Uncovering the past and future climate drivers of wheat yield shocks in Europe with machine learning

Peng Zhu, Rose Z Abramoff, David Makowski, Philippe Ciais

Earth's Future

RESEARCH ARTICLE
10.1029/2020EF001815

Key Points:

- Water limitation is the dominant climate driver for yield shock during 1980–2018
- Future warming causes higher yield shock probability and a paradigm shift from water limitation to extreme warming dominated yield shock
- Crop models can capture the negative impact of water limitation but largely missed excess water induced yield shock

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

P. Zhu,
zhuyp678@gmail.com

Citation:

Zhu, P., Abramoff, R., Makowski, D., & Ciais, P. (2021). Uncovering the past and future climate drivers of wheat yield shocks in Europe with machine learning. *Earth's Future*, *9*, e2020EF001815. https://doi.org/10.1029/2020EF001815

Received 24 SEP 2020
Accepted 3 APR 2021

Uncovering the Past and Future Climate Drivers of Wheat Yield Shocks in Europe With Machine Learning

Peng Zhu[1], Rose Abramoff[1], David Makowski[2,3], and Philippe Ciais[1]

[1]Laboratoire des Sciences du Climat et de l'Environnement, Institut Pierre-Simon Laplace (IPSL), Gif sur Yvette, France, [2]Université Paris-Saclay, AgroParisTech, INRAE, Unit Applied Mathematics and Computer Science (MIA 518), Paris, France, [3]CIRED—Centre International de Recherche sur l'Environnement et le Développement, Nogent-Sur-Marne, France

Abstract Recently, yield shocks due to extreme weather events and their consequences for food security have become a major concern. Although long yield time series are available in Europe, few studies have been conducted to analyze them in order to investigate the impact of adverse climate events on yield shocks under current and future climate conditions. Here we designated the lowest 10th percentile of the relative yield anomaly as yield shock and analyzed subnational wheat yield shocks across Europe during the last four decades. We applied a data-driven attribution framework to quantify primary climate drivers of wheat yield shock probability based on machine learning and game theory, and used this framework to infer the most critical climate variables that will contribute to yield shocks in the future, under two climate change scenarios. During the period 1980–2018, our attribution analysis showed that 32% of the observed wheat yield shocks were primarily driven by water limitation, making it the leading climate driver. Projection to future climate scenarios RCP4.5 and RCP8.5 suggested an increased risk of yield shock and a paradigm shift from water limitation dominated yield shock to extreme warming induced shocks over 2070–2099: 46% and 54% of areas were primarily driven by extreme warming under RCP4.5 and RCP8.5, respectively. A similar analysis conducted on yields simulated by an ensemble of crop models showed that models can capture the negative impact of low water supply but missed the impact of excess water. These discrepancies between observed and simulated yield data call for improvement in crop models.

Plain Language Summary Crop production is inherently sensitive to climate hazards including droughts, heat waves, and extreme precipitation. Many studies have investigated the impact of extreme events on crop yields, however, there is little conclusive evidence of the most frequent and significant causes of yield shocks. Here we applied a data-driven attribution framework to quantify the primary climate drivers of wheat yield shock probability in Europe. The results show that the overall dominant climate driver of yield shock during 1980–2018 is water limitation, however, it was surpassed by extreme warming in the late period (2000–2018). The future projection revealed an increasing yield shock probability and a paradigm shift from water limitation to extreme warming dominated yield shock across all regions. Crop model evaluation suggests that crop models can capture the negative impact of low water supply but missed the one of excess water. These discrepancies call for improvement in crop models. Our data-driven attribution framework can be readily applied to larger areas and other crop types to explore the leading driver of yield shock, therefore, the attribution analyses we presented here have broad implications and would be beneficial for targeted climate adaptation policy design in agricultural system.

1. Introduction

Food security is a global concern in a changing climate. With population growth and diet shift to high calories, it is predicted that food demand will be roughly doubled by 2050 (Hertel et al., 2010; Tilman et al., 2011). In particular, the demand for wheat is expected to increase over the next few years. According to FAO (Alexandratos & Bruinsma, 2012), wheat consumption will increase in many countries. Additionally, global warming is causing an increase in the frequency and severity of extreme events (Diffenbaugh et al., 2017; Field et al., 2012) resulting in yield reductions that have a cascade of socio-economic impacts as well as lower mean decadal yields (Asseng et al., 2015; Cottrell et al., 2019).

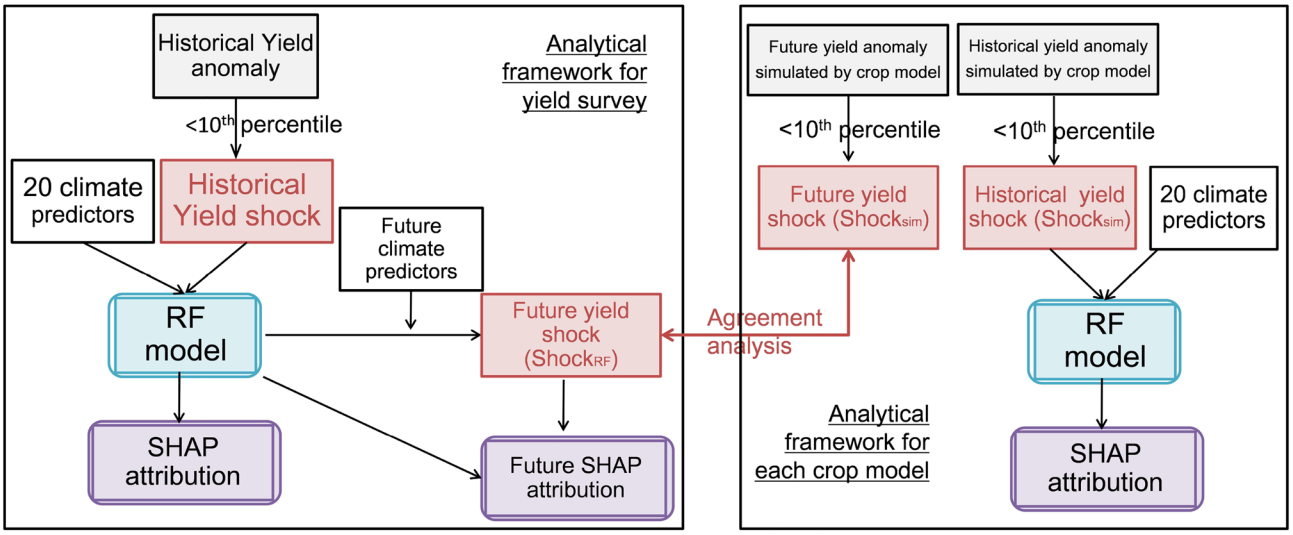

Figure 1. Schematic diagram of the workflow for yield shock attribution analysis in historical and future climate conditions.

This study focused on yield shock, which is often accompanied by food price spikes and can have dramatic influence on food security and social conflicts (Chen & Villoria, 2019; Crost et al., 2018). Yield shock is normally interrelated with single or compound extreme weather events, pest and disease outbreaks or geopolitical events (Beillouin et al., 2020; Ben-Ari et al., 2018; Ciais et al., 2005; Cottrell et al., 2017; Webber et al., 2020). Yet, the influential climate variables are diverse and not fully identified, and interactions between them matter as well. For example, heat stress accelerates crop leaf senescence (Lobell et al., 2012); excessive precipitation causes lodging or increased risk of pests and diseases (Zampieri et al., 2017). Several studies investigated the impact of specific climate variation or extreme climate events on crop yields (Beillouin et al., 2020; Jin et al., 2017; Lesk et al., 2016; Rötter et al., 2018) but no study has been conducted to attribute yield shocks to different types of weather events in Europe over several decades. This information is essential to identify the most influential climatic factors and the most sensitive periods of crop development. Considering that compound extreme events are more frequent in warmer climate (AghaKouchak et al., 2020; Zscheischler et al., 2018) a better understanding of the leading drivers of yield shock will help design a targeted climate adaptation policy for agricultural production systems.

