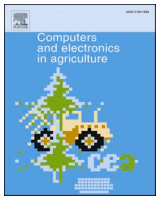
Computers and Electronics in Agriculture 218 (2024) 108688

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag



# Compound drought and hot stresses projected to be key constraints on maize production in Northeast China under future climate

Chuanwei Zhang [a,b], Jiangbo Gao [a,*], Lulu Liu [a], Shaohong Wu [a,b]

[a] Key Laboratory of Land Surface Pattern and Simulation, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, China
[b] University of Chinese Academy of Sciences of Resources and Environment, Beijing, China

ARTICLE INFO

*Keywords:*
Compound drought and hot
Climate change
Maize
APSIM
Northeast China

ABSTRACT

Climate change and the increasing frequency of climate extremes associated with warming have been the most important climatic stressors for maize production. However, crop-model based assessments of the major determinants of yield variability at regional scale under future climate conditions are still underrepresented. In this study, we simulated maize yield in Northeast China at a reference period (1986–2005), and two future periods (2030 s: 2020–2039, 2050 s: 2040–2059) of Shared Socioeconomic Pathways (SSPs) of SSP1-2.6, SSP2-4.5 and SSP5-8.5 using Agricultural Production sIMulator (APSIM). We first characterized the variations of maize yield under climate change based on the simulations, and further investigated the key determinants of yield variability using machine learning techniques. The results suggest that maize yield would decrease by 14.8 % to 19.6 % (depending on the climate scenarios) compared to the reference period without adaption. Random forest performed best in explaining yield variability in a suite of machine learning models (extreme gradient boosting, gradient boosting, classification and regression tree, ridge and lasso regression and random forest), with a mean $R^2$ of 0.77, a mean RMSE of 1239.2 kg/ha, a mean MAE of 885.1 kg/ha. Extreme climate indicators show a greater ability to explain yield variability in over half agro-ecological regions under higher warming levels such as SSP5-8.5. Cumulative precipitation (CPR) and compound drought and hot days (CDH) during growing seasons are the most important mean and extreme climate indicators affecting yield variability, respectively. Moreover, the effect sizes of CPR and CDH on yield variability are 1898 kg/ha and 5596 kg/ha, respectively. Therefore, CDH will be a key constraint on maize production. Future adaptive measures such as irrigation, breeding hot- or drought-tolerant cultivars should be implemented to enhance the resilience of maize crops in the face of climate change.

## 1. Introduction

Global warming, along with the increasing frequency of climate extremes associated with this warming trend, is projected to continue in 21st century (Van et al., 2021). Climate change has far-reaching impacts on various sectors such as agriculture, water, and energy, of which agriculture is the highly sensitive one as it is directly exposed to the shifting climate conditions (Feng et al., 2020). There is growing evidence suggesting the negative impacts of global warming and climate extremes on crop yield, both in the observations and projections (Zhao et al., 2017; Lesk et al., 2016). Therefore, it is of paramount importance to understand the impacts of climate variability on crop yield, which will provide essential evidence for policymakers, farmers and breeders to make informed adaptions. Simulating crop yield is a prerequisite for exploring the climate-yield relationship under future climate conditions. Process-based crop models, statistical models, and machine learning techniques are commonly used approaches for simulating crop yield (Xiao et al., 2021; Feng et al., 2020; Stevens et al., 2016). Crop models can offer us in-depth physiological explanations for yield variability compared to statistical models and machine learning techniques (Meroni et al., 2021). However, the extensive data requirements, including soil characteristics, meteorological variables, and eco-physiological parameters needed to describe crop variability for crop models, impose limitations on their application at regional scale (Zhang et al., 2020). Once properly calibrated and validated with observed data, crop models can offer insights into the complex interactions between soil, crop, weather and management, which can contribute to develop effective strategies for mitigating the impacts of climate change on

* Corresponding author at: No. 11A Datun Road, Chaoyang District, Beijing, China.
*E-mail address:* gaojiangbo@igsnrr.ac.cn (J. Gao).

https://doi.org/10.1016/j.compag.2024.108688
Received 16 February 2023; Received in revised form 31 December 2023; Accepted 29 January 2024
Available online 9 February 2024

agriculture (Laux et al., 2010). The comprehensive impacts of climate variability on crop yield have primarily been investigated using statistical models (Ceglar et al., 2016). However, statistical models generally focused on establishing the regression relationships between meteorological variables and yield variability, but they may overlook the issue of multicollinearity among these variables. Furthermore, few studies have taken into consideration the effects of compound events on crop yield in constructing statistical and machine learning models. The effects of meteorological variables on crop yield will possibly be exaggerated or weakened without accounting for the multicollinearity of meteorological variables and compound climate extremes. Recently, machine learning has gained widespread usage in disentangling complex relationships within large datasets, which can be challenging for conventional statistical methods (Nayak et al., 2022).

Climate variability can impact crop yield by limiting photosynthesis, accelerating phenological development, and reducing seed number or nutrient absorption during different phases of crop growth (Cohen et al., 2021). Crop sensitivity to climate variability is contingent upon the crop's eco-physiological parameters and growing environment. In other words, the response of a crop to climate variability are crop-specific and region-specific (Leng and Huang, 2017). However, there is insufficient investigation into the climate-crop relationship at regional perspective to meet the demand for precise agricultural management. The Northeast China (NEC) is one of the important grain producing areas in China. Over the years, the grain yield in this region has shown remarkable growth, increasing from $3.70 \times 10^7$ tons in 1980 to $16.54 \times 10^7$ tons in 2019, which accounted for 25 % of the total grain yield in the country (Li et al., 2016). And the area of maize occupies approximately 52 % of total agriculture land in NEC (Li et al., 2021), highlighting the crucial role of maize production in the region in ensuring food security in China.

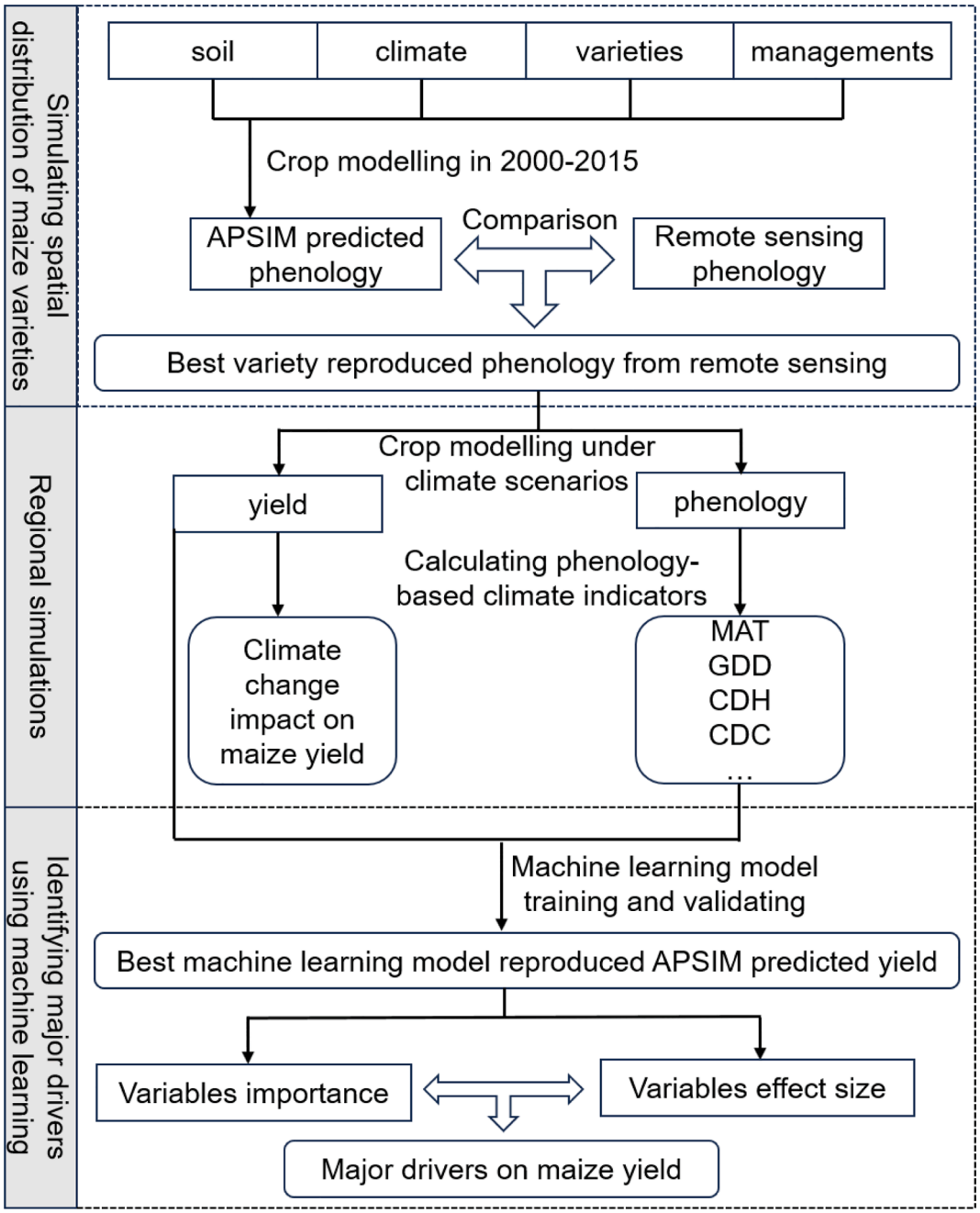



**Fig. 1.** Flowchart of assessing the impact of climate change on maize yield and identifying the major drivers on maize yield in Northeast China.

However, maize production in NEC is highly sensitive to climate change as it is mainly a rain-fed crop in the region, more serious effect on maize will occur without the mitigating effect of irrigation (Wang et al., 2018b).

Overall, there are still research gaps in investigating the key determinants of yield variability under future climate conditions at regional scale using methods that ignoring collinearity. In this study, we simulated maize yield at 0.25° ×0.25° resolution at a reference period (1986–2005), and two future periods (2030 s: 2020–2039, 2050 s:2040–2059) of Shared Socioeconomic Pathways (SSPs) of SSP1-2.6, SSP2-4.5, and SSP5-8.5 using Agricultural Production sIMulator (APSIM). As shown in Fig. 1, our major objectives were: two-fold: (1) investigate the maize yield variability under future climate conditions at regional perspective, (2) identify the key meteorological variables affecting maize yield variability. Our results will provide valuable regional information on the impact of meteorological variables on maize production in NEC, which can be instrumental in guiding agricultural management practices to achieve a high and stable yield level in the region.

## 2. Material

### 2.1. Study area

In our study, we focused on Northeast China (NEC), a prominent grain production region in the country (Fig. 2A). The prevailing climate type in the region is continental monsoon climate, with average annual temperature ranging from -1 °C to 9 °C and average annual precipitation ranging from 500 mm to 800 mm (Li et al., 2014). NEC is a base for commodity grain in China and boasts the third-largest black soil area globally. The total black soil area is about 1.09 million km$^2$, constituting 12 % of the global black soil expanse. Approximately 20 % of China's total grain yield is produced in the black soil regions. Nevertheless, the elevated intensity of land use, inadequate agricultural management, and the impact of climate change have led to soil erosion, posing great threat to the sustainable development of agricultural production (Xie et al., 2019). Research has shown that the annual losses of soil organic carbon in NEC reached up to 2.05 Mg ha$^{-1}$ (Qiu et al., 2004).

