# Weatherproofing for a smarter, resilient and more sustainable agri-sector

A thesis submitted to the School of Environmental Sciences of the University of East Anglia in partial fulfilment of the requirements for the degree of Doctor of Philosophy

By Joanna Louise Raymond

100244494

3rd February 2023

Word count: 54,513

# Abstract


Agricultural production is highly vulnerable to climate change. Warming global temperatures, increased frequency of extreme weather events, such as flash flooding, and shifts in the ranges of crop pathogens highlight current agroclimate challenges faced in the 21st century. In the mid-1990s, the UK observed plateaus in yields of staple crops including wheat (*Triticum aestivum* L.), and in recent years, strong yield impacts of high interannual weather variability have been observed nationally. To meet the growing demand for food and increase domestic production, there is an urgent need to increase the climate resilience of key cereal crops and the broader agri-food sector.

This thesis explores how crop breeding and changes in weather and climate variables important to agriculture, i.e. the agroclimate, have contributed to cereal yield variability in the UK. In the first *State of the UK Agroclimate*, trends and variability in national yields and key agroclimate metrics are quantified for 1981-2020 to allow growers and farmers to make climate-informed decisions on crop and variety choices. Incorporating historical time-series records of these agroclimate metrics into statistical models with variety trials data establishes their relative importance in determining winter wheat yields and enables variety sensitivity to each metric to be investigated. In doing so, methods of identifying climate-resilient crop varieties are presented. The contribution of plant breeding to national cereal yields is quantified through the calculation of genetic gain. Given the importance of this metric for the evaluation of success of plant breeding programmes and for funding allocation, the sensitivity and robustness of this metric is explored.

Changes in the UK agroclimate provide both risks and opportunities for cereal growers. The increase in solar radiation during grain fill observed in the South-East of the UK was shown to be beneficial for winter wheat yields, whilst increasing interannual yield variability has contributed to overall stagnation in yields in the last decade. Long-term warming trends have contributed to cereal drilling dates getting earlier, extending the growing season. However, in years of high autumn rainfall, delays in planting, a shortening of the growing season and lower yields were observed. Comparing national and variety trial yield trends shows that crop breeding is responsible for over 95% of yield increases over the last 30 years, however, to further optimise yield increases there is a need to grow varieties that perform best based on local climatic conditions. Use of case study datasets extracted from the variety trials data showed that genetic gain is sensitive to a number of factors, including the choice and number of long-term “check” varieties included in a variety trials programme. Recommendations are made on how best to calculate the metric.

# Contents

Abstract .......... 2
Contents .......... 3
Figures .......... 7
Tables .......... 17
Acknowledgements .......... 23
1 The role of breeding and climate in UK cereal production, and the need to combine the two in analysis .......... 24
1.1 Threats and opportunities for agriculture in changing climates .......... 25
1.1.1 The challenges of achieving global food security .......... 25
1.1.2 Understanding recent yield trends .......... 26
1.1.3 Methods of overcoming yield plateaus .......... 29
1.1.4 The role of crop breeding in increasing yields .......... 29
1.1.5 The influence of farm management on crop production .......... 31
1.1.6 The carbon dioxide fertilisation effect .......... 32
1.2 The influence of a changing climate on UK agriculture .......... 32
1.2.1 The importance of crop breeding .......... 33
1.2.2 Genotype-by-environment interaction .......... 34
1.2.3 UK growing climate .......... 34
1.2.4 Recent climatic changes .......... 37
1.2.5 Crop-climate relationships in wheat and barley .......... 38
1.2.6 Changing crop-climate relationships .......... 42
1.3 Using climate information to support crop breeding and agriculture in a changing climate44
1.3.1 State of agroclimate services in the UK .......... 44
1.3.2 Agroclimate metrics as tools to monitor climatic changes affecting agriculture . 45
1.3.3 Opportunities provided by combined multi-environment trials and climate data 48
1.3.4 Isolating the contribution of breeding and climate to crop yields .......... 49
1.4 Summary and research aims .......... 50
2 Materials and Methods .......... 52
2.1 Datasets and methods for Chapter 3 .......... 52
2.1.1 Irish spring barley data .......... 52
2.1.2 Irish climate data .......... 53
2.1.3 Modelling Irish spring barley and climate data .......... 60

2.1.4 Variable selection methods ... 62
2.1.5 Principal Component Analysis ... 65
2.1.6 Pearson's correlation analysis ... 65
2.1.7 Akaike Information Criterion ... 65
2.1.8 Model assumptions ... 65
2.1.9 Comparison of standard error of difference between means ... 66
2.2 Datasets and methods for Chapter 4 ... 66
2.2.1 UK variety trials data ... 66
2.2.2 UK on-farm data ... 67
2.2.3 UK temperature and precipitation dataset selection ... 67
2.2.4 Additional UK climate data ... 70
2.2.5 Quality control of variety trials datasets ... 72
2.2.6 Crop yield anomaly analysis ... 78
2.2.7 Agroclimate metrics ... 79
2.3 Datasets and methods for Chapter 5 ... 83
2.3.1 Winter wheat certified seed area statistics ... 83
2.3.2 Modelling phenotype trends in UK variety trial data ... 83
2.3.3 Modelling the breakdown of disease resistance ... 84
2.3.4 Estimating uncertainty in genetic gain estimates ... 85
2.3.5 Quantifying the contribution of genetic improvement to national yields ... 86
2.4 Datasets and methods for Chapter 6 ... 87
2.4.1 Reclassifying soil texture ... 88
2.4.2 Modelling yield impacts of seasonal climate at a regional level ... 89
2.4.3 Modelling yield impacts of seasonal climate using site-specific climate data ... 91
2.4.4 Modelling yield impacts of monthly climate using site-specific climate data ... 91
2.4.5 Modelling the effect of the UK agroclimate on winter wheat yields ... 91
2.4.6 Identifying climate-resilient varieties ... 92
3 Using the case of variability in early 20th Century Irish barley yields to identify best suited statistical modelling methods ... 94
3.1 Spring barley yields show high variation ... 95
3.2 Spring barley price shows similar variation to yield ... 97
3.3 Irish climate analysis ... 97
3.3.1 Long-term climate reveals anomalous years ... 97
3.3.2 NAO index shows extremes in the period 1901-1906 ... 99
3.3.3 1891-1920 climatology reveals extreme wetness in 1903 ... 101
3.3.4 Climate anomalies encompassed parts of Europe ... 104

3.4 Modelling Irish spring barley and climate ... 107
3.4.1 Variable selection methods failed to significantly reduce model complexity ... 108
3.4.2 Insufficient independent data for Principal Component Analysis ... 112
3.4.3 Highest correlations between July temperature and rainfall and yield ... 114
3.4.4 Use of Akaike Information Criterion reveals important climate variables ... 114
3.5 Comparison of standard error of difference between means ... 115
3.6 Discussion ... 116
3.6.1 Climatic causes of yield variability ... 116
3.6.2 Statistical methods ... 117
3.6.3 Limitations ... 119
3.7 Conclusion ... 119
4 The State of the UK Agroclimate ... 121
4.1 Long-term production trends show increasing variability ... 121
4.2 There are regional deviations to national yield anomalies ... 125
4.3 Climate anomalies help explain anomalous yields ... 127
4.4 The Changing UK Agroclimate ... 137
4.4.1 Drilling day of year is earlier but harvest dates haven't changed ... 137
4.4.2 The Start of the Growing Season is getting earlier ... 140
4.4.3 The available Growing Degree Days has increased ... 141
4.4.4 Vernalisation Degree Days reveal coastal effect ... 143
4.4.5 Widespread decrease in April air frost days ... 146
4.4.6 Seasonal variation in the water balance ... 147
4.4.7 The distribution of wet days across the growing season has changed ... 148
4.4.8 Heat stress during anthesis and grain fill is still rare ... 151
4.4.9 Total solar radiation received during grain fill has increased in the East ... 152
4.4.10 Disease prevalence shows high interannual variability ... 157
4.5 Discussion ... 159
4.5.1 The value of agroclimate metrics ... 160
4.5.2 Creating an agroclimate tool for growers ... 163
4.6 Conclusion ... 163
5 Quantifying genetic drivers of yield variability of UK cereal crops ... 165
5.1 Crop breeding has continued to contribute to cereal yield increases ... 165
5.2 The breakdown of disease resistance is seen in untreated variety trial yields ... 171
5.3 Genetic gain estimates are susceptible to bias ... 176
5.4 Variety improvement contributes over 95% of improvements in national yields ... 178
5.5 Discussion ... 179

5.5.1 Disease resistance as a driver for treated and untreated yield trends .............. 179
5.5.2 Uncertainty in genetic gain estimates................................................................ 182
5.5.3 Limitations of using the UK National List/Recommended List dataset .............. 184
5.6 Conclusion........................................................................................................ 185
6 Identifying the key climate drivers of interannual yield variability in winter wheat in the UK 187
6.1 Summer rainfall significantly affects yield................................................................ 187
6.2 Site-specific data shows winter rainfall affects yields ................................................ 188
6.3 Monthly rainfall is more important than temperature in determining yield ............. 190
6.4 Growing Degree Days accounts for most winter wheat yield variation ..................... 192
6.5 Statistical modelling can be used to identify climate-resilient varieties .................... 195
6.6 Discussion........................................................................................................ 200
6.6.1 Site-specific climate data reveals additional winter wheat yield drivers ............ 200
6.6.2 Univariate agroclimate analysis as a tool for climate variable selection ............ 201
6.6.3 Combined agroclimate multivariate analysis...................................................... 202
6.6.4 Individual varietal response to climate variability .............................................. 203
6.7 Conclusion........................................................................................................ 205
7 Conclusions and recommendations...................................................................... 206
7.1 Thesis summary ................................................................................................ 206
7.2 Addressing the overarching research questions.................................................... 208
7.2.1 After the 'yield plateau' of the 1990s and 2000s, what do we now see emerging in yield records?........................................................................................................ 208
7.2.2 How can we use variety trial records to quantify the genetic contribution to recent yield trends?........................................................................................................ 208
7.2.3 Which of an array of new high-resolution climate datasets should we synthesize into our analysis in order to most effectively isolate the confounding impact of spatial and temporal climate variability?........................................................................ 209
7.3 Recommendations for stakeholders.................................................................... 210
7.4 Future research ................................................................................................ 213
7.5 Concluding remarks.......................................................................................... 215
Appendix................................................................................................................ 216
Supplementary figures............................................................................................ 216
Supplementary tables.............................................................................................. 239
Glossary ................................................................................................................ 246
References.............................................................................................................. 247

# Figures

Figure 1.1: Mean temperature, total rainfall and total sunshine duration for the UK in 2018. The maps have a 1km x 1km spatial resolution and can be downloaded from the Met Office (https://www.metoffice.gov.uk/research/climate/maps-and-data/uk-actual-and-anomaly-maps). ..................... 35

Figure 1.2: UK (black) wheat yields (t/ha) from 1999-2020 relative to the 1980-2020 mean. England (yellow), Wales (red), Scotland (blue) and Northern Ireland (green) yields also given for 1999-2019. Data from DEFRA (2021)..................... 37

Figure 1.3: The major disease and weather and climate risks to winter wheat and winter barley yields at various stages of the growing season. The diseases affect both winter crops, except for those indicated by WW = winter wheat and WB = winter barley. Growth stages from AHDB Cereals & Oilseeds (2018b)..................... 40

Figure 1.4: Agroclimate and food system indicators used by the U.S. Department of Agriculture to show how climate change is influencing U.S. agriculture (Walsh et al., 2020). ..................... 46

Figure 2.1: **a.** The location of Ireland relative to Europe and North Africa. **b.** Weather stations open between 1901-1906 and closest to barley trial sites (+). Stations with rainfall only (blue circle), temperature only (red x) and both rainfall and temperature data are shown. Growing season (March-August) average temperature (°C) (**c.**), total rainfall (mm) (**d.**) and surface photosynthetically active radiation (MJ $m^{-2}$) (**e.**) for 1901-1930, calculated using ERA-20C (Poli et al., 2016). ..................... 57

Figure 2.2: Birr Castle temperature station comparison for daily maximum (left) and minimum (right) temperature (°C), for the period of overlap 1880-1911. The degree of correlation between the two stations is shown as well as any bias relative to the 1:1 line (black). ..................... 58

Figure 2.3: Glasnevin Dublin temperature station comparison for daily maximum (left) and minimum (right) temperature (°C), for the period of overlap 1882-1952. The degree of correlation between the two stations is shown as well as any bias relative to the 1:1 line (black). ..................... 58

Figure 2.4: Met Office stations used for validating gridded weather datasets. ..................... 69

Figure 2.5: Pre-processing and quality control steps undertaken prior to any statistical analysis. The National Institute of Agricultural Botany (NIAB) provided the 1982-2006 trials data, whilst Agriculture and Horticulture Development Board (AHDB) provided the 2007-2018 data. .......... 73

Figure 2.6: Trial locations (red) for 2007-2018 for winter wheat (left), winter barley (centre) and spring barley (right) relative to their respective growing areas (ha per 5 km x  km square) in the 2010 Agricultural Census (EDINA 2022). The regions shown in England are the government regions, whilst for Scotland these are the Met Office regions. Data on growing areas of all three crops in Northern Ireland and Wales was not available. ............................................................... 76

Figure 2.7: Regions used in the analysis are based on DEFRA regions for England and Met Office regions for Scotland. ......................................................................................................... 78

Figure 2.8: The Vernalisation Effectiveness Factor for a range of mean temperatures ($^{o}$C), as used by Wu et al. (2017), to calculate Vernalisation Degree Days (VDD). .................................... 83

Figure 2.9: Classification of mineral soils into soil texture classes  (AHDB, 2019). ........................ 89

Figure 2.10: The 10 regions of the UK used in the regional seasonal analysis. Regions are defined by the standard areas used by the Met Office (Met Office, 2020). ............................................. 90

Figure 3.1: Barley trials yields (t/ha) (black dots) for 51 trials across 18 farms between 1901-1906, for two varieties: Archer (blue) and Goldthorpe (yellow). Outliers (diamonds) represent trial yields in the 5$^{th}$ and 95$^{th}$ percentiles. There were 51 trials per variety, increasing from 4 in 1901 to 12 in 1906. Data from Student (1923). ............................................................................ 96

Figure 3.2: Spring barley trial mean yields (t/ha) at the three farms with data for the entire period 1901-1906. Error bars show the difference between Archer and Goldthorpe variety yields. Data from Student (1923). ................................................................................... 96

Figure 3.3: Barley trial price per hectare (£/ha) (black dots) for 1901-1906 for two varieties: Archer (blue) and Goldthorpe (yellow). Outliers (diamonds) represent trial yields in the 5$^{th}$ and 95$^{th}$ percentiles. There were 51 trials per variety, increasing from 4 in 1901 to 12 in 1906. Data from Student (1923). ......................................................................................................... 97

Figure 3.4: Growing season (March-August) rainfall anomalies (mm) for Birr Castle, Foulkesmill and Roches Point stations and nationally for 1850-2010. Years 1901-1906 are shown in red. Here the national average anomaly is calculated using the Island of Ireland precipitation series from 25 stations. ........................................................................................................................ 98

Figure 3.5: Growing season Growing Degree Days (°C days) for Birr Castle, Roches Point and Glasnevin stations for 1874-2020. Growing Degree Days is the sum of the mean temperature on days when mean temperature is above 5.6°C from March to August. ........................................ 99

Figure 3.6: Winter NAO index (top), spring NAO index (middle) and summer NAO index (bottom) for the period 1824-2020. Period 1901-1906 highlighted in grey. Data from Jones et al. (1997). ........................................ 100

Figure 3.7: Late winter temperatures compared to the 1981-2010 average when the North Atlantic Oscillation (NAO) was strongly negative (top, Jan-March 2010) and when it was strongly positive (bottom, January-March 1990). When the NAO is negative, winters are often cooler than average across mid-latitudes. When the NAO is positive, they can be warmer than average. Image from NOAA Climate.gov (Lindsey and Dahlman, 2021). ........................................ 101

Figure 3.8: Cumulative monthly rainfall (mm) for Ardee, Birr Castle, Foulkesmill, Greenore, Roches Point stations and the national average across 25 stations for 1901, 1902, 1903, 1904, 1905 and 1906. The 1891-1920 average is shown (solid black line) along with the period 10th and 90th percentile values (grey lines) and the period minimum and maximum values (dashed black line). Note: Ardee station only has data for 1891-1913 therefore the averages are for this period instead. Data from Ryan et al. (2020). ........................................ 102

Figure 3.9: 1891-1920 monthly mean minimum temperatures (°C) for Birr Castle, Roches Point and Dublin (Glasnevin) stations for the growing season. The range in temperatures for the coldest 5%, coldest 10%, warmest 10% and warmest 5% mean minimum temperatures are shown. Monthly mean minimum temperatures for 1901, 1902, 1903, 1904, 1905 and 1906 are also presented. Data from (Mateus et al., 2020). ........................................ 103

Figure 3.10: 1891-1920 monthly mean maximum temperatures (°C) for Birr Castle, Roches Point and Dublin (Glasnevin) stations for the growing season. The range in temperatures for the coldest 5%, coldest 10%, warmest 10% and warmest 5% mean maximum temperatures are shown. Monthly mean maximum temperatures for 1901, 1902, 1903, 1904, 1905 and 1906 are also presented. Data from (Mateus et al., 2020). ........................................ 104

Figure 3.11: Growing season (March to August) rainfall anomalies (%) relative to the 1901-1930 average. Brown corresponds to drier than average and blue corresponds to wetter than average. 100 km x 100 km resolution data from ERA-20C (Poli et al., 2016). ........................................ 105

Figure 3.12**:** Growing season (March to August) total photosynthetically active radiation (PAR) anomalies (%) relative to 1901-1930 average. Purple corresponds to less PAR than average

whilst orange corresponds to more PAR than average. 100km x 100km resolution data from ERA-20C (Poli et al. 2016). .......... 106

Figure 3.13: Growing season (March to August) mean temperature anomalies (°C) relative to 1901-1930 average. Blue corresponds to colder than average whilst red corresponds to warmer than average. 100km x 100km resolution data from ERA-20C (Poli et al., 2016). .......... 106

Figure 3.14: Correlation between each monthly climate variable and yield for the 1901-1906 barley trials in Ireland. The dashed black lines show significant correlation ($p<0.05$, $n=102$). .. 114

Figure 4.1: National annual wheat yields (green dots), area (blue) and production total (black) for 1961-2020 (top) and 1980-2020 (bottom). The 5-year running mean yield (green dash) and quadratic (green) give an indication of long-term UK wheat yield trends. Spring and winter wheat are combined these plots due to the very low usage of spring wheat varieties. Data from FAOSTAT and DEFRA. .......... 122

Figure 4.2: National annual combined winter and spring barley yields (green dots), combined area (blue) and production total (black) for 1961-2020 (top). Spring barley (grey dashed) and winter barley (grey) area and combined yield (green) are shown for 1984-2020 (bottom). The 5-year running mean yield (green dash) and quadratic (green) give an indication of long-term UK barley yield trends. Data from FAOSTAT and DEFRA. .......... 123

Figure 4.3: Winter wheat, winter barley and spring barley median annual variety trial yields for 1982-2018 (from 1983 for spring barley) for treated and untreated trials combined. .......... 124

Figure 4.4: National yield anomalies for **a**. wheat and **b.** barley, after removing the long-term trends for 1982-2020 and 1984-2020, respectively. Years in red represent anomalies in the top and bottom 10%. Data from DEFRA. .......... 125

Figure 4.5: Regional wheat yield anomalies for 1999-2019, after removing the regional trend from 1999-2019. Years in red represent anomalies in the top and bottom 10%. Data from DEFRA. .......... 126

Figure 4.6: Variety trial yield anomalies for 1982-2018 for winter wheat and winter barley and 1983-2018 for spring barley. Fungicide treated and untreated variety trials were combined. Harvest years in red represent anomalies in the top and bottom 10%. Data provided by AHDB. .......... 127

Figure 4.7: Climate anomalies in years of high yields (top 10%). All anomalies are compared to the reference period 1991-2020. Temperature and precipitation data from HadUK has a 1km x

1km resolution (Hollis et al., 2019). SIS = surface incoming solar radiation from CMSAF has a resolution of 0.05° x 0.05° (Pfeifroth, Trentmann, et al., 2018). ............................................... 132

Figure 4.8: Climate anomalies in years of low yields (bottom 10%). All anomalies are compared to the reference period 1991-2020. Temperature and precipitation data from HadUK has a 1km x 1km resolution (Hollis et al., 2019). SIS = surface incoming solar radiation from CMSAF has a resolution of 0.05° x 0.05° (Pfeifroth, Trentmann, et al., 2018). ............................................... 133

Figure 4.9: Monthly temperature, sunshine duration and rainfall anomalies for anomalous yield growing seasons relative to the 1991-2020 averages for the growing season for the UK (red), England (blue), Wales (green), Scotland (purple) and Northern Ireland (NI) (orange). 'Good' harvest years are displayed first, followed by 'bad' harvest years. Data from Met Office year ordered time-series............................................................................................................ 137

Figure 4.10: Cumulative **a**. drilling day of year and **b**. harvest day of year for winter wheat variety trials in Wales, England, Scotland and Northern Ireland for 1988-2018. The median drilling day of year is 283 (10th October) and harvest day of year is 233 (21st August)................................ 138

Figure 4.11: The annual ranges in a. drilling day of year and b. harvest day of year for winter wheat variety trials, 1988-2018. The linear trend in drilling day of year is statistically significant (p<0.001). ......................................................................................................................... 139

Figure 4.12: Start of the growing season (SOGS) day of year averages for 1981-1990, 1991-2000, 2001-2010 and 2011-2020, calculated from 1st January each year. Created using HadUK 1km x 1km gridded temperature data (Hollis et al., 2019)..................................................................... 140

Figure 4.13: Start of Growing Season against the harvest day of year for winter wheat trial sites across the UK for 1988-2018. Calculated using the method described in Table 2.14. ............... 141

Figure 4.14: Mean Growing Degree Days (GDD) (°C days) from September to August for 1982-1990, 1991-2000, 2001-2010 and 2011-2020. Created using 1km x 1km gridded temperature data from HadUK (Hollis et al., 2019)............................................................................................ 142

Figure 4.15: Median Growing Degree Days (GDD) (°C days) at winter wheat trial sites from drilling date to harvest date for 1988-2018. The increase in growing degree days is significant (p<0.05) (black). ............................................................................................................................ 143

Figure 4.16: Vernalisation Degree Days (VDD) (°C days) from September to April for 1982-1990, 1991-2000, 2001-2010 and 2011-2020. Created using equations [2.5] and [2.6] on the 1km x 1km gridded temperature data from HadUK (Hollis et al., 2019). ............................................. 144

Figure 4.17: Estimated anthesis day of year calculated using a thermal time of 2100 $^{o}$C days after drilling date for winter wheat trial sites, calculated using 1 km x 1km gridded HadUK temperature data (Hollis et al., 2019), extracted for each trial site. .......... 145

Figure 4.18: The mean and rate of change in number of air frost days in the UK for 1981-2020 for March, April and May. Percentage change is calculated using the 1981-1985 and 2016-2020 periods. Created using 1km x 1km gridded HadUK temperature data (Hollis et al., 2019). ....... 147

Figure 4.19: Interannual variability in the 30-day accumulated national water balance (mm) for 1981-2017. The running annual mean is also shown (red). Created using CHESS-met and CHESS-PE (Robinson et al., 2020a, 2020b). .......... 148

Figure 4.20: The percentage change in the monthly number of heavy rain (>10mm) days in the 2011-2020 decade relative to the 1981-1990. An increase in heavy rain days is indicated in blue, a decrease in brown. Data from HadUK 1km x 1km gridded precipitation dataset (Hollis et al., 2019). .......... 149

Figure 4.21: The percentage change in the monthly number of heavy rain (>20mm) days in the 2011-2020 decade relative to the 1981-1990. An increase in heavy rain days is indicated in blue, a decrease in brown. Data from HadUK 1km x 1km gridded precipitation dataset (Hollis et al., 2019). .......... 150

Figure 4.22: The number of mild heat stress ($T_{max}$ > 31$^{o}$C) days during grain fill for 1981-2020 at (52.22,0.1) near a trial site in Cambridge. Calculated using gridded 1 km x 1km temperature data from HadUK (Hollis et al., 2019). .......... 152

Figure 4.23: Mean total grain fill surface incoming solar (SIS) radiation (MJ/m$^2$) (left) and percentage change in grain fill SIS from 1987-1991 to 2016-2020 (right). Grain fill incorporates the period 16$^{th}$ June-31$^{st}$ July. Created using gridded 0.05$^{o}$ x 0.05$^{o}$ degrees CMSAF-SIS data (Pfeifroth, Trentmann, et al., 2018). .......... 153

Figure 4.24: Regional and national trends and variability in total surface incoming solar radiation (SIS) (MJ/m2) for the grain fill period 16$^{th}$ June-31$^{st}$ July. Calculated using gridded 0.05$^{o}$ x 0.05$^{o}$ degrees CMSAF-SIS data (Pfeifroth, Trentmann, et al., 2018). .......... 155

Figure 4.25: Total grain fill surface incoming solar radiation (SIS) (MJ/m$^2$) and winter wheat variety trial yield (t/ha) averaged across each trial site each year for each region of the UK. The grain fill period encompasses 16$^{th}$ June-31$^{st}$ July. Total SIS was calculated using gridded 0.05$^{o}$ x 0.05$^{o}$ degrees CMSAF-SIS data (Pfeifroth, Trentmann, et al., 2018). .......... 156

Figure 4.26: National Flag (top) and second leaf (bottom) prevalence (area of leaf affected, %) of foliar diseases of winter wheat on-farm for 1976-2019. No disease data was recorded for harvest years 1983 and 1984, or for tan spot pre-1999. Data from Hardwick et al. (2001), Polley & Thomas (1991) and Turner et al. (2021). ........ 158

Figure 5.1: Median winter wheat, winter barley and spring barley for fungicide treated (o) and fungicide untreated (x) trial yields for 1982-2018 harvest years. The linear increase in median yield is significant ($p<0.05$) for all 3 crops. ........ 166

Figure 5.2: Trends in variety $v_i$ and year $r_j$ effect for fungicide treated (i) and untreated (ii) winter wheat (a.), winter barley (b.) and spring barley (c.) trial yields from 1982-2018 (1983-2018 for spring barley), modelled using equation [2.7]. Variety effects (red squares) were plotted against the first year they entered the trials. Year effects (line) are plotted against calendar years. ........ 168

Figure 5.3: Genetic gain estimates and standard error for fungicide treated variety trials for four periods within 1982-2018: 1982-1991, 1991-2000, 2000-2009 and 2009-2018. The 1982-2018 genetic gain and standard error are shown by the horizontal line in black, and grey, respectively. ........ 171

Figure 5.4: Relative median yield difference (%) between fungicide treated and untreated trial yields for winter wheat (WW, blue), winter barley (WB, orange) and spring barley (SB, green). ........ 172

Figure 5.5: Variety age against the treated-untreated variety trials yield difference for a. winter wheat, b. winter barley and c. spring barley. The effect of variety age on the yield difference between fungicide treated and untreated varieties was modelled using equation [5.8]. Variety age indicates the number of years since the variety entered the trials system. The red line shows the linear relationship between the two variables. ........ 173

Figure 5.6: Variety age against treated and untreated trial yields for winter wheat, winter barley and spring barley. Here the effect of variety age on varieties under full fungicide treatment and no fungicide treatment was modelled using equation [2.8] separately. Variety age indicates the number of years since the variety entered the trials system. The red line shows the linear relationship between the two variables. ........ 174

Figure 5.7: The effect of variety age on yield difference between fungicide treated and untreated variety trials, on treated variety trial yields and on untreated variety trial yields for winter wheat, winter barley and spring barley. The effect of variety age on the yield difference,

treated and untreated variety yields was modelled using equation [2.8] and restricted to varieties aged up to 10 years. Variety age indicates the number of years since the variety entered the trials system. The red line shows the linear relationship between the two variables. ........................................................................................................................................ 175

Figure 5.8: Winter wheat genetic gain estimates for six case study periods and varying numbers of checks extracted from the 1982-2018 NL/RL dataset. Checks refer to varieties present in the trials system for a minimum of 10 consecutive years. Varieties with more than three years of data were restricted to their first three years in trial. ................................................................ 177

Figure 5.9: Estimated contribution of variety improvements to national winter wheat yield increases for 1983-2018. DEFRA national yield data (black dots) are fitted with a spline (black line), as used in (Mackay et al., 2011). The contribution of variety effects $\mathbf{vi}$ to national yield is shown (red line) and was calculated using equation [2.11] on the AHDB variety trials data and DEFRA national yield data. ................................................................................................. 178

Figure 5.10: Model demonstrating the effect of loss of disease resistance on estimation of variety and year effects. Here variety effects increase by 0.1 t/ha per year and then decline linearly due to breakdown in disease resistance. The difference in variety performance for those present in successive years is -0.1 t/ha, which gives an estimated year effect of -0.1 t/ha, as opposed to an actual year effect of 0. Differences in variety performance are estimated by calculating the differences between a variety and its predecessor(s) within years, +0.2 t/ha. This results in variety effects biased upwards and year effects biased downwards. Figure taken from Mackay et al. (2011). ......................................................................................................... 180

Figure 5.11: Model to demonstrate the effect of using a low yielding check (c1) vs. a high yielding check (c2) on the estimate of genetic gain. Genetic gain is calculated by regressing the adjusted variety mean yields on year of entry across all varieties (diamonds). If the check variety has a low adjusted mean yield, then the estimated genetic gain will be higher (G1 line) than the genetic gain estimate (G2 line) of a dataset including a high yielding check (c2) instead. ......... 183

Figure 6.1: Adjusted means and standard error for each soil type calculated from the localised seasonal climate model using equation [2.13]. H = heavy soils i.e. clay or deep clay, M = medium e.g. clay loam and sandy clay, O = organic, P = peat, S = sandy and light. .................................. 190

Figure 6.2: Yield responses (% of average) of each variety to low (white), medium (light blue) and high (dark blue) (as defined in Table 6.9) Growing Degree Days (GDD), growing season 10mm+ rainfall days (rain10), surface incoming solar radiation during grain fill (grainfillSIS) and

Vernalisation Degree Days from November to February (vdd_novfeb). Varieties included here are those present at least 10 years in the variety trials dataset from 1988-2017. The statistical significance of the interaction between variety and the agroclimate metric is given, except for grainfillSIS due to the interaction term not being included in the final model. .......................... 197

Figure 6.3: Yield responses (% of average) of each variety to zero (white) or at least one (dark blue) occurrences of mild heat stress during grain fill (grainfill31) and air frost in May (frost05). Grain fill corresponds to the period 16th June-31st July. ............................................................. 198

Figure A1: Correlation matrix for Irish spring barley growing season (March-August) monthly mean maximum and mean minimum temperature and total monthly rainfall. Only significant correlations ($p < 0.05$) are shown. The larger the square the stronger the correlation. Dark red corresponds to strong negative correlations, dark blue corresponds to strong positive correlations.................................................................................................................. 216

Figure A2: Correlation matrix for Irish spring barley growing season (March-August) monthly maximum and minimum temperature and total monthly rainfall. Only significant correlations ($p < 0.05$) are shown. The larger the square the stronger the correlation. Dark red corresponds to strong negative correlations, dark blue corresponds to strong positive correlations. ............... 217

Figure A3: Observed versus fitted values from using lmerTest on the extreme maximum and minimum monthly temperature, and total monthly rainfall data with the Irish spring barley trials data from 1901-1906. All the linear mixed model plots of observed vs. fitted values in Chapter 3 showed very similar graphs, so were excluded to avoid repetition. ........................................... 218

Figure A4: Growing season rainfall anomaly (% 1991-2020 average) plots for the UK (red), England (blue), Wales (green), Scotland (purple) and Northern Ireland (NI) (orange). Data from Met Office year ordered time-series............................................................................................ 222

Figure A5: Growing season sunshine duration anomaly (% 1991-2020 average) plots for the UK (red), England (blue), Wales (green), Scotland (purple) and Northern Ireland (NI) (orange). Data from Met Office year ordered time-series. ................................................................................... 226

Figure A6: Growing season mean temperature anomaly (°C difference to 1991-2020 average) plots for the UK (red), England (blue), Wales (green), Scotland (purple) and Northern Ireland (NI) (orange). Data from Met Office year ordered time-series (Met Office, 2022c). ......................... 230

Figure A7: Change in the percentage of winter wheat trials following oilseed rape (OSR) (red) and the average winter wheat trials yield. Calculated using national yield data from DEFRA. .. 231

Figure A8: Vernalisation Degree Days (VDD) (°C days) from 1st November to 28th February for 1982-1990, 1991-2000, 2001-2010 and 2011-2020. Calculated using equations [2.5] and [2.6] using 1 km x 1km gridded HadUK temperature data (Hollis et al., 2019). .................................. 232

Figure A9: Annual surface incoming solar radiation (SIS) anomalies (%) relative to 1991-2020 average. Calculated using gridded 0.05° x 0.05° degrees CMSAF SIS data (Pfeifroth, Trentmann, et al., 2018). .......................................................................................................................... 234

Figure A10: National Septoria blotch flag leave prevalence (area affected %) (Polley and Thomas, 1991; Hardwick et al., 2001; Turner et al., 2021) and regional wheat yields (t/ha) (DEFRA, 2021a) for 1999-2019. ......................................................................................................................... 235

Figure A11: Observed versus fitted winter wheat yield values (t/ha). Fitted values are from the seasonal climate model on winter wheat (sum of squares shown in Table 6.1), which used [2.12] on the UK National List/Recommended List treated variety trials data for 1988-2018 and regional climate data from the Met Office (Met Office, 2022c). The subsequent models in Chapter 6 had very similar distributions of observed vs. fitted values, hence these plots haven't been included. ........................................................................................................................ 236

Figure A12: Pearson's correlation coefficient between each pair of climate variables. Only significant correlations (p<0.05) are shown. The stronger and darker the correlation the larger the square, with positive correlations shown in blue and negative correlations shown in red. 237

Figure A13: Yield responses (% of average) of each variety to low (white), medium (light blue) and high (dark blue) (as defined in Table 6.9)numbers of frost days in April (frost04) and March (frost03), 20mm+ rainfall days (rain20), precipitation-evapotranspiration balance (pe_balance) and Vernalisation Degree Days from planting to anthesis (vdd_p2a). ........................................ 238

# Tables

Table 2.1: Number of Irish spring barley trial sites per year. Data from Student (1923). ............ 52

Table 2.2: Temperature datasets used for each Irish weather station. ........................................ 54

Table 2.3: Rainfall datasets used for each Irish weather station. ................................................. 55

Table 2.4: The closest weather station to the early 20th century Irish barley trials sites with daily data for 1901-1906. ........................................................................................................ 56

Table 2.5: Variables of interest and their range of values for Irish barley yield modelling. .......... 61

Table 2.6: Methods of variable selection used to select climate covariates. ................................ 64

Table 2.7: Gridded weather datasets used to identify the most suitable dataset for use in the agroclimate analysis ........................................................................................................ 68

Table 2.8: Synoptic weather stations used for validation of the gridded weather datasets HadUK, ERA5-Land and MÉRA. ........................................................................................................ 68

Table 2.9: Daily maximum and minimum temperature comparison statistics for three gridded temperature datasets. ........................................................................................................ 70

Table 2.10: Daily rainfall comparison statistics for three gridded rainfall datasets. ..................... 70

Table 2.11: Common site information issues in the variety trials dataset. GR = grid reference. *The AHDB Harvest Results archive (https://ahdb.org.uk/knowledge-library/harvest-results-archive) filled in the gaps for some of the missing grid reference for 2002 onwards. ................. 72

Table 2.12: Structure of winter wheat (WW), winter barley (WB) and spring barley (SB) trials data after quality control and restricting for varieties present for minimum 3 years. ................. 74

Table 2.13: Yearly structure of winter wheat trials data after quality control and restricting for varieties present for minimum 3 years showing large variations in the number of varieties, trial sites and consequently total observations. No. T and No. U show fluctuations in the number of fungicide treated and untreated trial observations, respectively. ................................................ 77

Table 2.14: Summary of selected agroclimate metrics used in Chapter 4's analysis of changes in the UK's agroclimate. ........................................................................................................ 81

Table 2.15: Genetic gain winter wheat case studies for various periods and check varieties. Check varieties are those present for at least 10 consecutive years in the NL/RL trials dataset.

Only the first three years of trials for varieties present for less than 10 years are included in each case study dataset. 'Observations' indicates the number of data points within the period for the 0-check model. Adding in checks increases the data points. ........................................................ 86

Table 2.16: Soil textures and their classification. *Soil textures not in AHDB (2019) and Landis Soil Portal and instead suggested by Haidee Philpott (2021, pers. comm.). ................................ 88

Table 2.17: Geographic location and climatology (1982-2018) of the 10 main regions used by the Met Office. Calculated using Met Office regional series annual average data. Here $T_{max}$ and $T_{min}$ correspond to the average annual maximum and minimum temperature ($^oC$) for each region over the period 1982-2018. Precipitation values represent the average annual total precipitation over the same period.................................................................................................. 89

Table 2.18: Average yields for varieties present at least 10 years in the UK National List/Recommended List winter wheat variety trials dataset. The first and last year each variety was present in the trials dataset is given. ......................................................................... 93

Table 3.1: Analysis of variance (ANOVA) of Irish barley model [2.2] including year as a factor and a variable, farm and variety and their interactions. Degrees of freedom (Df) and Sum of Squares (SS) are shown for each model term. The p-values and significance are also shown for the model with year as a variable but not for year as a factor as these cannot be calculated when the model has a perfect fit. *significant at the 95% confidence level. ............................................ 107

Table 3.2: ANOVA results for lm model including year (as a factor), farm and variety, and their two-way interactions. Degrees of freedom (Df), Sum of Squares (SS), F-value and p-value are shown for each model term. *significant at the 95% confidence level. ..................................... 107

Table 3.3: Estimated coefficients and their p-values and significance for 7 different frequentist variable selection methods (Table 2.6), as well as the full climate model (c.m.). cv.b = best subset selection with cross-validation (Method 1), s.bo = cross stepwise selection in both backwards and forwards directions (Method 2a), s.ba = 10-fold cross-validation backwards stepwise selection (Method 2b), s.f = 10-fold cross-validation forwards stepwise selection (Method 2c), l.m = cross-validation lasso using lambda that minimises the prediction error (Method 3a), l.1 = cross-validation lasso model using lambda for smallest model and within 1 standard error (Method 3b), e.n = cross-validation elastic net using lambda and alpha that minimises the prediction error (Method 4). Of the 7 models shown, only 4 give p-values and significance of each variable............................................................................................ 110

Table 3.4: Adjusted $R^2$ and RMSE for the 7 different variable selection methods (Table 2.6), as well as the full climate model (c.m.). cv.b = best subset selection with cross-validation (Method 1), s.bo = cross stepwise selection in both backwards and forwards directions (Method 2a), s.ba = 10-fold cross-validation backwards stepwise selection (Method 2b), s.f = 10-fold cross-validation forwards stepwise selection (Method 2c), l.m = cross-validation lasso using lambda that minimises the prediction error (Method 3a), l.1 = cross-validation lasso model using lambda for smallest model and within 1 standard error (Method 3b), e.n. = cross-validation elastic net using lambda and alpha that minimises the prediction error (Method 4). ................................ 111

Table 3.5: Estimated model parameters from the stan_glm model. Median is the median value computed from the model simulations. CI represents the Credible Interval, which quantifies the uncertainty about the regression coefficients. The 89% CI computes the Credible Interval with 89% probability that a coefficient lies above the lower value of the two values and below the higher value. If the CI doesn't straddle 0 then the coefficient is significant and is in bold. pd is the Probability of Direction, which is the probability the effect goes to the positive or negative direction (Bloggers, 2020). ................................ 111

Table 3.6: The 6 main principal components (PCs) contributing to variance in spring barley yields. The proportion of variance of each PC is shown, along with the degree of correlation of each mean climate variable with each PC. Variables correlating by more than |0.3| are in bold. ................................ 113

Table 3.7: Statistical significance and corresponding coefficient of each climate variable in the mixed model with year, variety and year:farm. Significant variables are shown in **bold**. The AIC of the overall model is given, with lower values corresponding to a better model fit. ................................ 115

Table 4.1: Pearson correlation coefficient between grain fill (16th June-31st July) surface incoming solar radiation (SIS) and trial yield within each region is shown below, with significant correlations ($p<0.05$) in **bold.** ................................ 157

Table 5.1: Linear trend over varieties $v_i$ and with the standard error (SE), and linear trend over years $r_j$, for winter wheat (WW), winter barley (WB) and spring barley (SB) for a range of periods within the trials datasets. The trend was calculated using equation [2.7]. T corresponds to fungicide treated trials; U corresponds to untreated trials. All trends are significant ($p<0.05$) unless denoted with (ns). ................................ 167

Table 5.2**:** Winter wheat disease incidence severity (average percentage area of leaf) of foliar diseases and fusarium ear blight combined with the difference between the detrended year

effects for untreated and treated variety trials with year effects <-0.5 shown in bold. Disease data from Hardwick et al. (2001), Polley & Thomas (1991) and Turner et al. (2021). No disease data was available for 1983 and 1984. ........ 170

Table 5.3: Adjusted variety means calculated using four linear models on a model yield dataset (Table A3). Model 1 includes all data and accounts for year and variety, model 2 doesn't include the check but accounts for year and variety, model 3 includes all data and accounts for year, variety and whether it's in the NL or RL year and model 4 doesn't include the check and accounts for year, variety and NL/RL year. ........ 184

Table 6.1: Sum of squares for fixed effects in the seasonal climate model of winter wheat, using [2.12] on the UK National List/Recommended List treated variety trials data for 1988-2018 and regional climate data from the Met Office (Met Office, 2022c). Fitted vs. observed values for this model are shown in Figure A11. *significant at the 95% confidence level. son = autumn, djf = winter, mam = spring and jja = summer. Coefficient estimates and standard error (SE) are given for the climate variables. ........ 188

Table 6.2: Sum of squares for fixed effects in the seasonal climate model of winter wheat, using [2.12] on the UK National List/Recommended List treated variety trials data for 1988-2018 and site-specific climate data, extracted from HadUK (Hollis et al., 2019). *significant at the 95% confidence level. son = autumn, djf = winter, mam = spring and jja = summer. Coefficient estimates and standard error (SE) are given for the climate variables. ........ 189

Table 6.3: Sum of squares for fixed effects in the seasonal climate model of winter wheat including soil texture, using [2.13] on the UK National List/Recommended List treated variety trials data for 1988-2018 and site-specific climate data, extracted from HadUK (Hollis et al., 2019). *significant at the 95% confidence level. son = autumn, djf = winter, mam = spring and jja = summer. Coefficient estimates and standard error (SE) are given for the climate variables. ........ 190

Table 6.4: Sum of squares for fixed effects in the monthly climate model of winter wheat, using [2.12] on the UK National List/Recommended List treated variety trials data for 1988-2018 and site-specific climate data, extracted from HadUK (Hollis et al., 2019). *significant at the 95% confidence level. Coefficient estimates and standard error (SE) are given for the climate variables. ........ 192

Table 6.5: Univariate climate model sum of squares (SS), p-value and significance for each climate variable and the respective climate x variety interaction term, calculated using [2.12] on

each single climate variable $C_{jk}$ paired with the treated variety trials data for 1988-2018. *significant at the 95% level. ..... 192

Table 6.6: Sum of squares for fixed effects in the multivariate agroclimate model of winter wheat, using [2.12] on the UK National List/Recommended List treated variety trials data for 1988-2018. *significant at the 95% confidence level. ..... 193

Table 6.7: Variable description of variables in the final agroclimate model [6.1] for winter wheat. Each variable is fitted as a fixed effect (F) or a random effect (R). ..... 194

Table 6.8: Sum of squares for fixed effects in the multivariate agroclimate model of winter wheat, using the optimised model [6.1]. *significant at the 95% confidence level. Coefficient estimates and standard error (SE) are given for the climate variables. ..... 195

Table 6.9: The low/medium/high categories for each agroclimate metric included in the final model [6.1]. ..... 195

Table 6.10: Variety x climate variable interaction coefficient for each variety with at least 10 years of data, giving the offset from the mean response of all varieties. Interactions with a false discovery rate (FDR) < 0.5 are highlighted in grey. The total number of interactions detected with an FDR < 0.5 for each climate variable and variety are shown at the end of each column and row, respectively. Climate variables included here were significant in the univariate climate models (Table 6.5). The variety 'Alchemy' is not included here as it's not possible to extract the interaction coefficient from the model's intercept term. ..... 199

Table A1: Estimated coefficients and their p-values and significance for 4 different variable selection methods, as well as the full climate model (c.m.) for the extreme monthly Irish climate dataset. cv.b = best subset selection with cross-validation (Method 1), s.bo = cross stepwise selection in both backwards and forwards directions (Method 2a), s.ba = 10-fold cross-validation backwards stepwise selection (Method 2b), s.f = 10-fold cross-validation forwards stepwise selection (Method 2c) model using lambda for smallest model and within 1 standard error. p-values and significance of each variable where available. ..... 240

Table A2: Coefficient estimates, p-values and their significance for three linear mixed model methods using the Irish monthly extremes climate data. lmer post selection uses the results from variable selection of climate variable, backwards elimination is the result of running step from lmerTest on the full linear mixed model, and Bayesian post selection represents the Bayesian mixed model run after climate covariate selection using lm. [1]Signif. codes: 0 *** 0.001 ** 0.01 * 0.05 . 0.1. ..... 241

Table A3: Model dataset for estimating genetic gain in the UK NL/RL trials. Seed means seed source, so whether its NL = national list or RL = recommended list. Here g is the check (control) variety. ........ 241

Table A4: The variance components of the random effects in the UK regional, seasonal climate model (Section 6.1). ........ 242

Table A5: The variance components of the random effects in the UK localised seasonal climate model (Section 6.2). ........ 242

Table A6: The variance components of the random effects in the UK localised, seasonal climate and soil model (Section 6.2). ........ 242

Table A7: The variance components of the random effects in the UK localised monthly climate model (Section 6.3). ........ 242

Table A8: The variance components of the random effects in the UK agroclimate model (Section 6.4). ........ 242

Table A9: Final UK multivariate agroclimate model fixed effects coefficient estimates and associated standard error, degrees of freedom (Df), t statistics and p value significance *significant at the 95% confidence level. ........ 245

# Acknowledgements

I am grateful to all those who have contributed to this thesis both intellectually and in a supporting role. First and foremost, I thank my supervisory team, who have provided knowledge, support and guidance throughout the PhD. Thank you to Professor Steve Dorling for giving me the opportunity to undertake this journey and encouraging me along the way. Thank you for the hours spent reading my work and for your excellent scientific advice on all things meteorology.

I am particularly grateful to Professor Ian Mackay, who went above and beyond his role on the supervisory team and provided much-needed support, guidance and advice in the depths of the pandemic when times were tough. I wouldn't have been able to finish this thesis without your support in the last two years. Your passion for science is inspirational and I hope that one day I can be half the scientist you are.

Thank you to Professor Steve Penfield, Professor Andrew Lovett and Haidee Philpott for imparting your vast array of knowledge on crops, agriculture, statistics and variety trials. Your input at meetings and comments on my writing have been incredibly valuable. Thanks to Ellie Marshall at AHDB for access to the variety trials data and to Sam Brooke at BSPB for the support and advice.

To the team at Weatherquest, thank you for allowing me into the family and for the many games of table tennis. Particular thanks go to Ian Gooch, for your help with data access and processing. Your friendship is one of the good things to come out of the pandemic.

I would like to thank all my friends, particularly the triathlon gang, who helped me get through the bumpy rollercoaster that is a PhD. I look forward to many more sporting adventures with you. A special mention to my trusty bikes: you allowed me to see the beauty of Norfolk and beyond, and the headspace you provided kept me (almost) sane.

To my wonderful family. I cannot put into words how grateful I am for your unwavering love and support, without which I wouldn't be where I am today. You never stopped believing in me.

And finally, to my partner Conor. I am so lucky you have been with me on this journey. Thank you for all you have done for me, particularly these last few months. I cannot wait for the next chapter of our lives.

This work would not have been possible without the financial support Natural Environment Research Council and the EnvEast Doctoral Training Partnership [grant number NE/L002582/1]', and the British Society of Plant Breeders (BSPB).

# 1 The role of breeding and climate in UK cereal production, and the need to combine the two in analysis

Challenges faced by agriculture in the 21st century are dominated by climate change and increasingly extreme weather events (Parolini, 2022). This has contributed to changes in the distributions of viable cropping areas and ranges of pests and diseases. Combined with an increasing global population, the goal of eliminating world hunger seems more and more unobtainable (FAO *et al.*, 2021). As one of the main contributors to climate change, the agricultural industry needs to decrease resource usage, greenhouse gas emissions and biodiversity destruction whilst increasing productivity. Despite the continued contribution of plant breeding to yield increases (Peltonen-Sainio, Jauhiainen and Laurila, 2009; Mackay *et al.*, 2011; Noleppa and Cartsburg, 2021), in recent decades agricultural productivity has been limited by stagnating yields in vital crops.

Crop yield is intrinsically linked with climate variability (Ray *et al.*, 2015), along with management decisions and agronomic changes. Recent global events have demonstrated how vulnerable crop production is to external shocks. The COVID pandemic disrupted vital supply chains, such as through labour shortages (Hobbs, 2020), and the Russian invasion of Ukraine has highlighted the risks associated with high dependency on one or two countries for specific foods, in this case wheat (*Triticum aestivum* L.). There is a need to strengthen the resilience of domestic supply chains and food production as well as diversify international food supply to enhance the UK's food security and overall national resilience (Berry and Brown, 2021; DEFRA, 2022).

In recent years, the UK has seen significant yield fluctuations in staple food crops, including cereals (DEFRA, 2021a). This creates high instability in income for farmers and in prices for consumers. Identifying climatic causes of production variability is important for future-proofing UK agriculture as the changes in climate affect the likelihood of challenging weather events (Arnell and Freeman, 2021). Several studies have identified the risks associated with future climate change for UK crop production (Semenov, 2009; Cho *et al.*, 2012; Arnell and Freeman, 2021) but very few have looked at how observed climate has affected historical yields, and the focus of these has largely been on one region (Addy *et al.*, 2020, 2021a) or using national yield and climate data which can mask significant local weather and climate variability (Knight *et al.*, 2012). Given the array of weather and climate risks to crops yields, there is an urgent need to utilise the detailed climate information now available to identify the best suited existing crops and varieties to grow locally, as well as to help develop climate resilient varieties.

This thesis investigates the genetic and environmental causes of recent variability and trends in UK cereal yields. Here, yield is defined as grain yield at 15% moisture content and measured in tonnes per hectare. The thesis explores how agroclimate information can be better incorporated into breeding programmes and made more accessible to growers. This analysis demonstrates the drivers of recent barley (*Hordeum vulgare* L.) and wheat yield variability and identifies wheat varieties with greatest sensitivity to the main climate drivers of yield. Whilst it's not possible to totally "weatherproof" agriculture as the title suggests, this thesis seeks to provide methods for improving the climate resilience of UK crops in the face of climate change.

This chapter commences with a discussion on the challenges faced in achieving global food security and addresses the genetic, agronomic, and climatic causes of observed yield plateaus. Section 1.2 explores the importance of agriculture and crop breeding in the UK, how the UK's growing climate is changing and the current literature on crop-climate relationships. Section 1.3 identifies the opportunities to use climate data to support crop breeding and agriculture in the changing climate.

## 1.1 Threats and opportunities for agriculture in changing climates

### 1.1.1 The challenges of achieving global food security

The term "food security" first appeared in the mid-1970s at the 1974 World Food Conference (FAO, 2006), and has evolved from focussing on food availability and price stability to incorporate four key dimensions: availability, access, utilization and stability. The challenge to achieve global food security and meet the Sustainable Development Goal of Zero Hunger by 2030 is multi-faceted and has become increasingly difficult in recent years. The decline in world hunger seen since 2005 came to an end in 2014 (FAO *et al.*, 2021). This is due, in part, to population growth, constraints on land availability for agriculture, an observed increase in weather variability and extreme events, and greater climatic uncertainty due to climate change (Cassman *et al.*, 2011; Mandryk *et al.*, 2015; Pörtner *et al.*, 2022), with the COVID-19 pandemic contributing to recently rising levels of malnutrition (FAO *et al.*, 2021). This has been compounded by the 2022 Russian invasion of Ukraine; together these two countries produce nearly 30% of the world's traded wheat and so the military conflict has resulted in great uncertainty surrounding the export of this grain (Behnassi and El Haiba, 2022). This has driven up food prices and has highlighted the fragility of globalised food systems and risks associated with high dependency on one or two countries for specific foods.

Observed impacts of climate change on agricultural productivity, documented in over 150 articles since the last Intergovernmental Panel on Climate Change (IPCC) Assessment Report in 2014, indicates that climate change has had an overall negative impact of productivity thus far (IPCC, 2022). Under current pledges for climate change mitigation, an anticipated 3°C of warming would result in more than 50% of agricultural area in China, Brazil, Egypt, Ethiopia, Ghana and India projected to be exposed to severe droughts of more than one year within a 30-year period (Price *et al.*, 2022). GDP and welfare in all of these countries, except for China, are projected to be negatively impacted by declining crop yields due to this level of warming (Wang *et al.*, 2021). Feeding the projected population of 9.7 billion in 2050 (Roser, 2013) with nutritious food and reduced environmental impact is an incredibly daunting task. Modelling studies have shown it is in theory possible (KC *et al.*, 2018; Springmann *et al.*, 2018), but would require drastic dietary changes (Vermeulen *et al.*, 2020), which can take significant time. To prevent widespread starvation and loss of livelihoods, there is a need to understand how agriculture on a global scale can adapt to changing climates as quickly and efficiently as possible.

### 1.1.2 Understanding recent yield trends

A further challenge for increasing food production to meet global demand has been the slow yield increase observed in several staple crops in many parts of the world. The intensification and mechanisation of agriculture, along with genetic improvements, saw major crop yield increases in the 1960s, 1970s and 1980s. The green revolution resulted in a unique period in human history when food supply consistently outstripped demand (Cassman *et al.*, 2011). While world food production has continued to increase, and yield gains have continued for some crops and countries, a decline in the rate of yield increase, often referred to as a ‘yield plateau’, arose at the turn of the new millennium in several crops and regions (Hafner, 2003; Cassman *et al.*, 2011; Grassini *et al.*, 2013). These include wheat in northwest Europe, Australia and India, maize (*Zea mays* L.) in China and rice (*Oryza sativa* L.) in China, Indonesia and India (Brisson *et al.*, 2010; Cassman *et al.*, 2011; Hochman *et al.*, 2017; Espe *et al.*, 2018). Globally, during 1990-2010 there was strong evidence for widespread deceleration in the rate of increase in average yields for 31% of total global rice, wheat and maize production (Grassini *et al.*, 2013). There is limited scope to increase crop-growing areas, however closing the yield gap between the potential yield of a crop variety at a specific location, and the average actual yield achieved by farmers, could be a way of increasing crop production (Senapati and Semenov, 2019). Possible causes of yield plateaus and yield gaps can be loosely classified as genetic, agronomic or climatic (Brisson *et al.*, 2010). Whilst each crop and location are unique, common causes for a yield plateau may apply.

*Genetic factors*

A handful of studies have indicated that there may be a limit to genetic improvement which breeders have begun to reach, and as a consequence genetic yield improvement has not increased (Calderini and Slafer, 1998; Espe *et al.*, 2018). However, genetic improvement through plant breeding has been widely shown to have contributed to continued increases in yield potential (Peltonen-Sainio, Jauhiainen and Hakala, 2009; Brisson *et al.*, 2010; Mackay *et al.*, 2011). For example, Brisson *et al.* (2010) showed that despite a yield plateau in wheat in France, yield potential, as defined as the improvement in variety trial yields, was still increasing in the order of 0.1 t/ha/yr, indicating genetics has not contributed to observed yield stagnations. Furthermore, the emerging field of phenomics, in which whole-plant phenotypes (traits) are broken down into separate ones that are controlled by a smaller number of genes, is enabling targeted breeding focused on enhancing fundamental plant processes such as photosynthesis, which is expected to lead to further yield increases (Flood *et al.*, 2011; van Bezouw *et al.*, 2019; Simpson *et al.*, 2022).

One hypothesis addressing the observed yield plateaus, is that average national yields plateau when they reach 70-80% of the genetic yield potential ceiling (Lobell *et al.*, 2009). The gap between average yields achieved by farmers and yield potential depends on the extent to which crop and soil management practices remove abiotic and biotic stresses, as well as the yielding capability of available crop varieties (Cassman, 1999; Grassini *et al.*, 2013). Yield potential can be seen as a biophysical limit to the attainable yield at a given location (Knight *et al.*, 2012; Grassini *et al.*, 2013). It is neither cost-effective nor physically possible to achieve near perfect management on an industrial scale: as farmers' yields approach the yield potential ceiling, incremental yield increases come from finer tuning of different management techniques, the costs of which can offset any monetary gain from small yield increases. Therefore, it should be expected that average yields stagnate when they approach a high fraction of the yield potential. For example, Ray *et al.* (2012) suggest wheat yields may have stagnated in Bangladesh and parts of India because current cultivars are approaching their yield potentials. The estimated global genetic yield gap, defined as the gap between the genetic yield potential of an optimized local wheat cultivar and the potential yield of the current local cultivar, is 51% (Senapati *et al.*, 2022).

On a local scale, this highlights the importance of regular cultivar replacement with new, better adapted and higher yielding varieties. At a global scale, substantial increases in food grain production will need to come from yield increases in countries in which the yield gap between obtained yield and yield potential is greatest, which will be a challenge given the low development levels of these countries (Grassini *et al.*, 2013).

*Agronomic factors*

There have been significant changes in agronomy since the 1960s, including changes in cropping regimes, fertiliser application and the introduction of stricter environmental policies that have limited the protection of crops from pests and diseases. For example, since 2000, the UK has seen a reduction in inputs of fertilizers (1.5%/year), plant protection product (1.2%/year), labour (0.8%/year) and capital (0.3%/year) (Noleppa and Cartsburg, 2021).

Changes in crop rotations and stricter environmental policies have also impacted cereal yields. To increase food supply, intensive agriculture has used increasing quantities of nitrogen fertilizer (Gu *et al.*, 2023), however more than half of the cropland nitrogen inputs are lost to the air and water. This has led to an array of negative environmental impacts, including air and soil pollution, whilst also contributing to climate change through the potent greenhouse gas nitrous oxide (Erisman *et al.*, 2013; Steffen *et al.*, 2015). As a consequence, a rise in environmental regulations for restricted nitrogen use, such as in Nitrate Vulnerable Zones, has resulted in a plateau in nitrogen application in the UK (Knight *et al.*, 2012). Combined with an observed decrease in wheat crops following nitrogen-fixing legumes, this has contributed to a 24% decrease in soil nitrogen from 2000 to 2019 (DEFRA, 2021a), and a reduction in UK wheat yields (Knight *et al.*, 2012). The influence of more recent policy changes on yields, such as the ban on outdoor use of the slug pesticide metaldehyde in April 2022 (DEFRA *et al.*, 2020), is yet to be quantified, but could well contribute to further stagnation.

*Climate factors*

Climate is widely recognised as being responsible for interannual variability in yield (Brisson *et al.*, 2010; Ray *et al.*, 2015), however the effect on yield stagnation is less well understood, partly because interannual variability can mask the trend. The IPCC Sixth Assessment Report indicates wheat, soybean and maize production have been negatively impacted on by recent climate trends (IPCC, 2022). In Asia and Australia, yields of wheat and rice may have stagnated due to climate-change related heat stress and increased night time temperatures (Ray *et al.*, 2012; Sadok and Jagadish, 2020): from 1979-2003, rice grain yield in the Philippines declined by 10% for each 1$^{o}$C increase in growing season (the period from planting to harvest) minimum temperature (Peng *et al.*, 2004). In Europe, the effects of climate change may have partly counteracted the genetic progress made in wheat. The depressive effect of climate is greatest in areas of intensive cereal growing (Brisson *et al.*, 2010). In Finland, an increase in mean temperatures reduced seed yield of newer rapeseed cultivars (Peltonen-Sainio *et al.*, 2007). There is limited capacity of a single crop genotype to perform well under climatic variability, thus a set of cultivars with diverse responses to weather conditions is required in order to spread risk (Kahiluoto *et al.*, 2019).

*Competing objectives*

Whilst in the 1970s and 1980s, farmers' main objective was often to maximise crop yield, in recent years farmers have had many other competing objectives, which means that high yields has not always been the end goal. Wheat and barley can be grown for different end uses, including bread-making, biscuit-making, malt for brewing and for feed (AHDB, 2023), with each market a trade-off in terms of quantity i.e. yield, and quality. Increased volatility in commodity prices in the world market and the lack of control over market prices for grain (Läänemets *et al.*, 2011) can mean that a grower may seek lower risk, such as maximise gross margins for any one field but investing less in other fields, resulting in lower average yields. Changes in government policy and the introduction of more incentives for environmental protection (Knight *et al.*, 2012) have encouraged a reduction in inputs, such as fertilisers, again having yield trade-offs.

### 1.1.3 Methods of overcoming yield plateaus

Overcoming yield plateaus and increasing production with reduced inputs and environmental impact is required to achieve food security. Production is a function of both yield and area, thus production can be increased by improving yields and/or increasing cultivation area (Bradshaw, 2017). In the UK, 71% of land is dedicated to agriculture, such that there is little suitable additional land to further increase food production in this way (DEFRA, 2018; Downing and Coe, 2018). Rather, the most likely avenue for increasing self-sufficiency and reducing the yield gap is by increasing land productivity (AHDB Cereals & Oilseeds, 2018b), which is achieved by the optimisation of many factors, including using well-adapted cultivars, and good crop and soil management practices.

### 1.1.4 The role of crop breeding in increasing yields

Breeding has made a major contribution to increasing global agricultural productivity. In the UK, at least 88% of cereal and oilseed rape crop yield improvement from 1982-2007 was attributed to genetic improvement, as opposed to agronomic changes (Mackay *et al.*, 2011). However, there is concern that major breeding efforts in the last century have caused a reduction in crop genetic diversity. This narrowing of the crop gene pool leaves crops at greater risk to strains of diseases. For example, there is a risk of the resurgence of stem rust (*Puccinia graminis*) in the UK, with the potential for widespread wheat and barley yield losses due to low genetic diversity and a lack of resistance to new strains (Lewis *et al.*, 2018). There is conflicting evidence for changes in diversity level, even for the same crop and same regions (Roussel *et al.*, 2004; Huang *et al.*, 2007). A recent assessment of the diversity of wheat responses to weather events in nine different European countries shows a decline in response diversity of wheat in farmer's fields (Kahiluoto *et al.*, 2019)

Breeding cultivars better adapted to the changing climatic conditions will be pivotal to overcoming yield plateaus and closing yield gaps. In the most recent Assessment Report of the Intergovernmental Panel on Climate Change, cultivar improvement was identified as one of the effective adaptation options for enhancing food security (Pörtner *et al.*, 2022). Current breeding programmes and cultivar selections don't prepare for climatic uncertainty and variability: for example, there is a lack of positive responses of European wheat to abundant on-farm precipitation, yet in variety trials the negative yield response is less (Bradshaw, 2017). This suggests there is an unexplored potential to draw upon tested cultivars. For UK oilseed rape, better uptake of new varieties helped close the yield gap and overcome the yield plateau of 1994-2004 (Knight *et al.*, 2012). An *in silico* experiment showed that even in high productive countries, designing crop ideotypes, or virtual idealized crops expected to produce greater grain quality and quantity, can close the gap further and increase land productivity by providing key traits for crop improvement (Senapati and Semenov, 2019).

In many developing countries, in particular where extreme weather events such as droughts are common, locally adapted, domesticated varieties, known as landraces, are the backbone of agricultural production. These traditional varieties of plants are well-adapted and can be genetically diverse (Azeez *et al.*, 2018). They demonstrate a range of grain yield responses to these conditions, with all landraces producing some yield, whereas some modern cultivars fail. Research into barley has shown that landraces yield more than modern cultivars in low-input and stress conditions (Ceccarelli *et al.*, 2007), whilst integrating material from landraces into spring wheat has been shown to increase yield under heat stress compared to elite lines, whilst leading to no significant yield penalty under favourable conditions (Molero *et al.*, 2022). Landraces and wild relatives are the best source of resistance to abiotic and biotic stresses. Thus, replacing numerous landraces and traditional cultivars with few modern varieties is a huge threat to global food security (Upadhyaya *et al.*, 2013; Bradshaw, 2017). There is a need to assess the existing collection of landraces for maintaining system resilience, and overcoming current yield stagnation and future climatic challenges (Mäkinen *et al.*, 2015).

Traditional breeding techniques are a form of artificial selection and thus far have been largely limited to naturally occurring varieties. However, in recent decades the direct manipulation of genes through DNA transfer has become possible, allowing new genes from one species to be incorporated into a completely unrelated species through genetic engineering, creating genetically modified (GM) crops (Phillips, 2008). The capacity for genetic modification to combine desirable traits has already been demonstrated for several crops, including the transfer of a sunflower gene that encodes a stress-responsive transcription factor into soya bean, increasing its drought stress-tolerance (Cabello and Chan, 2012; Cabello *et al.*, 2012). In 2015, GM crops were already being grown on over 10% of the world's arable land (The Royal Society, 2016). There is great potential for GM crops to alleviate some future yield losses due to climatic changes, however widespread opposition delayed uptake in Europe and there has been a wide gap between the rapid acceptance of cultivating GM crops by farmers, and the often-limited acceptance by consumers (Falloon *et al.*, 2015; Lucht, 2015). More recently, EU legislation has changed giving individual governments more power to decide whether to grow GM crops. Furthermore, gene editing, in which a small genetic change is induced often mimicking what could be produced through breeding, is becoming more widely accepted and in the UK, the new Genetic Technology (Precision Breeding) Bill is making its way through Parliament and seeks to make provision about the growth and sale of such plants (Vaughan, 2022).

### 1.1.5 The influence of farm management on crop production

To close the yield gap between theoretical yield potential and actual attained yield, optimal crop and soil management is required to alleviate more abiotic and biotic stresses that can limit crop growth and yield (Cassman *et al.*, 2003, 2011). There is considerable opportunity for the yield gap between the current, local climatic yield potential and actual yield to be closed through suitable management practices (Licker *et al.*, 2010; Liu *et al.*, 2022). Chen *et al.* (2014) showed that implementing soil-crop system management practices that account for crop ecophysiology and soil biogeochemistry can substantially increase average yields for rice, wheat and maize without needing increases in nitrogen fertilizer. Furthermore, simulation of the combined effect of optimal farm management, and the breeding and growing of well-adapted varieties, suggests that crop-level management adaptations could increase global yields in a 2$^{o}$C warmer world by an average of 7-15% relative to no-adaptation scenarios (Challinor *et al.*, 2014).

Precision agriculture offers additional methods of achieving more production with more sustainable agronomy, through the automation and precision application enabled through technology (Bhakta *et al.*, 2019). This application of technology, however, may not be feasible for poorer farmers in developing countries.

Ultimately, there is a need for an integrated approach to achieving sustainable food security in a changing climate (Lipper *et al.*, 2014). Practices need to be adapted to fit local contexts and actions both on-farm and beyond the farm. Incorporating weather and climate information into all aspects of farm management is an important step for decreasing the yield gap. In the North China Plain, varietal changes in both wheat and maize have helped stabilise the length of the pre-flowering period and extend the length of the grain-filling period (Liu *et al.*, 2010). Using climate information to support crop breeding and growers' decisions in this way has great potential to increase yields (Falloon *et al.*, 2015).

### 1.1.6 The carbon dioxide fertilisation effect

The current atmospheric carbon dioxide concentrations (417 mm) are projected to double by the end of the century, under the high emissions Representative Concentration Pathway (RCP) 8.5 (Stocker *et al.*, 2013). Elevated $CO_2$ has been shown to increase biomass and yields in $C_3$ plants, such as wheat, (Jablonski *et al.*, 2002; Högy *et al.*, 2009; O'Leary *et al.*, 2015; Fitzgerald *et al.*, 2016), but decrease protein content and overall nutritional value (Taub *et al.*, 2008; Myers *et al.*, 2014; Blandino *et al.*, 2020; Carreras Navarro *et al.*, 2020). Crop modelling studies using future climate projections, have found that the $CO_2$ fertilisation effect often outweighs the negative effects associate with climate change, such as increase in drought and heat stress (Cho *et al.*, 2012; Putelat *et al.*, 2021; Leung *et al.*, 2022). However, as has been observed in the UK, yield stagnation in cereal crops has occurred, despite the increase in carbon dioxide. This indicates that on real farms, these projected yield benefits are not always seen. The overall strength of the $CO_2$ fertilisation effect on future crop growth in the UK, and globally, contributes significant uncertainty to future crop production estimates and requires further research (Ritchie *et al.*, 2019).

## 1.2 The influence of a changing climate on UK agriculture

In 2020, agriculture covered 71% of land in the UK and 19% of land was used for arable crops (DEFRA, 2021a). Arable farming is typically focussed in the drier east whilst livestock farming is more common in the west (DEFRA, 2020). Wheat and barley are the predominant cereal crops, covering 1.4 million ha each in 2020. By comparison, the next most widely grown crop was oilseed rape, covering 0.38 million ha the same year (DEFRA, 2021a). Agriculture in 2020 contributed 0.49% to the national economy and its share of employment was 1.44% (DEFRA, 2021a), whilst the broader agri-food sector accounts for around 13% of the UK total workforce (Downing and Coe, 2018). The agri-food sector is therefore a vital part of the UK economy and livelihoods.

The UK has faced uncertainty in recent years due to changes resulting from Brexit (Lang *et al.*, 2018). These include the transition from the European Common Agricultural Policy (CAP), which previously contributed to the majority of farm business income for many farms (Downing and Coe, 2018; Environment Food and Rural Affairs Committee, 2018) to the UK's Agricultural Act, and the influence leaving the EU has had on trade. Prior to Brexit, the UK sourced 30% of its food from the EU, along with a further 11% via deals negotiated by the EU with other countries. Whilst leaving the EU has created many challenges, it has also provided the opportunity for the UK to import foods from elsewhere and encourages increased domestic food production to improve the declining self-sufficiency, which is currently about 60% (Lang *et al.*, 2018). In the most recent UK Food Security Report (DEFRA, 2021b), the biggest medium to long-term risk to domestic production was identified as climate change and other environmental pressures, such as soil degradation. The current cost of soil degradation, erosion and compaction on food production is estimated to be £1.2 billion each year (DEFRA, 2021b). Hence, the challenge of increasing domestic food production is not trivial.

### 1.2.1 The importance of crop breeding

Crop breeding is a means of increasing yields and overall production, as well as adapting to the changing climate. Crop improvement through breeding relies on producing genetic combinations that result in cultivars with novel phenotypes that can differ substantially to parental varieties (Mackay et al. 2019). These cultivars can be better adapted to local growing environments and demonstrate improved biotic and abiotic stress tolerance (Lüttringhaus *et al.*, 2020; Senapati and Semenov, 2020). Since 2000, without breeding-induced crop improvement, the EU would have become a net importer of all major arable crops in 2020 (Noleppa and Cartsburg, 2021). In Wales, without further genetic improvement programmes (Bell *et al.*, 2019), suitable area for most crop species is projected to decrease due to increased drought risk.

In the UK, the typical breeding cycle for winter wheat can take at least 10 years. Variety trials arise from the need to evaluate newly bred varieties for their value in cultivation and use. Testing the relative genetic potential of the new varieties ensures the release of only the best proportion for commercial use (Laidig *et al.*, 2008). Varieties can be dropped at any stage of testing if they are outclassed by new varieties, or offer no improvement on old varieties (Mackay *et al.*, 2019).

Varieties are initially tested in single and then multilocation trials by breeders, and these stages can typically last two to three years. If a variety is successful in these trials, it is then entered into the National List (NL) trials. These multi-environment trials allow breeders to test their varieties across current climates, soil types and locations within the UK's growing area. After at least two years in the NL trials, the Recommended List committee will review variety performance and if successful, a variety will then move into the Recommended List (RL) trials until outclassed, which is typically six years, but can be over 20 years (Austin, 1999; Mackay *et al.*, 2011; Berry *et al.*, 2015).

There are deviations from the approximately 10-year breeding cycle timeline. Firstly, the release of the variety is also governed by how quickly sufficient seed can be multiplied up to sell, which is dependent on when the breeder is willing to make the commitment to this. For example, a breeder may wait to see if their variety's performance is good enough in NL1 before multiplying up the seed. Secondly, although it may take 10 years to get a variety fully listed, the rate of breeding is faster than this. A breeder can commit to using a new line as a parent at any stage, depending on their view of its potential as a parent. In spring barley, there have been cycle times of three years from at least one breeder. Finally, there is expectation that genomic prediction and selection will reduce the breeding time further (Hickey *et al.*, 2017). This has the potential to reduce the length of both breeder's trials and NL/RL trials and can also reduce the time required before new parents are selected. As such, this 10-year release time may well decrease.

### 1.2.2 Genotype-by-environment interaction

Unlike animals, plants are sessile, hence they are highly affected by the weather they experience and their environment. This has resulted in the development of complex and highly varied genetic systems that allow plants to adapt to changes in the environment in order to complete their life cycle (Miflin, 2000). The phenotype of a given plant is dependent on both its genotype and environment, along with the interaction between the two. The genotype-by-environment interaction (GxE) is important and has a crucial impact on the phenotype observed (Gillberg *et al.*, 2019). The greatest cause of variability in crop yields is variation in the environment, hence the GxE variance component of traits such as yield is often larger than that of genetic variance (Seyedsadr *et al.*, 1996; Mackay *et al.*, 2019). Consequently, variability in weather and climate can result in highly varying yield responses from year to year.

### 1.2.3 UK growing climate

The growing climate is defined as the climatic conditions experienced in the growing season, during which conditions permit plant growth to occur. The UK climate is defined as a humid temperate oceanic climate, or *Cfb*, on the Köppen climate classification system (Beck *et al.*, 2018). Its location between the relatively warm waters of the Atlantic Ocean and continental mainland Europe in the mid-latitude westerly wind belt contribute to frequent changes in air mass and variable weather that the UK is renowned for (Met Office, 2022d).

There are large regional variations in climate, with the east and south of the UK tending to be drier, warmer, less windy, and sunnier than the north and west (Figure 1.1), hence the distribution in arable and livestock farming described earlier. One of the major drivers of weather patterns in the UK, Ireland and Western Europe is the North Atlantic Oscillation (NAO). The NAO index is defined as the normalized pressure difference between Iceland and the Azores (Jones *et al.*, 1997). Stations at these sites are located close to the centres of action for the NAO and have provided homogenous pressure series from which a monthly or seasonal NAO index can be calculated. A positive NAO corresponds to a large pressure difference (with Azores pressure being higher and/or Icelandic pressure lower than average), causing westerly winds to dominate and more frequent Atlantic storms. A negative NAO corresponds to a weaker than usual pressure difference, causing easterly winds, reducing the influence of Atlantic weather systems in north-west Europe (Kendon *et al.*, 2019). NAO impacts vary seasonally: a positive winter NAO (WNAO) brings mild and wet conditions, whilst a negative WNAO brings cold and dry conditions (Maisey *et al.*, 2018). By contrast, a positive summer NAO (SNAO) is usually associated with higher temperatures but lower rainfall, while negative SNAO are associated with the opposite (Met Office, 2021b). The NAO has been shown to have impacts on cereal yields and grain quality throughout Europe (Atkinson *et al.*, 2005, 2008; Ceglar *et al.*, 2017; Heino *et al.*, 2018).

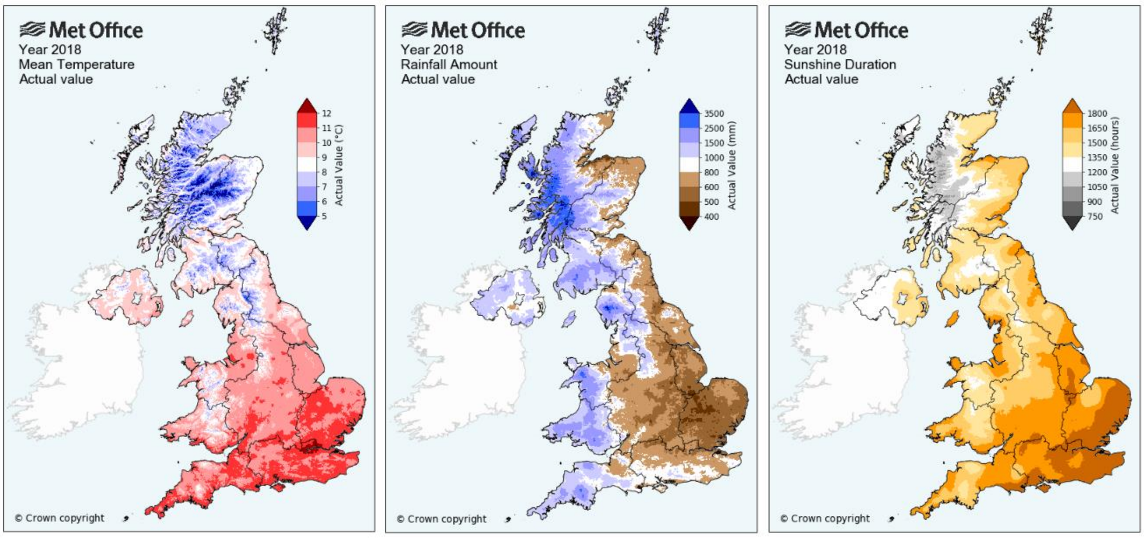



Figure 1.1*: Mean temperature, total rainfall and total sunshine duration for the UK in 2018. The maps have a 1km x 1km spatial resolution and can be downloaded from the Met Office (https://www.metoffice.gov.uk/research/climate/maps-and-data/uk-actual-and-anomaly-maps).*

Winter wheat accounts for over 95% of the wheat grown in the UK, as the temperate climate allows the crop to grow throughout the winter, contributing to higher yields than seen in spring wheat (Cho *et al.*, 2012). Wheat is often grown in a rotation with oilseed rape and sugar beet, although this has become less common given the concentration of sugar beet processing on four sites, with none north or west of Newark-on-Trent. In recent years, more spring barley has been grown than winter barley, particularly in growing seasons when there have been difficulties in establishing winter crops. In the UK, winter barley is typically sown slightly earlier in autumn than winter wheat, but drilling dates can vary widely both year-to-year as well as across the country (Turner *et al.*, 2021), depending largely on the amount of rainfall in early autumn. Harvest dates also vary widely across the country and interannually, largely due to the variation in weather and climate. Heavy rainfall in mid-summer can prevent cereal grains from drying to their required moisture content, as well as creating difficulties getting onto the land to harvest. There are a range of markets for wheat and barley grain, including feed wheat, bread wheat, biscuit wheat, feed barley and malting barley. Therefore, there is a trade off in terms of yield potential amongst the varieties to each specific market, and as such average yields reflect the diversity in variety types.

Agricultural production is intrinsically linked with climate due to its influence on viability. Globally, climatic variability is estimated to account for roughly one third of the observed global yield variability, with the explained variance exceeding 50% for maize and rice in some countries (Ray *et al.*, 2015). The sensitivity of cereals to the high UK weather variability is reflected in the interannual yield variation, with wheat yields varying by up to 50% in the last decade (Figure 1.2) (Berry and Brown, 2021). Ray *et al.* (2015) found that either precipitation variability or both temperature and precipitation variability explained approximately 45% of UK wheat yield variability. The 2020 growing season exemplified the vulnerabilities: a very wet autumn resulted in a reduction in winter cereal drilling and production dipped to a 30-year low. This was followed by difficulties establishing spring crops as a result of several storm events in February 2020 and subsequently a protracted spring dry period from late March through to end of May (DEFRA, 2021a). The media documented the impacts of the drop in yields and production totals, which resulted in sharp increases in the price of bread (Rowlatt, 2020; Tasker, 2020).

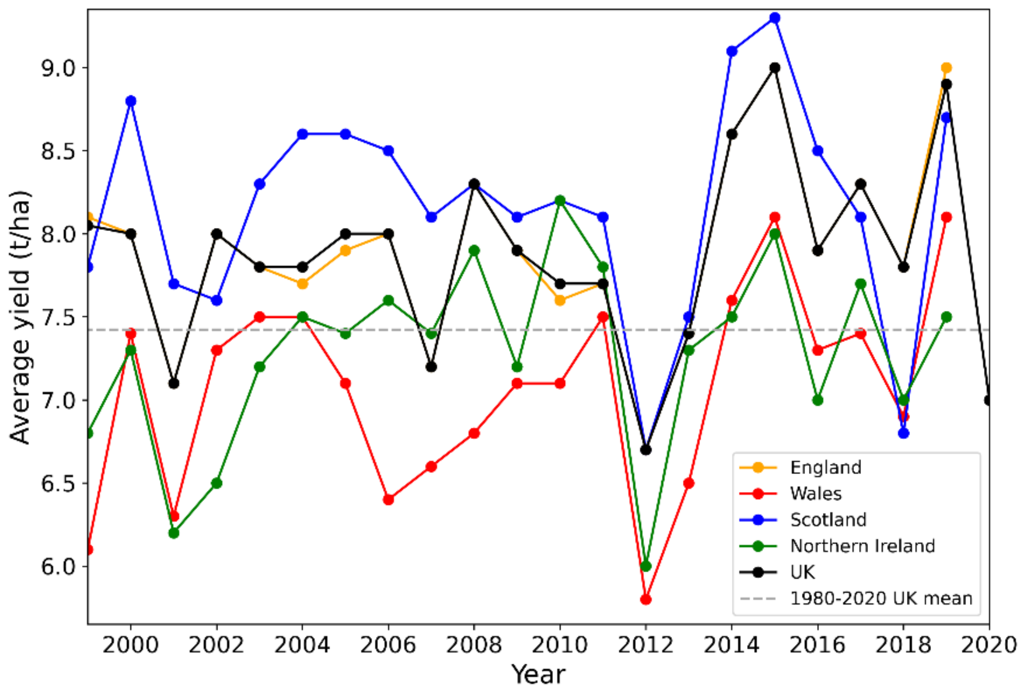



Figure 1.2*: UK (black) wheat yields (t/ha) from 1999-2020 relative to the 1980-2020 mean. England (yellow), Wales (red), Scotland (blue) and Northern Ireland (green) yields also given for 1999-2019. Data from* DEFRA (2021)*.*

### 1.2.4 Recent climatic changes

Globally, average temperatures have risen by 1.1°C since 1880 with the majority of the warming occurring since 1975 (Lenssen *et al.*, 2019; GISTEMP Team, 2022). An increase in global heatwaves, cold events, heavy rain and drought has been attributed to the changing climate (Met Office, 2022b). Analysis of global heatwave occurrences shows that since the 1950s heatwave frequency, duration and cumulative heat have all accelerated (Perkins-Kirkpatrick and Lewis, 2020). The changing climate has included a northward migration in European agroclimate zones (Ceglar *et al.*, 2019), which has resulted in an average poleward shift in the latitudinal ranges of crop pests and pathogens of 2.7 km/year since 1960 (Bebber *et al.*, 2013).

The UK has already seen an increase in the frequency and length of warm and hot spells, shorter and less frequent cold spells, less frost and snow, and several high temperature records broken (Met Office, 2022a). Notably, on the 19th July 2022, 40.3°C was recorded in Lincolnshire, setting a new UK temperature record 1.6°C above the previous record. This marked the first time temperatures have exceeded 40°C in England, and 35°C in Scotland (Kendon, 2022). The State of the UK Climate 2021 (Kendon *et al.*, 2022) showed that all of the top-ten warmest years on record for the UK have occurred since 2002. 2020 was the first year to have sunshine duration, total rainfall and temperature all in the top 10 on record (Kendon *et al.*, 2021).

### 1.2.5 Crop-climate relationships in wheat and barley

Climate change provides a multitude of challenges for UK agriculture, from longer-term changes in climate affecting crop viability, to an increase in extreme weather events, to the indirect impacts of changing pests and diseases. To understand the impacts of these changes both in recent years and in the future, the influence of weather and climate on wheat and barley at different growth stages is first explored.

Winter barley and winter wheat are drilled in autumn and harvested in late summer, whilst spring barley is sown from December until late April, although due to frost sensitivity it's typically sown later in the period in the colder North (AHDB Cereals & Oilseeds, 2018a). Development of crops sown on different dates becomes more synchronised by the increasing daylength in spring (Kettlewell *et al.*, 2003).

Barley and wheat go through the following cereal growth stages: germination (GS00-GS09), seedling growth (GS10-GS19), tillering (GS20-GS29), stem elongation (GS30-GS39), booting (GS40-GS49), ear emergence (GS50-GS59), flowering (GS60-GS69), milk development (GS70-GS79), dough development (GS80-89) and ripening (GS90-GS99) (AHDB Cereals & Oilseeds 2018a) (Figure 1.3). Speed of development differs between varieties, less so for spring barley. If there is adequate soil moisture, seeds will germinate, with germination rate controlled by soil temperature.

Temperature also drives leaf emergence during seedling growth. The first leaf emerges soon after drilling and leaves then emerge continuously on the main tillers (shoots) and stem until the flag leaf emerges. Tillering occurs after leaf three emerges and continues until stem extension begins. Tiller production and survival are affected by both the climate and husbandry. Spring barley generally produces fewer leaves and tillers than winter barley. Grain number per ear is determined between flag leaf and ear emergence (AHDB Cereals & Oilseeds, 2018a).

Once the canopy is established, it expands rapidly from the beginning of stem elongation (GS30) in spring until shortly after ear emergence in late May/early June (Kettlewell *et al.*, 2003). Grain growth starts at flowering (anthesis), which is an important transition time from vegetative to reproductive growth. Timing of flowering should maximise radiation intake but also avoid adverse biotic and abiotic stresses (Bentley *et al.*, 2013; Sheehan and Bentley, 2021). Flowering is followed by grain filling, during which reserves in the stems and leaves are redistributed to the developing grain. Grain fill depends on stem reserves and photosynthesis. When grain filling and grain growth cease, grain ripening takes a further two to three weeks, during which dry matter content increases and moisture decreases before harvest (TEAGASC, 2017; AHDB Cereals & Oilseeds, 2018a).

The influence of climate on growth and yields of crops is centralised around temperature. For winter wheat and winter barley, warmer temperatures only have a limited effect on the period between the crop emerging and the beginning of stem elongation as they must undergo vernalisation. Vernalisation is an important process in which prolonged, cold-exposure (0-10$^{o}$C for about one month or more) enables flowering in the warmer spring (Xu and Chong, 2018). For spring barley, which has little or no vernalisation requirement, development is accelerated at higher temperatures during all growth stages (Knight *et al.*, 2012). There is a risk that rising temperatures will prevent vernalisation taking place in current winter varieties, impacting negatively on grain yields. Introduction of varieties with lower vernalisation demands could be an effective adaptation strategy (Zhang *et al.*, 2013), and there is evidence that some newer cultivars already have lower vernalisation requirements (Grogan, Anderson, *et al.*, 2016; Grogan, Brown-Guedira, *et al.*, 2016; Rezaei *et al.*, 2018).

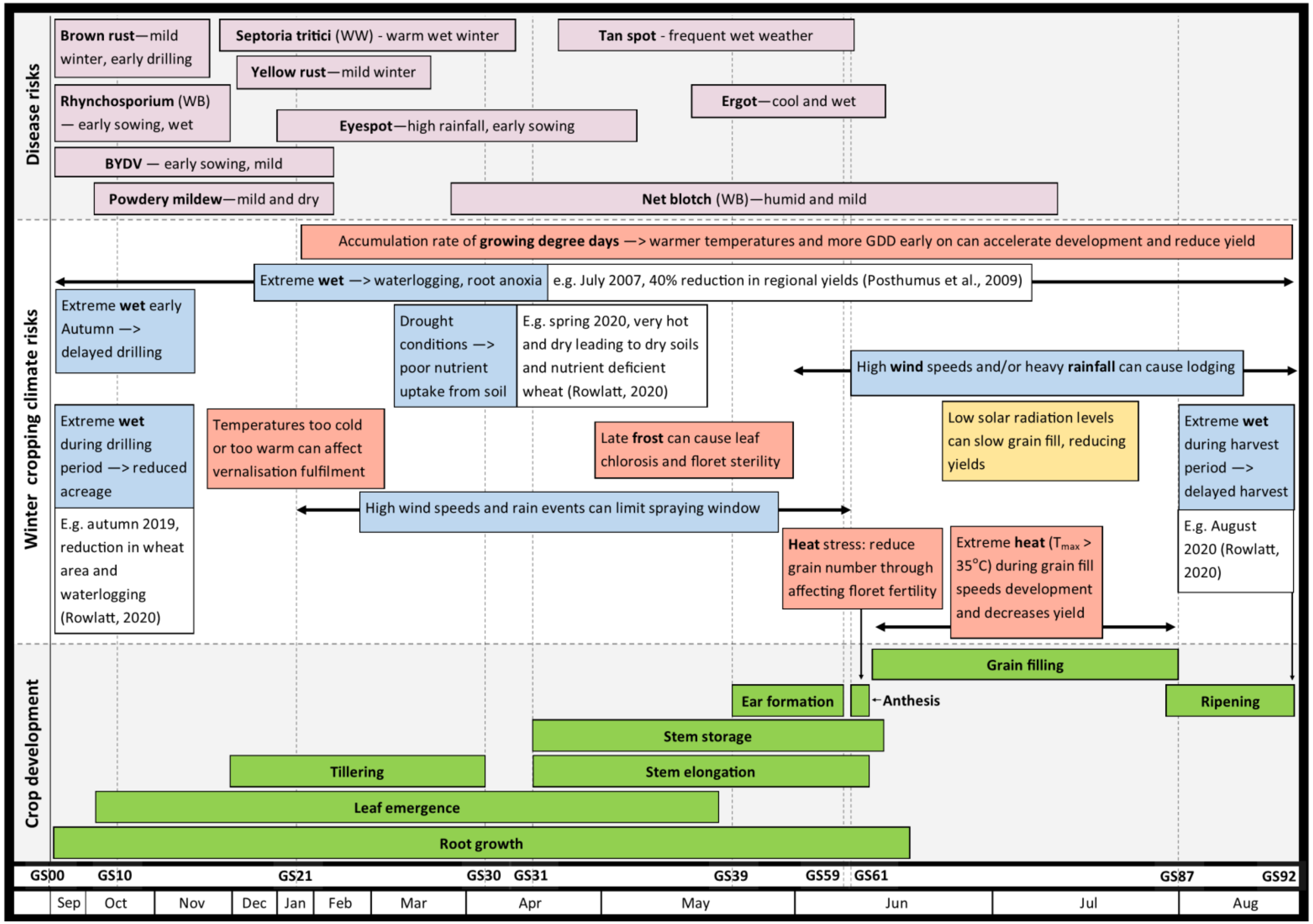



Figure 1.3*: The major disease and weather and climate risks to winter wheat and winter barley yields at various stages of the growing season. The diseases affect both winter crops, except for those indicated by WW = winter wheat and WB = winter barley. Growth stages from AHDB Cereals & Oilseeds (2018b).*

The influence of climate and disease pressure varies across the growing season for both spring and winter cereals (Figure 1.3). Meanwhile timing of weather events, such as heavy rainfall and extreme heat, also influences overall yield (Powell and Reinhard, 2016). Heavy rainfall in early autumn limits feasible drilling days due to land access issues, contracting the sowing time period, which can reduce variability in development growth stages later on, increasing the vulnerability of the crop to further weather impacts (Sheehan and Bentley, 2021). Drilling late shortens the growing season length, reducing potential yield, as well as increasing the risk of exposure to extreme heat during anthesis. In extreme years, like the 2020 growing season, extreme wet during autumn reduced crop acreage (Rowlatt, 2020). Even short periods (e.g. three days) of waterlogging can be detrimental for wheat in the first few weeks of growth (Malik *et al.*, 2002). Soil moisture availability in September influences the development of a crop's root system. A wet early autumn and therefore high soil moisture content can lead to shallower roots as water is easily accessible at that time. A larger, deeper and more efficient root system can develop in drier conditions, which can be beneficial later in the growing season in the event of drought and heat stress (Manschadi *et al.*, 2006; Senapati and Semenov, 2020). With a poor root system, drought conditions in spring can create issues with nutrient uptake from soil.

In the winter and spring months, frost is common in the UK, with the risk of occurrence decreasing later into the growing season. Late frosts in spring and summer, when winter-hardiness is lost, can cause leaf chlorosis and floret sterility, as well as grain damage, including shrunken kernels, leading to a reduction in yield (Gusta and Fowler, 1976; Cromey *et al.*, 1998; Barlow *et al.*, 2015). Heat stress ($\sim T_{max}>32^{o}C$) around anthesis can induce abnormal ovary development, reduce floret fertility and pollen viability, and drastically reduce the primary grain setting number and final yield (Saini *et al.*, 1983; Wheeler *et al.*, 2000; Grant *et al.*, 2011; Prasad and Djanaguiraman, 2014; Sage *et al.*, 2015). Drought around this time can cause premature abortion of florets and reduced viability, reducing spike fertility and seed setting (Dong *et al.*, 2017; Senapati *et al.*, 2021). Heat stress ($\sim T_{max}>35^{o}C$) during grain fill poses risk to the crop, reducing grain number and therefore final yield (Dreccer *et al.*, 2018; Rezaei *et al.*, 2018). Solar radiation during grain fill is important as it facilitates rapid dry weight growth where starch and protein are deposited in expanded grain cells, supplied by photosynthesis and the redistribution of stem reserves (AHDB Cereals & Oilseeds, 2018c). As such, rainfall and more cloud cover during grain filling reduces photosynthesis and decrease grain yields. During harvest time, heavy rainfall and waterlogging delays access to land and prevents the grain from drying (it is costly to have to dry grain once harvested), as well as causing direct damage to the crop through bending the stem, known as lodging (Posthumus *et al.*, 2009; AHDB Cereals & Oilseeds, 2018c). High wind and heavy rain any time after stem elongation increases the risk of lodging, with potential yield losses of up to 75%, and decreases the spraying window, making it difficult for a farmer to treat the crop with fungicides and pesticides.

There are many diseases that affect wheat and barley across the growing season in the UK, with varying degrees of impact on grain yield and quality. Some of the significant diseases for winter wheat and barley, and their main risk windows, are outlined in Figure 1.3, along with their respective main weather risk factors. Several of the diseases can affect both wheat and barley, albeit by slightly different fungal and bacterial strains. Additional factors affect disease susceptibility including variety, sowing date and cultivation, as well as crop rotations, so whether the crop follows a cereal or something else. For example, delayed sowing of winter wheat reduces Septoria leaf blotch pressure regardless of the variety (AHDB Cereals & Oilseeds, 2018b, 2018c).

### 1.2.6 Changing crop-climate relationships

The UK's climate is projected to see an increase in warmer, wetter winters and hotter drier summers. By mid-century, hot summers like 2018 could occur, on average, every other year and hot spells ($T_{max}$ >30°C), particularly in the south-east, are projected to increase (Met Office, 2021c). This will likely be accompanied by an increase in the intensity of heavy summer rainfall events and of short duration rainfall intensity in the autumn (Met Office, 2021c). The projected increase in intensity of rainfall events is likely to increase the risk of waterlogging in wheat (Malik *et al.*, 2002). In winter, all regions of the UK will have increased cloud cover and decreased solar radiation, whilst in summer the south will see a decrease in cloud coverage and the north-west an increase (Burnett *et al.*, 2014).

Incorporating future climate projections into process-based crop models, such as SIRIUS (Jamieson *et al.*, 1998) and DSSAT (Hoogenboom *et al.*, 2010), is a widely used method of forecasting potential impacts of climate change on future agricultural productivity (White *et al.*, 2011). Despite the projected lower total summer precipitation in the UK, relative wheat yield losses from drought (without accounting for any effects of increased atmospheric carbon dioxide) are predicted to be moderated by higher temperatures accelerating the growing season and bringing the maturity date earlier (Semenov and Shewry, 2011). For UK barley production, climate change could be beneficial, with the largest increases in yields expected to be seen in western UK, though also accompanied by increased yield variability (Yawson *et al.*, 2016).

