Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

AGROMINDS - Historic and projected global agro-meteorologic indicators for major summer crops

Thomas Oberleitner

oberleitner@iiasa.ac.at

International Institute for Applied Systems Analysis

Nikolay Khabarov

International Institute for Applied Systems Analysis

Christian Folberth

International Institute for Applied Systems Analysis

data-descriptor

Keywords:

Posted Date: August 24th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10591337/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

AGROMINDS - Historic and projected global agro-meteorologic indicators for major summer crops

Thomas Oberleitner*[1], Nikolay Khabarov[1], and Christian Folberth[1]

[1]International Institute for Applied Systems Analysis (IIASA), Laxenburg, Austria

Abstract

AGROMINDS is a global dataset of annually resolved agro-meteorological indicators aggregated over crop-specific growing seasons for maize, soybean, rice, and spring wheat. The dataset combines daily climate forcing with spatially explicit crop calendars and dynamic crop development to derive seasonal summaries of temperature, precipitation, radiation, atmospheric moisture demand, and weather extremes for the full growing season and phenologically defined sub-periods. It covers both historical conditions using reanalysis products and climate projections from five bias-adjusted CMIP6 global climate models (GCMs) under four emission pathways for the period 1971-2100 at a spatial resolution of 0.5° x 0.5°. By aligning temporal aggregation with dynamically estimated crop phenology rather than fixed calendar periods, AGROMINDS provides a more realistic representation of crop-relevant climate conditions across diverse environments. The dataset is designed to support analyses of historical and future climate impacts on agriculture, including crop yield modeling, climate risk assessment, compound event analysis, and model intercomparison studies. Accompanying open-source code enables reproducible generation and customization of the indicators using alternative crops, growing seasons and aggregation schemes.

1 Background & Summary

Agroclimatic indicators are required for a wide range of applications in science and practice, including analyses of historical weather impacts on crop yields and associated externalities [1, 2, 3], estimation of future climate risks [4], and the training of analytical and predictive models for crop yields [5, 6], among others. Broadly, such indicators represent an abstraction of meteorological data over a defined time period. These indicators can vary substantially with respect to (i) the selection of climate variables, (ii) the mode of aggregation or abstraction, and (iii) the temporal reference period. For example, common aggregation methods include sums and means, while threshold exceedances are frequently used to capture extremes [7]. Temporal references can range from calendar units such as months or years to fixed or dynamic crop-specific growing seasons and may be based on transient time steps or long-term climatologies [8, 9]. Equally, climate variables range from standard measures such as temperature and precipitation to derived variables such as potential evapotranspiration (PET).

These methodological choices strongly influence both the information content of resulting estimates and the types of research questions that can be addressed. Long temporal aggregation periods, for instance, are generally unable to capture the impacts of weather variability on crops

*Correspondence: oberleitner@iiasa.ac.at

in individual years. Indicators based on long-term averages primarily reflect typical growing conditions and provide insights on progressing climatologic change, whereas annual extreme indicators are better suited to represent the occurrence and frequency of weather shocks. Finally, the aggregation period can be crucial in reflecting general growth conditions over the year [10], climatic conditions during growing seasons defined by calendar seasons or fixed planting and harvest dates [11], or dynamic growing seasons based on planting dates and crops' temperature accumulation until maturity [8]. The latter is the approach commonly implemented in state-of-the-art cropping system models, and reflects that crops mature earlier in warmer seasons and later in cooler ones [12]. Furthermore, certain impacts depend on specific climatic variables; for example, pest and disease dynamics are often closely linked to relative humidity [13, 14].

Here, we present a new global dataset of agroclimatic indicators, AGROMINDS, tailored to growing seasons of major summer crops. The dataset builds on most recent climate and crop data provided by the Inter-Sectoral Impact Model Intercomparison Project (ISIMIP) and the Global Gridded Crop Model Intercomparison (GGCMI) initiative [15, 16], the two leading initiatives in global crop climate impact studies. AGROMINDS comprises a diverse set of general and extreme-oriented indicators, aggregated over annually varying crop-specific growing (sub-)periods defined by reported planting dates and dynamically estimated crop maturation. The dataset covers both historical and projected periods, including a reanalysis climate dataset and simulations from five GCMs.

Growing season-specific indicators provide a more precise representation of crop-relevant climate conditions but are rarely available at large spatial scales. The dataset presented herein addresses this gap and is designed to enable consistent and reproducible analyses across regions and time periods. It supports a wide range of applications, as illustrated in the Usage Notes.

These include, for example, analyses of global crop yield shocks and multiple breadbasket failures [17, 18, 19, 20], the development of crop yield prediction tools such as statistical or machine learning models [21], broader agroclimatic analyses including crop suitability assessments [22] and climate risk evaluation [23], and the interpretation of crop model simulation studies [8]. The accompanying Python code further facilitates customization of the indicators, for example through the use of regionally specific crop calendars or alternative aggregation schemes.

2 Methods

2.1 Estimation of crop growing seasons and basic crop phenology

Agro-meteorological indicators were calculated on annually resolved, crop-specific growing seasons defined by planting and harvest timing information. The pipeline combines daily climate forcing with spatially explicit crop-calendar and auxiliary information on a $0.5^\circ \times 0.5^\circ$ grid. Crop- and season specific planting days (PD) and harvest days (HD) are sourced from a state-of-the-art global crop calendar dataset [24]. These are used to determine the baseline crop- and location-specific potential heat unit (PHU) requirement, the sum of growing degree days (GDD; see below) over the reported growing season for the baseline climatology 1981-2010. In the subsequent climate indicator estimation, PD and PHU are used for annual dynamic growing season estimation including basic crop phenology parameters 3 required for the estimation of potential evapotranspiration (PET) following the Penman–Monteith approach. The methods for dynamically estimating growing season timing, crop phenology, and PET are largely based on Folberth et al. [9].

For the baseline and for each subsequent season, daily growing degree days are accumulated from planting onward as

$$\mathrm{GDD}_d = \max\left(T_{\mathrm{avg},d} - T_{\mathrm{base}}, 0\right),$$

where $T_{\mathrm{avg},d}$ is daily mean temperature and T_{base} is the crop-specific base temperature. Crop development is tracked using a heat unit index (HUI), defined as cumulative GDD normalized by PHU. Harvest is assigned to the first day on which HUI reaches 1. The simulated harvest date is capped at 21 days after the input calendar-based harvest date to avoid unrealistically long seasons in locations where maturity is not reached.

Crop development parameters follow the default parameterization of the Environmental Policy Integrated Climate (EPIC) model, which has been a popular core model in various global crop-climate impact studies, including all prior GGCMI and ISIMIP phases, and has shown good skill in evaluation [25, 8]. These parameters include the base temperature for heat-unit accumulation and shape parameters governing the evolution of leaf area index (LAI) and crop height (CHT) over the growing season.

