

Refined crop modelling reveals intensifying impacts of droughts and heatwaves on global wheat productivity

In the format provided by the
authors and unedited

Contents

Text S1: Comparisons of model skills in simulating leaf development.

Text S2: Comparisons of model skills in capturing yield variations with the GGCM Phase 3 using the same climate inputs.

Text S3: Comparisons of model skills in capturing yield variations and composite yield impacts from droughts, heatwaves and CHDE with the GGCM Phase 3 using other independent data for calibration.

Text S4: Comparisons of model skills in capturing yield variations and composite yield impacts from droughts, heatwaves and CHDE with the GGCM Phase 3 using another calibration strategy.

Text S5: Investigation of other extreme weather event identify approach.

Text S6: Frequency and intensity of droughts, heatwaves and CHDE.

Text S7: Comparisons of model skills in capturing composite yield impacts from droughts, heatwaves and CHDE with the GGCM Phase 3 using the same climate inputs.

Text S8: Introduction of CERES-Wheat model.

Text S9: Wheat yield data.

Text S10: Comparisons of model skills in capturing yield variations with the GGCM Phase 1.

Text S11: Comparisons of model skills in capturing composite yield impacts from droughts, heatwaves and CHDE with GGCM Phase 1.

Supplementary Figures

Fig.S1: Comparison of the simulated national and subnational yields with observed ones by different model versions of GGCM Phase 1.

Fig.S2: Out-of-sample validation of our different model versions.

Fig.S3: The RRMSE of maturity date (a) and yield (b) for each grid simulated by GWM-NCS.

Fig.S4: Spatial patterns of PARUE (a) and PARU2 (b) from GGCM-NCS-ECO.

Fig.S5: Comparison of simulated and reported yields by GWM-NCS-ECO-WE.

Fig.S6: The RRMSE of maturity date (a) and yield (b) for each grid simulated by GWM-NCS-ECO-WE.

Fig.S7: Spatial patterns of average R^2 between simulated LAI by GWM-NCS-ECO (a-c) or GWM-NCS-ECO-WE (d-f) and satellite-retrieved values during 2001-

2015 in the North China Plain.

Fig.S8: Spatial patterns of average R^2 between simulated LAI by GWM-NCS-ECO (a-c) or GWM-NCS-ECO-WE (d-f) and satellite-retrieved values during 2001-2015 in the United States of America.

Fig.S9: Comparison of simulated and reported yields by Global CERES-Wheat-NCS-ECO (GWM-NCS-ECO) at 0.5-degree resolution.

Fig.S10: Spatial distribution of simulation errors (RRMSE) of maturity date (a) and yield (b) for GWM-NCS-ECO calibrated with GlobalCropYield5min at 0.5-degree resolution.

Fig.S11: Taylor diagram comparing model performances at subnational (a) and national (b) scales.

Fig.S12: Spatial distribution of simulation errors (RRMSE) of maturity date (a) and yield (b) for GWM-NCS-ECO calibrated with SPAM at 0.5-degree resolution.

Fig.S13: Locations of agro-meteorological stations in China (a) and comparison of simulations and observations for anthesis date (ADAP, b), maturity date (MDAP, c), and yield (d) by GWM-NCS-ECO for validation.

Fig.S14: Comparison of simulated and reported yields by GWM-NCS-ECO-WE at 0.1° resolution calibrated by statistical data at national (a) and subnational (b) scales.

Fig.S15: Taylor diagram of simulated yields by GGCMI Phase 3 models and GWM-NCS-ECO-WE calibrated by statistical data at subnational (a) and national (b) scales.

Fig.S16: RRMSE of (a) maturity date and (b) yield for GWM-NCS-ECO-WE calibrated by statistical data at 0.1° resolution.

Fig.S17: Comparison of simulated and reported yields by GWM-NCS-ECO-WE at 0.1° resolution calibrated under irrigated/rainfed conditions (from MIRCA data) at national (a) and subnational (b) scales.

Fig.S18: Taylor diagram of simulated yields by GGCMI Phase 3 models and GWM-NCS-ECO-WE calibrated under irrigated or rainfed conditions (from MIRCA data) at subnational (a) and national (b) scales.

Fig.S19: RRMSE of (a) maturity date and (b) yield for GWM-NCS-ECO-WE calibrated under actual irrigated/rainfed conditions at 0.1° resolution.

Fig.S20: Changes in frequency (a, b), intensity (c, d), and duration (e, f) of

heatwaves (left) and CHDE (right) between 1981–1998 and 1999–2015.

Fig.S21: Statistically significant yield changes (1981–2015) for (a) heatwaves and (b) CHDEs.

Fig.S22: Influences of droughts (a), heatwave (b) and CHDE (c) on rainfed and irrigated wheat yields.

Fig.S23: Spatial patterns of rainfed and irrigated harvested area for spring (c, d, e) and winter wheat (a, b, f).

Fig.S24: Yield changes due to droughts, heatwaves and CHDE for spring and winter wheat.

Fig.S25: Relationships between wheat yield change and $-SPI$ (a), STI (b), and $-SCEI$ (c).

Fig.S26: Changes in wheat yield due to droughts, heatwaves and CHDE for different growth periods.

Fig.S27: Observed and simulated impacts of droughts on global wheat yields from GGCM Phase 3 and ours.

Fig.S28: Observed and simulated impacts of heatwaves on global wheat yields from GGCM Phase 3 and ours.

Fig.S29: Observed and simulated impacts of CHDE on global wheat yields from GGCM Phase 3 and ours.

Fig.S30: Observed and simulated impacts of 2003 European drought on wheat yields from GGCM Phase 3 and ours.

Fig.S31: Observed and simulated impacts of 2003 European heatwave on wheat yields from GGCM Phase 3 and ours.

Fig.S32: Observed and simulated impacts of droughts on global wheat yields from GGCM Phase 1 and ours.

Fig.S33: Observed and simulated impacts of heatwaves on global wheat yields from GGCM Phase 1 and ours.

Fig.S34: Observed and simulated impacts of CHDE on global wheat yields from GGCM Phase 1 and ours.

Fig.S35: Observed and simulated impacts of 2003 European drought on wheat yield from GGCM Phase 1 and ours.

Fig.S36: Observed and simulated impacts of 2003 European heatwave on wheat

yield from GGCM Phase 1 and ours.

Fig.S37: Global observed and simulated impacts of droughts, heatwaves, CHDEs, 2003 European drought and heatwave on wheat.

Fig.S38: Observed and simulated yield losses from droughts, heatwaves, CHDEs, 2003 European drought and heatwave at 0.5° resolution.

Fig.S39: Comparison of drought impact estimation errors across spatial resolutions in GWM-NCS-ECO-WE.

Fig.S40: Yield changes due to droughts (a), heatwaves (b) and CHDE (c) from GWM-NCS-ECO at 0.5° resolution (GlobalCropYield5min calibration).

Fig.S41: Yield changes due to droughts (a), heatwaves (b) and CHDE (c) from GWM-NCS-ECO at 0.5° resolution (SPAM calibration).

Fig.S42: Yield changes due to droughts (a), heatwaves (b) and CHDE (c) from GWM-NCS-ECO at 0.1° resolution (GlobalWheatYield4km calibration).

Fig.S43: Extreme weather event impacts simulated by GWM-NCS-ECO-WE calibrated under actual irrigated/rainfed conditions at 0.1° resolution against FAO observations.

Fig.S44: Statistically significant yield changes (1981–2015) for (a) droughts, (b) heatwaves, and (c) CHDEs at 0.1° resolution simulated by the GWM-NCS-ECO-WE under actual irrigated or rainfed conditions.

Fig.S45: Extreme weather event impacts simulated by GWM-NCS-ECO-WE calibrated to statistical data at 0.1° resolution against FAO observations.

Fig.S46: Changes in impacts of droughts, heatwaves and CHDE.

Fig.S47: Global patterns of frequency and intensity of extreme weather events during the study period.

Fig.S48: Global distribution of wheat cultivation areas.

Fig.S49: Accuracy of yield estimations using random forest model in 13 countries.

Fig.S50: Schematic of the New Calibration Scheme.

Fig.S51: Comparison of wheat development rate in response to temperature between the original CERES-Wheat and Wang-Engel function for (a) vegetative growth, (b) reproductive growth, (c) vernalization, and (d) photosynthesis.

Fig.S52: Extreme event impacts from raw FAO statistics.

Supplementary Tables

Table S1: Three categories of key parameters in the CERES-Wheat model.

Table S2: Accuracy of wheat yield simulations from different model versions of GGCM Phase 3 and ours.

Table S3: Accuracy of wheat yield simulations from different model versions of GGCM Phase 1 and ours.

Table S4: Summarization for information on subnational statistics.

Table S5: The details of temporal configuration of FAO data.

Table S6: Full statistical details for the agreement between statistical and simulated yields by three crop models.

Table S7: Full statistical details for the trend in the affected area caused by three disasters and the irrigation contribution.

Text S1: Comparisons of model skills in simulating leaf development.

To evaluate the impact of improved temperature response functions on other simulation outputs such as leaf area index (LAI), we conducted additional comparisons between simulations from GWM-NCS-ECO-WE and GWM-NCS-ECO with three mainstream remote sensing-based LAI products, including the Global Land Surface Satellite (GLASS) LAI, the GLOBMAP LAI, the Global Inventory Monitoring and Modeling Studies (GIMMS) LAI4g after consolidation with the reprocessed MODIS LAI. These LAI products showed good spatiotemporal consistency with observations (Liu et al., 2012; Cao et al., 2023; Xiao et al., 2016). We did not compare our improved model versions with GGCM Phase 3 outputs as it did not provide LAI simulation results.

We selected two main wheat producing area, that is, the North China Plain (NCP) in China and the United States of America (USA) for further analyses. More specifically, for the GLASS LAI and GLOBMAP datasets, we paired the simulated LAI time series from GWM-NCS-ECO and GWM-NCS-ECO-WE during the growing season with satellite-retrieved LAI values at 8-day intervals. For the GIMMS LAI4g dataset, we calculated the maximum LAI values for the first and second half of each month from the simulated LAI time series and paired them with the corresponding half-month composites of satellite-retrieved values. Finally, we computed the average R^2 between the simulated and satellite-based LAI time series from 2001 to 2015 to assess the extent to which the model captures the variability in satellite-observed LAI.

The results confirmed that the improvement in temperature response function can contribute to enhance model skill of simulating leaf development. Our GWM-NCS-ECO-WE framework achieved superior performance (average R^2 ranged from 0.46–0.52 and 0.48–0.54 for NCP and USA respectively; see [Fig.S7-S8](#)) compared to GWM-NCS-ECO (average R^2 : 0.37–0.41 and 0.4–0.47), with temporal consistency improved by 27% and 18%, respectively.

Text S2: Comparisons of model skills in capturing yield variations with the GGCM Phase 3 using the same climate inputs

We first chose the GSWP3-W5E5 (0.5°) meteorological data from 1981 to 2015 to drive crop model simulation, which was identical to the inputs in GGCM Phase 3 historical simulations. Meanwhile, we tested whether using different yield data affect the improvement in model skill by adopting our parameterization scheme. On the one hand, we adopted a global yield dataset at 5 arc-minutes resolution during 1982-2015 (i.e., GlobalCropYield5min) to calibrate crop model parameters, which was generated by machine learning approaches based on sub-national statistics covering ~12,000 administrative units (Cao et al., 2025). GlobalCropYield5min had a high accuracy with R^2 ranging from 0.7 to 0.95, demonstrating superior performance than other well-known global yield datasets. On the other hand, we used SPAM yield data for all production systems to optimize model parameters, which was identical to the inputs in pDSSAT model of GGCM Phase 3, while the soil data differs from that used in GGCM Phase 3 ensembles. Most of crop models in GGCM Phase 3 used soil input derived from the Harmonized World Soil Database (HWSD), whereas pDSSAT model utilized the Global Soil Data set for Earth system modelling (GSDE) and DSSAT-Pythia adopted the Global High-Resolution Soil Profile Database for Crop Modelling Applications (GHSPD) due to difficulties in retrieving all soil parameters from HWSD (Jägermeyr et al., 2021). We used the GHSPD dataset as same as DSSAT-Pythia model in GGCM Phase 3, because it provides soil files formatted for DSSAT usage with wide applications in global change studies. Furthermore, some critical soil properties required by DSSAT model such as the soil albedo, runoff potential and drainage rate cannot obtain from GSDE and HWSD. In addition, we used a global crop type maps at 0.05° resolution to separate the spatial distribution of winter and spring wheat (Becker-Reshef et al., 2023). Other data such as fertilizer application, crop calendar, as well as calibration settings (e.g., under fully irrigated conditions, the fertilization date and amount), are consistent with those of the main text.

Even at the coarser 0.5° resolution with harmonized climate data, our GWM-NCS-

ECO framework calibrated with GlobalCropYield5min achieved superior performance (National $R^2=0.68$, CRMSD=1.23 t/ha; see [Fig.S9 a-b](#)) compared to most GGCM Phase 3 models. The RRMSE for maturity date and yield were 5.8% and 25.1% respectively ([Fig.S10](#)), comparable to our high-resolution results. When calibrated with the same SPAM data as pDSSAT model in GGCM Phase 3, our framework still reduced simulation errors by 12.4% compared to the pDSSAT model ([Fig.S9 c-d](#), [Fig.S11](#)), with RRMSE values of 5.1% and 21.9% ([Fig.S12](#)). In summary, improvements in the model's ability to capture historical yield changes when leveraging the GWM-NCS-ECO framework were not significant at coarse spatial scales (0.5°) compared with high spatial scales (0.1°).

Text S3: Comparisons of model skills in capturing yield variations and composite yield impacts from droughts, heatwaves and CHDE with the GGCM Phase 3 using other independent data for calibration.

We also examine the performances of high-resolution calibration framework by using independent observational data from field experiments and high-resolution census statistics (Table S4). The following analyses are independent supplementary validation experiments and are separate from the primary calibration and evaluation framework described in the main text.

On the one hand, wheat experimental data (1981–2018) are obtained from 137 agro-meteorological stations possessing at least six years of records governed by the Chinese Meteorological Administration (CMA), covering phenology, cultivar characteristics, management, and yields. Data are recorded in situ by agricultural technicians and quality-assured through the national monitoring system (Fig.S13-a). The associated meteorological stations provided daily weather data, with sunshine hours converted to solar radiation by using the Angstrom-Prescott equation (<http://data.cma.cn>). Soil data were collected from the Global High-Resolution Soil profile database at 5 arc-minute resolution (GHSPD). Calibration and validation were conducted for GWM-NCS-ECO model with NSGA-II, using a 4:2 data split from stations records.

The validation results showed that GWM-NCS-ECO performed well in reproducing observed yield time series at field scale, with spatial patterns of simulated phenology and yields better correlated with the observations (R^2 of 0.99, 0.98 and 0.78 for days to anthesis date [ADAP], days to maturity date [MDAP], and yields, Fig.S13 b-d) and with lower biases (RMSE of 6.1, 6.7, and 865.7 kg/ha; RRMSE of 3.9%, 3.6%, and 15.6%). Therefore, the high-resolution calibration scheme is proved to be a trustworthy approach for crop model simulation at large scale.

On the other hand, we combined the national- and subnational-scale yield statistics to calibrate GWM-NCS-ECO-WE model at 0.1-degree resolution (hereinafter GWM-NCS-ECO-WE-STA). Subnational data were prioritized and assigned to all grids within

the relevant administrative unit. For grids lacking such data in some years or in all the years, national statistics were used. Automatic parameter calibration was conducted for parallel computations during 2001-2010 based on the NSGA-II algorithm on high performance super-cloud computer systems. The 2001–2010 calibration period was selected for both data availability and computational considerations. Because most countries in our compiled database have relatively complete subnational yield statistics during 2001–2010 ([Table S4](#)), allowing us to maximize the use of statistical data at the finest available administrative level; missing country–year records were supplemented with national FAO statistics. Moreover, the supplementary census-based calibration was conducted at the 0.1° grid scale, a shorter calibration window substantially improved computational efficiency. Other data such as soil properties, fertilizer application, crop calendar, as well as calibration settings (e.g., under fully irrigated conditions, the fertilization date and amount), are consistent with those of this study.

This model, calibrated independently from any ML-based yield product, still outperformed the GGCM Phase 3 ensemble (national $R^2 = 0.77$, CRMSD=1.02 t/ha; [Figs.S14-S15](#)). The RRMSE values for days to maturity date (MDAP) and yields are averagely 9.4% and 38.9%, respectively, which were relatively large especially in highly heterogeneous areas with only coarse statistics such as Russian Federation (large amount of grids at administrative unit level 1) ([Fig.S16](#)). While it captured extreme event impacts less accurately than our primary model (which benefits from the spatiotemporal detail of GlobalWheatYield4km), it still demonstrated a credible ability to simulate yield losses ([Fig.S45](#)).

Text S4: Comparisons of model skills in capturing yield variations and composite yield impacts from droughts, heatwaves and CHDE with the GGCM Phase 3 using another calibration strategy.

We investigated whether optimizing model parameters at actual irrigated or rainfed conditions can affect model skill of capturing historical yield variations and reproducing disaster-induced yield losses. More specifically, we first determined the cropping system of a specific grid cell according to the differences in irrigated and rainfed harvested areas from MIRCA data. Then we calibrated GWM-NCS-ECO-WE model version at 0.1-degree resolution under corresponding irrigated or rainfed conditions (hereinafter GWM-NCS-ECO-WE-IR/RF), with other settings kept consistent.

The version calibrated under actual conditions (GWM-NCS-ECO-WE-IR/RF) showed improved correlation with observed yields (national $R^2 = 0.7$) but introduced larger biases (CRMSD of 2.24 t/ha) and higher relative errors (RRMSE of 45.7% for yield) compared to our primary model versions calibrated under irrigated conditions ([Figs.S17-19](#)).

The model calibrated under actual conditions consistently underestimated the observed yield losses from extreme events. It explained less than half of the losses recorded in the EM-DAT database for droughts (2.9% vs. observed ~6.5%), heatwaves (1.9% vs. ~4%), and CHDE (4.3% vs. ~10.6%), and similarly underestimated the impacts of the 2003 European heatwave and drought ([Figs.S43-44](#)).

Text S5: Investigation of other extreme weather event identify approach.

To further investigate whether using STI can capture the spatiotemporal characteristics of heatwaves, we also used the other approach to identify extreme weather events and compared the results. More specially, we first calculated 1-month SPI for each grid during 1981-2015 and identified drought events as $SPI < -0.8$. Then, a heatwave event was defined as a spell of at least three consecutive days with daily maximum temperature exceeding the heat thresholds, which were determined as 30 °C and 34 °C for winter and spring wheat (Lobell et al., 2012; Anderson et al., 2023). In addition, two successive heatwave events were merged into a single event once they were separated by less than two days. We use heating degree days (HDD, °C·day) to assess the severity of heat stress for each growing season. CHDEs were identified when a heatwave coincided with a monthly drought event during wheat growing season.

$$HDD = \sum_{i=1}^n (T_{max,i} - T_{high}), T_{max,i} > T_{high} \quad (1)$$

where $T_{max,i}$ denotes daily maximum temperature of day i , T_{high} represents the heat thresholds, and n is the total number of days in a wheat growing season.

The severity for a CHDE (CHDE_s) was given as (Yin et al., 2023):

$$CHDE_s = \sum_{i=1}^{i=CHDE_d} [(-1 \times SPI_m) \times \frac{T_{max,i} - T_{high}}{T_{high} - T_{base}}] \quad (2)$$

where CHDE_d was the duration of a CHDE event, SPI_m was the SPI value for day i at corresponding month m , T_{base} was the base temperature representing the minimum biology temperature for wheat, here we use 5.5°C according to existing studies (Zhu and Troy, 2018; He et al., 2024).

The frequency, intensity and duration of heatwaves (CHDEs) during wheat growth period increased from 1981–1998 to 1999–2015 in 73.4% (65.5%), 76.1% (62.2%) and 77.2% (64.4%) of the grids, respectively (Fig.S20). The frequency and duration of heatwaves between two periods increased most on average by 0.26 yr⁻¹ and 3.5 d yr⁻¹, approximately two times the rate of CHDE (0.11 yr⁻¹ and 1.4 d yr⁻¹), respectively. Heatwave and CHDE intensified significantly at the rate of 17.2 °C·d yr⁻¹ and 0.26 yr⁻¹. These conclusions were consistent with those from analyses using STI.

In addition, the grid-scale EWE impact assessment results showed that global

average losses of wheat yield caused by heatwaves, and CHDE during 1981-2015 were 5.6% and 11.6%, respectively, which were comparable with those assessed by STI (5.8% and 13.3%) ([Fig.S21](#)). Both methods consistently identified CHDE as the most damaging event type and highlighted mid- and high-latitude regions as vulnerability hotspots.

Text S6: Frequency and intensity of droughts, heatwaves and CHDE.

The frequency of droughts, heatwaves and CHDE during wheat growth period increased from 1981–1998 to 1999–2015 in 59.3%, 92.2% and 78.9% of the grids (Fig.S47 a-c). The frequency of heatwaves between two periods increased most on average by 0.2 yr^{-1} , followed by CHDE (0.12 yr^{-1}) and droughts (0.02 yr^{-1}). Heatwaves and CHDE occurred more frequently in recent period across continents, while the frequency of drought events changed similarly but decreased slightly in Europe. The three kinds of EWEs increased most in Australia, and moderately in Asia, South America and Africa.

The magnitude of standardized precipitation index (SPI), standardized temperature index (STI), and standardized compound event indicator (SCEI) reflected the intensity of droughts, heatwaves and CHDE, with a larger STI, a lower SPI and SCEI indicating more severe disasters. Heatwave and CHDE intensified significantly in the past three decades at the rate of 0.04 yr^{-1} across almost all grids, compared to a weak increase in drought severity by $< 0.01 \text{ yr}^{-1}$ in 58.9% of grids (Fig.S47 d-f). At continent scale, the intensity of heatwaves had generally amplified across continents, with Australia and Africa ranked first. The case was same for CHDE, but the difference was that Australia and Asia experienced an intensified trend while Europe and North America undergone a decreased drought severity.

Rainfed systems were considerably more vulnerable to all EWEs than irrigated systems (Fig.S22). Spring wheat, which occupies ~96% of global rainfed harvested area (Fig.S23), experienced nearly twice the yield loss of winter wheat under extremes (Fig.S24). Droughts, heatwaves and CHDE caused average yield reductions of 6.9%, 2.7% and 7.4%, respectively, in rainfed areas, while irrigated wheat showed much smaller losses (Fig.S22)—consistent with previous findings (Jägermeyr & Frieler, 2018; Vogel et al, 2019). Irrigation buffers yield by alleviating soil water deficit, with an additional benefit from transpiration cooling (Troy et al., 2015). Winter wheat, with 48% of its area under irrigation, was thus less affected (Figs.S22-24).

The largest wheat yield losses due to drought and CHDE occurred in Australia

with an average of 21% and 32.2%, respectively (Fig.3). Africa, North America and South America also suffered severely from droughts (>10%), particularly for Morocco, Algeria, Canada and Argentina with average wheat yield losses ranging from 15%-21.6%. (Fig.3). CHDE generally caused almost two times yield losses across continents compared to single extreme events (Fig.3). North America, Australia and Europe were similarly affected most by heatwaves with yield dropped by an average of 7.6%, including Denmark, the United Kingdom, France and Canada (>10%) (Fig.3).

Text S7: Comparisons of model skills in capturing composite yield impacts from droughts, heatwaves and CHDE with the GGCM Phase 3 using the same climate inputs.

Using the EWE events recorded in the EM-DAT database, together with the FAO yield statistics, GWM-NCS-ECO-GlobalCropYield5min estimate that global droughts, heatwaves and CHDE resulted in national yield losses of 3.6%, 3.4%, and 7.1% respectively (Fig.S37). In addition, wheat yield losses by 19.7% and 12% due to 2003 European drought and heatwave were estimated. In comparison, GWM-NCS-ECO-SPAM showed relatively poor performance in reproducing observed yield losses, especially for heatwaves, CHDE and 2003 European heatwaves with explanatory ability decreased by approximately 10-15% (Fig.S38). We ascribed the superior performance of GWM-NCS-ECO-GlobalCropYield5min to that GlobalCropYield5min better captured spatiotemporal variations in crop yield compared to SPAM. However, both model versions showed enhanced skills in reproducing EWEs impact on yields compared to GGCM phase 3 models. For heatwaves, GWM-NCS-ECO-SPAM increased explanatory ability by 45% compared to pDSSAT model of GGCM Phase 3, respectively (Figs.S28, S38). In terms of 2003 European droughts and heatwaves, the modelling skill is improved by 85.2% and 44.4%.

Noticeably, the two model versions at 0.5° explains approximately half of the observed yield losses due to droughts, whereas the model at 0.1° increases the explanatory ability by an average of 27.7%. We further compared the differences between observed and simulated impacts of droughts on wheat yields at different spatial resolution for each EWE event and country, with the differences closer to 0 representing better model explanatory ability (Fig.S39). The results indicated that the high-resolution model's estimates were significantly closer to observations ($p < 0.05$). Therefore, we ascribed it to that high-resolution meteorological data and yield simulations could provide more heterogeneous estimates of EWE impacts on crop yield (Figs.S40-42), although both resolutions consistently identified CHDE as the most damaging and highlighted similar geographic hotspots.

