially nonlinear function of growing degree days, precipitation, trend, and farm fixed effects. Our findings show that temperature exposure above 33°C leads to yield damages. We estimate the effects of warming temperatures by 1°C increments and find an average 11% reduction in yield under a 1°C warming scenario to an average 67% reduction in yield under a 5°C warming scenario, with a large spatial heterogeneity of warming impacts. We consider the effectiveness that adjustment of the growing season can have in offsetting the effects of warming, by moving up growing season plant and harvest dates by one-day increments, while keeping the length of the growing season fixed at historical levels. Our findings show that under a 2°C warming scenario, a modest reduction in yield damages can be realized by planting 9 – 17 days earlier. We find similar results for other potential adjustments to planting and harvest dates that encapsulate changes in growing season length, in addition to cross-sectional regression that permits a wider scope for adaptation beyond growing season adjustments. Finally, we examine to what degree sorghum's ability to tolerate extreme heat (defined as temperatures in excess of 33°C) has changed over the study time period by re-estimating our initial model with the addition of an extreme heat-trend interaction term. We find that sorghum's ability to tolerate extreme heat has decreased by over 50%.

A growing body of literature uses econometric models that allow for nonlinear effects of temperature to examine the sensitivity of crop yields to extreme heat (Schlenker and Roberts 2009; Lobell et al. 2011; Roberts et al. 2012). These studies broadly argue that there exists a temperature threshold beyond which increased exposure will result in yield losses. Though the magnitude of these losses varies across crops and locations, the results of these studies highlight the need for adaptation strategies. This is consistent with the conclusions reached by the Fourth National Climate Assessment, which projected that in the absence of adaptation, increases in Midwest growing season temperatures would be the leading contributing factor to future declines in overall US agricultural productivity (Jay et al. 2018). Producer-driven strategies, such as switching seed varieties or crop mix have been found to reduce yield losses resulting from climate change (Burke and Lobell 2010; McCarl et al. 2016). Adjusting the growing season is another strategy that has the potential to mitigate yield losses (Ortiz-Bobea and Just 2012; Kawasaki and Uchida 2016). For example, Ortiz-Bobea and Just (2012) found that adjusting the plant date for corn in the Midwest would result in a reduction of 30 – 70% of yield losses caused by climate change. Kawasaki (2018) similarly found that adjusting the rice planting date in Japan would lead to large reductions in yield and revenue losses. In both studies, adjustment of the growing season came only through altering the plant date, with the harvest date held fixed. In this respect, our paper offers a methodological contribution by accounting for flexibility in both plant and harvest dates simultaneously.

To the best of our knowledge there has been only one prior application of the KFMA data to modeling the impact of weather on Kansas farms (Lambert 2014). However, this article used county-level average weather in estimating the impact of temperature and precipitation on crop and livestock output indices for a small subsample of KFMA farms. Recent literature has critiqued the inclusion of aggregated weather variables in modeling climate change impacts, showing that this can introduce bias into the estimates (Fezzi and Bateman 2015; D'Agostino and Schlenker 2016). Our article avoids this problem by using disaggregated weather at the PRISM grid-cell level matched to farm locations. This allows us to more precisely capture the heterogeneous warming impacts within Kansas and avoid potential measurement error problems that can bias-downward estimates of weather effects (Aufhammer et al. 2013). Additionally, our article raises new questions regarding sorghum's decreased ability to tolerate heat over time. Similar effects have been documented in other crops, however our result is the first of its kind for sorghum using farm-level data (Lobell 2008).

The location of our study, Kansas, provides an ideal setting for studying the effects of warming on sorghum production. Kansas represents by far the largest share of sorghum acreage and production in the United States, accounting in 2017 for more than 46% of total planted acreage, and more than 55% of total production (USDA-NASS). A large amount of the state's sorghum is grown alongside corn, which makes the argument of crop switching particularly relevant. We also observe sorghum production spread across all nine of Kansas' USDA Crop Reporting Districts (CRD). The wide range of climate across Kansas provides a considerable amount of weather variation in our sample.

In what follows, we describe the yield and weather data used in the study, the definition of growing season dates, and the construction of growing degree day variables from the underlying daily temperature data. Next, we specify the preferred and alternative models used to determine the effects of heat on sorghum yields. In our result section, we detail the impacts of incremental warming on sorghum yields, we show the bias in our estimates that occurs when using alternative temperature measurements, and we show how inclusion of other climate and economic variables does little to improve model performance or alter warming impacts. We then describe the simulation of three alternative adjustments to the growing season, and the reduction in warming-induced yield losses that occurs from changing plant and harvest dates. We also describe the lack of substantial improvement in yield impacts that results from re-estimation of our preferred model using a cross-sectional approach. Finally, we demonstrate the declining heat effects of sorghum over time.

Data

This essay made use of three datasets. Annual dryland sorghum yields were taken from the Kansas Farm Management Association (KFMA) database, which contains production and cost information on Kansas farms. A total of 45,971 farm-level observations on dryland sorghum production, spread across the 1978-2015 period were used in this analysis. Summary statistics for variables used are provided in **Table 1**. There is an average of 1,348 farms per year in the sample, and this varies across years from a low of 449 to a high of 1,676 farms. Due to entry and exit of farms from participation in the KFMA, the data form an unbalanced panel with some farms occurring in a single year and others spanning all years; farms are observed in the data, on average, for 16 years[1]. Farm locations are shown in **Figure 6**.

Daily minimum and maximum temperatures and daily precipitation gridded at the 4x4 kilometer level for the state of Kansas were gathered from the PRISM Climate Group dataset. Finally, we collected historic sorghum growing season data for each of Kansas' nine USDA CRDs from

USDA National Agricultural Statistics Service (NASS) crop progress reports (see **Figure 7** for a map of the CRD locations). A farm's plant (harvest) date for each year was the date at which percent planted (harvested) in the CRD's crop progress report was greater than or equal to 50%. **Table 2** provides summary statistics for yields, growing season length, and weather by CRD. Average growing seasons across CRDs range from 122 to 140 days and do vary from year to year as standard deviations ranges from 13 to 20 days. Planting dates exhibit less annual variation with standard deviations ranging from 5 to 11 days (versus a range of 14 to 19 days for harvest dates). Although not reported in the table, the earliest observed planting date in the full sample was in early May and the latest harvest date was in late December.

Mailing addresses for KFMA sorghum farms were geocoded and the associated yield data matched to PRISM weather grid cell observations.[2] Previous studies (Tack et al. 2015) have shown that growing degree days correct for bias present in average daily temperatures and minimum and maximum temperature variables that smooth out intra-season weather fluctuations. We calculated growing degree days (summed over days within each year's growing season) using an interpolation of daily minimum and maximum temperatures, following Schlenker and Roberts (2009). We also construct a cumulative precipitation variable by summing across daily precipitation. The production and weather data exhibit a large amount of cross-sectional and temporal variation within our sample (**Table 1, Table 2, and Figure 8**). Average growing degree days above 33°C and seasonal precipitation for the selected grid cells are shown in **Figure 9**, which provides a snapshot of the baseline (average weather) climate and demonstrates the substantial within-state climate variability for Kansas.

Methods

We model sorghum yields as a function of growing degree days and precipitation,

$$(1)\ y_{it} = \alpha_i + \beta_1 t + \beta_2 prec_{it} + \beta_3 prec_{it}^2 + \beta_4 DDlow_{it} + \beta_5 DDmed_{it} + \beta_6 DDhigh_{it} + \varepsilon_{it}$$

where y_{it} is log yield of farm *i* in year *t*. Farm fixed effects are included to control for time-invariant features of the farm such as soil quality, and a linear time trend captures changes in technology over time. We cluster standard errors by year to account for likely spatial correlation of the error terms across farms.

The three temperature variables, $DDlow_{it}$, $DDmed_{it}$, and $DDhigh_{it}$ capture respectively, degree days between zero and the lower threshold, the lower threshold and the upper threshold, and above the upper threshold. Essentially, these variables admit a piecewise linear relation between temperature and log-yield over three ranges of temperatures defined endogenously (see below) by the placement of cut points (thresholds). We also include a quadratic function of cumulative precipitation to capture the nonlinear effects of rainfall. We do not consider within-season heterogeneous effects of weather as recent research has suggested similar climate change impacts across both heterogeneous and homogeneous models (Tack et al. 2017; Ortiz-Bobea et al., 2019).

Climate change impacts were calculated by shifting the underlying daily temperature data by the intended scenario (i.e. 1 – 5°C warming), re-interpolating these data to calculate growing degree days, and calculating yield impacts by incorporating the shifted degree day variables. That is, we shifted daily minimum and maximum temperature variables (for the growing season) incrementally by 1°C for all farm locations. This process allows us to simulate warming by shifting the entire distribution of observed historic temperatures. For each warming scenario we then reinterpolated the shifted data to form growing degree days, and recalculated the three temperature variables, $DDlow_{it}$, $DDmed_{it}$, and $DDhigh_{it}$. In order to capture climate shifted yield impacts, we subtracted these new temperature variables (for the shifted data) from the original temperature variables and calculated impacts using the differenced temperature variables and original parameter estimates. Specifically, the yield impact for farm *i* when the average weather variables shift from the historically observed average, x_{i0}, to the new values, x_{i1}, is, $impact_i = 100[e^{\beta\{x_{i1}-x_{i0}\}} - 1]$. Precipitation is held fixed at historical averages for all warming scenario simulations.

Results

The results are divided into three sections. In the first section, we discuss the response of KFMA sorghum yields to extreme heat, consider impacts to yield from a warming climate, and compare estimates from our preferred model to ones with alternative specifications. In the second section we discuss to what degree farmers can mitigate yield loss through adjustment of the growing season. Finally, we report and discuss results for three alternative specifications: including year fixed effects, heat/precipitation interactions, and allowing the effect of heat to vary across time.

Sorghum Yields Drop Sharply with Exposure to Temperatures over 33°C

We follow the model laid out by Schlenker and Roberts (2009) to examine the effects of temperature on yield. We specify two separate temperature thresholds for growing degree days, and estimate the model over all possible thresholds. The thresholds that provided the best fit (i.e. highest r-squared) were selected as optimal, 10°C and 33°C. These thresholds admit a piecewise linear relationship between yield and temperature, as shown in **Figure 1.** This is similar to the temperature response functions estimated in Tack et al. (2017), which relied on field-trial and county-level aggregate producer data. Parameter estimates are reported in column 1 of **Table 3**. We consider how sensitive the parameter estimates are to the selected temperature cut point (i.e. 33ºC) for extreme heat and find the same piecewise relationship. Higher cut points for extreme heat produce larger marginal effects for $DDhigh_{it}$, however this is counterbalanced by the diminishing level of the variable (i.e. time exposure to temperatures greater than 35ºC is strictly less than 33ºC). Parameter estimates for the model with alternative extreme heat cut points (31-35ºC) are reported in **Table 4**.

Warming Scenarios Produce Heterogeneous Yield Impacts across Kansas

Under the most moderate warming scenario (1°C), all farms show yield losses relative to baseline climate. The effects of warming at the grid-cell level are presented in **Figure 2**. To measure aggregate effects at the state-level, we use a 5-year average (years 2012-2016) of total harvested sorghum acres at the district level (as reported by USDA-NASS) to calculate a weighted average. A 1°C warming scenario produces an average 11% decrease in yield; a 2°C warming scenario produces an average 24% decrease in yield; a 3°C warming scenario produces an average 39% decrease in yield; a 4°C warming scenario produces an average 54% decrease in yield; a 5°C warming scenario produces an average 67% decrease in yield.

Our results suggest that there is extensive spatial heterogeneity of the warming impacts as the severity of yield loss is highly geographic in nature: the most severe losses (for all warming scenarios) are found in the south-central and southwestern regions of the state; the most moderate losses are found in the northeast region. Under a 1°C warming scenario, yield losses range from 3-17% in the southwestern district while in the northeast district losses range from 1–11%. Under a 2°C warming scenario the range of yield losses increase to 10-35% in the southwestern district and to 6-26% in the northeastern district. However, as we move from the mildest warming scenario (1°C), to the most severe (5°C), we see severe levels of damage (>50%) appear across the entire state.

Alternative Specifications of Temperature and Precipitation

Our modeling approach included the creation of growing degree variables through interpolation of daily temperatures and a quadratic function of precipitation. We find that extending the precipitation effect to include a cubic term produces similar model performance (in terms of in- and out-of-sample fit) and does not alter the estimated warming impacts (**Table 5, Figure 11**). Conversely, we find that replacing degree day variables with average temperatures substantially reduces model performance and leads to understated warming impacts across all warming scenarios (**Table 5, and Figure 11**). This is most likely due to the dilution of daily fluctuations in temperature that occurs when using an average temperature variable. The use of minimum and maximum temperatures does not remedy this issue (**Table 5, Figure 11**). Parameter estimates for these models are provided in columns 2, 3, and 4 of **Table 3**.

Warming Impacts Are Robust to Inclusion of Additional Weather and Economic Variables as Controls

We weigh the consequences of omitting other weather variables (widely encountered in the climate change literature) and input and output prices (often included in output supply modeling) from our preferred specification. We do this through estimation of three additional alternative specifications that include additional control variables. The first two add an additional weather covariate to the preferred specification: vapor pressure deficit as measured in Roberts et al. (2013) and Lobell et al. (2014); and (ii) time-exposure to freezing temperatures below 0°C. The third focuses on economic variables through the inclusion of two measures of relative prices: the ratio of expected output price over sorghum seed price, and the ratio of expected output price over an index of all input prices.[3] None of these alternative models substantially improve model performance (both for the case in which we use the optimal degree day knots from the preferred model and for the case in which we re-optimize these knots under the alternative specifications) (**Table 6**), nor alter estimated warming impacts (**Figure 12**). Parameter estimates for these models are provided in columns 5, 6, and 7 of **Table 3**.

Re-optimizing Plant and Harvest Dates Partially Offsets Damages from Warming

A possible adaptation strategy that producers could use when faced with an increased warming climate is to advance the growing season to an earlier part of the year. This would potentially allow producers to escape some of the adverse effects of heat exposure. We consider three approaches for adjusting the growing season under a 2°C warming scenario, all of which focus on advancing the growing season earlier in the year.[4]

The first approach considers moving up the plant and harvest dates by equivalent one-day increments (i.e. keeping the length of the growing season fixed) for each farm in the sample. This is done up unto the point at which the growing season has been advanced by three weeks. The second approach is the same as the first, except that we allow the length of the growing season to be shortened based upon degree day accumulation above 5°C.[5] Specifically, we shorten a farm's growing season by advancing the harvest date forward to the date at which accumulated degree days over 5°C (under warming) is less than the maximum accumulated degree days over 5°C observed historically (the effect of this procedure on growing season length, under the 1 - 5°C warming scenarios, is shown in **Figure 13**). If this point is never reached, then the farm's growing season length is left unaltered and the second approach collapses into the first approach. The third approach moves up planting by one-day increments while holding the harvest date fixed (effectively lengthening the growing season).

It is important to consider the soil temperature requirements for sorghum germination (Anda and Pinter 1994) as shifting all of the farms' growing season by a full three weeks may not be feasible. To address this concern we construct a grid-specific threshold limiting the number of days that we can advance the growing season – that is we advance the growing season until the average minimum temperature in the first two-week period after planting under 2°C warming is equal to the average minimum temperature in the first two-week period under the original baseline climate.

Our results suggest that shifting the growing season reduces warming impacts across nearly the entire sample. This is shown for each USDA-CRD in **Figure 3**. The slopes of the blue, red, and green lines represent the acreage-weighted district-level mean reduction in yield damages associated with the three alternative growing season adjustments: a shifted, fixed-length growing season, a shifted, shorter growing season, and a shifted, longer growing season respectively. We see that allowing for adjustments to the growing season can reduce yield damages, however the difference in damages under the new growing seasons are small - generally not more than a few percentage points apart.

The black vertical lines show the district-wide average maximum number of days that planting can be advanced without violating sorghum's soil germination temperature constraint (which is proxied by the average minimum temperature in the first two weeks of the growing season). On average, farms in northeastern Kansas (i.e. Districts 40, 70, and 80) are the ones most likely to benefit from moving up the growing season. These districts are the ones that had the most moderate yield damages under a 2°C warming scenario, suggesting that shifting the growing season is most effective for farms that face shorter periods of extreme heat and less severe yield damages. In contrast, farms located in western and southwestern Kansas (i.e. Districts 10, 20, and 30) are less likely to be able to offset yield damages by moving the growing season. Recent agronomic research has identified sorghum varieties that are highly resistant to early season chilling, which if utilized might allow for sorghum planting to be pushed back further, up to the point at which sorghum could be planted simultaneously with corn – which would allow for more flexibility in shifting the growing season beyond two to three weeks and potentially allow for a greater reduction in yield damages (Chiluwal et al. 2018). We document the distribution of the historical difference in sorghum and corn planting dates across all nine CRDs in **Figure 14**.

The optimal, grid cell-specific number of days to advance the growing season (under the fixed-length approach) and the resulting reductions in yield damages are shown in **Figure 4a – 4d** (similar results are documented for the shorter and longer growing season approaches in **Figure 15 and Figure 16**). The optimal number of days to move up planting (and harvesting) for the sample ranges from between 9 to 17 days (**Figure 4b**). We see that these are similar within each CRD – this may well be a function of what district the grid cell lies in (as our data on planting and harvesting dates are reported at the district level). However, there is heterogeneity in the optimal number of days within each district as well.

Across the state, adjusting the growing season allows farms to only partially offset losses they would face under 2°C warming scenario. Average acreage-weighted yield damages for the sample are 24.18% without reoptimization and 22.29% under a re-optimized growing season. Similar insights are supported by the model that includes exposure to freezing temperatures, suggesting that it is not likely to play a more prominent role under growing season adjustments (**Figure 17**). The yield impact estimates after reoptimization of the growing season are shown in **Figure 4c**, and the percentage point difference is shown in **Figure 4d**. We again see a great geographic disparity in terms of impacts of the growing season shift, with farms in northern and eastern Kansas receiving greater reductions in damages from shifting the growing season earlier.

Alternative Forms of Adaptation and Controls for Time-Varying Omitted Variables

Reoptimizing the growing season is just one mechanism by which producers can offset warming impacts. Recently, Hsiang (2016) summarized more general approaches focused on long averages (Burke and Emmerick 2016) and cross-sectional regression. The cross-sectional approach collapses the panel by re-estimating the model using averages of all covariates across time, thereby accounting for a wider range of adaptation versus the long differences approach. Similar to the growing season adaptation results above, the cross-sectional approach suggests that adaptation is not likely to greatly reduce warming impacts with yield losses spanning from 7% to 59% across the 1 - 5°C warming scenarios (**Figure 18**). Parameter estimates for this model are provided in column 6 of **Table 7**.

It is also important to consider whether there might be additional omitted variables not considered above that are biasing the warming impacts. First we consider replacing the trend variable with year fixed effects which control for a much wider set of technology trends, price effects, and also common weather shocks across farms. Next, we consider year-by-CRD fixed effects which accomplishes the same objective but at the CRD level, i.e. CRD specific technology trends, price effects, and weather shocks. Our results suggest much larger warming impacts under these alternative specifications (**Figure 18**), however caution should be taken when considering the results of these models as the additional levels of fixed effects greatly reduce the variation of the weather data (**Table 8**). For example, the standard deviation of extreme heat (precipitation) reduces by 65% (39%) when adding year fixed effects relative to the preferred model with only farm fixed effects. Reductions increase to 79% (66%) when adding year-by-CRD fixed effects. Parameter estimates for these models are provided in columns 2 and 3 of **Table 7**.

Inclusion of a Temperature-Precipitation Interaction Term Suggests an Important Role for Adaptation via Irrigation

While our focus here is on dryland production, we consider the potential role that irrigation could play in adapting to warming temperatures by including an interaction between the precipitation and growing degree days above 33°C variables. Although this interaction is not statistically significantly different from zero, the coefficient estimate suggests that increased levels of precipitation are associated with reduced damages from heat (column 4 of **Table 7**). Incorporating this interaction into the warming impacts suggests that while low levels of precipitation (defined as the 10^{th} percentile of all sample observations) do not largely increase impacts relative to our preferred model, high levels (defined as the 90^{th} percentile of all sample observations) are associated with much smaller warming damages (**Figure 19**). Thus, while high levels of precipitation do not perfectly reflect the advantages of irrigated sorghum production where water applications can be perfectly timed as needed, these results do suggest a strong role that irrigation might play in adapting to warming temperatures.

Inclusion of a Trend-Extreme Heat Interaction Term Reveals Increased Heat Effects over Time

While irrigation is associated with reduced damages, it is not likely to be available in all areas and – due to increased pressures on water aquifers – it might play an even more limited role in crop production in the future. Thus, it is important to consider alternative management strategies within the dryland context that might currently be available to producers. We do not consider specific practices here but rather interact the trend and growing degree days above 33°C variables to measure whether the effect of heat has changed over time. Parameter estimates are reported in column 5 of **Table 7**. The negative coefficient on the trend-growing degree days above 33ºC term suggests that damages from heat are actually increasing over time. This term's coefficient is statistically significant at the 95% confidence level. Compared to the first year in the data, the effect of an additional unit of degree days above 33°C has increased by a factor of 1.7 compared to the most recent year (-0.010 vs -0.017). This discrepancy is also reflected in the warming impacts (**Figure 5**). The downside to this finding is that current production practices are associated with larger damages compared to our preferred model, while the upside is that there is likely to exist some historical practices that could help offset the effects of warming.

Conclusion

Our results show that the on-farm impact of a warming climate on dryland sorghum yields can be highly damaging. This has important implications for producers, policymakers, and researchers interested in the effects of, and optimal farm response to, climate change. For producers looking to avoid losses in other crops, sorghum has the benefit of being better able to withstand higher temperatures than other grain crops. But the argument that producers can fully adapt to a warming climate through diversification of crops - by switching from a "heat intolerant" crop to sorghum (Burke and Lobell 2010; Brown and Funk 2008; Cairns et al. 2013) – is not supported by our results. While sorghum has a higher temperature threshold for productive growth than other crops, it also suffers from the same damaging effects once its hotter growing season is properly accounted for.[6] An important issue raised in the literature is that the Schlenker and Roberts (2009) modeling approach, in not considering additional economic (Miao et al. 2016) and weather covariates (Chen et al. 2016; Zhang et al. 2017), could potentially produce biased climate change estimates. We find however, that our result is robust across a variety of alternative model specifications that explicitly control for these measures.

Our analysis also shows that losses associated with 2°C warming can be partially offset by adjusting the growing season, however the impact of adjustment is relatively small. For a farm with limited capital and labor resources, the benefits of planting earlier in the year may not be substantial enough to incentivize adoption of this strategy.[7] The feasibility of adjusting sorghum's growing season is most likely contingent, among other things, on a farm's planting and harvesting decisions for other crops and rotations. Producers will have to weigh the tradeoffs involved with adjusting the growing season and how this effects overall farm profitability. More generally, we also consider other econometric models that permit a wider scope for adaptation potential and find that warming impacts remain large.

Along with this management constraint there is an additional policy constraint that could potentially limit the ability of farms to realign their planting dates. Each year the USDA Risk Management Agency (RMA) publishes earliest planting dates for sorghum. These dates serve as part of a guideline for what is considered best management practices, and are important because if a farm chooses to plant prior to the RMA dates, they can lose their eligibility for insurance replant payments. For the last twenty years, RMA's earliest sorghum plant date for the entire state of Kansas has remained the same – April 26th (USDA – RMA, 2018). If policymakers are interested in promoting farm adaptation to climate change, they should consider incorporating more flexible planting dates when defining best management practices.

This study will benefit researchers interested in further investigating sorghum's response to extreme heat. Future research should focus on understanding better the factors driving sorghum's increased sensitivity to extreme heat over time. The agronomic literature has documented changes in both the types of hybrids and farm management practices over the past fifty years (Menz et al. 2004; Hamman, Dhuyvetter, and Boland 2001; Assefa and Staggenborg, 2010). Sorghum breeding has largely focused on drought tolerance and increasing average yields, rather than resilience to extreme heat (Menz et al. 2004). At the same time, row spacing and intrarow spacing decreased for dryland sorghum, suggesting increases in seeding rates and plant densities (Assefa and Staggenborg, 2010). More generally, our finding of increased sensitivity to heat somewhat contrasts with recent results suggesting decreased sensitivity to drought for genetically modified corn and soybeans (Yu and Babcock, 2010), thereby suggesting a potential upside for transgenic sorghum breeding. It is also possible that increased enrollment in heavily subsidized crop insurance has incentivized producers to take on more risk. While our data did not contain observations on the hybrids used, management practices, nor crop-specific crop insurance purchases of the farm, future researchers may try to collect this information in order to uncover the underlying dynamic of the results found in our study.

We close with some important caveats to our article. In conducting our analysis, we chose to exclude CO_2 from our study, despite that the beneficial effect of CO_2 on crop yields (i.e. the fertilization effect) has been the subject of a series of regional and global yield modeling studies (Reilly et al. 2003; McGrath and Lobell 2013). In view of this, our results should be considered as a pure counterfactual of temperature change rather than as a comprehensive climate change analysis. We also wish to emphasize that the simulated warming impacts presented in our study are short-run in nature; we do not account for any climate change adaptation behavior beyond changes in the plant and harvest dates (Yang and Shumway 2015; Kaminski et al. 2013) and we do not account for changes in area allocation between sorghum and other competing crops (Miao et al. 2016). Finally, our study does not examine risk effects stemming from changes in the higher order moments of crop yield, nor does it consider changes to the tails of the climate distribution that could result from more frequent extreme events.

Footnotes

[1] We cannot rule out sample selection concerns in general, but it is unlikely to be caused by farm exit after experiencing crop failure, which is a major concern for identifying weather effects on crop yields. Define an *exit* variable to take on a value of 1 for the farm's last year in the data (ignoring 2015) and a yield *shock* variable as the residual from a regression of log-yield on trend and farm fixed effects. The full sample correlation of these variables is very small, at -0.04. If we keep only negative yield shocks (i.e. so we just focus on yield losses) the correlation remains the same, and if we further keep only the largest yield losses (the largest 10%), the correlation remains small, at -0.09. The Psuedo-R-squared's for separate probits of *exit* on the alternative measures of *shock* are all less than 0.01. In addition, we re-estimate the preferred model (defined below) after dropping the farm's last year from the sample and find nearly identical warming impacts.

[2] KFMA farms that had mailing addresses in a different state were excluded from this analysis. Mailing addresses may be an imperfect measure of field location, but this is the only available location reference in the data. If this matching induces substantial measurement error, then we would expect attenuation bias in the estimated warming impacts, thereby suggesting larger impacts than those reported here.

[3] The expected output price is an average of the prior 5-years' real output prices; the sorghum seed price is a 1-year lag of seed price (seed prices are reported for a marketing year). Output and seed

prices were obtained from USDA-NASS and converted to real dollars using the U.S. Bureau of Labor's consumer price index for food; the price index for production items (relative to the parity base period, 1910 – 1914) was obtained from USDA-NASS.

[4] There is agronomic evidence, e.g. Carbone et al. (2003) and Singh et al. (1998), indicating that sorghum's growing season is shortened under warmer-than-average temperatures – from approximately 1-3 weeks depending on the severity of warming. Within our data, correlations between Julian (day-of-year) planting dates and average daily minimum and maximum temperatures are -0.26 and -0.11 respectively, thereby suggesting that warmer temperatures/climates are associated with earlier planting.

[5] 5°C has been identified in the agronomic literature as the threshold below which plant development will be inhibited (Ciampitti and Knapp 2013).

[6] The current study does not argue that the negative effects of warming on sorghum yields are greater in magnitude relative to the effects on corn yields. Such an argument would require a comparative examination of the asymmetric downside impacts of warming on corn and sorghum yield distributions. We leave this question open as an area of future research.

[7] In any given year Kansas many sorghum producers allocate only a share of total acreage to sorghum, with producers often occupied with corn planting in the weeks leading up to sorghum planting. Shifting up sorghum planting dates therefore may not be practical if it impedes other enterprises' planting decisions (Mishra et al. 2004).

Tables

Table 1 Summary Statistics for Regression Data

Variable	Mean	Std Dev	Min	Max
Number of farms by year	1,348.32	330.54	449	1,676
Number of years by farm	16.13	9.44	1	38
Sorghum yield	63.18	28.50	0	174.55
Cumulative precipitation	382.30	155.17	68.07	988.50
Average daily min temp	16.08	2.11	7.36	20.44
Average daily max temp	29.87	2.14	20.27	36.04
DDlow	1280.07	125.03	989.97	1844.27
DDmed	1690.30	149.22	1134.01	2294.05
DDhigh	25.26	21.27	0	140.59
Average daily vapor press deficit	2.49	0.39	1.43	3.83
Cumulative days of freeze exposure	0.74	2.69	0	25.44
Price ratio (output/seed)	0.0093	0.0053	0.0031	0.024
Price ratio (output/inputs)	0.062	0.021	0.034	0.12

Note: Total number of observations is 45,971 across 5,541 farms spanning 38 years (1978-2015). DDlow is defined as degree days above 0ºC minus degree days above 10ºC, DDmed is defined as degree days above 10ºC minus degree days above 33ºC, and DDhigh is defined as degree days above 33ºC.