After PD and HD are established for each pixel-year, all daily records outside the growing season are excluded from further aggregation. The final dataset therefore represents climate conditions experienced during the reported or dynamically estimated crop season rather than during fixed calendar intervals. All outputs are assigned to the year of planting to avoid inconsistencies that arise when harvest dates near the end of the calendar year extend into the following year for some growing seasons but not others [26]. Figure 1 provides an overview of the workflow.

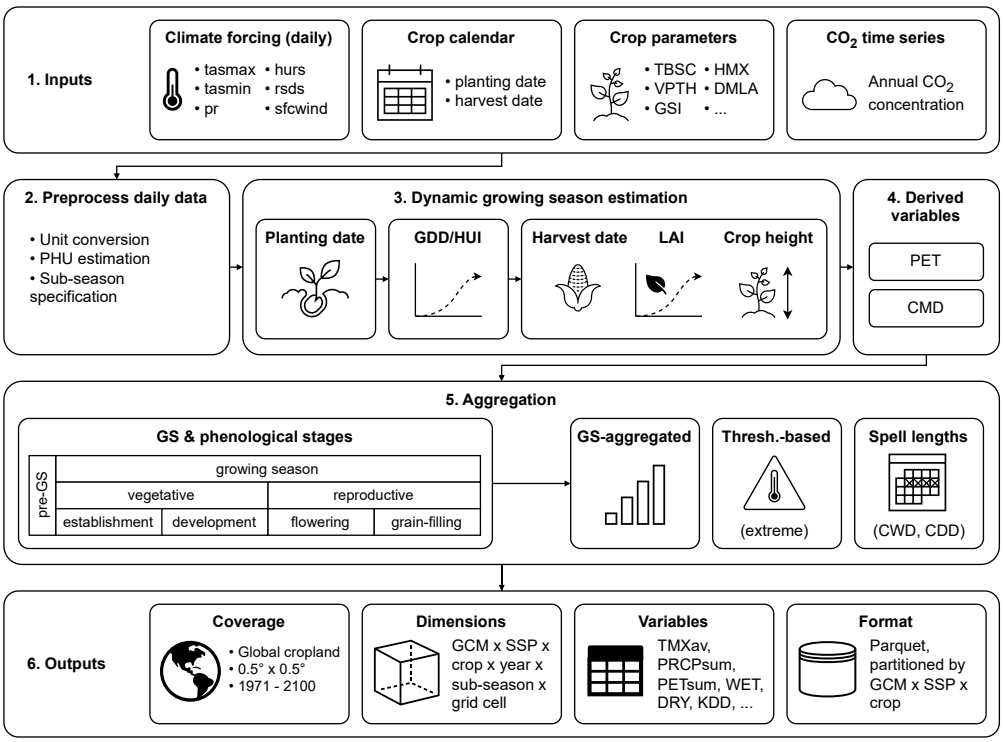


Figure 1: Overview of the AGROMINDS processing workflow. Daily climate forcing, crop calendars, crop parameters, and atmospheric CO_2 concentrations are combined to derive annually resolved agro-meteorological indicators. Climate variables are first preprocessed and used to dynamically estimate crop phenology and growing season timing based on heat-unit accumulation. Daily secondary variables, including PET and CMD, are subsequently derived and aggregated over the entire growing season and phenologically defined sub-periods. The resulting dataset comprises annual, crop-specific agro-meteorological indicators expressed as seasonal means, sums, threshold exceedances, and spell-length metrics.

2.2 Definition of agro-meteorologic indicators and crop growth phases

Daily climate input variables comprise near-surface relative humidity (`hurs`), total precipitation (`pr`), surface downwelling shortwave radiation (`rsds`), near-surface wind speed (`sfcwind`), daily maximum temperature (`tasmax`), and daily minimum temperature (`tasmin`). Prior to aggregation, variable units are converted to metrics commonly used in agricultural research: `hurs` from percent to fraction, `pr` from kg m^{-2} s^{-1} to mm day^{-1}, `rsds` from W m^{-2} to MJ m^{-2} day^{-1}, and `tasmax` and `tasmin` from K to °C. Daily mean temperature (`tas`) is calculated as the arithmetic mean of `tasmax` and `tasmin`.

Based on these daily inputs, the pipeline derives a set of growing season (GS)-aggregated and threshold-based agro-meteorological indicators (Table 1). Each GS-aggregated climate variable is identified by a lowercase name, with the suffix `_sum` denoting cumulative variables and `_av` denoting mean variables.. Seasonal mean indicators include average daily maximum temperature (`tmx_av`), minimum temperature (`tmn_av`), mean temperature (`tav_av`), wind speed (`ws_av`), and relative humidity (`rh_av`). Seasonal totals are calculated for precipitation (`prcp_sum`), solar radiation (`swr_sum`), potential evapotranspiration (`pet_sum`), growing degree days (`gdd_sum`), and cumulative moisture deficit (`cmd_sum`). Moisture deficit is defined at daily resolution as PET minus

precipitation and then summed over the aggregation period.

Potential evapotranspiration is estimated internally for each day using a Penman–Monteith approach that incorporates radiation, humidity, wind speed, air temperature, elevation, atmospheric CO_2, and dynamically evolving canopy properties [27]. Canopy properties are represented by daily LAI and CHT trajectories derived from the crop's GDD progression. This allows PET to reflect both meteorological forcing and the approximate phenological state of the crop. Atmospheric CO_2 concentrations are prescribed from an annual global time series provided by ISIMIP [28] and used in the canopy resistance term of the PET calculation.

In addition to seasonal means and totals, the dataset includes indicators of temperature, precipitation, and drought extremes expressed as counts of days within each aggregation period in which a climate variable crosses a defined threshold. Heat-related metrics include the number of days with daily maximum temperature $\geq 30°C$ (`hdd`) and $\geq 39°C$ (`kdd`), while cold-related metrics include frost days with daily minimum temperature $\leq 0°C$ (`frt`) and ice days with daily maximum temperature $\leq 0°C$ (`ice`). Precipitation extremes are quantified as the number of days with precipitation ≥ 10 mm (`r10`) and ≥ 20 mm (`r20`). Wet and dry day counts are defined using a threshold of 1 mm day^{-1} precipitation, with wet days (`wet`) exceeding this threshold and dry days (`dry`) at or below it. The longest runs of consecutive wet and dry days are additionally reported as `cwd` and `cdd`, respectively. As an integrated atmospheric drought indicator, we furthermore report the number of days with `pet_sum` exceeding `prcp_sum`, referred to as moisture deficit days (`mdd`). Finally, the dataset contains growing (sub-)season length (`len`) and the heat unit index at the end of the aggregation period (`hui`).