Text S8: Introduction of CERES-Wheat model

DSSAT includes growth models for over 42 crop types, including legumes (e.g., soybean) and cereals (e.g., maize, rice, wheat), and has been widely applied for global climate change impacts and adaptation, yield forecasting, and providing critical scientific support for policy-making (Jones et al., 2003; Xie et al., 2018). Here we utilize the CERES-Wheat models within the DSSAT framework to dynamically simulate the growth and yield formation processes of wheat, accounting influences from weather, soil, and field management practices.

The CERES-Wheat model simulates wheat growth, development, and yield formation through several key processes: phenological development, dry matter accumulation, leaf growth, yield formation, soil water balance, and soil carbon-nitrogen balance. The main mechanisms of key processes are described below.

(1) Phenology simulation

The model uses Accumulated Thermal Development Units (ATDU) to regulate progression through growth stages after emergence. ATDU is influenced by three factors: Daily Thermal Time (DTT), Vernalization Effect (VF), and Daylength Effect (DV):

$$TDU = \sum DTT * \min(VF, DV) \quad (1)$$

$$TT = \begin{cases} 0, & T_{mean} < T_{base} \text{ or } T_{mean} \geq T_{max} \\ \frac{T_{mean}}{T_{ol}} * T_{mean}, & T_{base} \leq T_{mean} < T_{ol} \\ T_{mean}, & T_{ol} \leq T_{mean} < T_{ou} \\ \frac{T_{max}-T_{mean}}{T_{max}-T_{ou}} * T_{mean}, & T_{ou} \leq T_{mean} < T_{max} \end{cases} \quad (2)$$

$$DV = 1 - \left[\left(\frac{P1D}{10,000} \right) * (20 - DL)^2 \right] \quad (3)$$

$$VF = VF_0 + \left(\frac{CUMVD}{P1V} \right) \quad (4)$$

$$VF_0 = 1 - \left(\frac{P1V}{50} \right) \quad (5)$$

Where T_{mean} and T_{max} are the daily mean and maximum temperatures, T_{base} is the base temperature, and T_{ou} and T_{ol} are the upper and lower bounds of the optimum temperature range, respectively. $P1D$ is the photoperiod sensitivity coefficient, and DL is the mean day length, computed as a function of latitude and day of the year (DOY).

VF0 denotes the relative development rate without vernalization, P1V is the vernalization sensitivity coefficient, and CUMVD is the cumulative number of vernalization days.

(2) Dry matter accumulation

The CERES-Wheat model calculates dry matter production based on daily solar radiation using a Radiation Use Efficiency (RUE) approach:

$$CARB = \begin{cases} PARUE \times \frac{PARAD}{PLTPOP} \times PARI, & \text{if } GP = VGP \\ PARU2 \times \frac{PARAD}{PLTPOP} \times PARI, & \text{if } GP = RGP \end{cases} \quad (6)$$

$$CARBO = AMAX1(PCARB \times CO2FP \times TFP \times WFP \times NFP \times RSFP \times SLPF) \quad (7)$$

where PCARB and CARBO represent potential and actual biomass production, PARUE and PARU2 are photosynthetically active radiation (PAR) utilization efficiency for vegetative and reproductive growth periods, respectively. PARAD is photosynthetically active radiation, PLTPOP is plant population. PARI is PAR interception fraction. CO2FP, TFP, WFP, NFP, RSFP, and SLPF are CO₂, temperature, water, nitrogen, reserve and soil factor related to photosynthesis functions. AMAX1 is a function that returns the maximum value of its arguments.

(3) Yield formation

After grain filling begins, yield is determined by the translocation and accumulation of dry matter into grains:

$$YIELD = GRNWT * 10 * EARS \quad (8)$$

$$GRNWT_i = GRNWT_{i-1} + GROGRN \quad (9)$$

$$GROGRN = RGFILL * GPP * G3 * 0.001 * (0.45 + 0.55 * SWFAC) \quad (10)$$

Where YIELD is wheat yield, GRNWT is grain weight per plant, EARS is the number of ears per m², GROGRN is the total daily weight gain of the kernels, GPP is grains per plant. G3 is a cultivar parameter representing potential stem weight at maturity under non-stressed conditions, SWFAC is water stress factors, RGFILL is the relative grain filling rate.

In its original form, the CERES-Wheat model employed a linear temperature response function, incorporating a range for the optimal temperature, with distinct

thresholds for different physiological processes (Fig.S51). For example, the temperature response of reproductive growth period defined the base and ceiling temperature thresholds as 0 °C and 60 °C, with the optimal temperature ranging from 30 to 50 °C. The original model assumed a constant maximum development rate of 1 across a broad optimal temperature range. However, recent studies demonstrated that crop development follows an Arrhenius-type temperature response (Parent et al., 2010). For wheat, this function defined base, optimum, and ceiling temperatures as 0 °C, 27.7 °C, and 40 °C, respectively (Parent et al., 2010). Wang et al. (2017) used this information to derive the temperature response of several key physiological processes for wheat. For instance, the new Wang-Engel function defined base, optimum, and ceiling temperatures for post-anthesis development as 0 °C, 33 °C, and 44 °C, respectively. Notably, the development rate declined progressively as temperature increased beyond 33 °C, reaching zero above 44 °C. This nonlinear response more accurately captures wheat phenological behavior compared to the original simplified representation. Consequently, the CERES-Wheat temperature function showed limited capacity under extreme temperatures, introducing significant uncertainty in simulations.

Text S9: Wheat yield data

The underlying wheat area map for our study is depicted in [Figure S48](#), which shows the full extent of all grids where wheat is cultivated in our study area (67 countries representing 98% of global wheat area and production). We first used wheat area maps and yield data from GlobalWheatYield4km (Zhang et al., 2024). The wheat maps were derived by first obtaining the wheat distribution maps from 2006 to 2014 using a phenology-based crop mapping methods, then integrating annual 1-km map as a single map on the basis of the grids with cultivation for at least 5 years to reduce the uncertainty of remote sensing data. Since GlobalWheatYield4km yield data was generated by a deep learning method based on the collection of subnational yield statistics covering 54 countries, its wheat maps were only generated for these countries ([Table S4](#)). For the remaining 13 countries with only FAO national-scale statistics available within our study area, including Algeria, Egypt, Ethiopia, Iraq, Libya, Morocco, Saudi Arabia, Sudan, Syria, Tunisia, Turkmenistan, Uzbekistan, and Yemen, gridded yield during 1981-2015 were predicted using the Random Forest (RF) approach. The LSTM model was not applied due to limited sample size. The gridded yield estimates for these 13 countries achieved an R^2 of 0.8, with a corresponding RMSE of 766.4 kg/ha ([Fig.S49](#)).

Text S10: Comparisons of model skills in capturing yield variations with the GGCM Phase 1

The simulated yields from different model versions were compared at different spatial scales. For the parallel Decision Support System for Agrotechnology Transfer (pDSSAT) from GGCM Phase 1 (i.e., pDSSAT-GGCM) that adopt site-specific calibration, the simulated and observed values exhibit poor agreement, with determination coefficient R^2 values of 0.29 and even only 0.05 for national- and subnational-level yields and disperse distribution along the 1:1 line (corresponding to the larger centered root mean squared differences CRMSDs of 2.24 t/ha and 3.43 t/ha) (Table S3). In comparison, simulated yields from MME approach are less biased (CRMSD of 1.31 t/ha and 1.47 t/ha in Table S3) than all GGCM Phase 1 models, which does provide an efficient way to reduce simulated error. Nevertheless, GGCM-MME notably underestimated observed yields, especially for subnational scales, with only 9% R^2 (Fig.S1).

Text S11: Comparisons of model skills in capturing composite yield impacts from droughts, heatwaves and CHDE with GGCM Phase 1

Among all the model versions, EPIC-BOKU, GEPIC, pDSSAT-GGCM and GWM-NCS failed to capture the sign of disaster effects on wheat yield for most cases (Figs.S32-34). Our analyses highlighted that improved model structure does improve GWM performance for reproducing EWEs impacts on yields. GGCM-MME approximately explained less than half of the observed yield losses, however, GWM-NCS-ECO (GWM-NCS-ECO-WE) increased the explanatory ability by 35.4% (41.5%), 60.1% (80.2%) and 24.5% (34.9%) from droughts, heatwaves and CHDE, respectively (Figs.S32-34). Such findings underscored the better performance of GWM-NCS-ECO and GWM-NCS-ECO-WE than GGCM-MME in better representing EWE impact on crop yield.

We further assessed the model performance in capturing wheat yield responses to certain catastrophic disasters, such as 2003 European drought and heatwave. Wheat yield losses were observed by 19.6% and 13.5% from 2003 drought and heatwave, which was well captured by GWM-NCS-ECO-WE (20.5% and 16.3%) and GWM-NCS-ECO (19.8% and 12.3%) yet by pDSSAT-GGCM (Figs.S35-36). Meanwhile, GGCM-MME could only capture less than quarter of yield losses (4.6% and 2.2% in Figs.S35-36).

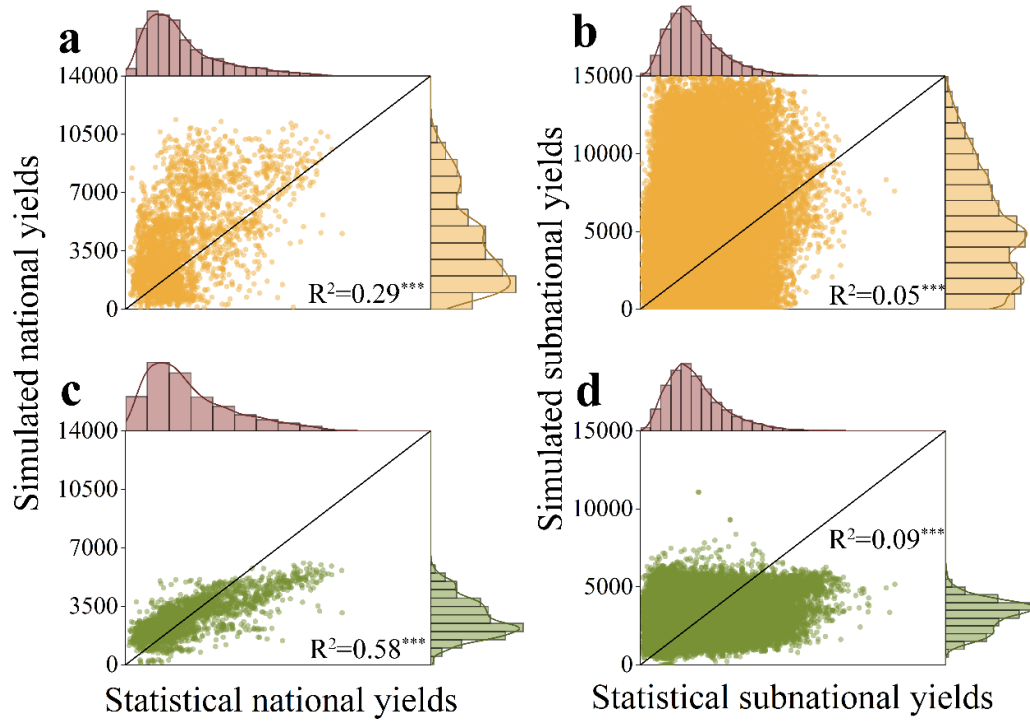


Fig.S1: Comparison of the simulated national and subnational yields with statistical ones by different model versions of GGCM Phase 1. Same as Fig. 2, but (a)-(b) and (c)-(d) are pDSSAT and multi-model ensemble (MME), respectively. *** ($P < 0.001$), ** ($P < 0.01$) and * ($P < 0.05$) indicate significant Pearson correlation (pDSSAT for national: $n = 2,944$; pDSSAT for subnational: $n = 93,385$; MME for national: $n = 2,147$; MME for subnational: $n = 106,291$). The statistical tests were two-sided, and no adjustments were made for multiple comparisons.

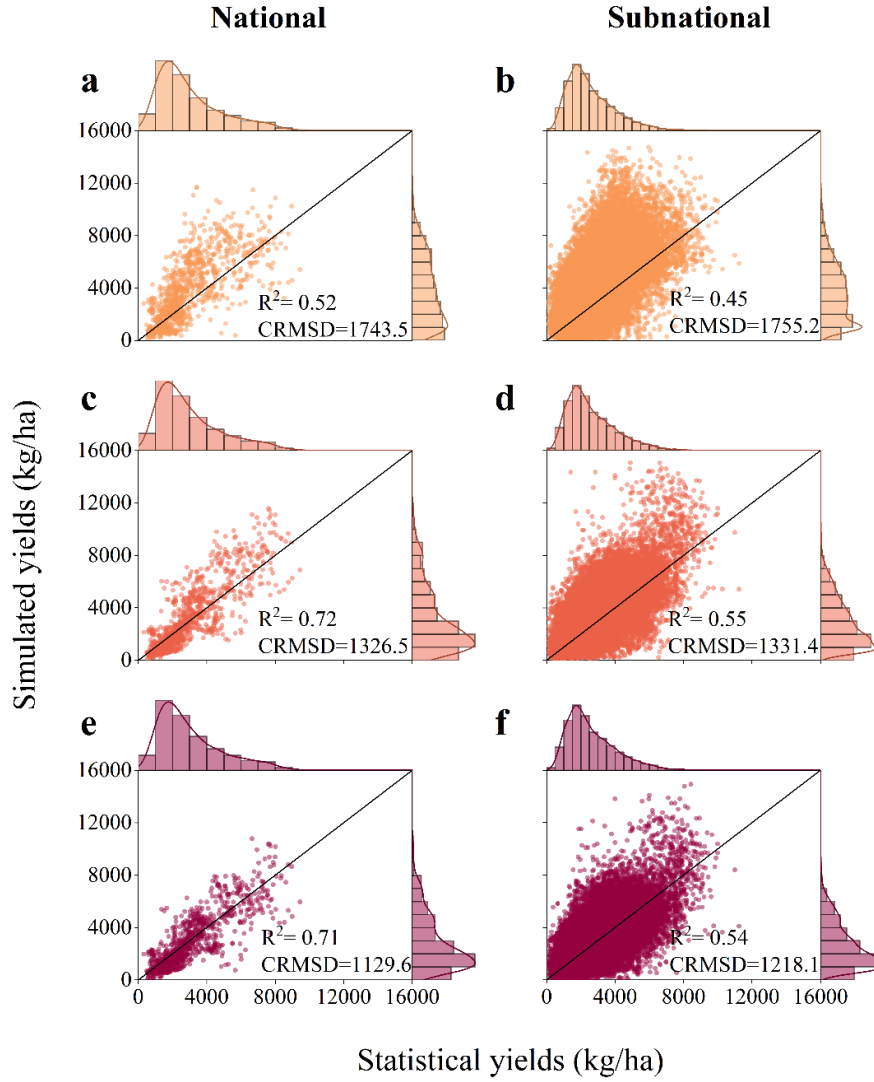


Fig.S2: Out-of-sample validation of our different model versions. (a, b) GWM-NCS, (c, d) GWM-NCS-ECO, and (e, f) GWM-NCS-ECO-WE at national (left) and subnational (right) scales (1981–1994 and 2011–2015). GWM-NCS for national: $n = 1,027$; GWM-NCS for subnational: $n = 49,235$; GWM-NCS-ECO for national: $n = 1,044$; GWM-NCS-ECO for subnational: $n = 49,975$; GWM-NCS-ECO-WE for national: $n = 1,023$; GWM-NCS-ECO-WE for subnational: $n = 49,598$. Histograms show probability density distributions.

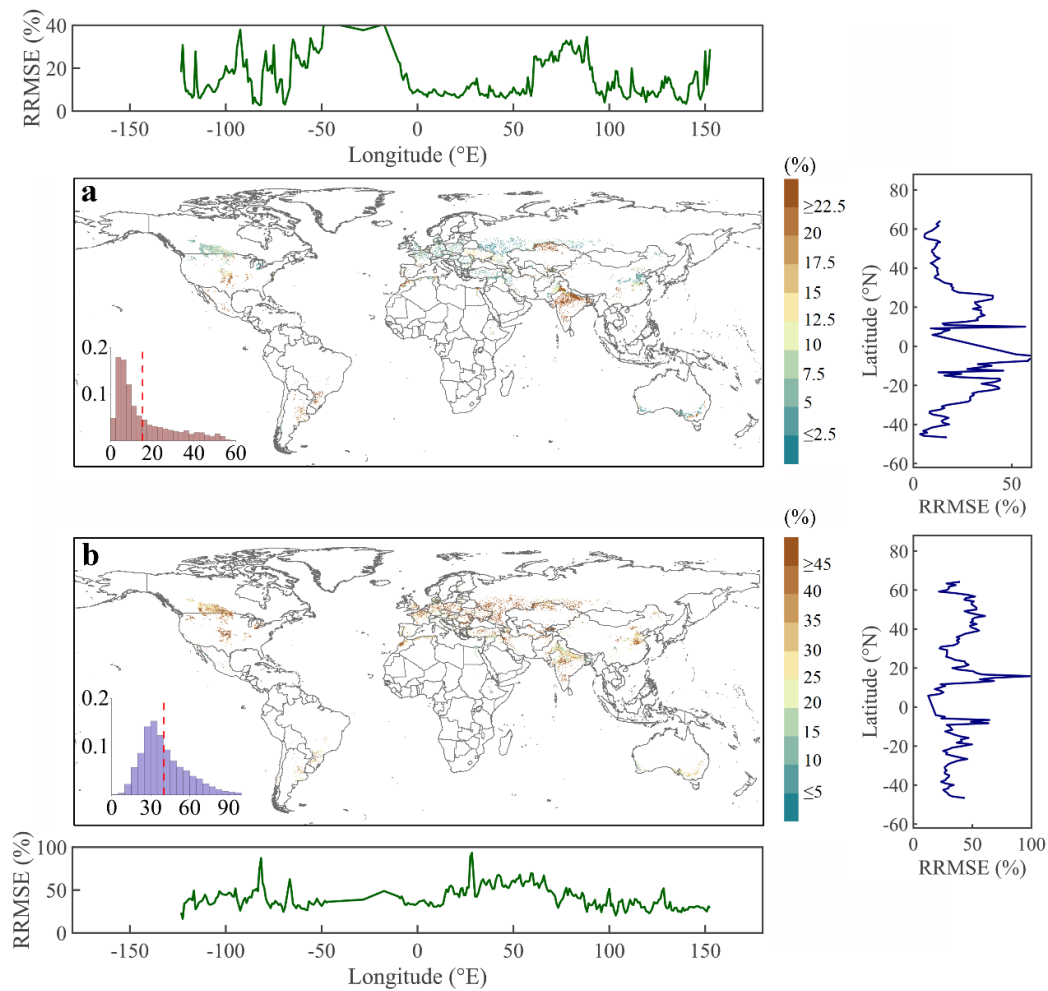


Fig.S3: The RRMSE of maturity date (a) and yield (b) for each grid simulated by GWM-NCS. Colored lines show the spatial patterns of RRMSE along longitudinal and latitudinal gradients. The inset in panel (a) and (b) depicts the frequency distributions of RRMSE for MDAP and yield.

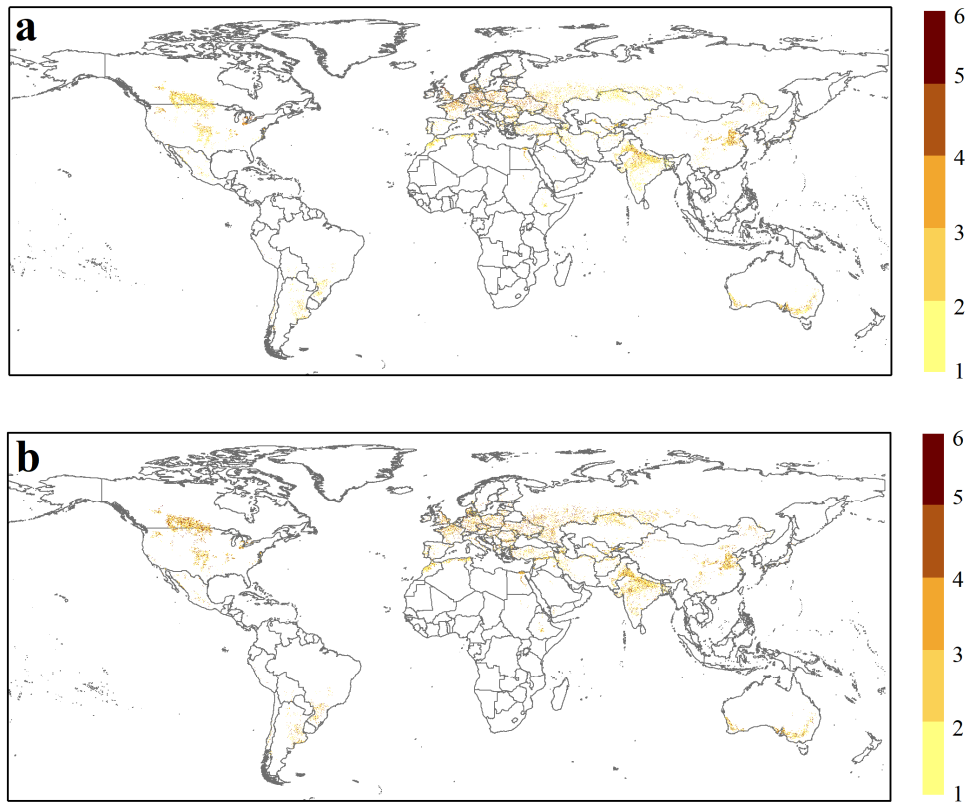


Fig.S4: Spatial patterns of PARUE (a) and PARU2 (b) from GWM-NCS-ECO. PARUE and PARU2 are ecotype-parameters from the model, depicting the radiation use efficiency.

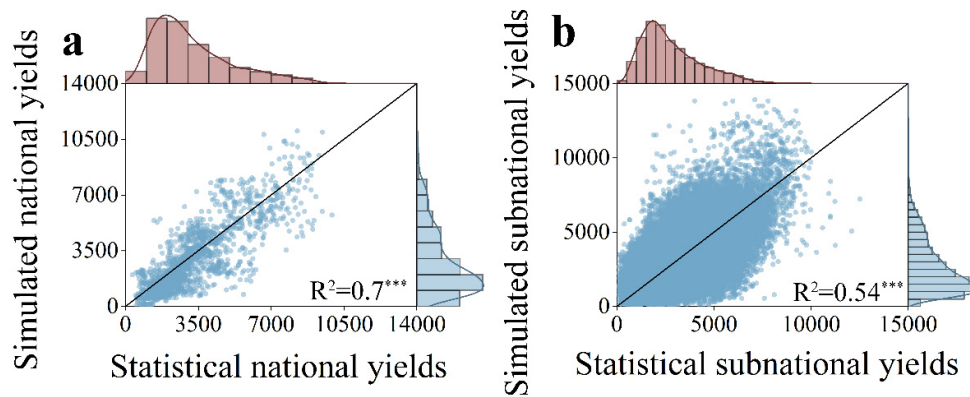


Fig.S5: Comparison of simulated and statistical yields by GWM-NCS-ECO-WE.

Yields were aggregated to national **(a)** and subnational **(b)** scale of spring and winter wheat under both rainfed and irrigated conditions; *** ($P < 0.001$), ** ($P < 0.01$) and * ($P < 0.05$) indicate significant Pearson correlation (national: $n = 1,674$; subnational: $n = 83,655$). The statistical tests were two-sided, and no adjustments were made for multiple comparisons. The histograms and solid lines for the probability density of the simulated and statistical data.

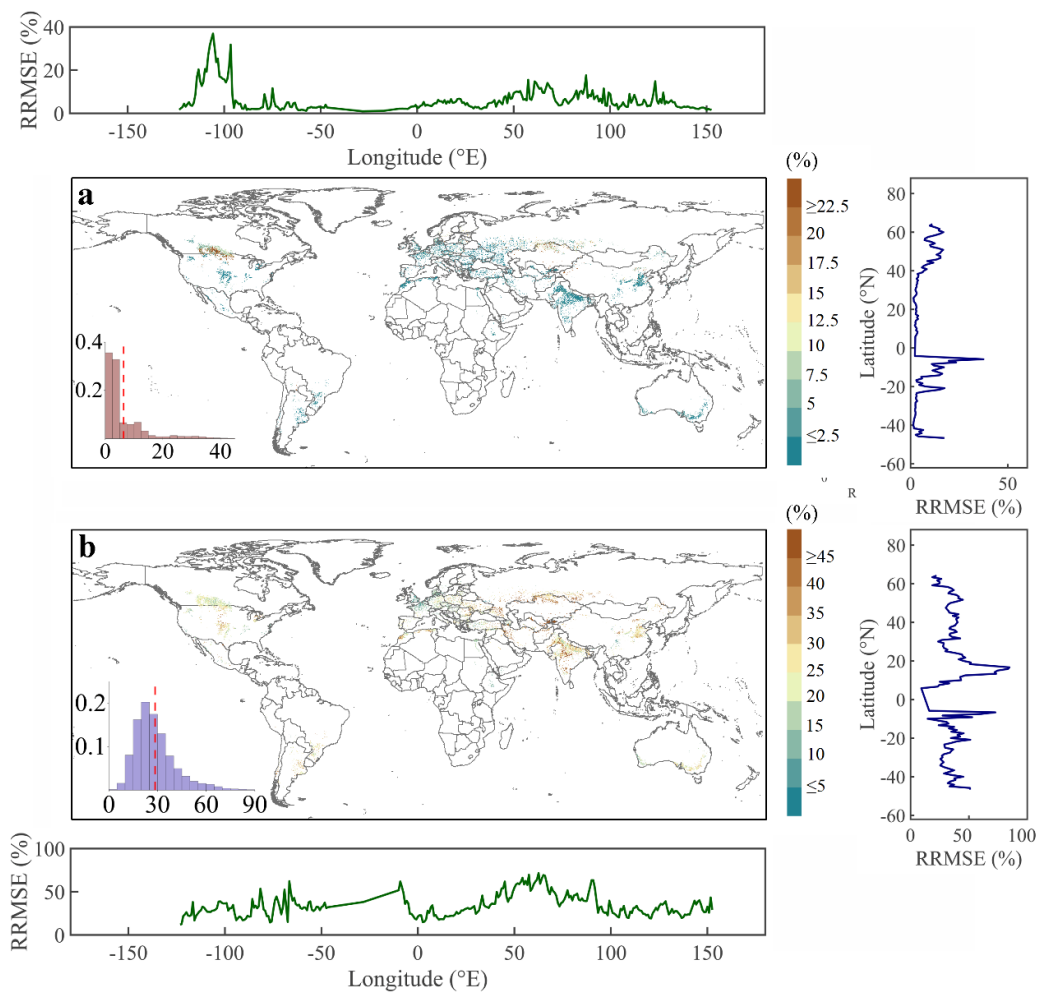


Fig.S6: The RRMSE of maturity date (a) and yield (b) for each grid simulated by GWM-NCS-ECO-WE. Colored lines show the spatial patterns of RRMSE along longitudinal and latitudinal gradients. Gridded observed maturity dates and yields were derived from GGCM Phase 3 crop calendar and GlobalWheatYield4km; The inset in panel (a) and (b) depicts the frequency distributions of RRMSE for maturity date and yield.