Winter wheat is widely planted in Europe and highly sensitive to climatic warming (Asseng et al., 2015; Porter & Semenov, 2005). Compared to spring crops, winter wheat has a longer growing season and thus higher probability of being exposed to extreme events, making it more difficult to be simulated by crop models (Guarin et al., 2020). For example, until approaching the time of harvest, all forecasting systems failed to predict the 2016 yield shock that occurred in the breadbasket of France. In contrast, a statistical model was able to attribute this extreme yield shock to a compound extreme of abnormal autumn warming and spring wet conditions (Ben-Ari et al., 2018). Therefore, a systematic analysis of the climate drivers underlying different types of shocks beyond well-studied drought events and an evaluation of the performances of current crop models to capture the drivers of those shocks is required to improve model credibility and help agricultural decision making (Rötter et al., 2011; Wang et al., 2017).

Here we built a data-driven framework with long term sub-national European winter wheat yield survey during 1980–2018 to attribute the yield shocks to different climatic drivers. We first used the random forest (RF) machine-learning algorithm to associate climate stresses at different wheat growth stages with wheat yield shocks. We then used an attribution analysis based on game theory (Shapley additive explanations) to assess the relative contribution of extreme weather to each yield shock event (Figure 1). Specifically, in the following content, we try to address the following questions: (1) How have the wheat yield shocks evolved over the historic period? (2) What were the primary climate drivers for the identified yield shocks? (3) How will the yield shock occurrence probability change under future climate scenarios? (4) How well are the current crop models able to predict wheat yield shocks under current and future climate?

2. Methods and Data Sets

2.1. Wheat and Climate Data

Winter wheat yield and cropping areas data in Europe over 1980–2018 were obtained based on survey data collected in a previous study (Beillouin et al., 2020) which covered 17 European countries at subnational scale (i.e., at nomenclature of Territorial Units for Statistics from EUROSTAT 2 and 3) and 1,435 administrative units across the 17 countries. To fill the missing yield data in some units, we employed a gridded yield data set at a spatial resolution of 0.5° (Iizumi and Sakai 2020). This global-scale gridded yield data was created through disaggregating national level yield data to finer spatial resolution based on grid cell crop harvest area and satellite estimated crop productivity. Correlation analysis suggests a good consistency between the two yield data sets (Figure S1) in most units. Given the spatial disparity in climate and potentially locally adapted farmer management practices in wheat production system, the whole Europe yield data set was divided into four regions: North, South, West, and East Europe, as done by Beillouin et al. (2020) (details in Table S1).

Yield shocks were identified by fitting long-term yield trends through locally weighted scatterplot smoothing ("loess," Cleveland, 1979) and then estimating relative yield anomalies $\bar{\alpha}_{i,t}$ as:

$$\bar{\alpha}_{i,t} = \frac{Y_{i,t} - \bar{Y}_{i,t}}{\bar{Y}_{i,t}}, \tag{1}$$

where $Y_{i,t}$ is wheat yield and $\bar{Y}_{i,t}$ is the expected yield value in the i administrative unit at year t. The expected yield value, $\bar{Y}_{i,t}$ corresponds to the long-term yield trend estimated using loess. Finally, yield shocks were defined as the lowest 10th percentile of relative yield anomalies in each region (Figure S2). For those administrative units having missing values in the yield survey, relative yield anomalies based on yield survey and gridded yield by Iizumi and Sakai (2020) were estimated separately. Then yield shock was identified based on the two combined yield anomalies. The yield shock events identified in Iizumi and Sakai data constitute 18% total yield shock events. We also calculated the area fraction of yield shock for each year across all administrative units as:

$$\text{Fraction}_t = \frac{\sum_{i=1}^{M} I_{t,i} * \text{Area}_i}{\sum_{i=1}^{M} \text{Area}_i}, \tag{2}$$

where Fraction_t is the area fraction of yield shock for year t; $I_{t,i} = 1$, if unit i in year t is identified as yield shock, otherwise $I_{t,i} = 0$; M is the number of administrative units (M = 1,435), $Area_i$ is the harvested area of administrative unit i.

Gridded climate data during 1979–2018 were obtained from the Joint Research Center's MARS meteorological database (Toreti, 2014). This database was selected to be consistent with the climate forcing data used in the following crop model simulations. This climate data included daily minimum ($T_{\min}$) and maximum ($T_{\max}$) surface air temperature, precipitation, 10-m wind speed, global radiation, vapor pressure deficit and potential evapotranspiration from a crop canopy at 25 km resolution covering Europe and neighboring countries.

The winter wheat growing season in Europe generally starts with sowing in autumn followed by a dormancy period in winter and plant development in spring, and ends with maturity in the following late spring or early summer. Considering the different response of the plant to climate stresses in different growing stages (Duncan et al., 2015; Siebers et al., 2017; Zhu et al., 2019), we used an observational crop phenology dataset to classify climate stressors according to their occurrence during each development stage. The crop phenology data were taken from Eurostat (http://ec.europa.eu/eurostat/web/main), including winter wheat sowing date, anthesis and harvest/maturity date. These observed crop phenology dates were aggregated for each administrative unit and then averaged over time. Four growing stages were defined: sowing date to November 30 as autumn; December 1 to February 28 as winter; March 1st to anthesis as vegetative period; and anthesis to maturity as reproductive period. For each of the four stages, five climate variables were defined (Table 1) (Ben-Ari et al., 2016, 2018) in order to cover four distinct categories of stress: warm

Table 1
Climate Variables Defined for Each Growth Stage to Capture Different Climatic Stress Effects on Wheat Yield

f_{wd}	Fraction of warm days (days with T_{max} >30°C or its 90th percentile in each growth stage)
f_{cd}	Fraction of cold days (days with T_{min} < its 10th percentile in each growth stage)
Prec	Mean daily precipitation in each stage (mm/day)
f_{rd}	Fraction of rainy day (rainy day is a day with precipitation > 1 mm)
PET	Mean daily potential evapotranspiration amount in each stage (mm/day)

stress, cold stress, water supply, and water demand. Specifically, we used two variables (mean precipitation and fraction of rainy days) to characterize the water availability accounting for precipitation distribution, as a recent study suggested that uneven precipitation distribution will offset the yield benefit of precipitation (Fishman, 2016). The 20 climate variables (five types of variable for each of the four growth stages considered) were used as the predictors of yield shock probability.

2.2. Crop Models and Future Crop Model Projections

Crop models have been used to study the climate change impacts on crop productivity, given their potential to simulate the large-scale spatiotemporal variations of crop yield (Mueller et al., 2017; Webber et al., 2018). Here we used simulated yield data produced by eight-crop models (4M, FASSET, HERMES, MONICA, SIMPLACE-Lintul5, SIRIUS 2015, SiriusQuality v3, and SSM-Wheat) which simulate wheat yield at 25 km resolution across Europe in the historical period 1980–2010 and future climate scenarios (RCP4.5 and RCP8.5). From the multiple model experiments, we selected the model simulation with responsiveness to mean temperatures, drought, and heat stress with simulated canopy temperature. Although other global-scale crop model simulation experiments were available (Mueller et al., 2017), the eight-crop models selected here were designed specifically for Europe and assimilated observed crop phenology data from Eurostat (http://ec.europa.eu/eurostat/web/main) to better match crop growth seasonality. Additionally, the simulated yield showed a good performance in reproducing the interannual variation of wheat yield at national level.

For the historic period simulation (1980–2010), all eight crop models were forced by the climate data obtained from the Joint Research Center's meteorological database (Toreti, 2014). For the climate projections under the two future scenarios (RCP4.5 and RCP8.5), 6 out of 8 crop models reported the outputs forced with elevated CO_2 and future climate models (GFDL-CM3, GISS-E2-R, HadGEM2-ES, MIROC5, and MPI-ESM-MR) during the period 2070–2099. The climate projections were downscaled to the resolution of the baseline data by assigning values to each 25 km grid. The climate projections were created using the enhanced delta change method by applying changes in simulated temperature and precipitation variability to historic mean climate (Ruane et al., 2015). This climate data can be accessed at: http://open-research-data-zalf.ext.zalf.de/ResearchData/DK_59.html. The same climate variables in Table 1 were derived for each climate model under RCP4.5 and RCP8.5, and were also used as predictors to predict future yield shock probability. In future climate scenarios, PET was calculated using Penman-Monteith equation (Allen et al., 1998).