Long name	**Short name**	**Description**
Climate indicators expressed as (sub-)growing season sums or averages		
Relative humidity	rh_av	Relative humidity [-]
Precipitation	prcp_sum	Sum of precipitation [mm]
Solar shortwave radiation	swr_sum	Sum of shortwave solar radiation [MJ m-2]
Windspeed	ws_av	Wind speed at 2m height [m s-1]
Max. temperature	tmx_av	Average maximum temperature [°C]
Min. temperature	tmx_av	Average minimum temperature [°C]
Average temperature	tav_av	Average temperature [°C]
Potential evapotranspiration	pet_sum	Potential evapotranspiration [mm]
Cumulative moisture deficit	cmd_sum	Cumulative moisture deficit summed over growing season [mm]
Climate indicators expressed as number of (sub-)growing season days exceeding a threshold [d]		
Wet days	wet	Number of wet days (precipitation >1mm)
Dry days	dry	Number of dry days (precipitation <= 1mm)
Moisture deficit days	mdd	Numer of days where cumulative moisture deficit > 0
Heating degree days	hdd	Heating degree days (maximum daily temperature >= 30°C)
Killing degree days	kdd	Killing degree days (maximum daily temperature >= 39°C)
Frost days	frt	Frost days (minimum daily temperature <= 0°C)
Ice days	ice	Ice days (maximum daily temperature <= 0°C)
Strong precipitation	r10	Number of days with heavy precipitation (precipitation >= 10mm)
Heavy precipitation	r20	Number of days with heavy precipitation (precipitation >= 20mm)
Consecutive wet days	cwd	Number of consecutive wet days (precipitation >1mm)
Consecutive dry days	cdd	Number of consecutive dry days (precipitation <= 1mm)
Growing season features		
(Sub)-Growing season length	len	Growing (sub-)season length [d]
Heat unit index end of period	hui	Heat unit index at the end of each period [-]

Table 1: Agro-meteorologic indicators quantified in the dataset based on [9, 7, 29, 10, 20, 30, 31].

Indicators are generated for the full growing season and for phenologically defined sub-periods

based on the fraction of achieved heat-unit development rather than on fixed calendar windows. Specifically, the vegetative phase (`gs_v`) spans HUI values from 0.0 to 0.5 and the reproductive phase (`gs_r`) spans HUI values from 0.5 to 1.0 (maturity). These are further subdivided into establishment (`gs_e`, HUI 0–0.25), development (`gs_d`, 0.25–0.5), flowering (`gs_f`, 0.5–0.75), and grain-filling (`gs_g`, 0.75–1.0 or harvest) periods. Defining sub-seasons through normalized crop development rather than fixed day counts ensures that the aggregation windows remain comparable across environments with differing season lengths and temperature regimes. An overview of GS fractions is provided in Table 2. We acknowledge that this is a rather coarse approach due to limited information on the one hand and to ensure global comparability on the other. In practice, cardinal development phases will vary by crops and cultivars. Where such information is available, the values can be adjusted in the code provided aongside the data.

Abbreviation	**GS Fraction**	**Description**
gs_p	-	Pre-growing season, i.e., the 30 days prior to sowing
gs_v	0-0.5	Vegetative phase, i.e., first half of the growing season
└ gs_e	0-0.25	Establishment phase, i.e., first quarter of the growing season
└ gs_d	0.25-0.5	Development phase, i.e., second quarter of the growing season
gs_r	0.5-1	Reproductive phase, i.e., second half of the growing season
└ gs_f	0.5-0.75	Flowering phase, i.e., third quarter of the growing season
└ gs_g	0.75-1	Grain-filling phase, i.e., fourth quarter of the growing season
gs	0-1	Total growing season

Table 2: Temporal aggregation periods of the whole growing season and sub-seasons.

All indicators are first computed at daily resolution and then aggregated by grid cell, year, and season segment. This design yields one annual observation per pixel and period, producing a compact but information-rich dataset suitable for global analyses of crop-relevant climate conditions under historical and projected climates.

2.3 Crop parameters

We include herein the four major annual grain crops maize, soybean, rice, and wheat that are most commonly used in climate impact assessments and provide the major share in vegetal calorie supply from global agriculture [16]. Besides differences in growing season timing, the estimation of agro-meteorologic indicators accounts for differences in cardinal temperatures and phenological development through crop-specific parameters. These are listed in Table 3.

	Maize	Rice	Soy	Spring Wheat
TO	25	25	25	20
TBSC	8	10	10	5
VPTH	0.5	0.5	0.5	0.5
VPD2	0.071	0.071	0.071	0.071
GSI	0.007	0.008	0.007	0.007
HMX	2	0.8	0.8	1
DMLA	6	6	5	6
DLAP1	3.055136	5.413224	3.055136	5.413224
DLAP2	13.38548	18.10165	13.38548	18.10165
DLAI	0.8	0.8	0.9	0.6
RLAD	1	0.5	0.1	1
GMHU	100	100	100	100

Table 3: Crop parameters used for phenologic development. TO = optimal growing temperature [°C], TBSC = Minimum temperature for plant growth [°C], VPTH = Vapor pressure deficit threshold above which the crop starts closing stomata [kPa], VPD2 = Second point on the leaf conductance decline curve, GSI = Maximum Stomatal Conductance [m/s], HMX = Max crop height [m], DMLA = Maximum potential leaf area index, DLAP1 = Growth curve parameter 1, DLAP2 = Growth curve parameter 2, DLAI = Fraction of growing season where leaf area declines, RLAD = Leaf area index decline parameter, GMHU = Heat units required for gemination.

2.4 Input data

We produce agro-climatic indicators for the ISIMIP phase 3a reanalysis dataset and the phase 3b priority GCMs, a core set of state-of-the-art, widely used climate datasets processed and provided by ISIMIP [15, 28].

The phase 3a climate reanalysis GSWP3-W5E5 is a merged product based on the Global Soil Wetness Project v3 (GSWP3) and the W5E5 reanalysis product [15, 32]. The latter consists of WATCH Forcing Data methodology version 2.0 applied to ERA5 data [33, 34] and ERA5 reanalysis data directly [35]. Details are provided in Frieler et al. [15] and Lange & Büchner [36]. The scenario of the phase 3a reanalysis is termed `obsclim` as an equivalent to the SSP-RCP labels in the climate projections (see below). GSWP3-W5E5 also served for estimating baseline PHU for each crop and location (see section 2.1).

The ISIMIP phase 3b climate projections include the GCMs GFDL-ESM4, IPSL-CM6A-LR, MPI-ESM1-2-HR, MRI-ESM2-0, and UKESM1-0-LL based on their contributions to the sixth phase of the Coupled Model Intercomparison Project (CMIP6). For each GCM, we include the complete set of core SSP-RCP scenarios, i.e., `historical`, `ssp126`, `ssp245`, `ssp370`, and `ssp585` forcing pathways. GCM simulation outputs have been bias-adjusted and statistically downscaled to 0.5° x 0.5° resolution using ISIMIP3 bias-adjustment and spatial downscaling (ISIMIP3BASD) version 2.5 methodology [37, 28, 38].