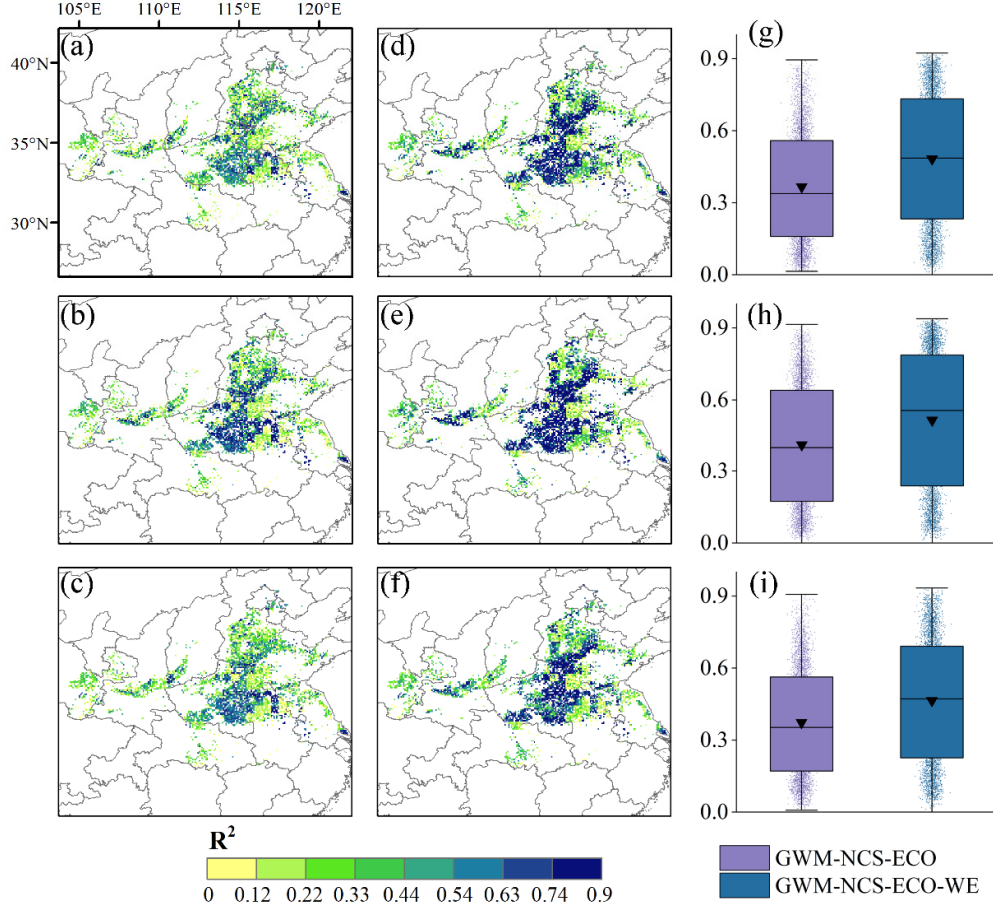


Fig.S7: Spatial patterns of average R^2 between simulated LAI by GWM-NCS-ECO (a-c) or GWM-NCS-ECO-WE (d-f) and satellite-retrieved values during 2001-2015 in the North China Plain. Satellite-retrieved LAI are derived from GLASS LAI (top row), GIMMIS LAI4g (medium row), and GLOBMAP products (bottom row). Spatial distributions are shown as boxplots (g-i), which indicate the median (center line), mean (triangle), interquartile range (IQR, box bounds), whiskers extending up to minimum and maximum values within $1.5 \times \text{IQR}$. The sample sizes (n) of independent spatial grid cells are: GWM-NCS-ECO (GLASS LAI: 4085; GIMMIS LAI4g: 4046; GLOBMAP: 4082) and GWM-NCS-ECO-WE (GLASS LAI: 4091; GIMMIS LAI4g: 4055; GLOBMAP: 4088).

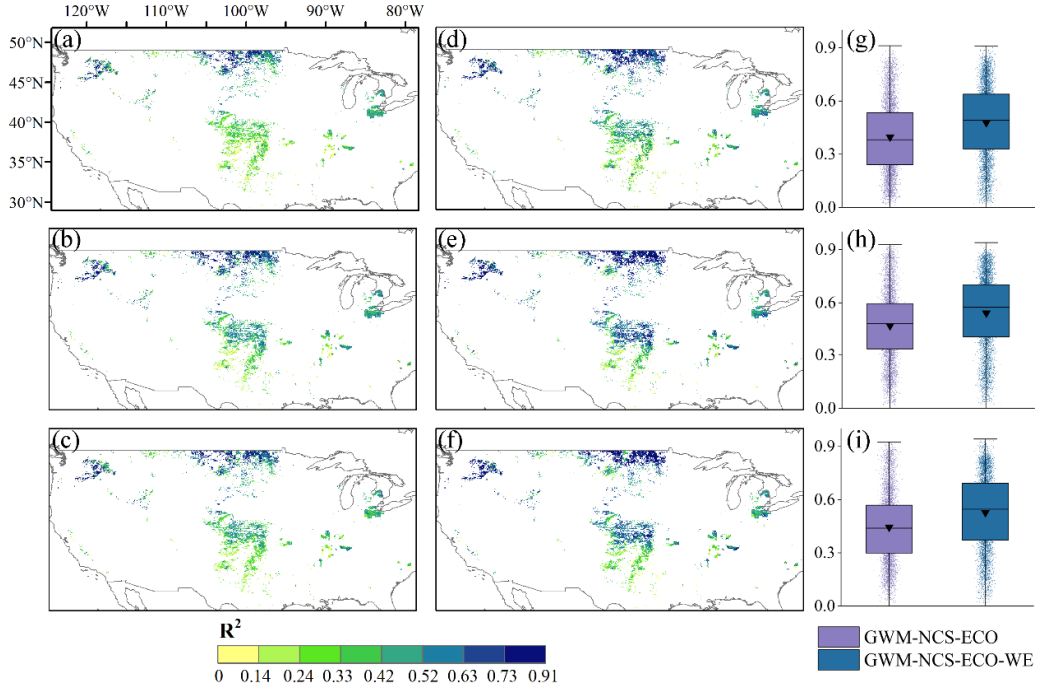


Fig.S8: Spatial patterns of average R^2 between simulated LAI by GWM-NCS-ECO (a-c) or GWM-NCS-ECO-WE (d-f) and satellite-retrieved values during 2001-2015 in the United States of America. Satellite-retrieved LAI are derived from GLASS LAI (top row), GIMMIS LAI4g (medium row), and GLOBMAP products (bottom row). Spatial distributions are shown as boxplots (g-i), which indicate the median (center line), mean (triangle), interquartile range (IQR, box bounds), whiskers extending up to minimum and maximum values within $1.5 \times \text{IQR}$. The sample sizes (n) of independent spatial grid cells are: GWM-NCS-ECO (GLASS LAI: 6358; GIMMIS LAI4g: 6356; GLOBMAP: 6359) and GWM-NCS-ECO-WE (GLASS LAI: 6438; GIMMIS LAI4g: 6434; GLOBMAP: 6438).

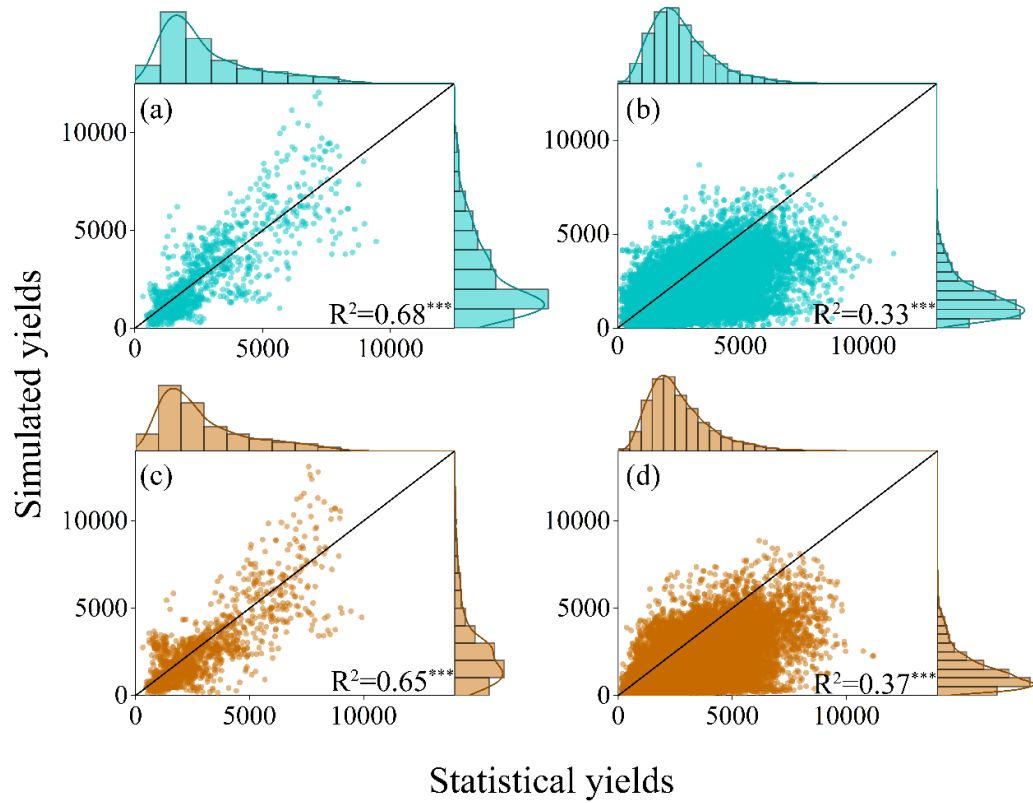


Fig.S9: Comparison of simulated and statistical yields by Global CERES-Wheat-NCS-ECO (GWM-NCS-ECO) at 0.5-degree resolution. Yields calibrated to GlobalCropYield5min (top row, a-b) and SPAM (bottom row, c-d) were aggregated to national (a, c) and subnational (b, d) scales, with histograms and solid lines for the probability density distributions of the simulated and statistical data. *** ($P < 0.001$), ** ($P < 0.01$) and * ($P < 0.05$) indicate significant Pearson correlation (GlobalCropYield5min for national: $n = 1092$; GlobalCropYield5min for subnational: $n = 38,184$; SPAM for national: $n = 1359$; SPAM for subnational: $n = 46,103$). The statistical tests were two-sided, and no adjustments were made for multiple comparisons.

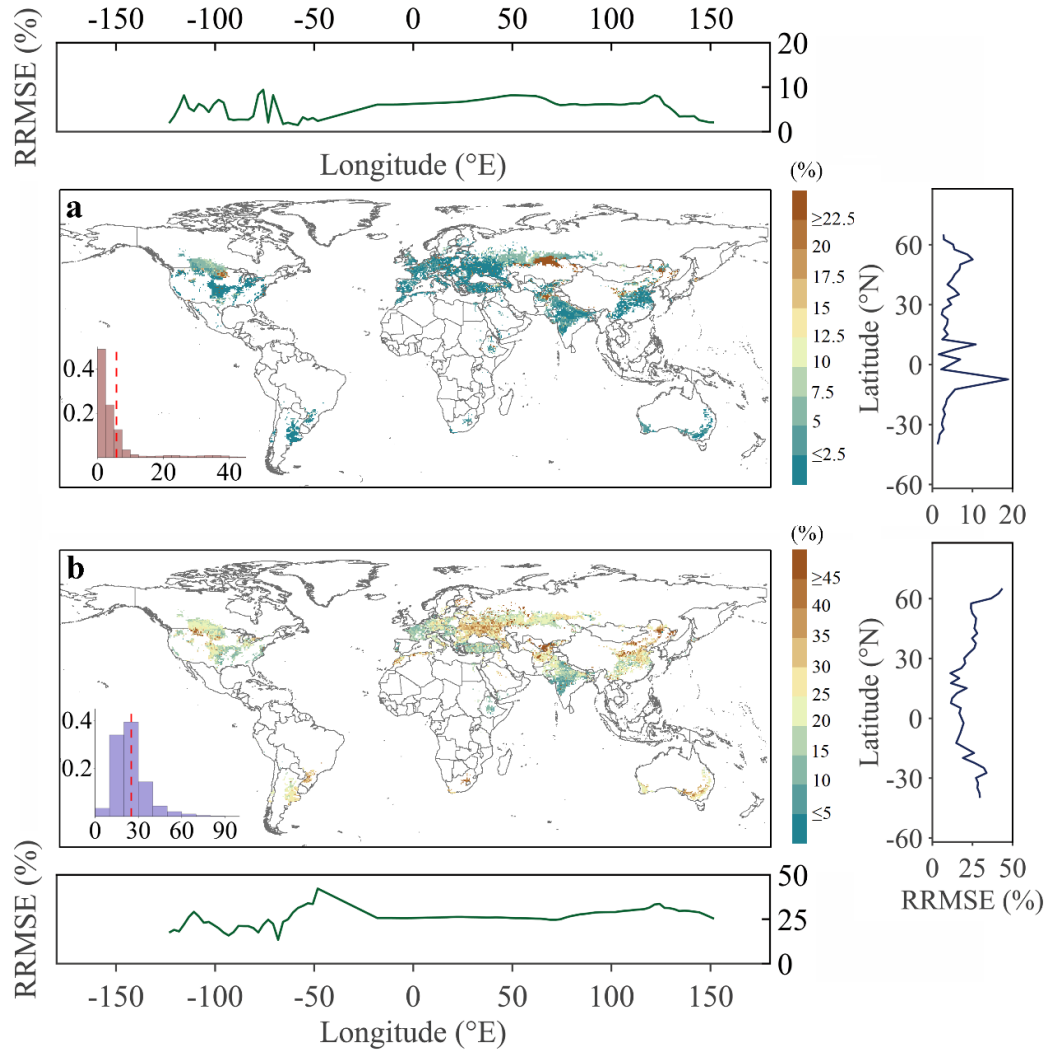


Fig.S10: Spatial distribution of simulation errors (RRMSE) of maturity date (a) and yield (b) for GWM-NCS-ECO calibrated with GlobalCropYield5min at 0.5-degree resolution. Colored lines show latitudinal and longitudinal gradients of RRMSE. Reference data are GGCM Phase 3 crop calendar (maturity) and GlobalCropYield5min (yield). Insets show the frequency distribution of RRMSE values.

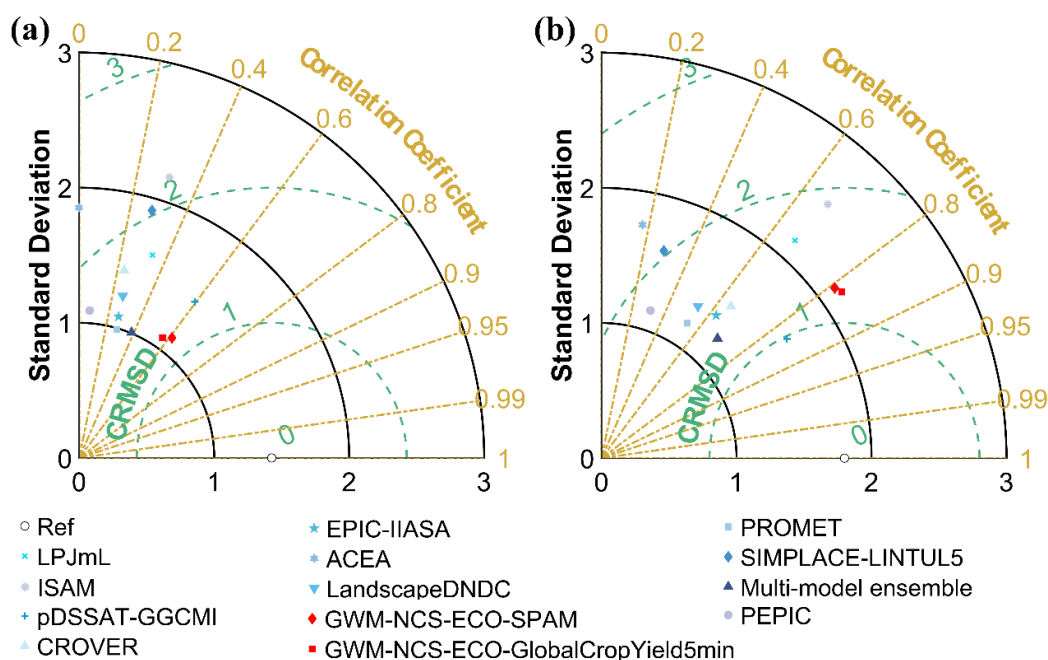


Fig.S11: Taylor diagram comparing model performances at subnational (a) and national (b) scales. Results for GGCM Phase 3 models are shown alongside our GWM-NCS-ECO run at 0.5° resolution. Proximity of symbols to the reference point indicates better agreement with observations, characterized by higher correlation (angle) and lower centered root mean square difference (contour).

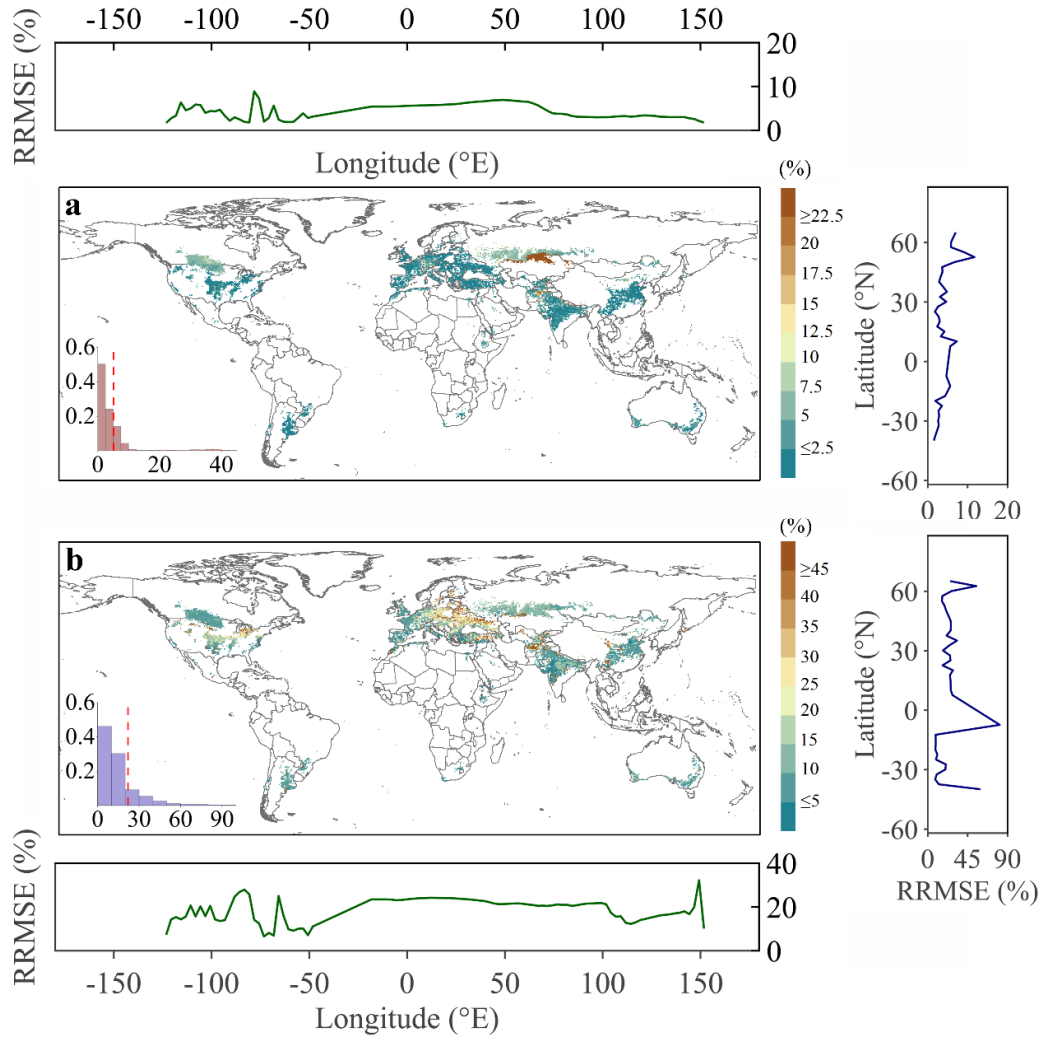


Fig.S12: Spatial distribution of simulation errors (RRMSE) of maturity date (a) and yield (b) for GWM-NCS-ECO calibrated with SPAM at 0.5-degree resolution. Reference dates and yields are from GGCMI Phase 3 crop calendar and SPAM. Colored lines and insets follow Fig.S10.

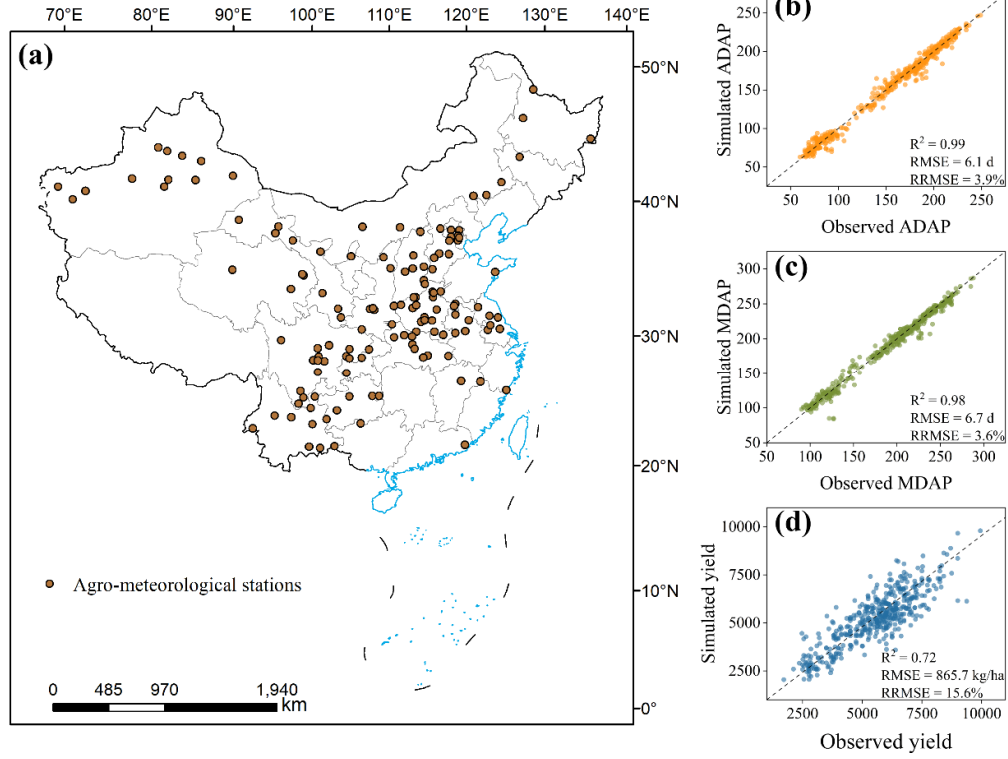


Fig.S13: Locations of agro-meteorological stations in China (a) and comparison of simulations and observations for anthesis date (ADAP, b), maturity date (MDAP, c), and yield (d) by GWM-NCS-ECO for validation.

