Changes in agricultural climate in South-Eastern England from 1892 to 2016 and differences in cereal and permanent grassland yield

Title: Changes in agricultural climate in South-Eastern England from 1892 to 2016 and differences in cereal and permanent grassland yield

Authors: John W. G. Addy[a, b, *], Richard H. Ellis[b], Andy J. Macdonald[c], Mikhail A. Semenov[d], Andrew Mead[a]

[a] Computational and Analytical Sciences, Rothamsted Research, UK

[b] School of Agriculture, Policy and Development, University of Reading, UK

[c] Sustainable Agriculture Sciences, Rothamsted Research, UK

[d] Plant Sciences, Rothamsted Research, UK

*Corresponding author address: John Addy,

Computational and Analytical Sciences, Rothamsted Research, Harpenden, Herts., AL5 2JQ, UK

Email: john.addy@rothamsted.ac.uk

Abstract

The long-term increasing trend of annual mean temperature is only one aspect of recent climate change. Other changes in climate, seen in within-year weather patterns relevant to crop production, have also occurred since the late-19th Century. Multivariate analysis combining Prinipal Components Analysis and K-means clustering applied to temporal meteorological datasets (monthly summaries of rainfall, temperature and sunlight duration at Rothamsted Research, UK, between 1892 and 2016) identified ten distinct clusters of years, each with different annual weather patterns. The frequency of occurrence of the years within each cluster altered considerably during this period, with the late 20th and early 21st Century distinctly different to earlier in the 20th Century, providing clear evidence of climate change with regard to the whole weather profile rather than just warming alone. The most-frequently represented cluster of the 21st Century to date had warmer temperatures with more intense rainfall but a dry June, compared to all other clusters. Half of the clusters identified were not represented in the most-recent 25-year period. Analysis of the total biomass yield of winter wheat (Triticum aestivum L.), spring barley (Hordeum vulgare L.), and grassland amongst the different weather clusters showed that years in clusters typical of the 20th Century climate provided greater off-take than those from the early-21st Century, but this impact was less for the pasture than for the two cereal crops implying herbage production was the more resilient to the changing climate at this site.

Keywords: Climate change; Wheat (Triticum aestivum L.); Barley (Hordeum vulgare L.); pasture; Multivariate analysis

1. Introduction

Climate change, as reflected in increases in mean annual temperature, is highly likely to exceed 1.5°C by 2081–2100 relative to 1850–1900, and possibly 2.0°C, with more frequent hot, and fewer cold, temperature extremes over most land areas globally and greater precipitation likely at higher latitudes (IPCC, 2014). But considerable global warming has already occurred. In 2016, the average global temperature was 1.43°C above the 20th Century average (NOAA, 2017). Warming is expected to be greater at higher latitudes (IPCC, 2014), and the changes recorded at different sites do vary. In a study of diverse vegetable-growing sites globally over 37 years or more, 24 sites showed a warming trend, five showed little evidence of an increasing or decreasing trend in mean annual air temperature, whilst only one site showed a cooling trend (Keatinge et al., 2014).

Some effects of weather on crops can be assessed through changes in mean environmental conditions. For example, the effect of variation in temperature and photoperiod on the duration of the vegetative phase of development in annual crops, and so on genotypic adaptation to conditions in different locations globally (Roberts et al., 1996). On the other hand, variation in crop yield and quality can depend greatly upon weather conditions at particular, highly sensitive periods within the growing season. In wheat (Triticum aestivum L.), grain yield was damaged greatly by brief periods of high temperature at early booting (Barber et al., 2017) and also at around the time of flowering (Barber et al., 2017; Ferris et al., 1998; Wheeler et al., 1996); the response of grain yield to nitrogen fertilizer was especially sensitive to the weather at a few discrete times of the year (Addy et al., 2020); yield was reduced but grain quality enhanced by drought in early- or mid-grain filling, respectively (Gooding et al., 2003); and high temperature imposed early in seed development reduced both seed quality and protein quality but increased nitrogen and sulphur contents in the grain, whereas high temperature imposed late in the seed maturation phase improved seed quality (Nasehzadeh and Ellis, 2017). Hence the pattern of weather variables during the year as well as the average changes amongst years are important to crop production and to assessing climate change impacts on agriculture and food supply (Porter and Semenov, 2005; Trnka et al., 2014).

Understanding simultaneous changes across multiple weather variables provides an understanding of how the climate has changed for agriculture, and potentially how this influences yield. A recent machine-learning and process-based modelling approach showed how understanding these climate interactions was complex with no single answer to yield failures (Webber et al., 2020). The use of multivariate analysis methods, in contrast to univariate analysis approaches, can provide insight into how the whole climate system has changed over several variables. Moreover, the pattern of variation in each weather variable during a year, the variation in these patterns amongst years, and whether or not the frequency of the different patterns is changing, is highly relevant to crop production and plant productivity more widely. Multivariate analysis methods have been applied before in climate studies. Cluster analysis was used to partition climate zones of the conterminous United States over temperature and precipitation variables from 1931 to 1980 (Fovell and Fovell, 1993). Further, using data from 1950 to 2002, cluster analysis was used to describe cyclone trajectories in the western North Pacific (Camargo et al., 2007).

The Rothamsted Meteorological Station (RMS) has recorded daily rainfall, temperature and sunlight together since 1892, whilst Rothamsted’s Long-Term Experiments (LTE) with winter wheat, spring barley (Hordeum vulgare L.), and pasture grasses began even earlier in the 19th Century (Macdonald et al., 2018). The objective of this study was to apply multivariate analyses to monthly summarised rainfall, temperature and sunlight duration data from the RMS to test the hypothesis that annual weather patterns could be objectively categorised into separate weather clusters, representing characteristic climates throughout history over several variables, and to assess whether the frequency of membership of these weather clusters changed over time. Subsequently, we tested if the weather clusters identified were associated with differences in the total biomass off-take for winter wheat, spring barley (where biomass is grain plus straw), and forage yields on long-term pasture (two herbage cuts per year; one in mid June and another in Oct/Nov) from the LTEs at the Rothamsted site (Harpenden, Hertfordshure, UK). Total biomass of winter wheat and spring barley were used to compare the off-take with forage yields from long-term pasture.

2. Methods

2.1 Rothamsted Meteorological Data

Data from the Rothamsted Meteorological Station comprised of daily rainfall, temperature and sunlight records from 1892 to 2016. The Rothamsted Meteorological Station has one of the longest meteorological records in the world and is used for all field data from Rothamsted Research. These data were summarised into monthly values for every crop growing year (October to September) during this period for each of total rainfall (mm), rainfall intensity (mm/day) (calculated as total rainfall divided by the number of days with > 1mm of rainfall), mean daily maximum temperature (°C), mean daily minimum temperature (°C), and total sunlight hours. Including mean daily maxima and minima captured the range between extreme mean temperatures, but to also consider impacts of the monthly extreme temperatures in the analyses the overall minimum temperature (°C) and the number of days for which maximum temperatures exceeded 31°C were also included (the latter for the four months June, July, August and September only each year, as these were the only months 31°C was exceeded). These monthly summaries of weather data were derived to capture both within-year variability and seasonal differences amongst years. In total 76 variables ($12 \times 6 + 4$) were calculated for each of the 125 years.

