LETTER • OPEN ACCESS

Mid-20th century warming hole boosts US maize yields

To cite this article: Trevor F Partridge *et al* 2019 *Environ. Res. Lett.* **14** 114008

View the article online for updates and enhancements.

Recent citations

- Brazilian maize yields negatively affected by climate after land clearing
  Stephanie A. Spera *et al*

- An empirical, Bayesian approach to modelling crop yield: Maize in USA
  Raphael Shirley *et al*

Environmental Research Letters

LETTER

Mid-20th century warming hole boosts US maize yields

Trevor F Partridge[1,4] ⓘ, Jonathan M Winter[1,2] ⓘ, Lin Liu[3] ⓘ, Anthony D Kendall[3] ⓘ, Bruno Basso[3] ⓘ and David W Hyndman[3] ⓘ

[1] Department of Earth Sciences, Dartmouth College, Hanover, NH, United States of America
[2] Department of Geography, Dartmouth College, Hanover, NH, United States of America
[3] Department of Earth and Environmental Sciences, Michigan State University, East Lansing, MI, United States of America
[4] Author to whom any correspondence should be addressed.

E-mail: trevor.f.partridge.gr@dartmouth.edu
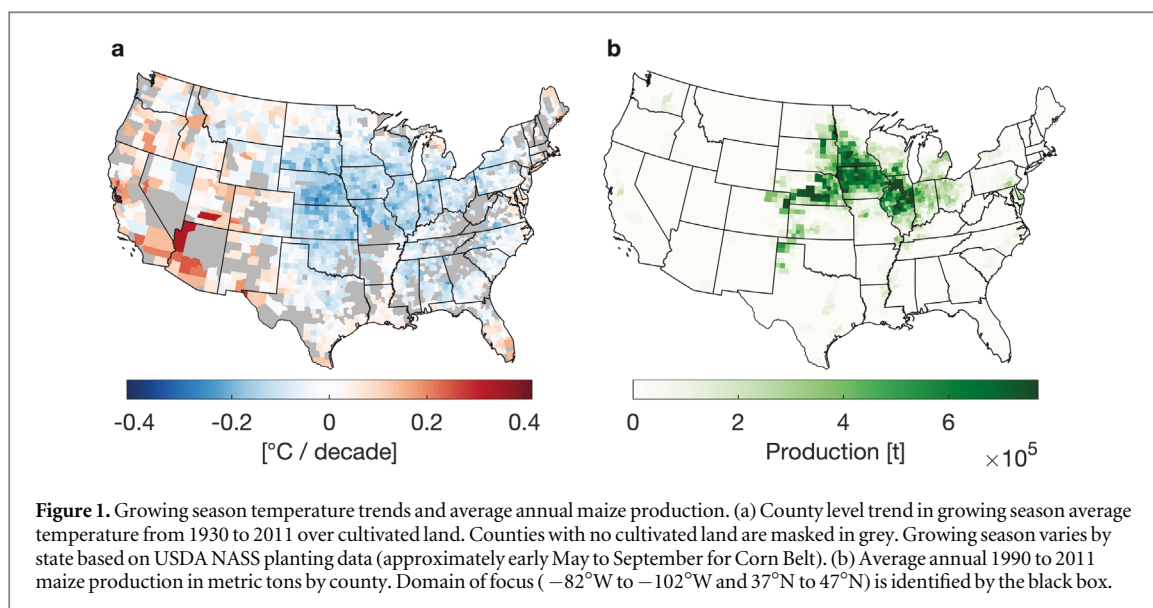
Abstract

The Corn Belt of the United States, one of the most agriculturally productive regions in the world, experienced a globally anomalous decrease in annual temperatures and a concurrent increase in precipitation during the mid- to late-20th century. Here, we quantify the impact of this 'warming hole' on maize yields by developing alternative, no warming hole, climate scenarios that are used to drive both statistical and process-based crop models. We show that the warming hole increased maize yields by 5%–10% per year, with lower temperatures responsible for 62% of the simulated yield increase and greater precipitation responsible for the rest. The observed cooling and wetting associated with the warming hole produced increased yields through two complementary mechanisms: slower crop development which resulted in prolonged time to maturity, and lower drought stress. Our results underscore the relative lack of climate change impacts on central US maize production to date, and the potential compounded challenge that a collapse of the warming hole and climate change would create for farmers across the Corn Belt.

Introduction

The remarkable increase in the United States' agricultural productivity during the 20th century (Dimitri *et al* 2005) occurred largely against the backdrop of a long-term global climate anomaly. During the second half of the twentieth century, growing season temperatures throughout most of the central US experienced an anomalously cool period associated with a phenomenon commonly referred to as the US warming hole (Hartmann *et al* 2013) as well as up to a 35% increase in summer precipitation (Alter *et al* 2018). This cooling and wetting may have created a more agriculturally favorable climate over the central US as the 20th century progressed. Given the global importance of US agriculture, the uncertainty in how future climate change may impact agricultural productivity, and the likelihood that the cooling trends associated with US 'warming hole' will disappear (Kumar *et al* 2013, Meehl *et al* 2015), understanding the historic relationship between yield and climate variability is

critical to current and future adaptation strategies. Here, we quantify the effects of the US warming hole on maize, which is one of the most important crops grown in the central US (i.e. the *Corn Belt*).

The US warming hole is characterized by an abrupt shift toward cooler annually averaged temperatures in 1958, followed by a prolonged cool period over much of the central or southeastern US, depending on season (Rogers 2013, Mascioli *et al* 2017, Partridge *et al* 2018). The exact timing and spatial extent of the warming hole varies with season, shifting from the southeastern US in winter and spring to the central US in summer and autumn. From 1980 until present, minimum temperatures throughout the US have trended upward, however there has been no significant trend in maximum temperatures within the warming hole region (USGCRP 2017). The mechanisms responsible for the observed regime shift and resultant lack of warming trend also vary by season and region. Studies focusing on the southeastern US suggest that the observed cooling trend is likely a result

Figure 1. Growing season temperature trends and average annual maize production. (a) County level trend in growing season average temperature from 1930 to 2011 over cultivated land. Counties with no cultivated land are masked in grey. Growing season varies by state based on USDA NASS planting data (approximately early May to September for Corn Belt). (b) Average annual 1990 to 2011 maize production in metric tons by county. Domain of focus ($-82°$W to $-102°$W and $37°$N to $47°$N) is identified by the black box.

of internal variability in the climate system, such as large-scale oscillations in the Pacific and Atlantic oceans; (Meehl *et al* 2012, Partridge *et al* 2018), increasing aerosol emissions during the early and mid-twentieth century (Leibensperger *et al* 2012, Mascioli *et al* 2017), and regional reforestation (Misra *et al* 2012, Ellenburg *et al* 2016). In the central US, agricultural practices have been linked to the observed summertime cooling trends. Increased irrigation (Lobell and Bonfils 2008, Pei *et al* 2016, Deines *et al* 2017) and agricultural intensification (Mueller *et al* 2015, Alter *et al* 2018) have both been shown to have remarkable influence on regional temperature and precipitation patterns. In this paper, we focus on the growing season of maize in the US Corn Belt, and thus use the term warming hole in reference to the summertime cooling pattern observed over the central US (figure 1).