### 2.2. Map of maize planting region and agro-ecological regions

We overlapped the boundary map of NEC onto the map of agricultural physical regionalization in China to obtain the Fig.S1. Subsequently, Region 2 (Fig.S1) was subdivided into three distinct sub-regions of II, III, and VI (Fig. 2B) using the provincial boundary to (1) obtain more detailed insights into the influence of climate change on yield variability; (2) account for the spatial heterogeneity of soil indicators across different provinces (Liu et al., 2009). The map depicting the maize planting region (Fig. 2C) was extracted from the dataset created by Luo et al. (2020a) (https://data.mendeley.com/datasets/jbs44b2hrk/2). The grids labelled as maize planting area for more than six consecutive years during 2000–2015 were used to generate the map.

### 2.3. Historical and future climate data

The China Meteorological Forcing Dataset (CMFD) used for historical yield projections (1986–2005) is a reanalyzed dataset based on observations from more than 2000 meteorological sites, along with GLDAS and TRMM precipitation data (Yang et al., 2018). Correspondingly, the newest version of the NASA Earth Exchange Global Daily Downscaled Projections named NEX-GDDP-CMIP6 (Thrasher et al., 2022), was employed to conduct future yield projections at periods of 2030 s (2020–2039) and 2050 s (2040–2059). This CMIP6 climate dataset was created using the statistical downscaling algorithm (Wood et al., 2002; Wood et al., 2004), and it has been bias-corrected using the newest version of the Global Meteorological Forcing Dataset (GMFD). The CMFD and NEX-GDDP-CMIP6 are both at daily scale. The CMFD was interpolated from 0.1°×0.1° resolution to 0.25°×0.25° resolution using bilinear interpolation in *Matlab* to make it comparable with NEX-GDDP-CMIP6. Four variables including maximum temperature (Tmax), minimum temperature (Tmin), mean temperature (Tas), precipitation (Pr), and shortwave solar radiation (Srad) were input into the crop model to simulate maize yield, and the four variables along with mean temperature (Tas) were used to calculate mean and extreme climate indicators (see section 3.3). Overall, we employed five General Circulation Moldes (GCMs) in each of the SSP scenario. The information of climate dataset used in this study are summarized in Table 1.

### 2.4. Soil data

A high-resolution gridded dataset of soil characteristics for use in the land surface modeling was employed to extract the soil variables such as PH, organic matter fraction, sand, silt, bulk density and clay fraction for APSIM-Maize (Shangguan et al., 2013). The resolution of the dataset is 30 arc-seconds and the variation of soil property was captured by eight layers to the depth of 2.3 m (i.e., 0—0.045, 0.045—0.091, 0.091—0.166, 0.166—0.289, 0.289—0.493, 0.493—0.829, 0.829—1.383 and 1.383—2.296 m).

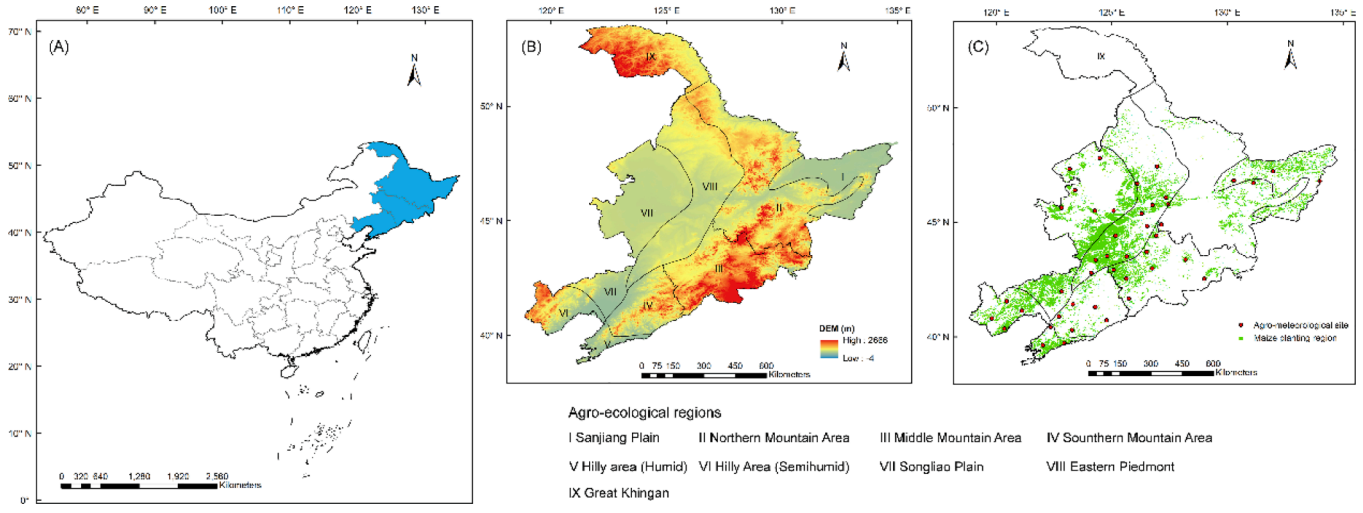



**Fig. 2.** Location (A), DEM and agro-ecological regions (B), maize planting regions and agro-meteorological sites (C) of Northeast China.

**Table 1**
Summary of climate data used in this study.

| Dataset | Model | Country | Resolution | Time | Experiment | Variables |
|---|---|---|---|---|---|---|
| CMFD | NA | China | 0.1°×0.1° | 1979–2018 | Observed | Tmax, Tmin, Tas, Pr, Srad |
| NEX-GDDP-CMIP6 | ACCESS-CM2 | Australia | 0.25°×0.25° | 1950–2100 | Historical, SSP1-2.6, SSP2-4.5, SSP5-8.5 | Tmax, Tmin, Tas, Pr, Srad |
| | CanESM5 | Canada | | | | |
| | MIROC6 | Japan | | | | |
| | MPI-ESM1-2-LR | Germany | | | | |
| | NorESM2-MM | Norway | | | | |

Note that the China Meteorological Forcing Dataset (CMFD) is a reanalyzed dataset based on observations, obtained from the Tibetan Plateau Data Center (https://data.tpdc.ac.cn); and the NASA Earth Exchange Global Daily Downscaled Projections (NEX-GDDP-CMIP6) is downscaled based on the General Circulation Model (GCM) runs under the Coupled Model Intercomparison Project Phase 6 (CMIP6), and it has been bias-corrected based on the Global Meteorological Forcing Dataset (GMFD). The NEX-GDDP-CMIP6 was obtained from NASA (https://doi.org/10.7917/OFSG3345). The historical period of NEX-GDDP-CMIP6 is 1950–2014, while the future period of SSPs is 2015–2100. Abbreviations: Maximum air temperature, Tmax; Minimum air temperature, Tmin; Mean air temperature, Tas; Precipitation, Pr; Shortwave solar radiation, Srad.

## 3. Method

### 3.1. Crop parameters for APSIM-Maize

APSIM (https://www.apsim.info/apsim-model/) is a process-based model that widely used in simulating climate change impacts on crop yield (Wang et al., 2018a). In APSIM, TTEE (thermal time required from emergence to end of juvenile, °C d) and TTFM (thermal time required from flower to maturity, °C d) are used to control maize development, while the GGR (potential grain growth rate, mg grain$^{-1}$ d$^{-1}$) and HGNM (maximum grain numbers per head, grain$^{-1}$) are related to maize productivity. Calibrating these parameters at each grid is unpractical due to limited on-farm observations. We simulated maize yield and phenology using eight typical varieties (Table 2) at each grid during period of 2000–2015, then variety reproduced the maturing dates from the maize phenology dataset from Luo et al. (2020b) best was selected as the default variety in the grid. We validated the results using yield records from agro-meteorological sites (Fig. 2C) and from city-level statistical yearbooks. As shown in Fig. 3, the $R^2$ values are 0.65 and 0.82 for point and regional scale, respectively. And RMSE values are 1011.1 kg/ha and 529.5 kg/ha for point and region scale, respectively.

### 3.2. Crop modelling

Several basic settings should be made before running the crop model: (1) the irrigation amount was set to zero as maize in NEC is major a rainfed crop (Li et al., 2022; Yang et al., 2015); (2) fertilizer was set to 300 kg/ha urea at sowing to fully meet nutrient need of maize growth; (3) sowing density was set to 67,500 plants/ha, (4) sowing date was fixed in each grid, which was acquired through interpolating the sowing dates recorded in agro-meteorological sites; (5) sowing depth and sowing space were set to 5 cm and 60 cm, respectively. The APSIM-maize doesn't consider the effect of changes in $CO_2$ concentration on maize yield. However, the transpiration rate of maize increased 37 % as the concentration of $CO_2$ increased from 350 ppm to 700 ppm (Lobell et al., 2015). Therefore, we modified the transpiration rate in the xml file of APSIM-maize according to the linear response of maize transpiration to $CO_2$. Calculations of $CO_2$ were referenced to Xiao et al. (2020), and the corresponding equations are shown in supplementary text 1.

**Table 2**
The cultivar parameters of the representative cultivars in Northeast China.

| Cultivar | Maturity | Genetic parameters | | | | Reference |
|---|---|---|---|---|---|---|
| | | TTEE | TTFM | GGR | HGNM | |
| Jidan120 | Early | 75 | 680 | 9 | 650 | Huang et al., 2020 |
| Haiyu6 | Early | 50 | 720 | 9.5 | 550 | Liu et al., 2012 |
| Sidan19 | Medium | 90 | 790 | 9 | 600 | Liu et al., 2012 |
| Jidan101 | Medium | 100 | 700 | 9.5 | 650 | Liu et al., 2012 |
| Danyu13 | Medium | 110 | 730 | 9.5 | 600 | Liu et al., 2012 |
| Baidan9 | Medium | 140 | 700 | 9.8 | 650 | Liu et al., 2012 |
| Zhengdan958 | Mid-late | 200 | 700 | 12 | 660 | Huang et al., 2020 |
| Danyu39 | Late | 200 | 1100 | 8 | 620 | Huang et al., 2020 |

Abbreviations: TTEE (tt_emerg_to_endjuv) is the thermal time required from emergence to end of juvenile (°C d); TTFM (tt_flower_to_matrurity) is the thermal time required from flower to maturity (°Cd); GGR (grain_gth_rate) is the potential grain growth rate (mg grain$^{-1}$ d$^{-1}$); HGNM (head_grain_no_max) is the maximum grain numbers per head (grain$^{-1}$).

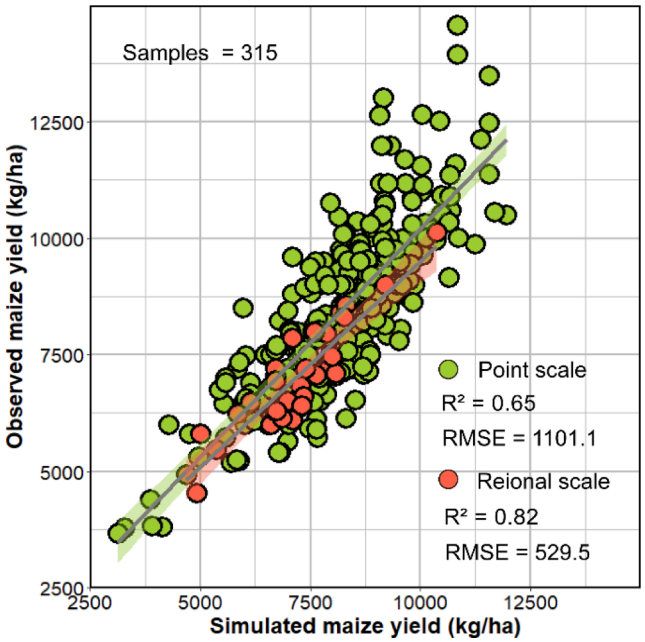



**Fig. 3.** Comparisons of APSIM simulated yields with observed yields from agro-meteorological sites (point scale) and city-level statistical yearbook (regional scale).