2.2 Rothamsted Long-Term Experiment Data

To make comparisons across the three experiments over multiple years, total biomass offtake (grain and straw off-take at 85% dry matter) data were used from the Broadbalk winter wheat experiment (continuous winter wheat, Section 1), the Hoosfield Continuous spring barley experiment (Series O), and the Park Grass Continuous herbage experiment ('a' subplots) from 1968 onwards (Macdonald et al., 2018). Note that within a genotype at this site the harvest index (grain yield /biomass yield) showed no trend over time in simulations (Addy et al., 2021). Park Grass yields were converted from 100% to 85% dry matter for this study. Yields from five different fertilizer treatments were used, including PKNaMg (Minerals), 48 kg N ha^{-1} + PKNaMg (N1 + Minerals), 96 kg N ha^{-1} + PKNaMg (N2 + Minerals), farmyard manure (35 t ha^{-1} fresh manure p.a. on Broadbalk and Hoosfield; manures alternate every two years between farmyard manure and poultry manure (supplying 65 kg N ha^{-1}) on Park Grass; poultry manure replaced fishmeal since 2003), or no inputs (Nil) because these covered a wide range of inputs and were applied consistently across all three experiments. However, the fertilizer forms differed between experiments. On Park Grass the N was applied as sodium nitrate, whilst on Hoosfield N was applied as calcium ammonium nitrate (Nitro-chalk). On Broadbalk N was applied as Nitro-chalk in 1968-85; since 1986 N has been applied as ammonium nitrate. P and K were applied to all experiments as triple superphosphate and potassium sulphate to ensure crop growth was not nutrient limited. However, in 2001 and 2003 P was withheld on some plots on Broadbalk and Hoosfield, respectively, to allow soil P concentrations to decline to more appropriate agronomic levels. In 2017 P inputs on Park Grass were decreased because many plots contained more than adequate levels of P for herbage production; further details are given in Macdonald et al. (2018). Yield data before 1968 was excluded due to a lack of homogeneity of agricultural practices prior to the introduction of short-strawed cereal cultivars in 1968, whilst total biomass was analysed to avoid cereal cultivar effects on partitioning to grain yield (Austin and Ford, 1989) in Broadbalk and Hoosfield and for comparability with pasture yield.

The soil of Broadbalk (wheat) is described as silty clay loam (Avery & Catt, 1995) with a top soil (0 – 23 cm) texture of 25% sand, 50% silt and 25% clay (Gregory et al., 2010). Hoosfield (barley) soil is described as flinty silty clay loam with a top soil (0 – 23 cm) texture of 28% sand, 52% silt and 20% clay (Blake et al., 2003). Park Grass (herbage) soil is described as silty clay loam (Avery & Catt, 1995), with a soil texture of 23% clay, 58% silt and 19% sand (Blake et al., 2003).

The Rothamsted Meteorological Station was representative of the environment of all three experiments due to their proximity (e-RA Rothamsted, 2021). The distance from each of the three field experiment sites to the Rothamsted Meteorological Station were 0.94, 1.24 and 0.88 km, respectively.

2.3 Statistical Analysis

Principal components analysis (PCA) was used as a dimension reduction tool to identify the key sources of variation amongst the weather variables. Due to the seven underlying variables being measured on different scales, the principal components (PCs) were constructed from the correlation matrix. The loadings for each of the 76 monthly summaries identify those variables that contribute most to each of the principal components (those with the largest absolute loadings). The communalities for each monthly summaried weather variable for the n^{th} dimension, calculated as the square root of the sum of squared loadings for the first n components, identify the overall contribution of each weather variable to the n-dimensional solution (larger values indicate a greater overall contribution). The scores then show the value for each year on each of the principal components, identifying those years with similar underlying weather patterns (those years having similar scores across multiple principal components). A K-means clustering procedure (Hartigan and Wong, 1979) was applied to the scores for the selected set of the most important principal components (those explaining the majority of the overall variation in the data) to group years together based on their weather patterns.

Multiple indices were considered to determine the optimum cluster number. The R package clusterCrit (Desgraupes, 2013) was used to investigate indices which could optimise cluster number. The within-cluster sum of squares and the C-Index were both used to select the optimum cluster number. The C-Index was used alongside the within-cluster-sum-of-squares value as an index to define optimum cluster size based on the range of within-cluster distances. The C-Index is defined by Desgraupes (2013) as

$$C = \frac{S_W - S_{min}}{S_{max} - S_{min}}$$

where, S_W is the sum of the distances between all pairs of points within each cluster; S_{min} is the sum of the smallest distances within each cluster; and S_{max} is the sum of the maximum distances within each cluster. The C-Index was calculated for all numbers of potential clusters from 2 to 50. The optimum cluster number was chosen from an elbow in both the plot of within-cluster-sum-of-squares and C-Index plot.

A linear mixed model (LMM) framework was used to analyse total biomass across weather clusters, the three experiments/crops (winter wheat in Broadbalk, spring barley in Hoosfield, and pasture in Park Grass), and five fertilizer treatments (PKNaMg, 48 kg N ha^{-1} + PKNaMg, 96 kg N ha^{-1} + PKNaMg, FYM, and Nil)

$$y = X\beta + U\gamma + \varepsilon$$

with y the response variable total biomass, X the fixed effects design matrix and U the random effects design matrix. The fixed model for cluster, experiment and fertilizer treatment represented the full factorial structure, and so terms (β) were fitted for the three main effects, three two-factor interaction effects and the three-way interaction effect. The random terms (γ) were the effect of plot (treatment within experiment) and the nested effect of year within plot. As years were clustered regarding their weather patterns, any additional variation associated with year within cluster was taken into account within the random model. The square root transformation of total biomass was analysed to satisfy the assumption of homogeneity of variance of the residuals across all treatment groups.

3. Results

3.1 Principal Components Analysis

The first 19 PCs explained 70% of the overall variation of the weather dataset. PC1 explained 10.00% of the variability in the weather dataset, with PC2 and PC3 explaining a further 6.53% and 5.80%, respectively. The loadings of total rainfall, mean daily maximum temperature, mean daily minimum temperature, total sunlight, rain intensity, minimum temperature, and days over 31°C for each month are shown in Figure 1 (a-f) for PC1 and PC2. The direction and magnitude (length) of the PCA loadings in Figure 1 illustrate how much of the variation of the weather dataset was explained by each variable. More information regarding the loadings for the first 30 PCs is given in Supplementary Table 1.

