Weather and Climate Extremes 9 (2015) 47–56

Contents lists available at ScienceDirect

Weather and Climate Extremes

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

Lengthening of the growing season in wheat and maize producing regions

Brigitte Mueller [a,*], Mathias Hauser [b], Carley Iles [c], Ruksana Haque Rimi [d], Francis W. Zwiers [e], Hui Wan [a]

[a] Climate Research Division, Environment Canada, Toronto, Canada
[b] Institute for Atmospheric and Climate Science, ETH Zurich, Switzerland
[c] University of Edinburgh, UK
[d] Environmental Change Institute, Oxford University, Oxford, UK
[e] Pacific Climate Impacts Consortium, University of Victoria, Victoria, Canada

ARTICLE INFO

Article history:
Received 1 December 2014
Received in revised form
15 April 2015
Accepted 23 April 2015
Available online 1 May 2015

Keywords:
Growing season length
Climate change
Detection
Attribution
Crop production

ABSTRACT

Human-induced increases in atmospheric greenhouse gas concentrations have led to rising global temperatures. Here we investigate changes in an annual temperature-based index, the growing season length, defined as the number of days with temperature above 5 °C. We show that over extratropical regions where wheat and maize are harvested, the increase in growing season length from 1956 to 2005 can be attributed to increasing greenhouse gas concentrations. Our analyses also show that climate change has increased the probability of extremely long growing seasons by a factor of 25, and decreased the probability of extremely short growing seasons. A lengthening of the growing season in regions with these mostly rain-fed crops could improve yields, provided that water availability does not become an issue. An expansion of areas with more than 150 days of growing season into the northern latitudes makes more land potentially available for planting wheat and maize. Furthermore, double-cropping can become an alternative to current practices in areas with very long growing seasons which are also shown to increase with a warming climate. These results suggest that there is a strong impact of anthropogenic climate change on growing season length. However, in some regions and with further exacerbated climate change, high temperatures may already be or may become a limiting factor for plant productivity.

1. Introduction

The global number of undernourished people has decreased from over a billion (or 18.7% of global population) in 1990–1992 to around 805 million people (11.3%) in 2012–2014. Reaching the Millennium Development Goal of halving the proportion of undernourished people in developing countries between 1990 and 2015 might be possible (FAO, IFAD, WFP, 2014). Several factors have contributed to reducing world hunger in the last decades. In general, an increase in agricultural production can be obtained through an increase in the area under production or an increase in the productivity on existing farmland. The most important factor for the decline in hunger over recent decades is increased crop yields (Foley, 2011), i.e. the productivity per unit area. Both developing new crop varieties and increasing planting densities can increase yields (McClung, 2014). Production can further be increased by harvesting two crops on the same field each year (double-cropping). Double-cropping is currently still relatively insignificant (Brochers et al., 2014) and mostly confined to the tropics (Siebert et al., 2010), but could have huge potential for food security as it can nearly double yields.

Wheat and maize are the crops with the largest area harvested and second only to sugarcane in their annual production (data for 2012, FAO statistic, 2014). Both agricultural output (production) and yields have increased steadily over the last decades, while the area harvested has stayed nearly constant (Fig. 1). In order to meet future demands, it is projected that production would need to reach 891 and 1343 million tonnes for wheat and maize, respectively, by 2050 (Ray et al., 2012).

Climate change can have either negative or positive impacts on crop production, depending on the region (Cheng et al., 2011; Porter et al., 2014). In high latitudes, warmer temperatures lead to longer growing seasons and an increase in potential agricultural land (Gornall et al., 2010). We here select areas where wheat and maize are grown to investigate whether growing season (GS) length (GSL) in such important agricultural areas has increased and whether it can potentially contribute to increased production. We therefore first analyze whether the area with GS long enough to grow wheat and maize or to do double-cropping has increased. We further investigate

* Corresponding author.
E-mail address: brigitte.mueller@ec.gc.ca (B. Mueller).

http://dx.doi.org/10.1016/j.wace.2015.04.001
2212-0947/

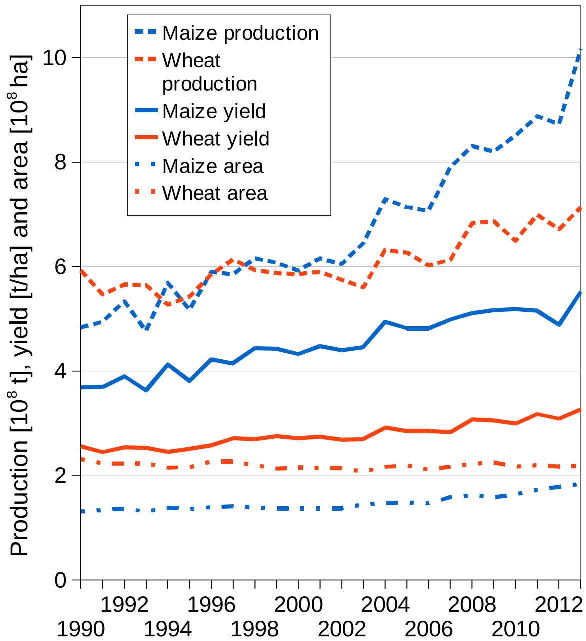

Fig. 1. Production, yield and area harvested for wheat (red) and maize (blue) over the last two decades. *Data source*: FAO homepage (2014). (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

whether changes in GSL in wheat and maize areas can be attributed to human activity. A significant lengthening in nearly-global GS has been found and attributed to an increase in anthropogenic greenhouse gas concentrations (Christidis et al., 2007). However, such an attribution has not been previously done specifically for wheat and maize regions. We also employ a larger set of climate model simulations and more recent data than Christidis et al. (2007).

We investigate changes in the occurrence of extremely long or short GSL since both types of extremes can substantially impact crop production. Based on an ensemble of possible GSLs from observationally constrained climate model simulations, we investigate changes in the probability of extreme GSL over recent years and estimate the effects that human-induced greenhouse gas emissions have had on these probabilities.

2. Regions with wheat and maize production

Due to their importance for global food production, our analyses are focused on wheat and maize areas. The definition of areas with wheat and maize is based on data from EarthStat (www.earthstat.org), a collaborative effort between the University of Minnesota's Institute on the Environment-Global Landscapes Initiative and McGill University's Land Use and the Global Environment lab. EarthStat provides harvested area and yields for 175 crops around the year 2000, which they obtained by combining national and sub-national agricultural census records with satellite imagery. The data are described in Monfreda et al. (2008).

Fig. 2 shows the areas where wheat and maize are produced (data for year 2000, interpolated to $2.5^\circ \times 3.75^\circ$ grid). For each crop, we consider grid cells that are more than 1% covered with the respective crop. Note that the areas partly overlap and our results for the two crop areas are thus not independent. We obtain 4.54% of all land area as our wheat area and 4.65% as maize area.

3. Observed growing season length

Even though some varieties of wheat can resist temperatures down to -20 °C in their early stages of growth, wheat and maize cannot tolerate frost during their main growth period. Minimum daily temperatures of about 5 °C are needed for measurable growth of both winter and spring wheat, and slightly higher temperatures are required for maize. The optimum temperature for wheat growth is between 15 to 20 °C. We select a temperature threshold for our growing season definition of 5 °C, above which wheat and maize both grow well and for which a global observational GSL dataset, the HadEX2 dataset (Donat, 2013, see below), is available. The annual length of the growing season is defined as the number of days between the first span of at least 6 days with daily mean temperature warmer than 5 °C and the first span of 6 days with daily mean temperature below 5 °C. For this calculation, a year lasts from 1st January to 31st December in the northern hemisphere and from 1st July to 30th June in the southern hemisphere (Frich et al., 2002).