Table 2 Growing Season and Weather Statistics by Crop Reporting District

CRD	Yield (bu/acre)	Plant Date (Julian)	Harvest Date (Julian)	Season Length (Days)	Precip (mm)	Min Temp (°C)	Max Temp (°C)
10	54.9	159.1	298.7	139.5	282.4	12.4	28.7
	(29.5)	(7.5)	(16.7)	(17.6)	(102.3)	(1.7)	(2.2)
20	48.4	158.6	296.3	137.6	265.2	13.2	29.5
	(26.6)	(5.4)	(14.1)	(13.9)	(87.1)	(1.5)	(1.9)
30	46.5	160.1	299.6	139.4	288.8	14.2	30.0
	(26.1)	(6.7)	(14.0)	(14.3)	(103.8)	(1.4)	(1.8)
40	68.9	162.9	296.8	133.8	347.8	14.8	29.3
	(29.9)	(7.2)	(15.9)	(16.8)	(129.1)	(1.8)	(2.2)
50	62.3	163.7	292.3	128.6	359.2	16.2	30.0
	(27.0)	(8.5)	(18.0)	(17.3)	(125.1)	(1.7)	(2.1)
60	54.3	162.5	293.0	130.5	362.1	16.5	30.3
	(26.5)	(11.4)	(18.3)	(15.9)	(126.5)	(1.6)	(2.0)
70	77.2	155.9	288.4	132.5	459.6	15.9	28.9
	(27.3)	(10.2)	(18.7)	(19.7)	(181.6)	(1.8)	(2.4)
80	69.1	154.9	284.4	129.4	451.7	16.6	29.5
	(26.8)	(10.0)	(13.8)	(14.0)	(179.9)	(1.4)	(1.9)
90	67.0	151.5	274.2	122.6	424.4	17.8	30.6
	(26.5)	(11.2)	(14.4)	(12.9)	(154.4)	(1.1)	(1.7)

Note: Mean values for variables by CRD are reported with standard deviations, in parentheses, below.

Table 3 Regression Model Parameter Estimates for Log of Sorghum Yields

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
cumulative precip	0.00328***	0.00592***	0.00536***	0.00404***	0.00336***	0.00346***	0.00308***
	(0.000583)	(0.00117)	(0.000847)	(0.000741)	(0.000628)	(0.000584)	(0.000570)
cumulative $precip^2$	-3.01e-06***	-9.25e-06***	-4.76e-06***	-3.94e-06***	-3.06e-06***	-3.13e-06***	-2.82e-06***
	(5.96e-07)	(2.35e-06)	(8.15e-07)	(7.60e-07)	(6.08e-07)	(5.98e-07)	(5.83e-07)
DDlow	-0.000427	-0.000414			-0.000312	-0.000650*	-0.000453
	(0.000291)	(0.000289)			(0.000455)	(0.000329)	(0.000302)
DDmed	0.000558**	0.000546**			0.000474	0.000617**	0.000510**
	(0.000224)	(0.000225)			(0.000367)	(0.000230)	(0.000229)
DDhigh	-0.0123***	-0.0122***			-0.0129***	-0.0123***	-0.0128***
	(0.00167)	(0.00167)			(0.00187)	(0.00163)	(0.00164)
cumulative $precip^3$		4.40e-09***					
		(1.51e-09)					
average temp			0.0318				
			(0.239)				
average $temp^2$			-0.00124				
			(0.00563)				
minimum temp				0.154***			
				(0.0428)			
maximum temp				-0.159***			
				(0.0395)			
vapor pressure deficit					0.0831		
					(0.213)		
freeze exposure						0.0141*	
						(0.00711)	
price ratio 1							10.68
							(12.35)
price ratio 2							2.113
							(3.539)
Observations	45,971	45,971	45,971	45,971	45,971	45,971	45,971
R-squared	0.488	0.489	0.442	0.458	0.488	0.489	0.490

Note: For all models DDlow is defined as degree days above 0ºC minus degree days above 10ºC, DDmed is defined as degree days above 10ºC minus degree days above 33ºC, and DDhigh is defined as degree days above 33ºC. Price ratio 1 is expected output price over seed price, while price ratio 2 is expected output price over an index of all input prices. Standard errors clustered by year are reported in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

Table 4 Regression Model Parameter Estimates with Alternative Cut Points

	(31C)	(32C)	(33C)	(34C)	(35C)
cumulative precip	0.00314***	0.00319***	0.00328***	0.00341***	0.00357***
	(0.000549)	(0.000566)	(0.000583)	(0.000599)	(0.000613)
$cumulative\ precip^2$	-2.94e-06***	-2.96e-06***	-3.01e-06***	-3.09e-06***	-3.21e-06***
	(5.76e-07)	(5.86e-07)	(5.96e-07)	(6.06e-07)	(6.17e-07)
DDlow	-0.000601*	-0.000514*	-0.000427	-0.000340	-0.000251
	(0.000300)	(0.000295)	(0.000291)	(0.000288)	(0.000284)
DDmed	0.000783***	0.000668***	0.000558**	0.000450*	0.000337
	(0.000228)	(0.000226)	(0.000224)	(0.000222)	(0.000217)
DDhigh	-0.00833***	-0.0100***	-0.0123***	-0.0155***	-0.0200***
	(0.00105)	(0.00128)	(0.00167)	(0.00230)	(0.00333)
Observations	45,971	45,971	45,971	45,971	45,971
R-squared	0.4868	0.4875	0.4878	0.4874	0.4863

Note: All regressions include farm fixed effects and a linear trend. For all models DDlow is defined as degree days above 0ºC minus degree days above 10C, however both DDmed and DDhigh are defined using alternative cut-points. DDmed is defined as degree days above 10C minus degree days above the noted cut-point, and DDhigh is defined as degree days above the noted cut-point. Standard errors clustered by year are reported in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

Table 5 Model Performance, Alternative Specifications of Temperature and Precipitation

Model	R-squared	RMSE
Preferred Model	0.4878	-11.39
Replace Quadratic with Cubic Precipitation	0.4893	-11.49
Replace Degree Days with Quadratic Average Temp	0.4419	-7.59
Replace Degree Days with Minimum/Maximum Temps	0.4575	-8.68

Note: R-squared reports the percentage of variation in log wheat yields explained by the model. Root-mean-squared prediction error (RMSE) is expressed as a percentage reduction relative to the baseline model with no weather variables. Each model is estimated 1,000 times, where each replication randomly selects 80% of the 45,971 observations in our full sample. Relative performance is measured according to the accuracy of each model's prediction for the omitted 20% of observations. The R-squared and RMSE for the baseline (no weather) model are 0.3461 and 0.6689, respectively. All models, including the baseline, have farm fixed effects and a linear trend. The optimal cut points for the degree day variables in the preferred model are 10º and 33ºC.

Table 6 Model Performance, Adding Additional Control Variables

Model	R-squared	Cut Points	RMSE
Optimal cut-points from the preferred model			
Preferred Model	0.4878	10, 33C	-11.39
Add Vapor Pressure Deficit	0.4879	10, 33C	-11.40
Add Freeze Exposure	0.4895	10, 33C	-11.53
Add Prices	0.4903	10, 33C	-11.61
Optimal cut-points for each model			
Preferred Model	0.4878	10, 33C	-11.39
Add Vapor Pressure Deficit	0.4879	10, 33C	-11.40
Add Freeze Exposure	0.4896	10, 32C	-11.56
Add Prices	0.4903	10, 33C	-11.61

Note: R-squared reports the percentage of variation in log wheat yields explained by the model. Root-mean-squared prediction error (RMSE) is expressed as a percentage reduction relative to the baseline model with no weather variables. Each model is estimated 1,000 times, where each replication randomly selects 80% of the 45,971 observations in our full sample. Relative performance is measured according to the accuracy of each model's prediction for the omitted 20% of observations. The R-squared and RMSE for the baseline (no weather) model are 0.3461 and 0.6689, respectively. All models, including the baseline, have farm fixed effects and a linear trend. All models in the first set use the optimal cut points for the degree day variables from the preferred model, while in the second set optimal cut points are re-optimized for each model.

Table 7 Regression Model Parameter Estimates for Log of Sorghum Yields

	(1)	(2)	(3)	(4)	(5)	(6)
cumulative precip	0.00328***	0.00306***	0.00246***	0.00258***	0.00322***	0.00339***
	(0.000583)	(0.000519)	(0.000410)	(0.000804)	(0.000572)	(0.000523)
$cumulative\ precip^2$	-3.01e-06***	-2.83e-06***	-2.21e-06***	-2.47e-06***	-2.96e-06***	-3.19e-06***
	(5.96e-07)	(5.11e-07)	(3.95e-07)	(6.92e-07)	(6.00e-07)	(5.28e-07)
DDlow	-0.000427	-0.000152	-3.61e-05	-0.000336	-0.000472	-0.000382*
	(0.000291)	(0.000298)	(0.00145)	(0.000305)	(0.000298)	(0.000203)
DDmed	0.000558**	0.000186	-0.000405	0.000473**	0.000560**	0.000642***
	(0.000224)	(0.000319)	(0.000395)	(0.000229)	(0.000219)	(0.000181)
DDhigh	-0.0123***	-0.0163***	-0.0114***	-0.0158***	-0.0101***	-0.0113***
	(0.00167)	(0.00308)	(0.00236)	(0.00312)	(0.00145)	(0.00139)
*DDhigh*precip*				1.42e-05		
				(1.00e-05)		
*DDhigh*trend*					-0.000185**	
					(7.40e-05)	
Observations	45,971	45,971	45971	45,971	45,971	5,541
R-squared	0.488	0.536	0.594	0.489	0.491	0.330
Farm FE	Y	Y	Y	Y	Y	N
County FE	N	N	N	N	N	Y
Year FE	N	Y	N	N	N	N
Year by CRD FE	N	N	Y	N	N	N

Note: All regressions include a linear trend and DDlow is defined as degree days above 0ºC minus degree days and below 10ºC, DDmed is defined as degree days above 10ºC minus degree days and below 33ºC, and DDhigh is defined as degree days above 33ºC. Model (6) is a cross-sectional regression using the time-averaged means of all variables with robust standard errors. All other models have standard errors clustered by year, *** p<0.01, ** p<0.05, * p<0.1.

Table 8 Variation of Weather Covariates Across Alternative Fixed Effects

Fixed Effects	*SD(Precip)*	*SD(DDlow)*	*SD(DDmed)*	*SD(DDhigh)*
None	155.2	125.0	149.2	21.3
Farm	131.7	106.9	125.8	18.5
Farm, Year	79.7	64.9	66.5	6.5
Farm, Year by CRD	45.4	3.6	17.1	3.9

Note: The first row reports the standard deviation (SD) of the four weather covariates for the full sample. The next three rows report the standard deviation of the residuals from a regression of the weather covariate on a trend and the noted fixed effects.

Table 9 Average Weather Variables by Growing Season (GS) Adjustment, Days Shifted

	Baseline (No Warming)			Fixed-Length GS (+2ºC)			Shorter GS (+2ºC)			Longer GS (+2ºC)		
shift	*freeze*	*dday5C*	*dday33C*	*freeze*	*dday5C*	*dday33C*	*freeze*	*dday5C*	*dday33C*	*freeze*	*dday5C*	*dday33C*
0	*1.13*	*2364.73*	*25.42*	*0.62*	*2627.25*	*54.49*	*0.12*	*2599.62*	*54.48*	*0.62*	*2627.25*	*54.49*
7	-	-	-	*0.43*	*2677.32*	*55.08*	*0.04*	*2633.04*	*55.02*	*0.62*	*2758.50*	*55.41*
14	-	-	-	*0.27*	*2709.86*	*54.80*	*0.01*	*2649.67*	*54.67*	*0.62*	*2880.72*	*55.80*
21	-	-	-	*0.20*	*2720.15*	*54.21*	*0.01*	*2651.29*	*54.00*	*0.62*	*2990.86*	*55.95*

Note: Each row reports time exposure to freeze, growing degree days greater than 5ºC, and growing degrees greater than 33ºC that are associated with growing season shifts for the baseline model (average climate and growing season), longer growing season adjustment (2ºC warmer than average climate, with only the plant date shifted, holding the harvest date fixed), fixed-length growing season adjustment (2ºC warmer than average climate, with both plant and harvest dates shifted by equal amounts) and shorter growing season adjustment (the same adjustment as the fixed-length adjustment, but with the harvest date shifted up further based on historical maximum exposure to degree days greater than 5ºC.

Figures

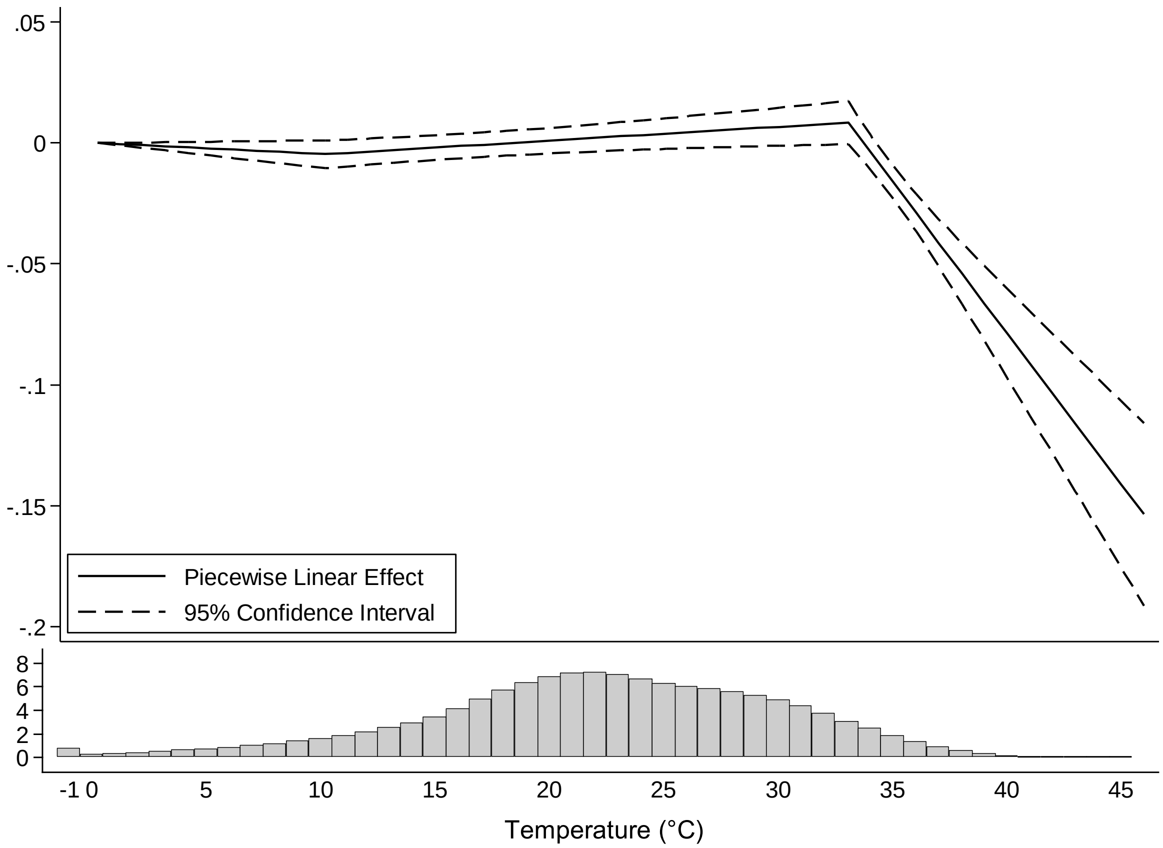

Figure 1 Piecewise linear relationship between temperature and sorghum yield

Note: The upper panel shows the effect on log yield when sorghum is exposed to a particular 1ºC temperature level. The dashed line denotes the 95% confidence interval calculated using standard errors clustered by year. The histogram in the lower panel shows the average temperature exposure across all farms in the dataset.

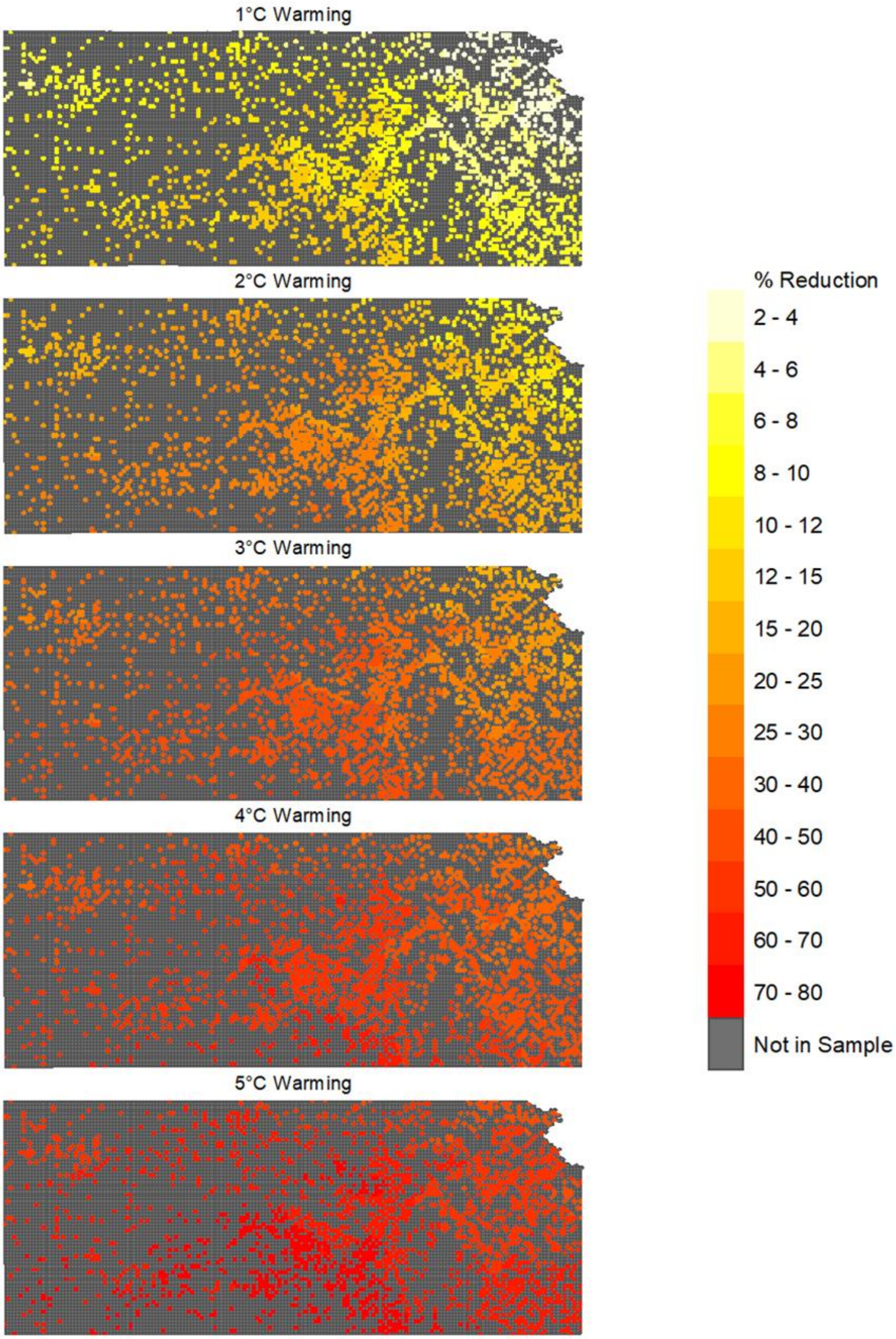

Figure 2 Predicted warming impacts on sorghum yields under 1-5ºC warming scenarios

Note: Each colored point represents a grid cell-specific yield reduction estimate for each warming scenario.

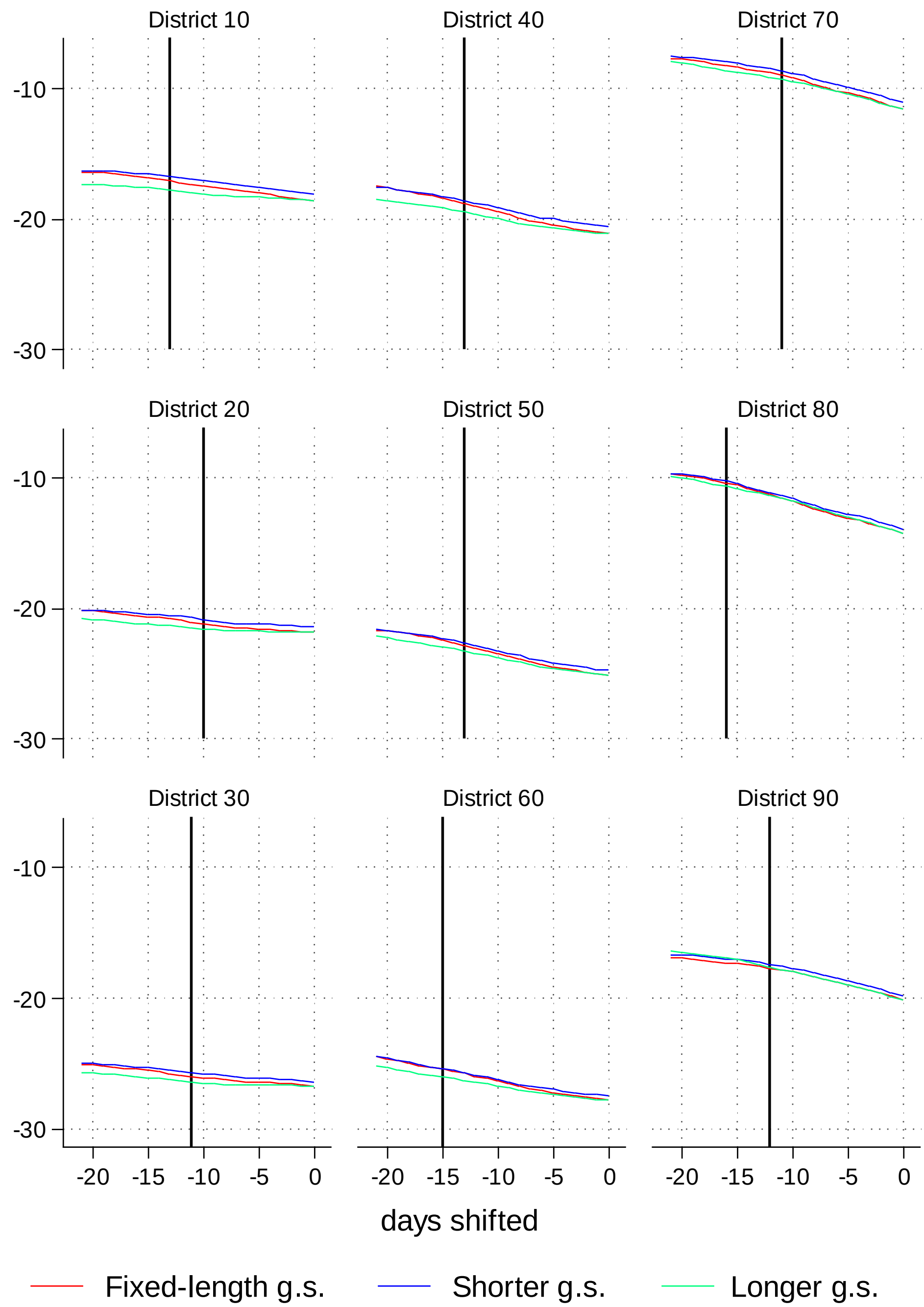

Figure 3 Predicted warming impacts on sorghum yields under 2ºC warming for various growing season adjustments

Note: Each red line shows the acreage-weighted average change in yield by advancing the growing season earlier, with both plant and harvest dates shifted by equal amounts (fixed-length g.s.); each blue line shows the acreage-weighted average change in yield by advancing the growing season earlier, allowing for shortening in growing season length based upon degree day accumulation above 5ºC (shorter g.s.); each green line shows the acreage-weighted average change in yield by advancing the growing season earlier, but keeping harvest date fixed (longer g.s.); each black line shows the district-wide average maximum number of days that planting can be advanced without violating sorghum's soil germination temperature constraint.

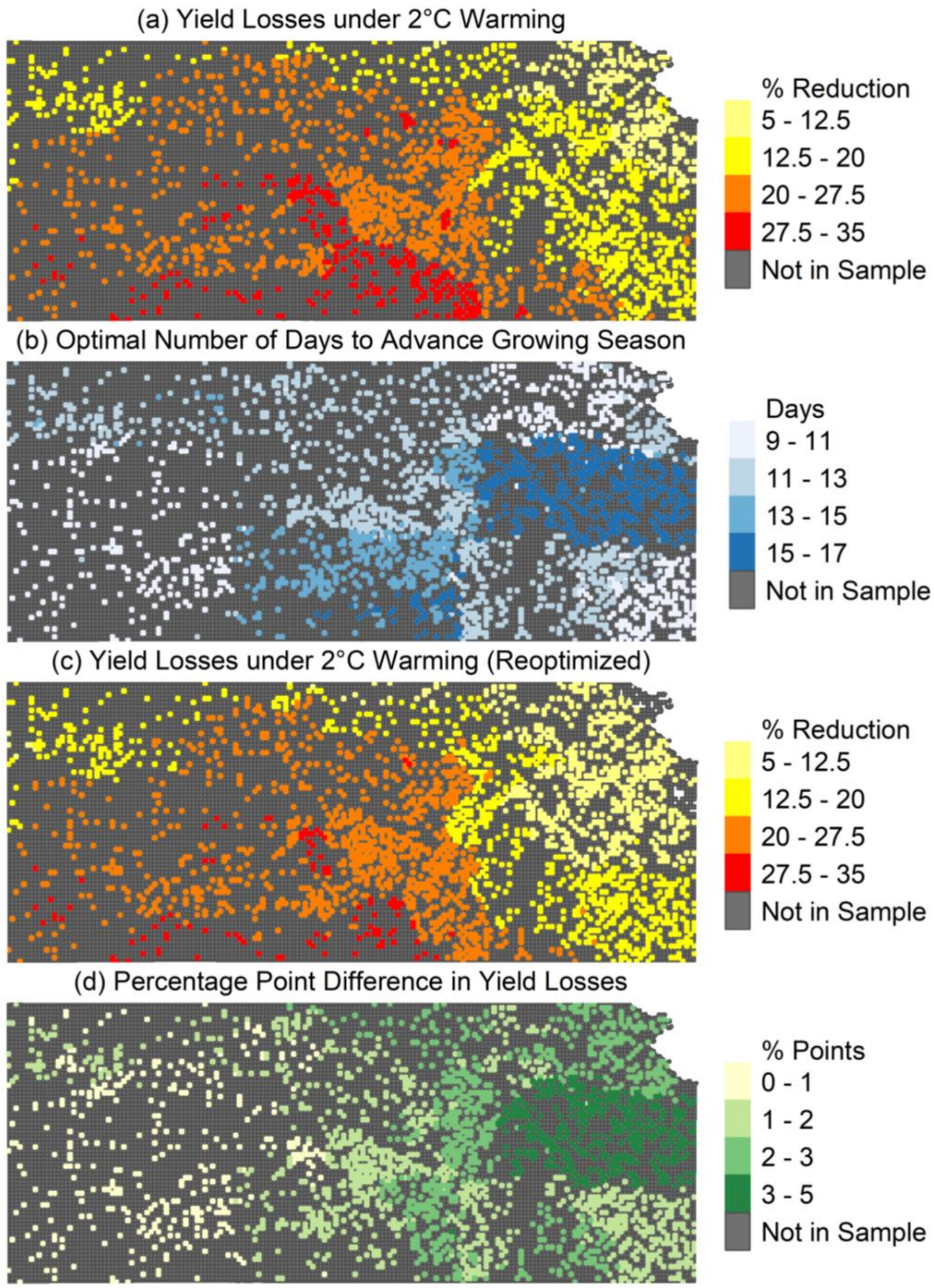

Figure 4 Predicted warming impacts on sorghum yields under 2°C warming when advancing both plant and harvest dates

Note: Panel (a) shows estimated impacts at the grid cell-level under fixed plant and harvest dates (baseline). Panel (b) shows the optimal number of days to advance the growing season when we advance the growing season, keeping the length of the growing season fixed (i.e. the first approach). Allowing for reoptimization of the growing season reduces predicted warming impacts (Panel (c)). The percentage point difference between warming impacts (baseline and reoptimized growing season; Panel (d)) shows a high degree of spatial heterogeneity

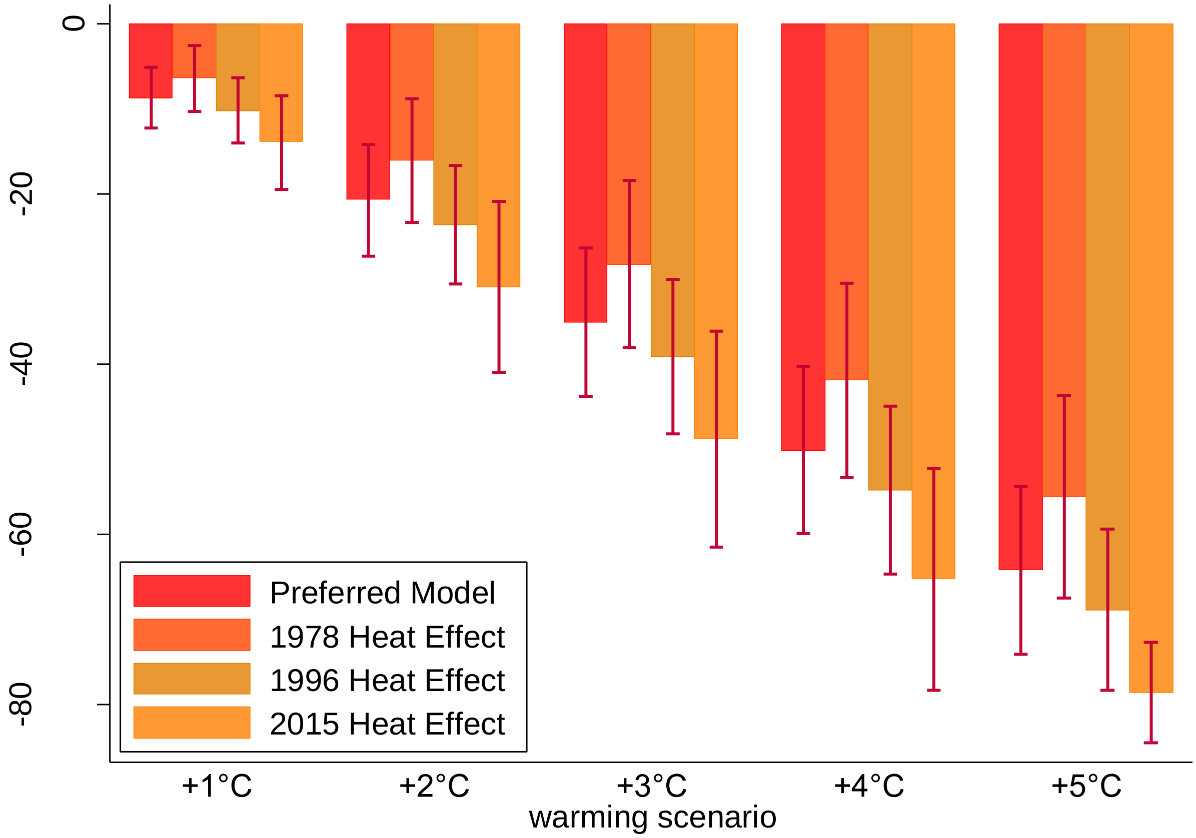

Figure 5 Predicted warming impacts on sorghum yields when including a heat-trend interaction term

Note: Each four-bar cluster shows estimates for the preferred model alongside three years from the heat-trend model.