To avoid excess data volumes, we produce indicators only for pixels with present cropland according to the Spatial Production Allocation Model (SPAM) 2020 v2r2 database [39]. This dataset provides harvested areas for 46 major crops and crop groups for circa the year 2020 combining remote sensing and census data products. It is among the most widely used crop-specific land use datasets in climate impact and land use modelling. The original 5 arcmin data were aggregated to 30 arcmin using the aggregate function from the R terra package [40] with factor six. Agro-climatic

indicators for pixels presently not including cropland can be produced using the publicly available code and data.

To define spatially explicit growing seasons, we use crop- and water supply regime-specific planting and harvest date calendars provided by GGCMI [16, 24]. This dataset combines various observational sources for major crops and, where applicable, by commonly rainfed (abbreviated rf in the data products) and commonly irrigated seasons separately. The dataset provides growing season information on single seasons per pixel for maize, soybean, and spring wheat. For rice as the crop grown most widely in multiple cropping globally, it provides information on the first and second season. Herein, we use these five calendars separately. We solely provide outputs for the rainfed calendars as this is the globally most dominant form of cultivation. Notably, for rice, the crop most comprehensively supplied with surface water, both calendars are identical. Users can generate outputs for the irrigated seasons of other crops using the provided code and publicly accessible data. The GGCMI calendar provides data for a total of 18 crops to which the agroclimatic indicators can be extended accordingly.

3 Data Records

3.1 Data archive

AGROMINDS is archived on Zenodo [41]. The archive contains globally gridded, annual agrometeorologic indicators derived for five crop-specific growing seasons from bias-adjusted climate forcing. Climate projections comprise a historical period (1971–2014) and four future Shared Socioeconomic Pathway (SSP) scenarios (2015–2100; SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5) for five CMIP6 global climate models (GFDL-ESM4, IPSL-CM6A-LR, MPI-ESM1-2-HR, MRI-ESM2-0, and UKESM1-0-LL). In addition, observational agro-meteorologic indicators derived from GSWP3 reanalysis data are provided for the historical period.

To facilitate downloading, the repository is distributed as 26 compressed ZIP archives. Of these, 25 archives correspond to the combination of five crops and five climate periods (historical, SSP126, SSP245, SSP370, and SSP585). Each archive contains five Apache Parquet files, one for each CMIP6 global climate model. Observational data for all five crops are distributed separately in a single archive. Individual Parquet files are named according to the convention `agrominds_<gcm>_<period>_<crop>_rf.parquet`, where `rf` denotes rainfed conditions.

The Zenodo record additionally includes a `README.md` file describing the archive organisation, a `variable_definitions.csv` file documenting all variables, units, and definitions, and a `file_inventory.csv` listing all distributed archives and their contents. A simplified representation of the archive structure is shown below.

```
AGROMINDS/
  README.md
  variable_definitions.csv
  file_inventory.csv
  agrominds_allcrop_obsclim_rf.zip
  agrominds_mai_historical_rf.zip
  agrominds_mai_ssp126_rf.zip
  agrominds_mai_ssp245_rf.zip
  ...
  agrominds_soy_ssp585_rf.zip
  agrominds_swh_ssp585_rf.zip
```

Data files are provided in columnar tabular format to support efficient storage and downstream analysis. Each row represents one grid-cell–year–period combination. Spatial referencing is provided through the columns `lat` and `lon`, temporal referencing through `yr`, and phenological aggregation period through `period`. The `period` field identifies whether the record corresponds to the entire growing season or to one of the crop-development sub-periods defined in Table 2. The main data files contain the agro-meteorological indicators listed in Table 1, whereas column names correspond to the short names in the table.

To reduce the size of the Parquet files, continuous variables are stored using a scaled integer representation. Each variable is multiplied by a variable-specific scaling factor (Table 4), rounded to the nearest integer, and stored using the smallest suitable integer type. The original values are recovered by applying the inverse scaling during decoding.

Variable	**Scaling Factor**	**Data Type**
yr	1	UInt16
lat	100	Int16
lon	100	Int16
len	1	UInt16
tmx_av	100	Int16
tmx_av	100	Int16
tav_av	100	Int16
prcp_sum	1	UInt16
swr_sum	1	UInt16
ws_av	100	UInt16
rh_av	1000	UInt16
pet_sum	10	UInt16
gdd_sum	1	UInt16
cmd_sum	1	Int16
hui	1000	UInt16
[**Thresh.-based**]	1	UInt16

Table 4: Variable-specific scaling factors used for fixed-point encoding of the dataset for compression. Continuous variables are multiplied by the scaling factor, rounded to the nearest integer, and stored using the indicated integer type. Original values are recovered by applying the inverse scaling.

Users should note that sub-season records are nested within the same annual growing season and therefore should not be interpreted as independent observations from the corresponding full-season record. Likewise, differences among files reflect both differences in climate forcing and differences in the timing and duration of crop-specific growing seasons derived by the processing pipeline. All externally sourced input datasets used to construct AGROMINDS, including climate forcing and crop-calendar information, are cited in the Methods section and reference list using formal data citations.

3.2 Exemplary indicator responses

Due to the extent of indicators, crops, sub-seasons, and climate scenarios only an exemplary presentation of representative parts of the dataset is feasible herein. We present for maize only and

for selected indicators (i) distributions of feature changes across cropland pixels, (ii) changes in global area-weighted mean indicator values, and (iii) spatial patterns in the form of maps. Figure 2 shows the distribution of a selected subset of features for future SSP projections (2070–2099) and the historical baseline (1985–2014).

Average maximum growing-season temperature increases consistently across all GCMs and locations, with an expectedly more pronounced shift under SSP3-7.0 than under SSP2-4.5. Under SSP3-7.0, UKESM1-0-LL projects the highest temperatures, whereas MPI-ESM1-2-HR projects the lowest, consistent with previous studies [9, 16]. Notably, for this GCM with very high transient climate response (TCR; [42]), i.e., sensitivity of temperature response to atm. CO_2 increase, GS temperature exceeds hardly that of other GCMs in absolute terms but shows spatially more widespread warming (Figure 2a,b). Changes in growing-season precipitation are comparatively subtle, although a shift towards lower values is evident for MPI-ESM1-2-HR, MRI-ESM2-0, and UKESM1-0-LL. This contrasts changes in global mean annual values that were shown to be proportional. I.e., precipitation increases on average for the selected GCMs with increasing temperature [43] (Figure 2c,d). Similarly to temperature, both radiation and KDD increase across all GCMs, with the largest changes projected by UKESM1-0-LL (Figure 2e-h). PET and the dryness indicator CMD also shift towards higher values, with the increase being more pronounced for PET and, again, for UKESM1-0-LL. As temperature and shortwave solar radiation are the key drivers of PET, this is expected due to the increase in both. For CMD, changes in precipitation may allow to partly meet the higher atmospheric water demand (Figure 2i-l).

