On the Timing of Relevant Weather Conditions in Agriculture

Presenter: Zhiyun Li, Cornell University, zl547@cornell.edu

Invited Paper prepared for presentation at the **2022 AEA/ASSA Annual Meeting VIRTUAL, January 7-9, 2022**

On the Timing of Relevant Weather Conditions in Agriculture

Zhiyun Li and Ariel Ortiz-Bobea*

Abstract

A growing empirical literature is analyzing the effects of weather fluctuations on a variety of economic outcomes with the goal of better understanding the potential impacts of climate change. In agricultural studies, constructing weather variables typically requires researchers to define a "season", a time period over which weather conditions are considered relevant to the agricultural outcome of interest. While researchers often have the background knowledge to make reasonable assumptions about seasonality in crop-specific analyses, these modeling choices are less obvious when dealing with aggregate agricultural data encompassing multiple crops or livestock. In this article, we explore the consequences of assuming an incorrect season in such analyses. We first provide a conceptual framework to show that imposing an incorrect season essentially introduces non-classical measurement error in weather regressors, causing unknown biases in weather impacts. We confirm this finding in simulations. We then propose a tractable data-driven approach to recover the "true" underlying season. The approach consists of a grid search with cross-validation that evaluates the fit of models based on a wide range of season definitions. In simulations, we find the approach is effective at recovering the "true" season under certain data generating processes. Finally, we apply our approach to a US state-level panel of agricultural Total Factor Productivity. We find, unsurprisingly, considerable differences in seasonality across regions. Importantly, our empirical findings suggest that imposing arbitrary seasons lead to substantially different estimates of weather effects in either direction, in line with our theoretical and simulated results. This work contributes to the development of more robust empirical studies of climate change impacts on agriculture and beyond.

JEL Codes: Q54, Q51, C52.

Keywords: weather, seasons, climate change, agriculture, non-classical measurement error, grid search, cross-validation.

*PhD candidate and Associate Professor, Charles H. Dyson School of Applied Economics and Management, Cornell University, Ithaca, NY 14850. Email:zl547@cornell.edu (Li). We would like to thank Ivan Rudik, Diego Cardoso, Todd Gerarden, Catherine Kling, Penny Sanders, Shanjun Li, and all seminar participants at Cornell University and the 4th Meeting of the Network on Agricultural Productivity (TFP) and the Environment for helpful comments. All errors are ours. The code and data necessary to replicate this study can be downloaded from: https://www.dropbox.com/s/f4f6ksaeokufa9n/timing_relevant_weather.zip?dl=0.

1 Introduction

A rapidly growing literature is documenting micro and macro-level impacts of weather shocks on a wide range of economic outcomes with the goal of characterizing the potential impacts of climate change. These studies have explored the effects of high temperature on energy demand (Dirks et al., 2015), labor productivity (Cachon et al., 2012), mortality (Deschênes and Greenstone, 2011), cognitive performance (Zivin et al., 2018), industrial output (Dell et al., 2012; Zhang et al., 2018), exports (Jones and Olken, 2010), conflict (Burke et al., 2014; Hsiang et al., 2011; Hsiang and Burke, 2014; Hsiang et al., 2013), migration (Cattaneo et al., 2019; Carleton and Hsiang, 2016; Stern, 2013) and corporate earnings (Addoum et al., 2019). There is an even larger literature exploring the effects of extreme weather on crop yields (Schlenker and Roberts, 2009; Tack et al., 2015) but also on more aggregate agricultural outcomes like Gross Domestic Product (GDP) (Dell et al., 2012; Burke et al., 2015) or Total Factor Productivity (TFP) (Liang et al., 2017; Ortiz-Bobea et al., 2018, 2021).

A seemingly innocuous but important empirical challenge in the literature is the choice of the "season" or the time window for constructing weather variables. Weather data are available at a much higher temporal frequency (e.g. daily) than the agricultural variable of interest (e.g. annual). As a result, researchers typically assume a temporal window over which to aggregate weather observations to construct weather variables for the regression analysis. Researchers often have the necessary background to make reasonable assumptions about these timing issues in crop-specific studies. For instance, it is common to assume a season spanning April through September when analyzing the effects of weather conditions on corn yields in the US Midwest. However, these modeling choices are more difficult to justify when dealing with aggregate agricultural outcomes like GDP or TFP, or when assessing the effects on livestock production, which do not have well-defined "seasons". In such cases it is not uncommon for researchers to consider the entire calendar year as the "season" in an attempt to capture all plausibly relevant weather conditions in a parsimonious way. To our knowledge, there is little formal guidance regarding these modeling choices and their implications for climate change impact assessments remain largely unexplored.

This article explores the consequences of imposing seasons that differ from the underlying data generating process (DGP) when estimating the effects of weather on aggregate outcomes. We see this "season selection" problem as a measurement error issue. We propose some terminology to clarify our exposition. We refer to the season in the DGP as the "true" or "correct" season. On the other hand, we refer to the "selected" season when referring to the one assumed by the researcher. Naturally, a selected season may be correct or incorrect depending on whether it matches the true season in the DGP. We also refer to "relevant" weather conditions when these occur within the true season, and as "irrelevant" to weather conditions occurring outside of the true season.

We illustrate the problem with an example. Suppose the researcher selects the annual maximum temperature as the weather regressor but the correct variable is the maximum temperature in July. The researcher assumed an incorrect season. This modeling choice essentially introduces irrelevant weather conditions into the weather variable. If these irrelevant weather conditions (occurring outside July) are *uncorrelated* with the relevant weather conditions (occurring in July) then the researcher introduced classical measurement error to the weather regressor. As is well known, this leads to attenuation bias. However, weather conditions occurring in the same year, particularly in adjacent time periods, may be correlated. This suggests that we may be dealing with nonclassical measurement error and thus with unknown biases.[1]

To explore the nature of the measurement error in weather regressors we conduct a simple exercise based on US weather data. Essentially we want to explore whether the measurement error –measured as the difference between weather variables defined over a "true" and a "selected" season– is correlated with weather over the "true" season. If that is the case then the measurement error is non-classical. We show the results of our exercise in figure 1. Each of the 20 panels shows the correlation of a weather variable (Tmax, Tmean, Tmin, Precipitation, shown by column) aggregated over a hypothetical "true" season (Winter, Spring, Summer, Fall, Annual, shown by row) and the measurement errors from aggregating the same variable over 78 different "selected" seasons defined over the calendar year (corresponding to each colored cell within a panel). White cells in the figure indicate no correlation between the weather variable aggregated over the hypothetical "true" season and the "selected" season, while red and blue cells indicate negative or positive correlations, respectively. Interestingly, most of the cells are either red or blue, suggesting that imposing incorrect seasons introduces non-classical measurement error and thus to biases in unknown direction. We later explore this issue more formally in the conceptual framework.

To address this issue we propose what a tractable data-driven approach to recover the "true" season. Our guiding principle is that the fit of an econometric model should improve when the selected season closely matches the true season in the underlying DGP. That is, accounting for relevant weather conditions should improve model fit, whereas incorporating irrelevant weather conditions should deteriorate it. Our strategy consist in a grid search over all possible calendar seasons within the year to identify the best-fitting season. This best-fitting season constitutes our estimate of the true season. Our measure of model fit is obtained from a cross-validation where we minimize out-of-sample mean squared error (MSE), a criterion that captures the tradeoff between goodness of fit and complexity of a model (Arlot and Celisse, 2010).

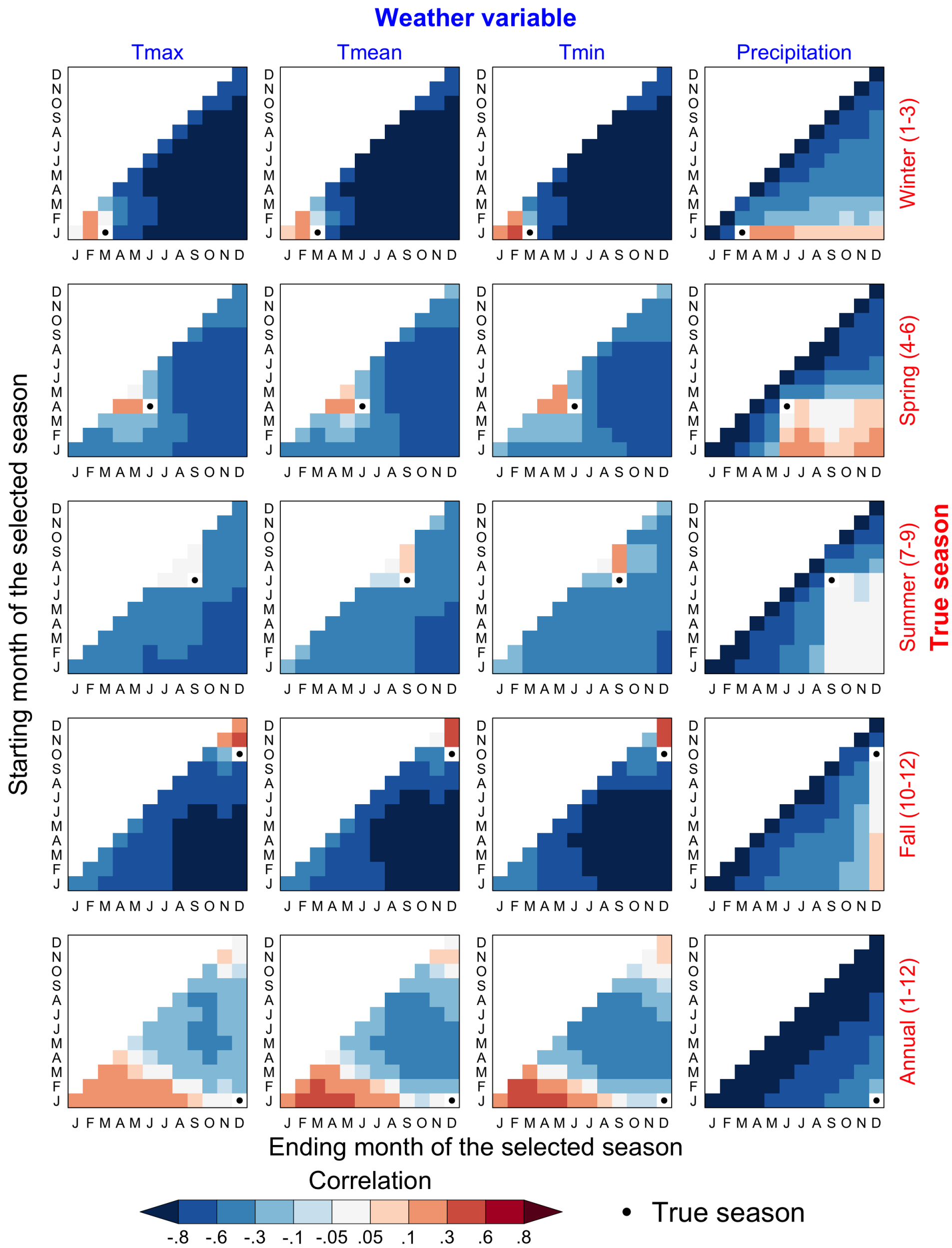
We evaluate our approach in a Monte Carlo simulation. Specifically, we generate data consistent with a particular "true" season and apply our approach to recover the underlying seasonality. We find that our approach is effective at recovering the underlying "true" season in the DGP even in some cases when the selected type of weather variable does not match the underlying DGP.[2] We also evaluate the impact of a uniform warming scenario when the selected season differs from the "true" underlying DGP to quantify the practical consequences of imposing an incorrect season. We find that the bias in impacts can be in either direction, as suggested by the non-classical nature of the measurement error.

Finally, we apply our season identification approach to a real-world setting. We analyze the effects of weather fluctuations on a US state-level panel of agricultural TFP growth rates over 1960-2004. In line with our simulations, our approach successfully recovers well-known growing seasons for the agricultural sector in various parts of the country dominated by field crops. Importantly, we find that the impact of a uniform warming can differ substantially, in either direction, when the selected season differs from our estimate of the "true" based on our grid search approach. Although we ignore the true DGP in this application, the results remain in line with the simulation results and show our approach makes a clear practical difference.

Our proposed approach has limitations the reader should keep in mind. First, our current approach

[1] Previous work has documented the presence of nonclassical measurement error in a wide range of settings especially labor (e.g.Borjas, 1980; Bound and Krueger, 1991; French, 2004; Kim and Solon, 2005; Arthi et al., 2018), consumer behavior (Gibson and Kim, 2010; Gibson et al., 2015), development (Baird and Özler, 2012; Beegle et al., 2012; Desiere and Jolliffe, 2018), health (Carletto et al., 2013; Larsen et al., 2019), and agriculture (Abay et al., 2019). There is relatively little work exploring the implications of measurement error in the climate-economy literature.