### 3.3. Calculating mean and extreme climate indicators

We chose 10 climate indicators categorized as mean and extreme climate to explain the yield variability under climate change (Table 3). The climate indicators including mean daily maximum temperature (MAT); mean daily minimum temperature (MIT); cumulative precipitation (CPR); cumulative shortwave radiation (CSR); cumulative effective growing degree days (GDD); cumulative compound drought and hot days (CDH); cumulative compound drought and cold days (CDC);cumulative days of daily maximum temperature higher than 30 °C (OCA); cumulative days of daily minimum temperature lower than 8 °C (OCI); and cumulative drought days (TDR). The drought and high

**Table 3**
Mean climate and extreme climate indicators during maize growing seasons. The climate indicators were calculated based on the APSIM-Maize simulated phenology.

| Variable name | Description | Type | Unit |
|---|---|---|---|
| MAT | Average daily maximum temperature | Mean climate | °C |
| MIT | Average daily minimum temperature | Mean climate | °C |
| CPR | Cumulative precipitation | Mean climate | mm |
| CSR | Cumulative shortwave radiation | Mean climate | MJ/m2/d |
| GDD | Growing degree days | Mean climate | °Cd |
| CDH | Cumulative compound drought and hot days | Extreme climate | day |
| CDC | Cumulative compound drought and cold days | Extreme climate | day |
| OCA | frequency of daily maximum temperature above 30 °C | Extreme climate | day |
| OCI | frequency of daily minimum temperature below 8 °C | Extreme climate | day |
| TDR | total drought days | Extreme climate | day |

temperature/low temperature occur simultaneous in a day was defined as a compound drought and hot/cold day (CDH/CDC). The time window for calculating the climate indicators was from sowing to maturing.

The Daily Evapotranspiration Deficit Index (DEDI) was used to assess the water deficit in each growing day, it was calculated as:

$$DEDI_i = \frac{D_i - D_{AVE}}{D_{STU}}, D_i = AET_i - PET_i \quad (1)$$

Where $D_{AVE}$, $D_{STU}$ are the daily average evapotranspiration deficit and standard deviation of daily evapotranspiration deficit during growing seasons of reference period, $D_i$ is the actual daily evapotranspiration deficit (mm/d). The 30th quantile of DEDI was set as the threshold for the drought day (Zhang et al., 2022).

GDD was calculated as:

$$GDD = \sum_{i=m}^{n} max(0, Tmean_i - 10) \quad (2)$$

where $m$ and $n$ are the beginning and ending dates for the growing season for maize, $Tmean_i$ is the daily average temperature on day $i$.

### 3.4. Statistical analysis

Maize yield variation was calculated using equation as follows:

$$\Delta Y(\%) = (Y_F - Y_R)/Y_R \times 100 \quad (3)$$

where $\Delta Y$ represents the maize yield changes under the climate scenarios of SSP1-2.6, SSP2-4.5 and SSP5-8.5 in 2030 s and 2050 s. $Y_F$ is the average yield in 2030 s or 2050 s, $Y_R$ is the average yield in reference period. We calculated $\Delta Y$ at each 0.25°×0.25° grid.

The daily data of the climate variables (Tmax, Tmin, Tas, Srad, and Pr) from NEX-GDDP-CMIP6 were validated for the period of 1979–2014 by comparing them with observed data from CMFD. The Mean Bias (MB, Eq. (5)) were used for validation.

$$MB = N^{-1} \sum_{i=1}^{N} (O_i - S_i) \quad (4)$$

where $N$ is number of simulations/observations, $S_i$ is the simulated values of five GCMs. Additionally, we calculated the Pearson's correlation coefficient (r) between average simulated and observed values of 1979–2014 period in all land pixels to evaluate the moder performance in capturing the spatial variations in climate variables.

### 3.5. Identifying the key determinants of yield variability

We first compared the abilities of a suit of machine learning models (lasso; ridge; recursive partitioning, rpart; gradient boosting machine, gbm; extreme gradient boosting, xgbTree; random forest, rf) in reproducing the yield variability to determine the best model. In this step, we initiated the optimization of hyperparameters for each model using the random search method within the *caret* R package. Following this, we employed the *train()* function from the same package to perform 10-fold cross-validation with data resampling to evaluate the model's performance in simulating yield variability through the coefficient of determination ($R^2$), the root mean square error (RMSE), and the mean absolute error (MAE), which were calculated as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(M_i - Pi)^2}{\sum_{i=1}^{n}(M_i - Pmean)^2} \quad (6)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(M_i - P_i)^2}{n}} \quad (7)$$

$$MAE = \frac{\sum_{i=1}^{n}|M_i - P_i|}{n} \quad (8)$$

where $M_i$ represents the APSIM simulated maize yield, $P_i$ represents the predicted maize yield by machine learning, and $P_{mean}$ represents the average predicted maize yield by machine learning. By doing this, the random forest was evaluated as the best model to explain yield variability.

To compare the relative importance of mean and extreme climate, we first trained random forest on a subset of climate indicators representing only mean or extreme climate (Table 3), respectively. Then we calculated the difference in explained variance ($R^2$) between mean climate and extreme climate of the random forest models. The procedures described above were repeated 10 times, and the average difference in explained variance was utilized to assess the relative significance of mean and extreme climate on yield variation.

To evaluate the key climate indicators affecting yield variability, we initially employed the *rfPermute* R package to measure variable importance. A significant advantage of *rfPermute* is its ability to provide us standardized importance for each climate indicator along with its corresponding significance level. The metirc named *increase in MSE* (%) was used to demonstrate the importance of climate indicators, that is, the more influential a climate indicator is, the higher MSE it can generate. Subsequently, we employed accumulated local effect (ALE) plots to compare the effect sizes of climate indicators on yield variation. These plots quantitatively illustrate their influence on yield variation. ALE plots offer several advantages in quantifying effect sizes: (1) ALE plots are unbiased even in the presence of a correlation feature space; (2) ALE plots exhibit faster computing speed; (3) ALE plots directly show the relative impacts of feature changes on yield variation; (4) ALE plots effectively identify the relationship between features and predictors, disregarding outliers. The *iml* R package was used to draw the ALE plots for the key determinants.

## 4. Results

### 4.1. Performance of climate model outputs and climate change characteristics

The performance of climate model outputs was evaluated by analyzing the MB and r values of climate variables of precipitation (Pr), shortwave solar radiation (Srad), maximum temperature (Tmax), minimum temperature (Tmin), and mean temperature (Tas), considering the monthly values (Fig. 4A, B, C, D, E), the spatial distribution of MB (Fig. 4F, G, H, I, J), and r between simulated and observed data for all land pixels (Fig. 4K, L, M, N, O). Results indicate that the five GCMs

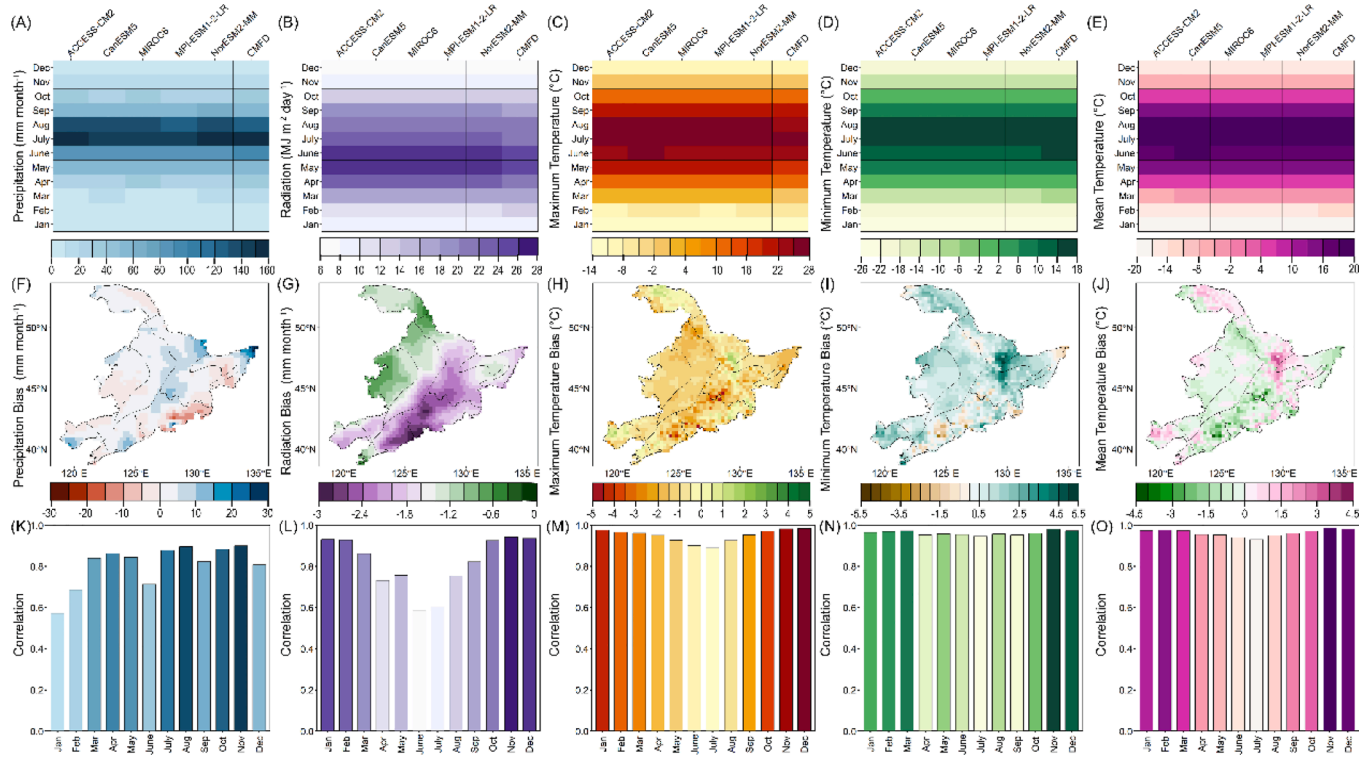



**Fig. 4. Top panels**: monthly accumulated precipitation (mm month-1), monthly average shortwave radiation (MJ/m$^{-2}$(−|-) day-1), monthly air temperature (maximum, minimum, and average temperature, °C) simulated by five General Circulation Models (GCMs), and the observed data for the period of 1979–2014. **Middle panels**: spatial distribution of statistical bias values between the multi-model ensemble mean and observed data. **Bottom panles**: correlation between simulated mean values and observations for all land pixels. **A, F and K**: precipitation; **B, G and L**: shortwave solar radiation; **C, H and M**: maximum temperature; **D, I and N**: minimum temperature; **E, J and O**: average temperature. Note that the correlation is the Pearson's correlation of simulated and observed values of all land pixels, and the bias was calculated as observations minus simulations.

exhibit similar accuracy in capturing monthly variations in climate variables, yet they tend to underestimate Pr in June, Tmin in March and June, and Tas in February. Conversely, these models overestimate Srad from April to July, Tmax in May and August. This overestimation of Tmax and underestimation of Tmin results in low bias in Tas. At regional level, climate projections obviously overestimate Pr and Tmax, but underestimate Srad and Tmin. But in most regions, the MB values fall within an acceptable range. The average MB of the region shows maximum monthly absolute MB for Pr, Srad, Tmax, Tmin, and Tas at 7.0 mm month$^{-1}$, 2.3 MJ m$^{-2}$ day$^{-1}$, 1.7 °C, 1.7 °C, and 1.2 °C. MB values for temperature are lower than the acceptable threshold of 2 °C (Flato et al., 2013). Underestimation of Srad was also found in Brazil, with MB for Srad falling below the acceptable threshold of 4 MJ m$^{-2}$ day$^{-1}$ (Dias et al., 2024). Monthly r values between simulated data and observed data for all land pixels are mostly greater than 0.8, indicating the strong performance in capturing spatial variations in climate variables within the region.