Observational GSL data on a $2.5^\circ \times 3.75^\circ$ latitude–longitude grid are obtained from HadEX2, which provides gridded land-based data of a variety of temperature and precipitation indices, many of which are pertinent to extremes. The data can be downloaded from www.climdex.org.

GSL can be seen as an indicator for potential plant productivity, and we evaluate its value for estimating changes in plant growth. Fig. 2 (bottom) shows correlations of observed annual GSL values and corresponding satellite-derived Normalized Difference Vegetation Index (NDVI) values from the Advanced Very High Resolution Radiometer (AVHRR), which is an indicator of vegetation greenness (see Supplementary Fig. A1 for correlations with an alternative NDVI dataset). In the extratropics, the correlations are relatively good. As the agreement is especially high north of 40°N, we use these areas (Asia, Europe, America) indicated with boxes in Fig. 2 for our temporo-spatial detection and attribution analysis (see Section 5). In temperate climates, wheat is usually grown as a rain-fed crop (FAO homepage, 2014). As such, water is usually not limited and warm temperatures are likely more important than rainfall for crop growth. We therefore select the northern and southern temperate zones (north of 25°N and south of 25°S, extratropics) for our analyses, i.e. the entire extratropical region.

Even though a relatively small area of the surface is used to grow wheat and maize (Section 2), from a temperature perspective alone, a much larger fraction of the land surface is suitable to grow wheat and maize. Fig. 3 shows the percentage of extratropical land area with GS longer than 100, 150, 200, 250 and 300 days. These GSL thresholds encompass a range of lengths needed to grow wheat and maize: 100–130 days for spring wheat, 180–250 days for winter wheat, and 100–200 days for maize (FAO homepage, 2014). The change in land area with GSL of 250–300 days is particularly important as it potentially allows double-cropping of winter wheat ($\sim$200 days) and summer maize ($\sim$100 days) if enough water and sun light are available. These areas increase from 35% of extratropical land areas to 38% (250 days) and 27% to 28% (300 days) over our 50-year analysis period. Areas for all thresholds show a positive trend over the time-period 1956–2005, but are largest for the shorter thresholds, with an increase from 96 to 97% and 53 to 60% for the 100 day and 200 day thresholds, respectively. This might be related to larger warming in cold areas (northern high-latitudes) than warmer areas (Seneviratne et al., 2014), which is confirmed by the extension of areas with GS longer than 150 days into the northern latitudes in Supplementary Fig. A2.

4. Model data and data processing

Climate model simulations were obtained from the Coupled Model Intercomparison Project Phase 5 (CMIP5) multi-model ensemble (Taylor et al., 2012) for the years 1956–2005. We considered temperature from CMIP5 simulations from three experiments: (1) changes in

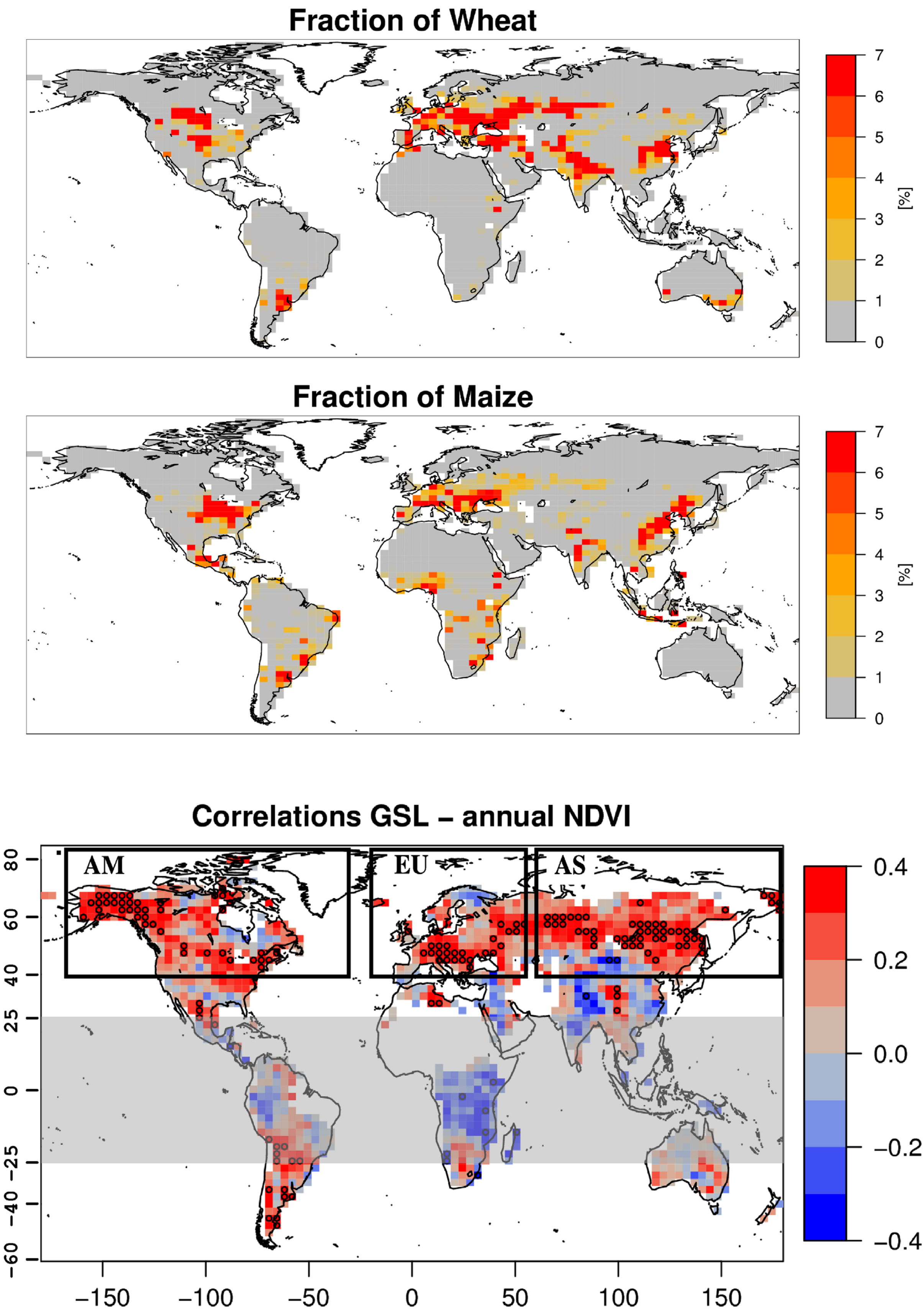

Fig. 2. Percentage of land-areas covered with wheat (top) and maize (middle, sum of maize for human consumption and animal feeding). Units are percentage of grid-cells. Data from Monfreda et al. (2008). Correlation of GSL-anomaly and annual (January to December in the northern hemisphere and July to June in the southern hemisphere) averaged NDVI-anomaly (AVHRR) from 1981 to 2002 (bottom). Pearson correlation (colors) and two-sided Pearson-test (stippling where significant at 5%-level). For the detection and attribution analyses, we consider the northern and southern hemispheric extratropics (regions not overlaid with gray) as well as the three regions indicated with boxes (Asia (AS), Europe (EU) and North America (AM)). (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