Climate change also affects production indirectly, through changes in distributions and impacts of plant pathogens. At high latitudes and for most crops, increasing yields are likely to be accompanied by an increase in infection risk as a result of increasing temperatures (Chaloner *et* *al.*, 2021). In the UK, a rise in Fusarium ear blight, which threatens wheat ear quality, has already been seen (Turner *et al.*, 2021). Changes in climate affect cultivar susceptibility to disease, such as though interactions with temperature-sensitive disease resistance genes (Steffenson *et al.*, 2009; Dawson *et al.*, 2015).

The changing climate is accompanied and partially caused by increases in atmospheric $CO_2$. Various field experiments show that grain yield increases under increased atmospheric $CO_2$ levels (Batts *et al.*, 1997; Hazra *et al.*, 2019). In the UK, simulated yields over 1892-2016 increased by up to 9.4% when accounting for increased atmospheric $CO_2$ (Addy *et al.*, 2021b). Sowing date adjustment and additional $CO_2$ fertilisation are projected to compensate for projected wheat yield losses as a result of temperature and precipitation changes at a national level (Cho *et al.*, 2012).

There is great value in using mechanistic models to look at the potential impacts of climate change on future crop production, in terms of adaptation and decision making. However, there are also some limitations. One of the difficulties in modelling extreme events is that the temporal and spatial scale varies significantly which can affect the impacts on yields: for example, in the isolated regions where floods occur, yields are zero but other areas may be unaffected (Okom *et al.*, 2017). Extreme events like these are also unaccounted for in some models (e.g. Cho *et al.*, 2012). Modelling future climate impacts on yields is inherently uncertain due to the large uncertainties in future greenhouse gas emissions and climate scenarios. For example, in projecting the yield impacts of climate change and enhanced $CO_2$ by the 2080s for soybean, Deryng *et al.* (2014) showed that soybean exhibits both positive and negative impacts due to the differences in climate model scenarios.

Fewer models consider effects of climate variability as well as mean variables (Wreford and Adger, 2010). Timing of extreme weather events is of great importance in determining yield effects and the effects can be positive or negative depending on the week in which they occur (Powell and Reinhard, 2016). Furthermore, model intercomparisons have shown substantial differences in the ways models respond to interannual variations, indicating there is still insufficient research on how yields are affected by interannual variability in commonly studied parameters such as temperature, rainfall and solar radiation, and how these can be modelled successfully (Ruane *et al.*, 2016).

## 1.3 Using climate information to support crop breeding and agriculture in a changing climate

Current crop breeding programmes do not always incorporate current or future climate information, despite the rise in open access, high resolution climate datasets. This likely contributes to the large genetic yield gap seen in European wheat, where current local cultivars are far from their optimum (Senapati and Semenov, 2019). In the UK, the Agriculture and Horticulture Development Board (AHDB) variety selection tool (https://ahdb.org.uk/variety-selection-wheat) has a distinct lack of detailed integrated weather and climate information, beyond splitting the UK into three regions, meaning farmers can easily select varieties that aren't well suited to their growing environment. Part of this lack of inclusion is due to insufficient research into how varieties perform within the different environments and how they respond to different weather events. Highest yielding varieties do not typically have stability or resistance to extreme weather events and climate variability (Redhead *et al.*, 2020). Hence there is a potential to deliver varieties more resilient to extreme weather and select those best suited to changing local climates, by combining meteorological data with agricultural data. Climate information has the potential to help define breeding targets in new crop breeding programmes (Falloon *et al.*, 2015). The Oklahoma Mesonet, launched in 1991, is an initiative that today consists of 121 environment-measuring stations, providing real-time agroclimate data. Farmers have made use of this open access resource and it provides an excellent example of where the use of weather information for agricultural decisions has resulted in significantly increased profitability (Ziolkowska and Zubillaga, 2018). Therefore, defining breeding targets and identifying meteorological parameters most useful to UK breeders is an important line of research.

### 1.3.1 State of agroclimate services in the UK

To enable breeders, growers and local policy makers to make climate-informed decisions on crops and varieties to grow, there needs to be an accessible agroclimate resource that provides agriculturally relevant information on how the climate is changing in their locality and recent variability that may help explain observed yield.

Addy *et al.* (2021) characterised the temporal patterns of key weather variables over crop production years into 10 distinct weather patterns. One weather cluster was shown to have dominated the 21$^{st}$ century: warmer temperatures and more intense rainfall with a dry June. Five frequently occurring weather clusters in the 20$^{th}$ century have not recurred in recent times. These included cold winter and early-spring (cluster 2), cold August to September (cluster 3) and cool and dry March (cluster 10). The clusters more typical of the 20$^{th}$ century, rather than 21$^{st}$ century, were associated with higher winter wheat yields. This highlights how changes in the UK agroclimate have already affected cereal production.

The Met Office provides a valuable summary of UK weather and climate each year, in the form of the *State of the UK Climate* report (e.g. Kendon *et al.* 2022). It highlights trends, variability, and extremes in many important climate variables, including temperature, precipitation, sunshine duration, and Growing Degree Days. However, its relevance to agriculture is limited by two main factors. Firstly, its restriction to the calendar year makes it difficult to understand the full impact of weather and climate across each growing season, which for winter crops typically starts in the autumn of the previous year and ends late summer. Secondly, it focusses on months and seasons, rather than important crop-specific periods, such as anthesis and grain fill, making it harder to reveal the impacts of weather and climate variability on recorded agricultural production.

### 1.3.2 Agroclimate metrics as tools to monitor climatic changes affecting agriculture

Use of agroclimate indicators has been shown to be useful for quantifying the effect of changes in weather and climate on agriculture, and can significantly improve crop model performance over simpler raw weather data (Mathieu and Aires, 2018). They provide valuable information for supporting specific farm management decisions. Internationally there has been much research into indicators of climate change in agricultural systems (Qian *et al.*, 2013; Piticar, 2019; Hatfield *et al.*, 2020) and the US Department of Agriculture has compiled and analysed a comprehensive range of metrics (Walsh *et al.*, 2020), encompassing physical indicators such as heat waves, biological indicators such as crop pathogens, crop and livestock indicators such as animal heat stress, phenological indicators such as winter chill units and socioeconomic indicators such as crop insurance payments (Figure 1.4).

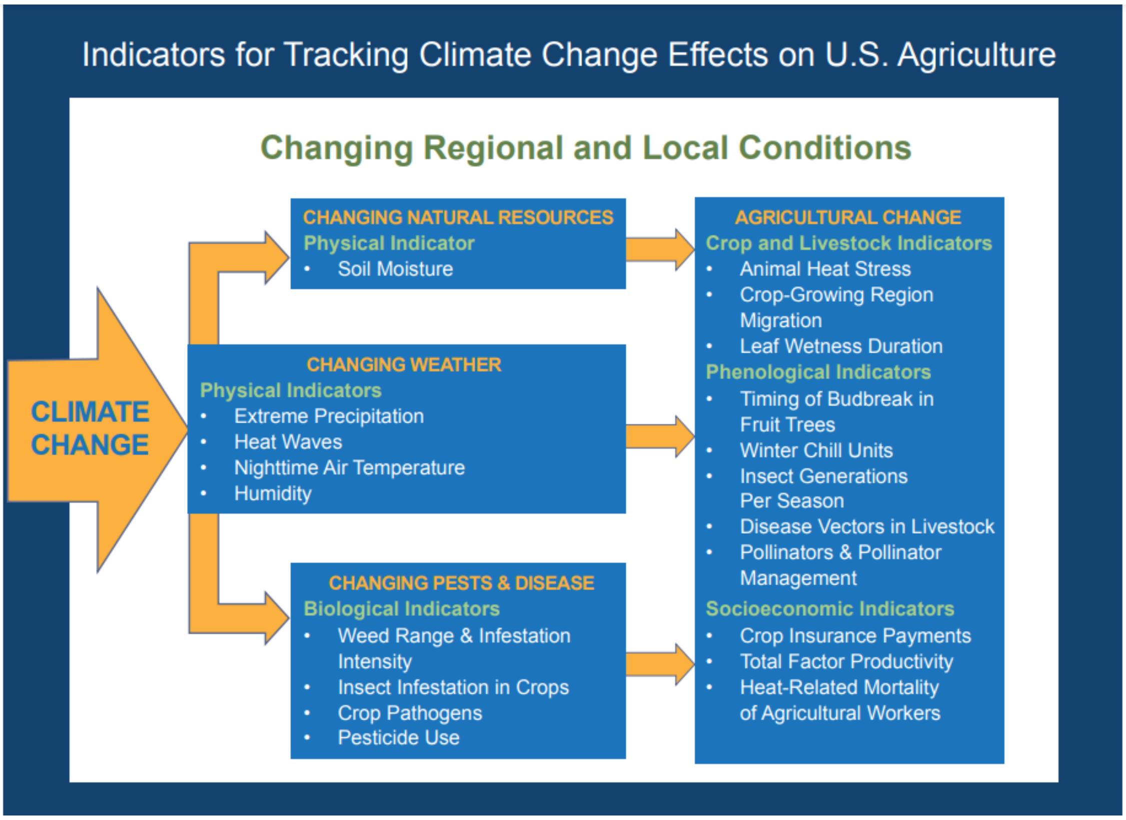



Figure 1.4: *Agroclimate and food system indicators used by the U.S. Department of Agriculture to show how climate change is influencing U.S. agriculture* (Walsh *et al.*, 2020).

In Europe, classification of climate stressors at different development stages showed that the primary climate driver of yield shocks in 2000-2018 was extreme warming in Western and Northern Europe, high water demand in Eastern Europe, and low water supply in Southern Europe (Zhu *et al.*, 2021). The critical growth stage for extreme warming driven yield shocks was the reproductive period. Furthermore, the occurrence of yield shocks due to extreme warming was projected to increase under both the moderate emissions RCP 4.5 and high emissions scenario RCP8.5. In parts of Europe, including the Mediterranean, agroclimate metrics showed that increased water deficit limited rainfed agriculture. Across Europe the risk of extremely unfavourable years and adverse weather events is likely to increase (Trnka *et al.*, 2010, 2014). The main adverse weather risk to the UK is increased field inaccessibility due to wetness in southern and eastern England (Trnka *et al.*, 2015), however this will likely be partly offset by the yield benefits of increased availability of solar radiation.

In the UK, indicators have focussed on future climate change impacts on agriculture. Arnell and Freeman (2021) project a range of agroclimate indicators using UK Climate Projections UKCP18 representing a range of greenhouse gas scenarios and both drought and heat risks for some production types will increase. These indicators contributed to the broader UK Climate Risk

Indicator tool (uk-cri.org), which enables future changes in various agriculture-related and other (e.g. wildfire) indicators to be explored at different spatial scales. By contrast, using two wheat indices, Semenov (2009) showed that despite the projected higher summer temperatures and lower precipitation, the impact of drought stress on simulated wheat yield is predicted to be smaller than present due to temperature-induced accelerated growth resulting in earlier maturity. The magnitude of impact due to increased drought risk on wheat yields is uncertain (Berry and Brown, 2021; Clarke *et al.*, 2021).

Projecting the yield impacts of late frost and heat stress during reproductive and grain fill periods under different greenhouse gas emissions demonstrated that the probability of occurrence will likely remain small by 2050 (Harkness *et al.*, 2020). A report by the Climate Change Committee provides a synthesis of the potential agricultural risks associated with exceeding extreme temperature thresholds (Jones *et al.*, 2020). A range of indices have been used to look at the observed agroclimate (Rivington *et al.*, 2013; Harding *et al.*, 2015; Arnell and Freeman, 2021). Whilst these provide an excellent summary of the agroclimate over 30-year periods, this masks the variability that is of interest. Identifying the specific causes of historical yield variability could help reduce the uncertainties associated with the projected increase in yield variability (Trnka *et al.*, 2010).

Analysis of the effect of weather on interannual variation in crop yield response to nitrogen fertilizer showed that wheat yields in the South-East of England are particularly sensitive to mean temperature in in November, April and May, and to total rainfall in October, February and June (Addy *et al.*, 2020). Using national weather and yield data in a multivariate analysis, Knight *et al.* (2012) showed that March rainfall, June sunshine and December sunshine are the weather variables most closely associated with variation in national wheat yields. Whilst these studies both provide useful insight on important monthly weather yield impacts, research on all growing regions with cereal growth-specific weather variables is still lacking.

Agroclimate metrics are often based on temperature and precipitation data, with some using sunshine duration as a proxy for solar radiation (Knight *et al.*, 2012; Ruane *et al.*, 2016). Few studies have used solar radiation data to create agroclimate metrics to specifically explain observed phenotypes, instead using monthly or growing season total solar radiation (Villegas *et al.*, 2016). Experimental techniques, such as variable shading have been used to investigate the relationship between solar radiation and yield (Kirkegaard *et al.*, 2018) however, solar radiation data is not regularly used with crop yield data. The availability of high-quality satellite solar radiation data, such as CM-SAF surface incoming solar radiation (SIS) (Pfeifroth, Trentmann, *et* *al.*, 2018), provides an excellent opportunity to explore the effect of solar radiation on important stages of the growing season and on final yields. This data will enable analysis at broader spatial and temporal scale than field experiments have thus far been capable of.

### 1.3.3 Opportunities provided by combined multi-environment trials and climate data

Identifying climatic causes of recent production variability is important for future-proofing UK agriculture as future changes in climate will affect the likelihood of challenging weather events (Arnell and Freeman, 2021). Numerous studies have utilised regional and national yield data (Peltonen-Sainio *et al.*, 2010; Olesen *et al.*, 2011; Cho *et al.*, 2012), while no known studies have used variety trials data for this purpose. Historical trials data is a valuable resource for high-quality yield data for multiple varieties in multiple environments (Smith *et al.*, 2005) and is therefore useful for understanding the drivers of GxE. The wide array of environments and varieties represented in trials can be used to see the impact of biotic and abiotic stresses on variety performance (Pidgeon *et al.*, 2006). Given that trials data is recorded every year, it reflects changes in climate over time.

Published studies using historical trials datasets date to as early as the start of the 20$^{th}$ century. Student (1923) analysed two spring barley varieties (*Archer* and *Goldthorpe*) grown in multi-environment trials from 1901-1906. He concluded that *Archer* was on average higher yielding and hypothesised that differences due to weather variability outweighed varietal differences. Other studies have revealed important results in terms of changes in crop performance over time, and have explored the environmental drivers of yield variation. Peltonen-Sainio *et al.* (2007) showed that the latest turnip rape (*B. r.* ssp. *rapa*) cultivars in Finnish variety trials were more sensitive to elevated temperatures at seed setting and filling. Use of variety trials data from across Europe showed that wheat in Slovakia has the greatest climate resilience while the Czech Republic had the least of the nine countries tested (Kahiluoto *et al.*, 2019). Peltonen-Sainio *et al.* (2009) showed that plant breeding has contributed to increased genetic yield potential of all cereal crops, despite levelling of national yields as cereal productions become less intensive.

In the UK, Mackay *et al.* (2011) combined the UK NL/RL dataset with national climate data and showed that UK winter wheat varieties had the greatest sensitivity to summer rainfall and winter temperature. However, the magnitude of the yield impacts of these climate variables was not quantified. Given the number of high resolution gridded observational and reanalyses datasets that now exist for the UK region, there is a great opportunity to build on this work based on national data to incorporate site-specific agroclimate data.

Analysis of variety performance in variety trials and national variety use can also give a good indication of variety uptake. In the USA, farmers have shown reluctance to change their crop mix or agricultural practices in response to rising temperatures, suggesting a lack of engagement in long-term production adaptation (Burke and Emerick, 2016). This reluctance to adapt will need to be overcome to ensure future food security and requires an integrated industry approach to adaptation, with research that demonstrates the value of growing locally adapted, more climate-resilient cultivars.

### 1.3.4 Isolating the contribution of breeding and climate to crop yields

Analysing the impacts of agroclimatic factors on yield is complex and requires robust statistical analysis to identify causal explanations (Shmueli, 2010). Various statistical models have been used in yield trend analyses, most frequently linear regression (Michel and Makowski, 2013; Takashima *et al.*, 2013; Laidig *et al.*, 2014; Weymann *et al.*, 2015). In particular, linear mixed models have been shown to be useful in dissecting genetic and non-genetic sources of yield variability (Mackay *et al.*, 2011; Piepho *et al.*, 2014). It is important to understand the masking effect of the environment relative to heritable traits, to quantify the effect of environment and GxE on phenotypic variation in traits (Nehe *et al.*, 2019).

In isolating the genetic effects from the environmental effects, it is possible to calculate the realized genetic gain of the trait of interest (Austin *et al.*, 1989; Austin, 1999). Genetic gain is a metric used to quantify the increase in performance of a trait achieved through selection (Jayaraman, 2000; Xu *et al.*, 2017; Sinha *et al.*, 2021). Realized genetic gain refers to the observed, rather than expected, gain due to selection over cycles (Jessica E. Rutkoski, 2019). In breeding programmes realized genetic gain is a valuable measure of success of the programme (Covarrubias-Pazaran, 2020; Covarrubias-Pazaran *et al.*, 2022) and can quantify return on investment.

Many programmes are typically more limited in length and contain many fewer varieties overall and per year than the UK NL/RL trials. Varieties can appear for a couple of years and their performance is measured against longer term control varieties to account for interannual weather variability which influences varietal performance, as well as longer term factors such as changes in agronomic practices and climate. A frequent problem in calculating genetic gain is the confounding genetic and year effects due to a lack of genetic connectivity when breeding materials are tested for just one or two years, hence the importance of the connectivity provided by control varieties (Jessica E. Rutkoski, 2019). Overall, the effectiveness and accuracy of genetic gain estimated from these programmes is not well known. Hence there is a need to quantify the influence of long-term varieties and connectivity on genetic gain calculations to understand the most robust method of calculation.

Combining trials and climate data can lead to high dimensional datasets, which can make statistical model output difficult to interpret, or even prevent models from converging. To overcome this issue, variable selection and shrinkage methods can be used on linear models and linear mixed models (Fan and Li, 2012). Identifying the best method to use on a large, high dimensional crop-climate dataset is not straightforward and few agroclimate modelling studies have utilised these methods (Gouache *et al.*, 2015; Mathieu and Aires, 2018; Addy *et al.*, 2020). As such, there is a need to explore the best-suited variable selection methods to use in agroclimate research to identify the optimal model for explaining crop traits such as yield.

## 1.4 Summary and research aims

The effects of climate change are already being observed across all aspects of society including agriculture. There is great need for breeders, climate scientists and crop modellers to collaborate to achieve future food security for the UK and beyond. This must also be achieved sustainably, minimising agriculture's environmental impact. Understanding the climatic conditions and extreme weather events plant breeders should consider and prepare for when looking for desirable traits to breed into crops is key to maximising future crop yields. Variety trials data is an underexploited industrial resource which can be combined with crop-specific agroclimate data to reveal how past crop yields have been affected by breeding programmes and interannual weather and climate variability. This can help us to understand how interannual variability affects yield, which can feed into process-based crop models. Quantifying agroclimate trends will highlight how the UK agroclimate is changing, enabling agronomists, growers and decision-makers to make climate-informed decisions, leading to variety-location combinations that will optimise yield.

The overall aim of this research was to reveal the individual and combined impacts of crop breeding and climate variability on UK cereal production. After the 'yield plateau' of the 1990s and 2000s, what do we now see emerging in yield records? How can we use variety trial records to quantify the genetic contribution to recent yield trends? Which of an array of new high-resolution climate datasets should we synthesize into our analysis in order to most effectively isolate the confounding impact of spatial and temporal climate variability? The research challenge required a strongly interdisciplinary approach, both in terms of science and methods.

To address this overall aim, the specific research objectives were as follows:

1. To select the best methods of utilising large crop-climate datasets to isolate crop-climate relationships (Chapter 3) by:
    i. Trialling variable selection methods on a combined dataset of early 20$^{th}$ century Irish spring barley variety trials data and historical weather data
    ii. Using this case study to demonstrate the extent to which interannual variation in yields can be partially explained by weather variability
2. To create the first *State of the UK Agroclimate* report as a periodic assessment of the changing UK agroclimate and its influence on production (Chapter 4) by:
    i. Analysing changes and variability in national and regional on-farm and variety trial data for the last four decades
    ii. Quantifying recent trends and variability in the UK climate that may have affected these yields through the use of carefully selected agroclimate metrics
3. To quantify the relative contribution of breeding and the environment to variety trial yield trends (Chapter 5) by:
    i. Calculating the genetic gain of UK cereals in variety trials using mixed effect modelling
    ii. Exploring the uncertainty in genetic gain estimates using case study periods extracted from the NL/RL dataset
4. To identify the significant agroclimate variables in determining winter wheat yields in the UK and the varieties with the greatest resilience in the changing climate (Chapter 6) by:
    i. Modelling historical winter wheat trials data and gridded historical weather data
    ii. Dissecting the genotype-by-environment component to identify varieties responding favourably to changing climate conditions

# 2 Materials and Methods

This thesis chapter introduces the datasets and methods used to:

1. Identify the best variable selection method to use on large crop-climate dataset and model the relationship between the agroclimate and spring barley yields in the early 20$^{th}$ century, presented in Chapter 3
2. Analyse variability and trends in the UK agroclimate since 1981, presented in Chapter 4
3. Model the genetic drivers of yield variability in UK cereal crop variety trials, presented in Chapter 5
4. Model the climate drivers of yield variability in UK winter wheat (*Triticum aestivum* L.) variety trials, presented in Chapter 6.

A combination of historical climate and agricultural datasets from a wide range of sources were downloaded and analysed. A subset of these datasets was then combined in a selection of statistical models to identify drivers of yield variability in both early 20$^{th}$ century Ireland spring barley (*Hordeum vulgare* L.) and recent UK cereal crop variety trials.

## 2.1 Datasets and methods for Chapter 3

In Chapter 3, a historical Irish barley trials dataset, documented by Student (1923), was combined with recently released climate data for the early 20$^{th}$ century to identify suitable statistical methods for quantifying the yield impacts of interannual climate variability and to explore the relative stability and resilience of the two barley varieties grown. This was a useful dataset for testing these methods before use on a much bigger, modern variety trials dataset, whilst also demonstrating that climate data from as early as the beginning of the 20$^{th}$ century can be used to help explain yield variability.

### 2.1.1 Irish spring barley data

Spring barley trials data was extracted from Student (1923) and consists of two varieties – *Archer* and *Goldthorpe* – in unreplicated 2-acre plots at 18 distinct farm locations across the barley-growing districts in Ireland (Figure 2.1a and 2.1b). Locations for each trial site are given by the town and district, from which a latitude and longitude has been estimated. The number of trial sites increased each year, from 4 in 1901 to 12 in 1906 (Table 2.1).

| Year | No. trial sites |
|---|---|
| 1901 | 4 |
| 1902 | 6 |
| 1903 | 8 |
| 1904 | 10 |
| 1905 | 11 |
| 1906 | 12 |

Table 2.1*: Number of Irish spring barley trial sites per year. Data from Student (1923).*

Yield data was recorded in barrels and stones per acre and price was recorded in £sd per acre. To give the values modern context, these was converted to tonnes/ha and £/ha, respectively.

Each trial site was paired with weather data for the growing season from the nearest weather station operating during the period (Figure 2.1b). Growing season was defined as 1st March to 31st August, based on present-day spring barley growing practices.

### 2.1.2 Irish climate data

Daily temperature data was downloaded from https://www.met.ie/climate/available-data/long-term-data-sets. The temperature data forms part of the recently released Ireland Long-term Maximum and Minimum Air Temperature dataset (ILMMT), for which raw daily observations from 12 long-term and 21 short-term maximum and minimum air temperature series were rescued from archives (Mateus *et al.*, 2020).

Daily rainfall data was obtained for the period 1901-1906 from Met Éireann and forms part of Ireland's pre-1940 rainfall records (Ryan *et al.*, 2021). Monthly rainfall totals for the Island of Ireland (IOI) were downloaded from a 305-year (1711-2016) rainfall data record (Murphy *et al.*, 2018) to allow for longer-term analysis of national rainfall trends.

Both daily climate datasets are the product of a large data rescue project by Met Éireann and Maynooth University, which also forms part of the worldwide data rescue effort I-DARE (https://www.idare-portal.org/). Part of this project involves digitising Met Éireann's pre-1960s rainfall and climate station records, including manuscripts and daily weather reports. Due to the large volume of data in need of rescue, school and undergraduate groups were involved in the transcription and helped double-key the data to reduce the risk of transcription errors (Mateus *et al.*, 2020; Ryan *et al.*, 2021).

To enable long term localised climate analysis, daily climate data for post-1960 was downloaded from the Met Éireann website (https://www.met.ie/climate/available-data/historical-data). This data was combined with the pre-1960 data (Tables 2.2 and 2.3). It is worth noting that the units change from inches to mm within some of the pre-1940 rainfall datasets; therefore, it is necessary to check the accompanying metadata prior to using the data to ensure unit consistency.

| Station | Years | Datasets used | Frequency | Units | Reference |
|---|---|---|---|---|---|
| Birr Castle | 01/01/1880-31/12/1920 | Birr Castle telegraphic reporting station (ILMMT) | Daily | °C | (Mateus *et al.*, 2020) |
| | 01/01/1921-31/12/1954 | Birr Castle telegraphic reporting station (ILMMT) | Daily | °C | (Mateus *et al.*, 2020) |
| | 01/10/1954-30/09/2009 | Station data downloaded from Met Éireann"dly4919" | Daily | °C | (Met Éireann, 2021) |
| Glasnevin | 01/01/1834-31/12/1958 | Botanic Gardens Dublin_1834-1958 (ILMMT) | Daily | °C | (Mateus *et al.*, 2020) |
| | 01/01/1961-30/11/2020 | Station data downloaded from Met Éireann | Daily | °C | (Met Éireann, 2021) |
| Phoenix Park | 18/01/1831-31/12/1958 | Phoenix Park Dublin_1831-1958 (ILMMT) | Daily | °C | (Mateus *et al.*, 2020) |
| | 01/01/1959-31/12/1959 | Phoenix Park Dublin_1959 (ILMMT) | Daily | °C | (Mateus *et al.*, 2020) |
| | 01/01/1961-31/08/2012 | Station data downloaded from Met Éireann | Daily | °C | (Met Éireann, 2021) |
| Roches Point | 14/01/1872-31/12/1920 | Roches Point_1872-1920 (ILMMT) | Daily | °C | (Mateus *et al.*, 2020) |
| | 01/01/1921-31/12/1956 | Roches Point_1921-1956 (ILMMT) | Daily | °C | (Mateus *et al.*, 2020) |
| | 01/01/1957-28/02/2021 | Station data downloaded from Met Éireann | Daily | °C | (Met Éireann, 2021) |

Table 2.2*: Temperature datasets used for each Irish weather station.*

| Station | Years | Datasets used | Frequency | Units | Reference |
| --- | --- | --- | --- | --- | --- |
| Ardee | 01/01/1886-31/07/1913 | Data_GDJ > ARDEE (LISRENNY) | Daily | inch | (Ryan *et al.*, 2021) |
| Birr Castle | 01/01/1875-31/12/1951 | Data_GDJ > BIRR CASTLE | Daily | inch [Jan 1875-Apr 1914]; mm [May 1914-Dec 1951] | (Ryan *et al.*, 2021) |
| | 01/10/1954-25/11/2009 | Station data downloaded from Met Éireann | Daily | mm | (Met Éireann, 2021) |
| Foulkesmill | 01/01/1874-31/12/1906 01/01/1914-31/12/1940 | Data_GDJ > FOULKESMILL (LONGRAIGUE) | Daily | inch | (Ryan *et al.*, 2021) |
| | 01/01/1941-31/12/2020 | Station data downloaded from Met Éireann | Daily | mm | (Met Éireann, 2021) |
| Greenore | 01/01/1876-31/12/1940 | Data_GDJ > GREENORE | Daily | inch | (Ryan *et al.*, 2021) |
| Roches Point | 01/07/1873-31/12/1940 | Data_GDJ > ROCHES POINT | Daily | inch [Jul 1873-Apr 1914]; mm [May 1914-Dec 1940] | (Ryan *et al.*, 2021) |
| | 01/01/1941-30/06/1996 01/04/2008-29/02/2016 | Station data downloaded from Met Éireann | Daily | mm | (Met Éireann, 2021) |
| Birr Castle | 1850-2010 | Long-Term-IIP-Network | Monthly | mm | (Noone *et al.*, 2016) |
| Foulkesmills | 1850-2010 | Long-Term-IIP-Network | Monthly | mm | (Noone *et al.*, 2016) |
| Phoenix Park | 1850-2010 | Long-Term-IIP-Network | Monthly | mm | (Noone *et al.*, 2016) |
| Roches Point | 1850-2010 | Long-Term-IIP-Network | Monthly | mm | (Noone *et al.*, 2016) |
| IIP_National | 1850-2010 | Long-Term-IIP-Network | Monthly | mm | (Noone *et al.*, 2016) |
| Island of Ireland monthly | 1711-2016 | IOI_1711 Series | Monthly | mm | (Murphy *et al.*, 2018) |

Table 2.3: *Rainfall datasets used for each Irish weather station.*

To locate and calculate the distance between the nearest rainfall and temperature station and each trial site, the Distance Matrix function in QGIS (QGIS Development Team, 2021) was used. The distance between a trial site and allocated precipitation station ranges from 0.6km to 58km, whereas the distance between each trial site and nearest temperature station ranges from 1km to 116km (Table 2.4, Figure 2.1b).

| **Station Number** | **Station Name** | **County** | **Rainfall** (Ryan *et al.*, 2021) | **Temperature** (Mateus *et al.*, 2020) |
|---|---|---|---|---|
| 438 | Ardee (Lisrenny) | Louth | Yes | |
| 119 | Birr Castle | Offaly | Yes | Yes |
| 1823 | Dublin (Glasnevin) | Dublin | | Yes |
| 108 | Foulkesmill (Longraigue) | Wexford | Yes | |
| 338 | Greenore | Louth | Yes | |
| 175 | Phoenix Park | Dublin | | Yes |
| 1004 | Roches Point | Cork | Yes | Yes |

Table 2.4: *The closest weather station to the early 20th century Irish barley trials sites with daily data for 1901-1906.*

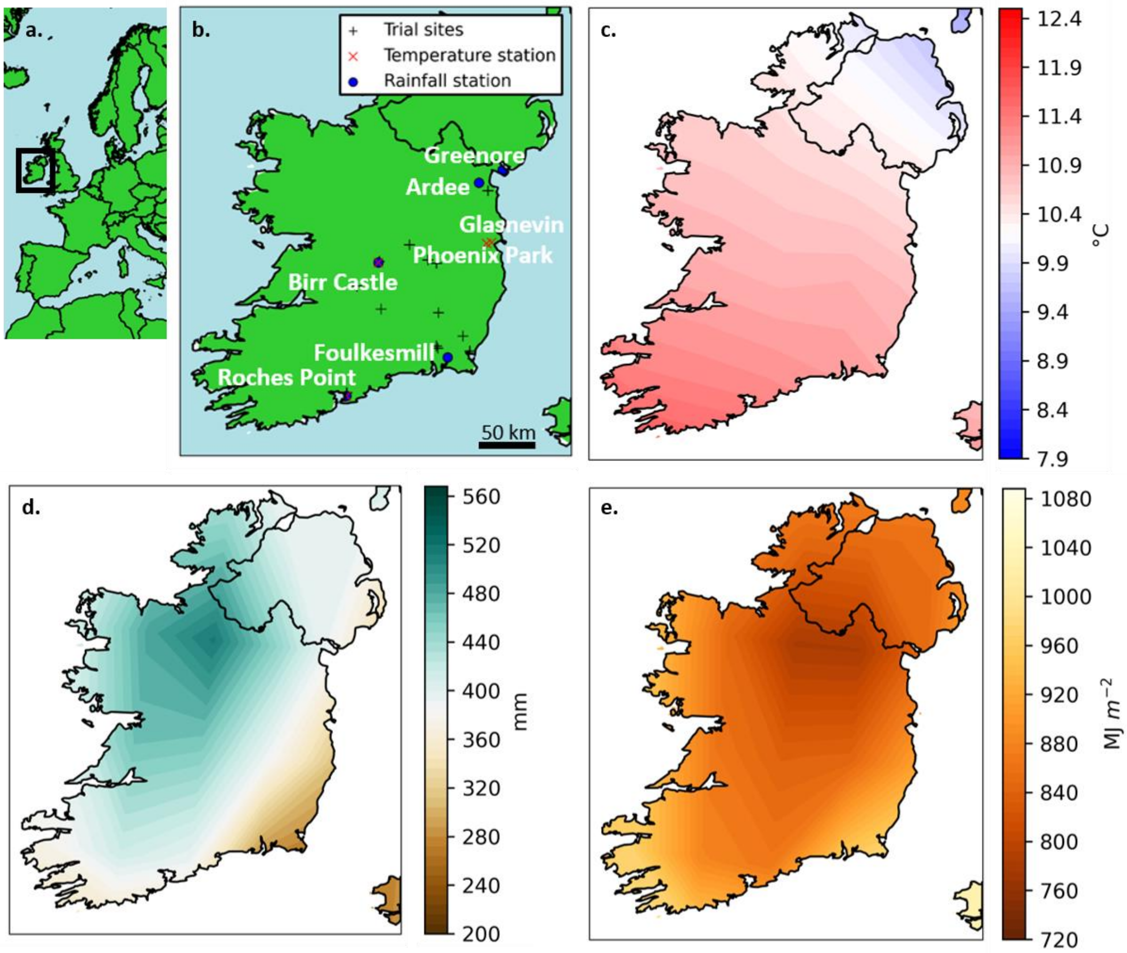



Figure 2.1: ***a.** The location of Ireland relative to Europe and North Africa.* ***b.** Weather stations open between 1901-1906 and closest to barley trial sites (+). Stations with rainfall only (blue circle), temperature only (red x) and both rainfall and temperature data are shown. Growing season (March-August) average temperature (°C) (**c.**), total rainfall (mm) (**d.**) and surface photosynthetically active radiation (MJ $m^{-2}$) (**e.**) for 1901-1930, calculated using ERA-20C* (Poli *et al.*, 2016).

There were two daily temperature datasets available for Birr Castle for the period 1901-1906. After comparing these datasets, the Birr Castle Telegraphic station was chosen over the Birr Castle second order station since it has a stronger correlation with the temperature station at Roches Point and temperature data was collected at the same times (08:00 and 18:00) as Roches Point. It also covers a significantly longer period (1880-1956 vs. 1872-1911). The two Birr stations have very high correlation for both maximum (r=0.983) and minimum (r=0.942) temperature (Figure 2.2). There were also two temperature datasets for the Glasnevin site. The Botanic Gardens Dublin dataset was chosen over the NLI dataset due to the greater length of record and the lack of missing data for 1901-1906. The two datasets also exhibit very high correlation with each other for maximum (r = 0.995) and minimum (r=0.994) temperature (Figure 2.3).

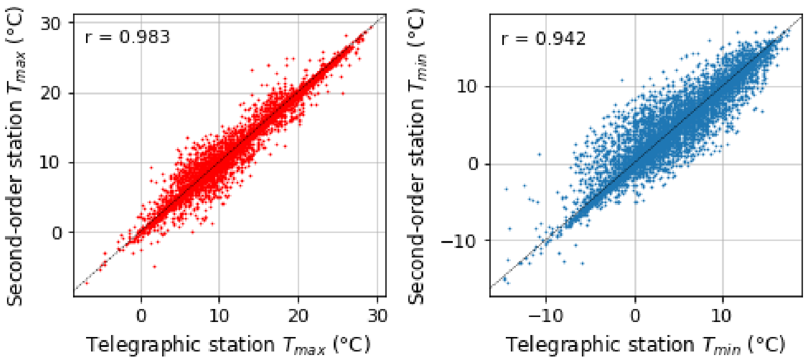



Figure 2.2: *Birr Castle temperature station comparison for daily maximum (left) and minimum (right) temperature (°C), for the period of overlap 1880-1911. The degree of correlation between the two stations is shown as well as any bias relative to the 1:1 line (black).*

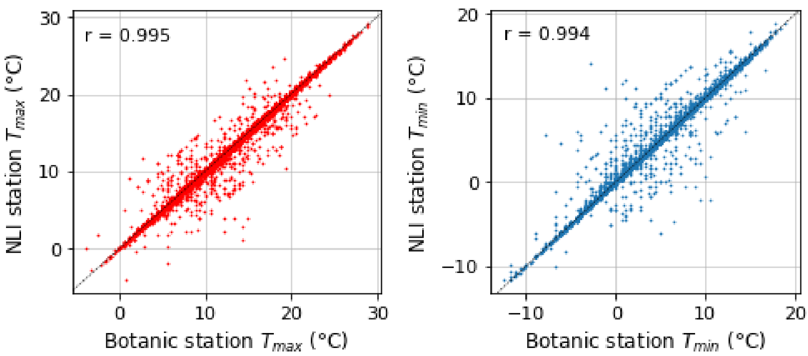



Figure 2.3: *Glasnevin Dublin temperature station comparison for daily maximum (left) and minimum (right) temperature (°C), for the period of overlap 1882-1952. The degree of correlation between the two stations is shown as well as any bias relative to the 1:1 line (black).*

In addition to station data, ECMWF's twentieth century reanalysis (ERA-20C) (Poli *et al.*, 2016) dataset was used to add a gridded and regional context to the weather stations and as a further quality control check. ERA-20C is a gridded dataset spanning 1900-2010, with a horizontal resolution of approximately 125km. It was produced with model IFS version Cy38r1 and it assimilates observations of surface pressure and surface marine winds. Daily, invariant and monthly mean data is available: http://apps.ecmwf.int/datasets/data/era20c-daily/. Here the Monthly Means of Daily Means for 2 metre temperature (K) and total precipitation (m) were used. Monthly Means of Daily Means for photosynthetically active radiation at the surface (J $m^{-2}$) were also downloaded - this was to supplement the lack of daily or monthly location-specific solar radiation data.

To download the 1° x 1°-gridded monthly ERA-20C data, the *server.retrieve()* function from the ***ecmwfapi*** module in Python (Van Rossum and Drake, 2009) was used. The files were then stitched together using the Climate Data Operator (CDO) (Schulzweida, 2021) *mergetime* command and the data type converted to double precision floating-point to overcome the common error "Numeric conversion not representable". The temperature data was converted from Kelvin to degrees Celsius and rainfall from m to mm.

As the precipitation data corresponds to mean daily total precipitation for each month, monthly total rainfall was calculated by multiplying the monthly mean of daily mean value by the number of days in the month using ***xarray*** (Hoyer and Hamman, 2017) in Python. Growing season average temperature (Figure 2.1c), total rainfall (Figure 2.1d) and total photosynthetically active radiation (PAR) (Figure 2.1e) for 1901-1930 confirm that the driest and sunniest region is the south-east of Ireland.

Growing Degree Days (GDD) were calculated using daily temperature data for growing season months March to August and the following equation:

$$GDD(5.6) = \begin{cases} \sum \dfrac{T_{min} + T_{max}}{2}, & \dfrac{T_{min} + T_{max}}{2} > 5.6 \\ 0, & \dfrac{T_{min} + T_{max}}{2} \leq 5.6 \end{cases}, \qquad [2.1]$$

where $T_{min}$ is daily minimum temperature and $T_{max}$ is daily maximum temperature. $GDD(5.6)$ gives a day-by-day sum of the number of degrees by which the mean temperature exceeds 5.6°C (Rivington *et al.*, 2013; Arnell and Freeman, 2021; Kendon *et al.*, 2022). To ensure that missing data did not incorrectly reduce the final GDD value, growing seasons with at least one day of missing maximum and/or minimum temperature data were dropped. Roches Point was missing data from 18 years (1872, 1994-2008, 2010, 2015), Dublin was missing 9 years (1959, 1960, 1963, 1965-1967, 1969, 1988, 2020) and Birr Castle 3 years (1952-1954). None of the years in 1901-1906 study period were missing data. Monthly and GDD values were then combined with the trials data using the ***pandas*** *join* function, to ensure each trial site had climate data for the nearest weather station for the months of the growing season (March to August).

In the climate analysis two averaging periods were used: the full record of selected weather datasets, to provide context for extreme events, and the relevant 30-year period, as the basis for the calculation of anomalies. For the weather station data, the 30-year period used was 1891-1920, however for the gridded reanalysis dataset ERA-20C, which begins in 1900, this was 1901-1930.

Given the known influence of the North Atlantic Oscillation (NAO) on the Irish climate, the WNAO, SNAO and spring NAO indexes were calculated for 1824-2020 using the NAO index series maintained by the University of East Anglia Climatic Research Unit (Jones *et al.*, 1997). Each index is calculated as the average NAO index across the three corresponding monthly index values. In this analysis winter consists of December to February, spring is March to May and summer is defined as June to August, to provide consistency with the rest of this work. The WNAO influence on temperature can persist into early spring, especially through its effect on surrounding sea temperatures, hence it is included in this analysis.

### 2.1.3 Modelling Irish spring barley and climate data

Total rainfall, mean maximum and mean minimum temperature for each month were calculated using the weather station data (Table 2.5). April maximum daily rainfall was also included in the analysis due to its low correlation with total rainfall in this month.

| **Variable** (units) | | **Range** |
|---|---|---|
| **Yield** (t/ha) | | [1.34,3.93] |
| **Value** (£/ha) | | [11.12,28.91] |
| **Year** | | [1901,1906] |
| **Variety** | | *Archer* or *Goldthorpe* |
| **Farm** | | 18 locations within latitude = [51.82,54.05] and longitude = [-8.23,-6.13] |
| **Maximum daily rainfall** (mm) | apr_rain_dmax | [5.3,36.1] |
| **Total monthly rainfall** (mm) | mar_rain_tot | [32.8,173.5] |
| | apr_rain_tot | [21.9,110] |
| | may_rain_tot | [15.9,95] |
| | jun_rain_tot | [30.3,143.7] |
| | jul_rain_tot | [16.2,128.6] |
| | aug_rain_tot | [53.3,202.2] |
| **Mean monthly maximum temperature** (°C) | mar_temp_max | [8.0,10.9] |
| | apr_temp_max | [10.2,13.2] |
| | may_temp_max | [12.7,17.8] |
| | jun_temp_max | [15.4,19.6] |
| | jul_temp_max | [17.6,21.1] |
| | aug_temp_max | [17.1,20.6] |
| **Mean monthly minimum temperature** (°C) | mar_temp_min | [0.5,5.0] |
| | apr_temp_min | [1.1,6.3] |
| | may_temp_min | [4.7,8.3] |
| | jun_temp_min | [7.4,11.2] |
| | jul_temp_min | [9.8,13.5] |
| | aug_temp_min | [8.1,12.6] |

Table 2.5*: Variables of interest and their range of values for Irish barley yield modelling.*

Prior to including the climate covariates, a linear model was run using *lm* in R (R Core Team, 2021) to understand the significance of year, variety and site effects, as well as their interactions:

$$y_{ijk} = \mu + v_i + r_j + f_k + vr_{ij} + fr_{jk} + vfr_{ijk} + e_{ijk} \qquad [2.2]$$

$y_{ijk}$ is the yield of variety $i$ in year $j$ at farm $k$, $\mu$ is the overall trial series mean, $v_i$ is the effect of variety $i$, $r_j$ is the effect of year $j$, $f_k$ is the effect of farm $k$, $vr_{ij}$ is the effect of variety $i$ in year $j$, $s_{jk}$ is the effect of the site, at farm $k$ in year $j$, $vfr_{ijk}$ is the interaction between variety $v_i$ with farm $k$ in year $j$ and $e_{ijk}$ is the residual term. Year was included in this model as a factor, using the *as.factor* function, and then as a variable to see if there is any specific year effect over the short six-year time span.

The spring barley variety trials are located across 18 different sites, creating a clustered dataset where trial yields are not independent. At a given site, the yields are all dependent on more similar environmental factors such as rainfall and soil type, as well as the same farmer and agronomy. Furthermore, not all farms were used each year (Table 2.1). Therefore, farm should be modelled as a random effect, creating the need for linear mixed-effect modelling. Year is also modelled as a random effect as the data is very incomplete, with some sites missing in some years.

The following model was fitted to the combined trials and climate dataset using REML through *lmer* from the ***lme4*** package (Bates *et al.*, 2020) in R:

$$y_{ijk} = \mu + T_{jk} + P_{jk} + v_i + r_j + vT_{ijk} + vP_{ijk} + s_{jk} + e_{ijk} \qquad [2.3]$$

$T_{jk}$ is the effect of monthly temperature in year $j$ at farm $k$, $P_{jk}$ is the effect of monthly precipitation in year $j$ at farm $k$, $vT_{ijk}$ is the interaction between variety $i$ and monthly temperature $T_{ijk}$ in year $j$ at farm $k$, and $vP_{ijk}$ is the interaction between variety $i$ and monthly precipitation $P$ in year $j$ at farm $k$. $s_{jk}$ is the effect of site within years, representing the interaction between year term $r_j$ and farm term $f_k$. This term has been included due to model convergence issues caused by including farm $f_k$ as a main effect in the mixed model. This also means each farm is treated as different each year, which is a more accurate representation, given the exact location of fields is unknown and may have varied, along with the corresponding environmental factors such as soil type.

The monthly variables $T_{jk}$ and $P_{jk}$ encompass the growing season (March-August) climate variables (Table 2.5). The site term $s_{jk}$ is fitted as a random effect, whilst the variety × temperature $vT_{ijk}$ and variety × rainfall $vP_{ijk}$ terms are fitted as fixed effects as the specific reaction of individual varieties (genotype) with the climate covariates is of interest.

### 2.1.4 Variable selection methods

To reduce the dimensionality of the data and identify the most significant monthly temperature and precipitation variables in determining yield to include in [2.3], best subset selection, forwards and backwards stepwise selection, the lasso (Tibshirani, 1996) and elastic net (Zou and Hastie, 2005) were used on the linear model run using *lm* in R:

$$y_{ijk} = \mu + T_{jk} + P_{jk} + e_{ijk} \qquad [2.4]$$

These were implemented in R using the functions and arguments detailed in Table 2.6. Significant variables ($p < 0.05$) in each of the selected models were identified using an analysis of variance (ANOVA). Here a type III sum of squares (SS) was used as the order of importance of the climate variables isn't known. For each method, the root mean square error (RMSE) and adjusted $R^2$ were calculated for the selected model.

Best subset selection, forwards and backwards stepwise selection were chosen as these are the most widely used variable selection methods. The elastic net and lasso were selected from the range of penalised regression methods due to their ability to reduce model complexity by setting some covariate coefficients to 0. Specifically, the lasso minimises the residual sum of squares by shrinking some variable coefficients and setting others to 0 (Zou and Hastie, 2005), whilst the elastic net encourages a grouping of strongly correlated predictors, such that they are in or out of the model together (Tibshirani, 1996). Ridge regression is often preferred over the lasso when there is a high degree of collinearity between covariates. It shrinks all coefficients towards zero by a tuning parameter, which means all variables are included in the final model. This is undesirable in this instance when the aim is to reduce the number of covariates to then include in further models.

| | Method | Functions | Packages | Arguments | Reference |
|---|---|---|---|---|---|
| 1 = *cv.b* | Best subset selection using cross-validation | *regsubsets* | **tidyverse caret leaps** | nvmax = 19 | (Wickham *et al.*, 2019) (Kuhn, 2020) (Lumley, 2020) |
| 2a = *s.bo* | Stepwise selection in both directions | *stepAIC* | **MASS** | direction = "both" | (Venables and Ripley, 2002) |
| 2b = *s.ba* | Backwards stepwise selection | *trainControl*<br>*train* | **leaps caret tidyverse** | number = 10<br>method = "leapBackward" | (Kuhn, 2020) (Lumley, 2020) (Wickham *et al.*, 2019) |
| 2c = *s.f* | Forwards stepwise selection | *trainControl*<br>*train* | **leaps caret tidyverse** | number = 10<br>method = "leapForward" | (Kuhn, 2020) (Lumley, 2020) (Wickham *et al.*, 2019) |
| 3a = *l.m* | Lasso using optimal λ that minimises the cross-validation error | *cv.glmnet*<br>*glmnet* | **tidyverse caret glmnet** | family = "gaussian", alpha = 1<br>family = "gaussian", alpha = 1 lambda = cv.lasso$lambda.min | (Wickham *et al.*, 2019) (Lumley, 2020) (Friedman *et al.*, 2010) |
| 3b = *l.1* | Lasso using λ with gives the simplest model and lies within one standard error of lambda.min | *cv.glmnet*<br>*glmnet* | **tidyverse caret glmnet** | family = "gaussian", alpha = 1<br>family = "gaussian", alpha = 1 lambda = cv.lasso$lambda.1se | (Wickham *et al.*, 2019) (Lumley, 2020) (Friedman *et al.*, 2010) |
| 4 = *e.n* | Elastic net using optimal λ and that minimise the cross-validation error | *trainControl*<br>*train*<br>*glmnet* | **tidyverse caret glmnet** | method = "repeatedcv", number = 10, repeats = 5<br>method = "glmnet", tuneLength = 10<br>family = "gaussian" | (Wickham *et al.*, 2019) (Lumley, 2020) (Friedman *et al.*, 2010) |

Table 2.6: *Methods of variable selection used to select climate covariates.*

Mixed-effect model backwards elimination was also carried out using *step* on equation [2.3] modelled using *lmer* from ***lmerTest*** package (Kuznetsova *et al.*, 2017). *glmmLasso* from ***glmmLasso*** (Groll, 2017) was used to explore the effect of using $L_1$-penalty shrinkage to the fixed effects in equation [2.3] (Groll and Tutz, 2014). Here the gaussian link function "identity" was used. A sequence of tuning parameter values λ was specified, ranging from 0 to 200 in steps of 5.

To enable comparison with the frequentist modelling methods described above, Bayesian methods were also used. Initially, a Bayesian linear regression model was fitted to the climate covariates using ***rstanarm*** (Goodrich *et al.*, 2020) packages *stan_glm* and *describe_posterior*, specifying the probability value ci = 0.95, with the equation taking the same form as [2.4]. The significant variables were then included in a Bayesian linear mixed model using *brm*, along with variety, year and site, as in [2.3].

### 2.1.5 Principal Component Analysis

A Principal Component Analysis was implemented using the *ggcorr* function from ***GGally*** (Schloerke *et al.*, 2021) and *prcomp* function from ***stats***.

### 2.1.6 Pearson's correlation analysis

Pearson's correlation analysis was used to identify the climate covariates with the highest correlation with yield as well as the degree of correlation between the climate covariates.

### 2.1.7 Akaike Information Criterion

Each climate variable was input into equation [2.3] iteratively and the significance of that variable and accompanying model Akaike Information Criterion (AIC) was calculated. The AIC was then compared with [2.3] without any climate variables to see if the additional variable improved the fit, using *anova(model1,model2)* in R. Additional climate variables were iteratively added and their significance and model AIC assessed.

### 2.1.8 Model assumptions

The first assumption is linearity, which is checked by plotting the model residuals against the predictor. If they appear random, then no mathematical transformation of the predictor or response is required (Palmeri, 2017).

The second assumption is the homogeneity of variance, which means that the variance of the residuals is equal across groups e.g. sites (Palmeri, 2017). This can be checked using *plot(model)* in R, which creates a fitted vs residual plot. If this assumption is satisfied there should be an even spread around the centred line.

The third assumption of the linear model is that the residuals have an approximately normal distribution (Palmeri, 2017). The Shapiro-Wilk test can be used to check the normality of residuals. The null-hypothesis of this test is that the population of residuals is normally distributed, therefore if the p-value is less than the selected alpha value (e.g. 0.05), the null hypothesis is rejected. For normal residuals, the Shapiro-Wilk test will return a p-value greater than the chosen alpha value.

Results of these tests are not discussed in the text as the model assumptions were not violated.

### 2.1.9 Comparison of standard error of difference between means

Student (1923) calculates the standard error of the mean difference in variety means. To understand if the models run in this analysis can improve on this value, this was first repeated using the equation $SE(d) = \frac{s_d}{\sqrt{n}}$, where $s_d$ is the standard deviation of the differences and $n$ is the number of paired trials. After checking this against Student (1923), the value was then converted to t/ha.

To find an estimate of the standard error of difference between the varieties in the selected model, the *emmeans* function in R was used. The model and variable of interest, variety $v_i$, were specified. The *contrast* function was then applied to this, using *method = "pairwise"*. This calculates the estimate of difference, standard error, degrees of freedom, t. ratio and p value for the variety pair. The statistical significance of the difference in mean values was then checked by calculating the t-statistic.

## 2.2 Datasets and methods for Chapter 4

In Chapter 4, national, regional and variety trial yield data was used to quantify recent yield trends and variability, to see if the 'yield plateau' still exists. Several climate datasets were used to create a set of agroclimate metrics to explore long-term changes and interannual variability in the UK agroclimate and how this has affected historical UK cereal production.

### 2.2.1 UK variety trials data

The National List (NL) and Recommended List (RL) field variety trial data for winter wheat, spring barley and winter barley were provided by the Agriculture and Horticulture Development Board

(AHDB) for 2007-2018. This was combined with earlier trials data for 1982-2006 (1983-2006 for spring barley) provided by National Institute of Agricultural Botany (NIAB). The AHDB Recommended Lists is managed by a project consortium of AHDB, BSPB, MAGB and UKFM. Full data is available at ahdb.org.uk/rl. Trials were located throughout England, Wales, Scotland and Northern Ireland, focused in the relevant growing areas. Since 1982, cereal trials have been split into untreated and treated trials, with the former receiving no fungicide treatment and the latter full fungicide treatment. Variables included in the trials data are yield (t/ha), site location and fungicide treatment. Drilling date, harvest date, soil type and previous cropping were also provided for 1988-2018 for winter wheat and 2008-2018 for winter and spring barley.

### 2.2.2 UK on-farm data

National annual wheat and barley yield and planted area data for 1984-2020 were downloaded from the Department for Environment, Food and Rural Affairs (DEFRA) Food and farming website (https://www.gov.uk/government/statistical-data-sets/structure-of-the-agricultural-industry-in-england-and-the-uk-at-june). Regional wheat yield data was also downloaded for 1999-2019. Very little spring wheat is grown in the UK, therefore national wheat yields were used as a proxy for national winter wheat yields (Mackay *et al.*, 2011).

To add further context to recent observed production trends, annual UK wheat and barley yield and area data for 1961-2020 was downloaded from FAOSTAT (https://www.fao.org/faostat/en/#data/QCL).

### 2.2.3 UK temperature and precipitation dataset selection

The importance of quality, accurate input weather data for crop models is highlighted in multiple studies (e.g. Battisti *et al.* 2019; Parkes *et al.* 2019). There is an array of gridded weather datasets available, with various spatial and temporal resolutions and coverage. Three high resolution datasets (Table 2.7) were downloaded and validated against observational weather station data to identify the most suitable dataset for use in this research: HadUK (Hollis *et al.*, 2019), MÉRA (Gleeson *et al.*, 2017) and ERA5-Land (Copernicus Climate Change Service (C3S), 2019). HadUK is obtained from interpolation of ground-based station data (Hollis *et al.*, 2019) while MÉRA and ERA5-Land are obtained from reanalysis of weather model runs, satellite and observed data (Gleeson *et al.*, 2017; Muñoz Sabater, 2019). The year 2017 was chosen as the validation year as it was the most recent year with available data from all of the gridded weather datasets at the time and there was minimal missing data that year.

| Name | Resolution | Time frame | Spatial coverage | Primary source type | References |
|---|---|---|---|---|---|
| HadUK | 1km; Daily | 1862-2018 | UK | Weather station data | Hollis *et al.* (2019) |
| ERA5-Land | 0.1° or ~9km; Hourly | 1981-present | Global | Reanalysis | Muñoz Sabater (2019) |
| MÉRA | 2.5km; Hourly | 1981-July 2019 | UK and Ireland | Reanalysis | Gleeson *et al.* (2017) |

Table 2.7: *Gridded weather datasets used to identify the most suitable dataset for use in the agroclimate analysis*

Observed weather data was retrieved from 21 stations around the UK. These stations belong to the Met Office weather station network (https://www.metoffice.gov.uk/) and were selected based on their proximity to important cropping regions and the completeness of the data available for daily maximum and minimum temperature and rainfall (Table 2.8, Figure 2.4). These stations also cover a wide distribution in weather conditions, with 2017 annual precipitation, for example, ranging from 483 mm (Wittering, Cambridgeshire) to 1447 mm (Cardinham, Cornwall).

| Station Number | Station Name | County | Latitude (°) | Longitude (°) |
|---|---|---|---|---|
| 3066 | Kinloss | Inverness | 57.65 | -3.56 |
| 3088 | Inverbervie | Aberdeenshire | 56.85 | -2.27 |
| 3144 | Strathallan | Stirling | 56.33 | -3.73 |
| 3158 | Charterhall | Berwickshire | 55.71 | -2.38 |
| 3171 | Leuchars | Fife | 56.38 | -2.86 |
| 3382 | Leconfield | Humberside | 53.87 | -0.44 |
| 3414 | Shawbury | Shropshire | 52.79 | -2.66 |
| 3462 | Wittering | Cambridgeshire | 52.61 | -0.46 |
| 3469 | Holbeach | Lincolnshire | 52.87 | 0.14 |
| 3590 | Wattisham | Suffolk | 52.12 | 0.96 |
| 3680 | Rothamsted | Hertfordshire | 51.81 | -0.36 |
| 3716 | St Athan | Cardiff | 51.41 | -3.44 |
| 3761 | Odiham | Hampshire | 51.24 | -0.94 |
| 3823 | Cardinham | Cornwall | 50.50 | -4.67 |
| 3872 | Thorney Island | West Sussex | 50.81 | -0.92 |
| 3882 | Herstmonceux | East Sussex | 50.89 | 0.32 |
| 3917 | Aldergrove | Belfast | 54.66 | -6.23 |
| 99008 | East Malling | Kent | 51.29 | 0.45 |
| 99025 | Sutton Bonnington | Nottinghamshire | 52.84 | -1.25 |
| 99080 | Wisley | Surrey | 51.31 | -0.48 |
| 99207 | Ross-on-Wye | Gloucestershire | 51.91 | -2.58 |

Table 2.8: *Synoptic weather stations used for validation of the gridded weather datasets HadUK, ERA5-Land and MÉRA.*

For validation purposes, 2017 daily $T_{min}$ and $T_{max}$ (°C) and rainfall (mm) nearest grid point values were extracted from MÉRA, ERA5-Land and HadUK for the 21 station locations, using the weather station latitudes and longitudes (Table 2.8).


Figure 2.4*: Met Office stations used for validating gridded weather datasets.*

*Validation of UK temperature and precipitation datasets*

For the three gridded temperature and precipitation datasets (Table 2.7), daily bias, RMSE, relative bias, relative RMSE, and Pearson's correlation coefficient were calculated against the 21 weather station data (Table 2.8, Figure 2.4). Relative bias and relative RMSE were calculated by dividing bias and RMSE by the mean observation for each station. The forecast accuracy of each dataset was also computed. For temperature, forecast accuracy of heat ($T_{max} \geq 20^oC$) days and forecast accuracy of cold ($T_{min} < 0^oC$) days was used. For rainfall, forecast accuracy was calculated for dry days (< 1 mm), where forecast accuracy is $\frac{\text{number of hits } + \text{number of correct negatives}}{\text{number forecasted events}}$.

Of the three gridded temperature datasets, HadUK estimates maximum and minimum temperature with the highest correlation, lowest bias and RMSE (Table 2.9). HadUK also has the highest forecast accuracy for heat ($T_{max} > 20^oC$) and cold ($T_{min} < 0^oC$) days.

| Statistic | Maximum temperature ($T_{max}$) | | | Minimum temperature ($T_{min}$) | | |
|---|---|---|---|---|---|---|
| | HadUK | ERA5-Land | MÉRA | HadUK | ERA5-Land | MÉRA |
| Correlation | 1.0 | 0.98 | 0.98 | 1.0 | 0.94 | 0.95 |
| Bias (°C) | 0.47 | -1.0 | -0.36 | -0.28 | 0.32 | -0.30 |
| RMSE (°C) | 0.64 | 1.63 | 1.3 | 0.46 | 1.7 | 1.5 |
| Forecast accuracy of temperatures | 0.98 | 0.93 | 0.95 | 0.98 | 0.95 | 0.96 |

Table 2.9*: Daily maximum and minimum temperature comparison statistics for three gridded temperature datasets.*

HadUK also performed the best at replicating rainfall observations. HadUK has very high correlation (0.99) with the station data (Table 2.10). ERA5-Land and MÉRA have weaker correlations (mean r = 0.79, 0.71, respectively), with much greater variability from station to station. Daily mean bias is significantly larger for ERA5-Land (0.12 mm) than MÉRA (0.027 mm), with both overestimating daily rainfall totals, but MÉRA has the highest daily RMSE. Again, HadUK has the lowest bias and RMSE, indicating better performance.

| Statistics | HadUK | ERA5-Land | MÉRA |
|---|---|---|---|
| Correlation | 0.99 | 0.79 | 0.71 |
| Bias (mm) | 0.003 | 0.12 | 0.027 |
| RMSE (mm) | 0.45 | 2.5 | 3.1 |
| Forecast accuracy of dry days (< 1mm) | 0.99 | 0.84 | 0.83 |

Table 2.10*: Daily rainfall comparison statistics for three gridded rainfall datasets.*

*Conclusion from validation*

HadUK performed best in both the temperature and rainfall comparisons. This is somewhat unsurprising given the very high resolution of the data (1 km) and that it has access to higher weather station density so doesn't need to fill in gaps using a model. It has also successfully been used in similar analyses (e.g. Arnell *et al.*, 2021). Hence HadUK has subsequently been used in this when daily UK temperature and rainfall data is required.

### 2.2.4 Additional UK climate data

Based on the results of the validation step, HadUK daily gridded weather data for 1981-2020 was subsequently downloaded from the Met Office (https://www.metoffice.gov.uk/research/climate/maps-and-data/data/haduk-grid/datasets) via the CEDA data archive using Linux (Hollis *et al.*, 2019). Three daily variables were extracted: maximum air temperature ($T_{max}$), minimum air temperature ($T_{min}$) and precipitation (P).

HadUK daily precipitation values represent the total precipitation measured between 0900 UTC on day D and 0900 on day D+1. Similarly, maximum air temperature is for the period 0900 UTC on day D to 0900 UTC on day D+1, whilst minimum air temperature is for the period 0900 UCT on day D-1 to 0900 UCT on day D (Hollis *et al.*, 2019; Tanguy *et al.*, 2019). These definitions were accounted for in calculating weekly and seasonal climate summaries used in the analysis.

*Surface solar radiation data*

Daily surface incoming shortwave radiation (SIS) data was downloaded for a UK domain from the EUMETSAT CM-SAF website for 1987-2020 (https://wui.cmsaf.eu/safira/action/viewDoiDetails?acronym=SARAH_V002). CM-SAF SIS data is derived from Meteosat satellite observations and is available for the region ± 65° longitude and ± 65° latitude (Pfeifroth, Trentmann, *et al.*, 2018).

*Hydrology data*

To calculate the water balance across the UK, gridded potential evapotranspiration and precipitation data was downloaded from the Centre for Ecology and Hydrology's CHESS-PE (Robinson *et al.*, 2020b) and CHESS-met (Robinson *et al.*, 2020a) data repositories, respectively. Both CHESS products cover 1961-2017 at 1km resolution.

The UK drought tool (Centre for Ecology and Hydrology, 2022) was also used to identify periods with an anomalous Standardized Precipitation Index (SPI). SPI values quantify the number of standard deviations by which an observed precipitation anomaly deviates from the long-term mean, accumulated over several months, for example, SPI-3 corresponds to SPI over three months (NCAR, 2020). Only SPI was available through the drought tool, hence it was used instead of the more comprehensive Standardized Precipitation and Evapotranspiration Index (SPEI).

*UK climate maps and data*

UK and regional time-series data for monthly and seasonal maximum temperature, minimum temperature, rainfall and sunshine duration were downloaded from the Met Office (https://www.metoffice.gov.uk/research/climate/maps-and-data). Temperature, rainfall and sunshine duration anomaly maps were also used for 2001-2022. Additional anomaly maps for the earlier 1981-2000 period were also created based on HadUK gridded data for rainfall and temperature anomalies and based on CMSAF-SIS data for solar radiation anomalies. Anomalies are given relative to 1991-2020, consistent with the current Met Office reference period.

Growing season (September to August) temperature, rainfall and sunshine anomaly graphs were created using the Met Office HadUK year ordered monthly time-series data (https://www.metoffice.gov.uk/research/climate/maps-and-data/uk-and-regional-series).

### 2.2.5 Quality control of variety trials datasets

Prior to statistical analysis, pre-processing and quality control of the trials data was required to amalgamate data from different databases, as summarised in Figure 2.5.

Variety names had to be collated so that variations in the use of capitals letters and misspellings did not result in individual varieties being treated as distinct. A handful of varieties included some starred "*" versions, which initially implied distinct varieties e.g. Solstice* different to Solstice, since variety names can be recycled over time. The variety IDs were found to be the same for both starred and unstarred varieties, hence all were included as the unstarred names.

Site information was only provided for 1988-2018 for winter wheat, and only 2007-2018 for winter and spring barley. For the 1988-2006 winter wheat data, this site information took the form of grid references. There were four common issues with the grid references provided (Table 2.11), relating to missing data, an incorrect main grid letter and Easting and Northing the wrong way around. In example 1 (Table 2.11), although grid reference has been used from the corresponding treated trial, soil type is not used as it is possible this untreated trial was grown in a nearby field that has a different soil type. Once corrected, grid references were then used to find latitude and longitudes for each site.

| Example | Issue | Decision |
|---|---|---|
| 1 | Untreated trial has no grid reference (GR), but there is a GR for the paired treated trial with the same sitename and other trial information for 1997. For years 1996 and 1995 treated and untreated same location | Use 1997 treated GR for untreated 1997 |
| 2 | 1992 GR is TQ whilst rest of this sitename are TL. TQ maps the site to Wales, not Essex as the sitename indicates. | Change TQ to TL |
| 3 | Eastings and Northings the wrong way around in 1994, given the sitename. TL502204 should be TL204502 | Change GR to TL204502 |
| 4 | GR, soil texture, sowing date, previous crop information missing for trial in 2006 | AHDB website* for GR |

Table 2.11: *Common site information issues in the variety trials dataset. GR = grid reference. *The AHDB Harvest Results archive (https://ahdb.org.uk/knowledge-library/harvest-results-archive) filled in the gaps for some of the missing grid reference for 2002 onwards.*

For the 2007-2018 winter wheat data, the provided latitude and longitude were used to obtain the grid references for each site for use in statistical models. All sites for all years were then mapped using QGIS, to ensure their locations were within the United Kingdom. Using the intersect function, only one winter wheat trial was found to lie outside the UK, in Ireland. Therefore, this trial site was removed.

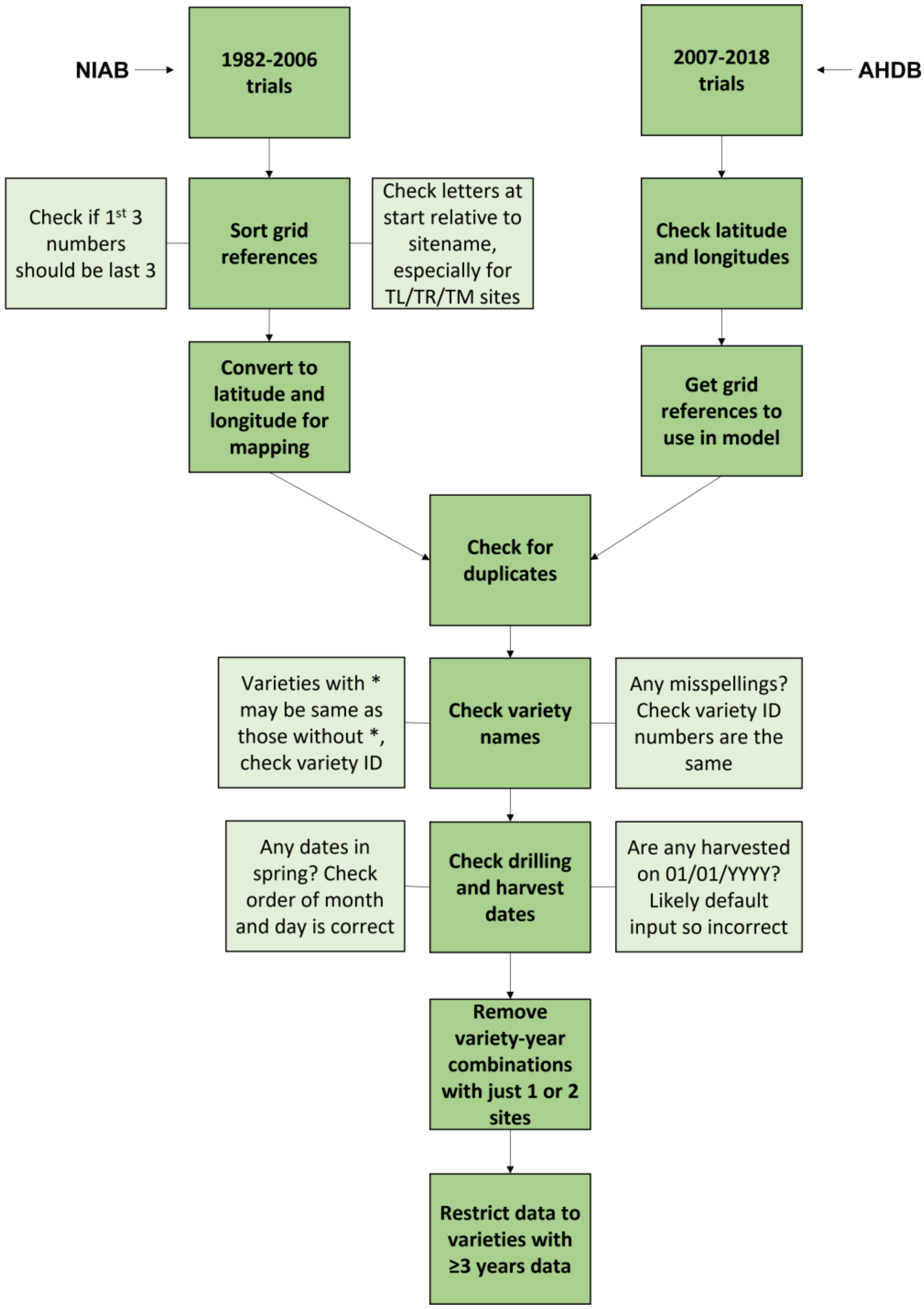



Figure 2.5*: Pre-processing and quality control steps undertaken prior to any statistical analysis. The National Institute of Agricultural Botany (NIAB) provided the 1982-2006 trials data, whilst Agriculture and Horticulture Development Board (AHDB) provided the 2007-2018 data.*

To maximise the trials included in the analysis, those with no site information were allocated unique site codes, based on the trial ID, and if this wasn't provided, sitename and harvest year were combined instead.

Data was then validated to remove duplicates. In Excel, this was achieved by sorting by years, then yield and then varieties. Variety-year combinations with only 1 or 2 sites were also removed to avoid potentially unrepresentative results due to insufficient replication. For example, the winter wheat variety *Consort* was only grown at one site in 2014. Not all years and sites are connected by common cultivars, limiting connectivity upon which to evaluate non-genetic factors. Furthermore, varieties in trial for only 1 or 2 years provide little information for trend analyses, therefore these varieties were also removed from the dataset (Mackay *et al.*, 2011; Piepho *et al.*, 2014; Laidig *et al.*, 2021).

Drilling and harvest dates were checked for potential errors. Several of the 2001 winter wheat trials were drilled in January and February 2001. The drilling of these trials was delayed because of the very wet autumn weather in 2000 (Figure 4.8) (Philpott, *pers. comm*), therefore these drilling dates were not discarded.

Several trials in 1988-1994 had harvest dates of the 1st January, which is the default date given if harvest date isn't provided, therefore these harvest dates were removed. A harvest date of 17th June 2011 was also removed as it appears abnormally early given the high yields recorded for this trial. Dates were also corrected for 4 additional trials. Specifically, WW2009ES821U had a supposed drilling date of 4th October 2010, despite being harvested 30th August 2009. The drilling date was subsequently corrected, based on additional supporting information to 29/09/08. 3 trials in 2007 had recorded harvest dates in March (WW2007ES104U) and June (WW2007NIOO9U and WW2007IS129T), which were corrected to 03/09/07, 06/09/07 and 06/08/07, respectively, upon consultation with Ellie Marshall, the data provider at AHDB.

The summary of the data structure, after quality control, can be seen in Table 2.12. The distribution of the 2007-2018 trials relative to the respective crop growing area (EDINA, 2022) can be seen graphically in Figure 2.6.

| Crop | No. observations | No. varieties | Start year | End year | Varieties per year | Treated trials | Untreated trials | T-U pairs |
|---|---|---|---|---|---|---|---|---|
| WW | 64,719 | 274 | 1982 | 2018 | 42 | 42,472 | 22,247 | 17,952 |
| WB | 38,786 | 197 | 1982 | 2018 | 29 | 21,610 | 17,176 | 11,565 |
| SB | 34,872 | 184 | 1983 | 2018 | 32 | 18,613 | 16,259 | 10,468 |

Table 2.12: *Structure of winter wheat (WW), winter barley (WB) and spring barley (SB) trials data after quality control and restricting for varieties present for minimum 3 years.*

The number of varieties and trials, and therefore total observations, fluctuate year on year for a variety of reasons (Table 2.13). Variety numbers depend on how many are submitted by breeders into NL trials and how many are then selected by the RL committees to progress through the system. Trial sites tend to remain relatively stable in their locations. Some sites are reviewed on an annual basis, whilst the core trial programme is reviewed every five years. This can cause large fluctuations when trials are moved to accurately reflect the UK crop area. Some trials can also be abandoned at an early stage, for example due to poor weather causing crop damage, hence won't be part of the data here (Marshall, *pers. comm.*). Changes in funding have also had a major influence on trial numbers. Given the large interannual variation in trial sites and varieties, this means not all varieties are grown on all sites and the trials data is unbalanced.

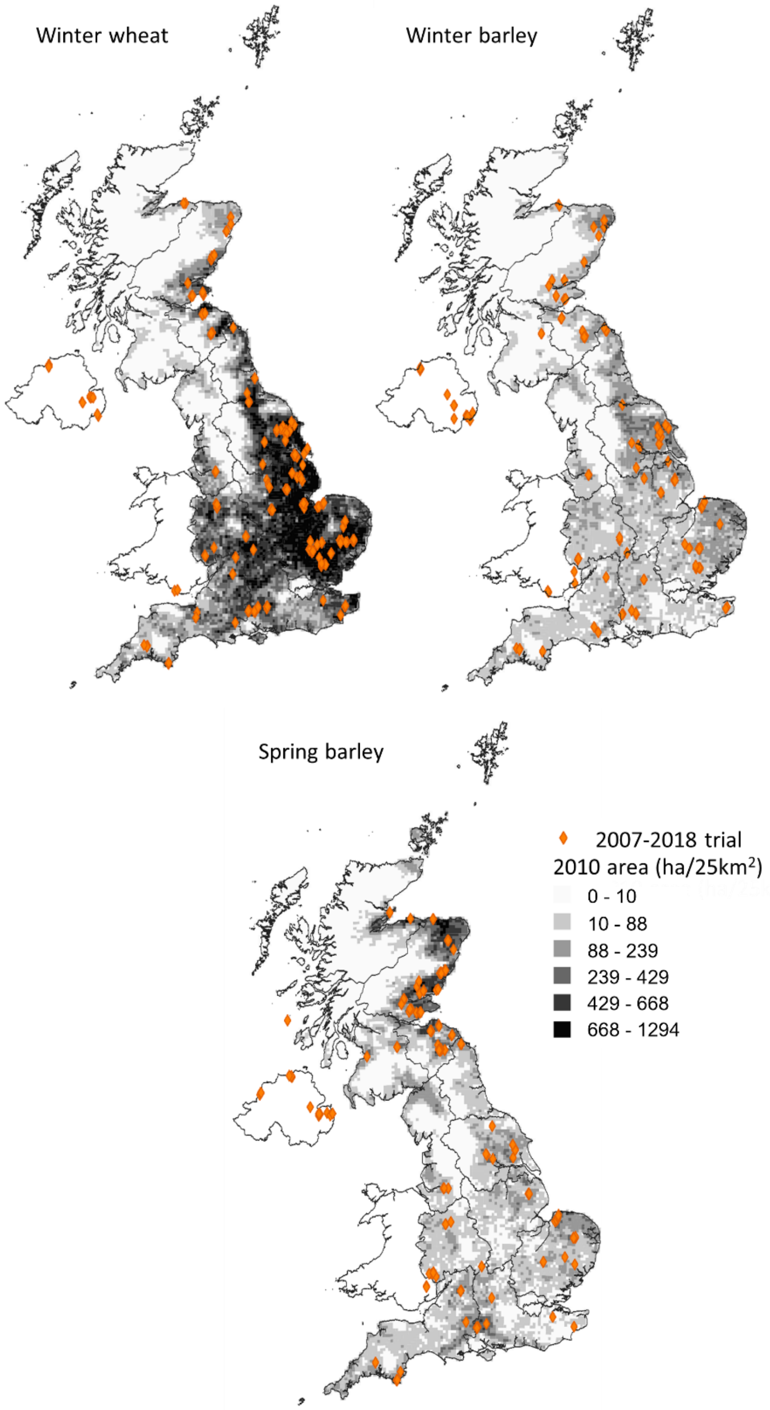



Figure 2.6*: Trial locations (red) for 2007-2018 for winter wheat (left), winter barley (centre) and spring barley (right) relative to their respective growing areas (ha per 5 km x km square) in the 2010 Agricultural Census (EDINA 2022). The regions shown in England are the government regions, whilst for Scotland these are the Met Office regions. Data on growing areas of all three crops in Northern Ireland and Wales was not available.*

| Year | Observations | Varieties | Trial sites | No. T | No. U |
|---|---|---|---|---|---|
| 1982 | 456 | 19 | 29 | 111 | 345 |
| 1983 | 1282 | 22 | 63 | 570 | 712 |
| 1984 | 1467 | 28 | 60 | 660 | 807 |
| 1985 | 1236 | 31 | 46 | 541 | 695 |
| 1986 | 1305 | 28 | 54 | 592 | 713 |
| 1987 | 893 | 25 | 43 | 398 | 495 |
| 1988 | 1704 | 33 | 84 | 995 | 709 |
| 1989 | 2049 | 41 | 81 | 1146 | 903 |
| 1990 | 1817 | 47 | 74 | 1011 | 806 |
| 1991 | 1511 | 43 | 66 | 760 | 751 |
| 1992 | 1952 | 40 | 75 | 1230 | 722 |
| 1993 | 2257 | 43 | 76 | 1367 | 890 |
| 1994 | 2527 | 44 | 68 | 1467 | 1060 |
| 1995 | 2788 | 54 | 69 | 1862 | 926 |
| 1996 | 2752 | 54 | 68 | 1852 | 900 |
| 1997 | 2764 | 51 | 59 | 1786 | 978 |
| 1998 | 2494 | 53 | 61 | 1609 | 885 |
| 1999 | 2423 | 47 | 59 | 1559 | 864 |
| 2000 | 2338 | 50 | 59 | 1505 | 833 |
| 2001 | 2300 | 56 | 49 | 1477 | 823 |
| 2002 | 2056 | 58 | 44 | 1391 | 665 |
| 2003 | 2295 | 59 | 48 | 1627 | 668 |
| 2004 | 1990 | 58 | 35 | 1448 | 542 |
| 2005 | 2073 | 56 | 53 | 1692 | 381 |
| 2006 | 2154 | 56 | 57 | 1695 | 459 |
| 2007 | 1272 | 36 | 30 | 979 | 293 |
| 2008 | 1078 | 34 | 28 | 800 | 278 |
| 2009 | 927 | 32 | 25 | 653 | 274 |
| 2010 | 1187 | 35 | 28 | 800 | 387 |
| 2011 | 1195 | 38 | 28 | 843 | 352 |
| 2012 | 1376 | 42 | 31 | 1130 | 246 |
| 2013 | 1214 | 44 | 30 | 970 | 244 |
| 2014 | 1462 | 42 | 37 | 1110 | 352 |
| 2015 | 1748 | 47 | 37 | 1388 | 360 |
| 2016 | 1630 | 47 | 36 | 1322 | 308 |
| 2017 | 1523 | 38 | 36 | 1181 | 342 |
| 2018 | 1224 | 31 | 36 | 945 | 279 |

Table 2.13*: Yearly structure of winter wheat trials data after quality control and restricting for varieties present for minimum 3 years showing large variations in the number of varieties, trial sites and consequently total observations. No. T and No. U show fluctuations in the number of fungicide treated and untreated trial observations, respectively.*

To facilitate regional analysis, the trials data was split into government regions, based on the regions shapefile ([https://geoportal.statistics.gov.uk/datasets/international-territorial-level-1-](https://geoportal.statistics.gov.uk/datasets/international-territorial-level-1-)

january-2021-uk-bgc/explore), using the zonal statistics function in QGIS. This splits England into the same regions used by DEFRA in their regional crop statistics. Met Office regions were used for Scotland (Figure 2.7).

### 2.2.6 Crop yield anomaly analysis

Trends in UK variety trial, national and regional yields were analysed for wheat and barley. Anomalous years were identified by first detrending the data and then subtracting the period mean to identify the top and bottom 10% of yields for each crop. Potential widespread climatic causes of anomalous yields were explored by looking at the ranked Met Office UK time series data and at climate anomaly maps.

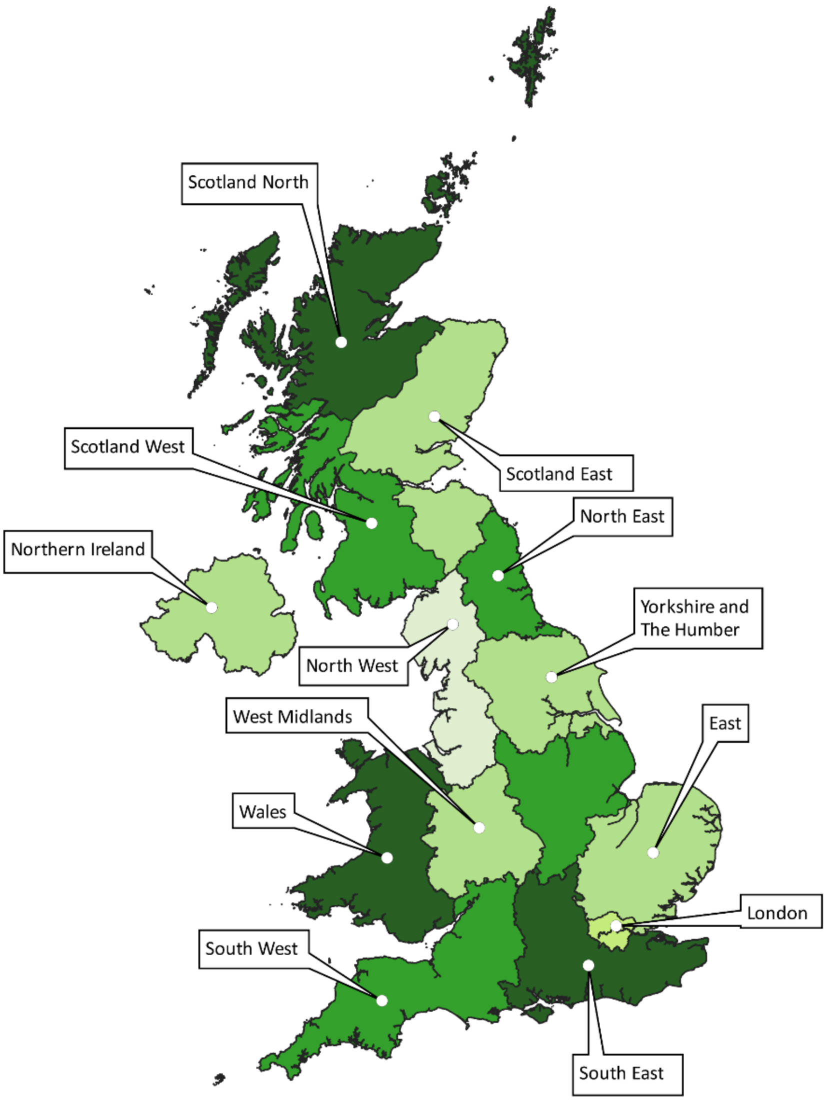



Figure 2.7: *Regions used in the analysis are based on DEFRA regions for England and Met Office regions for Scotland.*

### 2.2.7 Agroclimate metrics

After a thorough review of the literature on agroclimate indicators and influential weather on UK cereals, a list of agroclimate metrics was compiled (Table 2.14). These encompass national and regional climate-based metrics, as well as site-specific agricultural metrics, such as drilling day of year, extracted from the trials data. Drilling and harvest dates used are for winter wheat, unless otherwise indicated, due to data availability. Regional metrics are calculated for government regions for England, combined with Met Office regions for Scotland (Figure 2.7).

| Variable | Description | Data source | Motivation | Reference |
|---|---|---|---|---|
| Drilling Day of Year | - | NIAB/AHDB | Drilling date influences the number of GDD available to the crop. Drilling date is affected by autumn rainfall, as heavy rainfall can lead to a delay in drilling, as well as non-weather factors such as ploughing/direct drilling and weed pressure. Want to see if drilling day of year is getting earlier and what might be the driver of this. | |
| Harvest Day of Year | - | NIAB/AHDB | Identify trends in harvest date relative to changes in drilling date. | |
| Start of Growing Season *SOGS* | The first of five consecutive days with $T_{mean}$ > 5.6°C from 1st January | HadUK | This will show if the growing season is changing. Look at relationship between SOGS, GDD and harvest date. | (Rivington *et al.*, 2013; Harding *et al.*, 2015; Arnell and Freeman, 2021) |
| Growing Degree Days *gdd* | Sum $T_{mean}$ > 5.6°C for September to August | HadUK | GDD allows predictions of development stages. Look at variability and changes across the period. | (Harding *et al.*, 2015; Arnell and Freeman, 2021; Kendon *et al.*, 2021) |
| Vernalisation Degree Days *vdd* | Sum of vernalisation function x daily mean temperature from September to April | HadUK | Vernalisation is an important process for triggering reproductive growth for many winter crops. Increasing temperatures could affect the vernalisation fulfilment. Explore the relevance and value of this variable. | (Xiuchen Wu *et al.*, 2017) |
| Water balance *pe_balance* | 30-day running mean precipitation-evapotranspiration difference | CHESS-met CHESS-PE | Explore variability in water balance over the period to identify years with extreme rainfall and drought. | (Knox *et al.*, 2010; Daccache *et al.*, 2012; Arnell and Freeman, 2021) |
| Spring air frost days *frost03* *frost04* *frost05* | Days $T_{min}$ < 0°C per spring month | HadUK | Rather than focussing on total frosts across the year because the study is on winter crops the focus is on frosts during spring when the crop is more vulnerable. After loss of winter-hardiness frost can cause leaf chlorosis, lower stem damage and floret sterility (Gusta and Fowler, 1976; Barlow *et al.*, 2015). | (Harding *et al.*, 2015; Harkness *et al.*, 2020; Arnell and Freeman, 2021; Kendon *et al.*, 2021) |

| | | | | |
|---|---|---|---|---|
| Heavy rainfall<br>*<month>_rain10*<br>*<month>_rain20* | Days with >10mm or >20mm rainfall from September to August | HadUK | Mackay *et al.* (2011) found a sensitivity of winter wheat yield to summer rainfall. Knight *et al.* (2012) found a negative relationship between March rainfall and winter wheat yield.<br>Kettlewell *et al.* (2003) show a negative relationship between summer precipitation and winter wheat grain specific weight.<br>Heavy rain events can lead to delays in getting heavy machinery on the land to complete important jobs, such as spraying.<br>Look at changes in monthly rainfall >10mm and >20mm to see if any significant changes at critical times of growing season. | (Peltonen-Sainio, Jauhiainen and Hakala, 2009) |
| Extreme heat around anthesis<br>*anthesis32* | The number of days $T_{max} \geq 32^{o}C$ between 1st May-15th June | HadUK | Several studies categorise $32^{o}C$ as the threshold above which heat stress during the period prior to anthesis can be detrimental, causing sterility and reproductive damage. Want to see if the occurrence of these days has increased. | (Cammarano *et al.*, 2020; Arnell and Freeman, 2021) |
| Mild heat stress during grain fill<br>*grainfill31* | The number of days $T_{max} \geq 31^{o}C$ between 16th June-31st July | HadUK | Used as an index in Ceglar *et al.* (2019) and Bönecke *et al.* (2020). Significant association between number of days above 30 degrees during grain filling, grain number and therefore yield (Dreccer *et al.*, 2018; Rezaei *et al.*, 2018). | (Dreccer *et al.*, 2018; Rezaei *et al.*, 2018; Ceglar *et al.*, 2019; Bönecke *et al.*, 2020) |
| Extreme heat stress during grain fill<br>*grainfill35* | The number of days $T_{max} \geq 35^{o}C$ between 16th June-31st July | HadUK | Shortens development period and decreases yield (Nasehzadeh and Ellis, 2017; Savill *et al.*, 2018). Used as an adverse weather index (Harkness *et al.*, 2020). Heat stress during grain fill has been linked to yield stagnation in France (Knight *et al.*, 2012). | (Yang *et al.*, 2017; Harkness *et al.*, 2020; Jones *et al.*, 2020) |
| Total solar radiation during grain fill<br>*grainfillSIS* | Total SIS during the period 16th June-31st July ($MJ/m^2$) | CMSAF SIS | June sunshine shown to be important in determining wheat yield (Knight *et al.*, 2012). Solar radiation during grain fill rarely quantified, how has it changed and varied within the period? | (Knight *et al.*, 2012) |
| Septoria leaf blotch prevalence | Flag and second leaf prevalence (area %) of foliar diseases of winter wheat on-farm for 1976-2019. | Winter wheat survey data | Disease can cause widespread yield losses. It is highly dependent on the climate (temperature-humidity-rainfall during specific disease-sensitive periods), and, and as the climate changes the risk of individual diseases will change. Septoria blotch is a common wheat disease and can have significant yield impacts, hence its inclusion. | (Polley and Thomas, 1991; Hardwick *et al.*, 2001; Turner *et al.*, 2021) |

Table 2.14: Summary of selected agroclimate metrics used in Chapter 4's analysis of changes in the UK's agroclimate.

In the UK, anthesis in wheat typically occurs in early- to mid-June, at a thermal time of 2100°C days after drilling (AHDB Cereals & Oilseeds, 2018c). The "anthesis" period defined here (Table 2.14) is 1st May-15th June as it also encompasses the ~20 day period before anthesis, during which time increased temperatures have been shown to cause significant reduction of grain numbers (Yang *et al.*, 2017; Jones *et al.*, 2020). Meanwhile, grain fill is defined as the ~6-week period following on from this, 16th June-31st July.

A combination of Climate Data Operator (CDO) and Python were used to process the climate datasets to create these metrics and to explore trends and variability over time.

Growing Degree Days (GDD) were calculated using daily temperature data for the growing season months, defined nationally as September to August, and equation [2.1]. For individual trial sites, this metric has been calculated over the period defined by drilling and harvest dates.

Vernalisation Degree Days (VDD) were generated using a method described by Wu *et al.* (2017), which calculates a vernalisation effectiveness function of daily temperature (VF) extracted from the ORCHIDEE_crop model (Wu *et al.*, 2016) for the drilling to anthesis period:

$$VF = \max\left(1 - \left[\frac{T_{opt} - T_{mean}}{T_{amp}}\right]^2 ; 0\right) \qquad [2.5]$$

where $T_{mean}$ is the daily mean temperature, $T_{opt}$ is the optimal vernalisation temperature and $T_{amp}$ modifies the amplitude of the vernalising effect. Here $T_{opt}$ and $T_{amp}$ are parameterised as 6.5°C and 10°C, respectively, according to Brisson *et al.* (2009). For $T_{mean} > 16.5$°C and $T_{mean} < -3.4$°C VF is 0 (Figure 2.8). VDD is then calculated using:

$$VDD = \sum VF.T_{mean} \qquad [2.6]$$

At a national scale September-April is used as the drilling to anthesis period. At each trial site, the drilling date and daily mean temperature values were used to estimate the anthesis date, classifying it as a thermal time of 2100°C days from drilling (AHDB Cereals & Oilseeds, 2018c).

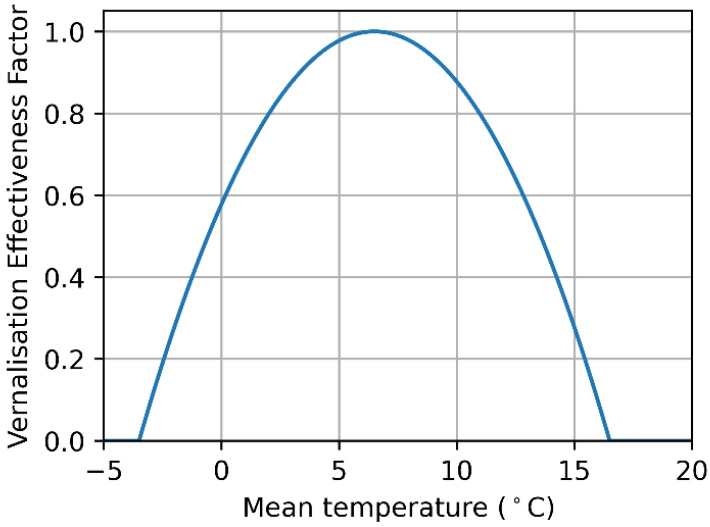



Figure 2.8: *The Vernalisation Effectiveness Factor for a range of mean temperatures ($^o$C), as used by Wu et al. (2017), to calculate Vernalisation Degree Days (VDD).*

## 2.3 Datasets and methods for Chapter 5

In Chapter 5, statistical modelling is used to quantify the contribution of crop breeding to UK cereal yields and explore the robustness of the genetic gain metric used to calculate this.

The yield datasets used in Chapter 5 are the UK NL/RL variety trials dataset after the quality control methods, as described in Sections 2.2.1 and 2.2.5, and on-farm national yields data introduced in Section 2.2.2.

### 2.3.1 Winter wheat certified seed area statistics

Winter wheat certified seed area statistics were also obtained from the British Society of Plant Breeders (BSPB) (2008-2018) and NIAB (1982-2007). These show how much winter wheat seed for each variety is sold each year as a percentage of total winter wheat seed purchased that year, and are a good representation of the proportion of national crop grown for each variety in the following year (Mackay *et al.*, 2011).

### 2.3.2 Modelling phenotype trends in UK variety trial data

To analyse changes in variety performance over time and to estimate adjusted means across locations and years for each cultivar, the following linear mixed model was initially fitted to the treated trials data and then to the untreated trials:

$$y_{ijk} = \mu + r_j + v_i + vr_{ij} + s_{jk} + e_{ijk} \quad [2.7]$$

$y_{ijk}$ is the yield of variety $i$ in year $j$ at site $k$, $\mu$ is the overall trial series mean, $v_i$ is the effect of variety $i$, $r_j$ is the effect of year $j$, $vr_{ij}$ is the effect of the interaction between variety $i$ and year $j$, $s_{jk}$ is the effect of site $k$ in year $j$ and $e_{ijk}$ is the residual term (Mackay *et al.* 2011).

Year $r_j$ is fitted as a factor and fixed effect due to the anticipated large non-linear effects of year and to provide consistency with Mackay *et al.* (2011). Likewise, variety $v_i$ is included as a fixed effect as it is historical data and individual varietal performance is of interest. The interaction terms varieties x years $vr_{ij}$ and sites within years $sr_{jk}$ are fitted as random effects as the data is incomplete.

Estimated variety effects and year effects calculated this way are known as the best linear unbiased estimators (BLUEs) for variety and year, respectively. The BLUEs for variety and year were then regressed on year of variety entry and calendar year, respectively, to arrive at estimates for genetic gain and the effect of the environment. The standard error of the regression model is also extracted to quantify the uncertainty in the genetic gain estimate.