To identify yield shock in the crop model simulation during both the historic period (1980–2010) and future scenarios (RCP4.5 and RCP8.5), simulated yield data were aggregated to the administrative unit and grouped into the same four regions as those considered in the yield survey data. Similarly, the relative yield anomaly was estimated with Equation 1 and then yield shocks ($Shock_{sim}$) were defined as the lowest 10th percentile of relative yield anomalies in each of the four regions.

2.3. Data-Driven Framework of Yield Shock Occurrence Modeling and Attribution

Yield shock events identified in the historical yield survey data were treated as a binary variable. We used a RF machine-learning algorithm to predict the shock occurrence probability using the 20 climate variables defined above (Table 1) as predictors. RF algorithm was selected as it was able to take into account non-linear responses and complex interactions between multiple predictor variables. The output of RF was expressed as a predicted probability of shock occurrence computed by averaging across an ensemble of

Table 2
Criteria Used to Attribute a Primary Driver to Each Yield Shock

Primary drivers	Criteria
Extreme warming	f_{wd} associated SHAP was the largest and f_{wd} exceeding its 80% percentile
High water demand	PET associated SHAP was the largest and PET exceeding its 80% percentile
Excessive water supply	Prec or f_{rd} associated SHAP was the largest and Prec or f_{rd} above its 80% percentile
Low water supply	Prec or f_{rd} associated SHAP was the largest and Prec or f_{rd} below its 20% percentile
Cold stress	f_{cd} associated SHAP was the largest and f_{cd} exceeding its 80% percentile

Note. Five different drivers are distinguished depending on the SHAP values associated to the climate variables f_{wd}, PET, P_{rec}, f_{rd}, and f_{cd}.

decision trees built with a training data set (Breiman, 2001; Cutler et al., 2012). One separate RF model was built for each of the four regions to account for the potential disparity in the response of yield to climate variation. The two hyperparameters—"mtry" and "ntree"—in the RF model, representing the numbers of predictors sampled for splitting at each node and number of trees grown in the RF model, were tuned using the available set of data. The "ntree" was set equal to 1,000, since we found that out-of-sample model accuracy plateaued after this threshold. The "mtry" was tuned through 5-fold cross-validation and set equal to 11, 14, 12, and 11 for North, West, East, and South Europe, respectively. In order to assess the RF model performance in discriminating yield shock and non-yield shock, we computed the area under receiver operating characteristic curve (AUC) value (Fawcett, 2006) for each of the four regions by five-fold cross validation (AUC equal to 1 means a perfect classification, while a value of 0.5 suggests a performance equal to a random classification). For each RF model, we used partial dependence plots (PDP) to analyze the relationship between the yield shock occurrence probability and each predictor variable, averaging over all other predictors (Friedman, 2001). Similarly, eight RF models were built to fit the yield shock derived from the eight sets of crop model simulations with the same climate predictors during 1980–2010, respectively. RF models were implemented using the "ranger" R package (Wright et al., 2017).

Although PDP provides a graphical description of how yield shock probability responds to climate stressors in their multi-dimensional space, this approach does not quantify how much each climate predictor has contributed to the predicted yield shock probability. An attribution analysis based on game theory (Shapley additive explanations, SHAP) allowed us to evaluate the relative contribution of each climate variable to the predicted yield shock probability and then to identify the most adverse types of stress. SHAP is a method of estimating the expected marginal contribution of a covariate across all possible combinations and has been used in recent study to interpret machine learning model (Padarian et al., 2020). For each observed yield shock event, we used SHAP to break down the RF model predicted yield shock probability into individual predictor contribution. We then selected the predictor with the largest SHAP value as the primary driver for a yield shock event. Based on the category of climate variables in Table 1 and the reported prevalent causes of wheat yield shock events (Webber et al., 2020; Zheng et al., 2012; Zampieri et al., 2017), we focused on the following five drivers: extreme warming, high water demand, excessive water supply, low water supply, and cold stress. The selection criteria for each primary driver is detailed in Table 2. SHAP value based attribution analysis was implemented using the "iml" R package (Molnar et al., 2018).

Based on the criteria listed in Table 2, all yield shocks associated with the same primary climate driver were grouped together, and the five resulting groups were used to calculate the area fraction of yield shocks with the same primary driver for the four regions and the whole of Europe. We also divided the area fraction of yield shocks with the same primary driver in 1980–2018 into two periods (1980–1999 and 2000–2018) to investigate how a certain climate stress driven yield shocks have changed over time.

To predict the probability of yield shock in the future, we applied the RF model trained over historic yield survey data to the two climate scenarios considered (RCP4.5 and RCP8.5). Then, shock probability above its top 90th percentile in each region was identified as yield shock. We then applied the SHAP method to decompose the yield shock probability into contributions of individual climate predictors and finally determined the primary driver for each shock in future climate scenarios. Future yield shocks predicted by RF models ($Shock_{RF}$) were also used as pseudo true values to evaluate the performance of crop models in

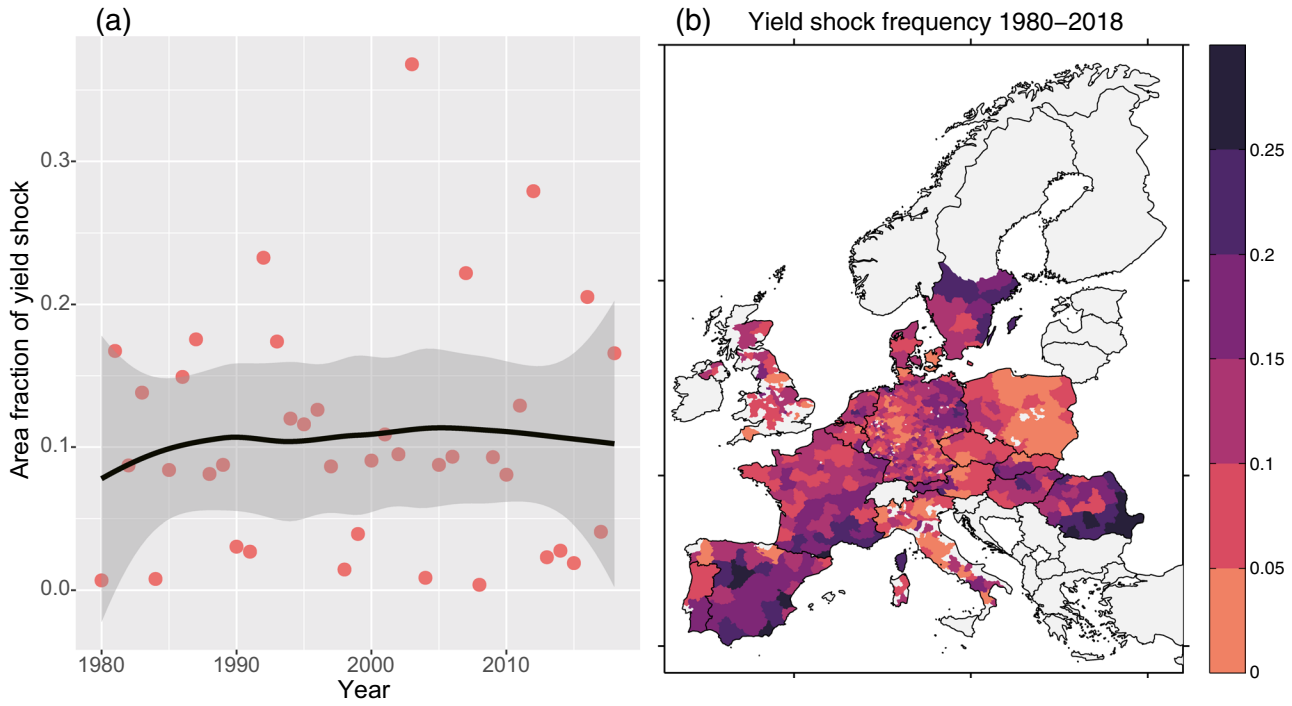

Figure 2. Yield shock area fraction for the studied European countries during 1980–2018 (a). Yield shock area fraction was estimated with Equation 2. The smooth curve is based on loess with shaded area as 95% confidence interval. The spatial pattern of yield shock frequency for each administrative unit (number of yield shock occurrence/total years).