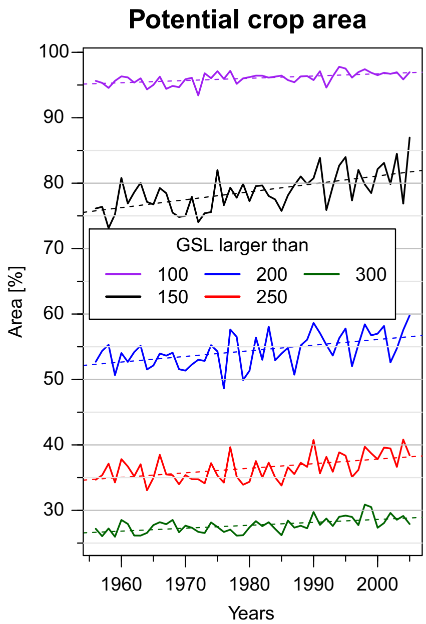

Fig. 3. Percentage of extratropical land area with observed GS longer than 100, 150, 200, 250 and 300 days. For necessary GSL for wheat and maize see the main text. Long GS (more than 250 and 300 days) would potentially allow double-cropping.

natural (NAT) forcing only, i.e. volcanic aerosols and solar output, (2) changes in greenhouse gases only (GHG), and (3) the combined effect of all forcings (ALL), i.e. natural and anthropogenic forcings, which include GHG, ozone and aerosols as well as land-use change for some of the models. We consider a total of 85 simulations under ALL, 36 under NAT and 38 under GHG forcing (see Supplementary Table A1). We further use 17,400 years of unforced control simulations from the 24 models listed in Supplementary Table A2 to obtain estimates of the internal (unforced) variability of GSL (see Section 5).

We select the time period 1956–2005 when data availability in the observations is relatively consistent over time (see Supplementary Fig. A3). Many of the forced model simulations that we use to evaluate observed GSL also end in 2005. We regrid all CMIP5 simulations and crop maps for wheat and maize onto the HadEX2-grid with a distance-weighted average remapping routine. In order to assure comparability between models and observations, we mask all model simulations with the data availability of the observations. We split the control simulations into 348 non-overlapping 50-year chunks and assign the years 1956–2005 to each 50-year chunk in order to mask it with observations. For each 50-year model segment (both forced and control), we calculate anomalies relative to the 1961–1990 climatology of that segment. Anomalies are also calculated for the observed record relative to the 1961–1990 climatology of the observations. We apply area-weighting when averaging the anomlies for individual grid boxes over regions. Lastly, we further reduce the size of our datasets for the detection and attribution analysis by calculating non-overlapping 5-year means of anomalies for each 50-year record (observations, forced runs and control segments). The dimension reduction is necessary to make the inversion of the co-variance matrices possible (see Section 5), thereby avoiding the need to use empirical orthogonal function (EOF) based dimension reductions. This helps us to increase the robustness of detection results and simplifies their interpretation.

The calculation of GSL describes the periods when individual annual cycles (as estimated from six-day moving averages of surface air temperature) remain persistently above a fixed threshold, 5 °C. Biases in model simulated surface air temperature, both in terms of the mean annual level and the progression of the annual cycle, would be expected to affect GSL calculated from model output. When evaluated relative to reanalysis data, the annual mean temperature bias of CMIP5 models (Flato et al., 2013) over land in the extratropical regions considered in this study is generally less than 2.5 °C, while the seasonality bias (bias in the contrast between summer and winter mean temperatures) is as large as 6 °C over parts of the northern hemispheric extratropical land area (larger than in reanalyses). Results are similar for a comparison of CMIP5 temperature to station data (Mueller and Seneviratne, 2014). CMIP5 models exhibit small cold biases over the year in all the regions relevant in our study, and temperatures are underestimated in cold months and slightly overestimated in warm months (Mueller and Seneviratne, 2014). The fact that models tend to have higher amplitude annual cycles than observational data suggests that they could slightly under-simulate GSL changes that result from external forcing, i.e. from an increase in annual mean temperatures. There is evidence that models may overestimate surface air temperature variability over extratropical land areas in summer (Cattiaux et al., 2013), suggesting that the natural internal variability of GSL may also be somewhat overestimated by models. Both aspects, potential signal and variance bias, have some implications for detection and attribution that are reflected in Sections 5 and 6.

5. Detection and attribution methodology

Our aim is to quantify the effects of external, anthropogenic influences on changes in GSL. To this end, a formal optimal detection technique can be used to assess how well model fingerprints from simulations with different forcings agree with observations. These model fingerprints are estimated by averaging 85 ALL, 36 GHG and 38 NAT simulations, which reduces the extent to which fingerprints are affected by internal climate variability. There is an implicit assumption that the inter-model differences have the same characteristics as the inter-simulation differences from the same model, i.e. that all differences between simulations are due to internal variability only. This is not completely true, and methods to account for these uncertainties remain an active area of research. To account for remaining errors in the fingerprints due to internal variability, we chose the total least squares regression algorithm in the optimal fingerprint detection method (Allen and Stott, 2003). We perform analyses using individual model fingerprints (one-signal analysis) and two fingerprints in combination (two-signal analysis).

In the one-signal analysis, the observations are separately regressed onto signals under ALL, GHG and NAT forcing estimated from CMIP5 simulations. In order to quantify the relative importance of different external forcings, we also perform a two-signal analysis, regressing the observations simultaneously onto signals under anthropogenic (ANT) and NAT forcing. We construct the ANT signal by subtracting the NAT signal from the ALL signal.

We employ the regression model

$$y = (X - V)*\beta + u, \tag{1}$$

where y is a vector of observations (anomalies), X a matrix composed of one or two model fingerprints (anomalies), and β a vector of scaling factors. The vector u contains regression residuals and represents internal climate variability, and the matrix V represents the remaining effects of internal climate variability on the model fingerprints. It is assumed that u and V are Gaussian random vectors. Their covariance structures differ only in magnitude, with the magnitude of the covariance of V being reduced relative to that of u in proportion to the number of climate simulations used to estimate X. To obtain the scaling factors, an estimate of the covariances of u and V and their inverses is needed. We estimate the covariance from unforced control simulations. To assess whether the model simulated variability from control simulations is similar to that which is observed (the residuals), we employ the residual consistency check (RCC) introduced by Allen and Tett (1999). The inversion of the covariance matrices can be problematic if the sample of internal variability is too small. Therefore, only the most important principal components of the covariance matrix are retained in many studies (EOF-truncation). The large number of samples of internal variability used in our study, however, ensures that the noise covariance matrix is invertible and thus an EOF-truncation is not necessary. The uncertainty of the scaling factor is estimated based on internal climate variability (Allen and Stott, 2003), using the method introduced by Ribes et al. (2013). We split the ensemble of control simulations into two halves for signal-to-noise optimization and for testing (see, e.g., Zhang et al., 2007). For an overview on detection and attribution methodologies, the reader is referred to Hegerl and Zwiers (2011).