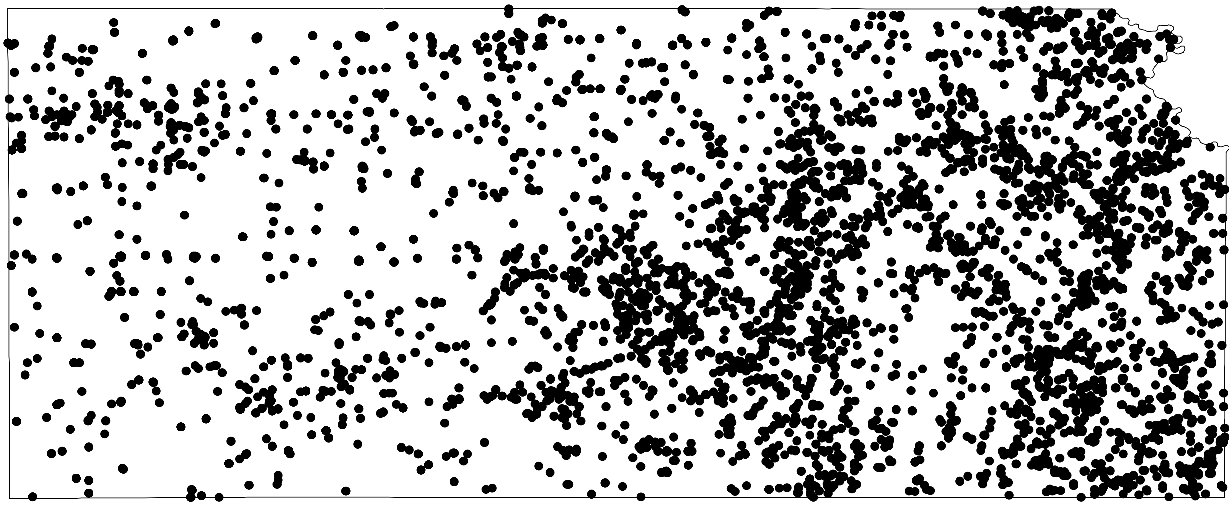

Figure 6 Locations of Kansas Farm Management Association (KFMA) dryland sorghum farms, 1978 – 2015

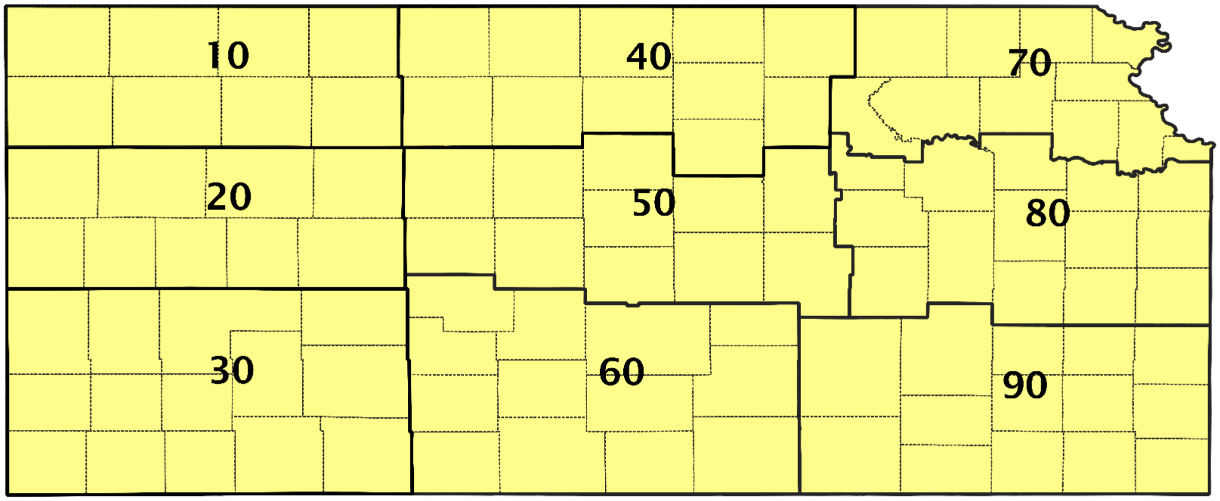

Figure 7 Locations of the nine USDA Crop Reporting Districts in Kansas

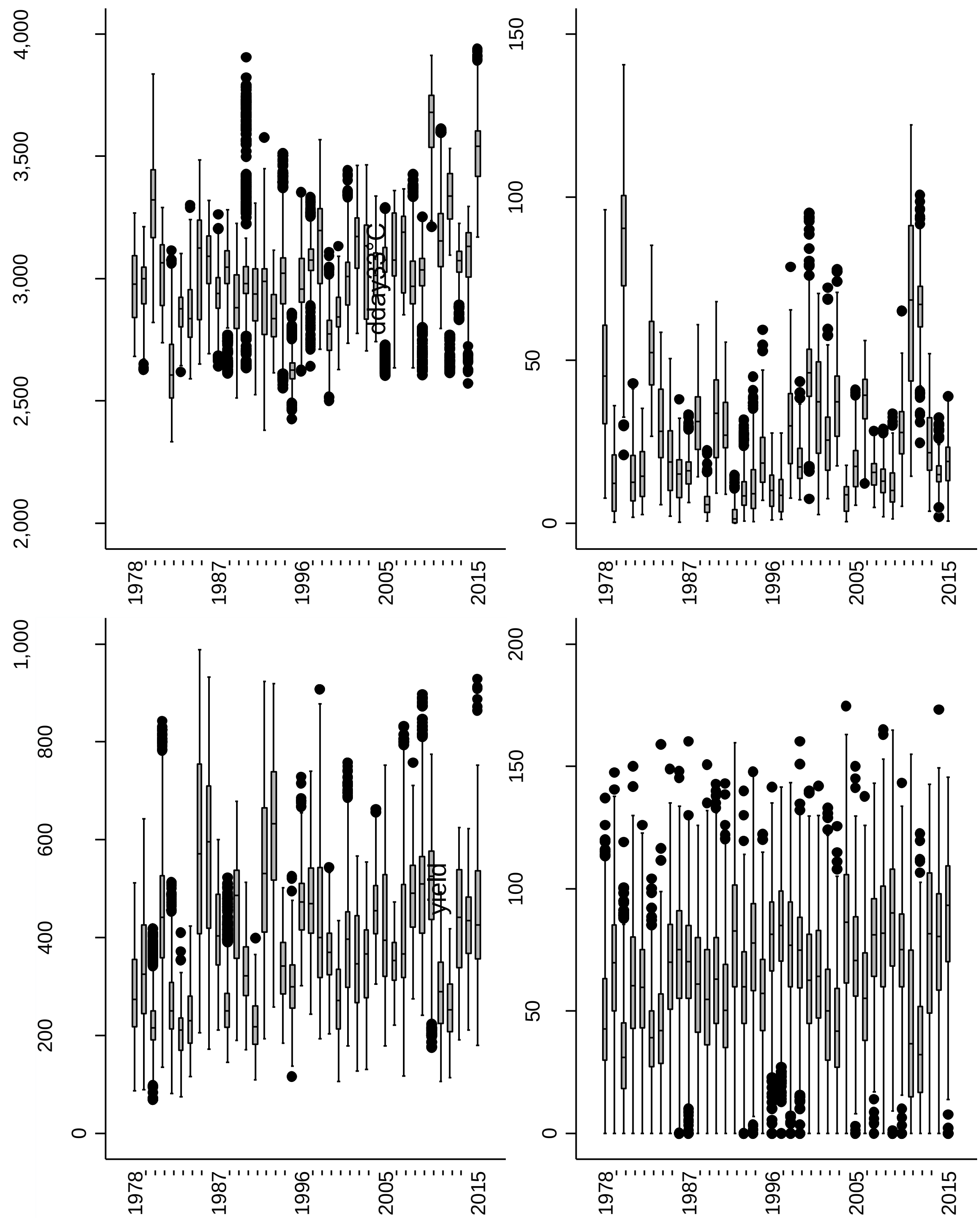

Figure 8 Box plots showing variation in annual dryland sorghum yields and growing season weather (total precipitation, growing degree days above 0°C, and growing degree days above 33°C) for KFMA farms

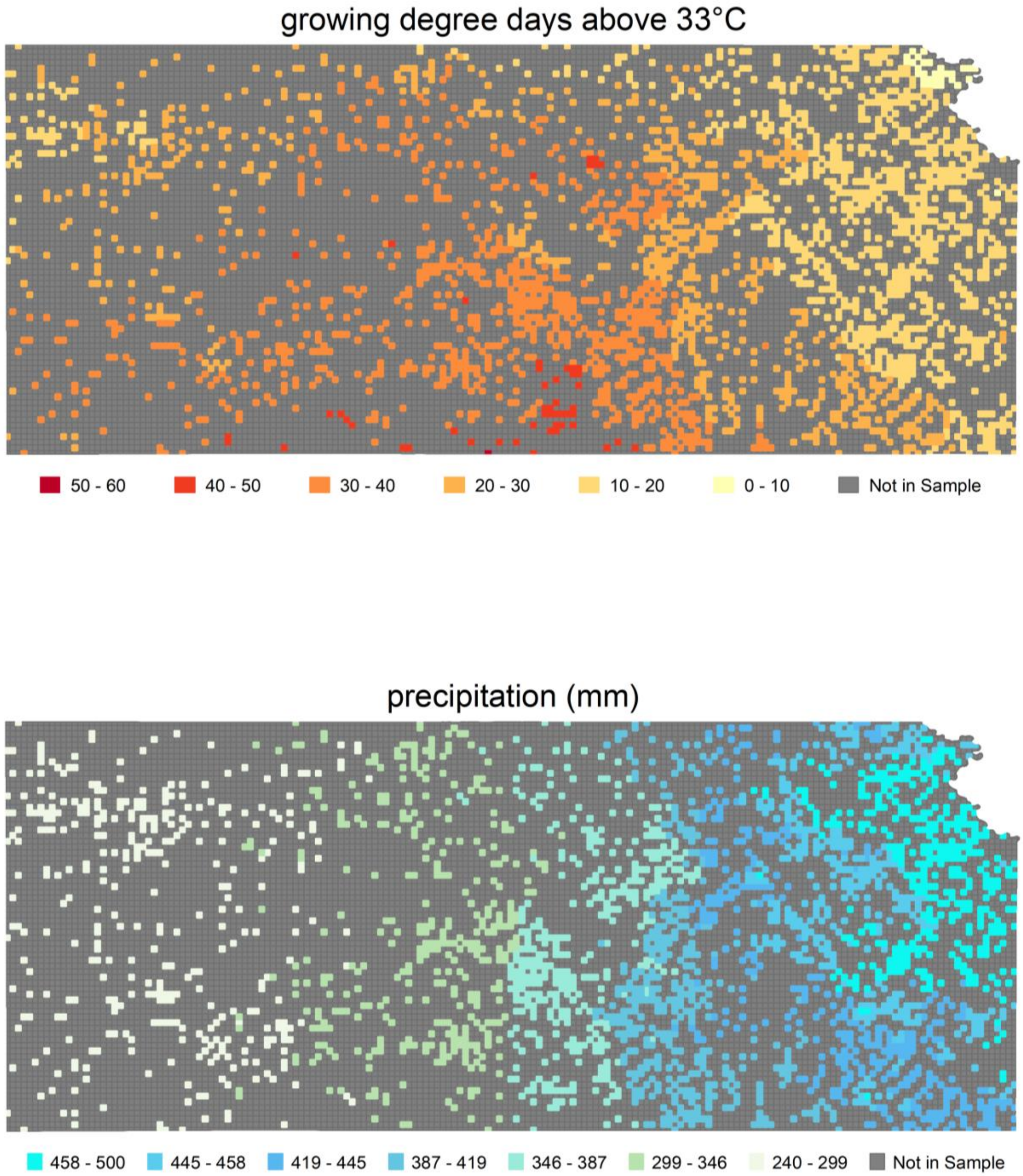

Figure 9 Average seasonal growing degree days above 33°C and precipitation for KFMA farms, with cooler and wetter weather in the northeast region of the state; warmer and dryer weather in the southwestern region of the state

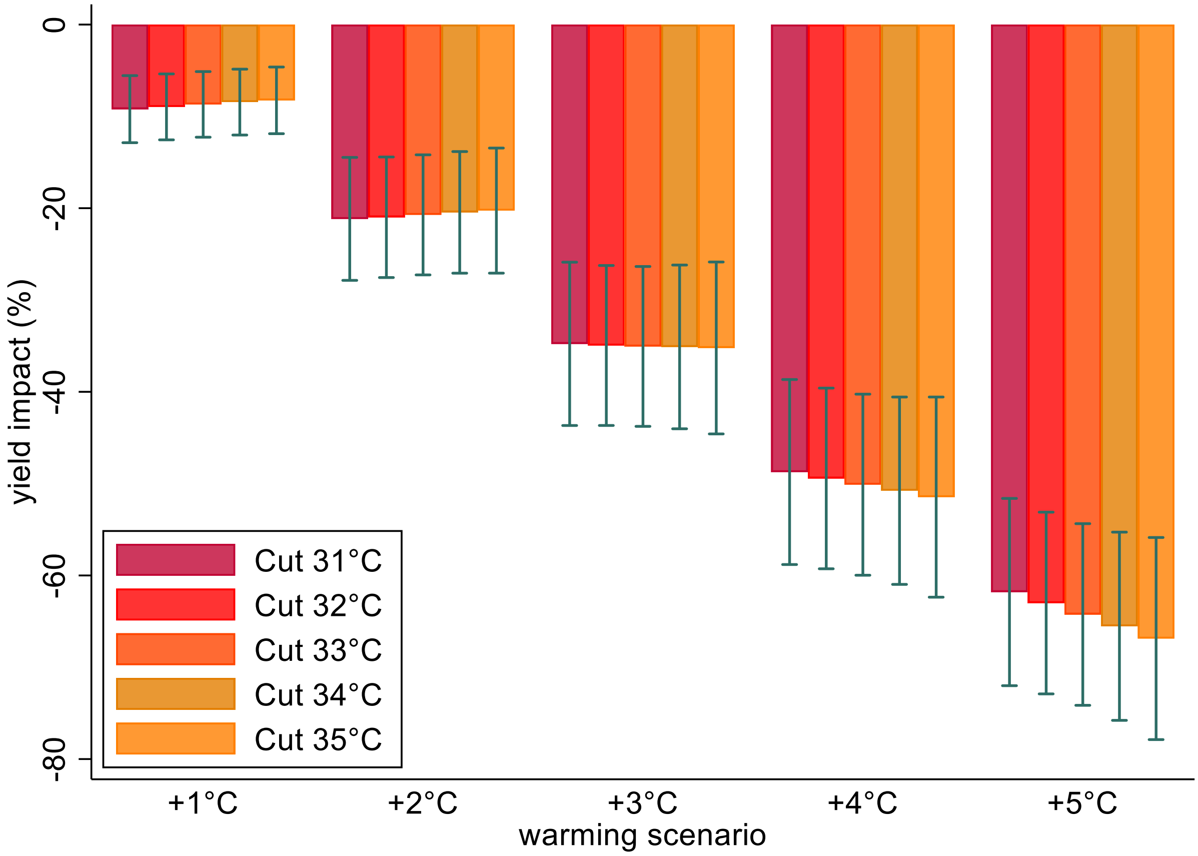

Figure 10 Predicted warming impacts on sorghum yields for alternative thresholds defining extreme heat

Note: Each five-bar cluster shows estimates for each warming scenario. The preferred model estimates impacts using 33ºC as the cut point.

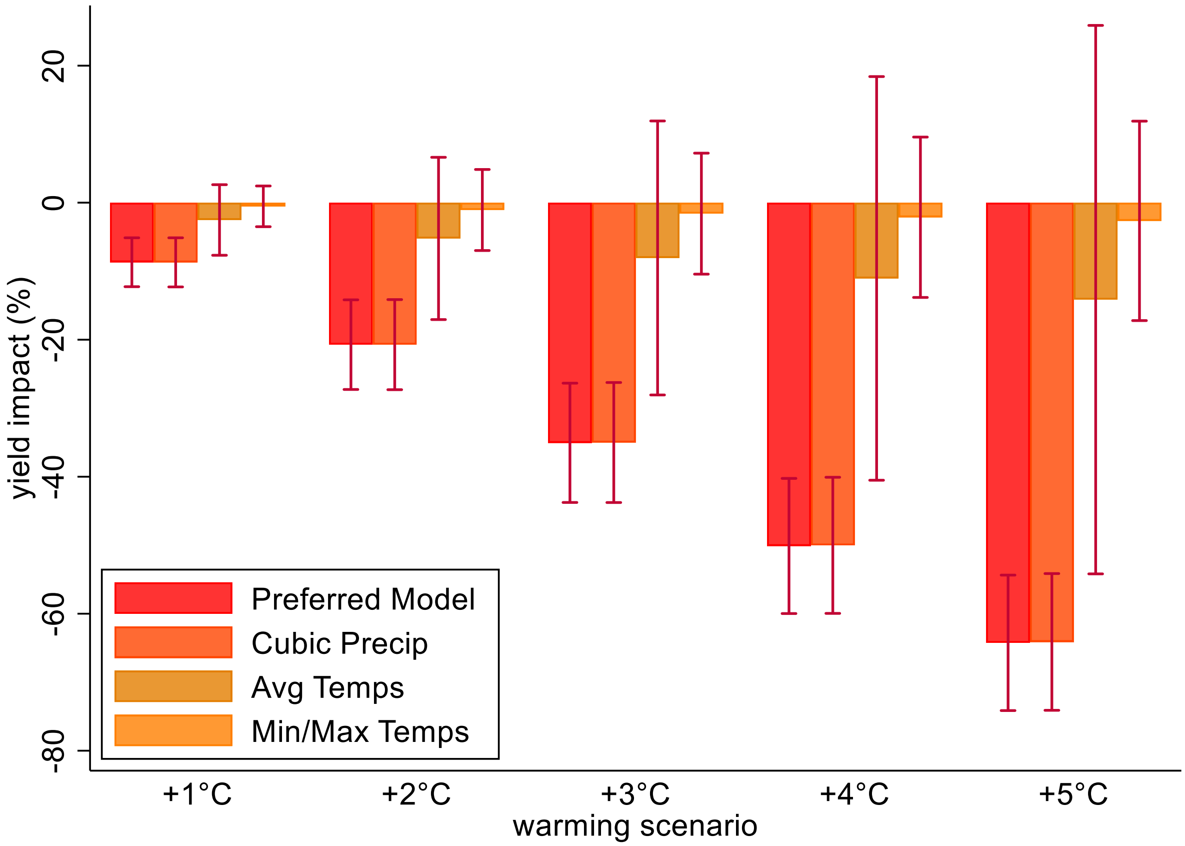

Figure 11 Predicted warming impacts on sorghum yields when using alternative precipitation and temperature specifications

Note: Each four-bar cluster shows estimates for each warming scenario. The preferred model estimates impacts using growing degree days and quadratic precipitation; alternative models include cubic precipitation or use average temperature or minimum and maximum temperatures in place of growing degree days.

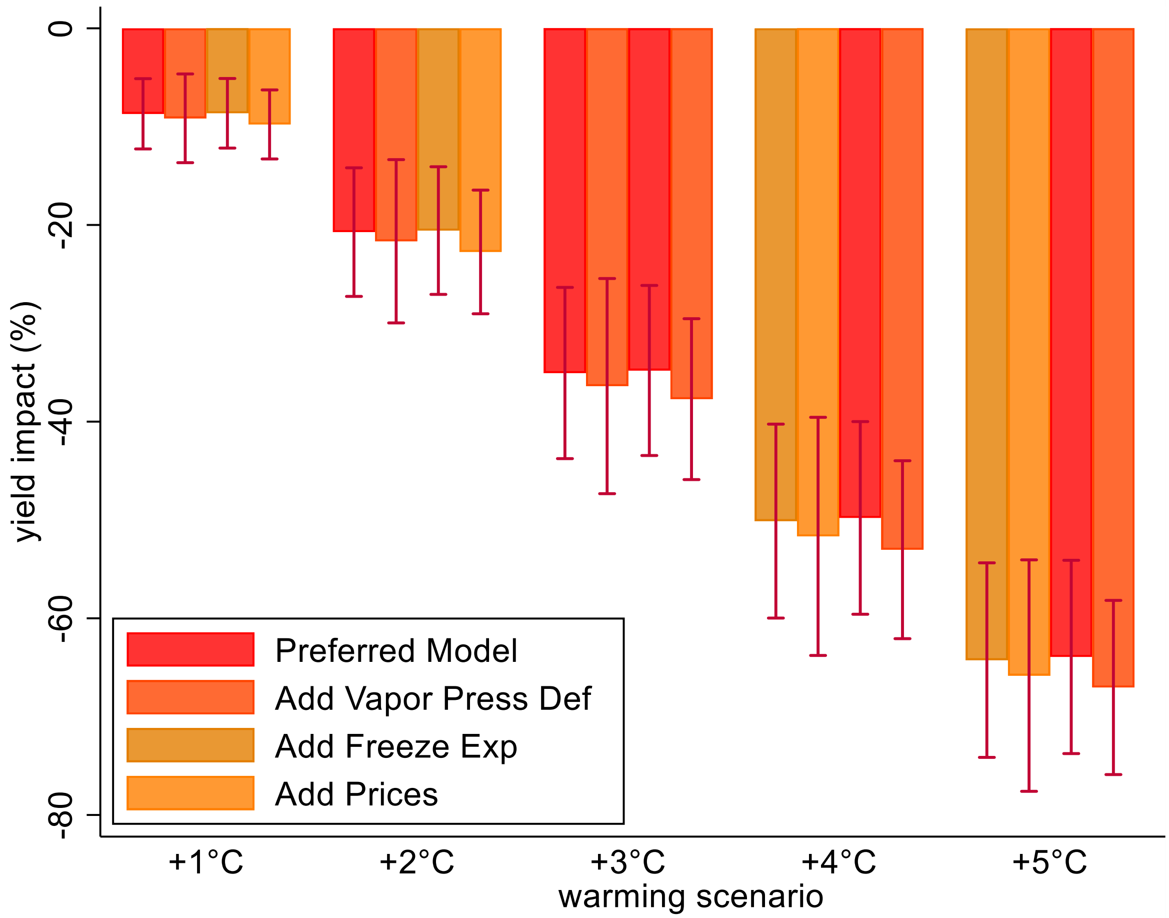

Figure 12 Predicted warming impacts on sorghum yields are robust when adding additional control variables

Note: Each four-bar cluster shows estimates using (i) the preferred specification, and specifications that add either (ii) vapor pressure deficit, (iii) time exposure below 0°C (i.e. freeze), or (iv) output-input price ratios.

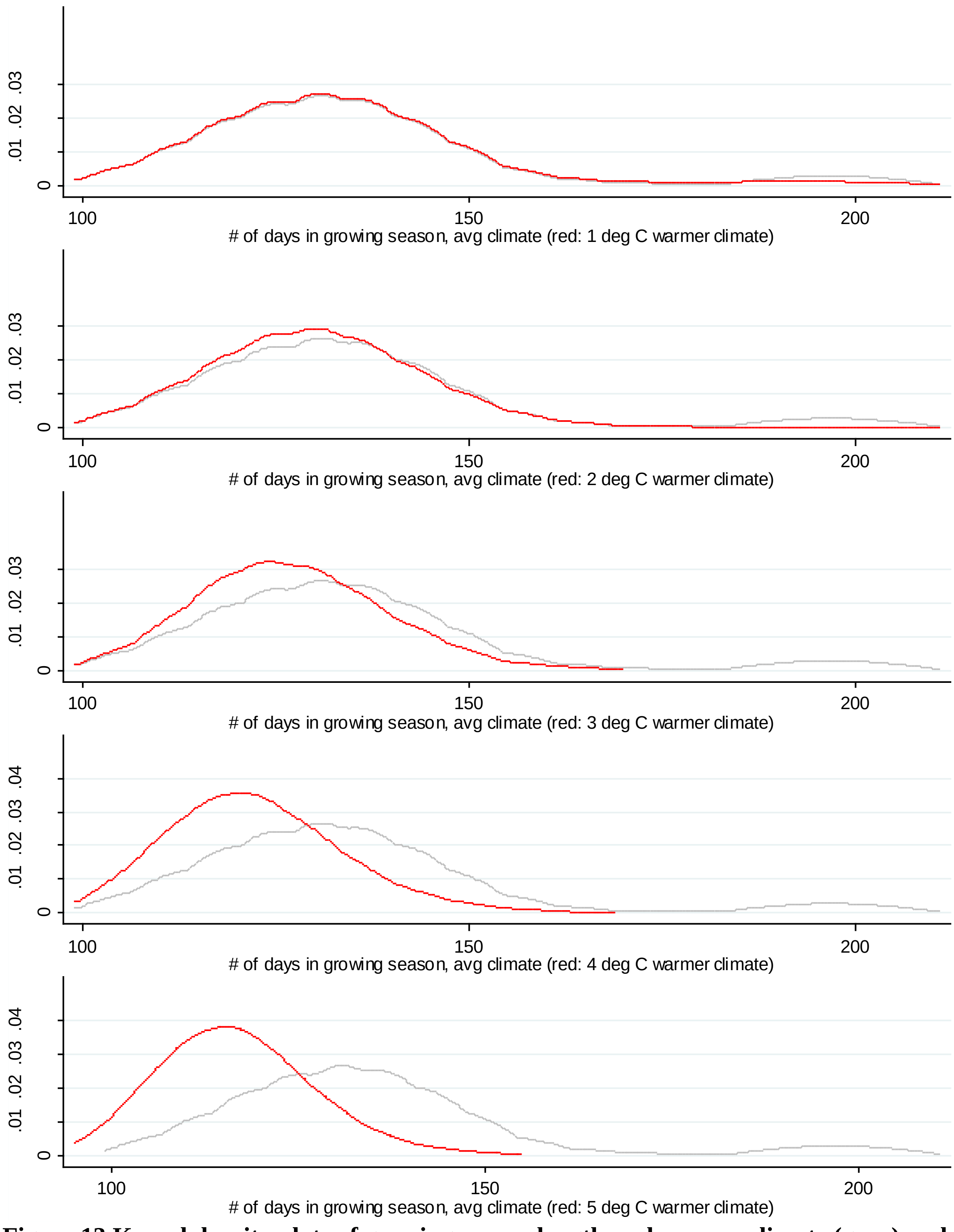

Figure 13 Kernel density plots of growing season length under mean climate (gray) and warmer than average climate (red)

Note: Capping growing season length based upon the historically observed maximum degree days over 5º C leads to progressively shorter growing seasons as warming increases in severity.