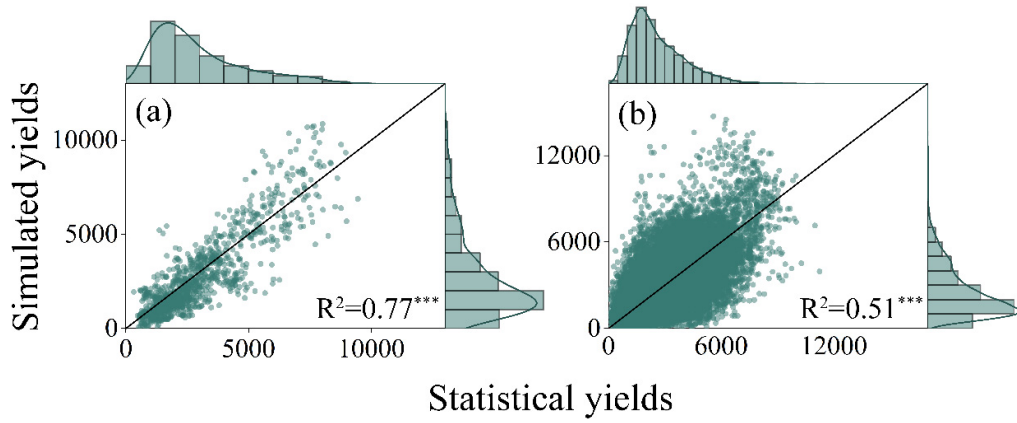


Fig.S14: Comparison of simulated and statistical yields by GWM-NCS-ECO-WE at 0.1° resolution calibrated by statistical data at national (a) and subnational (b) scales. Histograms show probability density distributions. *** ($P < 0.001$), ** ($P < 0.01$) and * ($P < 0.05$) indicate significant Pearson correlation (national: $n = 1,095$; subnational: $n = 51,082$). The statistical tests were two-sided, and no adjustments were made for multiple comparisons.

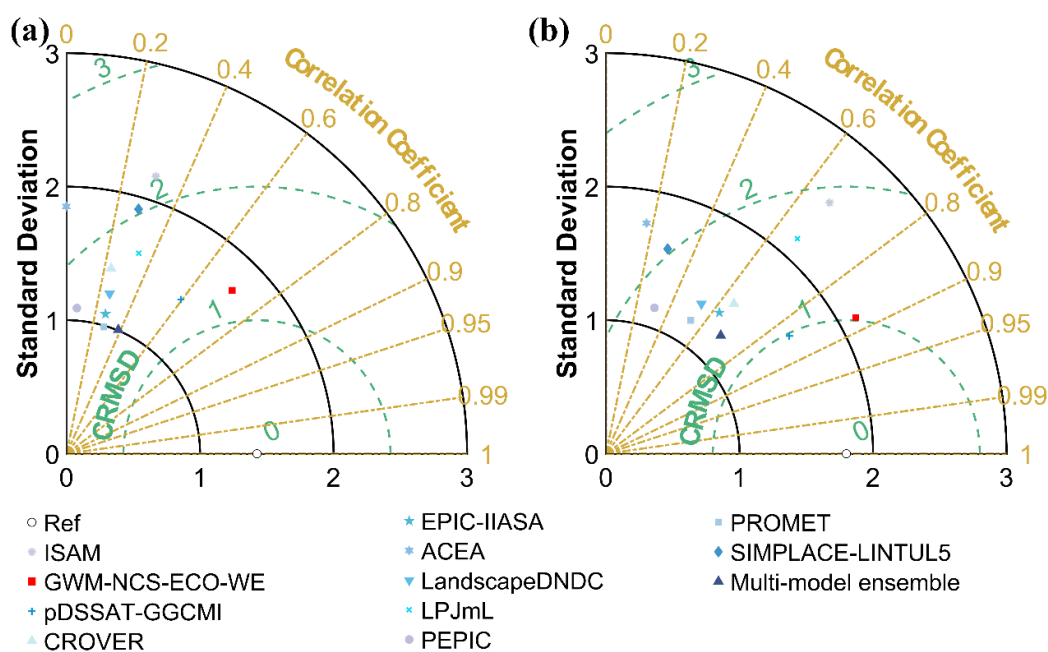


Fig.S15: Taylor diagram of simulated yields by GGCM Phase 3 models and GWM-NCS-ECO-WE calibrated by statistical data at subnational (a) and national (b) scales. Interpretation of symbols and contours follows Fig. S11.

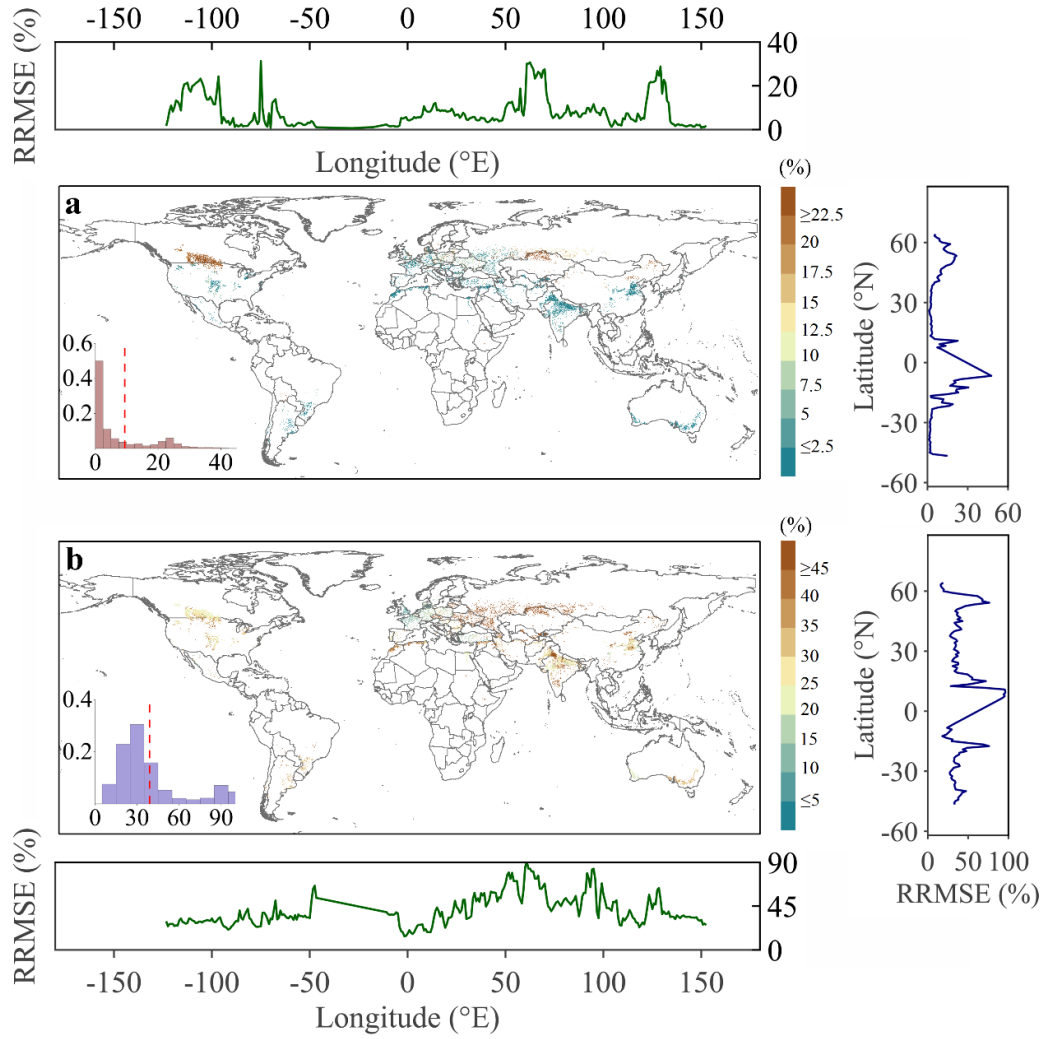


Fig.S16: RRMSE of (a) maturity date and (b) yield for GWM-NCS-ECO-WE calibrated by statistical data at 0.1° resolution. Colored lines show gradients; insets show frequency distributions. Reference data are from GGCM Phase 3 crop calendar and statistical data.

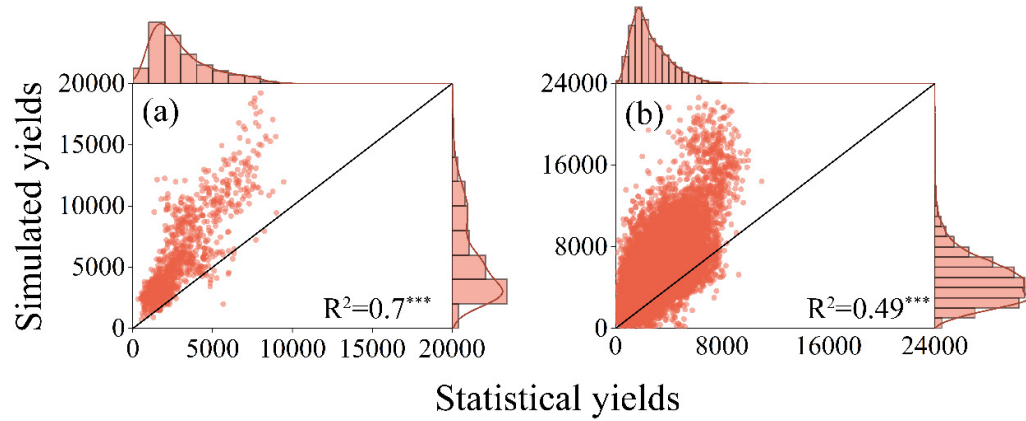


Fig.S17: Comparison of simulated and statistical yields by GWM-NCS-ECO-WE at 0.1° resolution calibrated under irrigated/rainfed conditions (from MIRCA data) at national (a) and subnational (b) scales. Histograms show probability density distributions. *** ($P < 0.001$), ** ($P < 0.01$) and * ($P < 0.05$) indicate significant Pearson correlation (national: $n = 2,530$; subnational: $n = 60,504$). The statistical tests were two-sided, and no adjustments were made for multiple comparisons.

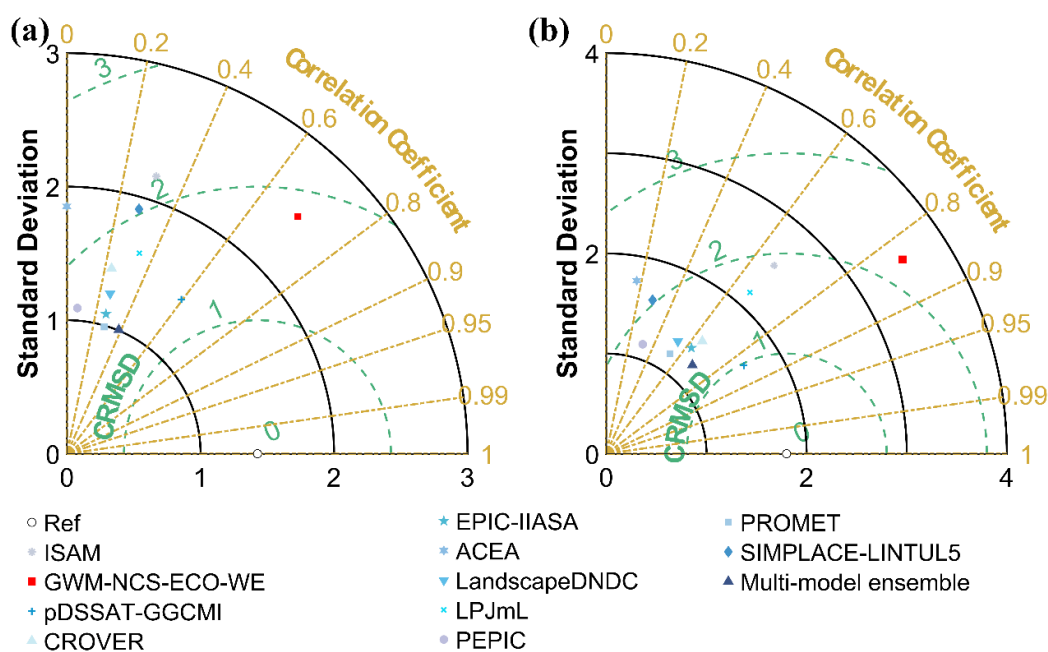


Fig.S18: Taylor diagram of simulated yields by GGCM Phase 3 models and GWM-NCS-ECO-WE calibrated under irrigated or rainfed conditions (from MIRCA data) at subnational (a) and national (b) scales. Interpretation of symbols and contours follows Fig. S11.

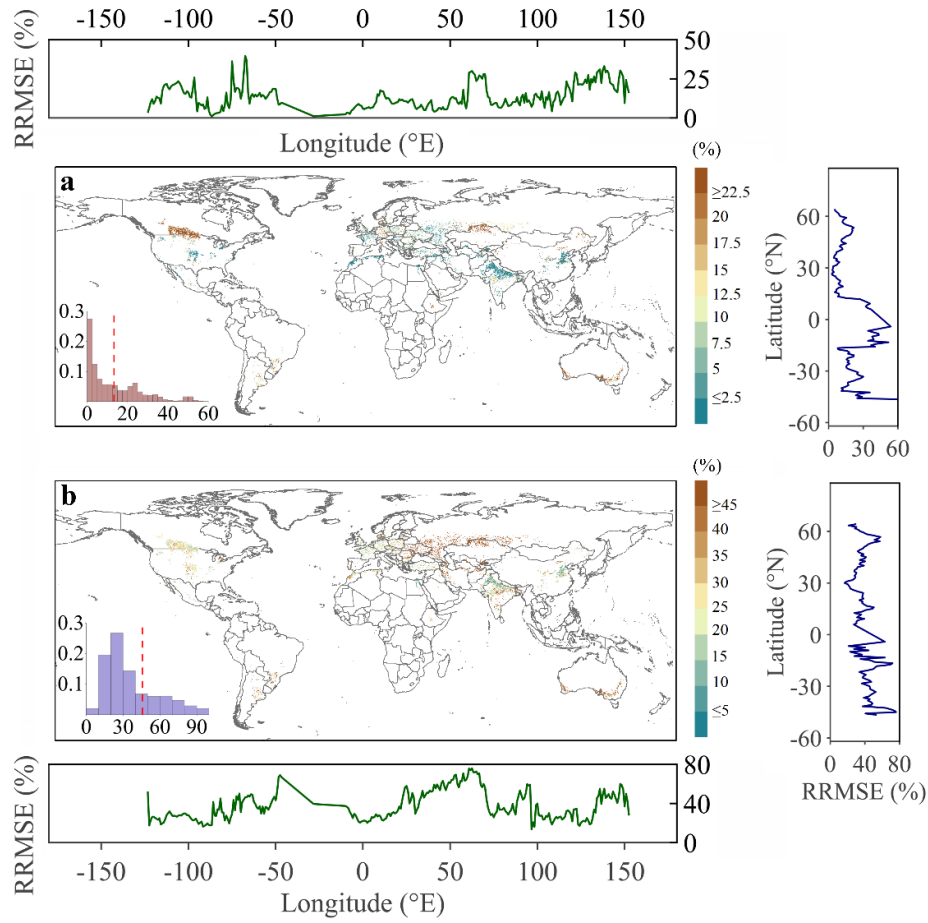


Fig.S19: RRMSE of (a) maturity date and (b) yield for GWM-NCS-ECO-WE calibrated under actual irrigated/rainfed conditions at 0.1° resolution. Colored lines show gradients; insets show frequency distributions. Reference data are from GGCM Phase 3 crop calendar and GlobalWheatYield4km.

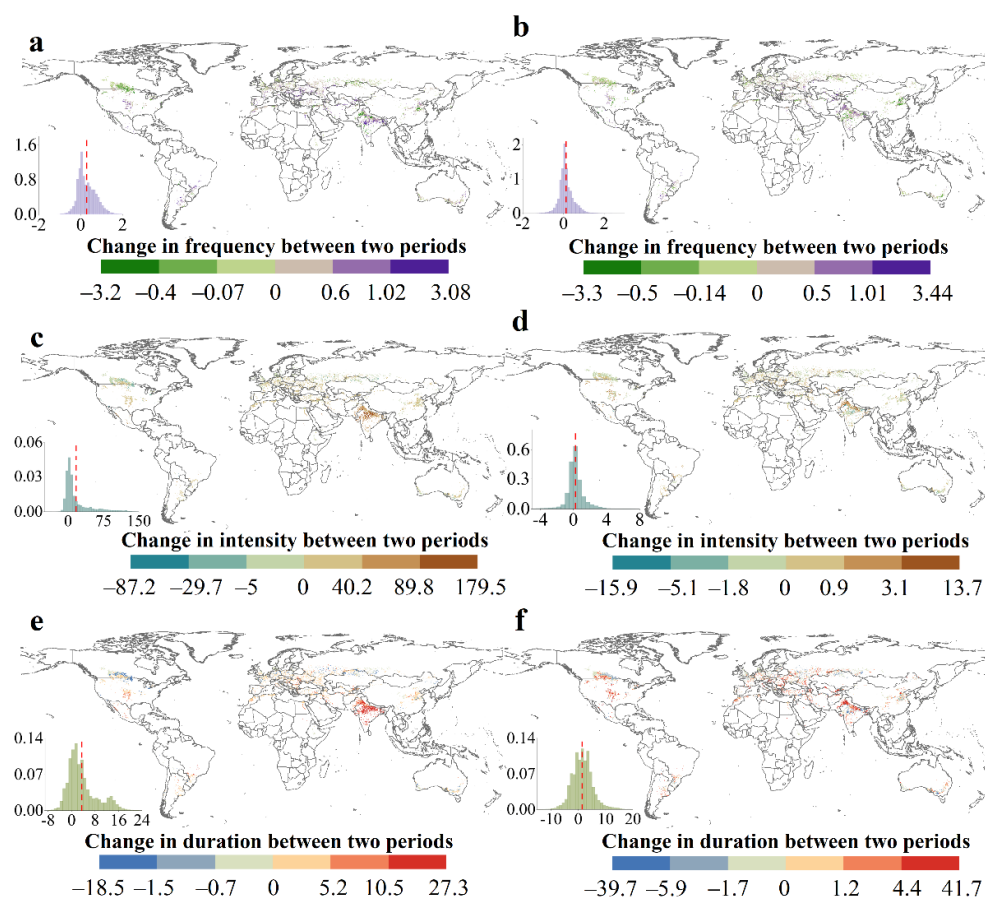


Fig.S20: Changes in frequency (a, b), intensity (c, d), and duration (e, f) of heatwaves (left) and CHDE (right) between 1981–1998 and 1999–2015. Insets show frequency distributions of changes (dashed lines for means).

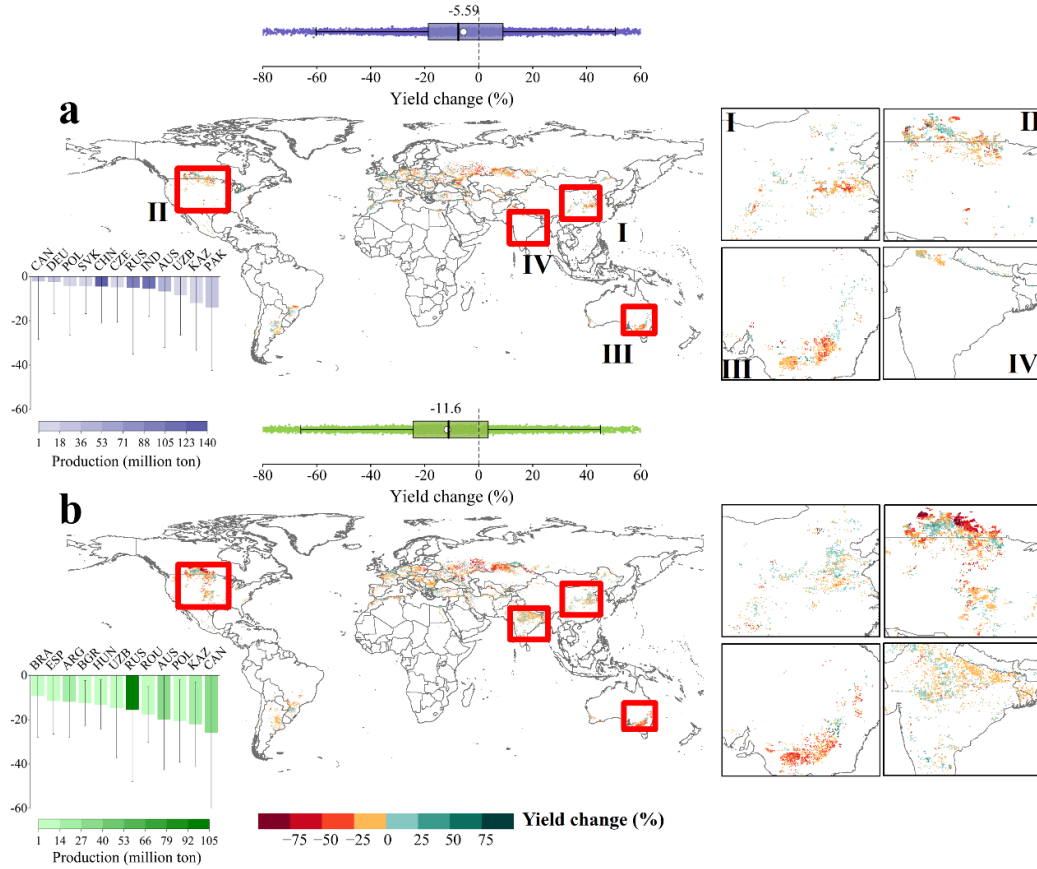


Fig.S21: Statistically significant yield changes (1981–2015) for (a) heatwaves and (b) CHDEs. Heatwave is defined as ≥ 3 consecutive days above temperature thresholds. Global distributions are shown as boxplots (box bounds for interquartile range, IQR; whiskers extending up to minimum and maximum values within $1.5 \times \text{IQR}$; white point for mean; center line for median). The sample sizes (n) of independent spatial grid cells are: heatwaves (13,998) and CHDE (20,406). National averages are shown as bar plots (error bars represent mean values $\pm$ standard error) for the twelve most affected countries (bar color indicates wheat production), with ISO-alpha3 country codes. The sample sizes (n) of independent spatial grid cells within each country are: heatwaves (PAK: 192; KAZ: 2,055; UZB: 446; AUS: 2,014; IND: 510; RUS: 5,760; CZE: 323; CHN: 3,144; SVK: 121; POL: 863; DEU: 1,556; CAN: 2,935) and CHDE (CAN: 2,508; KAZ: 1,403; POL: 457; AUS: 1,402; ROU: 324; RUS: 2,607; UZB: 195; HUN: 81; BGR: 159; ARG: 634; ESP: 100; BRA: 437).

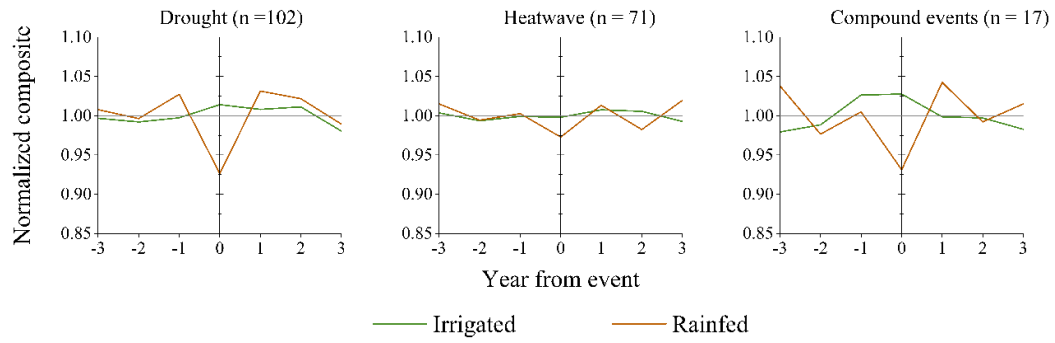


Fig.S22: Influences of droughts (a), heatwaves (b) and CHDE (c) on rainfed and irrigated wheat yields. Similar to Fig.4, but separated for rainfed and irrigated systems, with disasters recorded in EM-DAT during 1984–2007, simulated by the GWM-NCS-ECO-WE model.

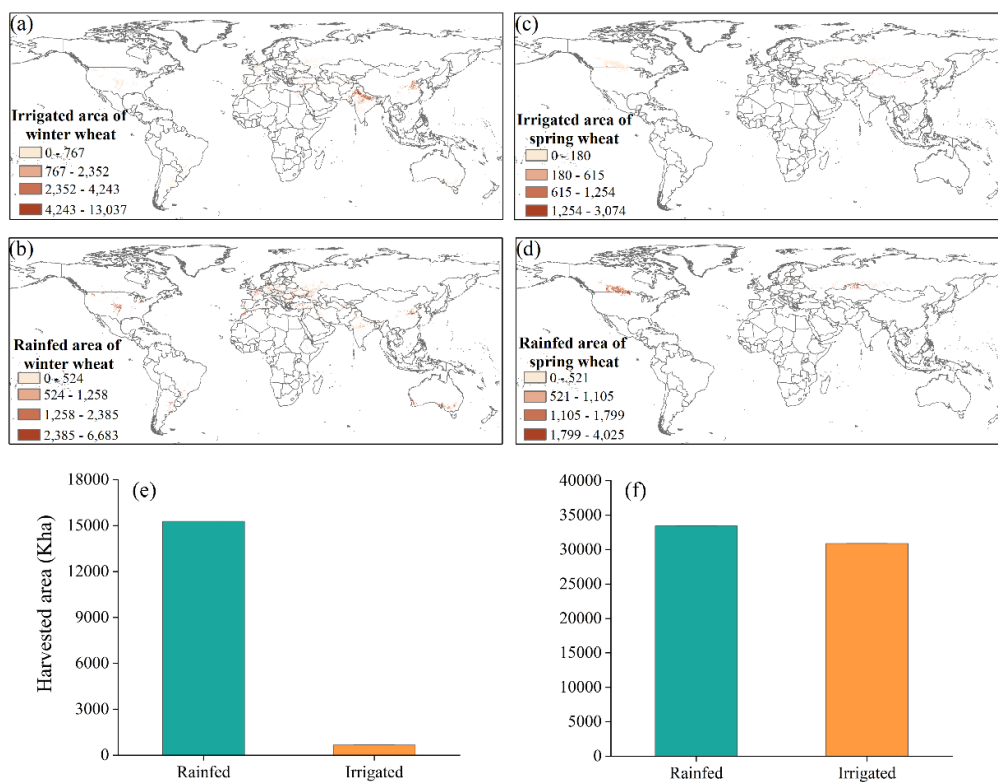


Fig.S23: Spatial patterns of rainfed and irrigated harvested area for spring (c, d, e) and winter wheat (a, b, f). The rainfed and irrigated harvested area were from MIRCA2000 dataset.