PC1 had negative loadings for both mean maximum and mean minimum temperature for every month (Figure 1b-c) and so highlighted the overall variation in average temperature across the cropping year, with generally warmer years having more negative scores for PC1. Since both mean maximum and mean minimum temperature loadings had high negative values in PC1, a large proportion of variation in the weather dataset from 1892 to 2016 was associated with changes in mean temperature. A contrast in seasonal temperatures (both mean maximum and mean minimum) was highlighted in PC2 (Fig. 1b-c), with positive loadings for April, May, June, July, August, September (weaker for mean minimum temperature than for mean maximum temperature), but negative loadings for December, January, February and March. Thus, years with higher than average winter temperatures and lower than average spring/summer temperatures will tend to have more negative scores, while those with lower than average winter temperatures and higher than average spring/summer temperatures will tend to have more positive scores. A similar seasonal contrast was observed in the loadings of minimum temperature and also the number of days over 31°C (Fig. 1f). The seasonal contrast for total rainfall and, to a lesser extent, rain intensity was observed for PC1 (Fig. 1a, e), with positive loadings for June, July and August, and negative loadings for October, January and February, reflecting differences between years with relatively wet winter and dry summers (more negative scores) and those with relatively dry winters and wet summers (more positive scores). Positive PC2 loadings for total rainfall in November and March possibly indicate that a warmer November tends also to be wetter, whilst a colder March tends also to be wetter. For PC1, most of the sunlight hours variables had negative loadings (October and May were the only months with positive loadings), with a weak summer-winter seasonal contrast in PC2 (positive loadings for July, August, September and December, negative loadings for October, November, February and March (Fig. 1d). The negative loading for November sunlight hours is an interesting contrast with the positive loadings for temperature and total rainfall (a warmer November associated with fewer sunlight hours), a similar contrast being seen for December (opposite signed loadings), reflecting contrasting weather patterns between November and December . These first two principal components therefore represent general patterns in average temperature and sunlight hours across the whole of the cropping season, and a range of winter-summer seasonal differences in temperature, rainfall and sunlight. Similar interpretations could be determined by plotting the loadings for other pairs of principal components (Supplementary Table 1), or by identifying those varables with larger absolute loadings for each principal component.

Communalities were used to assess the overall contribution of each weather variable to solutions including up to 30 dimensions (Supplementary Table 2), allowing identification of those variables contributing most to the selected 19-dimensional solution. Weather summaries which had communalities greater than 0.5 over the first 19 dimensions were from months in the late-spring and early-summer (April, May and June) and mid-late winter and early spring. Thus most of the weather variation captured in the first 19 dimensions (70% of variability) was explained by variation in weather in these periods. A further justification for the choice of the first 19 dimensions to describe the key sources of weather variability can be seen by considering the communalities for PC solutions with fewer or more dimensions (Supplementary Table 2). For example, for the 11 dimension solution (50% variation explained) not enough of the weather signal was captured, particularly within the temperature variables (too many communalities below 0.5) whilst for the 26 dimension solution (80% variation explained) too much noise was captured in the combination of variables (a lack of distinctness in the communalities, with values for almost all variables greater than 0.5 or even 0.6). Therefore, the 19 dimension solution (explaining 70% variation) was chosen as it captured enough of the long-term signal in the weather variables without being confounded by the inclusion of too much noise from the short-term variations in weather.

3.2 Selecting Cluster Number

There was no distinct elbow in the decline of the within-cluster sum of squares as cluster number increased (Fig. 2a). However, the rate of decline changed between a cluster number of 7 and 15 and there were also elbows in the decline of the C-Index at cluster numbers 5, 7 and 10, with a local minimum at 17 (Fig. 2b). Accordingly, a cluster number of 10 was selected, with a C-Index of 0.19. The temporal distribution of the membership of the more frequent of the 10 clusters is shown in Figure 3 and all 10 are listed in Table 1.

3.3 Cluster Summaries

Between 1900 and 1999, 25, 16 and 23 of the years were grouped into Clusters 2, 3 and 10, respectively (Fig. 3; Table 1), whereas since 2000 10 of the 16 years were within Cluster 1. Therefore, Clusters 2, 3 and 10 can be considered as characteristic climates of the 20th Century, representing 64 out of 100 years (64%). Cluster 1, on the other hand, can be considered as a characteristic climate of the early 21st Century, representing 10 out of the 16 years (63%). The membership of Clusters 7 and 9 span the 20th and 21st Century, but are more frequent towards recent periods, and so their respective memberships in the early to mid-20th century maybe considered anomalous years. Years within Cluster 8 span 70 central years of the 20th century, with cold January and February periods. The membership of all other clusters was small and these are therefore considered as outlier clusters or years.

Cluster 1 had the highest temperature, on average, across the whole crop production season (Fig. 4a-b); years within Cluster 1 also experienced very low levels of rainfall in June, but high rainfall from October until February. Mean maximum temperature for Cluster 2 between December and March and minimum temperature for February and March were all colder than for Clusters 3, 7, 9 and 10 (Fig. 4a-b). Years within Cluster 2 generally had a cold winter and early-spring. Years within Cluster 3 had low rainfall from April to August (Fig. 4c). Years within Cluster 7 had low hours of sunlight in February to April (Fig. 4d), with low rainfall in December, January, March, April and May (Fig. 4c). Years within Cluster 8 were generally cooler (Fig. 4a-b), with more between-month variation in sunlight duration (Fig. 4d). Cluster 9 provided warmer months in general and particularly in July to September (Fig. 4a-b), with low rainfall between May and August (Fig. 4c). Cluster 10 may be considered the typical 20th Century climate (Fig. 3). Its defining characteristics include low mean temperature in September and October (Fig. 4a-b), high rainfall in September to November (Fig. 4c), and low sunshine hours from April to September (Fig. 4d).

3.4 Effects on Crop Production

The Rothamsted LTE annual total biomass data from 1968-2016 were collated for the years in five of the ten clusters identified above. These were Clusters 1, 2, 7, 9 and 10 (12, 11, 6, 8, and 6 years' data, respectively) and were selected because each included a minimum of six years of results from 1968 onwards (Table 1). Hence, only six of the 49 years were omitted from the analysis.

The main effect of Cluster explained the largest amount of variation (F(4, 552, 37.74), P < 0.001; Supplementary Table 3), demonstrating a large impact of the five weather patterns on total biomass production across all three Experiments (i.e. crops). In contrast, the main effect of Experiment did not explain much model variation (F(2, 552, 0.85), P = 0.427) but that for Treatment (fertilizer regime) also explained large amounts of model variation (F(4, 552, 13.89), P < 0.001).