### 2.3.3 Modelling the breakdown of disease resistance

Changes in disease resistance of a variety can be observed by comparing treated and untreated trials grown at the same location. In this analysis, 17,952, 11,565 and 10,468 treated-untreated pairs for winter wheat, winter barley and spring barley, respectively, were available for varieties with a minimum of three years in the trials dataset (Table 2.12). In calculating the yield difference in treated and untreated pairs, environmental effects cancel, and it is possible to quantify loss of yield due to disease, as well as quantify contributions of breeding to the disease resistance of untreated varieties. The fitted model used is:

$$y_{ijk}^{d} = \mu^{d} + a_j + v_i^{d} + va_{ij} + sr_{jk}^{d} + e_{ijk}^{d} \qquad [2.8]$$

$y_{ijk}^{d}$ is the yield difference in treated and untreated trials for variety $i$ at site $k$ after $j$ years, $\mu^{d}$ is the mean difference, $v_i^{d}$ is the effect of variety $i$ on yield difference, $a_j$ is the effect of variety age $j$ on yield difference, $va_{ij}$ is the effect of the interaction between variety $i$ and variety age $j$, $sr_{jk}^{d}$ is the effect of site $k$ in year $j$ on yield difference and $e_{ijk}^{d}$ is the residual term.

The variety effect $v_i^{d}$ and variety age effects $a_j$ were fitted as fixed effects, whilst the variety x variety age $va_{ij}$ and site $s_{jk}^{d}$ terms are fitted as random effects. Estimated variety age effects were then regressed on yield difference to see the extent of disease resistance breakdown as varieties age. Variety effect $v_i^{d}$ represents an average resistance of a variety over its lifetime and variety age $a_j$ represents the average decay in resistance of all varieties over their lifetimes. The variety x variety age interaction term $va_{ij}$ is important for understanding whether different varieties lose resistance at different ages.

Variety age was calculated by subtracting the harvest year from the year of entry into the trials system. For varieties present pre-1982, year of entry was found using old trials datasets, courtesy of Ian Mackay. Otherwise, year of entry is defined as the first year a variety is present in the trials data provided.

### 2.3.4 Estimating uncertainty in genetic gain estimates

To explore the various influences on genetic gain estimates and their uncertainty in trials datasets, the NL/RL winter wheat treated trials data was subset into case study periods. To define these case study periods, control varieties present for at least 10 consecutive years were first identified. These were referred to as the "checks", which are frequently used in trials datasets to provide a comparison for new varieties (e.g. Fabio *et al.*, 2017; Rutkoski, 2019). A connectivity table was then created, to identify case study periods when at least four checks overlapped. These case study periods are summarised in Table 2.15.

Breeding programmes frequently trial varieties for just one or two years, therefore rather than excluding varieties present for less than three years in the trials data, here only data for the first, second and third year of each variety is included. Initially a Python script was written to ensure that the two or three years were consecutive, as would be the case in a breeding programme. However, after running the analysis, it became evident that for 2008 this excluded all but one variety grown that year from the dataset, preventing models encompassing this year from converging. 13 varieties that were introduced into the NL/RL trials system in 2006 and grown in 2008, were not present in the 2007 acquired data. These varieties are *Bantam, Cassius, Conqueror, Gallant, Grafton, Lear, Panorama, QPlus, Scout, Shogun, Viscount* and *Walpole*. Requests were made for this data without success. An exception was therefore made to these 13 varieties, allowing the first three years to not be consecutive to enable the models to run. The check varieties and number of observations for each case study are shown in Table 2.15.

| Case study (CS) | Period | Check varieties | Observations (0 checks) |
|---|---|---|---|
| 1 | 1982-1991 | *Norman, Galahad, Fenman, Longbow* | 2,890 |
| 2 | 1991-1999 | *Riband, Hereward, Soissons, Hussar* | 4,063 |
| 3 | 1996-2006 | *Claire, Riband, Hereward, Consort, Soissons* | 5,982 |
| 4 | 2000-2009 | *Claire, Consort, Solstice, Robigus* | 5,107 |
| 5 | 2005-2015 | *Claire, Cordiale, Solstice, JB Diego, Alchemy* | 6,249 |
| 6 | 2008-2017 | *JB Diego, Viscount, Cordiale, Grafton* | 7,263 |

Table 2.15*: Genetic gain winter wheat case studies for various periods and check varieties. Check varieties are those present for at least 10 consecutive years in the NL/RL trials dataset. Only the first three years of trials for varieties present for less than 10 years are included in each case study dataset. 'Observations' indicates the number of data points within the period for the 0-check model. Adding in checks increases the data points.*

For each case study (Table 2.15), the genetic gain is calculated as in Section 2.3.2, by first estimating the BLUEs for year and variety and then regressing adjusted variety means on year of entry but here excluding the check(s). Checks are removed for the regression estimate to avoid biasing the estimate down by older year of entries. Initially this is calculated for all checks and varieties. For a case study with five checks, each check is then dropped individually to calculate the genetic gain with a combination of four checks. Subsequently each combination of two checks are dropped, then three checks, four checks and finally all checks. Hence the effect of the number of checks and the checks chosen on genetic gain and its uncertainty can be investigated.

In addition to the case study periods, the genetic gain for the whole period 1982-2018 was recalculated, but this time using the dataset with the first 1-3 years for each variety to see the effect that this had on the calculated value.

### 2.3.5 Quantifying the contribution of genetic improvement to national yields

The certified seed statistics can be used to estimate the contribution of variety improvement to national yield. Winter wheat trials data was merged with the winter wheat seed statistics, taking care to match trials data year of harvest with the previous year's seed statistics. Initially, treated and untreated trial yields were correlated with national yield to see which correlate better. Treated trials are selected due to the higher correlation ($r_t = 0.94$, $r_u = 0.60$) and the bias in the untreated trials due to disease breakdown.

For this analysis, variety and year effects were taken from the treated phenotype trends analysis (Section 2.3.2). Many varieties purchased and present in the certified seed statistics in the 1980s entered the trials system pre-1982. Given these were excluded in the calculation of variety and year effects in Section 2.3.2, these were recalculated without a restriction on their year of entry. A high proportion of the varieties in the 1982 seed statistics have insufficient years of data (< 3) in the 1982-2018 trials data to allow for variety effects to be calculated, therefore 1982 is not included in the subsequent analysis to prevent national yield estimates being based only on a couple of varieties.

Following a three-step method adapted from Mackay *et al.* (2011), the deviation in variety effect was first calculated for each variety:

$$u_i = v_i - \mu_v \qquad [2.9]$$

where $u_i$ is deviation in variety effect for variety $i$, $v_i$ is the variety effect for variety $i$ and $\mu_v$ is the mean variety effect.

An estimate of national yield from the trials data is calculated using:

$$z_j = \mu + \frac{\sum_i w_{ij} u_i}{\sum_j w_{ij}} + y_j \qquad [2.10]$$

where $z_j$ is the national yield estimate for the $j$th year; $\mu$ is the trials mean; $y_j$ is the year effect of year $j$; $w_{ij}$ is the proportion of certified seed of variety $i$ grown in year $j-1$.

Given the estimated national yield in trials $z_j$ was greater than observed national yield, the estimate was scaled by the ratio of observed to estimated yield for each year. Variety contributions for each year were also adjusted to the reference year 1983 by subtracting $\frac{\sum_i w_i u_i}{\sum w_i}$ for the reference year from all $\frac{\sum_i w_{ij} u_i}{\sum_j w_{ij}}$. The resultant estimate of the contribution of variety improvement to national yield in each year is:

$$c_j = o_{ref} + \left( \frac{\sum_i w_{ij} u_i}{\sum_j w_{ij}} - \frac{\sum_i w_{iref} v_i}{\sum_i w_{iref}} \right) \times \frac{o_j}{z_j} \qquad [2.11]$$

where *ref* refers to the reference year 1983, and $o_j$ is the observed national yield in year $j$.

## 2.4 Datasets and methods for Chapter 6

In Chapter 6, the linear mixed modelling used in Chapter 5 was modified to initially include seasonal climate data and trial site soil texture to build on work by Mackay *et al.* (2011), and subsequently to include the agroclimate metrics ([2.14]) site-specific time-series data to identify and quantify key agroclimate drivers of interannual yield variability in UK NL/RL winter wheat variety trials. Individual variety sensitivity and responses to these variables was also explored to identify varieties with greatest resilience to the changing climate. To evaluate each model the conditional $R^2$, marginal $R^2$ and RMSE were calculated. The conditional $R^2$

takes both the fixed and random effects into account whilst the marginal $R^2$ considers only the variance of fixed effects (Nakagawa *et al.*, 2017).

### 2.4.1 Reclassifying soil texture

Across the NL/RL trials that had soil texture information, there were 19 soil texture abbreviations (Table 2.16). Using the LandIS Soil Portal (http://www.landis.org.uk/data/nmtopsoiltexture.cfm) and AHDB Principles of soil management (AHDB, 2019), the extended form of the abbreviations were found. Given the sheer number of textures, and that some were present less than 100 times, soil textures were reclassified using the AHDB soil texture triangle and peat rectangle (Figure 2.9). Given the composition of shallow (SHA) soils was not known, these trials were classified as having no soil texture data (Table 2.16).

| Soil texture | | Count | Classification |
|---|---|---|---|
| Abbreviation | Texture | | |
| C | Clay | 170 | Heavy |
| CL | Clay loam | 3592 | Medium |
| DC* | Deep clay | 4312 | Heavy |
| DF* | Deep fertile silty | 657 | Sandy and light |
| DS* | Deep sandy | 1435 | Sandy and light |
| LP | Loamy peat | 187 | Peat |
| LSD | Loamy sand | 2909 | Sandy and light |
| MED* | Medium | 10576 | Medium |
| ORG | Organic | 64 | Organic |
| PL | Peaty loam | 450 | Peat |
| PT | Peat | 31 | Peat |
| SC* | Sandy clay | 206 | Medium |
| SCL | Sandy clay loam | 5755 | Medium |
| SHA* | Shallow | 2028 | |
| SL | Sandy loam | 4805 | Sandy and light |
| SZL | Sandy silt loam | 723 | Sandy and light |
| ZC | Silty clay | 361 | Heavy |
| ZCL | Silty clay loam | 4106 | Medium |
| ZL | Silt loam | 2798 | Sandy and light |

Table 2.16*: Soil textures and their classification. *Soil textures not in AHDB (2019) and Landis Soil Portal and instead suggested by Haidee Philpott (2021, pers. comm.).*

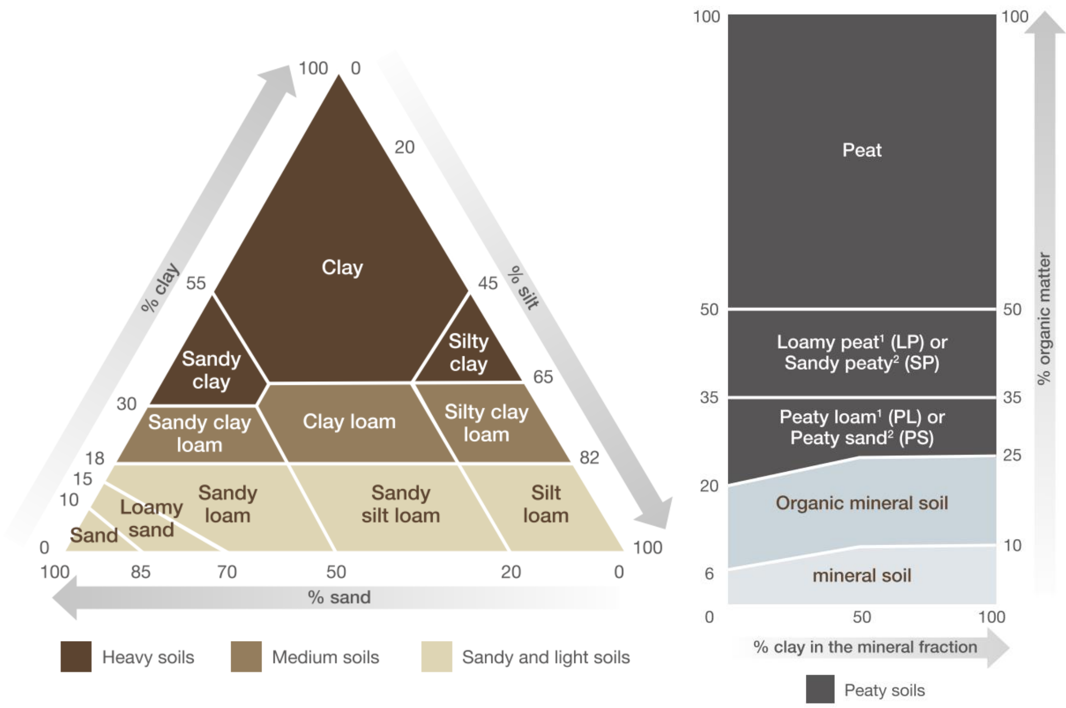



Figure 2.9: *Classification of mineral soils into soil texture classes* (AHDB, 2019).

### 2.4.2 Modelling yield impacts of seasonal climate at a regional level

To quantify the sensitivity of winter wheat yields to climate, first regional seasonal climate data (from Section 2.24) was combined with winter wheat variety trials data. This was to update the seasonal temperature and precipitation sensitivity analysis by Mackay *et al.* (2011) to include the period 2008-2018 and to quantify the magnitude of the yield impacts of significant seasonal variables. Regional data rather than national data was used due to the large climate differences (Table 2.17) between regions (Figure 2.10) and ease of access to the regional data.

| Region | $T_{max}$ (°C) | $T_{min}$ (°C) | Annual total precipitation (mm) |
|---|---|---|---|
| East Anglia | 14.4 | 6.3 | 619 |
| England E & NE | 12.8 | 5.3 | 776 |
| England NW & Wales N | 12.5 | 5.7 | 1313 |
| England SE & Central S | 14.5 | 6.4 | 788 |
| South Wales & England SW | 13.5 | 6.4 | 1254 |
| Northern Ireland | 12.5 | 5.5 | 1138 |
| Midlands | 13.6 | 5.7 | 792 |
| Scotland E | 11.0 | 3.9 | 1175 |
| Scotland N | 10.5 | 4.2 | 1695 |
| Scotland W | 11.5 | 4.9 | 1807 |

Table 2.17: *Geographic location and climatology (1982-2018) of the 10 main regions used by the Met Office. Calculated using Met Office regional series annual average data. Here $T_{max}$ and $T_{min}$ correspond*

*to the average annual maximum and minimum temperature (°C) for each region over the period 1982-2018. Precipitation values represent the average annual total precipitation over the same period.*

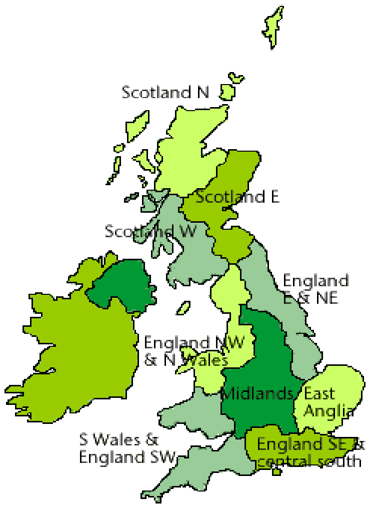



Figure 2.10*: The 10 regions of the UK used in the regional seasonal analysis. Regions are defined by the standard areas used by the Met Office (Met Office, 2020).*

Initially a linear mixed model was fitted to the seasonal climate and trials data taking the form:

$$y_{ijk} = \mu + \sum C_{jk} + v_i + r_j + vr_{ij} + \sum vC_{jk} + s_{jk} + e_{ijk} \qquad [2.12]$$

$y_{ijk}$ is the yield of variety $i$ in year $j$ at site $k$, $\mu$ is the overall trial series mean, $C_{jk}$ is the effect of the selected climate variable(s) in year $j$ at site $k$ (e.g. regional seasonal mean temperature), $v_i$ is the effect of variety $i$, $r_j$ is the effect of year $j$, $vr_{ij}$ is the effect of the interaction between variety $i$ and year $j$, $vC_{jk}$ is the interaction between variety $v_i$ and climate variable $C_k$, $s_{jk}$ is the effect of site $k$ in year $j$ and $e_{ijk}$ is the residual term. Here, the variety effects $v_i$, year effects $r_j$ and climate x variety interaction terms $vC_{jk}$ are fitted as fixed effects, whilst variety x years $vr_{ij}$ and sites within years $s_{jk}$ are fitted as random effects.

Soil texture was then added into the model. Here soil texture refers to the five reclassified soil textures (Table 2.16). Not all trials have soil types therefore this model uses a subset (30,227 trials) of the dataset used in [2.12]:

$$y_{ijkl} = \mu + \sum C_{jk} + v_i + r_j + l_l + vr_{ij} + vl_{il} + \sum vC_{jk} + s_{jk} + ls_{jkl} + e_{ijkl} \qquad [2.13]$$

where $l_l$ is the effect of soil texture $l$, $vl_{il}$ for the interaction between variety $i$ and soil texture $l$, $lr_{jl}$ for the interaction between soil texture $l$ and year $j$ and $ls_{jkl}$ for the effect of the interaction between soil texture $l$ and site $k$ in year $j$ on yield.

The effect of soil texture $l_l$ is included as a fixed effect and the variety x soil $vl_{il}$ and soil texture x site $ls_{jkl}$ interaction terms are included as random effects.

### 2.4.3 Modelling yield impacts of seasonal climate using site-specific climate data

To investigate the impact of using high resolution, localised climate data over regional data, site-specific seasonal temperature and precipitation data was combined with the treated trials data using the *join* function in Python and modelled using [2.12].

### 2.4.4 Modelling yield impacts of monthly climate using site-specific climate data

Significant climate variables in the localised seasonal analysis (2.4.3) were broken down into the corresponding monthly variables. These monthly covariates were then combined with the treated trials data and [2.12] fitted to the data to identify the most important months in determining yield variability.

### 2.4.5 Modelling the effect of the UK agroclimate on winter wheat yields

To better understand how climate affects winter wheat at different stages in the growing season, time-series data for each selected climate data derived (i.e. not drilling and harvest date nor Septoria leaf blotch incidence) agroclimate metric (Table 2.14) was downloaded for each winter wheat variety trial site location. Given the very low occurrence of the two extreme heat metrics – anthesis32 and grainfill35 – these were not included in the models as, thus far, they have not occurred frequently enough to be able to model yield impacts in the UK. Monthly rainfall days were combined to get a growing season total at each trial site each year from September to August. The agroclimate metrics were then combined with the winter wheat yield data. Given the CEH CHESS-MET and CHESS-PE datasets, and therefore the calculated potential-evapotranspiration balance (pe_balance), are only available to 2017, the combined dataset was restricted to 1988-2017 to avoid any missing climate data. The 1993 trials data also has no harvest dates which means VDD from planting to anthesis, GDD and pe_balance could not be calculated for this year. To allow easier comparisons between model performance containing different variables, 1993 was also dropped.

Prior to modelling the agroclimate metrics, a base model was run using [2.7] to allow comparison with the agroclimate model. The modelling was then split into two phases, as in Hakala *et al.* (2012). During the first phase, all varieties with at least three years of data and location information were included to establish the general relationship between specific agroclimate variables and winter wheat as a species under UK climatic conditions. After a univariate climate analysis, where each agroclimate metric and its interaction with variety were iteratively added to the base crop model ([2.7]) to test the significance of each variable's relationship with yield, the significant climate variables and their interactions with variety were then combined in multivariate agroclimate analysis. The first phase test data included 30,473 yield records. In the second phase, popular cultivars with at least 10 years of data were used (Table 2.18) which included 7,062 yield records.

Absolute yield anomaly for each yield record was also calculated by removing the linear trend over years. Yield anomaly was then modelled, as in Mathieu and Aires (2018), but due to the easier interpretation of coefficients from the yield models and similarity in significant variables between the yield and yield anomaly models, a focus on modelling yield was maintained.

A type III sum of squares (SS) was used as the order of importance of the climate variables isn't known.

### 2.4.6 Identifying climate-resilient varieties

To explore individual varietal responses to the significant climate variables, the climate covariates were classified into three categories of equal numbers of trials, using a similar method to Hakala *et al.* (2012). The average national yield for each variety across the period 1988-2017 was also calculated (Table 2.18). For each climate category, the percent yield of the average national yield for each variety was calculated. This was carried out for all the agroclimate metrics, not just significant ones.

| Variety | Average yield (t/ha) | First year | Last year |
|---|---|---|---|
| Alchemy | 10.34 | 2003 | 2015 |
| Claire | 10.24 | 1996 | 2016 |
| Consort | 10.20 | 1992 | 2014 |
| Cordiale | 10.17 | 2001 | 2017 |
| Deben | 10.48 | 1998 | 2007 |
| Einstein | 10.32 | 2000 | 2011 |
| Gallant | 10.03 | 2006 | 2016 |
| Grafton | 10.40 | 2006 | 2017 |
| Hereward | 9.09 | 1998 | 2008 |
| JB Diego | 10.61 | 2005 | 2017 |
| Malacca | 9.74 | 1994 | 2007 |
| Mercia | 8.40 | 1983 | 1998 |
| Riband | 9.68 | 1985 | 2006 |
| Robigus | 10.53 | 2000 | 2012 |
| Savannah | 10.59 | 1995 | 2004 |
| Scout | 9.91 | 2006 | 2016 |
| Soissons | 8.96 | 1991 | 2007 |
| Solstice | 10.05 | 1999 | 2015 |
| Viscount | 10.57 | 2006 | 2017 |
| Xi19 | 10.55 | 1999 | 2008 |

Table 2.18*: Average yields for varieties present at least 10 years in the UK National List/Recommended List winter wheat variety trials dataset. The first and last year each variety was present in the trials dataset is given.*

The genotype-by-environment interaction (GxE) was investigated further by extracting the variety x climate interaction terms from both the final multivariate model and the univariate climate model after it was rerun on the varieties present at least 10 years. For each climate variable, the numbers of varieties with significance exceeding a false discovery rate (FDR) threshold of <0.5 (Mackay *et al.*, 2011) was assessed to understand the sensitivity of UK winter wheat to these variables, as well as individual varietal sensitivity to them. Using FDR enables the detection of the most sensitive varieties, rather than just relying on statistical significance.

# 3 Using the case of variability in early 20th Century Irish barley yields to identify best suited statistical modelling methods

Spring barley (*Hordeum vulgare* L.) is the most widespread spring crop in Ireland as it is well suited to the Irish soils and long growing season, which offer high yield potential (TEAGASC, 2017). In 2020, yields averaged 7.1 t/ha and over 1 million tonnes of spring barley were produced (CSO, 2021). Barley has been grown in Ireland for centuries. In the late 1800s, yields were small and highly variable across farms and plants and, as today, very weather dependent. Over 80% of barley grown during the 19th century was *Chevalier* (Maxted *et al.*, 2014). In the 1890s John Bennett, a grower of barley for Guinness, began comparing qualities of different seed varieties extending to several farms across Ireland. As a consequence, the Department of Agriculture and Technical Instruction (DATI) founded a Cereal Breeding Station on Bennett's farm at Ballinacurra, County Cork, with financial support from Guinness in 1901 (West, 2006). Spring barley varieties grown in these trials during the early 20th century included *Spratt, Archer* and *Goldthorpe* (Student, 1923; Reid *et al.*, 1929; West, 2006).

Within barley's germplasm there are genotypes that can tolerate abiotic stresses, such as drought and heat (Ivandic *et al.*, 2000; Xiaojian Wu *et al.*, 2017; Bindereif *et al.*, 2021). Barley landraces can also grow well in biogeographical zones with reduced soil fertility in which modern elite barley varieties fail to reach maturity (Schmidt *et al.*, 2019). As environmental stresses become more frequent and there is a need for varieties demanding fewer resource inputs, heritage varieties, such as those grown by Bennett, may provide valuable genetic variation and a possible resource for these wild-type traits.

A well-documented set of multi-environment spring barley trials data for 1901-1906 exists, comparing two heritage spring barley varieties: *Archer* and *Goldthorpe*. *Archer* is a 2-row narrow-eared variety that originated in East of England and outperformed the long-running favourite *Chevalier* in yield, quality and straw strength (Hunter, 1913). *Goldthorpe* is a 2-row wide-eared barley known for its high malting quality. In 1889 a single wide ear was found in a field of *Chevalier* near Goldthorpe, Yorkshire and was selected and propagated to become *Goldthorpe* (Reid *et al.*, 1929; Gothard *et al.*, 1983; Malcolm, 1983). Analysis of these trials data by William Gosset in Student (1923) concluded that the chief difficulty in comparing variety performance was that differences between varieties are small compared with variations due to weather. Whilst weather was recorded during this period at various locations across Ireland, these data were not accessible to Student at that time.

A recent data rescue project has extended the temporal coverage of digitally available daily maximum and minimum air temperature and rainfall observations back to include this early 20$^{th}$ century period (Mateus *et al.*, 2020; Ryan *et al.*, 2021). In this chapter, these historical weather data are combined with the barley trials data to quantify the effect of interannual weather variability on the early 20$^{th}$ Century spring barley yields and understand the sensitivity of each variety to weather variability. Combining trials and climate data can lead to high dimensional datasets, which can result in overfitting by models and make statistical model output difficult to interpret. Therefore, this research explores the effectiveness of four main variable selection and shrinkage methods: best-subset selection, stepwise selection, the lasso (Tibshirani, 1996) and the elastic net (Zou and Hastie, 2005), as justified in Chapter 2, 2.1.4. The performance of these methods on highly correlated data is evaluated and compared to the use of other techniques including Principal Component Analysis and correlation analysis. This chapter is comprised of work accepted for publication in the *Annals of Applied Biology* (Raymond *et al.*, in press), with some additional material.

## 3.1 Spring barley yields show high variation

Median yields varied from year-to-year by up to 50% for *Archer* and up to 58% for *Goldthorpe* (Figure 3.1). For both varieties the lowest yields occurred in 1903 (combined mean 2.2 t/ha), and highest in 1905 (combined mean 3.2 t/ha). There was large variation in yields within years, in particular for *Archer* in 1903 (standard deviation, SD = 0.59 t/ha) and for *Goldthorpe* in 1901 (SD = 0.67 t/ha). This was despite the number of trials increasing each year.

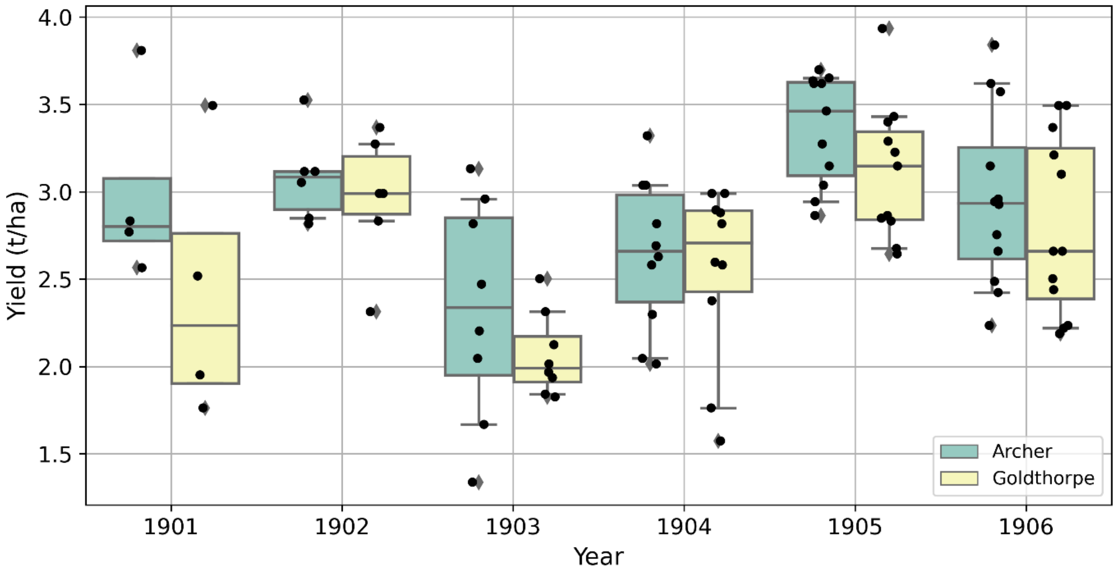



Figure 3.1*: Barley trials yields (t/ha) (black dots) for 51 trials across 18 farms between 1901-1906, for two varieties: Archer (blue) and Goldthorpe (yellow). Outliers (diamonds) represent trial yields in the 5$^{th}$ and 95$^{th}$ percentiles. There were 51 trials per variety, increasing from 4 in 1901 to 12 in 1906. Data from Student (1923).*

Only three farmers were involved in all six years of the trials: Hawkins, McCarthy and Wolfe (Figure 3.2). There were clear differences from farm to farm in yields reflecting the differences in climate, soil type, topography, farm management practices and years. All three farms showed similar interannual variability: 1903 was the lowest yielding year while 1902 and 1905 were the highest. Yields fluctuated by up to 50% with average yields increasing ~45% (1.8 t/ha) between 1903 and 1905, indicating low stability in these varieties.

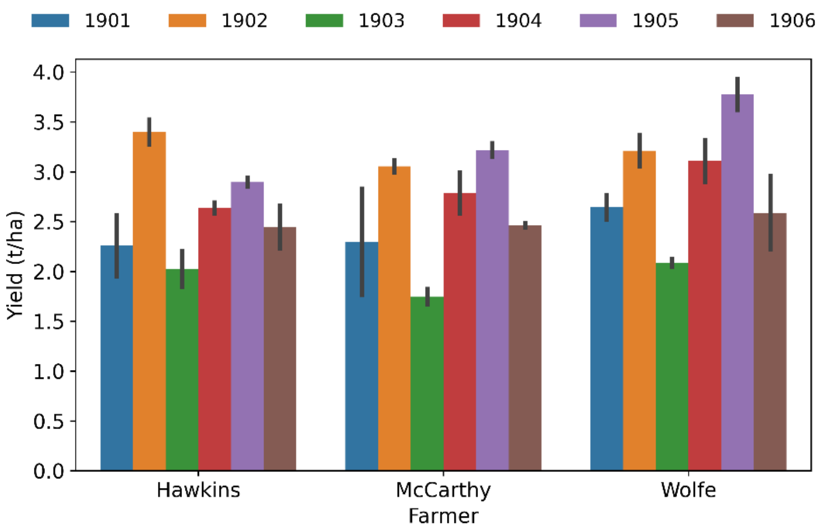



Figure 3.2*: Spring barley trial mean yields (t/ha) at the three farms with data for the entire period 1901-1906. Error bars show the difference between Archer and Goldthorpe variety yields. Data from Student (1923).*

## 3.2 Spring barley price shows similar variation to yield

Student used price as a measure of quality of the crop. The lowest quality of both varieties occurred in 1903 and highest in 1905 (Figure 3.3), as with yield (Figure 3.1). Student (1923) acknowledged that the value of the crop per acre was mostly dependent on the yield.

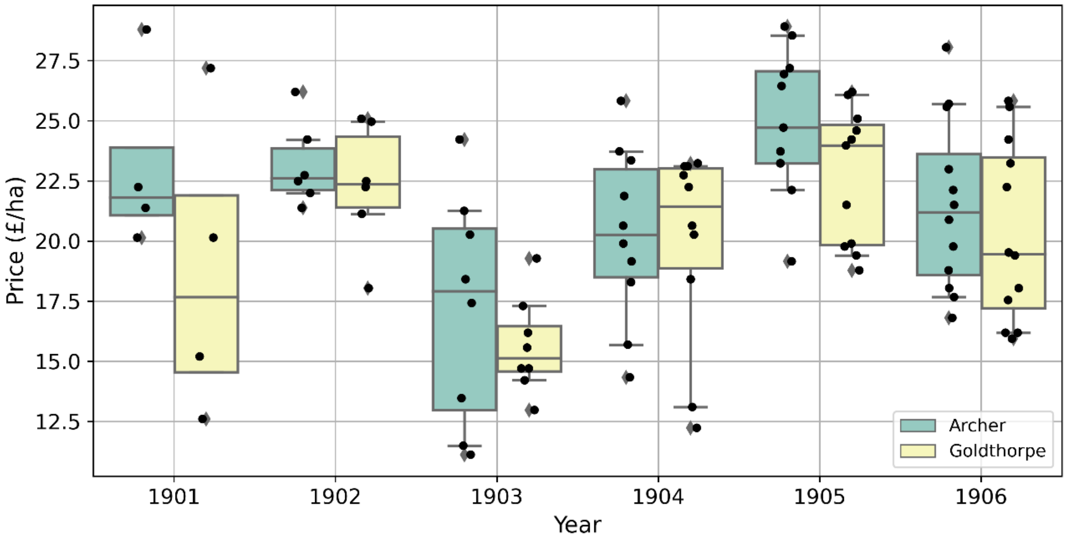



Figure 3.3: *Barley trial price per hectare (£/ha) (black dots) for 1901-1906 for two varieties: Archer (blue) and Goldthorpe (yellow). Outliers (diamonds) represent trial yields in the 5$^{th}$ and 95$^{th}$ percentiles. There were 51 trials per variety, increasing from 4 in 1901 to 12 in 1906. Data from Student (1923).*

## 3.3 Irish climate analysis

### 3.3.1 Long-term climate reveals anomalous years

On average, Ireland received 471 mm of rain during the March to August growing season across the period 1861-2010, but the period saw high variation from just 276 mm of rainfall in 1975, to 657 mm in 1986. Growing season rainfall anomalies show large interannual variability, with differences of up to 300 mm between neighbouring years (Figure 3.4). Averaged across all stations, the lowest yielding year 1903 was the wettest of the six barley trial years 1901-1906, with a large positive anomaly relative to the 1851-2010 average. Nationally the 1903 growing season received over 20% more rainfall than average. 1901 and 1902 were drier than average across the stations. Given the growing season in Ireland consistently receives high rainfall, water deficits are unlikely to be a yield-limiting factor.

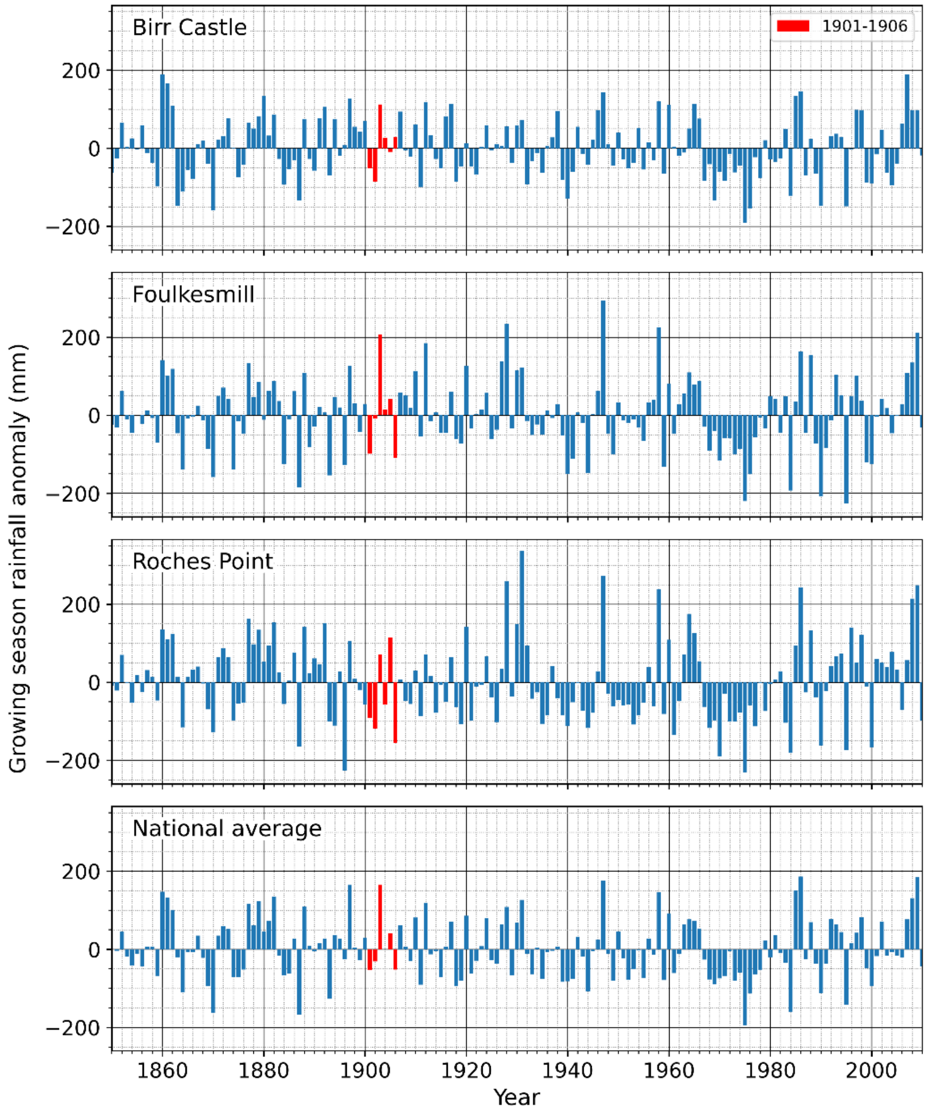



Figure 3.4: *Growing season (March-August) rainfall anomalies (mm) for Birr Castle, Foulkesmill and Roches Point stations and nationally for 1850-2010. Years 1901-1906 are shown in red. Here the national average anomaly is calculated using the Island of Ireland precipitation series from 25 stations.*

Over the 1874-2020 period, significant long-term increases in Growing Degree Days of 0.8°C days $yr^{-1}$ (r = 0.26, p=0.003) and 2.3°C days $yr^{-1}$ (r=0.66, p<0.001) were seen at Birr Castle and Glasnevin respectively (Figure 3.5). The more extreme increase in GDD seen at Dublin is likely due to increased urbanisation and industrialisation in the city (Dublin City Council, 2017), increasing absorption of solar radiation and enhancing the urban heat island effect. In addition to being the wettest of the six years, 1903 growing season had the 11th lowest GDD recorded at both Birr Castle and Glasnevin stations across the period.

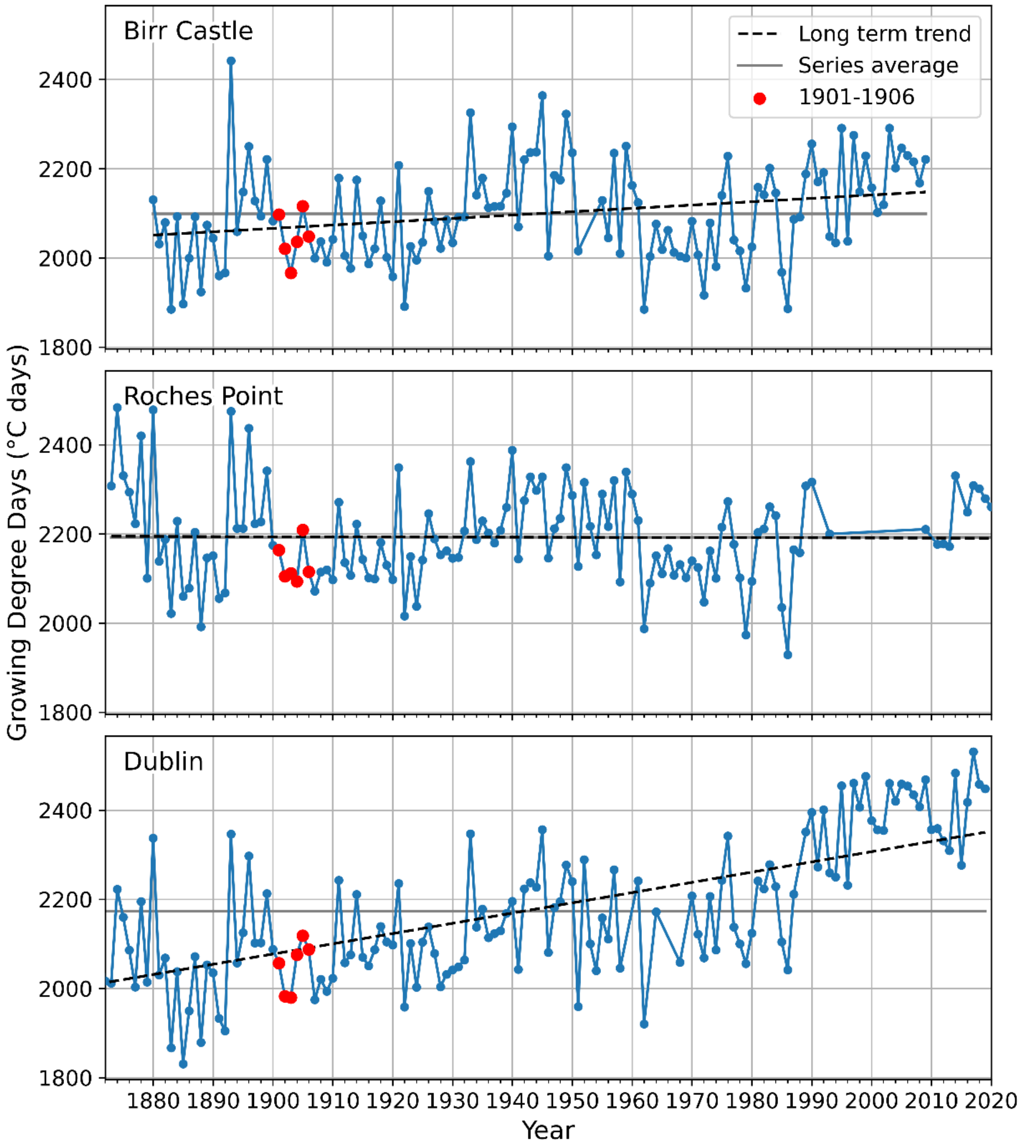



Figure 3.5*: Growing season Growing Degree Days (°C days) for Birr Castle, Roches Point and Glasnevin stations for 1874-2020. Growing Degree Days is the sum of the mean temperature on days when mean temperature is above 5.6°C from March to August.*

### 3.3.2 NAO index shows extremes in the period 1901-1906

Winter, spring and summer NAO indices all show large interannual variability (Figure 3.6). All six years in the period 1901-1906 except 1902 have positive WNAO index values. The WNAO index for 1903 is particularly large at 2.47, ranked in the highest 6% of WNAO values on record (1824-2020). 1903 is the only year in the six year period with a positive SNAO value. 1902 has a large negative SNAO index of -1.56, ranked in the lowest 10% of SNAO values. 1904 has a large positive spring NAO index ranked in the highest 9% of values.

Negative NAO in winter draws in colder, drier easterlies whereas a positive NAO in winter results in above average westerly winds from the mild, moist Atlantic Ocean (Figure 3.7) (Trigo *et al.*, 2002; Correia *et al.*, 2020; Met Office, 2021a). In the summer, lower NAO generally implies more meridional flow: South to North or North to South. When the flow is more South to North, this leads to warmer, moister air masses and heavier individual, slow moving rain events (Trigo *et al.*, 2002). Given 1903 had a large positive NAO in the winter and a large negative NAO in the summer, this may explain why 1903 was much wetter than the other years in the period (Figure 3.6).

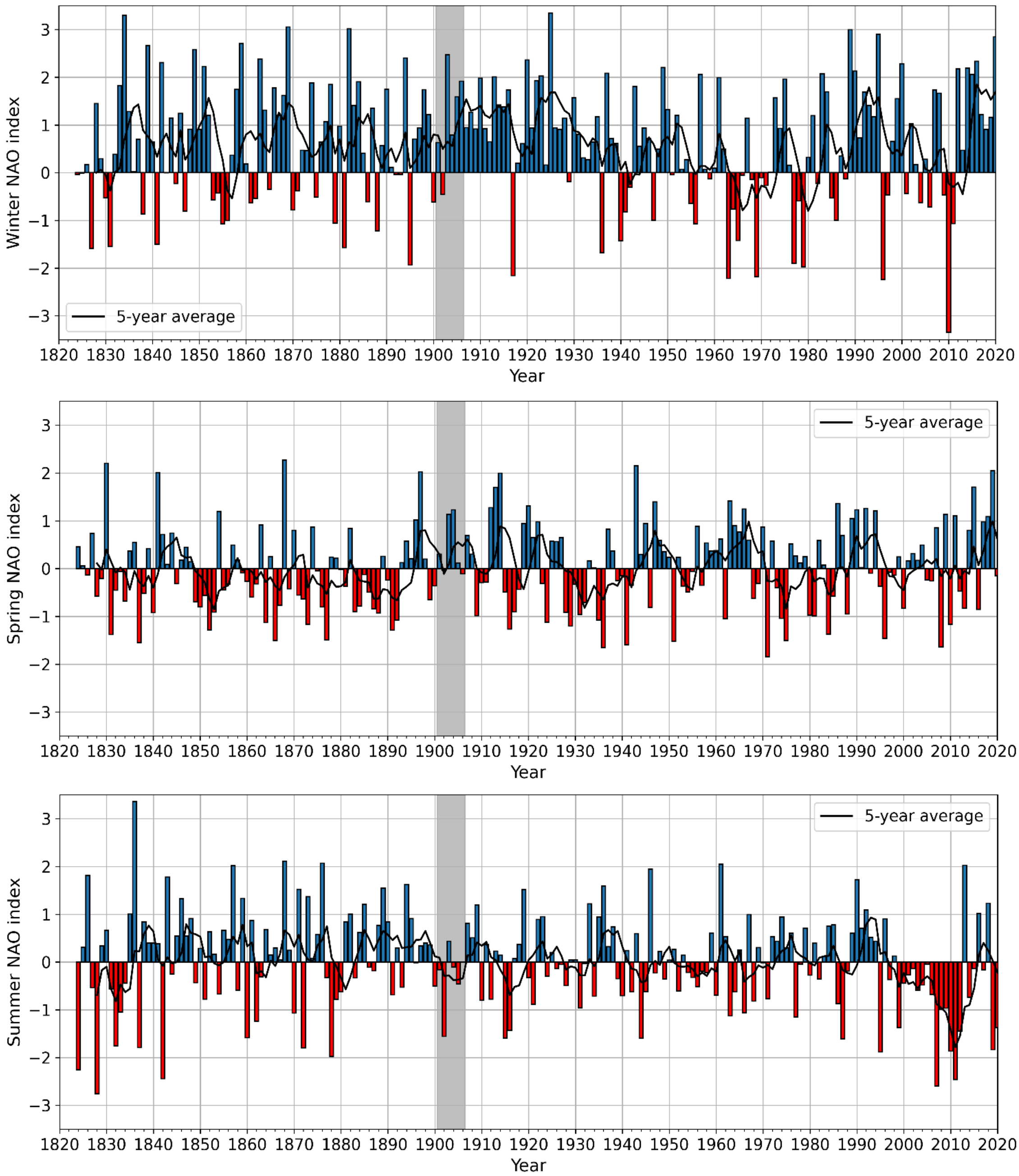



Figure 3.6*: Winter NAO index (top), spring NAO index (middle) and summer NAO index (bottom) for the period 1824-2020. Period 1901-1906 highlighted in grey. Data from Jones et al. (1997).*

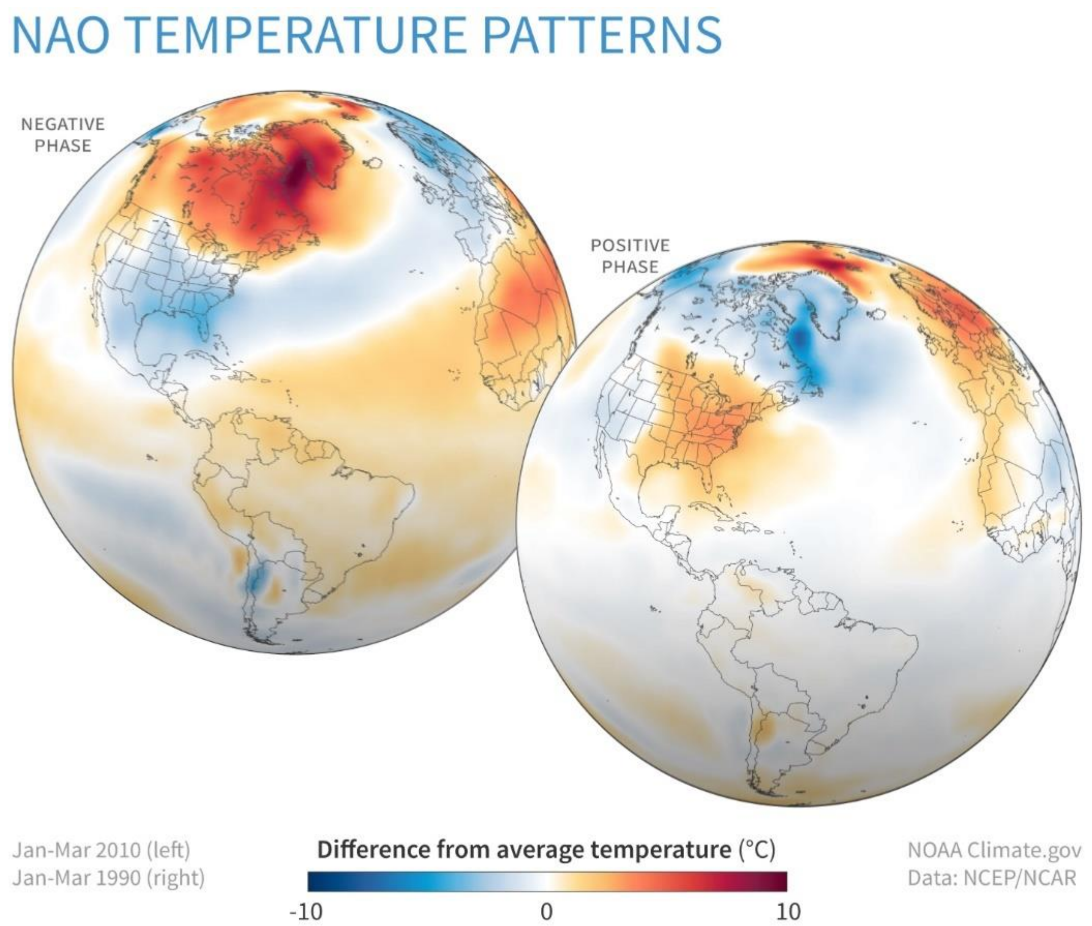



Figure 3.7*: Late winter temperatures compared to the 1981-2010 average when the North Atlantic Oscillation (NAO) was strongly negative (top, Jan-March 2010) and when it was strongly positive (bottom, January-March 1990). When the NAO is negative, winters are often cooler than average across mid-latitudes. When the NAO is positive, they can be warmer than average. Image from NOAA Climate.gov* (Lindsey and Dahlman, 2021).

### 3.3.3 1891-1920 climatology reveals extreme wetness in 1903

Comparing years 1901-1906 to the climate of 1891-1920 places the data in the context of the general climate at the time. The six-year period showed some extreme wetness and temperatures.

March 1903 was the wettest March in the 1891-1920 30-year period for Ardee, Birr and Foulkesmill stations and nationally (Figure 3.8). The coastal stations Roches Point and Greenore saw more 'normal' rainfall amounts, with the former recording its highest March rainfall for the 30-year period in 1905. Cumulatively, 1903 was the wettest growing season in the 30-year period at Foulkesmill and nationally, recording over 600 mm rainfall. It was also the wettest growing season in the period 1901-1906 at Ardee, Birr and Greenore.

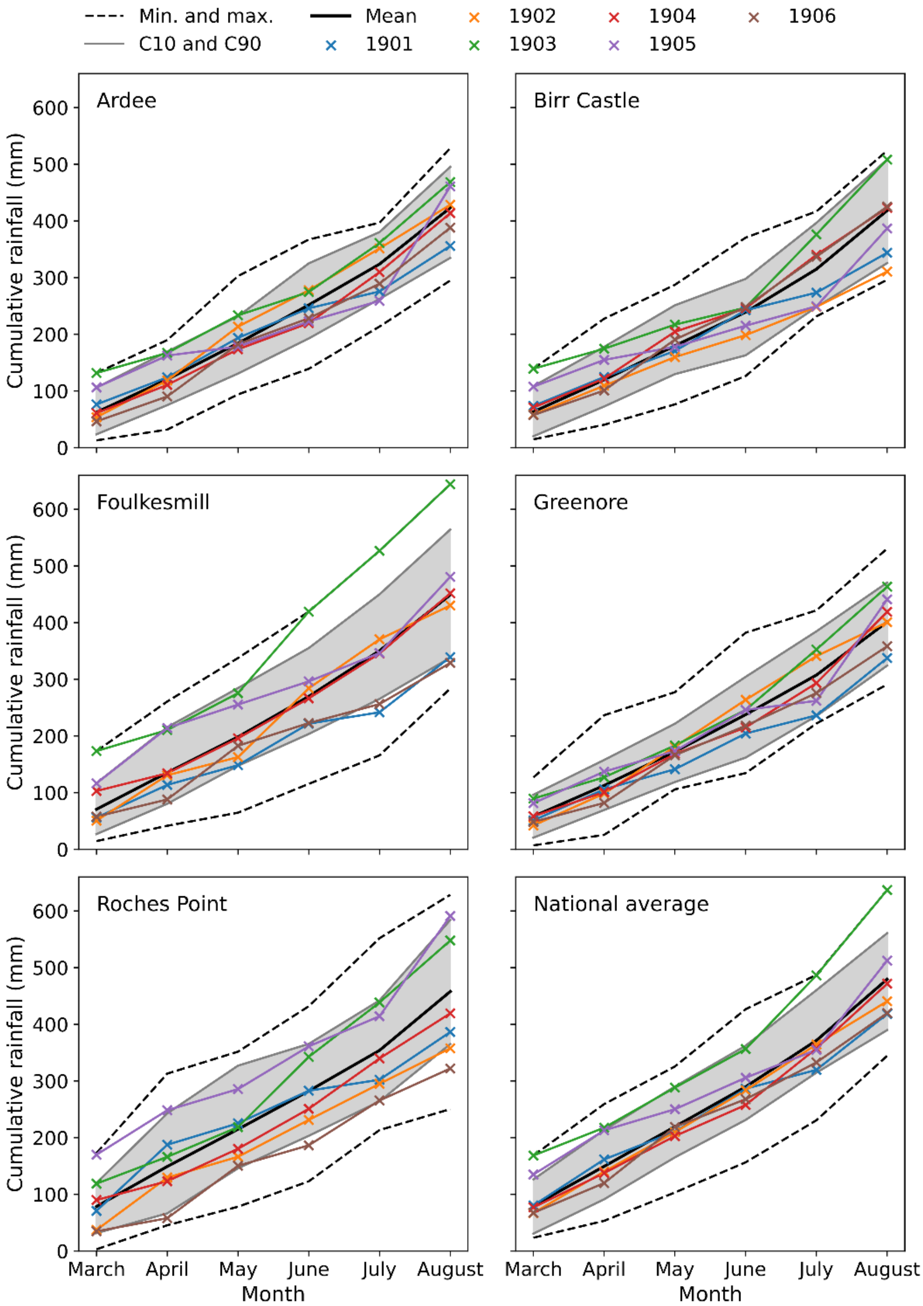



Figure 3.8: *Cumulative monthly rainfall (mm) for Ardee, Birr Castle, Foulkesmill, Greenore, Roches Point stations and the national average across 25 stations for 1901, 1902, 1903, 1904, 1905 and 1906. The 1891-1920 average is shown (solid black line) along with the period 10th and 90th percentile values (grey lines) and the period minimum and maximum values (dashed black line). Note: Ardee station only has data for 1891-1913 therefore the averages are for this period instead. Data from* Ryan *et al.* (2020).

Roches Point's coastal location is evident from the less extreme temperature values, with higher mean minimum temperatures and lower mean maximum temperatures (Figures 3.9 and 3.10). 1906 saw extremely low monthly mean minimum temperatures in April at all three stations, as well as the highest mean minimum temperature for August in the 30-years at

Dublin. The mean minimum temperatures at Roches Point and Birr Castle for this month were closer to the average highlighting that climate extremes vary spatially and can be localised, contributing to the range in observed yields. March 1902 saw relatively high mean minimum and maximum temperatures whilst May 1902 saw much lower-than-average mean maximum temperatures (Figure 3.10).

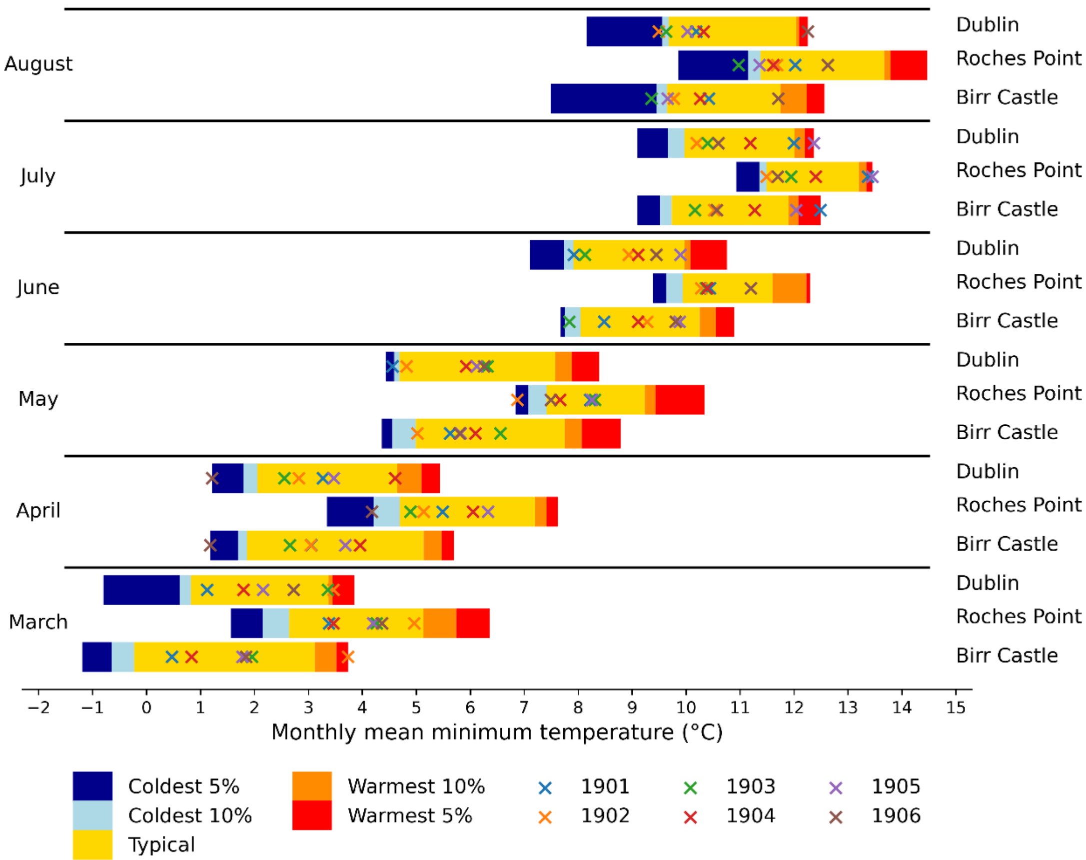



Figure 3.9: *1891-1920 monthly mean minimum temperatures (°C) for Birr Castle, Roches Point and Dublin (Glasnevin) stations for the growing season. The range in temperatures for the coldest 5%, coldest 10%, warmest 10% and warmest 5% mean minimum temperatures are shown. Monthly mean minimum temperatures for 1901, 1902, 1903, 1904, 1905 and 1906 are also presented. Data from* (Mateus *et al.*, 2020).

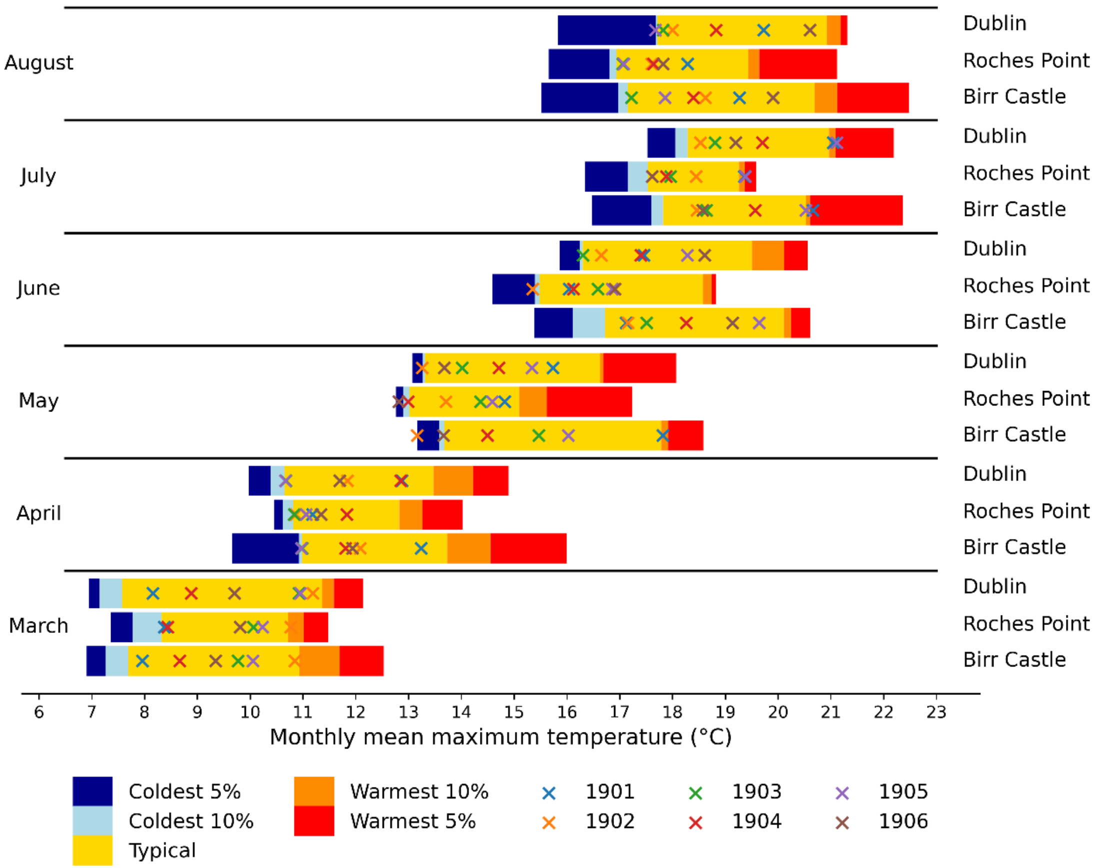



Figure 3.10: *1891-1920 monthly mean maximum temperatures (°C) for Birr Castle, Roches Point and Dublin (Glasnevin) stations for the growing season. The range in temperatures for the coldest 5%, coldest 10%, warmest 10% and warmest 5% mean maximum temperatures are shown. Monthly mean maximum temperatures for 1901, 1902, 1903, 1904, 1905 and 1906 are also presented. Data from* (Mateus *et al.*, 2020).

### 3.3.4 Climate anomalies encompassed parts of Europe

Analysis of growing season rainfall data from ERA-20C for 1901-1906 relative to the 1901-1930 averages shows that 1903 was much wetter than average across Ireland, the UK and much of Europe (Figure 3.11). 1906 was the driest year in the trials period. High rainfall is generally associated with a reduction in solar radiation and the 1903 growing season also received ~5% less photosynthetically active radiation (PAR) than the 1901-1930 average in Ireland (Figure 3.12). 1901 and 1904 show the most significant positive PAR anomalies over the growing season. Breaking this down into months, the four years 1901, 1902, 1904 and 1905 all have positive PAR anomalies in July in sync with the grain fill period. 1905 was the only growing season in the period when Ireland had a positive temperature anomaly, of ~0.3°C. Higher than average temperatures were also experienced across the UK and most of central, eastern and northern Europe (Figure 3.13). 1903 was the coldest growing season in Ireland, about 1°C below the 1901-1930 average.

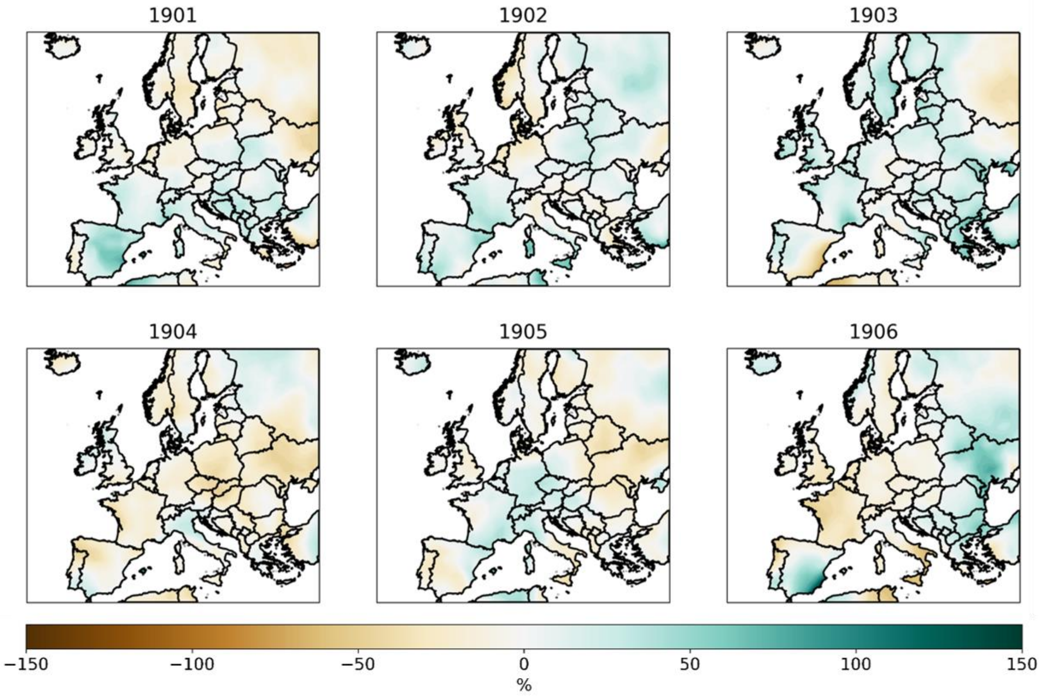



Figure 3.11*: Growing season (March to August) rainfall anomalies (%) relative to the 1901-1930 average. Brown corresponds to drier than average and blue corresponds to wetter than average. 100 km x 100 km resolution data from ERA-20C* (Poli *et al.*, 2016).

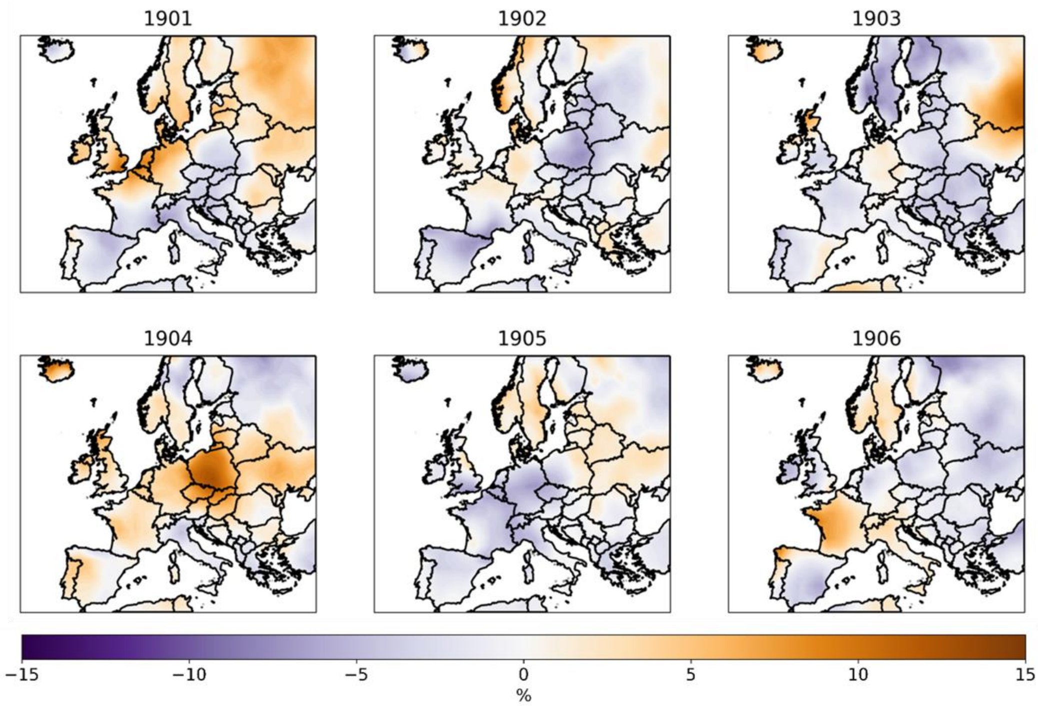



Figure 3.12*: Growing season (March to August) total photosynthetically active radiation (PAR) anomalies (%) relative to 1901-1930 average. Purple corresponds to less PAR than average whilst orange corresponds to more PAR than average. 100km x 100km resolution data from ERA-20C (Poli et al. 2016).*

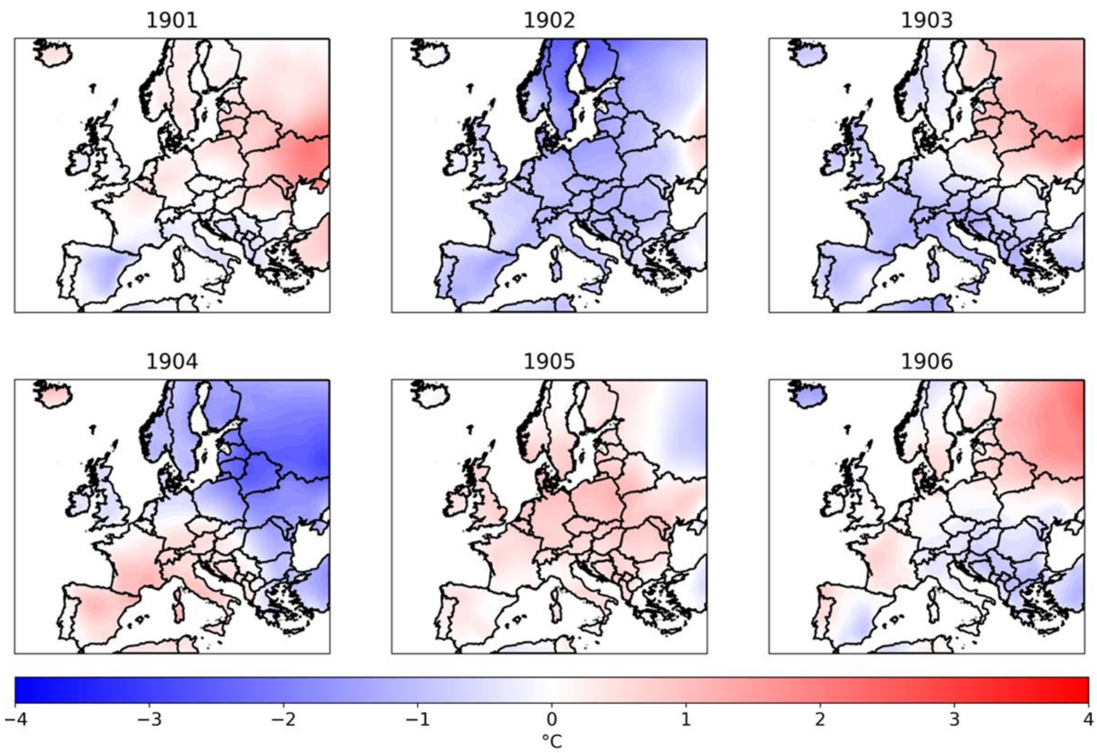



Figure 3.13*: Growing season (March to August) mean temperature anomalies (°C) relative to 1901-1930 average. Blue corresponds to colder than average whilst red corresponds to warmer than average. 100km x 100km resolution data from ERA-20C* (Poli *et al.*, 2016).

## 3.4 Modelling Irish spring barley and climate

A linear model was fit on equation [2.2] using *lm* in R, with the year term included first as a factor and then as a variable. Analysis of variance (ANOVA) between the two model show that the model including year as a factor results in better fit ($R^2$=1 and $R^2$=0.57, respectively) (Table 3.1). Furthermore, the interaction Year x Farm term of interest is not significant when year is a variable. Therefore, from herein year is included as a factor.

| Variable | *Year as a factor* | | *Year as a variable* | | |
|---|---|---|---|---|---|
| | Df | SS | Df | SS | p-value (sig.) |
| Year | 5 | 11.67 | 1 | 1.57 | 0.03 (*) |
| Farm | 17 | 9.13 | 17 | 9.34 | 0.09 |
| Variety | 1 | 1.13 | 1 | 1.13 | 0.07 |
| Year:Farm | 28 | 6.67 | 11 | 4.62 | 0.3 |
| Year:Variety | 5 | 0.36 | 1 | 0.07 | 0.6 |
| Farm:Variety | 17 | 1.03 | 17 | 0.98 | 1.0 |
| Year:Farm:Variety | 28 | 2.26 | 11 | 0.78 | 1.0 |
| Residuals | 0 | 0 | 42 | 13.76 | |

Table 3.1*: Analysis of variance (ANOVA) of Irish barley model [2.2] including year as a factor and a variable, farm and variety and their interactions. Degrees of freedom (Df) and Sum of Squares (SS) are shown for each model term. The p-values and significance are also shown for the model with year as a variable but not for year as a factor as these cannot be calculated when the model has a perfect fit. *significant at the 95% confidence level.*

Rerunning this model without the three-way interaction term and performing an ANOVA shows that the two variety interaction terms are not significant (p>0.05) (Table 3.2). The model has an adjusted $R^2$ of 0.747.

| Variable | Df | SS | F value | p-value (sig.) |
|---|---|---|---|---|
| Year | 5 | 11.67 | 28.89 | 3E-10 (*) |
| Farm | 17 | 9.13 | 6.64 | 6E-06 (*) |
| Variety | 1 | 1.13 | 13.99 | 0.0008 (*) |
| Year:Farm | 28 | 6.67 | 2.95 | 0.003 (*) |
| Year:Variety | 5 | 0.36 | 0.90 | 0.5 |
| Farm:Variety | 17 | 1.03 | 0.75 | 0.7 |
| Residuals | 28 | 2.26 | | |

Table 3.2: *ANOVA results for lm model including year (as a factor), farm and variety, and their two-way interactions. Degrees of freedom (Df), Sum of Squares (SS), F-value and p-value are shown for each model term. *significant at the 95% confidence level.*

Therefore, the base model is given by:

$$y_{ijk} = \mu + v_i + r_j + s_{jk} + e_{ijk} \quad [3.1]$$

where $y_{ijk}$ is the yield of variety $i$ in year $j$ at site $k$, $\mu$ is the overall trial series mean, $v_i$ is the effect of variety $i$, $r_j$ is the effect of year $j$, $s_{jk}$ is the effect of site within years and $e_{ijk}$ is the error term. This model had an adjusted $R^2$ of 0.771, which is higher than the previous model that includes the other interactions indicating a better model fit.

### 3.4.1 Variable selection methods failed to significantly reduce model complexity

The best subset selection method (cv.b), three forwards and backwards selection methods (s.bo, s.ba and s.f) and the elastic net method (e.n.) failed to sufficiently simplify the climate model to then include the selected climate covariates in the mixed model with year, variety and site [2.3] (Table 3.3). A combination of using too many highly correlated variables (Figure A1) and too few farm growing seasons likely contributed to this. The worst performing was backwards stepwise selection which did not drop any variables. The two lasso methods reduced the model complexity significantly from 19 to less than 7 climate variables, but these models had very low adjusted $R^2$ values of close to 0, indicating a poor model fit (Table 3.4).

Specifically, Table 3.3 shows:

- The magnitude of selected variable coefficients and the corresponding significance of the variable (where available)
- July maximum temperature (jul_temp_max) was included by all model selection methods
- Variables included in most of the models are April maximum rainfall (apr_rain_dmax), April total rainfall (apr_rain_tot), June total rainfall (jun_rain_tot), July total rainfall (jul_rain_tot) and May minimum temperature (may_temp_min.)
- August total rainfall (aug_rain_tot) was the most frequently dropped variable, followed by March total rainfall (mar_rain_tot) and July total rainfall (jul_rain_tot)

Using the mixed-model backwards elimination approach and *glmmLasso* algorithm, all climate variables were dropped. A Bayesian linear model was also run on the climate covariates, to select the significant variables (Table 3.5). The eight significant climate variables in the linear model were identified by the 0.89% Credible Intervals (CI) not straddling 0. The 89% is the default CI level as it is deemed to be more stable than higher intervals, such as 95%. Although not identical, the significant variables overlap with those from the frequentist variable selection models.

The linear models with the best performance (Table 3.3) were then incorporated into the mixed-effects model [2.3], but it was found that there were too many climate covariates to include resulting in issues with model convergence. A range of modifications were tried, including using joint regression analysis to represent the variety × climate interaction terms $vT_{ijk}$ and $vP_{ijk}$, and dropping these terms entirely. This did not solve the problem.

| Variable | Coefficients | | | | | | | | p-value | | | | sig. (p < 0.05) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | c.m | cv.b | s.bo | s.ba | s.f | l.m | l.1 | e.n | c.m | s.bo | s.ba | s.f | c.m | s.bo | s.ba | s.f |
| (Intercept) | 2.798 | 2.798 | 2.798 | 2.798 | 2.798 | 2.815 | 2.799 | 2.822 | | | | | | | | |
| apr_rain_dmax | 0.423 | 0.503 | 0.348 | 0.423 | 0.458 | 0.073 | 0 | 0.132 | 0.0002 | 0.3096 | 0.0002 | 0.0021 | * | | * | * |
| apr_rain_tot | -0.241 | -0.356 | -0.212 | -0.241 | -0.345 | 0 | 0 | -0.023 | 0.0423 | 0.0403 | 0.0423 | 0.0432 | * | * | * | * |
| apr_temp_max | -0.380 | -0.373 | -0.190 | -0.380 | -0.298 | 0 | 0 | 0 | 0.0131 | 0.0921 | 0.0131 | 0.1208 | * | | * | |
| apr_temp_min | 1.450 | 0.852 | 1.048 | 1.450 | 0.771 | 0 | 0 | 0.051 | 0.4955 | 0.0683 | 0.4955 | 0.1790 | | | | |
| aug_rain_tot | 0.174 | | | 0.174 | | 0 | 0 | -0.012 | 0.2541 | | 0.2541 | | | | | |
| aug_temp_max | 0.392 | 0.722 | | 0.392 | 0.603 | 0 | 0 | 0.082 | 0.0068 | | 0.0068 | 0.9660 | * | | * | |
| aug_temp_min | 0.732 | | 0.762 | 0.732 | 0.125 | 0 | 0 | 0.096 | 0.0323 | 0.5125 | 0.0323 | 0.2499 | * | | * | |
| jul_rain_tot | -0.169 | | | -0.169 | -0.017 | -0.154 | -0.095 | -0.152 | 0.0000 | | 0.0000 | 0.0000 | * | | * | * |
| jul_temp_max | -1.086 | -0.430 | -0.405 | -1.086 | -0.342 | 0.103 | 0.052 | 0.159 | 0.9697 | 0.0000 | 0.9697 | 0.0005 | | * | | * |
| jul_temp_min | 1.849 | 1.123 | 1.501 | 1.849 | 1.003 | 0 | 0 | -0.100 | 0.2139 | 0.6334 | 0.2139 | 0.0127 | | | | * |
| jun_rain_tot | -0.064 | -0.266 | | -0.064 | -0.225 | -0.108 | -0.010 | -0.104 | 0.0142 | | 0.0142 | 0.0145 | * | | * | * |
| jun_temp_max | 1.293 | 0.430 | 1.269 | 1.293 | 0.435 | 0 | 0 | 0.068 | 0.0134 | 0.1631 | 0.0134 | 0.2878 | * | | * | |
| jun_temp_min | -2.219 | -0.747 | -1.949 | -2.219 | -0.774 | 0 | 0 | 0.072 | 0.3766 | 0.3315 | 0.3766 | 0.3515 | | | | |
| mar_rain_tot | -0.234 | | -0.176 | -0.234 | | 0 | 0 | 0 | 0.1118 | 0.0193 | 0.1118 | | | * | | |
| mar_temp_max | 0.523 | 0.378 | 0.238 | 0.523 | 0.355 | 0.006 | 0 | 0 | 0.1178 | 0.1307 | 0.1178 | 0.3311 | | | | |
| mar_temp_min | 0.466 | | 0.794 | 0.466 | -0.038 | 0 | 0 | -0.017 | 0.7155 | 0.0333 | 0.7155 | 0.7727 | | * | | |
| may_rain_tot | 0.426 | 0.134 | 0.279 | 0.426 | | 0 | 0 | -0.102 | 0.8492 | 0.0179 | 0.8492 | | | * | | |
| may_temp_max | 0.173 | | | 0.173 | -0.099 | 0 | 0 | -0.079 | 0.4364 | | 0.4364 | 0.0210 | | | | * |
| may_temp_min | -0.969 | -0.524 | -0.845 | -0.969 | -0.447 | -0.035 | 0 | -0.087 | 0.1127 | 0.0000 | 0.1127 | 0.0787 | | * | | |

Table 3.3*: Estimated coefficients and their p-values and significance for 7 different frequentist variable selection methods (Table 2.6), as well as the full climate model (*c.m.*).* cv.b = *best subset selection with cross-validation (Method 1),* s.bo = *cross stepwise selection in both backwards and forwards directions (Method 2a),* s.ba = *10-fold cross-validation backwards stepwise selection (Method 2b),* s.f = *10-fold cross-validation forwards stepwise selection (Method 2c),* l.m = *cross-validation lasso using lambda that minimises the prediction error (Method 3a),* l.1 = *cross-validation lasso model using lambda for smallest model and within 1 standard error (Method 3b), e.n = cross-validation elastic net using lambda and alpha that minimises the prediction error (Method 4). Of the 7 models shown, only 4 give p-values and significance of each variable.*

| Model | Adjusted $R^2$ | RMSE |
|---|---|---|
| c.m | 0.449 | 0.376 |
| **cv.b** | **0.466** | **0.383** |
| **s.bo** | **0.459** | **0.384** |
| s.ba | 0.449 | 0.451 |
| s.f | 0.443 | 0.446 |
| l.m | -0.065 | 0.477 |
| l.1 | 0.024 | 0.496 |
| e.n | -1.08 | 0.428 |

Table 3.4*: Adjusted $R^2$ and RMSE for the 7 different variable selection methods (Table 2.6), as well as the full climate model (*c.m.*).* cv.b *= best subset selection with cross-validation (Method 1),* s.bo *= cross stepwise selection in both backwards and forwards directions (Method 2a),* s.ba *= 10-fold cross-validation backwards stepwise selection (Method 2b),* s.f *= 10-fold cross-validation forwards stepwise selection (Method 2c),* l.m *= cross-validation lasso using lambda that minimises the prediction error (Method 3a),* l.1 *= cross-validation lasso model using lambda for smallest model and within 1 standard error (Method 3b), e.n. = cross-validation elastic net using lambda and alpha that minimises the prediction error (Method 4).*

| Parameter | Median | 89% CI | pd |
|---|---|---|---|
| Intercept | 2.798 | **[ 2.735, 2.872]** | 100.00% |
| apr_rain_tot | -0.3 | **[-0.475, -0.118]** | 99.55% |
| aug_rain_tot | 0.011 | [-0.193, 0.252] | 52.78% |
| jul_rain_tot | -0.06 | [-0.274, 0.157] | 68.92% |
| jun_rain_tot | -0.176 | [-0.390, 0.020] | 91.47% |
| mar_rain_tot | -0.067 | [-0.278, 0.148] | 68.90% |
| may_rain_tot | 0.162 | [-0.119, 0.456] | 81.30% |
| apr_rain_dmax | 0.44 | **[ 0.196, 0.690]** | 99.75% |
| apr_temp_max | -0.264 | **[-0.530, -0.011]** | 94.67% |
| aug_temp_max | 0.421 | [-0.102, 0.966] | 88.55% |
| jul_temp_max | -0.37 | [-1.090, 0.285] | 79.85% |
| jun_temp_max | 0.63 | **[ 0.000, 1.230]** | 95.03% |
| mar_temp_max | 0.331 | [-0.105, 0.735] | 89.25% |
| may_temp_max | -0.041 | [-0.313, 0.254] | 58.43% |
| apr_temp_min | 0.839 | **[ 0.263, 1.395]** | 98.50% |
| aug_temp_min | 0.215 | [-0.357, 0.810] | 71.83% |
| jul_temp_min | 1.099 | **[ 0.403, 1.742]** | 99.33% |
| jun_temp_min | -1.045 | **[-2.034, -0.071]** | 95.30% |
| mar_temp_min | 0.175 | [-0.466, 0.810] | 66.47% |
| may_temp_min | -0.504 | **[-0.977, -0.030]** | 95.55% |

Table 3.5*: Estimated model parameters from the stan_glm model. Median is the median value computed from the model simulations. CI represents the Credible Interval, which quantifies the uncertainty about the regression coefficients. The 89% CI computes the Credible Interval with 89% probability that a coefficient lies above the lower value of the two values and below the higher value. If the CI doesn't straddle 0 then the coefficient is significant and is in bold. pd is the Probability of Direction, which is the probability the effect goes to the positive or negative direction* (Bloggers, 2020).

A similar analysis was also repeated using the monthly extreme maximum and minimum temperature values (Table A1), rather than monthly mean maximum and mean minimum temperature. Results from the lasso and elastic net methods were not shown due to their poor performance in the monthly mean model.

The variable selection methods reduced model complexity more successfully when using the maximum and minimum temperature values for each month. For example, for best subset selection with cross-validation (cv.b), only 10 variables were selected (Table A1), as opposed to 13 for the mean temperature model (Table 3.3), for stepwise selection (s.bo) 9 variables were selected (Table A1), compared to 14 previously (Table 3.3) and for backwards stepwise selection (s.ba) 13 variables were selected (Table A1) as opposed to all of them with the mean minimum and maximum temperature data (Table 3.3). Generally, there was lower correlation between monthly extreme temperature values (Figure A2) than monthly mean temperature values (Figure A1). A high degree of collinearity would make it more difficult to drop individual climate variables.

Use of the *step* function from the ***lmerTest*** package was also much more successful here and resulted in the selection of climate variables as well as Variety for the fixed effects (Table A2). The observed vs fitted plot for this model showed good agreement (Figure A3), as in all subsequent models in this chapter. Significant variables using this method overlapped well with the *lmer* model when it included the variables selected through stepwise selection (s.bo) and the Bayesian mixed model using the extreme temperature variables.

### 3.4.2 Insufficient independent data for Principal Component Analysis

Using PCA on the monthly mean temperature and total rainfall data found that the first six principal components (PC) explained ~90% of the observed variation in yield (Table 3.6). These were then input into equation [2.3], replacing the climate variables. PCs 2, 3, 4 and 5 were significant. However, the 4 PCs were not clearly defined by just one or two climate variables, rather several. Using PCA didn't therefore simplify the model.

| | | Principal Component | | | | | |
|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 |
| **Proportion of variance** | | **0.26** | **0.22** | **0.18** | **0.10** | **0.091** | **0.043** |
| Climate variable | Apr_rain_tot | 0.24 | 0 | -0.03 | **0.44** | -0.25 | 0.08 |
| | Aug_rain_tot | 0.19 | **0.34** | 0.1 | -0.09 | 0.26 | 0.16 |
| | Jul_rain_tot | -0.11 | -0.16 | 0.3 | **-0.4** | -0.04 | -0.09 |
| | Jun_rain_tot | 0.1 | -0.15 | **0.32** | -0.03 | -0.24 | **0.53** |
| | Mar_rain_tot | 0.22 | 0.23 | 0.24 | -0.29 | 0.07 | **0.33** |
| | May_rain_tot | **-0.35** | -0.18 | -0.09 | -0.17 | 0.12 | 0.25 |
| | Apr_rain_dmax | 0.03 | -0.24 | 0.19 | **0.47** | -0.24 | 0.23 |
| | Apr_temp_max | -0.2 | 0 | -0.26 | -0.11 | **-0.46** | -0.21 |
| | Aug_temp_max | **-0.38** | -0.02 | -0.2 | 0.16 | 0.06 | 0.3 |
| | Jul_temp_max | 0.03 | **0.43** | -0.14 | 0.2 | -0.01 | 0.12 |
| | Jun_temp_max | -0.2 | 0.29 | -0.14 | 0.09 | **0.35** | 0.05 |
| | Mar_temp_max | 0.17 | -0.02 | 0.28 | **0.31** | **0.38** | -0.26 |
| | May_temp_max | 0.12 | **0.38** | -0.11 | -0.03 | -0.24 | 0.12 |
| | Apr_temp_min | **0.35** | -0.06 | -0.17 | -0.22 | -0.21 | -0.23 |
| | Aug_temp_min | -0.03 | -0.27 | **-0.4** | 0.05 | 0.19 | 0.24 |
| | Jul_temp_min | **0.35** | 0.07 | **-0.31** | -0.02 | -0.08 | 0.11 |
| | Jun_temp_min | 0.24 | -0.17 | **-0.35** | 0.03 | 0.22 | -0.03 |
| | Mar_temp_min | 0.24 | **-0.36** | 0.05 | 0.07 | 0.21 | -0.07 |
| | May_temp_min | 0.28 | -0.21 | -0.21 | -0.25 | 0.12 | 0.3 |

Table 3.6*: The 6 main principal components (PCs) contributing to variance in spring barley yields. The proportion of variance of each PC is shown, along with the degree of correlation of each mean climate variable with each PC. Variables correlating by more than |0.3| are in bold.*

Overall, the conclusion from PCA and variable selection models is that there wasn't enough data to build a complex multiple linear regression model with highly correlated data. Therefore, two final methods were explored: yield-climate correlation analysis, and manual climate variable selection by iteratively adding each climate variable individually to equation [2.3].

### 3.4.3 Highest correlations between July temperature and rainfall and yield

In yield-climate correlation analysis July rainfall and July maximum temperature had the largest absolute correlation with yield: -0.49 and 0.45, respectively (Figure 3.14). These variables had a strong negative correlation (Figure A1).

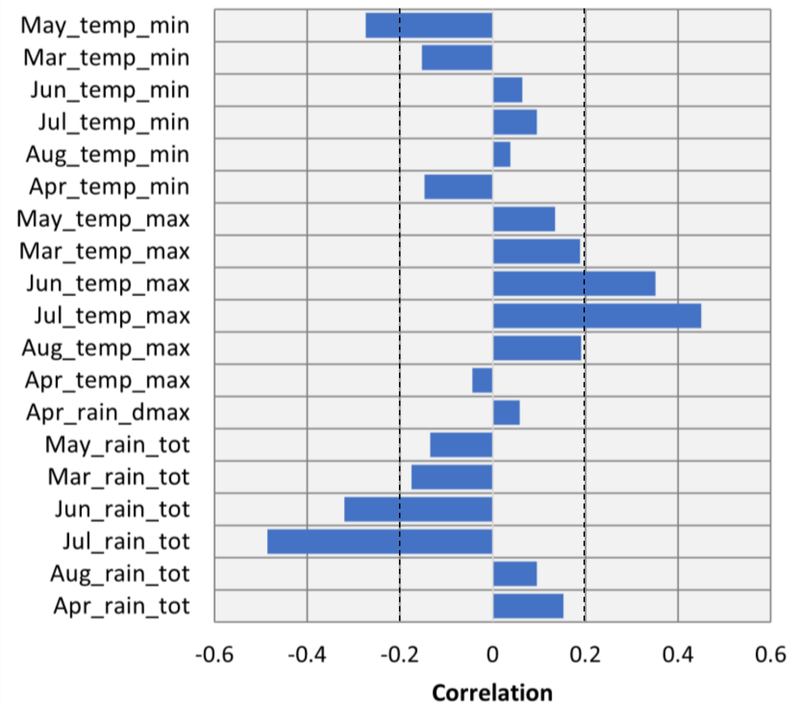



Figure 3.14*: Correlation between each monthly climate variable and yield for the 1901-1906 barley trials in Ireland. The dashed black lines show significant correlation (p<0.05, n=102).*

### 3.4.4 Use of Akaike Information Criterion reveals important climate variables

To understand if adding temperature or rainfall climate variables to the mixed model [2.3] improved the fit, first the AIC of the mixed model of [3.1] without any climate covariates was calculated (Table 3.7). Each climate variable along with its interaction with variety were added to the model one at a time. None of the interactions with variety were significant, so the variety x climate interaction term was dropped from the model and the models with each climate variable were looped through again.

Only three variables – July maximum temperature, August maximum temperature and July total rainfall – were significant when included in the model. The models which included either July maximum temperature or August maximum temperature improved the AIC and model fit. Notably all the models that contained temperature had a lower AIC and better fit than any of the rainfall models, including the significant July rainfall model (Table 3.7).

Both July maximum temperature and August maximum temperature had a positive relationship with yield (Table 3.7), such that yield increased by ~ 1/4 t/ha per 1$^{o}$C increase in July maximum temperature and by ~ 1/5 t/ha per 1$^{o}$C increase in August maximum temperature.

| Climate variable | Significance in model | Coefficient | AIC |
|---|---|---|---|
| - | - | - | 127.5 |
| Jul_temp_max | **0.004** | 0.27 | 123.6 |
| Aug_temp_max | **0.024** | 0.20 | 127.4 |
| Jul_rain_tot | **0.028** | -0.0069 | 135.0 |
| Apr_temp_min | 0.059 | -0.097 | 130.0 |
| Jun_temp_min | 0.090 | -0.11 | 130.4 |
| May_temp_min | 0.101 | -0.10 | 130.5 |
| Mar_temp_min | 0.110 | -0.095 | 130.7 |
| Jun_temp_max | 0.120 | 0.11 | 130.6 |
| Jun_rain_tot | 0.131 | -0.0037 | 137.4 |
| Mar_rain_tot | 0.190 | -0.0035 | 137.7 |
| May_temp_max | 0.212 | 0.098 | 131.2 |
| Jul_temp_min | 0.340 | -0.075 | 131.9 |
| May_rain_tot | 0.534 | 0.0034 | 137.8 |
| Aug_temp_min | 0.540 | -0.044 | 132.5 |
| Apr_rain_tot | 0.573 | 0.058 | 132.0 |
| Mar_temp_max | 0.591 | -0.0020 | 138.6 |
| Aug_rain_tot | 0.635 | 0.062 | 131.5 |
| Apr_rain_dmax | 0.693 | -0.0013 | 139.0 |
| Apr_temp_max | 0.842 | -0.026 | 131.7 |

Table 3.7*: Statistical significance and corresponding coefficient of each climate variable in the mixed model with year, variety and year:farm. Significant variables are shown in **bold**. The AIC of the overall model is given, with lower values corresponding to a better model fit.*

## 3.5 Comparison of standard error of difference between means

Student (1923) calculated the standard error of the mean difference in variety means to test whether there was a significant difference in varietal performance. To test if we get the same result using this method, first the mean difference in the variety values was calculated as £1.52/ha (12 shillings/acre) with a standard deviation of £2.95/ha (23.9 shillings/acre) and corresponding standard error of the mean difference £0.41/ha (3.3 shillings/acre), in accordance with Student (1923). This corresponds to a t-statistic of 3.68, which was statistically significant (p<0.001) at the 95% level (Df = 50). This provided strong evidence that there was a difference in varietal performance.

To test whether the addition of climate variables reduces the yield difference associated with variety, the standard error of difference between variety values in the three mixed models containing significant climate effects (Table 3.7) was calculated. They all gave identical values (to 2 s.f.) of £0.41/ha (3.3 shillings/acre). This indicates that the models do not reduce the standard error, which is expected given no variety x climate interactions were included in the final models. The climate variables simply partitioned the effects of Farm and Year and did not affect the Variety effect.

## 3.6 Discussion

### 3.6.1 Climatic causes of yield variability

Use of recently digitised weather data for the early 20$^{th}$ century has showed that contrasting climatic conditions in 1903, and 1905 coincided with variation in spring barley varietal performance. In 1903 a wet March (Figure 3.8) likely made it challenging to drill the crop, resulting in delayed planting shortening the growing season and compounded by potential difficulties in crop establishment. Nationally, the 1903 growing season received over 20% more rainfall than average (Figure 3.4), contributing to greater cloud coverage and lower than average growing season PAR (Figure 3.12), notably during April, May, June and July. Reduced solar radiation interception during the grain fill period in June and July constrains photosynthesis, reducing the final ear weight amassed in this period (TEAGASC, 2017). The 1903 growing season was also cooler than average (Figure 3.9) with low GDDs (Figure 3.5). This coincides with the year of lowest mean yields and greatest yield variability for *Archer*, but much lower variability for *Goldthorpe* (SD = 0.22 t/ha) (Figure 3.1).

To better understand the weather in this period, NAO indices were analysed. A large positive WNAO and large negative SNAO likely contributed to the higher precipitation received in 1903 (Figure 3.6).