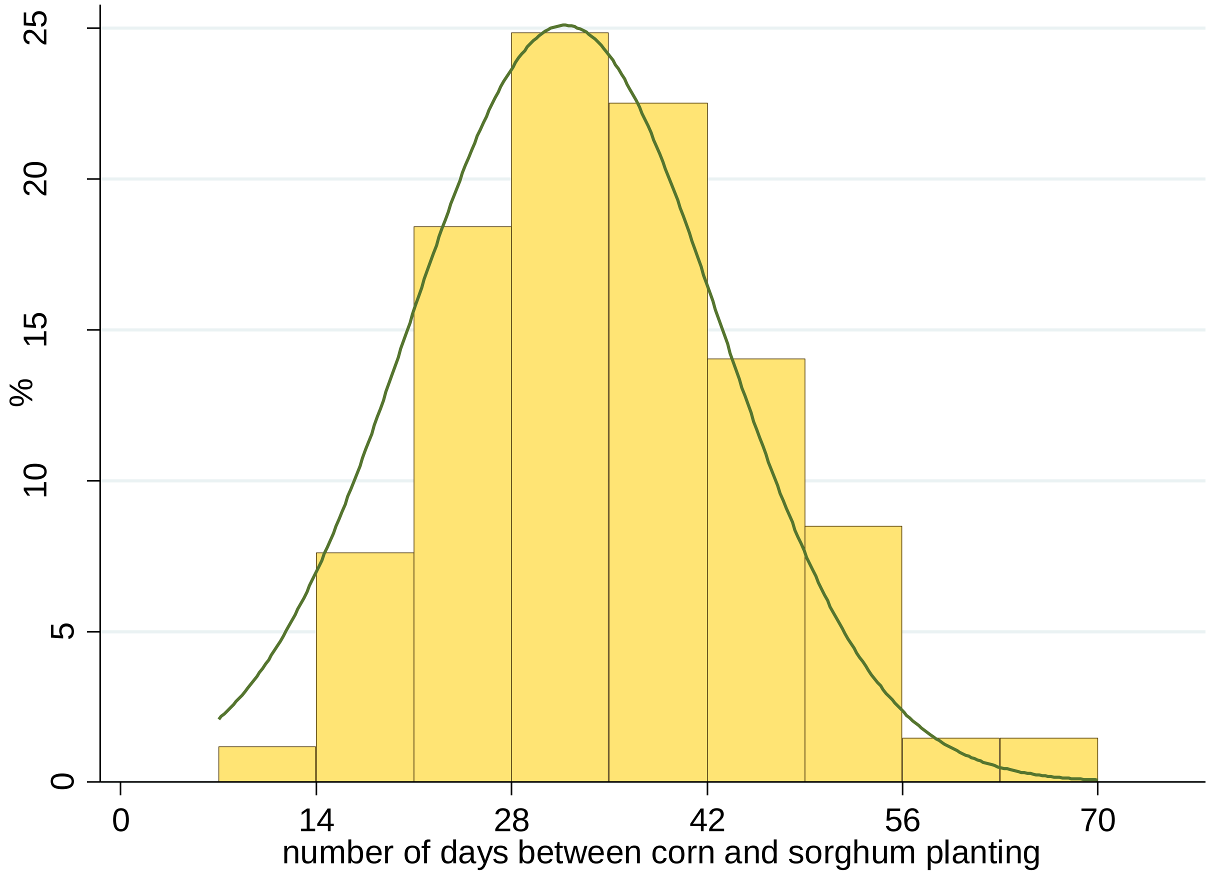

Figure 14 The distribution of the difference in planting of corn and planting of sorghum across all nine CRDs for the study time period, 1978 – 2015

Note: The average difference in planting dates is 32 days.

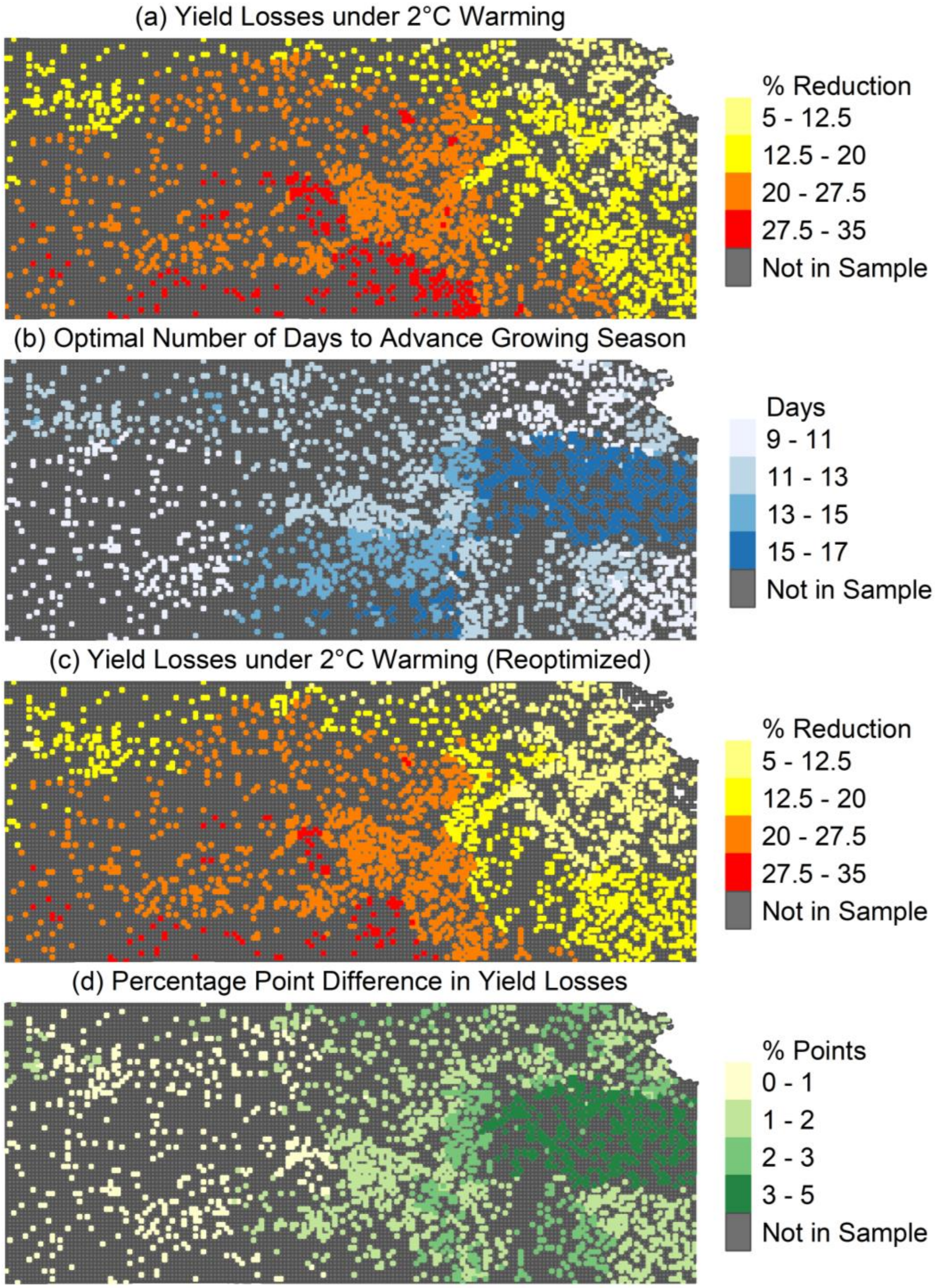

Figure 15 Predicted warming impacts on sorghum yields under 2°C warming when advancing the plant date and shortening the growing season

Note: Panel (a) shows estimated impacts at the grid cell-level under fixed plant and harvest dates (baseline). Panel (b) shows the optimal number of days to advance the growing season when we advance the plant date, shortening the growing season based upon degree day accumulation above 5ºC (i.e. the second approach). Allowing for reoptimization of the growing season reduces predicted warming impacts (Panel (c)). The percentage point difference between warming impacts (baseline and reoptimized growing season; Panel (d)) shows a high degree of spatial heterogeneity.

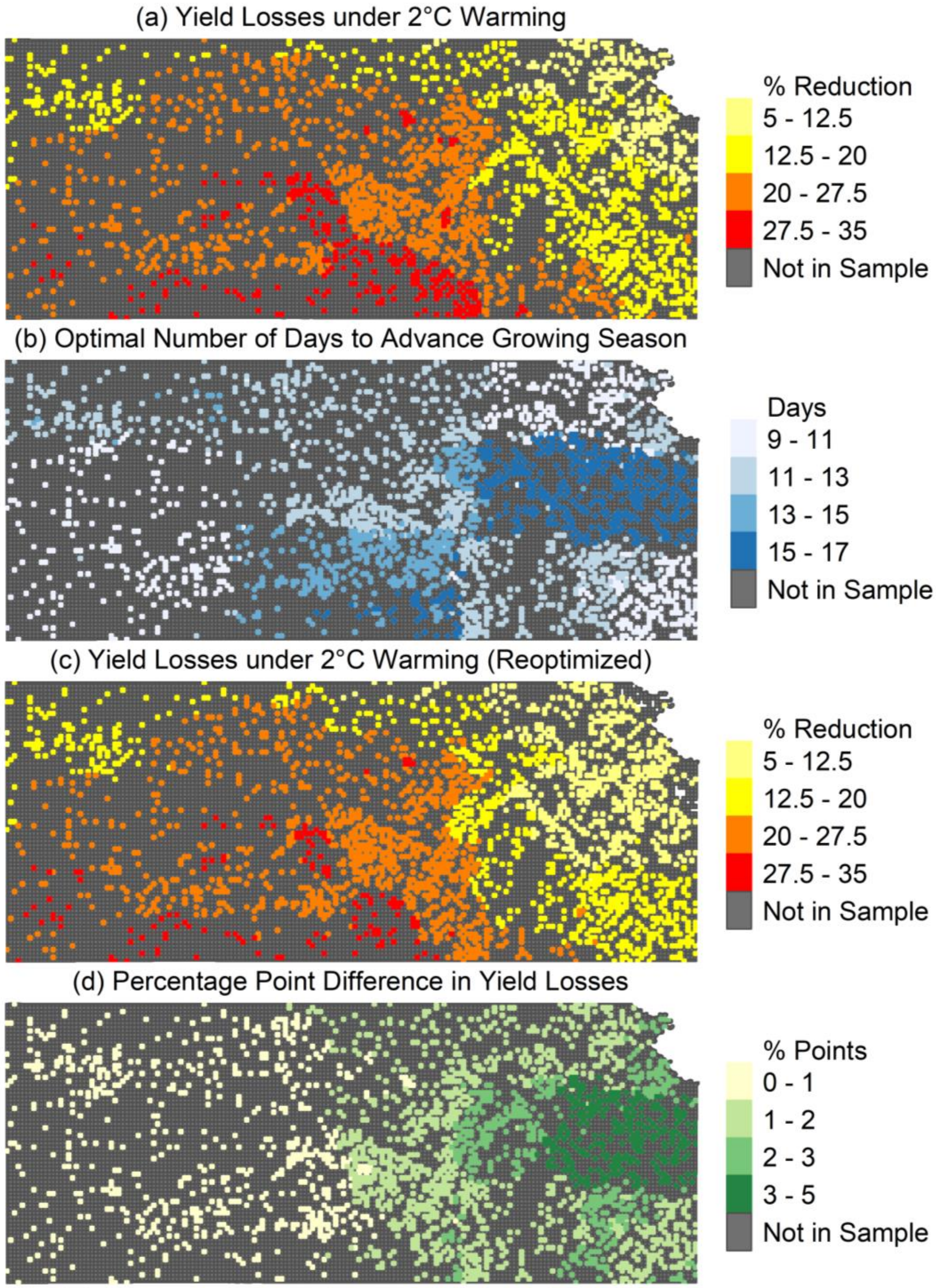

Figure 16 Predicted warming impacts on sorghum yields under 2°C warming advancing plant dates, keeping harvest dates fixed

Note: Panel (a) shows estimated impacts at the grid cell-level under fixed plant and harvest dates (baseline). Panel (b) shows the optimal number of days to advance the growing season when we advance the plant date, keeping harvest date fixed (i.e. the third approach). Allowing for reoptimization of the growing season reduces predicted warming impacts (Panel (c)). The percentage point difference between warming impacts (baseline and reoptimized growing season; Panel (d)) shows a high degree of spatial heterogeneity.

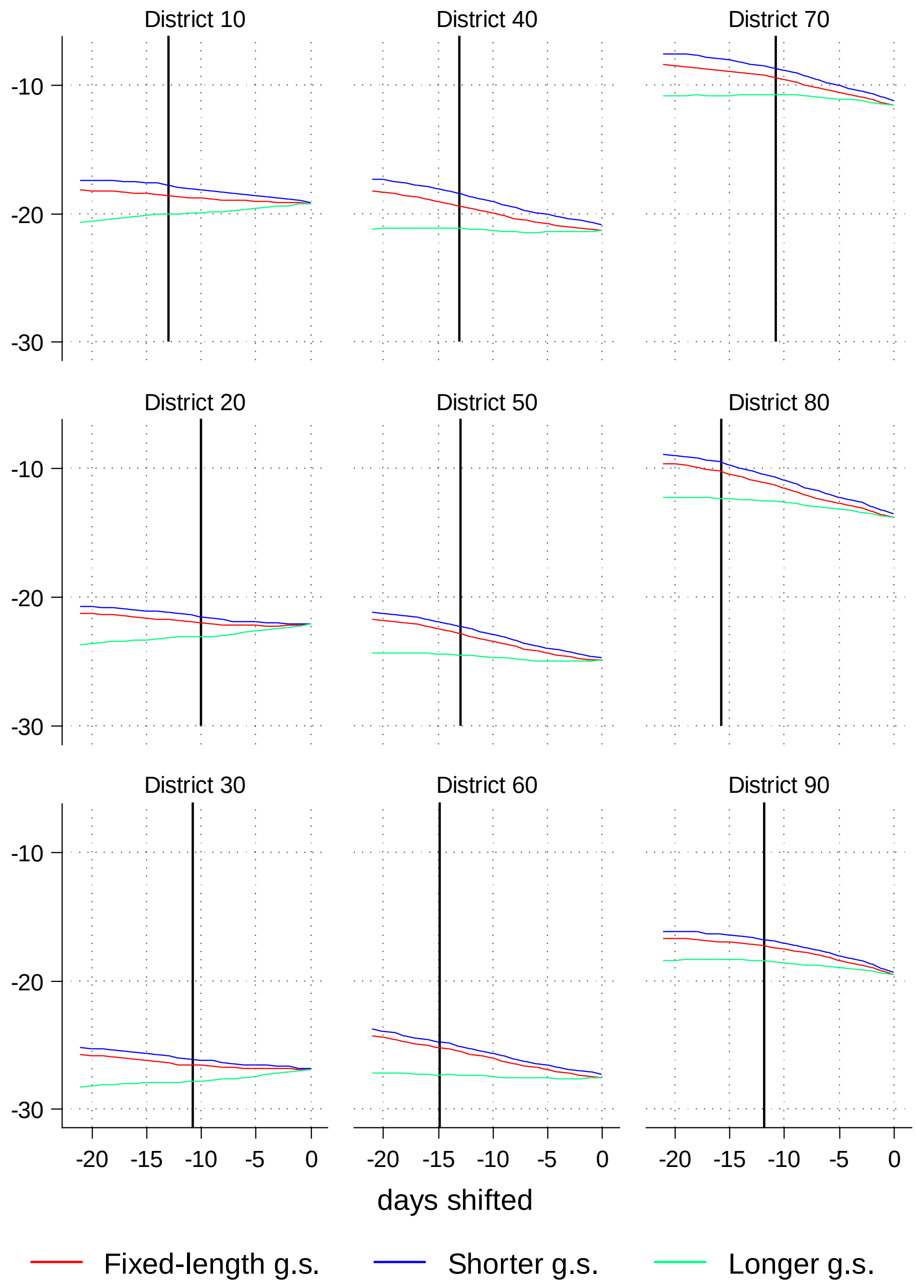

Figure 17 Predicted warming impacts on sorghum yields under 2ºC warming for various growing season adjustments (when controlling for freeze)

Note: Each red line shows the acreage-weighted average change in yield by advancing the growing season earlier, with both plant and harvest dates shifted by equal amounts (fixed-length g.s.); each blue line shows the acreage-weighted average change in yield by advancing the growing season earlier, allowing for shortening in growing season length based upon degree day accumulation above 5ºC (shorter g.s.); each green line shows the acreage-weighted average change in yield by advancing the growing season earlier, but keeping harvest date fixed (longer g.s.); each black line shows the district-wide average maximum number of days that planting can be advanced without violating sorghum's soil germination temperature constraint.

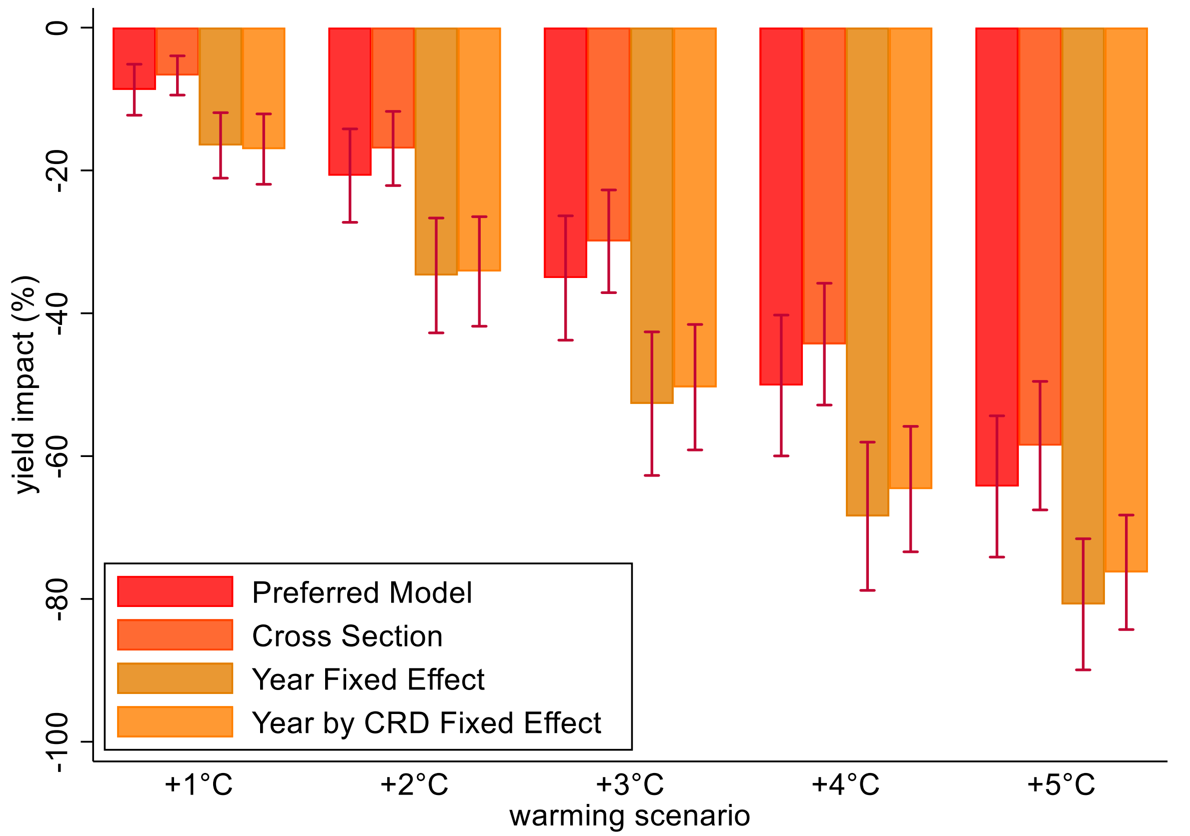

Figure 18 Predicted warming impacts on sorghum yields for cross-sectional and additional fixed effects regressions

Note: Each four-bar cluster shows estimates for each warming scenario. The preferred model estimates impacts using panel fixed-effects regression; the cross-section model collapses the panel and re-estimates the preferred model using averages of all covariates; the last two models replace the time trend with year fixed effects and year-by-CRD fixed effects.

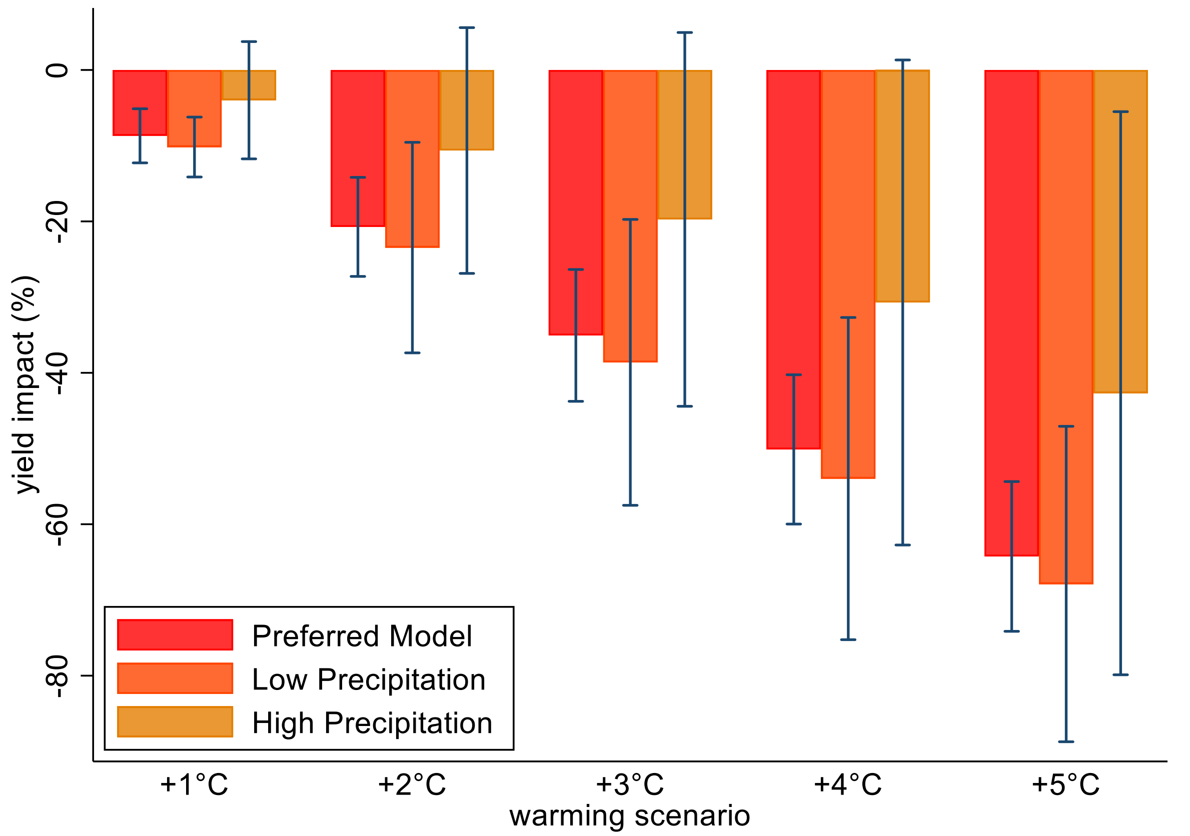

Figure 19 Predicted warming impacts on sorghum yields when including a heat-precipitation interaction term

Note: Each three-bar cluster shows estimates for each warming scenario. The preferred model estimates impacts under average precipitation; the low precipitation model and high precipitation models include a precipitation-extreme heat interaction term and estimate impacts under precipitation set respectively at 10% and 90% of the maximum historically observed precipitation in the sample data.

Chapter 2 - Forecasting Winter Wheat Harvest Basis using Soil Moisture Measurements

Introduction

Producers face a significant amount of price risk in marketing their grain at harvest. One way they can limit exposure to volatility in the cash markets is to ‘lock in’ prices with forward contracts, however, in so doing producers open themselves up to basis risk. Basis (defined as cash price minus futures price) levels have historically been relatively stable, due to similar co-movement patterns between cash and futures markets from year to year. The historic stability of interyear basis has allowed for simple harvest basis forecasts (in which next year’s harvest basis is solely a function of this and prior year’s basis) to be highly accurate (Hatchett et al. 2010). However, increased price volatility in cash and futures markets over the past fifteen years has been accompanied by a collapse in basis’ historic pattern (Adjemian et al. 2013; Taylor et al. 2014). Under these conditions, simple basis forecasts have performed poorly and have produced the need for forecasting models which can better handle the year-to-year oscillation in basis.

To meet this need we match soil moisture observations to historic elevator-level harvest basis to forecast basis using within growing-season weather. Here, soil moisture conditions around a grain elevator serve as a proxy for local supply (i.e. crop yields and production) which harvest basis is a function of. Specifically, we construct a panel model in which harvest basis is a defined as a function of prior harvest basis, elevator fixed-effects, and weekly, additive cubic soil moisture. We then compare the accuracy of our preferred model to a baseline model (in which weekly soil moisture readings are omitted) through an out-of-sample forecasting exercise. Our findings indicate that inclusion of weekly soil moisture in harvest basis forecasts made from mid-

April onward improves forecast accuracy by 14% for Kansas. We find that our preferred model is robust to inclusion of additional weather and price covariates and that model performance increases in years in which the initial ten weeks of the marketing week are abnormally dry or wet. When we expand the geographic scope of our analysis and try to forecast basis over larger, multi-state regions we find that the forecasting performance of our preferred model converges to the performance of the baseline model, suggesting that unsuitability of our model to forecasting across aggregated regions. Finally, we find that when we apply our preferred model to other individual states in the Great Plains (apart from Kansas) that forecasting accuracy over our baseline model increases dramatically (i.e. by 39%) as we move from north to south.

Research conducted prior to the onset of the recent surge in price volatility found (for a range of row crops) that the best predictor of future harvest basis was the prior year's basis or a historical moving average of prior years' harvest basis (Kastens et al. 1998; Garcia et al. 1986). A more recent study by Taylor et al. (2006) examined the effectiveness of incorporating current market information (in the form of basis deviation from the historical average) in harvest and post-harvest forecasts for Kansas corn, soybean, milo, and wheat – and found that this approach offered improvements, but only for post-harvest forecasts (for harvest forecasts the traditional approach using historical moving averages provided better forecasts). Thompson et al. (2019) similarly found that the accuracy of post-harvest (but not harvest) corn and soybean basis forecasting models for the eastern Corn Belt could be improved by including current basis information. Our research builds on this more recent work of Taylor et al. (2006) and Thompson et. al (2019) by (i) assessing alternative (i.e. weather) sources of within-season information regarding local markets that could potentially improve harvest basis forecasts, (ii) providing harvest basis forecasts for winter wheat for every possible weekly vantage point in the marketing year, and (iii) demonstrating the linkages between the geographic scope of the data and the accuracy of forecasting models.

Data

The data used in this analysis is divided into two types, price and weather data. Daily closing red hard winter wheat cash prices were collected from two sources: Cash Grain Bids (CGB) and Data Transmission Network (DTN) for years 2004 - 2018. The number of grain elevators total 858 and are spread across eight states (Texas, Oklahoma, Kansas, Colorado, Nebraska, Wyoming, South Dakota, and North Dakota) in the Great Plains Region (**Figure 20**). Wheat futures prices for July delivery at the Kansas City Board of Trade (KCBT) were obtained from Bloomberg for the same dates as the cash prices. **Figure 21** shows daily elevator-level basis prices were constructed by differencing futures prices from the corresponding cash prices.

In order to facilitate this analysis, the price data was manipulated in several ways. First, due to gaps in the daily price data, weekly basis prices were constructed by averaging daily prices across each five-day workweek (Monday through Friday) period. Because the number of weeks in any particular month alternated across years, in months that observed five weeks, the fifth week was averaged with the fourth week (resulting in 48 weeks for each year) (in keeping with the approach of Taylor et al. 2006). A marketing year comprised of 40 weeks (beginning the 1st week in September and ending the 4th week in June) was then defined and weekly basis prices that fell outside of this window were removed from the analysis. This left a total number of 549,120 total weekly basis price observations in the dataset.[8]

Weather data, was obtained from the European Space Agency (ESA) and the PRISM Climate Group (PRISM). The soil moisture data compiled for this analysis was taken from ESA's

Climate Change Initiative Soil Moisture (ESA CCI SM) product, which includes daily soil moisture (as a percentage) at a 27km x 27km resolution spanning the lower 48 states, from November 1st 1978 to December 30th 2018. Daily minimum and maximum temperatures at a 4km x 4km resolution were collected from PRISM, and interpolated following Schlenker and Roberts (2007) to form degree days. Using the degree day data, two degree day variables (i.e. degree days over 10ºC and degree days over 30ºC), were constructed as measures of moderate and extreme heat. Grain elevator addresses were geocoded and daily soil moisture and degree day data was collected for grids in which an elevator was located. The weather data was then manipulated in exactly the same manner as the price data, resulting in weekly average soil moisture and degree day readings for each week in each marketing year, 2004-2018. Despite the use of weekly averages, gaps in the soil moisture data remain. The majority of these gaps are due to weather conditions during the winter, including hard freezing of the soil and/or snow cover, which preclude satellite observation of soil moisture conditions. Boxplots of the key weather and price data for Kansas and for the Great Plains region that were used in this analysis are shown in **Figure 22, Figure 23, Figure 24, and Figure 25**.