A resulting scaling factor β that is consistent with unity implies a good match between model simulations under the respective forcing and observations. Detection of a change in GSL can be claimed if scaling factors are not consistent with zero and the RCC-test is passed.

Combining three regions (Asia, Europe, America) into one vector for the detection and attribution analysis allows detection of changes in the large-scale spatio-temporal evolution of GSL over wheat and maize areas. We also perform a temporal-only detection analysis by considering only the time evolution of the spatial mean calculated over the entire extratropical domain.

An analysis of the consistency of observed and model variability is given in Fig. 4. The variability of GSL simulated under ALL forcing is similar to the observed variability in extratropical wheat and maize regions, but slightly higher in the wheat region. Under GHG forcing, model simulated variability is overestimated. The slight overestimation of variability in the wheat region in simulations under ALL forcing is not unexpected (see Section 4) and can affect the RCC-test. We will therefore decrease the model variability slightly for the detection and attribution analyses but will also show results with the original variability.

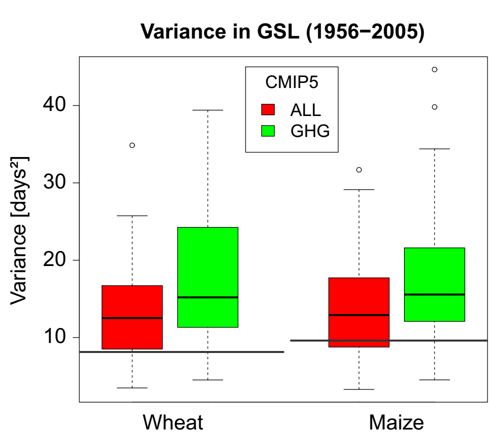

Fig. 4. Comparison of variances in models and observations: the observed variances of 5-year GSL anomalies in extratropical regions are indicated by long horizontal bars. Model simulated variances of 5-year GSL anomalies are indicated by box and whiskers plots under ALL (red) and GHG (green) forcing. The lower and upper ends of the box indicate the 25th and 75th percentiles of the variances, respectively, while the horizontal bar within the box indicates the median. The two horizontal bars indicate the range that covers 90% of the variances. Outliers are indicated with empty circles. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

6. Changes in growing season length and their detection and attribution

Global trends in GSL from 1956 to 2005 are shown in Fig. 5. The observations show a lengthening of the GS in most global land areas. The strong negative trend in observed GSL in South America is related to data-issues in that region (Lisa Alexander, UNSW, personal communication). CMIP5 simulations under ALL and GHG forcing agree well with observations. If only NAT forcing is considered, trends in GSL are not well reproduced by the models. Grid points with trends significant at the 5% level are stippled. Note, however, that the significance cannot be compared between observations and simulations as the model-trends were calculated from the model-mean, i.e. the variability is reduced and trends are more likely to be significant. Assessing the significance of trends from individual model simulations shows that observations exhibit significant trends at 34.3% of grid points, and that individual CMIP5 simulations under ALL, GHG and NAT forcing do so at 17.9, 32.3 and 5.1% of grid points on average, respectively (see Supplementary Fig. A4). The lower fraction of significant trends in ALL compared to GHG simulations reflect the counter-effect (cooling effect) of aerosols forcing which is only included in the ALL forcing, as well as possible modest underestimation of the forced response and overestimation of variability as discussed in Section 4.

As there is a good agreement of observational and model simulated trends when anthropogenic forcings are considered and poor agreement when they are not, we use a detection and attribution analysis based on Eq. (1) to test the hypothesis that changes in GSL can only be explained if anthropogenic forcing is considered. Fig. 6 shows scaling factors over the wheat area and the maize area, for the extratropical region (1-region, temporal detection) and for the 3-regions combined (Asia, Europe, North America, i.e. temporo-spatial detection). The best estimates of the scaling factors are displayed with dots and the 5 and 95% confidence intervals with vertical bars. Results of the RCC-test are also shown in Fig. 6: if the confidence intervals of the scaling factors are represented with dashed lines, the RCC-test is not passed, while full lines indicate that the RCC-test is passed. There is some evidence from the RCC-test for the two-signal analysis (see later) and the variance evaluation (Fig. 4) that the models somewhat overestimate internal variability (see also Section 4), and results are thus improved when reducing the internal variability by 20% (multiplying by 0.8). Results presented in the main manuscript reflect this modest reduction of internal variability. Supplementary Fig. A5 shows results obtained when using the original model variability.

Scaling factors and confidence intervals including one but excluding zero can be found in the extratropics for ALL forcing and in the 3-region-setup for ALL and GHG (Fig. 6). They indicate that the model simulated responses under ALL and GHG forcing are detected (distinguishable from internal variability) and consistent with observations. The scaling factors for the GHG signals are smaller than for the ALL signal, which reflects the lack of the negative impacts of aerosols and volcanic forcing on temperature and subsequently, GSL. In all regional setups, simulations under NAT forcing do not represent the observations well, which can be seen in scaling factors that are consistent with or below zero and large confidence intervals.

The detection and attribution analysis shows that the lengthening of the GS over wheat and maize regions are distinguishable from internal variability alone. The scaling factors further show that the changes cannot be explained by natural forcing alone, but can be explained by the combined effect of natural and anthropogenic forcing. Greenhouse gas forcing alone can also explain changes in GSL in most setups of the analysis except for extratropical wheat areas, where the RCC-test is not passed. The results