Between 1930 and 2011, the yields within the average maize producing county in the US increased from 3.02 to 12.43 t ha^{-1}, with annual growth rates peaking at 3%−5% during the 1960s (Kucharik and Ramankutty 2005). Although increased fertilizer applications, improved cultivar genetics, and innovations in agricultural technology are responsible for much of the yield improvement, decreased interannual climate variability (Thompson 1986, Baker *et al* 1993) and prolonged growing seasons with reduced heat extremes since 1980 (Butler *et al* 2018) have also contributed to the impressive maize yield gains.

Here, we investigate the relationship between maize yields and climate through a multi-model approach, combining results from a process-based and statistical crop model. Process-based crop models can capture the physiological response of crops to variations in climate and management conditions because they explicitly simulate plant growth and development throughout the growing season. Process-based models have been

extensively validated globally (Basso *et al* 2016), however, they require site-specific input on soil, climate, genotypes and management which can often be limited, making them more uncertain (Lobell and Asseng 2017). Statistical crop models are computationally inexpensive, require little to no field-based calibration data, and provide reasonable estimates of yield variability (Lobell and Burke 2010, Jeong *et al* 2016, Lobell and Asseng 2017). However, they lack physiological mechanisms and thus cannot provide the mechanistic insights offered by process models. We apply two separate models to simulate long-term rainfed and irrigated maize yields: the System Approach to Land Use Sustainability (SALUS) model (Basso *et al* 2006), and an empirical machine learning algorithm calibrated to county-level annual yields. We then statistically remove the warming hole from the climate record, producing counterfactual climate scenarios to drive the crop models. Finally, we analyze the scenario outputs and quantify the net yield benefit provided by the warming hole, as well as the proportional benefits provided individually by the reduced temperatures and increased precipitation within the warming hole.

Methods

A flowchart of the methodology described below is included in the supporting information as figure S1 is available online at stacks.iop.org/ERL/14/114008/mmedia.

Data

We acquire historical 1930–2011 county level yield, production, planting date, and plant growth stage data for maize from the US Department of Agriculture National Agricultural Statistical Service (USDA NASS). Climate data is derived from daily gridded

values of maximum temperature (Tmax), minimum temperature (Tmin), daily precipitation (Pr), and incoming shortwave radiation from the Livneh hydrometeorological dataset (Livneh *et al* 2013, 2015). The Livneh hydrometeorological dataset contains meteorological observations from ~20 000 weather stations across the US and flux estimates gridded to a 1/16° scale from 1915 to 2015 for the continental US and parts of Canada and Mexico. Within Livneh, shortwave radiation is estimated using the MTCLIM algorithm in the Variable Infiltration Capacity model (Maurer *et al* 2002). We note that our analysis ends in 2011 because Livneh shortwave radiation data are only available for 1915–2011. We approximate county irrigation by the fraction of average historical 1930–2005 area equipped for irrigation (AEI, Siebert *et al* 2014) relative to a county's cultivated area (cultivated area calculation explained below).

Gridded Tmax, Tmin, and Pr values are aggregated to the county level for the random forest. To reduce the influence of non-representative grid cells on county aggregates, such as cells covering mountainous or metropolitan areas, we overlay the gridded climate data with a cultivated area mask developed using the 2017 Cultivated Layer dataset from USDA-NASS (Boryan *et al* 2012). The 2017 National Cultivated Layer uses national cropland data layers from 2013 to 2017 to estimate cultivated area at 30 m binary resolution across the United States. We interpolate the cultivated layer to a 1/16° grid, and aggregate climate data to the county level using grid cells corresponding with cultivated land. As total land area devoted to agriculture has decreased by more than 400 000 km² during the 20th century (Brown *et al* 2005; World Bank 2019) we argue that using recent cultivated area estimates are likely conservative.

Removing the warming hole from the climate record
We develop counterfactual (no warming hole) scenarios using two techniques: a delta-change method (hereafter Delta) and a quantile mapping method (hereafter QMap). Both methods are commonly used to bias correct global climate model data for use in regional assessments (Maraun and Widmann 2018, Winter *et al* 2019). The QMap approach produces a nearly identical shift in the mean value as the Delta approach, but also adjusts the variance of the post-1958 distribution to pre-1958 values (figure S2). Thus, including both methods allows for a direct comparison of how a mean shift in climate affects maize yields relative to a shift in both the mean and variance.

The Delta approach effectively adjusts observations by some change factor or delta, calculated as the difference between a target and reference distribution, additively for temperature and multiplicatively for precipitation. We remove the observed 1958 regime shift associated with the onset of the warming hole (Rogers 2013, Partridge *et al* 2018) by adjusting daily

values of Tmax, Tmin, Pr, and SW by the difference or ratio between the pre-1958 (1915–1957) and post-1958 (1959–2011) long-term monthly means. This approach preserves interannual variability of the post-1958 data but adjusts the mean to pre-1958 values (figure S2).

The QMap scenario is constructed in four steps (Brocca *et al* 2011). First, we divide Tmax, Tmin, Pr, and SW into pre- and post-1958 distributions. Second, we find the difference between the pre- and post-1958 values for each post-1958 percentile. Third, we fit a polynomial function to the post-1958 values and the calculated differences between the pre- and post-1958 values by percentile. Finally, we adjust the post-1958 values using the calculated differences from the polynomial function. For both methods, we mandate that all adjusted values lie within the bounds of historical records. Any values (less) greater than the maximum (minimum) historical value are set to the maximum (minimum) historical value and any negative values of Pr and SW are set to 0.

In addition to the Delta and QMap scenarios, we create three additional scenarios to address the sensitivity of the model outcomes to the time period included. To account for the inclusion of the Dust Bowl, an extraordinarily warm time period, in the reference distribution, we create a separate counterfactual climate scenario with the Dust Bowl years (1930–1939) removed from the reference period. Additionally, we test the sensitivity of onset of the warming hole by creating two counterfactual scenarios with breakpoints at 1953 and 1963, 5 years earlier and later than the observed warming hole onset, respectively.