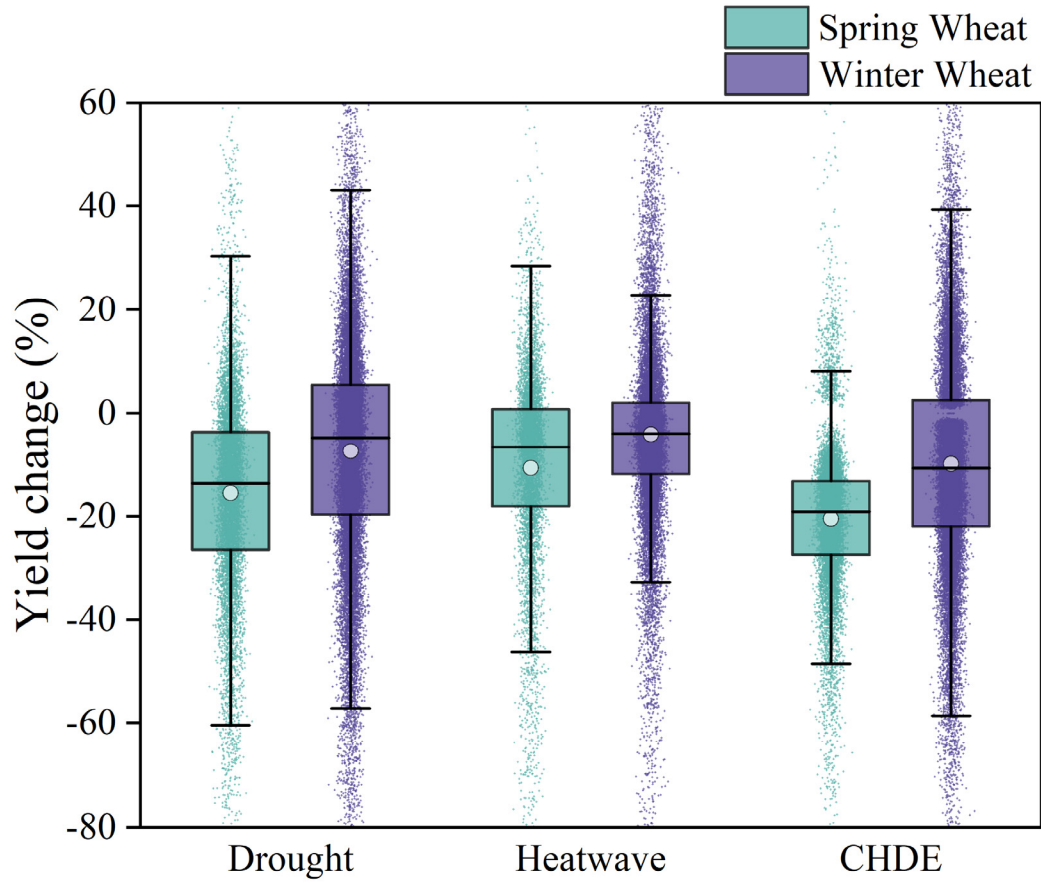


Fig.S24: Yield changes due to droughts, heatwaves and CHDE for spring and winter wheat. Box plots show the range of yield changes across all the grids, which indicate the median (center line), mean (circle), interquartile range (IQR, box bounds), whiskers extending up to minimum and maximum values within $1.5 \times \text{IQR}$. The sample sizes (n) of independent spatial grid cells are: droughts (spring wheat: 8,835; winter wheat: 21,988), heatwaves (spring wheat: 6,382; winter wheat: 17,817), and CHDE (spring wheat: 11,264; winter wheat: 23,342).

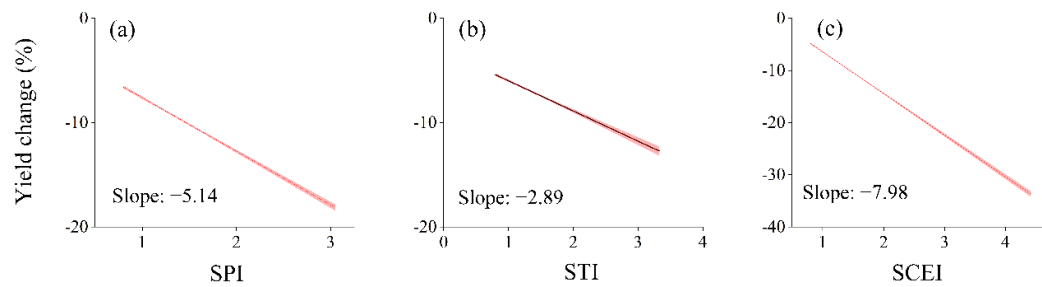


Fig.S25: Relationships between wheat yield change and $-SPI$ (a), STI (b), and $-SCEI$ (c). Red line denotes linear regression line with 95% confidence interval.

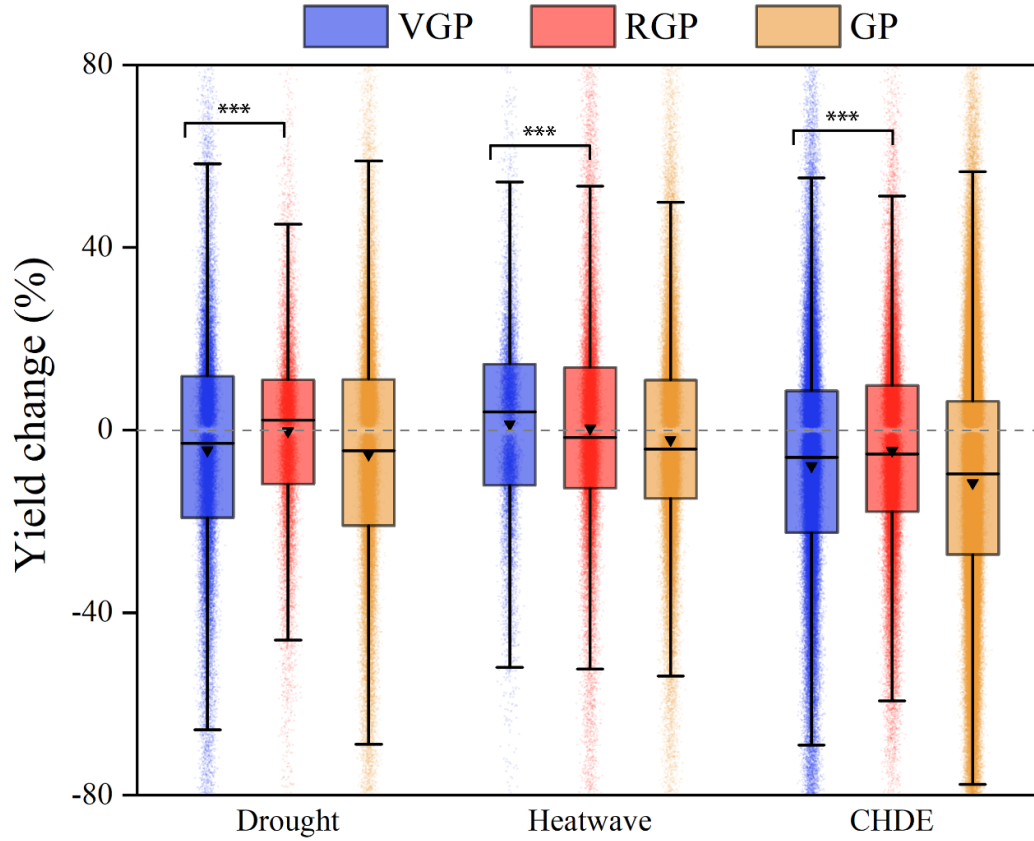


Fig.S26: Changes in wheat yield due to droughts, heatwaves and CHDE for different growth periods. Box plots show the range of yield changes across all the grids, which indicate the median (center line), mean (triangle); interquartile range (IQR, box bounds), whiskers extending up to minimum and maximum values within $1.5 \times \text{IQR}$; VGP vegetative growth period; RGP reproductive growth period; GP whole growth period; with significance of difference in median yield changes between VGP and RGP determined using a two-sided Wilcoxon test (***: $p < 0.001$; **: $p < 0.01$; *: $p < 0.05$; no adjustments were made for multiple comparisons). The sample sizes (n) of independent spatial grid cells for each year's event that occurs in corresponding GP are: drought (VGP: 35,463; RGP: 17,043; GP: 49,023), heatwave (VGP: 11,840; RGP: 33,805; GP: 58,410), and CHDE (VGP: 88,248; RGP: 50,899; GP: 220,909).

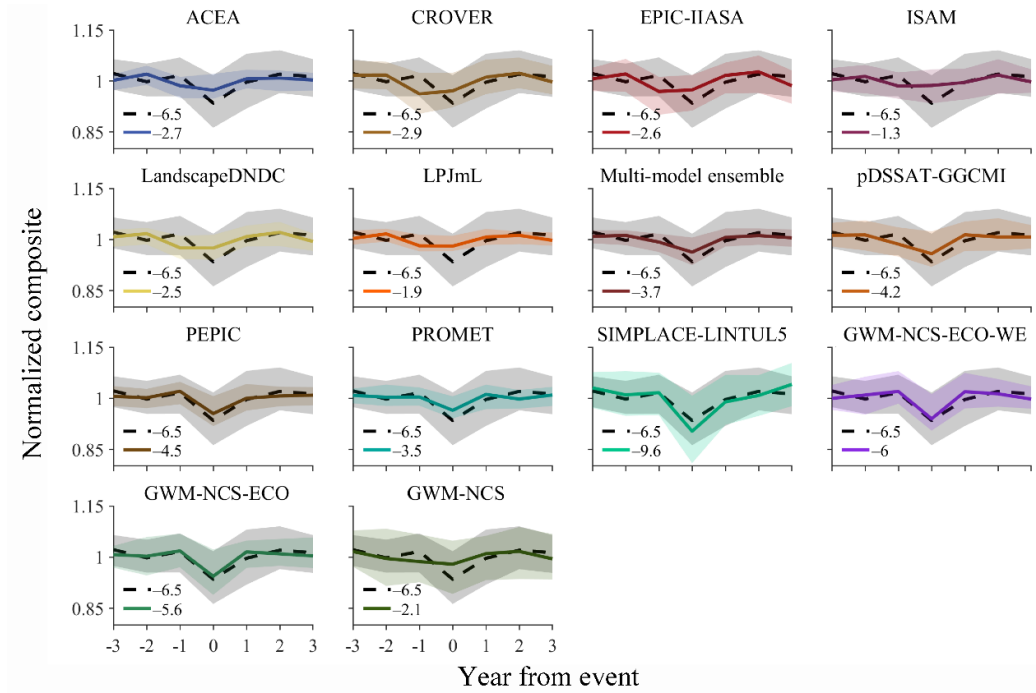


Fig.S27: Observed and simulated impacts of droughts on global wheat yields from GGCM Phase 3 and ours. Composites are based on 7-year time windows of country-level yields centered on the respective drought events; Data are presented as mean normalized wheat yield values across all identified disaster events, with shaded bands showing the 25th–75th percentile ranges across events, and legends for the mean event impacts.

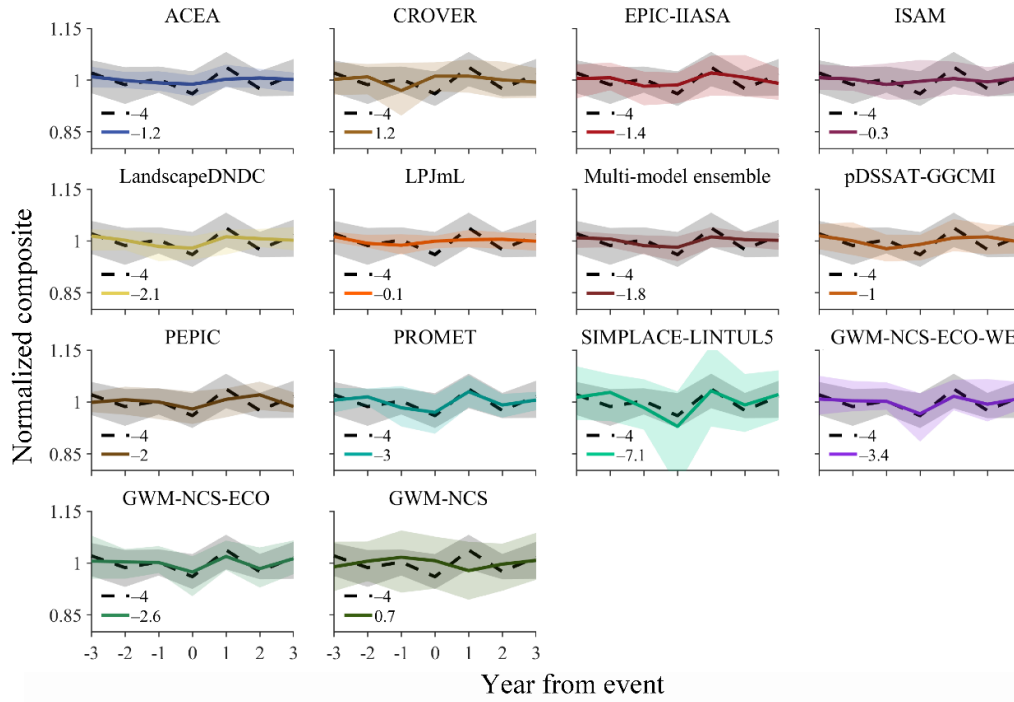


Fig.S28: Observed and simulated impacts of heatwaves on global wheat yields from GGCM Phase 3 and ours. Composites are based on 7-year time windows of country-level yields centered on the respective drought events; Data are presented as mean normalized wheat yield values across all identified disaster events, with shaded bands showing the 25th–75th percentile ranges across events, and legends for the mean event impacts.

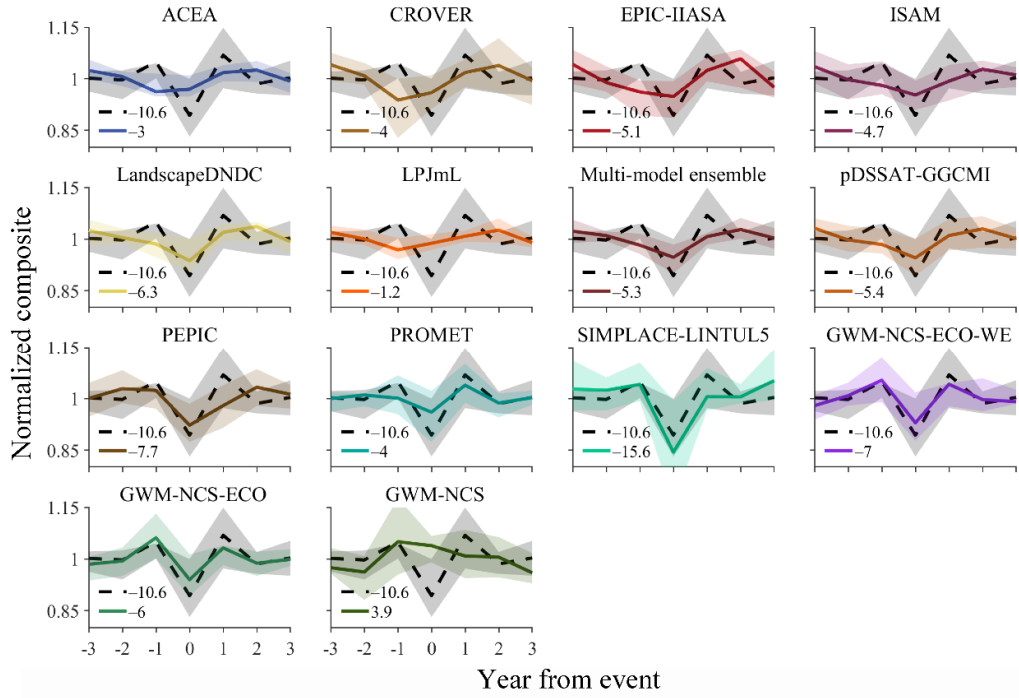


Fig.S29: Observed and simulated impacts of CHDE on global wheat yields from GGCM Phase 3 and ours. Composites are based on 7-year time windows of country-level yields centered on the respective drought events; Data are presented as mean normalized wheat yield values across all identified disaster events, with shaded bands showing the 25th–75th percentile ranges across events, and legends for the mean event impacts.

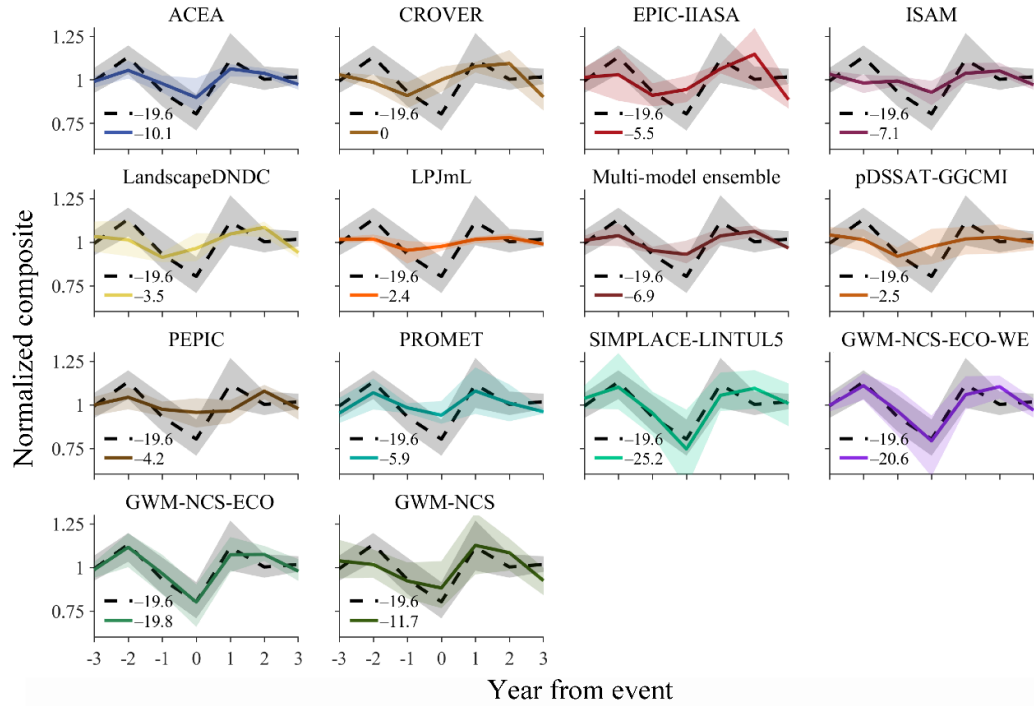


Fig.S30: Observed and simulated impacts of 2003 European drought on wheat yields from GGCMI Phase 3 and ours. Composites are based on 7-year time windows of country-level (5 countries listed in EM-DAT including Croatia, Hungary, Italy, Portugal, and Russia) yields centered on the respective drought events; Data are presented as mean normalized wheat yield values across all identified disaster events, with shaded bands showing the 25th–75th percentile ranges across events, and legends for the mean event impacts.

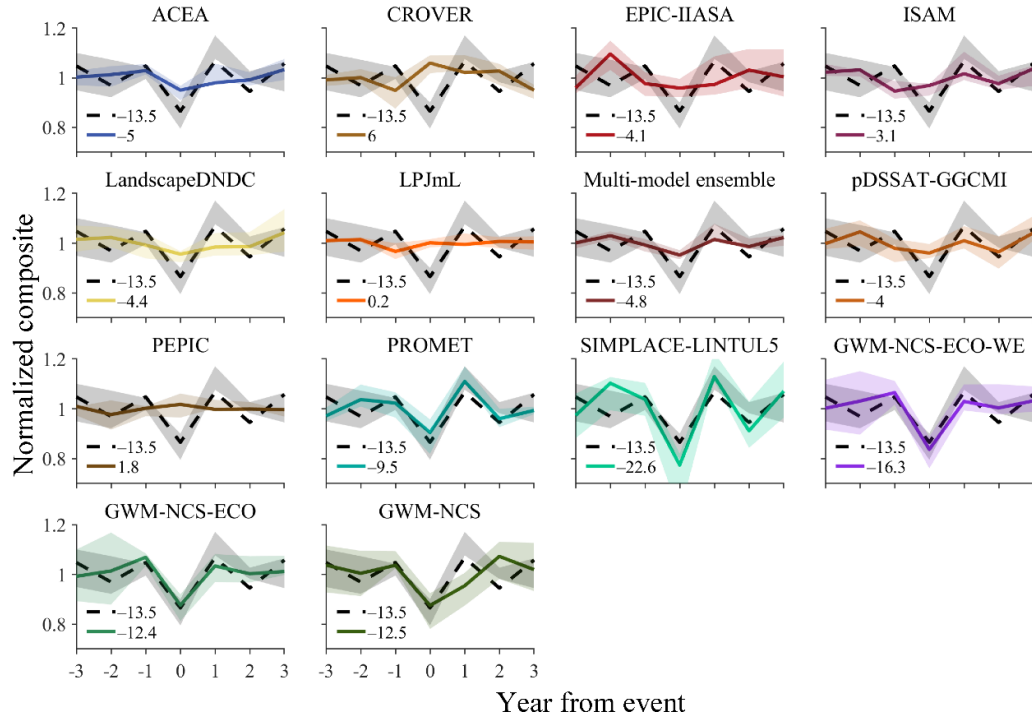


Fig.S31: Observed and simulated impacts of 2003 European heatwave on wheat yields from GGCM Phase 3 and ours. Composites are based on 7-year time windows of country-level (10 countries listed in EM-DAT: Austria, Croatia, Czech Republic, United Kingdom, France, Germany, Italy, Portugal, Slovakia, Spain) yields centered on the respective drought events; Data are presented as mean normalized wheat yield values across all identified disaster events, with shaded bands showing the 25th–75th percentile ranges across events, and legends for the mean event impacts.

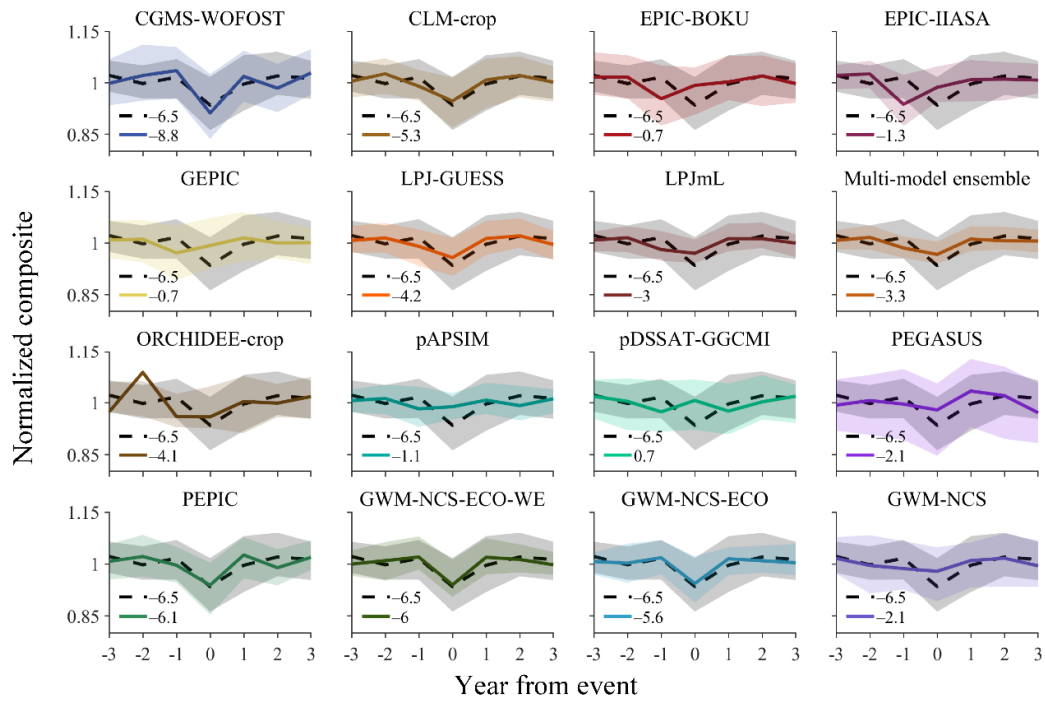


Fig.S32: Observed and simulated impacts of droughts on global wheat yields from GGCM Phase 1 and ours. Composites are based on 7-year time windows of country-level yields centered on the respective drought events; Data are presented as mean normalized wheat yield values across all identified disaster events, with shaded bands showing the 25th–75th percentile ranges across events, and legends for the mean event impacts.