[2] For instance, the true variable in the DGP is Tmax but we select Tmin.



Notes: Correlations are computed based on state-level data across the lower 48 US states over the 1961-2004 period. Temperature variables are averaged over time whereas precipitation variables are summed over time. Correlations are computed between 2 variables: a hypothetical "true" variable (e.g. Winter Tmax, shown in the top left panel) and the measurement error from 78 possible calendar "selected" seasons defined within the year (each of the colored cells in a panel). As indicated in the main text, the measurement error is measured as the difference between the weather variable aggregated over the hypothetical "true" season and the weather variable aggregated over a range of "selected" seasons. In each panel, the vertical axis indicates the starting month of a "selected" season, while the horizontal axis indicates the ending month of a "selected" season. We highlight with a solid black circle the "true" season.

Figure 1: Correlation of weather variables defined over a hypothetical "true" season with the measurement error of that variable when constructed over alternative "selected" seasons.

is primarily geared toward identifying the underlying season. However, issues of variable selection are not restricted to temporal matters. The researcher may be considering competing types of weather variables (temperature versus vapor pressure deficit, or precipitation). The true DGP could in fact reflect multiple types of variables each one with different seasons. In this regard, our grid search approach could be extended to accommodate such cases by expanding the models under consideration which would clearly increase its computational requirements. Second, our approach does not explicitly tackles functional form issues. Most of our analysis is based on a quadratic functional form, but this could be too strong of an assumption. Again, our approach could be expanded to accommodate such cases which increases computational demand. Third, our approach does not account for non-additivity of weather variables within the growing season. It is well known that weather could have varying effects throughout the season and there are approaches for harnessing these within-season nuances (e.g. Ortiz-Bobea et al., 2019). Despite these limitations, we think our approach is a step in the right direction and proposes an improvement relative to current practices. We hope this study motivates others to develop competing approaches to address these issues. [3]

The rest of the paper is organized as follows. Section 2 analyzes how non-classical measurement error arising from assuming the incorrect season introduces bias in estimated weather coefficients and associated climate change impact projections. In section 3 we describe the grid search approach to recover the "true" season. We then describe the data used in the subsequent simulation and real-world application in section 4. In section 5, we conduct a simulation to assess the consequences of assuming an incorrect season while also evaluating our proposed grid search approach. We present the real-world application in section 6 based on US state-level agricultural TFP growth rates. We conclude in section 7.

2 Conceptual framework

2.1 Set of relevant seasons

Suppose an agricultural outcome only responds to weather fluctuations during a particular time window, but remains insensitive to weather conditions during other periods of the year. As mentioned in the introduction, we define the former time period as the "relevant" season (e.g. one month or a few consecutive months). We denote the set of possible seasons as S which includes 78 elements. Each element is a season comprised of contiguous calendar months, defined by the starting and the ending month:

$$S = \{\{Jan\}, ..., \{Jan, ..., Dec\}, ..., \{Nov\}, \{Nov, Dec\}, \{Dec\}\}$$

Suppose the "true" season containing only relevant weather conditions is $T^* \in S$, and the researcher selects an arbitrary season $T \in S$ to aggregate the weather data. If the selected season $T = T^*$, there is no measurement error. However, a researcher could select an arbitrary season such that $T \neq T^*$. We consider four potential cases with incorrect seasons with some illustrative examples:

1. $T \subseteq T^*$: e.g. the researcher selects the summer but the relevant season is the entire year.

[3] Possible strategies include alternative model selection approaches based in LASSO and MARS, which could potentially recover the correct season and type of weather variable simultaneously. There is an emerging literature studying these model selection issues in the literature of climate econometrics, comparing the properties and trade-offs of existing selection procedures (Cui et al., 2018; Ghanem and Smith, 2020; Hendry).

2. $T \supseteq T^*$: e.g. the research selects the entire year but the relevant season is the summer.

3. $T \nsubseteq T^*$, $T \nsupseteq T^*$ and $T \cap T^* \neq \emptyset$: e.g. the researcher selects the period spanning May through August but the relevant season actually spans June through September.

4. $T \cap T^* = \emptyset$: e.g. the researcher selects the summer but the relevant season is the spring.

In each case, weather variables are mismeasured via aggregation using the selected time window. If irrelevant weather conditions were orthogonal to relevant weather conditions, this would introduce classical measurement error to weather variables and lead to attenuation bias in the estimation of weather effects. However, weather between adjacent or even distant periods of time within the same year are not uncorrelated, suggesting the measurement error is non classical as suggested by figure 1.

2.2 Bias in estimated coefficients

Here we show analytically how arbitrarily selecting a season such that $T \neq T^*$ biases estimates. Specifically, we derive a closed-form solution for a simple linear specification.

If the true relevant season is T^* for an agricultural outcome Y is known, then parameters could be consistently estimated by an OLS regression of Y on the weather variable W^* which is constructed by aggregating weather conditions over the time period T^*:

$$Y = \beta W^* + \epsilon \tag{1}$$

where β is the coefficient of interest and ϵ is an error term that is *iid* $[0, \sigma^2]$.

However, the true relevant season T^* is unknown and instead the research selects weather variable $W \neq W^*$ by aggregating weather conditions over the arbitrary season T, such that:

$$W = W^* + \nu \tag{2}$$

where the measurement error ν is potentially non-classical:

$$\text{plim } \frac{1}{n} W^{*'} \nu \neq 0 \tag{3}$$

$$\text{plim } \frac{1}{n} \epsilon' \nu \neq 0 \tag{4}$$

To assess the nature of the bias, consider the OLS estimator for β:

$$\hat{\beta} = \frac{cov(W^* + \nu, \beta W^* + \epsilon)}{var(W^* + \nu)} \tag{5}$$

and

$$\text{plim } \hat{\beta} = \frac{\beta(\sigma^2_{W^*} + \sigma_{W^*\nu})}{\sigma^2_{W^*} + \sigma^2_\nu + 2\sigma_{W^*\nu}} \tag{6}$$

$$= \left(1 - \frac{\sigma^2_\nu + \sigma_{W^*\nu}}{\sigma^2_{W^*} + \sigma^2_\nu + 2\sigma_{W^*\nu}}\right)\beta \tag{7}$$

$$= (1 - b_{\nu W})\beta \tag{8}$$

Notice that $b_{\nu W}$ is the regression coefficient of a regression of the measurement error ν on the selected weather variable W.

$$b_{\nu W} \equiv \frac{\sigma^2_\nu + \sigma_{W^*\nu}}{\sigma^2_{W^*} + \sigma^2_\nu + 2\sigma_{W^*\nu}} = \frac{Cov(\nu, W)}{Var(W)} \tag{9}$$

Classical measurement error is a special case when the measurement error is not correlated with true weather (i.e., $\sigma_{W^*\nu} = 0$) and equation boils down to:

$$1 - b_{\nu W} = \frac{\sigma^2_{W^*}}{\sigma^2_{W^*} + \sigma^2_\nu} \tag{10}$$

Since $1 - b_{\nu W} \in (0, 1)$, there will be attenuation bias in the estimates meaning the coefficient of interest $\hat{\beta}$ will be biased inconsistently towards 0.

When the measurement error is correlated with the true weather (i.e, $\sigma_{W^*\nu} \neq 0$) as we have shown in 1, the measurement error is non-classical. To analyze how the bias in the estimate $\hat{\beta}$ responds to the covariance between the measurement error and true weather variable $\sigma_{W^*\nu}$, we take the derivative of $1 - b_{\nu W}$ with respect to $\sigma_{W^*\nu}$, which has the sign of $\sigma^2_\nu - \sigma^2_{W^*}$.

If $\sigma^2_\nu > \sigma^2_{W^*\nu}$, which suggests more than half of the variance in the selected weather variable (i.e, $W = W^* + \nu$) is due to errors in measurement, $1 - b_{\nu W}$ is monotonically increasing with $\sigma_{W^*\nu}$. In this case, if the measurement error and the true weather are negatively correlated (i.e, $\sigma_{W^*\nu} < 0$), a smaller $\sigma_{W^*\nu}$ leads to a smaller attenuation factor $1 - b_{\nu W}$. This results in larger bias which may even reverse the sign of the estimated $\hat{\beta}$. If the measurement error and the true weather are positively correlated (i.e, $\sigma_{W^*\nu} > 0$), $1 - b_{\nu W} < 1$ where a more positive $\sigma_{W^*\nu}$ implies a larger attenuation factor $1 - b_{\nu W}$. This results in smaller attenuation bias.

If $\sigma^2_\nu < \sigma^2_{W^*\nu}$, indicating more than half of the variance in W comes from the true weather variable W^* instead of the measurement error, $1 - b_{\nu W}$ is monotonically decreasing with $\sigma_{W^*\nu}$. In this case, if the measurement error and the true weather are negatively correlated (i.e, $\sigma_{W^*\nu} < 0$), a more negative $\sigma_{W^*\nu}$ implies a larger attenuation factor $1 - b_{\nu W}$. However, we are uncertain about the direction of the bias since it depends on if the initial value of $1 - b_{\upsilon W^*}$ is smaller or greater than 1. If the measurement error and the true weather are positively correlated (i.e, $\sigma_{W^*\nu} > 0$), $1 - b_{\nu W} < 1$ and a more positive $\sigma_{W^*\nu}$ implies a lower attenuation factor $1 - b_{\nu W}$. This results in larger bias in the estimate which may reverse the sign of $\hat{\beta}$.

In summary, arbitrarily imposing a season essentially introduces non-classical measurement error in weather regressors, which causes unknown biases in weather impacts.

2.3 Bias in the estimated impact of climate change

Biases on coefficients of weather variables are difficult to interpret directly. In fact, it is conceivable that biased coefficients are not a sufficient condition for obtaining biased impact estimates. That is, it is possible that the projected changes in the distribution of weather (e.g. reflecting climate change) defined over incorrect seasons could somehow "compensate" for the aforementioned biases.We explore this problem in greater detail.

We first decompose the estimated climate change impact to understand the role of measurement error. Climate change impacts are typically computed based on the change in average weather conditions between two periods of time, a "reference" historical period, say 1990-2020, and a future "projection" period, say 2020-2050. For ease of notation we refer to these two periods as "1" and "2" respectively. W_1 and W_2 are weather variables defined over these two periods:

$$W_1 = W_1^* + \nu_1 \tag{11}$$

$$W_2 = W_2^* + \nu_2 \tag{12}$$

The change in climate can be expressed as:

$$\begin{aligned} E(\Delta W) &= E(W_2) - E(W_1) && (13)\\ &= E(W_2^* - W_1^*) + E(\nu_2 - \nu_1) && (14)\\ &= E(\Delta W^*) + E(\Delta\nu) && (15) \end{aligned}$$

The bias in the climate change impact estimate is:

$$E[\Delta\hat{Y}] - \Delta Y^* = E[\Delta W \cdot \hat{\beta}] - \Delta W^* \cdot \beta \tag{16}$$

Plugging equations (7) and (17) into (18), we further obtain:

$$\begin{aligned} E[\Delta\hat{Y}] - \Delta Y^* &= (\Delta W^* + E(\Delta\nu)) \cdot E(\hat{\beta}) - \Delta W^* \cdot \beta && (17)\\ &= (\Delta W^* + E(\Delta\nu)) \cdot (1 - b_{\nu_1 W})\beta - \Delta W^* \cdot \beta && (18)\\ &= [E(\Delta\nu) \cdot (1 - b_{\nu_1 W}) - \Delta W^* \cdot b_{\nu_1 W}] \cdot \beta && (19) \end{aligned}$$

Equation (21) suggests that the bias in the estimated climate change impact $\Delta\hat{Y}$ depends not only on the measurement error in the sample period used for estimation, but also on the measurement error in the projection period.

As an example, suppose we know the true season T_1^* in period 1. This means we can select the correct season and have no measurement error in W_1 so that $b_{\nu_1 W} = 0$. However, the future season in the projection period is unknown. Getting that future season wrong introduces measurement error that biases the projected impact. This illustrates the importance of incorporating timing considerations when computing projected climate change impacts.

Interestingly, observing the previous equation indicates there is one odd case when the measurement error in the reference and projection periods results in unbiased climate change impact estimates. That occurs when $E(\Delta\nu) \cdot (1 - b_{\nu_1 W}) - \Delta W^* \cdot b_{\nu_1 W} = 0$. In this case the bias in the coefficients is exactly compensated by the bias introduced by the mismatch of seasons between "reference" and "projection" periods.

3 A grid search approach to recover the true season

The guiding principle of our strategy is that a model relying on weather variables aggregated over the period matching the underlying DGP should provide a superior fit. We build on this idea to propose a grid search to identify the model with the best-fitting season out of a large set of calendar seasons within the year. We refer to the best-fitting season as the "optimal" season, which is effectively our estimate of the true season T^*. Our proposed approach is essentially an estimator of the true season.