Model framework
In this study, we use the well-validated SALUS model (Basso *et al* 2006) and the *Treebagger* random forest algorithm in Matlab (Breiman 2001) to simulate long-term rainfed and irrigated maize yield. Both models are calibrated to county-level reported yields from 1930–2011 (random forest) and 1956–2011 (SALUS). A detailed description of the SALUS model can be found in the supplemental text S1. SALUS simulations do not extend earlier than 1956 because this would require assumptions based on limited to no cultivar and management data. Our analysis spans the Corn Belt, from −82°W to −102°W and 37 °N to 47 °N. To ensure that this analysis includes predominantly agricultural counties, any county with less than 30 years of yield data or fewer than 10 000 harvested acres on average is excluded, resulting in 794 counties. For both models, we assume static planting dates within each state, set to the 1979–2011 average statewide median reported planting date. The length of the growing season is also static for the random forest, as described below. However, maize harvest dates in the SALUS model vary dynamically by state and year to capture the maize growth cycle at maturity.

For SALUS validation, we simulate both rainfed (counties with AEI fraction <5%; $n = 590$) and irrigated (counties with AEI fraction >60%; $n = 4$, all in Nebraska) maize yields. Automatic irrigation is triggered when plant available water content drops below 50% of maximum soil water holding capacity. After triggering, water is then added to the soil until plant available water content reaches 70% of maximum soil water holding capacity. We account for some of the technological changes in the simulated 56 years by varying management practices. The crop genetic coefficients are obtained using an inverse calibration procedure to match state-wide average reported yields to yields from maize cultivar and management practices used in the 56 year simulated period. Specifically, we vary maize cultivars from low yielding maize cultivars in 1950s to high yielding cultivars in recent decades, planting densities from 4 to 10 plants m^{-2}, and nitrogen fertilizer application rates from less than 10 to 120 kg ha^{-1}. We extract the dominant soil information for each map unit polygon in the Soil Survey Geographic Database as soil input for each studied county (USDA/NRCS 2018) and ran a unique SALUS simulation for each cell of the $1/16°$ climate grid. For each simulation year and grid cell, we calculate the AEI-weighted average yield of the simulated rainfed and irrigated yields. We then average the weighted annual yield across the grid cells in each county.

SALUS allows us to quantify the mechanistic effects of the warming hole on altered yields. Here we use both time to maturity, calculated directly by SALUS, as well as a drought stress metric that is the average of a daily drought stress factor over the 20-day period from vegetative-tasseling to early grain filling, as defined by SALUS. The drought stress factor is computed as $F_{drought} = 1 - W_{supply}/W_{demand}$, where W_{supply} is the total water available to the plant and W_{demand} is the potential evaporative demand, calculated using the Ritchie equations (Ritchie 1972). Thus, values of $F_{drought} = 1$ indicates complete drought stress, and $F_{drought} = 0$ indicates no drought stress.

Random forest is a non-parametric supervised ensemble-based machine learning algorithm consisting of a large number of individual decision trees trained on random subsets of the input data and predictor variables (Breiman 2001). Random Forest algorithms have been shown to consistently outperform multiple linear regression methods in applications including climate and agriculture (Jeong *et al* 2016). In this study, we train a random forest of 700 trees on 1930–2011 county yield data for the 794 counties using nine predictor variables; three non-climate and six climate. For each decision tree, we mandate the leaf size, or the number of observations per leaf (each split in the decision tree produces two leaves), to be more than 10 to avoid overfitting. The non-climate variables include: (1) year, (2) average county latitude, and (3) the average historical fraction of cultivated area that is equipped for irrigation (AEIFrac). The six climate variables are: (1) growing season growing degree days (GDD, see supplemental text S2), (2) growing season killing degree days (KDD; see supplemental text S2; Butler and Huybers 2015), (3) total growing season precipitation (Pr_{gs}), (4) average temperature during a critical period of the growing season ($Tavg_{Cr}$), (5) maximum Tmax during a critical period of the growing season (Txx_{Cr}), and (6) total Pr during a critical period of the growing season (Pr_{Cr}). These predictor variables are chosen from a full suite of climate metrics by highest permuted predictor importance.

For the random forest model, the growing season is defined as the period from the planting date (described above) to the last day of the week corresponding with at least 50% of a state's corn reaching maturity based on average 1979–2011 NASS survey data (USDA/NASS 2018). We identify a county-specific critical period within the growing season following a similar approach as Teasdale and Cavigelli (2017). We correlate 1930–2011 county yield timeseries with weekly aggregates of Tavg and Pr for all combinations of consecutive weeks from the first week of a county's growing season to the last week of the growing season, for each year in the dataset. A county-specific critical period is defined as the period of weeks that returns the highest absolute correlation between yield and either Tavg or Pr. We then define the critical period for the model as the average of those from the 794 counties; here weeks 12–17 post planting for $Tavg_{Cr}$ and Txx_{Cr} and 9–12 weeks post planting for Pr. A graphical representation of the correlation matrix identifying a county's critical period is shown in figure S3.

To evaluate both models' accuracy at simulating historical maize yields, we use the root mean square error and the mean absolute percentage error as well as the coefficient of determination (r^2), the *p*-value, the slope, and the intercept of a linear regression model between simulated and observed yield. For evaluating the random forest, we use out-of-bag predictions that produce unbiased metrics without needing to divide the observations into training and test sets (Breiman 1996). Correlation plots of the observed versus simulated yields are shown in figure S4.