The average cliamte conditions characterized by climate varibles of Pr, Srad, Tmax, Tmin, Tas are listed in Table 4. Climate change will incareae all the climate variables. Specifically, slightly increase of Pr (4.5 ~ 8.7 %) and relatively higher increase of Srad (12.0 %~14.5 %) are projected. Notably, the increase of Srad is possibly overestimated according to the comparison between simulated data and observed data (Fig. 4G). The increase of mean air temperature is higher than 1.4 °C, indicating more heat resources can be used for maize production in future. But the increase of Tmax also stress the higher risk of heat-related climate extremes.

**Table 4**
Regional multi-model ensemble mean of cliamte variables.

| Scenarios | Periods | Climate variables | | | | |
|---|---|---|---|---|---|---|
| | | Pr | Srad | Tmax | Tmin | Tas |
| Observed | reference | 606.1 | 4842.4 | 9.5 | −0.2 | 4.4 |
| SSP1-2.6 | 2030 s | 634.5 | 5486.5 | 11.7 | 0.4 | 6.0 |
| | 2050 s | 659.1 | 5544.2 | 12.3 | 1.0 | 6.6 |
| SSP2-4.5 | 2030 s | 633.1 | 5424.0 | 11.5 | 0.04 | 5.8 |
| | 2050 s | 655.5 | 5468.5 | 12.5 | 1.1 | 6.8 |
| SSP5-8.5 | 2030 s | 641.4 | 5454.0 | 11.8 | 0.25 | 6.0 |
| | 2050 s | 658.3 | 5465.3 | 13.1 | 1.6 | 7.4 |

Note that the reference period is from 1986 to 2005, 2030 s is from 2020 to 2039, and 2030 s is from 2040 to 2059. Abbreviations: Pr, precipitation, mm year$^{-1}$; Srad, shortwave solar radiation, MJ m$^{-2}$ year$^{-1}$; Tmax, maximum tepmerature, °C; Tmin, minimum temperature, °C; Tas, mean temperature, °C.

### 4.2. Impacts of climate change on yield

Climate change will exert negative impacts on maize yield in all agro-ecological regions without adaption such as irrigation, genetic improvemet (Fig. 5). On average, climate change will decrease maize yield of the whole region about 13.1 %~27.5 % across climate scenarios. Our results suggest moderate yield decrease under SSP1-2.6 (13.1 %~14.7 %), but severe decrease under SSP2-4.5 and SSP5-8.5 scenarios (25.7 % ~27.5 %), indicating the negative impact will increase with the warming level. The magnitude of climate change impacts on mazie yield show obviously spatial heterogenecity. The average yield losses of all scenarios in the agro-ecological regions range from 16.2 % to 26.6 %. Region III will be the most vulnerable to climate change, with the yield decrease of 13.6 %~34.1 % (the range describes the differences among climate scenarios). However, Regions IV and VI show higher resistance than other agro-ecological regions to climate change with the yield decrease of 7.5 %~21.0 % and 10.6 %~21.8 %, respectively. There is a obviously pattern that climate change impact on maize yield increase

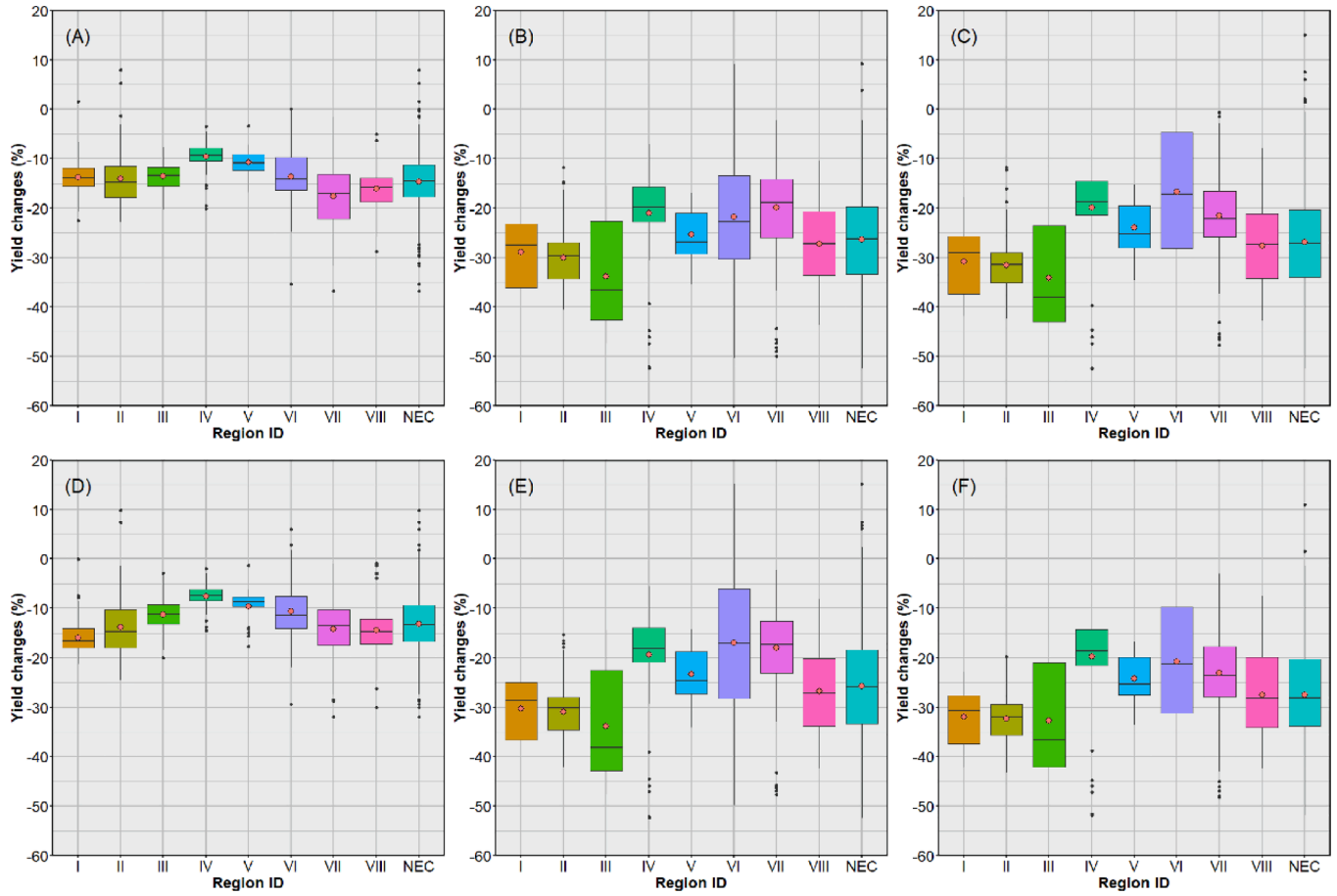



**Fig. 5.** Projected yield changes of eight agro-ecological regions in Northeast Chian under climate change. **A** SSP1-2.6 in 2030 s (2020–2039). **B** SSP2-4.5 in 2030 s. **C** SSP5-8.5 in 2030 s. **D** SSP1-2.6 in 2050 (2040–2059). **E** SSP2-4.5 in 2050 s. **F** SSP5-8.5 in 2050 s. The red points in the boxes denote the mean values. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

from south to north, and decrease from east to west.

### 4.3. Importance of mean climate and extreme climate to yield variability

The ability of six machine learning models in explaining maize yield variability was evaluated using $R^2$, RMSE and MAE (Fig. 6). The two regression-based methods, lasso and ridge, demonstrate comparable performance in simulating maize yield, with both methods achieving a mean $R^2$ of 0.54. Additionally, the Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) are nearly equal for both methods. The single tree-based method of rpart outperforms the regression-based methods, with a mean $R^2$ of 0.63, a mean RMSE of 1580.9 kg/ha, and a mean MAE of 1142.7 kg/ha. The multi-tree-based methods, gbm and xgbTree, outperform both the regression-based and single tree-based

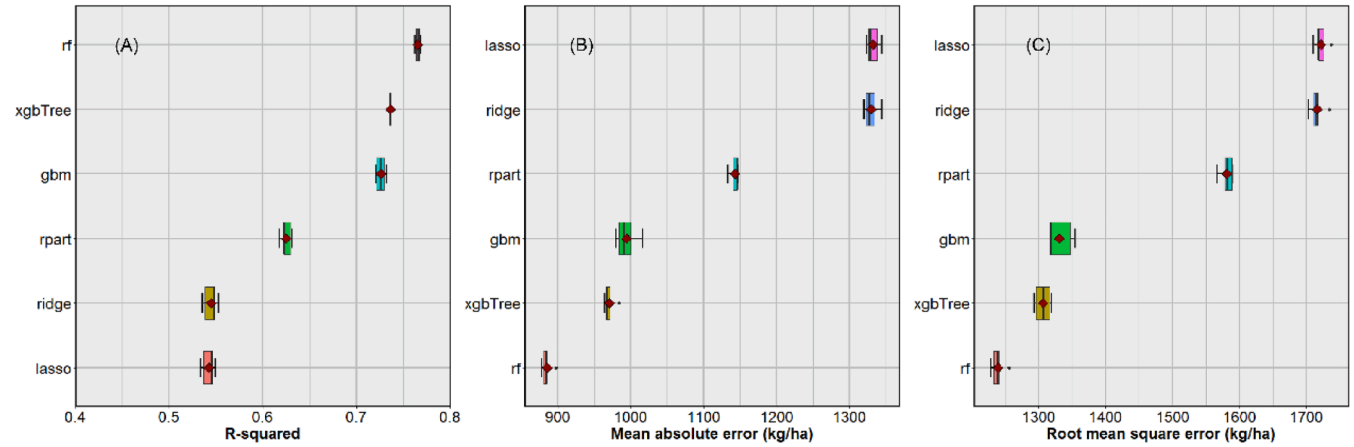



**Fig. 6.** Comparisons of six machine learning models in explaining maize yield variability in Northeast China using 10-fold cross-validation, considering (A) coefficient of determination ($R^2$), (B) root mean square error (RMSE), and (C) mean absolute error (MAE). The upper and lower whiskers represent the maximum and minimum R squared values, respectivaly. The red diamonds indicate the average R squared value. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

methods, achieving a mean $R^2$ of 0.73, a mean RMSE of 1318.6 kg/ha, and a mean MAE of 982.3 kg/ha across the two methods. Finally, the multi-tree-based method of rf exhibits the highest performance in explaining maize yield, with a mean $R^2$ of 0.77, a mean RMSE of 1239.2 kg/ha, a mean MAE of 885.1 kg/ha. Consequently, the rf method was utilized to assess the primary determinants of maize yield variability.