We evaluate model fit based on a 10-fold year-block cross-validation, in which whole years of data are sampled together in a block. This approach seems suitable for common applications with high spatial but low serial correlation such as agriculture.[4]

Our grid search for season identification consists of three steps:

1. Perform a cross-validation of models based on weather variables aggregated over a wide range of candidate "seasons" within the year and store the out-of-sample MSE for each model.[5]

2. Visualize the out-of-sample MSE of every model on a grid. This visualization helps better appreciate the overall seasonality of the outcome variable.[6]

3. Select the "optimal" season corresponding to the model with the best out-of-sample model fit.[7]

In the Monte Carlo simulation, we evaluate how effective this approach is at recovering the true season T^* in a realistic setting with a relatively short panel dataset.

4 Data

4.1 Data sources and data processing

For the real-world application, we rely on agricultural TFP growth rates for the lower 48 US states, which we obtain from USDA Economic Research Service (ERS).[8] TFP is defined as aggregate output divided by

[4] Needless to say, the cross-validation approach could be adjusted to accommodate serially dependent data (see Arlot and Celisse, 2010).

[5] If we restrict ourselves to seasons defined as periods of contiguous months over the 12-month calendar year, we have 78 possibilities. One could conceivably extend this approach to include weather from the previous calendar years.

[6] This could reveal if there is more than one portion of the year that is relevant for the outcome variable.

[7] This criteria could be adjusted to constrain seasons to be of a certain length (e.g. best 3-month-long seasons), allowing the researcher to incorporate background information about the timing of relevant weather conditions.

[8] Unfortunately, the dataset has not been updated after 2004 because a critical source of labor information, the Farm Labor Survey, was discontinued. The methodology for estimating TFP is described on the ERS website (`www.ers.usda.gov/data-products/agricultural-productivity-in-the-us/methods/`). Link to the data: `www.ers.usda.gov/data-products/agricultural-productivity-in-the-us/`

Table 1: Summary Statistics for Annual Weather Variables (1961-2004)

	Season	Max	Min	Mean	St.d
	Winter	23.5	-3.6	7.8	6.3
	Spring	33.8	16.9	23.1	3.8
Tmax (°C)	Summer	38.7	23.5	28.7	3.0
	Fall	27.4	8.1	16.1	4.7
	Annual	29.6	11.0	18.1	4.5
	Winter	16.7	-9.1	1.9	5.8
	Spring	24.3	10.9	16.3	3.8
Tmean (°C)	Summer	30.6	17.5	21.9	3.1
	Fall	21.4	2.1	9.6	4.3
	Annual	22.0	4.7	11.7	4.3
	Winter	10.0	-14.6	-4.1	5.5
	Spring	17.9	2.8	9.6	3.9
Tmin (°C)	Summer	22.5	8.5	15.1	3.7
	Fall	15.4	-4.0	3.2	4.2
	Annual	15.9	-1.6	5.3	4.2
	Winter	411.4	38.3	195.5	105.7
	Spring	387.2	12.6	256.5	96.2
PPT (mm)	Summer	537.8	9.6	245.6	107.6
	Fall	225.6	28.7	137.9	60.3
	Annual	1500.5	198.5	906.4	368.9

Notes: We rely on weather data aggregated to the state level based on cropland weights for these summary statistics. The seasons are defined as follows. Winter (Jan-Mar), Spring (Apr-Jun), Summer (Jul-Sep), Fall (Oct-Dec) and Annual (Jan-Dec).

aggregate input. There is no obvious "season" to such data given their aggregate nature encompassing most agricultural outputs and inputs. We conduct our analysis relying on this 44-year panel from 1960-2004 and estimate regression models by USDA Climate Hub regions (see figure A.1).

We rely on weather data for both the simulation and the real-world application. We obtain fine-scale gridded weather data from the PRISM Climate group at Oregon State University (`http://prism.oregonstate.edu`). The PRISM group provides weather data for the continental US since 1895 with a 4-km spatial resolution. For this study we obtain data for minimum and maximum temperature (Tmin, Tmax) and precipitation over the 1961-2004 period overlapping with the agricultural data.

Because weather conditions can vary substantially within a state, accounting for the location of agricultural activities may be critical when constructing appropriate state-level weather variables. We therefore aggregate the fine-scale weather data to the state level using cropland weights based on cropland counts within each PRISM grid cell based on USDA's Crop Data Layer (CDL, 30m) over 2008-2014. We depict the cropland weights in figure A.2.

4.2 Summary statistics

Table 1 provides summary statistics for annual and state weather conditions across the US over 1961-2004. As expected, climatic conditions vary considerably across seasons with winter exhibiting slightly more variable weather conditions. Also, summers appear to be wetter.

Table 2 summarizes statistics for the dependent variable, the annual TFP growth rate. The average

Table 2: Summary statistics for annual agricultural TFP growth rate by USDA Climate Hub region (1961-2004)

Region\ TFP growth rate(%)	Max	Min	Mean	St.d
Pacific Northwest	15.3	-9.5	2.0	5.0
Northern Plains	40.8	-76.4	0.8	11.0
Midwest	28.6	-40.2	1.5	9.6
Northeast	25.4	-18.7	1.6	6.3
Southwest	17.0	-19.1	1.3	5.3
Southern Plains	19.4	-27.6	0.7	7.3
Southeast	25.8	-31.8	1.3	8.0
National	40.8	-76.4	1.4	7.9

TFP growth rate across the lower 48 states is around 1.4% over the 1961-2004 sample period but there is substantial regional variation. For instance, agricultural TFP grew at a relatively modest rate of 0.8% and 0.7% in the Northern Plains and Southern Plains, respectively. In contrast, TFP grew at faster but more stable rates around 2.0% and 1.3% in the Pacific Northwest and Southwest region, respectively. We also see higher standard deviations in Eastern regions of the country where there is relatively less irrigation and agriculture is more sensitive to weather fluctuations.

5 Simulation

5.1 Description

The overall goal here is to study the consequences of selecting an incorrect season to aggregate weather variables. Specifically, we're interested in understanding how the non-classical measurement error resulting from this practice can bias estimates of climate change impacts in terms of magnitude and direction. Another goal here is to evaluate the effectiveness of our proposed grid search approach in recovering the true season T^* when we have no *a priori* knowledge about the timing of relevant weather conditions for the outcome variable. For this purpose, we conduct a Monte Carlo simulation in which we apply our grid search approach in a context where we obviously control the DGP. We propose a common setting with a relatively short panel of 10 years of weather data for 48 states (2000-2010).

We express the DGP in general terms as:

$$Y_{it} = \beta_1 W_{it}^* + \beta_2 (W_{it}^*)^2 + \alpha_i + \epsilon_{it} \tag{20}$$

where Y_{it} is the generated outcome variable in state i in year t, and W_{it}^* is the weather variable aggregated over a hypothetical "true" season. We consider DGPs with alternative weather variables where $W \in \{Tmin, Tmax, Tmean, PPT\}$ and the true relevant seasons $T^* \in \{Winter, Spring, Summer, Fall, Annual\}$. Considering all these variations leads to 4 variables $\times$ 5 seasons = 20 different DGPs.

The α_i term is a state-level fixed effect drawn from a uniform distribution $U(0, 1)$ and ϵ_{it} is a normally distributed error $N(0; 10^2)$. We choose parameters $(\beta_1, \beta_2) = (12, -0.5)$ which yields a response function with an "inverted U" shape peaking at $W^* = 12$. Increases in W^* beyond that point are detrimental to the outcome variable. For simplicity, we rely on these same parameter values for both temperature (measured in °C) and precipitation (measured in mm).

We subsequently conduct a regression analysis to explore the consequences of selecting incorrect seasons and to examine the effectiveness of the grid search approach. The general regression model can be expressed as:

$$Y_{it} = \beta_1 W_{it}^T + \beta_2 (W_{it}^T)^2 + a_i + u_{it} \tag{21}$$

where Y_{it} is the simulated outcome variable described above and W_{it}^T is the weather variable aggregated over an arbitrarily selected season T, with $W \in \{Tmax, Tmin, Tmean, PPT\}$. This means that the selected weather variable and season need not match those in the DGP. Also, a_i represents a state fixed effect and u_{it} is the error term. We also consider a "baseline" model that excludes weather variables for the purpose of comparing model fit.

For each DGP we consider models based on all 4 weather variables to evaluate the performance of our approach when the selected weather variable does not match the true variable type in the DGP. This amounts to 4 estimation variables × 20 DGPs which adds to 80 simulations. In addition, we perform 450 iterations for each simulation. In each iteration, we apply our grid search approach with the goal of deriving the sampling distribution of $\hat{T}$. Recall that our grid search approach consists of a 10-fold cross-validation for models based on every calendar period T in the choice set S with 78 possibilities as we defined in 2.[9] This means that we conduct 10 regressions × 78 seasons = 780 regressions for each iteration, or 780 regressions × 450 iterations = 351,000 regressions for each simulation. That amounts to around 30 million regressions for our entire set of 80 simulations.

5.2 Simulated impacts with correct type of weather variable

Here we explore the consequence of aggregating weather variables over incorrect seasons when the researcher selects the correct type of weather variable (e.g. researcher selects Tmax when the DGP is also based on Tmax).

To quantify this we compute the impact of a 1°C uniform warming on the simulated outcome variable. We do this for every model estimated based on a wide range of candidate seasons within the calendar year and for a model based on the "true" season. We compare these impacts estimates by computing a "bias ratio" that captures how the estimate $\Delta\hat{Y}$ deviates from the true impact ΔY^* in fractional terms.

$$\text{Bias ratio} = \frac{\Delta\hat{Y} - \Delta Y^*}{\Delta Y^*}$$

Figure 2 summarizes the results. Each column of panels indicate the type of weather variable used in both the DGP and the estimation (because the researcher selected the correct type of weather variable). Each row of panels corresponds to a different "true" season used in the DGP (Winter, Spring, etc). The color in each panel indicates the bias ratio for a model based on an arbitrary selected season.

Naturally, when the researcher selects the correct season, $T = T^*$, the bias ratio is close to 0 which is indicated by grey areas in figure 2. Interestingly, selected seasons that are "close" to the true season tend to

[9] Here it is a leave-on-year out cross-validation given we have 10 years of data.

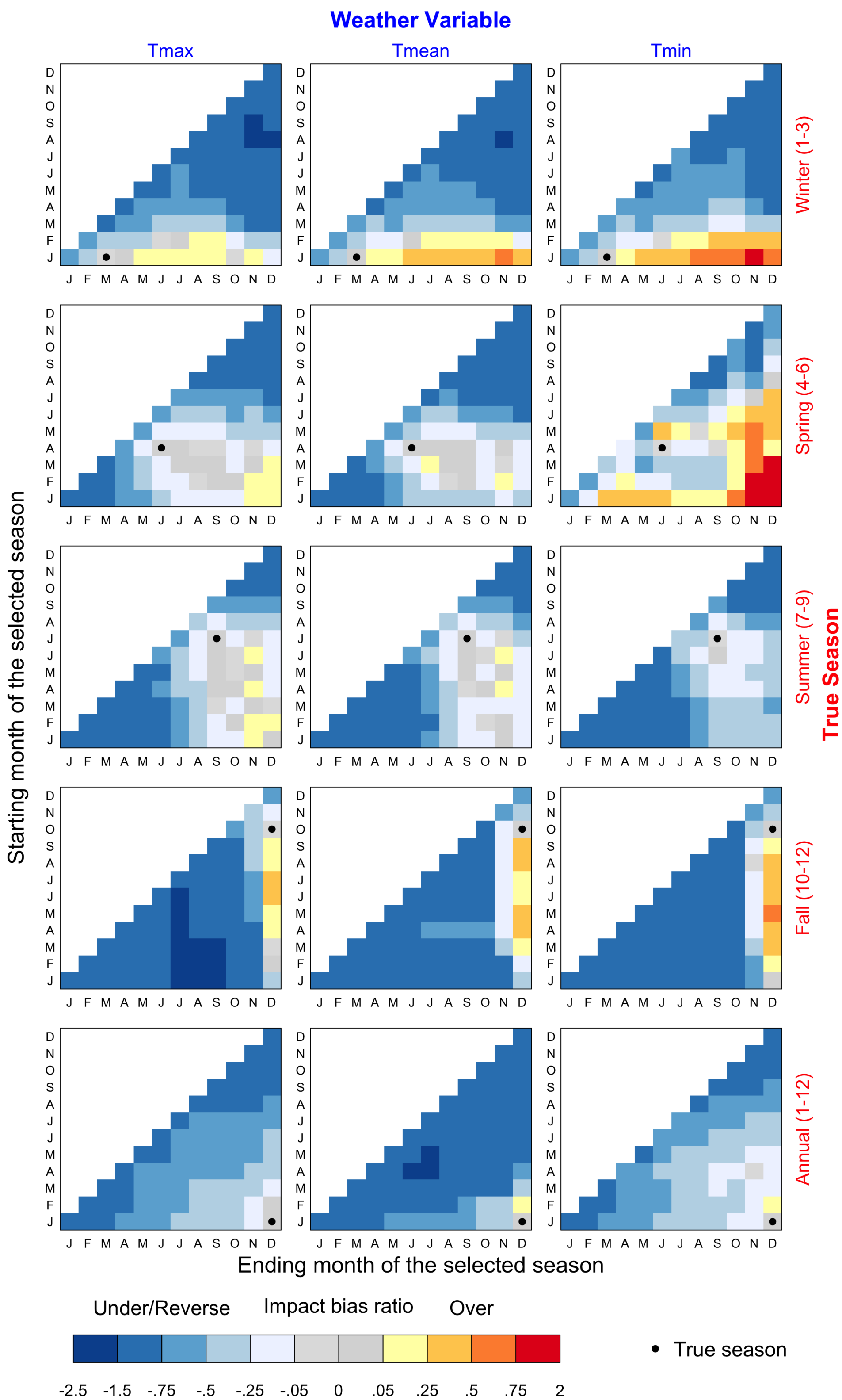


Figure 2: Simulated results for the impact of 1°C uniform warming

exhibit small impact bias ratios. This is encouraging when researchers have some background information about what the "true" season might be.