Differences in the total biomass of winter wheat, spring barley and pasture amongst years were exposed by a Cluster by Experiment interaction (F(8, 552, 8.09), P < 0.001), where large differences were detected (Supplementary Table 3), with the three crops responding differently to variation in annual weather patterns (Fig. 5). All three crops provided similar total biomass across the years within Cluster 7, but the cluster means of herbage off-take were less variable across the five weather clusters than cereal biomass yields (Fig. 5). The total biomass of winter wheat and spring barley grown at Rothamsted, UK, from years within Cluster 1 (years generally warmer and wetter, typical early 21st Century climate) were, on average, lower than those of herbage. On average, years within Cluster 2 (cooler winter and early-spring) provided the greatest winter wheat biomass, whilst years within Clusters 7 (warm early-summer and drier) and 10 (generally cooler, typical 20th Century climate) provided the highest spring barley and pasture off-take, on average, respectively.

4. Discussion

Climate change is often discussed principally in terms of the long-term rise of temperature (e.g. IPCC, 2019); and the temperature signal of climate change recorded globally long-term over land and ocean, or over land alone, is indeed unambiguous (IPCC, 2019). Nonetheless, weather and climate are much more than just temperature alone, whilst crop productivity is also dependent upon rainfall and solar radiation and the temporal patterns and extremes of their distribution during the growing season (and earlier during tillage and sowing or planting for non-perennial crops). The use of univariate methods to investigate the individual influence of temperature and rainfall on crop yield has limitations (Katz, 1977). Whilst multivariate cluster analysis has been applied to spatial weather data (Fovell and Fovell, 1993), we have shown here, that: the combination of Principal Components Analysis and K-means analysis can be applied to characterise the temporal patterns of key weather variables over crop production years into a limited number of distinct clusters (Figs 4-7); that the relative frequency of these clusters has changed markedly over the period 1892 to 2016 (Fig. 3), thereby providing clear evidence of a change in climate; and that these clusters explain differences in crop total biomass at this site (Fig. 5).

Although years do not naturally group into clusters depending on their weather, this multivariate approach of objectively detecting changes in climate has identified several years with similar weather conditions in the late-20th and early-21st centuries, conditions which differ markedly from those generally seen in earlier years. The methods of PCA and cluster analysis made comparisons between climates (clusters) more meaningful by considering the whole weather profile, rather than an average temperature being compared to a moving baseline climatology. PCA takes potentially correlated variables within a dataset and forms uncorrelated linear combinations of these variables. By constructing such uncorrelated linear combinations, the primary sources of between-year variability within the Rothamsted Meteorological Station data were identified.

For example, PC1, the linear combination of variables which explained the maximum proportion of the variability of the RMS dataset, separated out a temperature and sunlight effect, suggesting that the years from 1892 to 2016 may first be ordered from warmer to cooler. Further, the communalities for the 19 dimension solution (including the first 19 PCs) identified winter, early and late-spring, and summer as the most represented seasons across the different weather variable groups, responsible for explaining most of the between-year variability over the study period. A higher dimensional solution would have failed to identify these key seasons by introducing too much short-term weather variability (noise), whilst a lower dimensional solution would have failed to capture all of the important long-term weather variability (signal). The identification of these key periods was to be expected from previous reports of climate change globally (Hartmann et al., 2013; Kendon et al., 2017; Kirtman et al., 2013; Kovats et al., 2014; NOAA, 2017).

The membership of Cluster 1 (generally warmer temperatures than other clusters) featured 10 of the 16 years in the 21st Century analysed, two of the 1990s decade, and just a single year in the entire preceding period of 1892-1989. Therefore, Cluster 1 is the archetype of climate of the 21st Century to date. In comparison, in the early and mid-20th Century the Rothamsted climate varied mainly between cool to slightly warm with some cool and wet years (Clusters

2, 3 and 10), albeit with warmer years with some wet and dry periods (Clusters 4, 7, 8 and 9). Membership of clusters 3, 4, 8, and 10 have not occurred since 1984, 1964, 1990, 1986, and 1993, respectively, suggesting a low likelihood of these climates occurring in the 21st century. This change to a high frequency of warmer years is in accordance with predictions (IPCC, 2014).

Studies into crop variation from time-series yield data have long-shown associations between inter-annual variation in weather and that for the yield of winter wheat, barley and pastures (Addy et al., 2020; Chmielewski and Potts, 1995; Fisher, 1925; Hatfield and Dold, 2018; Hooker, 1907; Silvertown et al., 1994; Wishart and Mackenzie, 1930). The mean total biomass across all fertilizer treatments for both winter wheat and spring barley in years within Cluster 1 was lower than the total biomass for years in Clusters 2, 7, 9 and 10 (Fig. 5). Generally, Cluster 1 was warmer in April, May, June, July and August and experienced a drier June compared to clusters 2, 3 and 10. Previous analysis from studies of crop yield variation at Rothamsted showed that on average rainfall during short periods in the summer was beneficial to barley yields (Wishart and Mackenzie, 1930). Years which were warm and dry (Clusters 7 and 9) had lower winter wheat and spring barley total biomass compared to years which were generally cooler (Clusters 2 and 10). Total biomass from years which were warmer (Clusters 1, 7 and 9) were generally expected to be lower, even when years from Cluster 1 were generally drier than those from 7 and 9. Where warmer temperatures increase grain growth rate, the grain-filling duration is reduced and the overall effect is lower yield because the former is more than offset by the latter (Sofield et al., 1977). Moreover, if the crops experienced periods of heat stress around anthesis this will further reduce grain yield due to poor seed set (Barber et al., 2017; Ferris et al., 1998; Gooding et al., 2003; Nasehzadeh and Ellis, 2017; Wheeler et al., 1996). Hence the warmer temperatures at these key times, most common in years in the the early 21st Century (Cluster 1), were the likely cause of the lower winter wheat and spring barley off-take.

On average, drier years (Clusters 7 and 9) provided lower herbage off-take in Park Grass than warmer temperatures and wetter years (Cluster 1). This tallies with earlier studies of herbage yields from Park Grass which showed a positive relationship between rainfall and biomass (Silvertown et al., 1994) and that yields were associated negatively with high temperature in July and August but positively with higher rainfall (Sparks and Potts, 2003). The comparison amongst clusters for herbage and cereal biomass (Fig. 5) shows that herbage biomass have been the more stable to date with changes in climate over this 49-year period. This suggests that the positive effects of higher rainfall offset the negative effects of warmer temperatures on pasture performance. If rainfall levels are maintained, this indicates that pastures may be more resilient to future changes in climate this century than either wheat or barley.