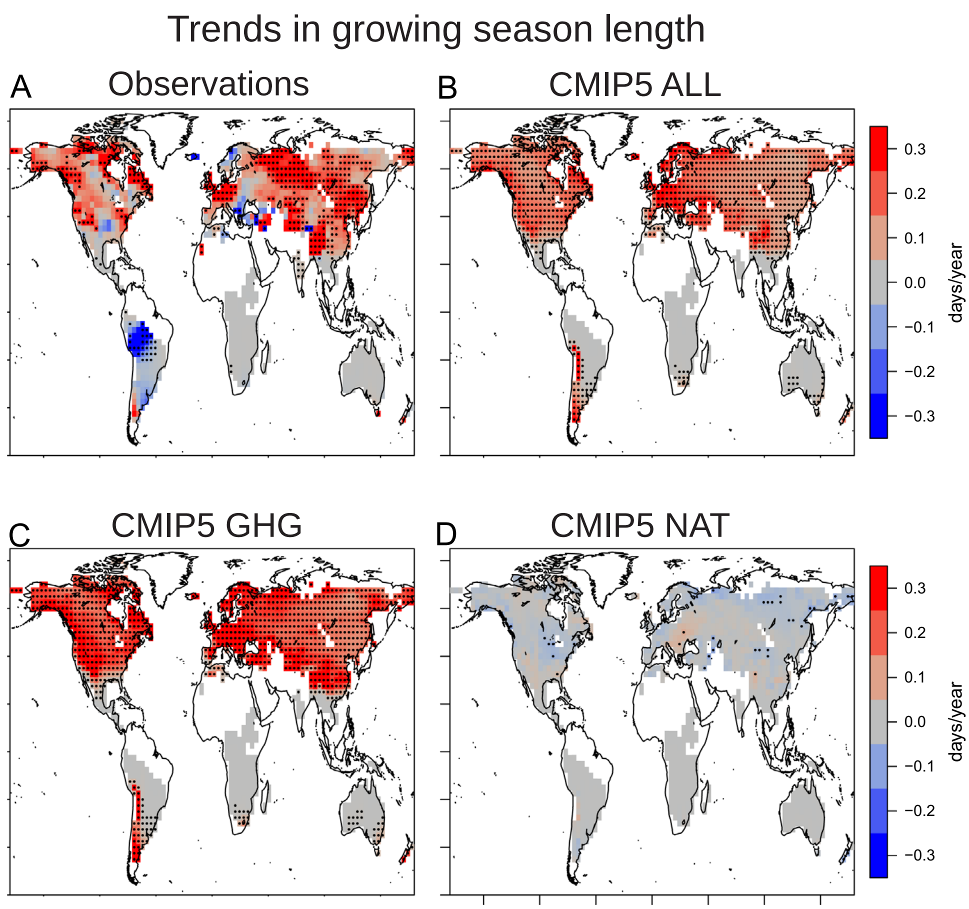

Fig. 5. Trend in GSL 1956–2005 in observations (A) and CMIP5-simulations under ALL forcing (B), GHG forcing (C) and NAT forcing (D). Trends significant at the 5%-level are stippled. For more analyses of trend significance, see Supplementary Fig. A4.

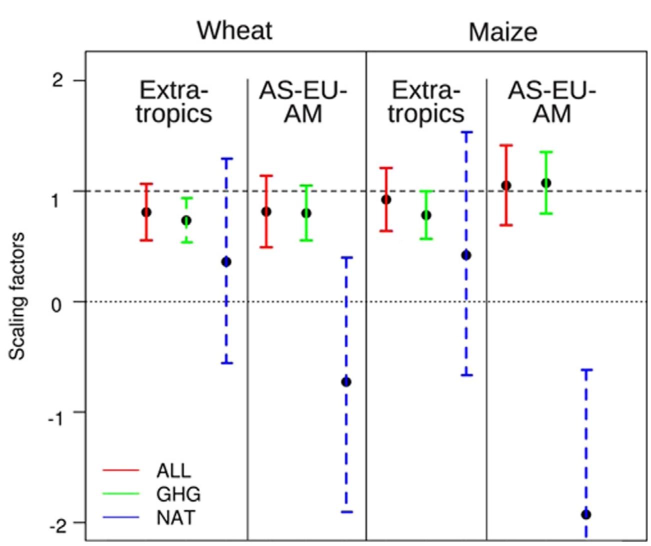

Fig. 6. Scaling factors from one-signal detection and attribution analyses: scaling factors for CMIP5-simulations with ALL, GHG and NAT forcing over wheat and maize regions for temporal detection (extratropical region) and temporo-spatial detection (3 northern hemispheric regions AS, EU, AM). Best estimates are displayed with dots and the 5–95%-confidence intervals with full lines if the residual consistency test is passed and with dashed lines if not. The internal variability of the models has been reduced by 20% (see main text). For results with original variability, see Supplementary Fig. A5.

are similar when using the original internal variability (Supplementary Fig. A5) except that the RCC is not passed in any of the setups with GHG alone. Failing the RCC-test is expected in the GHG analysis as the differences between the observations and the GHG signal in the residuals stem from components other than internal variability.

Fig. 7 displays scaling factors for the two-signal analysis for the NAT and the ANT signal. The best estimate and confidence intervals are again shown with dots and bars, respectively. The scaling factors from the two-signal analysis support the result from the one-signal analysis. For all regional setups, the ANT signal can be detected, while the NAT signal cannot be detected. Furthermore, all RCC-tests are passed. The RCC-tests are not passed in the one-signal analysis of NAT, indicating that while the NAT forcing alone is not a good predictor for our regression model, the ANT and NAT signal simultaneously are. Two-signal detection results with the original internal model variability are shown in Supplementary Fig. A6. The confidence intervals are slightly larger, as expected, and the extratropical wheat region does not pass the RCC test, which can be explained by the overestimation of variability (Fig. 4). The scaling factor for the ANT signal is very close to unity in all two-signal analyses, indicating that the bias in model simulated GSL does not have a material impact on the estimates of ANT-forced GSL changes that were used as the multi-model ANT fingerprint.

7. Probability of extreme growing season length anomalies

We are also interested in how climate change has altered the probability of an unusually long or short growing season occurring in any given year. We therefore define extreme short or long GS as an event with a return period of 100 years in the unperturbed climate simulated by control runs. This corresponds roughly to GSL with an occurrence probability of 1% in a climate without any external influences. The thresholds for extratropical regions are defined as

$$x_T = \pm(\bar{x} + K_T * s_x), \tag{2}$$

with x_T being the GSL-anomaly x estimated from the control simulations for the return period of $T = 100$ year, $\bar{x}$ the mean and s_x the standard deviation of GSL-anomaly averaged over the respective region, and K_T the 99% quantile for a unit normal (corresponding to a 1% probability). The resulting thresholds are 8.7 days for the extratropical wheat area and 10.1 days for the extratropical maize area. Our calculation assumes that GSL values are normally distributed, which is likely not the case. However, the 1% probability is just a way to define a threshold that can be easily understood, and results are qualitatively similar with 10 or 5 day thresholds.

Fig. 7. Scaling factors from two-signal detection and attribution analyses (ANT and NAT) for same regions as in Fig. 6 (extratropics and 3 northern hemispheric regions AS, EU, AM). The best estimate and the 5–95%-confidence intervals are shown. Full (dashed) lines indicate the residual consistency test is (is not) passed. The internal variability of the models has been reduced by 20%. For results with original variability, see Supplementary Fig. A6.

For the calculation of the exceedance probabilities in a world with anthropogenic forcing, we need an ensemble of GSL-timeseries. This ensemble is constructed in two steps. First, we calculate observationally constrained ALL and GHG signals by multiplying the multi-model ALL and GHG fingerprints by the best-estimates of the scaling factors obtained from the one-signal detection and attribution analyses (see Fig. 6, extratropical region). We use the scaling factor obtained from the ensemble-mean because including information from different models has been shown to decrease the error compared to individual simulations (Pierce et al., 2009). To account for the variability that is lost by averaging the timeseries to obtain the signals, we then add the 348 individual control simulations to the scaled ALL (or GHG, respectively) signals. With these sets of 348 observationally constrained timeseries under ALL and GHG forcing, we can calculate the exceedance probabilities displayed in Figs. 8 and 9. The shading denotes the confidence intervals of the probabilities based on GHG and ALL reconstructions obtained by multiplying the confidence interval of the scaling factors (5–95% intervals, bars in Fig. 6) with the GHG and ALL signal before adding the individual control simulations. The confidence intervals of the reconstruction are not exact because when estimating the scaling factors with the total least squares method, the signal itself is assumed to be uncertain. However, the error induced by using the original signal instead is expected to be minor as our estimation of the signal is relatively robust considering the number of models we use (see Supplementary Table A1). The probabilities of extremely long and short GS in the unforced runs (control simulations) and with NAT