However, there are plenty of blue areas in figure 2, indicating that a wide range of selected seasons lead to negative bias ratios. That is, these incorrect modeling decisions lead to underestimation of projected impacts of a uniform 1°C warming. There is also a non-negligible share of selected seasons exhibiting positive bias ratios. That is, selecting models based on those seasons tend to overestimate impacts of a uniform 1°C warming.

Importantly, selecting the whole year as the season, a common modeling decision, can lead to over or under estimation of projected impacts depending on the true season in the DGP.

These results are in line with our conceptual framework. We find that selecting the incorrect season can introduce non-classical measurement error which results in bias in weather impacts that can be in either directions. While the simulation results reflect weather correlations occurring in the US, such patterns could arise elsewhere in different ways. This means that there is no *a priori* way to determine the direction of the bias when selecting an incorrect season.

5.3 Simulated impacts with incorrect type of weather variable

We now explore the consequence of aggregating weather variables over incorrect seasons when the researcher selects the incorrect type of weather variable (e.g. researcher selects Tmin when the DGP is based on Tmax). Figure 3 summarizes the results for simulations based on a true season corresponding to the summer (July-September).[10]

The layout of this figure is similar to figure 2 but now the columns and rows indicate the type of weather variable used in the DGP and in the estimation, respectively. For instance, the top center panel corresponds to a simulation where the DGP is based on Tmean, but the estimation variable used is Tmax. Similarly, the colors in each cell within a panel indicates the impact bias ratio of the selected model relative to the true model.

As expected, the diagonal panels shows that when the researcher selects the correct type of weather variables and the true season, the impact bias ratio is close to zero. The off-diagonal panels, on the other hand, indicate the simulations when the type of variable selected in the estimation does not match the variable in the DGP. Selecting the incorrect type of weather variable can on itself, even when selecting the correct season, bias impact estimate in either direction. Not surprisingly, selecting seasons that are very different from the true season leads to biases, which can be positive or negative depending on the case.

5.4 Evaluating the proposed grid search approach

Our previous analysis shows that selecting the incorrect season can lead to substantial biases in directions that are virtually unpredictable. Therefore approaches that help researchers make better assumptions about the timing of relevant weather conditions could be particularly fruitful. Here we evaluate whether our proposed grid approach provides useful guidance regarding season selection.

[10]The results for other seasons such as spring, fall, winter and the whole year are similar as the one for the season of summer. We include these figures in the Appendix. See figuresA.3, A.4, A.5 and A.6.

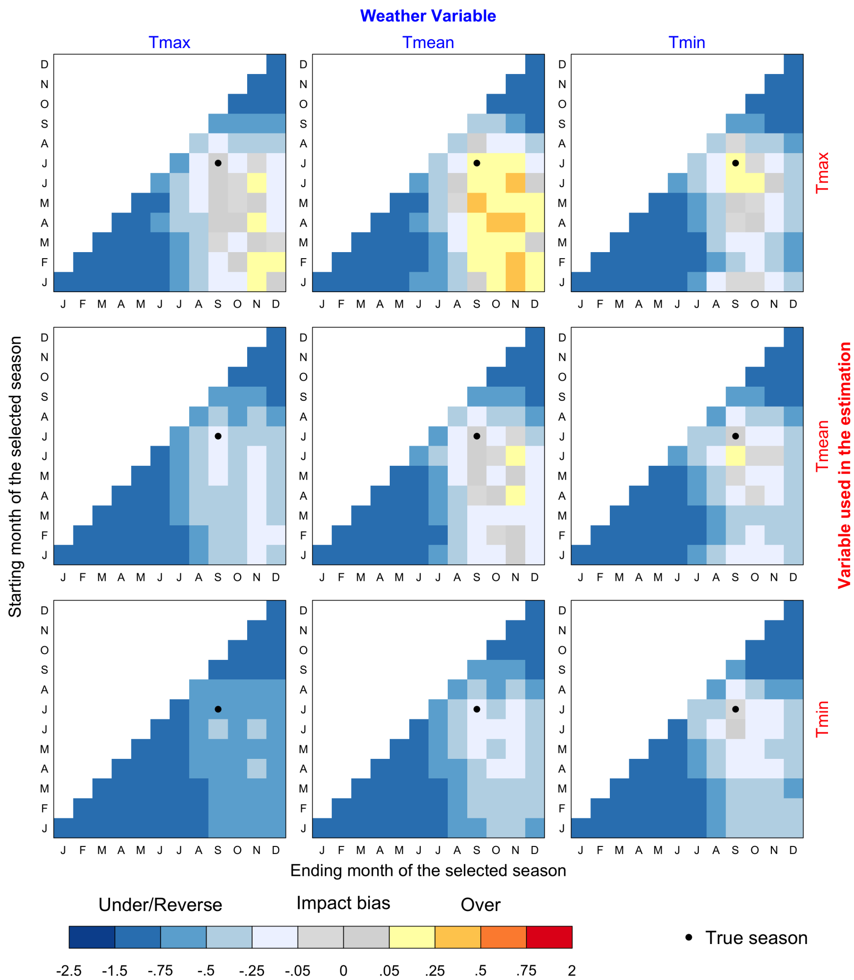


Figure 3: Simulated results for the impact of weather shock (cross variables and summer season)

Figure 4 shows the result of the grid search using the same type of weather variable in the DGP and in the estimation. Columns of panels indicate the type of weather variables used in both of the DGP and the estimation underlying the grid search. Rows of panels indicate the true season in the DGP. Similar to previous figures, the y-axis and x-axis in each panel indicate the season starting and ending month of seasons selected by the grid search algorithm, respectively. The color shown in each cell now corresponds to the density of the "optimal" season indicated by the grid search approach across the 450 iterations of the simulation. The true season is indicated with a solid circle while the most frequent "optimal" season obtained form the algorithm is shown with a hollow square symbol.

Figure 4 shows that, on average, our grid search approach correctly recovers the "true" season. Visually this can be seen by the coincidence of the square and the circle symbols in each panel. In other words, if the true season is part of the set of candidate seasons, models based on that season will tend to fit better out of sample. Note there is some density outside the true season, meaning that our approach does not always identify the correct season due to sampling variability.

Figure 5 summarizes results when the type of weather variable in the DGP and the estimate underlying the grid search do not match and the true season corresponds to the summer (July-September).[11] The columns of panels correspond to weather variables used in the DGP while the rows of panels correspond to the weather variables used in the estimation underlying the grid search. The panels on the diagonal correspond to cases for which the researcher selects the correct variable type matching the DGP. The panel off the diagonal corresponds to cases when the researcher selects an incorrect type of weather variable not matching the DGP.

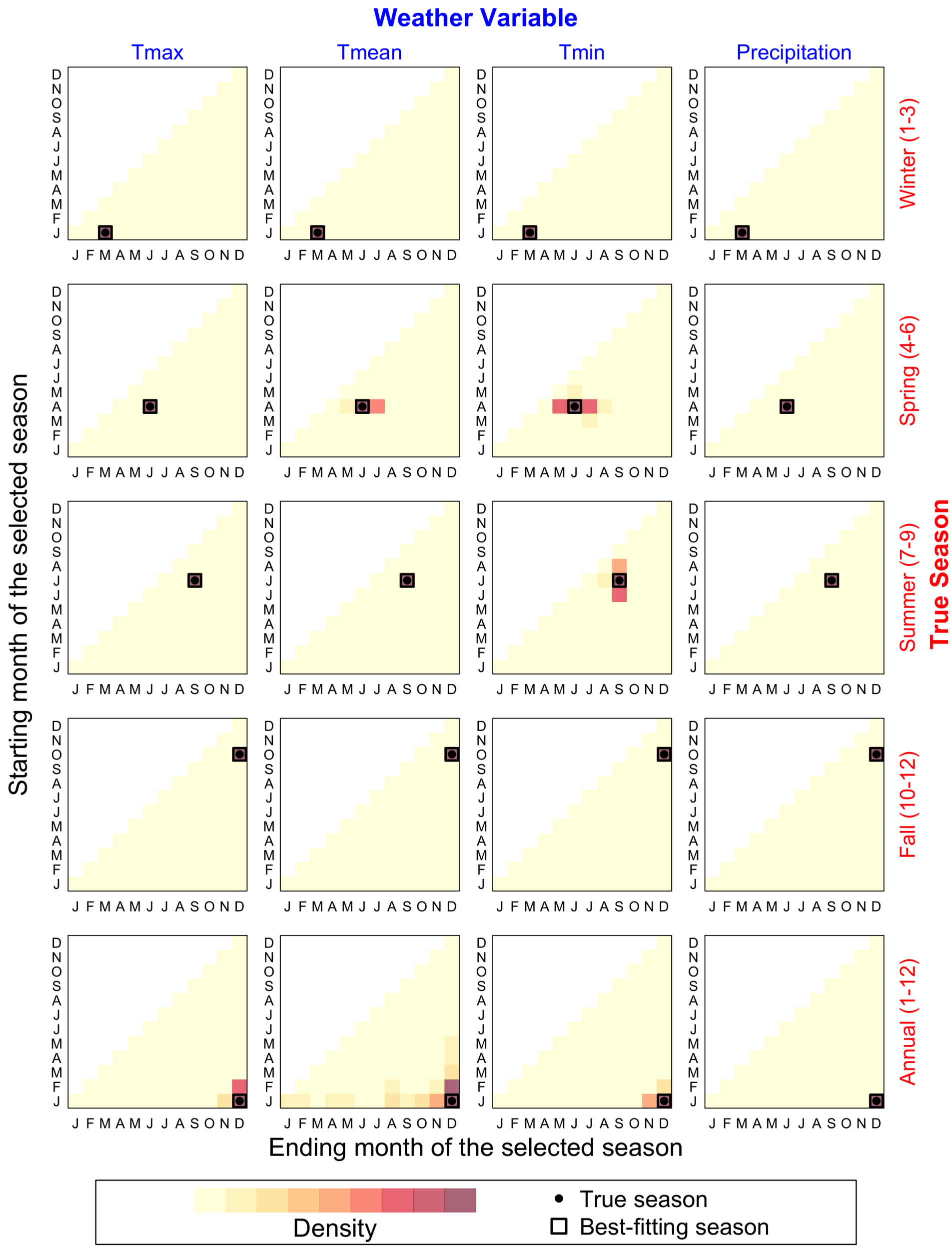
Figure 5 yields some additional insights. Our proposed grid search approach can often recover the true season even when the selected type of weather variable is incorrect. This suggests there are strong correlations between weather variables which allows our approach to still recover the underlying seasonality even with the incorrect type of weather variable. The less promising case is when one uses Tmin to recover the seasonality of a DGP based on precipitation. This suggests these variables are not strongly correlated in the summer.

In summary, our approach is capable of recovering, on average, the underlying "true" seasonality. This is the case even in cases when the selected weather variables do not exactly match the DGP. This encouraging because of our previous analyses show how selecting the incorrect season can lead to substantial biases in directions that are difficult to predict. We thus thing our approach provides a valuable contribution to empirical researchers in this literature.