forecasting yield shock. A yield shock occurrence predicted by the RF model during both historic and future periods might be identified as yield shock by 0, 1, 2 ,..., 6 crop models ($Shock_{sim}$). The level of agreement between $Shock_{RF}$ and $Shock_{sim}$ was assessed as a function of the type of climate variables considered as the main driver by RF. For example, for *N* yield shocks ($Shock_{RF}$) primarily driven by extreme warming, we could get *N*0, *N1* ,..., *N*5, *N*6 (*N*0 + *N*1 + ... *N*5 + *N*6 = *N*, with N the total number of $Shock_{RF}$ for this driver) yield shocks also identified as shock ($Shock_{sim}$) by 0, 1, 2 ,..., 6 crop models, respectively. A high value of N4 + N5 + N6 + N7 suggested a strong agreement between RF and crop models.

3. Results

3.1. Spatial and Temporal Distributions of Yield Shocks in Europe

Time series of wheat cropping area influenced by yield shock during 1980–2018 suggested that the year showing the highest area impacted by yield shock corresponded to year 2003 (Figure 2a). According to several previous studies, this year was characterized by a compound heat and drought stress (Ciais et al., 2005; van der Velde et al., 2010). When looking into the four regions, North Europe and East Europe show increasing trend in the area fractions of yield shock (Figure S3). Spatially, the yield shock frequency (yield shock occurrence numbers/total years) is higher in southern areas characterized by warmer conditions, especially in Romania, Spain, and Southern France (Figure 2b).

3.2. Response of Yield Shock to Climate Stresses

AUC values estimated with 5-fold cross validation for the four regions were all greater than 0.85, suggesting that the trained RF model had a high capacity to discriminate between yield shock and non-yield shock, with North Europe scoring the highest AUC (0.95, see Table S2). With the trained RF model, we also derived the PDP to show the marginal effect of each predictor on yield shock probability. Overall, the same predictor might have quite different PDPs across the four regions, suggesting spatially varying sensitivities of yield shocks to climate and also justifying our partitioning of Europe into several regions. Based on PDP,

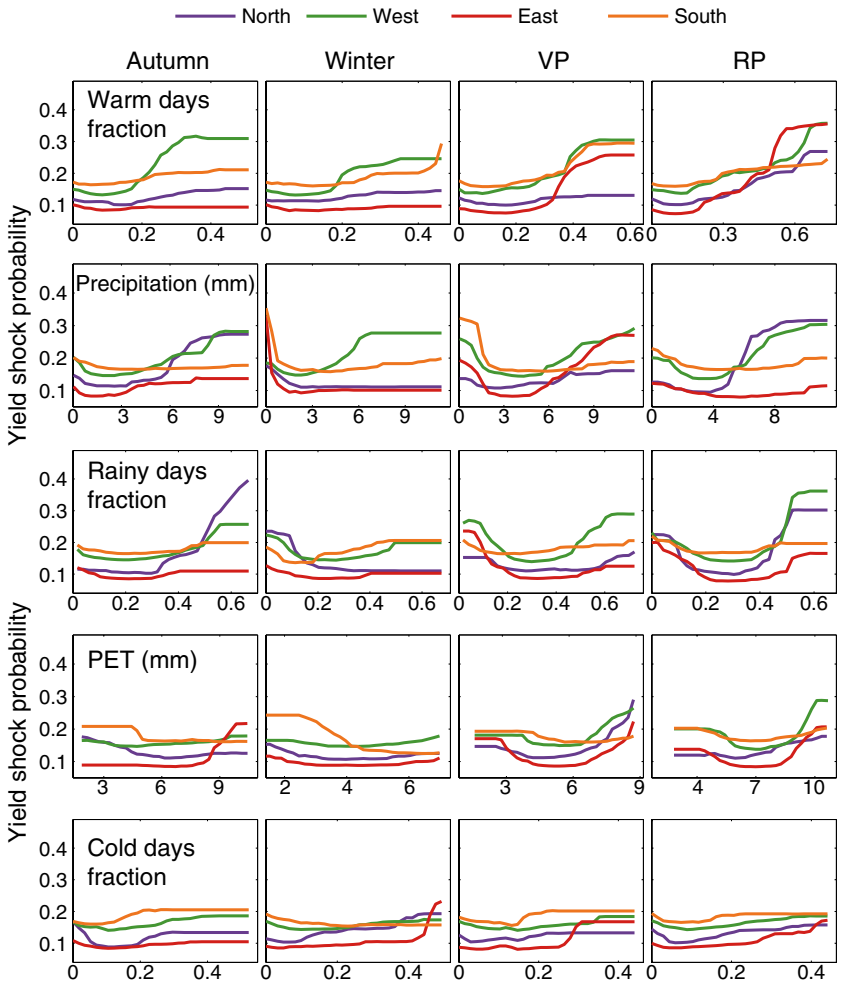

Figure 3. Response curves for yield shock probability generated through partial dependence plot with the RF models trained for the four regions. VP, vegetative period and RP, reproductive period.

we found a greater f_{wd} in a given region generally corresponded to a higher probability of yield shock, especially for extreme temperature in the reproductive period (Figure 3). As f_{wd} collectively accounted for >30°C heat stress or >90th percentile abnormal warming, the positive relationship suggested both heat stress and abnormal warming should be carefully considered in crop models to predict yield shock. For water supply (precipitation or f_{rd}), RF models suggested both low and excessive water supply can result in higher yield shock probability, especially for those in winter and in the reproductive period in North and West Europe (Figure 3). In terms of water demand (PET), higher PET generally corresponded to higher yield shock probability, especially in the vegetative and reproductive periods. We also noted that although water demand was usually highly correlated with warming, yield shock was less responsive to water demand relative to warming (Figure 3). The yield shock responded only weakly to cold stress, except for the winter cold stress in North and East Europe.

On the other hand, compared with the PDP of RF models based on official yield surveys, the crop model derived PDP suggested an overall positive but a weaker response of yield shock probability to f_{wd} (Figure S5). For water supply, lower water supply resulted in higher yield shock probability, especially for precipitation (Figure S5). However, the higher yield shock probability with excessive water supply as observed in Figure 3 cannot be reproduced by crop models. For PET and cold stress, crop model based PDP suggested a mild increase in yield shock probability with higher PET or cold stress.

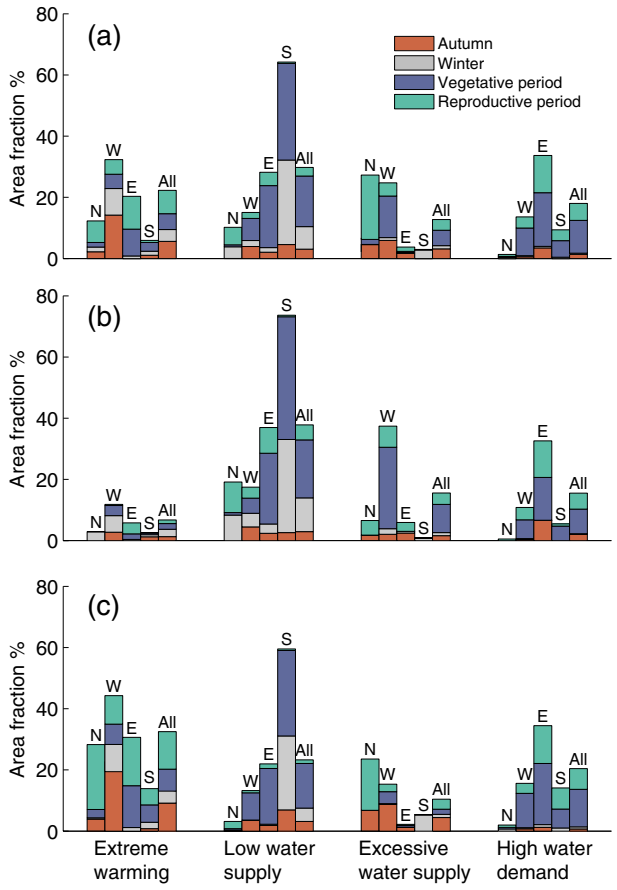

Figure 4. Area fractions of different primary climate driven yield shocks for the separate four regions (N: North Europe; W: West Europe; E: East Europe; S: South Europe) and the whole Europe (All). The first panel (a) is for the whole historic period (1980–2018) and the latter two panels are for divided two time period: 1980–1999 (b) and 2000–2018 (c).