Results

Both SALUS and the random forest simulate maize yields that closely reproduce the reported yields (figure S4). Table 1 lists the model performance evaluation metrics. The high evaluation metrics for the random forest are due to our decision to train the model on strongly-trending annual yield data while including year as a predictor variable. Training the random forest model with year as the only predictor variable illustrates the significance of this decision (figure S4(c)). However, importantly, aspects of interannual climate variability (e.g. large drought events) are still captured by this simplified model, and we argue that
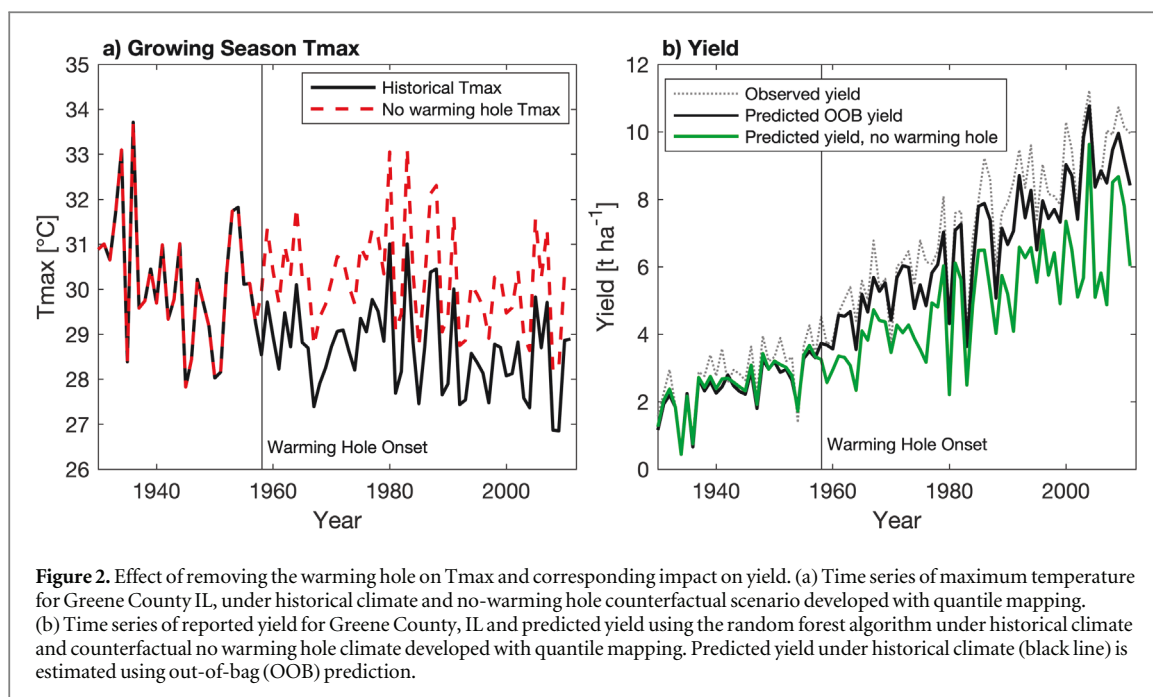
Figure 2. Effect of removing the warming hole on Tmax and corresponding impact on yield. (a) Time series of maximum temperature for Greene County IL, under historical climate and no-warming hole counterfactual scenario developed with quantile mapping. (b) Time series of reported yield for Greene County, IL and predicted yield using the random forest algorithm under historical climate and counterfactual no warming hole climate developed with quantile mapping. Predicted yield under historical climate (black line) is estimated using out-of-bag (OOB) prediction.

Table 1. Model evaluation metrics for SALUS and the random forest models. Metrics include root mean square error (RMSE) and mean annual percent error (MAPE).
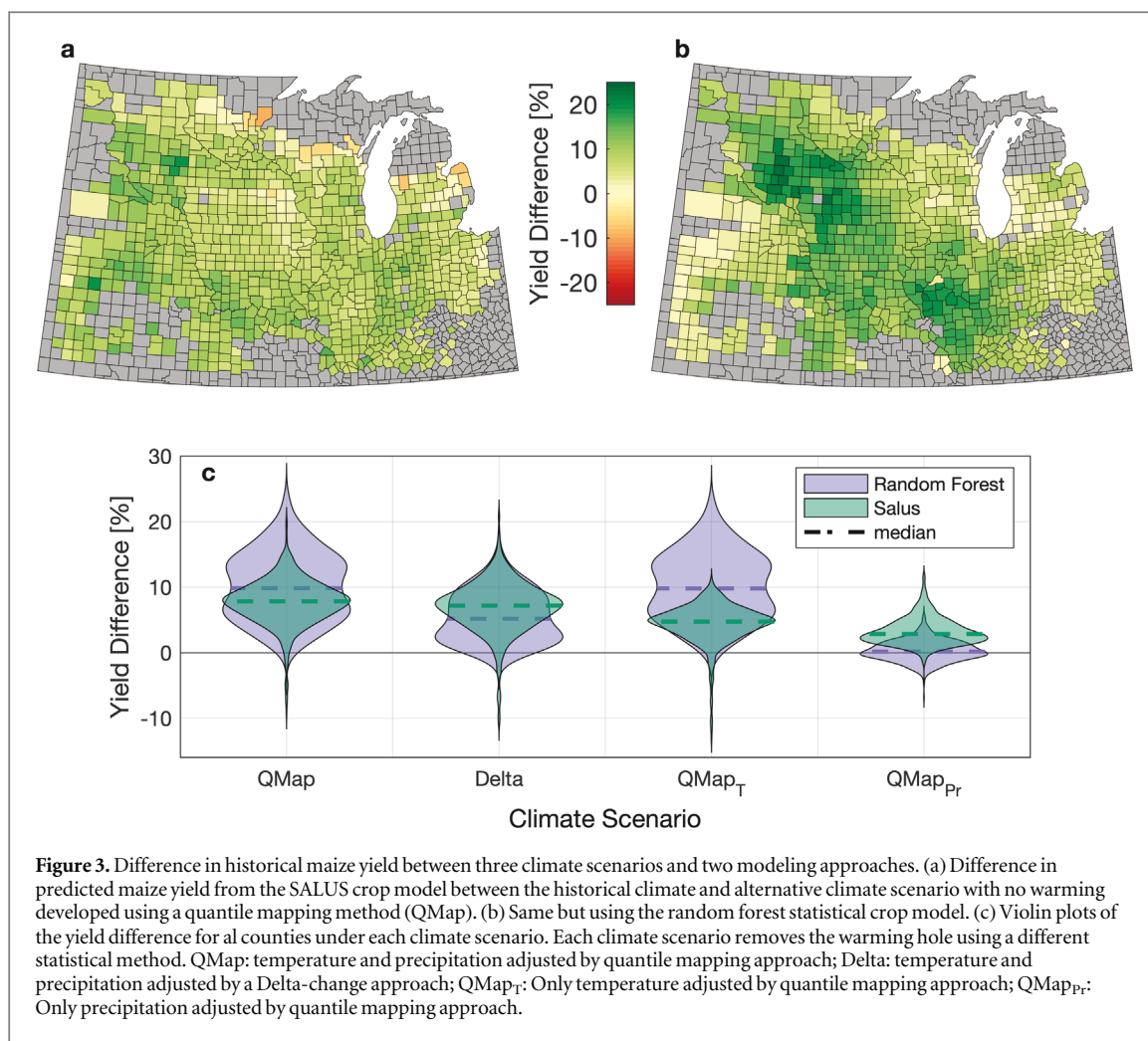
			Observed versus predicted linear regression			
	RMSE (t ha^{-1})	MAPE (%)	Intercept (t ha^{-1})	Slope	r^2	p-value
SALUS	1.44	19.3	1.22	0.81	0.67	<0.01
Random forest	0.71	15.9	−0.18	1.03	0.94	<0.01

this approach is analogous to the common approach of training a model on detrended yield data (e.g. Lobell and Field 2007) given that the relationship between year and yield within the full random forest model is nearly linear (figure S5). In addition, using yields that are not detrended allows for the direct comparison of results between the random forest and SALUS crop model.