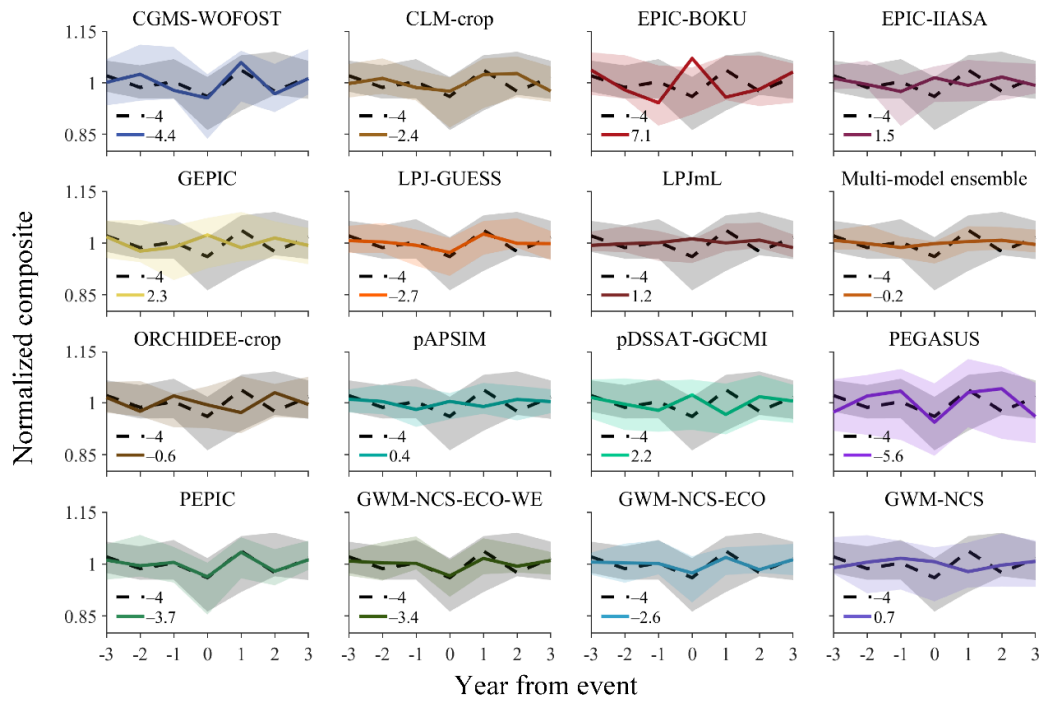


Fig.S33: Observed and simulated impacts of heatwaves on global wheat yields from GGCM1 Phase 1 and ours. Composites are based on 7-year time windows of country-level yields centered on the respective heatwave events; Data are presented as mean normalized wheat yield values across all identified disaster events, with shaded bands showing the 25th–75th percentile ranges across events, and legends for the mean event impacts.

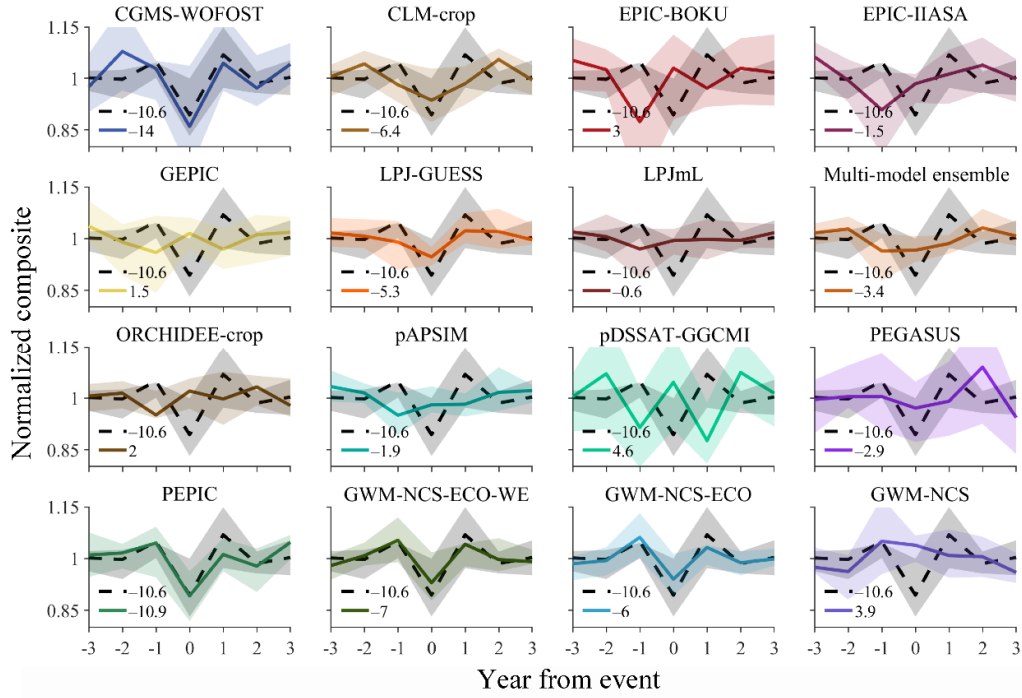


Fig.S34: Observed and simulated impacts of CHDE on global wheat yields from GGCM Phase 1 and ours. Composites are based on 7-year time windows of country-level yields centered on the respective CDHE events; Data are presented as mean normalized wheat yield values across all identified disaster events, with shaded bands showing the 25th–75th percentile ranges across events, and legends for the mean event impacts.

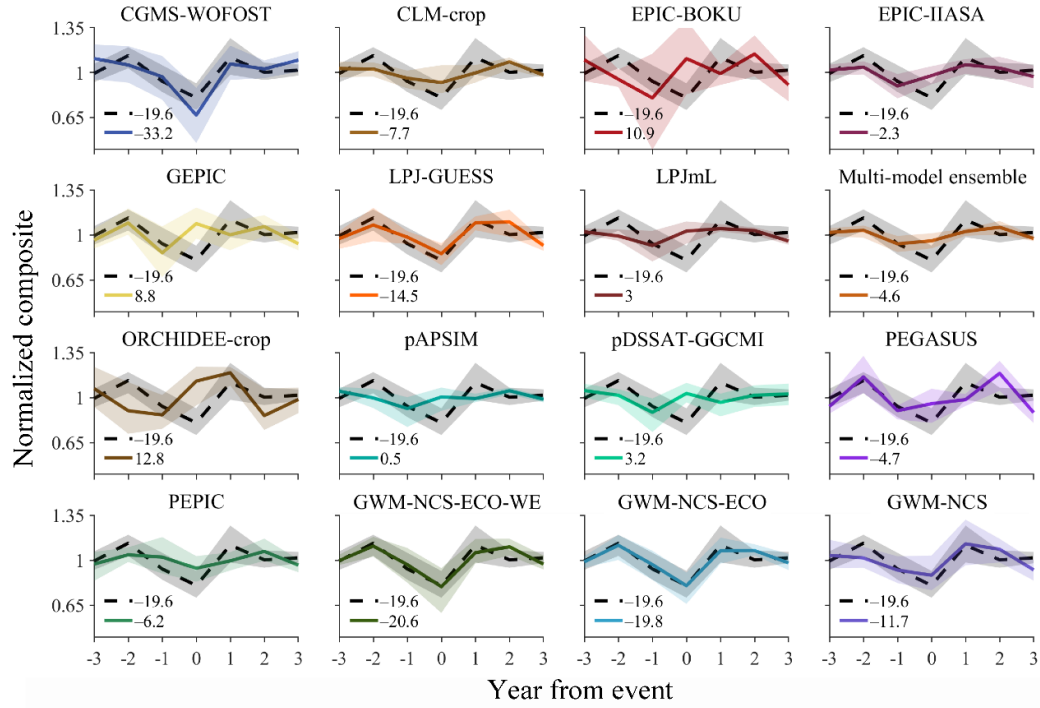


Fig.S35: Observed and simulated impacts of 2003 European drought on wheat yield from GGCM Phase 1 and ours. Composites are based on 7-year time windows of country-level yields centered on the respective drought events (5 countries listed in EM-DAT: Croatia, Hungary, Italy, Portugal, and Russia); Data are presented as mean normalized wheat yield values across all identified disaster events, with shaded bands showing the 25th–75th percentile ranges across events.

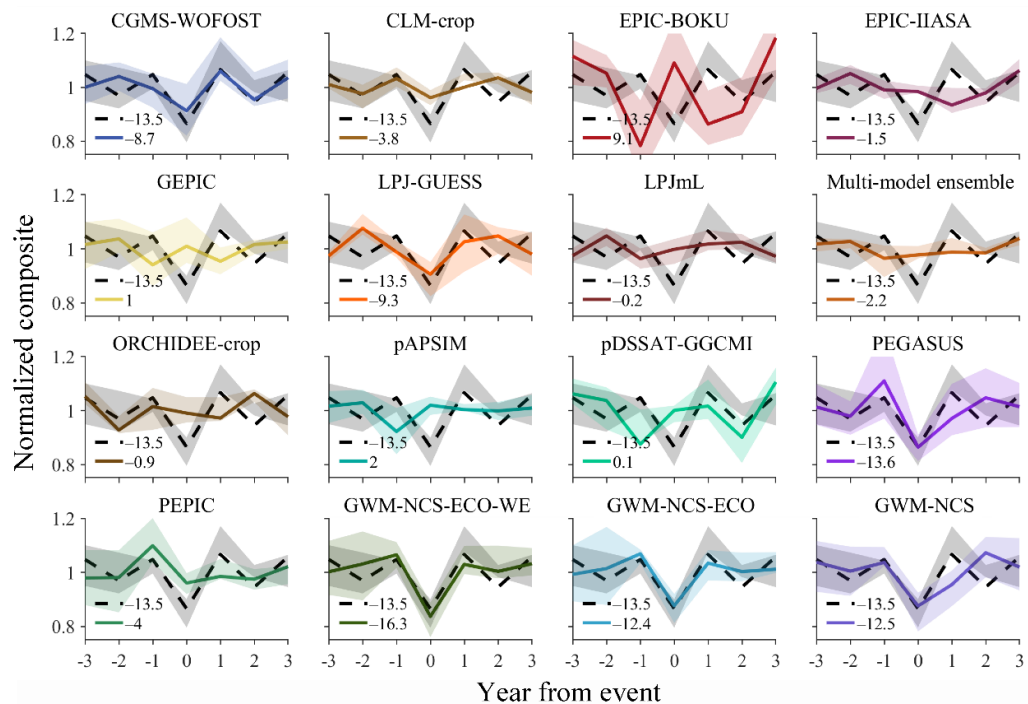


Fig.S36: Observed and simulated impacts of 2003 European heatwave on wheat yield from GGCM Phase 1 and ours. Composites are based on 7-year time windows of country-level yields centered on the respective drought events (10 countries listed in EM-DAT: Austria, Croatia, Czech Republic, United Kingdom, France, Germany, Italy, Portugal, Slovakia, Spain); Data are presented as mean normalized wheat yield values across all identified disaster events, with shaded bands showing the 25th–75th percentile ranges across events, and legends for the mean event impacts.

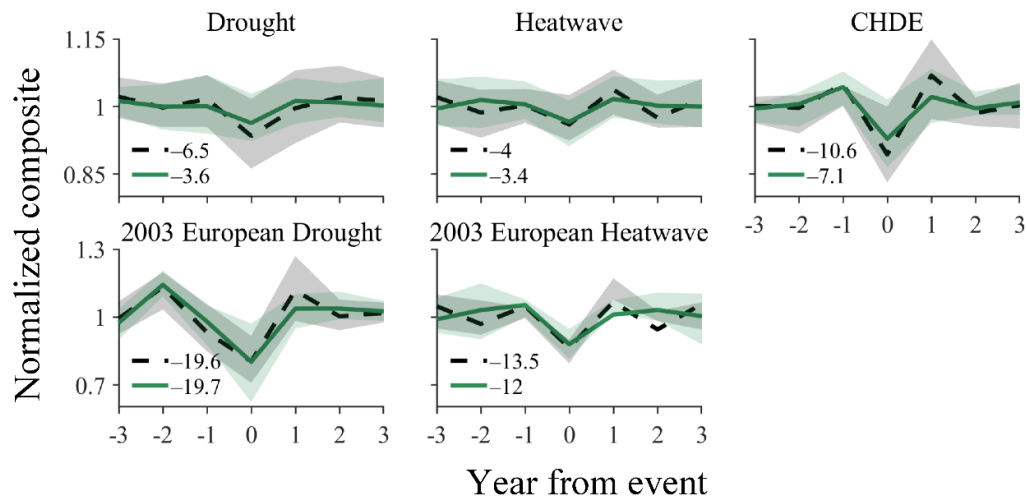


Fig.S37: Global observed and simulated impacts of droughts, heatwaves, CHDEs, 2003 European drought and heatwave on wheat. Normalized composites of wheat yield at 0.5° resolution were from the FAO statistics (dash line) and GWM-NCS-ECO calibrated by GlobalCropYield5min dataset (solid line). Data are presented as mean normalized wheat yield values across all identified disaster events, with shaded bands showing the 25th–75th percentile ranges across events, and legends for the mean event impacts.

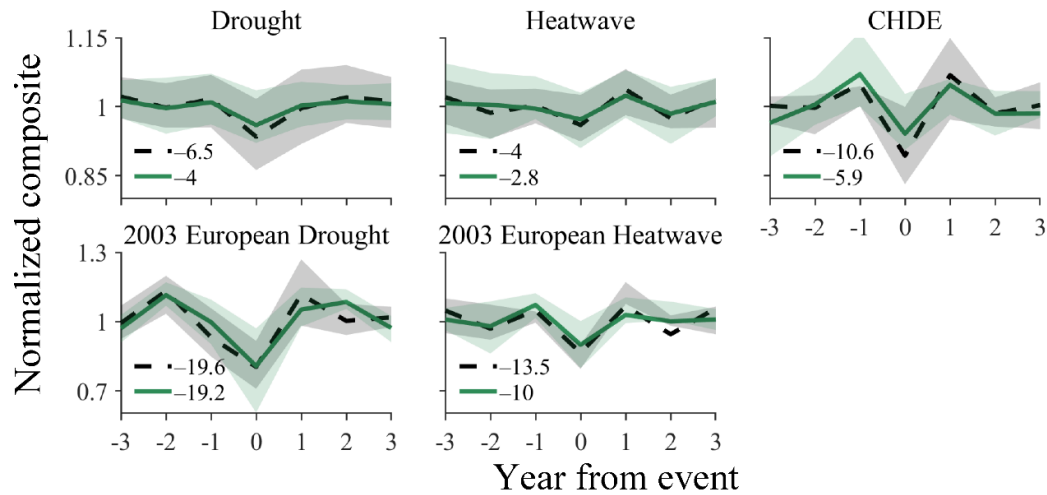


Fig.S38: Observed and simulated yield losses from droughts, heatwaves, CHDEs, 2003 European drought and heatwave at 0.5° resolution. FAO statistical composites are shown as dash lines and GWM-NCS-ECO calibrated with SPAM are in solid line. Data are presented as mean normalized wheat yield values across all identified disaster events, with shaded bands showing the 25th–75th percentile ranges across events, and legends for the mean event impacts.

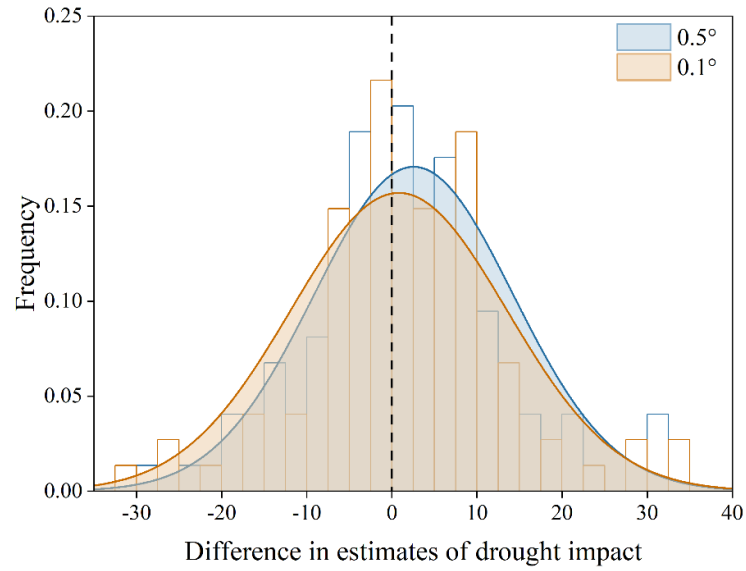


Fig.S39: Comparison of drought impact estimation errors across spatial resolutions in GWM-NCS-ECO-WE. The differences are calculated as estimates of 0.1°- and 0.5°- resolutions minus observes, respectively. Yield composites are within 7-year time windows for each drought event in a country.

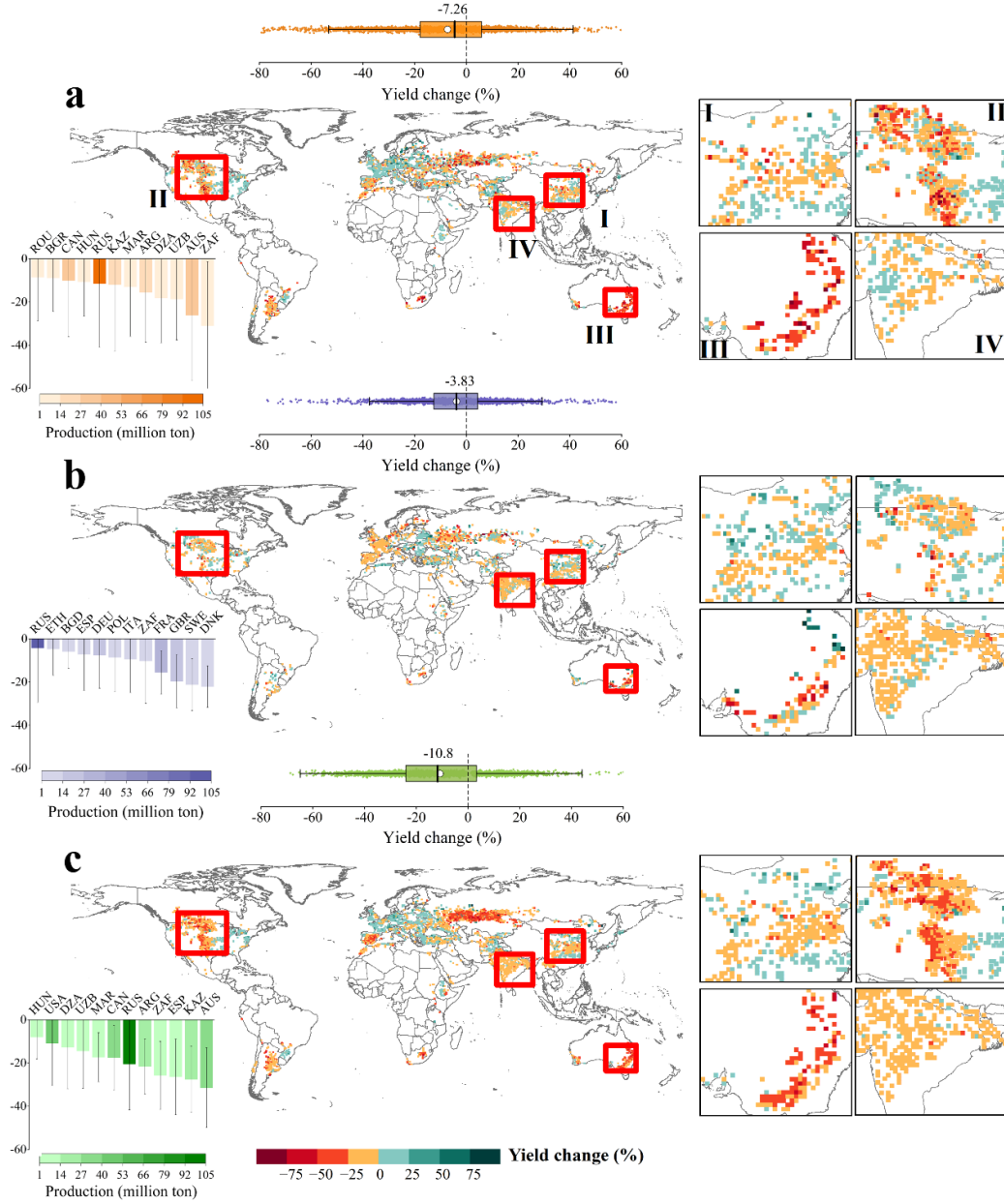


Fig.S40: Yield changes due to droughts (a), heatwaves (b) and CHDE (c) from GWM-NCS-ECO at 0.5° resolution (GlobalCropYield5min calibration). Only statistically significant mean yield changes (1981–2015) are shown in maps. Global distributions are shown as boxplots (box bounds for interquartile range, IQR; whiskers extending up to minimum and maximum values within $1.5 \times \text{IQR}$; white point for mean; center line for median). The sample sizes (n) of independent spatial grid cells are: drought (3,378), heatwaves (2,641) and CHDE (3,624). National averages are shown as bar plots (error bars represent mean values $\pm$ standard error) for the twelve most

affected countries (bar color indicates wheat production), with ISO-alpha3 country codes. The sample sizes (n) of independent spatial grid cells within each country are: droughts (ZAF: 38; AUS: 224; UZB: 48; DZA: 34; ARG: 192; MAR: 49; KAZ: 190; RUS: 902; HUN: 34; CAN: 274; BGR: 32; ROU: 73), heatwaves (DNK: 11; SWE: 21; GBR: 74; FRA: 195; ZAF: 37; ITA: 70; POL: 85; DEU: 127; ESP: 138; BGD: 16; ETH: 57; RUS: 763), and CHDE (AUS: 161; KAZ: 238; ESP: 101; ZAF: 37; ARG: 110; RUS: 546; CAN: 153; MAR: 34; UZB: 31; DZA: 24; USA: 561; HUN: 11).

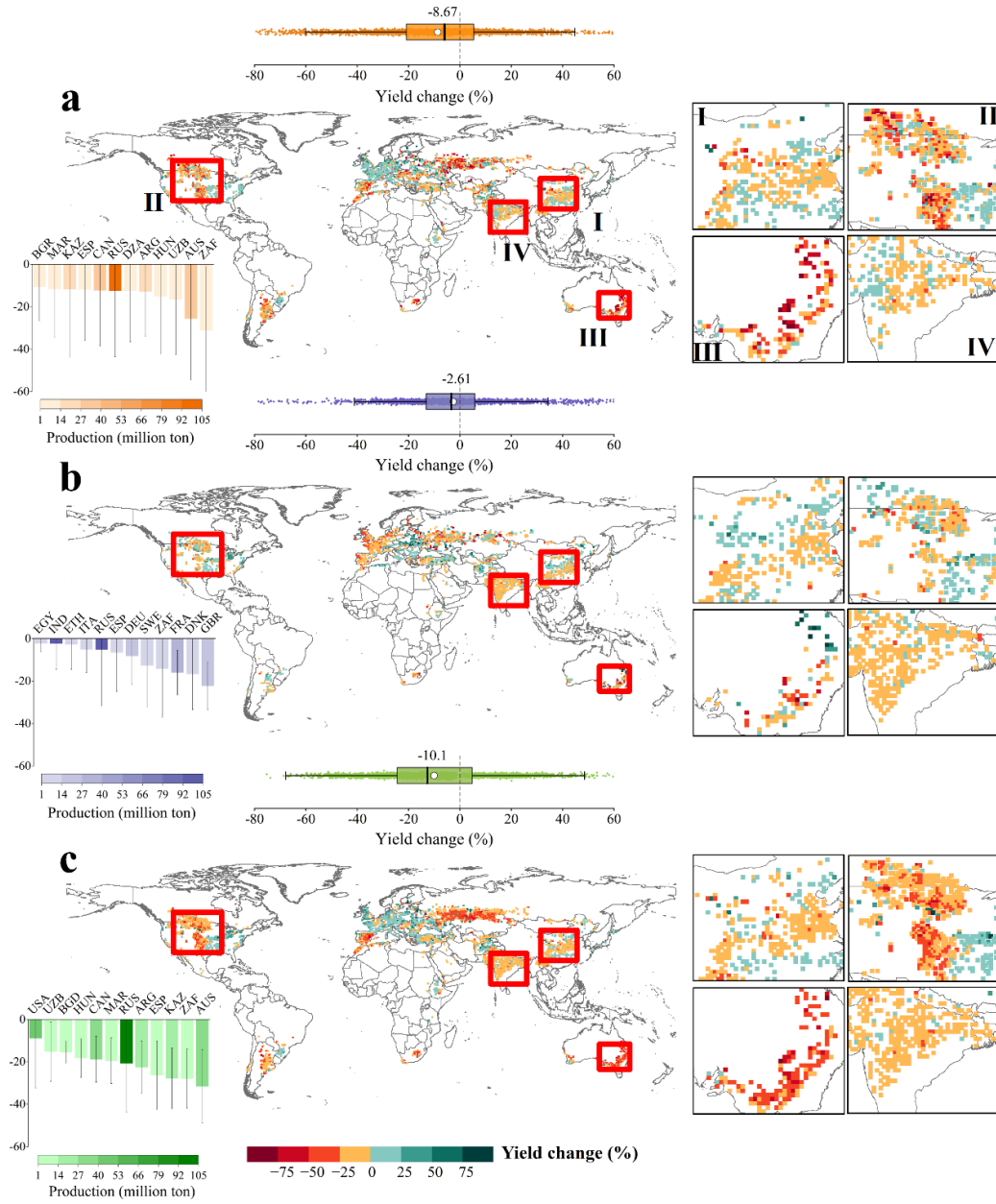


Fig.S41: Yield changes due to droughts (a), heatwaves (b) and CHDE (c) from GWM-NCS-ECO at 0.5° resolution (SPAM calibration). Global distributions are shown as boxplots (box bounds for interquartile range, IQR; whiskers extending up to minimum and maximum values within $1.5 \times \text{IQR}$; white point for mean; center line for median). The sample sizes (n) of independent spatial grid cells are: drought (3,318), heatwaves (2,651) and CHDE (3,693). National averages are shown as bar plots (error bars represent mean values $\pm$ standard error) for the twelve most affected countries (bar color indicates wheat production), with ISO-alpha3 country codes. The sample sizes (n)

of independent spatial grid cells within each country are: droughts (ZAF: 39; AUS: 223; UZB: 40; HUN: 35; ARG: 194; DZA: 34; RUS: 845; CAN: 263; ESP: 135; KAZ: 187; MAR: 48; BGR: 33), heatwaves (GBR: 74; DNK: 11; FRA: 196; ZAF: 34; SWE: 18; DEU: 132; ESP: 138; RUS: 719; ITA: 67; ETH: 51; IND: 513; EGY: 15), and CHDE (AUS: 173; ZAF: 35; KAZ: 232; ESP: 95; ARG: 107; RUS: 518; MAR: 29; CAN: 170; HUN: 9; BGD: 11; UZB: 29; USA: 578).