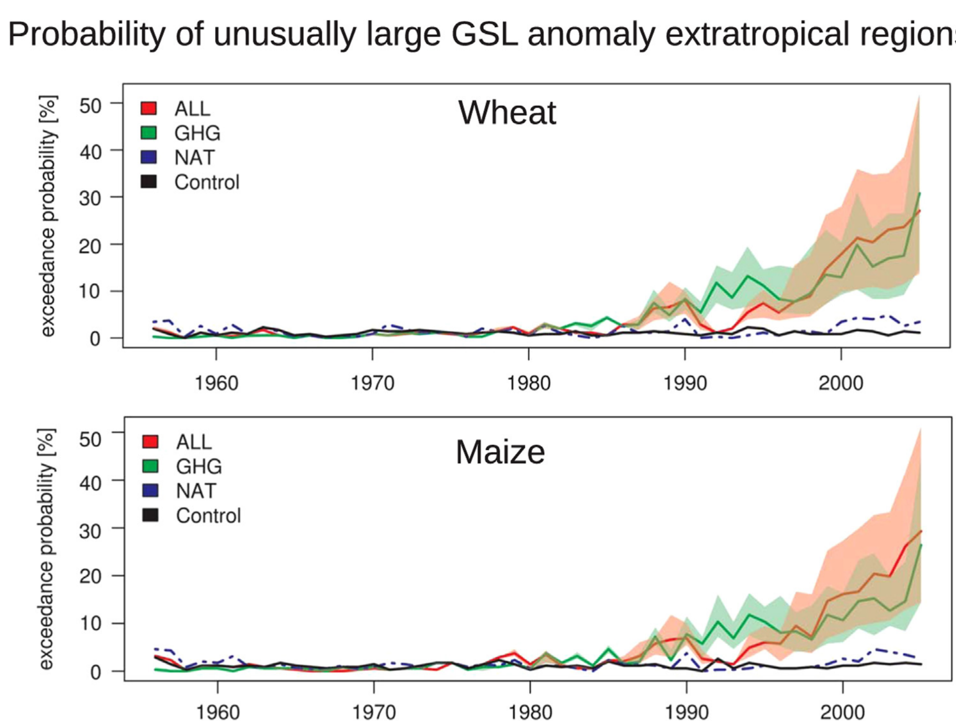

Fig. 8. Frequency of extremely long growing season length. The thresholds are defined as the 100-year return period in control simulations (8.7 days for extratropical wheat and 10.1 days for extratropical maize regions). Scaling factors from Fig. 6 are employed to reconstruct ALL and GHG.

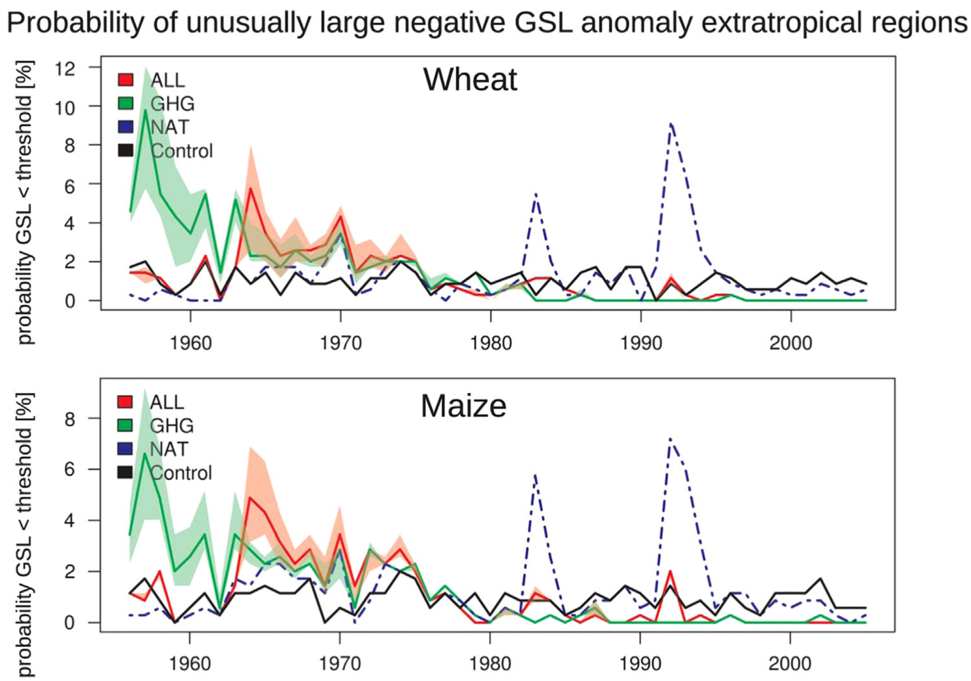

Fig. 9. Frequency of extremely short growing season. The thresholds are defined as the 100-year return period in control simulations (−8.7 days for extratropical wheat and −10.1 days for extratropical maize regions). Scaling factors from Fig. 6 are employed to reconstruct ALL and GHG.

forcing are also displayed in Figs. 8 and 9. They are calculated using the 36 original NAT simulations or the 348 55-year chunks from control simulations (neither is observationally constrained).

The probability of an extremely large GS-anomaly (Fig. 8) is similar in the control climate as in the climate with natural forcing (around 1%), which is expected. In a world with anthropogenic forcing (ALL, GHG), extremely long GS is much more frequent than in an undisturbed world after around 1985. In the early 1990s, the probability for long GS is again reduced in a climate with ALL forcing compared to GHG. The reason for this difference is likely the natural forcing from the eruption of Pinatubo in 1991, resulting in cooler temperatures and decreased GSL, which can be seen in the ALL world but not in the GHG world. The results show a very strong influence of human activity on the length of the GS. Our results suggest that over the 1956–2005 period, the probability of extremely long GS has changed from about 1% to 25% in both wheat and maize areas due to anthropogenic climate forcing.

Fig. 9 shows the probability of extremely short GS. The probability of short GS decreases over the analyzed time period in a world with only GHG forcing. In the more recent period, extremely short GS has become very unlikely due to increased atmospheric greenhouse gas concentrations (ALL and GHG), which can have a positive impact on wheat and maize yields. With ALL and NAT forcing, the probability is relatively small before 1963. Because GSL-anomalies are calculated relative to the base-period 1961–1990 corresponding to each forcing, the absolute probabilities cannot be compared between the individual forcings. Thus, the relatively low probabilities in ALL and NAT prior to 1963 could be due to the increase in probabilities after the Agung eruption in 1963 and the El Chichon eruption in 1982 if it affected the level of the entire time-series. Probabilities under GHG do not show this relative increase after 1963, indicating a natural influence on the climate system seen in ALL and NAT. The probabilities are consistent with GSL timeseries over extratropical regions (Supplementary Fig. A7). While simulations under ALL forcing show a slight decrease in GSL from 1956 to 1962 followed by a shortening and then lengthening of the GS, GHG simulations show a steady increase over the entire time period (Supplementary Fig. A7).

In a world with only natural external forcing, the likelihood of extremely short GS increases substantially after volcanic eruptions, which can be seen clearly after the major eruptions in 1982 (El Chichon) and 1991 (Pinatubo, Fig. 9). The effect of these volcanoes can also be seen in the ALL forcing simulations, although these probabilities are reduced compared to NAT due to the counter-effect from GHG.