SALUS and the random forest model results agree that maize yields benefited from the warming hole. Figure 2 illustrates the effect of removing the warming hole on Tmax and the corresponding impact on predicted maize yield for Green County, IL. Average growing season Tmax is approximately 1.5 °C higher under the no warming-hole scenario in Greene County and the corresponding predicted yield decreases by an average of 26%.

Without the warming hole, over fifty percent of counties in the Corn Belt would have experienced average annual maize yields at least 5.2%–9.8% below historical values, depending on the climate scenario and model (figure 3). We report the yield difference as the average annual percent difference between the predicted yield under the historical climate and a given counterfactual climate scenario. Thus, a positive yield difference represents a yield benefit from the observed warming hole. The largest yield benefits occur under

the QMap scenario, with median differences of 9.8% for the random forest and 7.8% for SALUS. Counties in the Dakotas and western Minnesota exhibit the largest yield differences of up to 20% for SALUS simulations and 24% according to the random forest, suggesting that this area may have benefited the most from the warming hole. The models forced with the Delta scenario indicate that fifty percent of counties would have experienced at least a 5.2% (random forest) and 7.2% (SALUS) reduction in yield without the warming hole (figure S6). The QMap and Delta scenarios have a similar shift in mean climate, so the difference in median yield response between the two scenarios is effectively the yield response due to a change in climate variability alone. Surprisingly, the random forest is more sensitive than SALUS to the increased variability within a growing season as imposed under the QMap scenario even though it uses primarily on growing season aggregate climate data. These results are qualitatively robust to the breakpoint used in creating the counterfactual scenarios and the exclusion of the Dust Bowl in the reference period. Perhaps unsurprisingly, removing the Dust Bowl from the reference period reduces the median yield benefit to 4.6% for the random forest forced by the QMap scenario, however the vast majority (>96%) of counties still exhibit increased yields relative to the

Figure 3. Difference in historical maize yield between three climate scenarios and two modeling approaches. (a) Difference in predicted maize yield from the SALUS crop model between the historical climate and alternative climate scenario with no warming developed using a quantile mapping method (QMap). (b) Same but using the random forest statistical crop model. (c) Violin plots of the yield difference for al counties under each climate scenario. Each climate scenario removes the warming hole using a different statistical method. QMap: temperature and precipitation adjusted by quantile mapping approach; Delta: temperature and precipitation adjusted by a Delta-change approach; $QMap_T$: Only temperature adjusted by quantile mapping approach; $QMap_{Pr}$: Only precipitation adjusted by quantile mapping approach.
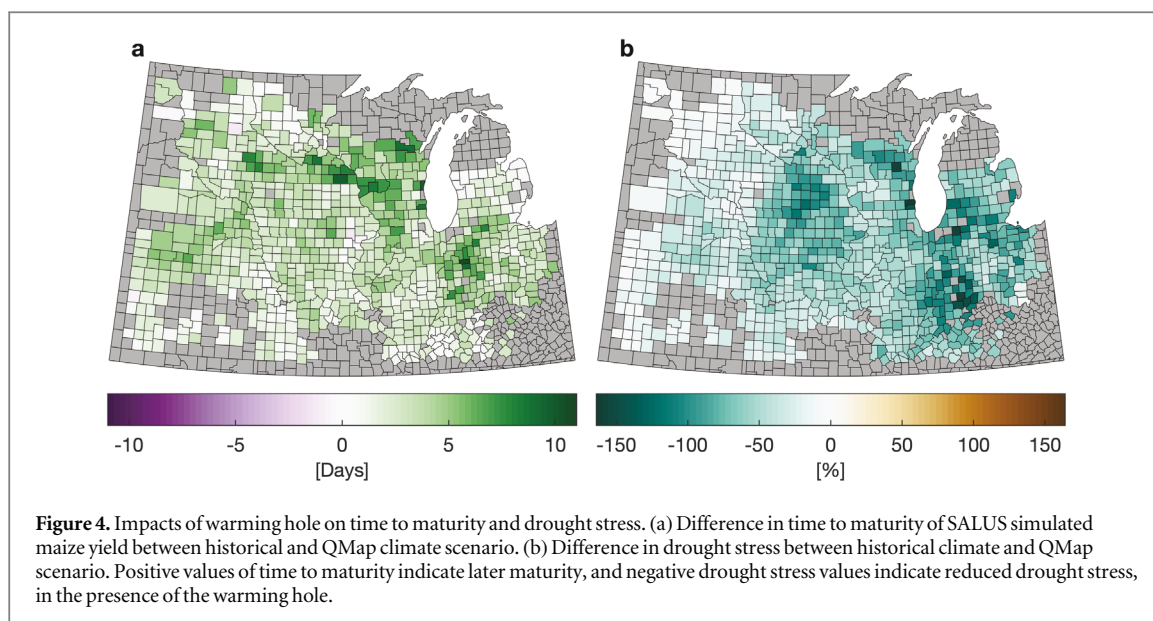
counterfactual scenario with the Dust Bowl removed (not shown).

In addition to forcing the models with the full counterfactual climate scenarios, we also explore the isolated effects of temperature and precipitation on maize yields (figure 3(c)). Crop models cannot attribute the simulated yield benefit to one particular factor, given the complex nature of the soil-plant-atmosphere interactions and feedbacks, so we quantify the relative importance of temperature and precipitation using the two additional counterfactual scenarios derived from the QMap scenario. $QMap_T$ and $QMap_{Pr}$ represent historical observations of shortwave radiation with either Tmin and Tmax or precipitation adjusted by quantile mapping, respectively. The random forest suggests that the yield difference under the QMap scenario is driven almost entirely by the temperature difference between the historical climate and the QMap climate. For the random forest, there was no statistical difference in maize yield when precipitation alone was adjusted under the QMap scenario ($p = 0.64$). These results suggest that the maize yields in the Corn Belt have benefited little from the increasing summer precipitation in the region, with the possible exception of the more arid and non-irrigated

northwest corner of the domain (figures S7 and S8). However, statistical models struggle to capture to importance of precipitation on plant growth and consequently yield (Lobell and Burke 2010, Ortiz-Bobea *et al* 2019).

By contrast, SALUS results suggest that potential yield gains as a result of the warming hole come from both lower temperatures and increased precipitation. SALUS results indicate that 62% of the median yield benefit resulted from the lower temperatures associated with the warming hole, while the remaining benefit was a result from increased precipitation (38%). We identify two potential causes for the difference in the importance of precipitation between the random forest and SALUS. First, the random forest was trained using yield observations from the whole corn belt, and therefore has a limited ability to simulate geographical variations in the sensitivity of maize to precipitation. Second SALUS is able to simulate the impacts of daily precipitation on crop growth, while the random forest aggregates precipitation.