The main effect of fertilizer treatment on biomass did not differ amongst clusters, suggesting the effect of fertilizer is consistent across clusters for these selected treatments (Supplementary Table 3). Temporal variation in yields from other long-term experiments have been shown to be influenced by fertilizers, however (Macholdt et al., 2021). Indeed, the quantitative response of winter wheat and spring barley grain yield to nitrogen fertilizer in the Rothamsted LTEs was shown to be particularly sensitive to temperature and rainfall in certain months of the year (Addy et al., 2020). We suggest that the current study with a reduced range of fertilizer treatments was not sufficiently sensitive to detect an interaction. The total biomass of short-strawed and long-strawed wheat cultivars were the same on low-fertilised plots in the Broadbalk Experiment (Austin and Ford, 1989), but problems including take-all disease (Etheridge, 1969) and managing the weed blackgrass (Fisher, 1921) on the Rothamsted LTEs are examples of other factors influencing inter-annual yield variation.

Clusters 2 and 10 (with cooler spring temperatures and memberships spanning the 20th Century) occurred with low frequency in the 21st Century (one and zero years, respectively) compared to any other period since 1892. If the observed climate trends continue, more years such as in Clusters 1, 7 and 9 (generally warmer and lower yielding) would be expected throughout the 21st Century with fewer years such as in Clusters 2 and 10 (generally cooler and higher yielding). The individual influences of temperature and rainfall variables on yield have been investigated previously using univariate analysis methods, but these approaches have been shown to have limitations for predicting future responses (Katz, 1977). However, this study has explored the multivariate effect of climate on several crop production systems through the identification of contrasting climates, albeit only at one site in the UK. This study demonstrates how the application of multivariate cluster analysis, not just to spatial weather data (Fovell and Fovell, 1993) but also to temporal weather datasets, can be used to objectively assess how climate has changed over multiple correlated variables and the consequent impact on the productivity of several key arable crops.

5. Conclusions

The major conclusions of this study are:

(1) Multivariate cluster analysis can be applied to temporal weather datasets to objectively identify long-term changes in weather patterns across multiple variables, including all the main meteorological inputs relevant to crop production (i.e. the whole weather profile not just temperature alone).

(2) Ten distinct weather clusters were identified,

(3) A weather cluster, comprising higher average temperature with more intense rainfall but a dry June, has dominated the most recent 25-year period.

(4) Five weather clusters which occurred frequently during much of the 20th Century have not recurred in recent times.

(5) The above provides unambiguous evidence of climate change in several more dimensions than mean temperature alone.

(6) Differences in the total biomass of crops of winter wheat, spring barley, and herbage was associated strongly with differences between the identified weather clusters.

Acknowledgements

We thank Margaret Glendining, and Sarah Perryman and Tony Scott for access to yield and weather data from the Electronic Rothamsted Archive (e-RA), and the Lawes Agricultural Trust for supporting the work of JWGA as part of a PhD Scholarship. The Rothamsted Long-Term Experiments National Capability (LTE-NCG) is supported by the UK Biotechnology and Biological Science Research Council (BBS/E/C/000J0300) and the Lawes Agricultural Trust.

References

Addy, J.W.G., Ellis, R.H., Macdonald, A.J., Semenov, M.A., 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. Agric. For. Meteorol. 284, 107898. https://doi.org/10.1016/j.agrformet.2019.107898

Addy, J.W.G., Ellis, R.H., Macdonald, A.J., Semenov, M.A., Mead, A. 2021 The impact of weather and increased atmospheric CO2 from 1892 to 2016 on simulated yields of UK wheat. J. Roy, Soc. Interface, 18, 20210250. https://doi.org/10.1098/rsif.2021.0250

Austin, R.B., Ford, M.A., 1989. Effects of nitrogen fertilizer on the performance of old and new varieties of winter wheat. Vor. fuer Pflanzenzuechtung 307–315.

Avery, B.W. and Catt, J.A. (1995). The Soil at Rothamsted. Lawes Agricultural Trust Co. Ltd, Harpenden UK.

Barber, H.M., Lukac, M., Simmonds, J., Semenov, M.A., Gooding, M.J., 2017. Temporally and Genetically Discrete Periods of Wheat Sensitivity to High Temperature. Front. Plant Sci. 8, 51. https://doi.org/10.3389/fpls.2017.00051

Blake, L., Johnston, A.E., Poulton, P.R., Goulding, K.W.T., 2003. Changes in soil phosphorus fractions following positive and negative phosphorus balances for long periods. Plant and Soil. 254(2), pp.245-261.

Camargo, S.J., Robertson, A.W., Gaffney, S.J., Smyth, P., Ghil, M., 2007. Cluster Analysis of Typhoon Tracks. Part I: General Properties. J. Clim. 20, 3635–3653. https://doi.org/10.1175/JCLI4188.1

Chmielewski, F.-M., Potts, J.M., 1995. The relationship between crop yields from an experiment in southern England and long-term climate variations. Agric. For. Meteorol. 73, 43–66. https://doi.org/10.1016/0168-1923(94)02174-I

Desgraupes, B., 2013. Vignettes: An R Package for Computing Clustering Quality Indices. clusterCrit: Clustering Indices.

e-RA Rothamsted, 2021. Rothamsted Research. Accessed 4 July 2021, <http://www.era.rothamsted.ac.uk/index.php>

Etheridge, J., 1969. Take-all Broadbalk wheat, 1958-1967. Rothamsted Exp. Stn. Report.

Ferris, R., Ellis, R.H., Wheeler, T.R., Hadley, P., 1998. Effect of High Temperature Stress at Anthesis on Grain Yield and Biomass of Field-grown Crops of Wheat. Ann. Bot. 82, 631–639.

Fisher, R.A., 1925. The Influence of Rainfall on the Yield of Wheat at Rothamsted. Philos. Trans. R. Soc. B Biol. Sci. 213, 89–142. https://doi.org/10.1098/rstb.1925.0003

Fisher, R.A., 1921. Studies in Crop Variation. I. An examination of the yield of dressed grain from Broadbalk. J. Agric. Sci. 11, 107–135.

Fovell, R.G., Fovell, M.-Y.C., 1993. Climate Zones of the Conterminous United States Defined Using Cluster Analysis. J. Clim. 6, 2103–2135.

Gooding, M.J., Ellis, R.H., Shewry, P.R., Schofield, J.D., 2003. Effects of restricted water availability and increased temperature on the grain filling, drying and quality of winter wheat. J. Cereal Sci. 37, 295–309. https://doi.org/10.1006/jcrs.2002.0501

Gregory, A.S., Watts, C.W., Griffiths, B.S., Hallett, P.D., Kuan, H.L., Whitmore, A.P., 2009. The effect of long-term soil management on the physical and biological resilience of a range of arable and grassland soils in England. Geoderma. 153(1-2), pp.172-185.