Mean climate will exert higher impact on maize yield under low warming scenario like SSP1-2.6. However, under higher warming scenario of SSP5-8.5, extreme climate will significantly influence maize yield, as its explained variance on yield variability is higher than that of mean climate in more than half of the agro-ecological regions across climate scenarios (Fig. 7). This effect is particularly evident in Regions I, II, III, and IV. The impacts of extreme and mean climate on yield variability show substantial spatial heterogeneity across agro-ecological regions. Moreover, the importance of extreme and mean climate in influencing yield variability is subject to changes with varying climate scenarios.

### 4.4. Variables importance and effect sizes on yield variability

The mean and extreme climate indicators (Table 3) all have significant effects ($p < 0.01$) on yield variability. Notably, significant spatial heterogeneity exists in the importance of these variables on yield variability across the agro-ecological regions (Fig.S2~Fig.S9). Fig. 8 illustrates the variable importance on maize yield variability across six climate scenarios. Among the mean climate indicators, CPR emerges as the most crucial factor governing yield variability across all climate scenarios. CSR and GDD are other important mean climate indicators significantly influencing yield variability. Drought-related extreme climate indicators play significant role in regulating yield variability. Furthermore, besides CDH, other extreme climate indicators such as CDC and TDR also play significant roles in regulating yield variability.

The effect sizes of the most important mean and extreme climate indicators, i.e., CPR and CDH on maize yield variability were compared using the accumulated local effects (ALE) plots (Fig. 9~Fig. 10). Overall, CDH has higher effect size on yield variability than CPR. On average, the effect sizes of CDH and CPR are 5596 and 1898 kg/ha, respectively. Specifically, the increase 28 CDH days can lead to average yield loss of 1363 kg/ha. On contrary, the decrease 41 CDH days can improve yield by 4233 kg/ha. As for CPR, sharply increase in yield variability is projected with the changes of CPR from −200 mm to 164 mm. Additional water supply of approximately 164 mm will bring about yield improvement of 1044 kg/ha. However, CPR decrease 199 mm can lead to yield loss of 558 kg/ha.

## 5. Discussions

### 5.1. Reliability of yield simulations

Predicting regional-scale crop yield using crop models is challenging due to limited observations required to calibrate the model parameters (Liu et al., 2019). Several upscaling strategies have been reported in previous literatures, including using one representative cultivar in an agro-ecological region (Luo et al., 2021), coupling crop model with machine learning techniques (Zhang et al., 2021), assimilating remote sensing data to calibrate parameters (Zhuo et al., 2022), and employing

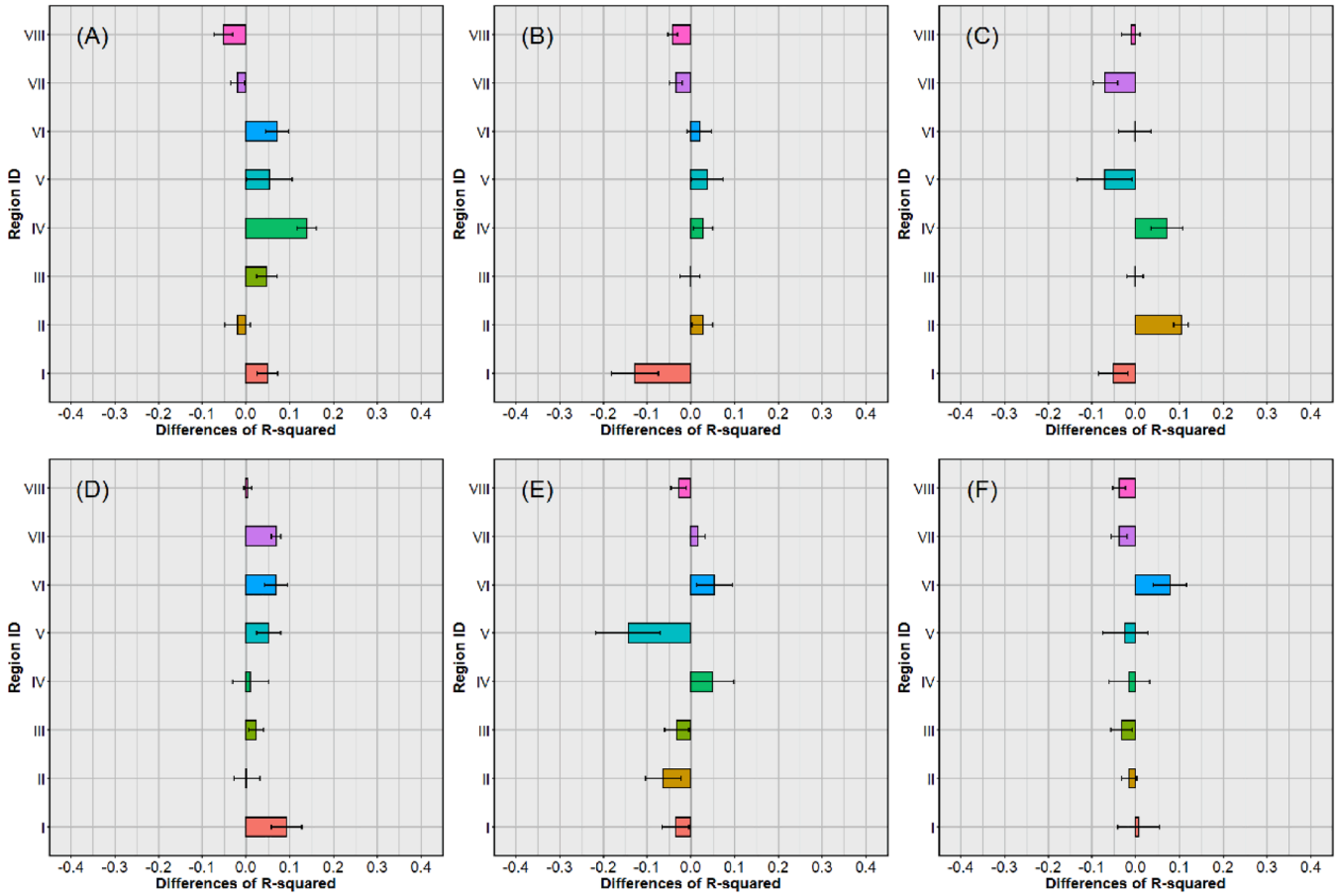



**Fig. 7.** Differences in explained variance ($R^2$) between mean and extreme climate on maize yield in Northeast China. **A** SSP1-2.6 in 2030 s (2020–2039). **B** SSP2-4.5 in 2030 s. **C** SSP5-8.5 in 2030 s. **D** SSP1-2.6 in 2050 (2040–2059). **E** SSP2-4.5 in 2050 s. **F** SSP5-8.5 in 2050 s. The negative $R^2$ values indicate that extreme climate has a greater influence on yield variability than mean climate, while positve $R^2$ values suggest the opposite.

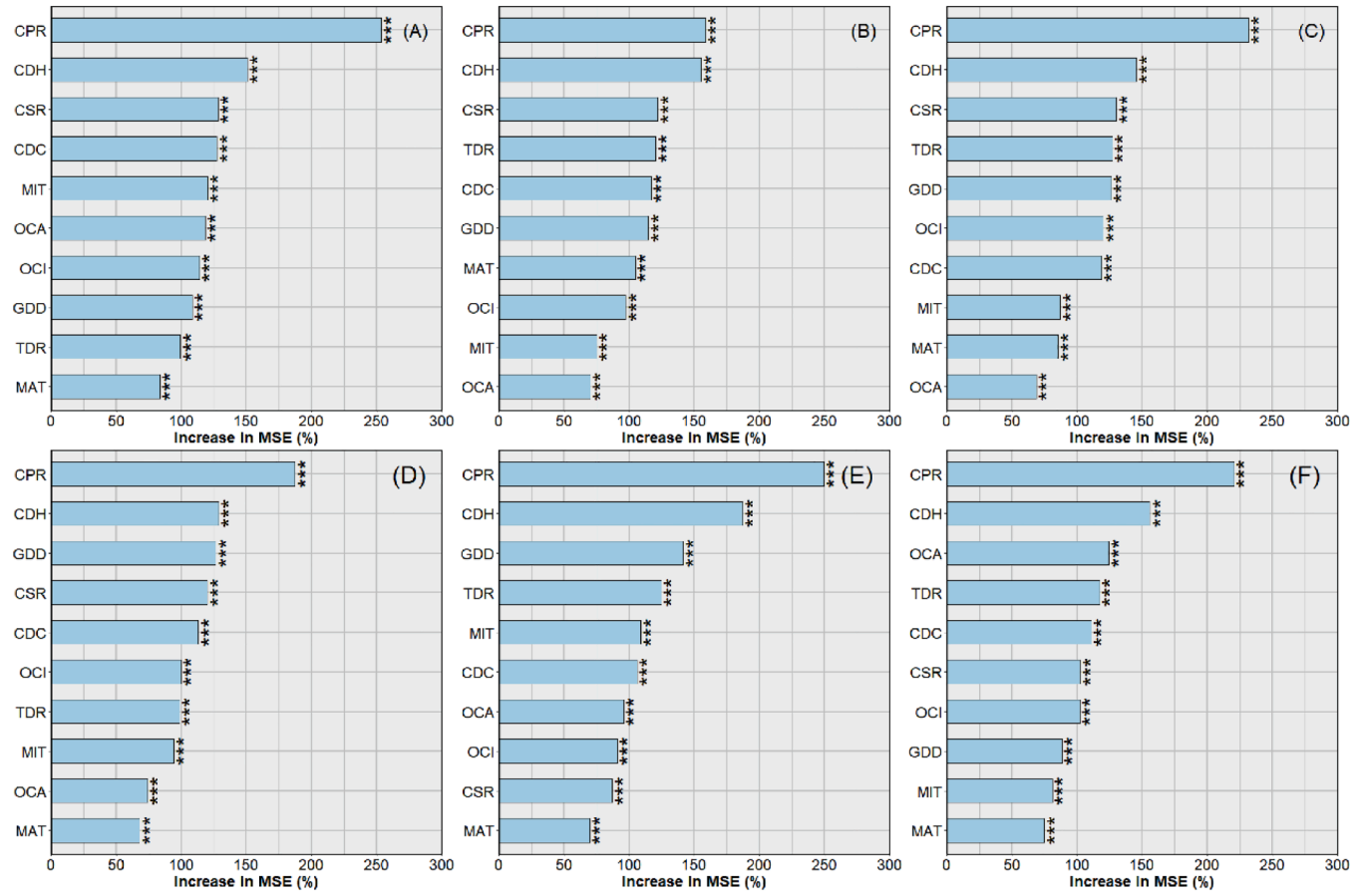



**Fig. 8.** Variables importance of mean and extreme climate indicators on yield variability in Northeast China. **CPR** Cumulative precipitation. **CSR** cumulative shortwave radiation. **CDH** total compound drought and hot days. **TDR** Total drought days. **GDD** Growing degree days. **CDC** Compound drought and cold days. **OCA** Frequency of daily maximum temperature above 30 °C. **OCI** Frequency of daily minimum temperature below 8 °C. **MAT** Average daily maximum temperature. **MIT** Average daily minimum temperature. Detailed information on the abbreviations can be found in Table 2. **A** SSP1-2.6 in 2030 s (2020–2039). **B** SSP2-4.5 in 2030 s. **C** SSP5-8.5 in 2030 s. **D** SSP1-2.6 in 2050 (2040–2059). **E** SSP2-4.5 in 2050 s. **F** SSP5-8.5 in 2050 s. *** denotes $p < 0.0$.

the average of virtual cultivars generated from cultivars planted in an agro-ecological region (Chen et al., 2021). In this study, the variety of eight typical maize varieties in NEC that reproduced mature dates from a remote sensing dataset developed by Luo et al. (2020b) best was determined as the originally planted variety in each grid. We used yield records from 44 agro-meteorological sites to validate the model accuracy. Results show strong correlation between simulations and observations with $R^2$ and RMSE are 0.65 and 1011.1 kg/ha, respectively (Fig. 3). Furthermore, we used city-level yield statistics to validate the model accuracy at regional scale. The validation revealed a mean $R^2$ value of 0.82 and a mean Root Mean Square Error (RMSE) of 529.5 kg/ha. The precision our study is similar to previous studies that predicting regional-scale maize yield. For instance, Zhang et al. (2021) combined machine learning with a crop model to simulate regional-scale maize yield in China and achieved an $R^2$ of 0.56 when comparing the simulations with records from agro-meteorological sites. Similarly, Cheng et al. (2022) utilized random forest to simulate wheat and maize yield at high resolution, obtaining an $R^2$ of 0.51 for wheat and an $R^2$ of 0.65 for maize when comparing the simulations with in-situ observations. Notably, precision of simulating crop yield at regional-scale is generally lower than that at site-scale. Huang et al. (2020) obtained $R^2$ values ranging from 0.74 to 0.98 when simulating maize yield after calibrating the model using in-situ observations.