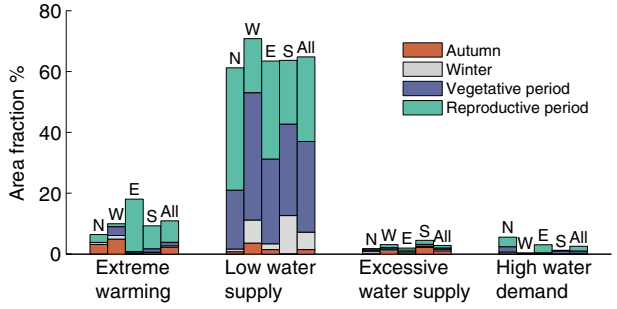

Figure 5. Area fractions of different primary climate driven yield shocks for the separate four regions and the whole Europe, based on the average of eight crop models. Results for each crop model can be found in Figure S7.

3.3. Attribution of Yield Shock Probability to Individual Climate Drivers

With yield shock probability being attributed to individual predictors (Figure 4a), the yield shocks primarily driven by the first four climate drivers in Table 2 comprised 83% of total yield shocks across Europe (cold stress was not presented, as cold stress driven yield shocks comprised less than 5% of total yield shocks). Across all studied countries, the attribution analysis suggested that low water supply was the most pervasive primary climate driver of shocks, followed with extreme warming, high water demand and excessive water supply (Figure 4a, although the dominant climate driver varied for each region). There was also a distinct spatial disparity about low water supply driven yield shock; the relatively drier and warmer regions (South and East Europe) were more dominated by low water supply than North and West Europe, and vice versa for the excessive water supply (Figure 4a). Seasonally, reproductive period was the critical stage for extreme warming driven yield shocks; however, for shocks driven by low water supply and high water demand the most important season was vegetative period, which was probably related with the lasting effect of water supply (Zhang & Oweis, 1999).

When the total time period was divided into two periods 1980–1999 and 2000–2018, for the whole Europe there was a distinct shift, such that the most pervasive emerging primary driver changed from low water supply to extreme warming from the earlier period to the later period (Figures 4b and 4c). However, low water supply in South Europe remained as the most pervasive primary driver. Seasonally, the two time periods had a similar pattern to the whole time period; reproductive period was the most important stage for extreme warming and vegetative period was the most important for low water supply and high water demand. Further, the crop models overrepresent the effect of low water supply on yield during the reproductive period compared to the yield survey data Figure 5.

3.4. Change in Climate Drivers of Wheat Yield Shocks due to Climate Change

We projected yield shock probability to two climate scenarios (RCP4.5 and RCP8.5) using the RF models trained with historical survey yield data. As expected, future warming resulted in higher yield shock probability and RCP8.5 showed even higher shock probability with RCP8.5 representing the higher emission scenario and warmer climate than RCP4.5 (Figure 6). Spatially, East European countries like Hungary and Romania were more prone to be influenced by yield shock in both scenarios. When the top 90th percentile of projected yield shock probability was identified as a yield shock event, the attribution analysis suggested that extreme warming was the most pervasive primary driver in both RCP4.5 and RCP8.5, while shocks driven by low water supply decreased significantly (Figure 7). This shift in yield shock drivers could be regarded as the extension of climate driver shift since 2000 as identified in Figure 4. Out of the 4 regions, East Europe had the most area with extreme warming as the primary driver in both scenarios (Figure 7), which might explain why East Europe had a higher yield shock probability in Figure 6.

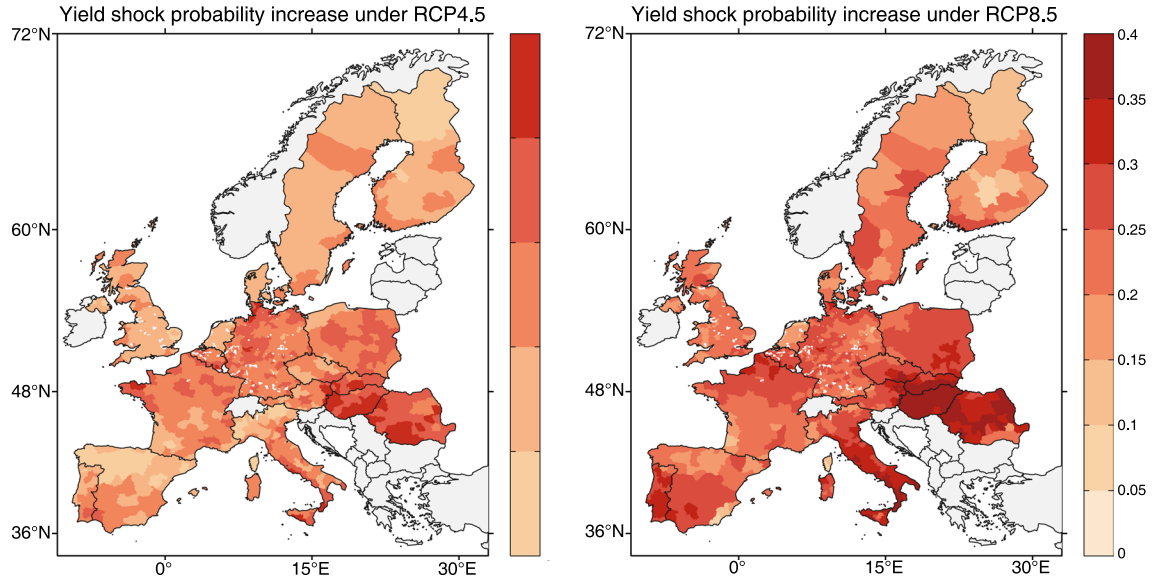

Figure 6. Spatial pattern of future yield shock probability increase (2070–2099) predicted with Random forest model under the scenarios of RCP4.5 and RCP8.5.

3.5. Yield Shock Prediction Agreement Between RF Models and Crop Models

Generally, crop models had a satisfactory performance for predicting yield shocks primarily driven by low water supply during 1980–2010. This category of yield shock constituted 32% of total yield shocks and more than 60% of them could be predicted by more than 3 out of 6 crop models during 1980–2010 (Figure 8a). However, there was a lower model agreement for other climate stress driven yield shocks, especially for the yield shock driven by autumn and winter (S1) stresses. For example, 73% yield shocks primarily driven by vegetative and reproductive period (S2) extreme warming could be predicted by more than 3 out of 6 crop model, while only 17% for yield shocks primarily driven by autumn and winter (S1) extreme warming (Figure 8a). Yield shock primarily driven by excessive water supply had a low model agreement for both stages (Figure 8a), which could be inferred from crop model response curves where yield shock probability was unresponsive to excessive water supply (Figure S5). Relative to the historic period, the model agreement became lower in future climate for all kinds of climate driven shock events (Figures 8b and 8c). This implies that in the future it might be even more challenging to forecast yield shock and design adaptation strategies accordingly using existing crop models as predictive tools.