To explore potential mechanisms responsible for the yield reductions under the no-warming hole climate scenarios, we examine differences in simulated drought stress and maturity time using the SALUS model. Under

Figure 4. Impacts of warming hole on time to maturity and drought stress. (a) Difference in time to maturity of SALUS simulated maize yield between historical and QMap climate scenario. (b) Difference in drought stress between historical climate and QMap scenario. Positive values of time to maturity indicate later maturity, and negative drought stress values indicate reduced drought stress, in the presence of the warming hole.

both no warming hole scenarios, drought stress increases, and crops mature earlier, leading to lower yields. Maize in the average county reaches maturity 2.7 d or 2.9 d earlier than the historical observed climate and experiences an 52% or 49% increase in drought stress for the QMap or the Delta scenarios, respectively (figure 4). The largest difference in crop maturity time occurs in the northern counties of Wisconsin where cooler late spring and early autumn temperatures otherwise slow plant development. Interestingly, the northern counties of Minnesota and the Dakotas do not exhibit the same reduction in time to maturity, likely due to later planting dates in this region. There is a longitudinal gradient in the reduction of drought stress between the historical and QMap scenarios. Counties in the eastern half of the domain appear to have substantially lower drought stress levels as a result of the warming hole and concurrent precipitation trends.
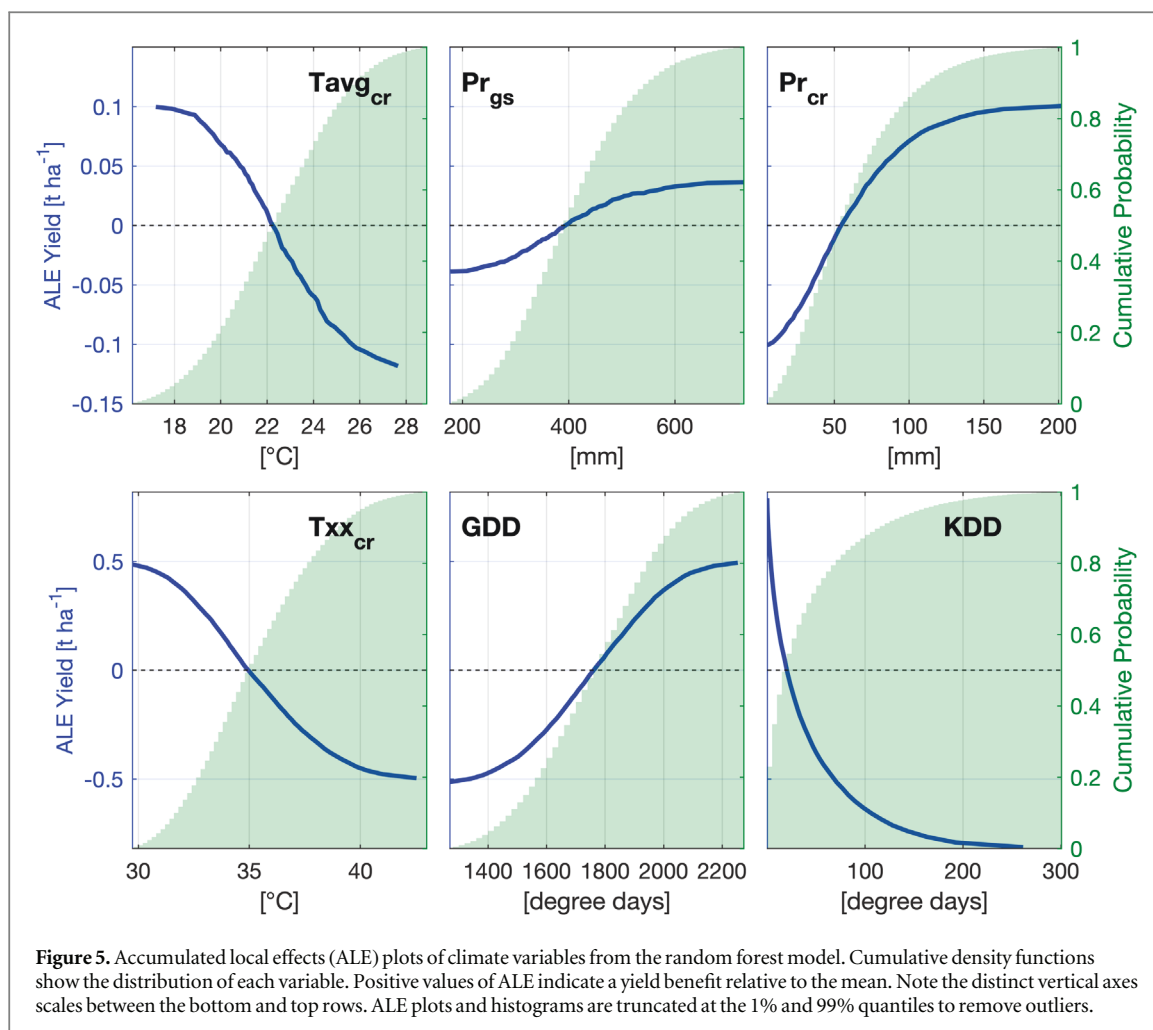
Random forest algorithms are often referred to as 'black boxes', as it can be difficult to interpret the relationships formed within the ensemble of decision trees. Accumulated local effect (ALE) plots provide a way to visualize the relationship between the response variable (yield) and each predictor variable in isolation (Apley 2016). Figure 5 shows ALE plots for climate variables used in the random forest as well as the total distribution of that variable within the domain. The plots are centered at zero, which represents the mean model prediction across all values of the variable. Values below zero have a net negative effect on yield (decreased simulated yield) and values above zero are beneficial (increased simulated yield). ALE plots for the non-climatic variables are shown in figure S5. GDD, the hottest daily maximum temperature during the critical period (Txx_{cr}), 12–17 weeks after planting, and KDD (Butler and Huybers 2015) have a similar effect on yield (see methods for description of random forest and variable calculation). The ALE plot of average temperature during the critical period ($Tavg_{cr}$)

suggests that days with average temperature above 22 °C can be detrimental to maize yields. However, this threshold should be interpreted with caution as Tavg is a statistical construct of daily Tmax and Tmin, which are more likely to physically affect plant growth. The ALE plots of Pr give insight into why forcing the random forest with $QMap_{pr}$ had little effect on simulated yield (figure 3(c)). Precipitation summed over the growing season (Pr_{gs}) and during the critical period (Pr_{cr}; weeks 9–12 post planting) exhibit relatively flat curves, and the corresponding precipitation distributions illustrate the general abundance of water in the region. The more arid northwestern portion of the domain did benefit from the historical increase in summer precipitation, so a similar analysis focused only on this region might find a stronger relationship between Pr and maize yields (figure S7).