A more recent experiment detailed by Gothard *et al.* (1983) found that *Goldthorpe* outperformed *Archer* when spring and summer rainfall was high. Combined, these results suggest *Goldthorpe* may be able to withstand much higher soil moisture and waterlogging. Hunter (1929) notes that *Goldthorpe* requires plenty of moisture to produce the best yields and quality, supporting this theory. Continual dampness can also increase disease pressures for diseases such as *Rhynchosporium* which prefer cool wet weather and which, if present early in the season, can reduce tiller survival and potential yields (TEAGASC, 2017). If *Rhynchosporium* was present, this may show greater resistance of *Goldthorpe* to the disease.

In contrast, the 1905 growing season was warmer than average (Figure 3.13) with high growing season GDDs (Figure 3.5). There was low growing season PAR (Figure 3.12), but high PAR in July, when high solar radiation is important for grain fill. The growing season was drier than average, starting wet but drying in June and July. These favourable conditions likely contributed to the relatively high yields seen in 1905 for both varieties.

Of the farms with six years of trials data, Farmer Wolfe performed the best on average (Figure 3.2). This farm was located ~30km south-west of Birr Castle and experienced higher summer temperatures and less summer rainfall than the other two farms. Other factors such as favourable agronomy, farm management and soil type may also have encouraged higher yields here.

### 3.6.2 Statistical methods

Through trialling various variable selection methods on both mean temperature and rainfall data and extreme temperature data, this research has highlighted the importance of identifying collinearity early on in analysis involving multiple covariates. The use of these methods and Principal Component Analysis was limited by the high correlation between covariates within a small dataset, but it was still possible to extract information on the most important variables using simple mixed models.

July maximum temperature and August maximum temperature had a positive relationship with yield and July total rainfall had a negative relationship with yield (Table 3.7). July rainfall can also be used as a proxy for solar radiation, so a wet July would usually be associated with more cloud cover, reducing solar radiation interception during grain fill. Likewise, wet weather during grain filling can encourage ear and grain diseases, such as fusarium ear blight and ergot, which can cause shrivelled grain and mycotoxins (AHDB Cereals & Oilseeds, 2018b). Hence the plant benefits from more solar radiation and less rainfall in July. Higher July maximum temperature implies less daytime cloud cover intercepting solar radiation, hence the correlation between these two July variables and yield is of opposite polarity. In future analysis of more recent crop yield data, inclusion of solar radiation data in the models would be desirable to directly quantify the relationship between solar radiation and yield.

The finding that July maximum temperatures were positively correlated with spring barley yield (Figure 3.14) contrasts with other published research which shows that warmer temperatures during anthesis and grain fill can have a detrimental effect (Hakala *et al.*, 2020; Addy *et al.*, 2021a). This result is highly likely due to July maximum temperatures in Ireland in the early 20$^{th}$ century falling well short of those more regularly seen today in some major UK spring barley growing areas. Specifically, maximum temperature did not exceed 28$^{o}$C during the six-year trials period whereas those in South-East England now regularly exceed 30$^{o}$C in summer months (Kendon *et al.*, 2022). This finding shows the importance of region-specific crop-climate research: despite the proximity of the UK to Ireland their climates differ and the same relationships between weather variables and yield cannot be assumed.

It wasn't possible to detect any GxE within the mixed models used (Table 3.7). The lack of significance throughout of climate variety interactions may well be related to the relatively small trials dataset, approximation of site locations and sometimes large distances to weather stations. However, it is clear from the more stable performance of *Goldthorpe* in 1903 relative to *Archer* coupled with wider evidence (Reid *et al.*, 1929; Gothard *et al.*, 1983) that GxE is a driver of performance here. This highlights the importance of considering the local climate in crop variety selection.

The last few years have seen a surge in the growing of heritage barley varieties from the early 20$^{th}$ century. *Goldthorpe*, its predecessor *Chevalier* and offspring *Irish Goldthorpe*, as well as hybrids of *Archer*, such as *Plumage Archer*, have been grown for breweries across the UK and Ireland and are currently being investigated by organisations such as New Heritage Barley. Some heritage varieties display highly desirable traits, such as Fusarium fungal disease resistance in *Chevalier* (BBSRC UKRI, 2016). How these varieties perform in the current and future climate is of interest given the performance of these varieties in the 1901-1906 trials. It is hoped that *Archer* and *Goldthorpe* will be trialled on large scale field plots to allow for comparisons with the yields from 1901-1906, but also to test the models in the current climate on larger datasets.

### 3.6.3 Limitations

Use of any data sources comes with a degree of error and uncertainty. A lot is unknown about the spring barley growing season in the early 1900s, such as planting and harvest dates and relative disease pressures. The exact location of the sites was also unknown. Differences in soils and agronomic practices across trial sites were accounted for by the inclusion of the random site term, representing the farmer x year interaction. However, this research does not account for sources of error in the climate data. There are two main issues here. The first is that data has been recorded manually, by many different people using varying equipment, which introduces the possibility of human error. The second is potential transcribing errors, however to minimise inconsistencies each record was input twice (Murphy *et al.*, 2018; Mateus *et al.*, 2020).

The distance of trial sites to the nearest temperature station was over 100 km for some sites. Whilst it was possible to compare station data with the ERA-20C reanalysis data to ensure station data was reasonably representative of the weather at each site, ERA-20C has very low spatial resolution itself (~125 km) therefore it is still possible that climate data assigned to a site is not reflective of the true climate at that location and time. This may in part explain why GxE was so difficult to detect.

Given what is known about the importance of solar radiation for the growth and yield of barley, it would be desirable in future analysis of this kind to include solar radiation, or PAR, as a variable in statistical models. This would allow the contribution of solar radiation to yield to be analysed and quantified.

There is a possibility that some spring barley was sown in February, as documented by Reid *et al.* (1929) in Lincolnshire on particular shallow soils in the late 1920s. This hasn't been accounted for in the models, as it was not possible to access any of the Irish farming diaries or calendars from the time. Sowing dates would be useful to confirm this.

## 3.7 Conclusion

Through combining recently published historical rainfall and temperature data with spring barley trials data, it was possible to identify climatic influences on spring barley yield variability seen in early twentieth century trials data in Ireland, building on the earlier findings of Student (1923). Variety was found to have a greater influence on spring barley yields in 1901-1906 than individual climate variables. July total rainfall, July maximum temperature and August maximum temperature were the most important climate variables, with the former having a negative effect on yield and the latter two temperature variables having positive relationships with yield.

The dataset allowed different variable selection and mixed modelling methods to be trialled and evaluated, with some encouraging results as well as providing some important lessons on modelling correlated covariates. These methods and their applications are now better understood and can be used on larger datasets, such as the UK variety trials data in Chapter 6, with more confidence.

# 4 The State of the UK Agroclimate

The grand challenge of producing more food for a growing population in a harsher, more variable climate requires an understanding of how past weather and climate variability has influenced crop production. Agroclimate indicators are a useful tool for quantifying the effect of changes in weather and climate on agriculture (Section 1.3.2) and aid interpretation of crop performance. This chapter addresses the limited historical analysis of agroclimate metrics and potential causes of recent cereal yield variability and trends in the UK. The output of this analysis can be used by farmers and growers to better understand performance, make climate-smart decisions on future crop and variety usage as well as inform farm management decisions. This information can be also integrated into national crop breeding and trial programmes to help identify and further the development of climate-resilient crop varieties.

In this chapter, the concept of a *State of the UK Agroclimate* is introduced as a periodic assessment of the changing UK agroclimate and its influence on crop production. Initially, the regional and national farm and variety trial yield trends and variability are analysed to confirm whether a 'yield plateau' persists in UK wheat (*Triticum aestivum* L.) and barley (*Hordeum vulgare* L.). Yield outliers are identified and their causes explored. Relevant agroclimate metrics identified in Chapter 2 (Table 2.14) are created using carefully selected climate datasets. Trends and variability in these metrics are analysed over the last four decades (1981-2020) and are related to farming practices.

## 4.1 Long-term production trends show increasing variability

Very little spring wheat is grown in the UK, therefore national wheat yields were used as a proxy for national winter wheat yields (Mackay *et al.*, 2011). From 1961 to 1990, wheat yields increased rapidly from 3.5 t/ha to 7 t/ha (~0.1 t/ha/yr) (Figure 4.1). From 1991 to 2010, the rate of growth slowed to 0.06 t/ha/yr, with 2001-2010 exhibiting a marked plateau (0.04 t/ha/yr). For 2011-2020, increased interannual variability consistent with the anticipated increase in extreme events associated with climate change, masks the longer-term trend in wheat yield. The highest yields were achieved in this period (e.g. 2015 9.0 t/ha), indicating higher yield potential, but also greater potential for large yield losses (e.g. 2012 6.7 t/ha). Consequently, the estimated rate of yield increase was only 0.01 t/ha/yr for the last decade, indicating further wheat yield stagnation.

Wheat area increased 2.5-fold from 1961 to 1990 to 2 Mha, but subsequent years have seen variations in area of up to 20%. Total production followed a similar trend increasing rapidly from 3 Mt in 1961 to 14 Mt in 1990 and then plateauing from 1990-2020 with large interannual variability of up to 4 Mt.

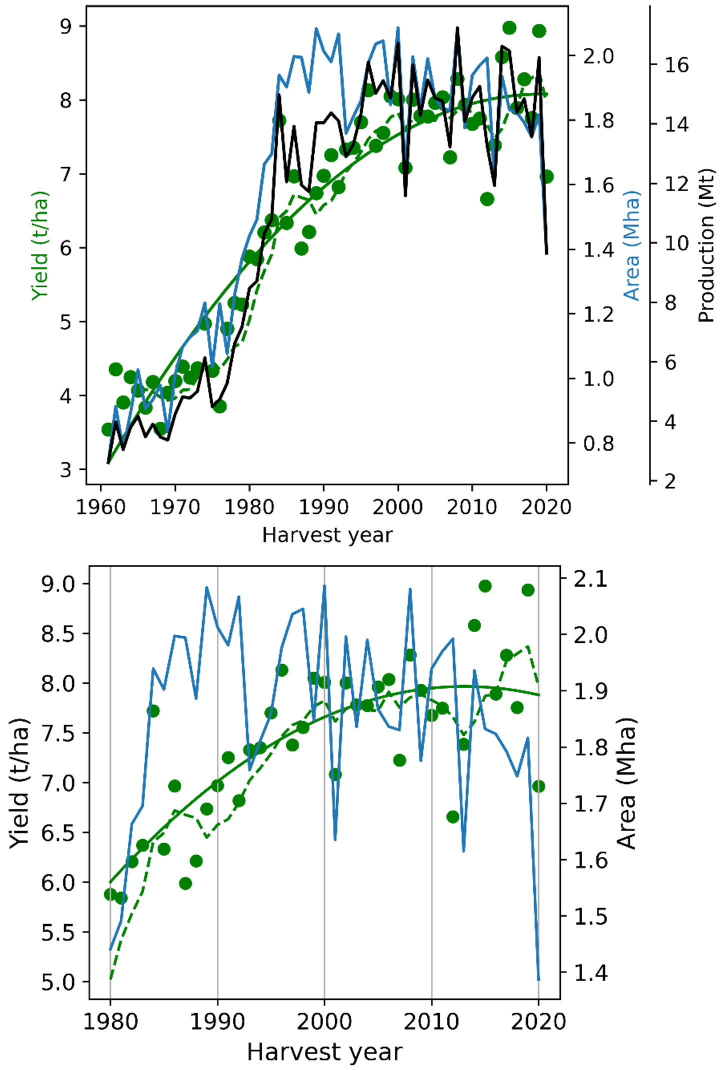



Figure 4.1: *National annual wheat yields (green dots), area (blue) and production total (black) for 1961-2020 (top) and 1980-2020 (bottom). The 5-year running mean yield (green dash) and quadratic (green) give an indication of long-term UK wheat yield trends. Spring and winter wheat are combined these plots due to the very low usage of spring wheat varieties. Data from FAOSTAT and DEFRA.*

National barley yields also showed significant growth from 1961 to 1990 of 2 t/ha (~0.07 t/ha/yr), before slowing across 1991-2010 to 0.04 t/ha/yr (Figure 4.2). In 2011-2020, there was again high variability of over 1.2 t/ha between years, indicating that whilst yields have still been increasing, there is greater potential for barley yield shocks. The falls in yield in 2013 and 2020 correspond to years of high spring barley and low winter barley area; more spring barley is grown in years of poor winter cropping conditions. As a result of this high interannual yield variability, the overall rate of yield increase in the most recent decade was very low at just 0.02 t/ha/yr. This indicates that on average, barley yields are still plateauing.

Combined barley area halved from the mid-1960s to 1990 (Figure 4.2), due to changes in subsidies which encouraged growing of oilseed rape. A corresponding decline was seen in barley production. In the most recent decade barley planted area, in particular spring barley, and production has begun to rise again, linked with a decrease in oilseed rape (OSR) as flea beetles become more problematic, in part due to the banning of neo-nicotinoids and milder winters (Thursfield, 2019; AHDB, 2022).

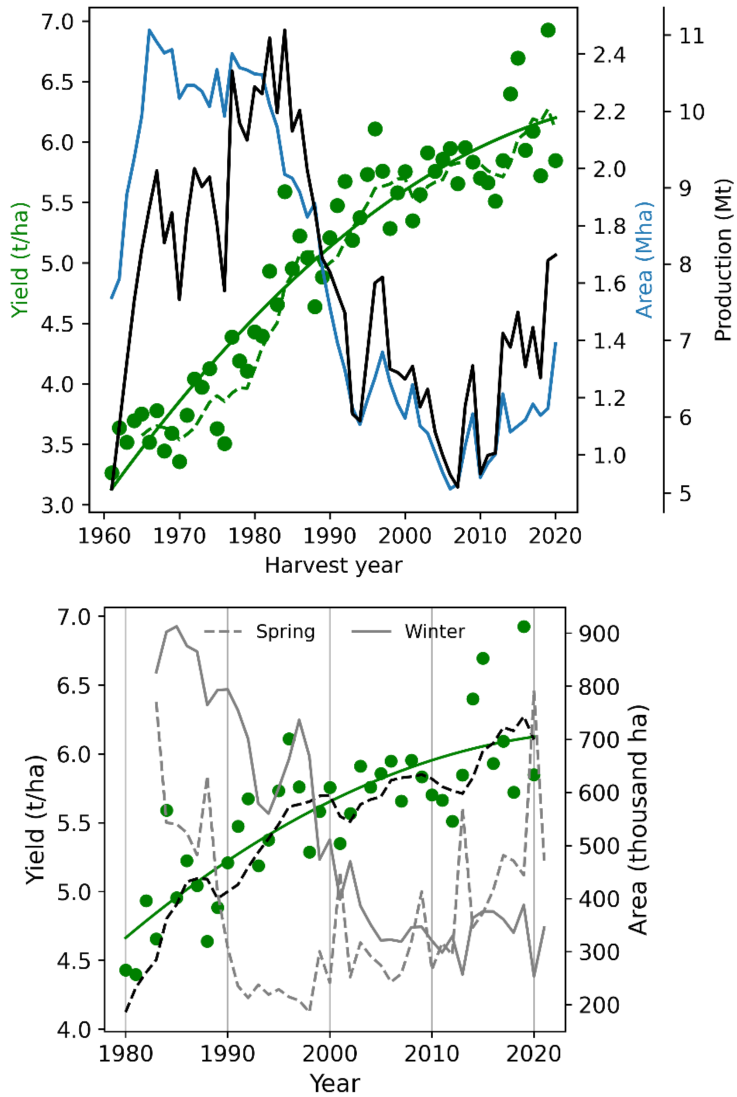



Figure 4.2*: National annual combined winter and spring barley yields (green dots), combined area (blue) and production total (black) for 1961-2020 (top). Spring barley (grey dashed) and winter barley (grey) area and combined yield (green) are shown for 1984-2020 (bottom). The 5-year running mean yield (green dash) and quadratic (green) give an indication of long-term UK barley yield trends. Data from FAOSTAT and DEFRA.*

Winter wheat, winter barley and spring barley variety trials all show linear increases in yield across 1982-2018, at a rate of 0.08, 0.09 and 0.04 t/ha/yr, respectively (Figure 4.3). In contrast to the quadratic relationships between harvest year and national yields (Figures 4.1 and 4.2), the linear increases in variety trial yields shows variety trial yields have not plateaued. However, there is still high interannual yield variability, particularly in the last decade for winter wheat.

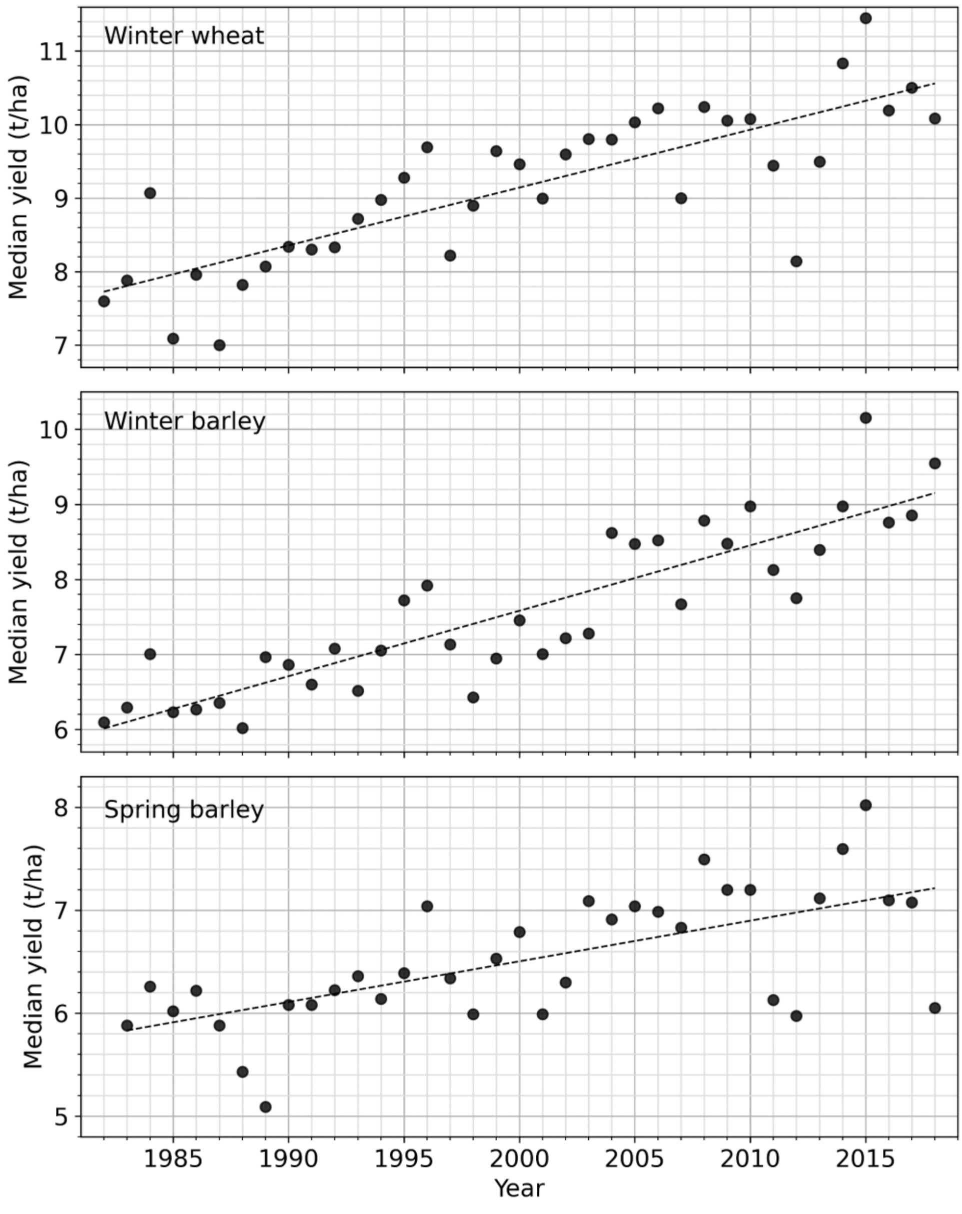



Figure 4.3: *Winter wheat, winter barley and spring barley median annual variety trial yields for 1982-2018 (from 1983 for spring barley) for treated and untreated trials combined.*

## 4.2 There are regional deviations to national yield anomalies

Abnormally high yields occurred in 1984, 1996, 2015 and 2019 for both national wheat and barley (Figure 4.4). For wheat, the negative yield outliers occurred in 1987, 1988, 2012 and 2020, whilst for barley they occurred in 1988, 1989, 2012 and 2018.

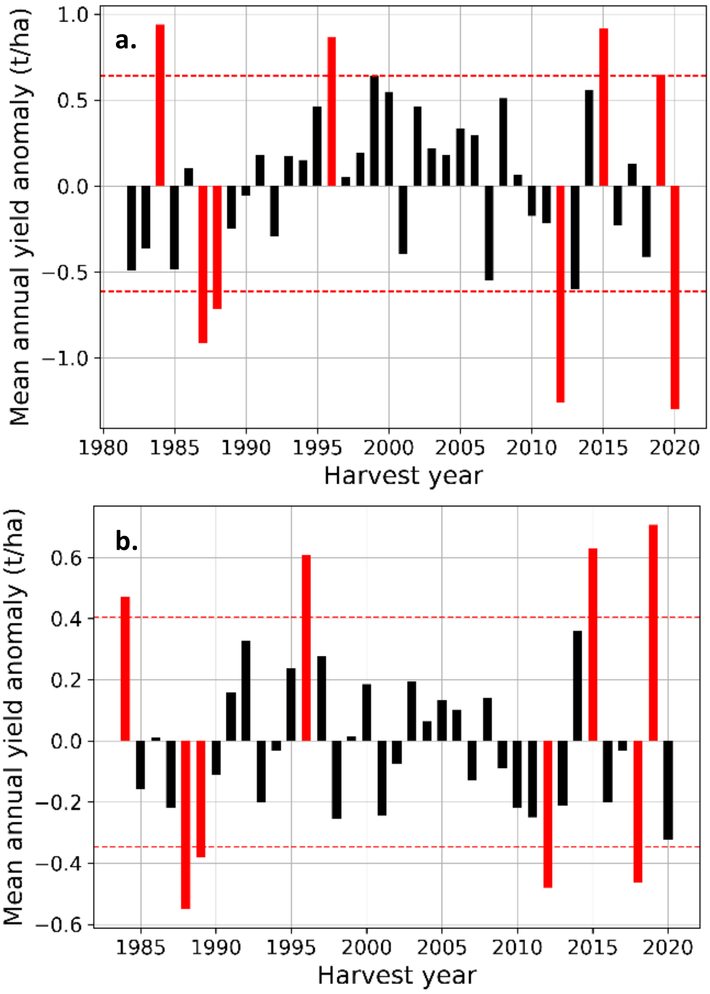



Figure 4.4: *National yield anomalies for **a**. wheat and **b.** barley, after removing the long-term trends for 1982-2020 and 1984-2020, respectively. Years in red represent anomalies in the top and bottom 10%. Data from DEFRA.*

Regionally, there were deviations from the national wheat yield anomalies (Figure 4.5). For example, in 2010 Northern Ireland had the largest positive yield anomaly of the period 1999-2019, whilst the rest of the country experienced very small or slightly negative anomalies. Then in 2019, when nationally and across many regions there was a large positive yield anomaly, Northern Ireland's wheat yields were very average. In 2018 Scotland experienced a large negative yield anomaly whilst the rest of the country only saw small negative deviations in yield. In 2001, large negative yield anomalies occurred in most regions, but this wasn't detected at a national level. This highlights how national yield analysis alone can hide variation within the country.

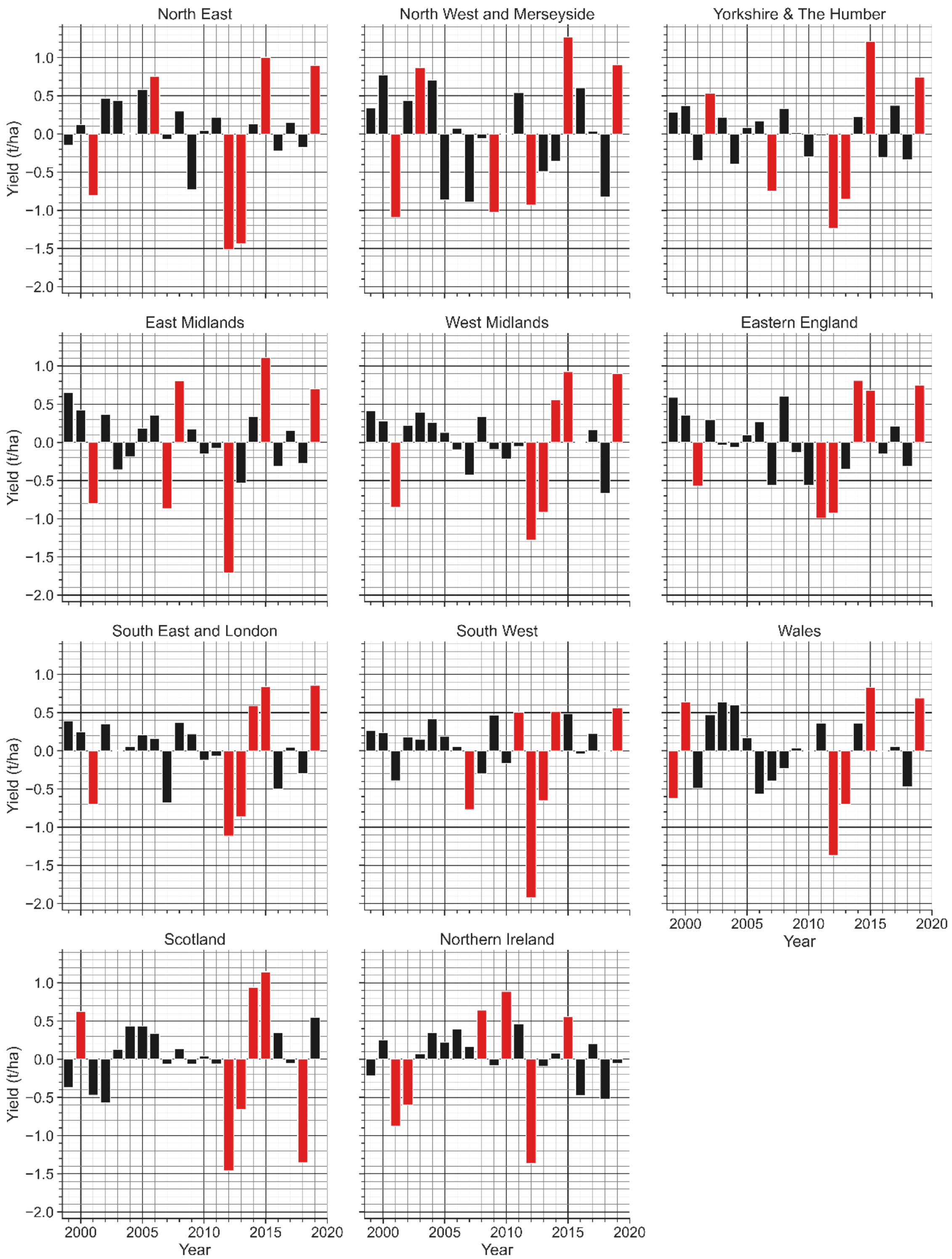



Figure 4.5: *Regional wheat yield anomalies for 1999-2019, after removing the regional trend from 1999-2019. Years in red represent anomalies in the top and bottom 10%. Data from DEFRA.*

There is overlap in the anomalous years in the on-farm data and trials data, particularly for positive yield anomalies, with 1996 and 2015 also identified in all three trials datasets as good years (Figure 4.6). 2012 was a particularly bad year across the variety trials and national farm data.

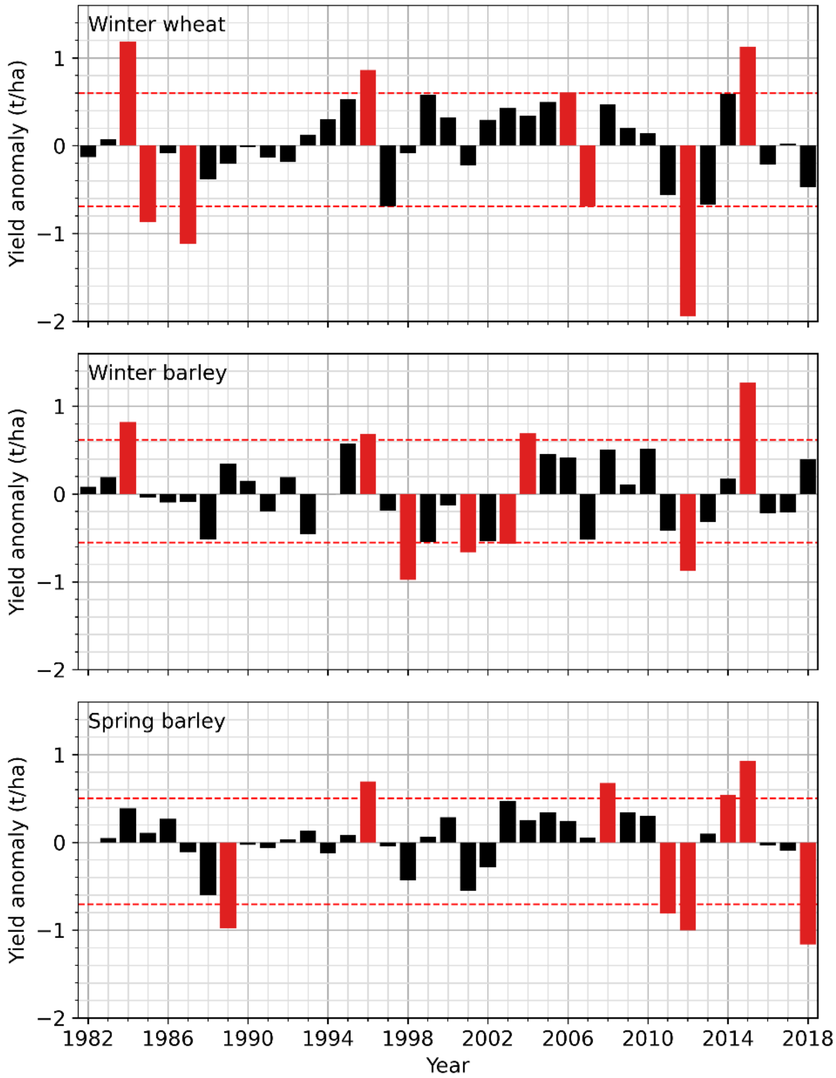



Figure 4.6*: Variety trial yield anomalies for 1982-2018 for winter wheat and winter barley and 1983-2018 for spring barley. Fungicide treated and untreated variety trials were combined. Harvest years in red represent anomalies in the top and bottom 10%. Data provided by AHDB.*

Given the available data on national, regional and variety trial yield anomalies for 1982-2020, the high yielding harvest years for wheat and barley were 1984, 1996, 2015 and 2019. The low yielding harvest years were 1988, 2001 and 2012. 2018 was also low yielding for spring barley and 2020 was for wheat (Figures 4.4, 4.5 and 4.6).

## 4.3 Climate anomalies help explain anomalous yields

Use of Met Office ranked data allowed anomalous months and seasons in outlying yield harvest years to be identified, and appropriate maps were then downloaded or created to see the spatial extent of likely influential weather patterns (Figure 4.7 and Figure 4.8). Anomaly plots (Figures 4.9 and A4-6) complement this spatial representation by giving an indication of the growing season climate from September to August. Winter North Atlantic Oscillation (WNAO) indices have also been included from Figure 3.6, to give a qualitative measure as to whether this may have contributed to the observed yield anomalies.

The first good harvest year, **1984**, experienced high sunshine (gridded data not available) and low rainfall in July across the whole of the UK (Figure 4.7), increasing the photosynthetic potential during the grain fill period. It was a relatively cool growing season, particularly January to April, encouraging an extended growth period and the spring and summer months were largely drier and sunnier than average (Figure 4.9). 1984 had a positive WNAO of 1.70.

With the exception of January, all of the months between December and July in the **1996** harvest year were also cool (Figures 4.7 and 4.9), reflected in a later than average harvest (~30$^{th}$ August). May maximum temperature was the 4$^{th}$ coldest on record (of 104 years) and December was the 8$^{th}$ coolest. June 1996 was in the top 10% for sunshine duration (Figure 4.9). 1996 had a strong negative WNAO of -2.24.

Low September rainfall was a feature of the **2015** harvest year (Figure 4.7), allowing farmers onto the land to drill earlier than normal (~2$^{nd}$ October), as well as encouraging deeper root establishment which can be important in times of subsequent drought or low precipitation during the harvest year. The months May to August were cooler than average (Figure 4.9). 2015 had a strong positive WNAO of 2.06.

The most recent good year, **2019**, consisted of several very warm months, including the warmest February maximum temperatures on record, and December, July, March and November all experienced top 10% warmest minimum temperatures (Figure 4.7). Summer was also the 11$^{th}$ wettest nationally, although this was mostly in the North and Midlands, rather than growing areas in the East. 2019 had a positive WNAO 1.15.

The first of the low yielding years, **1988** had the wettest July on record, with much of the country receiving over 200% of the 1991-2020 average (Figure 4.8). This was accompanied by low maximum temperatures and sunshine, receiving just ~85% of the period average for July. Drilling date was also late, with a median of 25$^{th}$ October. 1988 had a WNAO of -0.13.

Autumn 2000 was the wettest on record, with parts of the South-East receiving over 350% of the 1991-2020 average in October and over 200% in November (Figures 4.8 and 4.9). This coincided with delayed drilling of winter wheat trials (~20$^{th}$ October). By contrast, in Scotland November rainfall was less than 70% of the average in the North and West, and October rainfall was only slightly more than average. This regional variation in rainfall anomalies may help explain why Scotland did not have a strong negative yield anomaly in **2001**. 2001 had a WNAO of -0.44.

The most extreme yield loss occurred in **2012**, the year with the wettest April and June on record (Figure 4.8). The UK also received the lowest June sunshine on record and 3rd lowest in summer, leading to low photosynthesis rates during the grain fill period. Harvest dates were much later than average (~5th September), indicative of a long growing season. 2012 had a strong positive WNAO of 2.18.

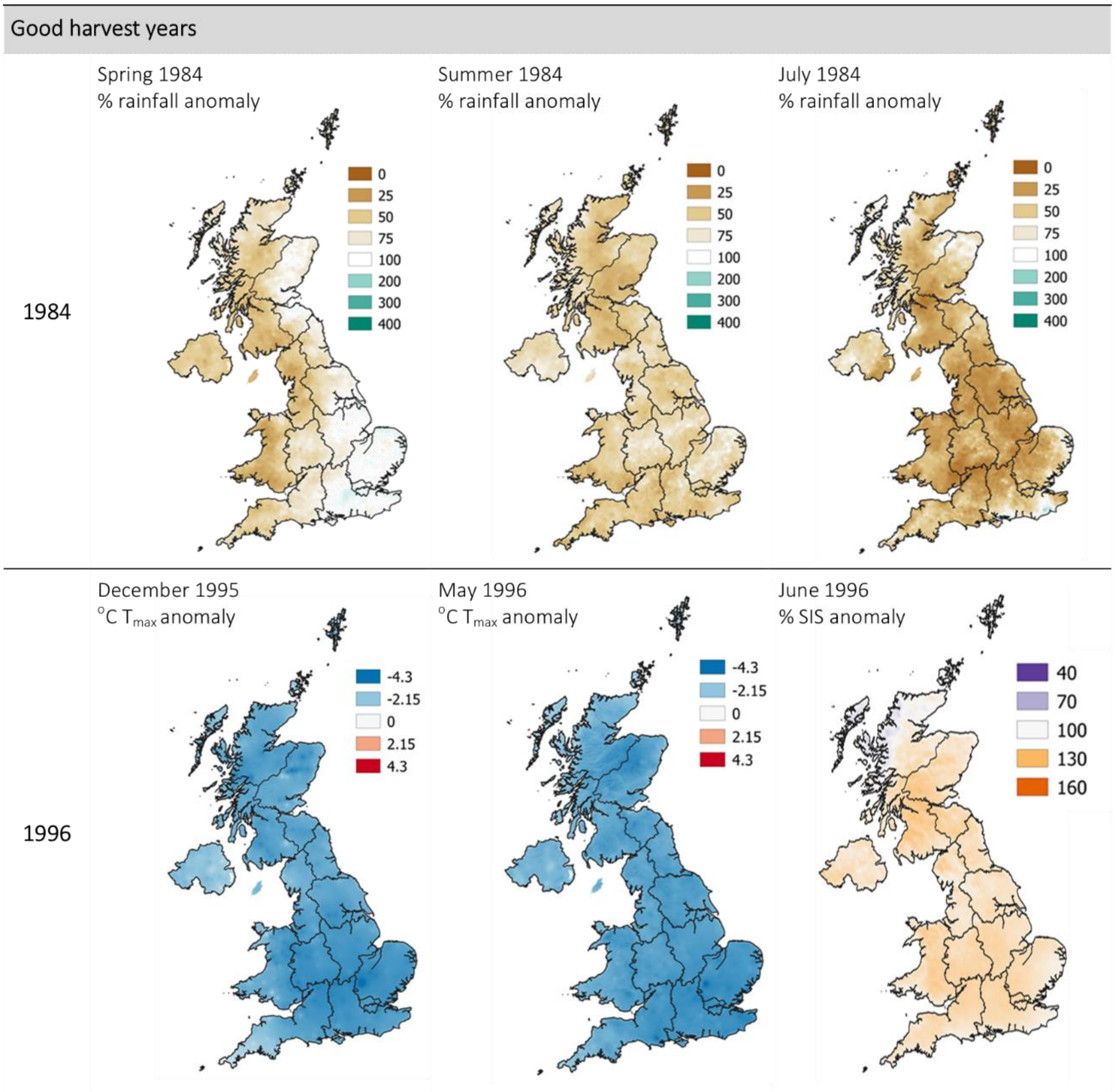
Spring barley was the only crop to yield in the lowest 10% in **2018**. The spring 2018 rainfall anomaly was very split between a drier than average North-West of the UK and a very wet South and East of the country (Figure 4.8). The latter encompasses much of the growing area for spring barley (Figure 2.7) and given the crop is typically drilled at this time of year, heavy rainfall could have delayed field activity, which may explain the delayed spring barley trials drilling date of ~13th April, relative to the 2007-2018 mean of 27th March. This was followed by the hottest May and 2nd hottest June and July on record, which encourages faster development, hence the earlier harvest date ~22nd August, the earliest in the 2007-2018 period. In Scotland, seven consecutive months (February to August) of below average rainfall (Figure 4.9) resulted in drought conditions, with a Standardized Precipitation Index (SPI) of <-2.0 i.e. "extremely dry" in some regions (Centre for Ecology and Hydrology, 2022). This was reported to have impacted both quality and yield of malt barley supply (Berry and Brown, 2021).

The **2018** winter barley harvest was less affected by the poor spring conditions as the crop is already established by this time of year. Hot summer temperatures did likely limit winter barley crop yields somewhat by accelerating growth, as reflected by an earlier median harvest date of ~7th August, nearly 2 weeks earlier than the 1988-2018 average. 2018 had a positive WNAO of 0.91.

The poor wheat (Figure 4.4) and spring barley (Figure 4.6) yields of **2020** came after a year of highly variable rainfall. For parts of the wheat growing area (e.g. East England), over 170% of the 1991-2020 rainfall average fell in autumn 2019 (Figure 4.8), making land access more difficult and potentially delaying drilling (drilling dates unavailable for this year). This was followed by a series of storm events in February 2020, which contributed to this being the wettest February on record and made it difficult for spring crops to establish. Given both winter wheat and spring barley were mostly drilled at a time of high soil moisture, shallow root systems may have formed, such that in the subsequent very dry spring, in which much of the growing area received less than 50% of the

1991-2020 average rainfall, intake of nutrients without sufficient moisture could have been limited. 2020 had one of the highest WNAO on record, of 2.85.

Combined, these show that monthly and seasonal climate anomalies can help explain some observed yields. However, for some years, such as 2019, the climatic contribution is less clear. Hence more specific agroclimate metric analysis is required. WNAO values showed some abnormal values in the anomalous yield years, for example high WNAO values in 2020 and 2012, which were both low yielding years. However, the high yielding year of 2015 also had a high WNAO and there was no obvious relationships between WNAO and average national yields.


Good harvest years
1984
Spring 1984
% rainfall anomaly
0
25
50
75
100
200
300
400
Summer 1984
% rainfall anomaly
0
25
50
75
100
200
300
400
July 1984
% rainfall anomaly
0
25
50
75
100
200
300
400
1996
December 1995
°C $T_{max}$ anomaly
-4.3
-2.15
0
2.15
4.3
May 1996
°C $T_{max}$ anomaly
-4.3
-2.15
0
2.15
4.3
June 1996
% SIS anomaly
40
70
100
130
160

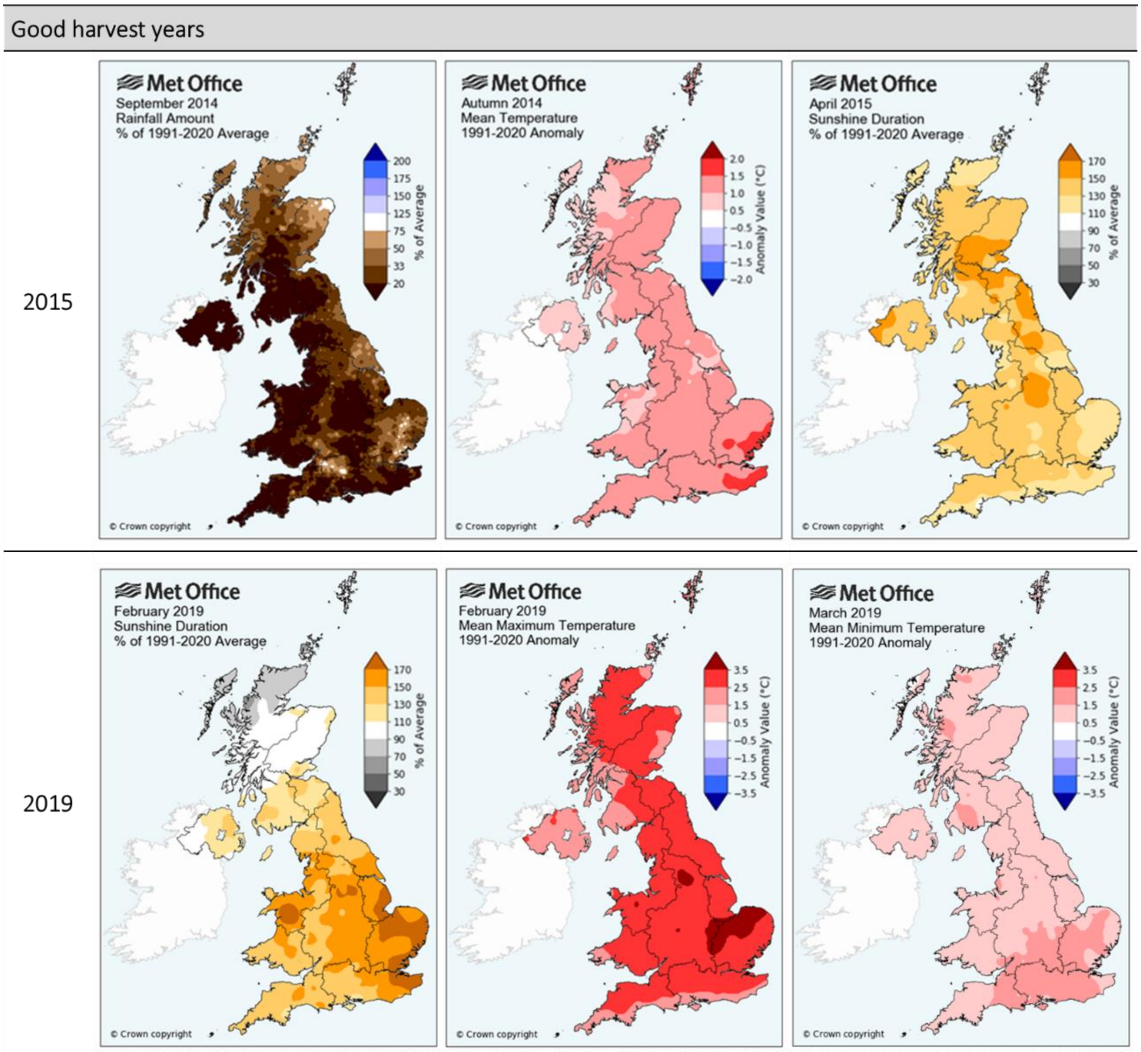
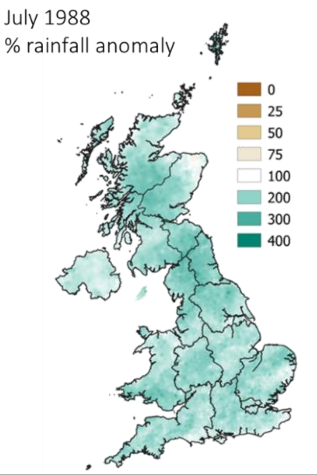
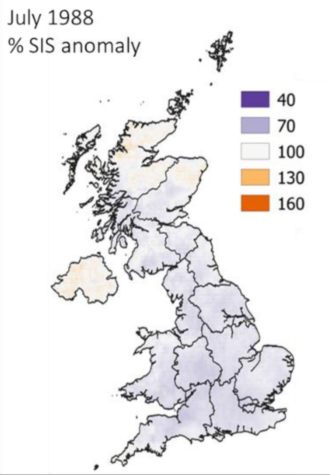


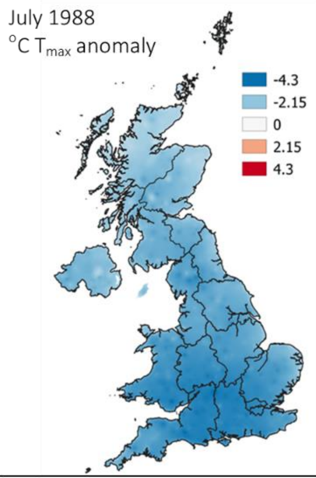

Figure 4.7*: Climate anomalies in years of high yields (top 10%). All anomalies are compared to the reference period 1991-2020. Temperature and precipitation data from HadUK has a 1km x 1km resolution* (Hollis *et al.*, 2019). *SIS = surface incoming solar radiation from CMSAF has a resolution of 0.05$^o$ x 0.05$^o$* (Pfeifroth, Trentmann, *et al.*, 2018).

## Bad harvest years

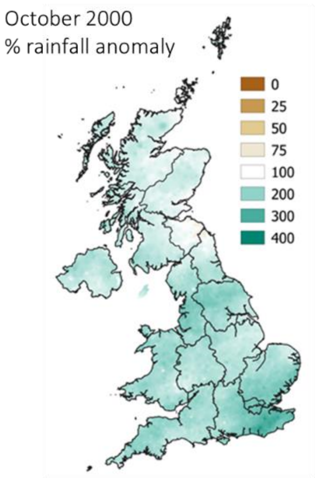
1988

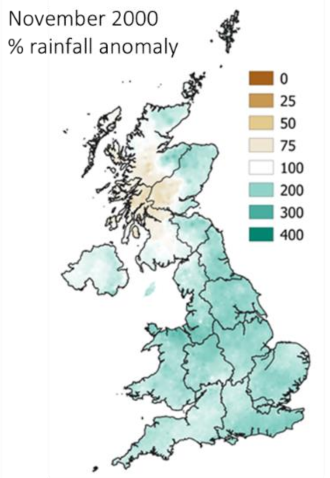


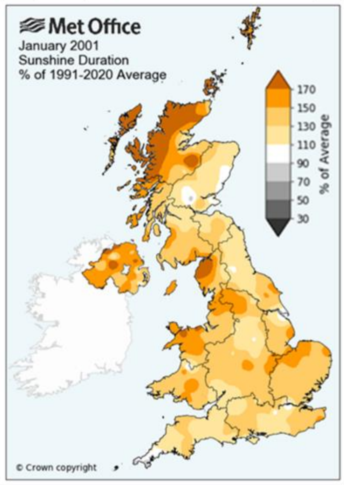



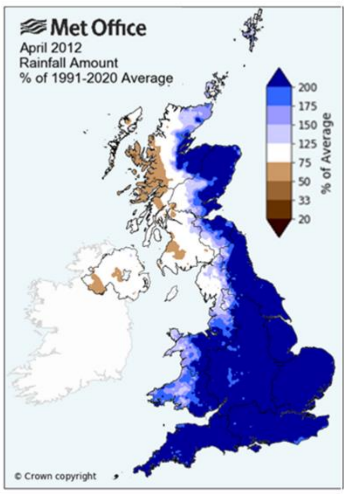

2001

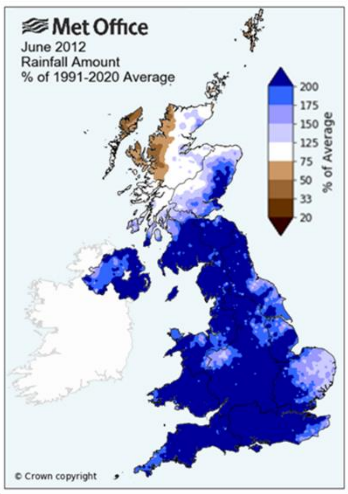


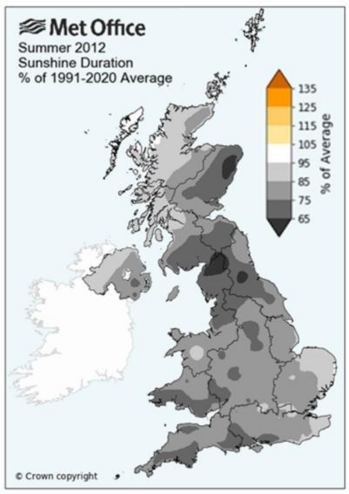




2012



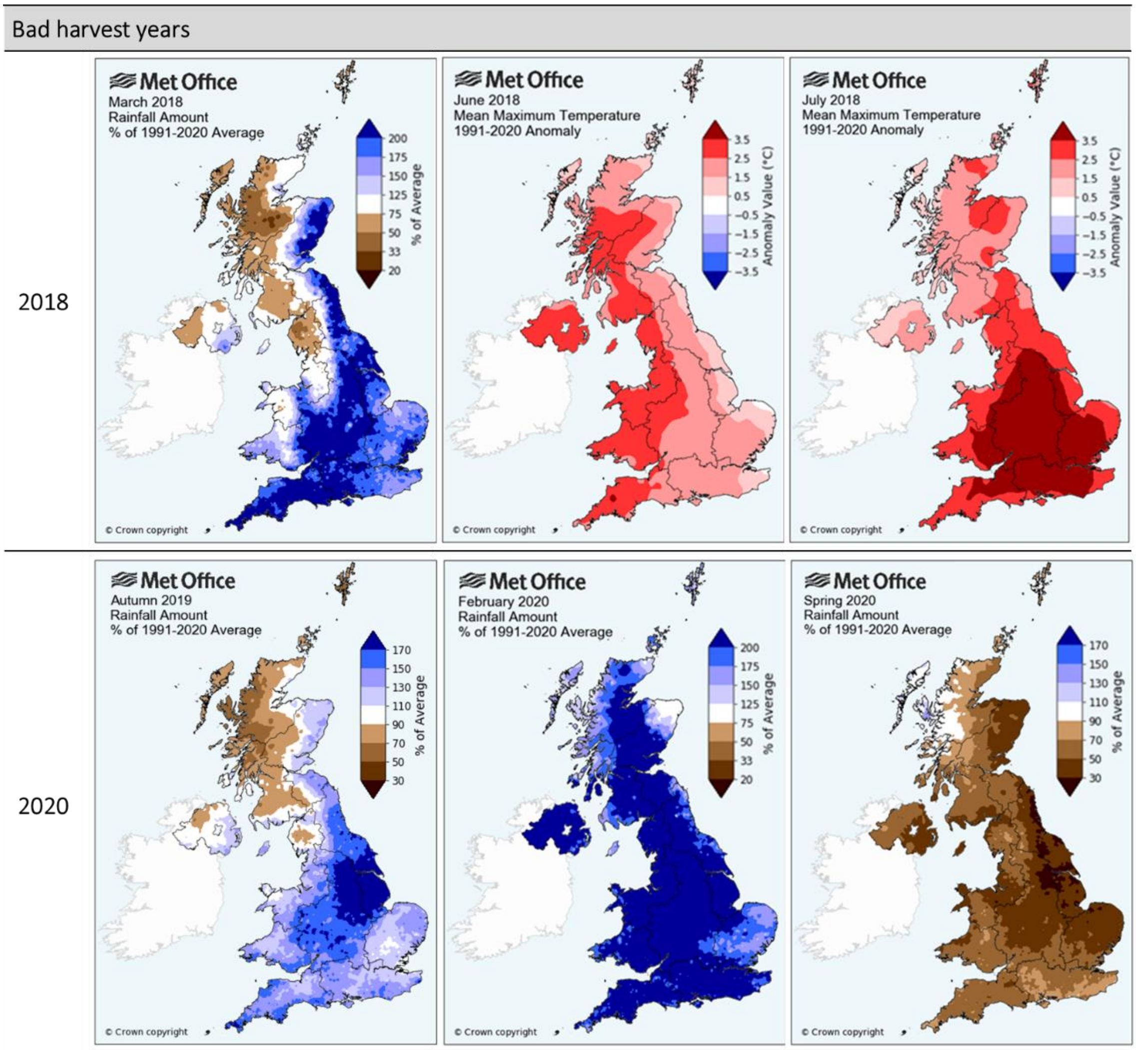
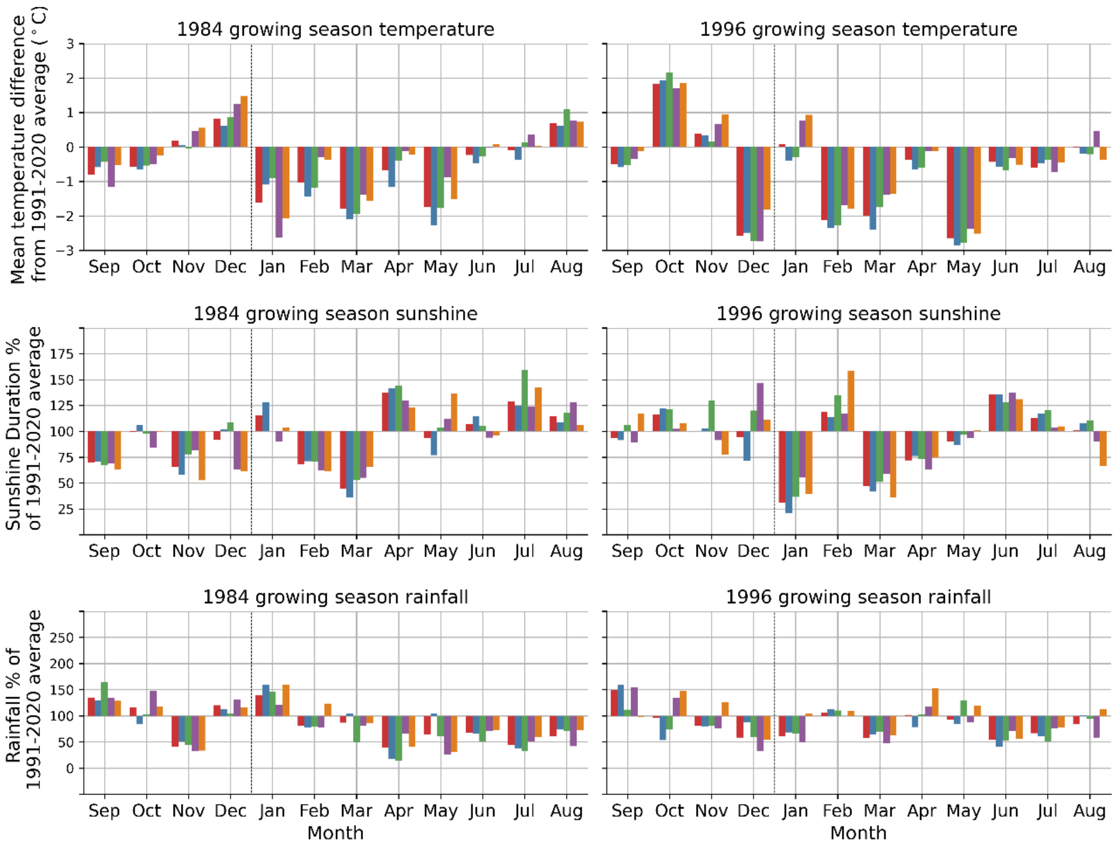
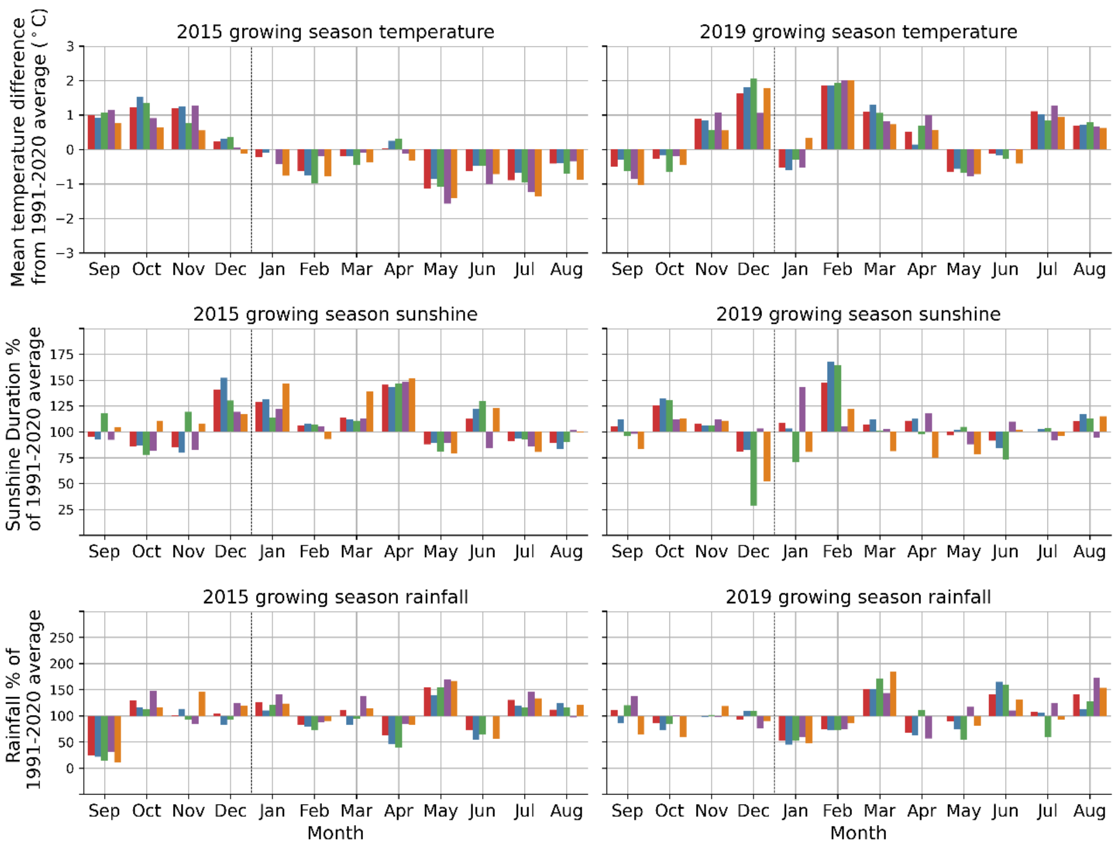
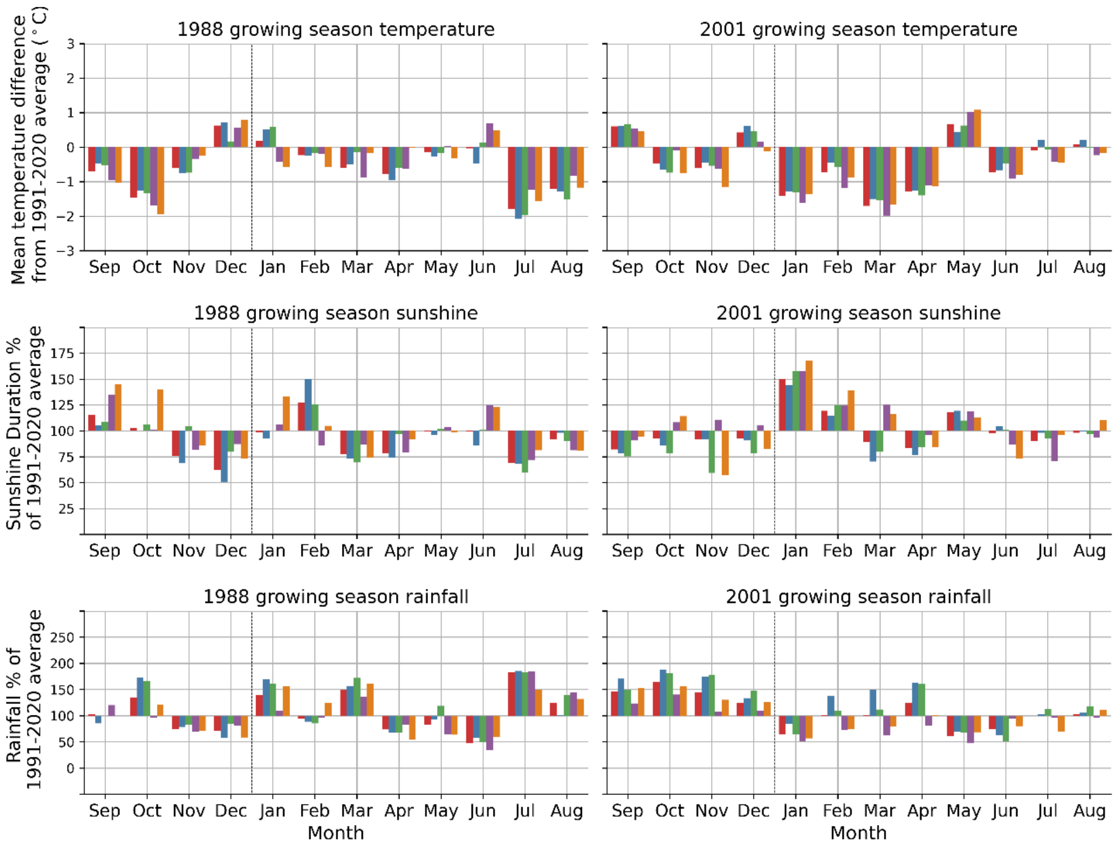
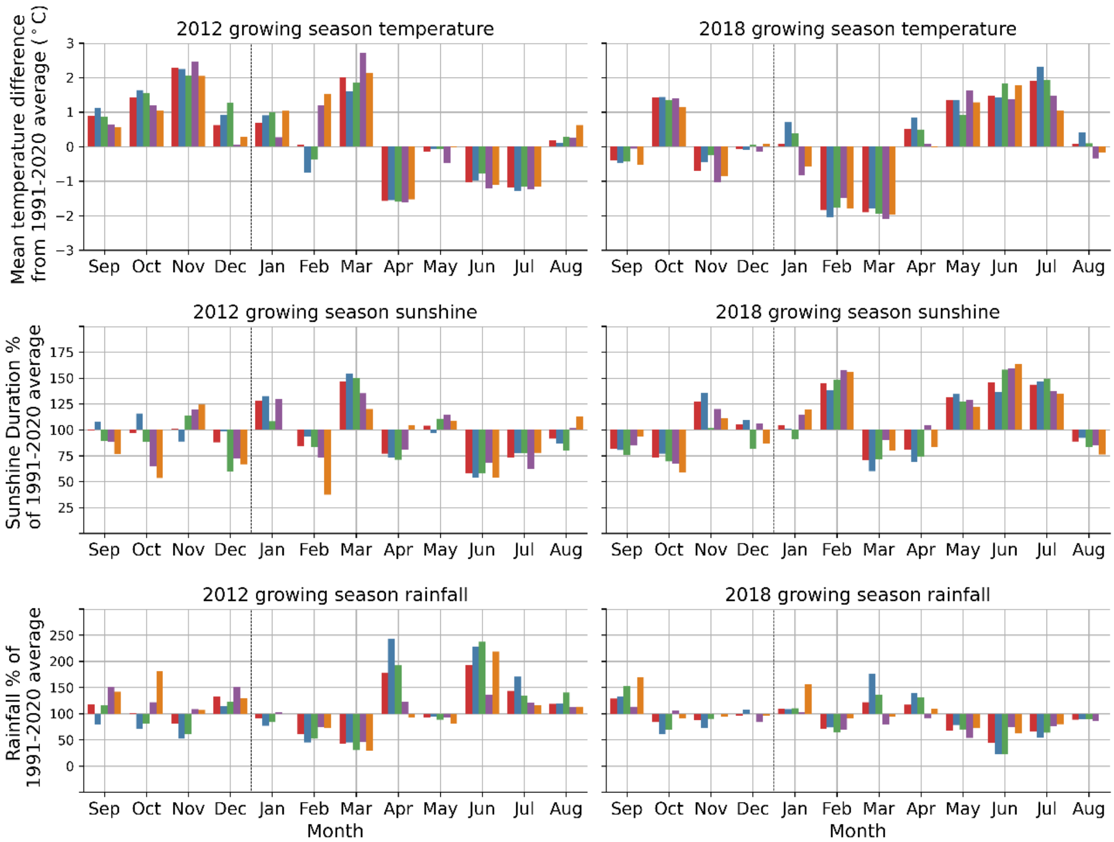



Figure 4.8: *Climate anomalies in years of low yields (bottom 10%). All anomalies are compared to the reference period 1991-2020. Temperature and precipitation data from HadUK has a 1km x 1km resolution (Hollis et al., 2019). SIS = surface incoming solar radiation from CMSAF has a resolution of 0.05º x 0.05º (Pfeifroth, Trentmann, et al., 2018).*


1984 growing season temperature
1996 growing season temperature
Mean temperature difference from 1991-2020 average (°C)
1984 growing season sunshine
1996 growing season sunshine
Sunshine Duration % of 1991-2020 average
1984 growing season rainfall
1996 growing season rainfall
Rainfall % of 1991-2020 average
Sep Oct Nov Dec Jan Feb Mar Apr May Jun Jul Aug
Month


2015 growing season temperature
2019 growing season temperature
Mean temperature difference from 1991-2020 average (°C)
2015 growing season sunshine
2019 growing season sunshine
Sunshine Duration % of 1991-2020 average
2015 growing season rainfall
2019 growing season rainfall
Rainfall % of 1991-2020 average
Sep Oct Nov Dec Jan Feb Mar Apr May Jun Jul Aug
Month


1988 growing season temperature
2001 growing season temperature
Mean temperature difference from 1991-2020 average (°C)
3
2
1
0
−1
−2
−3
Sep Oct Nov Dec Jan Feb Mar Apr May Jun Jul Aug
1988 growing season sunshine
2001 growing season sunshine
Sunshine Duration % of 1991-2020 average
175
150
125
100
75
50
25
Sep Oct Nov Dec Jan Feb Mar Apr May Jun Jul Aug
1988 growing season rainfall
2001 growing season rainfall
Rainfall % of 1991-2020 average
250
200
150
100
50
0
Sep Oct Nov Dec Jan Feb Mar Apr May Jun Jul Aug
Month


2012 growing season temperature
2018 growing season temperature
Mean temperature difference from 1991-2020 average (°C)
3
2
1
0
−1
−2
−3
Sep Oct Nov Dec Jan Feb Mar Apr May Jun Jul Aug
2012 growing season sunshine
2018 growing season sunshine
Sunshine Duration % of 1991-2020 average
175
150
125
100
75
50
25
Sep Oct Nov Dec Jan Feb Mar Apr May Jun Jul Aug
2012 growing season rainfall
2018 growing season rainfall
Rainfall % of 1991-2020 average
250
200
150
100
50
0
Sep Oct Nov Dec Jan Feb Mar Apr May Jun Jul Aug
Month

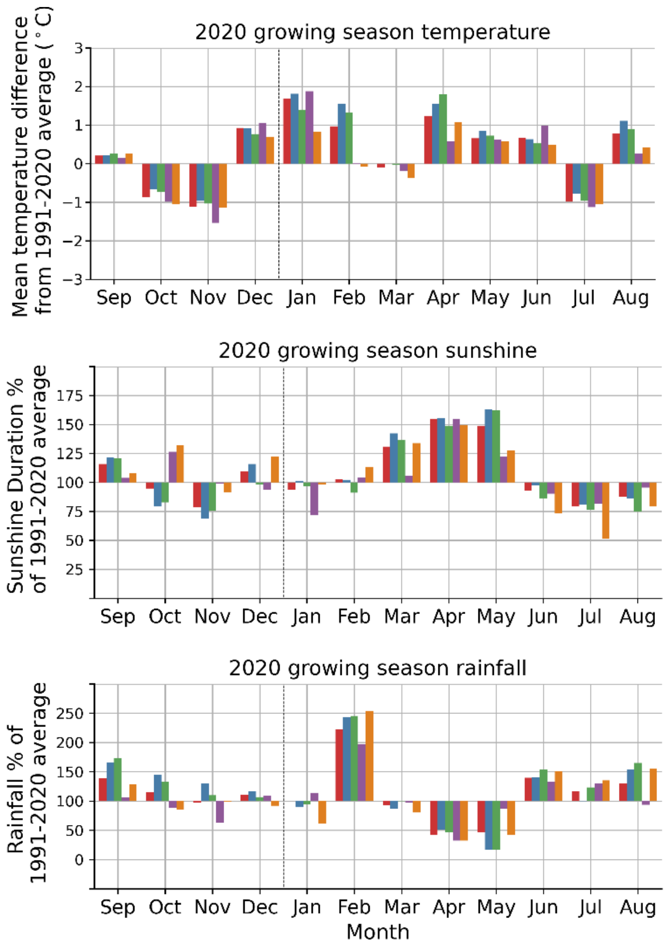



Figure 4.9*: Monthly temperature, sunshine duration and rainfall anomalies for anomalous yield growing seasons relative to the 1991-2020 averages for the growing season for the UK (red), England (blue), Wales (green), Scotland (purple) and Northern Ireland (NI) (orange). 'Good' harvest years are displayed first, followed by 'bad' harvest years. Data from Met Office year ordered time-series.*

## 4.4 The Changing UK Agroclimate

### 4.4.1 Drilling day of year is earlier but harvest dates haven't changed

Winter wheat variety trial drilling and harvest dates were split into the four UK nations to look at variation across the UK. The median drilling day of year for 1988-2018 was 283 i.e. the 10th October, but drilling dates varied from early September to as late as early February (Figure 4.10a.). Harvest dates were more dependent on the nation the trial site is in. Median harvest day of year across the UK was 233 i.e. 22nd August, but there is much variation between the nations (Figure 4.10b.). In Wales and England, where the latitude is lower, winter wheat trials are typically harvested much earlier, on day 228 (~16th August) and 231 (~19th August), respectively. In Northern Ireland and Scotland trials are harvested two weeks later, on day 245 (2nd September) and day 246 (3rd September). The higher latitude controls this, with cooler temperatures allowing for slower development. As a result, winter wheat in Scotland and Northern Ireland has a longer growing season on average.

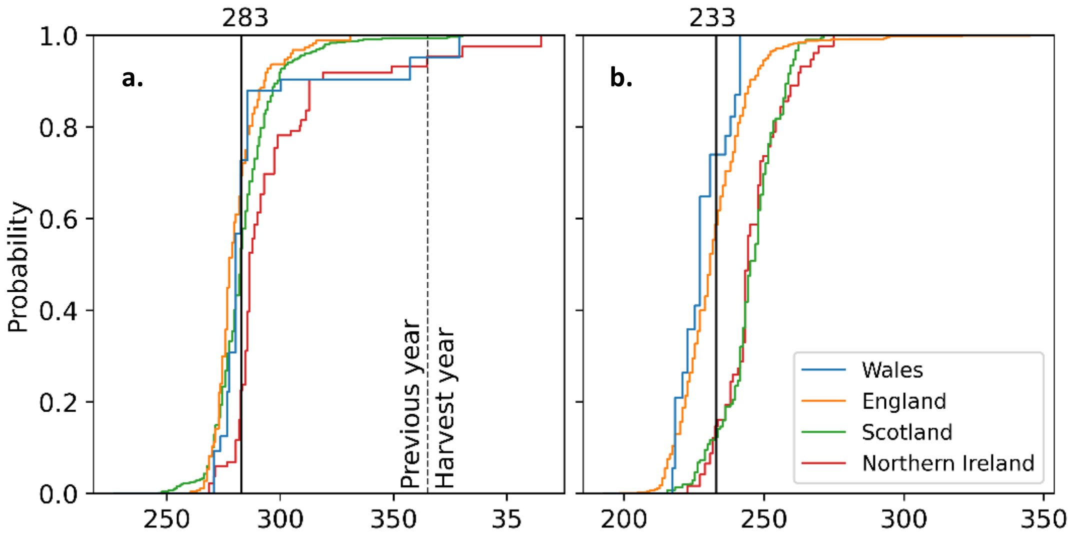



Figure 4.10*: Cumulative* a*. drilling day of year and* b*. harvest day of year for winter wheat variety trials in Wales, England, Scotland and Northern Ireland for 1988-2018. The median drilling day of year is 283 (10th October) and harvest day of year is 233 (21st August).*

To investigate the relationship between latitude and drilling and harvest day of year in England, regional dates were explored. However, the limited number of trials in several regions meant this analysis was inconclusive.

Analysis of time-series data shows that since 1988, winter wheat trials drilling day of year has been getting earlier by one day every three years (Figure 4.11a.). However, in the last few years, there was more variability and the trend may have started to reverse, so drilling is occurring slightly later again. This trend has been documented in on-farm winter wheat surveys, which found that in 2017-2019 there was an increased proportion of farms drilling in October rather than late September (Turner *et al.*, 2021).

The delayed drilling for harvest years 1988 and 2001 really stands out. October 1987 was in the top 9% of wettest Octobers and autumn 2000 was the wettest on record, therefore heavy rainfall and possible waterlogging was likely a contributing factor to the delay. Sowing dates are also highly dependent on previous cropping: early sowing typically follows OSR, whilst late November-sown crops often followed later harvested sugar beet or potatoes (Hardwick *et al.*, 2001). Given that the proportion of wheat crops following OSR almost doubled from 1991-2016 on-farm (Turner *et al.*, 2021) and in the trials data (Figure A7), long-term changes in cropping will also be affecting trends in sowing dates. Linked with previous cropping, drilling date has also been influenced by responses to black grass pressure as a confounding factor. The reduced blackgrass risk through the reduction in repeat wheat has been reversed by the increased risk due to a movement towards minimum tillage (Turner *et al.*, 2021).

Harvest dates showed considerable variation from year to year with no significant trend over time (Figure 4.11b.). Overall, the growing season length (from drilling to harvest) increased by one day every two years since 1988 (p=0.003).

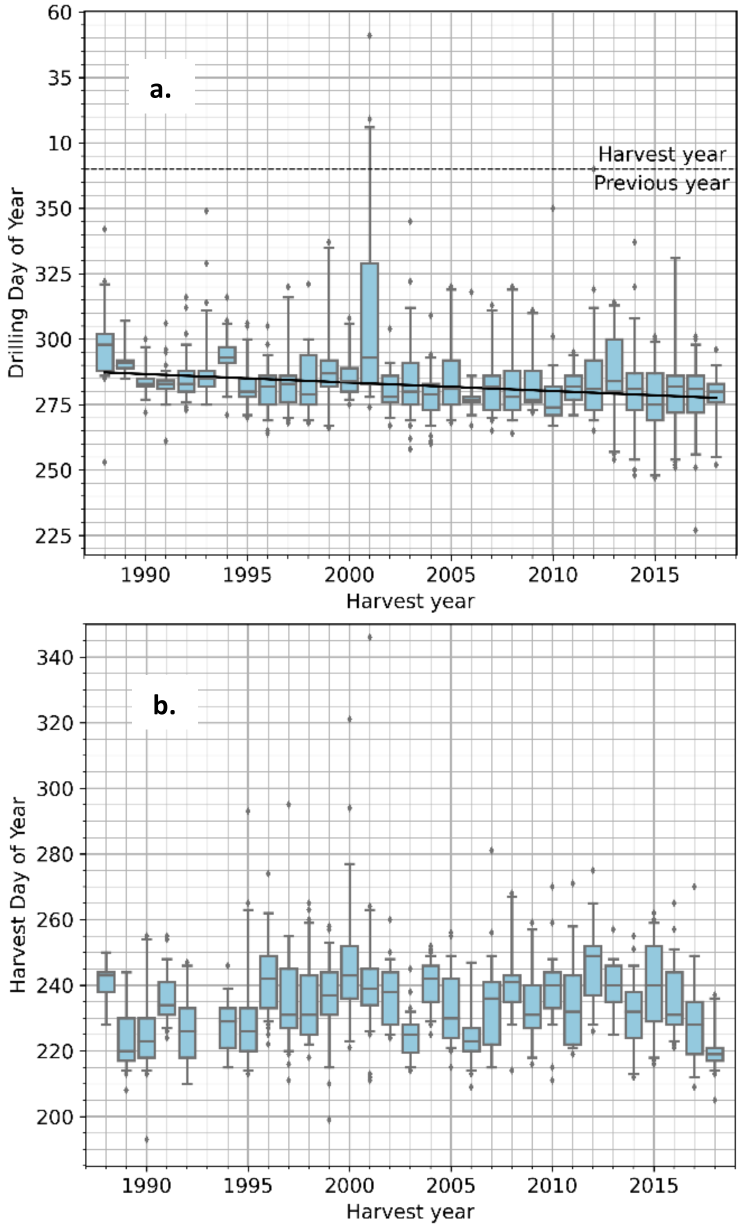



Figure 4.11*: The annual ranges in a. drilling day of year and b. harvest day of year for winter wheat variety trials, 1988-2018. The linear trend in drilling day of year is statistically significant (p<0.001).*

The delayed harvests of 1988, 1996, 2000 and 2012 are noticeable. Heavy summer rainfall likely delayed harvest in 1988 and 2012, whilst in 1996 several months of cooler weather may have slowed development, leading to a later maturity date and longer growing season.

### 4.4.2 The Start of the Growing Season is getting earlier

Decadal mean SOGS show that the growing season starts earliest in the South of England and nationwide SOGS has gotten earlier when comparing 1981-1990 and 2011-2020 (Figure 4.12). There is variability across the 40-year period, such that the average SOGS for 2001-2010 was later than the previous decade. In 2011-2020 in parts of the South West and South East SOGS was as early as the 6-10th January. Crops respond to many environmental cues including temperature

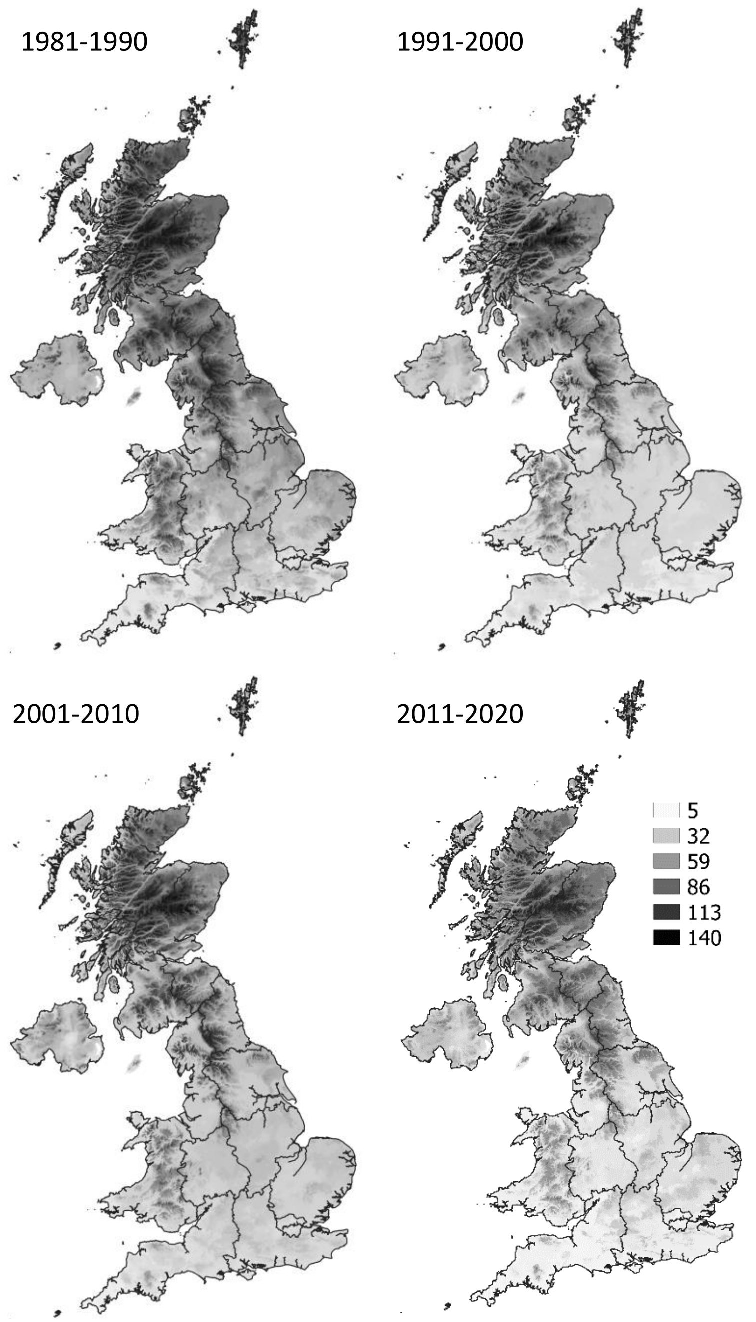



Figure 4.12*: Start of the growing season (SOGS) day of year averages for 1981-1990, 1991-2000, 2001-2010 and 2011-2020, calculated from 1st January each year. Created using HadUK 1km x 1km gridded temperature data* (Hollis *et al.*, 2019).

and day length, therefore an earlier SOGS may not necessarily mean that field tasks immediately follow.

It might be expected that earlier SOGS also means earlier harvest date, however, use of winter wheat trials data shows that this was not the case (Figure 4.13). There was no relationship between SOGS and harvest date for winter wheat trials in the UK. An earlier SOGS can still be followed by a cool spring/summer, delaying the accumulation of the required GDD.

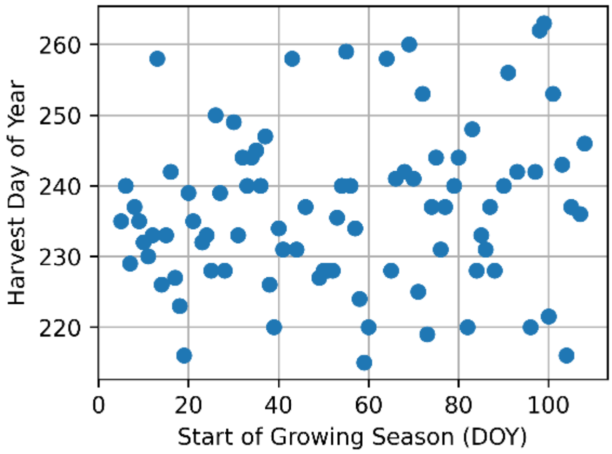



Figure 4.13*: Start of Growing Season against the harvest day of year for winter wheat trial sites across the UK for 1988-2018. Calculated using the method described in Table 2.14.*

### 4.4.3 The available Growing Degree Days has increased

Growing Degree Days (GDD) (calculated from 1st September-31st August; Table 2.14) for each decade within 1982-2020 was highest in the South-East and East Anglia, at over 1800 °C days, compared to less than 1000 °C days in parts of Scotland (Figure 4.14). GDD increased each decade, with an overall ~15% increase in GDD between 1982-1990 and 2011-2020, aligning with findings of Nesbitt *et al.* (2022).

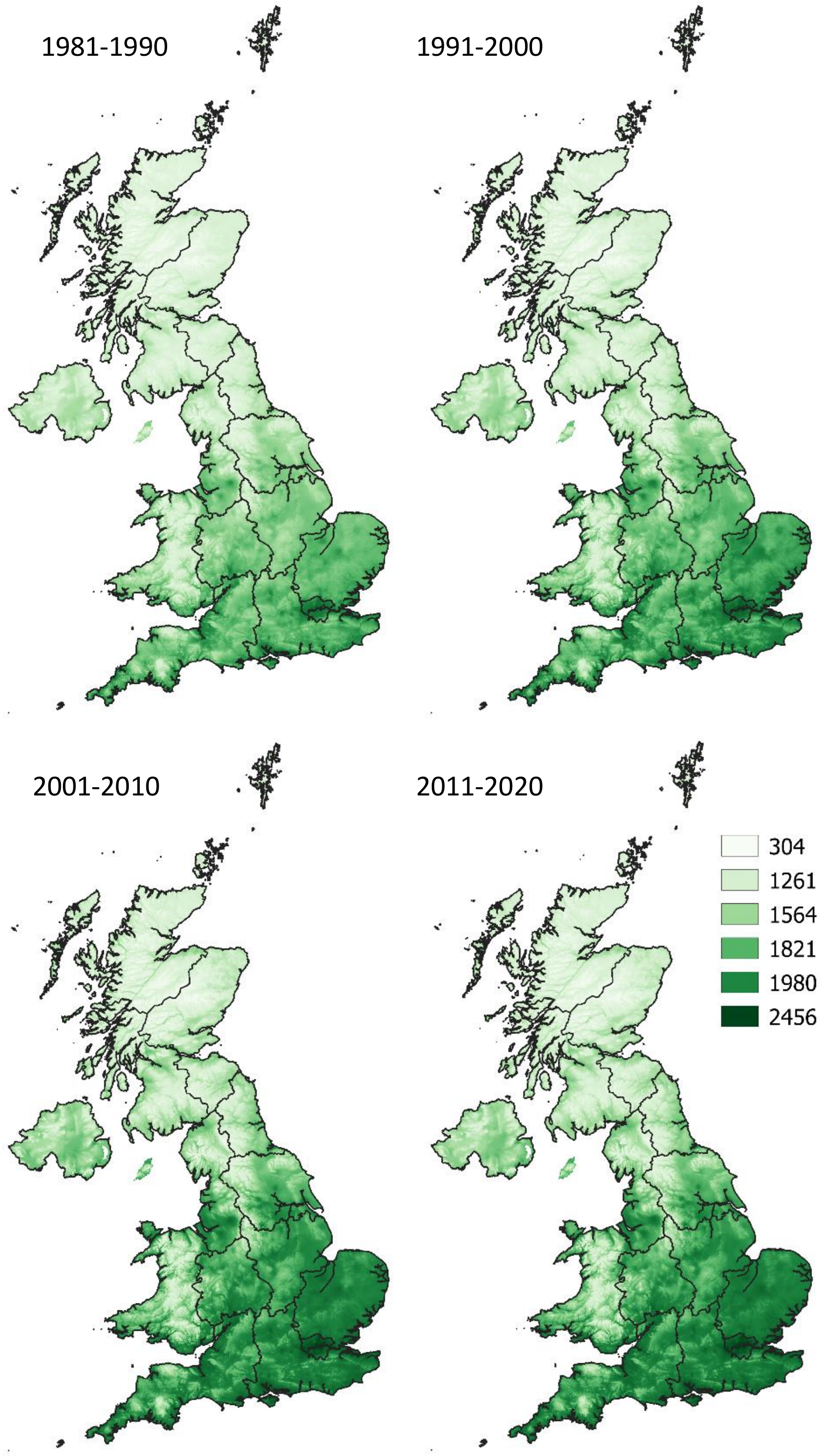



Figure 4.14*: Mean Growing Degree Days (GDD) ($^{o}$C days) from September to August for 1982-1990, 1991-2000, 2001-2010 and 2011-2020. Created using 1km x 1km gridded temperature data from HadUK* (Hollis *et al.*, 2019).

GDD was also calculated from drilling date to harvest date at winter wheat trial sites. The range in GDD values (785-1946 $^{o}$C days) calculated this way was smaller than those calculated from 1$^{st}$

September to 31st August (Figure 4.15), as drilling typically takes place a few weeks after 1st September from when the national values were calculated.

GDD at winter wheat trial sites significantly ($p<0.001$) changed from 1988-2018, increasing on average by 3 °C days every year (Figure 4.15). This could in part be due to the earlier drilling dates (Figure 4.11a.): drilling two days earlier increased GDD by 1 ($p<0.001$). However, GDD for winter wheat variety trials showed no significant relationship with yield.

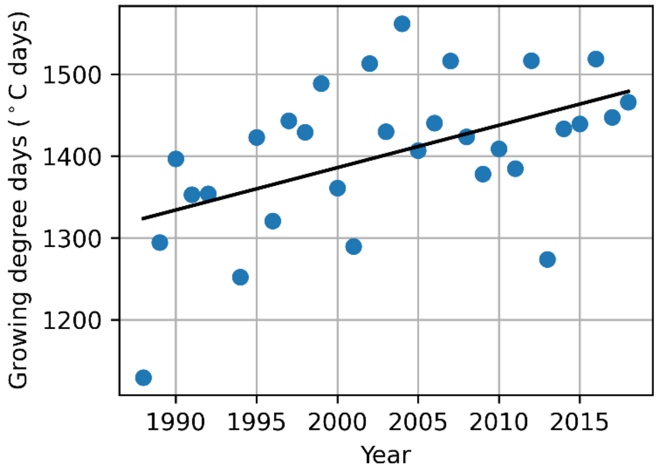



Figure 4.15*: Median Growing Degree Days (GDD) (°C days) at winter wheat trial sites from drilling date to harvest date for 1988-2018. The increase in growing degree days is significant ($p<0.05$) (black).*

### 4.4.4 Vernalisation Degree Days reveal coastal effect

Keeping in mind the vernalisation effectiveness factor (Figure 2.10), Vernalisation Degree Days (VDD) (Figure 4.16) showed very different spatial patterns to GDD (Figure 4.14). Across all decades, the South-West, coastal Wales and Northern Ireland had the highest VDD. Low VDD over much of Scotland across the four decades was due to $T_{mean}$ being suboptimal (optimal = 6.5°C). The effect of temperature moderation by the sea on coastal locations is evident around parts of the UK, where VDD was higher especially in the most recent decades, likely due to fewer extreme cold or warm days from autumn to spring. As temperatures continue to rise in the UK, it is likely that this trend in increase in VDD will peak and then reverse, as winter and early spring temperatures go above the optimal value for vernalisation. In the coldest parts of the country, such as northern Scotland, rising temperatures could be beneficial for vernalisation for longer.

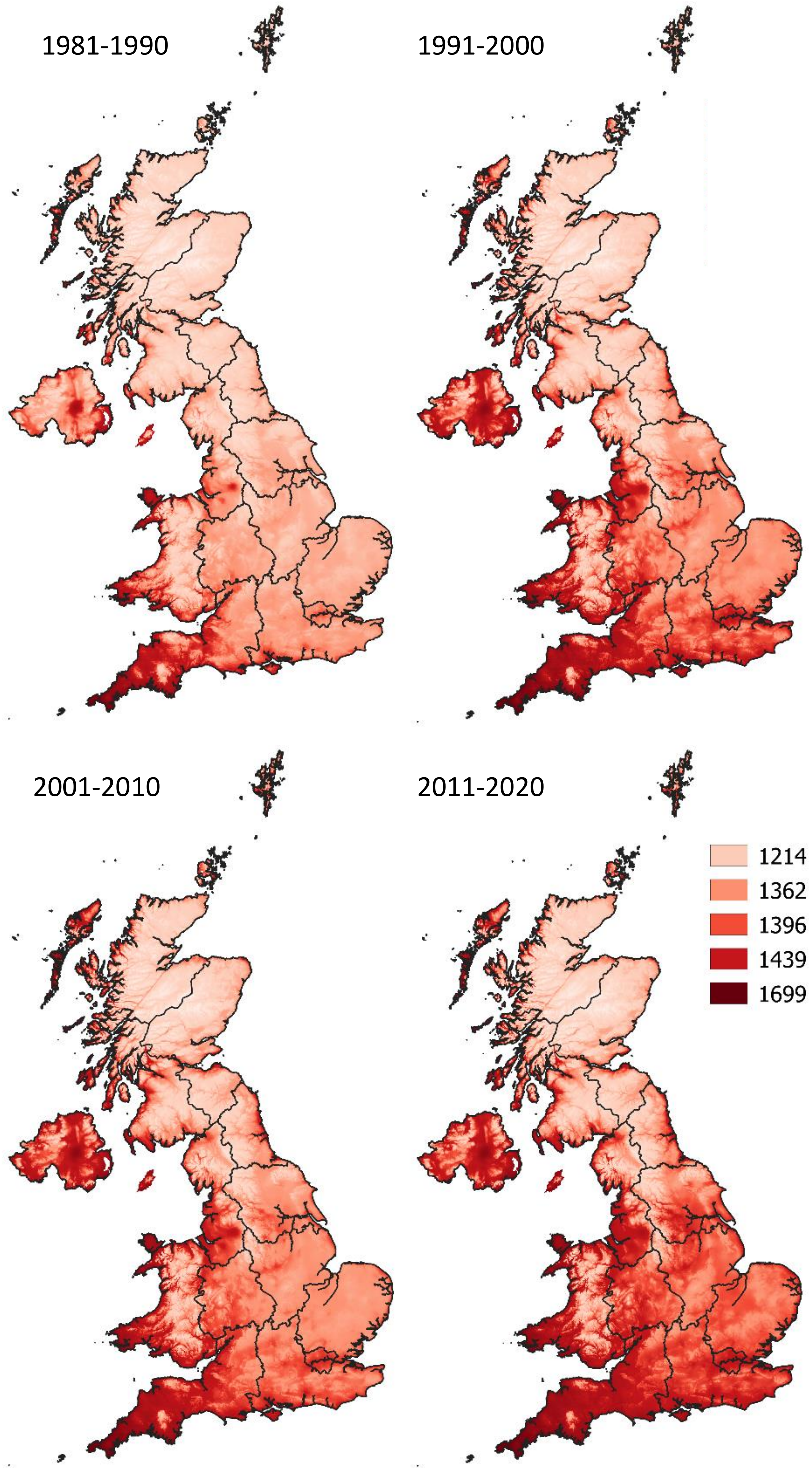



Figure 4.16: *Vernalisation Degree Days (VDD) (°C days) from September to April for 1982-1990, 1991-2000, 2001-2010 and 2011-2020. Created using equations [2.5] and [2.6] on the 1km x 1km gridded temperature data from HadUK* (Hollis *et al.*, 2019).

VDD was also calculated for each winter wheat trial, from drilling date to an estimated anthesis date. Across the period the VDD mean was 1502. Calculating anthesis based on a thermal time of 2100 °C days (AHDB Cereals & Oilseeds, 2018c) was flawed as for some sites and years the anthesis date was calculated to be in August (Figure 4.17), just days before being harvested: this growth timeline seems unlikely. The years with the most extreme estimated anthesis dates were 1988, 2001 and 2013, all of which had late drilling dates (Figure 4.11a.). Median harvest dates for these three years were at least a week later than the UK median day of year for 1988-2018 and all three years had negative yield anomalies (Figure 4.4a.)

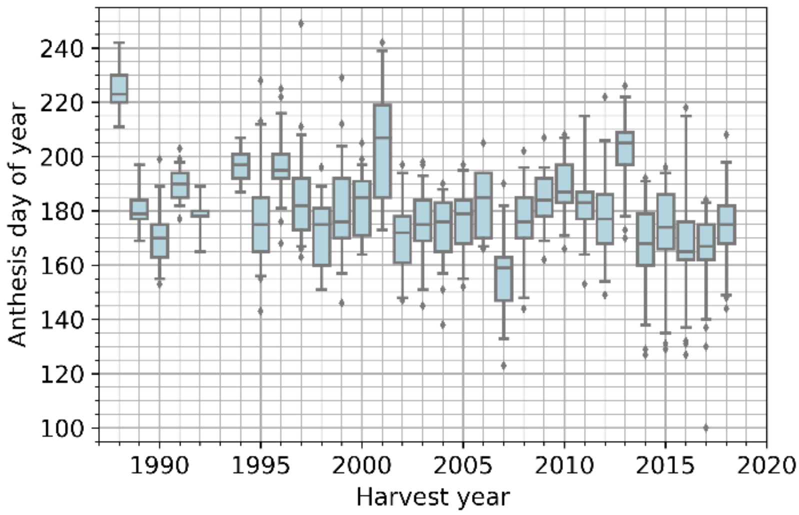



Figure 4.17*: Estimated anthesis day of year calculated using a thermal time of 2100 $^{o}$C days after drilling date for winter wheat trial sites, calculated using 1 km x 1km gridded HadUK temperature data* (Hollis *et al.*, 2019)*, extracted for each trial site.*

Estimated anthesis date (Figure 4.17) has been getting earlier at an approximate rate of a day every two years, which agrees with a previously calculated annual rate of change of flowering for UK winter wheat of -0.6 - -0.4 days/year for 1985-2014 (EEA, 2017). Anthesis date didn't have a significant relationship with yield, whilst VDD did, contributing 1.5 t/ha increase for every 1000 VDD, or 1.5 kg/ha/VDD. This lies within the range of values calculated by Wu *et al.* (2017) for Temperate Europe of 2.8 ± 1.5 kg/ha/VDD. Trial site VDD showed no statistically significant change over time at the trial sites, despite the observed earlier drilling dates. This could be influenced by the estimated anthesis date (Figure 4.17) used in the VDD calculation getting earlier.