6 Application

We now turn our attention to a real-world application. In this case we obviously ignore the underlying DGP but we can nonetheless apply our grid search approach to recover the plausible underlying seasonality of the data. Specifically, we estimate the nonlinear effects of temperature on agricultural TFP growth using a state-level panel with state fixed effects. This means that similar to our simulations, our identifying variation stems from the short-run year-to-year fluctuations in weather variables and TFP growth. For each USDA Climate Hub region, we estimate a general econometric model of the form:

[11] Results for other seasons (spring, fall, winter and annual) are similar. We include these results in the Appendix. See figures A.7, A.8, A.9 and A.10.



Notes: Columns of panels indicate the type of weather variable used in both of the DGP and the estimation underlying the grid search. Rows of panels indicate the true season in the DGP. The y-axis and x-axis in each panel indicate the season starting and ending month of seasons selected by the grid search algorithm, respectively. The color shown in each cell corresponds to the density of the "optimal" season indicated by the grid search approach across the 450 iterations of the simulation. The true season is indicated with a solid circle while the most frequent "optimal" season obtained form the algorithm is shown with a hollow square symbol.

Figure 4: Simulated results of the grid search approach

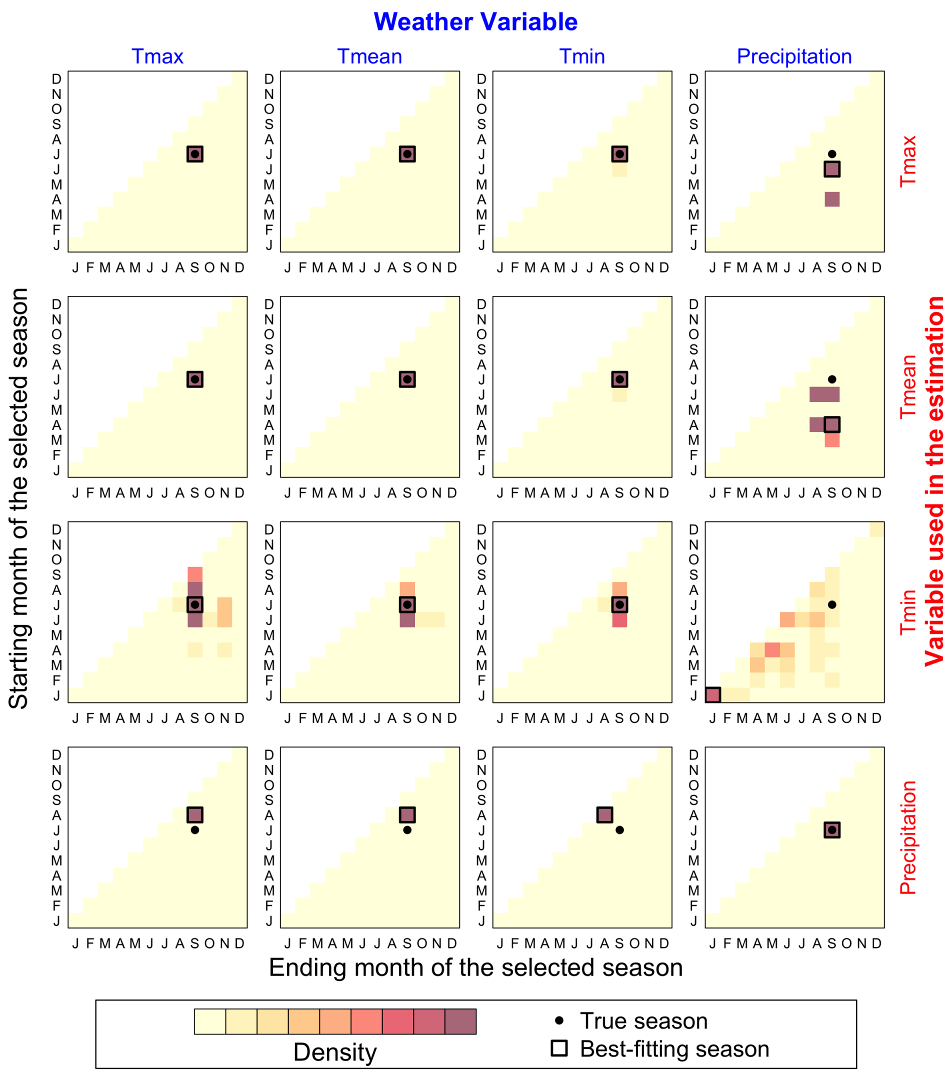


Figure 5: Simulated results of the grid search approach (cross variables and summer season)

$$Y_{it} = \beta_1 W_{it}^T + \beta_2 (W_{it}^T)^2 + a_i + u_{it}$$

where Y_{it} is TFP growth in state i in year t, W_{it}^T is maximum temperature variable over the T calendar period within the year t, α_i represents a state fixed effect and u_{it} is an error term.

6.1 Season grid search results

We illustrate our estimate for the true season in figure 6. Results for each region are presented in separate panels. Each panel shows the fit of models based on the 78 possible calendar seasons within contiguous months within the calendar year. Model fit is captured by the reduction in out-of-sample MSE reduction relative to a model without weather variables. That means, that higher reductions indicate better model fit. These higher levels of fit are indicated in blue.

For the Midwest region, the largest reduction in MSE corresponds to a model based on the the July-August season. This precisely coincides with the critical period of the growing season for corn and soybeans, which are field crops grown intensively in this region. Importantly, the set of seasons that appears to improve explanatory power relatively to a model without weather is relatively small. For instance, periods outside the key portions of the growing season (seasons starting after September or ending prior to April) provide little to no improvements in model fit.

These findings are encouraging and suggest our approach may be effective in identifying the timing of relevant weather conditions for agriculture in realistic settings.

6.2 Impact of uniform warming scenarios

Regression coefficients for weather variables are difficult to interpret so we focus on the sensitivity of TFP growth to a uniform warming. This approach also allows comparing projected impacts across regions.

Table 6.2 shows the impact of a 1°C warming for models based on weather variables aggregated according to various seasons. The first column corresponds to a model based on our estimate of the true season. The subsequent columns show results for models based on arbitrary seasons, corresponding to the winter, spring, summer, fall and the entire year. Rows in the table indicate regions. Entries in the table correspond to the impact estimate in log points for each model and region.

For models based on our estimate of the true season (first column), we detect a statistically significant effect of the 1°C warming in all regions except in the Pacific Northwest.[12] Unsurprisingly, the projected impacts are particularly noticeable in the Midwest region. A uniform 1°C warming appears to decrease TFP growth by about 4%.

Importantly, and as anticipated by our simulations, lengthening the season to the whole year tends to bias impact estimates. In this case the bias is towards 0. We also find that the models based on our estimate of the true season do not systematically point to larger or smaller impacts than models based on arbitrary seasons (that are presumably incorrect).

[12]Generally, estimates based on the arbitrary seasons tend to be similar if those seasons are close to our estimate of the true season.

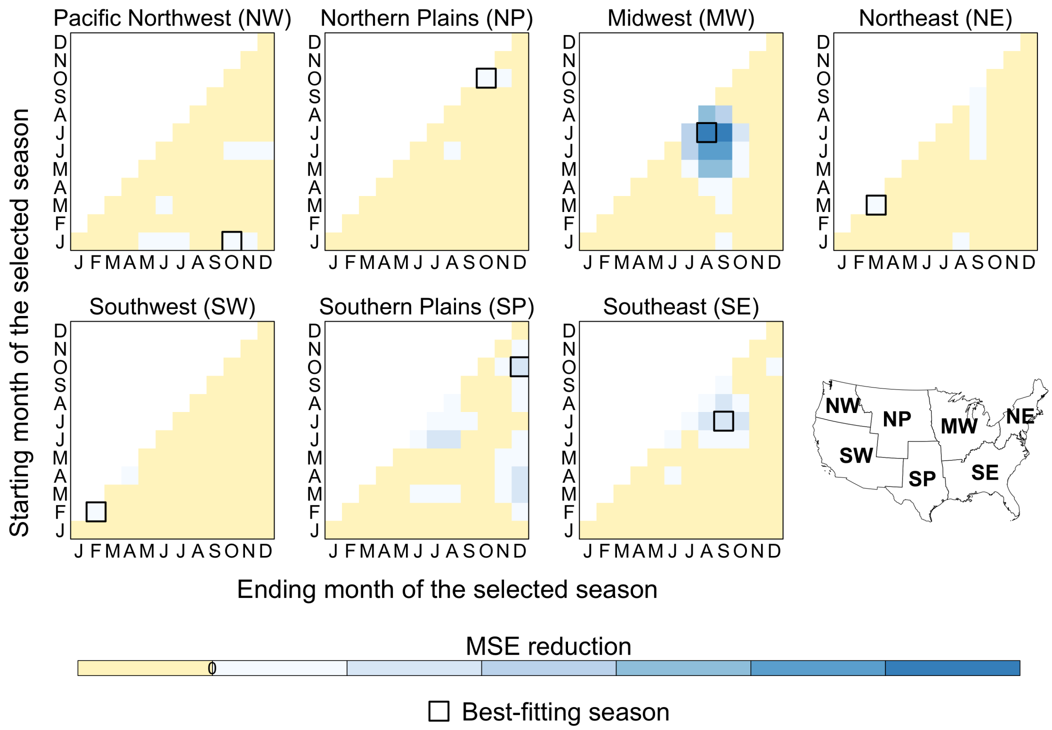


Figure 6: Estimates of the true season based on the proposed grid search by USDA Climate Hub region.

Table 3: Impact of a uniform 1°C warming on agriculture TFP growth rate

Region \Season	Optimal		Winter (1-3)		Spring (4-6)		Summer (7-9)		Fall (10-12)		Annual (1-12)
Pacific Northwest	0.0067		0		-0.0017		0.0032		0		0.0079
	(0.0056)		(0.0038)		(0.0041)		(0.0042)		(0.0048)		(0.0088)
Northern Plains	-0.0303	***	0.003		-0.0129	***	-0.0155	**	0.0041		-0.0035
	(0.0065)		(0.0033)		(0.0046)		(0.0065)		(0.0041)		(0.0082)
Midwest	-0.0438	***	-0.0038		0.0049		-0.0471	***	0.0128	***	-0.0126
	(0.0039)		(0.003)		(0.0044)		(0.0047)		(0.0035)		(0.0076)
Northeast	-0.0041	***	-0.0063	***	-0.0016		-0.0176	***	0.0084	***	-0.0066
	(0.0012)		(0.0021)		(0.0034)		(0.0036)		(0.0023)		(0.005)
Southwest	-0.0043	*	-0.004		0.0108	***	-0.0031		0.0014		0.0101
	(0.0025)		(0.0029)		(0.0029)		(0.0044)		(0.0029)		(0.0064)
Southern Plains	-0.0164	***	-0.0022		-0.0065		-0.0174	***	0.0105	*	-0.0044
	(0.0041)		(0.0038)		(0.0072)		(0.0057)		(0.0054)		(0.0113)
Southeast	-0.0396	***	0		0.0133	***	-0.0396	***	0.0177	***	0.0019
	(0.0038)		(0.0027)		(0.005)		(0.0038)		(0.0035)		(0.0074)

Notes: Symbols *, **, ***, correspond to significance levels of 10, 5, 1 percent. Optimal seasons for each region: New England (Sep-Dec), Mideast (Apr-Aug), Great Lakes (May-Sep), Plains (May-Aug), Southeast (May-Aug), Southwest (Jan-Apr), Rocky Mountain (May-Sep), Far West (Jun-Oct).

7 Conclusion

There is a growing literature documenting the impacts of weather shocks on agricultural outcomes. These studies require researchers to make assumptions about the timing of relevant weather conditions. These modeling choices are not always obvious.

In this study we show that these seemingly innocuous modeling decisions can have serious consequences. We show that selecting an incorrect season essentially introduces non-classical measurement error in weather regressors. This translates into biases in unknown direction in the estimation of weather effects and impacts. We confirm these issues analytically, in simulations and find suggestive evidence this is occurring in a real-world application to US agricultural TFP growth rates. This raises important questions about the robustness of impact estimates especially in contexts dealing with aggregate agricultural data in which timing issues are more elusive.

To deal with these issues, we propose a grid search approach to recover the true underlying seasonality from the data. The underlying principle is that if the selected season by the researcher matches the true season in the DGP, than this should confer a superior model fit. We harness this idea and develop a cross-validation grid search approach that assesses the out-of-sample fit of models based on a wide range of candidate seasons. We provide evidence based on simulations that our approach is effective at recovering the true seasons under certain assumptions, sometimes even when the type of weather variable selected does not that match the DGP.

Applying our approach to a state-level panel of US agricultural TFP growth rates, indicates that our approach points to substantially different impact estimates than models based on commonly used seasons. This suggest that our approach could have important practical implications in developing more robust estimates of climate change impacts on agriculture and possibly in other contexts.

Our approach, however, has several caveats that we think future research could address. One of the limitations is our current implementation imposes additivity of weather variables within the selected season. We also do not consider cases with multiple seasons or different set of weather variables. More work can certainly improve this attempt to bring some much needed guidance to the empirical analysis of weather effects on agricultural production.