Hartigan, J.A., Wong, M.A., 1979. Algorithm AS 136: A K-Means Clustering Algorithm. Appl. Stat. 28, 100. https://doi.org/10.2307/2346830

Hartmann, D.L., Klein Tank, A.M.G., Rusticucci, M., Alexander, L. V., Brönnimann, S., Charabi, Y., Dentener, F.J., Dlugokencky, E.J., Easterling, D.R., Kaplan, A., Soden, B.J., Thorne, P.W., Wild, M., Zhai, P.M., 2013. 2013: Observations: Atmosphere and Surface, in: Stocker, T.F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S.K., Boschung, J., Nauels, A., Xia, Y., Bex, V., Midgley, P.M. (Eds.), Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Hatfield, J.L., Dold, C., 2018. Agroclimatology and Wheat Production: Coping with Climate Change. Front. Plant Sci. 9, 224. https://doi.org/10.3389/fpls.2018.00224

Hooker, R.H., 1907. Correlation of the weather and crops. J. R. Stat. Soc. 70, 1–51.

IPCC, 2019. IPCC Special Report on Climate Change, Desertification, Land Degradation, Sustainable Land Management, Food Security, and Greenhouse gas fluxes in Terrestrial Ecosystems. Summary for Policymakers. Approved Draft. (2019).

IPCC, 2014. Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Core Writing Team, R.K. Pachauri and L.A. Meyer. https://doi.org/10.1017/CBO9781107415324.004

Katz, R.W., 1977. Assessing the impact of climatic change on food production. Clim. Change 1, 85–96.

Keatinge, J.D.H., Ledesma, D., Keatinge, F., & Hughes, J., 2014. Projecting annual air temperature changes to 2025 and beyond: Implications for vegetable production worldwide. J. Agric. Sci. 152, 38-57. doi:10.1017/S0021859612000913 Kendon, M., McCarthy, M., Jevrejeva, S., Legg, T., 2017. State of the UK Climate 2016.

Kendon, M., McCarthy, M., Jevrejeva, S., Matthews, A., Legg, T., 2019. State of the UK climate 2018, Int J Climatol. https://doi.org/10.1002/joc.6213

Kirtman, B., Power, S.B., Adedoyin, J.A., Boer, G.J., Bojariu, R., Camilloni, R.I., Doblas-Reyes, F.J., Fiore, A.M., Kimoto, M., Meehl, G.A., Prather, M., Sarr, A., Schär, C., Sutton, R., van Oldenborgh, G.J., Vecchi, G., Wang, H.J., 2013. Near-term Climate Change: Projections and Predictability. In:, in: Stocker, T.F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S.K., Boschung, J., Nauels, A., Xia, Y., Bex, V., Midgley, P.M. (Eds.), Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Kovats, R.S., Valentini, R., Bouwer, L.M., Georgopoulou, E., Jacob, D., Martin, E., Rounsevell, M., Soussana, J.-F., 2014. 2014: Europe, in: Barros, V.R., Field, C.B., Dokken, D.J., Mastrandrea, M.D., Mach, K.J., Bilir, T.E., Chatterjee, M., Ebi, K.L., Estrada, Y.O., Genova, R.C., Girma, B., Kissel, E.S., Levy, A.N., MacCracken, S., Mastrandrea, P.R., White, L.L. (Eds.), Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part B: Regional Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge

University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 1267–1326.

Macdonald, A.J., Poulton, P.R., Clark, I.M., Scott, T., Glendining, M.J., Perryman, S.M., Storkey, J., Bell, J.R., McMillan, V., Hawkings, J., 2018. Guide to the Classical and other Long-Term experiments, Datasets and Sample Archive. Rothamsted Research.

Macholdt, J., Hadasch, S., Piepho, H.P., Reckling, M., Taghizadeh-Toosi, A., 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 Res. 264, p.108083.

Nasehzadeh, M., Ellis, R.H., 2017. Wheat seed weight and quality differ temporally in sensitivity to warm or cool conditions during seed development and maturation. Ann. Bot. 120, 479–493. https://doi.org/10.1093/aob/mcx074

NOAA, 2017. Global Climate Report - Annual 2016. [WWW Document]. URL https://www.ncdc.noaa.gov/sotc/global/201613 (accessed 9.20.17).

Porter, J.R., Semenov, M.A., 2005. Crop responses to climatic variation. Philos. Trans. R. Soc. B Biol. Sci. 360, 2021–2035. https://doi.org/10.1098/rstb.2005.1752

Roberts, E.H., Qi, A., Ellis, R.H., Summerfield, R.J., Lawn, R.J., Shanmugasundaram, S., 1996. Use of field observations to characterise genotypic flowering responses to photoperiod and temperature: A soyabean exemplar. Theor. Appl. Genet. 93, 519–533. https://doi.org/10.1007/BF00417943

Silvertown, J., Dodd, M.E., McConway, K., Potts, J., Crawley, M., 1994. Rainfall, Biomass Variation, and Community Composition in the Park Grass Experiment. Ecology 75, 2430. https://doi.org/10.2307/1940896

Sofield, I., Evans, L.T., Cook, M.G., Wardlaw, I.F., 1977. Factors influencing the rate and duration of grain filling in wheat. Aust. J. Plant Physiol. 4, 785–97.

Sparks, T.H., Potts, J.M., 2003. Late summer grass production. In Indicators of climate change in the UK [WWW Document]. URL http://www.ecn.ac.uk/ICCUK/ (accessed 5.15.18).

Trnka, M., Rötter, R.P., Ruiz-Ramos, M., Kersebaum, K.C., Olesen, J.E., Žalud, Z., Semenov, M.A., 2014. Adverse weather conditions for European wheat production will become more frequent with climate change. Nat. Clim. Chang. 4, 637–643. https://doi.org/10.1038/nclimate2242

Webber, H., Lischeid, G., Sommer, M., Finger, R., Nendel, C., Gaiser, T., Ewert, F., 2020. No perfect storm for crop yield failure in Germany. Environ. 15(10), p.104012.

Wheeler, T.R., Batts, G.R., Ellis, R.H., Hadley, P., Morison, J.I.L., 1996. Growth and yield of winter wheat (Triticum aestivum) crops in response to CO2 and temperature. J. Agric. Sci. 127, 37. https://doi.org/10.1017/S0021859600077352

Wishart, J., Mackenzie, W.A., 1930. Studies in Crop Variation VII. The Influence of Rainfall on the Yield of Barley at Rothamsted. J. Agric. Sci. 20, 417–439.

Figures

Figure 1: A representation of the loadings for the first two principal components (PC1, PC2) from the Principal Components Analysis of summarised monthly weather variables at Rothamsted (1892 to 2016). For (a) total rainfall, (b) mean daily maximum temperature, (c) mean daily minimum temperature, (d) total sunlight, (e) rain intensity, and (f) minimum temperature (black) for each month of the year and days over 31°C in the summer months June, July, August and September (grey); PC1 and PC2 explained 10.00% and 6.53%, respectively, of weather dataset variability; the solid ellipse represents a magnitude of 0.25 in all directions from the origin, identifying those variables with a communality greater than 0.25 in the 2 dimension PCA solution.