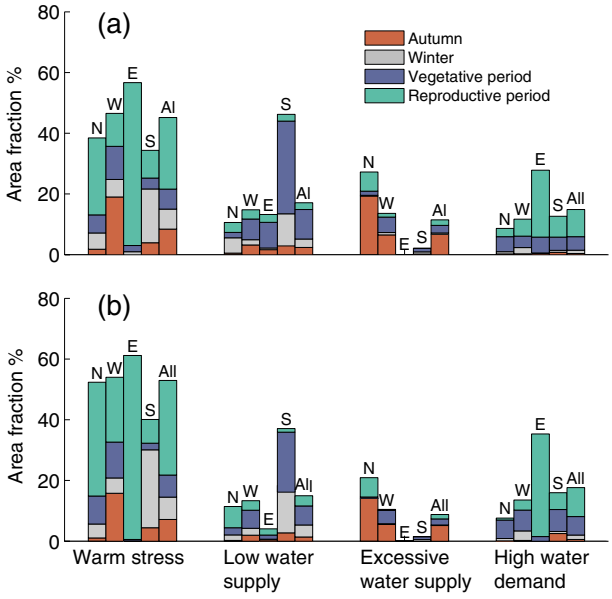

Figure 7. Area fractions of different primary climate driven yield shocks for the separate four regions (N: North Europe; W: West Europe; E: East Europe; S: South Europe) and the whole Europe (All). The yield shocks were predicted using Random forest models with future climate variables. The upper panel (a) is for RCP4.5 and the lower panel (b) is for RCP8.5.

4. Discussion and Conclusions

Here we built a data-driven attribution framework to investigate the primary climate drivers that may contribute to yield shocks in both historical and future climate scenarios. The attribution analysis suggests that for all studied areas low water supply was the dominant driver for yield shock. However, the most dominant climate drivers varied for each region: North Europe was dominated by excessive water, West Europe by excessive water, East Europe by high water demand, and South Europe by low water supply. This suggests that agricultural adaptation strategies should be tailored to meet the regional environment. Future warmer

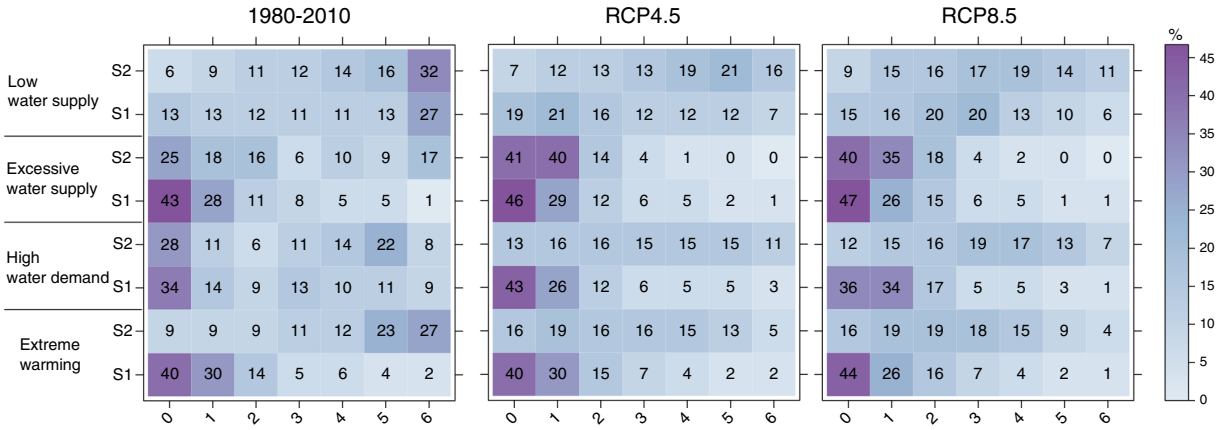

Figure 8. Yield shock prediction agreement between RF models and crop models. There are totally 6 crop models reported future yield simulations. Therefore, for each $Shock_{RF}$, it might be identified as yield shock by 0, 1, 2, ..., 6 crop models. $Shock_{RF}$ was grouped based on its primary driver on different seasons (S1 means the early season: autumn and winter; S2 means the later season: vegetative period and reproductive period). The numbers on the grid represents the percentage of yield shocks also identified as shock ($Shock_{sim}$) by 0, 1, 2, ..., 6 crop models, respectively. For example, for N yield shocks ($Shock_{RF}$) primarily driven by extreme warming in S1, there could be *N*0, *N*1, ..., *N*5, *N*6 (*N*0 + *N*1 + ... *N*5 + *N*6 = *N*) yield shocks also identified as shock ($Shock_{sim}$) by 0, 1, 2, ..., 6 crop models, respectively. A higher value of *N*3 + *N*4 +*N*5 + *N*6 suggested a higher confidence to predict the yield shock. RP, reproductive period.

climate may cause extreme warming to become the most pervasive climate driver of yield shock across all regions.

Unlike most studies where a machine learning algorithm was used to directly model yield variation (Crane-Droesch, 2018), here we classified yield data into yield shock and non-yield shock with the lowest tenth percentile yield as threshold. Then yield shock and non-yield shock was treated as a binary variable and a RF model was trained to classify them. This treatment allowed the RF model to focus on the extreme low yield cases while minimizing the influence of small yield variation on the estimation of yield response to climate. High AUC (>0.85) based on out-of-sample prediction suggested that the RF model produced a reliable classification. We also showed that the attribution analysis was robust to an alternative yield time series detrending method using Savitzky-Golay smoothing function (Figures S8–S9). We should note that although we divided Europe into 4 groups, our model was unable to capture the finer spatial variation of farmer-adapted management practices in each administrative unit, like shift in sowing date and irrigation implementation. For example, irrigation is expected to expand to maintain crop yield at future warmer climate (Peña-Arancibia et al., 2020); however, our projection did not consider the potential expansion of irrigation area in the future, which might overestimate the risk of future yield shock.

On the other hand, our selection of four regions is based on the relatively homogeneous weather and geographic patterns within each region. Dividing the whole data set into more regions might better account for the spatial heterogeneity but will also reduce the training sample size for each regional RF model. We also conducted a sensitivity test through using an alternative spatial aggregation method (Table S3). The attribution analysis based on this second spatial aggregation (Figure S10) also suggested that the major driver of yield shock during historical period is water limitation, although the climate drivers in each growing period differ between the two aggregation methods. For example, the extreme warming stress in Eastern Europe becomes more dominant yield shock driver in the alternative aggregation method (Figure S10). We looked into this and find it is probably because yield shock probability shows a steeper response to warming stress after removing Austria in the second spatial aggregation (Figure S11). This sensitivity test suggests that more attention need to be paid to spatial aggregation if more detailed administrative units become available in the future.

Since we did not account for the dynamics of crop phenology, which might reflect the farmer adaptation practices to climate change, we conducted a sensitivity test to see whether our conclusion will change with shifted crop phenology. When we shifted the crop phenological date by two weeks (±14 days), our main conclusion still holds, although a few differences were noticed in certain growing periods of climate drivers (Figure S12). Although our results are robust, we suggest that the dynamic information of crop phenology

should be considered when available to better account for the influence of farmer management practices on crop yield.

Our analysis confirmed that accounting for the accurate timing information of climate stress is important, which can be inferred from the different responsive curves for the same climate variable in different seasons. On the other hand, lower crop model predictive power for autumn and winter stress relative to the later growing season suggested that crop models need to improve the simulation of abnormal weather effects during these seasons. Studies have revealed that abnormal warm winter resulted in insufficient vernalizing days and affected wheat leaf and tiller number or flowing timing (Kirby, 1992; Wu et al., 2017), which is likely related to the recent yield shock in the breadbasket of France (Ben-Ari et al., 2018). In addition, the excess water stress was always poorly represented by crop models, which calls for incorporating water excess effects in crop models to improve their predictive performance (Li et al., 2019; Rosenzweig et al., 2002).