Conclusion

Understanding the dynamic relationship between agricultural productivity and regional climate is of critical importance given the challenges imposed by population growth, shifting diets, and climate change. Here we find that the 20th century trend in US maize yield was significantly aided by anomalously cool temperatures and increased growing season precipitation. These results expand upon those of Butler *et al* (2018), who use a statistical crop model to show that increased minimum temperatures and a reduction in maximum temperature extremes since 1980 have been beneficial to maize yields. We find that this agriculturally pleasant climate extends back to 1958 and also quantify the importance of a simultaneous increase in growing season precipitation. We further identify both prolonged maturity time and reduced water stress as the physiological mechanisms responsible for the yield benefit (figure 4). It is unlikely that the pleasant climate

Figure 5. Accumulated local effects (ALE) plots of climate variables from the random forest model. Cumulative density functions show the distribution of each variable. Positive values of ALE indicate a yield benefit relative to the mean. Note the distinct vertical axes scales between the bottom and top rows. ALE plots and histograms are truncated at the 1% and 99% quantiles to remove outliers.

observed during the latter half of the 20th century will substantially persist into the future. Although maximum growing season temperatures in the Corn Belt continue to increase at a slower rate than minimum temperatures, possibly due to agricultural expansion (Mueller *et al* 2015, Alter *et al* 2018, Nikiel and Eltahir 2019), annual temperatures are projected to increase by 2.3 °C by mid-century (RCP 4.5; Vose *et al* 2017). Summer precipitation in the region is projected to decrease by approximately 10% while extreme events will likely increase (Easterling *et al* 2017).

The dramatic improvements in technology and management practices across US Corn Belt produced remarkable increases in maize yield and productivity. However, we show that a contemporaneous climate anomaly referred to as the US 'warming hole', which translated to more advantageous weather in the region, also contributed to yield increases. We find that the warming hole resulted in an increased median annual maize yield of 5%–10%, primarily in response to decreased growing season temperatures and decreased heat extremes. SALUS driven with counterfactual scenarios isolating the increase in precipitation and decrease in temperature indicate that cooler temperatures were responsible for 62% (and thus increased precipitation was repsonsible for 38%) of

the simulated yield increase across most of the Corn Belt. As the central US prepares to adapt to substantial projected increases in temperatures by the end of the 21st century, it is essential to understand the extent that the US warming hole has historically benefitted maize yields. Our results underscore the relative lack of climate change impacts on central US maize production to date, and the potential compounded challenge that a collapse of the warming hole and climate change would create for farmers across the Corn Belt.

Acknowledgments

Author contributions

TFP, JMW, ADK and DWH designed the study and developed the random forest model. TFP and JMW developed the climate scenarios. LL and BB ran the SALUS model simulations and analyzed drivers of yield changes. All authors contributed to writing the manuscript.

Competing interests statement

We have no conflicts of interest to declare and none of the material has been published or is under consideration elsewhere.

ORCID iDs

Trevor F Partridge ⓘ https://orcid.org/0000-0003-1589-4783
Jonathan M Winter ⓘ https://orcid.org/0000-0003-1261-4774
Lin Liu ⓘ https://orcid.org/0000-0001-9078-9516
Anthony D Kendall ⓘ https://orcid.org/0000-0003-3914-9964
Bruno Basso ⓘ https://orcid.org/0000-0003-2090-4616
David W Hyndman ⓘ https://orcid.org/0000-0001-9078-9516

References

Alter R E, Douglas H C, Winter J M and Eltahir E A B 2018 Twentieth century regional climate change during the summer in the central united states attributed to agricultural intensification *Geophys. Res. Lett.* **45** 1586–94

Apley D W 2016 Visualizing the effects of predictor variables in black box supervised learning models arXiv:1612.08468

Baker D G, Ruschy D L and Skaggs R H 1993 Agriculture and the recent 'Benign Climate' in Minnesota *Bull. Am. Meteorol. Soc.* **74** 1035–40

Basso B, Dumont B, Cammarano D, Pezzuolo A, Marinello F and Sartori L 2016 Environmental and economic benefits of variable rate nitrogen fertilization in a nitrate vulnerable zone *Sci. Total Environ.* **545–546** 227–35

Basso B, Ritchie J T, Grace P R and Sartori L 2006 Simulation of tillage systems impact on soil biophysical properties using the SALUS model *Ital. J. Agronomy* **1** 677

Boryan C, Yang Z and Di L 2012 Deriving 2011 cultivated land cover data sets using usda national agricultural statistics service historic cropland data layers *2012 IEEE Int. Geoscience and Remote Sensing Symp.* (Piscataway, NJ: IEEE) pp 6297–300

Breiman L 1996 Out-Of-Bag Estimation *Technical Report* (Berkeley, CA: Department of Statistics, University of California, Berkeley) (https://www.stat.berkeley.edu/~breiman/OOBestimation.pdf)

Breiman L 2001 Random forests *Mach. Learn.* **45** 5–32

Brocca L, Hasenauer S, Lacava T, Melone F, Moramarco T, Wagner W and Bittelli M 2011 Soil moisture estimation through ASCAT and AMSR-E sensors: an intercomparison and validation study across Europe *Remote Sens. Environ.* **115** 3390–408

Brown D G, Johnson K M, Loveland T R and Theobald D M 2005 Rural land-use trends in the conterminous United Sates, 1950–2000 *Ecol. Appl.* **15** 1851–63

Butler E E and Huybers P 2015 Variations in the sensitivity of US maize yield to extreme temperatures by region and growth phase *Environ. Res. Lett.* **10** 034009

Butler E E, Mueller N D and Huybers P 2018 Peculiarly pleasant weather for US maize *PNAS* **115** 11935–40

Deines J, Kendall A and Hyndman D W 2017 Annual irrigation dynamics in the US Northern High Plains derived from Landsat satellite data *Geo. Res. Lett.* **44** 9350–60

Dimitri C, Effland A and Conklin N 2005 *The 20th Century Transformation of US Agriculture and Farm Policy* (*Economic Information Bulletin* No. 3) (Washington, DC: USDA Economic Research Service) (www.ers.usda.gov)

Easterling D R, Arnold J R, Knutson T, Kunkel K E, LeGrande A N, Leung L R and Wehner M F 2017 *Precipitation Change in the United States. Climate Science Special Report: Fourth National Climate Assessment* ed D J Wuebbles *et al* vol 1 (Washington, DC: US Global Change Research Program) pp 207–30