Methods

We forecast harvest basis for winter wheat using the prior year's harvest basis and within season weekly soil moisture readings from every vantage point ($j = 1, 2, \ldots 40$) in the marketing year,

(2) $b_{it} = \alpha_i + \beta_0 b_{it-1} + \sum_{w=1}^{J} I_j \sum_{k=1}^{3} \gamma_{wk} sm_{wit}^{k} + \varepsilon_{it}, \quad (I_j = 1 \text{ if } w \leq j \text{ and 0 otherwise})$

where b_{it} is harvest basis at elevator *i* in year *t*; α_i are fixed effects that capture variation in basis caused by time invariant, elevator-specific characteristics; b_{it-1} is harvest basis at elevator *i* in the prior year (*t-1*); and sm_{wit} is the soil moisture reading in week *w* at elevator *i* in year *t*. We incorporate an indicator variable (I_j) into the model such that at any vantage point only the current and prior week's soil moisture readings are admitted additively. Given the likely nonlinear effect of soil moisture on price, we also include both quadratic and cubic functions (k =2,3) of soil moisture.[9] An outline of model **(2)** (hereafter denoted as the 'preferred' model) by specific vantage points can be found in **Appendix A**.

To evaluate forecasting performance of our preferred model, an out-of-sample forecasting exercise was conducted in which one year's worth of data was dropped from the dataset at a time and the model was estimated from all forty weekly vantage points with the remaining years' data. The resulting parameters were then used to predict out-of-sample harvest basis and calculate root-mean-squared-error (RMSE). A baseline model (the same as the preferred model only with average weekly soil moisture omitted) was similarly estimated and RMSE calculated.[10] This process was repeated $t = T - 1$ times – that is, until every year except the first year in the dataset had been forecasted. Since the dataset includes 14 (i.e. $t = 13$) years of weekly observations from all forty vantage points in the growing season, a total of 520 forecasts were produced.[11]

Results

The results are divided into three sections. In the first section, we compare the RMSE resulting from our out-of-sample exercise for the baseline and preferred models for the state of Kansas. We also compare the RMSE that results when our preferred model is altered (through changes in the specification of soil moisture, through inclusion of additional weather and price covariates, and through exclusion of portions of early and mid-season soil moisture readings). In the second section, we examine how the forecasting performance of our preferred model is effected by different initial weather conditions at the start of the growing season. That is, we classify forecast years as having abnormally dry or wet or normal starting soil moisture readings and estimate our models and rerun our out-of-sample exercise for each weather classification. In the third section we consider the generalizability of the preferred model when we (i) aggregate our data to include other Great Plains states and forecast basis for larger, multi-state regions, and when we (ii) produce state-level forecasts for states other than Kansas.

Soil Moisture Improves Model Performance for Kansas Elevators from Mid-April Onward

Parameter values and standard errors resulting from estimation of our baseline and preferred models for five of the forty possible vantage points (i.e. the 1st, 10th, 20th, 30th, and 40th week) are provided in **Appendix B**. We perform our out-of-sample forecasting exercise for our baseline and preferred models for all elevator locations in Kansas. We also prepare two alternative versions of our preferred model (one which drops the set of all weekly cubic soil moisture variables, and one which drops the set of all quadratic and cubic soil moisture variables) and rerun our out-of-sample-forecasting exercise. Average RMSE for all forty vantage points for the baseline, preferred, and two alternative soil moisture models are reported in columns 2-5 of **Table 11** and in **Figure 26**. Basis forecasts that include soil moisture perform worse relative to our baseline model for the early and mid-parts of the marketing year. From the 31st week (i.e. the 3rd week in April) onward however, RMSE of the models that include soil moisture drops. The magnitude of this drop depends on the polynomial degree of soil moisture – RMSE for the preferred model (that includes linear, quadratic, and cubic soil moisture variables) drops by 25%; RMSE for the alternative model that includes only linear and quadratic soil moisture variables drops by 22%, and RMSE for the alternative model that includes only linear soil moisture variables drops by 18%. The reduction in RMSE is such that from week 31 onwards, both the alternative model that includes linear and quadratic soil moisture and the preferred model outperform the baseline model (at week 31, the preferred model outperforms the baseline by 14%).[12]

Inclusion of Additional Price and Weather Covariates Does Not Improve Model Performance

We examine the degree to which our preferred model can be improved upon through inclusion of additional price and weather covariates. We construct three alternative models. The first alternative model is identical to our preferred model except for an additional price variable that captures the weekly average Kansas City Board of Trade July futures contract for winter wheat. The second alternative model is identical to the preferred model except for two additional degree day variables that capture exposure to temperatures greater than 10ºC and 30ºC. The third alternative model is identical to the preferred model but includes both the futures price variable and the two degree day variables. We rerun our out-of-sample forecast using these three alternative specifications. Average RMSE for all forty vantage points for these alternative models are shown in columns 6-8 of **Table 11**. **Figure 27** also shows average RMSE by vantage point for the three alternative models, as well as average RMSE for the baseline and preferred models. The out-of-sample performance of the three alternative models is worse to both the baseline and preferred models for the first 30 weeks in the marketing year, with higher RMSEs reported for these alternative models at each vantage point. From week 31 onwards, the alternative models produce virtually identical RMSEs to the preferred model. Taken together, these results suggest that inclusion of additional weather and price covariates does not improve our preferred model's forecasting ability.

Exclusion of Early and Mid-Season Soil Moisture Negatively Impacts Model Performance

We have shown that on average our preferred model's forecasting ability is worse than the baseline model for the early and middle parts of the marketing year. Winter wheat's growth cycle is such that after initial root formation in the fall, the plant enters dormancy for the winter – meaning that the majority of winter wheat's water requirements occur in the final months leading up to harvest after dormancy is ended (Rogers 1997). Given this, we consider the tradeoffs of excluding soil moisture readings from the early and mid-part of the growing season with two additional alternative models. The first of these alternative models excludes all soil moisture terms for weeks 1 - 30 from all forecast vantage points (in practice this collapses our preferred model into to our baseline model for the first 30 weeks). The second alternative model excludes all soil moisture terms for weeks 11-30 from all forecast vantage points. Average RMSE for out-of-sample exercises for these alternative models are reported in columns 9 and 10 of **Table 11** and shown in **Figure 28.**

The first of these alternative models produces the same RMSE as the baseline for weeks 1 – 30; from week 31 onwards it performs worse than our baseline model for the majority of weeks and it performs worse than our preferred model for all of the remaining weeks. The second of these alternative models suffers from poor performance also. It produces the same RMSE as the preferred model for the first ten weeks; from week 11 onward however it performs worse than our preferred model for the majority of the remaining weeks (i.e. weeks 11 – 40); it also performs worse than our baseline model for all of the remaining weeks. The poor performance of these models indicates that including early and midseason soil moisture is crucial to improving forecast performance relative to the baseline, but (as demonstrated above) only when forecasting from a mid-April (i.e. week 31) vantage point onwards.

Forecast Performance of Preferred Model Increases During Abnormal Weather Years

The results we have presented so far have been average results across all years. How well does our forecasting model perform in abnormal weather years – in which the year being forecasted is abnormally dry or abnormally wet? We answer this question by classifying years as either being abnormally dry, normal, or abnormally wet and reproducing our average RMSE statistics conditional on weather classification. Our classification procedure is as follows. We calculate averages of soil moisture observations at elevator *i* in the year *t* for the first ten weeks in each marketing year (i.e. the average soil moisture observed between the first week in September and the second week in November), and denote this as sm_ten_{it}. Next, we calculate the average of sm_ten_{it} across all elevators for a given year and denote this as $sm\_ten_{\bar{\imath}t}$. Finally, we calculate the average of sm_ten_{it} across all elevators and all years and denote this as μ. Abnormally dry years are classified as years in which $sm\_ten_{\bar{\imath}t} < \mu - std.dev.(sm\_ten_{it})$, abnormally wet years are classified as years in which $sm\_ten_{\bar{\imath}t} > \mu + std.dev.\left(sm_{ten_{it}}\right)$, and normal years are classified as years that do not meet the criteria for being either abnormally dry nor abnormally wet.[13]

Average RMSE for our baseline and preferred models for abnormally dry, abnormally wet, and normal years are shown in **Table 12** and in **Figure 29**. Overall week-to-week changes in the preferred model's RMSE are similar across all three weather conditions – the preferred model performs better from vantage points at the end of the marketing year than it does in the early and middle parts of the year. However, there are two important differences in model performance between normal and abnormal weather year forecasts. Forecasts made (during abnormally dry or wet years) using the preferred model from the first seven weeks' vantage points are associated with RMSE very close to the baseline model - and in the case of forecasts made during abnormally dry years, the preferred model shows improvements in RMSE beyond the baseline for the majority of early weeks. We also see that during abnormally dry years the preferred model's forecasting performance relative to the baseline model at week 31 is much larger than during normal weather years (i.e. the preferred model generates an RMSE 27.45% lower than the baseline model in dry years, while the preferred model generates an RMSE 9.77% lower in normal years).

Forecast Performance of Preferred Model for Multi-State Regions Approaches the Baseline

We seek to understand how expanding the geographic scope of our analysis impacts the forecasting performance of our preferred model. We do this by constructing four increasingly large, multi-state elevator-level datasets (i.e. data drawn from: (i) Kansas and Nebraska; (ii) Kansas, Nebraska, and Oklahoma; (iii) Kansas, Nebraska, Oklahoma, and South Dakota; and (iv) all states in the Great Plains – Kansas, Nebraska, Oklahoma, South Dakota, North Dakota, Colorado, Texas, and Wyoming) and rerunning our out-of-sample forecast exercise with our baseline and preferred models.[14] Average RMSE for the baseline model and for the preferred model by region are reported in **Table 13** and shown in **Figure 30**. These results show that as the geographic scope of the data increases, the resulting RMSE produced by forecasting using the preferred model approaches the average baseline RMSE. This result holds across all weekly vantage point except for the earliest week vantage points.[15]

Forecast Performance of Preferred Model Increases in Southern Great Plains States

We also seek to understand how model performance is impacted when applying the preferred model for Kansas to other, individual state-level analyses. We do this by constructing datasets for each of the states in the Great Plains (other than Wyoming) and rerunning our out-of-sample forecast exercise with our baseline and preferred models. Average RMSE for the baseline model and for the preferred model by region are reported in **Table 14** and **Table 15** and shown in **Figure 31**, **Figure 32**, and **Figure 33**. Comparing the preferred model across states at week 31 of the marketing year shows a large discrepancy in performance relative to the baseline, with states south of Kansas showing a substantial increase in performance; conversely states north of Kansas show a decrease in performance. For example, the preferred model outperforms the baseline by 31% and 36% when forecasting basis for Texas and Oklahoma, respectively. In South Dakota the preferred model outperforms the baseline by only 7%, while in North Dakota, the preferred model performs worse than the baseline by 3%.

Conclusion

Recent literature has used time series methods to determine the value of including current market information, typically in the form of basis deviation from historical averages, on the accuracy of harvest basis forecasts. Our analysis seeks to complement this work in two ways: (i) by providing a fixed effects panel regression approach to forecasting basis, and (ii) by demonstrating how inclusion of another indicator of current market conditions (weekly soil moisture at grain elevator locations) impacts the accuracy of our harvest basis forecasts. Our results show that including soil moisture in basis forecasts leads to a decline in accuracy (as measured by RMSE) relative to a baseline model for the early and middle parts of the growing seasons, and to an increase in accuracy relative to the baseline model for the last part of the growing season. This suggests that soil moisture only impacts harvest basis after a sufficiently long period of time (i.e. 30 weeks or more) and that potentially larger temporal resolution (e.g. monthly or bimonthly) would be sufficient to capture the effect of local weather on elevator prices. We find the results of our model are robust to inclusion of additional weather and price covariates and across various early growing season conditions.

The performance of our preferred model is however sensitive to the geographic extent of our data. We find our preferred model produces more accurate (relative to the baseline) forecasts, in general, for states in the Southern Great Plains, and that aggregating up to multi-state regions produces forecasts that converge to the baseline. This result indicates that soil moisture is a good predictor for some but not all elevators and that the optimal forecasting model will vary across space. This is in keeping with the panel forecasting literature (Baltagi 2013) which argues that pooling observations at a large geographic scale can lower forecasting performance and that construction of models (e.g. spatial autoregressive error model) that explicitly account for spatial dependencies can improve performance. For now, we leave the tasks of determining the selection of the geographic scope of our data and comparing our preferred model to an explicitly spatial econometric model for future research.

We hope the results of our analysis can be useful to extension professionals, but we must advise caution in applying our preferred model to elevator-specific forecasts. As mentioned earlier, determination of the set of elevators' price data to include in training the model is critical to its forecasting accuracy. It is also very important to not assume that the preferred model can be used in forecasting basis at some point in the year other than at harvest and that our model can accurately forecast harvest basis for crops other than winter wheat. The implicit assumption underlying our analysis was that a growing season's soil moisture will have the largest impact on local supply, and thus the basis level, at harvest-time when crop production for a year is realized. Therefore, we recommend using our preferred model for end-of-season harvest basis forecasts and using already established models (i.e. a 1-year or 5-year moving average; Taylor et al. 2006) for early and mid-season harvest basis forecasts and for all post-harvest basis forecasts.

Footnotes

[8] The marketing year follows the growing season for winter wheat (planting occurs at the beginning of September and harvest occurs at the end of June) – and these terms are used interchangeably throughout the paper.

[9] Prior to estimation all covariates were demeaned – that is the levels of the covariates were subtracted by their averages over years (i.e. $\tilde{x}_{it} = x_{it} - \bar{x}_i$), such that fixed effects parameters were not directly estimated.

[10] Though Taylor et al. (2006) show that 5-year moving averages of prior years' basis are the most accurate for current harvest basis forecasts for winter wheat, Hatchett et al. (2010) conclude that: (i) the selected length of a moving average does not produce substantial differences in forecasted basis levels, and (ii) if there exists significant structural change across the forecast data, using the most recent basis observation is preferable to a moving-average approach.

[11] The first year's harvest basis (2005) could not be forecasted because the preferred model requires a prior year's harvest basis observation.

[12] Inclusion of higher degree soil moisture polynomials beyond cubic terms did not provide any noticeable further reductions in RMSE.

[13] Following this criteria, years 2011, 2013, and 2015 were classified as abnormally dry, years 2006, 2007, and 2010 were classified as abnormally wet, and years 2008, 2009, 2012, ,2014, 2016, 2017, and 2018 were classified as normal weather years.

[14] The sensitivity of our analysis to choice of states to include in our incrementally larger aggregated datasets is not one that is fully explored in the current paper – our choice was based on geographic proximity to our initial geographic area of interest (Kansas) and the average number of elevators observed by year (this information is reported in **Table 10**).

[15] Here the average baseline RMSE as reported in **Table 13** and **Figure 30** are the average RMSE for the baseline model across all forecasting years and across all datasets, however there was very little variation in this measure across the different datasets (the minimum and maximum baseline RMSE was 0.281 and the 0.286 respectively).

Tables

Table 10 Summary Statistics for Forecasting Data

Variable	State	Mean	Std Dev	Min	Max
# of elevators by year	*Colorado*	28.87	7.05	11	36
	Kansas	361.54	76.11	140	478
	Nebraska	101.16	14.97	11	112
	North Dakota	19.15	7.44	5	35
	Oklahoma	34.80	10.21	6	54
	South Dakota	54.22	11.91	10	70
	Texas	25.73	7.71	4	37
	Wyoming	3.56	1.10	1	5
# of years by elevator	*Colorado*	10.99	2.57	3	14
	Kansas	10.82	2.64	1	14
	Nebraska	10.78	2.31	1	14
	North Dakota	8.34	3.05	1	12
	Oklahoma	9.29	2.92	1	14
	South Dakota	10.66	2.30	2	14
	Texas	9.36	2.38	2	14
	Wyoming	9.47	2.57	5	12
Harvest basis ($)		-0.49	0.38	-2.25	2.18
KBOT July futures – week 1 ($)		6.30	1.59	3.47	8.84
KBOT July futures - week 9 ($)		6.17	1.51	3.38	9.29
KBOT July futures - week 29 ($)		6.19	1.83	3.22	9.95
KBOT July futures - week 37 ($)		6.07	1.50	3.32	9.06
Soil moisture - week 1		0.20	0.05	0.04	0.37
Soil moisture - week 9		2.49	0.06	0.04	0.37
Soil moisture - week 29		0.74	0.06	0.04	0.39
Soil moisture - week 37		0.21	0.05	0.04	0.35
10ºC dday - week 1		193.66	143.99	27.98	1141.26
10ºC dday - week 9		1023.26	771.74	148.81	6375.672
10ºC dday - week 29		1364.40	1052.71	148.81	8765.35
10ºC dday - week 37		2,199.33	1668.41	232.43	12,878.43
30ºC dday - week 1		7.05	9.61	0	103.07
30ºC dday - week 9		16.92	19.61	0	173
30ºC dday - week 29		17.16	20.15	0	173
30ºC dday - week 37		24.70	28.15	0	249.36

Table 11 Average RMSE for Baseline, Preferred, and Alternative Models (KS)

week	*baseline*	*sm*	sm^2	sm^3	sm^3	sm^3	sm^3 +fut.	sm^3	sm^3
1	0.286	0.284	0.285	0.285	0.299	0.307	0.328	0.286	0.285
2	0.286	0.292	0.299	0.299	0.304	0.322	0.323	0.286	0.299
3	0.286	0.3	0.307	0.309	0.313	0.33	0.336	0.286	0.309
4	0.286	0.31	0.321	0.321	0.316	0.339	0.338	0.286	0.321
5	0.286	0.318	0.333	0.335	0.333	0.37	0.365	0.286	0.335
6	0.286	0.333	0.339	0.336	0.34	0.367	0.366	0.286	0.336
7	0.286	0.341	0.342	0.344	0.353	0.369	0.376	0.286	0.344
8	0.286	0.347	0.359	0.356	0.37	0.372	0.383	0.286	0.356
9	0.286	0.354	0.372	0.37	0.385	0.383	0.397	0.286	0.37
10	0.286	0.359	0.376	0.373	0.388	0.378	0.396	0.286	0.37
11	0.286	0.37	0.393	0.392	0.419	0.394	0.415	0.286	0.37
12	0.286	0.365	0.394	0.39	0.418	0.393	0.414	0.286	0.37
13	0.286	0.359	0.378	0.374	0.406	0.383	0.408	0.286	0.37
14	0.286	0.359	0.377	0.374	0.41	0.384	0.413	0.286	0.37
15	0.286	0.364	0.385	0.381	0.413	0.39	0.418	0.286	0.37
16	0.286	0.362	0.365	0.362	0.384	0.37	0.392	0.286	0.37
17	0.286	0.361	0.365	0.358	0.382	0.367	0.389	0.286	0.37
18	0.286	0.356	0.357	0.354	0.377	0.361	0.383	0.286	0.37
19	0.286	0.342	0.34	0.335	0.355	0.345	0.364	0.286	0.37
20	0.286	0.339	0.338	0.332	0.352	0.342	0.361	0.286	0.37
21	0.286	0.34	0.34	0.332	0.352	0.342	0.363	0.286	0.37
22	0.286	0.344	0.34	0.332	0.352	0.342	0.363	0.286	0.37
23	0.286	0.345	0.345	0.337	0.357	0.346	0.366	0.286	0.37
24	0.286	0.342	0.341	0.331	0.351	0.34	0.361	0.286	0.37
25	0.286	0.353	0.352	0.342	0.363	0.348	0.37	0.286	0.37
26	0.286	0.356	0.347	0.338	0.36	0.343	0.364	0.286	0.37
27	0.286	0.362	0.345	0.335	0.351	0.342	0.36	0.286	0.37
28	0.286	0.369	0.346	0.334	0.347	0.341	0.355	0.286	0.37
29	0.286	0.363	0.34	0.328	0.327	0.332	0.333	0.286	0.37
30	0.286	0.367	0.344	0.331	0.33	0.333	0.333	0.286	0.37
31	0.286	0.301	0.268	0.247	0.254	0.25	0.257	0.26	0.303
32	0.286	0.304	0.273	0.251	0.251	0.254	0.255	0.271	0.312
33	0.286	0.305	0.268	0.25	0.252	0.249	0.252	0.287	0.319
34	0.286	0.312	0.27	0.251	0.252	0.25	0.252	0.303	0.32
35	0.286	0.304	0.261	0.246	0.249	0.245	0.248	0.282	0.305
36	0.286	0.309	0.264	0.248	0.249	0.248	0.248	0.303	0.316
37	0.286	0.311	0.264	0.248	0.248	0.248	0.247	0.316	0.313
38	0.286	0.309	0.262	0.246	0.252	0.245	0.251	0.33	0.314
39	0.286	0.31	0.264	0.248	0.256	0.248	0.256	0.339	0.317
40	0.286	0.307	0.264	0.249	0.257	0.25	0.257	0.346	0.303

Table 12 Avg. RMSE for Baseline and Preferred Models, by Avg. Initial Weather (KS)

week	*baseline (dry)*	*sm^3(dry)*	*baseline (wet)*	*sm^3(wet)*	*baseline (norm.)*	*sm^3(norm.)*
1	0.255	0.268	0.365	0.4	0.266	0.243
2	0.255	0.225	0.365	0.38	0.266	0.297
3	0.255	0.223	0.365	0.376	0.266	0.316
4	0.255	0.237	0.365	0.393	0.266	0.327
5	0.255	0.243	0.365	0.402	0.266	0.346
6	0.255	0.263	0.365	0.369	0.266	0.353
7	0.255	0.253	0.365	0.387	0.266	0.365
8	0.255	0.222	0.365	0.471	0.266	0.364
9	0.255	0.211	0.365	0.464	0.266	0.397
10	0.255	0.219	0.365	0.467	0.266	0.398
11	0.255	0.274	0.365	0.464	0.266	0.412
12	0.255	0.299	0.365	0.458	0.266	0.401
13	0.255	0.329	0.365	0.44	0.266	0.366
14	0.255	0.321	0.365	0.465	0.266	0.359
15	0.255	0.319	0.365	0.489	0.266	0.36
16	0.255	0.324	0.365	0.413	0.266	0.355
17	0.255	0.31	0.365	0.42	0.266	0.353
18	0.255	0.296	0.365	0.419	0.266	0.35
19	0.255	0.272	0.365	0.409	0.266	0.33
20	0.255	0.269	0.365	0.415	0.266	0.323
21	0.255	0.283	0.365	0.408	0.266	0.32
22	0.255	0.278	0.365	0.403	0.266	0.324
23	0.255	0.272	0.365	0.417	0.266	0.331
24	0.255	0.271	0.365	0.411	0.266	0.322
25	0.255	0.279	0.365	0.428	0.266	0.332
26	0.255	0.271	0.365	0.424	0.266	0.33
27	0.255	0.26	0.365	0.413	0.266	0.333
28	0.255	0.261	0.365	0.416	0.266	0.33
29	0.255	0.255	0.365	0.42	0.266	0.32
30	0.255	0.257	0.365	0.423	0.266	0.324
31	0.255	0.185	0.365	0.326	0.266	0.24
32	0.255	0.191	0.365	0.325	0.266	0.246
33	0.255	0.183	0.365	0.335	0.266	0.242
34	0.255	0.188	0.365	0.331	0.266	0.245
35	0.255	0.217	0.365	0.311	0.266	0.231
36	0.255	0.216	0.365	0.313	0.266	0.234
37	0.255	0.219	0.365	0.31	0.266	0.233
38	0.255	0.22	0.365	0.308	0.266	0.231
39	0.255	0.223	0.365	0.311	0.266	0.232
40	0.255	0.237	0.365	0.312	0.266	0.228

Table 13 Avg. RMSE for Baseline and Preferred Models, by Region

week	*avg. baseline*	*KS*	*KS+NE*	*KS+NE+OK*	*KSN+NE+OK+SD*	*all states*
1	0.284	0.285	0.283	0.285	0.288	0.291
2	0.284	0.299	0.293	0.295	0.295	0.298
3	0.284	0.309	0.304	0.305	0.303	0.306
4	0.284	0.321	0.317	0.318	0.318	0.32
5	0.284	0.335	0.327	0.327	0.327	0.328
6	0.284	0.336	0.326	0.326	0.326	0.329
7	0.284	0.344	0.332	0.334	0.331	0.336
8	0.284	0.356	0.345	0.343	0.333	0.334
9	0.284	0.37	0.355	0.352	0.34	0.342
10	0.284	0.373	0.356	0.353	0.34	0.342
11	0.284	0.392	0.367	0.366	0.348	0.347
12	0.284	0.39	0.362	0.361	0.343	0.34
13	0.284	0.374	0.348	0.347	0.331	0.328
14	0.284	0.374	0.348	0.347	0.332	0.329
15	0.284	0.381	0.351	0.351	0.335	0.332
16	0.284	0.362	0.337	0.339	0.324	0.323
17	0.284	0.358	0.337	0.339	0.324	0.323
18	0.284	0.354	0.336	0.336	0.322	0.322
19	0.284	0.335	0.328	0.328	0.317	0.318
20	0.284	0.332	0.328	0.327	0.317	0.319
21	0.284	0.332	0.327	0.326	0.316	0.318
22	0.284	0.332	0.327	0.326	0.318	0.319
23	0.284	0.337	0.329	0.327	0.319	0.321
24	0.284	0.331	0.326	0.325	0.317	0.32
25	0.284	0.342	0.335	0.333	0.323	0.325
26	0.284	0.338	0.327	0.326	0.32	0.322
27	0.284	0.335	0.326	0.326	0.32	0.323
28	0.284	0.334	0.329	0.33	0.322	0.324
29	0.284	0.328	0.322	0.323	0.316	0.318
30	0.284	0.331	0.323	0.326	0.318	0.32
31	0.284	0.247	0.267	0.269	0.279	0.285
32	0.284	0.251	0.272	0.273	0.283	0.289
33	0.284	0.25	0.268	0.269	0.28	0.287
34	0.284	0.251	0.272	0.274	0.279	0.286
35	0.284	0.246	0.266	0.269	0.275	0.282
36	0.284	0.248	0.269	0.273	0.277	0.283
37	0.284	0.248	0.27	0.273	0.278	0.285
38	0.284	0.246	0.268	0.271	0.276	0.281
39	0.284	0.248	0.27	0.273	0.278	0.283
40	0.284	0.249	0.271	0.273	0.28	0.286

Table 14 Avg. RMSE for Baseline and Preferred Models by State (KS, TX, OK, CO)

week	*KS*	*KS*	*TX*	*TX*	*OK*	*OK*	*CO*	*CO*
1	0.286	0.285	0.258	0.264	0.302	0.304	0.281	0.289
2	0.286	0.299	0.258	0.285	0.302	0.329	0.281	0.289
3	0.286	0.309	0.258	0.31	0.302	0.338	0.281	0.292
4	0.286	0.321	0.258	0.308	0.302	0.36	0.281	0.314
5	0.286	0.335	0.258	0.304	0.302	0.376	0.281	0.328
6	0.286	0.336	0.258	0.336	0.302	0.386	0.281	0.32
7	0.286	0.344	0.258	0.327	0.302	0.396	0.281	0.318
8	0.286	0.356	0.258	0.251	0.302	0.386	0.281	0.324
9	0.286	0.37	0.258	0.264	0.302	0.399	0.281	0.319
10	0.286	0.373	0.258	0.271	0.302	0.407	0.281	0.315
11	0.286	0.392	0.258	0.264	0.302	0.395	0.281	0.298
12	0.286	0.39	0.258	0.238	0.302	0.412	0.281	0.304
13	0.286	0.374	0.258	0.236	0.302	0.336	0.281	0.301
14	0.286	0.374	0.258	0.241	0.302	0.327	0.281	0.314
15	0.286	0.381	0.258	0.262	0.302	0.329	0.281	0.319
16	0.286	0.362	0.258	0.263	0.302	0.315	0.281	0.333
17	0.286	0.358	0.258	0.268	0.302	0.295	0.281	0.32
18	0.286	0.354	0.258	0.23	0.302	0.283	0.281	0.322
19	0.286	0.335	0.258	0.224	0.302	0.275	0.281	0.33
20	0.286	0.332	0.258	0.233	0.302	0.273	0.281	0.33
21	0.286	0.332	0.258	0.219	0.302	0.27	0.281	0.285
22	0.286	0.332	0.258	0.222	0.302	0.268	0.281	0.297
23	0.286	0.337	0.258	0.232	0.302	0.248	0.281	0.289
24	0.286	0.331	0.258	0.224	0.302	0.257	0.281	0.285
25	0.286	0.342	0.258	0.224	0.302	0.26	0.281	0.255
26	0.286	0.338	0.258	0.23	0.302	0.261	0.281	0.263
27	0.286	0.335	0.258	0.242	0.302	0.251	0.281	0.234
28	0.286	0.334	0.258	0.227	0.302	0.262	0.281	0.237
29	0.286	0.328	0.258	0.189	0.302	0.257	0.281	0.238
30	0.286	0.331	0.258	0.181	0.302	0.238	0.281	0.236
31	0.286	0.247	0.258	0.177	0.302	0.194	0.281	0.227
32	0.286	0.251	0.258	0.172	0.302	0.191	0.281	0.229
33	0.286	0.25	0.258	0.183	0.302	0.193	0.281	0.239
34	0.286	0.251	0.258	0.186	0.302	0.183	0.281	0.231
35	0.286	0.246	0.258	0.196	0.302	0.162	0.281	0.228
36	0.286	0.248	0.258	0.201	0.302	0.157	0.281	0.221
37	0.286	0.248	0.258	0.207	0.302	0.156	0.281	0.228
38	0.286	0.246	0.258	0.192	0.302	0.152	0.281	0.224
39	0.286	0.248	0.258	0.183	0.302	0.144	0.281	0.203
40	0.286	0.249	0.258	0.134	0.302	0.145	0.281	0.208