The VDD metric calculated here and by Wu *et al.* (2017) encompasses a long period of the growing season (September to April). In the UK, vernalisation in winter wheat typically takes place in November and December (Steve Penfield, *pers. comm.*), hence the national VDD metric was recalculated for November to February (Figure A8). The effect of modifying this metric on the relationship of VDD with yield is explored in Chapter 6 (Section 6.4).

### 4.4.5 Widespread decrease in April air frost days

The number of air frost days in the UK in the spring months varied widely across the country, with the highest amount of frost in the North-West of the country and the least in the South-East (Figure 4.18). March typically had the most, with some parts of Scotland enduring on average over 20 frost days.

The overall trend in the number of air frost days varied each month. In March, there was an East-West divide, such that the West saw an increase in air frost days and the East a decrease. In April there was a widespread decrease in the number of air frost days, whilst in May there was very little change. This is significant for all growers, particularly those in the East of the country, who will likely have seen accelerated growth due to warmer temperature in March and April preventing frost formation but come May the risk of a frost day is still there, along with the potential of damage to sterility and abortion of formed grains around the earlier-occurring anthesis (Barlow *et al.*, 2015).

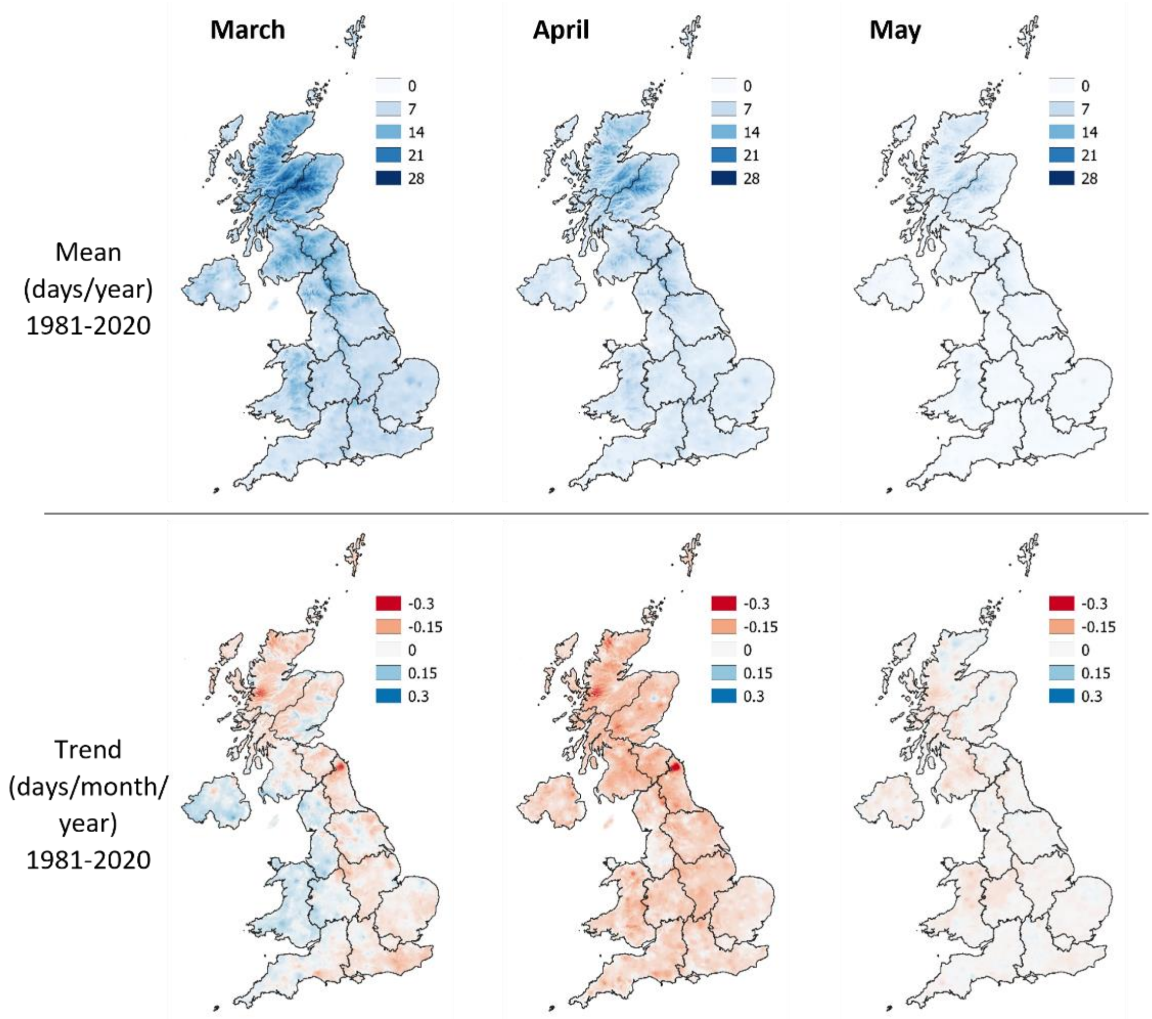



Figure 4.18*: The mean and rate of change in number of air frost days in the UK for 1981-2020 for March, April and May. Percentage change is calculated using the 1981-1985 and 2016-2020 periods. Created using 1km x 1km gridded HadUK temperature data* (Hollis *et al.*, 2019).

### 4.4.6 Seasonal variation in the water balance

The national water balance gives a good indication of the seasonal cycle in water availability and years of extremes (Figure 4.19). Peaks correspond to winter when there is a water surplus and troughs correspond to summer when there is frequently a water deficit. The summer of 1995 stands out as having the largest water deficit, likely due to national rainfall at less than 50% of the 1991-2020 average for June and August (Figure A4). This resulted in widespread drought across the country in the summer months, with an SPI-3 of <-2 i.e. “extremely dry” across most areas of England, Wales and Northern Ireland (Centre for Ecology and Hydrology, 2022). In 1995, wheat and barley yields were generally above average, indicating it wasn’t detrimental for these crops.

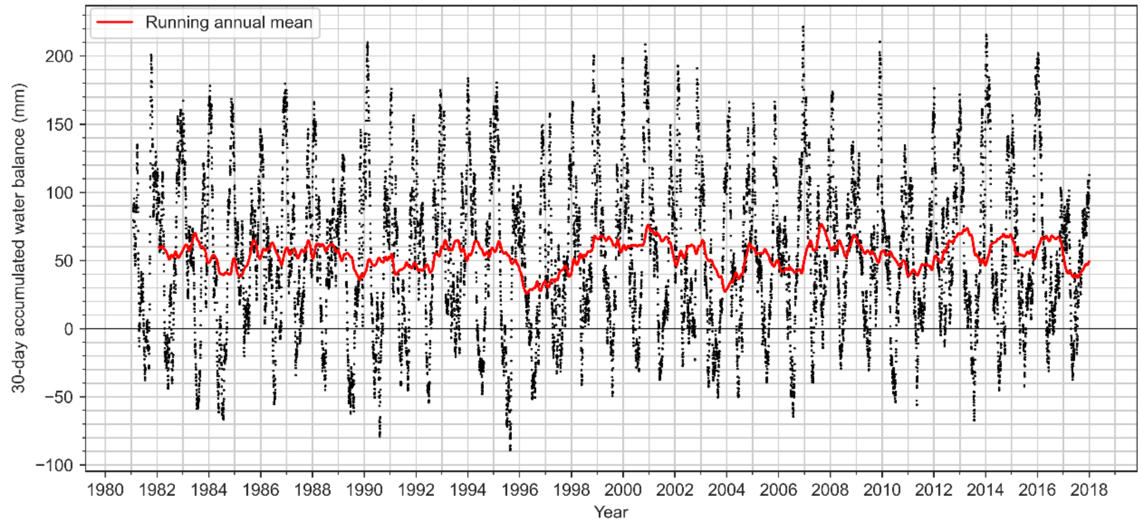



Figure 4.19*: Interannual variability in the 30-day accumulated national water balance (mm) for 1981-2017. The running annual mean is also shown (red). Created using CHESS-met and CHESS-PE* (Robinson *et al.*, 2020a, 2020b).

### 4.4.7 The distribution of wet days across the growing season has changed

Across the growing season (September to August) the number of heavy rain days (>10mm) has changed since 1981 (Figure 4.20). September saw widespread reduction of up to 50%, which could be beneficial for field operations, such as winter cereal drilling, at that time of year. This may partially explain the trend in earlier winter wheat drilling (Figure 4.11a.) February saw the biggest increase in heavy rain days, by over 500% for some parts of the East. This is followed by a reduction for much of the UK in March, when spring crops are often drilled and the first round of winter crop spraying takes place (T0). A very wet February may still delay field operations in March if the ground does become severely waterlogged (Berry and Brown, 2021).

In June and July, much of the country saw an increase in 10mm+ rain days (Figure 4.20) and 20mm+ rain days (Figure 4.21). This is potentially problematic, as this can extend periods of cloud cover, limiting photosynthesis during grain fill and damp, mild conditions with high relative humidity can also encourage diseases such as Fusarium ear blight to spread (Bayer, 2020). In August, increasing heavy rain days will have made it more difficult to time harvest, as the crop may have reached maturity, but waterlogged land can make it too difficult for field operations and it is costly to have to dry grain once harvested wet.

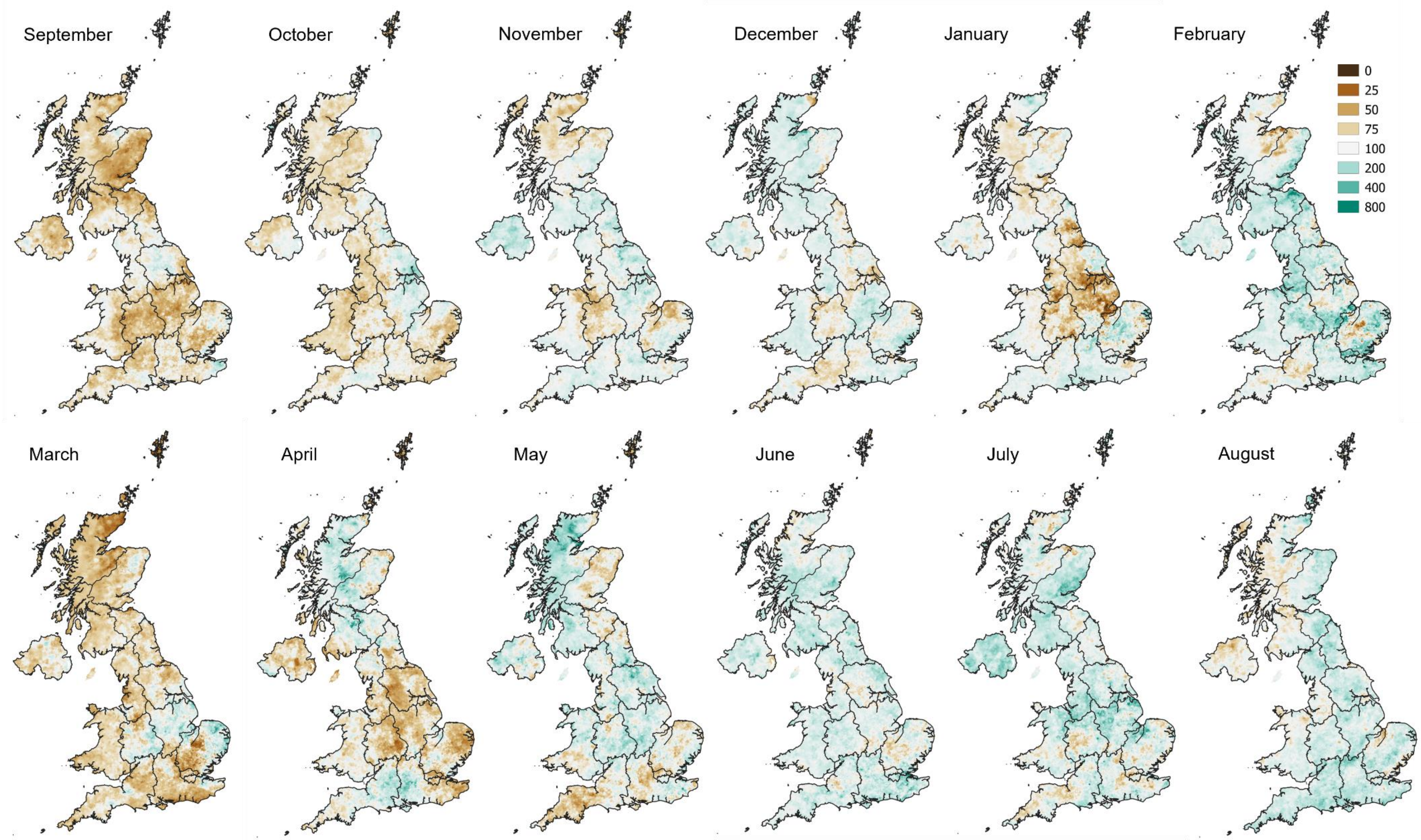



Figure 4.20*: The percentage change in the monthly number of heavy rain (>10mm) days in the 2011-2020 decade relative to the 1981-1990. An increase in heavy rain days is indicated in blue, a decrease in brown. Data from HadUK 1km x 1km gridded precipitation dataset* (Hollis *et al.*, 2019).

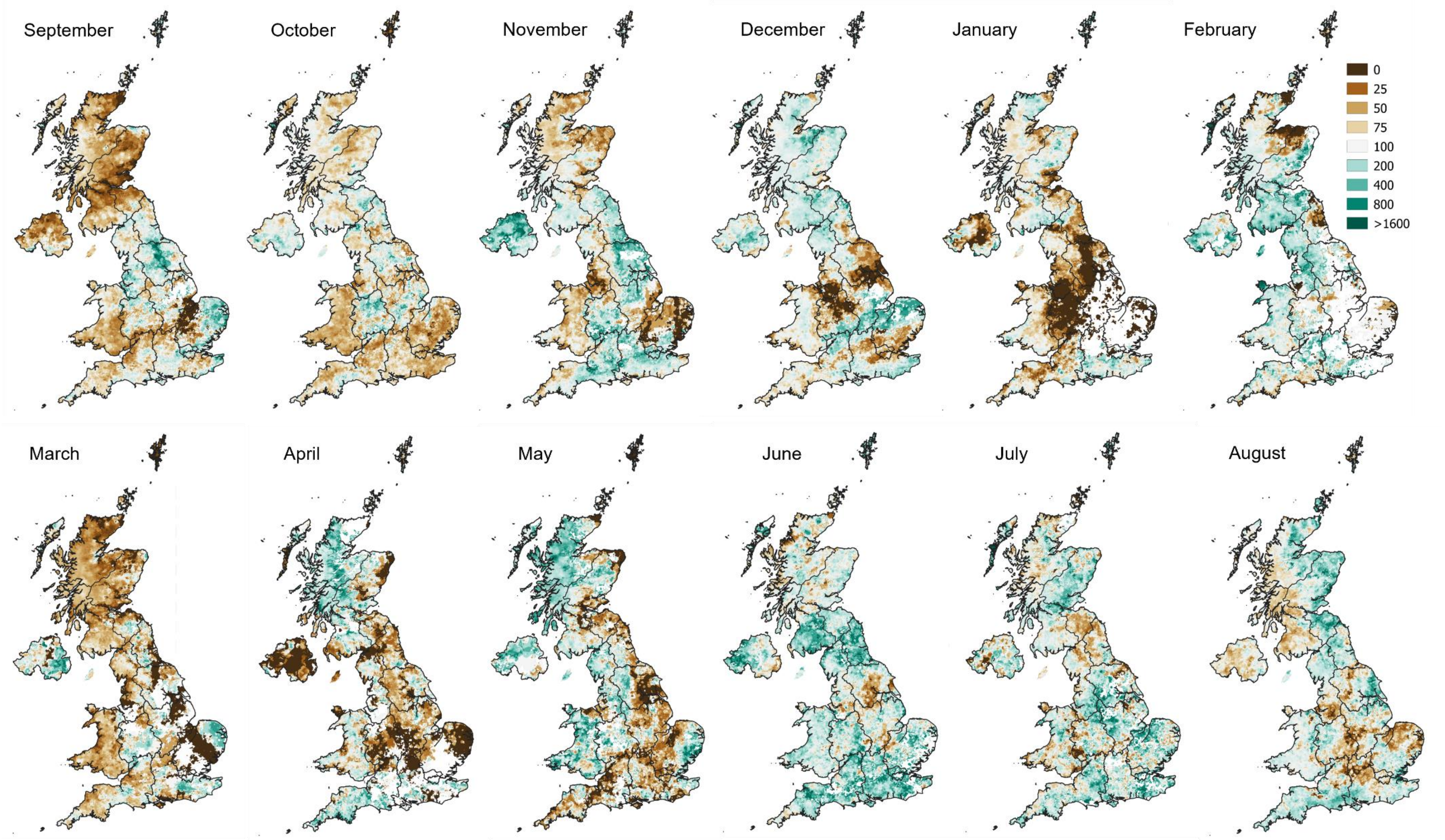



Figure 4.21*: The percentage change in the monthly number of heavy rain (>20mm) days in the 2011-2020 decade relative to the 1981-1990. An increase in heavy rain days is indicated in blue, a decrease in brown. Data from HadUK 1km x 1km gridded precipitation dataset* (Hollis *et al.*, 2019).

### 4.4.8 Heat stress during anthesis and grain fill is still rare

Across the 40-year period, extreme heat ($T_{max}$ > 32°C) was very rare. Looking at the trial sites, during the anthesis period $T_{max}$ > 32°C occurred only twice, one day in 1996 and one day in 2005 at a few locations in East Anglia and South-East. Evidently, extreme heat during anthesis has not been a major risk in the past. Rising temperatures could increase this risk, however model simulations of 2°C of warming indicate exceedance of this threshold in only two days in a decade in the warmest region (London) and no exceedance for 4°C of warming in the 2050s (2040-2059) and 2080s (2070-2089) (Jones *et al.*, 2020). This is due to the faster rate of thermal time accumulation due to warmer temperatures, such that heat stress is avoided by changes in crop phenology. In the UK, mean anthesis dates in 2050 are projected to be 10-11 days earlier using midrange emissions (RCP4.5) and 12-14 days earlier using high emissions (RCP8.5) (Harkness *et al.*, 2020). This shift in anthesis dates has already been seen in the trials data (Figure 4.17). In Germany, an earlier heading date and a shift of the flowering period to cooler spring has compensated for any increase in risk (Rezaei *et al.*, 2015).

Similarly, extreme heat during grain fill ($T_{max}$ > 35°C) only occurred on one day in 2006, 2015 and 2019 and two days in 2018, at a handful of trial sites in the East and South-East. Exceeding a $T_{max}$ of 35°C for at least 3 consecutive days has been shown to be detrimental to grain size and yield, therefore these events in the past 20 years are unlikely to have caused significant yield loss. The recent heatwave in July 2022, in which temperatures were above 35°C for two days in many areas of England, has shown the risk of extreme heat damage to grains during grain fill is possible in coming years.

In a national screening assessment on likelihood of exceedance of various critical temperature thresholds in the natural environment, Jones *et al.* (2020) found that that the grain-filling threshold is exceeded under a 4°C scenario in four of the major wheat producing regions in England. Climate data to run the national screening assessments were extracted from UKCP18 12 km resolution projections for a high emissions RCP8.5 pathway and used in the HADGEM Perturbed Physics Ensemble Model ID 7, with 2001-2010 as the baseline scenario. Threshold exceedance only occurs once or twice per decade per region, however the resulting economic losses in these events could be very high, with estimates of up to £19 million for a single event in the East Midlands, and £12 million in South East England.

The frequency of mild heat stress during grain fill ($T_{max}$ > 31°C) was much higher (Figure 4.22). An anomalous year within the period 1981-2020 was 2006, which had the hottest July on record and saw the grain fill threshold exceeded 10 times at several locations in Cambridgeshire. Whilst less detrimental than extreme heat, mild heat stress has been shown to affect grain size and yields (Dreccer *et al.*, 2018). 2006 had a positive yield anomaly across most regions, including East Anglia and the South-East (Figure 4.5), suggesting that whilst it could have been yield limiting, several days of mild heat stress did not ruin the crop.

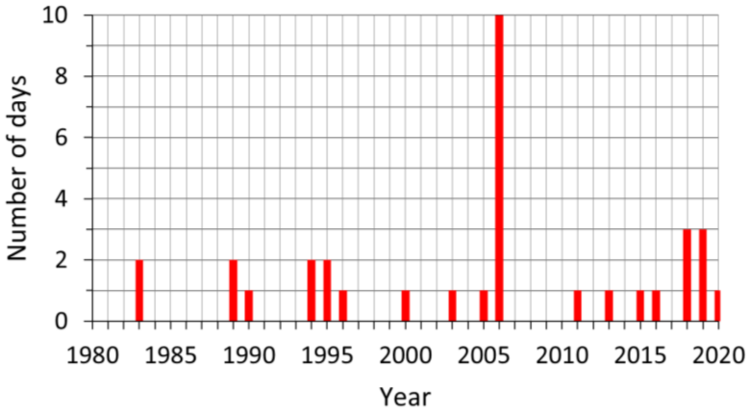



Figure 4.22*: The number of mild heat stress ($T_{max}$ > 31°C) days during grain fill for 1981-2020 at (52.22,0.1) near a trial site in Cambridge. Calculated using gridded 1 km x 1km temperature data from HadUK* (Hollis *et al.*, 2019).

### 4.4.9 Total solar radiation received during grain fill has increased in the East

For the period 1987-2020, solar radiation received during grain fill was largely dependent on the latitude. The south of the UK received the most solar radiation during grain fill with over 800 MJ/m$^2$ on average received in the South-West England, South-East England and East Anglia (Figure 4.23). In Scotland, particularly in the north, crops received over 25% less grain fill SIS. This pattern of greater surface solar radiation in the South of the UK is not only seen during grain fill but also across the year (Pfeifroth, Sanchez-Lorenzo, *et al.*, 2018).

The change in grain fill SIS across the UK is highly dependent on longitude. Regions in the East saw increases in grain fill SIS of 5-13%, whereas those in the West saw decreases of up to 15%. A reduction in total summer surface solar radiation in the South West, Wales and Scotland was also observed by Pfeifroth *et al.* (2018a). For reference, annual grain fill SIS anomalies relative to the 1991-2020 average are in Figure A9. Analysis of trends in grain fill SIS at a national level indicates no change over the period, when evidently at a regional level there are significant and important changes for growers to consider (Figure 4.24).

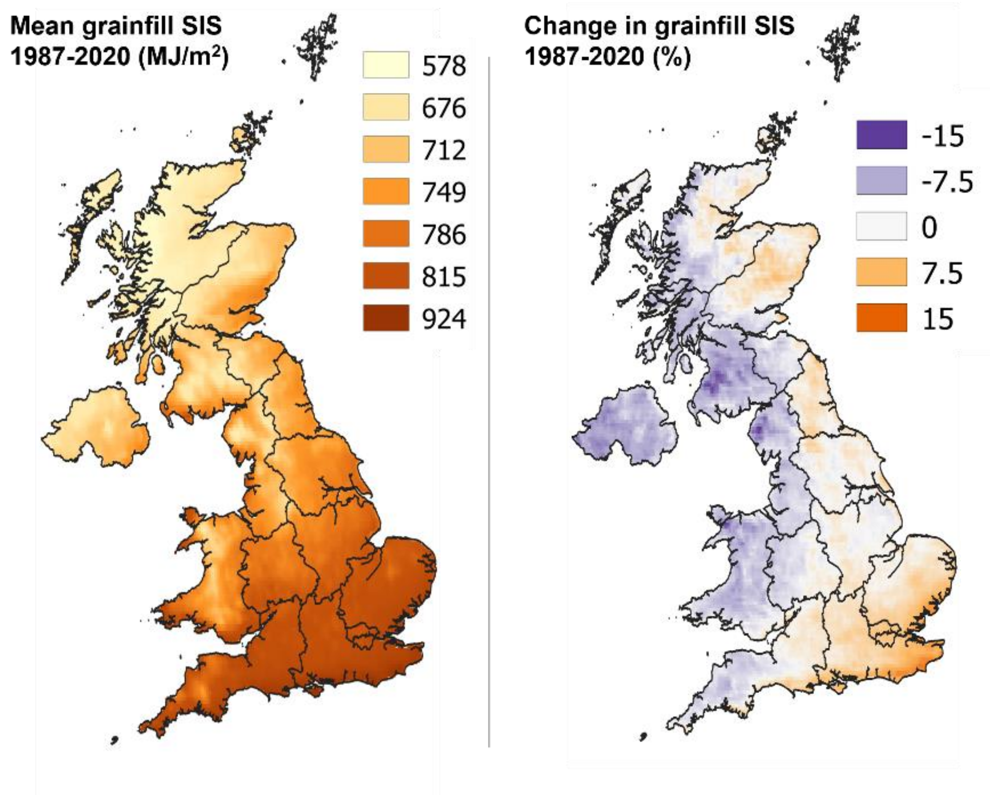
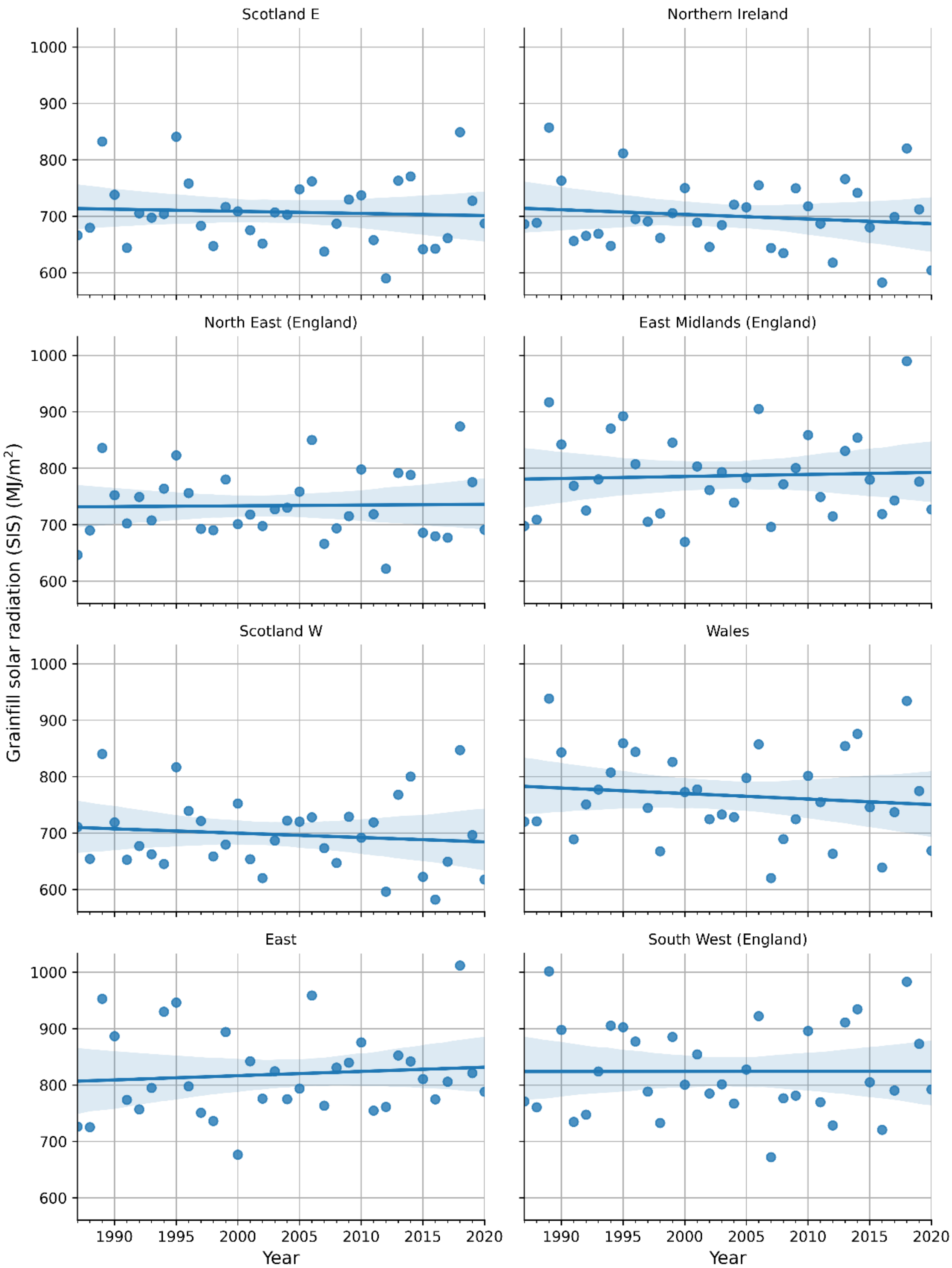



Figure 4.23*: Mean total grain fill surface incoming solar (SIS) radiation (MJ/m$^2$) (left) and percentage change in grain fill SIS from 1987-1991 to 2016-2020 (right). Grain fill incorporates the period 16$^{th}$ June-31$^{st}$ July. Created using gridded 0.05$^o$ x 0.05$^o$ degrees CMSAF-SIS data* (Pfeifroth, Trentmann, *et al.*, 2018).

Regression analysis between regional grain fill SIS and regional wheat yields showed no significant relationship between the two variables. However, when site specific grain fill SIS data was paired with winter wheat variety trial yield data (Figure 4.25), significant positive correlations exist between grain fill SIS and winter wheat yield in the North-East, Northern Ireland, Scotland East, South-East and Wales (Table 4.1), highlighting the value of localised climate information.


Scotland E
Northern Ireland
North East (England)
East Midlands (England)
Scotland W
Wales
East
South West (England)
Grainfill solar radiation (SIS) (MJ/m²)
1000
900
800
700
600
1990
1995
2000
2005
2010
2015
2020
Year

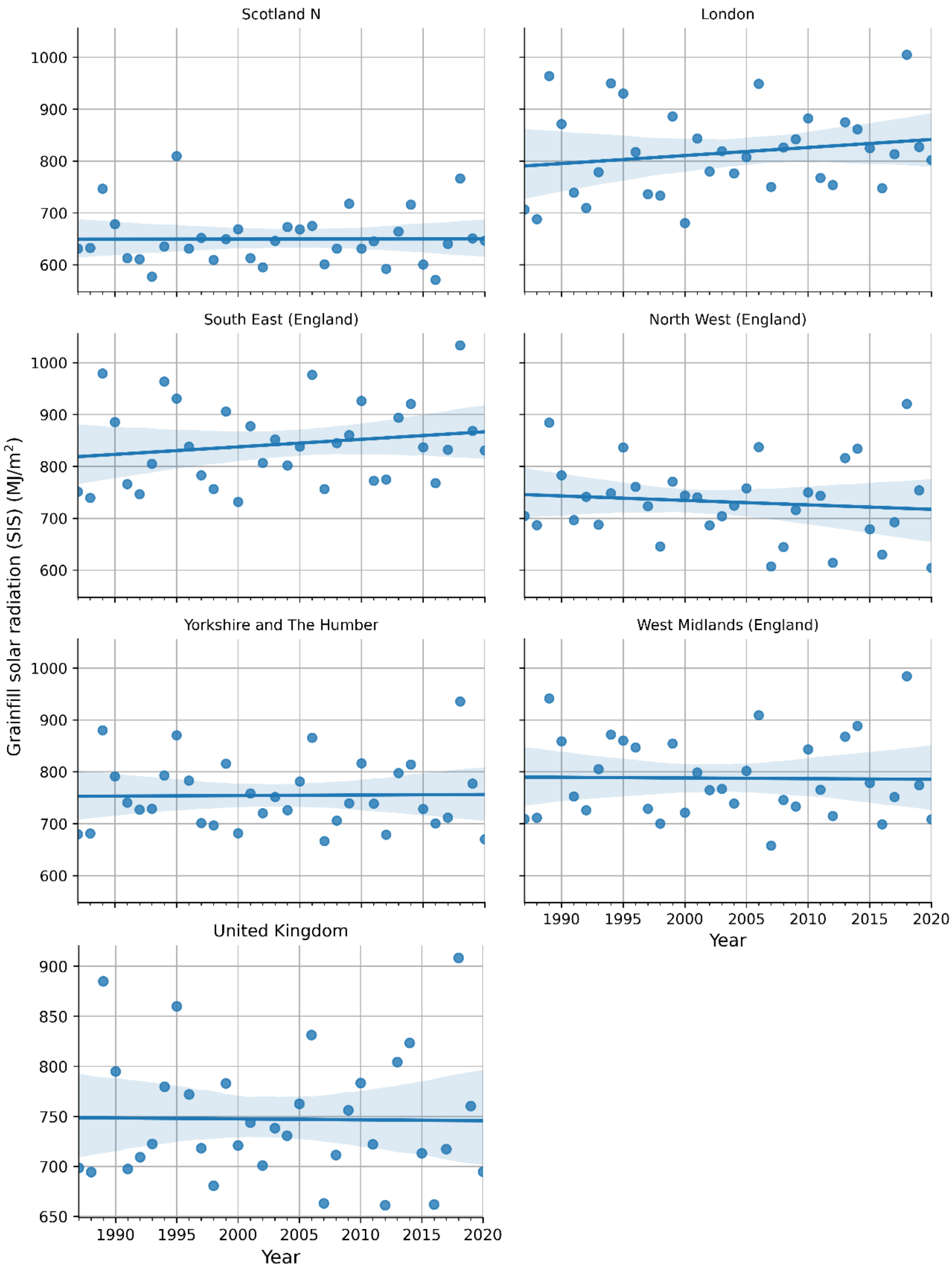



Figure 4.24: *Regional and national trends and variability in total surface incoming solar radiation (SIS) (MJ/m2) for the grain fill period 16th June-31st July. Calculated using gridded 0.05° x 0.05° degrees CMSAF-SIS data (Pfeifroth, Trentmann, et al., 2018).*

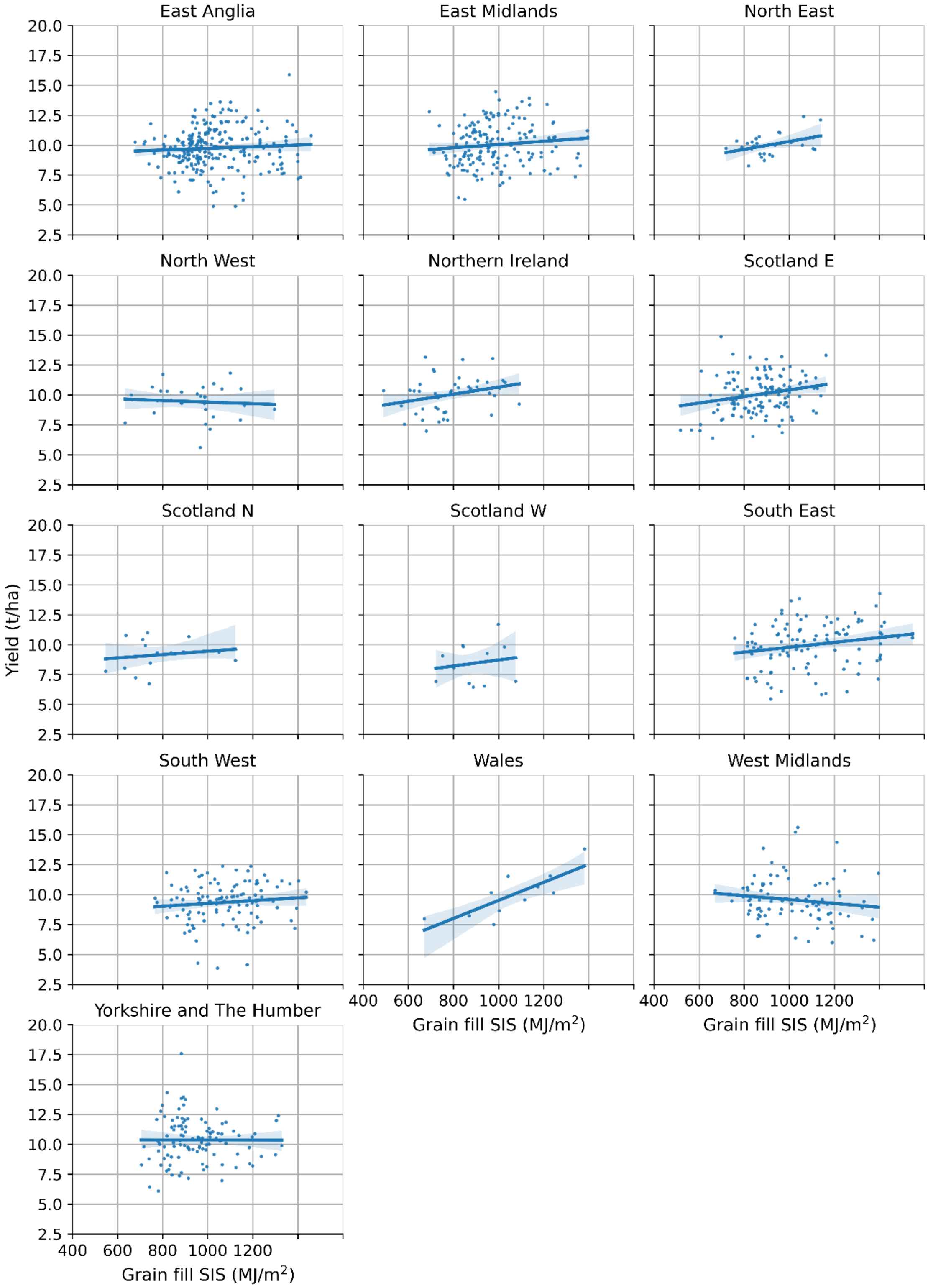



Figure 4.25: *Total grain fill surface incoming solar radiation (SIS) (MJ/$m^2$) and winter wheat variety trial yield (t/ha) averaged across each trial site each year for each region of the UK. The grain fill period encompasses 16$^{th}$ June-31$^{st}$ July. Total SIS was calculated using gridded 0.05$^o$ x 0.05$^o$ degrees CMSAF-SIS data (Pfeifroth, Trentmann, et al., 2018).*

| Region | coef | p |
|---|---|---|
| East Anglia | 0.067 | 0.3 |
| East Midlands | 0.11 | 0.1 |
| **North East** | **0.41** | **0.03** |
| North West | -0.079 | 0.7 |
| **Northern Ireland** | **0.29** | **0.05** |
| **Scotland E** | **0.22** | **0.01** |
| Scotland N | 0.17 | 0.5 |
| Scotland W | 0.16 | 0.6 |
| **South East** | **0.20** | **0.04** |
| South West | 0.11 | 0.3 |
| **Wales** | **0.79** | **0.004** |
| West Midlands | -0.14 | 0.2 |
| Yorkshire and The Humber | -0.004 | 1.0 |

Table 4.1: *Pearson correlation coefficient between grain fill (16$^{th}$ June-31$^{st}$ July) surface incoming solar radiation (SIS) and trial yield within each region is shown below, with significant correlations (p<0.05) in* ***bold.***

### 4.4.10 Disease prevalence shows high interannual variability

Disease prevelance is highly dependent on climatic conditions. National disease survey data shows that there is large interannual variability in foliar disease prevalence in the flag and second leaf of the winter wheat crop (Polley and Thomas, 1991; Hardwick *et al.*, 2001; Turner *et al.*, 2021) (Figure 4.26).

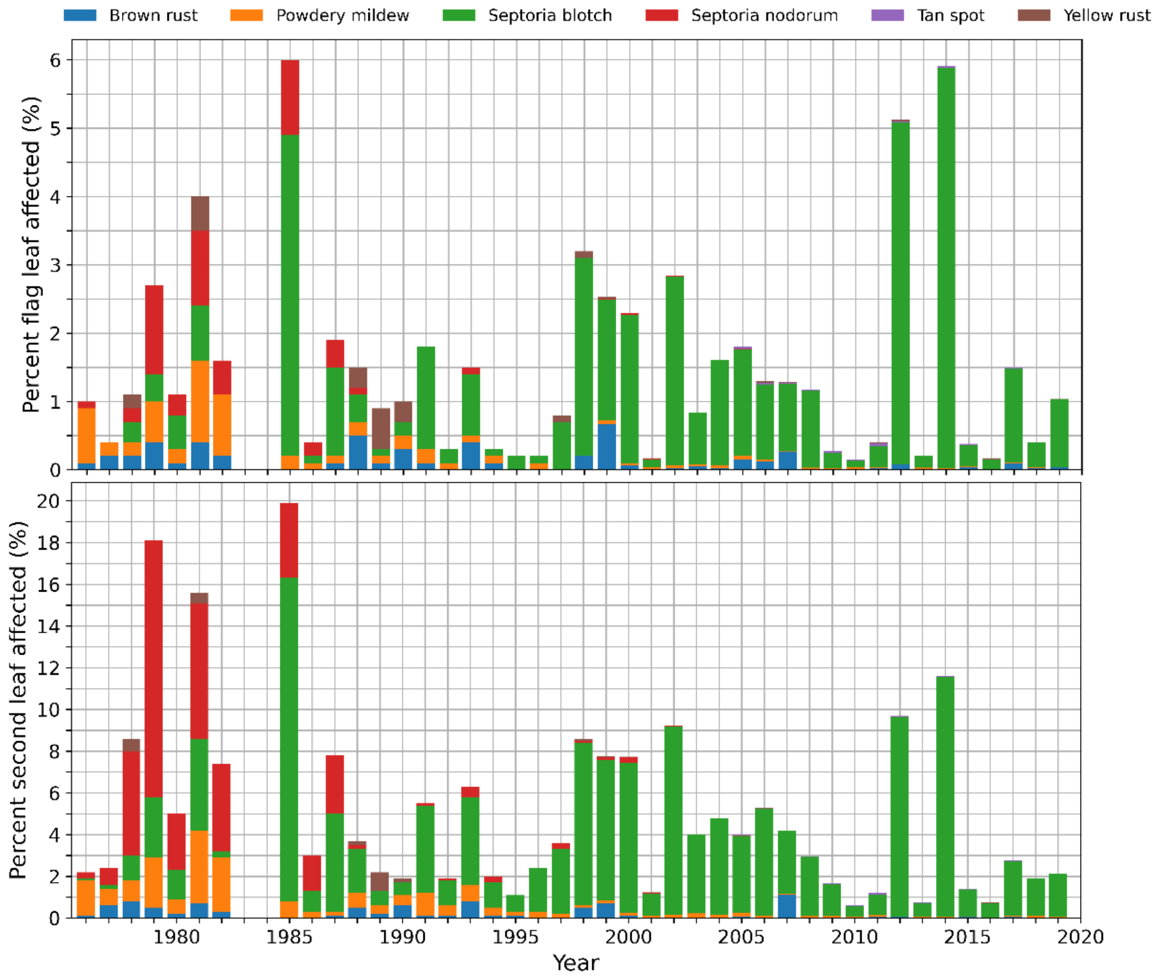



Figure 4.26*: National Flag (top) and second leaf (bottom) prevalence (area of leaf affected, %) of foliar diseases of winter wheat on-farm for 1976-2019. No disease data was recorded for harvest years 1983 and 1984, or for tan spot pre-1999. Data from Hardwick* et al. *(2001), Polley & Thomas (1991) and Turner* et al. *(2021).*

From 1976-1982, Septoria nodorum blotch was the most widespread disease recorded in the disease survey (Figure 4.26) (Polley and Thomas, 1991). Septoria leaf blotch was the most widespread foliar disease in most years since 1985, although it is more severe in the west and south-east and least in the East. Favourable conditions are warmer temperatures in May and June, the occurrence of which has increased, and rain-splash events (Turner *et al.*, 2021), which have also increased in June (Figures 4.20 and 4.21). The risk of incidence of this disease will likely increase as the climate changes.

Combining the national anomaly temperature and rainfall time-series graphs (Figures A4 and A6) with national disease prevalence statistics for highly affected years 1985, 2012 and 2014, it is evident that there was no relationship between national Septoria leaf blotch prevalence and national temperature and rainfall data for May and June. As reported in Turner *et al.* (2021), this is likely due to localised rainfall events being more important here which national and regional climate analysis will mask. This is exemplified by looking at changes in 10 mm and 20 mm rainfall events in May (Figures 4.20 and 26), where neighbouring locations within the same region and even county show opposing trends over time.

There was also no significant relationship between Septoria leaf blotch flag leaf prevalence and national yields (p=0.98). Given the regional dependence of the severity of the disease, the relationship between regional yields and national disease prevalence was also investigated. Whilst no significant relationships were found, there were weak associations between yield and Septoria leaf blotch flag leaf prevalence in the higher severity regions of the South-West and Wales, but not in the South-East where it is also more severe (Figure A10).

The trend in earlier drilling (Figure 4.11a.) has increased the risk of infection, as there is shorter time between harvest and drilling so more disease life cycles can carry over to the next growing season (Polley and Thomas, 1991; Gladders *et al.*, 2001).

Fusarium ear blight is an emerging threat to grain quality (AHDB Cereals & Oilseeds, 2018b; Turner *et al.*, 2021). There have been several epidemics in the last two decades, such as 2007 and 2014, and the incidence of the fungus has increased dramatically (Turner *et al.*, 2021). Risk factors include warm dry springs allowing spore production followed by rain-splash events in June which spread the spores onto ears. The fungus favours warm humid conditions during flowering and high summer humidity and/or rainfall can allow the infection to spread (Bayer, 2020). As such, rising summer temperatures and increase heavy rainfall days in the summer months (Figures 4.20 and 4.21), combined with the movement towards minimum tillage, may well encourage this disease in future.

## 4.5 Discussion

Increasing variability in wheat and barley yields has made it difficult to quantify the trend in recent yields after the documented yield 'plateau' of the 1990s and 2000s (Figures 4.1 and 4.2). Yield potential has increased, with the highest yields in the period of analysis in 2015, however the gap in yields between good and bad years has also increased, such that on average both barley and wheat yields have continued to stagnate and the yield plateau still exists. This contrasts with the decrease in yield variability seen in Finland since the 1990s (Peltonen-Sainio, Jauhiainen and Hakala, 2009). This lack of yield stability increases the risk of income fluctuations for farmers, who are also vulnerable to fluctuations in world prices, potentially inducing financial hardships. Increasing the stability of UK cereal yields is highly important.

The plateau in national yields (Figures 4.1 and 4.2) contrasts the continuous yield increases seen in the variety trials data (Figure 4.3), indicating that the plateau is not due to limited genetic improvement. A similar result was also seen in France in the 2000s (Brisson *et al.*, 2010). In the USA, farmers have shown a reluctance to change their crop mix or agricultural practices in response to rising temperatures (Burke and Emerick, 2016). Therefore, it is possible that in the UK there is also a reluctance to adopt newer varieties that are better suited to the warmer, more variable climate, hence the observed plateau and greater variability in on-farm yields. The contribution of variety improvement to national yields is explored further in Chapter 5.

Identification of yield anomalies and analysis of national climate time-series provided insight into the climate drivers of this variability. In some years, such as 2012 and 1984, the climatic causes of low and high yields respectively, were evident by looking at monthly temperature, rainfall and sunshine duration data. However, in other years (e.g. 2019), there were no clear explanations, suggesting that national monthly data is not sufficient to explain yield anomalies and/or these yield anomalies have non-climatic causes. Furthermore, this analysis highlighted that the climatic influences on yield are multi-faceted and favourable weather conditions at certain times of the year, such as high sunshine duration during June and July, do not always mean high yields, as the weather in the rest of the growing season may have been poor. Growing season (September to August) anomaly graphs for each harvest year provided a useful view of the growing season as a whole for the UK and individual nations (Figures 4.9 and A4-6).

There have been significant changes in agronomy in the past 40 years, including an increase in minimum till and changes in cropping to more OSR and less barley. In addition to the reduction in heavy rain in September (Figures 4.20 and 4.21), these agronomic changes will have influenced the trend towards earlier drilling date (Figure 4.11a.). Earlier sowing is associated with higher GDD, however potential yield benefits have to be balanced with the increased risk of blackgrass and other weed pressure that come from a reduction in tilling.

### 4.5.1 The value of agroclimate metrics

The creation and analysis of various agroclimate metrics has been useful for quantifying changes in the climate that affect agriculture over the past 40 years, as well as variability within the period. However, some metrics could be modified in future to increase their potential usefulness.

Decadal analysis showed that the trend in SOGS to earlier in the year was not linear (Figure 4.12). 2001-2010 was seemingly colder, at least at the start of the year, than the previous decade. From 2011-2020 SOGS in a few areas of the South was the 6$^{th}$ January, just a day after the earliest possible SOGS. Long-term projections of this metric show that it will continue to get earlier at a national rate of ~5-10 days each decade (https://uk-cri.org/), which may well result in a saturation of the SOGS threshold, especially in the South. It also raises the question as to whether the metric could be calculated differently to incorporate earlier winter months, as it is possible the SOGS threshold is reached before this. This may also lead to a significant relationship between SOGS and harvest date. Furthermore, it is very possible that after the SOGS threshold has been met, a cold spell follows essentially pausing the growing season once again.

National analysis of GDD showed that there is an increasing availability of GDD in the September-August period (Figure 4.14). Combined with earlier drilling, this means that winter wheat trials had 3 $^{o}$C days more per year from 1988-2018 (Figure 4.15). Increase in GDD early on in the growing season due to changes in drilling date and higher temperatures likely contributed to the changes in estimated anthesis date (Figure 4.17). This indicates farmers are making adaptations to cope with the changing climate, which should reduce the risk of extreme heat events, like that of July 2022, in future decades. The risk of heat stress during grain fill has thus far been very small in the UK, however expected increases in the exceedance of the threshold in a 4$^{o}$C-warmer world could make it an emerging threat (Jones *et al.*, 2020) and a common occurrence by the end of the century (Arnell *et al.*, 2021). From a UK cereals perspective it is important that warming is limited to 1.5$^{o}$C or at most 2$^{o}$C. Breeding cultivars that tolerate heat stress is a clear priority. There is also the opportunity to share cultivars and knowledge with wheat-growing countries that already have a warmer climate, such as southern France and Australia.

Increasing temperatures also have the potential to affect the vernalisation requirement of winter crops. At a national level, VDD was highest in the South-West, Wales, Northern Ireland and coastal areas and these areas also saw the biggest increases in available VDD (Figure 4.16) indicating an increase in optimal vernalisation temperatures. By contrast, Cho *et al.* (2012) showed that by the 2080s, warmer winter temperatures are projected to cause a reduction in optimal vernalisation temperatures in the South West. At trial sites, VDD showed no change over time.

The observed positive relationship between VDD and yield agreed with Wu *et al.* (2017), however it is not clear how useful this metric is. Whilst it is known that chilling is very important for winter crops, the amount of time crops require in these temperatures is not well understood. Different varieties also have different vernalisation requirements (number of VDD), therefore using one metric for all varieties is somewhat oversimplified. Given vernalisation in winter wheat typically takes place in November and December in the UK (Steve Penfield, *pers. comm.*), it is possible the modification to the time period for VDD November-February (Figure A8) is a better representation of the available VDD to winter crops than that used initially and by Wu *et al.* (2017). Evidently, there is an urgent need for crop modellers and physiologists to work together to define more suitable vernalisation metrics. Further investigations could also look at changing the upper temperature limit $T_{amp}$ on VDD, and the change in the number of non-vernalisation days.

The longitudinal dependence of change in grain fill SIS (Figure 4.23) was masked by an insignificant change nationally (Figure 4.24). This change is important for growers. Those in the South-East and East can be reassured that they will, on average, receive ample solar radiation in June and July for cereals and soft fruits, but those in the West may want to consider crops less dependent on solar radiation at this important time. There was a significant positive correlation between grain fill SIS and winter wheat yield for some regions in the UK when using site specific data, which was masked when using regional yield and solar radiation data. This relationship is explored further through statistical modelling in Chapter 6.

All of the agroclimate metrics can be modified as required for other crops and specific periods of interest. One limitation here is the use of static dates that in the national analysis, when year-to-year drilling dates, and subsequently growth periods, can vary by a few weeks. 2022 had record-breaking early harvests, with Suffolk farmers beginning as early as 29$^{th}$ June (Henderson, 2022), suggesting a shift in the growing season with a much earlier anthesis and grain fill period than used in this analysis. Drilling and harvest dates from the trials data were used when possible, however more information of trial and on-farm crop growth stages would be useful to increase the explanatory power of the agroclimate metrics for yield.

A major limitation with the winter wheat disease data used in Section 4.4.10 (Figure 4.26), taken from the annual Cereal Disease Survey, is that samples of the crops are collected too late for yellow rust (James Brown, *pers. comm.*). This is reflected in the data for 2014, when there were "trace" amounts of yellow rust, but this was in fact the worst yellow rust year for about 20 years. Hence, the relative incidence of wheat diseases is not very accurate in Figure 4.26. Other more accurate disease data, collected across the growing season, would be useful here to gain a better understanding of what disease severity was each year.

An additional metric that would be valuable here is the change in the number of days suitable for spraying. This is a vital part of crop management and allows the application of fungicides and pesticides to the crop, therefore understanding if the distribution of dry hours/days, with wind speeds low enough for this activity, is changing is important.

### 4.5.2 Creating an agroclimate tool for growers

There is great potential to incorporate greater climate data into the UK breeding process, as well as make agroclimate metrics relevant to agronomists and growers. The current AHDB Variety Selection Tool (https://ahdb.org.uk/variety-selection-tool) has successfully incorporated disease resistance ratings, agronomic features and market options to help growers identify the best variety for their farm, however there is no inclusion of climate data. Variety trials involve testing the varieties across multiple environments, including climates, due to the importance of genotype-by-environment interactions in determining yield. Observed performance of varieties in these different environments can be used to help recommend growing climates for each variety, such that if a grower inputs their location into the tool, historical climate records will indicate which varieties may be most climate-smart at their farm location, in addition to the disease resistance ratings and agronomic factors.

In addition to a variety recommendation tool that accounts for climate, it is also hoped that the *State of the UK Agroclimate* can be regularly updated and that an agroclimate tool can be created which allows farmers and growers to see at their farm how the agroclimate has varied from year to year and how it is changing to inform future management decisions. This can complement the data on projected changes in Climate Risk Indicators at uk-cri.org.

## 4.6 Conclusion

This analysis has allowed the identification of significant changes over the past 40 years, as well as variability within the period. Combining these agroclimate metrics with yield data has given an indication of whether they may have influenced cereal yields. Specifically, this research has found:

- On-farm yield potential is increasing for both wheat and barley, but due to increasing yield variability, actual farm yields are still plateauing
- Variety trial yields have increased linearly, suggesting the on-farm yield plateau is due to agronomic or management factors, rather than a limit of genetic improvement
- The trend in earlier drilling dates has started to reverse in the last few years
- Changes in heavy rain days across the growing season could make it easier for farmers to get on the land in September but more difficult to harvest in August
- Heat stress during grain fill has not yet been a problem but is an emerging threat
- Solar radiation during grain fill has increased in the East and decreased in the West

There is great potential to use these metrics and other climate information in breeding programmes, in making crop and variety choices and management decisions. To understand the relative roles of the agroclimate metrics, statistical modelling is required to incorporate multiple agroclimate covariates in the same model, and to account for other factors such as changes and variability in genotype and agronomy. This is explored in Chapter 6.

# 5 Quantifying genetic drivers of yield variability of UK cereal crops

Multi-environment trials allow investigation of varietal yield performance across a range of locations and years (Smith *et al.*, 2005). In the UK, the National List/Recommended List (NL/RL) variety trials allows breeders to test new varieties in different growing environments, inducing a genotype-by-environment interaction (GxE) as the crop responds to the changes in its environment. To dissect genetic and non-genetic sources of yield variability and quantify the contribution of breeding to yield increases (i.e. the genetic gain), linear mixed models have been used (Mackay et al. 2011; Piepho et al. 2014). Using UK NL/RL winter cereal trials, Mackay *et al.* (2011) showed that from 1982-2007, 88% of the improvement in winter wheat (*Triticum aestivum* L.) and winter barley (*Hordeum vulgare* L.) yield was attributable to genetic improvement, indicating crop breeding in the UK has been fundamental to increasing the maximum attainable yields. Genetic gain estimates are therefore a valuable measure of success of a breeding programme (Covarrubias-Pazaran, 2020; Covarrubias-Pazaran *et al.*, 2022) and contribute to funding decisions. However, the effectiveness and accuracy of genetic gain estimated from these programmes is not well known.

In this Chapter, the extent to which crop breeding is contributing to the stagnating wheat and barley yields seen in Chapter 4 (Figures 4.1 and 4.2) is quantified to update analysis by Mackay *et al.* (2011). The effect of variety age on variety trial yields is modelled to understand how long-term trends in individual variety performance can contribute to estimates of genetic gain. By subsetting NL/RL treated variety trials data into case study periods, introducing an upper limit on the number of years varieties are present for and varying the number of long-term varieties (checks), the strength of using genetic gain as a measure of success of breeding programmes is also investigated.

## 5.1 Crop breeding has continued to contribute to cereal yield increases

Winter wheat, winter barley and spring barley all had significant long-term increases in median treated and untreated yields (Figure 5.1). Treated trials are those that have received full fungicide treatment whilst untreated receive none. Median treated and untreated variety yields for spring barley had the smallest interannual variability (~2 t/ha), whilst winter wheat yields varied by up to 4 t/ha between 2012 and 2015, indicating lower stability.

Across all time periods, treatments and crops, the contribution of the variety effects to linear increases in yield was positive, indicating breeding has had positive contributions to yield increases (Table 5.1, Figure 5.2). In the spring barley and winter wheat data treated and untreated plots there are no variety effect data points for 2011 (Figure 5.2a,c). In the original dataset, varieties introduced this year were only present for just one or two years and/or had just one or two sites per year and were therefore removed prior to analysis.

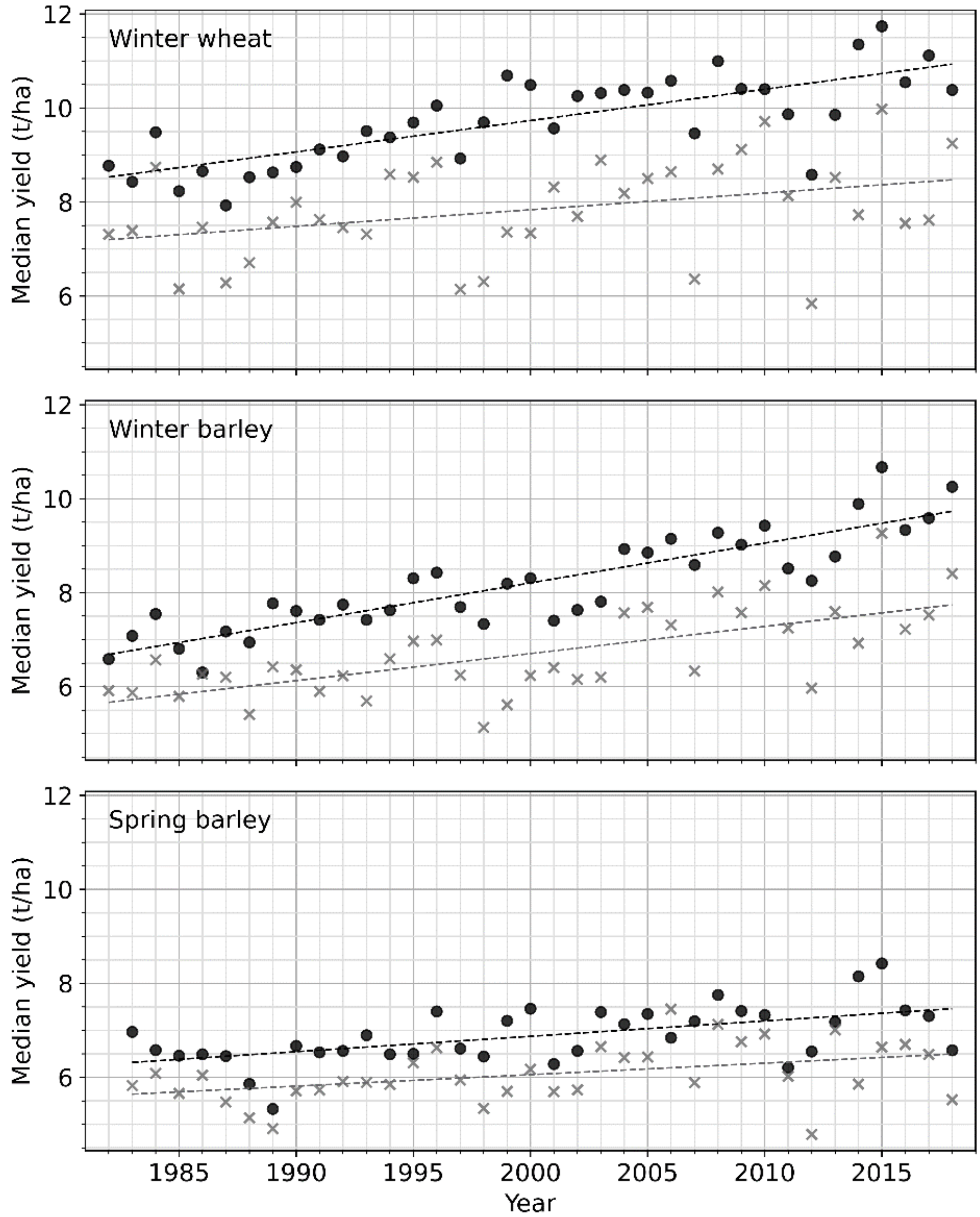



Figure 5.1: *Median winter wheat, winter barley and spring barley for fungicide treated (o) and fungicide untreated (x) trial yields for 1982-2018 harvest years. The linear increase in median yield is significant (p<0.05) for all 3 crops.*

| Years | Treatment | Crop | Observations | Linear trend over varieties (SE) (t/ha/yr) | Linear trend over years (t/ha/yr) |
|---|---|---|---|---|---|
| 1982-2018 | T | WW | 42472 | 0.063 (0.002) | -0.0030 (ns) |
| 1982-2018 | U | WW | 22247 | 0.109 (0.003) | -0.076 |
| 1982-2007 | T | WW | 31260 | 0.076 (0.004) | -0.013 |
| 1982-2007 | U | WW | 18667 | 0.108 (0.005) | -0.087 |
| 2007-2018 | T | WW | 11509 | 0.063 (0.01) | 0.025 |
| 2007-2018 | U | WW | 3535 | 0.220 (0.03) | -0.046 |
| 1982-2018 | T | WB | 21306 | 0.054 (0.002) | 0.035 |
| 1982-2018 | U | WB | 17034 | 0.057 (0.003) | 0.0034 (ns) |
| 1982-2007 | T | WB | 17959 | 0.068 (0.004) | 0.017 |
| 1982-2007 | U | WB | 15454 | 0.078 (0.003) | -0.032 |
| 2007-2018 | T | WB | 3546 | 0.063 (0.03) | 0.075 |
| 2007-2018 | U | WB | 1672 | 0.053 (0.01) | -0.0032 (ns) |
| 1983-2018 | T | SB | 18613 | 0.058 (0.001) | -0.019 |
| 1983-2018 | U | SB | 16529 | 0.068 (0.002) | -0.027 |
| 1983-2007 | T | SB | 14387 | 0.059 (0.002) | -0.013 |
| 1983-2007 | U | SB | 14430 | 0.079 (0.003) | -0.048 |
| 2007-2018 | T | SB | 4512 | 0.074 (0.01) | -0.083 |
| 2007-2018 | U | SB | 1978 | 0.081 (0.02) | -0.17 |

Table 5.1*: Linear trend over varieties $v_i$ and with the standard error (SE), and linear trend over years $r_j$, for winter wheat (WW), winter barley (WB) and spring barley (SB) for a range of periods within the trials datasets. The trend was calculated using equation [2.7]. T corresponds to fungicide treated trials; U corresponds to untreated trials. All trends are significant (p<0.05) unless denoted with (ns).*

For winter wheat, there was a strong positive trend of 0.063 t/ha/yr (SE = 0.002, $p<0.001$) between estimated variety yield (best linear unbiased estimator, or BLUE) and the year of origin for treated varieties in the period 1982-2018 (Figure 5.2ai). This indicates that breeders have successfully increased yields of new treated varieties by ~ 1 t/ha every 15-16 years in this period. The linear trend over years was insignificant (-0.0030 t/ha/yr, p=0.3). Using the BLUEs for both variety and year, it was possible to calculate the contribution of genetic effects to any linear increases in yields using the equation $\frac{Varieties\ regression}{Varieties\ regression + Years\ regression}$ (Mackay *et al.*, 2011). For treated varieties genetic effects contribute ~105% to the linear increases in yields. Looking at the most recent decade in the trials period, 2007-2018, and comparing the genetic gain to that for 1982-2007, also calculated by Mackay *et al.* (2011), it is evident that the rate of genetic gain slowed for treated winter wheat, from 0.076 t/ha/yr to 0.063 t/ha/yr.

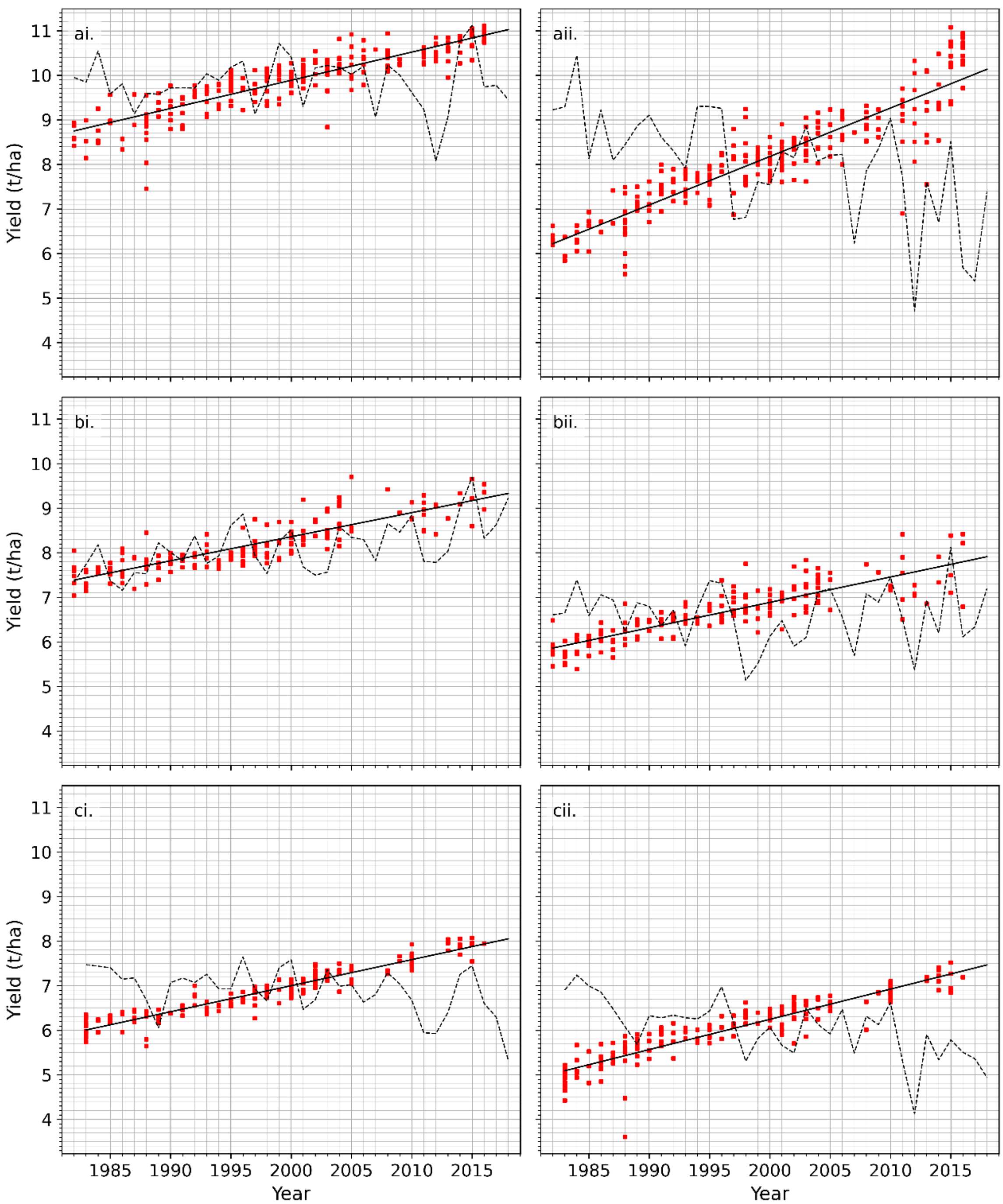



Figure 5.2*: Trends in variety* $\mathrm{v_i}$ *and year* $\mathrm{r_j}$ *effect for fungicide treated (i) and untreated (ii) winter wheat (a.), winter barley (b.) and spring barley (c.) trial yields from 1982-2018 (1983-2018 for spring barley), modelled using equation [2.7]. Variety effects (red squares) were plotted against the first year they entered the trials. Year effects (line) are plotted against calendar years.*

Untreated winter wheat variety trials (Figure 5.2aii) had a much greater genetic gain of 0.109 t/ha/yr (SE = 0.003, p<0.001), suggesting that untreated variety yields increased by ~ 1 t/ha every 10 years. Furthermore, the rate of increase in untreated variety effects appeared to increase from 2012 compared to treated variety effects. The linear trend over years was significantly negative for untreated variety trials at -0.076 t/ha/yr (SE = 0.005, $p<0.001$). Untreated winter wheat genetic gain values were substantially higher than the treated genetic gain values for all calculated periods within 1982-2018 and the corresponding linear trend over years was significantly negative.

Treated winter barley varieties experienced a lower genetic gain in yield, at 0.054 t/ha/yr for 1982-2018 (Figure 5.2bi). In contrast to winter wheat, the linear trend over year was also positive, at 0.035 t/ha/yr, indicating year contributed ~40% to the linear increase in treated trial yields. Furthermore, untreated varieties genetic gain was only marginally higher (0.057 t/ha/yr) and the effect of year was not significant. For the most recent decade, both treated and untreated genetic gain values decreased relative to 1982-2007.

Spring barley genetic gain values were more consistent with the pattern seen in winter wheat. The effect of year was negative for each time period and untreated variety trials consistently displayed higher genetic gain than treated variety trials (Table 5.1). For 1982-2018 the treated variety trials had a genetic gain of 0.058 t/ha/yr, but in the most recent period 2007-2018 the genetic gain increased to 0.074 t/ha/yr, suggesting that spring barley is the only crop of the three for which breeding has increased the rate of genetic gain and contribution to yield.

Year effects showed large variability from to year-to-year in both the treated and untreated trials for all crops (Figure 5.2). Where these variations coincided in both sets of trials likely reflects abiotic effects of the environment, such as climate variability. For example, all six trials have peaks in 1984 and 2015, both of which are known for being years with favourable weather and troughs in 2012, a year known for having detrimental weather (Section 4.3). The year effects varied considerably more in the untreated trials, which was likely due to variability in biotic effects i.e. diseases.

To test this hypothesis, winter wheat disease data was extracted from three national winter wheat disease surveys (Polley and Thomas, 1991; Hardwick *et al.*, 2001; Turner *et al.*, 2021) and combined with the yield difference between detrended untreated and treated year effects (Table 5.2). This yield difference can loosely be defined as the yield loss due to disease. There were several years within the period when untreated yields were >0.5 t/ha lower than treated and these typically coincided with years of higher disease prevalence. For example, the large dip in the untreated year effect (Figure 5.2aii) in 2012 occurred in a year when over 9% of second leaf and 5% of the flag leaf areas were affected by Septoria leaf blotch. 1993 untreated yield loss coincided with the highest area of Septoria nodorum blotch on both the flag and second leaf.

| Disease | Septoria nodorum blotch | | Septoria leaf blotch | | Powdery mildew | | Yellow rust | | Brown rust | | Tan spot | | Fusarium ear blight | U-T difference |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Leaf** | Flag | 2nd | Flag | 2nd | Flag | 2nd | Flag | 2nd | Flag | 2nd | Flag | 2nd | | |
| **1982** | 0.5 | 4.2 | t | 0.3 | 0.9 | 2.6 | t | 0 | 0.2 | 0.3 | | | | -0.15 |
| **1985** | 1.1 | 3.6 | 4.7 | 15.5 | 0.2 | 0.8 | 0 | 0 | t | t | | | | **-0.67** |
| **1986** | 0.2 | 1.7 | 0.1 | 1 | 0.1 | 0.3 | 0 | 0 | 0 | 0 | | | | 0.26 |
| **1987** | 0.4 | 2.8 | 1.3 | 4.7 | 0.1 | 0.2 | 0 | 0 | 0.1 | 0.1 | | | | -0.13 |
| **1988** | 0.1 | 0.2 | 0.4 | 2.1 | 0.2 | 0.7 | 0.3 | 0.2 | 0.5 | 0.5 | | | | -0.17 |
| **1989** | t | t | 0.1 | 0.7 | 0.1 | 0.4 | 0.6 | 0.9 | 0.1 | 0.2 | | | t | 0.33 |
| **1990** | t | t | 0.2 | 0.6 | 0.2 | 0.5 | 0.3 | 0.2 | 0.3 | 0.6 | | | 0.1 | 0.51 |
| **1991** | t | 0.1 | 1.5 | 4.2 | 0.2 | 1.1 | t | t | 0.1 | 0.1 | | | 0.1 | 0.08 |
| **1992** | t | 0.1 | 0.2 | 1.2 | 0.1 | 0.5 | t | t | t | 0.1 | | | 0.1 | -0.15 |
| **1993** | 0.1 | 0.5 | 0.9 | 4.2 | 0.1 | 0.8 | t | t | 0.4 | 0.8 | | | 0.1 | **-0.80** |
| **1994** | t | 0.3 | 0.1 | 1.2 | 0.1 | 0.4 | t | t | 0.1 | 0.1 | | | t | 0.81 |
| **1995** | t | t | 0.2 | 0.8 | t | 0.2 | t | 0 | t | 0.1 | | | t | 0.58 |
| **1996** | t | t | 0.1 | 2.1 | 0.1 | 0.3 | 0 | 0 | t | t | | | t | 0.47 |
| **1997** | t | 0.3 | 0.7 | 3.1 | t | 0.2 | 0.1 | t | t | t | | | 0.1 | **-0.77** |
| **1998** | t | 0.1 | 2.9 | 7.8 | t | 0.1 | 0.1 | 0.1 | 0.2 | 0.5 | | | 0.6 | **-1.28** |
| **1999** | 0.02 | 0.13 | 1.77 | 6.74 | 0.05 | 0.15 | 0.02 | 0.04 | 0.67 | 0.7 | 0 | 0 | | **-1.37** |
| **2000** | 0.03 | 0.27 | 2.18 | 7.22 | 0.03 | 0.11 | t | 0.01 | 0.06 | 0.13 | 0 | 0 | | **-1.07** |
| **2001** | 0.01 | 0.06 | 0.11 | 1.05 | 0.03 | 0.11 | 0.01 | 0.01 | 0.01 | 0.01 | 0 | 0 | | 0.86 |
| **2002** | 0.01 | 0.03 | 2.76 | 9.01 | 0.05 | 0.14 | t | t | 0.02 | 0.03 | 0 | t | | -0.07 |
| **2003** | t | 0.01 | 0.76 | 3.78 | 0.03 | 0.19 | 0 | 0 | 0.05 | 0.04 | 0 | t | | 0.63 |
| **2004** | t | 0.01 | 1.54 | 4.6 | 0.05 | 0.17 | 0 | t | 0.02 | t | T | t | | -0.02 |
| **2005** | 0.01 | 0.01 | 1.55 | 3.69 | 0.06 | 0.17 | t | t | 0.15 | 0.08 | 0.03 | 0.04 | | 0.32 |
| **2006** | t | t | 1.1 | 5.11 | 0.03 | 0.09 | 0.03 | 0.02 | 0.12 | 0.03 | 0.02 | 0.03 | | 0.24 |
| **2007** | t | t | 0.98 | 3.02 | 0.01 | 0.02 | 0.01 | 0.01 | 0.27 | 1.14 | 0.01 | 0.01 | | **-0.55** |
| **2008** | t | t | 1.13 | 2.84 | 0.02 | 0.06 | t | t | 0.01 | 0.03 | 0.01 | 0.02 | | -0.03 |
| **2009** | t | t | 0.23 | 1.53 | 0.02 | 0.08 | t | t | t | 0.01 | 0.03 | 0.05 | | 0.76 |
| **2010** | 0 | 0 | 0.1 | 0.51 | 0.04 | 0.06 | t | t | 0 | 0 | 0.01 | 0.03 | | 1.90 |
| **2011** | 0 | 0 | 0.31 | 1 | 0.02 | 0.09 | 0.01 | t | 0.02 | 0.05 | 0.04 | 0.07 | | 1.03 |
| **2012** | 0 | 0 | 5 | 9.59 | t | t | 0.02 | 0.02 | 0.08 | 0.05 | 0.02 | 0.04 | | **-0.76** |
| **2013** | 0 | 0 | 0.18 | 0.65 | 0.03 | 0.08 | t | t | t | t | T | 0.01 | | 1.22 |
| **2014** | 0 | 0 | 5.86 | 11.5 | 0.01 | 0.03 | t | t | 0.01 | 0.02 | 0.03 | 0.04 | | **-1.30** |
| **2015** | t | t | 0.31 | 1.3 | 0.01 | 0.02 | 0 | 0 | 0.04 | 0.05 | 0.02 | 0.03 | | 0.21 |
| **2016** | 0 | 0 | 0.15 | 0.69 | t | 0.01 | 0.02 | 0.02 | t | t | T | 0.02 | | **-1.16** |
| **2017** | 0 | 0 | 1.37 | 2.61 | 0.02 | 0.06 | t | t | 0.09 | 0.05 | 0.02 | 0.06 | | **-1.44** |
| **2018** | 0 | 0 | 0.36 | 1.76 | 0.02 | 0.09 | t | t | 0.02 | 0.04 | T | t | | 0.96 |

Table 5.2*: Winter wheat disease incidence severity (average percentage area of leaf) of foliar diseases and fusarium ear blight combined with the difference between the detrended year effects for untreated and treated variety trials with year effects <-0.5 shown in bold. Disease data from Hardwick* et al. *(2001), Polley & Thomas (1991) and Turner* et al. *(2021). No disease data was available for 1983 and 1984.*

Linear regression showed that there is a negative relationship (β=-2.42 t/ha/%, p<0.001) between the difference in detrended untreated-treated year effects and Septoria leaf blotch such that higher incidence of Septoria leaf blotch was associated with a greater yield loss in the untreated varieties trials. However, there were significant deviations from this, such as in 2016 when there was little Septoria leaf blotch but large U-T differences.

Genetic gain estimates based on limited observations generate much larger standard errors and show the regression model is less precise. This is an important consideration in making this calculation. Looking at the four periods within 1982-2018 with the winter wheat trials data, there were large fluctuations in the estimated genetic gain of treated varieties (Figure 5.3). 1982-1991 had the lowest genetic gain (0.022 t/ha/yr) and greatest standard error (SE = 0.04 t/ha/yr) which reflects the lower number of observations (6498 compared to 10,000+) contributing to it. 1991-2000 and 2009-2018 had the same genetic gain (0.087 t/ha/yr), but 2000-2009 dipped to 0.065 t/ha/yr. Selection by breeders at more sites per variety and more trials overall would reduce this standard error.

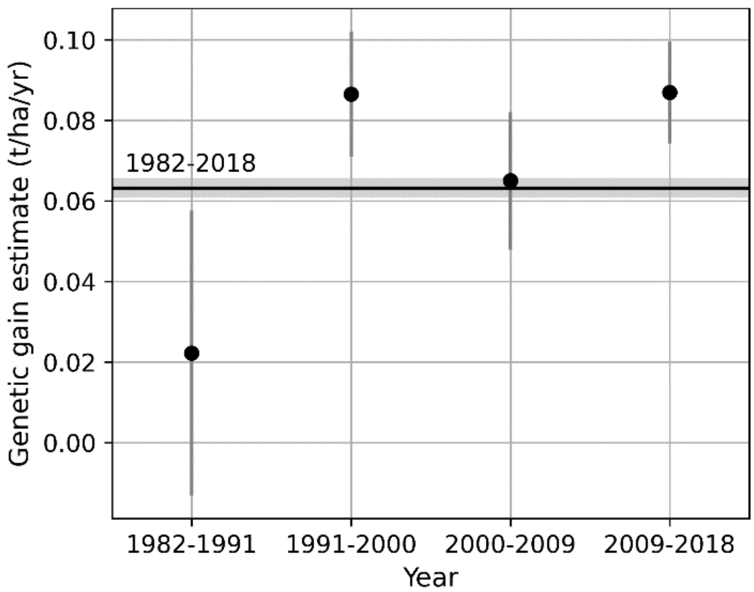



Figure 5.3*: Genetic gain estimates and standard error for fungicide treated variety trials for four periods within 1982-2018: 1982-1991, 1991-2000, 2000-2009 and 2009-2018. The 1982-2018 genetic gain and standard error are shown by the horizontal line in black, and grey, respectively.*

## 5.2 The breakdown of disease resistance is seen in untreated variety trial yields

Yield difference in treated and untreated trials at the same site varied significantly year to year (Figure 5.4). Comparisons between the crops show that on average spring barley exhibited the lowest median yield difference each year relative to mean treated and untreated trial yields and winter wheat the largest yield differences. This suggests winter wheat was the most susceptible of the three crops to yield impacts from disease.

Some of the peaks in yield difference overlapped across the crops, for example in 2012 treated trials significantly outperformed untreated for all three crops. For winter wheat, 1997-2000, 2007, 2014, 2016 and 2017 also had large relative yield differences of over 30%. Some of these years corresponded to high disease pressure in the winter wheat disease severity surveys (Table 5.2) and suggest fungicide treatment was particularly important. In 1984, 1994 and 1995 treated trials average yields were less than 10% higher than untreated yields across the UK. There was a significant positive correlation between year and yield difference for winter wheat (r = 0.49, p=0.002), suggesting that treated varieties were outperforming untreated varieties by an increasing amount over the period 1982-2018.

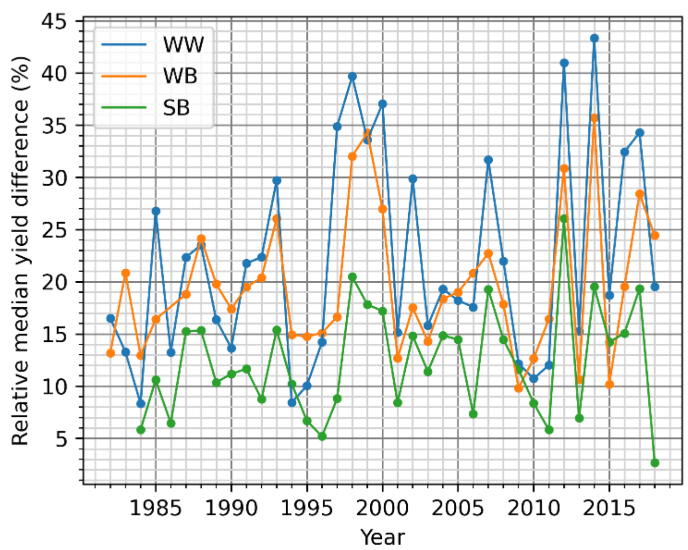



Figure 5.4*: Relative median yield difference (%) between fungicide treated and untreated trial yields for winter wheat (WW, blue), winter barley (WB, orange) and spring barley (SB, green).*

There was a positive linear relationship of 0.064 t/ha/yr (SE = 0.006, p<0.001) (Figure 5.5a), 0.032 t/ha/yr (SE = 0.002, p<0.001) (Figure 5.5b) and 0.015 t/ha/yr (SE = 0.004, p<0.001) (Figure 5.5c) between treated and untreated yield difference and variety age for winter wheat, winter barley and spring barley, respectively. As variety age increased, the mean estimated yield difference between treated and untreated trials increased, suggesting that after varieties were first introduced into the trials network, disease resistance broke down. To investigate this further, the effect of variety age on treated and untreated trial yields were also modelled separately (Figure 5.6).

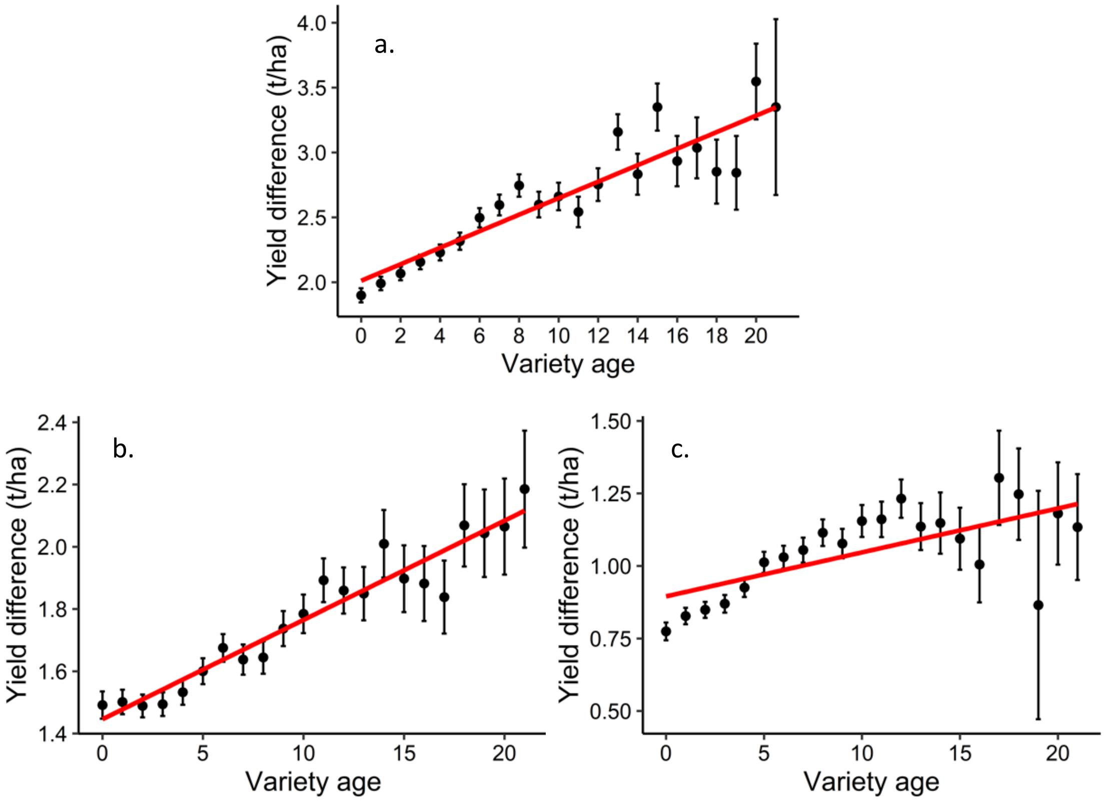



Figure 5.5: *Variety age against the treated-untreated variety trials yield difference for* a. *winter wheat,* b. *winter barley and* c. *spring barley. The effect of variety age on the yield difference between fungicide treated and untreated varieties was modelled using equation [5.8]. Variety age indicates the number of years since the variety entered the trials system. The red line shows the linear relationship between the two variables.*

The difference in treated to untreated variety trial yields increased as the varieties aged until at least 12 years when there was greater variability (Figure 5.5). In the two winter crops, after the first two years, treated variety trial yields increased as the variety aged (Figure 5.6). By contrast, spring barley treated variety trial yields decreased as the variety aged until it was in the trials system for 10 years, when its yields increased.

All three crops exhibited yield decreases as untreated varieties aged (Figure 5.6). This was attributed to loss of disease resistance and as a consequence the effect of year is overestimated downwards resulting in variety means being overestimated, giving much higher biased estimates of untreated genetic gain. After ~12 years in trial, the longer lasting untreated varieties showed slight yield increases again.

The relationship between variety age and yield was less clear as the variety age increased past 10 years, when looking at treated varieties, untreated varieties and the yield difference between the two. This is likely because there were fewer varieties as the variety age increased. For example, in the winter wheat analysis (Figure 5.5a., 5.6) there were only eight varieties present for at least

15 years, with only one spanning 21 years and two spanning 20 years, which means that rather than being a good estimate of the average effect of variety age it moved instead to being the average effect of one or two varieties.

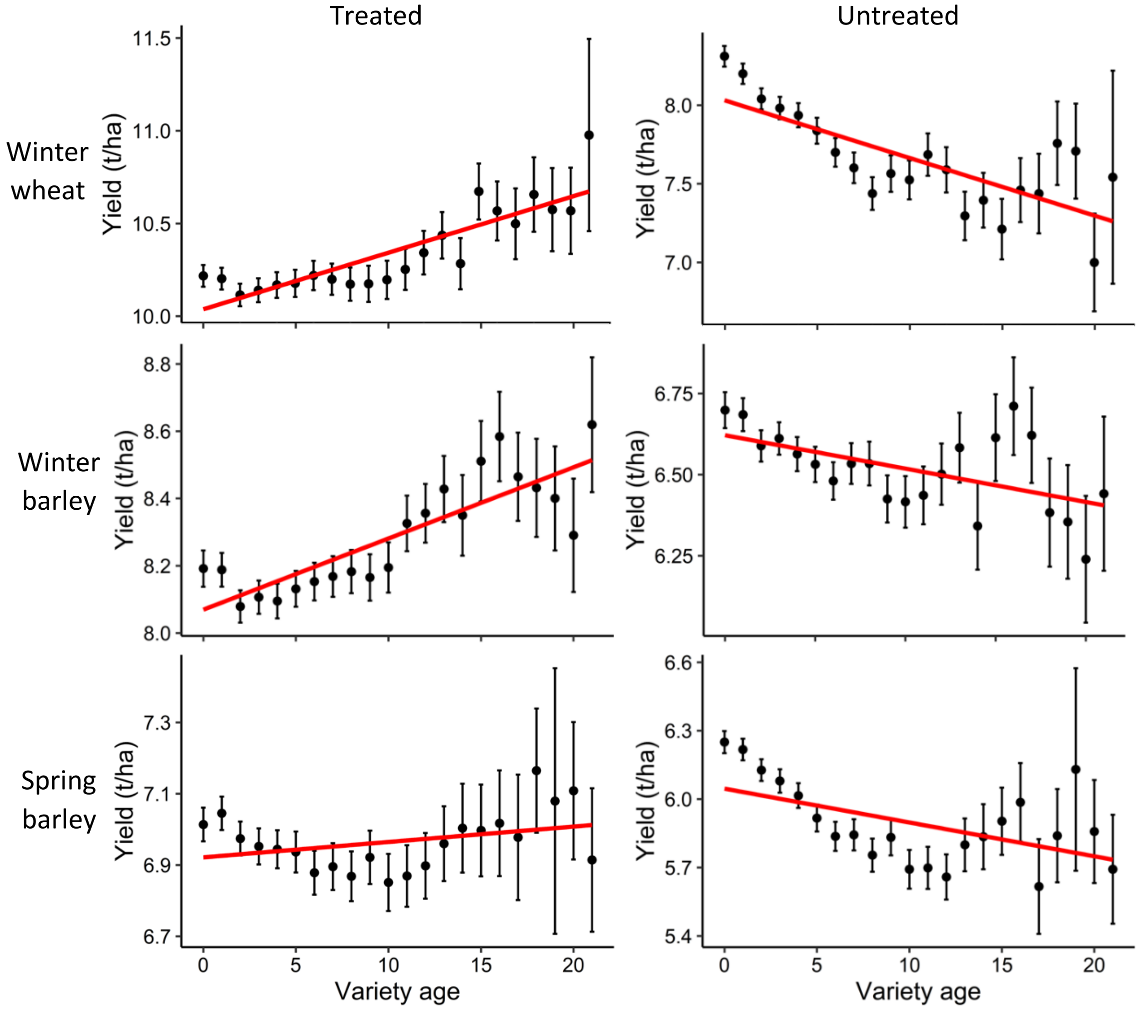



Figure 5.6*: Variety age against treated and untreated trial yields for winter wheat, winter barley and spring barley. Here the effect of variety age on varieties under full fungicide treatment and no fungicide treatment was modelled using equation [2.8] separately. Variety age indicates the number of years since the variety entered the trials system. The red line shows the linear relationship between the two variables.*

The analysis was repeated restricting varieties to 10 years old to reduce the bias of longer lasting varieties (Figure 5.7). The linear relationship between variety age and yield difference was clearer across all three crops, which stems from the greater overlap of varieties for varieties aged up to 10 years. As varieties aged in the untreated trials, their yield decreased (Figure 5.7). The effect is biggest in winter wheat, which shows yield losses of 1 t/ha for a variety present for 10 years in the trials system. Spring barley untreated varieties experienced yield losses of 0.6 t/ha and winter barley yield losses were only 0.3 t/ha over 10 years.

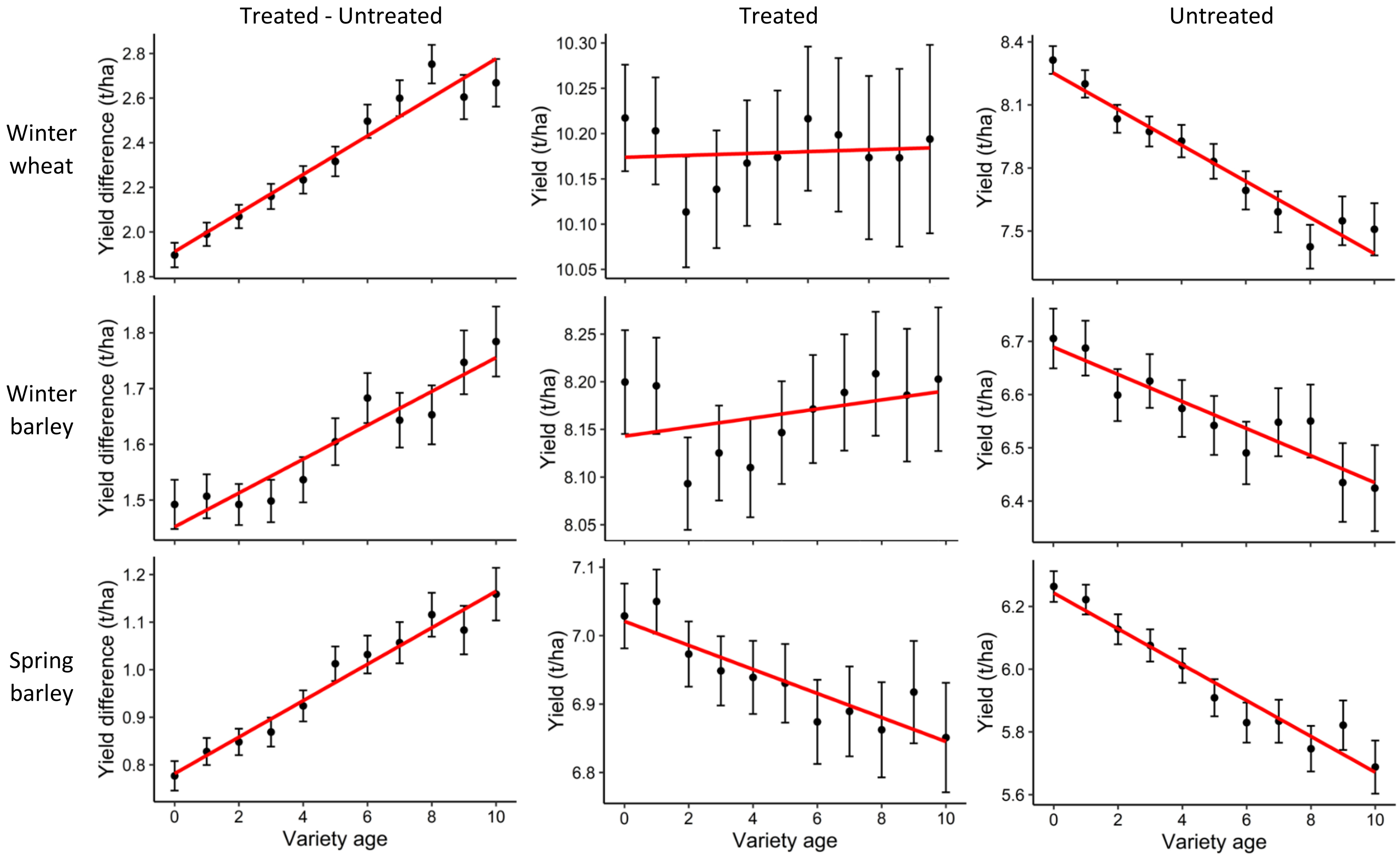



Figure 5.7*: The effect of variety age on yield difference between fungicide treated and untreated variety trials, on treated variety trial yields and on untreated variety trial yields for winter wheat, winter barley and spring barley. The effect of variety age on the yield difference, treated and untreated variety yields was modelled using equation [2.8] and restricted to varieties aged up to 10 years. Variety age indicates the number of years since the variety entered the trials system. The red line shows the linear relationship between the two variables.*

For treated varieties, there was an initial drop in yield after two years (Figure 5.7). Winter wheat and winter barley yields then gradually increased, but there was no overall clear long-term trend. In spring barley, there was an overall negative relationship between variety age and yield. Across the 10 years the yields only decreased by 0.3 t/ha, half the rate of the untreated varieties, hence the increasing yield difference as a variety ages still occurred.