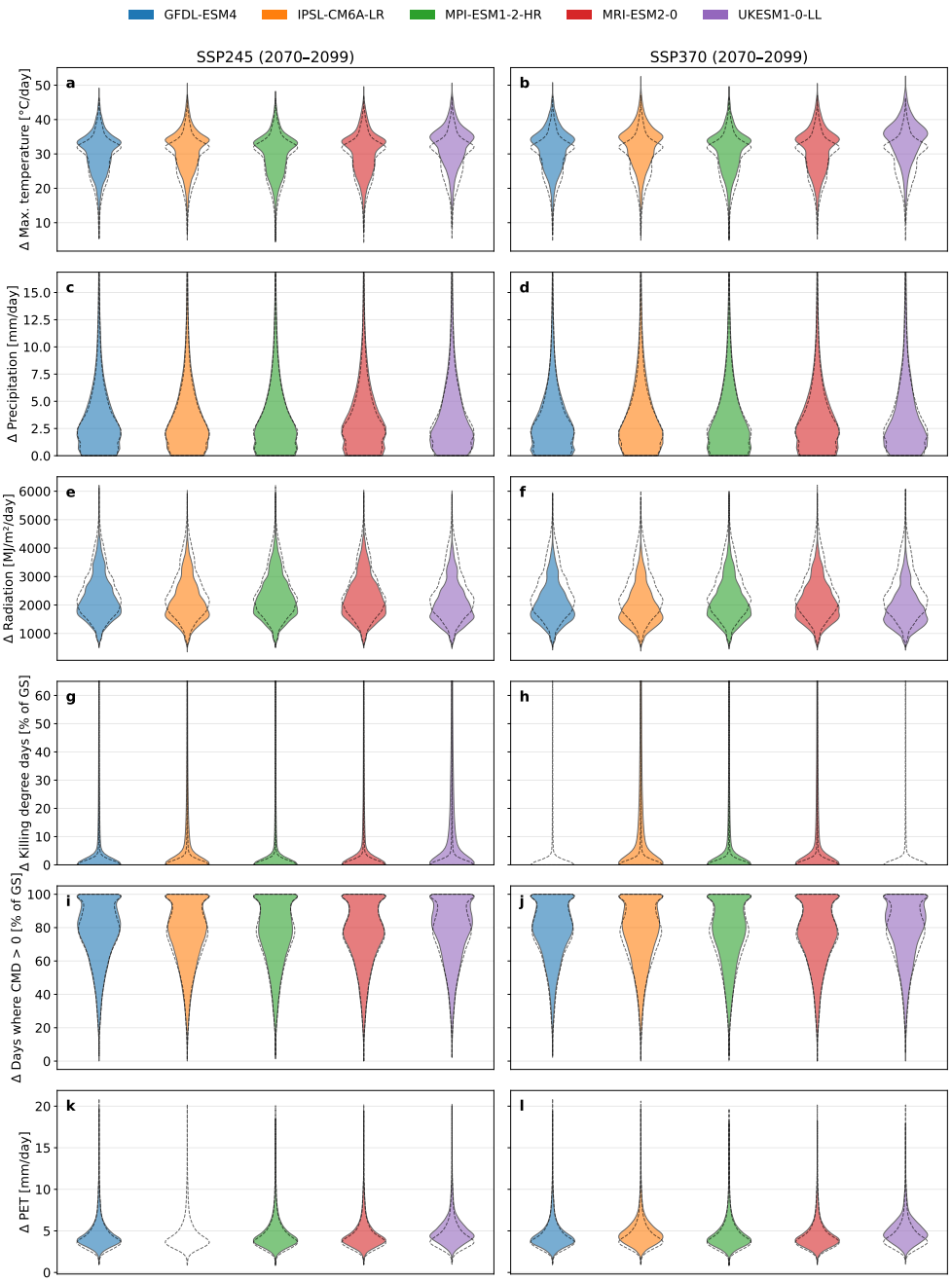


Figure 2: Distributions of selected features across GCMs, with individual violins representing different GCMs. The left column shows future projections under SSP2-4.5, while the right column shows projections under SSP3-7.0 for the time period 2070–2099. For reference, values for the historical period (1985–2014) are indicated by the dashed outlines. For improved visualization, the precipitation and KDD axes were truncated to exclude the most extreme values (above the 99th and 95th percentiles, respectively); all other variables are shown over their full range.

Figure 3 shows the global evolution of the selected climate features over time as annual values weighted by the harvested area of rainfed maize using the SPAM2020 v2r2 database [39]. Consistent

with the distributional analysis, temperature and temperature-related features, including killing degree days, exhibit a clear increasing trend towards the end of the century (Figure 3a,b,g,h). The relative ordering of the GCMs remains largely consistent throughout time, with MPI-ESM1-2-HR projecting the lowest values, MRI-ESM2-0, GFDL-ESM4, and IPSL-CM6A-LR occupying the intermediate range, and UKESM1-0-LL consistently projecting the highest values. Solar radiation behaves similarly but interestingly IPSL-CM6A-LR is here at the lower end of change rates in SSP2-4.5 (Figure 3e,f). Precipitation exhibits primarily large interannual variability compared to the prior indicators, although a statistically significant increasing trend is observed across all GCMs and for both SSPs (maximum p-value = 0.01 for the slope of a linear regression; Figure 3c,d). The fraction of days with atmospheric dry conditions (CMD greater zero; Figure 3i,j) does apparently not increase for UKESM1-0-LL substantially more than for other GCMs despite its large increases in growing season temperature and radiation. This is likely due to the simultaneous increase in precipitation that can in part compensate the higher atmospheric water demand.