8. Summary and conclusions

Wheat and maize are two of the most important crops worldwide. Here, we analyze how the length of the growing season (calculated based on temperature only) has changed over areas where wheat and maize are grown. We compare observations with global climate model simulations driven with different forcings, either natural forcing only (NAT), greenhouse gas forcing only (GHG), or all anthropogenic forcing combined with natural forcing (ALL). A spatial comparison of the trends in observations and in model simulations shows a lengthening of the growing season from 1956 to 2005 and strong agreement in most areas of the world for observations and simulations under ALL and GHG forcing. When model simulations are only forced with natural forcing, the growing season length does not show a trend. We test the hypothesis that anthropogenic forcing contributed to changes in growing season length with a formal detection and attribution analysis. We find that in the selected areas over the extratropics as well as over three northern hemispheric regions where wheat and maize are grown, model simulations under ALL and GHG forcing can explain changes in observed growing season length well, while NAT forcing alone cannot.

These results from the detection and attribution analysis strengthen previous results from Christidis et al. (2007) by using a more comprehensive climate model ensemble (CMIP5) and focusing on areas with particular crops. Our analyses are performed over regions that cover relatively small land areas in total but are dispersed relatively widely within continents. The narrow confidence intervals and the fact that the residual consistency test is passed in all regional setups when using simulations under ALL forcing provides strong confidence that detection and attribution of temperature-based indices may be possible for such areas, provided that a large number of model simulations are considered. It should be noted, however, that detection and attribution of changes in small geographically contiguous regions remains a considerable challenge.

We further estimate how much anthropogenic climate change has increased the probability of extremely long growing seasons. We define unusually long growing season length as that which corresponds to a 1 in a 100 year event (probability of 1%) in a world without any external forcing. We estimate that the probability of such an event had increased to about 25% by 2005 due to increasing anthropogenic greenhouse gas concentrations for areas where wheat and maize are grown. Correspondingly, the probabilities of an extremely short growing season are strongly reduced with increasing greenhouse gas concentrations.

Our results suggest that anthropogenic greenhouse gases have changed growing season length significantly. Whether yields increase or decrease with longer growing seasons is still unclear (Porter et al., 2014). Some locations that are currently too cold to grow wheat and maize may benefit from higher temperatures and longer growing seasons. Another important aspect is the possibility of double-cropping, i.e. having multiple harvests from the same field each year. We studied the changes in the size of areas with growing seasons longer than 250 or 300 days and found that they have increased from 35 to 38% and 27 to 28%, respectively, of extratropical land-area over the last 50 years. Provided that water and light availability are not limiting factors, double-cropping could be possible over larger areas than currently implemented (e.g., only around 0.4% of US land or 2% of US crop area, Brochers et al., 2014). In warmer regions, the growing season of crops can, however, end prematurely with very high temperatures. The maximum temperature for crop growth depends on the stage of growth as well as the cultivars (Porter and Gawith, 1999; Luo, 2011). For wheat, Porter and Gawith (1999) report a lethal maximum temperature of 47.5 °C and a temperature of 37 °C beyond which growth stops. This upper threshold for growth has so far been exceeded in only a very small percentage of the northern hemispheric (polewards of 40°N) wheat and maize areas (see Supplementary Fig. A8). The corresponding area in extratropical wheat and maize regions, polewards of 25° South and North, is substantially larger and nearly reaches 50% in some years. Thus longer growing seasons in these regions do not necessarily translate into an increase in yield, and with further climate change, extremely high temperatures may limit crop growth in more areas in the future. An earlier sowing date can help mitigate these negative impacts of climate change since it allows harvesting before summer temperatures reach their maximum, and thus, crops can actually take advantage of the longer growing season. Earlier sowing dates have been proposed to ensure continued high yields (Liu et al., 2013). Losses due to extreme temperatures may also be offset by a reduced risk of crop losses due to reduced probabilities of extremely short growing seasons as found in our analyses.

An increase in mean temperatures has been shown to reduce the period of actual growth from sowing date to crop maturity due to faster crop development (Wang et al., 2009). Our definition of growing season length is based only on temperature and therefore does not account for changes in the speed of crop phenological processes, nor other physiological effects.

In addition to high temperatures, negative impacts on agricultural output may also result from decreases in soil moisture and water availability in many important crop areas (Dorigo et al., 2012; Greve et al., 2014). A further reduction in water availability would decrease the viability of rain-fed crops and increase the demand for irrigation or reduce productivity in regions where not enough water is allocated or no infrastructure is available for irrigation. The productivity of rain-fed crops will additionally be affected by drought related hot temperatures due to the lack of evaporative cooling (see, e.g., Lobell and Gourdji, 2012).

Whether high yields can be sustained or increased in a warming world also depends largely on the crop sensitivity to heat days (Lobell et al., 2012). Conventional breeding approaches for heat and drought tolerant plants as well as the introduction of genes with relevant traits can complement agricultural strategies to offset the negative impacts of climate change (Reynolds et al., 2010; Bita and Gerats, 2013). An increase in productivity due to the increasing length of the growing season based on temperature is only possible with major adaptation which includes moving crop fields to colder climates or introducing heat- and drought tolerant varieties.

Acknowledgments

We acknowledge the World Climate Research Programme (WCRP) and the International Centre for Theoretical Physics (ICTP), which have supported this research through the WCRP-ICTP Summer School on Extremes (2014). We are very grateful to Qiuzi Han Wen from Environment Canada for helping to start this project and providing support for the analyses. We also thank Yang Feng from Environment Canada for helping with data preparation and R-codes and Xuebin Zhang for his useful comments. We acknowledge the WCRP's Working Group on Coupled Modelling, which is responsible for CMIP, and we thank the climate modelling groups for producing and making available their model output. For CMIP the US Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals. MODIS NDVI data are originally distributed by the Land Processes Distributed Active Archive Center (LP DAAC), located at the U.S. Geological Survey (USGS) Earth Resources Observation and Science (EROS) Center (lpdaac.usgs.gov), and were distributed in netCDF format by the Integrated Climate Data Center (ICDC, http://icdc.zmaw.de) University of Hamburg, Hamburg, Germany. All data have been analyzed using the software R version 3.1.0 (R Development Core Team, 2008).

Appendix A. Supplementary data

Supplementary data associated with this paper can be found in the online version at http://dx.doi.org/10.1016/j.wace.2015.04.001.

References

Allen, M.R., Stott, P.A., 2003. Estimating signal amplitudes in optimal fingerprinting, part I: theory. Clim. Dyn. 21 (5–6), 477–491. http://dx.doi.org/10.1007/s00382-003-0313-9.

Allen, M.R., Tett, S.F.B., 1999. Checking for model consistency in optimal fingerprinting. Clim. Dyn. 15 (6), 419–434. http://dx.doi.org/10.1007/s003820050291.

Bita, C., Gerats, T., 2013. Plant tolerance to high temperature in a changing environment: scientific fundamentals and production of heat stress-tolerant crops. Front. Plant. Sci. 4 (273), http://dx.doi.org/10.3389/fpls.2013.00273.

Brochers, A., Truex-Powell, E., Wallander, S., Nickerson, C., 2014. Multi-cropping practices: recent trends in double-cropping. Econ. Inf. Bull. 125,22, 22.

Cattiaux, J., Douville, H., Peings, Y., 2013. European temperatures in CMIP5: origins of present-day biases and future uncertainties. Clim. Dyn. 41 (11–12), 2889–2907.