References

Abay, K. A., Abate, G. T., Barrett, C. B., and Bernard, T. (2019). Correlated non-classical measurement errors, 'Second best' policy inference, and the inverse size-productivity relationship in agriculture. *Journal of Development Economics*, 139(C):171–184. Publisher: Elsevier.

Addoum, J. M., Ng, D. T., and Ortiz-Bobea, A. (2019). Temperature Shocks and Establishment Sales. SSRN Scholarly Paper ID 3411225, Social Science Research Network, Rochester, NY.

Arlot, S. and Celisse, A. (2010). A survey of cross-validation procedures for model selection. *Statistics Surveys*, 4(0):40–79.

Arthi, V., Beegle, K., De Weerdt, J., and Palacios-López, A. (2018). Not your average job: Measuring farm labor in Tanzania. *Journal of Development Economics*, 130:160–172.

Baird, S. and Özler, B. (2012). Examining the reliability of self-reported data on school participation. *Journal of Development Economics*, 98(1):89–93. Publisher: Elsevier.

Beegle, K., De Weerdt, J., Friedman, J., and Gibson, J. (2012). Methods of household consumption measurement through surveys: Experimental results from Tanzania. *Journal of Development Economics*, 98(1):3–18. Publisher: Elsevier.

Borjas, G. J. (1980). The Relationship between Wages and Weekly Hours of Work: The Role of Division Bias. *The Journal of Human Resources*, 15(3):409.

Bound, J. and Krueger, A. B. (1991). The Extent of Measurement Error in Longitudinal Earnings Data: Do Two Wrongs Make a Right? *Journal of Labor Economics*, 9(1):1–24. Publisher: [University of Chicago Press, Society of Labor Economists, NORC at the University of Chicago].

Burke, M., Hsiang, S. M., and Miguel, E. (2014). Climate and Conflict. Working Paper 20598, National Bureau of Economic Research.

Burke, M., Hsiang, S. M., and Miguel, E. (2015). Global non-linear effect of temperature on economic production. *Nature*, 527(7577):235–239.

Cachon, G., Gallino, S., and Olivares, M. (2012). Severe Weather and Automobile Assembly Productivity. *SSRN Electronic Journal*.

Carleton, T. A. and Hsiang, S. M. (2016). Social and economic impacts of climate. *Science*, 353(6304):aad9837.

Carletto, C., Zezza, A., and Banerjee, R. (2013). Towards better measurement of household food security: Harmonizing indicators and the role of household surveys. *Global Food Security*, 2(1):30–40.

Cattaneo, C., Beine, M., Fröhlich, C. J., Kniveton, D., Martinez-Zarzoso, I., Mastrorillo, M., Millock, K., Piguet, E., and Schraven, B. (2019). Human Migration in the Era of Climate Change. *Review of Environmental Economics and Policy*, 13(2):189–206.

Cui, X., Ghanem, D., and Kuffner, T. (2018). On model selection criteria for climate change impact studies. *arXiv:1808.07861 [stat]*. arXiv: 1808.07861.

Dell, M., Jones, B. F., and Olken, B. A. (2012). Temperature Shocks and Economic Growth: Evidence from the Last Half Century. *American Economic Journal: Macroeconomics*, 4(3):66–95.

Deschênes, O. and Greenstone, M. (2011). Climate Change, Mortality, and Adaptation: Evidence from Annual Fluctuations in Weather in the US. *American Economic Journal: Applied Economics*, 3(4):152–185.

Desiere, S. and Jolliffe, D. (2018). Land productivity and plot size: Is measurement error driving the inverse relationship? *Journal of Development Economics*, 130(C):84–98. Publisher: Elsevier.

Dirks, J. A., Gorrissen, W. J., Hathaway, J. H., Skorski, D. C., Scott, M. J., Pulsipher, T. C., Huang, M., Liu, Y., and Rice, J. S. (2015). Impacts of climate change on energy consumption and peak demand in buildings: A detailed regional approach. *Energy*, 79:20–32.

French, E. (2004). The Labor Supply Response to (Mismeasured but) Predictable Wage Changes. *The Review of Economics and Statistics*, 86(2):602–613.

Ghanem, D. and Smith, A. (2020). What are the Benefits of High-frequency Data for Fixed Effects Panel Models? Publisher: Harvard Dataverse type: dataset.

Gibson, J., Beegle, K., Weerdt, J. D., and Friedman, J. (2015). What does Variation in Survey Design Reveal about the Nature of Measurement Errors in Household Consumption? *Oxford Bulletin of Economics and Statistics*, 77(3):466–474. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/obes.12066.

Gibson, J. and Kim, B. (2010). Non-Classical Measurement Error in Long-Term Retrospective Recall Surveys*: Non-classical measurement error. *Oxford Bulletin of Economics and Statistics*, 72(5):687–695.

Hendry, D. F. First in, First out: Econometric Modelling of UK Annual CO2 Emissions, 1860–2017. page 28.

Hsiang, S. M. and Burke, M. (2014). Climate, conflict, and social stability: what does the evidence say? *Climatic Change*, 123(1):39–55.

Hsiang, S. M., Burke, M., and Miguel, E. (2013). Quantifying the Influence of Climate on Human Conflict. *Science*, 341(6151):1235367.

Hsiang, S. M., Meng, K. C., and Cane, M. A. (2011). Civil conflicts are associated with the global climate. *Nature*, 476(7361):438–441.

Jones, B. F. and Olken, B. A. (2010). Climate Shocks and Exports. *American Economic Review*, 100(2):454–459.

Kim, B. and Solon, G. (2005). Implications of Mean-Reverting Measurement Error for Longitudinal Studies of Wages and Employment. *The Review of Economics and Statistics*, 87(1):193–196.

Larsen, A. F., Headey, D., and Masters, W. A. (2019). Misreporting Month of Birth: Diagnosis and Implications for Research on Nutrition and Early Childhood in Developing Countries. *Demography*, 56(2):707–728.

Liang, X.-Z., Wu, Y., Chambers, R. G., Schmoldt, D. L., Gao, W., Liu, C., Liu, Y.-A., Sun, C., and Kennedy, J. A. (2017). Determining climate effects on US total agricultural productivity. *Proceedings of the National Academy of Sciences*, 114(12):E2285–E2292.

Ortiz-Bobea, A., Ault, T. R., Carrillo, C. M., Chambers, R. G., and Lobell, D. B. (2021). The Historical Impact of Anthropogenic Climate Change on Global Agricultural Productivity. *Nature Climate Change*, 11(4):306–312. arXiv: 2007.10415.

Ortiz-Bobea, A., Knippenberg, E., and Chambers, R. G. (2018). Growing climatic sensitivity of U.S. agriculture linked to technological change and regional specialization. *Science Advances*, 4(12):eaat4343.

Ortiz-Bobea, A., Wang, H., Carrillo, C. M., and Ault, T. R. (2019). Unpacking the climatic drivers of US agricultural yields. *Environmental Research Letters*, 14(6):064003. Publisher: IOP Publishing.

Schlenker, W. and Roberts, M. J. (2009). Nonlinear temperature effects indicate severe damages to U.S. crop yields under climate change. *Proceedings of the National Academy of Sciences*, 106(37):15594–15598.

Stern, N. (2013). The Structure of Economic Modeling of the Potential Impacts of Climate Change: Grafting Gross Underestimation of Risk onto Already Narrow Science Models. *Journal of Economic Literature*, 51(3):838–859.

Tack, J. B., Barkley, A., and Nalley, L. (2015). Effect of warming temperatures on US wheat yields. *Proceedings of the National Academy of Sciences*.

Zhang, P., Deschenes, O., Meng, K., and Zhang, J. (2018). Temperature effects on productivity and factor reallocation: Evidence from a half million chinese manufacturing plants. *Journal of Environmental Economics and Management*, 88:1–17.

Zivin, J. G., Hsiang, S. M., and Neidell, M. (2018). Temperature and Human Capital in the Short and Long Run. *Journal of the Association of Environmental and Resource Economists*, 5(1):77–105.

Appendix

Figures

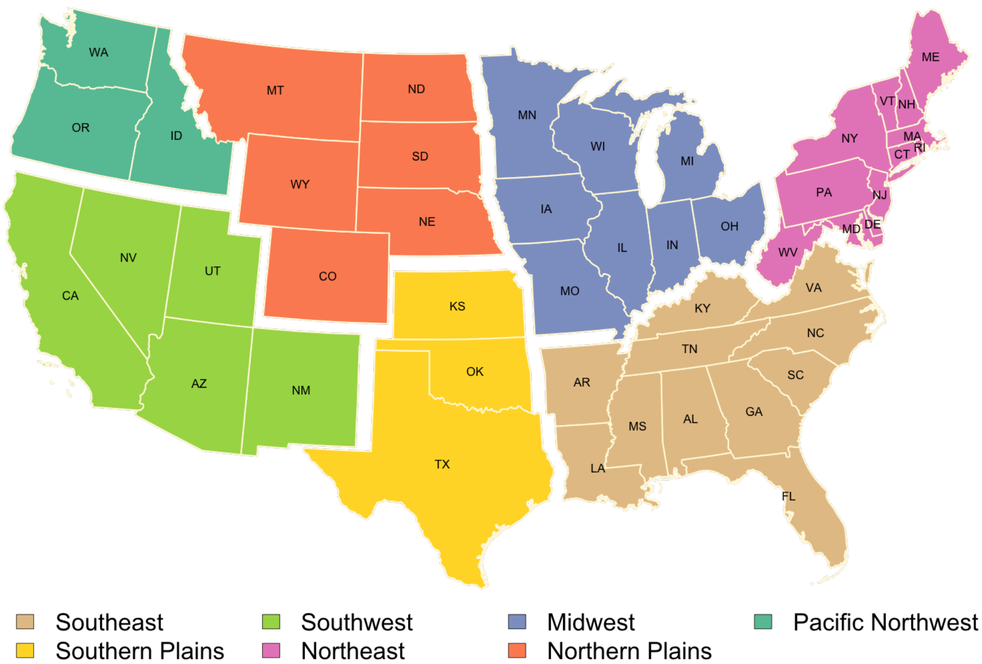
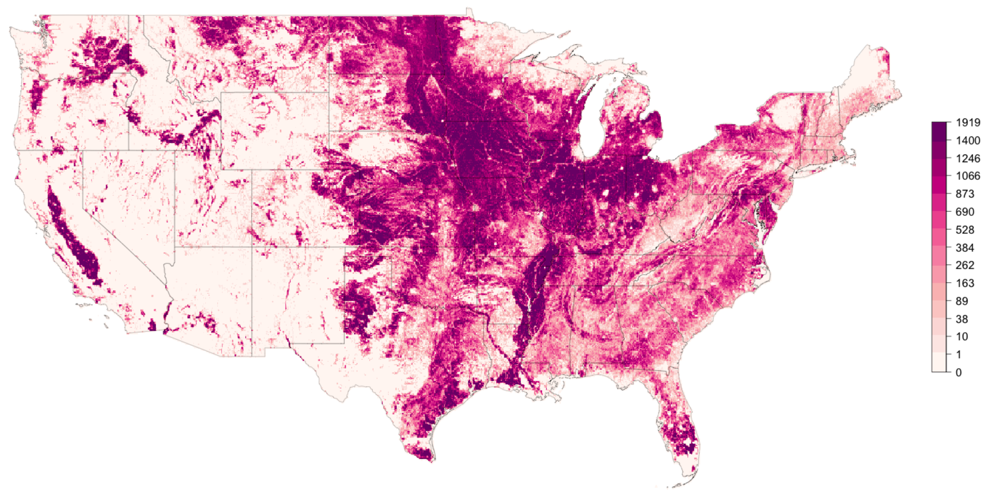


Figure A.1: Map of USDA Climate Hub regions for the lower 48 states



Notes: The color scale reflect the cropland weights associated with each PRISM grid cell.