### 5.2. *Extreme climate explain yield variability more than mean climate under higher warming levels*

The impact of climate change climate extremes on crop yield has gained increased attention in the last decades (Feng et al., 2019b). We find that under SSP1-2.6 and SSP2-4.5 scenarios in 2030 s, yield variability in the majority of agro-ecological regions will be predominantly controlled by mean climate. The pattern is expected to change under SSP2-4.5 scenario in 2050 s and SSP5-8.5 scenario in both 2030 s and 205 s, with extreme climate indicators explaining yield variability more than mean climate indicators. The findings suggest a stronger impact of extreme climate on crop yields as temperature rise (Fig. 7). A previous study also found that climate extremes exert significantly negative impacts on soybean yield in NEC (Guo et al., 2022). The reasons for the increasing importance of climate extremes on yield variability are twofold: (1) the frequency of climate extremes will increase with rising temperatures; (2) crop yield responds more significantly to climate extremes compared to mean climate. For instance, Shi et al. (2021) highlighted that crops grown in NEC were more susceptible to the impacts of droughts. Our results demonstrate the importance of climate extremes on yield variability, and indicate this importance will increase with rising temperatures. The local maize community should prioritize efforts to mitigate the damage caused by climate extremes on maize yield.

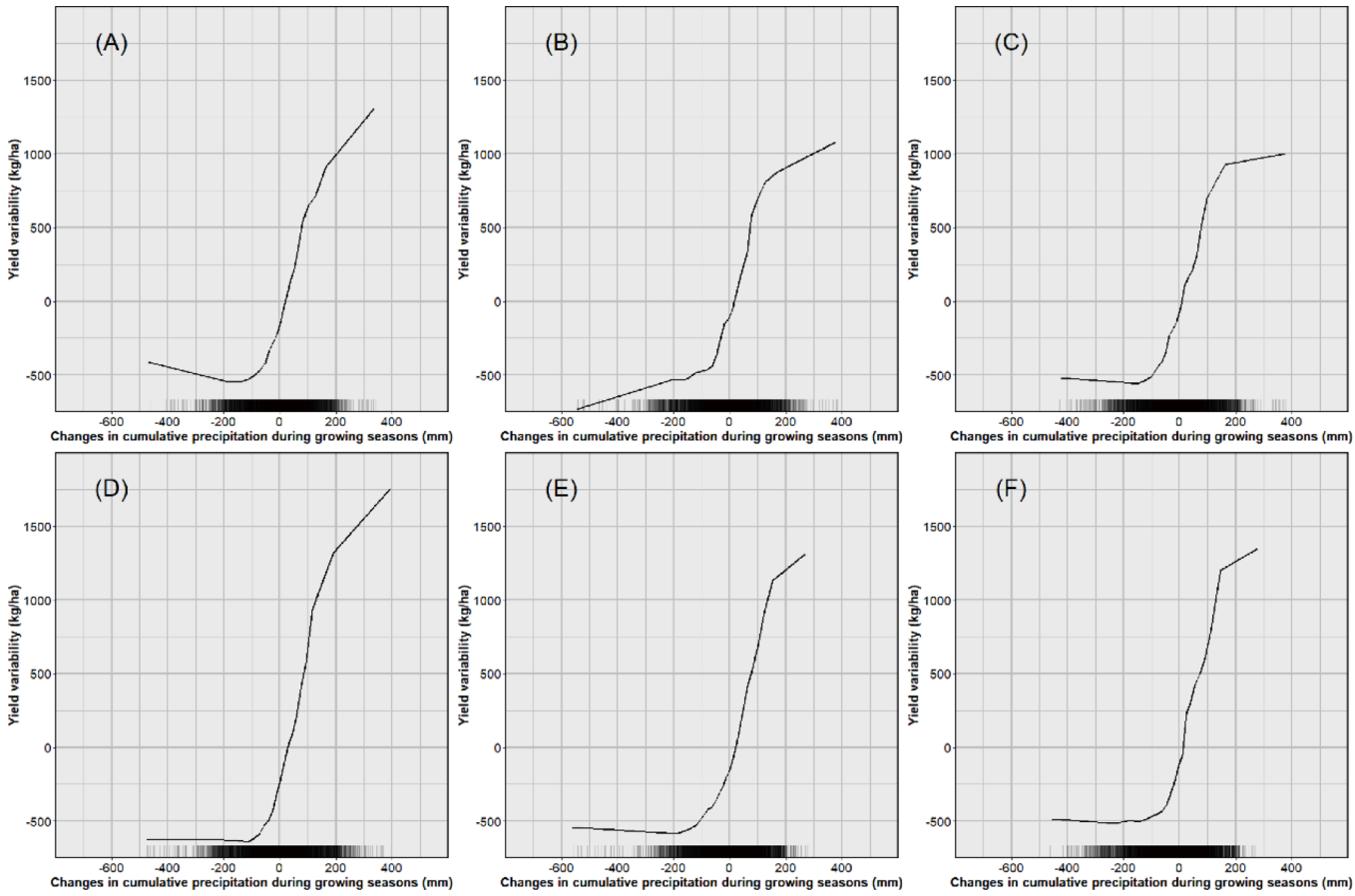



**Fig. 9.** Effect size of changes in cumulative precipitation during growing season explaining yield variability in Northeast China. **A** SSP1-2.6 in 2030 s (2020–2039). **B** SSP2-4.5 in 2030 s. **C** SSP5-8.5 in 2030 s. **D** SSP1-2.6 in 2050 (2040–2059). **E** SSP2-4.5 in 2050 s. **F** SSP5-8.5 in 2050 s.

### *5.3. Compound drought and hot days during growing seasons play important role in governing yield variability*

Maize yield projected to decrease in the future without adaption (Fig. 5), which aligns with previous studies that focused on the impact of climate change on crop yield (Feng et al., 2019a; Stevens et al., 2016). On average, maize yield in NEC is projected to decrease by 13.1 % to 27.5 % under future cliamte. The yield decrease can probably be attributed to (1): climate warming shortens maize growing priod, particularly from flowering to maturing, which is a critical phase for achieing high yield (Asseng et al., 2015); (2): the increasing frequency of climate extremes, such as compound hot and drought days, affecting the formation of organic matter. Our results demonsrate that extreme climate significantly impacts yield variability in NEC (Fig. 6 and Fig. 7), consistent with studies by Guo et al. (2022) and Chen et al. (2020).

Notably, the most crucial extreme climate indicator affecting yield variability is CDH (Fig. 8), followed by TDR. As the frequency of CDH stresses is incrreasing in observations and future climate projections (Li et al., 2023; Lu et al., 2018), the impact of CDH stresses on maize yield will increase in the future. For example, Wu et al. (2023) investigate the impacts of different types of cliamte extremes on maize yield in China, and found that the negative impacts of agriculture compound drought and hot events on maize yield increased during the past decades at both natianl and regional scales. Feng et al., (2019b) found that compound drought and hot events caused highest risk on global maize yield reduction than individual drought/hot events during 1961–2014.

The compound drought and hot days during growing seasons play a significant role in governing the yield variability (Fig. 7), with its effect size on yield variability being the highest (Fig. 10). For example, an increase of 28 additional compound drought and hot days during maize growing seasons, compared to reference period, can result in yield loss of 1363 kg/ha. Indeed, various studies have demenstrated the negative effects of cliamte extremes on yield stability (Leng et al., 2019; Li et al., 2019; Lesk et al., 2016), and recently, there has been growing attention towards investigating the impacts of compound climate extremes on crop.

### *5.4. Implications and uncertainties*

Currently, maize production in NEC relies primarily on rainfall. However, providing an additional 164 mm of water can significantly increase yield by approximately 1044 kg/ha (Fig.9). Given the potential yield benefit of providing moderate water to maize, it is critical to carefully plan and prioritize irrigation, particular in regions where rainfall is decreasing. Our findings suggest that, under higher warming levels, extreme climate will play a more significant role than mean climate in determining yield variability. Moreover, we have highlighted the crucial role of CDH in maize yield, for instance, an increase of 28 CDH days could lead to yield loss of 1363 kg/ha (Fig. 10). This emphasizes the need to breed resilient cultivars that can withstand drought and heat in order to adapt to climate change. Furthermore, the increase in air temperature enables the potential for planting varieties with longer growth duration to achieve higher yield, particular in mountainous and high latitude regions.

In this study, we focused on the impact of climate change on maize yield, as well as the key drivers of yield variability in future, based on yield simulations from crop model rather than historical observations. To enhance the reliability of our projections, we validated our findings by comparing them with historical observations at both site and regional scales (Fig. 3). However, uncertainties persist due to climate projections