Table 15 Avg. RMSE for Baseline and Preferred Models, by State (KS, NE, SD, ND)

week	*KS*	*KS*	*NE*	*NE*	*SD*	*SD*	*ND*	*ND*
1	0.286	0.285	0.251	0.271	0.317	0.32	0.353	0.344
2	0.286	0.299	0.251	0.283	0.317	0.33	0.353	0.359
3	0.286	0.309	0.251	0.293	0.317	0.332	0.353	0.326
4	0.286	0.321	0.251	0.299	0.317	0.328	0.353	0.366
5	0.286	0.335	0.251	0.286	0.317	0.355	0.353	0.351
6	0.286	0.336	0.251	0.271	0.317	0.357	0.353	0.374
7	0.286	0.344	0.251	0.281	0.317	0.368	0.353	0.348
8	0.286	0.356	0.251	0.275	0.317	0.29	0.353	0.355
9	0.286	0.37	0.251	0.26	0.317	0.282	0.353	0.365
10	0.286	0.373	0.251	0.257	0.317	0.286	0.353	0.365
11	0.286	0.392	0.251	0.263	0.317	0.286	0.353	0.361
12	0.286	0.39	0.251	0.249	0.317	0.288	0.353	0.356
13	0.286	0.374	0.251	0.249	0.317	0.284	0.353	0.368
14	0.286	0.374	0.251	0.251	0.317	0.287	0.353	0.359
15	0.286	0.381	0.251	0.253	0.317	0.29	0.353	0.366
16	0.286	0.362	0.251	0.252	0.317	0.28	0.353	0.38
17	0.286	0.358	0.251	0.252	0.317	0.283	0.353	0.332
18	0.286	0.354	0.251	0.25	0.317	0.287	0.353	0.344
19	0.286	0.335	0.251	0.248	0.317	0.288	0.353	0.337
20	0.286	0.332	0.251	0.247	0.317	0.292	0.353	0.329
21	0.286	0.332	0.251	0.236	0.317	0.292	0.353	0.332
22	0.286	0.332	0.251	0.237	0.317	0.284	0.353	0.332
23	0.286	0.337	0.251	0.234	0.317	0.285	0.353	0.344
24	0.286	0.331	0.251	0.235	0.317	0.287	0.353	0.358
25	0.286	0.342	0.251	0.236	0.317	0.282	0.353	0.369
26	0.286	0.338	0.251	0.227	0.317	0.282	0.353	0.372
27	0.286	0.335	0.251	0.229	0.317	0.284	0.353	0.353
28	0.286	0.334	0.251	0.234	0.317	0.291	0.353	0.352
29	0.286	0.328	0.251	0.23	0.317	0.294	0.353	0.337
30	0.286	0.331	0.251	0.226	0.317	0.295	0.353	0.343
31	0.286	0.247	0.251	0.227	0.317	0.295	0.353	0.362
32	0.286	0.251	0.251	0.233	0.317	0.298	0.353	0.374
33	0.286	0.25	0.251	0.232	0.317	0.3	0.353	0.395
34	0.286	0.251	0.251	0.227	0.317	0.269	0.353	0.376
35	0.286	0.246	0.251	0.231	0.317	0.267	0.353	0.335
36	0.286	0.248	0.251	0.228	0.317	0.268	0.353	0.329
37	0.286	0.248	0.251	0.233	0.317	0.279	0.353	0.353
38	0.286	0.246	0.251	0.235	0.317	0.281	0.353	0.273
39	0.286	0.248	0.251	0.216	0.317	0.285	0.353	0.0966
40	0.286	0.249	0.251	0.214	0.317	0.288	0.353	1.21E-08

Figures

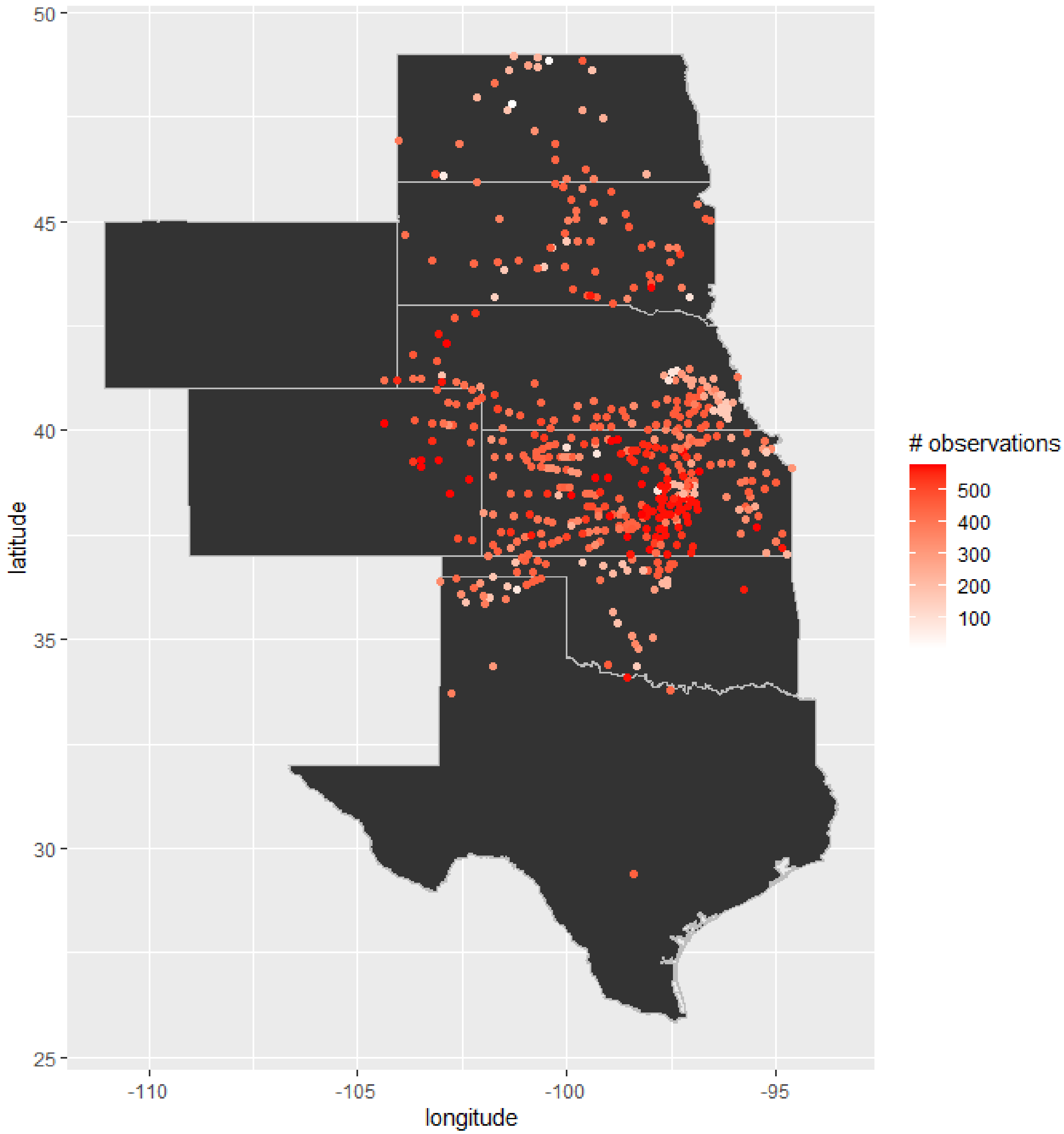

Figure 20 Number of Total Weekly Basis Observations by Elevator Location, Years 2005-2019

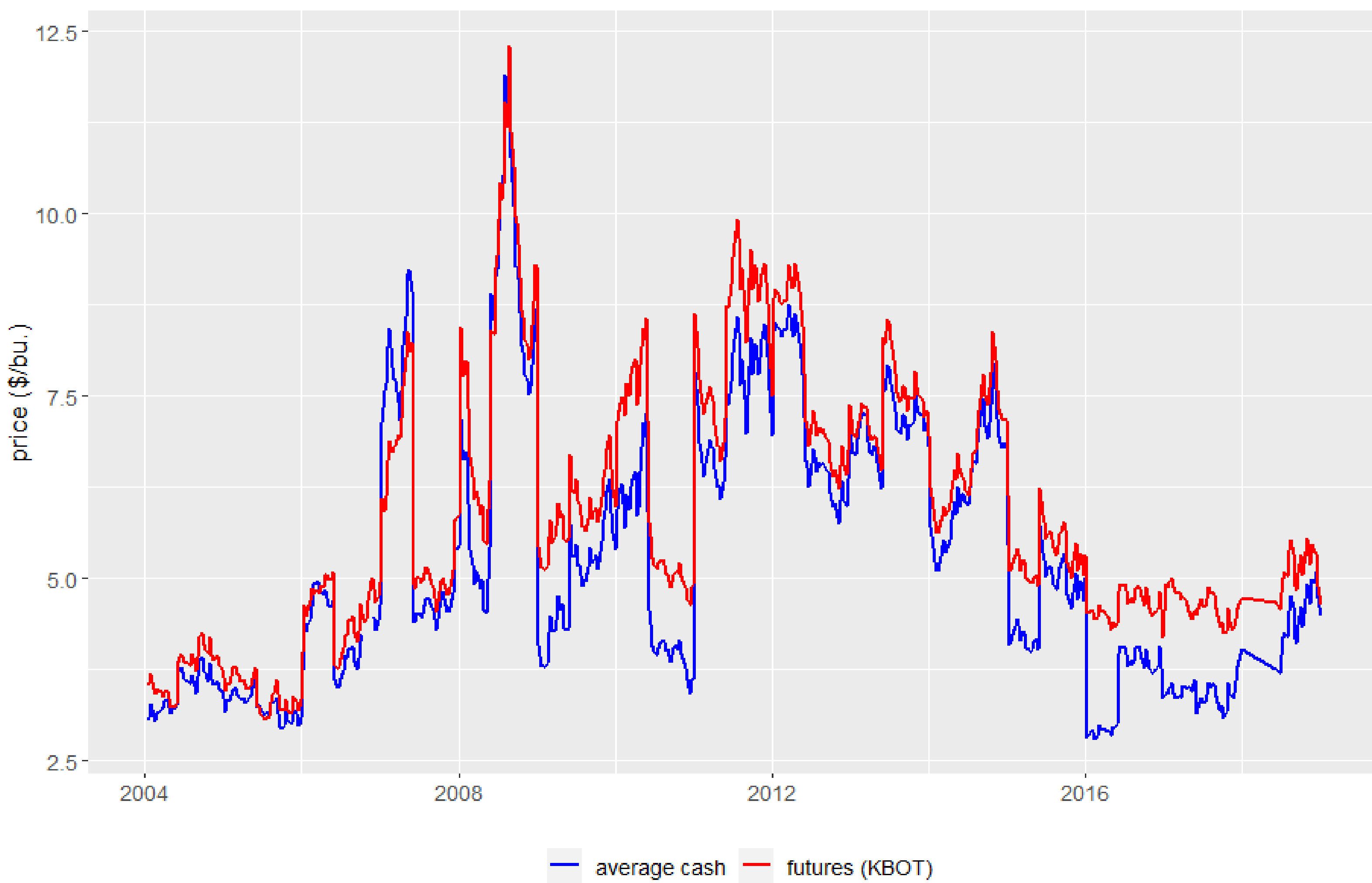

Figure 21 Avg. Kansas Cash and Kansas City Board of Trade (KBOT) July Futures Hard Red Winter Wheat Contract Prices, 2004-2019.

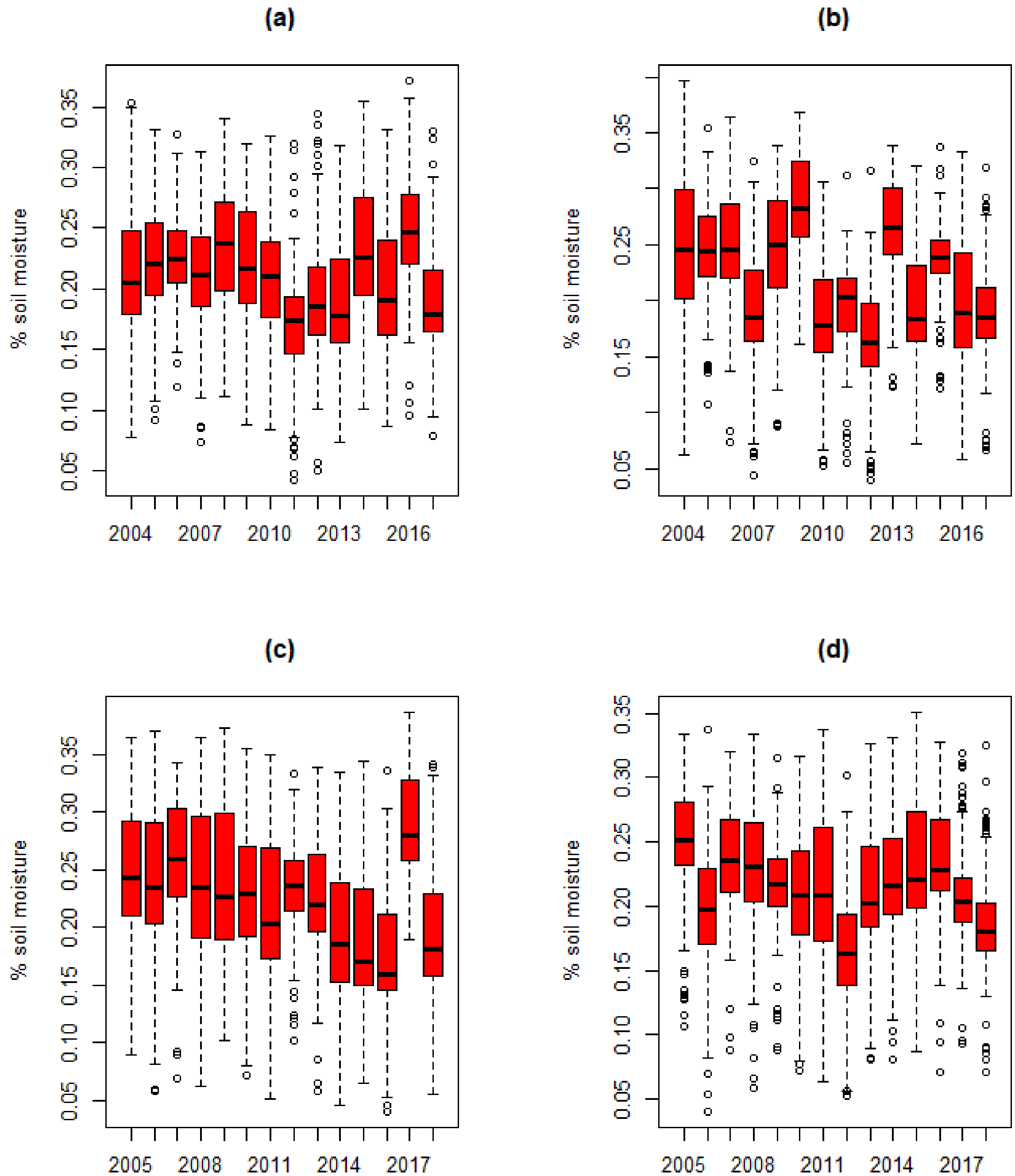

Figure 22. Boxplots of soil moisture at Kansas elevator locations, years 2004-2018, at different points in the marketing year.

Note: Figures 3a, 3b, 3c, and 3d show respectively soil moisture readings on the first week in September, November, April, and June.

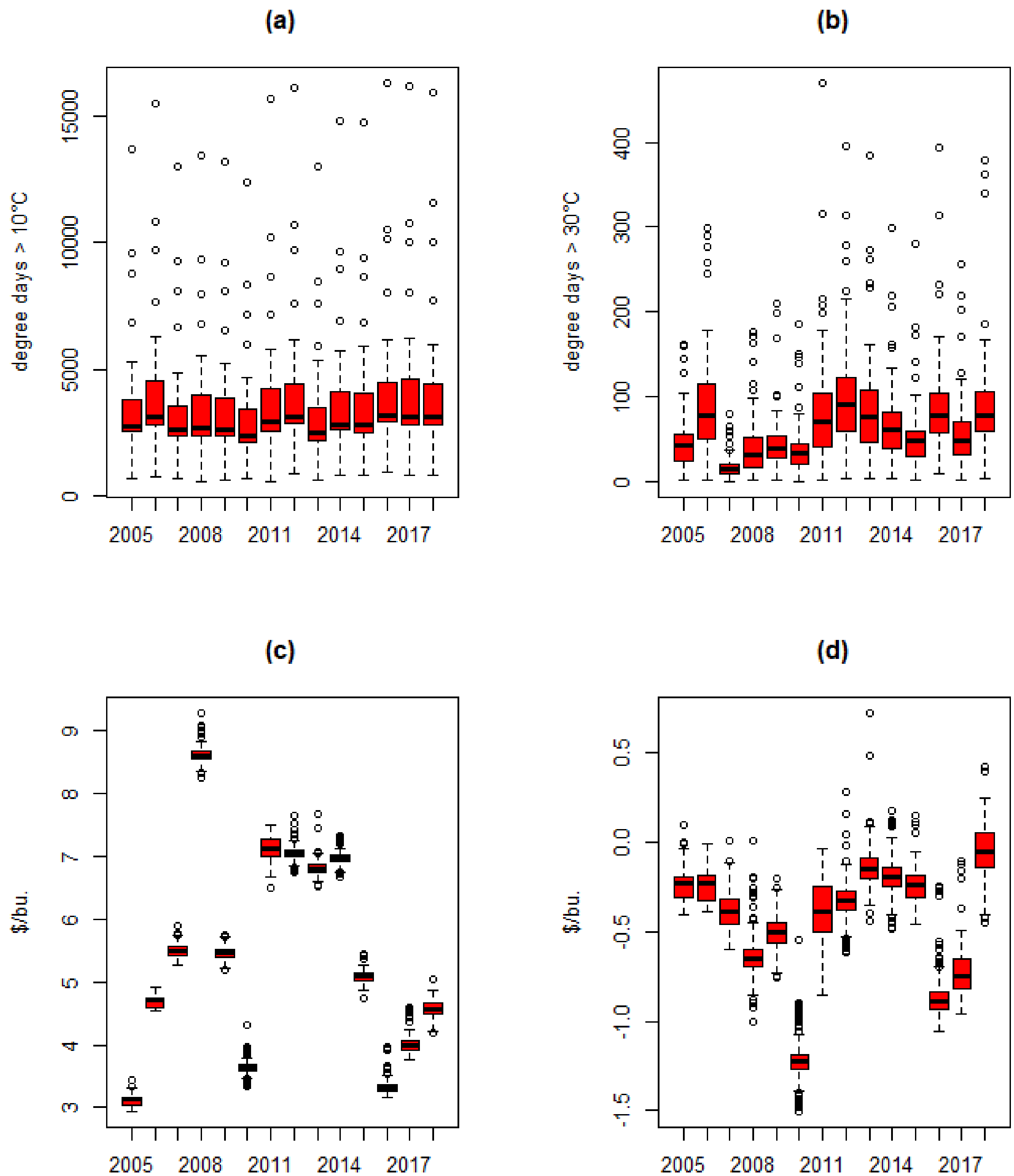

Figure 23. Boxplots of degree days and prices at harvest (the 40th week in the marketing year) at Kansas elevators, years 2005-2018.

Note: Figures 4a and 4b show respectively degree days in excess of 10ºC and degree days in excess of 30ºC; figures 4c and 4d show respectively spot prices and basis levels.

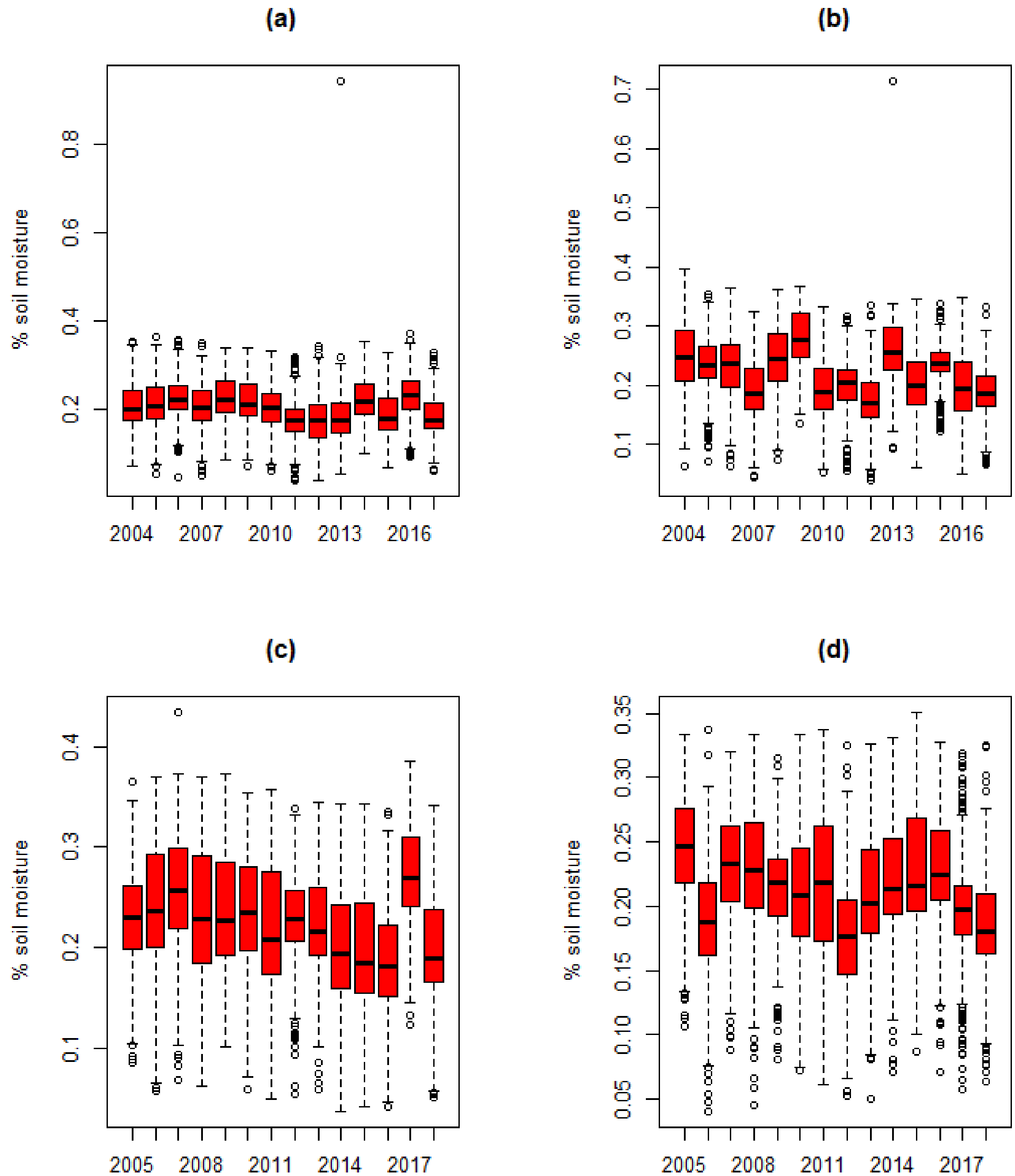

Figure 24 Boxplots of soil moisture at all Great Plains elevator locations, years 2004-2018, at different points in the marketing year

Note: Figures 5a, 5b, 5c, and 5d show respectively soil moisture readings on the first week in September, November, April, and June.

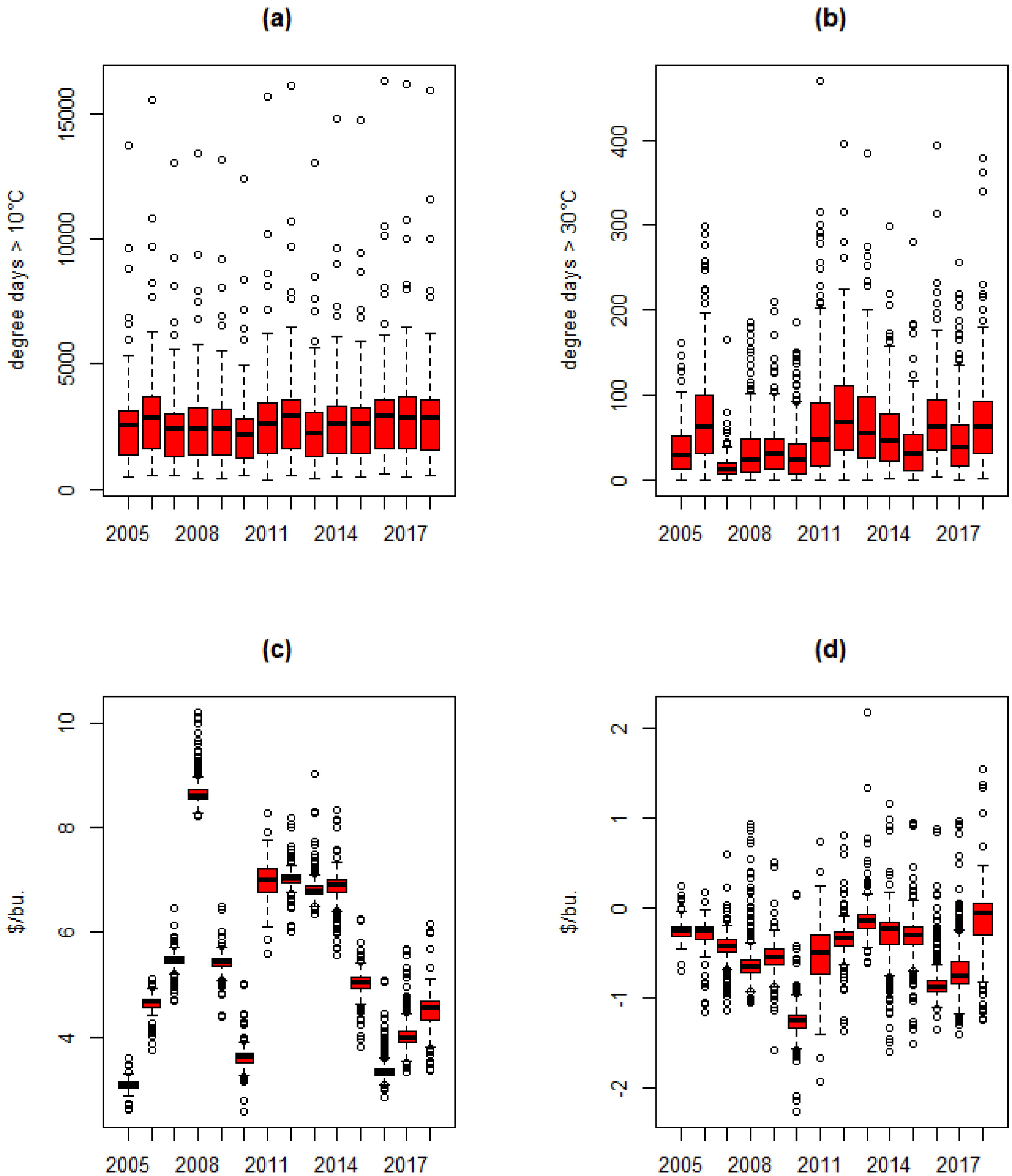

Figure 25 Boxplots of degree days and prices at harvest (the 40th week in the marketing year) at all Great Plains elevator locations, years 2005-2018.

Note: Figures 6a and 6b show respectively degree days in excess of 10ºC and degree days in excess of 30ºC; figures 6c and 6d show respectively spot prices and basis levels.