Cheng, C., Lei, C., Deng, A., Qian, C., Hoogmoed, W., Zhang, W., 2011. Will higher minimum temperatures increase corn production in Northeast China? An analysis of historical data over 1965–2008. Agric. For. Meteorol. 151, 1580–1588.

Christidis, N., Stott, P., Brown, S., Karoly, D., Caesar, J., 2007. Human contribution to the lengthening of the growing season during 1950–99. J. Clim. 20, 5441–5454. http://dx.doi.org/10.1175/2007JCLI1568.

Donat, M.G., et al., 2013. Updated analyses of temperature and precipitation extreme indices since the beginning of the twentieth century: the HadEX2 dataset. J. Geophys. Res. Atmos. 118 (5), 2098–2118. http://dx.doi.org/10.1002/jgrd.50150.

Dorigo, W., de Jeu, R., Chung, D., Parinussa, R., Liu, Y., Wagner, W., Fernandez-Prieto, D., 2012. Evaluating global trends (1988–2010) in harmonized multi-satellite surface soil moisture. Geophys. Res. Lett. 39, 18. http://dx.doi.org/10.1029/2012GL052988, L18,405.

FAO homepage, 2014. ⟨http://www.fao.org/nr/water/cropinfo.html⟩ (accessed 15 September 2014).

FAO, IFAD, WFP, 2014. The State of Food Insecurity in the World 2014. Strengthening the Enabling Environment for Food Security and Nutrition. Technical Report, Rome, FAO.

FAO statistic, 2014. ⟨http://faostat.fao.org ⟩ (accessed 26 September 2014).

Flato, G., et al., 2013. Evaluation of climate models. In: Stocker, T.F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S.K., Boschung, J., Nauels, A., Xia, Y., Bex, V., Midgley, P. M. (Eds.), Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom, New York, NY, USA.

Foley, J.A., et al., 2011. Solutions for a cultivated planet. Nature 478 (7369), 337–342. http://dx.doi.org/10.1038/nature10452.

Frich, P., Alexander, L.V., Della-Marta, P., Gleason, B., Haylock, M., Tank, A.M.G.K., Peterson, T., 2002. Observed coherent changes in climatic extremes during the second half of the twentieth century. Clim. Res. 19 (3), 193–212. http://dx.doi.org/10.3354/cr019193.

Gornall, J., Betts, R., Burke, E., Clark, R., Camp, J., Willett, K., Wiltshire, A., 2010. Implications of climate change for agricultural productivity in the early twenty-first century. Philos. Trans. R. Soc. B: Biol. Sci. 365 (1554), 2973–2989.

Greve, P., Orlowsky, B., Mueller, B., Sheffield, J., Reichstein, M., Seneviratne, S., 2014. Global assessment of trends in wetting and drying over land. Nat. Geosci. 7, 716–721. http://dx.doi.org/10.1038/ngeo2247.

Hegerl, G., Zwiers, F., 2011. Use of models in detection and attribution of climate change. Wiley Interdiscip. Rev. Clim. Change 2 (4), 570–591. http://dx.doi.org/10.1002/wcc.121.

Liu, Z., Hubbard, K.G., Lin, X., Yang, X., 2013. Negative effects of climate warming on maize yield are reversed by the changing of sowing date and cultivar selection in Northeast China. Glob. Change Biol. 19 (11), 3481–3492.

Lobell, D.B., Sibley, A., Ivan Ortiz-Monasterio, J., 2012. Extreme heat effects on wheat senescence in India. Nat. Clim. Change 2 (3), 186–189. http://dx.doi.org/10.1038/nclimate1356.

Lobell, D.B., Gourdji, S., 2012. The influence of climate change on Global crop productivity. Plant Physiol. 160, 1686–1697. http://dx.doi.org/10.1104/pp.112.208298.

Luo, Q., 2011. Temperature thresholds and crop production: a review. Clim. Change 109 (3–4), 583–598. http://dx.doi.org/10.1007/s10584-011-0028-6.

McClung, C.R., 2014. Making hunger yield. Science 344 (6185), 699–700. http://dx.doi.org/10.1126/science.1254135.

Monfreda, C., Ramankutty, N., Foley, J.A., 2008. Farming the planet: 2. geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000. Glob. Biogeochem. Cycles 22 (1), GB1022. http://dx.doi.org/10.1029/2007GB002947.

Mueller, B., Seneviratne, S., 2014. Systematic land climate and evapotranspiration biases in CMIP5 simulations. Geophys. Res. Lett. 41 (1), 128–134. http://dx.doi.org/10.1002/2013GL058055.

Pierce, D.W., Barnett, T.P., Santer, B.D., Gleckler, P.J., 2009. Selecting global climate models for regional climate change studies. Proc. Natl. Acad. Sci. USA 106 (21), 8441–8446. http://dx.doi.org/10.1073/pnas.0900094106.

Porter, J.R., Gawith, M., 1999. Temperatures and the growth and development of wheat: a review. Eur. J. Agronomy 10 (1), 23–36.

Porter, J.R., Xie, L., Challinor, A.J., Cochrane, K., Howden, S.M., Iqbal, M.M., Lobell, D.B., Travasso, M.I., 2014. Food security and food production systems. In: Field, C.B., et al. (Eds.), Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel of Climate Change. Cambridge University Press, Cambridge, UK and New York, USA, pp. 485–533.

R Development Core Team, 2008. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0.

Ray, D.K., Ramankutty, N., Mueller, N.D., West, P.C., Foley, J.A., 2012. Recent patterns of crop yield growth and stagnation. Nat. Commun. 3, 1293. http://dx.doi.org/10.1038/ncomms2296.

Reynolds, M., Hays, D., Chapman, S., 2010. Breeding for adaptation to heat and drought stress. In: Reynolds, M.P. (Ed.), Climate Change and Crop Production. CAB International, London, UK, pp. 71–91.

Ribes, A., Planton, S., Terray, L., 2013. Application of regularised optimal fingerprinting to attribution. Part I: method, properties and idealised analysis. Clim. Dyn. 41 (11–12), 2817–2836. http://dx.doi.org/10.1007/s00382-013-1735-7.

Seneviratne, S.I., Donat, M.G., Mueller, B., Alexander, L.V., 2014. No pause in the increase of hot temperature extremes. Nat. Clim. Change 4 (3), 161–163. http://dx.doi.org/10.1038/nclimate2145.

Siebert, S., Portmann, F., Doell, P., 2010. Global patterns of cropland use intensity. Remote Sens. 2, 1625–1643. http://dx.doi.org/10.3390/rs2071625.

Taylor, K.E., Stouffer, R.J., Meehl, G.A., 2012. An overview of CMIP5 and the experiment design. Bull. Am. Meteorol. Soc. 93 (4), 485–498. http://dx.doi.org/10.1175/BAMS-D-11-00094.1.

Wang, J., Wang, E., Luo, Q., Kirby, M., 2009. Modelling the sensitivity of wheat growth and water balance to climate change in Southeast Australia. Clim. Change 96 (1–2), 79–96.

Zhang, X.B., Zwiers, F.W., Hegerl, G.C., Lambert, F.H., Gillett, N.P., Solomon, S., Stott, P.A., Nozawa, T., 2007. Detection of human influence on twentieth-century precipitation trends. Nature 448 (7152), 461-U4. http://dx.doi.org/10.1038/nature06025.