We note that yield shock could be influenced by single or compound climate extremes as well as non-extreme variability (Anderson et al., 2019; Beillouin et al., 2020; Ben-Ari et al., 2018; Ciais et al., 2005). The traditional yield loss attribution analysis based on crop models requires a set of model experiments wherein certain crop stress schemes are turned on or off and then their outputs are compared (Webber et al., 2018). However, climate stresses are often interrelated and the interaction effect between water and heat stress could be also important for determining crop yield (Jin et al., 2017; Zandalinas et al., 2018). Omitting either of them could lead to unintended results. Our data-driven attribution framework provides a flexible alternative, as it does not require a priori knowledge on crop physiological response to climate stress and can handle nonlinear responses. The interaction effects of climate stresses were implicitly considered in the RF model through the hierarchical classification tree of predictors within the structure of the model. Then, the game theory based attribution gave a reasonable quantification of the relative importance of climate stresses for either single or compound extremes.

Overall, our study analyzed the spatial-temporal pattern of wheat yield shock and applied a data-driven attribution framework to quantify primary climate drivers of yield shock during the historical period. Our attribution analysis suggests that the overall dominant climate driver of yield shock during 1980–2018 is water limitation. However, it was surpassed by extreme warming in the late period (2000–2018). The future projection based on a calibrated RF model revealed enhanced yield shock probability and a paradigm shift from water limitation dominated yield shock to extreme warming dominated yield shock with warmer climate. However, crop models generally attributed low water supply as the primary climate driver of yield shock and were unresponsive to excessive water supply. The subsequent yield shock agreement analysis suggested that yield shocks driven by autumn and winter season extreme warming, high water demand, and excessive water supply were not as easily predicted by the crop models compared to the RF model. The future warmer climate resulted in even lower performance of crop models in predicting yield shocks. This analysis suggests that these climate stress effects during specific growth stages should be improved in future crop models to better develop climate adaptation strategies in cropping systems.

Data Availability Statement

Simulated wheat yield data used in this study is available at https://zcloud2.zalf.de/s/ZfsYAQ54nFdFJEX. The climate data can be accessed at: http://open-research-data-zalf.ext.zalf.de/ResearchData/DK_59.html.

Acknowledgments
The authors would like to thank the anonymous reviewers for their very valuable feedback. This work has been funded by the ANR CLAND Convergence Institute.

References

Aghakouchak, A., Chiang, F., Huning, L. S., Love, C. A., Mallakpour, I., Mazdiyasni, O., et al. (2020). Climate extremes and compound hazards in a warming world. *Annual Review of Earth and Planetary Sciences*, *48*, 519–548. https://doi.org/10.1146/annurev-earth-071719-055228

Alexandratos, N., & Bruinsma, J. (2012). World agriculture towards 2030/2050. *Land Use Policy*, *20*(4), 375. https://doi.org/10.1016/S0264-8377(03)00047-4

Allen, R. G., Pereira, L. S., Raes, D., & Smith, M. (1998). Crop evapotranspiration: Guidelines for computing crop requirements. *Irrigation and Drainage Paper* No. 56 (Vol. 300). FAO.

Anderson, W. B., Seager, R., Baethgen, W., Cane, M., & You, L. (2019). Synchronous crop failures and climate-forced production variability. *Science Advances*, *5*(7), eaaw1976. https://doi.org/10.1126/sciadv.aaw1976

Asseng, S., Ewert, F., Martre, P., Rötter, R. P., Lobell, D. B., Cammarano, D., et al. (2015). Rising temperatures reduce global wheat production. *Nature Climate Change*, *5*, 143–147. https://doi.org/10.1038/nclimate2470

Beillouin, D., Schauberger, B., Bastos, A., Ciais, P., & Makowski, D. (2020). Impact of extreme weather conditions on European crop production. *Philosophical Transactions of the Royal Society B*, *375*(1810), 20190510. https://doi.org/10.1098/rstb.2019.0510

Ben-Ari, T., Adrian, J., Klein, T., Calanca, P., Van der Velde, M., & Makowski, D. (2016). Identifying indicators for extreme wheat and maize yield losses. *Agricultural and Forest Meteorology*, *220*, 130–140. https://doi.org/10.1016/j.agrformet.2016.01.009

Ben-Ari, T., Boé, J., Ciais, P., Lecerf, R., Van Der Velde, M., & Makowski, D. (2018). Causes and implications of the unforeseen 2016 extreme yield loss in the breadbasket of France. *Nature Communications*, *9*(1), 1–10. https://doi.org/10.1038/s41467-018-04087-x

Breiman, L. (2001). Random forests. *Machine learning*, *45*, 5–32.

Chen, B., & Villoria, N. B. (2019). Climate shocks, food price stability and international trade: Evidence from 76 maize markets in 27 net-importing countries. *Environmental Research Letters*, *14*(1), 014007. https://doi.org/10.1088/1748-9326/aaf07f

Ciais, P., Reichstein, M., Viovy, N., Granier, A., Ogée, J., Allard, V., et al. (2005). Europe-wide reduction in primary productivity caused by the heat and drought in 2003. *Nature*. *437*(7058), 529–533. https://doi.org/10.1038/nature03972

Cleveland, W. S. (1979). Robust locally weighted regression and smoothing scatterplots. *Journal of the American statistical association*, *74*(368), 829–836. https://doi.org/10.1080/01621459.1979.10481038

Cottrell, R. S., Nash, K. L., Halpern, B. S., Remenyi, T. A., Corney, S. P., Fleming, A., et al. (2019). Food production shocks across land and sea. *Nature Sustainability*, *2*(2), 130–137. https://doi.org/10.1038/s41893-018-0210-1

Crane-Droesch, A. (2018). Machine learning methods for crop yield prediction and climate change impact assessment in agriculture. *Environmental Research Letters*, *13*(11), 114003. https://doi.org/10.1088/1748-9326/aae159

Crost, B., Duquennois, C., Felter, J. H., & Rees, D. I. (2018). Climate change, agricultural production and civil conflict: Evidence from the Philippines. *Journal of Environmental Economics and Management*, *88*, 379–395. https://doi.org/10.1016/j.jeem.2018.01.005

Cutler, A., Cutler, D. R., & Stevens, J. (2012). *Ensemble machine learning: Random forests* (pp. 225–250). Springer.

Diffenbaugh, N. S., Singh, D., Mankin, J. S., Horton, D. E., Swain, D. L., Touma, D., et al. (2017). Quantifying the influence of global warming on unprecedented extreme climate events. *Proceedings of the National Academy of Sciences of the United States of America*, *114*, 4881–4886. https://doi.org/10.1073/pnas.1618082114

Duncan, J. M. A., Dash, J., & Atkinson, P. M. (2015). Elucidating the impact of temperature variability and extremes on cereal croplands through remote sensing. *Global Change Biology*, *21*, 1541–1551. https://doi.org/10.1111/gcb.12660

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, *27*, 861–874. https://doi.org/10.1016/j.patrec.2005.10.010

Field, C. B., Barros, V., Stocker, T. F., & Dahe, Q. (Eds.). (2012). *Managing the risks of extreme events and disasters to advance climate change adaptation: Special report of the intergovernmental panel on climate change*. Cambridge University Press.

Fishman, R. (2016). More uneven distributions overturn benefits of higher precipitation for crop yields. *Environmental Research Letters*, *11*(2), 024004.