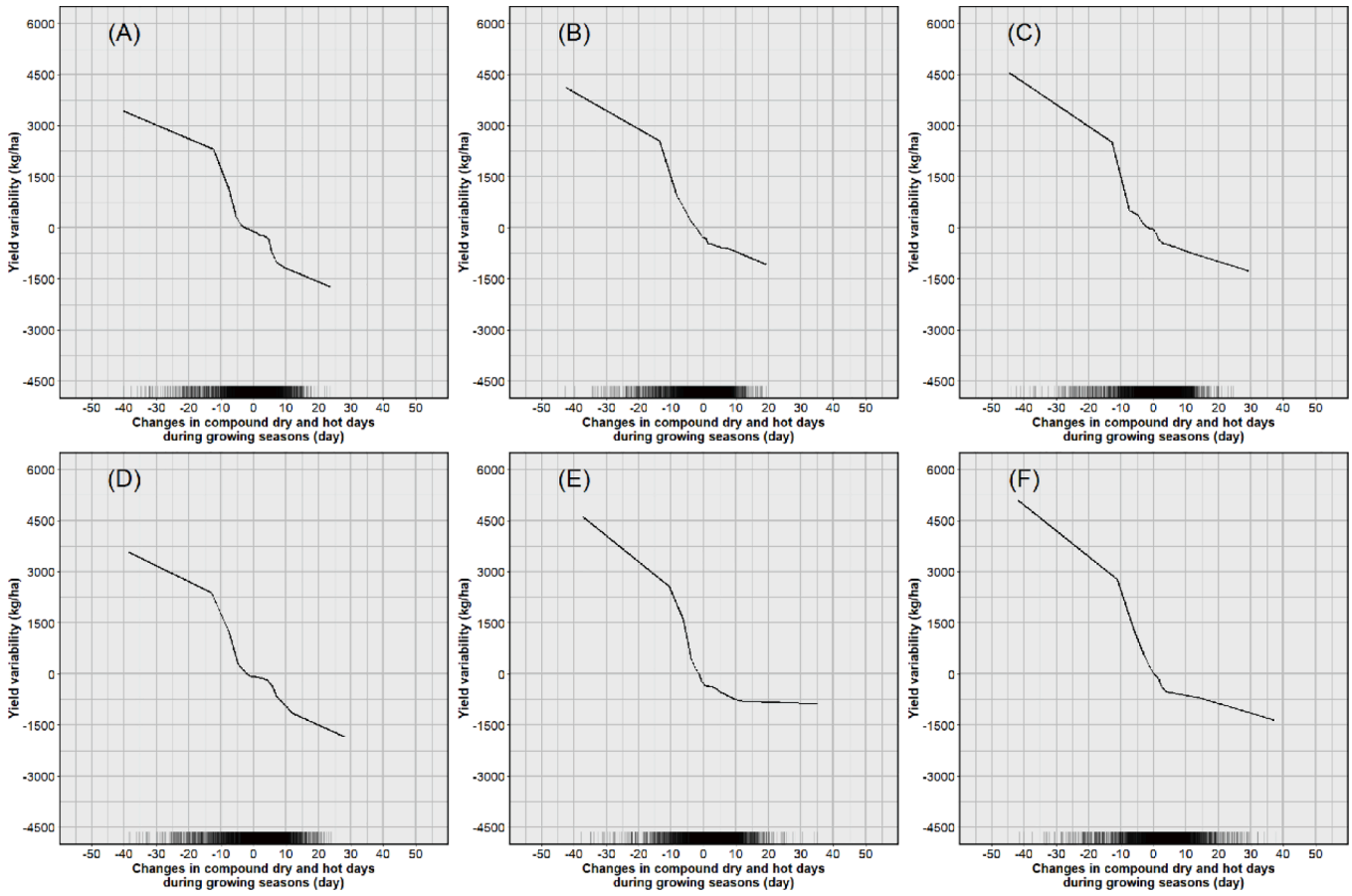



**Fig. 10.** Effect size of changes in compound drought and hot days during growing season explaining yield variability in Northeast China. **A** SSP1-2.6 in 2030 s (2020–2039). **B** SSP2-4.5 in 2030 s. **C** SSP5-8.5 in 2030 s. **D** SSP1-2.6 in 2050 (2040–2059). **E** SSP2-4.5 in 2050 s. **F** SSP5-8.5 in 2050 s.

(Dokoohaki et al., 2021; Mueller et al., 2021). For example, the overestimation of Tmax in climate projections may lead to an overestimation of the effect size of CDH on yield variability. Moreover, the overestimation of Srad can result in overestimation of future yield. Additionally, uncertainties can be attributed to the structures of crop models (Dokoohaki et al., 2021; Wang et al., 2017), as crop models demonstrate limited ability in accounting for factors like pests, weeds, and climate extremes such as drought, excessive rainfall, and heatwaves, which impact crop yield (Shahhosseini et al., 2021). Using an ensemble of crop models can help quantify uncertainties arising from differences in model structures (Li et al., 2015). Nevertheless, even with model ensemble mean, complete reduction of uncertainties may not achievable. To improve model performance, a possible approach involves optimizing the relationship between environmental factors with crop growth (Asseng et al., 2019). The response of crops to climate is closely related to their growth parameters, but the eight typical listed in Table 2 may not fully represent the possible responses of maize. Therefore, we propose using extensive field observations to refine the outputs of crop models to improve their accuracy in predicting regional-scale crop yield.

## 6. Conclusions

The dataset generated by APSIM-Maize, which characeterizes maize production at a 0.25 degrees horizontal resolution under three Socio-economic Pathways- SSP1-2.6, SSP2-4.5, and SSP5-8.5- in Northeast China (NEC) was emplyed to investigate the impacts of cliamte change and the determinants of yield variability. The major conclusions are as follows:

(1) Climte change is projected to negatively affect maize yield in NEC, with a decline ranging from 13.1 % to 27.5 % (considering differences in climate scenarios) compared to the reference period of 1986–2005.
(2) Random forest demonstrated superior performance among the regression- or tree-based machine learning methods used for predicting maize yield, with a mean $R^2$ of 0.77, a mean RMSE of 1239.2 kg/ha, a mean MAE of 885.1 kg/ha.
(3) Under lower warming levels, the mean climate exhibits a stronger ability to explain yield variability. However, this pattern changes adversely with rising temperatures.
(4) CPR and CDH were identified as the most important mean and extreme climate indicators, repectively, influencing yield variability. Moreover, CDH has the highest effect size on yield variability, making it a key determinant of maize yield in the future. Other climate indicators such as CSR and TDR also play important role in regulating yield variability.

## CRediT authorship contribution statement

**Chuanwei Zhang:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft. **Jiangbo Gao:** Funding acquisition, Project administration, Writing – review & editing. **Lulu Liu:** Data curation, Software, Writing – review & editing. **Shaohong Wu:** Supervision, Writing – review & editing.

## Declaration of competing interest

The authors declare that they have no known competing financial

interests or personal relationships that could have appeared to influence the work reported in this paper.

## Data availability

Data will be made available on request.

## Acknowledgement

This work was financially supported by the Strategic Priority Research Program of the Chinese Academy of Sciences (Grant No. XDA28130104).

## Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.compag.2024.108688.

## References

Asseng, S., Martre P., Maiorano A., Roetter R. P., O'Leary G. J., Fitzgerald G. J., Girousse C., Motzo R., Giunta F., Babar M. A., Reynolds M. P., Kheir A. M. S., Thorburn P. J., Waha K., Ruane A. C., Aggarwal P. K., Ahmed M., Balkovic J., Basso B., Biernath C., Bindi M., Cammarano D., Challinor A. J., De Sanctis G., Dumont B., Rezaei E. E., Fereres E., Ferrise R., Garcia-Vila M., Gayler S., Gao Y., Horan H., Hoogenboom G., Izaurralde R. C., Jabloun M., Jones C. D., Kassie B. T., Kersebaum K.-C., Klein C., Koehler A.-K., Liu B., Minoli S., San Martin M. M., Mueller C., Kumar S. N., Nendel C., Olesen J. E., Palosuo T., Porter J. R., Priesack E., Ripoche D., Semenov M. A., Stockle C., Stratonovitch P., Streck T., Supit I., Tao F., Van der Velde M., Wallach D., Wang E., Webber H., Wolf J., Xiao L., Zhang Z., Zhao Z., Zhu Y. & Ewert F., 2019. Climate change impact and adaptation for wheat protein. *Global Change Biology,* 25, 155-173.

Asseng, S., Ewert, F., Martre, P., Roetter, R.P., Lobell, D.B., Cammarano, D., Kimball, B. A., Ottman, M.J., Wall, G.W., White, J.W., Reynolds, M.P., Alderman, P.D., Prasad, P.V.V., Aggarwal, P.K., Anothai, J., Basso, B., Biernath, C., Challinor, A.J., De Sanctis, G., Doltra, J., Fereres, E., Garcia-Vile, M., Gayler, S., Hoogenboom, G., Hunt, L.A., Izaurralde, R.C., Jabloun, M., Jones, C.D., Kersebaum, K.C., Koehler, A. K., Mueller, C., Kumar, S.N., Nendel, C., O'Leary, G., Olesen, J.E., Palosuo, T., Priesack, E., Rezaei, E.E., Ruane, A.C., Semenov, M.A., Shcherbak, I., Stoeckle, C., Stratonovitch, P., Streck, T., Supit, I., Tao, F., Thorburn, P.J., Waha, K., Wang, E., Wallach, D., Wolf, I., Zhao, Z., Zhu, Y., 2015. Rising temperatures reduce global wheat production. Nat. Clim. Chang. 5, 143–147.

Ceglar, A., Toreti A., Lecerf R., Van der Velde M. & Dentener F. Impact of meteorological drivers on regional inter-annual crop yield variability in France. *Agricultural and Forest Meteorology,* 216, 58-67.

Chen, S., He, L., Cao, Y., Wang, R., Wu, L., Wang, Z., Zou, Y., Siddique, K.H.M., Xiong, W., Liu, M., Feng, H., Yu, Q., Wang, X., He, J., 2021. Comparisons among four different upscaling strategies for cultivar genetic parameters in rainfed spring wheat phenology simulations with the DSSAT-CERES-Wheat model. Agric Water Manag 258, 107181.

Chen, X., Wang, L., Niu, Z., Zhang, M., Li, C.a., Li, J., 2020. The effects of projected climate change and extreme climate on maize and rice in the Yangtze River Basin China. Agric. Forest Meteorol. 282, 107867.

Cheng, M., Jiao, X., Shi, L., Penuelas, J., Kumar, L., Nie, C., Wu, T., Liu, K., Wu, W., Jin, X., 2022. High-resolution crop yield and water productivity dataset generated using random forest and remote sensing. Sci. Data 9, 641.

Cohen, I., Zandalinas, S.I., Huck, C., Fritschi, F.B., Mittler, R., 2021. Meta-analysis of drought and heat stress combination impact on crop yield and yield components. Physiol. Plant. 171, 66–76.

Dias, C.G., Martins, F.B., Martins, M.A., 2024. Climate risks and vulnerabilities of the Arabica coffee in Brazil under current and future climates considering new CMIP6 models. In: Science of the Total Environment, p. 167753.

Dokoohaki, H., Kivi, M.S., Martinez-Feria, R., Miguez, F.E., Hoogenboom, G., 2021. A comprehensive uncertainty quantification of large-scale process-based crop modeling frameworks. Environ. Res. Lett. 16, 084010.

Feng, S., Hao, Z., Zhang, X., Hao, F., 2019b. Probabilistic evaluation of the impact of compound drought-hot events on global maize yields. Sci. Total Environ. 689, 1228–1234.

Feng, P., Wang, B., Liu, D.L., Waters, C., Yu, Q., 2019a. Incorporating machine learning with biophysical model can improve the evaluation of climate extremes impacts on wheat yield in south-eastern Australia. Agric. For. Meteorol. 275, 100–113.

Feng, P., Wang, B., Liu, D.L., Waters, C., Xiao, D., Shi, L., Yu, Q., 2020. Dynamic wheat yield forecasts are improved by a hybrid approach using a biophysical model and machine learning technique. Agric. For. Meteorol. 285, 107922.

Flato, G., Marotzke, J., Abiodun, B., Braconnot, P., Chou, S.C., Collins, W., Cox, P., Driouech, F., Emori, S., Eyring, V., Forest, C., Gleckler, P., Guilyardi, E., Jakob, C., Kattsov, V., Reason, C., Rummukainen, M., 2013. Evaluation of climate models. In: Stocker, T.F., Qin, D., Plattner, G.-K., et al. (Eds.), Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, pp. 741–866.

Guo, S., Guo, E., Zhang, Z., Dong, M., Wang, X., Fu, Z., Guan, K., Zhang, W., Zhang, W., Zhao, J., Liu, Z., Zhao, C., Yang, X., 2022. Impacts of mean climate and extreme climate indices on soybean yield and yield components in Northeast China. Sci. Total Environ. 838, 156284.

Huang, M., Wang, J., Wang, B., Liu, D.L., Yu, Q., He, D., Wang, N., Pan, X., 2020. Optimizing sowing window and cultivar choice can boost China's maize yield under 1.5 degrees C and 2 degrees C global warming. Environ. Res. Lett. 15, 024015.