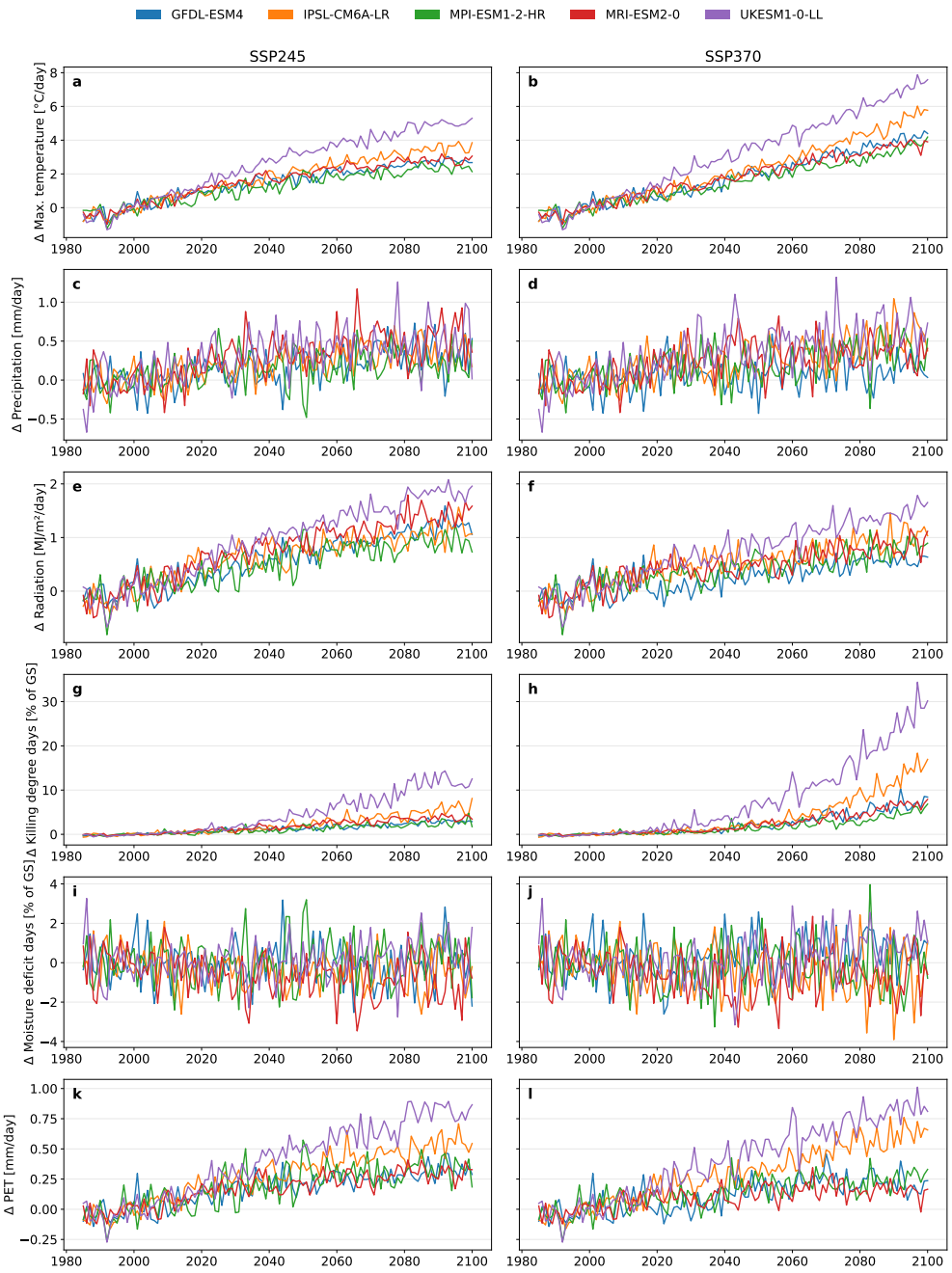


Figure 3: Timeseries of area-weighted feature values accross all pixels, shown as deviations from the 1995-2014 historical baseline.

Figure 4 illustrates the projected spatial changes under SSP3-7.0 for 2070–2099 relative to the historical baseline (2000–2014). Both GCMs project widespread warming, accompanied by increased PET and a substantial rise in killing degree days, particularly in already warm regions where temperatures are more likely being pushed beyond the threshold with additional warming (Figure 4a,b,g-j). Precipitation changes are more heterogeneous, with both increases and decreases projected depending on location, and UKESM1-0-LL often projecting changes of greater magnitude

than MPI-ESM1-2-HR. Yet, there are also several regions with opposite changes in growing-season precipitation among the two GCMs, such as North-Western Mexico, Eastern Australia, Colombia, Eastern EU, and Northern Madagascar (Figure 4c,d). Especially where changes in precipitation are large, they typically go along with decreases in solar radiation (Figure 4e,f). Consequently, PET tends to increase mostly where precipitation decreases and solar radiation increases (Figure 4i,j).

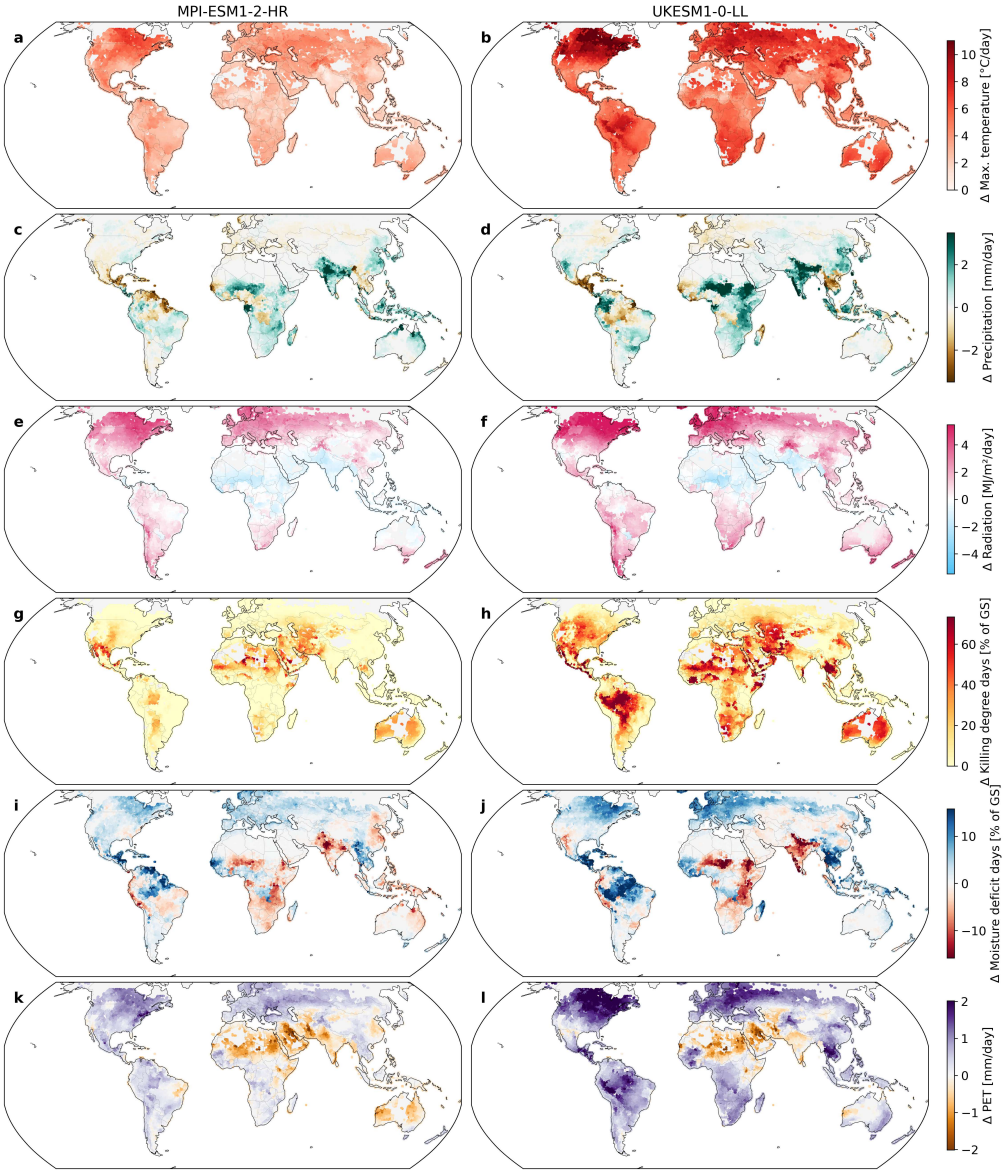


Figure 4: Spatial changes in selected climate features under SSP3-7.0 for the period 2070–2099 relative to the historical reference period (2000–2014). Rows correspond to individual climate features, and columns show results from the MPI-ESM1-2-HR and UKESM1-0-LL GCMs. Color scales are shared within each row to facilitate comparison between models.