Ellenburg W L, McNider R T, Cruise J F and Christy J R 2016 Towards an understanding of the twentieth-century cooling trend in the Southeastern United States: biogeophysical impacts of land-use change *Earth Interact.* **20** 1–31

Hartmann D J, Klein Tank A M G, Rusticucci M, Alexander L V, Brönnimann S, Charabi Y A R and Zhai P 2013 Observations: Atmosphere and Surface *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* pp 159–254

Jeong J H, Resop J P, Mueller N D, Fleisher D H, Yun K, Butler E E and Kim S-H 2016 Random forests for global and regional crop yield predictions *PLoS One* **11** e0156571

Kucharik C J and Ramankutty N 2005 Trends and variability in US corn yields over the twentieth century *Earth Interact.* **9** 1–29

Kumar S, Merwade V, Kinter J L and Niyogi D 2013 Evaluation of temperature and precipitation trends and long-term persistence in CMIP5 twentieth-century climate simulations *J. Clim.* **26** 4168–85

Leibensperger E M, Mickley L J, Jacob D J, Chen W T, Seinfeld J H, Nenes A and Rind D 2012 Climatic effects of 1950–2050 changes in US anthropogenic aerosols: II. Climate response *Atmos. Chem. Phys.* **12** 3349–62

Livneh B, Bohn T J, Pierce D W, Munoz-Arriola F, Nijssen B, Vose R and Brekke L 2015 A spatially comprehensive, hydrometeorological data set for Mexico, the US, and Southern Canada 1950–2013 *Sci. Data* **2** 150042

Livneh B, Rosenberg E A, Lin C, Nijssen B, Mishra V, Andreadis K M and Lettenmaier D P 2013 A long-term hydrologically based dataset of land surface fluxes and states for the conterminous United States: update and extensions *J. Clim.* **26** 9384–92

Lobell D B and Asseng S 2017 Comparing estimates of climate change impacts from process-based and statistical crop models *Environ. Res. Lett.* **12** 015001

Lobell D B and Bonfils C 2008 The effect of irrigation on regional temperatures: a spatial and temporal analysis of trends in California, 1934–2002 *J. Clim.* **21** 2063–71

Lobell D B and Burke M B 2010 On the use of statistical models to predict crop yield responses to climate change *Agric. For. Meteorol.* **150** 1443–52

Lobell D B and Field C B 2007 Global scale climate–crop yield relationships and the impacts of recent warming *Environ. Res. Lett.* **2** 014002

Maraun D and Widmann M 2018 *Statistical Downscaling and Bias Correction for Climate Research* (Cambridge: Cambridge University Press)

Mascioli N R, Previdi M, Fiore A M and Ting M 2017 Timing and seasonality of the United States 'warming hole' *Environ. Res. Lett.* **12** 034008

Maurer E P, Wood A W, Adam J C, Lettenmaier D P and Nijssen B 2002 A long-term hydrologically based dataset of land surface fluxes and states for the conterminous United States *J. Clim.* **15** 3237–51

Meehl G A, Arblaster J M and Branstator G 2012 Mechanisms contributing to the warming hole and the consequent US east-west differential of heat extremes *J. Clim.* **25** 6394–408

Meehl G A, Arblaster J M and Chung C T Y 2015 Disappearance of the southeast US 'warming hole' with the late 1990s transition of the Interdecadal Pacific Oscillation *Geophys. Res. Lett.* **42** 5074–81

Misra V, Michael J P, Boyles R, Chassignet E P, Griffin M and O'brien J J 2012 Reconciling the spatial distribution of the surface temperature trends in the Southeastern United States *J. Clim.* **25** 3610–8

Mueller N D, Butler E E, McKinnon K A, Rhines A, Tingley M, Holbrook N M and Huybers P 2015 Cooling of US Midwest summer temperature extremes from cropland intensification *Nat. Clim. Change* **6** 317–22

Nikiel C A and Eltahir E A B 2019 Summer climate change in the midwest and great plains due to agricultural development during the twentieth century *J. Clim.* **32** 5583–99

Ortiz-Bobea A, Wang H, Carrillo C M and Ault T R 2019 Unpacking the climatic drivers of US agricultural yields *Environ. Res. Lett.* **14** 064003

Partridge T F, Winter J M, Osterberg E C, Hyndman D W, Kendall A D and Magilligan F J 2018 Spatially distinct seasonal patterns and forcings of the US Warming Hole *Geophys. Res. Lett.* **45** 2055–63

Pei L, Moore N, Zhong S, Kendall A D, Gao Z and Hyndman D W 2016 Effects of Irrigation on summer precipitation over the United States *J. Clim.* **29** 3541–58

Ritchie J T 1972 Model for predicting evaporation from a row crop with incomplete cover *Water Resour. Res.* **8** 1204–13

Rogers J C 2013 The 20th century cooling trend over the southeastern United States *Clim. Dyn.* **40** 341–52

Siebert S, Kummu M, Porkka M, Döll P, Ramankutty N and Scanlon B R 2014 A global dataset of the extent of irrigated land from 1900 to 2005 *Hydrol. Earth Syst. Sci. Discuss.* **11** 13207–58

Teasdale J R and Cavigelli M A 2017 Meteorological fluctuations define long-term crop yield patterns in conventional and organic production systems *Sci. Rep.* **7** 688

Thompson L M 1986 Climatic change, weather variability, and corn production1 *Agron. J.* **78** 649

USDA NASS 2018 USDA—National Agricultural Statistics Service —Quick Stats [Data file] (https://nass.usda.gov/Quick_Stats/) (Retrieved: 26 March 2019)

USDA NRCS 2017 Soil Survey Geographic Database (SSURGO). NRCS [Data file]. (Retrieved: 26 March 2018)

USGCRP 2017 *Climate Science Special Report: Fourth National Climate Assessment* ed D J Wuebbles *et al* vol 1 (Washington, DC: US Global Change Reseacrh Program) (https://doi.org/10.7930/J0J964J6)

Vose R S, Easterling D R, Kunkel K E, LeGrande A N and Wehner M F 2017 Temperature changes in the United States ed D J Wuebbles *et al Climate Science Special Report: Fourth National Climate Assessment* vol 1 (Washington, DC: US Global Change Research Program) pp 185–206

Winter J M, Bowen F L, Partridge T F and Chipman J W 2019 Future extreme event risk in the rural Northeastern United States *Ann. Am. Assoc. Geogr.* **109** 1110–30

World Bank 2019 Agricultural Land (sq. km). [Data file] (https://data.worldbank.org/indicator/AG.LND.AGRI.K2?locations=US) (Retrieved: 22 March 2019)