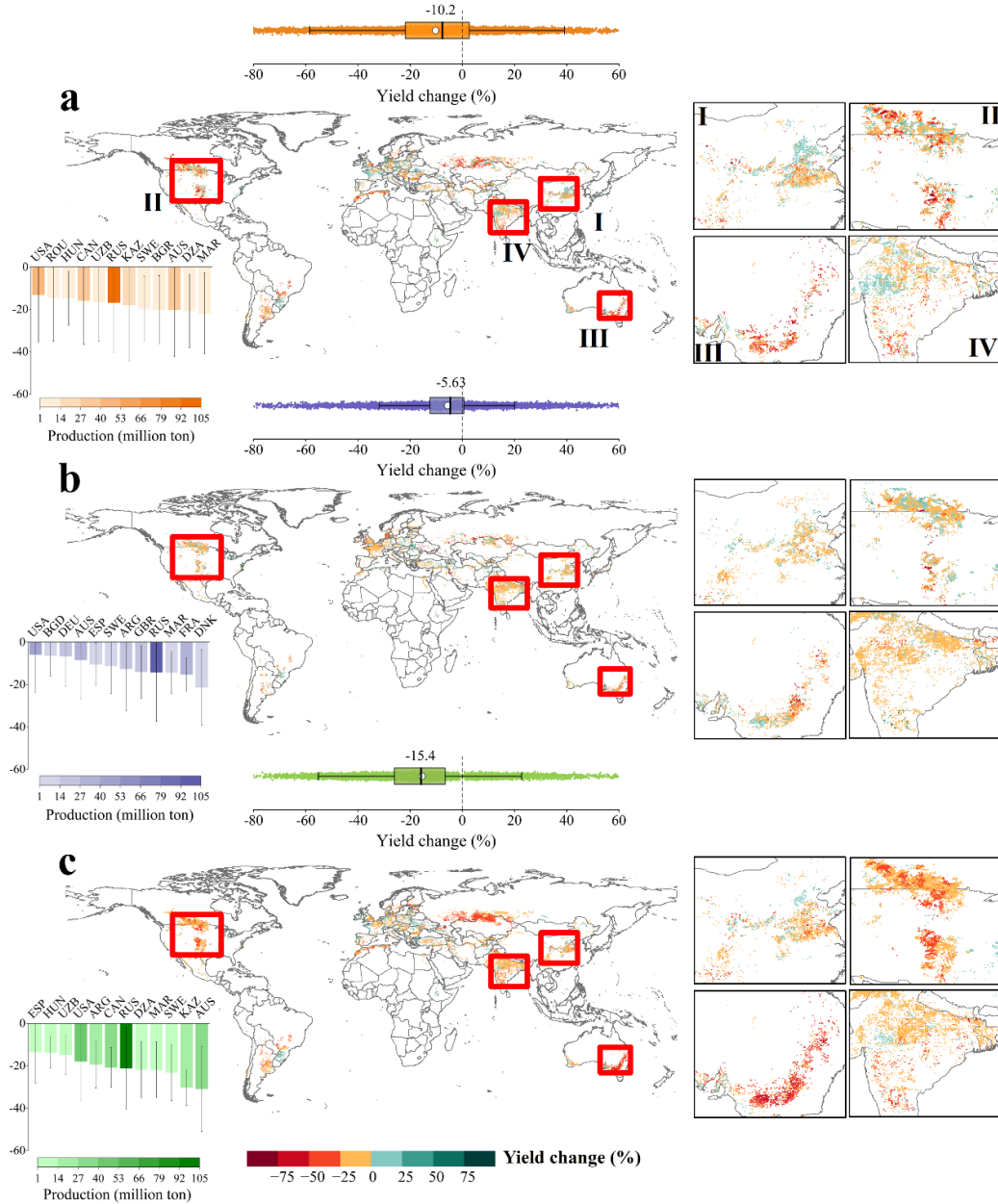


Fig.S42: Yield changes due to droughts (a), heatwaves (b) and CHDE (c) from GWM-NCS-ECO at 0.1° resolution (GlobalWheatYield4km calibration). Global distributions are shown as boxplots (box bounds for interquartile range, IQR; whiskers extending up to minimum and maximum values within $1.5 \times \text{IQR}$; white point for mean; center line for median). The sample sizes (n) of independent spatial grid cells are: drought (31,351), heatwaves (26,565) and CHDE (35,671). National averages are shown as bar plots (error bars represent mean values $\pm$ standard error) for the twelve

most affected countries (bar color indicates wheat production), with ISO-alpha3 country codes. The sample sizes (n) of independent spatial grid cells within each country are: droughts (MAR: 789; DZA: 543; AUS: 1,776; BGR: 112; SWE: 134; KAZ: 1,485; RUS: 2,331; UZB: 606; CAN: 3,361; HUN: 213; ROU: 367; USA: 3,317), heatwaves (DNK: 250; FRA: 1,206; MAR: 206; RUS: 1,282; GBR: 428; ARG: 351; SWE: 87; ESP: 63; AUS: 1,681; DEU: 453; BGD: 302; USA: 2,756), and CHDE (AUS: 2,319; KAZ: 2,553; SWE: 136; MAR: 773; DZA: 484; RUS: 2,863; CAN: 4,202; ARG: 1,009; USA: 3,915; UZB: 561; HUN: 157; ESP: 196).

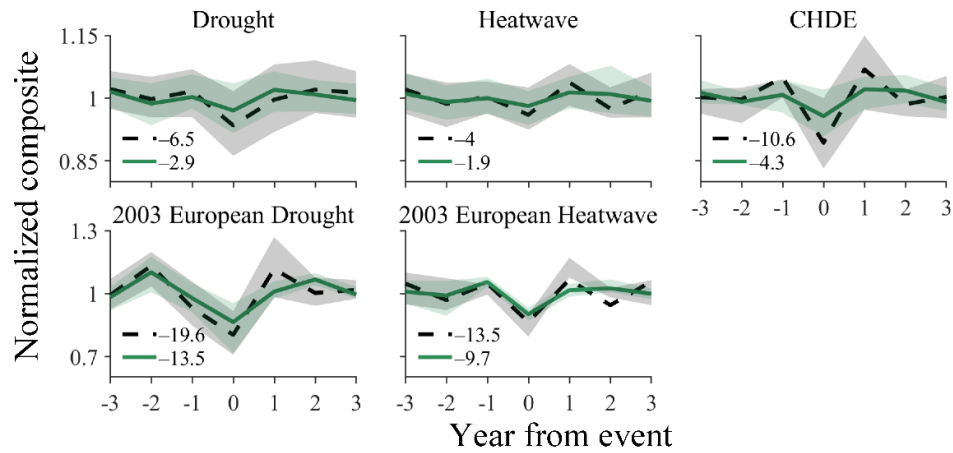


Fig.S43: Extreme weather event impacts simulated by GWM-NCS-ECO-WE calibrated under actual irrigated/rainfed conditions at 0.1° resolution against FAO observations. Shaded areas and legends follow Fig. S37. Data are presented as mean normalized wheat yield values across all identified disaster events, with shaded bands showing the 25th–75th percentile ranges across events, and legends for the mean event impacts.

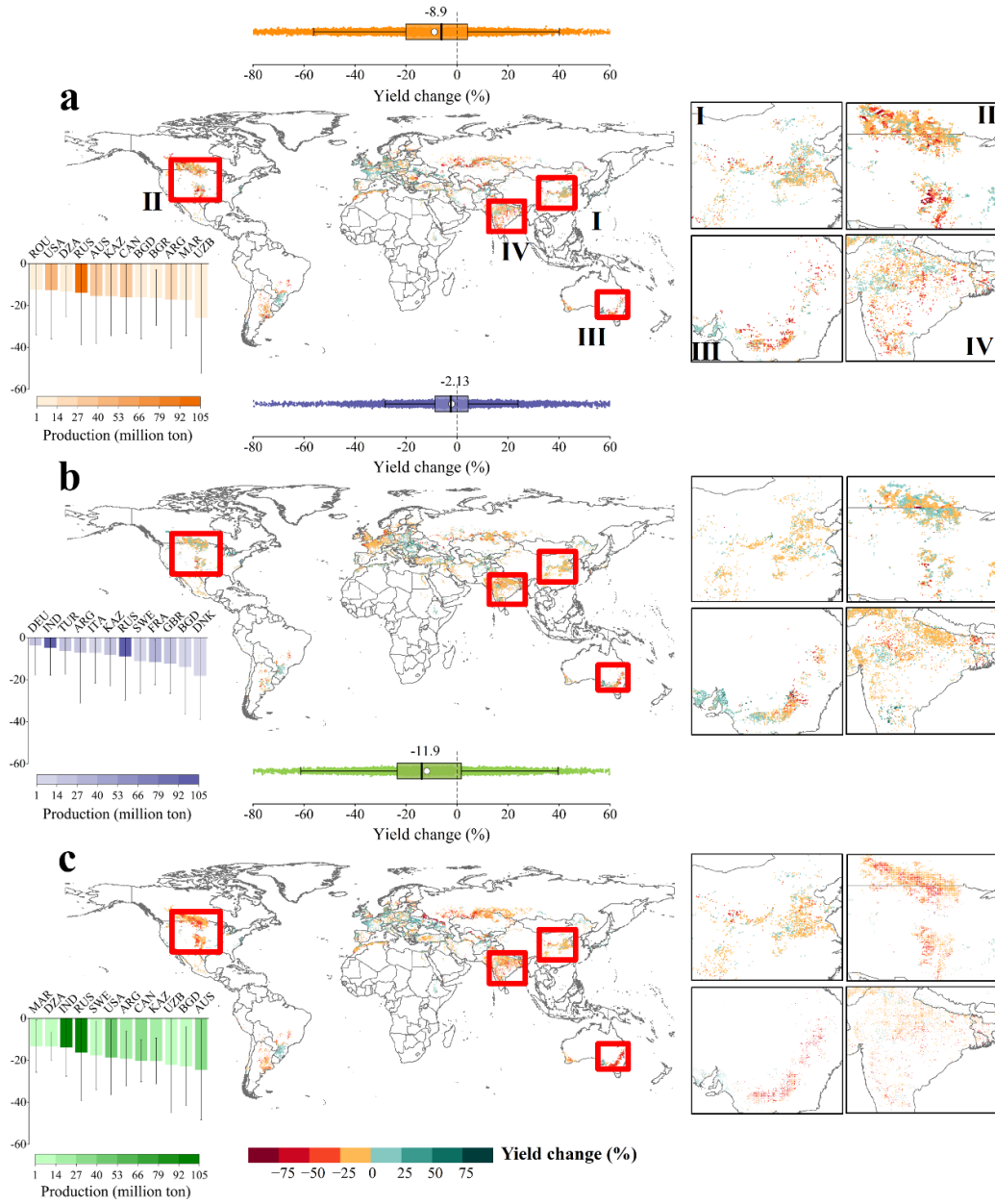


Fig.S44: Statistically significant yield changes (1981–2015) for (a) droughts, (b) heatwaves, and (c) CHDEs at 0.1° resolution simulated by the GWM-NCS-ECO-WE under actual irrigated or rainfed conditions. Global distributions are shown as boxplots (box bounds for interquartile range, IQR; whiskers extending up to minimum and maximum values within $1.5 \times \text{IQR}$; white point for mean; center line for median). The sample sizes (n) of independent spatial grid cells are: drought (30,826), heatwaves (23,087) and CHDE (33,991). National averages are shown as bar plots (error bars represent mean values $\pm$ standard error) for the twelve most affected countries (bar color

indicates wheat production), with ISO-alpha3 country codes. The sample sizes (n) of independent spatial grid cells within each country are: droughts (UZB: 582; MAR: 681; ARG: 1,004; BGR: 168; BGD: 168; CAN: 3,962; KAZ: 1,318; AUS: 1,573; RUS: 2,272; DZA: 455; USA: 3,583; ROU: 270), heatwaves (DNK: 153; BGD: 215; GBR: 264; FRA: 717; SWE: 99; RUS: 1,073; KAZ: 842; ITA: 91; ARG: 290; TUR: 463; IND: 3,273; DEU: 432), and CHDE (AUS: 2,107; BGD: 238; UZB: 623; KAZ: 1,710; CAN: 4,204; ARG: 908; USA: 4,262; SWE: 116; RUS: 2,722; IND: 3,714; DZA: 443; MAR: 722).

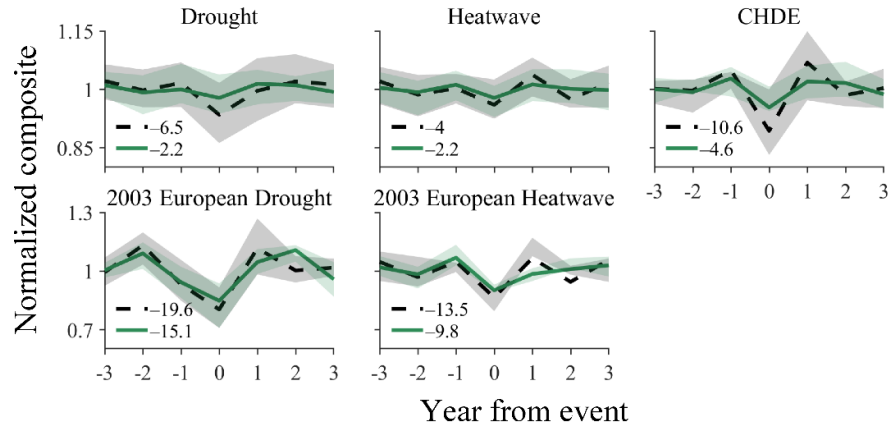


Fig.S45: Extreme weather event impacts simulated by GWM-NCS-ECO-WE calibrated to statistical data at 0.1° resolution against FAO observations. Shaded areas and legends follow Fig. S37. Data are presented as mean normalized wheat yield values across all identified disaster events, with shaded bands showing the 25th–75th percentile ranges across events, and legends for the mean event impacts.

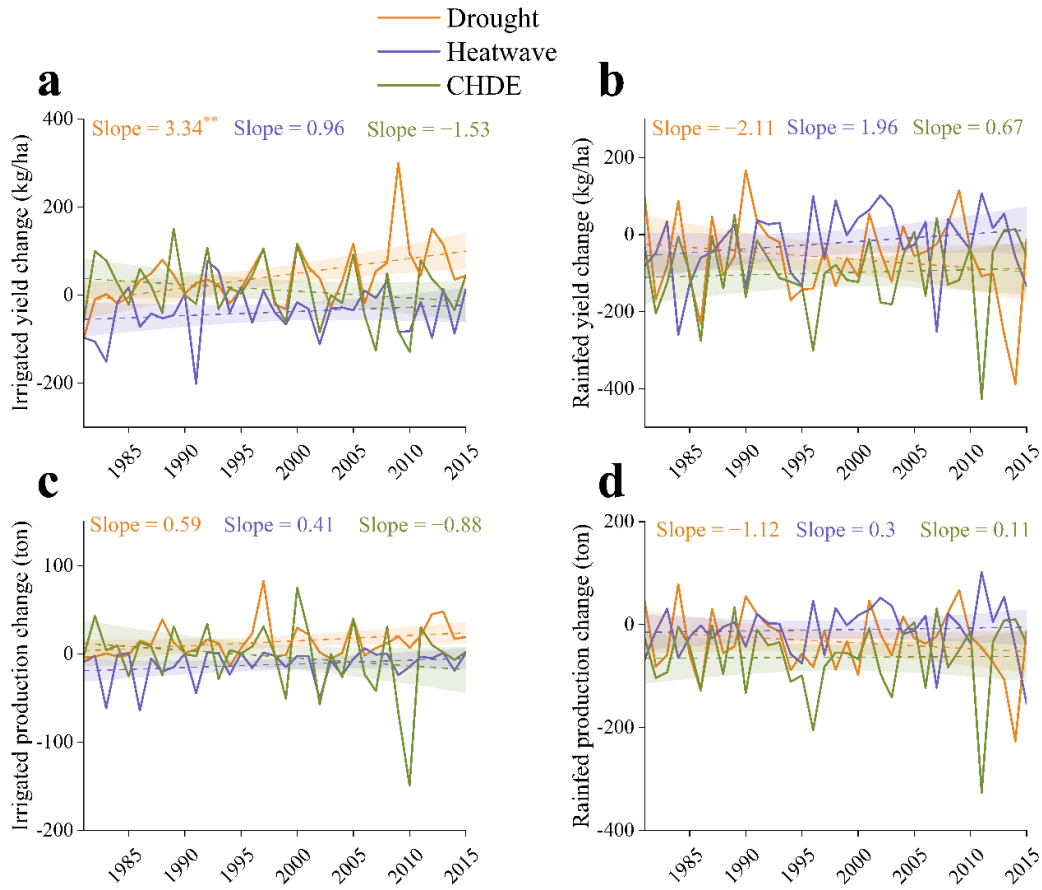


Fig.S46: Changes in impacts of droughts, heatwaves and CHDE. Annual mean yield (a-b) and production (c-d) of affected grid-cell areas caused by droughts, heatwaves and CHDE for irrigated and rainfed systems; Trend lines are based on linear regression with the shaded polygons indicating the 95% confidence intervals, and the significance of slope determined by a two-sided Student's *t* test (** $p < 0.01$; * $p < 0.05$; $^+p < 0.1$, no adjustments for multiple comparisons). Exact *p*-values are as follows: **a** ($p = 0.002, 0.328$, and 0.187 for drought, heatwave, and CHDE); **b** ($p = 0.263, 0.203$, and 0.697 for drought, heatwave, and CHDE); **c** ($p = 0.132, 0.178$, and 0.186 for drought, heatwave, and CHDE); **d** ($p = 0.285, 0.344$, and 0.929 for drought, heatwave, and CHDE).

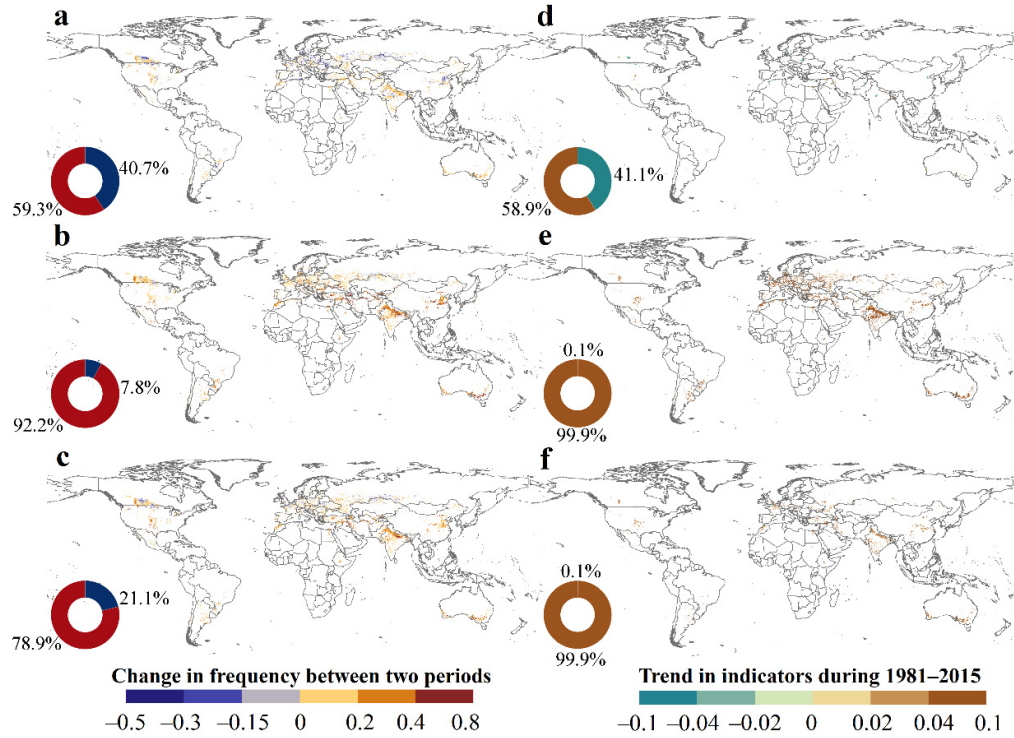


Fig.S47: Global patterns of frequency and intensity of extreme weather events during the study period. Change in frequency of drought (a), heatwave (b), and CHDE (c) between 1981–1998 and 1999–2015, calculated by the differences between two periods; Trend in $-SPI$ (d), STI (e), and $-SCEI$ (f) during 1981–2015; The pie diagrams for the percentages of grids with a significant increase and decrease trend at a significance level of 0.05 (the significance is determined by a two-sided Student' t test with no adjustments for multiple comparisons).

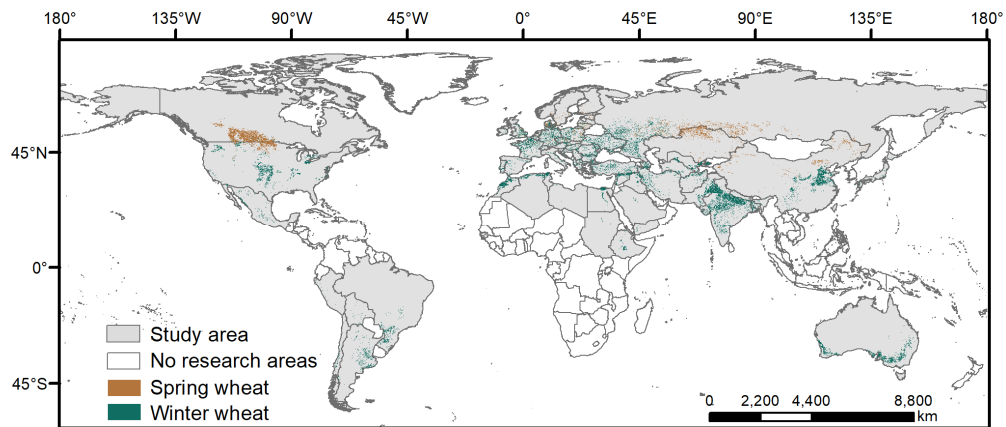


Fig.S48: Global distribution of wheat cultivation areas.

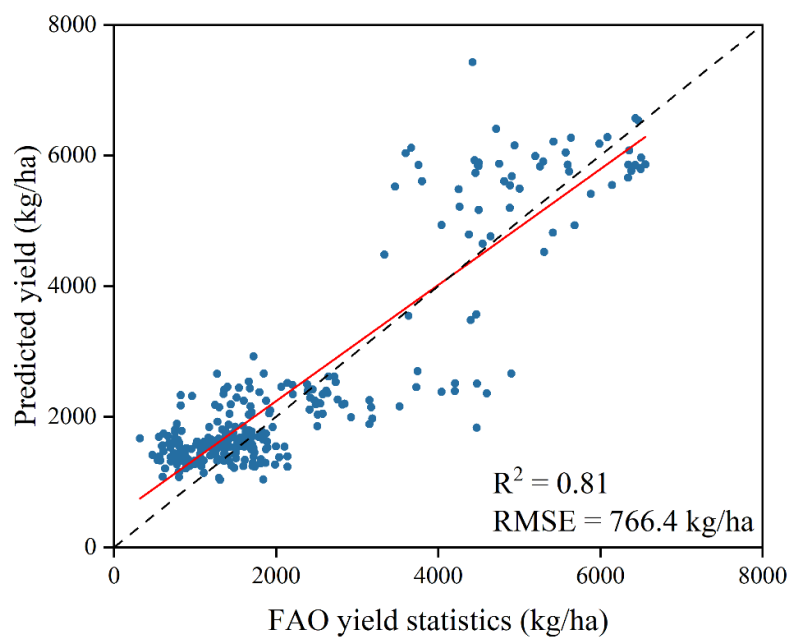


Fig.S49: Accuracy of yield estimations using random forest model in 13 countries.

Countries include Algeria, Egypt, Ethiopia, Iraq, Libya, Morocco, Saudi Arabia, Sudan, Syrian Arab Republic, Tunisia, Turkmenistan, Uzbekistan, and Yemen.