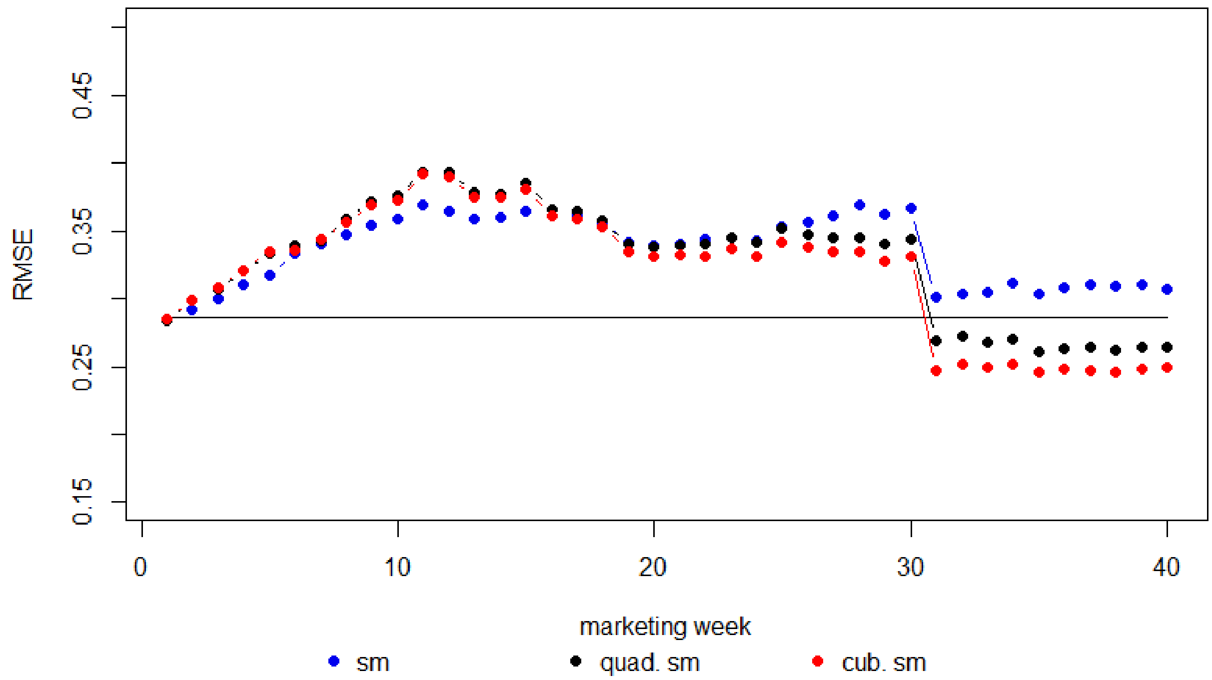

Figure 26 Out-of-sample performance (measured as RMSE) of the baseline model (the gray horizontal line) and models that include weekly soil moisture (the dotted lines) using Kansas elevator data.

Note: The dotted blue and black lines are the out-of-sample performance of models that add (respectively) linear and quadratic weekly soil moisture terms. The dotted red line is the out-of-sample performance of a model that adds cubic soil moisture terms (i.e. the 'preferred' model). Performance of these models improves substantially in the final weeks of the marketing year, with models that admit nonlinear soil moisture terms out-performing the baseline model from weeks 31 – 40.

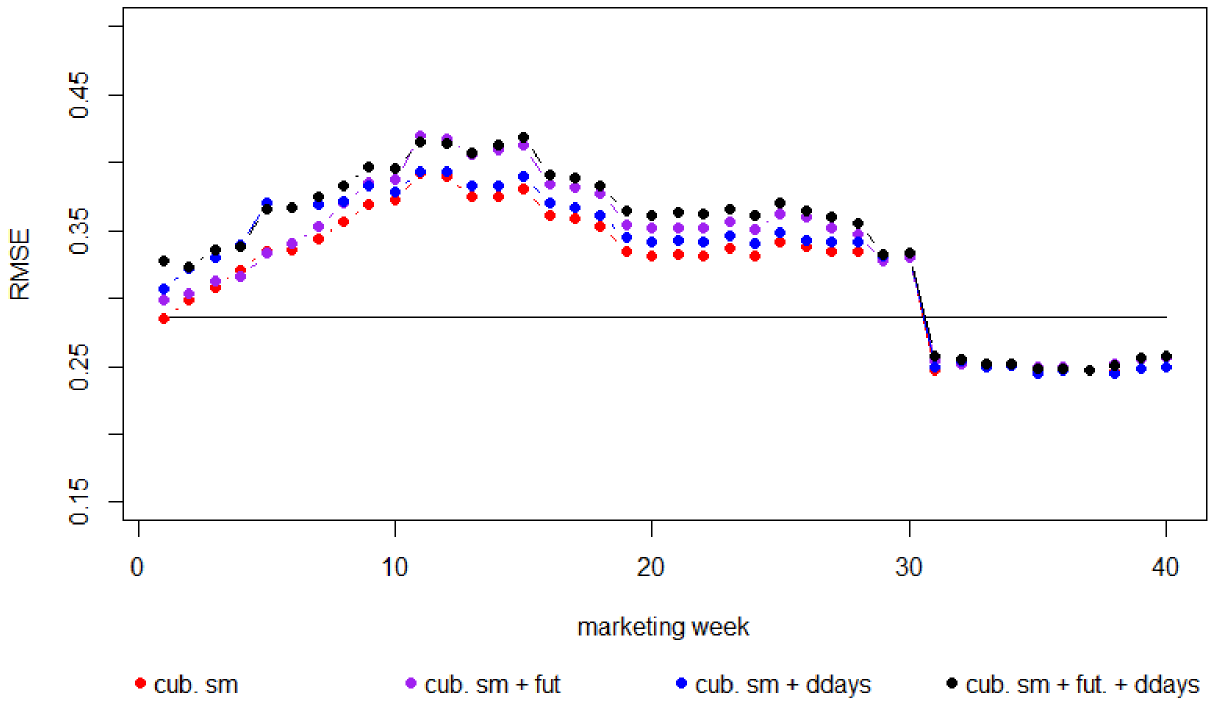

Figure 27 Out-of-sample performance (measured as RMSE) of the baseline model (the gray horizontal line), the 'preferred' model (the dotted red line), and alternative models (the other dotted lines) which admit additional weather and price covariates using Kansas elevator data.

Note: The model associated with the dotted purple line adds current KBOT July futures winter wheat price, the model associated with the dotted blue line adds two degree day variables (one measuring degree days greater than 10ºC and one measuring degree days greater than 30ºC), and the model associated with the dotted black line adds both the current KBOT July futures winter wheat price and the two degree day variables.

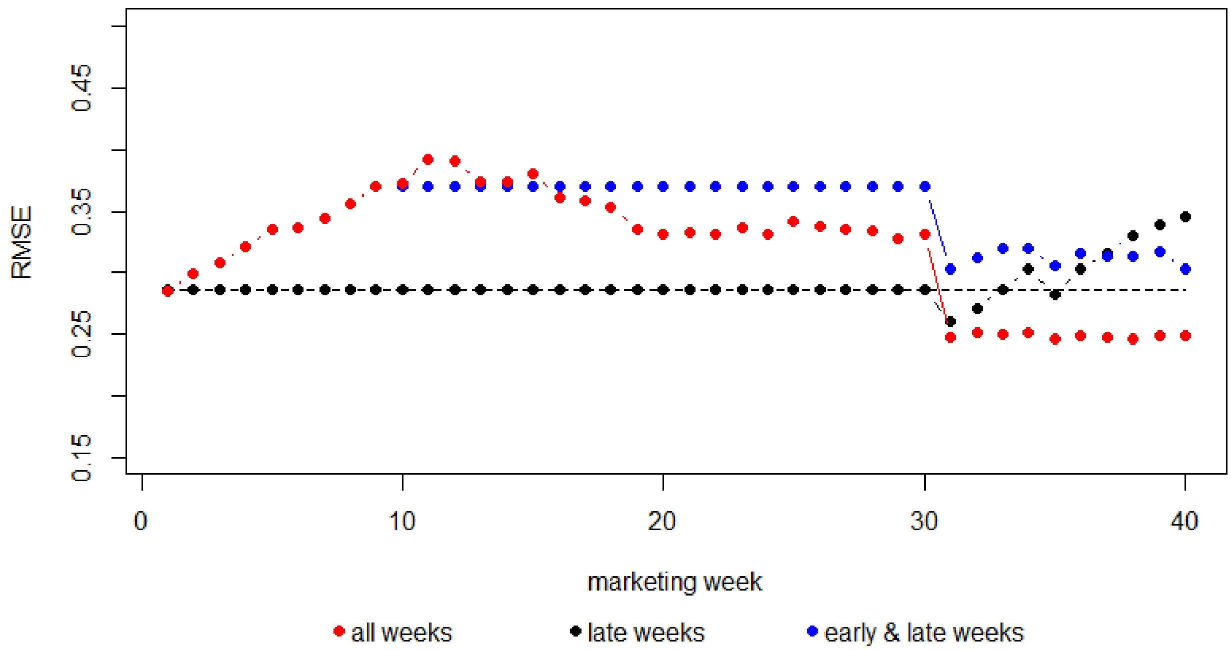

Figure 28 Out-of-sample performance (measured as RMSE) of the baseline model (the gray horizontal line), the preferred model (the dotted red line) and alternative versions of the 'preferred' model in which portions of weekly soil moisture are restricted (i.e. that is, they are dropped as covariates in estimation of the 'preferred' model) using Kansas elevator data.

Note: The model associated with the dotted black line is a version of the preferred specification in which only soil moisture from the last ten weeks of the marketing year are included. The model associated with the dotted blue line is a version of the preferred specification in which only soil moisture from the first and last ten weeks of the marketing year are included.

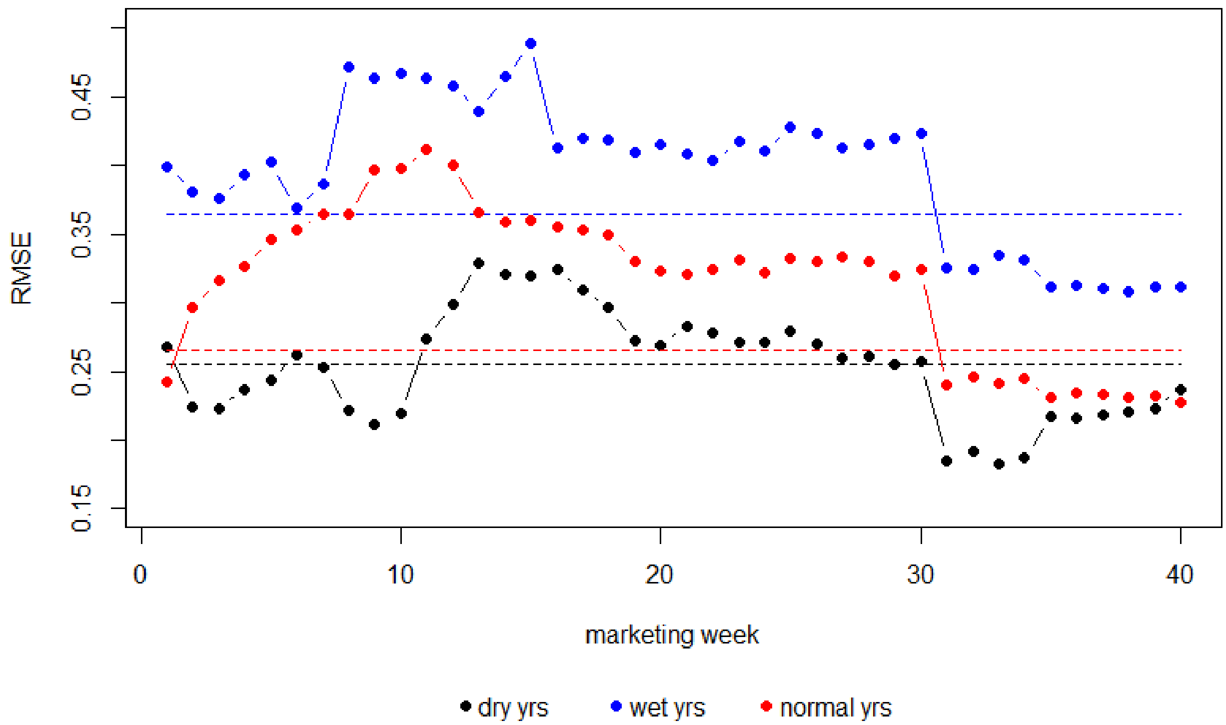

Figure 29 Out-of-sample performance (measured as RMSE) of the preferred specification (the dotted lines) relative to the baseline model (the dashed horizontal lines) under three different weather regimes using Kansas elevator data.

Note: The black dashed and dotted lines show respectively the baseline and the 'preferred' specification's out-of-sample performance under drier-than-average growing seasons; the blue dashed and dotted lines show respectively the baseline and the 'preferred specification's out-of-sample performance under wetter-than-average growing seasons; the red dashed and dotted lines show respectively the baseline and the 'preferred specification's out-of-sample performance under normal weather growing seasons.

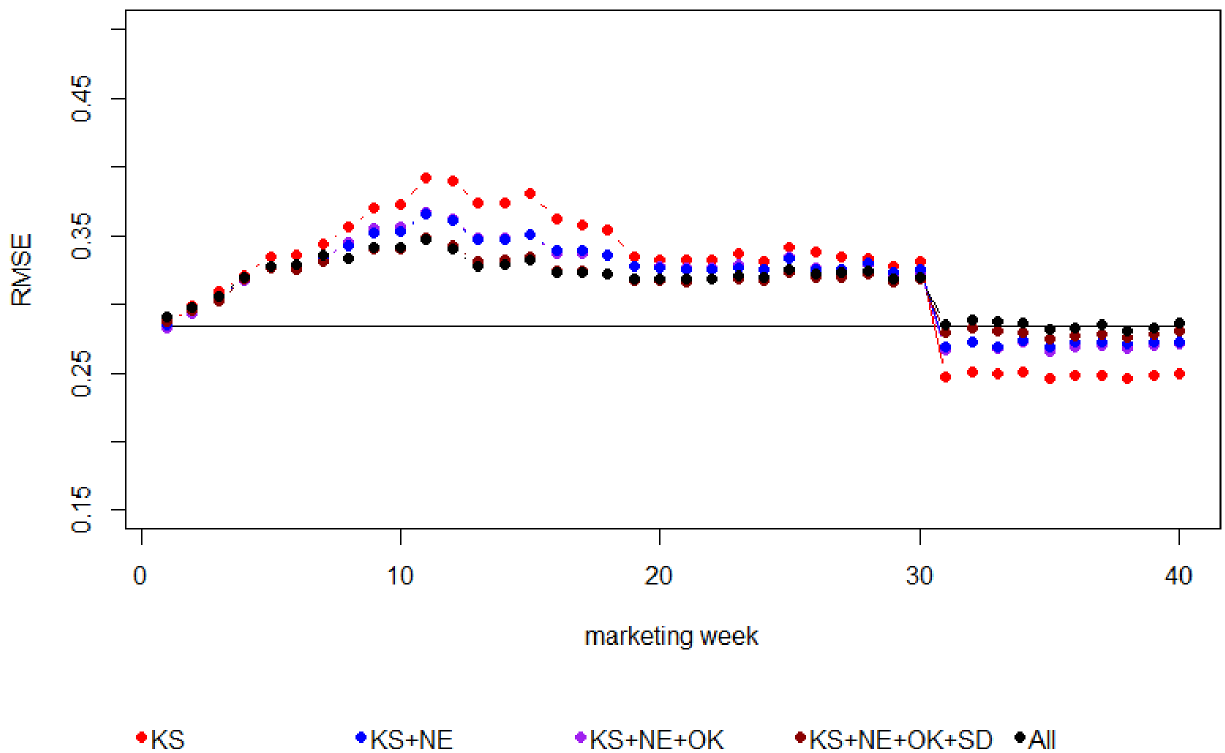

Figure 30 Out-of-sample performance (measured as RMSE) of the avg. baseline model across all datasets (the gray horizontal line), and the 'preferred' model (the dotted colored lines) are reported by different multi-state regions: for Kansas alone (KS), for Kansas and Nebraska (KS+NE), for Kansas, Nebraska, and Oklahoma (KS+NE+OK), for Kansas, Nebraska, Oklahoma, and South Dakota (KS+NE+OK+SD), and for Kansas, Nebraska, Oklahoma, South Dakota, North Dakota, Texas, and Wyoming (all).

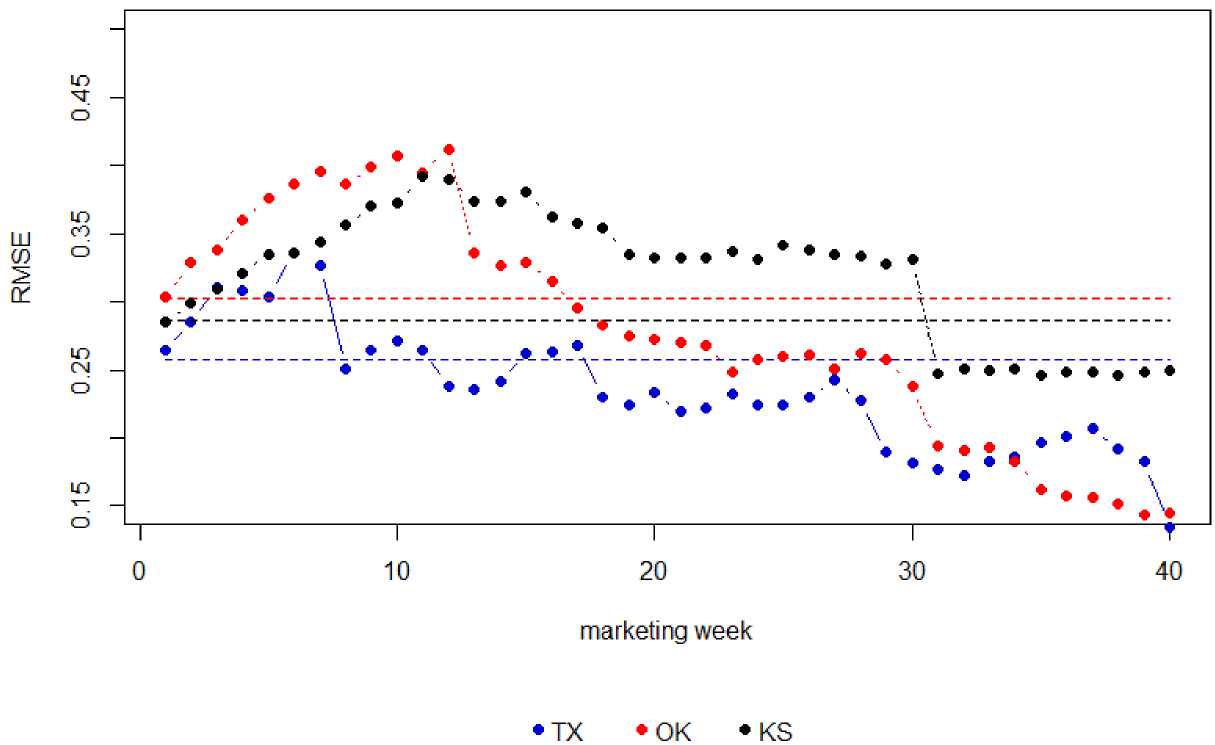

Figure 31 Out-of-sample performance (measured as RMSE) of the baseline model (the dashed lines), and the 'preferred' model (the dotted colored lines) are reported by state (i.e. the dataset was restricted to each individual state and the two models were estimated and RMSE calculated using this restricted data).

Note: The blue-colored lines show the baseline and preferred models' RMSE for Texas; the red-colored lines show the baseline and preferred models' RMSE for Oklahoma; the black-colored lines show the baseline and preferred models' RMSE for Kansas.

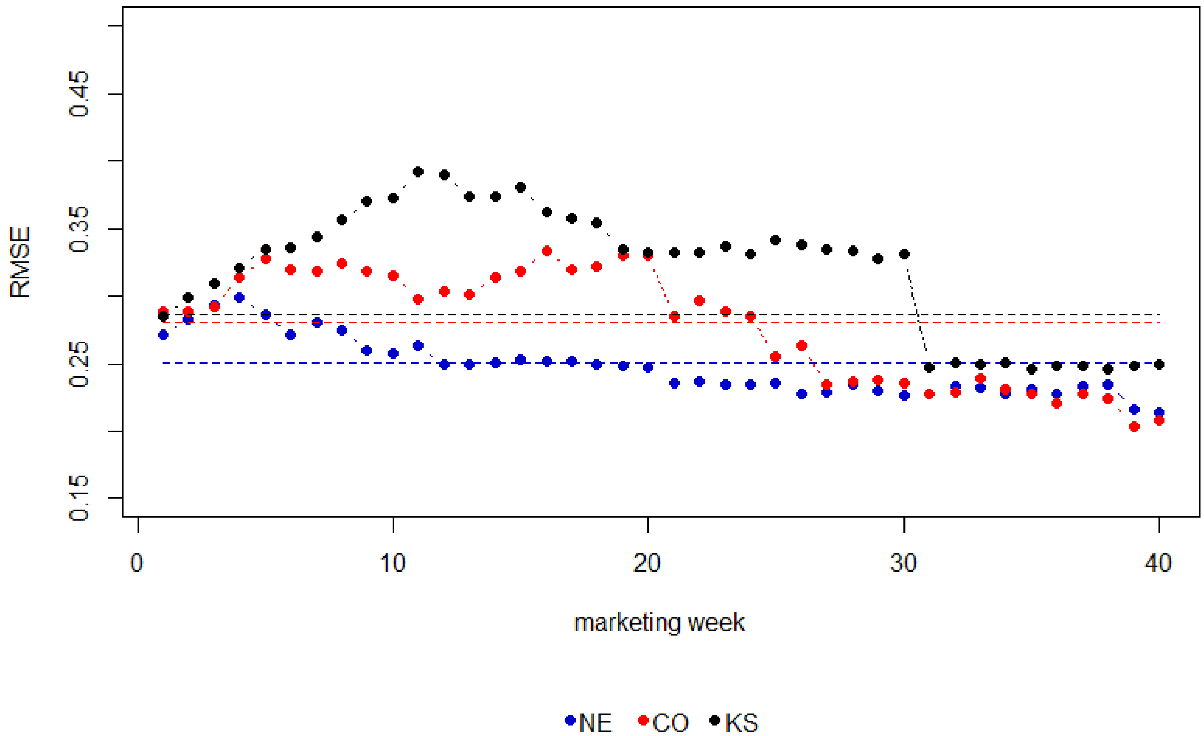

Figure 32 Out-of-sample performance (measured as RMSE) of the baseline model (the dashed lines), and the 'preferred' model (the dotted colored lines) are reported by state (i.e. the dataset was restricted to each individual state and the two models were estimated and RMSE calculated using this restricted data).

Note: The blue-colored lines show the baseline and preferred models' RMSE for Nebraska; the red-colored lines show the baseline and preferred models' RMSE for Colorado; the black-colored lines show the baseline and preferred models' RMSE for Kansas.

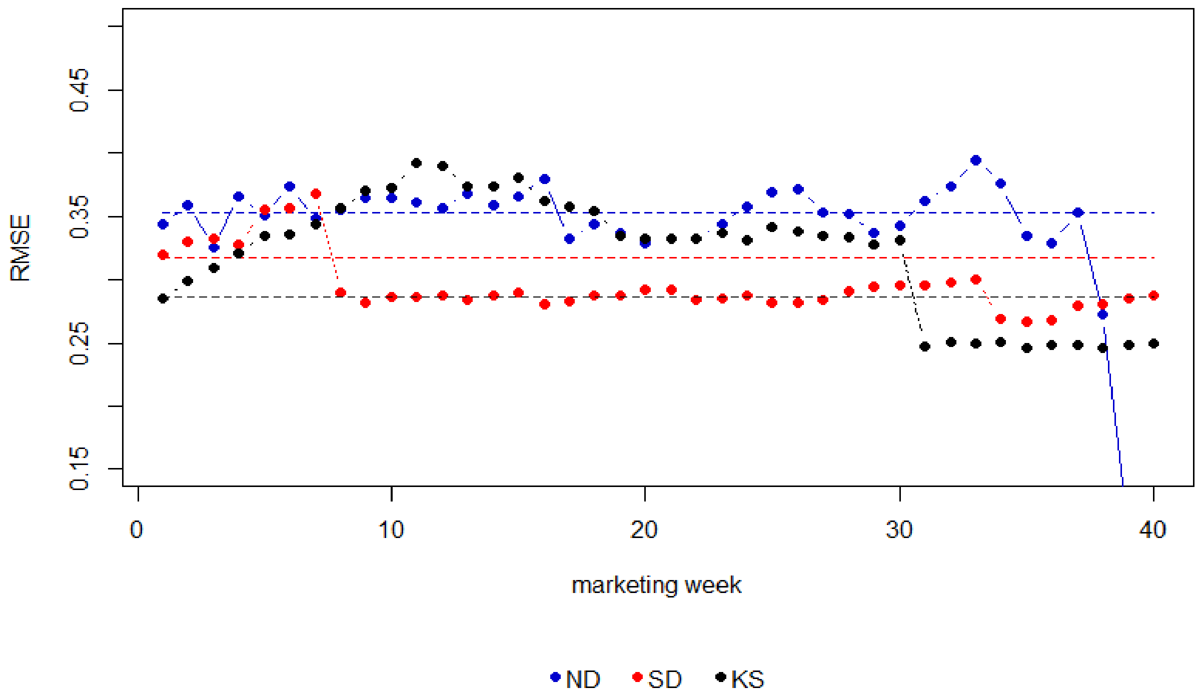

Figure 33 Out-of-sample performance (measured as RMSE) of the baseline model (the dashed lines), and the 'preferred' model (the dotted colored lines) are reported by state (i.e. the dataset was restricted to each individual state and the two models were estimated and RMSE calculated using this restricted data).

Note: The blue-colored lines show the baseline and preferred models' RMSE for North Dakota; the red-colored lines show the baseline and preferred models' RMSE for South Dakota; the black-colored lines show the baseline and preferred models' RMSE for Kansas.

References

Adjemian, M.K., P. Garcia, S. Irwin, and A. Smith. 2013. "Non-Convergence in Domestic Commodity Futures Markets: Causes, Consequences, and Remedies." *Economic Research Service: Economic Information Bulletin Number 115*.

Anda, A. and L. Pinter. 1994. "Sorghum Germination and Development as Influenced by Soil Temperature and Water Content." *Agronomy Journal*, 86(4).

Assefa, Y. and S.A. Staggenborg. 2010. "Grain Sorghum Yield with Hybrid Advancement and Changes in Agronomic Practices from 1957 through 2008." *Agronomy Journal*, 102(2).

Aufhammer, M., S.M. Hsiang, W. Schlenker, and A. Sobel. 2013. "Using Weather Data and Climate Model Output in Economic Analyses of Climate Change. *Review of Environmental Economics and Policy*, 7(2).

Baltagi, B.H. 2013. "Panel Data Forecasting." In *Handbook of Economic Forecasting, Volume 2B* [Elliott, G. and A. Timmermann (eds.)]. Amsterdam, Elsevier.

Belton, P.S. and J.R.N. Taylor. 2004. "Sorghum and Millets: Protein Sources for Africa." *Trends in Food Science & Technology*, 15(2).

Brown, M.E. and C.C. Funk. 2008. "Food Security under Climate Change." *Science*, 319(5863).

Burke, M., and D.B. Lobell. 2010. "Food Security and Adaptation to Climate Change: What Do We Know?" *Climate Change and Food Security*. Springer, Dordrecht.

Burke, M. and K. Emerick. 2016. "Adaptation to Climate Change: Evidence from US Agriculture." *American Economic Journal: Economic Policy*, 8(3).

Cairns, J.E., J. Hellin, K. Sonder, J.L. Araus, J.F. MacRobert, C. Thierfelder, and B.M. Prasanna. 2013. "Adapting Maize Production to Climate Change in Sub-Saharan Africa." *Food Security*, 5(3).

Carbone, G.J., W. Kiechle, C. Locke, L.O. Mearns, L. McDaniel, and M.W. Downton. 2003. "Response of Soybean and Sorghum to Varying Spatial Scales of Climate Change Scenarios in the Southeastern United States". *Climatic Change*, Vol. 60.

Challinor, A.J., J. Watson, D.B. Lobell, S.M. Howden, D.R. Smith, and N. Chhetri. 2014. "A Meta-Analysis of Crop Yield under Climate Change and Adaptation." *Nature Climate Change*, 4(4).

Chen, S., X. Chen, and J. Xu. 2016. "Impacts of Climate Change on Agriculture: Evidence from China." *Journal of Environmental Economics and Management*, 76.

Chiluwal, A., R. Bheemanahalli, R. Perumal, A.R. Asebedo, E. Bashir, A. Lamsal, D. Sebela, N.J. Shetty, and S.V.K. Jagadish. 2018. "Integrated Aerial and Destructive Phenotyping Differentiates Chilling Stress Tolerance during Early Seedling Growth in Sorghum." *Field Crops Research*, 227.