## 5.3 Genetic gain estimates are susceptible to bias

Breakdown in disease resistance can result in biased estimates of untreated trials datasets. Factors affecting genetic gain estimates of treated variety trial yields were explored using subsets of the NL/RL dataset (Figure 5.8), specifically the inclusion of check varieties (i.e. varieties present at least 10 consecutive years) to increase connectivity and the number of checks included.

The number of checks involved in a breeding programme significantly influenced the genetic gain estimate (Figure 5.8). This was particularly clear for the first four case studies, in which the genetic gain estimate was highest when there were no checks included and decreased by ~30-40% upon the inclusion of one check, ~10-20% when there were two checks, ~5-10% when there were three checks. The values also begun to converge as the checks increased, so that estimates with four and five checks were similar to three checks, but with small standard errors. This effect on standard error was as expected given the standard error is proportional to the number of data points, which increases as more checks are included in the dataset.

For the two most recent time periods, the 5$^{th}$ and 6$^{th}$ case study (Figure 5.8), the decay pattern was less clear. This is partly due to the size of genetic gain values, which were all much lower for these two periods. For 2005-2015 the values decreased by 20% on average from zero checks to one check, and then by a few percent between subsequent increases in checks. Unfortunately, for 2008-2017 it was also not possible to calculate the genetic gain for zero checks as there was insufficient data. When checks were added this was no longer an issue.

The 1982-2018 treated genetic gain for winter wheat was recalculated using just the first three years of data for each variety with no checks. The estimated genetic gain was 0.158 t/ha/yr (SE = 0.003), which is 2.5 times larger than the original estimate of 0.063 t/ha/yr (SE = 0.002). This suggests that having zero checks results in an increased genetic gain, as seen in the case studies (Figure 5.8).

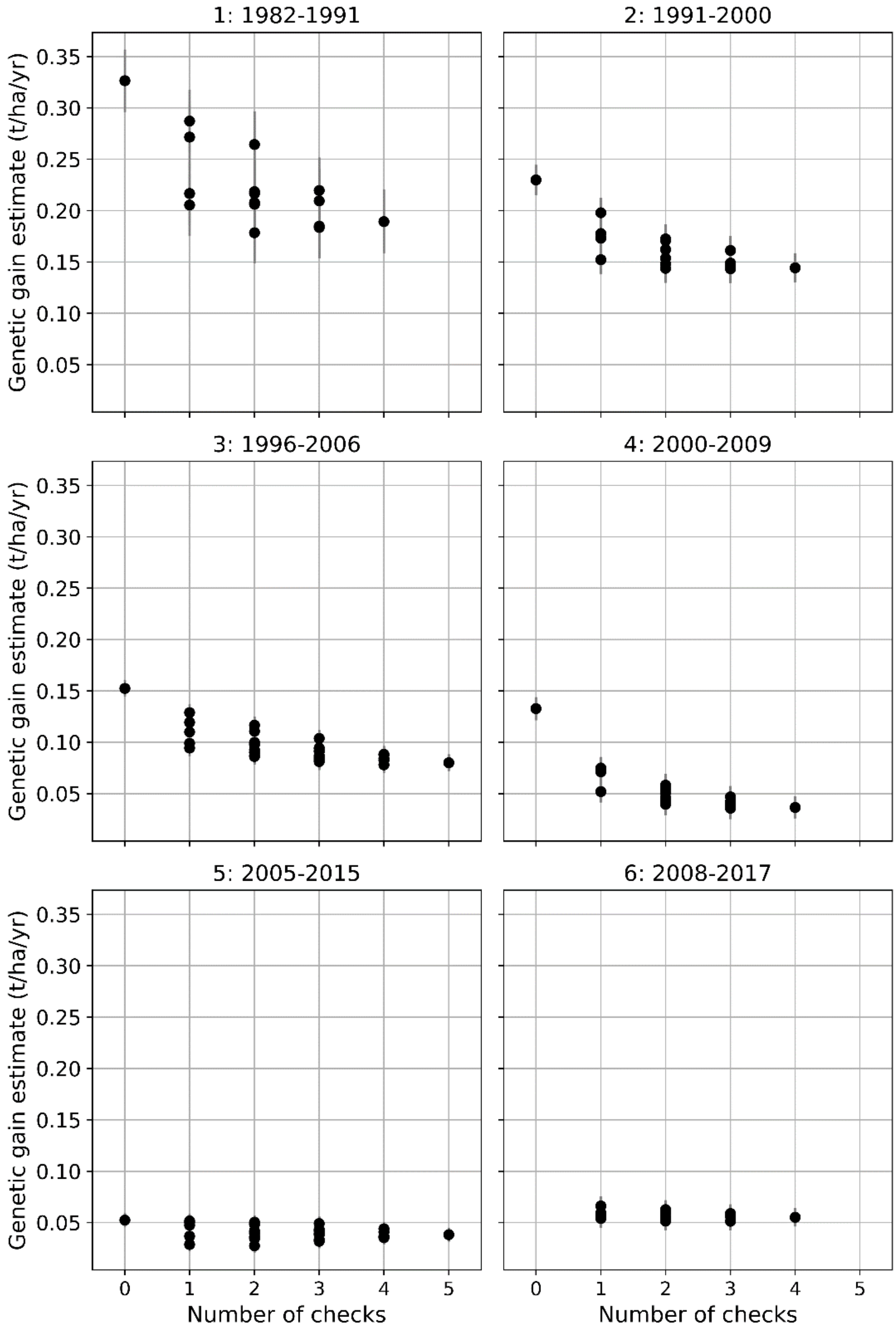



Figure 5.8*: Winter wheat genetic gain estimates for six case study periods and varying numbers of checks extracted from the 1982-2018 NL/RL dataset. Checks refer to varieties present in the trials system for a minimum of 10 consecutive years. Varieties with more than three years of data were restricted to their first three years in trial.*

Another finding is the effect of the choice of check on the genetic gain estimates (Figure 5.8). This is shown by the spread in genetic gain estimates for all case study periods with just one check. For example, for 2005-2015 with the variety *Alchemy* as a check, the period had a genetic gain of 0.050 t/ha/yr, compared to almost half that at 0.029 t/ha/yr when *Claire* was the check variety. For 1982-1991, the values ranged from 0.21 t/ha/yr (*Galahad*) to nearly 50% larger at 0.29 t/ha/yr (*Fenman*).

## 5.4 Variety improvement contributes over 95% of improvements in national yields

Estimated national yields $z_j$ on trials are higher than observed national yields $o_j$ by an average of 1.28 t/ha. Variety trials routinely outperform on-farm yields, likely because trials are located on better soils and are typically located in the middle of fields so are unaffected by edge-effects that reduce yields (Ian Mackay, *pers. comm.*).

National winter wheat yields increased over the 36-year period from less than 6.5 t/ha to over 8 t/ha (Figures 5.9 and 4.1). A spline (black) was fitted to the national yield data using the *smooth.spline* function from the ***stats*** base package in R. There was a clear decrease in the rate of improvement in winter wheat yields from around 1998 which continued to 2018, also seen in Section 4.1. Variety improvement contributed almost all (95.5-99.8%) the improvements in yield observed over the period 1983-2018, in agreement with the findings in trials yields in Section 5.1 and hypothesis made in Chapter 3.

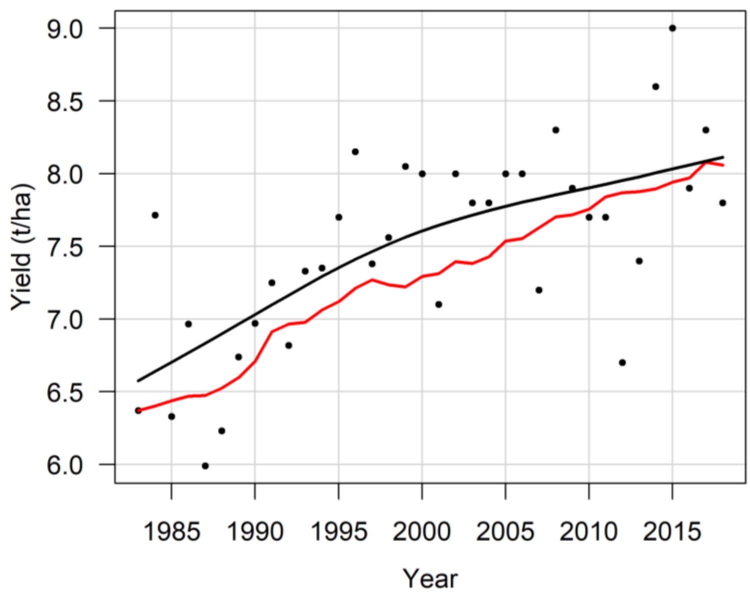



Figure 5.9*: Estimated contribution of variety improvements to national winter wheat yield increases for 1983-2018. DEFRA national yield data (black dots) are fitted with a spline (black line), as used in* (Mackay *et al.*, 2011)*. The contribution of variety effects $v_i$ to national yield is shown (red line) and was calculated using equation [2.11] on the AHDB variety trials data and DEFRA national yield data.*

## 5.5 Discussion

### 5.5.1 Disease resistance as a driver for treated and untreated yield trends

Variety trial yields have increased over time for all three crops (Figure 5.1). Winter wheat showed the highest interannual variability in yields (Figure 5.1), indicating lower yield stability than barley, which has also been observed in Denmark (Macholdt *et al.*, 2021). Winter wheat also had the largest relative yield differences between treated and untreated yields (Figure 5.4) which were partially attributed to the yield effects of disease such as Septoria leaf blotch (Table 5.2). As discussed in Section 4.5.1, the winter wheat disease survey data used in Table 5.2 and Figure 4.26 (Polley and Thomas, 1991; Hardwick *et al.*, 2001; Turner *et al.*, 2021) does not record the full extent of some diseases, in particular yellow rust, as the survey is often undertaken once the disease has left the crop. In addition to 2014 (Section 4.5.1), 2016 was also a bad year for yellow rust (James Brown, *pers. comm.*), as well as 1998-2000 because of the breakdown in the resistance gene *Yr17* (Bayles *et al.*, 2000), none of which were picked up in the survey data (Table 5.2). In all five of these years untreated yields were at least 1 t/ha less than treated yields. The significant correlation between untreated-treated yields and Septoria leaf blotch incidence is likely confounded by the occurrence of yellow rust that hasn't been recorded and years with greater untreated-treated yield differences may instead be due to yellow rust. The analysis should be repeated with a more appropriate dataset to confirm this hypothesis and reveal the relative disease impacts on the untreated-treated yield differences.

Barley diseases such as *Rhynchosporium*, mildew and rusts have smaller but persistent effects, which may help explain the lower yield differences in treated and untreated yields in barley compared to wheat (Figure 5.4) (Steve Hoad, *pers. comm.*). Winter barley had larger relative yield differences between treated and untreated variety trial yields compared to spring barley variety trials. Two important diseases - *Rhynchosporium* and net blotch - affect winter barley more than spring barley, likely contributing to this difference. Furthermore, the development of resistant spring barley varieties to powdery mildew has reduced the impacts of the disease in recent years (AHDB Cereals & Oilseeds, 2018a), whereas few winter barley varieties have the resistance genes and are more susceptible.

Comparison of genetic gain values suggests that winter wheat, and to some extent spring barley, untreated genetic gain was overestimated and this is thought to be due to loss of disease resistance shown by the effect of variety age on yield (Figure 5.5). The effect of breakdown in disease resistance on the estimated year effect is a negative bias resulting in the downward trend in year effect (Figure 5.2aii and 5.2cii), which the estimated variety effect compensates for by being overestimated upwards (Figure 5.10). It is unlikely that winter wheat varieties have experienced a break down in Septoria leaf blotch resistance, as varieties' Septoria ratings remained largely stable over the study period, although they have since declined (AHDB, 2021). Rather its more likely to do with rust resistance and occasionally mildew resistance, some of which are known to be non-durable (James Brown, *pers. comm.*). To explore the contribution of individual diseases further, a more in-depth analysis focussing on individual varieties and their relative disease resistances is required.

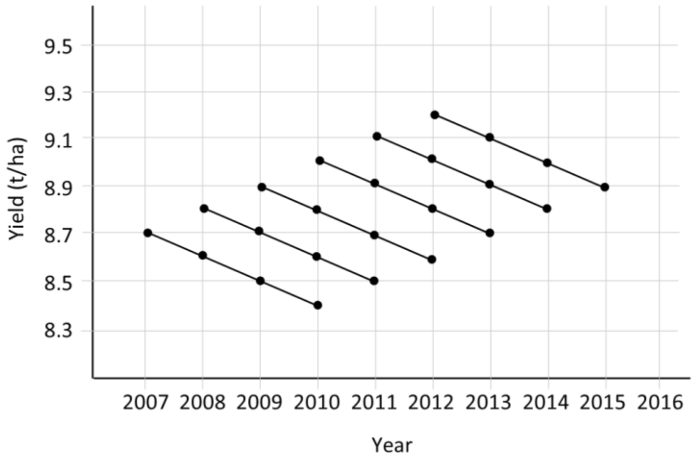



Figure 5.10*: Model demonstrating the effect of loss of disease resistance on estimation of variety and year effects. Here variety effects increase by 0.1 t/ha per year and then decline linearly due to breakdown in disease resistance. The difference in variety performance for those present in successive years is -0.1 t/ha, which gives an estimated year effect of -0.1 t/ha, as opposed to an actual year effect of 0. Differences in variety performance are estimated by calculating the differences between a variety and its predecessor(s) within years, +0.2 t/ha. This results in variety effects biased upwards and year effects biased downwards. Figure taken from Mackay et al. (2011).*

In France, this higher rate of genetic gain in untreated wheat trials was also observed, but instead was attributed to improvement in resistance to fungal disease (Brisson *et al.*, 2010). This has also been suggested as an explanation in the UK (Shorinola *et al.*, 2022) but it seems unlikely as the yield difference between treated and untreated yields was found to be increasing for winter wheat (Figure 5.4), suggesting either improved fungicide treatments, reducing the outbreak of disease and yield impacts in treated varieties, or an increase in disease susceptibility in untreated varieties or disease pressure. If it's the former and resistance to fungal disease has also improved but at a slower rate, this would still not explain why untreated varieties see such a significant decline in yield as they age.

Winter barley untreated yield losses were much smaller (Figure 5.6) and this is reflected in the smaller untreated genetic gain values (Table 5.1). This is possibly also explained by lower susceptibility of barley to disease. Untreated variety trial yields still showed long-term decreases because two important barley diseases – *Rhynchosporium* and net blotch – have a system of variety-specificity which isn't well understood and varieties' ratings for both diseases decline over time.

The increase in the variety age effect of both treated and untreated yields after ~10 years (Figure 5.6) for all three crops has not previously been documented. Long-term varieties in this dataset were introduced at different points in the period of interest, therefore it does not appear to be the effect of increasing yields at some point in the timeseries which is independent of genotype. Some varieties may have become more resistant to disease as different disease races come to dominate. For example, a variety is normally more susceptible to one or more races of yellow rust rather than all races of yellow rust. If, over time, the dominant race isn't the one it's susceptible to, its resistance could improve and this may explain the observed uptick seen. It may also be possible that these longer standing varieties end up benefitting from the effect of being surrounded by newer resistant varieties, so they get less disease than they would if older, less resistant varieties were nearby. A more sophisticated analysis at the plot level, taking into account the effects of neighbouring plots could be a way of testing this.

In the treated variety age plots (Figure 5.7), there were distinct drops in yield after the first and second year a variety is in trial. The first two years (age 0 and 1) correspond to the NL trial years, during which breeders submit only modest quantities of seed. It is in the breeders' best interest to select the best seed, on size, density and even germination tests, for trial which will yield higher than the average seed. In the RL trials (age 2 onwards), much larger quantities are required, lowering average yields. Hence seed quality could be an influence on the initial yield drop observed.

Selection bias could also have contributed to the initial yield drop. In a simplified process, there is a first stage to get onto the NL and a second stage to get through and onto the RL. If the heritability (i.e. the degree of variation in yield due to genetic variation) is 1, there should be no difference in the phenotype between the two years and if there is no selection from stage 1 to 2 there should be no difference, on average. However, with some selection and with heritability less than 1, the lines selected in the first stage will have better performance than the population mean for two reasons: they are genetically better and they tend to have positive environmental deviations. Re-testing in the second stage, the lines will still be genetically better, but their environmental deviations will be zero on average. Hence the mean of the selected group in stage 1 will look worse than the mean of the same lines retested.

With the NL/RL, this selection bias will only prevail if there is little selection between NL1 and NL2 and some selection from NL2 to RL1. In subsequent years of the RL, there is selection but because of the many sites used with a year, the heritability is higher and selection also incorporates results from the previous year. Testing this through a simulation of the original dataset could indicate whether selection bias is contributing here.

Whilst the explanation of this drop is not clear, it is a notable strength of the historical datasets that there is sufficient power to pick up small effects. The difference between years one and two and subsequent years for winter wheat is <1%. This is clearly visible in the graphs, but the power to detect a different this small in a typical bespoke experiment is close to zero.

### 5.5.2 Uncertainty in genetic gain estimates

Estimates of genetic gain for different case study periods within the 1982-2018 NL/RL trials data were dependent on the number of check varieties included in the dataset and the specific checks chosen. Specifically, increasing the number of long-running check varieties lowered the genetic gain estimates (Figure 5.8). Absence of check varieties results in low connectivity which means the estimates of genetic gain can be confounded with the year effect, hence it is recommended that checks are used to improve estimates (Covarrubias-Pazaran, 2020). Having multiple checks makes it easier to identify the effects of year. However, it also means an increased proportion of the estimate of genetic gain comes from the difference in age and yield between the checks themselves. If these yields are increasing at a lower rate than the new varieties, this could be dragging the estimate down.

Genetic gain estimates were also highly dependent on the checks chosen, particularly when only one check was included. This is likely because the checks weren't completely stable and behaved differently within the case study periods. They had different mean yields and some showed slight increases over time whilst others didn't, all of which will have influenced the genetic gain estimate.

Additional model runs that included checks in the regression estimate showed that this lowers the genetic gain estimate further. The extent to which the estimate is lowered is dependent on the mean yield of that check. For example, a check with a higher adjusted mean yield (c2 in Figure 5.11) will lower the genetic gain estimate (G2) compared to a check (c1) that behaves the same across the period but with a lower adjusted mean yield introduced in the same year, as newer varieties with higher adjusted mean yield won't show as large relative increases in comparison. When checks are included in the final regression estimate of genetic gain, it could therefore be possible to, knowingly or not, bias a genetic gain estimate upwards by using a consistently low yielding check in a breeding programme (Figure 5.11). Use of unstable checks can also influence this estimate. Evidently this method of calculating genetic gain needs refining to reduce the vulnerability of the estimate to the choice and number of checks.

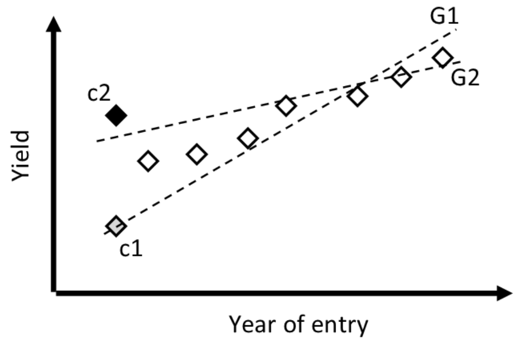



Figure 5.11*: Model to demonstrate the effect of using a low yielding check (c1) vs. a high yielding check (c2) on the estimate of genetic gain. Genetic gain is calculated by regressing the adjusted variety mean yields on year of entry across all varieties (diamonds). If the check variety has a low adjusted mean yield, then the estimated genetic gain will be higher (G1 line) than the genetic gain estimate (G2 line) of a dataset including a high yielding check (c2) instead.*

A large inflation in the genetic gain estimate is also seen in the 1982-2018 dataset when it is restricted to the first three trial years for each variety (0.158 t/ha/yr vs. 0.062 t/ha/yr). In this particular winter wheat dataset it was found that variety yields decreased in the first three years they were present in the system, before showing overall long-term increases (Figures 5.6 and 5.7). Therefore, it is possible that the treated genetic gain value here is biased in the same way the untreated variety trials were (Figure 5.10).

Furthermore, a long-term increase in yield was found in long-lasting winter wheat varieties (Figure 5.6). This means in the full dataset analysis (Figure 5.1ai) the adjusted mean yields for long lasting varieties were higher than in the analysis restricted to the first three years during which yields generally declined, resulting in a lower genetic gain estimate as explained in Figure 5.11.

The effect of inclusion of a check can be demonstrated using a model dataset which includes a check variety present across all years whose yield increases linearly each year (variety g) (Table A3). This is equivalent to the variety effect being constant but year effects increasing linearly. Using a simple linear model of yield including year and variety, there is a slight overestimate in the relative yield increase between the oldest (variety a) and newest (variety e) by 4.5 (Table 5.3, model 1), but when the check is dropped from the model this overestimate is accentuated with a relative yield increase of 6 as the effect of year increasing yields by 1 each year is not accounted for (Table 5.3, model 2).

| Variety | Model | | | |
|---|---|---|---|---|
| | 1 | 2 | 3 | 4 |
| a | 4.4 | 3.67 | 4.5 | 4.5 |
| b | 5.54 | 5.17 | 5.5 | 5.5 |
| c | 6.67 | 6.67 | 6.5 | 6.5 |
| d | 7.8 | 8.17 | 7.5 | 7.5 |
| e | 8.93 | 9.67 | 8.5 | 8.5 |
| g | 4 | - | 4.5 | - |

Table 5.3: *Adjusted variety means calculated using four linear models on a model yield dataset (Table A3). Model 1 includes all data and accounts for year and variety, model 2 doesn't include the check but accounts for year and variety, model 3 includes all data and accounts for year, variety and whether it's in the NL or RL year and model 4 doesn't include the check and accounts for year, variety and NL/RL year.*

The drop in yield when the variety moves from NL to RL can also be simulated. By adding in a seed term to models 1 and 2 (Table 5.3), results show that the effect of the drop from NL to RL seen in all three crop trials datasets results in an overestimate of genetic gain (Table 5.3, models 3 and 4). When a new variety enters the system in year three of an older candidate, it will appear to have a higher relative yield compared to the same variety entering in year two, increasing the apparent rate of genetic gain. In models 3 and 4, where seed source has been included as a term, the variety means were estimated perfectly, therefore recovering the true (simulated) values. This suggests that the age of the variety or seed source factor may need to be included in these analyses routinely.

### 5.5.3 Limitations of using the UK National List/Recommended List dataset

The NL/RL trials data used in this analysis is thought to be biased as each year some trials don't make it to the harvest results as they are scrapped, whether that be due to extreme weather, difficulties planting, large losses due to disease etc. As a result, the worst trial yield results of the year are not seen. Theoretically, the sites that fail should be failing at random, but this is not likely at all as it will depend on, for example, location.

There were several variety-year combinations with very few sites, particularly for winter wheat. These were dropped from the analysis to prevent false representation of variety performance across multiple environments. It is not clear why there is such a large reduction in sites in these cases. It is possible these varieties survived with a regional recommendation only so weren't tested outside of the region, but there would still be more than one site per region if this were the case. By dropping this data, it did result in interruptions in the continuity of some variety series which was problematic and disappointing.

There were significant data access issues, such that it took over a year to be granted access to this important dataset. Extensive quality control and cleaning of the data was required before any analysis could take place as there were a lot of mistakes in variables such as grid references and drilling dates. Furthermore, even after the data was re-extracted by the data holders, there was still missing data for 2007 for several varieties, which made it difficult to estimate genetic gain for 2007-2018. There are several varieties that were introduced in 2006 and present for several years, but not 2007 (i.e. *Viscount, Grafton, Gallant, Scout*). All four of these varieties are in seed statistics data from 2008 suggesting the NL1, NL2 and RL years were 2006, 2007 and 2008, then they could be bought for 2009 harvest.

One assumption made when comparing the NL/RL variety trials data and the national on-farm crop data is that the agronomy of UK farms and variety trials have changed in the same way over this period. Until the introduction of separate treated and untreated trial series in 1982, the agronomy of the trials followed "best practice". After that change, untreated trials received no fungicide treatments and treated trials were treated prophylactically. The protocol for trials management can be found at [https://ahdb.org.uk/ahdb-recommended-lists-for-cereals-and-oilseeds-2021-2026](https://ahdb.org.uk/ahdb-recommended-lists-for-cereals-and-oilseeds-2021-2026). Aside from fungicide treatments, it continues to adopt UK best practice and is under the control of the grower in whose field the trial is located. Since trials are located on uniform ground and away from headlands, hedges and trees, the average yield of the treated trials is always higher than the national yield, as seen here. However, annual variation in yield in the two trial series remains very similar, hence this assumption on same agronomy seems valid.

## 5.6 Conclusion

This research has shown that breeding is still contributing to increases in yield in winter wheat, winter barley and spring barley for both treated and untreated variety trials in the UK. Statistical modelling has shown that an increasing yield difference between fungicide treated and untreated variety trials as varieties age is driven by both a breakdown in disease resistance of untreated varieties and previously unobserved long-term yield increases of varieties in treated trials.

Use of NL/RL trials data has allowed potential sources of uncertainty in genetic gain estimates to be explored. Varying the number of long-term check varieties in the data has shown that inclusion of checks leads to a less biased estimate of year effects. However, the genetic gain estimate is highly sensitive to the check chosen and is influenced by the initial drop in yield associated with moving from NL to RL. This raises important questions in terms of how much emphasis should be put on genetic gain estimates and how best to calculate them.

To reduce the risk of bias it is recommended that:

- check varieties are included to increase connectivity between varieties
- checks are used to calculate best linear unbiased estimators (BLUEs) for variety and year effects but then removed for the regression estimate
- the effect of variety age or seed source on yield is considered prior to estimating genetic gain, and if there is a distinct yield drop as seen in the NL/RL data, a term is included in the mixed model to account for this

It is also recommended that a similar analysis is redone using those varieties present in years 3, 4 and 5 to remove the effect of the yield drop in the first two years. Further research is required into reducing bias in genetic gain is required to gain a fuller understanding on the causes of variation in the estimates seen here.

# 6 Identifying the key climate drivers of interannual yield variability in winter wheat in the UK

This chapter builds on the linear mixed modelling in Chapter 5 and seasonal sensitivity analysis by Mackay *et al.* (2011). Since Mackay *et al.* (2011)'s analysis of UK National List/Recommend List (NL/RL) variety sensitivity to seasonal climate, there has been a surge in availability of high-resolution climate data and an additional decade of variety trial results. Here NL/RL data from 1988-2017 is combined with site-specific influential climate variables identified in Chapter 4 and quantify the impact of key climate drivers of interannual yield variability in UK winter wheat (*Triticum aestivum* L.) variety trials.

Initially, regional seasonal climate data is used in mixed modelling to explore the sensitivity of winter wheat yields to seasonal temperature and rainfall for the most recent decade. Regional data is then replaced by site-specific seasonal data to demonstrate the value of using high resolution, localised climate data. Significant seasonal temperature and rainfall data is dissected into monthly data to identify the most influential months within important seasons. Specially selected agroclimate metrics are then incorporated into the analysis to quantify winter wheat sensitivity to more specific variables, such as solar radiation during grain fill and April frost. Individual variety sensitivity and responses to these variables are explored to identify varieties with greater resilience to the changing climate.

## 6.1 Summer rainfall significantly affects yield

The effect of regional seasonal mean temperature and rainfall on winter wheat variety trial yields were modelled ([2.12]). Of these, only summer rainfall had a statistically significant relationship with yield, such that higher summer rainfall was associated with lower yields with a coefficient estimate of $\beta$=-0.8 (± 0.2) t/ha/100 mm ($p<0.001$) (Table 6.1) The genotype-by-environment interaction (GxE) of varieties with summer rainfall was also found to be significant and explained ~20% of the overall variability associated with the fixed effects, substantially outweighing the variation explained by climate (0.7%) (Table 6.1). This indicates winter wheat varieties had a diverse response to summer rainfall. The conditional $R^2$ and marginal $R^2$ for this model are 0.25 and 0.94, respectively with a Root Mean Squared Error (RMSE) of 0.96. The variance associated with each random effect term are in Table A4 and fitted vs. observed yield values are shown in Figure A11.

| | Sum Sq | Df | F value | p (sig.) | coef | coef SE |
|---|---|---|---|---|---|---|
| son_temp | 0.04 | 1 | 0.20 | 0.7 | -0.1 | 0.2 |
| djf_temp | 0.76 | 1 | 3.62 | 0.06 | 0.3 | 0.2 |
| mam_temp | 0.05 | 1 | 0.22 | 0.6 | -0.1 | 0.3 |
| jja_temp | 0.34 | 1 | 1.62 | 0.2 | -0.2 | 0.2 |
| son_rain | 0 | 1 | 0.00 | 1.0 | 0.0 | 0.001 |
| djf_rain | 0.3 | 1 | 1.44 | 0.2 | -0.001 | 0.0009 |
| mam_rain | 0.05 | 1 | 0.22 | 0.6 | -0.0007 | 0.002 |
| jja_rain | 3.65 | 1 | 17.37 | 0.000 (*) | -0.008 | 0.002 |
| Year | 31.9 | 30 | 5.06 | 0.000 (*) | | |
| Variety | 568.19 | 246 | 10.98 | 0.000 (*) | | |
| jja_rain:Variety | 150.13 | 246 | 2.90 | 0.000 (*) | | |

Table 6.1: *Sum of squares for fixed effects in the seasonal climate model of winter wheat, using [2.12] on the UK National List/Recommended List treated variety trials data for 1988-2018 and regional climate data from the Met Office* (Met Office, 2022c). *Fitted vs. observed values for this model are shown in Figure A11. *significant at the 95% confidence level. son = autumn, djf = winter, mam = spring and jja = summer. Coefficient estimates and standard error (SE) are given for the climate variables.*

## 6.2 Site-specific data shows winter rainfall affects yields

Inclusion of site-specific seasonal climate data into the crop-climate model [2.12] found that both winter rainfall (β=-0.4 ± 0.1 t/ha/100 mm, p=0.008) and summer rainfall (β=-0.4 ± 0.2 t/ha/100 mm, p=0.02) had significant negative relationships with yield (Table 6.2), such that higher rainfall corresponded to lower yields. GxE contributed 24% of the total variation in yield, with more variation in varietal response to summer rainfall than winter rainfall. Again, the GxE sum of squares is much higher than that for climate variability (0.7%) (Table 6.2). The conditional $R^2$ and marginal $R^2$ for this model were 0.26 and 0.94, respectively, the former a slight increase on the regional climate model with an RMSE of 0.96. The variance associated with each random effect term is shown in Table A5.

| | Sum Sq | Df | F-value | p (sig.) | coef | coef SE |
|---|---|---|---|---|---|---|
| son_temp | 0.71 | 1 | 3.35 | 0.07 | 0.3 | 0.2 |
| djf_temp | 0.32 | 1 | 1.52 | 0.2 | 0.2 | 0.1 |
| mam_temp | 0.26 | 1 | 1.25 | 0.3 | -0.2 | 0.2 |
| jja_temp | 0.58 | 1 | 2.75 | 0.1 | -0.2 | 0.1 |
| son_rain | 0 | 1 | 0.00 | 1.0 | -0.00005 | 0.001 |
| djf_rain | 1.47 | 1 | 6.98 | 0.008 (*) | -0.004 | 0.001 |
| mam_rain | 0.23 | 1 | 1.09 | 0.3 | 0.001 | 0.001 |
| jja_rain | 2.88 | 1 | 13.69 | 0.000 (*) | -0.004 | 0.002 |
| Year | 32.78 | 30 | 5.19 | 0.000 (*) | | |
| Variety | 626.38 | 246 | 12.09 | 0.000 (*) | | |
| djf_rain:Variety | 103.61 | 246 | 2.00 | 0.000 (*) | | |
| jja_rain:Variety | 110.27 | 246 | 2.13 | 0.000 (*) | | |

Table 6.2*: Sum of squares for fixed effects in the seasonal climate model of winter wheat, using [2.12] on the UK National List/Recommended List treated variety trials data for 1988-2018 and site-specific climate data, extracted from HadUK* (Hollis *et al.*, 2019). **significant at the 95% confidence level. son = autumn, djf = winter, mam = spring and jja = summer. Coefficient estimates and standard error (SE) are given for the climate variables.*

Inclusion of soil type in the model, as in [2.13], reduced the variation in yield attributed to variety (Table 6.3) and assigned it instead to the random interaction term between variety and soil type. In doing so, summer temperature also becomes significant when the variety x climate interaction terms aren't present. Adding in the variety x climate interaction terms for the significant climate terms summer and winter rainfall, and summer temperature, the variation in yield attributed to summer temperature was partially absorbed into the interaction with variety, resulting in the sum of squares in (Table 6.3). The marginal $R^2$ and conditional $R^2$ for this model were 0.27 and 0.94 respectively and the model had an RMSE of 0.95. The variance associated with each random effect term is shown in Table A6.

| | Sum Sq | Df | F-value | p (sig.) | coef | coef SE |
|---|---|---|---|---|---|---|
| son_temp | 0.25 | 1 | 1.30 | 0.3 | 0.2 | 0.2 |
| djf_temp | 0.56 | 1 | 2.91 | 0.09 | 0.2 | 0.1 |
| mam_temp | 0.02 | 1 | 0.08 | 0.8 | -0.06 | 0.2 |
| jja_temp | 0.72 | 1 | 3.70 | 0.06 | -0.2 | 0.2 |
| son_rain | 0.01 | 1 | 0.07 | 0.8 | 0.0003 | 0.001 |
| djf_rain | 1.44 | 1 | 7.44 | 0.006 (*) | -0.004 | 0.001 |
| mam_rain | 0.12 | 1 | 0.64 | 0.4 | 0.001 | 0.001 |
| jja_rain | 2.48 | 1 | 12.81 | 0.000 (*) | -0.003 | 0.002 |
| Soil_class | 2.15 | 4 | 2.77 | 0.03(*) | | |
| Year | 28.64 | 30 | 4.92 | 0.000 (*) | | |
| Variety | 510.93 | 246 | 10.71 | 0.000 (*) | | |
| jja_rain:Variety | 82.62 | 246 | 1.73 | 0.000 (*) | | |
| jja_temp:Variety | 244.36 | 246 | 5.12 | 0.000 (*) | | |
| djf_rain:Variety | 111.62 | 246 | 2.34 | 0.000 (*) | | |

Table 6.3*: Sum of squares for fixed effects in the seasonal climate model of winter wheat including soil texture, using [2.13] on the UK National List/Recommended List treated variety trials data for 1988-2018 and site-specific climate data, extracted from HadUK* (Hollis *et al.*, 2019). **significant at the 95% confidence level. son = autumn, djf = winter, mam = spring and jja = summer. Coefficient estimates and standard error (SE) are given for the climate variables.*

Winter wheat yields were highest on heavy, clay soils, averaging 10.2 t/ha, and lowest on organic soils (7.8 t/ha) (Figure 6.1), however there were so few variety trials on organic soils (64) that the overall variation in yield explained by soil was small (~0.2%), therefore in future models it was not included.

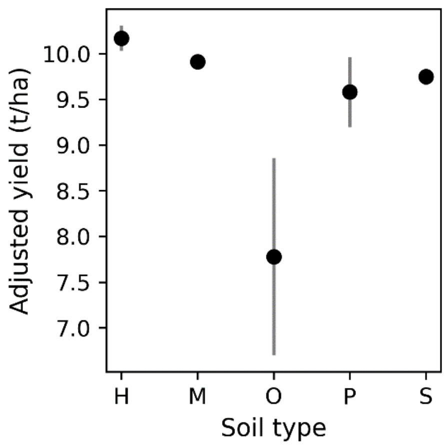



Figure 6.1*: Adjusted means and standard error for each soil type calculated from the localised seasonal climate model using equation [2.13]. H = heavy soils i.e. clay or deep clay, M = medium e.g. clay loam and sandy clay, O = organic, P = peat, S = sandy and light.*

## 6.3 Monthly rainfall is more important than temperature in determining yield

The significant seasonal variables in the site-specific analysis (Table 6.2) were then split into monthly variables to understand which months were contributing most to yield variation. Given the significance of summer temperature in the seasonal soil analysis (Table 6.3), the monthly summer temperature variables were included to test their significance individually. July, December and January rainfall were all found to have a significant relationship with yield, with higher July and January rainfall associated with yield decreases of -0.5 (± 0.3) t/ha per 100 mm of rainfall and higher December rainfall contributing to increases in yield of 0.1 (± 0.3) t/ha per 100 mm (Table 6.4). The large standard error relative to the coefficient estimate for December rainfall suggests low precision in the estimated yield impact, which may help explain why January and December rainfall seem to have conflicting yield impacts. Overall, these results suggest that in the UK, particularly in the summer and winter, rainfall is a more important determinant of yield than mean temperature. The marginal $R^2$ and conditional $R^2$ for this model were 0.29 and 0.94, respectively, with an RMSE of 0.95. The variance associated with each random effect term are in Table A7.

| | Sum Sq | Df | F-value | p (sig.) | coef | coef SE |
|---|---|---|---|---|---|---|
| Jun_rain | 0.13 | 1 | 0.615 | 0.4 | -0.006 | 0.003 |
| Jul_rain | 0.84 | 1 | 3.994 | 0.05 (*) | -0.005 | 0.003 |
| Aug_rain | 0.46 | 1 | 2.183 | 0.1 | -0.003 | 0.002 |
| Jun_tmean | 0.21 | 1 | 1.027 | 0.3 | -0.11 | 0.1 |
| Jul_tmean | 0.29 | 1 | 1.389 | 0.2 | -0.1 | 0.1 |
| Aug_tmean | 0.27 | 1 | 1.310 | 0.3 | 0.1 | 0.1 |
| Dec_rain | 1.01 | 1 | 4.848 | 0.03 (*) | 0.001 | 0.003 |
| Jan_rain | 1 | 1 | 4.800 | 0.03 (*) | -0.005 | 0.003 |
| Feb_rain | 0.28 | 1 | 1.351 | 0.2 | -0.002 | 0.002 |
| Year | 31.34 | 30 | 4.996 | 0.000 (*) | | |
| Variety | 325.02 | 246 | 6.319 | 0.000 (*) | | |
| Jul_rain:Variety | 90.56 | 246 | 2.178 | 0.000 (*) | | |
| Dec_rain:Variety | 70.67 | 246 | 1.761 | 0.000 (*) | | |
| Jan_rain:Variety | 91.47 | 246 | 1.374 | 0.000 (*) | | |

Table 6.4*: Sum of squares for fixed effects in the monthly climate model of winter wheat, using [2.12] on the UK National List/Recommended List treated variety trials data for 1988-2018 and site-specific climate data, extracted from HadUK* (Hollis *et al.*, 2019). **significant at the 95% confidence level. Coefficient estimates and standard error (SE) are given for the climate variables.*

## 6.4 Growing Degree Days accounts for most winter wheat yield variation

Agroclimate metrics from Table 2.14 for each variety trial site were combined with the NL/RL yield data and incorporated individually, along with their interaction with variety, into the linear mixed crop model [2.7]. April frost (frost04), May frost (frost05), grain fill surface incoming solar radiation (grainfillSIS), the number of 10 mm+ rain days (rain10), the two Vernalisation Degree Days (VDD) metrics (VDD from November to February, vdd_novfeb, and VDD from planting to anthesis, vdd_p2a) and Growing Degree Days (GDD) had significant effects on yields (Table 6.5).

| $C_{jk}$ | Coefficient | SS clim | p clim (sig.) | SS var x clim | p var x clim (sig.) |
|---|---|---|---|---|---|
| frost03 | 0.022 | 0.010 | 0.8 | 88.4 | 0.000 (*) |
| frost04 | -0.006 | 5.060 | 0.000 (*) | 78.8 | 0.000 (*) |
| frost05 | -0.103 | 1.976 | 0.004 (*) | 99.9 | 0.000 (*) |
| grainfill31 | -0.042 | 0.056 | 0.6 | 71.6 | 0.007 (*) |
| grainfillSIS | 0.003 | 0.885 | 0.05 (*) | 205.4 | 0.000 (*) |
| rain10 | -0.023 | 1.042 | 0.04 (*) | 121.3 | 0.000 (*) |
| rain20 | -0.065 | 0.385 | 0.2 | 137.0 | 0.000 (*) |
| SOGS | -0.006 | 0.522 | 0.1 | 170.0 | 0.000 (*) |
| vdd_novfeb | 0.002 | 4.485 | 0.000 (*) | 237.4 | 0.000 (*) |
| gdd | -0.002 | 10.072 | 0.000 (*) | 243.0 | 0.000 (*) |
| vdd_p2a | 0.004 | 20.061 | 0.000 (*) | 148.7 | 0.000 (*) |
| pe_balance | -0.002 | 0.021 | 0.8 | 158.4 | 0.000 (*) |

Table 6.5*: Univariate climate model sum of squares (SS), p-value and significance for each climate variable and the respective climate x variety interaction term, calculated using [2.12] on each single climate variable* $C_{jk}$ *paired with the treated variety trials data for 1988-2018. *significant at the 95% level.*

Significant variables in the univariate analysis (Table 6.5) were combined in a single model to identify the most influential climate variables. Due to the questions raised previously in Chapter 4, Section 4.5 about the calculation of VDD from planting to anthesis, only vdd_novfeb was included in the model. In addition to the harvest year and variety, yield variation was attributed to variation in grain fill surface incoming solar radiation, the number of 10 mm+ rainfall days, GDD and the range in responses of each variety to the different climatic conditions (Table 6.6). The marginal $R^2$ and conditional $R^2$ for this model were 0.22 and 0.92, respectively and the RMSE was 0.92.

| | Sum Sq | Df | F-value | p (sig.) |
|---|---|---|---|---|
| frost04 | 0.074 | 1 | 0.287 | 0.6 |
| frost05 | 0.590 | 1 | 2.300 | 0.1 |
| grainfillSIS | 1.121 | 1 | 4.368 | 0.04 (*) |
| rain10 | 2.121 | 1 | 8.264 | 0.004 (*) |
| gdd | 4.040 | 1 | 15.742 | 0.000 (*) |
| vdd_novfeb | 0.295 | 1 | 1.149 | 0.3 (*) |
| Year | 24.469 | 28 | 3.406 | 0.000 (*) |
| Variety | 23.233 | 19 | 4.765 | 0.000 (*) |
| frost04:Variety | 3.079 | 19 | 0.632 | 0.9 |
| frost05:Variety | 6.635 | 19 | 1.361 | 0.1 |
| grainfillSIS:Variety | 7.288 | 19 | 1.495 | 0.08 |
| rain10:Variety | 13.713 | 19 | 2.813 | 0.000 (*) |
| gdd:Variety | 19.436 | 19 | 3.986 | 0.000 (*) |
| vdd_novfeb:Variety | 10.736 | 19 | 2.202 | 0.002 (*) |

Table 6.6*: Sum of squares for fixed effects in the multivariate agroclimate model of winter wheat, using [2.12] on the UK National List/Recommended List treated variety trials data for 1988-2018. *significant at the 95% confidence level.*

To obtain the multivariate model in the simplest form, backwards elimination using *step* from ***lmerTest*** was used. This method was used as it works on mixed models, unlike several of the methods trialled in Chapter 3. Due to the low correlation between the covariates and the much larger dataset, this approach was far more successful at reducing model complexity than in the Irish barley trials analysis in Chapter 3. Specifically, backwards elimination removed both the frost terms and their interactions with variety, as well as the grainfillSIS x variety interaction term. Hence the optimal model is:

$$y_{ijk} = \mu + o_{jk} + p_{10jk} + g_{jk} + d_{jk} + v_i + r_j + vp_{10jk} + vg_{jk} + vd_{jk} + vr_{ij} + s_{jk} + e_{ijk} \quad [6.1]$$

with model terms defined in Table 6.7.

| Variable | Variable description | Fixed (F) or random (R) |
|---|---|---|
| $y_{ijk}$ | yield of variety $i$ in growing season $j$ at site $k$ | |
| $\mu$ | $\mu$ is the overall trial series mean | F |
| $o_{jk}$ | the effect of grain fill surface solar radiation in growing season $j$ at site $k$ | F |
| $p_{10jk}$ | the effect of the number of 10 mm+ rain days in growing season $j$ at site $k$ | F |
| $g_{jk}$ | the effect of the available GDD in growing season $j$ at site $k$ | F |
| $d_{jk}$ | the effect of available VDD from November to February in growing season $j$ at site $k$, | F |
| $v_i$ | the effect of variety $i$ | F |
| $r_j$ | the effect of growing season $j$ | F |
| $vp_{10jk}$ | the interaction between variety $v_i$ and the number of 10 mm+ rain days $p_{10jk}$ | F |
| $vg_{jk}$ | the interaction between variety $v_i$ and the available GDD $g_{jk}$ | F |
| $vd_{jk}$ | the interaction between variety $v_i$ and the available VDD from November to February $d_{jk}$, | F |
| $vr_{ij}$ | the effect of the interaction between variety $i$ and growing season $j$ | R |
| $s_{jk}$ | the effect of site $k$ in growing season $j$ | R |
| $e_{ijk}$ | residual term | R |

Table 6.7: *Variable description of variables in the final agroclimate model [6.1] for winter wheat. Each variable is fitted as a fixed effect (F) or a random effect (R).*

The marginal $R^2$ and conditional $R^2$ for this model were 0.22 and 0.92, respectively and RMSE was 0.92, so very similar values to the more complex model (Table 6.6) before using backwards elimination. Comparisons of this model to the base model that doesn't contain any climate variable or climate x variety interaction term ([2.7]) shows that the agroclimate model [6.1] is significantly better at capturing the data than the simpler model [2.7] ($\chi^2$=291.4, Df=61, p<0.001). To provide context to the results and model coefficients (Tables 6.8 and A7), the mean and categorised values for each variable for the varieties and sites in this model are given (Table 6.9).

Analysis of variance of the model covariates show that of the four climate covariates, GDD accounts for the most variation in yield (Table 6.8). This was accompanied by a very small negative model coefficient, indicating that a big increase in GDD is associated with small decreases in yield. There was large variation in yield responses to GDD, with varietal response to different GDD availability accounting for ~18% in overall yield variation in this model. Increases in the number of 10 mm+ rain days decreased yields by ~0.13 t/ha per 10 extra 10 mm+ rain days. Higher solar radiation during grain fill increased yields by ~0.3 t/ha per extra 100 MJ/$m^2$. However, the interaction term with variety was dropped in the backwards elimination process, suggesting the yield response between different varieties to solar radiation during grain fill is not significantly different to each other, rather they largely respond in a similar, positive way. VDD from November to February was included in this final model but proved not to significantly affect yield (p=0.1). The variance associated with each random effect term in model [6.1] are in Table A8.

| | SS | Df | F value | p (sig.) | coef | coef SE |
|---|---|---|---|---|---|---|
| grainfillSIS | 1.043 | 1 | 4.071 | 0.04 (*) | 0.0030 | 0.001 |
| rain10 | 1.987 | 1 | 7.753 | 0.005 (*) | -0.013 | 0.008 |
| gdd | 4.123 | 1 | 16.086 | 0.000 (*) | -0.000050 | 0.0004 |
| vdd_novfeb | 0.676 | 1 | 2.637 | 0.1 | 0.00049 | 0.001 |
| Year | 26.709 | 28 | 3.722 | 0.000 (*) | *See Table A9* | |
| Variety | 36.927 | 19 | 7.584 | 0.000 (*) | | |
| rain10:Variety | 14.16 | 19 | 2.908 | 0.000 (*) | | |
| gdd:Variety | 20.968 | 19 | 4.306 | 0.000 (*) | | |
| vdd_novfeb:Variety | 11.838 | 19 | 2.431 | 0.001 (*) | | |

Table 6.8*: Sum of squares for fixed effects in the multivariate agroclimate model of winter wheat, using the optimised model [6.1]. *significant at the 95% confidence level. Coefficient estimates and standard error (SE) are given for the climate variables.*

## 6.5 Statistical modelling can be used to identify climate-resilient varieties

There was a range in yield responses (Figure 6.2, Figure A13) by each winter wheat variety (Table 2.18) to the different agroclimate conditions (low/medium/high; Table 6.9). The interaction between grain fill solar radiation and variety was included here to explore individual varietal responses to varying levels of solar radiation, despite the interaction term not being included in the final agroclimate model. The majority of varieties yielded higher under high GDD, for example *Gallant* yielded 6% more than average in years with high GDD and 7% lower than average in years with low GDD suggesting it may yield well on farms in the East and South-East where the UK receives on average the most GDD each year (Figure 4.23). There were several exceptions which contribute to the overall negative relationship between GDD and yield. For example, both *Mercia* and *Soissons* yielded highest in years of low GDD.

| Variable | Units | Mean | Categories |
|---|---|---|---|
| grainfillSIS | MJ/m$^2$ | 786.6 | 559-755, 755-815, 815-1002 |
| rain10 | Number of days | 16.2 | 3-12, 12-17, 17-89 |
| gdd | $^o$C days | 1442.7 | 861-1369, 1369-1513, 1513-2186 |
| vdd_novfeb | $^o$C days | 973 | 323-925, 925-1042, 1042-1301 |

Table 6.9*: The low/medium/high categories for each agroclimate metric included in the final model [6.1].*

Varietal yield responses to the number of 10 mm+ rain days suggest that between 12-17 10 mm+ rain days across the growing season is preferable for winter wheat (Figure 6.2). All 20 varieties, except *Mercia*, *Soissons* and *Savannah* produced their highest yields in this “medium” rainfall category, whilst these three varieties preferred less rainfall. *Mercia* also produced higher yields during growing seasons with fewer rain days of 20 mm+ and reduced water availability (pe_balance; Figure A13). Several varieties had a strong positive response to high solar radiation during grain fill, particularly *Claire*, *Cordiale*, *Deben* and *XI19* which yielded at least 5% higher than average. Given the demonstrated increase in grain fill solar radiation in the South-East and East in Chapter 4 (Figure 4.24), these varieties may benefit from growing in this area. Varieties such as *Soissons* which had less yield variation in response to solar radiation conditions could be considered for growing in the West of the UK where decreased solar radiation is expected (Figure 4.24).

Despite the insignificant relationship between Vernalisation Degree Days from November to February and yield (Table 6.8), there was a distinct yield penalty across most varieties when available VDD was in the "low" category. Yield responses for medium and high levels of VDD were very similar across most varieties, suggesting there is a minimum VDD requirement and once this is reached more VDD will not increase yields further.

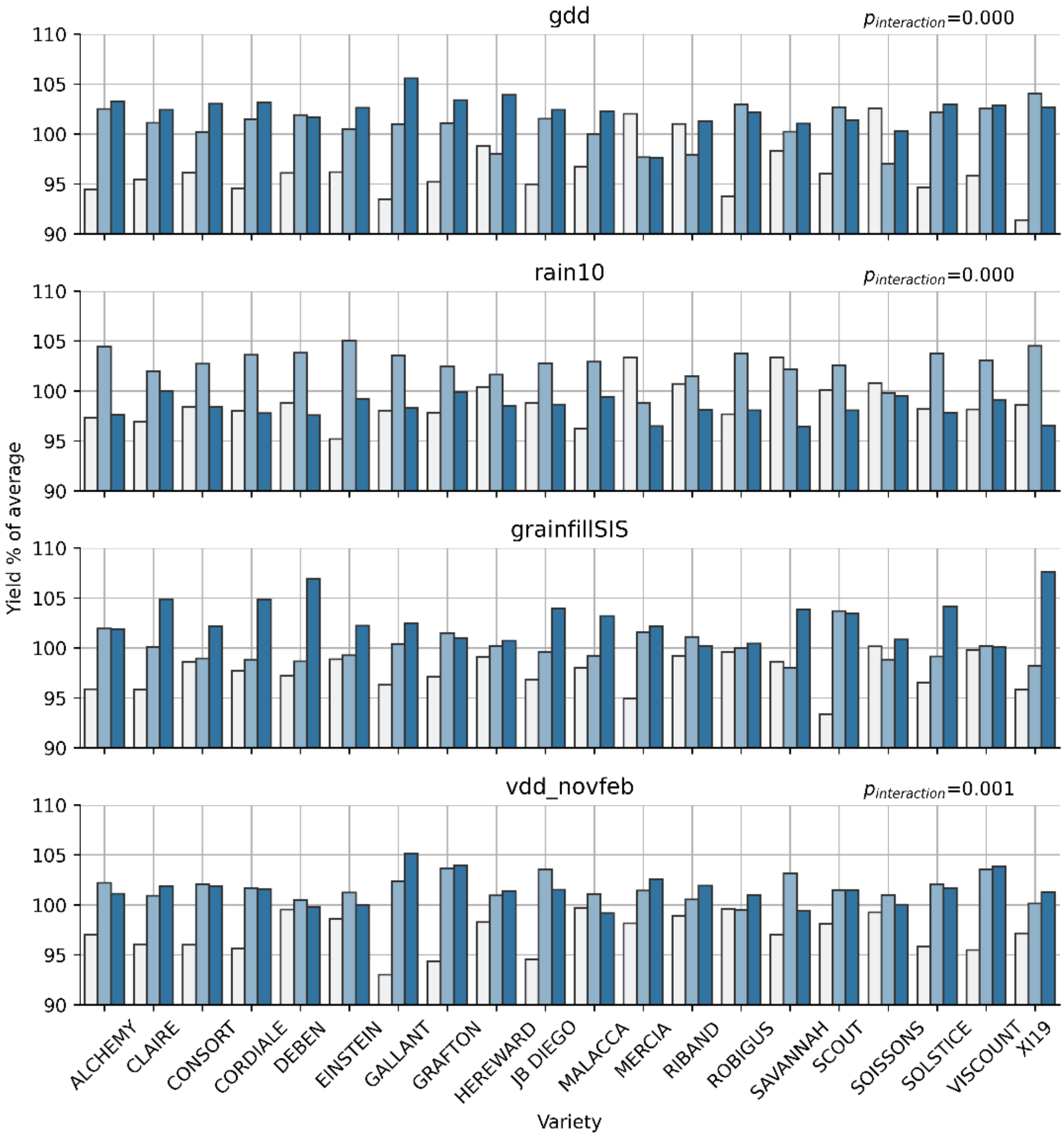



Figure 6.2*: Yield responses (% of average) of each variety to low (white), medium (light blue) and high (dark blue) (as defined in Table 6.9) Growing Degree Days (GDD), growing season 10mm+ rainfall days (rain10), surface incoming solar radiation during grain fill (grainfillSIS) and Vernalisation Degree Days from November to February (vdd_novfeb). Varieties included here are those present at least 10 years in the variety trials dataset from 1988-2017. The statistical significance of the interaction between variety and the agroclimate metric is given, except for grainfillSIS due to the interaction term not being included in the final model.*

Given the low frequency in occurrence of mild heat stress during grain fill ($T_{max}$>31°C) and air frost days in May, the yield response to these climate variables was split into two: whether they occurred or not. The influence of granfill31 on yield is minimal here, however late spring frost is more detrimental to yield (Figure 6.3). Across all varieties, the occurrence of at least one May frost day resulted in yield losses, with *Mercia* showing the smallest yield loss (<2%) and *XI19* the largest (>10%). Frost earlier on in spring (frost03 and frost04 in Figure A13) had a much less detrimental yield impact.

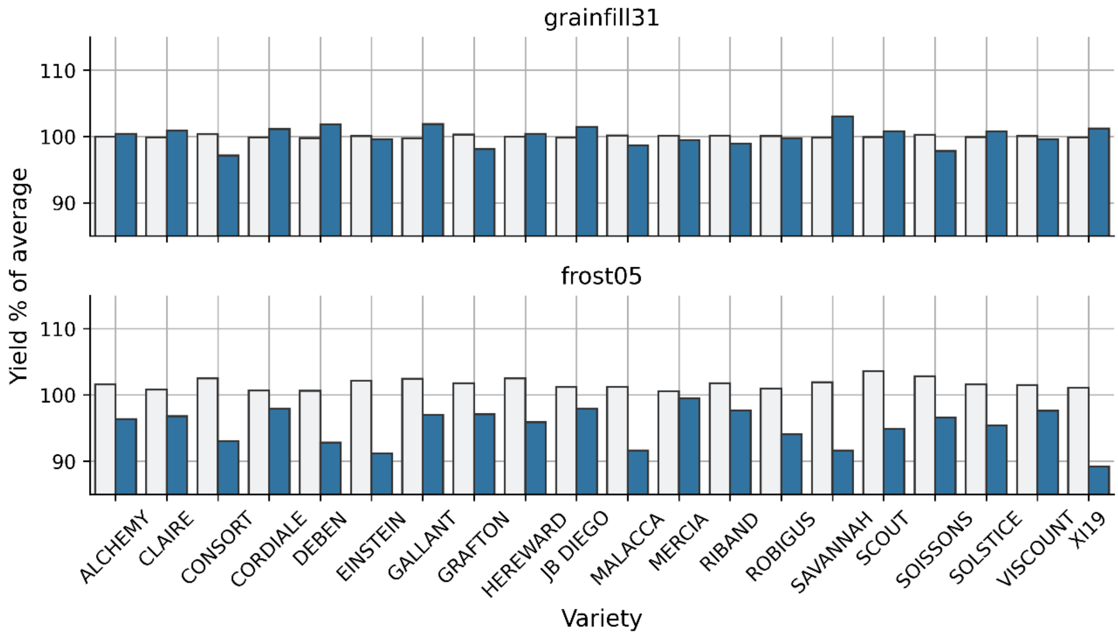



Figure 6.3*: Yield responses (% of average) of each variety to zero (white) or at least one (dark blue) occurrences of mild heat stress during grain fill (grainfill31) and air frost in May (frost05). Grain fill corresponds to the period 16th June-31st July.*

Using the statistical significance of each variety x climate interaction in the univariate climate models (Table 6.5) and a false discovery rate (FDR) < 0.5, it was possible to detect varieties which were more sensitive to specific climate variables. In total, 74 interactions were detected across 19 varieties (Table 6.10). Four varieties had FDR < 0.5 for 6 variables: *Riband*, *Scout*, *Soissons* and *XI19*. GDD was of greatest importance here, aligning with the results of the multivariate ANOVA (Table 6.8).

| Variety | frost04 (days) | frost05 (days) | grainfillSIS ($MJ/m^2$) | rain10 (days) | gdd (°C days) | vdd_novfeb (°C days) | vdd_p2a (°C days) | FDR<0.5 |
|---|---|---|---|---|---|---|---|---|
| CLAIRE | 0.02 | 0.03 | 0.0004 | 0.005 | -0.0002 | 0.0007 | -0.0002 | 2 |
| CONSORT | 0.03 | 0.1 | -0.003 | -0.01 | -0.001 | -0.0007 | 0.00002 | 5 |
| CORDIALE | 0.04 | -0.05 | -0.0004 | -0.006 | -0.0006 | 0.0006 | -0.0006 | 3 |
| DEBEN | 0.04 | 0.05 | -0.00004 | -0.004 | -0.0008 | -0.0003 | -0.0005 | 2 |
| EINSTEIN | 0.03 | -0.06 | -0.001 | 0.003 | -0.001 | -0.00008 | -0.0002 | 3 |
| GALLANT | 0.03 | -0.1 | 0.0003 | -0.003 | -0.0001 | 0.002 | -0.001 | 3 |
| GRAFTON | 0.03 | -0.04 | -0.0003 | 0.01 | -0.0006 | 0.0007 | 0.0002 | 4 |
| HEREWARD | 0.03 | 0.08 | -0.001 | -0.01 | -0.0007 | 0.00002 | -0.001 | 5 |
| JB DIEGO | 0.03 | -0.1 | 0.0003 | -0.002 | -0.0002 | 0.0009 | -0.001 | 3 |
| MALACCA | 0.03 | 0.06 | -0.002 | -0.004 | -0.001 | -0.001 | 0.0002 | 3 |
| MERCIA | 0.04 | 0.1 | -0.002 | -0.006 | -0.002 | -0.001 | -0.00004 | 4 |
| RIBAND | 0.07 | 0.07 | -0.003 | -0.009 | -0.001 | -0.001 | 0.0009 | 6 |
| ROBIGUS | 0.03 | -0.003 | -0.0009 | -0.005 | -0.0002 | 0.0003 | -0.0001 | 2 |
| SAVANNAH | 0.07 | 0.1 | -0.004 | -0.03 | -0.002 | -0.002 | -0.001 | 6 |
| SCOUT | 0.02 | -0.05 | -0.0004 | -0.007 | -0.0007 | 0.0001 | -0.001 | 2 |
| SOISSONS | 0.04 | 0.1 | -0.003 | -0.02 | -0.002 | -0.003 | -0.0008 | 6 |
| SOLSTICE | 0.04 | -0.00008 | -0.001 | -0.002 | -0.0008 | 0.00005 | -0.001 | 4 |
| VISCOUNT | 0.02 | 0.06 | -0.002 | -0.001 | -0.001 | -0.0006 | 0.0006 | 5 |
| XI19 | 0.03 | -0.2 | 0.001 | -0.01 | -0.0002 | 0.001 | -0.0009 | 6 |
| **FDR<0.5** | **9** | **6** | **12** | **7** | **16** | **13** | **12** | **74** |

Table 6.10: *Variety x climate variable interaction coefficient for each variety with at least 10 years of data, giving the offset from the mean response of all varieties. Interactions with a false discovery rate (FDR) < 0.5 are highlighted in grey. The total number of interactions detected with an FDR < 0.5 for each climate variable and variety are shown at the end of each column and row, respectively. Climate variables included here were significant in the univariate climate models (Table 6.5). The variety 'Alchemy' is not included here as it's not possible to extract the interaction coefficient from the model's intercept term.*

## 6.6 Discussion

The use of multi-environment trials data, specifically the NL/RL variety trials dataset, enabled the identification and quantification of specific agroclimatic influences on UK winter wheat yields. The inclusion of trial site location in the dataset allowed the data to be paired with historical site-specific seasonal climate and agroclimate data which has not previously been undertaken for the whole UK.

### 6.6.1 Site-specific climate data reveals additional winter wheat yield drivers

An initial analysis of yield responses to seasonal climate demonstrated the value that using higher resolution, localised climate data as opposed to regional or national data, can have on identifying crop-climate interactions.

The significant negative relationship between summer precipitation and yield (Tables 6.1-6.3) aligns with the results of Kettlewell *et al.* (2003) who found that higher summer precipitation negatively affected winter wheat grain specific weight, one of the determinants of grain yield. Summer rainfall can reduce specific grain weight through alternate wetting and drying causing wrinkling of the grain surface and reducing the packing efficiency of the grain (Bracken and Bailey, 1928; Kettlewell *et al.*, 2003). Furthermore, summer rainfall is inversely related to solar radiation, such that higher rainfall is generally associated with greater cloud cover, reducing available solar radiation the crop can photosynthesise during grain fill. In the multivariate agroclimate model lower grain fill solar radiation was associated with decreased yields (Table 6.8, Figure 6.2). Wetter summers also linked to increased disease risk: for example, rain-splash events in June can encourage the spread of Septoria leaf blotch (Turner *et al.*, 2021), as discussed in Section 4.4.10.

Using localised seasonal climate data instead of regional, it was possible to identify winter rainfall as an additional significant climate variable. Winters in the UK are typically the wettest season and therefore higher winter rainfall could cause lower yields through waterlogging, rather than low rainfall inducing a water deficit for the crop. As discussed in Section 1.2.5, higher than average rainfall in autumn/early winter can also encourage shallow root development, leaving the crop vulnerable to dry conditions in spring and summer. The negative relationship between winter rainfall and yield (Table 6.2) contradicts findings by Lopes (2022), who observed positive correlations between winter rainfall and wheat yield in Northern Europe. Possible explanations for the difference in results could include the difference in climates of countries within Northern Europe to the oceanic climate of the UK. Furthermore, Lopes (2022) use both national yield data and national climate summaries as opposed to site-specific variety trial yield data and climate data which may be oversimplifying the relationship. At some range, the rainfall relationships must stop being linear, as there is an optimum between being completely submerged and completely dry.

Summer temperature had a significant effect on yield when soil texture was included in the model (Table 6.3). This could be due to different soils warming up and cooling down at different rates, affecting crop growth rates, however this would need investigating further in additional models that include a climate x soil interaction term. This model showed yield differences between the soil textures, however deciphering these textures was somewhat difficult due to the wide-ranging acronyms used in recorded trials data. Soil texture is typically included in the dataset to ensure the correct quota of trials that represent farms on the same soil type (David Schafer, *pers comm.*), therefore some soil types only featured a handful of times and categorising them was challenging. Given the very few trials grown on organic soils, an additional model run would be useful to see whether soil texture still varies significantly between the more common soil types which showed more similar yields (Figure 6.1). A more reliable soil dataset would enable the interaction between soil type and the agroclimate metrics to be investigated, and how varieties respond on different soils.

In Germany, the Ackerzahl metric is often used for soil rating to quantify the fertility of the soil and is available for locations across Germany (Schachtschabel *et al.*, 1976; Piepho *et al.*, 1998). It incorporates several variables including soil type, geological age, the stage of degradation and average temperature and rainfall. It has been valuable for calculating yield potential at specific sites (Piepho *et al.*, 1998; Bönecke *et al.*, 2020). A similar metric in the UK would allow suitability mapping of varieties, as well as the calculation of yield potential, using a more comprehensive and reliable metric for soil than has been available for this analysis.

### 6.6.2 Univariate agroclimate analysis as a tool for climate variable selection

Running a mixed model for each climate variable and its variety interaction gave an initial indication as to the most important agroclimate yield drivers, and their relationship with yield (Table 6.5). Early spring frost (frost03), the start of the growing season (SOGS), the number of mild heat stress days during grain fill, the number of growing season heavy rain days (rain20) and growing season precipitation-evapotranspiration balance (pe_balance) were not significant (Table 6.5). Whilst there was variation in yield responses to March frost amongst cultivars (Table 6.5, SS var x clim = 88.4), winter wheat yields were evidently not overly sensitive to early spring frost, which is perhaps unsurprising given the high frequency of frost days in March across the UK (Figure 4.18). This contrasts the univariate analysis findings for later in spring, when more April and May frost days had significant detrimental effects on yield, with May frost decreasing yields by 0.1 t/ha per frost day. May typically coincides with the reproductive stage of winter wheat, when the size of the yield impact from frost damage is greater than any other stage (Frederiks *et al.*, 2012).

The lack of relationship of SOGS with winter wheat yield is also expected, given the definition of the metric which growing could only begin from the 6th January despite being in the ground from as early as the preceding September. When the number of 20 mm+ rain days increased, yields decreased but this effect was statistically insignificant despite a significant effect for the 10mm+ rain days. Splitting the growing season into shorter periods e.g. growth stages if phenological data was available could have facilitated a more useful and successful analysis when looking at the number of rain days and pe_balance. Whilst mild heat stress ($T_{max}$>31°C) during grain fill occurs more frequently than extreme heat stress ($T_{max}$>35°C), it perhaps hasn't occurred often enough or for long enough periods to have detectable yield effects found in other studies (Dreccer *et al.*, 2018; Ceglar *et al.*, 2019).

### 6.6.3 Combined agroclimate multivariate analysis

In the context of a warming climate, the significant negative relationship between GDD and yield (Table 6.8) is concerning given the increase in GDD availability shown in Chapter 4 across the last four decades. To fully understand the exact cause of this relationship, it would be useful to break this term down into shorter time periods to see if it's a rise in temperature, and so GDD, at a particular time of year contributing to the negative yield effects, or more the accumulation across the growing season due to accelerated development. In the individual varietal responses, winter wheat varieties yielded higher, on the whole, when above the low threshold of 1369 GDD and then yields didn't increase in the higher category, seemingly plateauing (Figure 6.2). *Soissons* was one exception that showed higher yields in years of low GDD. This early flowering French variety might be performing better in years of low GDD because it flowers and matures early and can escape a cold, wet summer which suppresses yields of other varieties.

A similar pattern is seen with Vernalisation Degree Days in November-February, such that after the low threshold, yield doesn't increase further between medium and high vdd_novfeb (Figure 6.2). The fact that this vernalisation term is not significant in the final multivariate model (Table 6.8) despite the very consistent yield response seen (Figure 6.2) could be because of the positive correlation between VDD and GDD (0.5, Figure A12) such that GDD might be explaining the same variation. This could be explored further by partitioning GDD into shorter time periods as previously described. The positive response of winter wheat to solar radiation during grain fill complements the findings in Figure 4.25 that in most areas of the country there is also a positive relationship with yield, regardless of the genotype grown.

### 6.6.4 Individual varietal response to climate variability

The investigation of individual varietal yield response to climate variability wasn't just restricted to current varieties on the Recommended List, but also included older varieties no longer grown. Breeders can still use the sensitivity results from this research and use old varieties as parent material if they have demonstrated a desirable trait. The individual varietal responses to agroclimate variability (Section 6.6.4) are very dependent on the range of conditions experienced by each variety. Given that varieties are trialled across different sites in different years, no two varieties are likely to have experienced the same range of weather and climate variability. Unlike Mackay *et al.* (2011) who used varieties with just three years of data or more, here varieties had at least 10 years of data to ensure they had experienced a wider range of weather and climate variability which improved the chances of the rarer agroclimate events occurring e.g. May frost.

Categorising agroclimate metrics into low, medium and high (Hakala *et al.*, 2012) was a useful way of seeing how different varieties responded to variation in the metrics. For example, *XI19* yielded higher under low and medium March frost days (Figure A13), indicating it may be better suited to the East of the country, where March frost days were shown to be decreasing (Figure 4.18). *Mercia* yielded higher in growing seasons with fewer rain days of 20 mm+ and lower overall water availability (Figure A13), suggesting it is better suited to drier parts of the country and could be more resilient to future drier summers. Yield impact of late spring frost was substantial across all varieties except for *Mercia* (Figure 6.3) which combined with the yield responses to March and April frost (Figure A13) implies it could be frost resistant. To fully understand individual variety responses to the agroclimate metrics, these results should be shared with breeders whose greater knowledge of variety characteristics should give insight into why they are responding in these ways.

To check the stability of the yield response method it would be interesting to increase the number of weather categories, whilst being mindful of the limited amount of yield data that would be contributing to each category and see if the observed relationship remains the same. This could reveal any thresholds above which yields no longer increase/decrease. Furthermore, it would be useful to test the significance of the difference in yield responses observed in Figures 6.2, 6.3 and A11 for each variety to each agroclimate metric.

The false discovery rate (FDR) method was useful in identifying the most sensitive varieties and their significant relationships with yield. However, care must be taken when interpreting the variety x climate interaction coefficients (Table 6.10). For example, there was a significant positive yield response of winter wheat to increasing solar radiation during grain fill in both the multivariate analysis (Table 6.8) and univariate analysis (Table 6.5). However, the individual varietal responses were mostly negative (Table 6.10), which alone would incorrectly imply that increases in solar radiation during grain fill decreases their yields. Rather, these individual varietal responses are relative to the overall response of winter wheat to solar radiation during grain fill so if combined for each variety, show that most varieties still respond positively to this variable.

One of the limitations of this research is the use of static periods e.g. for grain fill. Given the wide range in drilling dates seen (Figure 4.11) and the range in climates and GDD experienced by crops across the country, periods such as anthesis and grain fill can occur several weeks apart at different sites in different years. Furthermore, the benchmark period for grain filling is 45 days, but can be as short as 28 days in severe drought conditions (AHDB Cereals & Oilseeds, 2018c). Given the lack of historical phenological observations for crops in the UK, an attempt was made to calculate anthesis for each variety trial (Figure 4.17) using a thermal degree day value given by AHDB. However, this was shown to have major weaknesses, giving estimated anthesis dates that were close to the actual harvest date. Hence it was decided to use a static period in this modelling.

The implication of the results from analysis of trials data should be applicable to on-farm winter wheat yields as well. Trials data is useful as it removes some of the noise created by management, allowing these crop-climate relationships to be extracted. A similar analysis could be repeated using on-farm yield data to check this.

To understand variety sensitivity and yield response to additional climate variables, this modelling can be repeated with these variables. From this, a picture should build up on exactly how different varieties respond to different weather and climate events, and therefore where they might be best suited across the UK.

The AHDB variety selection tool (https://ahdb.org.uk/variety-selection-tool) is an excellent resource for helping growers to select varieties based on various disease and agronomic factors, such as desired disease resistance, end-use of the crop and latest safe sowing date. However, there is no way of accounting for the grower's local climate in the selection tool. Ultimately results from agroclimate sensitivity analyses such as those presented here can enhance selection tools to allow improved variety suitability mapping that considers varietal preference to different climates. A useful additional piece of research would be to rerun these agroclimate models on quality traits, such as protein content, to allow the effect of climate on these metrics to be further explored and subsequently used in the tool.

How future climate will play out is uncertain: to be prepared we need genetic diversity at hand and improved modelling that considers the interactions between genotype and environment, such as has been demonstrated in this chapter. Furthermore, as the climate changes, varieties need to be trialled in environments where projected UK weather already occurs e.g. southern France, and current varieties grown successfully there could be seen as potential future varieties for the UK, especially the South-East. International collaboration is essential in achieving future food security.

## 6.7 Conclusion

This research has gone beyond the simplistic seasonal climate sensitivity analysis by Mackay *et al.* (2011) and demonstrated the opportunities provided by historical variety trials datasets when combined with local or site-specific weather and climate data in determining the most important climate drivers of crop performance. Use of linear mixed models allows the response of each variety to the climate metric of interest to be extracted and interpreted to help understand how individual varieties respond in different environments. There is great potential for this work to be done on a more regular basis, with an array of climate variables, to help growers and breeders identify climate-resilient varieties to grow in their area.

# 7 Conclusions and recommendations

## 7.1 Thesis summary

In 2014 Professor Tim Benton, the 2011-2016 UK Champion for Global Food Security, wrote "*We are used to having plentiful food, and at an affordable price. But the crisis in Ukraine should remind us that we cannot take Britain's food security for granted*" (Benton, 2014). Eight years later and the situation feels all too familiar. Except that in 2022, this is also in the context of post-Brexit supply chain disruptions and increased food insecurity induced by the covid-19 pandemic (Ipsos Mori and Food Standards Agency, 2022). Not to mention a more variable climate, exemplified by the record-breaking heatwave in July 2022 during which temperatures went well above several critical physiological thresholds for UK crops, resulting in enormous irrigation demand and associated (expensive) energy use for pumping at a time of drought and competition for water, culminating in substantial crop losses for crops such as potatoes and sugar beet (Riley, 2022). These have all contributed to a 40-year high in overall inflation, with food prices in June 2022 inflated by 9.8% relative to 12 months previous (Gurung *et al.*, 2022).

To reduce the UK's vulnerability to international supply chains in the future requires increased domestic production that is resilient to a changing, increasingly volatile climate. There has already been a northward migration in agroclimate zones (Ceglar *et al.*, 2019), which has resulted in a poleward shift in the latitudinal ranges of crop pests and diseases (Bebber *et al.*, 2013) as well as an increase in the frequency and length of warm and hot spells (Met Office, 2022a). There has been significant research into the effect of climate change on important food crops in the UK, but largely focused on projected, rather than observed, impacts. These have included studies on how the UK agroclimate is projected to change (e.g. Arnell & Freeman 2021; Harding *et al.* 2015; Harkness *et al.* 2020; Rivington *et al.* 2013) and how this translates into yield impacts (e.g. Cho *et al.* 2012; Semenov 2009). The limited work on past changes in the agricultural climate and how interannual variability has affected yields focused on one region of the UK (Addy *et al.*, 2020, 2021a) and on using monthly weather variables (Knight *et al.*, 2012).

This thesis has addressed the gap in historical agroclimate analysis across the UK and has documented how the UK agroclimate has changed from 1981-2020. High resolution gridded weather data has enabled a spatial analysis, revealing local variability and trends across the country rather than constrained to the UK as a whole (Chapter 4). Using this data, it is now possible to look at how the agroclimate has varied across the period at any location in the UK and assess the overall trend in each agroclimate variable (Table 2.14), paving the way toward a *'State of the UK Agroclimate'* report.

Variety trials records have been used in several countries (e.g. France, UK, Finland) to quantify the contribution of plant breeding to yield improvements and yield variability (e.g. Brisson *et al.*, 2010; Mackay *et al.*, 2011; Peltonen-Sainio, Pirjo *et al.*, 2009; Shorinola *et al.*, 2022). Building on Mackay *et al.* (2011), Section 5.1 presented a quantification of the genetic gain for 2007-2018 for winter wheat (*Triticum aestivum* L.), winter barley (*Hordeum vulgare* L.) and spring barley in the UK and evidenced how breeding has continued to positively contribute to linear increases in yields for both treated and untreated variety trials (Table 5.1, Figure 5.2). The inflated untreated variety trial genetic gain values discussed in in the literature (Mackay *et al.*, 2011; Shorinola *et al.*, 2022) was also seen in Chapter 5 and discussed in Section 5.5.1. Given the importance of genetic gain in determining the success of breeding programmes and future funding allocation (Covarrubias-Pazaran, 2020; Covarrubias-Pazaran *et al.*, 2022), the robustness of the metric was subsequently tested and demonstrated (Section 5.3) in a novel analysis using subsets of UK National List/Recommended List (NL/RL) variety trials data and showed high sensitivity of the metric to the connectivity of the varieties, seed source (i.e. NL or RL) and the specific long-term varieties used.

Combining multi-environment variety trials data with climate data also provides great opportunity to understand how crops respond at a species level to different climatic conditions, as well as how individual varieties interact with the environment and the overall climate resilience of current crops (Kahiluoto *et al.*, 2019; Peltonen-Sainio, P. *et al.*, 2007). In Chapter 3, pairing the best available weather and climate data of the period with barley trials data documented by Student (1923) demonstrated that much can be learnt about varietal response to different weather, even using variety trials data from as early as the start of the 20$^{th}$ century. Whilst the limited number of sites and years restricted model performance, analysis of variability in the early 20$^{th}$ century Irish agroclimate (Section 3.1.3) helped explain up to 58% of the high interannual variation in yield and price (Figures 3.4 and 3.6). The results of this chapter have been submitted for publication (Raymond *et al.*, in press). This initial mixed modelling exercise was also a good opportunity to test several variable selection methods on a small dataset before modelling the much larger UK NL/RL variety trials dataset.

The NL/RL data had previously only been correlated with national seasonal temperature data (Mackay *et al.*, 2011), therefore the crop-climate statistical modelling in Chapter 6 (Section 6.1.1) is the first documented statistical model combining NL/RL data with climate data in the UK. Comparing the results of regional (Section 6.1.1) and site-specific seasonal (Section 6.1.2) climate data highlighted the value of using site-specific climate data in analyses like this. This was developed further by incorporating site-specific agroclimate variables to identify which of these most strongly determine historical winter wheat yields in the UK, as well as the sensitivity of specific varieties. The results of this modelling exercise, along with analysis of the changing UK agroclimate, can be used to inform breeders and growers of the best growing climate for each variety included in the analysis. This research was co-funded by the British Society of Plant Breeders (BSPB) who therefore provide a direct route for impact from this research. The extent to which the work in this thesis can help improve climate resilience of UK cereals will likely be revealed in future years in the observed yield variability and depend on the extent to which the information is used to make climate-informed decisions on crop and variety choice.

## 7.2 Addressing the overarching research questions

### 7.2.1 After the 'yield plateau' of the 1990s and 2000s, what do we now see emerging in yield records?

From the 1990s, as documented in Section 1.1.2, yields began stagnating in key crops globally including wheat, maize and rice (Brisson *et al.*, 2010; Cassman *et al.*, 2011; Hochman *et al.*, 2017; Espe *et al.*, 2018), as well as nationally (Knight *et al.*, 2012). National and regional yield trend analysis showed that high wheat and barley yield variability, of up to 50%, has dominated the most recent decade 2011-2020 (Figures 4.1 and 4.2; Section 4.1), with increased yield potential realised in some years but also greater yield losses in others contributing overall to further decadal yield stagnation, on average. In Chapter 4, national and variety trial yield anomalies (Section 4.2) were combined with national climate data to help explain some of the observed yield variation (Section 4.3). 'Good' high yielding years were characterised by higher-than-average June and July sunshine, a cooler growing season and a drier autumn (Figure 4.7) 'Bad' low yielding years typically had either a very wet autumn, or a wet summer (Figure 4.8). This further highlighted the increasing challenges which farmers are facing in coping with both inter- and intra-annual weather and climate variability.

### 7.2.2 How can we use variety trial records to quantify the genetic contribution to recent yield trends?

The discrepancy between the national yield stagnation and linear increase in variety trial yields seen in Section 4.1 indicated that the yield plateau was not driven by a lack of genetic improvement. Instead, using a selection of statistical models presented in Chapter 5, it was shown that genetic improvement was responsible for over 95% of yield increases seen in national wheat yields (Figure 5.9). Indeed, crop breeding was shown to have contributed to yield increases in winter wheat, spring barley and winter barley, all three of which had positive genetic gain estimates across the period 1982/3-2018 (Figure 5.2, Table 5.1). This suggests that the genetic resource made available to farmers by breeders is underutilised, shining a light on the impact of climate variability and on aspects of agronomic practice.

Use of the UK NL/RL variety trial records also made it possible to investigate the sensitivity of this genetic gain metric. The identified sources of uncertainty went beyond the loss of disease resistance that has previously been shown to contribute to overestimates in untreated variety trial genetic gain estimates. Specifically, Chapter 5 showed that lack of connectivity between varieties (e.g. when all varieties are present <3 years in a breeding programme trials) leads to overestimates in variety effects and that inclusion of long-term "check" varieties in both steps of the genetic gain calculation (Section 2.3.3) can bias the estimate. Modelling the effect of variety age on yield revealed a small (<1%) but distinct yield difference between years one and two and subsequent years for winter wheat, indicative of a variety going from NL1 to NL2 to RL where more seed is progressively required. Thus, seed source, due to its influence on the effect of yield as a variety ages, was also found to bias the estimate. Without this modelling, the influence of seed source was masked. Hence Chapter 5 also demonstrated the power that large trials datasets offer for revealing small but important yield trends. The power to detect a difference this small (<1%) in a typical bespoke experiment is close to zero.

### 7.2.3 Which of an array of new high-resolution climate datasets should we synthesize into our analysis in order to most effectively isolate the confounding impact of spatial and temporal climate variability?

In this research, several high-resolution gridded climate datasets were used to disentangle the complex interaction between crops and their environment. The gridded nature of these datasets enabled changes in the UK climate to be seen both spatially and temporally. Of the three temperature and precipitation datasets compared in Section 2.2.3, HadUK (Hollis *et al.*, 2019) performed best, but ERA5-Land (Muñoz Sabater, 2019) and MÉRA (Gleeson *et al.*, 2017) could also have been used to create the temperature- and precipitation-derived agroclimate metrics. Gridded satellite surface incoming solar radiation (SIS) data enabled the first known spatial analysis of changes in grain fill solar radiation. These datasets underpinned Chapter 4, *The State of the UK Agroclimate,* which showed some important changes in the UK agroclimate from 1981-2020. Specifically, there has been a widespread increase in available GDD (Figure 4.14) and there are now fewer 10 mm+ rain days in September but more in August (Figure 4.20), which has likely contributed to the trend in earlier drilling dates, although this has started to reverse in recent years (Figure 4.11a.). The number of early spring frost days (Figure 4.18) and total solar radiation during grain fill (Figure 4.23) have both decreased in the East and both increased in the West.

Statistical modelling in Chapter 6 showed that high-resolution gridded climate datasets from a range of sources can be used to create agroclimate time-series for site-specific locations, which when combined with geolocated crop trait data, isolate the impacts of climate variability on both crop and individual variety performance. For example, multivariate analysis showed that winter wheat yields respond positively to increased solar radiation during grain fill (Figure 6.7). In this large-scale analysis, it was also possible to detect the genotype-by-environment interaction (GxE), which was subsequently dissected to reveal variety sensitivity to individual metrics which can be used to make localised variety recommendations. For example, both *Deben* and *XI19* showed very strong positive responses to the highest amounts of solar radiation (Figure 6.2), indicating they could be best suited to the South-East of England where grain filling solar radiation has historically been highest and has increased (Figure 4.23). There is great potential to use the methods here to explore yield responses to additional agroclimate variables and for different crops, as well as the response of other important traits, such as quality and protein content, based on the interests of a farmer or breeder.

In addition to using gridded datasets, yield and climate anomaly analysis in Chapter 4 (Section 4.3) showed that several climate resources exist that are easily accessible which require minimal processing and can aid in yield variation interpretation. The array of climate tools on the Met Office website (https://www.metoffice.gov.uk/research/climate/maps-and-data) can be integrated into yield analysis and provide a high-level indication of how interannual climate variability may have contributed to observed yield variability.

The crop-climate modelling in Chapter 3 highlighted the value of data rescue projects, such as that led by Met Éireann (Mateus *et al.*, 2020; Ryan *et al.*, 2021). By digitising and releasing historical weather records from the early 20th century, their work enabled the investigation into individual and combined impacts of variety and climate variability on spring barley trial yields in Ireland from 1901-1906. Whilst no variety x climate interaction was detected in the model, this climate data showed that *Goldthorpe* was more stable than *Archer* in a year of heavy rainfall and more soil moisture, which could be useful for breeders interested in introducing heritage varieties such as these, into their breeding programmes. Analysis like this can be completed on any multi-environment trials dataset with location data, provided both the climate and trials datasets are accessible.

## 7.3 Recommendations for stakeholders

There are several relevant audiences for whom the results presented in this thesis are of value. Primarily, growers and farmers of UK cereals can use the results from Chapter 4 *State of the UK*

*Agroclimate* to understand how the agroclimate in their area has changed to make climate-informed decisions in the near future, as well as to help explain past crop production on their farms. To make this information accessible beyond the academic community, it is intended that the results will be incorporated into an online agroclimate tool such as the UK Climate Risk Indicator Tool (uk-cri.org), with the option for stakeholders to click on a location of interest and see the 40-year time series for the agroclimate metrics used here, as well as the trend over the period. It will also be used to create an accessible report which is updated every five years. By integrating farmers and breeders into the evaluation process of the first *State of the UK Agroclimate* report, further issues can be modified to ensure it best meets stakeholder needs.

There is great potential for this agroclimate analysis to also be incorporated into the UK breeding process, as well as to encourage use of this data by agronomists and growers. As discussed in Section 6.6.4, the current Agriculture and Horticulture Development Board (AHDB) Variety Selection Tool (https://ahdb.org.uk/variety-selection-tool) does not include climate data. Combining what is known about the changing agroclimate (Chapter 4) with individual variety sensitivity to the selected agroclimate variables (Chapter 6) into the tool would enable recommendations on growing climates and therefore regions for each variety, such that if a grower inputs their location, historical climate records will indicate which varieties may be most climate-smart at their farm location, complementing the existing disease resistance ratings and agronomic factors. To best facilitate this analysis, the crop-climate modelling would need to be repeated each year to include new varieties, keep the results current and could include additional agroclimate variables as seen necessary for winter wheat and other crops. This could help tackle the difference in trends in on-farm yields and variety trial yields by encouraging use of newer, locally adapted varieties.

The combined results of the winter wheat agroclimate modelling and variety sensitivity analysis in Sections 6.1-6.5 could be used by breeders to make decisions on which traits to breed into future varieties. Furthermore, through identifying varietal resilience to different agroclimate conditions, this has also provided a possible resource from which to select parent varieties to use to create new, better adapted and more resilient varieties. Given the projected increase in extreme events, it is important farmers grow varieties with yield response diversity, therefore it might be necessary to increase the number of varieties on the AHDB Recommended List to facilitate increased diversity or create more extensive regional Recommended Lists. The rapidly changing climate will give rise to novel climates in the UK; therefore, incorporating a range of international variety trial sites in areas of countries that are already experiencing similar climates to those projected for the UK is needed urgently.

The observed changes in the agroclimate can also be used by DEFRA and government to make agroclimate-informed projections on future cropping and production totals, as well as to encourage farmers and growers to adopt new varieties which grow best in their growing environment.

The continued contribution of breeding to variety trial yield increases for winter wheat, winter barley and spring barley will be fed back to breeding companies via BSPB, the co-funders of this research, as an indicator of their recent success. However, the question raised about the robustness of genetic gain estimates in Chapter 5 also needs to reach breeding programmes and their funders. Further research should be undertaken on the least biased way of calculating genetic gain such that it doesn't over- or underestimate the value. Based on the results in Chapter 5 and discussion in Section 5.5.3, to reduce the risk of bias it is recommended that:

- Long-term "check" varieties are included to increase connectivity between varieties
- Checks are used to calculate the best linear unbiased estimators (BLUEs) for variety and year effects but then removed for the regression estimate
- The effect of variety age or seed source on yield is considered prior to estimating genetic gain, and if there is a distinct yield drop as seen in the NL/RL data, a term is included in the mixed model to account for this.

As the collators and providers of the NL/RL variety trials data, AHDB should improve the pipeline from variety trials data collection to data storage to data release. Accessing this fundamental dataset for use in this thesis took over a year. Subsequent months were dedicated to carrying out quality control as described in Section 2.2.5, such as checking for duplicates and ensuring sowing dates, harvest dates and trial site locations were realistic. There was almost double the amount of data for 2008 as every other year, and 2007 and 2009 were seemingly lacking data. Thus, half-way through the research, it was necessary for AHDB to re-extract the data for these years and for this data to then be combined with the longer record. It is estimated that the time spent during the PhD project on quality control, pre-processing and chasing the data amounts to over 1.5 years. To streamline for future researchers, this refined dataset for 1982-2018 will be shared with AHDB and published (with AHDB approval) alongside journal paper submissions. Future variety trials data should be added to this dataset once the same quality control methods have been completed. The ease of access to the Irish barley trials data made this dataset very attractive and easy to use and showed that despite its age, there is still much to learn from historical datasets. It is recommended that today's large-scale multi-environment field trial datasets are made widely available to facilitate and encourage research like this to take place to enhance our knowledge of crop-climate interactions in a changing climate.

It would be useful if AHDB kept records of cultivar trials which don't make it to trial, often as a result of weather events. As in much of Europe (Kahiluoto *et al.*, 2019), large yield losses due to weather events are currently not recorded in cultivar trials. This makes it difficult to see in future years the diversity of responses to past weather variability. For example, it can mask the effect of localised heavy rainfall causing waterlogging and crop abandonment and suggest crop yields are less vulnerable to certain weather events than they are if yields don't reflect this loss. By documenting yield losses and sharing this information with NL/RL variety trials dataset users, this can provide additional context to the data and may help explain years when the agroclimate cannot seemingly explain the yields.

The final recommendation to the AHDB and DEFRA would be to make phenological dates more available, even if it's at a regional level rather than site-specific. Records of on-farm planting and harvest dates would enable a similar analysis of on-farm data, as well as allow comparisons between trends in dates in variety trials and on-farm. A global analysis of crop planting dates does not include the UK as this data wasn't available (Sacks *et al.*, 2010). Other dates, such as the start of anthesis, would make date ranges in agroclimate metrics more representative supporting a more accurate analysis of cereal phenology.

## 7.4 Future research

This thesis has demonstrated methods to improve the climate resilience of UK cereals through better incorporation of weather and climate data into breeding programmes and UK cereal growing and highlighted to ways to improve the robustness of genetic gain estimates. Several of the issues raised warrant further investigation. Firstly, building on the results of the case study examples used to estimate the uncertainty of genetic gain in Section 5.3, further research is required into reducing bias in genetic gain to gain a fuller understanding of the causes of variation in the estimates seen here. Using a different variety trials dataset from another breeding programme (i.e. not NL/RL trials) could highlight additional factors to consider when calculating the metric.

Secondly, soil quality plays an important role in enhancing resilience to climate change and climate variability (Qiao *et al.*, 2022). Soil texture varies considerably across the country (National Soil Resources Institute, 2022) and variation in soil texture was shown to contribute significantly to yield variation (Section 6.2; Table 6.3). Due to the subjectivity in soil texture classification and the uncertainty in some of the soil texture acronyms recorded at trial sites, soil texture was dropped from the statistical models without examining any variety-climate-soil interactions. Hence with the use of accurate, site-specific soil type data it would be useful to re-run this analysis and explore the roles of soil further, incorporating the interaction between soil and climate, and soil, climate and variety into the model. This could show which areas of the UK may be more vulnerable to yield shocks in response to variability in specific agroclimate variables, as well as which varieties are better suited to specific soils in different climatic regions. By overlaying the distribution of soils and the agroclimate, suitability maps could be developed for different varieties.

Additional analysis that incorporates management options beyond drilling dates could give greater insight into options available to farmers to improve their resilience to weather and climate variability. To do so would require use of on-farm crop data or on-farm experimentation data, which has been proposed could help transform agriculture globally by allowing these interactions between management, genotype and the environment to be investigated in a real-world setting (Lacoste *et al.*, 2022).

Work in this thesis has been constrained to analysis of temperature, precipitation, and solar radiation or sunshine duration. Analysis of wind, solar radiation at other times in the growing season, and soil moisture would aid in a more complete analysis of the changing UK agroclimate. Likewise, breaking metrics, such as the number of 10 mm+ rainfall days, down into shorter sub-monthly periods would be beneficial for identifying the most vulnerable periods of the growing season to heavy rainfall.

Further research should also utilise the Standardised Precipitation and Evapotranspiration Index (SPEI) instead of the Standardised Precipitation Index (SPI). The SPEI metric incorporates a measure of evaporative demand, which is especially dependent on temperature. SPEI is sensitive to the rising global temperatures and would be a more comprehensive metric to analyse the impact of the changing climate on crop yields (Vicente-Serrano *et al.*, 2010; Price *et al.*, 2022). SPEI-12 at harvest would provide a good measure of loss (or gain) of ground water over the cereal growing season. Spatial and temporal changes in this metric could be analysed to better understand how climate change has affected crop water availability in recent decades. This metric could also be included as a variable in the statistical modelling to identify the varieties least sensitive to drought and waterlogging. Incorporation of an NAO metric in to the statistical modelling could provide an indication of some of the larger scale processes influencing historical yield variability. Furthermore, other teleconnections could also be considered, such as the El Niño

Southern Oscillation, which has already been shown to have yield effects in several countries (Heino *et al.*, 2018)

Finally, to create a more comprehensive report on *The State of the UK Agroclimate*, yield trends and anomalies for other crops grown in the UK should be analysed along with additional relevant agroclimate metrics. Due to site-specific data availability, it was only possible to complete the crop-climate statistical modelling for winter wheat. Therefore, if it was possible to access site-specific information on variety trials data for other crops, including spring and winter barley, this analysis could be repeated to improve understanding of agroclimate drivers of yield variability for other crops in the UK. This analysis does also not need to be limited to yield, but could be applied to other desirable traits, such as protein content and quality.