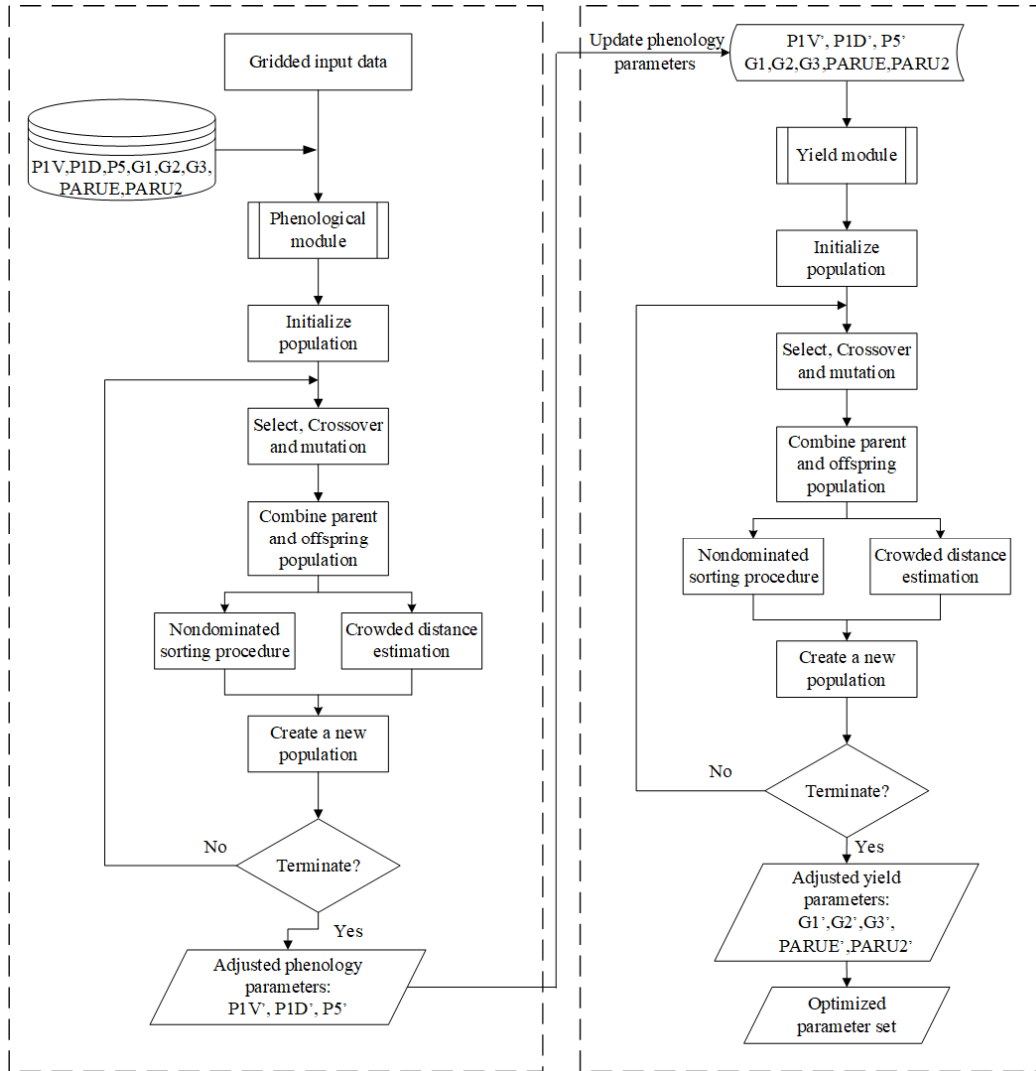


Fig.S50: Schematic of the New Calibration Scheme. Genotype- (P1V, P1D, P5, G1, G2, G3) and ecotype-specific (PARUE, PARU2) parameters were sequentially calibrated to referenced gridded dataset using NSGA-II algorithm.

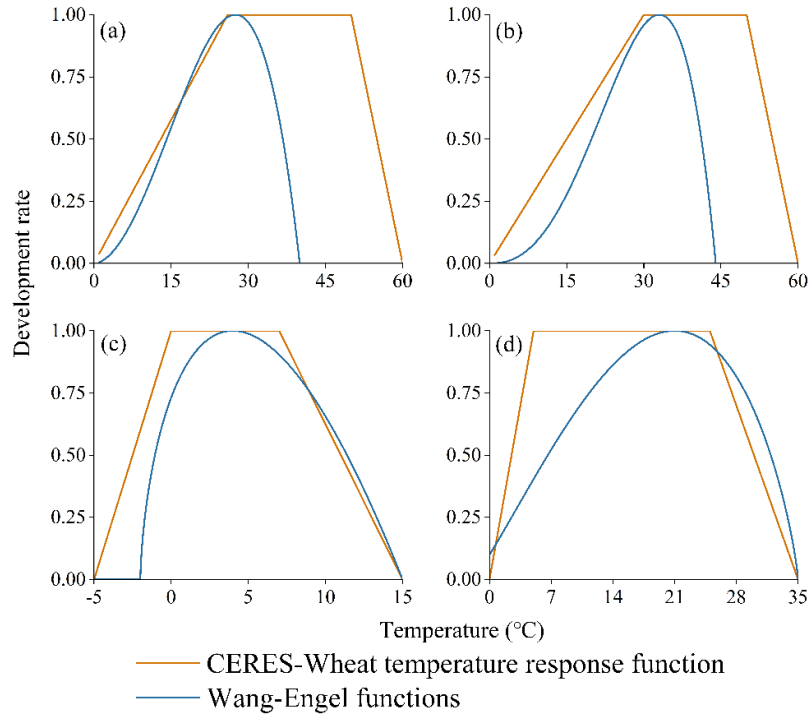


Fig.S51: Comparison of wheat development rate in response to temperature between the original CERES-Wheat and Wang-Engel function for (a) vegetative growth, (b) reproductive growth, (c) vernalization, and (d) photosynthesis.

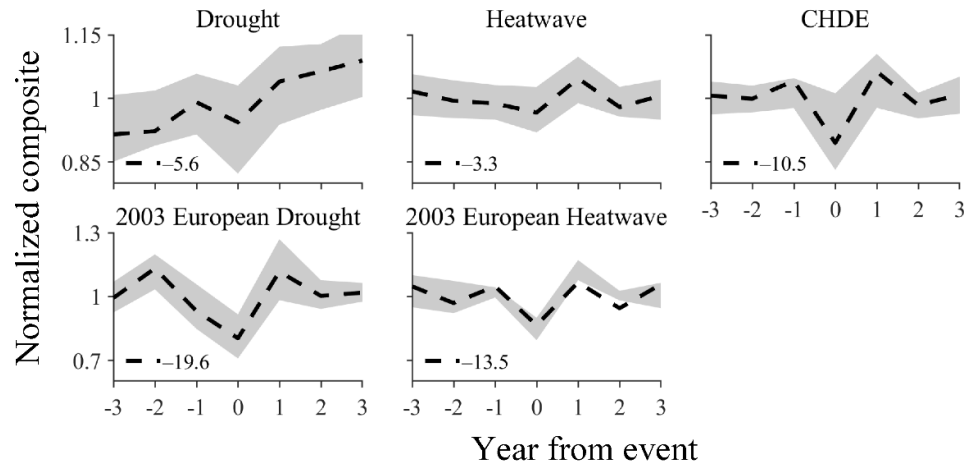


Fig.S52: Extreme event impacts from raw FAO statistics. Shaded areas and legends follow Fig. S37.

Table S1: Three categories of key parameters in the CERES-Wheat model.

Parameter	Definition	Unit	Ranges	Control growth process
Cultivar coefficients				
P1V	Days required to complete vernalization at optimum temperature	Days	0-60	Phenology
P1D	Photoperiod sensitivity coefficient	%	0-200	
P5	Thermal time from the onset of linear fill to maturity	°C • day	100-999	
G1	Kernel number per unit canopy weight at anthesis	kernels • g ⁻¹	10-50	Yield
G2	Standard kernel size under optimum conditions	Mg	10-80	
G3	Standard, non-stressed dry weight of a single tiller at maturity	g	0.5-8	
Ecotype coefficients				
PARUE	Conversion rate from PAR to dry matter before the end of leaf growth	g • MJ ⁻¹	1-6	Yield
PARU2	Conversion rate from PAR to dry matter after the end of leaf growth	g • MJ ⁻¹	1-6	
Species coefficients				
TRDV1	Temperature response, vegetative growth period	°C	(0, 27.5, 40) ¹	Physiological responses to temperature
TRDV2	Temperature response, reproductive growth period	°C	(0, 33, 44) ¹	
TRPHS	Temperature response, photosynthesis	°C	(-2, 21, 35) ²	
TRVRN	Temperature response, vernalization	°C	(-2, 4, 15) ¹	

Note:

¹ Temperature response follows Wang-Engel functions with shape factor (β) of 1 and base temperature (T_b), optimum temperature (T_o), and maximum temperature (T_c) corresponding to the values in the parentheses.

² Temperature response follows Wang-Engel functions with shape factor (β) of 0.8 and T_b , T_o , and T_c corresponding to the values in the parentheses.

Table S2: Accuracy of wheat yield simulations from different model versions of GGCM Phase 3 and ours.

Models	Subnational scale			National scale		
	R	CRMSD (t/ha)	SDEVR	R	CRMSD (t/ha)	SDEVR
GWM-NCS-ECO-WE	0.74	1.24	1.17	0.83	1.17	1.1
GWM-NCS-ECO	0.74	1.34	1.29	0.84	1.29	1.31
GWM-NCS	0.67	1.74	1.56	0.73	1.74	1.33
ACEA	0.01	2.34	1.3	0.17	2.29	0.97
CROVER	0.23	1.76	1	0.65	1.42	0.81
EPIC-IIASA	0.27	1.54	0.76	0.63	1.42	0.75
LandscapeDNDC	0.26	1.63	0.87	0.54	1.56	0.74
LPJmL	0.34	1.74	1.12	0.66	1.65	1.2
PEPIC	0.07	1.74	0.77	0.31	1.81	0.64
PROMET	0.28	1.49	0.7	0.54	1.54	0.66
SIMPLACE-LINTUL5	0.28	2.03	1.34	0.29	2.04	0.89
ISAM	0.31	2.21	1.53	0.67	1.88	1.4
pDSSAT-GGCM	0.6	1.29	1.01	0.84	0.98	0.9
Multi-model ensemble	0.39	1.39	0.71	0.7	1.29	0.69

Note: Correlation coefficients (R), centered root mean squared difference (CRMSD) and the standard deviation relative to that of the observations (SDEVR). A perfect match with official statistics would be having a SDEV of 1 and a R of 1 as well as a CRMSD of 0.

Table S3: Accuracy of wheat yield simulations from different model versions of GGCM Phase 1 and ours.

Models	Subnational scale			National scale		
	R	CRMSD (t/ha)	SDEVR	R	CRMSD (t/ha)	SDEVR
GWM-NCS-ECO-WE	0.74	1.24	1.17	0.83	1.17	1.1
GWM-NCS-ECO	0.74	1.34	1.29	0.84	1.29	1.31
GWM-NCS	0.67	1.74	1.56	0.73	1.74	1.33
GEPIC	0.2	1.51	0.61	0.69	1.3	0.62
LPJmL	0.29	2.1	1.4	0.7	1.43	1.06
CGMS-WOFOST	0.27	2.3	1.55	0.5	1.97	1.18
CLM-crop	0.2	1.83	1.03	0.51	1.73	0.95
EPIC-BOKU	0.1	2.18	1.06	0.3	1.85	0.7
EPIC-IIASA	0.21	1.67	0.86	0.73	1.24	0.86
LPJ-GUESS	0.09	2.02	1.1	0.18	1.88	0.54
ORCHIDEE-crop	0.33	1.67	0.99	0.61	1.45	0.77
pAPSIM	0.28	1.93	1.23	0.49	1.71	0.86
pDSSAT-GGCM	0.23	3.43	2.52	0.54	2.24	1.49
PEGASUS	0.18	2.35	1.51	0.6	1.53	0.89
PEPIC	0.1	1.57	0.58	0.58	1.47	0.54
Multi-model ensemble	0.3	1.47	0.7	0.76	1.31	0.58

Note: The same as Table S2.

Table S4: Summarization for information on subnational statistics.

Country	Wheat type	Political unit of statistical data (Number)	Time frame	Source
Afghanistan	Winter wheat	ADM1 (34)	2011-2017	CountrySTAT, http://afghanistan.countrystat.org/
Albania	Winter wheat	ADM0 (1)	1987-2017	EUROSTAT, https://ec.europa.eu/eurostat/web/main
Argentina	Winter wheat	ADM2 (503)	1981-2019	Ministerio de Agroindustria, http://www.siiia.gov.ar/
Australia	Winter wheat	Natural Resource Management Region (50) + ADM1 (9)	2005-2018	Australian Bureau of Statistics, http://www.abs.gov.au/
Austria	Winter wheat	ADM1 (9)	1981-2018	EUROSTAT, https://ec.europa.eu/eurostat/web/main
Azerbaijan	Winter wheat	ADM2 (79)	2000-2018	CountrySTAT, http://azerbaijan.countrystat.org/home/en/
Bangladesh	Winter wheat	ADM2 (65)	1981-2014	Bangladesh Bureau of Statistics, https://bbs.portal.gov.bd
Belgium	Winter wheat	ADM1 (11)	1981-2017	EUROSTAT, https://ec.europa.eu/eurostat/web/main
Bolivia	Winter wheat	ADM1 (9)	1984-2019	Instituto Nacional de Estadística, https://www.ine.gob.bo/
Brazil	Winter wheat	ADM2 (5504)	1981-2017	Companhia Nacional de Abastecimento, www.conab.gov.br/index.php
Bulgaria	Winter wheat	ADM1 (6)	1985-2018	EUROSTAT, https://ec.europa.eu/eurostat/web/main
Canada	Winter wheat Spring wheat	ADM2 (82)	1981-2019	Statistics Canada, https://www.statcan.gc.ca/eng/start
Chile	Winter wheat	ADM1 (13)	1981-2019	Oficina de Estudios y Políticas Agrarias, https://www.odepa.gob.cl/

China	Winter wheat Spring wheat	County (2008)	1981- 2015	National Bureau of Statistics of China, http://www.stats.gov.cn
Croatia	Winter wheat	ADM0 (1)	2006- 2018	EUROSTAT, https://ec.europa.eu/eurostat/web/main
Czechia	Winter wheat	ADM1 (8)	1981- 2017	EUROSTAT, https://ec.europa.eu/eurostat/web/main
Denmark	Winter wheat Spring wheat	ADM1 (5)	1981- 2018	EUROSTAT, https://ec.europa.eu/eurostat/web/main
Estonia	Winter wheat Spring wheat	ADM0 (1)	1987- 2018	EUROSTAT, https://ec.europa.eu/eurostat/web/main
Finland	Spring wheat	ADM1 (5)	1981- 2018	EUROSTAT, https://ec.europa.eu/eurostat/web/main
France	Winter wheat	ADM1 (27)	1981- 2018	EUROSTAT, https://ec.europa.eu/eurostat/web/main
Germany	Winter wheat	ADM1 (38)	1981- 2018	EUROSTAT, https://ec.europa.eu/eurostat/web/main
Greece	Winter wheat	ADM1 (13)	1981- 2018	EUROSTAT, https://ec.europa.eu/eurostat/web/main
Hungary	Winter wheat	ADM1 (19)	1981- 2017	EUROSTAT, https://ec.europa.eu/eurostat/web/main
India	Winter wheat	ADM2 (506)	1997- 2015	Department of Planning, Economics & Statistics, https://data.gov.in/
Iran	Winter wheat	ADM1 (31)	2004- 2020	Iran's Agricultural Ministry, www.maj.ir
Ireland	Winter wheat	ADM1 (3)	2017- 2018	EUROSTAT, https://ec.europa.eu/eurostat/web/main
Italy	Winter wheat	ADM1 (21)	1981- 2018	EUROSTAT, https://ec.europa.eu/eurostat/web/main
Japan	Winter wheat	ADM1 (47)	1981- 2018	Statistics of Japan, https://www.e-stat.go.jp/
Kazakhstan	Spring wheat	ADM1 (14)	2014- 2020	Ministry of National economy of the Republic of Kazakhstan Committee on Statistics, https://stat.gov.kz/

Latvia	Winter wheat Spring wheat	ADM0 (1)	1987- 2018	EUROSTAT, https://ec.europa.eu/eurostat/web/main
Lithuania	Winter wheat Spring wheat	ADM1 (2)	2017- 2018	EUROSTAT, https://ec.europa.eu/eurostat/web/main
Luxembourg	Winter wheat	ADM0 (1)	1981- 2018	EUROSTAT, https://ec.europa.eu/eurostat/web/main
Macedonia	Winter wheat	ADM0 (1)	2005- 2018	EUROSTAT, https://ec.europa.eu/eurostat/web/main
Mexico	Winter wheat	ADM2 (2442)	2003- 2018	Servicio de Informacion Agroalimentaria y Pesquera, http://infosiap.siap.gob.mx/
Nepal	Winter wheat	ADM2 (75)	1981- 2018	NEPAL IN DATA, https://nepalindata.com/
Netherlands	Winter wheat	ADM1 (12)	1981- 2018	EUROSTAT, https://ec.europa.eu/eurostat/web/main
Norway	Spring wheat	ADM1 (7)	2015- 2017	EUROSTAT, https://ec.europa.eu/eurostat/web/main
Pakistan	Winter wheat	ADM3 (113)	1981- 2014	Directorate of Agriculture Punjab, http://www.amis.pk/
Peru	Winter wheat	ADM1 (26)	1981- 2018	Ministerio de Agricultura, https://www.gob.pe/
Poland	Winter wheat	ADM1 (17)	1995- 2016	EUROSTAT, https://ec.europa.eu/eurostat/web/main
Portugal	Winter wheat	ADM1 (7)	1999- 2015	EUROSTAT, https://ec.europa.eu/eurostat/web/main
Republic of Moldova	Winter wheat	ADM1 (37)	2007- 2018	National Bureau of Statistics of the Republic of Moldova, https://statbank.statistica.md/PxWeb/pxweb/en/
Romania	Winter wheat	ADM1 (7)	1995- 2017	EUROSTAT, https://ec.europa.eu/eurostat/web/main
Russia Federation	Winter wheat Spring wheat	ADM1 (62)	1990- 2016	Федеральная служба государственной статистики, https://data.gov.ru/?language=en
Serbia	Winter wheat	ADM1 (4)	2017- 2018	EUROSTAT, https://ec.europa.eu/eurostat/web/main

Slovakia	Winter wheat	ADM1 (4)	1995-2017	EUROSTAT, https://ec.europa.eu/eurostat/web/main
Slovenia	Winter wheat	ADM1 (2)	2007-2017	EUROSTAT, https://ec.europa.eu/eurostat/web/main
Spain	Winter wheat	ADM1 (18)	1981-2018	EUROSTAT, https://ec.europa.eu/eurostat/web/main
Sweden	Spring wheat	ADM1 (8)	1992-2018	EUROSTAT, https://ec.europa.eu/eurostat/web/main
Switzerland	Winter wheat	ADM1 (7)	2017-2018	EUROSTAT, https://ec.europa.eu/eurostat/web/main
Turkey	Winter wheat	ADM1 (26)	1995-2016	EUROSTAT, https://ec.europa.eu/eurostat/web/main
Ukraine	Winter wheat	ADM1 (27)	2004-2017	EUROSTAT, https://ec.europa.eu/eurostat/web/main
United Kingdom	Winter wheat	ADM1 (40)	1987-2018	EUROSTAT, https://ec.europa.eu/eurostat/web/main
United States	Winter wheat Spring wheat	ADM2 (2846)	1981-2016	USDA, https://quickstats.nass.usda.gov/

Table S5: The details of temporal configuration of FAO data.

Country	The direction of shift	R ² improvement
Australia	Backward	221.8%
Brazil	Backward	66.3%
Chile	Backward	91.9%
China	Forward	81.1%
Russia	Forward	136.1%
Bolivia	Backward	85.5%
Macedonia	Backward	113.1%

Note: The relative improvement in R² was calculated as the difference between improved and original R² divided by the original R², where R² indicated the explained variance between simulated and observed yield anomaly.

Table S6: Full statistical details for the agreement between statistical and simulated yields by three crop models.

Crop model	GGCMI-MME		GWM-NCS		GWM-NCS-ECO	
	National	Subnational	National	Subnational	National	Subnational
N	3024	106296	1669	82900	1674	83655
Pearson correlation	0.696	0.385	0.728	0.669	0.841	0.739
p	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001

Table S7: Full statistical details for the trend in the affected area caused by three disasters and the irrigation contribution.

	Slope	95% CI	df	t	p-value
DR_AFFE_IR	0.074	[-0.046, 0.193]	33	1.25	0.219
HW_AFFE_IR	0.247	[0.055, 0.438]	33	2.62	0.013
CHDE_AFFE_IR	0.176	[0.036, 0.317]	33	2.55	0.016
DR_AFFE_RF	0.049	[-0.042, 0.141]	33	1.10	0.281
HW_AFFE_RF	0.326	[0.194, 0.458]	33	5.04	<0.001
CHDE_AFFE_RF	0.184	[0.074, 0.294]	33	3.39	0.002
DR_CONT_Y	-0.87	[-4.06, 2.31]	33	-0.56	0.581
HW_CONT_Y	-3.53	[-6.78, -0.29]	33	-2.21	0.034
CHDE_CONT_Y	-2.93	[-6.29, 0.44]	33	-1.77	0.086
DR_CONT_P	0.43	[-4.5, 5.36]	33	0.18	0.859
HW_CONT_P	-6.23	[-14.29, 1.82]	33	-1.58	0.125
CHDE_CONT_P	-3.45	[-8.61, 1.72]	33	-1.36	0.184

Note: DR: Drought; HW: Heatwave; CHDE: Compound heatwave and drought events; AFFE: Annual sums of affected grid-cell areas due to disasters; IR: Irrigated systems; RF: Rainfed systems; CONT: The irrigation contribution to alleviating yield (or production) losses by the three types of EWEs; Y: Yield; P: Production.

References:

- Anderson, W. et al. Climate variability and simultaneous breadbasket yield shocks as observed in long-term yield records. *Agric. For. Meteorol.* **331**, 109321, doi: 10.1016/j.agrformet.2023.109321 (2023).
- Becker-Reshef, I. et al. Crop Type Maps for Operational Global Agricultural Monitoring. *Sci. Data* **10**, 172, doi: 10.1038/s41597-023-02047-9 (2025).
- Cao, J. et al. Mapping global yields of four major crops at 5-minute resolution from 1982 to 2015 using multi-source data and machine learning. *Sci. Data* **12**, 357, doi: 10.1038/s41597-025-04650-4 (2025).
- Cao, S. et al. Spatiotemporally consistent global dataset of the GIMMS leaf area index (GIMMS LAI4g) from 1982 to 2020. *Earth Syst. Sci. Data* **15**, 4877-4899, doi:10.5194/essd-15-4877-2023 (2023).
- He, Y. et al. Global warming determines future increase in compound dry and hot days within wheat growing seasons worldwide. *Clim. Change* **177**, 70, doi: 10.1007/s10584-024-03718-1 (2024).
- Jägermeyr, J. & Frieler, K. Spatial variations in crop growing seasons pivotal to reproduce global fluctuations in maize and wheat yields. *Sci. Adv.* **4**, eaat4517, doi: 10.1126/sciadv.aat4517 (2018).
- Jägermeyr, J. et al. Climate impacts on global agriculture emerge earlier in new generation of climate and crop models. *Nat. Food* **2**, 873–885 (2021).
- Jones, J. W. et al. The DSSAT cropping system model. *Eur. J. Agron.* **18**, 235-265 (2003).
- Liu, Y., Liu, R. G. & Chen, J. M. Retrospective retrieval of long-term consistent global leaf area index (1981-2011) from combined AVHRR and MODIS data. *J. Geophys. Res.-Biogeosci.* **117**, G04003, doi:10.1029/2012JG002084 (2012).
- Lobell, D. B., Sibley, A. & Ortiz-Monasterio, J. I. Extreme heat effects on wheat senescence in India. *Nat. Clim. Change* **2**, 186-189, doi:10.1038/Nclimate1356 (2012).
- Parent, B., Turc, O., Gibon, Y., Stitt, M. & Tardieu, F. Modelling temperature-compensated physiological rates, based on the co-ordination of responses to temperature of developmental processes. *J. Exp. Bot.* **61**, 2057-2069 (2010).
- Troy, T. J., Kipgen, C. & Pal, I. The impact of climate extremes and irrigation on US crop yields. *Environ. Res. Lett.* **10**, 054013 (2015).
- Vogel, E. et al. The effects of climate extremes on global agricultural yields. *Environ. Res. Lett.* **14**, 054010 (2019).
- Wang, E. et al. The uncertainty of crop yield projections is reduced by improved temperature response functions. *Nat. Plants* **3**, 17102, doi: 10.1038/nplants.2017.102 (2017).
- Xiao, Z. Q. et al. Long-Time-Series Global Land Surface Satellite Leaf Area Index Product Derived From MODIS and AVHRR Surface Reflectance. *IEEE Trans. Geosci. Remote Sensing* **54**, 5301-5318, doi:10.1109/Tgrs.2016.2560522 (2016).
- Xie, W. et al. Decreases in global beer supply due to extreme drought and heat. *Nat. Plants* **4**, 964-973, doi: 10.1038/s41477-018-0263-1 (2018).
- Yin, J. B. et al. Future socio-ecosystem productivity threatened by compound drought-

- heatwave events. *Nat. Sustain* **6**, 259-272, doi: 10.1038/s41893-022-01024-1 (2023).
- Zhang, Z., Luo, Y. C., Han, J. C., Xu, J. L. & Tao, F. L. Estimating Global Wheat Yields at 4 km Resolution during 1982-2020 by a Spatiotemporal Transferable Method. *Remote Sens.* **16**, 2342, doi: 10.3390/rs16132342 (2024).
- Zhu, X. & Troy, T. J. Agriculturally Relevant Climate Extremes and Their Trends in the World's Major Growing Regions. *Earth Future* **6**, 656-672, doi: 10.1002/2017ef000687 (2018).