Laux, P., Jaecket, G., Tingem, R.M., Kunstmann, H., 2010. Impact of climate change on agricultural productivity under rainfed conditions in Cameroon-A method to improve attainable crop yields by planting date adaptations. Agric. For. Meteorol. 150, 1258–1271.

Leng, G., Hall, J., 2019. Crop yield sensitivity of global major agricultural countries to droughts and the projected changes in the future. Sci. Total Environ. 654, 811–821.

Leng, G., Huang, M., 2017. Crop yield response to climate change varies with crop spatial distribution pattern. Sci. Rep. 7, 1463.

Lesk, C., Rowhani, P., Ramankutty, N., 2016. Influence of extreme weather disasters on global crop production. Nature 529, 84–87.

Li, Y., Guan, K., Schnitkey, G.D., DeLucia, E., Peng, B., 2019. Excessive rainfall leads to maize yield loss of a comparable magnitude to extreme drought in the United States. Glob. Chang. Biol. 25, 2325–2337.

Li, T., Hasegawa, T., Yin, X., Zhu, Y., Boote, K., Adam, M., Bregaglio, S., Buis, S., Confalonieri, R., Fumoto, T., Gaydon, D., Marcaida III, M., Nakagawa, H., Oriol, P., Ruane, A.C., Ruget, F., Singh, B., Singh, U., Tang, L., Tao, F., Wilkens, P., Yoshida, H., Zhang, Z., Bouman, B., 2015. Uncertainties in predicting rice yield by current crop models under a wide range of climatic conditions. Glob. Chang. Biol. 21, 1328–1341.

Li, L., Li, X., Zheng, X., Li, X., Jiang, T., Ju, H., Wan, X., 2022. The effects of declining soil moisture levels on suitable maize cultivation areas in Northeast China. J. Hydrol. 608, 127636.

Li, B., Liu, Z., Huang, F., Yang, X., Liu, Z., Wan, W., Wang, J., Xu, Y., Li, Z., Ren, T., 2021. Ensuring National Food Security by Strengthening High-productivity Black Soil Granary in Northeast China. Bull. Chin. Acad. Sci. 36, 1184–1193.

Li, C., Wang, R., Ning, H., Luo, Q., 2016. Changes in climate extremes and their impact on wheat yield in Tianshan Mountains region, northwest China. Environ. Earth Sci. 75, 1228.

Li, H., Yao, F., Zhang, J., Hao, C., 2014. Analysis on Climatic Maize Yield and Its Sensitivity to Climate Change in Northeast China. Chin. J. Agrometeorol. 35, 423–428.

Li, E., Zhao, J., Zhang, W., Yang, X., 2023. Spatial-temporal patterns of high-temperature and drought during the maize growing season under current and future climate changes in northeast China. J. Sci. Food Agric. 103, 5709–5716.

Liu, Y., Wu, S., Zheng, D., Dai, E., 2009. Soil indicators for eco-geographic regionalization: A case study in mid-temperate zone of eastern China. J. Geog. Sci. 19, 200–212.

Liu, Y., Chen, Q., Tan, Q., 2019. Responses of wheat yields and water use efficiency to climate change and nitrogen fertilization in the North China plain. Food Security 11, 1231–1242.

Lobell, D.B., Hammer, G.L., Chenu, K., Zheng, B., McLean, G., Chapman, S.C., 2015. The shifting influence of drought and heat stress for crops in northeast Australia. Glob. Chang. Biol. 21, 4115–4127.

Lu, Y., Hu, H., Li, C., Tian, F., 2018. Increasing compound events of extreme hot and drought days during growing seasons of wheat and maize in China. Sci. Rep. 8, 16700.

Luo, Y., Zhang, Z., Li, Z., Chen, Y., Zhang, L., Cao, J., Tao, F., 2020a. Identifying the spatiotemporal changes of annual harvesting areas for three staple crops in China by integrating multi-data sources. Environ. Res. Lett. 15, 074003.

Luo, Y., Zhang, Z., Chen, Y., Li, Z., Tao, F., 2020b. ChinaCropPhen1km: a high-resolution crop phenological dataset for three staple crops in China during 2000–2015 based on leaf area index (LAI) products. Earth Syst. Sci. Data 12, 197–214.

Luo, Y., Zhang, Z., Zhang, L., Cao, J., 2021. Spatiotemporal patterns of winter wheat phenology and its climatic drivers based on an improved pDSSAT model. Sci. China-Earth Sci. 64, 2144–2160.

Meroni, M., Waldner, F., Seguini, L., Kerdiles, H., Rembold, F., 2021. Yield forecasting with machine learning and small data: What gains for grains? Agric. For. Meteorol. 308, 108555.

Mueller, C., Franke, J., Jaegermeyr, J., Ruane, A.C., Elliott, J., Moyer, E., Heinke, J., Falloon, P.D., Folberth, C., Francois, L., Hank, T., Izaurralde, R.C., Jacquemin, I., Liu, W., Olin, S., Pugh, T.A.M., Williams, K., Zabel, F., 2021. Exploring uncertainties in global crop yield projections in a large ensemble of crop models and CMIP5 and CMIP6 climate scenarios. Environ. Res. Lett. 16, 034040.

Qiu, J., Wang, L., Tang, H., Li, H., Li, C., 2004. Study on the Situation of Soil Organic Carbon Storage in Arable Lands in Northeast China. Sci. Agric. Sin. 37, 1166–1171.

Shahhosseini, M., Hu, G., Huber, I., Archontoulis, S.V., 2021. Coupling machine learning and crop modeling improves crop yield prediction in the US Corn Belt. Sci. Rep. 11, 1606.

Shangguan, W., Dai, Y., Liu, B., Zhu, A., Duan, Q., Wu, L., Ji, D., Ye, A., Yuan, H., Zhang, Q., Chen, D., Chen, M., Chu, J., Dou, Y., Guo, J., Li, H., Li, J., Liang, L., Liang, X., Liu, H., Liu, S., Miao, C., Zhang, Y., 2013. A China data set of soil properties for land surface modeling. J. Adv. Model. Earth Syst. 5, 212–224.

Shi, W., Wang, M., Liu, Y., 2021. Crop yield and production responses to climate disasters in China. Sci. Total Environ. 74, 569–583.

Stevens, T., Madani, K., 2016. Future climate impacts on maize farming and food security in Malawi. Sci. Rep. 6, 36241.

Thrasher, B., Wang, W., Michaelis, A., Melton, F., Lee, T., Nemani, R., 2022. NASA Global Daily Downscaled Projections, CMIP6. Sci. Data 9, 262.
Van der Wiel, K., Bintanja, R., 2021. Contribution of climatic changes in mean and variability to monthly temperature and precipitation extremes. Communic. Earth Environm. 2, 1.
Wang, Y., Huang, J., Wang, J., Findlay, C., 2018b. Mitigating rice production risks from drought through improving irrigation infrastructure and management in China. Aust. J. Agric. Resour. Econ. 62, 161–176.
Wang, E., Martre, P., Zhao, Z., Ewert, F., Maiorano, A., Roetter, R.P., Kimball, B.A., Ottman, M.J., Wall, G.W., White, J.W., Reynolds, M.P., Alderman, P.D., Aggarwal, P. K., Anothai, J., Basso, B., Biernath, C., Cammarano, D., Challinor, A.J., De Sanctis, G., Doltra, J., Fereres, E., Garcia-Vila, M., Gayler, S., Hoogenboom, G., Hunt, L.A., Izaurralde, R.C., Jabloun, M., Jones, C.D., Kersebaum, K.C., Koehler, A.-K., Liu, L., Mueller, C., Kumar, S.N., Nendel, C., O'Leary, G., Olesen, J.E., Palosuo, T., Priesack, E., Rezaei, E.E., Ripoche, D., Ruane, A.C., Semenov, M.A., Shcherbak, I., Stockle, C., Stratonovitch, P., Streck, T., Supit, I., Tao, F., Thorburn, P., Waha, K., Wallach, D., Wang, Z., Wolf, J., Zhu, Y., Asseng, S., 2017. The uncertainty of crop yield projections is reduced by improved temperature response functions. Nat. Plants 3, 17102.
Wang, N., Wang, E., Wang, J., Zhang, J., Zheng, B., Huang, Y., Tan, M., 2018a. Modelling maize phenology, biomass growth and yield under contrasting temperature conditions. Agric. For. Meteorol. 250, 319–329.
Wood, A.W., Maurer, E.P., Kumar, A., Lettenmaier, D.P., 2002. Long-range experimental hydrologic forecasting for the eastern United States. J. Geophys. Res.-Atmos. 107, 4429.
Wood, A.W., Leung, L.R., Sridhar, V., Lettenmaier, D.P., 2004. Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs. Clim. Change 62, 189–216.
Wu, X., Jiang, D., Zhang, F., 2023. Increasing impact of compound agricultural drought and hot events on maize yield in China. Climate Res. 90, 17–29.
Xiao, D., Liu, D.L., Wang, B., Feng, P., Waters, C., 2020. Designing high-yielding maize ideotypes to adapt changing climate in the North China Plain. Agr. Syst. 181, 102805.
Xiao, D., Liu, D.L., Feng, P., Wang, B., Waters, C., Shen, Y., Qi, Y., Bai, H., Tang, J., 2021. Future climate change impacts on grain yield and groundwater use under different cropping systems in the North China Plain. Agric. Water Manag. 246, 106685.
Xie, Y., Lin, H., Ye, Y., Ren, X., 2019. Changes in soil erosion in cropland in northeastern China over the past 300 years. Catena 176, 410–418.
Yang, K., 2018. China Meteorological Forcing Data (1979–2018). Big Data System for Pan-Third Pole. https://doi.org/10.11888/AtmosphericPhysics.tpe.249369.file.
Yang, X., Ming, B., Tao, H., Wang, P., 2015. Spatial distribution characteristics and impact on spring maize yield of drought in Northeast China. Chin. J. Eco-Agric. 23, 758–767.
Zhang, X., Duan, Y., Duan, J., Jian, D., Ma, Z., 2022. A daily drought index based on evapotranspiration and its application in regional drought analyses. Sci. China-Earth Sci. 65, 317–336.
Zhang, Z., Li, Z., Chen, Y., Zhang, L., Tao, F., 2020. Improving regional wheat yields estimations by multi-step-assimilating of a crop model with multi-source data. Agric. For. Meteorol. 290, 107993.
Zhang, L., Zhang, Z., Tao, F., Luo, Y., Cao, J., Li, Z., Xie, R., Li, S., 2021. Planning maize hybrids adaptation to future climate change by integrating crop modelling with machine learning. Environ. Res. Lett. 16, 124043.
Zhao, C., Liu B., Piao S., Wang X., Lobell D. B., Huang Y., Huang M., Yao Y., Bassu S., Ciais P., Durand J.-L., Elliott J., Ewert F., Janssens I. A., Li T., Lin E., Liu Q., Martre P., Mueller C., Peng S., Penuelas J., Ruane A. C., Wallach D., Wang T., Wu D., Liu Z., Zhu Y., Zhu Z. & Asseng S., 2017. Temperature increase reduces global yields of major crops in four independent estimates. *Proceedings of the National Academy of Sciences of the United States of America,* 114**,** 9326-9331.
Zhuo, W., Huang, J., Xiao, X., Huang, H., Bajgain, R., Wu, X., Gao, X., Wang, J., Li, X., Wagle, P., 2022. Assimilating remote sensing-based VPM GPP into the WOFOST model for improving regional winter wheat yield estimation. Eur. J. Agron. 139, 126556.