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, *29*, 1189–1232. https://doi.org/10.1214/aos/1013203451

Guarin, J. R., Asseng, S., Martre, P., & Bliznyuk, N. (2020). Testing a crop model with extreme low yields from historical district records. *Field Crops Research*, *249*, 107269. https://doi.org/10.1016/j.fcr.2018.03.006

Hertel, T. W., Burke, M. B., & Lobell, D. B. (2010). The poverty implications of climate-induced crop yield changes by 2030. *Global Environmental Change*, *20*(4), 577–585. https://doi.org/10.1016/j.gloenvcha.2010.07.001

Iizumi, T., & Sakai, T. (2020). The global dataset of historical yields for major crops 1981–2016. *Scientific Data*, *7*(1), 1–7. https://doi.org/10.1038/s41597-020-0433-7

Jin, Z., Zhuang, Q., Wang, J., Archontoulis, S. V., Zobel, Z., & Kotamarthi, V. R. (2017). The combined and separate impacts of climate extremes on the current and future US rainfed maize and soybean production under elevated CO2. *Global Change Biology*, *23*(7), 2687–2704. https://doi.org/10.1111/gcb.13617

Kirby, E. J. M. (1992). A field study of the number of main shoot leaves in wheat in relation to vernalization and photoperiod. *Journal of Agricultural Science*, *118*, 271–278. https://doi.org/10.1017/s0021859600070635

Lesk, C., Rowhani, P., & Ramankutty, N. (2016). Influence of extreme weather disasters on global crop production. *Nature*, *529*, 84–87. https://doi.org/10.1038/nature16467

Li, Y., Guan, K., Schnitkey, G. D., DeLucia, E., & Peng, B.K. (2019). Excessive rainfall leads to maize yield loss of a comparable magnitude to extreme drought in the United States. *Global Change Biology*, *25*(7), 2325–2337. https://doi.org/10.1111/gcb.14628

Lobell, D. B., Sibley, A., & Ivan Ortiz-Monasterio, J. (2012). Extreme heat effects on wheat senescence in India. *Nature Climate Change*, *2*, 186–189. https://doi.org/10.1038/nclimate1356

Molnar, C., Casalicchio, G., & Bischl, B. (2018). IML: An R package for interpretable machine learning. *Journal of Open Source Software*, *3*, 786. https://doi.org/10.21105/joss.00786

Müller, C., Elliott, J., Chryssanthacopoulos, J., Arneth, A., Balkovic, J., Ciais, P., et al. (2017). Global gridded crop model evaluation: Benchmarking, skills, deficiencies and implications. *Geoscientific Model Development*, *10*(4), 1403–1422. https://doi.org/10.5194/gmd-10-1403-2017

Padarian, J., McBratney, A. B., & Minasny, B. (2020). Game theory interpretation of digital soil mapping convolutional neural networks. *Soil*, *6*, 389–397. https://doi.org/10.5194/soil-6-389-2020

Peña-Arancibia, J. L., Mainuddin, M., Ahmad, M. D., Hodgson, G., Murad, K. F. I., Ticehurst, C., et al. (2020). Groundwater use and rapid irrigation expansion in a changing climate: Hydrological drivers in one of the world's food bowls. *Journal of Hydrology*, *581*, 124300.

Porter, J. R., & Semenov, M. A. (2005). Crop responses to climatic variation. *Philosophical Transactions of Royal Society B*, *360*(1463), 2021–2035. https://doi.org/10.1098/rstb.2005.1752

Rosenzweig, C., Tubiello, F. N., Goldberg, R., Mills, E., & Bloomfield, J. (2002). Increased crop damage in the US from excess precipitation under climate change. *Global Environmental Change*, *12*(3), 197–202. https://doi.org/10.1016/s0959-3780(02)00008-0

Rötter, R. P., Appiah, M., Fichtler, E., Kersebaum, K. C., Trnka, M., & Hoffmann, M. P. (2018). Linking modelling and experimentation to better capture crop impacts of agroclimatic extremes—A review. *Field Crops Research*, *221*, 142–156. https://doi.org/10.1016/j.fcr.2018.02.023

Rötter, R. P., Carter, T. R., Olesen, J. E., & Porter, J. R. (2011). Crop-climate models need an overhaul. *Nature Climate Change*, *1*(4), 175–177. https://doi.org/10.1038/nclimate1152

Ruane, A. C., Goldberg, R., & Chryssanthacopoulos, J. (2015). Climate forcing datasets for agricultural modeling: Merged products for gap-filling and historical climate series estimation. *Agricultural and Forest Meteorology*, *200*, 233–248. https://doi.org/10.1016/j.agrformet.2014.09.016

Siebers, M. H., Slattery, R. A., Yendrek, C. R., Locke, A. M., Drag, D., Ainsworth, E. A., et al. (2017). Simulated heat waves during maize reproductive stages alter reproductive growth but have no lasting effect when applied during vegetative stages. *Agriculture, Ecosystems & Environment*, *240*, 162–170. https://doi.org/10.1016/j.agee.2016.11.008

Tilman, D., Balzer, C., Hill, J., & Befort, B. L. (2011). Global food demand and the sustainable intensification of agriculture. *Proceedings of the National Academy of Sciences*, *108*(50), 20260–20264. https://doi.org/10.1073/pnas.1116437108

Toreti, A. (2014). *Gridded agro-meteorological data in Europe*. European Commission, Joint Research Centre (JRC) PID. Retrieved from http://data.europa.eu/89h/jrc-marsop4-7-weather_obs_grid_2019

van der Velde, M., Wriedt, G., & Bouraoui, F. (2010). Estimating irrigation use and effects on maize yield during the 2003 heatwave in France. *Agriculture, Ecosystems & Environment*, *135*(1–2), 90–97. https://doi.org/10.1016/j.agee.2009.08.017

Wang, E., Martre, P., Zhao, Z., Ewert, F., Maiorano, A., Rötter, R. P., et al. (2017). The uncertainty of crop yield projections is reduced by improved temperature response functions. *Nature Plants*, *3*(8), 1–13. https://doi.org/10.1038/s41477-017-0032-6

Webber, H., Ewert, F., Olesen, J. E., Müller, C., Fronzek, S., Ruane, A. C., et al. (2018). Diverging importance of drought stress for maize and winter wheat in Europe. *Nature Communications*, *9*(1), 1–10. https://doi.org/10.1038/s41467-018-06525-2

Webber, H., Lischeid, G., Sommer, M., Finger, R., Nendel, C., Gaiser, T., & Ewert, F. (2020). No perfect storm for crop yield failure in Germany. *Environmental Research Letters 15*, 104012.

Wright, M. N., Wager, S., & Probst, P. (2017). *CRAN - Package ranger: A Fast Implementation of Random Forests, Version 0.8.0*.

Wu, X., Liu, H., Li, X., Tian, Y., & Mahecha, M. D. (2017). Responses of winter wheat yields to warming-mediated vernalization variations across temperate Europe. *Frontiers in Ecology and Evolution*, *5*, 126. https://doi.org/10.3389/fevo.2017.00126

Zampieri, M., Ceglar, A., Dentener, F., & Toreti, A. (2017). Wheat yield loss attributable to heat waves, drought and water excess at the global, national and subnational scales. *Environmental Research Letters*, *12*(6), 064008. https://doi.org/10.1088/1748-9326/aa723b

Zandalinas, S. I., Mittler, R., Balfagón, D., Arbona, V., & Gómez-Cadenas, A. (2018). Plant adaptations to the combination of drought and high temperatures. *Physiologia Plantarum*, *162*, 2–12. https://doi.org/10.1111/ppl.12540

Zhang, H., & Oweis, T. (1999). Water-yield relations and optimal irrigation scheduling of wheat in the Mediterranean region. *Agricultural Water Management*, *38*(3), 195–211. https://doi.org/10.1016/s0378-3774(98)00069-9

Zheng, B., Chenu, K., Fernanda Dreccer, M., & Chapman, S. C. (2012). Breeding for the future: What are the potential impacts of future frost and heat events on sowing and flowering time requirements for Australian bread wheat (*Triticum aestivium*) varieties? *Global Change Biology*, *18*, 2899–2914. https://doi.org/10.1111/j.1365-2486.2012.02724.x

Zhu, P., Zhuang, Q., Archontoulis, S. V., Bernacchi, C., & Müller, C. (2019). Dissecting the nonlinear response of maize yield to high temperature stress with model-data integration. *Global Change Biology*, *25*(7), 2470–2484. https://doi.org/10.1111/gcb.14632

Zscheischler, J., Westra, S., Van Den Hurk, B. J. J. M., Seneviratne, S. I., Ward, P. J., Pitman, A., et al. (2018). Future climate risk from compound events. *Nature Climate Change*, *8*(6), 469–477. https://doi.org/10.1038/s41558-018-0156-3