## 7.5 Concluding remarks

This thesis has demonstrated the relative contributions of plant breeding and the changing agroclimate to UK cereal yields. It is envisaged this work can encourage increased integration of crop-specific weather and climate data into breeding programmes, variety trial evaluation and when making crop and variety recommendations and selections. To help secure the future of UK food security and reduce the financial burden induced by high interannual yield variability, this integration will be necessary.

# Appendix

## Supplementary figures

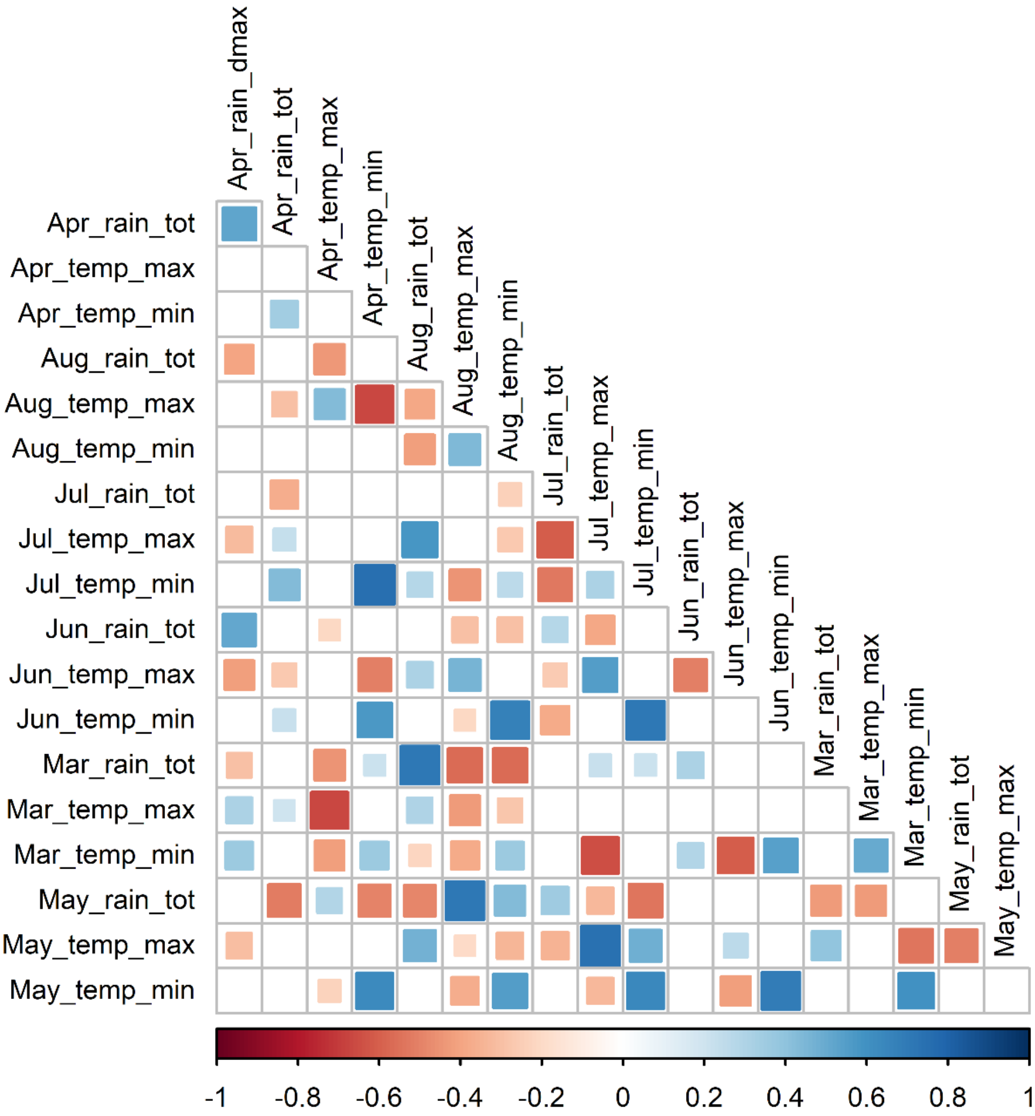



Figure A1: *Correlation matrix for Irish spring barley growing season (March-August) monthly mean maximum and mean minimum temperature and total monthly rainfall. Only significant correlations (p < 0.05) are shown. The larger the square the stronger the correlation. Dark red corresponds to strong negative correlations, dark blue corresponds to strong positive correlations.*

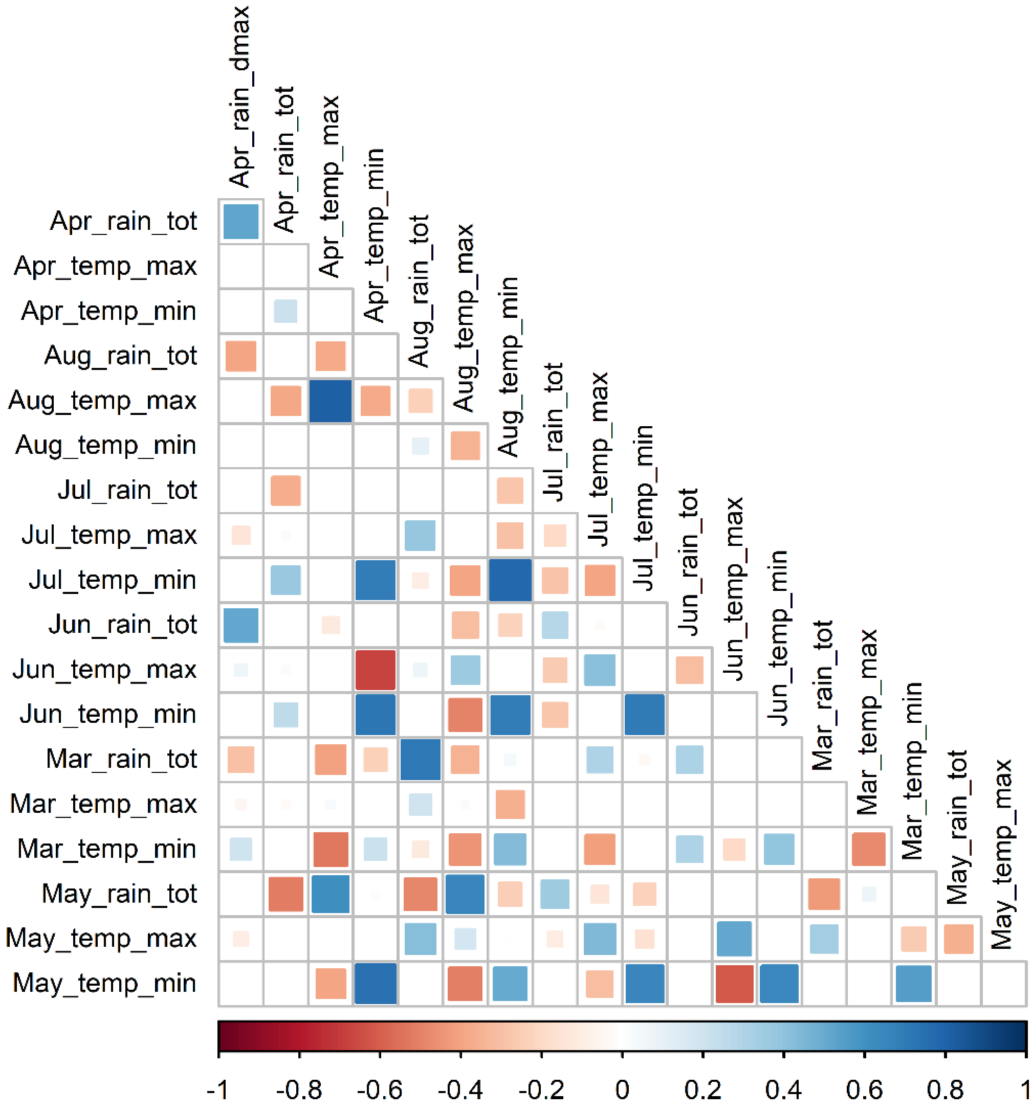



Figure A2: *Correlation matrix for Irish spring barley growing season (March-August) monthly maximum and minimum temperature and total monthly rainfall. Only significant correlations (p < 0.05) are shown. The larger the square the stronger the correlation. Dark red corresponds to strong negative correlations, dark blue corresponds to strong positive correlations.*

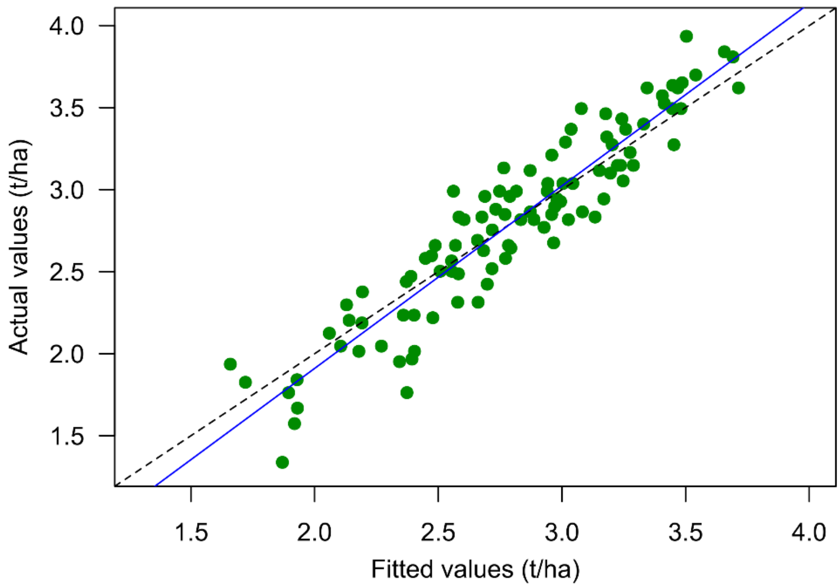
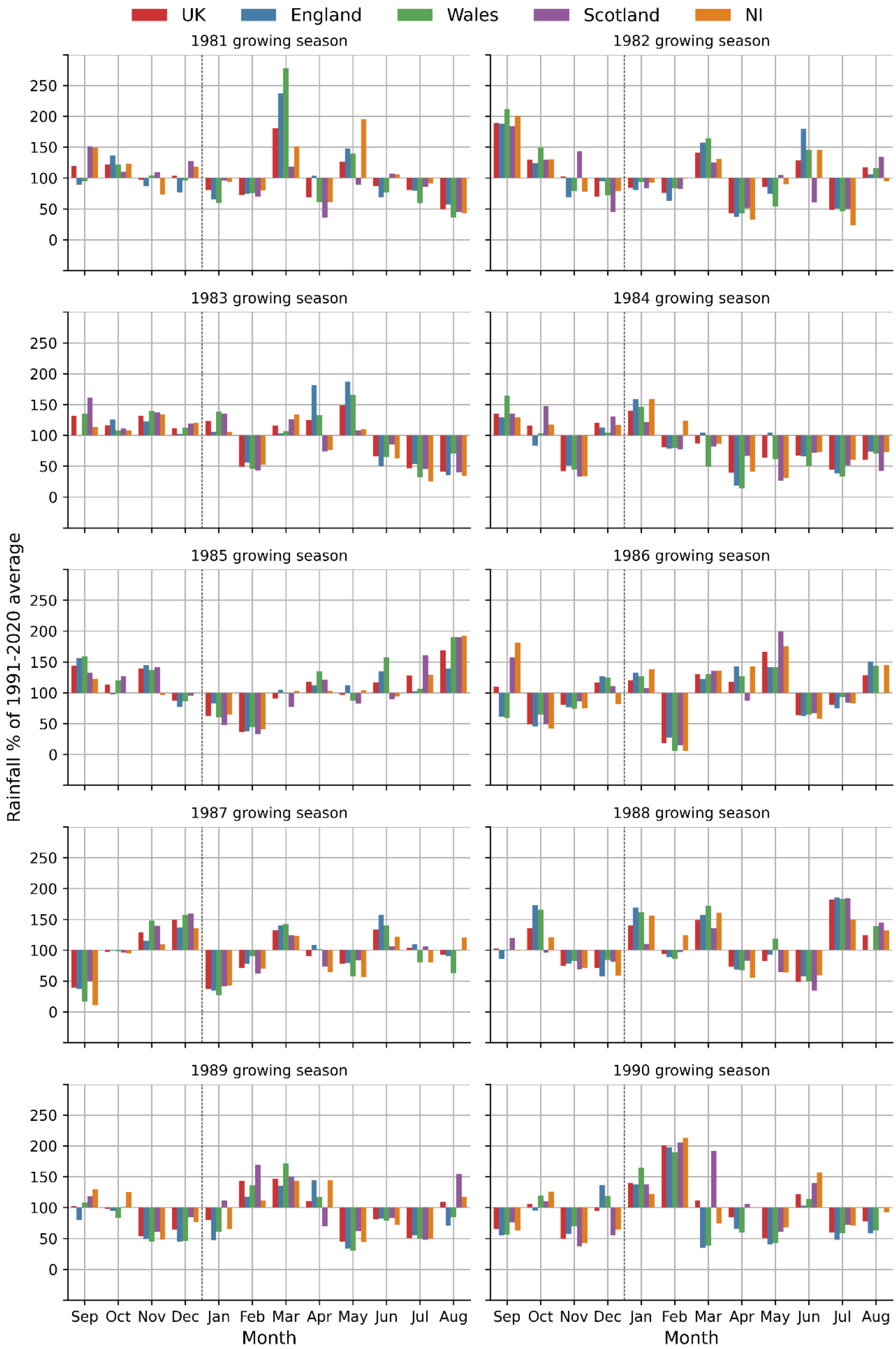
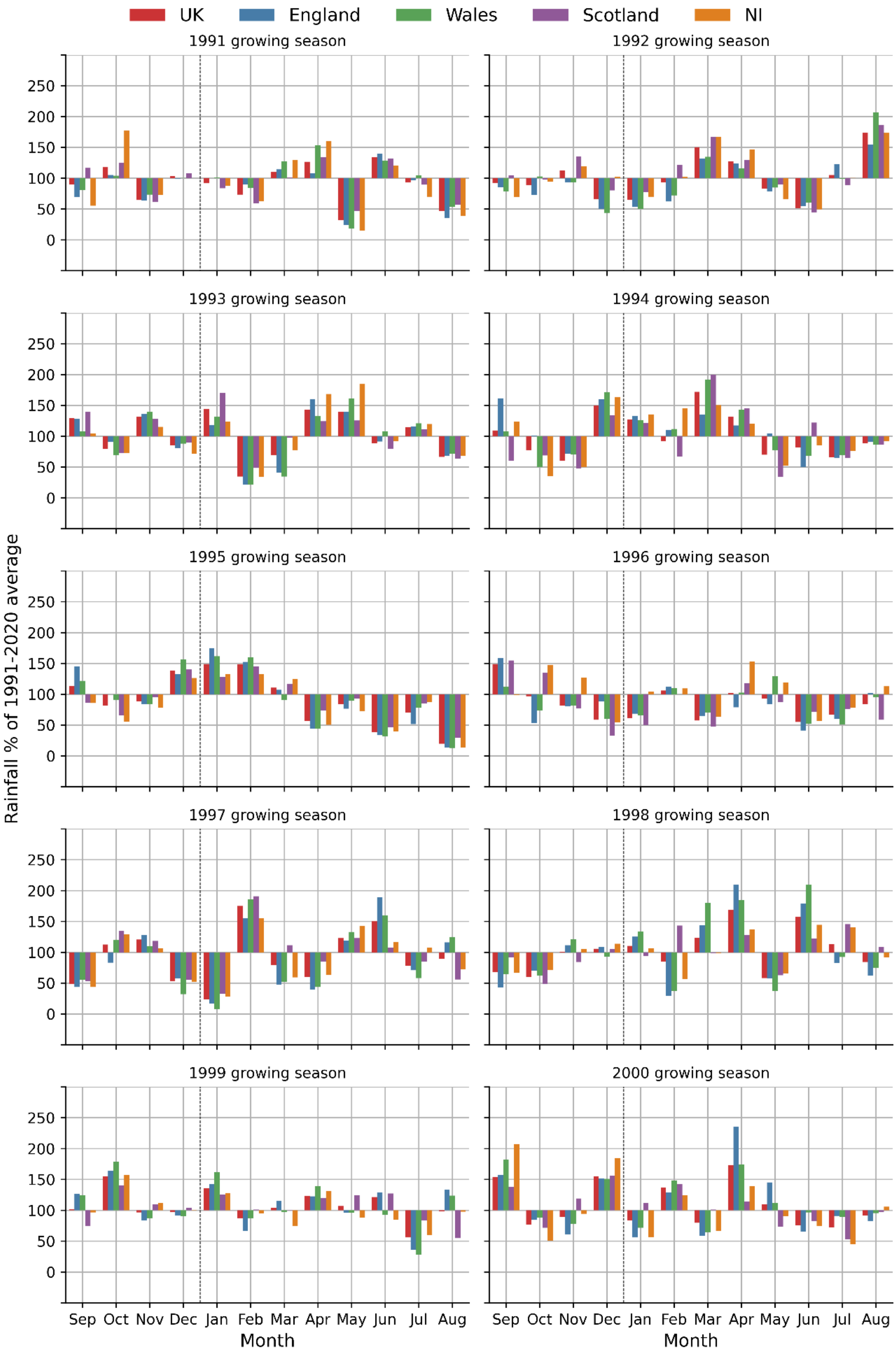
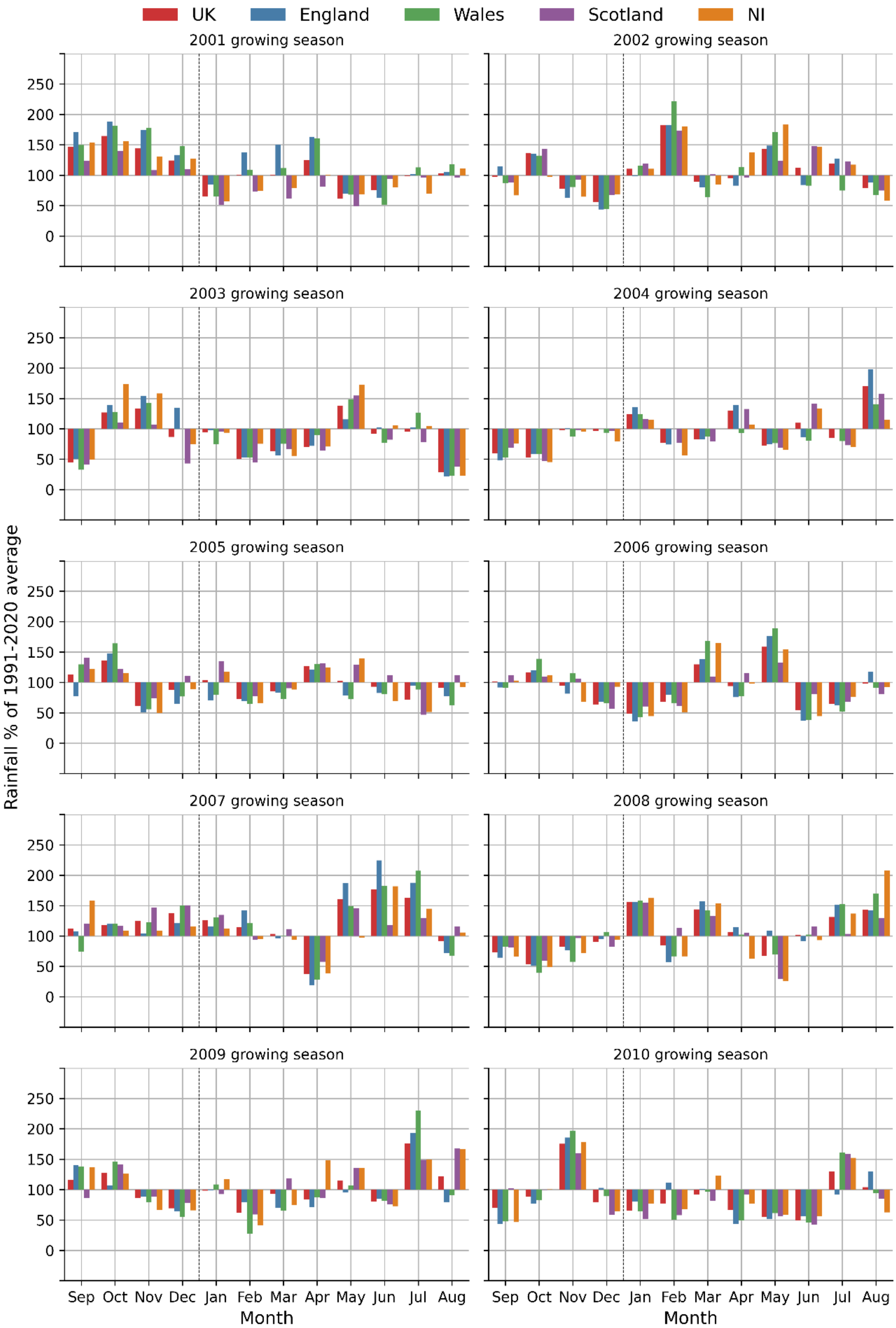



Figure A3*: Observed versus fitted values from using lmerTest on the extreme maximum and minimum monthly temperature, and total monthly rainfall data with the Irish spring barley trials data from 1901-1906. All the linear mixed model plots of observed vs. fitted values in Chapter 3 showed very similar graphs, so were excluded to avoid repetition.*


UK
England
Wales
Scotland
NI
1981 growing season
1982 growing season
1983 growing season
1984 growing season
1985 growing season
1986 growing season
1987 growing season
1988 growing season
1989 growing season
1990 growing season
Rainfall % of 1991-2020 average
250
200
150
100
50
0
Sep Oct Nov Dec Jan Feb Mar Apr May Jun Jul Aug
Month



UK
England
Wales
Scotland
NI
1991 growing season
1992 growing season
1993 growing season
1994 growing season
1995 growing season
1996 growing season
1997 growing season
1998 growing season
1999 growing season
2000 growing season
Rainfall % of 1991-2020 average
0
50
100
150
200
250
Sep Oct Nov Dec Jan Feb Mar Apr May Jun Jul Aug
Month


UK England Wales Scotland NI
2001 growing season
2002 growing season
2003 growing season
2004 growing season
2005 growing season
2006 growing season
2007 growing season
2008 growing season
2009 growing season
2010 growing season
Rainfall % of 1991-2020 average
0 50 100 150 200 250
Sep Oct Nov Dec Jan Feb Mar Apr May Jun Jul Aug
Month

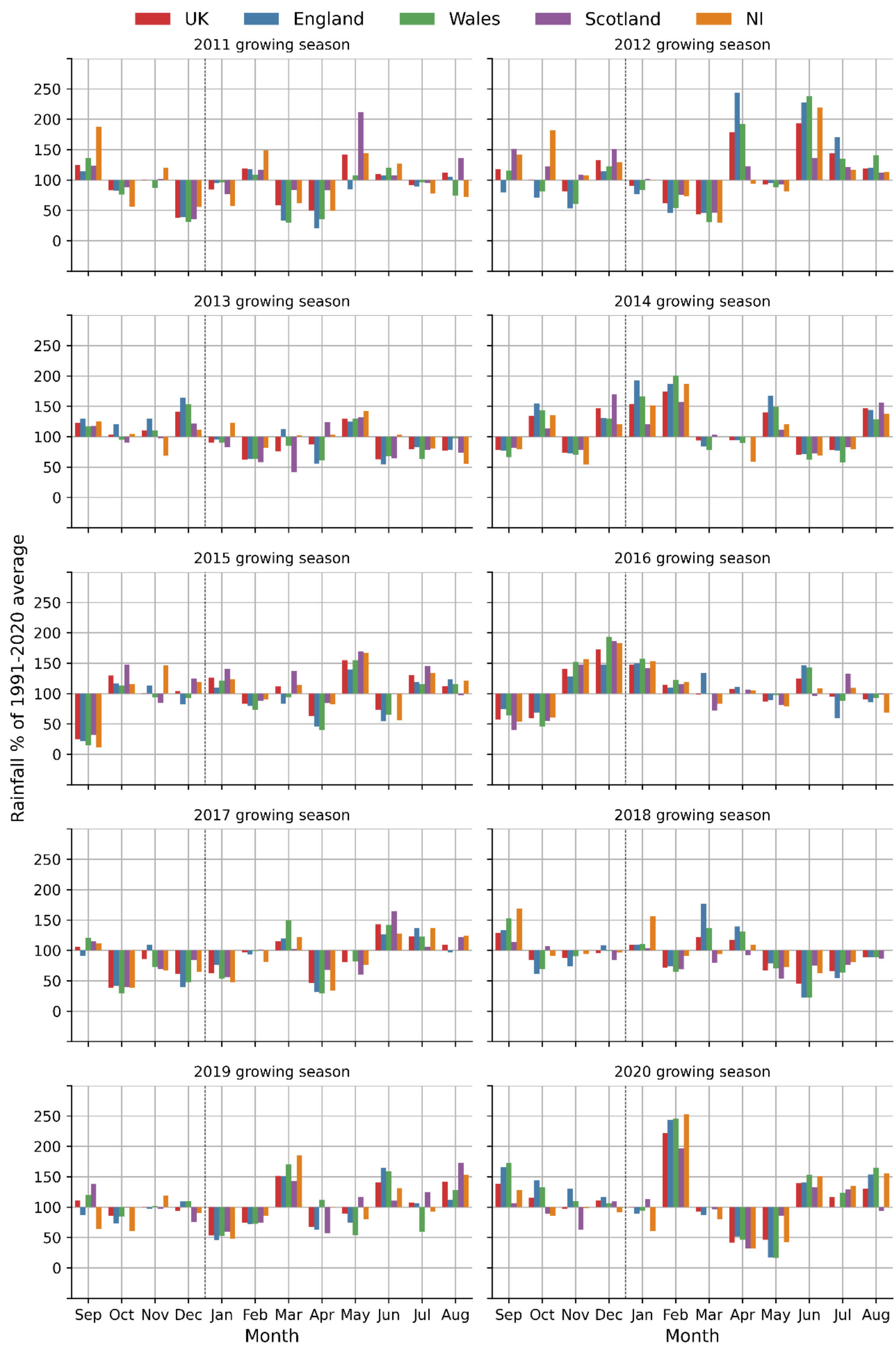
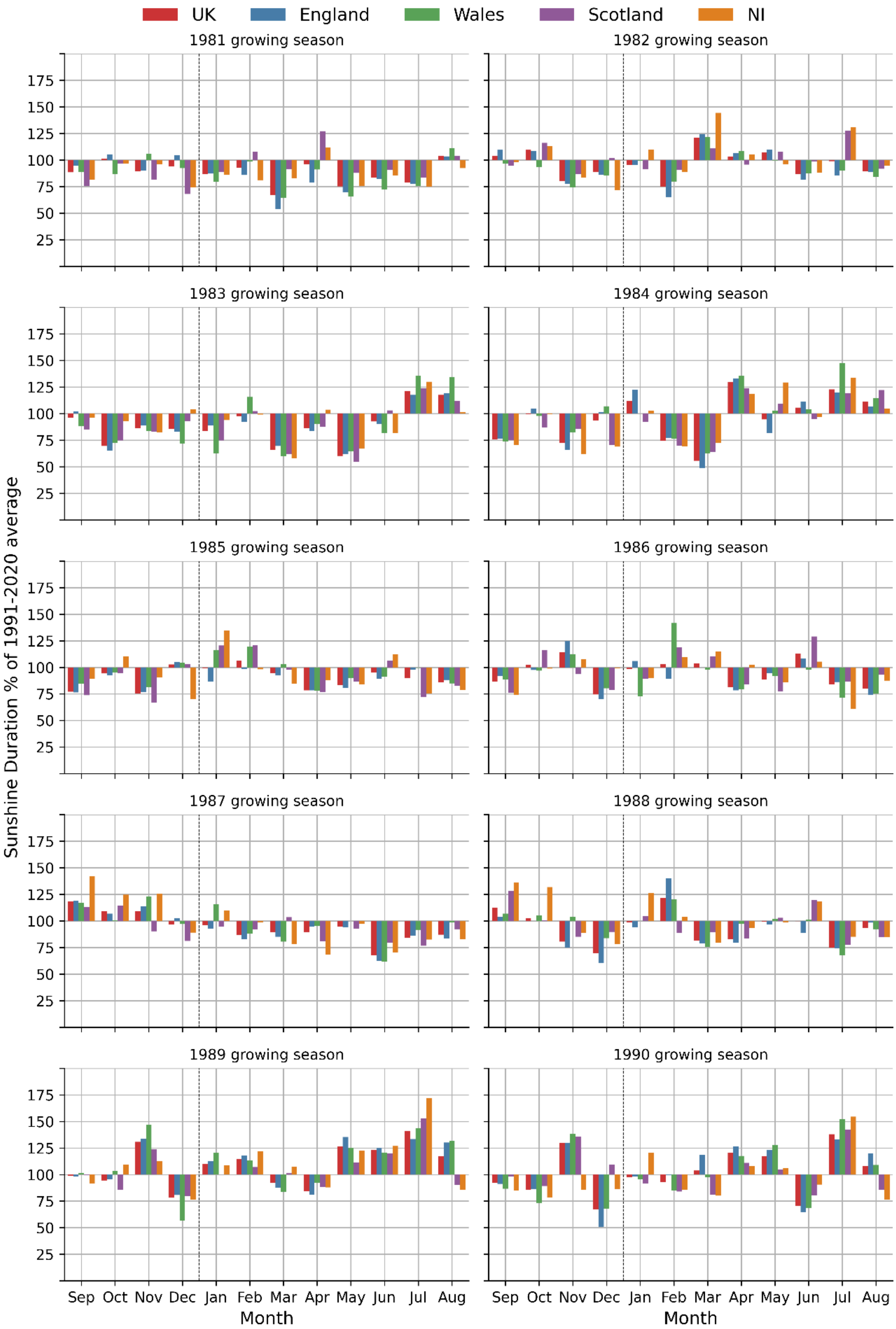
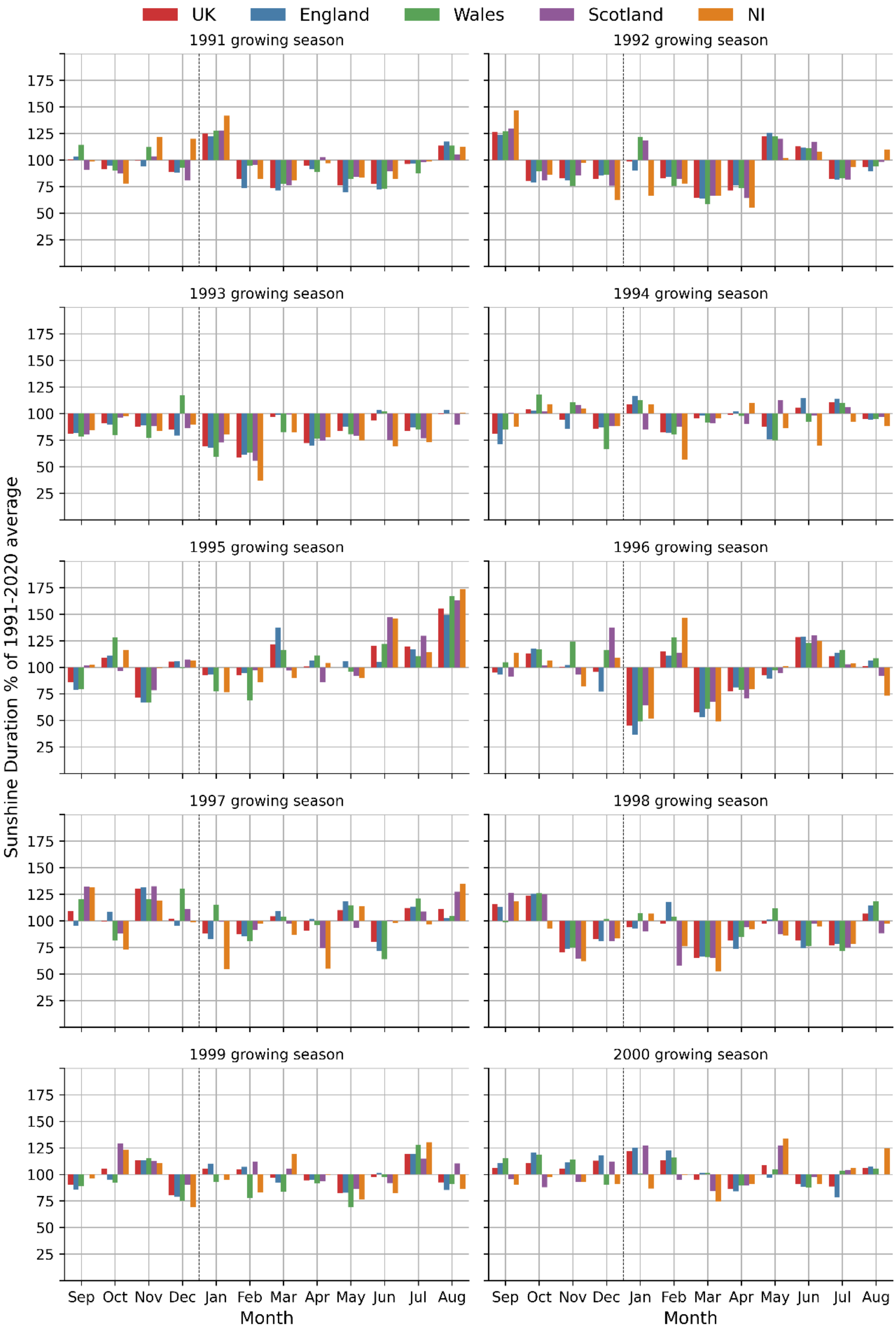


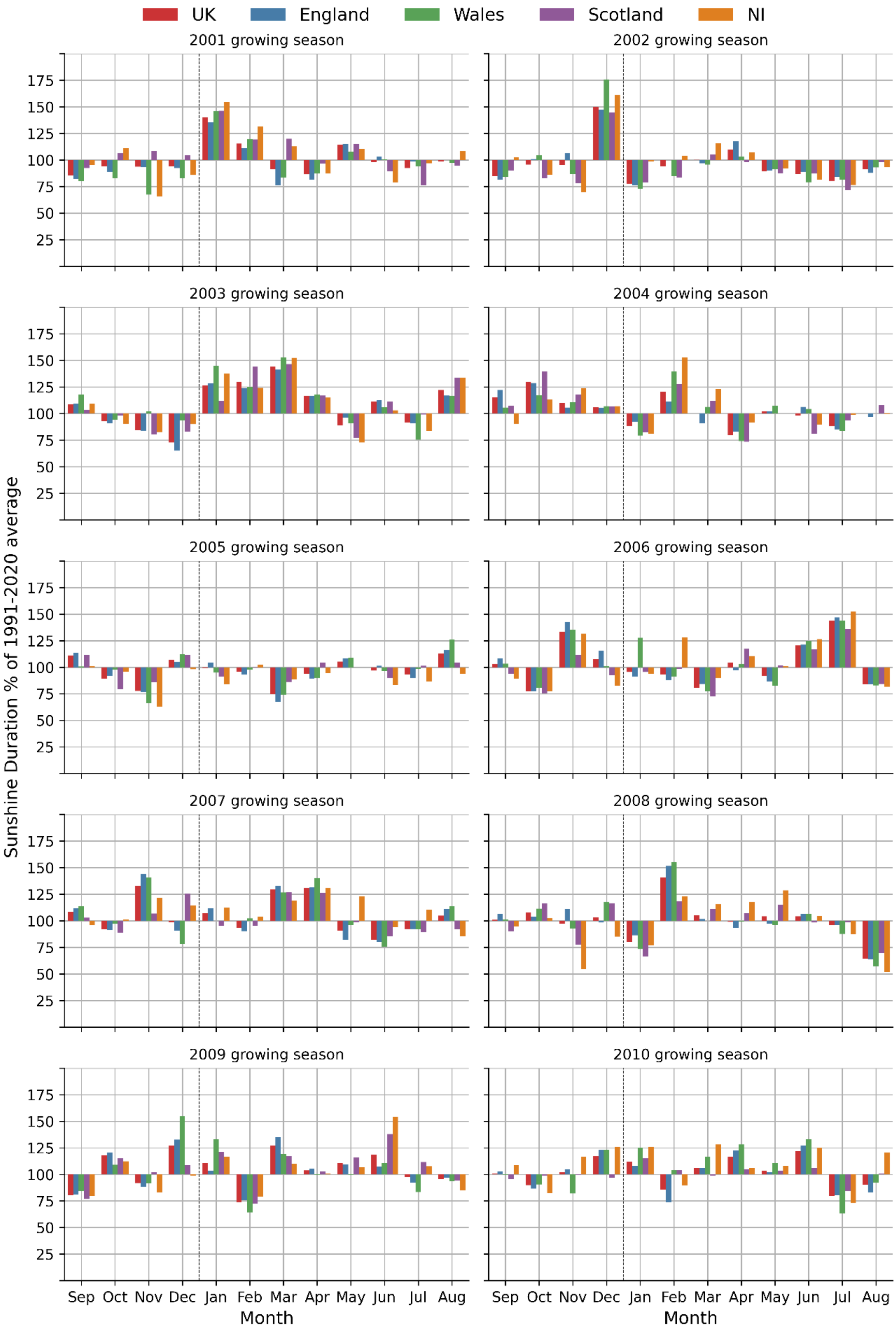

Figure A4: *Growing season rainfall anomaly (% 1991-2020 average) plots for the UK (red), England (blue), Wales (green), Scotland (purple) and Northern Ireland (NI) (orange). Data from Met Office year ordered time-series.*



UK
England
Wales
Scotland
NI
1981 growing season
1982 growing season
1983 growing season
1984 growing season
1985 growing season
1986 growing season
1987 growing season
1988 growing season
1989 growing season
1990 growing season
Sunshine Duration % of 1991-2020 average
175
150
125
100
75
50
25
Sep Oct Nov Dec Jan Feb Mar Apr May Jun Jul Aug
Month



UK
England
Wales
Scotland
NI
1991 growing season
1992 growing season
1993 growing season
1994 growing season
1995 growing season
1996 growing season
1997 growing season
1998 growing season
1999 growing season
2000 growing season
Sunshine Duration % of 1991-2020 average
175
150
125
100
75
50
25
Sep Oct Nov Dec Jan Feb Mar Apr May Jun Jul Aug
Month


UK
England
Wales
Scotland
NI
2001 growing season
2002 growing season
2003 growing season
2004 growing season
2005 growing season
2006 growing season
2007 growing season
2008 growing season
2009 growing season
2010 growing season
Sunshine Duration % of 1991-2020 average
175
150
125
100
75
50
25
Sep Oct Nov Dec Jan Feb Mar Apr May Jun Jul Aug
Month

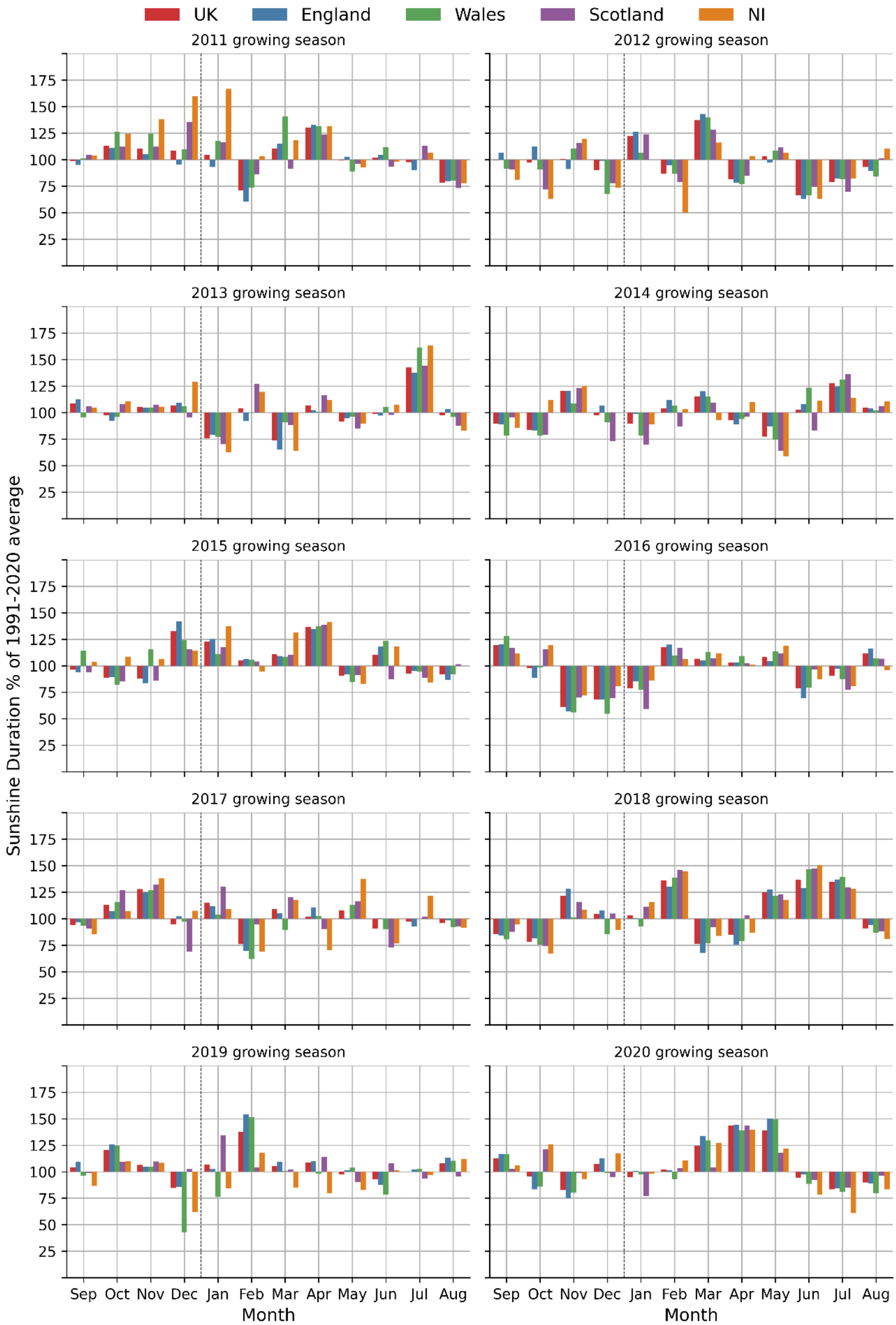
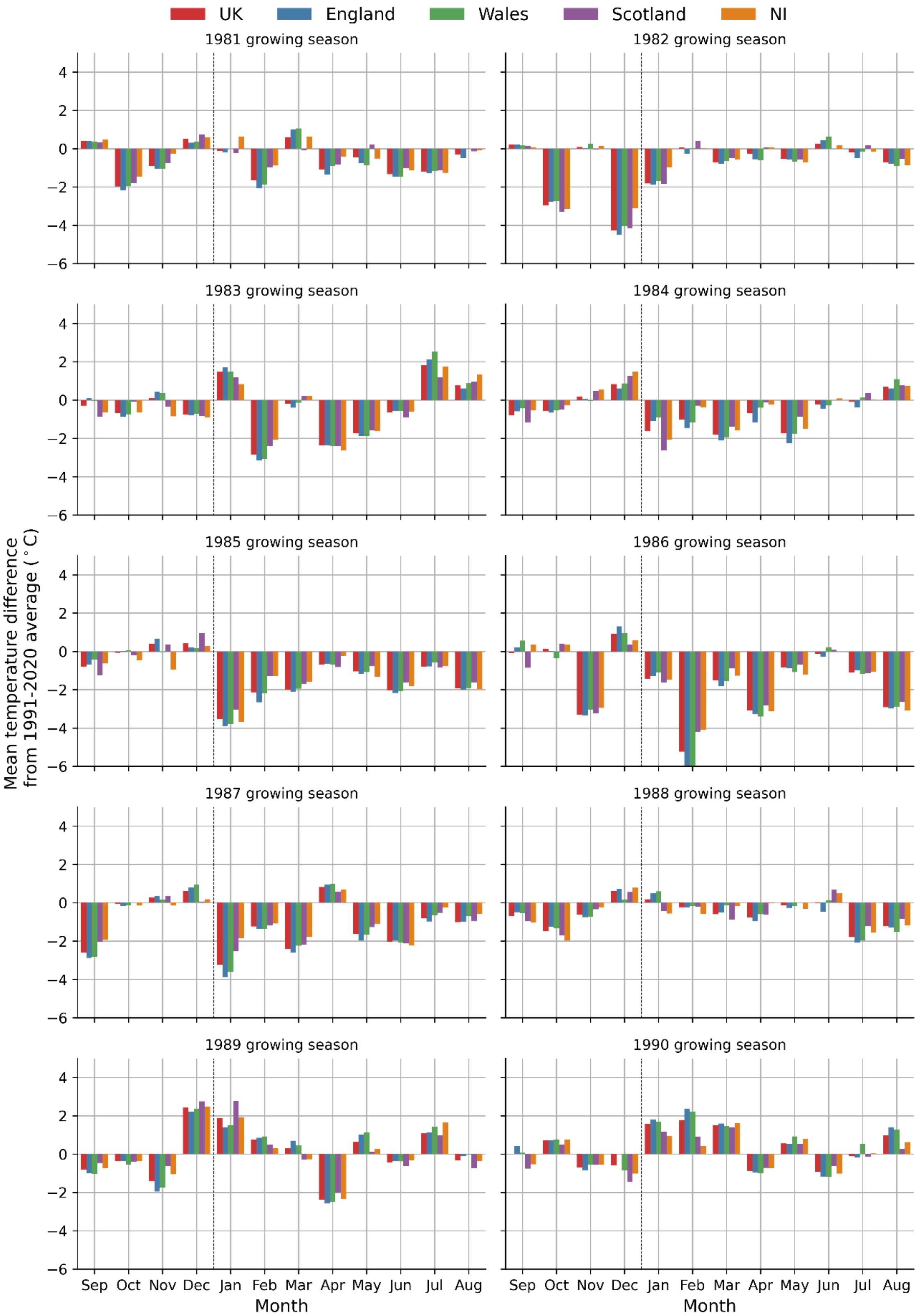
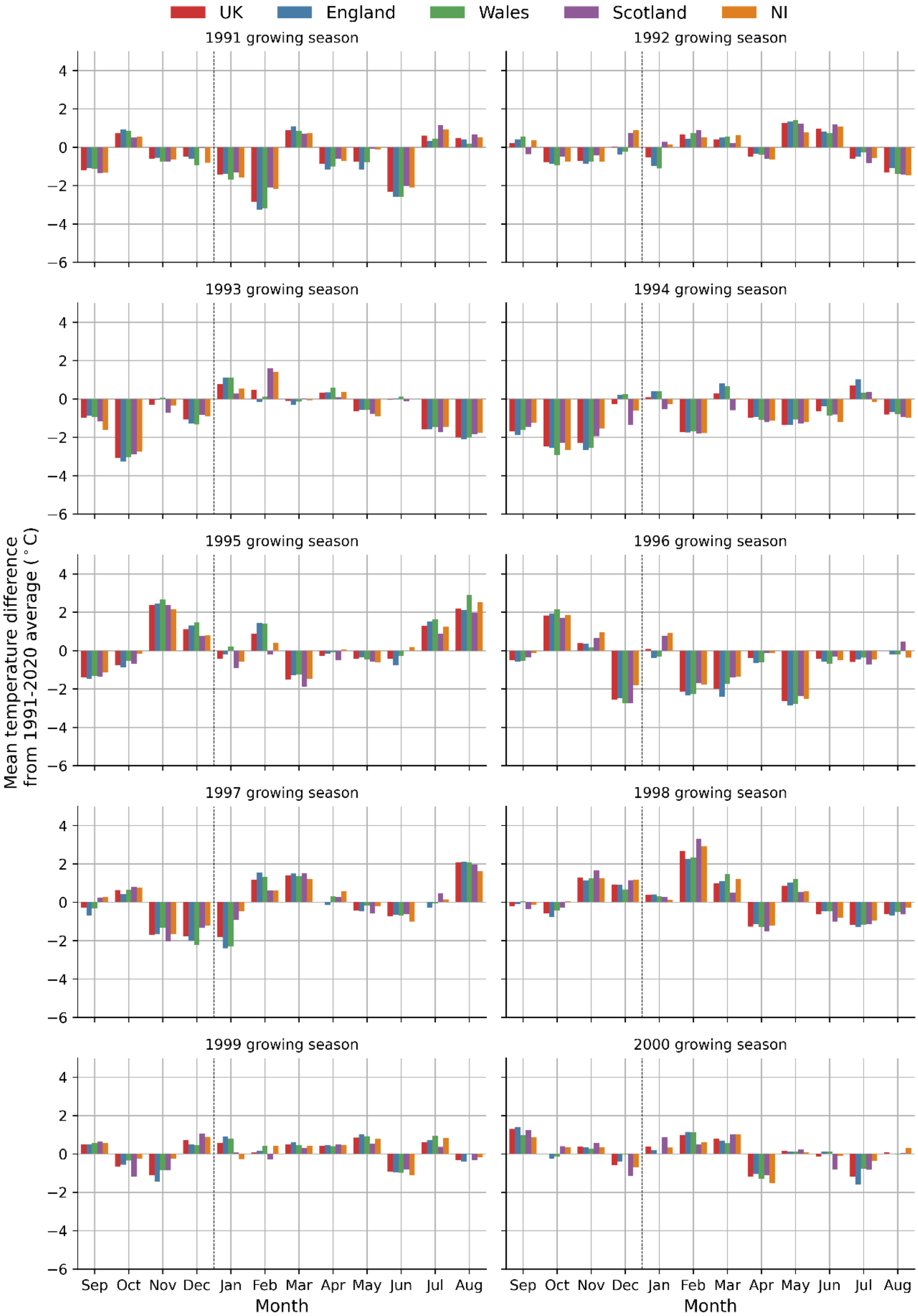
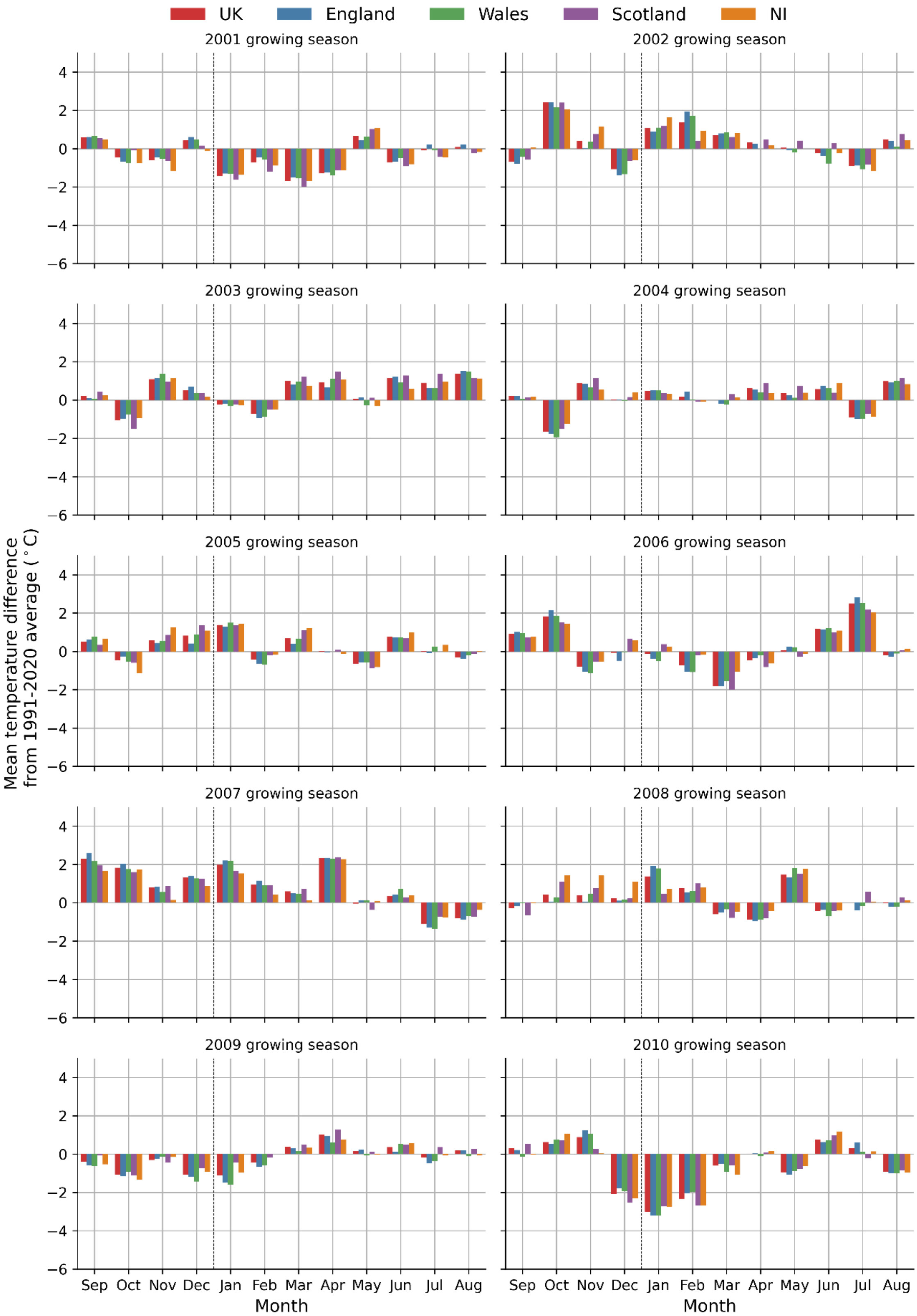



Figure A5: *Growing season sunshine duration anomaly (% 1991-2020 average) plots for the UK (red), England (blue), Wales (green), Scotland (purple) and Northern Ireland (NI) (orange). Data from Met Office year ordered time-series.*



UK
England
Wales
Scotland
NI
1981 growing season
1982 growing season
1983 growing season
1984 growing season
1985 growing season
1986 growing season
1987 growing season
1988 growing season
1989 growing season
1990 growing season
Mean temperature difference from 1991-2020 average (°C)
4
2
0
−2
−4
−6
Sep Oct Nov Dec Jan Feb Mar Apr May Jun Jul Aug
Month


UK
England
Wales
Scotland
NI
1991 growing season
1992 growing season
1993 growing season
1994 growing season
1995 growing season
1996 growing season
1997 growing season
1998 growing season
1999 growing season
2000 growing season
Mean temperature difference from 1991-2020 average (°C)
4
2
0
−2
−4
−6
Sep Oct Nov Dec Jan Feb Mar Apr May Jun Jul Aug
Month



UK
England
Wales
Scotland
NI
2001 growing season
2002 growing season
2003 growing season
2004 growing season
2005 growing season
2006 growing season
2007 growing season
2008 growing season
2009 growing season
2010 growing season
Mean temperature difference from 1991-2020 average (°C)
4
2
0
−2
−4
−6
Sep Oct Nov Dec Jan Feb Mar Apr May Jun Jul Aug
Month

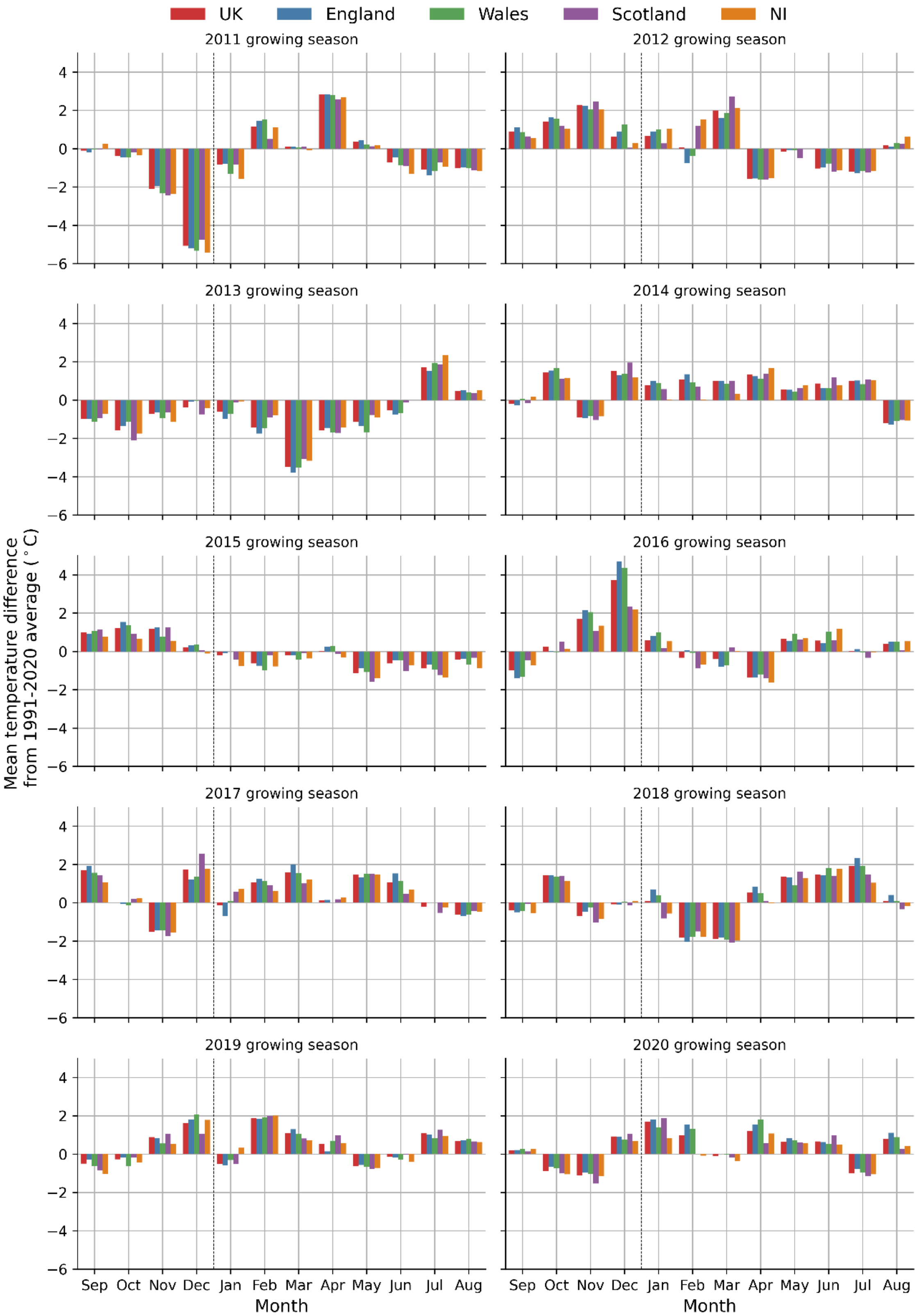



Figure A6: *Growing season mean temperature anomaly (°C difference to 1991-2020 average) plots for the UK (red), England (blue), Wales (green), Scotland (purple) and Northern Ireland (NI) (orange). Data from Met Office year ordered time-series* (Met Office, 2022c).

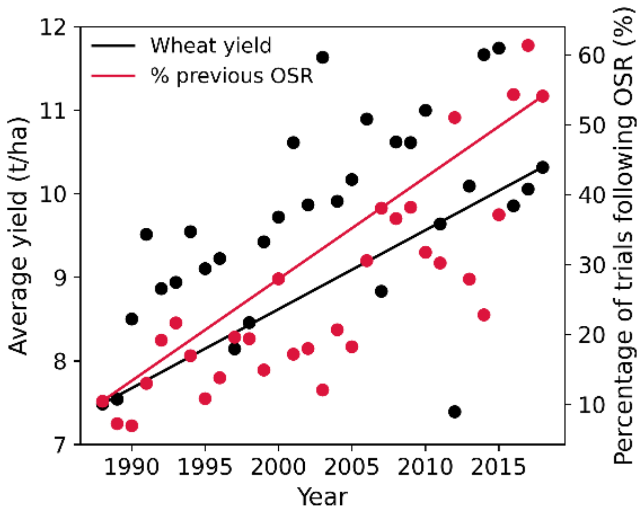



Figure A7: *Change in the percentage of winter wheat trials following oilseed rape (OSR) (red) and the average winter wheat trials yield. Calculated using national yield data from DEFRA.*

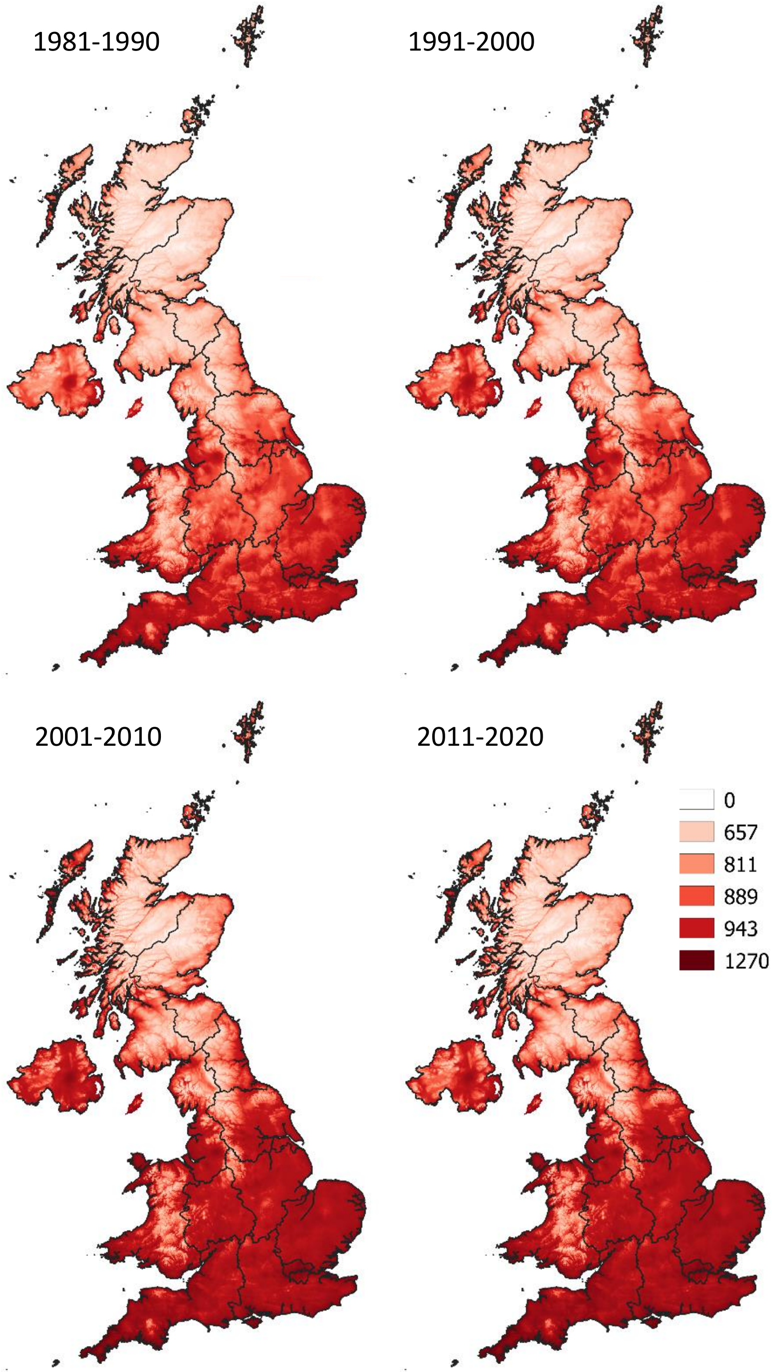
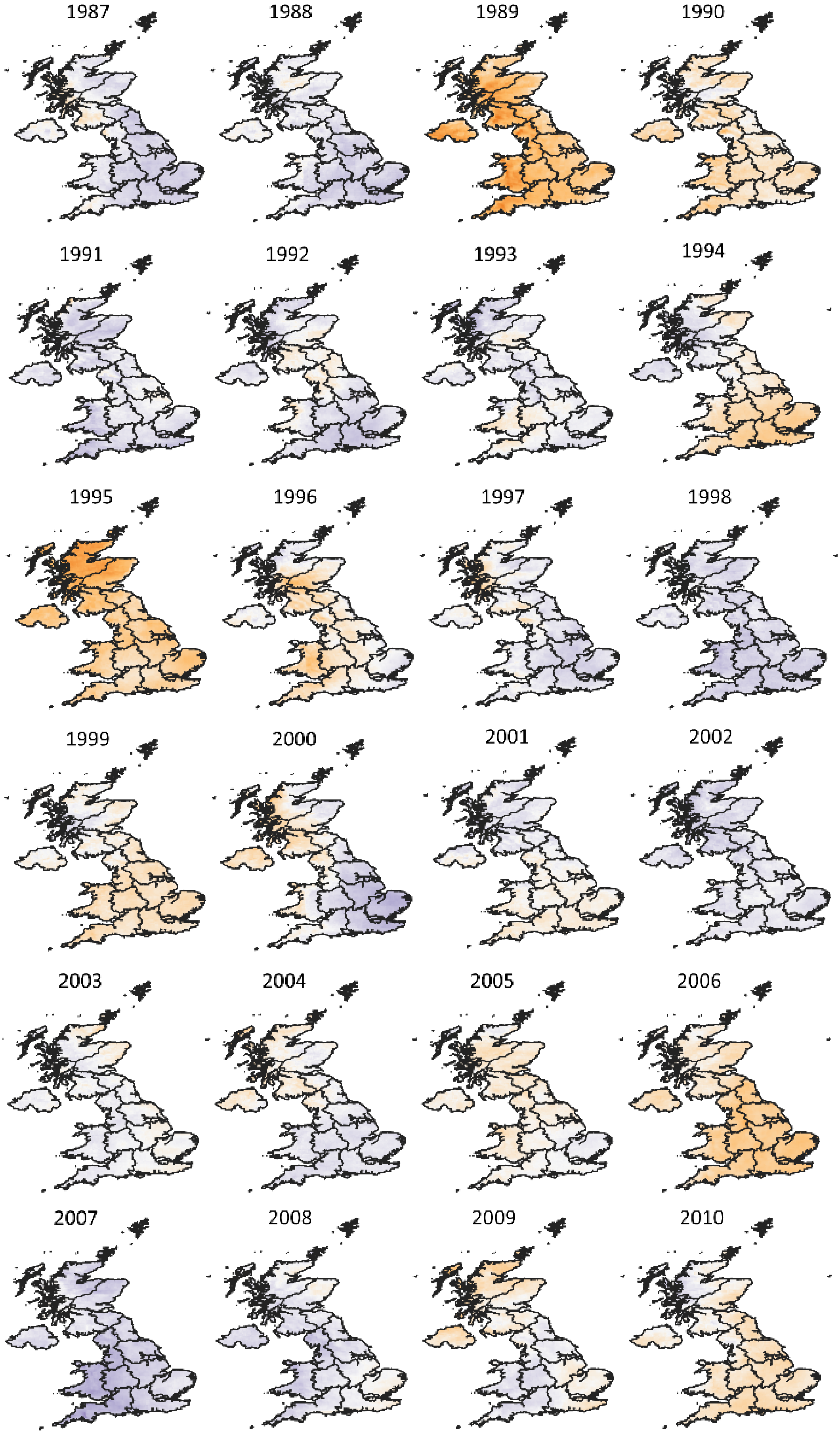



Figure A8: *Vernalisation Degree Days (VDD) (°C days) from 1st November to 28th February for 1982-1990, 1991-2000, 2001-2010 and 2011-2020. Calculated using equations [2.5] and [2.6] using 1 km x 1km gridded HadUK temperature data* (Hollis *et al.*, 2019).


1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010

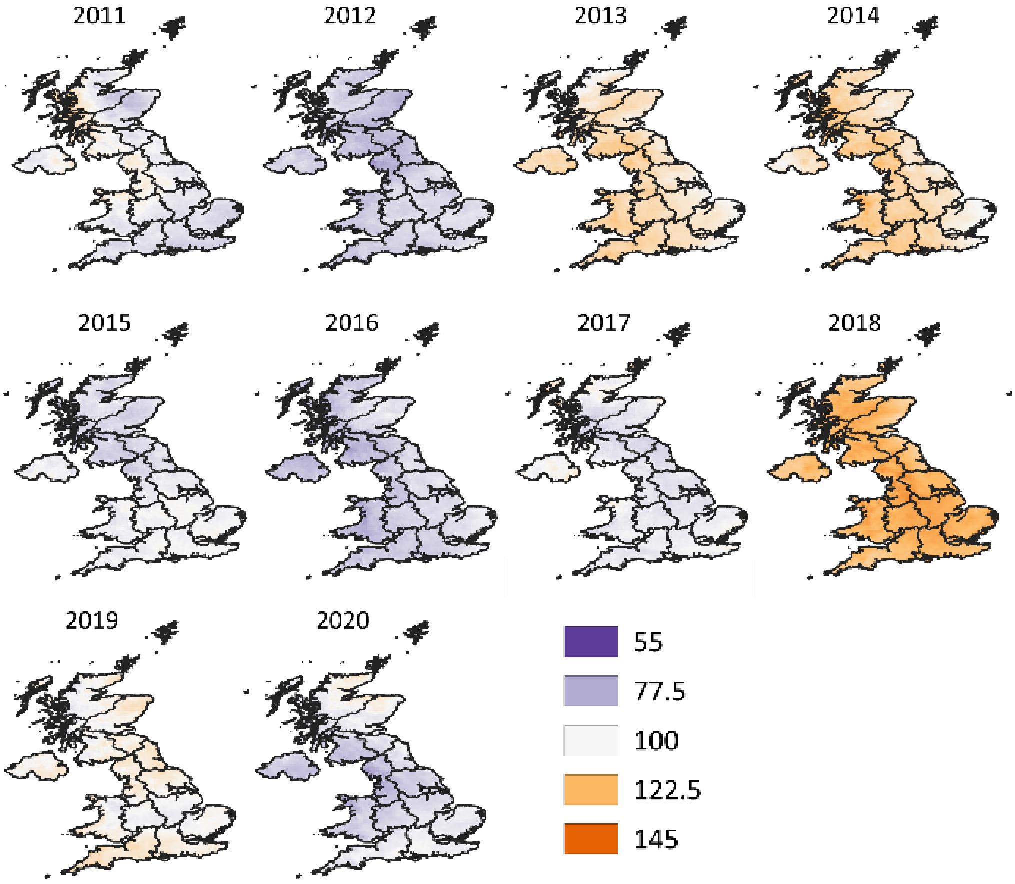



Figure A9: *Annual surface incoming solar radiation (SIS) anomalies (%) relative to 1991-2020 average. Calculated using gridded 0.05º x 0.05º degrees CMSAF SIS data* (Pfeifroth, Trentmann, *et al.*, 2018).

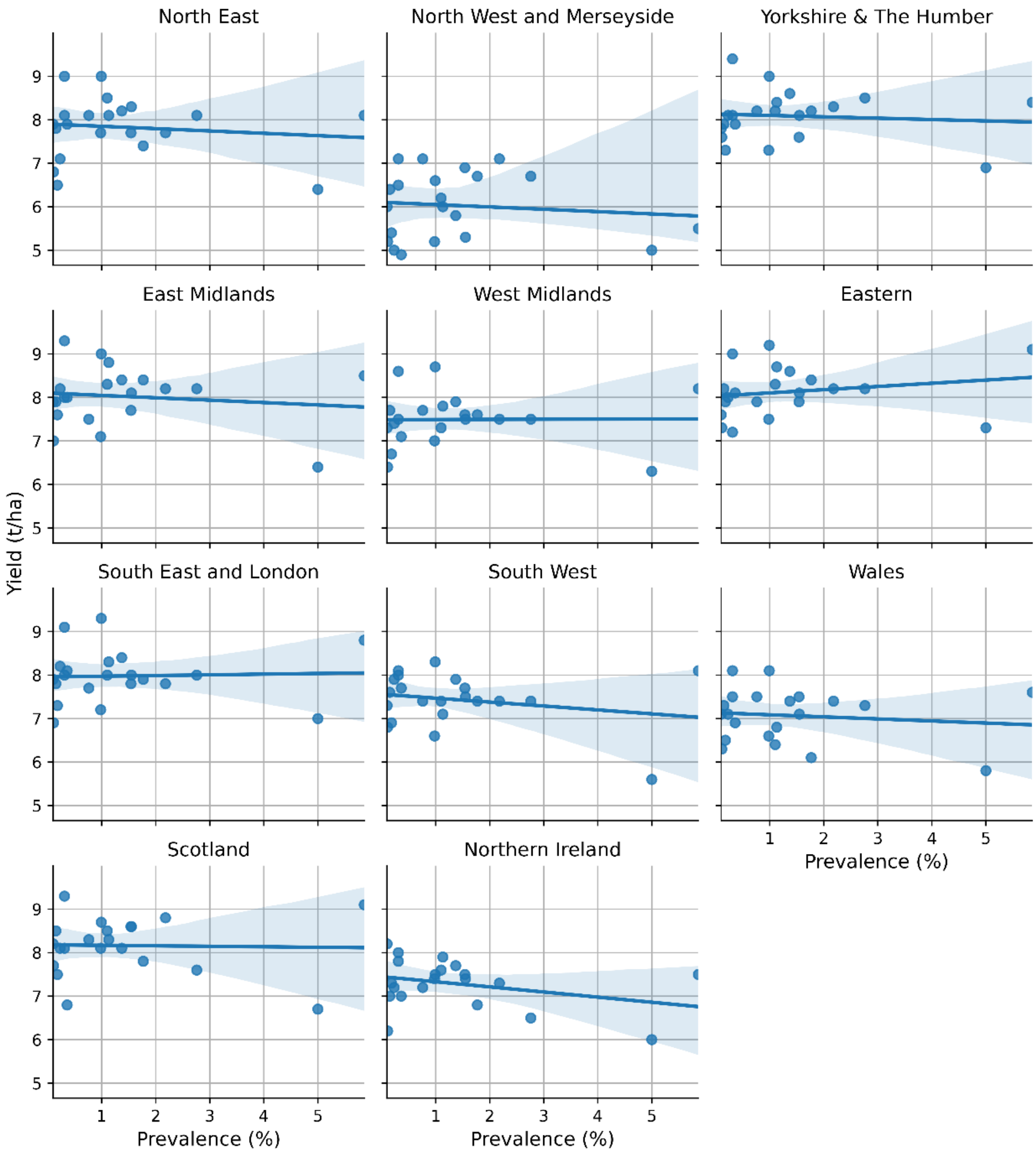



Figure A10: *National Septoria blotch flag leave prevalence (area affected %)* (Polley and Thomas, 1991; Hardwick *et al.*, 2001; Turner *et al.*, 2021) *and regional wheat yields (t/ha)* (DEFRA, 2021a) *for 1999-2019.*

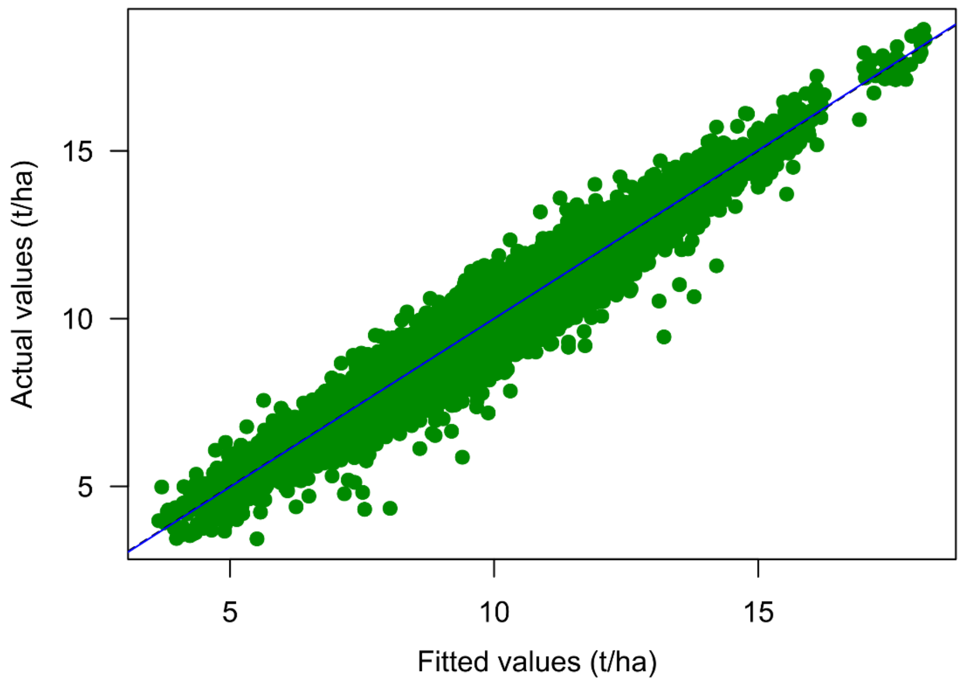



Figure A11*: Observed versus fitted winter wheat yield values (t/ha). Fitted values are from the seasonal climate model on winter wheat (sum of squares shown in Table 6.1), which used [2.12] on the UK National List/Recommended List treated variety trials data for 1988-2018 and regional climate data from the Met Office* (Met Office, 2022c)*. The subsequent models in Chapter 6 had very similar distributions of observed vs. fitted values, hence these plots haven't been included.*

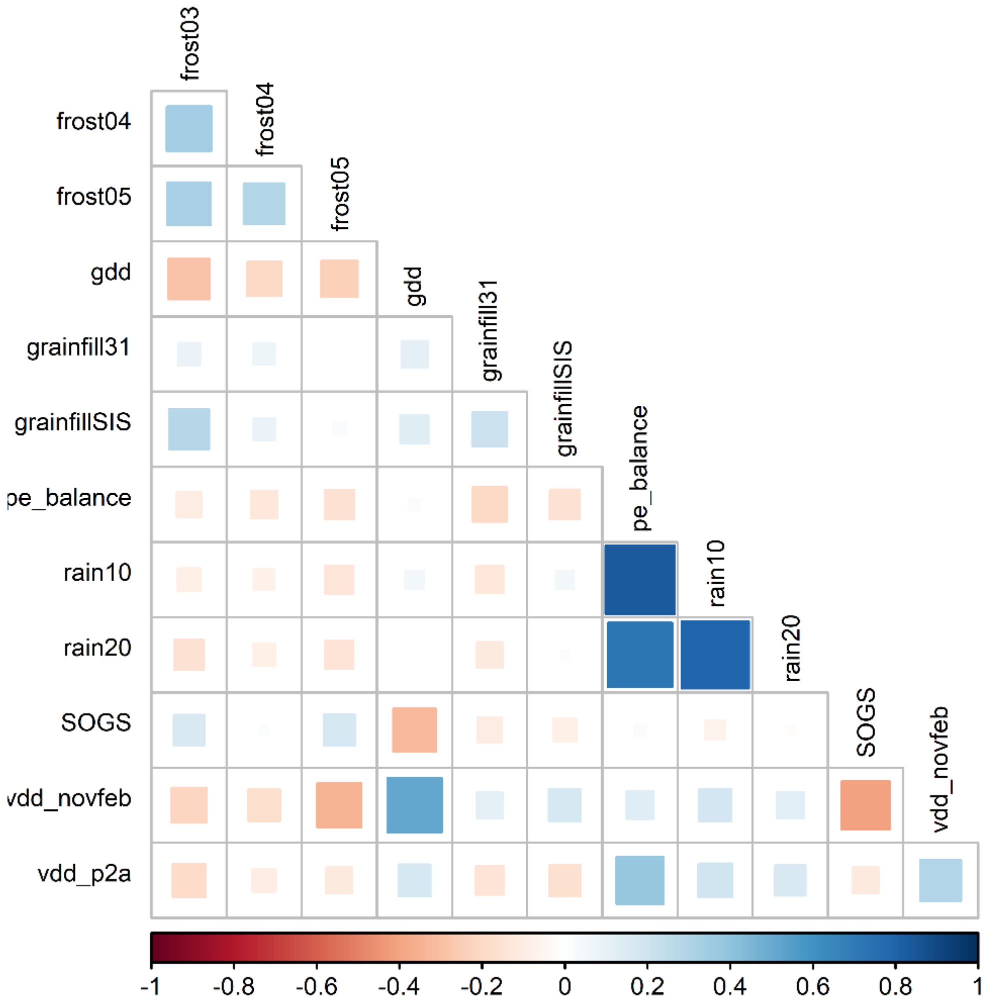



Figure A12: *Pearson's correlation coefficient between each pair of climate variables. Only significant correlations (p<0.05) are shown. The stronger and darker the correlation the larger the square, with positive correlations shown in blue and negative correlations shown in red.*

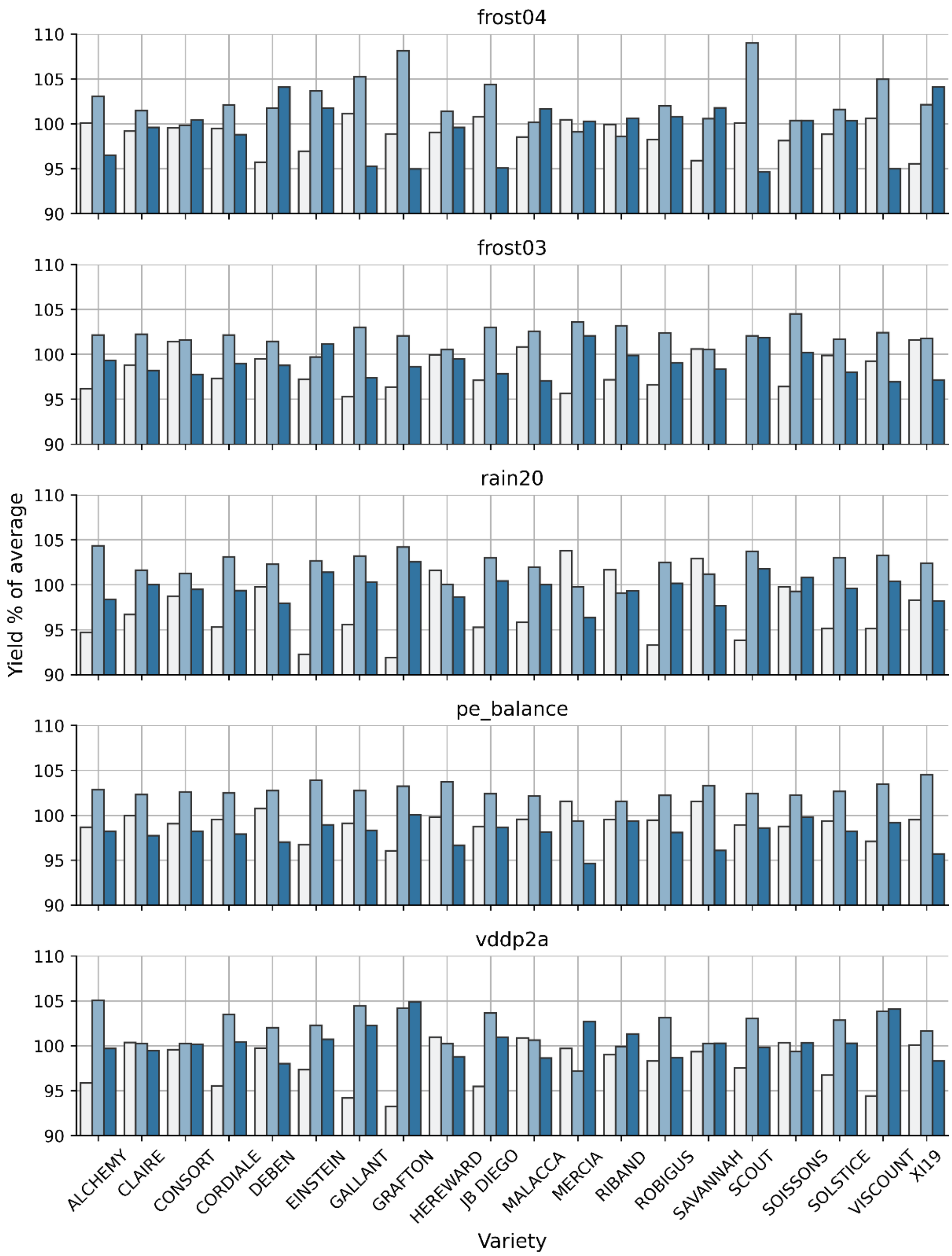



Figure A13: *Yield responses (% of average) of each variety to low (white), medium (light blue) and high (dark blue) (as defined in Table 6.9)numbers of frost days in April (frost04) and March (frost03), 20mm+ rainfall days (rain20), precipitation-evapotranspiration balance (pe_balance) and Vernalisation Degree Days from planting to anthesis (vdd_p2a).*

Supplementary tables

| Variable | Coefficients | | | | | p-value | | | | sig. ($p < 0.05$) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | c.m | cv.b | s.bo | s.ba | s.f | c.m | s.bo | s.ba | s.f | c.m | s.bo | s.ba | s.f |
| (Intercept) | 2.798 | 2.798 | 2.798 | 2.798 | 2.798 | | | | | | | | |
| apr_rain_dmax | 0.219 | 0.068 | 0.214 | 0.284 | 0.168 | 0.0003 | 0.0007 | 0.0003 | 0.0008 | * | * | * | * |
| apr_rain_tot | -0.117 | | -0.113 | -0.157 | -0.077 | 0.0433 | 0.0377 | 0.0376 | 0.0389 | * | * | * | * |
| apr_temp_max | -0.042 | | | | | 0.0146 | | | | * | | | |
| apr_temp_min | 0.182 | 0.283 | 0.240 | 0.181 | 0.168 | 0.1802 | 0.6958 | 0.6657 | 0.9031 | | | | |
| aug_rain_tot | 0.015 | | | | | 0.2565 | | | | | | | |
| aug_temp_max | 0.249 | 0.367 | | 0.153 | 0.233 | 0.7813 | | 0.4569 | 0.2586 | | | | |
| aug_temp_min | 0.085 | 0.144 | | | 0.101 | 0.1100 | | | 0.1985 | | | | |
| jul_rain_tot | -0.069 | | -0.116 | -0.083 | -0.071 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | * | * | * | * |
| jul_temp_max | 0.247 | 0.303 | 0.267 | 0.242 | 0.244 | 0.2188 | 0.0000 | 0.0001 | 0.0000 | | * | * | * |
| jul_temp_min | -0.030 | | | | -0.074 | 0.0030 | | | 0.1132 | * | | | |
| jun_rain_tot | -0.115 | | -0.208 | -0.210 | -0.097 | 0.0147 | 0.0119 | 0.0119 | 0.0125 | * | * | * | * |
| jun_temp_max | 0.035 | | | | | 0.4634 | | | | | | | |
| jun_temp_min | 0.092 | | | 0.124 | 0.094 | 0.0526 | | 0.3337 | 0.2968 | | | | |
| mar_rain_tot | 0.031 | | | 0.079 | | 0.1136 | | 0.4793 | | | | | |
| mar_temp_max | 0.060 | 0.162 | | | 0.064 | 0.3643 | | | 0.2599 | | | | |
| mar_temp_min | 0.090 | 0.216 | 0.180 | 0.111 | 0.115 | 0.2404 | 0.4417 | 0.8278 | 0.8076 | | | | |
| may_rain_tot | -0.231 | -0.306 | -0.085 | -0.207 | -0.250 | 0.8500 | 0.9611 | 0.6170 | 0.9613 | | | | |
| may_temp_max | -0.144 | -0.142 | | -0.107 | -0.128 | 0.2740 | | 0.0042 | 0.0063 | | | * | * |
| may_temp_min | -0.265 | -0.334 | -0.354 | -0.334 | -0.268 | 0.0732 | 0.0004 | 0.0045 | 0.0378 | | * | * | * |

Table A1*: Estimated coefficients and their p-values and significance for 4 different variable selection methods, as well as the full climate model* (c.m.*) for the extreme monthly Irish climate dataset.* cv.b *= best subset selection with cross-validation (Method 1),* s.bo *= cross stepwise selection in both backwards and forwards directions (Method 2a),* s.ba *= 10-fold cross-validation backwards stepwise selection (Method 2b),* s.f *= 10-fold cross-validation forwards stepwise selection (Method 2c) model using lambda for smallest model and within 1 standard error. p-values and significance of each variable where available.*

| Parameter | Coefficient estimate | | | p-value and significance[1] | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | *lmer* post selection | Backwards elimination | Bayesian post selection | *lmer* post selection | | Backwards elimination | | Bayesian post selection | |
| Intercept | 2.90 | 2.90 | 2.88 | 0.000 | *** | 0.000 | *** | 0.000 | * |
| apr_rain_tot | -0.11 | | -0.12 | 0.198 | | | | 0.117 | |
| jul_rain_tot | -0.12 | -0.18 | -0.12 | 0.098 | . | 0.004 | ** | 0.098 | |
| jun_rain_tot | -0.21 | -0.21 | -0.20 | 0.005 | ** | 0.003 | ** | 0.007 | * |
| may_rain_tot | -0.08 | | -0.08 | 0.244 | | | | 0.241 | |
| apr_rain_dmax | 0.21 | 0.17 | 0.23 | 0.014 | * | 0.011 | * | 0.022 | * |
| jul_temp_max | 0.27 | 0.19 | 0.27 | 0.000 | *** | 0.001 | ** | 0.001 | * |
| apr_temp_min | 0.24 | | 0.11 | 0.033 | * | | | 0.284 | |
| mar_temp_min | 0.18 | | 0.15 | 0.027 | * | | | 0.080 | |
| may_temp_min | -0.35 | | -0.18 | 0.005 | ** | | | 0.236 | |
| varietygoldthorpe | -0.21 | -0.21 | -0.21 | 0.000 | *** | 0.000 | *** | 0.000 | * |

Table A2*: Coefficient estimates, p-values and their significance for three linear mixed model methods using the Irish monthly extremes climate data. lmer post selection uses the results from variable selection of climate variable, backwards elimination is the result of running step from lmerTest on the full linear mixed model, and Bayesian post selection represents the Bayesian mixed model run after climate covariate selection using lm.* [1]Signif. codes: 0 *** 0.001 ** 0.01 * 0.05 . 0.1.

| Year | Variety | Seed | Yield |
|---|---|---|---|
| 1 | a | NL | 2 |
| 2 | a | NL | 3 |
| 3 | a | RL | 3 |
| 2 | b | NL | 4 |
| 3 | b | NL | 5 |
| 4 | b | RL | 5 |
| 3 | c | NL | 6 |
| 4 | c | NL | 7 |
| 5 | c | RL | 7 |
| 4 | d | NL | 8 |
| 5 | d | NL | 9 |
| 6 | d | RL | 9 |
| 5 | e | NL | 10 |
| 6 | e | NL | 11 |
| 7 | e | RL | 11 |
| 1 | g | RL | 1 |
| 2 | g | RL | 2 |
| 3 | g | RL | 3 |
| 4 | g | RL | 4 |
| 5 | g | RL | 5 |
| 6 | g | RL | 6 |
| 7 | g | RL | 7 |

Table A3*: Model dataset for estimating genetic gain in the UK NL/RL trials. Seed means seed source, so whether its NL = national list or RL = recommended list. Here g is the check (control) variety.*

| Groups | Variance |
|---|---|
| Variety:Year | 0.037 |
| Year: Site | 2.22 |
| Residual | 0.210 |

Table A4: *The variance components of the random effects in the UK regional, seasonal climate model (Section 6.1).*

| Groups | Variance |
|---|---|
| Variety:Year | 0.033 |
| Year:Site | 2.22 |
| Residual | 0.211 |

Table A5: *The variance components of the random effects in the UK localised seasonal climate model (Section 6.2).*

| Groups | Variance |
|---|---|
| Variety:Year | 0.18 |
| Year:Site | 1.50 |
| Variety:Soil_class | 0.043 |
| Residual | 0.44 |

Table A6: *The variance components of the random effects in the UK localised, seasonal climate and soil model (Section 6.2).*

| Groups | Variance |
|---|---|
| Variety:Year | 0.17 |
| Year:Site | 1.49 |
| Residual | 0.46 |

Table A7: *The variance components of the random effects in the UK localised monthly climate model (Section 6.3).*

| Groups | Variance |
|---|---|
| Variety:Year | 2.37 |
| Year:Site | 0.040 |
| Residual | 0.26 |

Table A8: *The variance components of the random effects in the UK agroclimate model (Section 6.4).*

| | Estimate | Std. Error | Df | t value | p (sig.) |
|---|---|---|---|---|---|
| (Intercept) | 6.9280 | 1.0615 | 998.9 | 6.526 | 0.000 (*) |
| grainfillSIS | 0.0030 | 0.0015 | 750.3 | 2.018 | 0.044 (*) |
| rain10 | -0.0134 | 0.0085 | 1342.6 | -1.579 | 0.115 |
| gdd | 0.0000 | 0.0004 | 4958.7 | -0.122 | 0.903 |
| vdd_novfeb | 0.0005 | 0.0011 | 1211.2 | 0.447 | 0.655 |
| Year1989 | -0.0659 | 0.7373 | 867.2 | -0.089 | 0.929 |
| Year1990 | 0.2179 | 0.6386 | 891.0 | 0.341 | 0.733 |
| Year1991 | 0.5017 | 0.6167 | 914.2 | 0.814 | 0.416 |
| Year1992 | 0.3492 | 0.6435 | 914.4 | 0.543 | 0.588 |
| Year1994 | 0.4987 | 0.8036 | 869.8 | 0.621 | 0.535 |
| Year1995 | 0.1664 | 0.6400 | 889.9 | 0.260 | 0.795 |
| Year1996 | 0.9874 | 0.6234 | 909.2 | 1.584 | 0.114 |
| Year1997 | 0.5827 | 0.6427 | 987.9 | 0.907 | 0.365 |
| Year1998 | 1.0585 | 0.6162 | 904.1 | 1.718 | 0.086 |
| Year1999 | 1.8059 | 0.6523 | 896.4 | 2.768 | 0.006 (*) |
| Year2000 | 1.8509 | 0.6219 | 889.9 | 2.976 | 0.003 (*) |
| Year2001 | 0.2195 | 0.6408 | 876.9 | 0.343 | 0.732 |
| Year2002 | 1.2752 | 0.6098 | 919.8 | 2.091 | 0.037 (*) |
| Year2003 | 1.0236 | 0.6000 | 908.0 | 1.706 | 0.088 |
| Year2004 | 1.3383 | 0.6094 | 932.7 | 2.196 | 0.028 (*) |
| Year2005 | 0.5656 | 0.6113 | 894.8 | 0.925 | 0.355 |
| Year2006 | 0.7783 | 0.7019 | 866.3 | 1.109 | 0.268 |
| Year2007 | 0.2420 | 0.6451 | 888.1 | 0.375 | 0.708 |
| Year2008 | 1.3373 | 0.6261 | 902.3 | 2.136 | 0.033 (*) |
| Year2009 | 0.8528 | 0.6679 | 897.6 | 1.277 | 0.202 |
| Year2010 | 0.3294 | 0.6852 | 894.2 | 0.481 | 0.631 |
| Year2011 | 0.7845 | 0.7676 | 952.8 | 1.022 | 0.307 |
| Year2012 | -0.5507 | 0.6288 | 910.5 | -0.876 | 0.381 |
| Year2013 | -0.0813 | 0.6747 | 898.7 | -0.120 | 0.904 |
| Year2014 | 1.5490 | 0.6206 | 908.3 | 2.496 | 0.013 (*) |
| Year2015 | 1.6953 | 0.6040 | 916.3 | 2.807 | 0.005 (*) |
| Year2016 | 0.9117 | 0.6138 | 938.3 | 1.485 | 0.138 |
| Year2017 | 0.8714 | 0.6136 | 945.8 | 1.420 | 0.156 |
| VarietyCLAIRE | -0.5134 | 0.4262 | 2309.0 | -1.205 | 0.228 |
| VarietyCONSORT | 0.8722 | 0.4693 | 2937.8 | 1.859 | 0.063 |
| VarietyCORDIALE | 0.0487 | 0.4425 | 2224.7 | 0.110 | 0.912 |
| VarietyDEBEN | 0.9702 | 0.5568 | 3032.4 | 1.742 | 0.082 |
| VarietyEINSTEIN | 0.7159 | 0.4580 | 2078.9 | 1.563 | 0.118 |
| VarietyGALLANT | -0.9041 | 0.4924 | 2518.8 | -1.836 | 0.066 |
| VarietyGRAFTON | 0.1822 | 0.5376 | 2503.9 | 0.339 | 0.735 |
| VarietyHEREWARD | -0.8191 | 0.5093 | 2517.2 | -1.608 | 0.108 |
| VarietyJB DIEGO | 0.0231 | 0.4629 | 2254.0 | 0.050 | 0.960 |
| VarietyMALACCA | 0.7411 | 0.4945 | 2743.1 | 1.499 | 0.134 |
| VarietyMERCIA | 0.3012 | 0.6138 | 3266.0 | 0.491 | 0.624 |
| VarietyRIBAND | 1.1334 | 0.5238 | 3127.6 | 2.164 | 0.031 (*) |

| | | | | | |
|---|---|---|---|---|---|
| VarietyROBIGUS | 0.3835 | 0.5072 | 2573.0 | 0.756 | 0.450 |
| VarietySAVANNAH | 2.9004 | 0.5155 | 3029.4 | 5.627 | 0.000 (*) |
| VarietySCOUT | 0.4761 | 0.5345 | 2685.6 | 0.891 | 0.373 |
| VarietySOISSONS | 1.8207 | 0.5729 | 2665.4 | 3.178 | 0.001 (*) |
| VarietySOLSTICE | 0.3802 | 0.4303 | 2311.4 | 0.884 | 0.377 |
| VarietyVISCOUNT | 1.5487 | 0.5038 | 2836.7 | 3.074 | 0.002 (*) |
| VarietyXI19 | -0.2191 | 0.5890 | 2782.9 | -0.372 | 0.710 |
| rain10:VarietyCLAIRE | 0.0025 | 0.0052 | 6178.9 | 0.477 | 0.633 |
| rain10:VarietyCONSORT | -0.0134 | 0.0057 | 6264.9 | -2.336 | 0.020 (*) |
| rain10:VarietyCORDIALE | -0.0089 | 0.0054 | 6170.9 | -1.627 | 0.104 |
| rain10:VarietyDEBEN | -0.0058 | 0.0061 | 6184.2 | -0.952 | 0.341 |
| rain10:VarietyEINSTEIN | 0.0013 | 0.0058 | 6165.8 | 0.230 | 0.818 |
| rain10:VarietyGALLANT | -0.0083 | 0.0059 | 6185.3 | -1.413 | 0.158 |
| rain10:VarietyGRAFTON | 0.0071 | 0.0073 | 6176.1 | 0.970 | 0.332 |
| rain10:VarietyHEREWARD | -0.0126 | 0.0057 | 6281.3 | -2.211 | 0.027 (*) |
| rain10:VarietyJB DIEGO | -0.0041 | 0.0054 | 6179.9 | -0.758 | 0.448 |
| rain10:VarietyMALACCA | -0.0050 | 0.0056 | 6204.1 | -0.894 | 0.372 |
| rain10:VarietyMERCIA | -0.0052 | 0.0073 | 6394.5 | -0.715 | 0.475 |
| rain10:VarietyRIBAND | -0.0092 | 0.0062 | 6353.3 | -1.478 | 0.139 |
| rain10:VarietyROBIGUS | -0.0065 | 0.0059 | 6168.6 | -1.104 | 0.270 |
| rain10:VarietySAVANNAH | -0.0258 | 0.0062 | 6242.3 | -4.158 | 0.000 (*) |
| rain10:VarietySCOUT | -0.0089 | 0.0067 | 6170.8 | -1.322 | 0.186 |
| rain10:VarietySOISSONS | -0.0137 | 0.0068 | 6295.0 | -1.997 | 0.046 (*) |
| rain10:VarietySOLSTICE | -0.0042 | 0.0052 | 6161.9 | -0.797 | 0.425 |
| rain10:VarietyVISCOUNT | -0.0019 | 0.0071 | 6176.8 | -0.267 | 0.789 |
| rain10:VarietyXI19 | -0.0173 | 0.0063 | 6172.8 | -2.755 | 0.006 (*) |
| gdd:VarietyCLAIRE | -0.0007 | 0.0003 | 3512.7 | -2.236 | 0.025 (*) |
| gdd:VarietyCONSORT | -0.0015 | 0.0003 | 3620.3 | -4.637 | 0.000 (*) |
| gdd:VarietyCORDIALE | -0.0014 | 0.0003 | 3452.9 | -4.144 | 0.000 (*) |
| gdd:VarietyDEBEN | -0.0011 | 0.0004 | 4105.2 | -2.790 | 0.005 (*) |
| gdd:VarietyEINSTEIN | -0.0018 | 0.0003 | 3553.3 | -5.253 | 0.000 (*) |
| gdd:VarietyGALLANT | -0.0012 | 0.0004 | 3870.1 | -3.290 | 0.001 (*) |
| gdd:VarietyGRAFTON | -0.0014 | 0.0004 | 4460.0 | -3.404 | 0.001 (*) |
| gdd:VarietyHEREWARD | -0.0011 | 0.0003 | 3798.2 | -3.224 | 0.001 (*) |
| gdd:VarietyJB DIEGO | -0.0008 | 0.0003 | 4253.9 | -2.391 | 0.017 (*) |
| gdd:VarietyMALACCA | -0.0012 | 0.0003 | 3952.3 | -3.460 | 0.001 (*) |
| gdd:VarietyMERCIA | -0.0021 | 0.0004 | 3943.8 | -4.984 | 0.000 (*) |
| gdd:VarietyRIBAND | -0.0019 | 0.0004 | 3794.0 | -5.230 | 0.000 (*) |
| gdd:VarietyROBIGUS | -0.0005 | 0.0004 | 4246.8 | -1.290 | 0.197 |
| gdd:VarietySAVANNAH | -0.0020 | 0.0004 | 3680.4 | -5.456 | 0.000 (*) |
| gdd:VarietySCOUT | -0.0012 | 0.0004 | 4227.8 | -3.022 | 0.003 (*) |
| gdd:VarietySOISSONS | -0.0022 | 0.0004 | 3592.0 | -5.798 | 0.000 (*) |
| gdd:VarietySOLSTICE | -0.0013 | 0.0003 | 3634.1 | -4.123 | 0.000 (*) |
| gdd:VarietyVISCOUNT | -0.0013 | 0.0004 | 3610.7 | -3.424 | 0.001 (*) |
| gdd:VarietyXI19 | -0.0008 | 0.0004 | 4260.7 | -2.092 | 0.037 (*) |
| vdd_novfeb:VarietyCLAIRE | 0.0012 | 0.0005 | 732.2 | 2.243 | 0.025 (*) |

| | | | | | |
|---|---|---|---|---|---|
| vdd_novfeb:VarietyCONSORT | 0.0012 | 0.0006 | 1217.0 | 1.905 | 0.057 |
| vdd_novfeb:VarietyCORDIALE | 0.0019 | 0.0005 | 592.7 | 3.732 | 0.000 (*) |
| vdd_novfeb:VarietyDEBEN | 0.0007 | 0.0007 | 1571.3 | 0.935 | 0.350 |
| vdd_novfeb:VarietyEINSTEIN | 0.0017 | 0.0005 | 577.1 | 3.123 | 0.002 (*) |
| vdd_novfeb:VarietyGALLANT | 0.0027 | 0.0006 | 604.4 | 4.867 | 0.000 (*) |
| vdd_novfeb:VarietyGRAFTON | 0.0019 | 0.0006 | 565.9 | 3.263 | 0.001 (*) |
| vdd_novfeb:VarietyHEREWARD | 0.0014 | 0.0006 | 1087.3 | 2.126 | 0.034 (*) |
| vdd_novfeb:VarietyJB DIEGO | 0.0016 | 0.0005 | 638.4 | 3.072 | 0.002 (*) |
| vdd_novfeb:VarietyMALACCA | 0.0002 | 0.0007 | 1456.6 | 0.343 | 0.732 |
| vdd_novfeb:VarietyMERCIA | 0.0011 | 0.0008 | 1234.5 | 1.432 | 0.152 |
| vdd_novfeb:VarietyRIBAND | 0.0012 | 0.0007 | 1277.0 | 1.840 | 0.066 |
| vdd_novfeb:VarietyROBIGUS | 0.0005 | 0.0007 | 1226.8 | 0.698 | 0.485 |
| vdd_novfeb:VarietySAVANNAH | 0.0003 | 0.0007 | 1336.4 | 0.477 | 0.633 |
| vdd_novfeb:VarietySCOUT | 0.0013 | 0.0006 | 550.3 | 2.265 | 0.024 (*) |
| vdd_novfeb:VarietySOISSONS | 0.0001 | 0.0007 | 1149.2 | 0.142 | 0.887 |
| vdd_novfeb:VarietySOLSTICE | 0.0013 | 0.0005 | 641.5 | 2.591 | 0.010 (*) |
| vdd_novfeb:VarietyVISCOUNT | 0.0007 | 0.0006 | 579.6 | 1.227 | 0.220 |
| vdd_novfeb:VarietyXI19 | 0.0017 | 0.0008 | 1539.5 | 2.229 | 0.026 (*) |

Table A9: *Final UK multivariate agroclimate model fixed effects coefficient estimates and associated standard error, degrees of freedom (Df), t statistics and p value significance *significant at the 95% confidence level.*

# Glossary

AHDB = Agriculture and Horticulture Development Board

AIC = Akaike Information Criterion

BLUE = best linear unbiased estimator

BSPB = British Society of Plant Breeders

DEFRA = Department for Environment, Food and Rural Affairs

FDR = false discovery rate

GDD = Growing Degree Days

GxE = genotype-by-environment interaction

ILMMT = Ireland Long-term Maximum and Minimum Air Temperature

IOI = Island of Ireland

IPCC = Intergovernmental Panel on Climate Change

NAO = North Atlantic Oscillation

NIAB = National Institute of Agricultural Botany

NL = National List

RCP = Representative Concentration Pathways

RL = Recommended List

RMSE = root mean square error

SNAO = Summer North Atlantic Oscillation

SPI = Standardised Precipitation Index

SPEI = Standardised Precipitation and Evapotranspiration Index

VDD = Vernalisation Degree Days

WNAO = Winter North Atlantic Oscillation

# References


Addy, J. W. G., Ellis, R. H., Macdonald, A. J., Semenov, M. A. and Mead, A. (2020) ‘Investigating the effects of inter-annual weather variation (1968–2016) on the functional response of cereal grain yield to applied nitrogen, using data from the Rothamsted Long-Term Experiments’, *Agricultural and Forest Meteorology*, 284(December 2019). doi: 10.1016/j.agrformet.2019.107898.