Figure 2: Scree plots of the within-cluster sums of squares (a) and C-Index (Desgraupes, 2013) (b) as cluster number varied from 0 to 50. The vertical line at cluster number 10 is discussed in the text: the C-Index value is 0.19 this point.

Figure 3: The smoothed relative frequency of occurrence (smoothed histograms) of years in Clusters 1, 2, 3, 7, 8, 9 and 10 from 1892-2016 at Rothamsted. The vertical lines identify the individual years in each cluster. The three weather clusters omitted (4, 5, 6) occurred only infrequently (Table 1).

Figure 4: Monthly summaries of mean daily maximum (a), mean daily minimum (b), mean rainfall (c), and Mean total sunlight duration (d) for weather clusters 1 (black solid line), 2 (grey dashed line), 3 (grey dotted line), 7 (grey dot-dashed line), 8 (black dashed line), 9 (black dotted line), and 10 (black dot-dashed line) at Rothamsted (Table 1) identified by multivariate analysis.

Figure 5: Mean annual biomass yields (t ha−1 at 85% dry matter, transformed to square root left-hand axis, since analysis was conducted after this transformation; untransformed

scale shown on the right-hand axis), averaged over all fertilizer treatments, of winter wheat (■), spring barley (●), and pasture crops (two cuts per year) (▲) for those years over the period 1968-2016 within each of five different weather clusters at Rothamsted (Table 1) identified by multivariate analysis. The vertical line is the standard error of the difference (SED). The five weather clusters omitted from analysis occurred only once (5, 6), twice (3, 8), or not at all (4) during the period 1968-2016.

Tables

Table 1: The weather cluster membership of years between 1892 and 2016 for the selected 10-cluster solution

Supplementary Tables

Supplementary Table 1: The loadings for each of the 76 weather variables for the first 30 principal components from the Principal Components Analysis.

Supplementary Table 2: The communalities for each of the 76 weather variables for PCA solutions including up to 30 dimensions.

Supplementary Table 3: F-statistics with estimated denominator degrees of freedom (DDF) and observed p-values for the cluster (weather) by experiment (crop) by treatment (fertilizer) analysis.

Table 1 & Figures follow

Table 1

Cluster	Number of Years (1968 to 2016)	Year	Defining Characteristics
1	13 (12)	1943, 1994, 1999, 2000, 2001, 2002, 2004, 2005, 2006, 2007, 2008, 2014, 2015	High temperatures, drier June
2	27 (11)	1892, 1900, 1901, 1904, 1907, 1908, 1909, 1915, 1928, 1929, 1941, 1946, 1954, 1955, 1956, 1962, 1969, 1971, 1977, 1978, 1979, 1980, 1985, 1987, 1991, 1996, 2013	Cold Winter and early-Spring
3	18 (2)	1896, 1897, 1902, 1903, 1905, 1912, 1913, 1916, 1920, 1921, 1923, 1935, 1938, 1948, 1957, 1960, 1974, 1984	Cold August to September
4	6 (0)	1895, 1940, 1942, 1947, 1963, 1964	Cold Winter and years with a general drought
5	4 (1)	1898, 1906, 1911, 1990	Cold March to June and a warm July to September. General absence of rainfall and more hours of direct sunlight
6	1 (1)	1976	Drought Summer with warm temperatures
7	11 (6)	1893, 1914, 1934, 1945, 1952, 1982, 1992, 1997, 2009, 2010, 2011	Warm March to June, dry Spring
8	9 (2)	1917, 1919, 1922, 1924, 1931, 1965, 1966, 1970, 1986	Cold January and February, similar to Cluster 2, but years were generally warmer
9	12 (8)	1899, 1933, 1949, 1959, 1975, 1983, 1989, 1995, 1998, 2003, 2012, 2016	Warm July to September, dry July and August
10	24 (6)	1894, 1910, 1918, 1925, 1926, 1927, 1930, 1932, 1936, 1937, 1939, 1944, 1950, 1951, 1953, 1958, 1961, 1967, 1968, 1972, 1973, 1981, 1988, 1993	Cool and dry March.