4 Technical Validation

AGROMINDS was technically validated through a set of structural and plausibility checks applied to both intermediate and final outputs. Input datasets were screened for coordinate consistency, valid year ranges, and availability of the variables required for season construction and indicator calculation. Growing seasons and sub-periods were then checked for internal consistency, including valid planting and harvest timing, correct handling of cross-year seasons, and plausible growing season lengths.

The resulting agro-meteorological indicators were further evaluated for numerical plausibility. In particular, the distributions of non-negative variables, including precipitation, potential evapotranspiration, growing degree days, and season length, were evaluated for consistency with their expected physical constraints, while count-based indicators were verified not to exceed the length of the corresponding aggregation period. Descriptive statistics and visual inspection of gridded outputs were additionally used to identify potential artifacts such as missing values, duplicated records, or spatial misalignment. These checks confirmed that the dataset is structurally coherent and suitable for downstream analysis.

5 Usage Notes

AGROMINDS is designed for applications requiring spatially explicit, annually resolved summaries of crop-relevant climate conditions over reported or dynamically estimated growing seasons. Because the indicators are provided on a common global grid and are calculated consistently across years, they are suitable for comparative analyses across regions, time periods, and climate scenarios. Potential applications include statistical analyses of historical yield variability, the development of statistical or machine-learning crop yield models, assessments of compound or concurrent climate risks across breadbasket regions, and broader agroclimatic studies such as crop suitability or exposure mapping. For example, closely aligned indicators have recently been applied by the authors for the evaluation of climate sensitivity of global crop models and observational yields [8] and for the training of crop model emulators [9] showing high explanatory power in both cases.

The dataset is particularly useful where fixed calendar aggregations would inadequately represent crop-relevant weather conditions. By aligning temporal aggregation windows with crop-specific planting and maturation dynamics, the indicators better capture the timing of heat, water stress, and precipitation extremes experienced during crop development. Users may therefore combine the data with gridded crop model output, sub-national production statistics, remote-sensing products, or socioeconomic datasets to investigate climate impacts on agricultural systems.

Several considerations are important for interpretation. First, the indicators represent meteorological conditions and derivatives based on common estimation methods (e.g., PET); they do not by themselves account for management, soils, pests, nutrient limitations, or direct crop damage processes. Second, the growing seasons are based on a well-established but stylized crop maturity and phenology approach. Thus, they reflect a harmonized representation of crop timing rather than locally observed management. Third, some indicators are derived quantities with explicit dependencies on upstream calculations, especially those related to crop development and PET. Users interested only in a subset of variables using other data sources may therefore adapt the accompanying code to skip unnecessary downstream calculations and reduce processing time.

For reproducible downstream use, each record should be interpreted as one pixel-year-period combination, where `period` identifies the whole growing season or one of its phenological subperiods. Analyses combining multiple periods should account for the fact that sub-season indicators

are not independent of the full-season aggregates but are nested within them. When comparing absolute values across crops, models, or regions, users should further consider that differences may arise both from climate forcing and from differences in season length and timing. The accompanying code base enables recalculation of the indicators with alternative crop calendars, dynamic-season settings, or customized feature subsets for user-specific applications.

6 Code Availability

All code required to generate the AGROMINDS dataset is publicly available at `https://github.com/iiasa/agrominds`. The repository contains the processing workflow for deriving growing season-specific agro-meteorological indicators from daily climate input data, including routines for season definition, crop-development tracking, potential evapotranspiration, and temporal aggregation. Software dependencies are documented in the provided `requirements.txt` file. The code is released under the MIT License.

7 Acknowledgements

This research was funded in whole by the Austrian Science Fund (FWF) (10.55776/P36220). For the purpose of open access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

References

[1] Lesk, C. *et al.* Compound heat and moisture extreme impacts on global crop yields under climate change. *Nature Reviews Earth & Environment* **3**, 872–889 (2022). URL `https://doi.org/10.1038/s43017-022-00368-8`.

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

[3] Mistry, M. N., Sue Wing, I. & De Cian, E. Simulated vs. empirical weather responsiveness of crop yields: Us evidence and implications for the agricultural impacts of climate change. *Environmental Research Letters* **12**, 075007 (2017). URL `https://doi.org/10.1088/1748-9326/aa788c`.

[4] Goulart, H. M. D., Van Der Wiel, K., Folberth, C., Balkovic, J. & Van Den Hurk, B. Storylines of weather-induced crop failure events under climate change. *Earth System Dynamics* **12**, 1503–1527 (2021).

[5] Goulart, H. M. D., Wiel, K. V. D., Folberth, C., Boere, E. & Hurk, B. V. D. Increase of simultaneous soybean failures due to climate change (2022).

[6] Lischeid, G., Webber, H., Sommer, M., Nendel, C. & Ewert, F. Machine learning in crop yield modelling: A powerful tool, but no surrogate for science. *Agricultural and Forest Meteorology* **312**, 108698 (2022). URL `https://www.sciencedirect.com/science/article/pii/S0168192321003841`.

[7] Nobakht, M., Beavis, P., O'Hara, S., Hutjes, R. & Supit, I. Agroclimatic indicators: Product user guide and specification (pugs) (2019). URL `https://confluence.ecmwf.`

int/spaces/CKB/pages/278550972/Agroclimatic+Indicators+Product+User+Guide+and+Specification+PUGS. Copernicus Climate Change Service (C3S), ECMWF Confluence Knowledge Base.

[8] Oberleitner, T. *et al.* Explaining climatic drivers of yield anomalies in global crop models through metamodel-based attribution. *ESS Open Archive* **2025** (2025). URL https://essopenarchive.org/doi/abs/10.22541/essoar.176071967.73925210/v1.

[9] Folberth, C. *et al.* Cromes v1.0: a flexible crop model emulator suite for climate impact assessment. *Geoscientific Model Development* **18**, 5759–5779 (2025). URL https://gmd.copernicus.org/articles/18/5759/2025/.

[10] Mistry, M. N. A high-resolution global gridded historical dataset of climate extreme indices. *Data* **4** (2019). URL https://www.mdpi.com/2306-5729/4/1/41.

[11] Schillerberg, T. & Tian, D. Changes in crop failures and their predictions with agroclimatic conditions: Analysis based on earth observations and machine learning over global croplands. *Agricultural and Forest Meteorology* **340**, 109620 (2023). URL https://www.sciencedirect.com/science/article/pii/S0168192323003118.

[12] Zabel, F. *et al.* Large potential for crop production adaptation depends on available future varieties. *Global Change Biology* **27**, 3870–3882 (2021).