Ciampitti, I. and M. Knapp. 2013. "Sorghum Development and Potential Freeze Injury." Agronomy eUpdate, 425. Accessed at: https://webapp.agron.ksu.edu/agr_social/m_eu_article.throck?article_id=78

D'Agostino, A.L. and W. Schlenker. 2016. "Recent Weather Fluctuations and Agricultural Yields: Implications for Climate Change." *Agricultural Economics,* 47(S1).

European Space Agency. 2018. European Space Agency Climate Change Initiative Soil Moisture (ESA CCI SM) v04.4. Available online at: https://www.esa-soilmoisture-cci.org/.

Fezzi, C. and I. Bateman. 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).

Garcia, P. R. Hauser, and A. Tumblin. 1986. "Corn and Soybean Basis Behavior: An Intertemporal Cross-Sectional Analysis." *Proceedings of the NCR-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management.*

Hamman, L., K.C. Dhuyvetter, and M. Boland. 2001. "Economic Issues with Grain Sorghum." Kansas State University Extension Publication, MF-2513. Accessed at: https://www.bookstore.ksre.ksu.edu/pubs/MF2513.pdf.

Hatchett, R.B., B.W. Brorsen, K.B. Anderson. 2010. "Optimal Length of Moving Average to Forecast Future Basis." *Journal of Agricultural and Resource Economics*, 35(1).

Hsiang, Solomon M. 2016. "Climate Econometrics." *Annual Review of Resource Economics*, 8.

Jay, A., D.R. Reidmiller, C.W. Avery, D. Barrie, B.J. DeAngelo, A. Dave, M. Dzaugis, M. Kolian, K.L.M. Lewis, K. Reeves, and D. Winner. 2018. "Overview." In *Impacts, Risks, and Adaptation in the United States: Fourth National Climate Assessment, Volume II* [Reidmiller, D.R., C.W. Avery, D.R. Easterling, K.E. Kunkel, K.L.M. Lewis, T.K. Maycock, and B.C. Stewart (eds.)]. U.S. Global Change Research Program, Washington, DC, USA. doi: 10.7930/NCA4.2018.CH1.

Kaminski, J., I. Kan, and A. Fleischer. 2013. "A Structural Land-Use Analysis of Agricultural Adaptation to Climate Change: A Proactive Approach." *American Journal of Agricultural Economics*, 95(1).

Kastens, T.L., R. Jones, and T.C. Schroeder. 1998. "Futures-Based Price Forecasts for Agricultural Producers and Businesses." *Journal of Agricultural and Resource Economics*, 23(1).

Kawasaki, K. 2018. “Two Harvests Are Better than One: Double Cropping as a Strategy for Climate Change Adaptation.” *American Journal of Agricultural Economics*, 101(1).

Kawasaki, K. and S. Uchida. 2016. “Quality Matters More Than Quantity: Asymmetric Temperature Effects on Crop Yield and Quality Grade.” *American Journal of Agricultural Economics*, 98(4).

Lambert, D.K. 2014. “Historical Impacts of Precipitation and Temperature on Farm Production in Kansas.” *Journal of Agricultural and Applied Economics*, 46(4).

Lobell, D.B. and C.B. Field. 2007. “Global Scale Climate-Crop Yield Relationships and the Impacts of Recent Warming.” *Environmental Research Letters*, 2(1).

Lobell, D.B., M.B. Burke, C. Tebaldi, M.D. Mastrandrea, W.P. Falcon, R.L. Naylor. 2008. “Prioritizing Climate Change Adaptation Needs for Food Security in 2030.” *Science*, 319.

Lobell, D.B., W. Schlenker, and J. Costa-Roberts. 2011. “Climate Trends and Global Crop Production since 1980.” *Science*, 1204531.

Lobell, D.B., M. Bänziger, C. Magorokosho, and B. Vivek. 2011. “Nonlinear Heat Effects on African Maize as Evidenced by Historical Yield Trials.” *Nature Climate Change*, 1(1), 42.

Lobell, D.B., M.J. Roberts, W. Schlenker, N. Braun, B.B. Little, R.M. Rejesus, G.L. Hammer. 2014. “Greater Sensitivity to Drought Accompanies Maize Yield Increase in the U.S. Midwest.” *Science*, 344(6183).

Lobell, D.B. and S. Asseng. 2017. “Comparing Estimates of Climate Change Impacts from Process-Based and Statistical Crop Models.” *Environmental Research Letters*, 12(1), 015001.

McCarl, B.A., A.W. Thayer, and J.P.H. Jones. 2016. “The Challenge of Climate Change Adaptation for Agriculture: An Economically Oriented Review.” *Journal of Agricultural and Applied Economics*, 48(4).

McGrath, J.B. and D.B. Lobell. 2013. “Regional Disparities in the CO_2 Fertilization Effect and Implications for Crop Yields.” *Environmental Research Letters*, 8(1).

Menz, M.A., R.R. Klein, N.C. Unruh, W.L. Rooney, P.E. Klein, and J.E. Mullet. 2004. “Genetic Diversity of Public Inbreds of Sorghum Determined by Mapped AFLP and SSR Markers.” *Crop Science*, 44(4).

Miao, R., M. Khanna, and H. Huang. 2016. “Responsiveness of Crop Yield and Acreage to Prices and Climate.” *American Journal of Agricultural Economics*, 98(1).

Mishra, A.K., H.S. El-Osta, and C.L. Sandretto. 2004. “Factors Affecting Farm Enterprise Diversification.” *Agricultural Finance Review,* 64(2).

National Oceanic and Atmospheric Administration. 2019. Soil Moisture Operational Products System (SMOPS). Available online at: https://www.ospo.noaa.gov/Products/land/smops/.

Ortiz-Bobea, A., H. Wang, C.M. Carrillo, and T.R. Ault. 2019. “Unpacking the Climatic Drivers of US Agricultural Yields.” *Environmental Research Letters,* 14(6).

Ortiz-Bobea, A., and R.E. Just. 2012. “Modeling the Structure of Adaptation in Climate Change Impact Assessment.” *American Journal of Agricultural Economics,* 95(2).

Pachauri, R.K., M.R. Allen, F.R. Barros, J. Broome, W. Cramer, R. Christ, J.A. Church, L. Clarke, Q. Dahe, P. Dasgupta, and N.K. Dubash. 2014. *Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II, and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. IPCC.

Reilly, J., F. Tubiello, B. McCarl, D. Abler, R. Darwin, K. Fuglie, S. Hollinger, C. Izaurralde, S. Jagtap, J. Jones, L. Mearns, D. Ojima, E. Paul, K. Paustian, S. Riha, N. Rosenberg, and C. Rosenzweig. 2003. “U.S. Agriculture and Climate Change: New Results.” *Climatic Change* 57(1-2).

Roberts, M.J., W. Schlenker, and J. Eyer. 2012. “Agronomic Weather Measures in Econometric Models of Crop Yield with Implications for Climate Change.” *American Journal of Agricultural Economics*, 95(2).

Rogers, D. 1997. “Irrigation Management” in *Wheat Production Handbook, Kansas State University Research & Extension Publication C-529*. Available online at: https://www.bookstore.ksre.ksu.edu/pubs/c529.pdf

Schlenker, W. and M.J. Roberts. 2009. “Nonlinear Temperature Effects Indicate Severe Damages to U.S. Crop Yields under Climate Change.” *Proceedings of the National Academy of Sciences*, 106(37).

Singh, B., M. El Maayer, P. Andre, C.R. Bryant, and J.P. Thouez. 1998. “Impacts of a GHG-Induced Climate Change on Crop Yields: Effects of Acceleration in Maturation, Moisture Stress and Optimal Temperature”. *Climatic Change*, Vol. 38.

Snyder, R.L. 1985. “Hand Calculating Degree Days.” 1985. *Agricultural and Forest Meteorology*, 35(1-4).

Tack, J., A. Barkley, and L.L. Nalley. 2015. “Effect of Warming Temperatures on US Wheat Yields.” *Proceedings of the National Academy of Sciences,* 112(22).

Tack, J., J. Lingenfelser, and S.V.K. Jagadish. 2017. "Disaggregating Sorghum Yield Reductions under Warming Scenarios Exposes Narrow Genetic Diversity in US Breeding Programs." *Proceedings of the National Academy of Sciences*, 114(35).

Taylor, M.R., K.C. Dhuyvetter, and T.L. Kastens. 2006. "Forecasting Crop Basis Using Historical Averages Supplemented with Current Market Information." *Journal of Agricultural and Resource Economics,* 31(3).

Taylor, M.R., G. Tonsor, and K. Dhuyvetter. 2014. "Structural Change in Forward Contracting Costs for Kansas Wheat." *Journal of Agricultural and Resource Economics*, 39(2).

Thompson, N.M., A.J. Edwards, J.R. Mintert, and C.A. Hurt. 2019. "Practical Alternatives for Forecasting Corn and Soybean Basis in the Eastern Corn Belt throughout the Crop-Marketing Year." *Journal of Agricultural and Resource Economics*, 44(3).

United Sorghum Checkoff Program. 2016. "Renewables." Available at: http://www.sorghumcheckoff.com/market-opportunities/renewables (accessed: 07/25/2018).

Yang, S. and C.R. Shumway. 2015. "Dynamic Adjustment in US Agriculture under Climate Change." *American Journal of Agricultural Economics*, 98(3).

Yu, T. and B. Babcock. 2010. "Are U.S. Corn and Soybeans Becoming More Drought Tolerant?" *American Journal of Agricultural Economics*, 92(5).

Zhang, P., J. Zhang, and M. Chen. 2017. "Economic Impacts of Climate Change on Agriculture: The Importance of Additional Climatic Variables Other Than Temperature and Precipitation." *Journal of Environmental Economics and Management*, 83.

Appendix A - Preferred Model by Vantage Point

Our preferred harvest basis forecasting model is from every vantage point (*j = 1, 2, …40*) in the marketing year,

(A1) $b_{it} = \alpha_i + \beta_0 b_{it-1} + \sum_{w=1}^{J} I_j \sum_{k=1}^{3} \gamma_{wk} sm_{wit}^{k} + \varepsilon_{it}$, ($I_j = 1$ if $w \leq j$ and 0 otherwise)

where b_{it} is harvest basis at elevator *i* in year *t*; α_i are fixed effects that capture variation in basis caused by time invariant, elevator-specific characteristics; b_{it-1} is harvest basis at elevator *i* in the prior year (*t-1*); and sm_{wit} is the soil moisture reading in week *w* at elevator *i* in year *t*. We incorporate an indicator variable (I_j) into the model such that at any vantage point only the current and prior week's soil moisture readings are admitted additively.

From vantage point *j = 1* (i.e. the first week in September), our harvest basis forecast is,

(A2) $b_{it} = \alpha_i + \beta_0 b_{it-1} + \gamma_{1,1} sm_{1it} + \gamma_{1,2} sm_{1it}^2 + \gamma_{1,3} sm_{1it}^3 + \varepsilon_{it}$.

From vantage point *j = 2* (i.e. the second week in September), our harvest basis forecast is,

(A3) $b_{it} = \alpha_i + \beta_0 b_{it-1} + \gamma_{1,1} sm_{1it} + \gamma_{1,2} sm_{1it}^2 + \gamma_{1,3} sm_{1it}^3 + \gamma_{2,1} sm_{2it} + \gamma_{2,2} sm_{2it}^2 +$
$\gamma_{2,3} sm_{2it}^3 + \varepsilon_{it}$.

We see that as we move from one vantage point to the next, the only change in the forecasting model is the introduction of the next week's soil moisture observations (as three variables - a linear, quadratic, and cubic term).

From vantage point *j = 40* (i.e. the fourth week in June), our harvest basis forecast is,

(A4) $b_{it} = \alpha_i + \beta_0 b_{it-1} + \gamma_{1,1} sm_{1it} + \gamma_{1,2} sm_{1it}^2 + \gamma_{1,3} sm_{1it}^3 + \gamma_{2,1} sm_{2it} + \gamma_{2,2} sm_{2it}^2 +$
$\gamma_{2,3} sm_{2it}^3 + \gamma_{3,1} sm_{3it} + \gamma_{3,2} sm_{3it}^2 + \gamma_{3,3} sm_{3it}^3 + \gamma_{4,1} sm_{4it} + \gamma_{4,2} sm_{4it}^2 + \gamma_{4,3} sm_{4it}^3 +$
$\gamma_{5,1} sm_{5it} + \gamma_{5,2} sm_{5it}^2 + \gamma_{5,3} sm_{5it}^3 + \gamma_{6,1} sm_{6it} + \gamma_{6,2} sm_{6it}^2 + \gamma_{6,3} sm_{6it}^3 + \gamma_{7,1} sm_{7it} +$
$\gamma_{7,2} sm_{7it}^2 + \gamma_{7,3} sm_{7it}^3 + \gamma_{8,1} sm_{8it} + \gamma_{8,2} sm_{8it}^2 + \gamma_{8,3} sm_{8it}^3 + \gamma_{9,1} sm_{9it} + \gamma_{9,2} sm_{9it}^2 +$

$$\gamma_{9,3}sm_{9it}^3 + \gamma_{10,1}sm_{10it} + \gamma_{10,2}sm_{10it}^2 + \gamma_{10,3}sm_{10it}^3 + \gamma_{11,1}sm_{11it} + \gamma_{11,2}sm_{11it}^2 +$$

$$\gamma_{11,3}sm_{11it}^3 + \gamma_{12,1}sm_{12it} + \gamma_{12,2}sm_{12it}^2 + \gamma_{12,3}sm_{12it}^3 + \gamma_{13,1}sm_{13it} + \gamma_{13,2}sm_{13it}^2 +$$

$$\gamma_{13,3}sm_{13it}^3 + \gamma_{14,1}sm_{14it} + \gamma_{14,2}sm_{14it}^2 + \gamma_{14,3}sm_{14it}^3 + \gamma_{15,1}sm_{15it} + \gamma_{15,2}sm_{15it}^2 +$$

$$\gamma_{15,3}sm_{15it}^3 + \gamma_{16,1}sm_{16it} + \gamma_{16,2}sm_{16it}^2 + \gamma_{16,3}sm_{16it}^3 + \gamma_{17,1}sm_{17it} + \gamma_{17,2}sm_{17it}^2 +$$

$$\gamma_{17,3}sm_{17it}^3 + \gamma_{18,1}sm_{18it} + \gamma_{18,2}sm_{18it}^2 + \gamma_{18,3}sm_{18it}^3 + \gamma_{19,1}sm_{19it} + \gamma_{19,2}sm_{19it}^2 +$$

$$\gamma_{19,3}sm_{19it}^3 + \gamma_{20,1}sm_{20it} + \gamma_{20,2}sm_{20it}^2 + \gamma_{20,3}sm_{20it}^3 + \gamma_{21,1}sm_{21it} + \gamma_{21,2}sm_{21it}^2 +$$

$$\gamma_{21,3}sm_{21it}^3 + \gamma_{22,1}sm_{22it} + \gamma_{22,2}sm_{22it}^2 + \gamma_{22,3}sm_{22it}^3 + \gamma_{23,1}sm_{23it} + \gamma_{23,2}sm_{23it}^2 +$$

$$\gamma_{23,3}sm_{23it}^3 + \gamma_{24,1}sm_{24it} + \gamma_{24,2}sm_{24it}^2 + \gamma_{24,3}sm_{24it}^3 + \gamma_{25,1}sm_{25it} + \gamma_{25,2}sm_{25it}^2 +$$

$$\gamma_{25,3}sm_{25it}^3 + \gamma_{26,1}sm_{26it} + \gamma_{26,2}sm_{26it}^2 + \gamma_{26,3}sm_{26it}^3 + \gamma_{27,1}sm_{27it} + \gamma_{27,2}sm_{27it}^2 +$$

$$\gamma_{27,3}sm_{27it}^3 + \gamma_{28,1}sm_{28it} + \gamma_{28,2}sm_{28it}^2 + \gamma_{28,3}sm_{28it}^3 + \gamma_{29,1}sm_{29it} + \gamma_{29,2}sm_{29it}^2 +$$

$$\gamma_{29,3}sm_{29it}^3 + \gamma_{30,1}sm_{30it} + \gamma_{30,2}sm_{30it}^2 + \gamma_{30,3}sm_{30it}^3 + \gamma_{31,1}sm_{31it} + \gamma_{31,2}sm_{31it}^2 +$$

$$\gamma_{31,3}sm_{31it}^3 + \gamma_{32,1}sm_{32it} + \gamma_{32,2}sm_{32it}^2 + \gamma_{32,3}sm_{32it}^3 + \gamma_{33,1}sm_{33it} + \gamma_{33,2}sm_{33it}^2 +$$

$$\gamma_{33,3}sm_{33it}^3 + \gamma_{34,1}sm_{34it} + \gamma_{34,2}sm_{34it}^2 + \gamma_{34,3}sm_{34it}^3 + \gamma_{35,1}sm_{35it} + \gamma_{35,2}sm_{35it}^2 +$$

$$\gamma_{35,3}sm_{35it}^3 + \gamma_{36,1}sm_{36it} + \gamma_{36,2}sm_{36it}^2 + \gamma_{36,3}sm_{36it}^3 + \gamma_{37,1}sm_{37it} + \gamma_{37,2}sm_{37it}^2 +$$

$$\gamma_{37,3}sm_{37it}^3 + \gamma_{38,1}sm_{38it} + \gamma_{38,2}sm_{38it}^2 + \gamma_{38,3}sm_{38it}^3 + \gamma_{39,1}sm_{39it} + \gamma_{39,2}sm_{39it}^2 +$$

$$\gamma_{39,3}sm_{39it}^3 + \gamma_{40,1}sm_{40it} + \gamma_{40,2}sm_{40it}^2 + \gamma_{40,3}sm_{40it}^3 + \varepsilon_{it}.$$

Appendix B - Baseline and Preferred Model Regression Results

Table B1. Regression Results for the Baseline and Preferred Model

	baseline	***week 1***	***week 10***	***week20***	***week30***	***week40***
prior basis	0.04**	-0.04**	0.1***	0.1***	0.02	0.04**
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
sm week 1		1.3	-0.9	-5.6*	0.1	1.5
		(3.4)	(3.5)	(3.2)	(3.1)	(2.7)
sm week 2			-1.1	1.0	-12.3***	-10.2***
			(3.9)	(3.3)	(3.3)	(2.8)
sm week 3			-4.2	-6.6**	-4.5	-3.3
			(3.4)	(2.9)	(3.0)	(2.6)
sm week 4			-12.7***	-0.7	4.8	-3.3
			(4.5)	(3.8)	(3.5)	(2.9)
sm week 5			15.7***	-1.0	-10.6***	-2.1
			(4.1)	(3.5)	(3.4)	(2.8)
sm week 6			-3.2	-5.2**	-3.2	-1.4
			(2.8)	(2.4)	(2.5)	(2.2)
sm week 7			6.2*	15.4***	24.9***	15.3***
			(3.3)	(3.3)	(3.4)	(3.0)
sm week 8			-25.9***	-17.9***	-23.0***	-9.4***
			(4.4)	(4.1)	(3.9)	(3.3)
sm week 9			3.4	-2.4	1.7	4.0
			(3.4)	(2.9)	(3.0)	(2.6)
sm week 10			3.4	-1.0	-1.5	-3.3*
			(2.6)	(2.2)	(2.2)	(2.0)
sm week 11				11.3***	20.0***	10.3***
				(2.4)	(2.4)	(2.1)
sm week 12				0.3	-6.1**	-7.1***
				(2.6)	(2.6)	(2.2)
sm week 13				-4.1	-4.9*	-2.0
				(2.8)	(2.6)	(2.2)
sm week 14				10.1***	6.5**	1.1
				(2.9)	(2.7)	(2.3)
sm week 15				-20.4***	-19.0***	-19.6***
				(3.7)	(3.5)	(3.0)
sm week 16				6.6**	3.3	7.5***
				(3.1)	(2.9)	(2.4)

sm week 17	-0.6	2.1	-3.6**
	(2.3)	(2.2)	(1.8)
sm week 18	-7.4***	-3.8	7.7***
	(2.7)	(2.7)	(2.3)
sm week 19	13.9***	16.5***	2.8
	(3.5)	(3.4)	(2.8)
sm week 20	-20.8***	-21.1***	-14.1***
	(4.1)	(4.2)	(3.4)
sm week 21		10.4***	10.2***
		(2.4)	(2.0)
sm week 22		-6.7*	-4.8*
		(3.4)	(2.9)
sm week 23		5.1	-2.5
		(3.5)	(3.1)
sm week 24		-3.1	-7.5***
		(3.3)	(2.6)
sm week 25		1.0	-3.1
		(3.0)	(2.5)
sm week 26		1.8	0.02
		(2.8)	(2.3)
sm week 27		-6.5***	2.1
		(2.1)	(1.9)
sm week 28		-18.8***	-9.8***
		(3.1)	(2.5)
sm week 29		6.4**	2.9
		(2.7)	(2.4)
sm week 30		2.4	1.8
		(2.2)	(1.9)
sm week 31			6.4***
			(2.3)
sm week 32			-15.4***
			(2.6)
sm week 33			-5.4**
			(2.4)
sm week 34			-0.1
			(2.6)
sm week 35			-1.3
			(2.8)

sm week 36					11.4***
					(3.0)
sm week 37					8.8***
					(2.8)
sm week 38					-9.3***
					(2.3)
sm week 39					3.1
					(2.9)
sm week 40					-0.4
					(2.5)
sm^2 *week 1*	-19.2	-6.4	29.2*	2.6	1.5
	(16.0)	(16.2)	(15.0)	(14.4)	(12.3)
sm^2 *week 2*		-7.3	-10.3	35.3**	39.7***
		(16.8)	(14.0)	(14.1)	(11.9)
sm^2 *week 3*		12.1	24.2*	17.0	-2.0
		(14.8)	(12.6)	(12.8)	(11.0)
sm^2 *week 4*		70.5***	14.0	-15.8	29.4**
		(20.2)	(16.8)	(15.8)	(13.1)
sm^2 *week 5*		-104.0***	-29.4*	15.3	-28.5**
		(18.7)	(15.8)	(15.6)	(12.7)
sm^2 *week 6*		38.6***	39.3***	36.5***	17.9*
		(12.3)	(10.5)	(11.1)	(9.4)
sm^2 *week 7*		6.0	-54.5***	-90.9***	-52.4***
		(14.3)	(13.8)	(14.5)	(12.5)
sm^2 *week 8*		107.3***	72.0***	91.6***	45.4***
		(19.2)	(17.9)	(17.2)	(14.5)
sm^2 *week 9*		-28.7*	8.9	-10.8	-18.9
		(15.3)	(13.3)	(13.2)	(11.6)
sm^2 *week 10*		-30.4**	-9.2	-10.3	7.2
		(11.9)	(10.4)	(10.2)	(9.0)
sm^2 *week 11*			-56.4***	-81.9***	-41.7***
			(10.9)	(11.0)	(9.3)
sm^2 *week 12*			-4.1	19.2*	19.9**
			(11.8)	(11.6)	(9.7)
sm^2 *week 13*			13.5	19.4*	17.2*
			(11.8)	(11.2)	(9.4)
sm^2 *week 14*			-48.5***	-31.2***	-1.1
			(12.4)	(11.8)	(10.0)

sm^2 week 15	88.5***	82.9***	91.0***
	(15.6)	(14.9)	(12.6)
sm^2 week 16	-26.6**	-13.6	-36.6***
	(13.0)	(12.2)	(10.1)
sm^2 week 17	-2.0	-11.3	13.8*
	(10.9)	(10.4)	(8.4)
sm^2 week 18	22.1*	5.4	-41.4***
	(12.1)	(12.3)	(10.3)
sm^2 week 19	-56.2***	-67.3***	-20.3*
	(14.5)	(14.0)	(11.7)
sm^2 week 20	80.7***	84.0***	53.0***
	(17.4)	(17.6)	(14.6)
sm^2 week 21		-40.1***	-39.0***
		(9.3)	(7.5)
sm^2 week 22		3.2	12.2
		(14.7)	(12.3)
sm^2 week 23		-5.6	16.5
		(14.9)	(13.0)
sm^2 week 24		6.5	27.9**
		(13.9)	(11.2)
sm^2 week 25		-9.3	7.2
		(13.5)	(11.2)
sm^2 week 26		3.0	6.6
		(12.1)	(9.9)
sm^2 week 27		24.2***	-6.8
		(9.2)	(8.5)
sm^2 week 28		52.0***	21.6*
		(13.4)	(11.1)
sm^2 week 29		-8.3	-1.4
		(12.0)	(10.6)
sm^2 week 30		-13.5	1.5
		(10.3)	(9.1)
sm^2 week 31			-62.5***
			(10.3)
sm^2 week 32			75.6***
			(11.5)
sm^2 week 33			30.7***
			(11.1)

sm^2 *week 34*					-13.5
					(12.0)
sm^2 *week 35*					3.9
					(12.5)
sm^2 *week 36*					-61.8***
					(13.2)
sm^2 *week 37*					-29.2**
					(13.2)
sm^2 *week 38*					43.8***
					(10.9)
sm^2 *week 39*					-13.1
					(13.1)
sm^2 *week 40*					7.7
					(11.9)
sm^3 *week 1*	20.8	12.6	-49.6**	-8.3	-11.7
	(24.6)	(24.8)	(22.7)	(21.9)	(18.5)
sm^3 *week 2*		9.8	8.3	-41.6**	-55.7***
		(23.9)	(19.8)	(19.7)	(16.6)
sm^3 *week 3*		0.3	-22.0	-11.7	20.1
		(20.9)	(17.7)	(17.9)	(15.5)
sm^3 *week 4*		-109.8***	-21.1	24.1	-47.5**
		(29.5)	(24.6)	(23.1)	(19.3)
sm^3 *week 5*		203.3***	86.1***	20.5	83.0***
		(28.2)	(23.7)	(23.2)	(19.0)
sm^3 *week 6*		-104.0***	-91.2***	-87.1***	-45.1***
		(18.2)	(15.4)	(16.1)	(13.7)
sm^3 *week 7*		-34.8*	75.5***	121.2***	69.7***
		(20.5)	(19.3)	(20.2)	(17.3)
sm^3 *week 8*		-150.6***	-110.4***	-138.8***	-82.7***
		(28.1)	(26.1)	(25.1)	(20.9)
sm^3 *week 9*		44.8**	-10.7	14.4	17.4
		(22.7)	(19.7)	(19.4)	(17.2)
sm^3 *week 10*		62.9***	29.8*	40.9***	4.9
		(17.8)	(15.7)	(15.4)	(13.4)
sm^3 *week 11*			85.6***	106.0***	50.3***
			(16.4)	(16.1)	(13.6)
sm^3 *week 12*			-0.2	-27.9	-21.9
			(17.6)	(17.3)	(14.3)

sm³ week 13	-22.1	-30.0*	-35.2***
	(16.5)	(15.6)	(13.0)
sm³ week 14	66.5***	40.5**	-6.4
	(17.5)	(16.6)	(14.0)
sm³ week 15	-134.0***	-125.4***	-133.7***
	(21.8)	(20.8)	(17.4)
sm³ week 16	55.4***	36.8**	65.6***
	(18.1)	(17.1)	(14.1)
sm³ week 17	3.0	14.5	-15.9
	(16.7)	(15.9)	(12.7)
sm³ week 18	-21.0	7.5	66.7***
	(17.9)	(18.4)	(15.4)
sm³ week 19	61.8***	79.7***	30.3*
	(19.6)	(18.8)	(15.6)
sm³ week 20	-108.1***	-111.9***	-67.2***
	(24.0)	(24.3)	(20.1)
sm³ week 21		45.0***	46.0***
		(11.5)	(9.3)
sm³ week 22		17.7	-9.5
		(20.6)	(17.3)
sm³ week 23		-0.1	-21.5
		(20.8)	(17.9)
sm³ week 24		-9.2	-34.6**
		(19.3)	(15.5)
sm³ week 25		28.2	0.2
		(19.5)	(16.2)
sm³ week 26		-22.5	-18.7
		(17.0)	(14.2)
sm³ week 27		-40.5***	-1.4
		(13.0)	(12.0)
sm³ week 28		-25.7	-2.0
		(19.4)	(16.1)
sm³ week 29		-20.9	-10.9
		(17.0)	(15.1)
sm³ week 30		23.3	-16.8
		(15.4)	(13.7)
sm³ week 31			113.1***
			(15.2)

sm^3 week 32						-112.1***
						(16.6)
sm^3 week 33						-60.2***
						(16.7)
sm^3 week 34						46.6**
						(18.1)
sm^3 week 35						-16.2
						(18.5)
sm^3 week 36						101.9***
						(19.2)
sm^3 week 37						26.0
						(20.4)
sm^3 week 38						-61.0***
						(16.8)
sm^3 week 39						13.4
						(19.9)
sm^3 week 40						-11.4
						(18.7)
Observations	3,956	3,956	3,956	3,956	3,956	3,956
R^2	0.0	0.1	0.5	0.7	0.8	0.9

Note: Standard errors are reported in parentheses, *** p<0.01, ** p<0.05, * p<0.1.