Addy, J. W. G., Ellis, R. H., Macdonald, A. J., Semenov, M. A. and Mead, A. (2021a) ‘Changes in agricultural climate in South-Eastern England from 1892 to 2016 and differences in cereal and permanent grassland yield’, *Agricultural and Forest Meteorology*, 308–309(July). doi: 10.1016/j.agrformet.2021.108560.

Addy, J. W. G., Ellis, R. H., Macdonald, A. J., Semenov, M. A. and Mead, A. (2021b) ‘The impact of weather and increased atmospheric CO 2 from 1892 to 2016 on simulated yields of UK wheat’, *Journal of the Royal Society Interface*, 18(179). doi: 10.1098/rsif.2021.0250.

AHDB (2019) *Principles of soil management*. doi: 10.1201/b14972.

AHDB (2021) *Septoria disease rating dip revealed by early RL dataset release*. Available at: https://ahdb.org.uk/news/septoria-disease-rating-dip-revealed-by-early-rl-dataset-release.

AHDB (2022) *Cabbage stem flea beetle (CSFB) and its management in oilseed rape*. Available at: https://ahdb.org.uk/knowledge-library/cabbage-stem-flea-beetle-csfb-and-its-management-in-oilseed-rape.

AHDB (2023) *AHDB Recommended Lists (RL) for cereals and oilseeds: Cereal Trials protocols (2022 – 26)*.

AHDB Cereals & Oilseeds (2018a) *Barley growth guide*. Agriculture and Horticulture Development Board.

AHDB Cereals & Oilseeds (2018b) *Wheat and barley disease management guide*. Agriculture and Horticulture Development Board.

AHDB Cereals & Oilseeds (2018c) *Wheat growth guide*. Agriculture and Horticulture Development Board.

Arnell, N. W. and Freeman, A. (2021) ‘The effect of climate change on agro-climatic indicators in the UK’, *Climatic Change*, 165(1–2), pp. 1–26. doi: 10.1007/s10584-021-03054-8.

Arnell, N. W., Kay, A. L., Freeman, A., Rudd, A. C. and Lowe, J. A. (2021) 'Changing climate risk in the UK: A multi-sectoral analysis using policy-relevant indicators', *Climate Risk Management*, 31(December 2020). doi: 10.1016/j.crm.2020.100265.

Atkinson, M. D., Kettlewell, P. S., Hollins, P. D., Stephenson, D. B. and Hardwick, N. V. (2005) 'Summer climate mediates UK wheat quality response to winter North Atlantic Oscillation', *Agricultural and Forest Meteorology*, 130(1–2), pp. 27–37. doi: 10.1016/j.agrformet.2005.02.002.

Atkinson, M. D., Kettlewell, P. S., Poulton, P. R. and Hollins, P. D. (2008) 'Grain quality in the Broadbalk Wheat Experiment and the winter North Atlantic Oscillation', *The Journal of Agricultural Science*, 146(5), pp. 541–549. doi: 10.1017/S0021859608007958.

Austin, R. B. (1999) 'Yield of wheat in the United Kingdom: Recent advances and prospects', *Crop Science*, 39(6), pp. 1604–1610. doi: 10.2135/cropsci1999.3961604x.

Austin, R. B., Ford, M. A. and Morgan, C. L. (1989) 'Genetic improvement in the yield of winter wheat: A further evaluation', *The Journal of Agricultural Science*, 112(3), pp. 295–301. doi: 10.1017/S0021859600085749.

Azeez, M. A., Adubi, A. O. and Durodola, F. A. (2018) 'Landraces and Crop Genetic Improvement', in Grillo, O. (ed.) *Rediscovery of Landraces as a Resource for the Future*. InTech, p. 13. doi: 10.5772/intechopen.75944.

Barlow, K. M., Christy, B. P., O'Leary, G. J., Riffkin, P. A. and Nuttall, J. G. (2015) 'Simulating the impact of extreme heat and frost events on wheat crop production: A review', *Field Crops Research*, 171, pp. 109–119. doi: 10.1016/j.fcr.2014.11.010.

Bates, D., Mächler, M., Bolker, B. M. and Walker, S. C. (2020) *lme4: linear mixed-effects models. R package version 1.1.21*.

Battisti, R., Bender, F. D. and Sentelhas, P. C. (2019) 'Assessment of different gridded weather data for soybean yield simulations in Brazil', *Theoretical and Applied Climatology*, 135(1–2), pp. 237–247. doi: 10.1007/s00704-018-2383-y.

Batts, G. R., Morison, J. I. L., Ellis, R. H., Hadley, P. and Wheeler, T. R. (1997) 'Effects of CO2 and temperature on growth and yield of crops of winter wheat over four seasons', *European Journal of Agronomy*, 7(1–3), pp. 43–52. doi: 10.1016/S1161-0301(97)00022-1.

Bayer (2020) *Wheat Diseases*, *Crop Science*. Available at: https://cropscience.bayer.co.uk/threats/diseases/wheat-diseases/.

Bayles, R. A., Flath, K., Hovmøller, M. S. and de Vallavieille-Pope, C. (2000) 'Breakdown of the Yr17 resistance to yellow rust of wheat in northern Europe', *Agronomie*, 20(7), pp. 805–811. doi: 10.1051/agro:2000176.

BBSRC UKRI (2016) *New Heritage Barley Ltd – reviving a Victorian barley variety for modern brewers*.

Bebber, D. P., Ramotowski, M. A. T. and Gurr, S. J. (2013) 'Crop pests and pathogens move polewards in a warming world', *Nature Climate Change*, 3(11), pp. 985–988. doi: 10.1038/nclimate1990.

Beck, H. E., Zimmermann, N. E., McVicar, T. R., Vergopolan, N., Berg, A. and Wood, E. F. (2018) 'Present and future Köppen-Geiger climate classification maps at 1-km resolution', *Scientific Data*, 5(1), p. 180214. doi: 10.1038/sdata.2018.214.

Behnassi, M. and El Haiba, M. (2022) 'Implications of the Russia–Ukraine war for global food security', *Nature Human Behaviour*, 6(6), pp. 754–755. doi: 10.1038/s41562-022-01391-x.

Bell, G., Naumann, E.-K. and Medcalf, K. (2019) *Application of ALC and UKCP18 Data for Modelling Crop Suitability. Report to the Welsh Government*.

Bentley, A. R., Jensen, E. F., Mackay, I. J., Hönicka, H., Fladung, M., Hori, K., Yano, M., Mullet, J. E., Armstead, I. P., Hayes, C., Thorogood, D., Lovatt, A., Morris, R., Pullen, N., Mutasa-Göttgens, E. and Cockram, J. (2013) 'Flowering Time', in *Genomics and Breeding for Climate-Resilient Crops*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 1–66. doi: 10.1007/978-3-642-37048-9_1.

Benton, T. (2014) *The sanctions fight with Russia could make Britain go hungry*, *Telegraph*. Available at: https://www.telegraph.co.uk/news/worldnews/europe/ukraine/11025555/The-sanctions-fight-with-Russia-could-make-Britain-go-hungry.html.

Berry, P. and Brown, I. (2021) 'National environment and assets', in Betts, R. A., Haward, A. B., and Pearson, K. V. (eds) *The Third UK Climate Change Risk Assessment Technical Report*. London: Climate Change Committee.

Berry, P. M., Kendall, S., Rutterford, Z., Orford, S. and Griffiths, S. (2015) 'Historical analysis of the effects of breeding on the height of winter wheat (Triticum aestivum) and consequences for lodging', *Euphytica*, 203(2), pp. 375–383. doi: 10.1007/s10681-014-1286-y.

van Bezouw, R. F. H. M., Keurentjes, J. J. B., Harbinson, J. and Aarts, M. G. M. (2019)

‘Converging phenomics and genomics to study natural variation in plant photosynthetic efficiency’, *The Plant Journal*, 97(1), pp. 112–133. doi: 10.1111/tpj.14190.

Bhakta, I., Phadikar, S. and Majumder, K. (2019) ‘State-of-the-art technologies in precision agriculture: a systematic review’, *Journal of the Science of Food and Agriculture*, pp. 1–11. doi: 10.1002/jsfa.9693.

Bindereif, S. G., Rüll, F., Kolb, P., Köberle, L., Willms, H., Steidele, S., Schwarzinger, S. and Gebauer, G. (2021) ‘Impact of global climate change on the european barley market requires novel multi-method approaches to preserve crop quality and authenticity’, *Foods*, 10(7). doi: 10.3390/foods10071592.

Blandino, M., Badeck, F.-W., Giordano, D., Marti, A., Rizza, F., Scarpino, V. and Vaccino, P. (2020) ‘Elevated CO 2 Impact on Common Wheat ( Triticum aestivum L.) Yield, Wholemeal Quality, and Sanitary Risk’, *Journal of Agricultural and Food Chemistry*, 68(39), pp. 10574–10585. doi: 10.1021/acs.jafc.0c02975.

Bloggers, R. (2020) *Bayesian linear regression*, *R bloggers*. Available at: https://www.r-bloggers.com/2020/04/bayesian-linear-regression/ (Accessed: 12 April 2021).

Bönecke, E., Breitsameter, L., Brüggemann, N., Chen, T. W., Feike, T., Kage, H., Kersebaum, K. C., Piepho, H. P. and Stützel, H. (2020) ‘Decoupling of impact factors reveals the response of German winter wheat yields to climatic changes’, *Global Change Biology*, 26(6), pp. 3601–3626. doi: 10.1111/gcb.15073.

Bracken, A. F. and Bailey, C. H. (1928) ‘Effect of delayed harvesting on the quality of wheat.’, *Cereal Chem.*, 5, pp. 128–145.

Bradshaw, J. E. (2017) ‘Plant breeding: past, present and future’, *Euphytica*. doi: 10.1007/s10681-016-1815-y.

Brisson, N., Gate, P., Gouache, D. and Charmet, G. (2010) ‘Why are wheat yields stagnating in Europe ? A comprehensive data analysis for France’, *Field Crops Research*, 119, pp. 201–212. doi: 10.1016/j.fcr.2010.07.012.

Brisson, N., Launay, M., Mary, B. and Beaudoin, N. (2009) ‘Conceptual Basis, Formalisations and Parameterization of the Stics Crop Model’, *Editons Quae*.

Burke, M. and Emerick, K. (2016) ‘Adaptation to Climate Change: Evidence from US Agriculture’, *American Economic Journal: Economic Policy*, 8(3), pp. 106–140.

Burnett, D., Barbour, E. and Harrison, G. P. (2014) 'The UK solar energy resource and the impact of climate change', *Renewable Energy*, 71(November), pp. 333–343. doi: 10.1016/j.renene.2014.05.034.

Cabello, J., Arce, A. and Chan, R. (2012) 'The homologous HD-Zip I transcription factors HaHB1 and AtHB13 confer cold tolerance via the induction of pathogenesis-related and glucanase proteins.', *Plant Journal*, 69, pp. 141–153.

Cabello, J. and Chan, R. (2012) 'The homologous homeodomain-leucine zipper transcription factors HaHB1 and AtHB13 confer tolerance to drought and salinity stresses via the induction of proteins that stabilize membranes.', *Plant Biotechnology Journal*, 10, pp. 815–825.

Calderini, D. F. and Slafer, G. A. (1998) 'Changes in yield and yield stability in wheat during the 20th century', *Field Crops Research*, 57, pp. 335–347.

Cammarano, D., Valdivia, R. O., Beletse, Y. G., Durand, W., Crespo, O., Tesfuhuney, W. A., Jones, M. R., Walker, S., Mpuisang, T. N., Nhemachena, C., Ruane, A. C., Mutter, C., Rosenzweig, C. and Antle, J. (2020) 'Integrated assessment of climate change impacts on crop productivity and income of commercial maize farms in northeast South Africa', *Food Security*, 12(3), pp. 659–678. doi: 10.1007/s12571-020-01023-0.

Carreras Navarro, E., Lam, S. K. and Trębicki, P. (2020) 'Elevated Carbon Dioxide and Nitrogen Impact Wheat and Its Aphid Pest', *Frontiers in Plant Science*, 11. doi: 10.3389/fpls.2020.605337.

Cassman, K. G. (1999) 'Ecological intensification of cereal production systems: Yield potential, soil quality, and precision agriculture', *PNAS*, 96, pp. 5952–5959. doi: DOI 10.1073/pnas.96.11.5952.

Cassman, K. G., Dobermann, A., Walters, D. T. and Yang, H. (2003) 'Meeting cereal demand while protecting natural resources and improving environmental quality', *Annual Review of Environment and Resources*, 28(1), pp. 315–358. doi: 10.1146/annurev.energy.28.040202.122858.

Cassman, K. G., Grassini, P. and van Wart, J. (2011) 'Crop Yield Potential, Yield Trends, and Global Food Security in a Changing Climate', in *Handbook Of Climate Change And Agroecosystems: Impacts, Adaptation, And Mitigation*. London: Imperial College Press, pp. 37–51. doi: 10.1142/9781848166561_0004.

Ceccarelli, S., Grando, S. and Baum, M. (2007) ‘Participatory plant breeding in water-limited environments’, *Experimental Agriculture*, 43(4), pp. 411–435. doi: 10.1017/S0014479707005327.

Ceglar, A., Turco, M., Toreti, A. and Doblas-Reyes, F. J. (2017) ‘Linking crop yield anomalies to large-scale atmospheric circulation in Europe’, *Agricultural and Forest Meteorology*, 240–241, pp. 35–45. doi: 10.1016/j.agrformet.2017.03.019.

Ceglar, A., Zampieri, M., Toreti, A. and Dentener, F. (2019) ‘Observed northward migration of agro-climate zones in Europe will further accelerate under climate change’, *Earth’s Future*. doi: 10.1029/2019EF001178.

Centre for Ecology and Hydrology (2022) *CEH Droughts Portal*. Available at: https://eip.ceh.ac.uk/apps/droughts/.

Challinor, A. J., Watson, J., Lobell, D. B., Howden, S. M., Smith, D. R. and Chhetri, N. (2014) ‘A meta-analysis of crop yield under climate change and adaptation’, *Nature Climate Change*, 4(4), pp. 287–291. doi: 10.1038/nclimate2153.

Chaloner, T. M., Gurr, S. J. and Bebber, D. P. (2021) ‘Plant pathogen infection risk tracks global crop yields under climate change’, *Nature Climate Change*, 11(8), pp. 710–715. doi: 10.1038/s41558-021-01104-8.

Chen, X., Cui, Z., Fan, M., Vitousek, P., Zhao, M., Ma, W., Wang, Zhenlin, Zhang, Weijian, Yan, X., Yang, J., Deng, X., Gao, Q., Zhang, Q., Guo, S., Ren, J., Li, S., Ye, Y., Wang, Zhaohui, Huang, J., Tang, Q., Sun, Y., Peng, X., Zhang, J., He, M., Zhu, Y., Xue, J., Wang, G., Wu, Liang, An, N., Wu, Liangquan, Ma, L., Zhang, Weifeng and Zhang, F. (2014) ‘Producing more grain with lower environmental costs’, *Nature*, 514(7253), pp. 486–489. doi: 10.1038/nature13609.

Cho, K., Falloon, P., Gornall, J., Betts, R. and Clark, R. (2012) ‘Winter wheat yields in the UK: Uncertainties in climate and management impacts’, *Climate Research*, 54(1), pp. 49–68. doi: 10.3354/cr01085.

Clarke, D., Hess, T. M., Haro-Monteagudo, D., Semenov, M. A. and Knox, J. W. (2021) ‘Assessing future drought risks and wheat yield losses in England’, *Agricultural and Forest Meteorology*, 297(April 2020), p. 108248. doi: 10.1016/j.agrformet.2020.108248.

Copernicus Climate Change Service (C3S) (2019) *C3S ERA5-Land reanalysis*, *Copernicus Climate Change Service*. Available at: https://cds.climate.copernicus.eu/cdsapp#!/home.

Correia, J. M., McDermott, F., Sweeney, C., Doddy, E. and Griffin, S. (2020) ‘An investigation of the regional correlation gradients between euro-atlantic atmospheric teleconnections and winter solar short wave radiation in Northwest Europe', *Meteorological Applications*, 27(2), pp. 1–17. doi: 10.1002/met.1892.

Covarrubias-Pazaran, G. (2020) *Genetic gain as a high-level key performance indicator*.

Covarrubias-Pazaran, G., Gebeyehu, Z., Gemenet, D., Werner, C., Labroo, M., Sirak, S., Coaldrake, P., Rabbi, I., Kayondo, S. I., Parkes, E., Kanju, E., Mbanjo, E. G. N., Agbona, A., Kulakow, P., Quinn, M. and Debaene, J. (2022) 'Breeding Schemes: What Are They, How to Formalize Them, and How to Improve Them?', *Frontiers in Plant Science*, 12(January), pp. 1–15. doi: 10.3389/fpls.2021.791859.

Cromey, M. G., Wright, D. S. C. and Boddington, H. J. (1998) 'Effects of frost during grain filling on wheat yield and grain structure', *New Zealand Journal of Crop and Horticultural Science*, 26, pp. 279–290. doi: https://doi.org/10.1080/01140.

CSO (2021) *Area, Yield and Production of Crops*. Available at: http://www.cso.ie/en/releasesandpublications/er/aypc/areayieldandproductionofcrops2014/.

Daccache, A., Keay, C., Jones, R. J. A., Weatherhead, E. K., Stalham, M. A. and Knox, J. W. (2012) 'Climate change and land suitability for potato production in England and Wales: Impacts and adaptation', *Journal of Agricultural Science*, 150(2), pp. 161–177. doi: 10.1017/S0021859611000839.

Dawson, I. K., Russell, J., Powell, W., Steffenson, B., Thomas, W. T. B. and Waugh, R. (2015) 'Barley: A translational model for adaptation to climate change', *New Phytologist*, 206(3), pp. 913–931. doi: 10.1111/nph.13266.

DEFRA (2018) *Agriculture in the United Kingdom 2017*.

DEFRA (2020) 'Defra Statistics: Agricultural Facts, England Regional Profiles', (February), p. 21.

DEFRA (2021a) '2020 Agriculture in the United Kingdom'.

DEFRA (2021b) *UK Food Security Report 2021*, *Presented to Parliament pursuant to Section 19 of the Agriculture Act 2020*.

DEFRA (2022) *Government food strategy*.

DEFRA, HSE and Prentis, V. (2020) *Press Release: Outdoor use of metaldehyde to be banned* *to protect wildlife*, *gov.uk*. Available at: https://www.gov.uk/government/news/outdoor-use-of-metaldehyde-to-be-banned-to-protect-wildlife.

Deryng, D., Conway, D., Ramankutty, N., Price, J. and Warren, R. (2014) 'Global crop yield response to extreme heat stress under multiple climate change futures', *Environmental Research Letters*, 9(3). doi: 10.1088/1748-9326/9/3/034011.

Dong, B., Zheng, X., Liu, H., Able, J. A., Yang, H., Zhao, H., Zhang, M., Qiao, Y., Wang, Y. and Liu, M. (2017) 'Effects of Drought Stress on Pollen Sterility, Grain Yield, Abscisic Acid and Protective Enzymes in Two Winter Wheat Cultivars', *Frontiers in Plant Science*, 8. doi: 10.3389/fpls.2017.01008.

Downing, E. and Coe, S. (2018) *Brexit: Future UK agriculture policy*.

Dreccer, M. F., Fainges, J., Whish, J., Ogbonnaya, F. C. and Sadras, V. O. (2018) 'Comparison of sensitive stages of wheat, barley, canola, chickpea and field pea to temperature and water stress across Australia', *Agricultural and Forest Meteorology*, 248(October 2017), pp. 275–294. doi: 10.1016/j.agrformet.2017.10.006.

Dublin City Council (2017) *Dublin City Parks Strategy 2017-2022*. doi: 10.13140/RG.2.2.20732.90247.

EDINA (2022) *Agricultural Census 2010*, *AgCensus*. Available at: https://agcensus.edina.ac.uk/.

EEA (2017) *Climate change impacts and vulnerability in Europe 2016: An indicator-based report*.

Environment Food and Rural Affairs Committee (2018) *The future for food, farming and the environment*.

Erisman, J. W., Galloway, J. N., Seitzinger, S., Bleeker, A., Dise, N. B., Roxana Petrescu, A. M., Leach, A. M. and de Vries, W. (2013) 'Consequences of human modification of the global nitrogen cycle', *Philosophical Transactions of the Royal Society B: Biological Sciences*, 368(1621). doi: 10.1098/rstb.2013.0116.

Espe, M. B., Hill, J. E., Leinfelder-Miles, M., Espino, L. A., Mutters, R., Mackill, D., van Kessel, C. and Linquist, B. A. (2018) 'Rice yield improvements through plant breeding are offset by inherent yield declines over time', *Field Crops Research*, 222, pp. 59–65. doi: 10.1016/j.fcr.2018.03.017.

Fabio, E. S., Kemanian, A. R., Montes, F., Miller, R. O. and Smart, L. B. (2017) 'A mixed model approach for evaluating yield improvements in interspecific hybrids of shrub willow, a dedicated bioenergy crop', *Industrial Crops and Products*, 96, pp. 57–70. doi: 10.1016/j.indcrop.2016.11.019.

Falloon, P., Bebber, D., Bryant, J., Bushell, M., Challinor, A. J., Dessai, S., Gurr, S. and Koehler, A.-K. (2015) 'Using climate information to support crop breeding decisions and adaptation in agriculture', *World Agriculture*, 5(1), pp. 25–43.

Fan, Y. and Li, R. (2012) 'Variable selection in linear mixed effects models', *Annals of Statistics*, 40(4), pp. 2043–2068. doi: 10.1214/12-AOS1028.

FAO (2006) *Food Security: Policy Brief*. doi: 10.1016/j.jneb.2010.12.007.

FAO, IFAD, UNICEF, WHO and WFP (2021) *The State of Food Security and Nutrition in the World 2021*. Rome, Italy: FAO.

Fitzgerald, G. J., Tausz, M., O'Leary, G., Mollah, M. R., Tausz-Posch, S., Seneweera, S., Mock, I., Löw, M., Partington, D. L., McNeil, D. and Norton, R. M. (2016) 'Elevated atmospheric [CO 2 ] can dramatically increase wheat yields in semi-arid environments and buffer against heat waves', *Global Change Biology*, 22(6), pp. 2269–2284. doi: 10.1111/gcb.13263.

Flood, P. J., Harbinson, J. and Aarts, M. G. M. (2011) 'Natural genetic variation in plant photosynthesis', *Trends in Plant Science*, 16(6), pp. 327–335. doi: 10.1016/j.tplants.2011.02.005.

Frederiks, T. M., Christopher, J. T., Harvey, G. L., Sutherland, M. W. and Borrell, A. K. (2012) 'Current and emerging screening methods to identify post-head-emergence frost adaptation in wheat and barley', *Journal of Experimental Botany*, 63(15), pp. 5405–5416. doi: 10.1093/jxb/ers215.

Friedman, J., Hastie, T. and Tibshirani, R. (2010) 'Regularization Paths for Generalized Linear Models via Coordinate Descent', *Journal of Statistical Software*, 33(1), pp. 1–22.

Gillberg, J., Marttinen, P., Mamitsuka, H. and Kaski, S. (2019) 'Modelling G×E with historical weather information improves genomic prediction in new environments', *Bioinformatics*, 35(20), pp. 4045–4052. doi: 10.1093/bioinformatics/btz197.

GISTEMP Team (2022) *GISS Surface Temperature Analysis (GISTEMP), version 4*, *NASA Goddard Institute for Space Studies*. Available at: https://data.giss.nasa.gov/gistemp/.

Gladders, P., Paveley, N. D., Barrie, I. A., Hardwick, N. V., Hims, M. J., Langton, S. and Taylor, M. C. (2001) 'Agronomic and meteorological factors affecting the severity of leaf blotch caused by Mycosphaerella graminicola in commercial wheat crops in England', *Annals of Applied Biology*, 138(3), pp. 301–311. doi: 10.1111/j.1744-7348.2001.tb00115.x.

Gleeson, E., Whelan, E. and Hanley, J. (2017) 'Met Éireann high resolution reanalysis for Ireland', *Advances in Science and Research*, 14, pp. 49–61. doi: 10.5194/asr-14-49-2017.

Goodrich, B., Gabry, J., Ali, I. and Brillerman, S. (2020) 'rstanarm: Bayesian applied regression modeling via Stan'.

Gothard, P. G., Riggs, T. J. and Smith, D. B. (1983) 'The Malting Quality of Some Spring Barley Varieties Grown in England and Wales Between 1880 and 1980', *Journal of the Institute of Brewing*, 89(5), pp. 344–348. doi: 10.1002/j.2050-0416.1983.tb04200.x.

Gouache, D., Bouchon, A. S., Jouanneau, E. and Le Bris, X. (2015) 'Agrometeorological analysis and prediction of wheat yield at the departmental level in France', *Agricultural and Forest Meteorology*, 209–210, pp. 1–10. doi: 10.1016/j.agrformet.2015.04.027.

Grant, R. F., Kimball, B. A., Conley, M. M., White, J. W., Wall, G. W. and Ottman, M. J. (2011) 'Controlled Warming Effects on Wheat Growth and Yield: Field Measurements and Modeling', *Agronomy Journal*, 103(6), pp. 1742–1754. doi: 10.2134/agronj2011.0158.

Grassini, P., Eskridge, K. M. and Cassman, K. G. (2013) 'Distinguishing between yield advances and yield plateaus in historical crop production trends', *Nature Communications*, 4, pp. 1–11. doi: 10.1038/ncomms3918.

Grogan, S. M., Anderson, J., Baenziger, P. S., Frels, K., Guttieri, M. J., Haley, S. D., Kim, K., Liu, S., McMaster, G. S., Newell, M., Prasad, P. V. V., Reid, S. D., Shroyer, K. J., Zhang, G., Akhunov, E. and Byrne, P. F. (2016) 'Phenotypic Plasticity of Winter Wheat Heading Date and Grain Yield across the US Great Plains', *Crop Science*, 56(5), pp. 2223–2236. doi: 10.2135/cropsci2015.06.0357.

Grogan, S. M., Brown-Guedira, G., Haley, S. D., McMaster, G. S., Reid, S. D., Smith, J. and Byrne, P. F. (2016) 'Allelic variation in developmental genes and effects on winter wheat heading date in the U.S. Great Plains', *PLoS ONE*, 11(4), pp. 1–23. doi: 10.1371/journal.pone.0152852.

Groll, A. (2017) 'glmmLasso: Variable Selection for Generalized Linear Mixed Models by L1-Penalized Estimation'.

Groll, A. and Tutz, G. (2014) 'Variable selection for generalized linear mixed models by L1-penalized estimation', *Statistics and Computing*, 24(2), pp. 137–154. doi: 10.1007/s11222-012-9359-z.

Gu, B., Zhang, X., Lam, S. K., Yu, Y., van Grinsven, H. J. M., Zhang, S., Wang, X., Bodirsky, B. L., Wang, S., Duan, J., Ren, C., Bouwman, L., de Vries, W., Xu, J., Sutton, M. A. and Chen, D. (2023) 'Cost-effective mitigation of nitrogen pollution from global croplands', *Nature*, 613(7942), pp. 77–84. doi: 10.1038/s41586-022-05481-8.

Gurung, I., McIntyre, Z. and Hughes, I. (2022) *Food Prices Tracking: July Update*, *The Food Foundation*. Available at: https://foodfoundation.org.uk/news/food-prices-tracking-july-update.

Gusta, L. V. and Fowler, D. B. (1976) 'DEHARDENING AND REHARDENING OF SPRING-COLLECTED WINTER WHEATS AND A WINTER RYE', *Canadian Journal of Plant Science*, 56(4), pp. 775–779. doi: 10.4141/cjps76-126.

Hafner, S. (2003) 'Trends in maize, rice, and wheat yields for 188 nations over the past 40 years: a prevalence of linear growth', *Agriculture, Ecosystems & Environment*, 97(1–3), pp. 275–283. doi: 10.1016/S0167-8809(03)00019-7.

Hakala, K., Jauhiainen, L., Himanen, S. J., Rötter, R., Salo, T. and Kahiluoto, H. (2012) 'Sensitivity of barley varieties to weather in Finland', *The Journal of Agricultural Science*, 150(2), pp. 145–160. doi: 10.1017/S0021859611000694.

Hakala, K., Jauhiainen, L., Rajala, A. A., Jalli, M., Kujala, M. and Laine, A. (2020) 'Different responses to weather events may change the cultivation balance of spring barley and oats in the future', *Field Crops Research*, 259. doi: 10.1016/j.fcr.2020.107956.

Harding, A. E., Rivington, M., Mineter, M. J. and Tett, S. F. B. (2015) 'Agro-meteorological indices and climate model uncertainty over the UK', *Climatic Change*, 128(1–2), pp. 113–126. doi: 10.1007/s10584-014-1296-8.

Hardwick, N. V., Jones, D. R. and Slough, J. E. (2001) 'Factors affecting diseases of winter wheat in England and Wales, 1989-98', *Plant Pathology*, 50(4), pp. 453–462. doi: 10.1046/j.1365-3059.2001.00596.x.

Harkness, C., Semenov, M. A., Areal, F., Senapati, N., Trnka, M., Balek, J. and Bishop, J. (2020) 'Adverse weather conditions for UK wheat production under climate change', *Agricultural and Forest Meteorology*, 282–283(August 2019), p. 107862. doi:

10.1016/j.agrformet.2019.107862.

Hatfield, J. L., Antle, J., Garrett, K. A., Izaurralde, R. C., Mader, T., Marshall, E., Nearing, M., Philip Robertson, G. and Ziska, L. (2020) 'Indicators of climate change in agricultural systems', *Climatic Change*, 163(4), pp. 1719–1732. doi: 10.1007/s10584-018-2222-2.

Hazra, S., Swain, D. K. and Bhadoria, P. B. S. (2019) 'Wheat grown under elevated CO2 was more responsive to nitrogen fertilizer in Eastern India', *European Journal of Agronomy*, 105, pp. 1–12. doi: 10.1016/j.eja.2019.02.001.

Heino, M., Puma, M. J., Ward, P. J., Gerten, D., Heck, V., Siebert, S. and Kummu, M. (2018) 'Two-thirds of global cropland area impacted by climate oscillations', *Nature Communications*, 9(1), pp. 1–10. doi: 10.1038/s41467-017-02071-5.

Henderson, E. (2022) 'Earliest ever barley cut for farms as harvest picks up pace', *Farmers Weekly*.

Hickey, J. M., Chiurugwi, T., Mackay, I. and Powell, W. (2017) 'Genomic prediction unifies animal and plant breeding programs to form platforms for biological discovery', *Nature Genetics*, 49(9), pp. 1297–1303. doi: 10.1038/ng.3920.

Hobbs, J. E. (2020) 'Food supply chains during the COVID-19 pandemic', *Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie*, 68(2), pp. 171–176. doi: 10.1111/cjag.12237.

Hochman, Z., Gobbett, D. L. and Horan, H. (2017) 'Climate trends account for stalled wheat yields in Australia since 1990', *Global Change Biology*, 23(5), pp. 2071–2081. doi: 10.1111/gcb.13604.

Högy, P., Wieser, H., Köhler, P., Schwadorf, K., Breuer, J., Franzaring, J., Muntifering, R. and Fangmeier, A. (2009) 'Effects of elevated CO 2 on grain yield and quality of wheat: results from a 3-year free-air CO 2 enrichment experiment', *Plant Biology*, 11, pp. 60–69. doi: 10.1111/j.1438-8677.2009.00230.x.

Hollis, D., McCarthy, M., Kendon, M., Legg, T. and Simpson, I. (2019) 'HadUK-Grid—A new UK dataset of gridded climate observations', *Geoscience Data Journal*, 6(2), pp. 151–159. doi: 10.1002/gdj3.78.

Hoogenboom, G., Jones, J. W., Porter, C. H., Wilkens, P. W., Boote, K. J., Hunt, L. a. and Tsuji, G. Y. (2010) *Decision Support System for Agrotechnology Transfer Version 4.5. Volume 1: Overview.*, *Agricultural Systems*.

Hoyer, S. and Hamman, J. J. (2017) 'xarray: N-D labeled Arrays and Datasets in Python', *Journal of Open Research Software*, 5. doi: 10.5334/jors.148.

Huang, X. Q., Wolf, M., Ganal, M. W., Orford, S., Koebner, R. M. D. and Röder, M. S. (2007) 'Did modern plant breeding lead to genetic erosion in European winter wheat varieties?', *Crop Science*, 47(1), pp. 343–349. doi: 10.2135/cropsci2006.04.0261.

Hunter, H. (1913) 'Irish Barley Growing Experiments', *Journal of the Institute of Brewing*, pp. 547–577.

IPCC (2022) *Food, Fibre and Other Ecosystem Products*, *IPCC Sixth Assessment Report*. doi: 10.1017/9781009325844.007.714.

Ipsos Mori and Food Standards Agency (2022) *The COVID-19 consumer research*, *food.gov.uk*. Available at: https://www.food.gov.uk/print/pdf/node/4406.

Ivandic, V., Hackett, C. A., Zhang, Z. J., Staub, J. E., Nevo, E., Thomas, W. T. B. and Forster, B. P. (2000) 'Phenotypic responses of wild barley to experimentally imposed water stress', *Journal of Experimental Botany*, 51(353), pp. 2021–2029.

Jablonski, L. M., Wang, X. and Curtis, P. S. (2002) 'Plant reproduction under elevated $CO_2$ conditions: a meta-analysis of reports on 79 crop and wild species', *New Phytologist*, 156(1), pp. 9–26. doi: 10.1046/j.1469-8137.2002.00494.x.

Jamieson, P. D., Semenov, M. A., Brooking, I. R. and Francis, G. S. (1998) 'Sirius: a mechanistic model of wheat response to environmental variation', *European Journal of Agronomy*, 8(3–4), pp. 161–179. doi: 10.1016/S1161-0301(98)00020-3.

Jayaraman, K. (2000) '6.1 Genetics and plant breeding', in *A Statistical Manual For Forestry Research*.

Jones, L., Fitch, A., Evans, C., Stephen, T., Spears, B., Gunn, I., Carvalho, L., May, L., Schonrogge, K., Clilverd, H., Mitchell, Z., Garbutt, A., Taylor, P., Fletcher, D., Gorst, A., Smale, R., Giam, G., Aron, J., Elliott, J., Illman, H., Ray, D., Fung, F., Tinker, J., Berenice-Wilmes, S., King, N. and Malham, S. (2020) 'Climate driven threshold effects in the natural environment. Report to the Climate Change Committee', (May), p. 231.

Jones, P. D., Jonsson, T. and Wheeler, D. (1997) 'Extension to the North Atlantic Oscillation using early instrumental pressure observations from gibraltar and south-west Iceland', *International Journal of Climatology*, 17(13), pp. 1433–1450. doi: 10.1002/(sici)1097-0088(19971115)17:13<1433::aid-joc203>3.3.co;2-g.

Kahiluoto, H., Kaseva, J., Balek, J., Olesen, J. E., Ruiz-Ramos, M., Gobin, A., Kersebaum, K. C., Takáč, J., Ruget, F., Ferrise, R., Bezak, P., Capellades, G., Dibari, C., Mäkinen, H., Nendel, C., Ventrella, D., Rodríguez, A., Bindi, M. and Trnka, M. (2019) 'Decline in climate resilience of European wheat', *Proceedings of the National Academy of Sciences*, 116(1), pp. 123–128. doi: 10.1073/pnas.1804387115.

KC, K. B., Dias, G. M., Veeramani, A., Swanton, C. J., Fraser, D., Steinke, D., Lee, E., Wittman, H., Farber, J. M., Dunfield, K., McCann, K., Anand, M., Campbell, M., Rooney, N., Raine, N. E., Acker, R. Van, Hanner, R., Pascoal, S., Sharif, S., Benton, T. G. and Fraser, E. D. G. (2018) 'When too much isn't enough: Does current food production meet global nutritional needs?', *PLOS ONE*, 13(10). doi: 10.1371/journal.pone.0205683.

Kendon, M. (2022) *Unprecedented extreme heatwave, July 2022*.

Kendon, M., McCarthy, M., Jevrejeva, S., Matthews, A. and Legg, T. (2019) 'State of the UK climate 2018', *International Journal of Climatology*, 39(S1), pp. 1–55. doi: 10.1002/joc.6213.

Kendon, M., McCarthy, M., Jevrejeva, S., Matthews, A., Sparks, T. and Garforth, J. (2021) ' State of the UK Climate 2020 ', *International Journal of Climatology*, 41(S2), pp. 1–76. doi: 10.1002/joc.7285.

Kendon, M., McCarthy, M., Jevrejeva, S., Matthews, A., Sparks, T., Garforth, J. and Kennedy, J. (2022) 'State of the UK Climate 2021', *International Journal of Climatology*, 42(S1), pp. 1–80. doi: 10.1002/joc.7787.

Kettlewell, P. S., Stephenson, D. B., Atkinson, M. D. and Hollins, P. D. (2003) 'Summer rainfall and wheat grain quality: Relationships with the North Atlantic Oscillation', *Weather*, 58(4), pp. 155–164. doi: 10.1256/wea.38.02.

Kirkegaard, J. A., Lilley, J. M., Brill, R. D., Ware, A. H. and Walela, C. K. (2018) 'The critical period for yield and quality determination in canola (Brassica napus L.)', *Field Crops Research*, 222, pp. 180–188. doi: 10.1016/j.fcr.2018.03.018.

Knight, S., Kightley, S., Bingham, I., Hoad, S., Lang, B., Philpott, H., Stobart, R., Thomas, J., Barnes, A. and Ball, B. (2012) *Desk study to evaluate contributory causes of the current 'yield plateau' in wheat and oilseed rape*.

Knox, J. W., Morris, J. and Hess, T. M. (2010) 'Identifying future risks to UK agricultural crop production: Putting climate change in context', *Outlook on Agriculture*, 39(4), pp. 245–248. doi: 10.5367/oa.2010.0015.

Kuhn, M. (2020) ‘caret: Classification and Regression Training.’

Kuznetsova, A., Brockhoff, P. B. and Christensen, R. H. B. (2017) ‘lmerTest Package: Tests in Linear Mixed Effects Models’, *Journal of Statistical Software*, 82(13). doi: 10.18637/jss.v082.i13.

Läänemets, O., Viira, A. H. and Nurmet, M. (2011) ‘Price, yield, and revenue risk in wheat production in Estonia’, *Agronomy Research*, 9(Special Issue II), pp. 421–426.

Lacoste, M., Cook, S., McNee, M., Gale, D., Ingram, J., Bellon-Maurel, V., MacMillan, T., Sylvester-Bradley, R., Kindred, D., Bramley, R., Tremblay, N., Longchamps, L., Thompson, L., Ruiz, J., García, F. O., Maxwell, B., Griffin, T., Oberthür, T., Huyghe, C., Zhang, W., McNamara, J. and Hall, A. (2022) ‘On-Farm Experimentation to transform global agriculture’, *Nature Food*, 3(1), pp. 11–18. doi: 10.1038/s43016-021-00424-4.

Laidig, F., Drobek, T. and Meyer, U. (2008) ‘Genotypic and environmental variability of yield for cultivars from 30 different crops in German official variety trials’, *Plant Breeding*, 127(6), pp. 541–547. doi: 10.1111/j.1439-0523.2008.01564.x.

Laidig, F., Feike, T., Klocke, B., Macholdt, J., Miedaner, T., Rentel, D. and Piepho, H. P. (2021) ‘Long-term breeding progress of yield, yield-related, and disease resistance traits in five cereal crops of German variety trials’, *Theoretical and Applied Genetics*. doi: 10.1007/s00122-021-03929-5.

Laidig, F., Piepho, H. P., Drobek, T. and Meyer, U. (2014) ‘Genetic and non-genetic long-term trends of 12 different crops in German official variety performance trials and on-farm yield trends’, *Theoretical and Applied Genetics*, 127(12), pp. 2599–2617. doi: 10.1007/s00122-014-2402-z.

Lang, T., Millstone, E., Lewis, T. and Marsden, T. (2018) *Feeding Britain: Food Security after Brexit – Food Research Collaboration*.

Lenssen, N. J. L., Schmidt, G. A., Hansen, J. E., Menne, M. J., Persin, A., Ruedy, R. and Zyss, D. (2019) ‘Improvements in the GISTEMP Uncertainty Model’, *Journal of Geophysical Research: Atmospheres*, 124(12), pp. 6307–6326. doi: 10.1029/2018JD029522.

Leung, F., Sitch, S., Tai, A. P. K., Wiltshire, A. J., Gornall, J. L., Folberth, G. A. and Unger, N. (2022) ‘CO 2 fertilization of crops offsets yield losses due to future surface ozone damage and climate change’, *Environmental Research Letters*, 17(7), p. 074007. doi: 10.1088/1748-9326/ac7246.

Lewis, C. M., Persoons, A., Bebber, D. P., Kigathi, R. N., Maintz, J., Findlay, K., Bueno-Sancho, V., Corredor-Moreno, P., Harrington, S. A., Kangara, N., Berlin, A., García, R., Germán, S. E., Hanzalová, A., Hodson, D. P., Hovmøller, M. S., Huerta-Espino, J., Imtiaz, M., Mirza, J. I., Justesen, A. F., Niks, R. E., Omrani, A., Patpour, M., Pretorius, Z. A., Roohparvar, R., Sela, H., Singh, R. P., Steffenson, B., Visser, B., Fenwick, P. M., Thomas, J., Wulff, B. B. H. and Saunders, D. G. O. (2018) 'Potential for re-emergence of wheat stem rust in the United Kingdom', *Communications Biology*, 1(1), p. 13. doi: 10.1038/s42003-018-0013-y.

Licker, R., Johnston, M., Foley, J. A., Barford, C., Kucharik, C. J., Monfreda, C. and Ramankutty, N. (2010) 'Mind the gap: How do climate and agricultural management explain the "yield gap" of croplands around the world?', *Global Ecology and Biogeography*, 19(6), pp. 769–782. doi: 10.1111/j.1466-8238.2010.00563.x.

Lindsey, R. and Dahlman, L. (2021) *Climate Variability: North Atlantic Oscillation*, *Climate.gov*. Available at: https://www.climate.gov/news-features/understanding-climate/climate-variability-north-atlantic-oscillation (Accessed: 18 December 2021).

Lipper, L., Thornton, P., Campbell, B. M., Baedeker, T., Braimoh, A., Bwalya, M., Caron, P., Cattaneo, A., Garrity, D., Henry, K., Hottle, R., Jackson, L., Jarvis, A., Kossam, F., Mann, W., McCarthy, N., Meybeck, A., Neufeldt, H., Remington, T., Sen, P. T., Sessa, R., Shula, R., Tibu, A. and Torquebiau, E. F. (2014) 'Climate-smart agriculture for food security', *Nature Climate Change*, 4(12), pp. 1068–1072. doi: 10.1038/nclimate2437.

Liu, Y., Wang, E., Yang, X. and Wang, J. (2010) 'Contributions of climatic and crop varietal changes to crop production in the North China Plain, since 1980s', *Global Change Biology*, 16(8), pp. 2287–2299. doi: 10.1111/j.1365-2486.2009.02077.x.

Liu, Y., Zhang, H., Li, G., Sun, X. and Wang, M. (2022) 'A comprehensive method to increase yield and narrow the yield gap of winter wheat for sustainable intensification', *Journal of the Science of Food and Agriculture*, 102(10), pp. 4238–4249. doi: 10.1002/jsfa.11775.

Lobell, D. B., Cassman, K. G. and Field, C. B. (2009) 'Crop Yield Gaps: Their Importance, Magnitudes, and Causes', *Annual Review of Environment and Resources*, 34(1), pp. 179–204. doi: 10.1146/annurev.environ.041008.093740.

Lopes, M. S. (2022) 'Will temperature and rainfall changes prevent yield progress in Europe?', *Food and Energy Security*, (March 2021), pp. 1–12. doi: 10.1002/fes3.372.

Lucht, J. M. (2015) 'Public acceptance of plant biotechnology and GM crops', *Viruses*, 7(8), pp. 4254–4281. doi: 10.3390/v7082819.

Lumley, T. (2020) 'leaps: Regression Subset Selection'.

Lüttringhaus, S., Gornott, C., Wittkop, B., Noleppa, S. and Lotze-campen, H. (2020) 'The Economic Impact of Exchanging Breeding Material : Assessing Winter Wheat Production in Germany', 11(December), pp. 1–14. doi: 10.3389/fpls.2020.601013.

Macholdt, J., Hadasch, S., Piepho, H.-P., Reckling, M., Taghizadeh-Toosi, A. and Christensen, B. T. (2021) 'Yield variability trends of winter wheat and spring barley grown during 1932–2019 in the Askov Long-term Experiment', *Field Crops Research*, 264(October 2020). doi: 10.1016/j.fcr.2021.108083.

Mackay, I., Horwell, A., Garner, J., White, J., McKee, J. and Philpott, H. (2011) 'Reanalyses of the historical series of UK variety trials to quantify the contributions of genetic and environmental factors to trends and variability in yield over time', *Theoretical and Applied Genetics*, 122(1), pp. 225–238. doi: 10.1007/s00122-010-1438-y.

Mackay, I., Piepho, H. and Garcia, A. A. F. (2019) 'Statistical Methods for Plant Breeding', in *Handbook of Statistical Genomics*, pp. 501–20. doi: 10.1002/9781119487845.ch17.

Maisey, P., Thornton, H., Fung, F., Harris, G., Lowe, J., McSweeney, C., Mitchell, J., Murphy, J., Rostron, J., Sexton, D. and Yamazaki, K. (2018) *UKCP18 Fact Sheet: Weather Types*. Exeter.

Mäkinen, H., Kaseva, J., Virkajärvi, P. and Kahiluoto, H. (2015) 'Managing resilience of forage crops to climate change through response diversity', *Field Crops Research*, 183, pp. 23–30. doi: 10.1016/j.fcr.2015.07.006.

Malcolm, J. P. (1983) 'The History of Barley in New Zealand', *Agronomy Society of NZ Special Publication*, pp. 3–15.

Malik, A. I., Colmer, T. D., Lambers, H., Setter, T. L. and Schortemeyer, M. (2002) 'Short-term waterlogging has long-term effects on the growth and physiology of wheat', *New Phytologist*, 4, pp. 225–236.

Mandryk, M., Doelman, J. and Stehfest, E. (2015) *Assessment of global land availability: land supply for agriculture*, *FOODSECURE*.

Manschadi, A. M., Christopher, J., DeVoil, P. and Hammer, G. L. (2006) 'The role of root architectural traits in adaptation of wheat to water-limited environments', *Functional Plant Biology*, 33(9), p. 823. doi: 10.1071/FP06055.

Mateus, C., Potito, A. and Curley, M. (2020) ‘Reconstruction of a long-term historical daily maximum and minimum air temperature network dataset for Ireland (1831-1968)’, *Geoscience Data Journal*, 7(2), pp. 102–115. doi: 10.1002/gdj3.92.

Mathieu, J. A. and Aires, F. (2018) ‘Assessment of the agro-climatic indices to improve crop yield forecasting’, *Agricultural and Forest Meteorology*, 253–254, pp. 15–30. doi: 10.1016/j.agrformet.2018.01.031.

Maxted, N., Scholten, M., Allender, C., Astley, D., Vincent, H. and Kell, S. P. (2014) *Landrace conservation strategy for the United Kingdom*. Birmingham.

Met Éireann (2021) *Historical Data*. Available at: https://www.met.ie/climate/available-data/historical-data (Accessed: 19 March 2021).

Met Office (2021a) *North Atlantic Oscillation*, *Met Office*. Available at: https://www.metoffice.gov.uk/weather/learn-about/weather/atmosphere/north-atlantic-oscillation (Accessed: 16 December 2021).

Met Office (2021b) *The North Atlantic Oscillation and the UK*. Available at: https://www.metoffice.gov.uk/research/climate/seasonal-to-decadal/gpc-outlooks/ens-mean/nao-uk.

Met Office (2021c) *UK Climate Projections: Headline Findings*.

Met Office (2022a) *Climate change in the UK*. Available at: https://www.metoffice.gov.uk/weather/climate-change/climate-change-in-the-uk.

Met Office (2022b) *Effects of climate change*. Available at: https://www.metoffice.gov.uk/weather/climate-change/effects-of-climate-change.

Met Office (2022c) *UK and regional series*. Available at: https://www.metoffice.gov.uk/research/climate/maps-and-data/uk-and-regional-series.

Met Office (2022d) *UK regional climates*. Available at: https://www.metoffice.gov.uk/research/climate/maps-and-data/regional-climates/index.

Michel, L. and Makowski, D. (2013) ‘Comparison of statistical models for analyzing wheat yield time series.’, *PloS one*, 8(10), pp. 1–11. doi: 10.1371/journal.pone.0078615.

Miflin, B. (2000) ‘Crop improvement in the 21st century’, *Journal of Experimental Botany*, 51(342), pp. 1–8. doi: 10.1093/jxb/51.342.1.

Molero, G., Coombes, B., Joynson, R., Pinto, F., Piñera-, F. J., Rivera-amado, C., Hall, A. and Reynolds, M. P. (2022) 'Exotic alleles contribute to heat tolerance in wheat under field conditions', *bioRxiv*, pp. 1–26. doi: https://doi.org/10.1101/2022.02.09.479695.

Muñoz Sabater, J. (2019) *ERA5-Land hourly data from 1981 to present.*, *Copernicus Climate Change Service (C3S) Climate Data Store (CDS)*. doi: 10.24381/cds.e2161bac.

Murphy, C., Broderick, C., Burt, T. P., Curley, M., Duffy, C., Hall, J., Harrigan, S., Matthews, T. K. R., Macdonald, N., McCarthy, G., McCarthy, M. P., Mullan, D., Noone, S., Osborn, T. J., Ryan, C., Sweeney, J., Thorne, P., Walsh, S. and Wilby, R. L. (2018) 'A 305-year continuous monthly rainfall series for the island of Ireland (1711-2016)', *Climate of the Past*, 14(3), pp. 413–440. doi: 10.5194/cp-14-413-2018.

Myers, S. S., Zanobetti, A., Kloog, I., Huybers, P., Leakey, A. D. B., Bloom, A. J., Carlisle, E., Dietterich, L. H., Fitzgerald, G., Hasegawa, T., Holbrook, N. M., Nelson, R. L., Ottman, M. J., Raboy, V., Sakai, H., Sartor, K. A., Schwartz, J., Seneweera, S., Tausz, M. and Usui, Y. (2014) 'Increasing CO2 threatens human nutrition', *Nature*, 510(7503), pp. 139–142. doi: 10.1038/nature13179.

Nakagawa, S., Johnson, P. C. D. and Schielzeth, H. (2017) 'The coefficient of determination R2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded', *Journal of the Royal Society Interface*, 14(134). doi: 10.1098/rsif.2017.0213.

Nasehzadeh, M. and Ellis, R. H. (2017) 'Wheat seed weight and quality differ temporally in sensitivity to warm or cool conditions during seed development and maturation', *Annals of Botany*, 120(3), pp. 479–493. doi: 10.1093/aob/mcx074.

National Soil Resources Institute (2022) *The soils of England and Wales*, *UKSO*. Available at: http://www.ukso.org/static-maps/soils-of-england-and-wales.html.

NCAR (2020) *Climate Data: Standardized Precipitation Index (SPI)*, *ClimateDataGuide*. Available at: https://climatedataguide.ucar.edu/climate-data/standardized-precipitation-index-spi.

Nehe, A., Akin, B., Sanal, T., Evlice, A. K., Ünsal, R., Dinçer, N., Demir, L., Geren, H., Sevim, I., Orhan, Ş., Yaktubay, S., Ezici, A., Guzman, C. and Morgounov, A. (2019) 'Genotype x environment interaction and genetic gain for grain yield and grain quality traits in Turkish spring wheat released between 1964 and 2010', *PLoS ONE*, 14(7), pp. 1–18. doi:

10.1371/journal.pone.0219432.

Nesbitt, A., Dorling, S., Jones, R., Smith, D. K. E., Krumins, M., Gannon, K. E., Dorling, L., Johnson, Z. and Conway, D. (2022) 'Climate change projections for UK viticulture to 2040: a focus on improving suitability for Pinot Noir', *OENO One*, 56(3), pp. 69–87. doi: 10.20870/oeno-one.2022.56.3.5398.

Noleppa, S. and Cartsburg, M. (2021) 'The socio-economic and environmental values of plant breeding in the EU and selected EU member states', *hffa Research*.

Noone, S., Murphy, C., Coll, J., Matthews, T., Mullan, D., Wilby, R. L. and Walsh, S. (2016) 'Homogenization and analysis of an expanded long-term monthly rainfall network for the Island of Ireland (1850–2010)', *International Journal of Climatology*, 36(8), pp. 2837–2853. doi: 10.1002/joc.4522.

O'Leary, G. J., Christy, B., Nuttall, J., Huth, N., Cammarano, D., Stöckle, C., Basso, B., Shcherbak, I., Fitzgerald, G., Luo, Q., Farre-Codina, I., Palta, J. and Asseng, S. (2015) 'Response of wheat growth, grain yield and water use to elevated <scp>CO</scp> $_2$ under a Free-Air <scp>CO</scp> $_2$ Enrichment ( <scp>FACE</scp> ) experiment and modelling in a semi-arid environment', *Global Change Biology*, 21(7), pp. 2670–2686. doi: 10.1111/gcb.12830.

Okom, S., Russell, A., Chaudhary, A. J., Scrimshaw, M. D. and Francis, R. A. (2017) 'Impacts of projected precipitation changes on sugar beet yield in eastern England', *Meteorological Applications*, 24(1), pp. 52–61. doi: 10.1002/met.1604.

Olesen, J. E., Trnka, M., Kersebaum, K. C., Skjelvåg, A. O., Seguin, B., Peltonen-Sainio, P., Rossi, F., Kozyra, J. and Micale, F. (2011) 'Impacts and adaptation of European crop production systems to climate change', *European Journal of Agronomy*, 34(2), pp. 96–112. doi: 10.1016/j.eja.2010.11.003.

Palmeri, M. (2017) *Chapter 18: Testing the Assumptions of Multilevel Models*, *A Language, not a Letter: Learning Statistics in R*. Available at: https://ademos.people.uic.edu/Chapter18.html (Accessed: 18 December 2021).

Parkes, B., Higginbottom, T. P., Hufkens, K., Ceballos, F., Kramer, B. and Foster, T. (2019) 'Weather dataset choice introduces uncertainty to estimates of crop yield responses to climate variability and change', *Environmental Research Letters*, 14(12). doi: 10.1088/1748-9326/ab5ebb.

Parolini, G. (2022) ‘Weather, climate, and agriculture: Historical contributions and perspectives from agricultural meteorology’, *WIREs Climate Change*, 13(3), pp. 1–19. doi: 10.1002/wcc.766.

Peltonen-Sainio, P., Jauhiainen, L. and Hakala, K. (2009) ‘Are there indications of climate change induced increases in variability of major field crops in the northernmost European conditions?’, *Agricultural and Food Science*, 18, pp. 206–222.

Peltonen-Sainio, P., Jauhiainen, L. and Hannukkala, A. (2007) ‘Declining rapeseed yields in Finland: How, why and what next?’, *Journal of Agricultural Science*, 145(6), pp. 587–598. doi: 10.1017/S0021859607007381.

Peltonen-Sainio, P., Jauhiainen, L. and Laurila, I. P. (2009) ‘Cereal yield trends in northern European conditions: Changes in yield potential and its realisation’, *Field Crops Research*, 110(1), pp. 85–90. doi: 10.1016/j.fcr.2008.07.007.

Peltonen-Sainio, P., Jauhiainen, L., Trnka, M., Olesen, J. E., Calanca, P., Eckersten, H., Eitzinger, J., Gobin, A., Kersebaum, K. C., Kozyra, J., Kumar, S., Dalla Marta, A., Micale, F., Schaap, B., Seguin, B., Skjelvåg, A. O. and Orlandini, S. (2010) ‘Coincidence of variation in yield and climate in Europe’, *Agriculture, Ecosystems and Environment*, 139(4), pp. 483–489. doi: 10.1016/j.agee.2010.09.006.

Peng, S., Huang, J., Sheehy, J. E., Laza, R. C., Visperas, R. M., Zhong, X., Centeno, G. S., Khush, G. S. and Cassman, K. G. (2004) ‘Rice yields decline with higher night temperature from global warming’, *Proceedings of the National Academy of Sciences of the United States of America*, 101(27), pp. 9971–9975. doi: 10.1073/pnas.0403720101.

Perkins-Kirkpatrick, S. E. and Lewis, S. C. (2020) ‘Increasing trends in regional heatwaves’, *Nature Communications*, 11(1), pp. 1–8. doi: 10.1038/s41467-020-16970-7.

Pfeifroth, U., Sanchez-Lorenzo, A., Manara, V., Trentmann, J. and Hollmann, R. (2018) ‘Trends and Variability of Surface Solar Radiation in Europe Based On Surface- and Satellite-Based Data Records’, *Journal of Geophysical Research: Atmospheres*, 123(3), pp. 1735–1754. doi: 10.1002/2017JD027418.

Pfeifroth, U., Trentmann, J., Hollmann, R., Selbach, N., Werscheck, M. and Meirink, J. F. (2018) ‘ICDR SEVIRI Radiation - based on SARAH-2 methods, Satellite Application Facility on Climate Monitoring’,.

Phillips, T. (2008) ‘Genetically Modified Organisms (GMOs): Transgenic Crops and

Recombinant DNA Technology', *Nature Education*, 1(1), p. 213.

Pidgeon, J. D., Ober, E. S., Qi, A., Clark, C. J. A., Royal, A. and Jaggard, K. W. (2006) 'Using multi-environment sugar beet variety trials to screen for drought tolerance', *Field Crops Research*, 95(2–3), pp. 268–279. doi: 10.1016/j.fcr.2005.04.010.

Piepho, H. P., Denis, J. B. and Van Eeuwijk, F. A. (1998) 'Predicting Cultivar Differences Using Covariates', *Journal of Agricultural, Biological, and Environmental Statistics*, 3(2), pp. 151–162. doi: 10.2307/1400648.

Piepho, H. P., Laidig, F., Drobek, T. and Me Yer, U. (2014) 'Dissecting genetic and non-genetic sources of long-term yield trend in german official variety trials', *Theoretical and Applied Genetics*, 127(5), pp. 1009–1018. doi: 10.1007/s00122-014-2275-1.

Piticar, A. (2019) 'Changes in agro-climatic indices related to temperature in Central Chile', *International Journal of Biometeorology*, 63(4), pp. 499–510. doi: 10.1007/s00484-019-01681-6.

Poli, P., Hersbach, H., Dee, D. P., Berrisford, P., Simmons, A. J., Vitart, F., Laloyaux, P., Tan, D. G. H., Peubey, C., Thépaut, J. N., Trémolet, Y., Hólm, E. V., Bonavita, M., Isaksen, L. and Fisher, M. (2016) 'ERA-20C: An atmospheric reanalysis of the twentieth century', *Journal of Climate*, 29(11), pp. 4083–4097. doi: 10.1175/JCLI-D-15-0556.1.

Polley, R. W. and Thomas, M. R. (1991) 'Surveys of diseases of winter wheat in England and Wales, 1976–1988', *Annals of Applied Biology*, 119(1), pp. 1–20. doi: 10.1111/j.1744-7348.1991.tb04839.x.

Pörtner, H.-O., Roberts, D. C., Poloczanska, E. S., Mintenbeck, K., Tignor, M., Alegría, A., Craig, M., Langsdorf, S., Löschke, S., Möller, V. and Okem, A. (2022) 'Summary for policymakers', in Pörtner, H.-O., Roberts, D. C., Tignor, M., Poloczanska, E. S., Mintenbeck, K., Alegría, A., Craig, M., Langsdorf, S., Löschke, S., Möller, V., Okem, A., and Rama, B. (eds) *Climate Change 2022: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge, UK and New York, NY, USA: Cambridge University Press, pp. 3–33. doi: 10.1017/9781009325844.001.

Posthumus, H., Morris, J., Hess, T. M., Neville, D., Phillips, E. and Baylis, A. (2009) 'Impacts of the summer 2007 floods on agriculture in England', *Journal of Flood Risk Management*, 2(3), pp. 182–189. doi: 10.1111/j.1753-318X.2009.01031.x.

Powell, J. P. and Reinhard, S. (2016) 'Measuring the effects of extreme weather events on yields', *Weather and Climate Extremes*, 12, pp. 69–79. doi: 10.1016/j.wace.2016.02.003.

Prasad, P. V. V. and Djanaguiraman, M. (2014) 'Response of floret fertility and individual grain weight of wheat to high temperature stress: sensitive stages and thresholds for temperature and duration', *Functional Plant Biology*, 41(12), p. 1261. doi: 10.1071/FP14061.

Price, J., Warren, R., Forstenhäusler, N., Wallace, C., Jenkins, R., Osborn, T. and Van Vuuren, D. P. (2022) 'Quantification of meteorological drought risks between 1.5°C and 4°C of global warming in six countries', *Climatic Change*, p. ISSN 0165-0009 (In Press).

Putelat, T., Whitmore, A. P., Senapati, N. and Semenov, M. A. (2021) 'Local impacts of climate change on winter wheat in Great Britain', *Royal Society Open Science*, 8(6). doi: 10.1098/rsos.201669.

QGIS Development Team (2021) *QGIS Geographic Information System, version 3.10.1*. QGIS Association.

Qian, B., De Jong, R., Gameda, S., Huffman, T., Neilsen, D., Desjardins, R., Wang, H. and McConkey, B. (2013) 'Impact of climate change scenarios on Canadian agroclimatic indices', *Canadian Journal of Soil Science*, 93(2), pp. 243–259. doi: 10.4141/CJSS2012-053.

Qiao, L., Wang, X., Smith, P., Fan, J., Lu, Y., Emmett, B., Li, R., Dorling, S., Chen, H., Liu, S., Benton, T. G., Wang, Y., Ma, Y., Jiang, R., Zhang, F., Piao, S., Müller, C., Yang, H., Hao, Y., Li, W. and Fan, M. (2022) 'Soil quality both increases crop production and improves resilience to climate change', *Nature Climate Change*, 12(6), pp. 574–580. doi: 10.1038/s41558-022-01376-8.

R Core Team (2021) 'R: A Language and Environment for Statistical Computing, version 4.0.4'. Vienna, Austria: R Foundation for Statistical Computing.

Ray, D. K., Gerber, J. S., Macdonald, G. K. and West, P. C. (2015) 'Climate variation explains a third of global crop yield variability', *Nature Communications*, 6, pp. 1–9. doi: 10.1038/ncomms6989.

Ray, D. K., Ramankutty, N., Mueller, N. D., West, P. C. and Foley, J. A. (2012) 'Recent patterns of crop yield growth and stagnation', *Nature Communications*, 3, pp. 1293–1297. doi: 10.1038/ncomms2296.

Raymond, J., Mackay, I., Penfield, S., Lovett, A., Philpott, H. and Dorling, S. (no date) 'Combining historical agricultural and climate datasets sheds new light on early 20th century barley performance. [in press]', *Annals of Applied Biology*.

Redhead, J. W., Oliver, T. H., Woodcock, B. A. and Pywell, R. F. (2020) 'The influence of landscape composition and configuration on crop yield resilience', *Journal of Applied Ecology*, 57(11), pp. 2180–2190. doi: 10.1111/1365-2664.13722.

Reid, R., Hunter, H., Stewart, J., Joyce, J., Neville, G. H., Russell, J. and Keen, B. A. (1929) *Malting Barley*, *Rothamsted Conferences*. London.

Rezaei, E. E., Siebert, S. and Ewert, F. (2015) 'Intensity of heat stress in winter wheat—phenology compensates for the adverse effect of global warming', *Environmental Research Letters*, 10(2), p. 024012. doi: 10.1088/1748-9326/10/2/024012.

Rezaei, E. E., Siebert, S., Hüging, H. and Ewert, F. (2018) 'Climate change effect on wheat phenology depends on cultivar change', *Scientific Reports*, 8(1), pp. 1–10. doi: 10.1038/s41598-018-23101-2.

Riley, J. (2022) *Worst drought-hit farms face yield losses of 10-50%*, *Farmers Weekly*. Available at: https://www.fwi.co.uk/news/worst-drought-hit-farms-face-yield-losses-of-10-50.

Ritchie, P. D. L., Harper, A. B., Smith, G. S., Kahana, R., Kendon, E. J., Lewis, H., Fezzi, C., Halleck-Vega, S., Boulton, C. A., Bateman, I. J. and Lenton, T. M. (2019) 'Large changes in Great Britain's vegetation and agricultural land-use predicted under unmitigated climate change', *Environmental Research Letters*, 14(11). doi: 10.1088/1748-9326/ab492b.

Rivington, M., Matthews, K. B., Buchan, K., Miller, D. G., Bellocchi, G. and Russell, G. (2013) 'Climate change impacts and adaptation scope for agriculture indicated by agro-meteorological metrics', *Agricultural Systems*, 114, pp. 15–31. doi: 10.1016/j.agsy.2012.08.003.

Robinson, E. L., Blyth, E. M., Clark, D. B., Comyn-Platt, E. and Rudd, A. C. (2020a) 'Climate hydrology and ecology research support system meteorology dataset for Great Britain (1961-2017) [CHESS-met]'. NERC Environmental Information Data Centre. (Dataset). doi: https://doi.org/10.5285/2ab15bf0-ad08-415c-ba64-831168be7293.

Robinson, E. L., Blyth, E. M., Clark, D. B., Comyn-Platt, E. and Rudd, A. C. (2020b) 'Climate hydrology and ecology research support system potential evapotranspiration dataset for Great Britain (1961-2017) [CHESS-PE]'. NERC Environmental Information Data Centre. (Dataset). doi: https://doi.org/10.5285/9116e565-2c0a-455b-9c68-558fdd9179ad.

Roser, M. (2013) *Future Population Growth*, *Published online at OurWorldInData.org*. Available at: https://ourworldindata.org/future-population-growth.

Van Rossum, G. and Drake, F. L. (2009) *Python 3 Reference Manual, version 3.7.3*. Scotts Valley, CA: CreateSpace.

Roussel, V., Koenig, J., Beckert, M. and Balfourier, F. (2004) 'Molecular diversity in French bread wheat accessions related to temporal trends and breeding programmes', *Theoretical and Applied Genetics*, 108(5), pp. 920–930. doi: 10.1007/s00122-003-1502-y.

Rowlatt, J. (2020) 'Bread price may rise after dire UK wheat harvest', *BBC*.

Ruane, A. C., Hudson, N. I., Asseng, S., Camarrano, D., Ewert, F., Martre, P., Boote, K. J., Thorburn, P. J., Aggarwal, P. K., Angulo, C., Basso, B., Bertuzzi, P., Biernath, C., Brisson, N., Challinor, A. J., Doltra, J., Gayler, S., Goldberg, R., Grant, R. F., Heng, L., Hooker, J., Hunt, L. A., Ingwersen, J., Izaurralde, R. C., Kersebaum, K. C., Kumar, S. N., Müller, C., Nendel, C., O'Leary, G., Olesen, J. E., Osborne, T. M., Palosuo, T., Priesack, E., Ripoche, D., Rötter, R. P., Semenov, M. A., Shcherbak, I., Steduto, P., Stöckle, C. O., Stratonovitch, P., Streck, T., Supit, I., Tao, F., Travasso, M., Waha, K., Wallach, D., White, J. W. and Wolf, J. (2016) 'Multi-wheat-model ensemble responses to interannual climate variability', *Environmental Modelling and Software*, 81, pp. 86–101. doi: 10.1016/j.envsoft.2016.03.008.

Rutkoski, Jessica E. (2019) 'A practical guide to genetic gain', in *Advances in Agronomy*. Elsevier Inc., pp. 217–249. doi: 10.1016/bs.agron.2019.05.001.

Rutkoski, J. E. (2019) 'Estimation of realized rates of genetic gain and indicators for breeding program assessment', *Crop Science*, 59(3), pp. 981–993. doi: 10.2135/cropsci2018.09.0537.

Ryan, C., Murphy, C., McGovern, R., Curley, M. and Walsh, S. (2021) 'Ireland's pre-1940 daily rainfall records', *Geoscience Data Journal*, 8(1), pp. 11–23. doi: 10.1002/gdj3.103.

Sacks, W. J., Deryng, D., Foley, J. A. and Ramankutty, N. (2010) 'Crop planting dates: An analysis of global patterns', *Global Ecology and Biogeography*, 19(5), pp. 607–620. doi: 10.1111/j.1466-8238.2010.00551.x.

Sadok, W. and Jagadish, S. V. K. (2020) 'The Hidden Costs of Nighttime Warming on Yields', *Trends in Plant Science*, 25(7), pp. 644–651. doi: 10.1016/j.tplants.2020.02.003.

Sage, T. L., Bagha, S., Lundsgaard-Nielsen, V., Branch, H. A., Sultmanis, S. and Sage, R. F. (2015) 'The effect of high temperature stress on male and female reproduction in plants', *Field Crops Research*, 182, pp. 30–42. doi: 10.1016/j.fcr.2015.06.011.

Saini, H., Sedgley, M. and Aspinall, D. (1983) 'Effect of Heat Stress During Floral Development on Pollen Tube Growth and Ovary Anatomy in Wheat (Triticum aestivum L.)', *Functional Plant Biology*, 10(2), p. 137. doi: 10.1071/PP9830137.

Savill, G. P., Michalski, A., Powers, S. J., Wan, Y., Tosi, P., Buchner, P. and Hawkesford, M. J. (2018) 'Temperature and nitrogen supply interact to determine protein distribution gradients in the wheat grain endosperm', *Journal of Experimental Botany*, 69(12), pp. 3117–3126. doi: 10.1093/jxb/ery127.

Schachtschabel, P., Blume, H. P., Hartge, K. H. and Schwertmann, U. (1976) *Lehrbuch der Bodenkuncle*. Stuttgart: Enke.

Schloerke, B., Cook, D., Larmarange, J., Briatte, F., Marbach, M., Thoen, E., Elberg, A. and Crowley, J. (2021) *GGally: Extension to 'ggplot2'. Version 2.1.2*.

Schmidt, S. B., George, T. S., Brown, L. K., Booth, A., Wishart, J., Hedley, P. E., Martin, P., Russell, J. and Husted, S. (2019) 'Ancient barley landraces adapted to marginal soils demonstrate exceptional tolerance to manganese limitation', *Annals of Botany*, 123(5), pp. 831–843. doi: 10.1093/aob/mcy215.

Schulzweida, U. (2021) 'CDO User Guide (Version 2.0.0)'. doi: http://doi.org/10.5281/zenodo.5614769.

Semenov, M. A. (2009) 'Impacts of climate change on wheat in England and Wales', *Journal of The Royal Society Interface*, 6(33), pp. 343–350. doi: 10.1098/rsif.2008.0285.

Semenov, M. A. and Shewry, P. R. (2011) 'Modelling predicts that heat stress, not drought, will increase vulnerability of wheat in Europe', *Scientific Reports*, 1(1), p. 66. doi: 10.1038/srep00066.

Senapati, N., Halford, N. G. and Semenov, M. A. (2021) 'Vulnerability of European wheat to extreme heat and drought around flowering under future climate', *Environmental Research Letters*, 16(2). doi: 10.1088/1748-9326/abdcf3.

Senapati, N. and Semenov, M. A. (2019) 'Assessing yield gap in high productive countries by designing wheat ideotypes', *Scientific Reports*, 9(1), pp. 1–12. doi: 10.1038/s41598-019-40981-0.

Senapati, N. and Semenov, M. A. (2020) 'Large genetic yield potential and genetic yield gap estimated for wheat in Europe', *Global Food Security*, 24(September 2019). doi: 10.1016/j.gfs.2019.100340.

Senapati, N., Semenov, M. A., Halford, N. G., Hawkesford, M. J., Asseng, S., Cooper, M., Ewert, F., van Ittersum, M. K., Martre, P., Olesen, J. E., Reynolds, M., Rötter, R. P. and Webber, H. (2022) 'Global wheat production could benefit from closing the genetic yield gap', *Nature Food*, 3(7), pp. 532–541. doi: 10.1038/s43016-022-00540-9.

Seyedsadr, M., Cornelius, P. and Crossa, J. (1996) 'Statistical Tests and Estimators of Multiplicative Models for Genotype-by-Environment Interaction', in *Genotype-by-Environment Interaction*. CRC Press, pp. 199–234. doi: 10.1201/9781420049374.ch8.

Sheehan, H. and Bentley, A. (2021) 'Changing times: Opportunities for altering winter wheat phenology', *Plants People Planet*, 3(2), pp. 113–123. doi: 10.1002/ppp3.10163.

Shmueli, G. (2010) 'To explain or to predict?', *Statistical Science*, 25(3), pp. 289–310. doi: 10.1214/10-STS330.

Shorinola, O., Simmonds, J., Wingen, L. U. and Uauy, C. (2022) 'Trend, population structure, and trait mapping from 15 years of national varietal trials of UK winter wheat', *G3 Genes|Genomes|Genetics*, 12(2). doi: 10.1093/g3journal/jkab415.

Simpson, C. J. C., Reeves, G., Tripathi, A., Singh, P. and Hibberd, J. M. (2022) 'Using breeding and quantitative genetics to understand the C4 pathway', *Journal of Experimental Botany*, 73(10), pp. 3072–3084. doi: 10.1093/jxb/erab486.

Sinha, P., Singh, V. K., Bohra, A., Kumar, A., Reif, J. C. and Varshney, R. K. (2021) 'Genomics and breeding innovations for enhancing genetic gain for climate resilience and nutrition traits', *Theoretical and Applied Genetics*, 134(6), pp. 1829–1843. doi: 10.1007/s00122-021-03847-6.

Smith, A. B., Cullis, B. R. and Thompson, R. (2005) 'The analysis of crop cultivar breeding and evaluation trials: An overview of current mixed model approaches', *Journal of Agricultural Science*, 143(6), pp. 449–462. doi: 10.1017/S0021859605005587.

Springmann, M., Clark, M., Mason-D'Croz, D., Wiebe, K., Bodirsky, B. L., Lassaletta, L., de Vries, W., Vermeulen, S. J., Herrero, M., Carlson, K. M., Jonell, M., Troell, M., DeClerck, F., Gordon, L. J., Zurayk, R., Scarborough, P., Rayner, M., Loken, B., Fanzo, J., Godfray, H. C. J., Tilman, D., Rockström, J. and Willett, W. (2018) 'Options for keeping the food system within environmental limits', *Nature*, 562(7728), pp. 519–525. doi: 10.1038/s41586-018-0594-0.

Steffen, W., Richardson, K., Rockström, J., Cornell, S. E., Fetzer, I., Bennett, E. M., Biggs, R., Carpenter, S. R., de Vries, W., de Wit, C. A., Folke, C., Gerten, D., Heinke, J., Mace, G. M., Persson, L. M., Ramanathan, V., Reyers, B. and Sörlin, S. (2015) 'Planetary boundaries: Guiding human development on a changing planet', *Science*, 347(6223). doi: 10.1126/science.1259855.

Steffenson, B. J., Jin, Y., Brueggeman, R. S., Kleinhofs, A. and Sun, Y. (2009) 'Resistance to Stem Rust Race TTKSK Maps to the rpg4 / Rpg5 Complex of Chromosome 5H of Barley', *Phytopathology®*, 99(10), pp. 1135–1141. doi: 10.1094/PHYTO-99-10-1135.

Stocker, T. F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S. K., Boschung, J. and Al., E. (2013) *Climate change 2013: the physical science basis. Contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change.*

Student (1923) 'On Testing Varieties of Cereals', *Biometrika*, 15(3/4), pp. 271–293.

Takashima, N. E., Rondanini, D. P., Puhl, L. E. and Miralles, D. J. (2013) 'Environmental factors affecting yield variability in spring and winter rapeseed genotypes cultivated in the southeastern Argentine Pampas', *European Journal of Agronomy*, 48, pp. 88–100. doi: 10.1016/j.eja.2013.01.008.

Tanguy, M., Dixon, H., Prosdocimi, I., Morris, D. G. and Keller, V. D. J. (2019) *Gridded estimates of daily and monthly areal rainfall for the United Kingdom (1890-2017) [CEH-GEAR]*, *NERC Environmental Information Data Centre. (Dataset)*. Available at: https://doi.org/10.5285/ee9ab43d-a4fe-4e73-afd5-cd4fc4c82556.

Tasker, J. (2020) 'Cereal yields set to hit 30-year low as weather takes toll', *Farmers Weekly*.

Taub, D. R., Miller, B. and Allen, H. (2008) 'Effects of elevated CO 2 on the protein concentration of food crops: a meta-analysis', *Global Change Biology*, 14(3), pp. 565–575. doi: 10.1111/j.1365-2486.2007.01511.x.

TEAGASC (2017) *The Spring Barley Guide*, *Crops Environment & Land Use Programme*.

The Royal Society (2016) *GM plants. Questions and answers*. Available at: https://royalsociety.org/-/media/policy/projects/gm-plants/gm-plant-q-and-a.pdf.

Thursfield, L. (2019) *Oilseed Rape and the Cabbage Stem Flea Beetle*, *JIC*. Available at: https://www.jic.ac.uk/blog/oilseed-rape-and-the-cabbage-stem-flea-beetle/.

Tibshirani, R. (1996) 'Regression Shrinkage and Selection Via the Lasso', *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1), pp. 267–288. doi: 10.1111/j.2517-6161.1996.tb02080.x.

Trigo, R. M., Osborn, T. J. and Corte-Real, J. M. (2002) 'The North Atlantic Oscillation influence on Europe: Climate impacts and associated physical mechanisms', *Climate Research*, 20(1), pp. 9–17. doi: 10.3354/cr020009.

Trnka, M., Eitzinger, J., Dubrovský, M., Semerádová, D., Štěpánek, P., Hlavinka, P., Balek, J., Skalák, P., Farda, A., Formayer, H. and Žalud, Z. (2010) 'Is rainfed crop production in central Europe at risk? Using a regional climate model to produce high resolution agroclimatic information for decision makers', *Journal of Agricultural Science*, 148(6), pp. 639–656. doi: 10.1017/S0021859610000638.

Trnka, M., Hlavinka, P. and Semenov, M. A. (2015) 'Adaptation options for wheat in Europe will be limited by increased adverse weather events under climate change', *Journal of The Royal Society Interface*, 12(112), p. 20150721. doi: 10.1098/rsif.2015.0721.

Trnka, M., Rötter, R. P., Ruiz-Ramos, M., Kersebaum, K. C., Olesen, J. E., Žalud, Z. and Semenov, M. A. (2014) 'Adverse weather conditions for European wheat production will become more frequent with climate change', *Nature Climate Change*, 4(7), pp. 637–643. doi: 10.1038/nclimate2242.

Turner, J. A., Chantry, T., Taylor, M. C. and Kennedy, M. C. (2021) 'Changes in agronomic practices and incidence and severity of diseases in winter wheat in England and Wales between 1999 and 2019', *Plant Pathology*, 70(8), pp. 1759–1778. doi: 10.1111/ppa.13433.

Upadhyaya, H. D., Sharma, S., Reddy, K. N., Saxena, R., Varshney, R. K. and Gowda, C. L. L. (2013) 'Pigeonpea', in *Genetic and Genomic Resources of Grain Legume Improvement*. Elsevier, pp. 181–202. doi: 10.1016/B978-0-12-397935-3.00008-6.

Vaughan, A. (2022) *The Genetic Technology (Precision Breeding) Bill being introduced to Parliament will allow gene-edited crops to be developed and grown in England and sold in Great Britain*, *New Scientist*. Available at: https://www.newscientist.com/article/2321556-uk-to-relax-law-on-gene-edited-food-in-post-brexit-change-from-eu/.

Venables, W. N. and Ripley, B. D. (2002) *Modern Applied Statistics with S*. Fourth. New York: Springer.

Vermeulen, S. J., Park, T., Khoury, C. K. and Béné, C. (2020) 'Changing diets and the transformation of the global food system', *Annals of the New York Academy of Sciences*, 1478(1), pp. 3–17. doi: 10.1111/nyas.14446.

Vicente-Serrano, S. M., Beguería, S. and López-Moreno, J. I. (2010) 'A Multiscalar Drought

Index Sensitive to Global Warming: The Standardized Precipitation Evapotranspiration Index', *Journal of Climate*, 23(7), pp. 1696–1718. doi: 10.1175/2009JCLI2909.1.

Villegas, D., Alfaro, C., Ammar, K., Cátedra, M. M., Crossa, J., García del Moral, L. F. and Royo, C. (2016) 'Daylength, Temperature and Solar Radiation Effects on the Phenology and Yield Formation of Spring Durum Wheat', *Journal of Agronomy and Crop Science*, 202(3), pp. 203–216. doi: 10.1111/jac.12146.

Walsh, M. K., Backlund, P., Buja, L., DeGaetano, A., Melnick, R., Prokopy, L., Takle, E., Todey, D. and Ziska, L. (2020) 'Climate Indicators for Agriculture. USDA Technical Bulletin 1953', pp. 1–75.

Wang, D., Jenkins, K., Forstenhäusler, N., Lei, T., Price, J., Warren, R., Jenkins, R. and Guan, D. (2021) 'Economic impacts of climate-induced crop yield changes: evidence from agri-food industries in six countries', *Climatic Change*, 166(3), pp. 1–19. doi: 10.1007/s10584-021-03062-8.

West, T. T. (2006) *Malting the Barley: John H. Bennett, the Man and His Firm - 200 Years of Malting Barley in Ballinacurra*. Cork: Charleston House.

Weymann, W., Böttcher, U., Sieling, K. and Kage, H. (2015) 'Effects of weather conditions during different growth phases on yield formation of winter oilseed rape', *Field Crops Research*, 173, pp. 41–48. doi: 10.1016/j.fcr.2015.01.002.

Wheeler, T. R., Craufurd, P. Q., Ellis, R. H., Porter, J. R. and Vara Prasad, P. V. (2000) 'Temperature variability and the yield of annual crops', *Agriculture, Ecosystems and Environment*, 82(1–3), pp. 159–167. doi: 10.1016/S0167-8809(00)00224-3.

White, J. W., Hoogenboom, G., Kimball, B. A. and Wall, G. W. (2011) 'Methodologies for simulating impacts of climate change on crop production', *Field Crops Research*, 124(3), pp. 357–368. doi: 10.1016/j.fcr.2011.07.001.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T., Miller, E., Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D., Spinu, V., Takahashi, K., Vaughan, D., Wilke, C., Woo, K. and Yutani, H. (2019) 'Welcome to the Tidyverse', *Journal of Open Source Software*, 4(43). doi: 10.21105/joss.01686.

Wreford, A. and Adger, W. N. (2010) 'Adaptation in agriculture: Historic effects of heat waves and droughts on UK agriculture', *International Journal of Agricultural Sustainability*, 8(4), pp. 278–289. doi: 10.3763/ijas.2010.0482.

Wu, Xiaojian, Cai, K., Zhang, G. and Zeng, F. (2017) ‘Metabolite Profiling of Barley Grains Subjected to Water Stress: To Explain the Genotypic Difference in Drought-Induced Impacts on Malting Quality’, *Frontiers in Plant Science*, 8(September), pp. 1–12. doi: 10.3389/fpls.2017.01547.

Wu, Xiuchen, Liu, H., Li, X., Tian, Y. and Mahecha, M. D. (2017) ‘Responses of winter wheat yields to warming-mediated vernalization variations across temperate europe’, *Frontiers in Ecology and Evolution*, 5(OCT), pp. 1–10. doi: 10.3389/fevo.2017.00126.

Wu, X., Vuichard, N., Ciais, P., Viovy, N., De Noblet-Ducoudré, N., Wang, X., Magliulo, V., Wattenbach, M., Vitale, L., Di Tommasi, P., Moors, E. J., Jans, W., Elbers, J., Ceschia, E., Tallec, T., Bernhofer, C., Grünwald, T., Moureaux, C., Manise, T., Ligne, A., Cellier, P., Loubet, B., Larmanou, E. and Ripoche, D. (2016) ‘ORCHIDEE-CROP (v0), a new process-based agro-land surface model: Model description and evaluation over Europe’, *Geoscientific Model Development*, 9(2), pp. 857–873. doi: 10.5194/gmd-9-857-2016.

Xu, S. and Chong, K. (2018) ‘Remembering winter through vernalisation’, *Nature Plants*, 4(12), pp. 997–1009. doi: 10.1038/s41477-018-0301-z.

Xu, Y., Li, P., Zou, C., Lu, Y., Xie, C., Zhang, X., Prasanna, B. M. and Olsen, M. S. (2017) ‘Enhancing genetic gain in the era of molecular breeding’, *Journal of Experimental Botany*, 68(11), pp. 2641–2666. doi: 10.1093/jxb/erx135.

Yang, X., Tian, Z., Sun, L., Chen, B., Tubiello, F. N. and Xu, Y. (2017) ‘The impacts of increased heat stress events on wheat yield under climate change in China’, *Climatic Change*, 140(3–4), pp. 605–620. doi: 10.1007/s10584-016-1866-z.

Yawson, D. O., Ball, T., Adu, M. O., Mohan, S., Mulholland, B. J. and White, P. J. (2016) ‘Simulated regional yields of spring barley in the United Kingdom under projected climate change’, *Climate*, 4(4), pp. 1–21. doi: 10.3390/cli4040054.

Zhang, Y., Feng, L. P., Wang, J., Wang, E. L. and Xu, Y. L. (2013) ‘Using APSIM to explore wheat yield response to climate change in the North China Plain: the predicted adaptation of wheat cultivar types to vernalization’, *The Journal of Agricultural Science*, 151(6), pp. 836–848. doi: 10.1017/S0021859612000883.

Zhu, P., Abramoff, R., Makowski, D. and Ciais, P. (2021) ‘Uncovering the Past and Future Climate Drivers of Wheat Yield Shocks in Europe With Machine Learning’, *Earth’s Future*, 9(5), pp. 1–13. doi: 10.1029/2020ef001815.

Ziolkowska, J. R. and Zubillaga, J. (2018) ‘Importance of weather monitoring for agricultural decision-making – an exploratory behavioral study for Oklahoma Mesonet’, *Journal of the Science of Food and Agriculture*, 98(13), pp. 4945–4954. doi: 10.1002/jsfa.9027.

Zou, H. and Hastie, T. (2005) ‘Regularization and variable selection via the elastic net’, *Journal of the Royal Statistical Society. Series B: Statistical Methodology*, 67(2), pp. 301–320. doi: 10.1111/j.1467-9868.2005.00503.x.