Figure A.2: Map of weights used for aggregation of fine-scale weather data to the state level

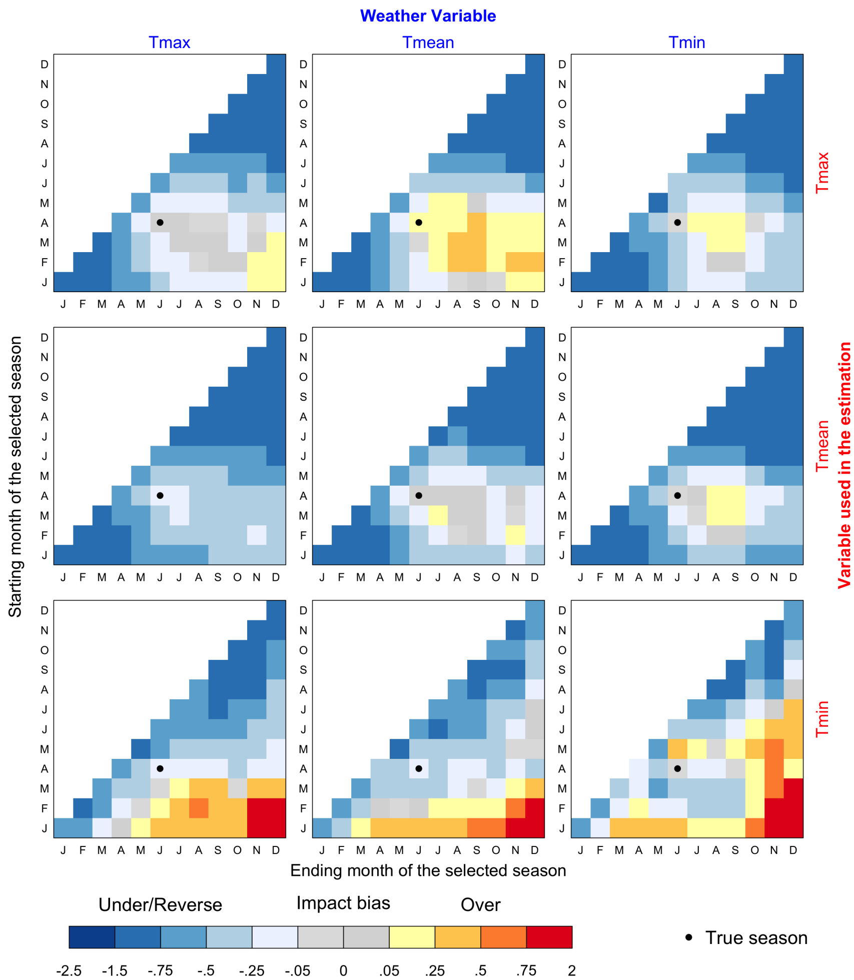


Figure A.3: Simulated results for the impact of a 1°C uniform warming (cross variables and spring season)

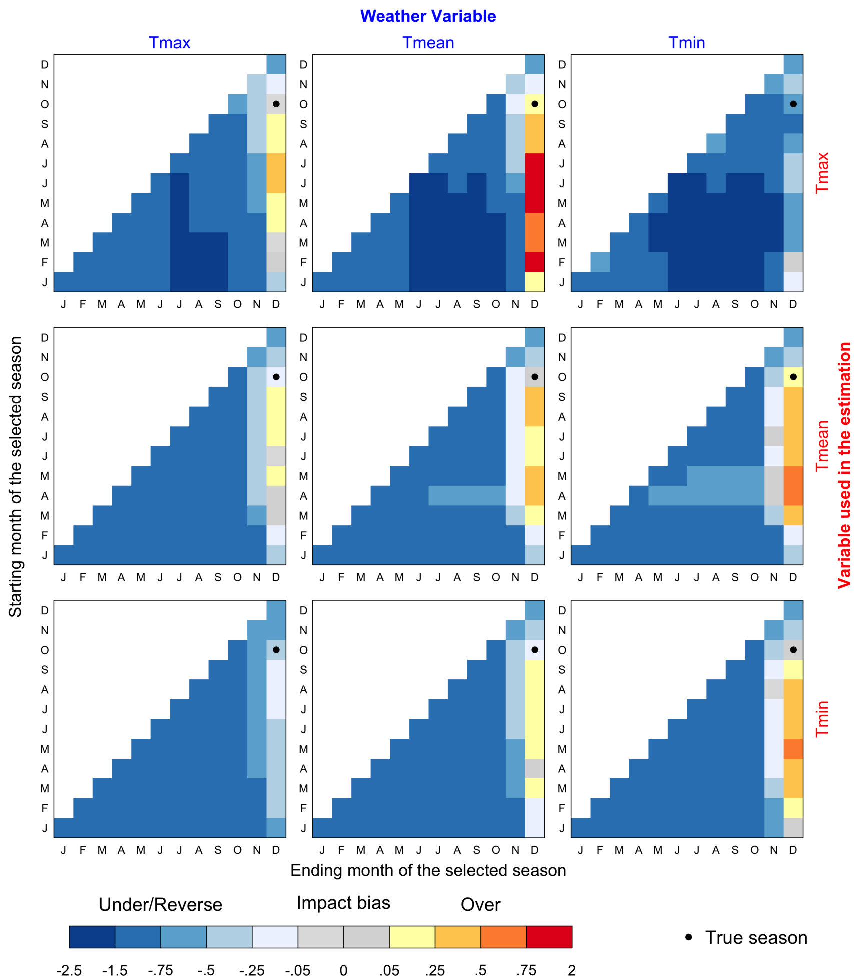


Figure A.4: Simulated results for the impact of a 1°C uniform warming (cross variables and fall season)

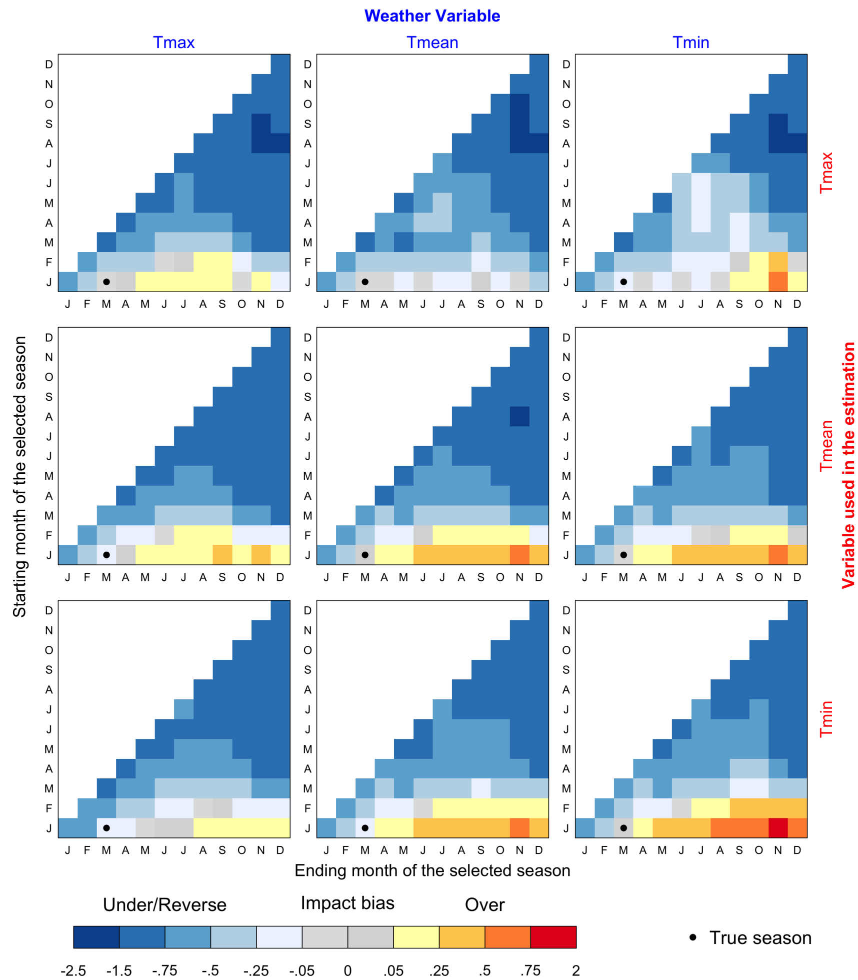


Figure A.5: Simulated results for the impact of a 1°C uniform warming (cross variables and winter season)

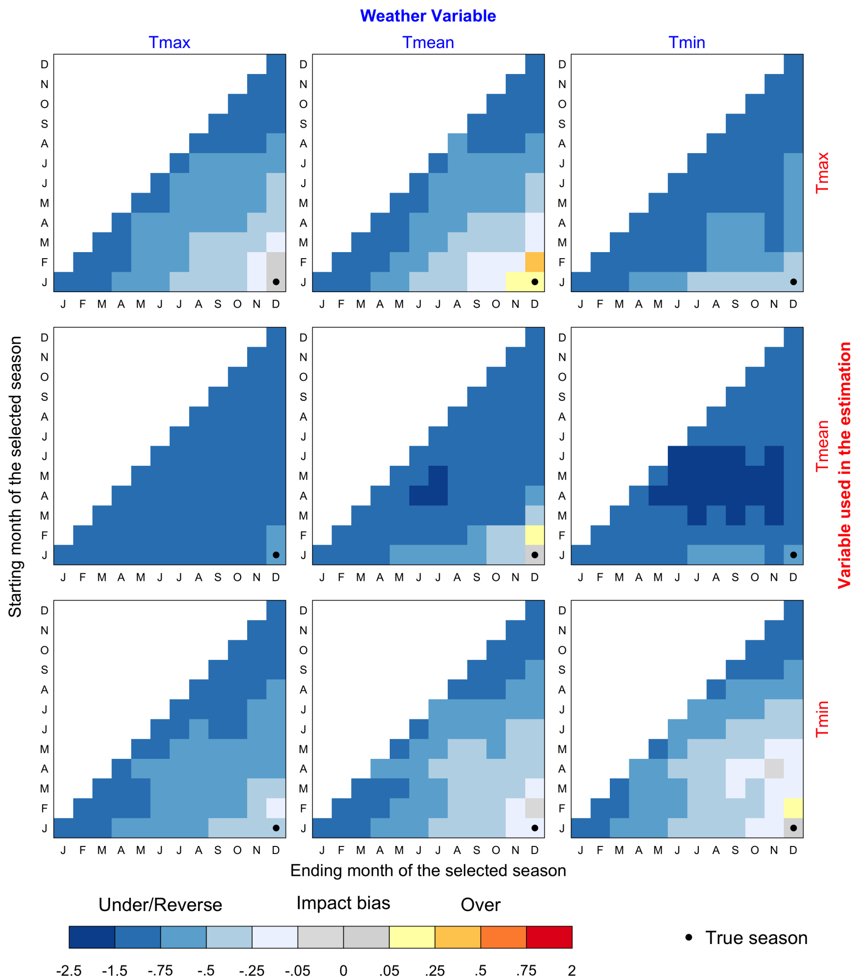


Figure A.6: Simulated results for the impact of a 1°C uniform warming (cross variables and annual)

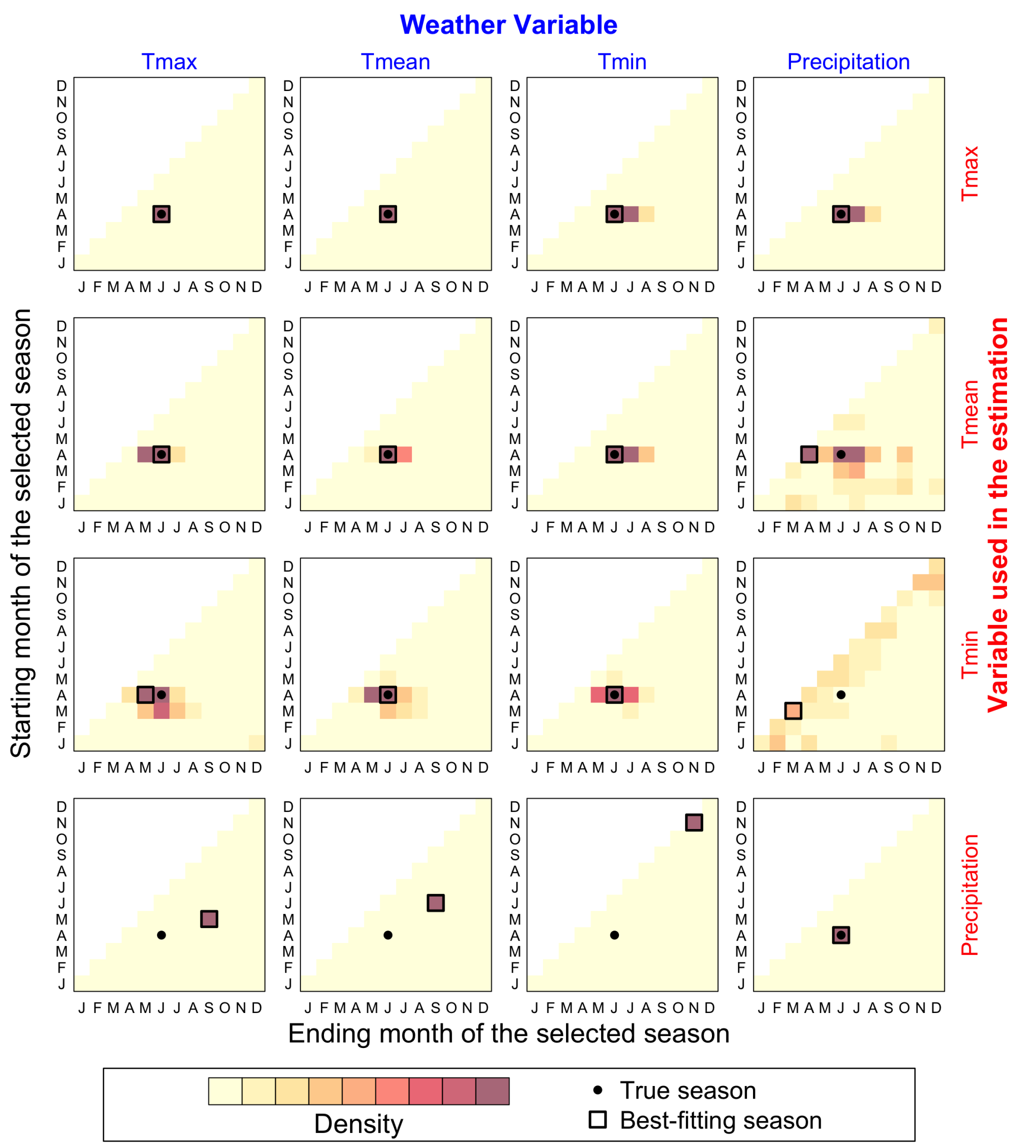


Figure A.7: Simulated results of the grid search approach (cross variables and spring season)