[13] Donatelli, M. *et al.* Modelling the impacts of pests and diseases on agricultural systems. *Agricultural systems* **155**, 213–224 (2017).

[14] Pequeno, D. N. *et al.* Production vulnerability to wheat blast disease under climate change. *Nature Climate Change* **14**, 178–183 (2024).

[15] Frieler, K. *et al.* Scenario setup and forcing data for impact model evaluation and impact attribution within the third round of the Inter-Sectoral Impact Model Intercomparison Project (ISIMIP3a). *Geoscientific Model Development* **17**, 1–51 (2024). URL https://gmd.copernicus.org/articles/17/1/2024/.

[16] Jägermeyr, J. *et al.* Climate impacts on global agriculture emerge earlier in new generation of climate and crop models **2**, 873–885. URL https://www.nature.com/articles/s43016-021-00400-y.

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

[18] Chen, X., Anderson, W., You, L. & Pope, E. Observed trends in multiple breadbasket yield shocks. *Environmental Research Letters* **19**, 104005 (2024). URL https://doi.org/10.1088/1748-9326/ad7040.

[19] Gaupp, F., Hall, J., Hochrainer-Stigler, S. & Dadson, S. Changing risks of simultaneous global breadbasket failure. *Nature Climate Change* **10**, 54–57 (2020).

[20] Hasegawa, T., Wakatsuki, H. & Nelson, G. C. Evidence for and projection of multi-breadbasket failure caused by climate change. *Current Opinion in Environmental Sustainability* **58**, 101217 (2022).

[21] Burney, J., McIntosh, C., Lopez-Videla, B., Samphantharak, K. & Maia, A. G. Empirical modeling of agricultural climate risk. *Proceedings of the National Academy of Sciences* **121**, e2215677121 (2024). URL https://www.pnas.org/doi/abs/10.1073/pnas.2215677121. https://www.pnas.org/doi/pdf/10.1073/pnas.2215677121.

[22] Sloat, L. L. *et al.* Climate adaptation by crop migration. *Nature Communications* **11**, 1243 (2020). URL https://doi.org/10.1038/s41467-020-15076-4.

[23] Biess, B., Gudmundsson, L., Windisch, M. G. & Seneviratne, S. I. Future changes in spatially compounding hot, wet or dry events and their implications for the world's breadbasket regions. *Environmental Research Letters* **19**, 064011 (2024). URL https://doi.org/10.1088/1748-9326/ad4619.

[24] Jägermeyr, J., Müller, C., Minoli, S., Ray, D. & Siebert, S. GGCMI phase 3 crop calendar (2021). URL https://doi.org/10.5281/zenodo.5062513.

[25] Müller, C. *et al.* Global gridded crop model evaluation: Benchmarking, skills, deficiencies and implications **10**, 1403–1422. URL https://gmd.copernicus.org/articles/10/1403/2017/.

[26] Elliott, J. *et al.* The Global Gridded Crop Model Intercomparison: data and modeling protocols for Phase 1 (v1.0). *Geoscientific Model Development* **8**, 261–277 (2015). URL https://gmd.copernicus.org/articles/8/261/2015/.

[27] Stockle, C. O., Williams, J. R., Rosenberg, N. J. & Jones, C. A method for estimating the direct and climatic effects of rising atmospheric carbon dioxide on growth and yield of crops: Part I—Modification of the EPIC model for climate change analysis. *Agricultural Systems* **38**, 225–238 (1992). URL https://www.sciencedirect.com/science/article/pii/0308521X9290067X.

[28] Frieler, K. *et al.* Scenario set-up and the new CMIP6-based climate-related forcings provided within the third round of the Inter-Sectoral Model Intercomparison Project (ISIMIP3b, group I and II). *Geoscientific Model Development* **19**, 4095–4135 (2026).

[29] McErlich, C. *et al.* Positive correlation between wet-day frequency and intensity linked to universal precipitation drivers. *Nature Geoscience* **16**, 410–415 (2023).

[30] Schauberger, B. *et al.* Consistent negative response of us crops to high temperatures in observations and crop models. *Nature Communications* **8**, 13931 (2017).

[31] Sweet, L., Müller, C., Jägermeyr, J. & Zscheischler, J. A data-driven method for identifying climate drivers of agricultural yield failure from daily weather data (2025). URL https://egusphere.copernicus.org/preprints/2025/egusphere-2025-3006/.

[32] Mengel, M., Treu, S., Lange, S. & Frieler, K. ATTRICI v1.1 – counterfactual climate for impact attribution. *Geoscientific Model Development* **14**, 5269–5284 (2021). URL https://gmd.copernicus.org/articles/14/5269/2021/.

[33] Weedon, G. P. *et al.* The wfdei meteorological forcing data set: Watch forcing data methodology applied to era-interim reanalysis data. *Water Resources Research* **50**, 7505–7514 (2014). URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2014WR015638. https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/2014WR015638.

[34] Cucchi, M. *et al.* WFDE5: bias-adjusted ERA5 reanalysis data for impact studies. *Earth System Science Data* **12**, 2097–2120 (2020). URL https://essd.copernicus.org/articles/12/2097/2020/.

[35] Lange, S. *et al.* WFDE5 over land merged with ERA5 over the ocean (W5E5 v2.0) (2021). URL https://doi.org/10.48364/ISIMIP.342217.

[36] Lange, S. & Büchner, M. ISIMIP3b bias-adjusted atmospheric climate input data (2021). URL https://doi.org/10.48364/ISIMIP.842396.1.

[37] Lange, S. Trend-preserving bias adjustment and statistical downscaling with ISIMIP3BASD (v1.0). *Geoscientific Model Development* **12**, 3055–3070 (2019). URL https://gmd.copernicus.org/articles/12/3055/2019/.

[38] Lange, S. ISIMIP3BASD (2021). URL https://doi.org/10.5281/zenodo.4686991.

[39] International Food Policy Research Institute (IFPRI). Global Spatially-Disaggregated Crop Production Statistics Data for 2020 Version 2.0 Release 2 (2026). URL https://doi.org/10.7910/DVN/SWPENT.

[40] Hijmans, R. J. *et al.* terra: Spatial Data Analysis (2026). URL https://cran.r-project.org/web/packages/terra/index.html.

[41] Oberleitner, T., Khabarov, N. & Folberth, C. AGROMINDS - Historic and projected global agro-meteorologic indicators for major summer crops (2026). URL https://doi.org/10.5281/zenodo.21639801 (preview).

[42] Hausfather, Z., Marvel, K., Schmidt, G. A., Nielsen-Gammon, J. W. & Zelinka, M. Climate simulations: recognize the 'hot model' problem. *Nature* **605**, 26–29 (2022). URL https://www.nature.com/articles/d41586-022-01192-2.

[43] Lange, S. ISIMIP3b bias adjustment fact sheet (2021). URL https://www.isimip.org/documents/413/ISIMIP3b_bias_adjustment_fact_sheet_Gnsz7CO.pdf.