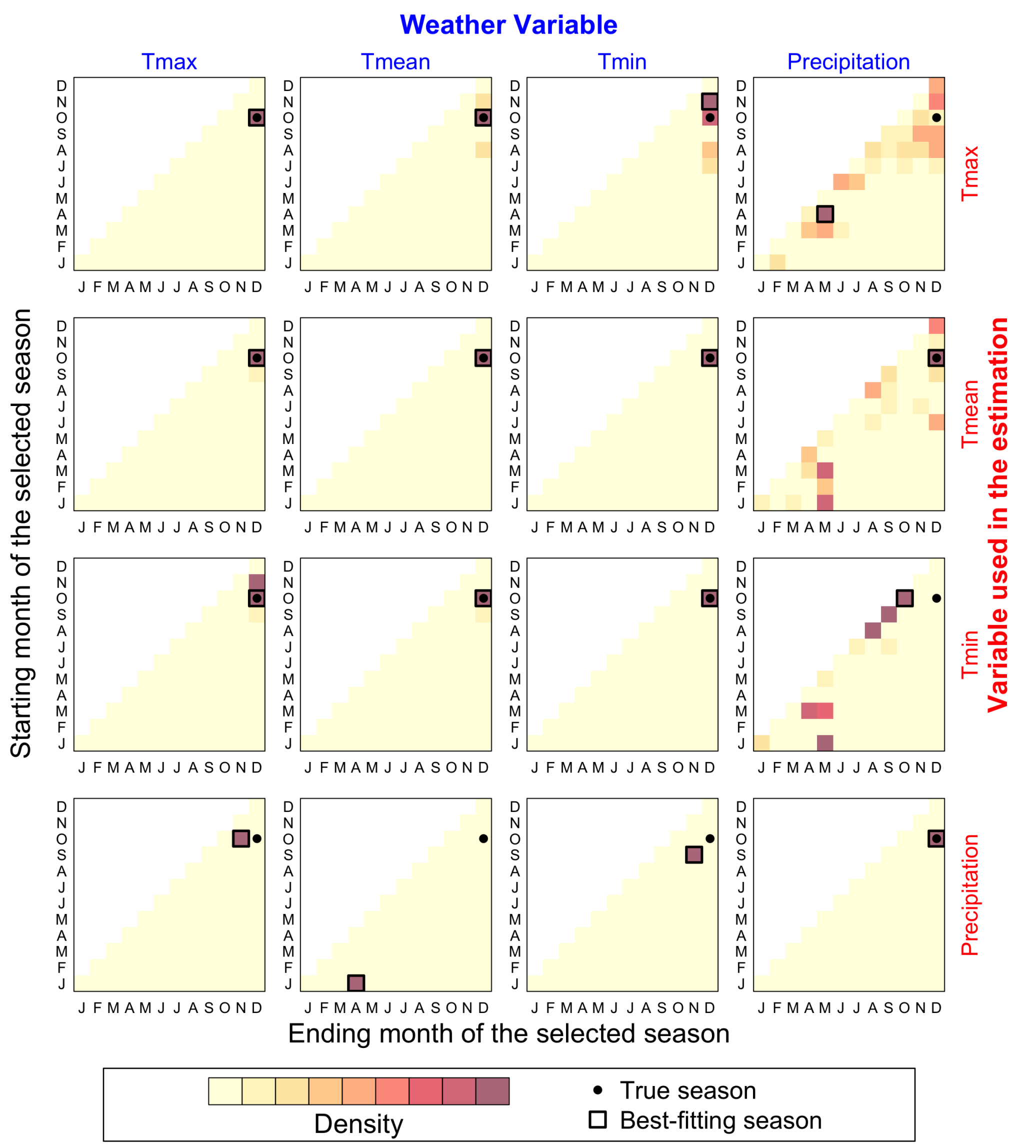


Figure A.8: Simulated results of the grid search approach (cross variables and fall season)

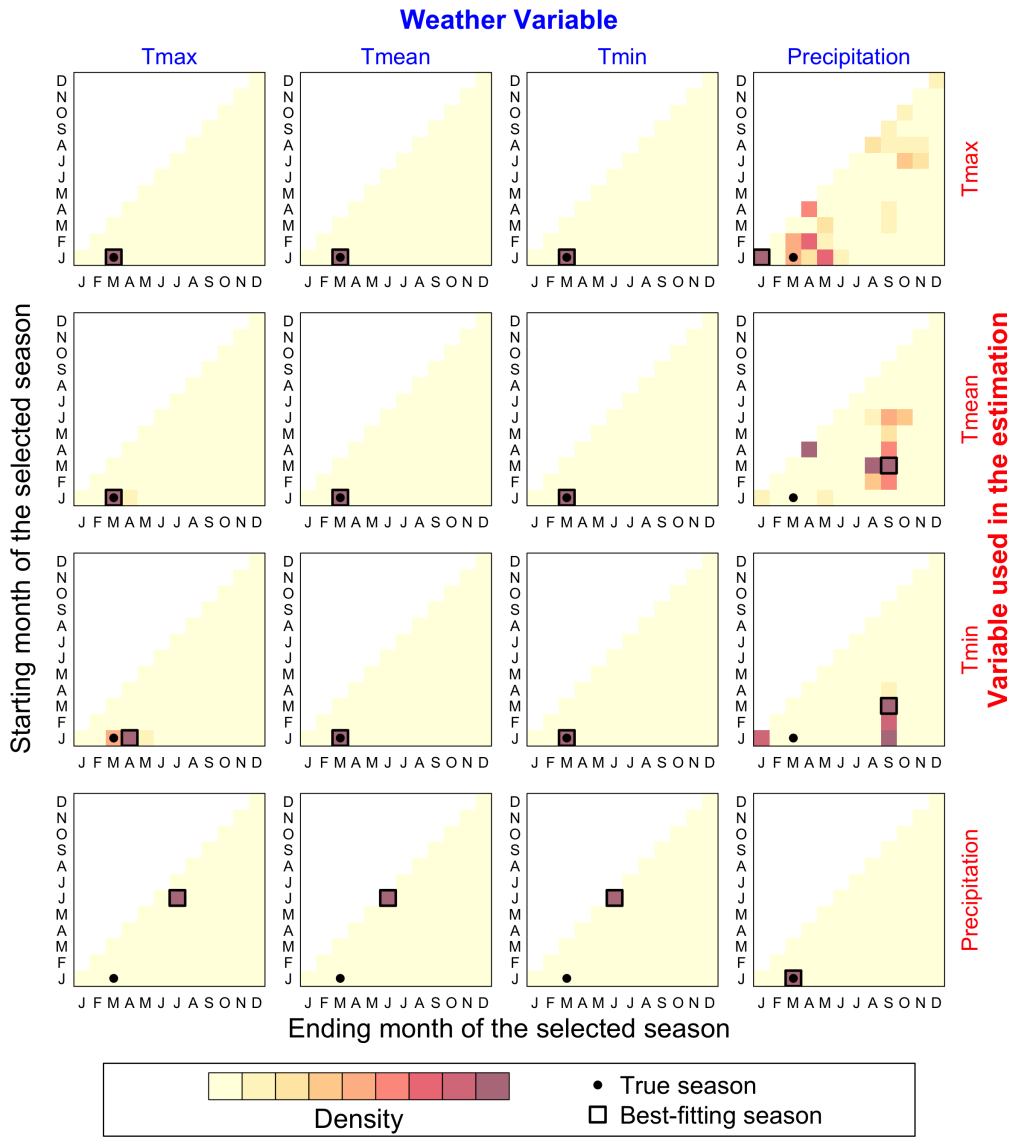


Figure A.9: Simulated results of the grid search approach (cross variables and winter season)

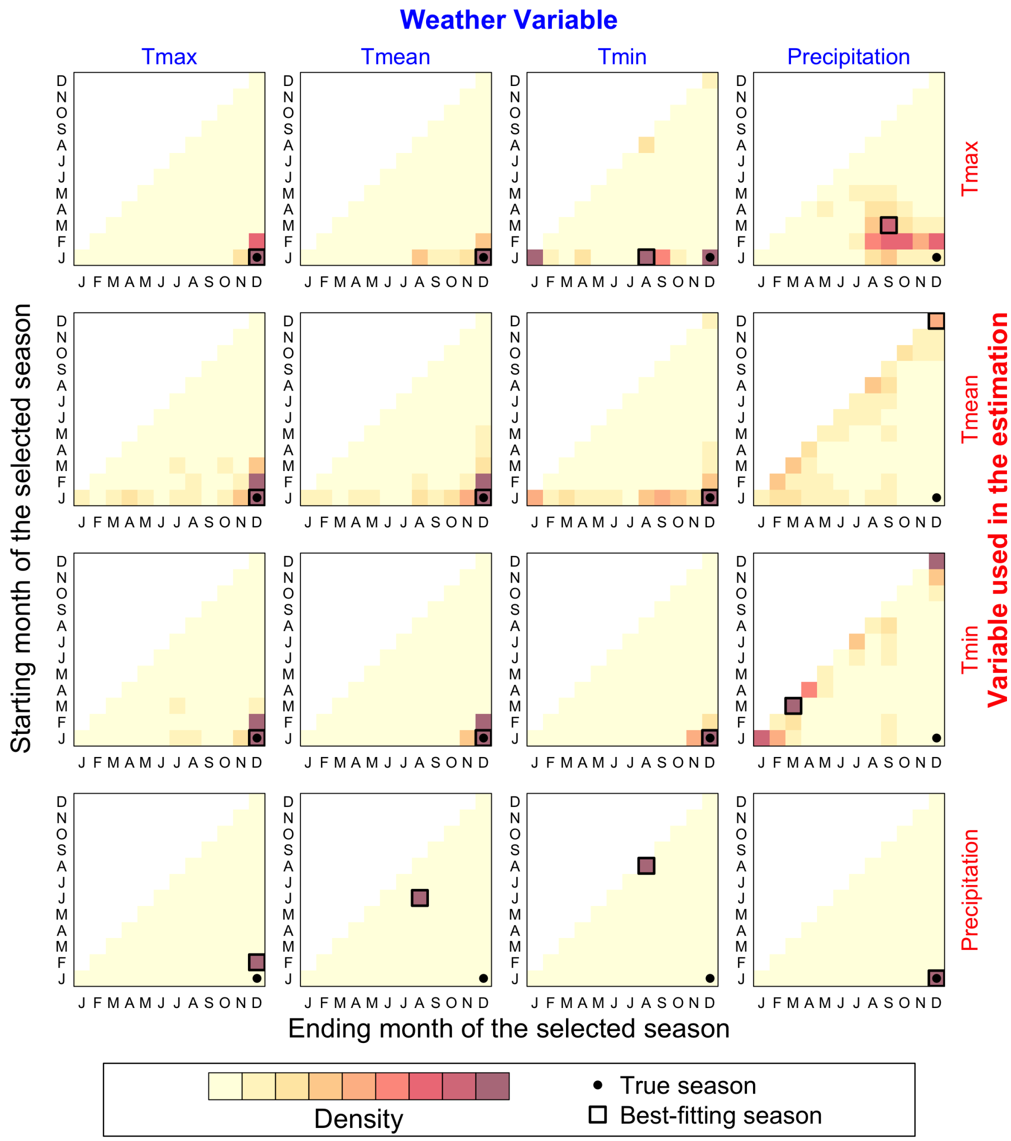


Figure A.10: Simulated results of the grid search approach (cross variables and annual)