Figure 1

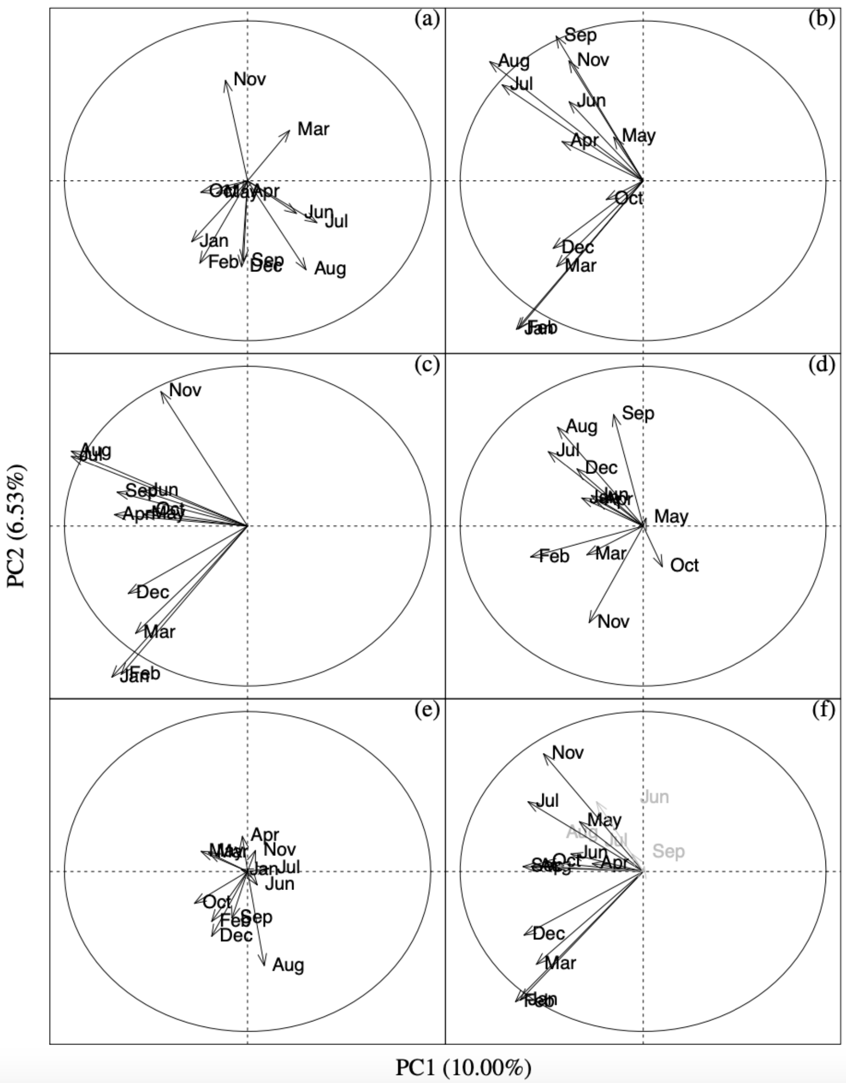


Figure 2

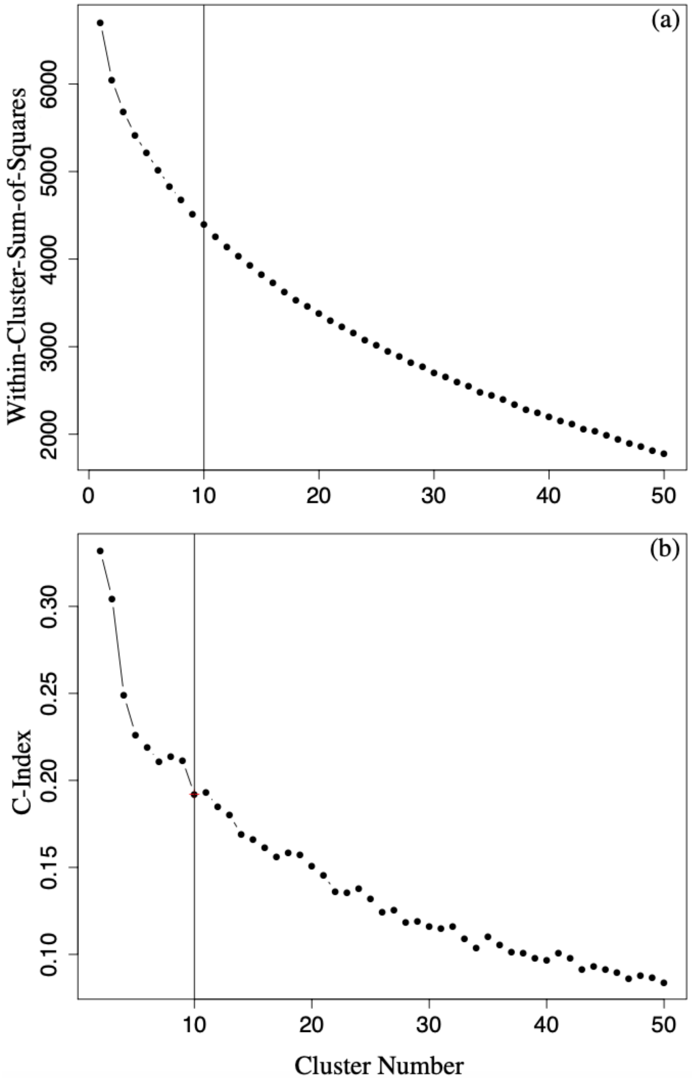
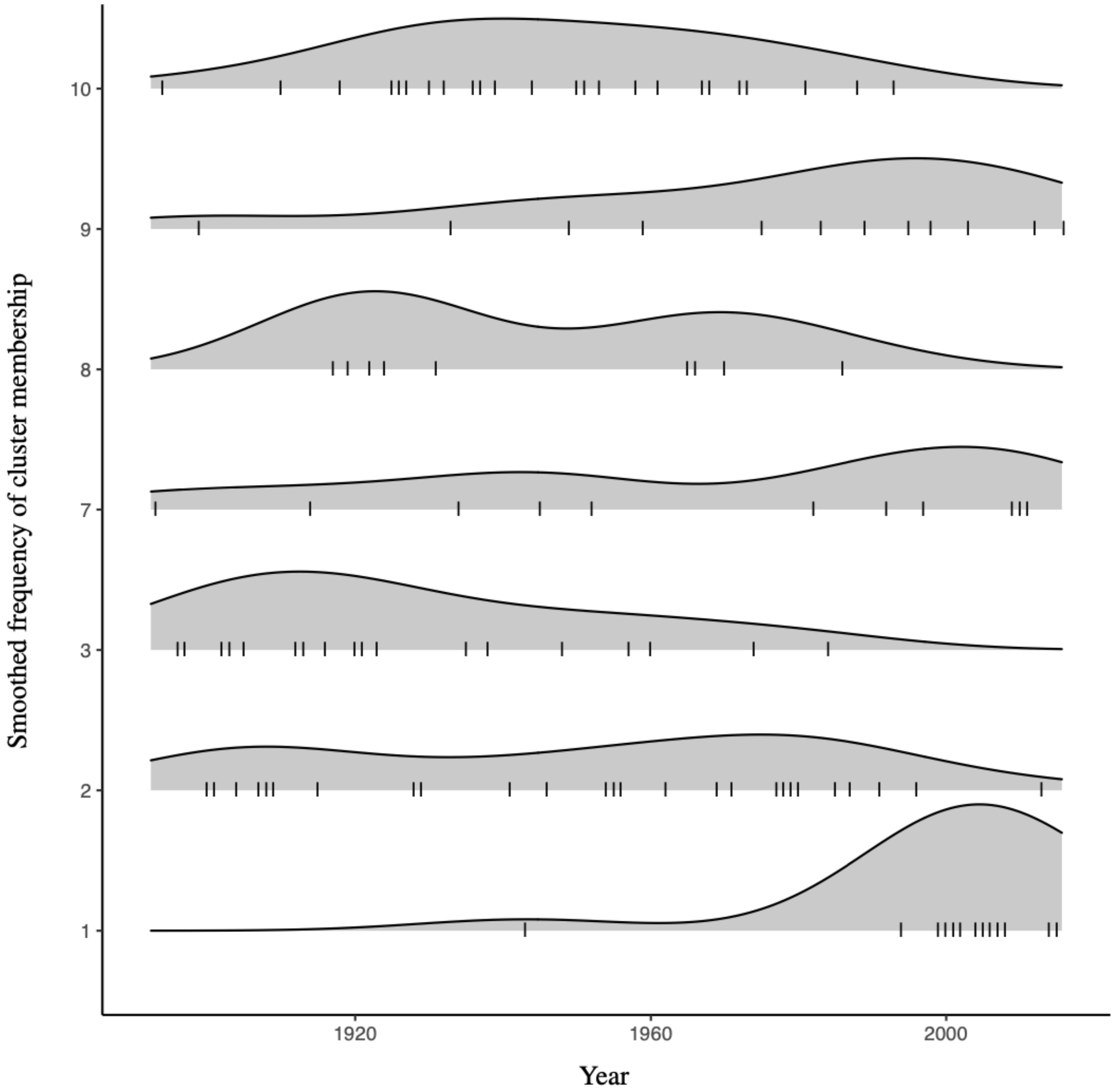

(a)
Within-Cluster-Sum-of-Squares
2000
3000
4000
5000
6000
0
10
20
30
40
50
(b)
C-Index
0.10
0.15
0.20
0.25
0.30
10
20
30
40
50
Cluster Number

Figure 3



Smoothed frequency of cluster membership
10
9
8
7
3
2
1
1920
1960
2000
Year

Figure 4

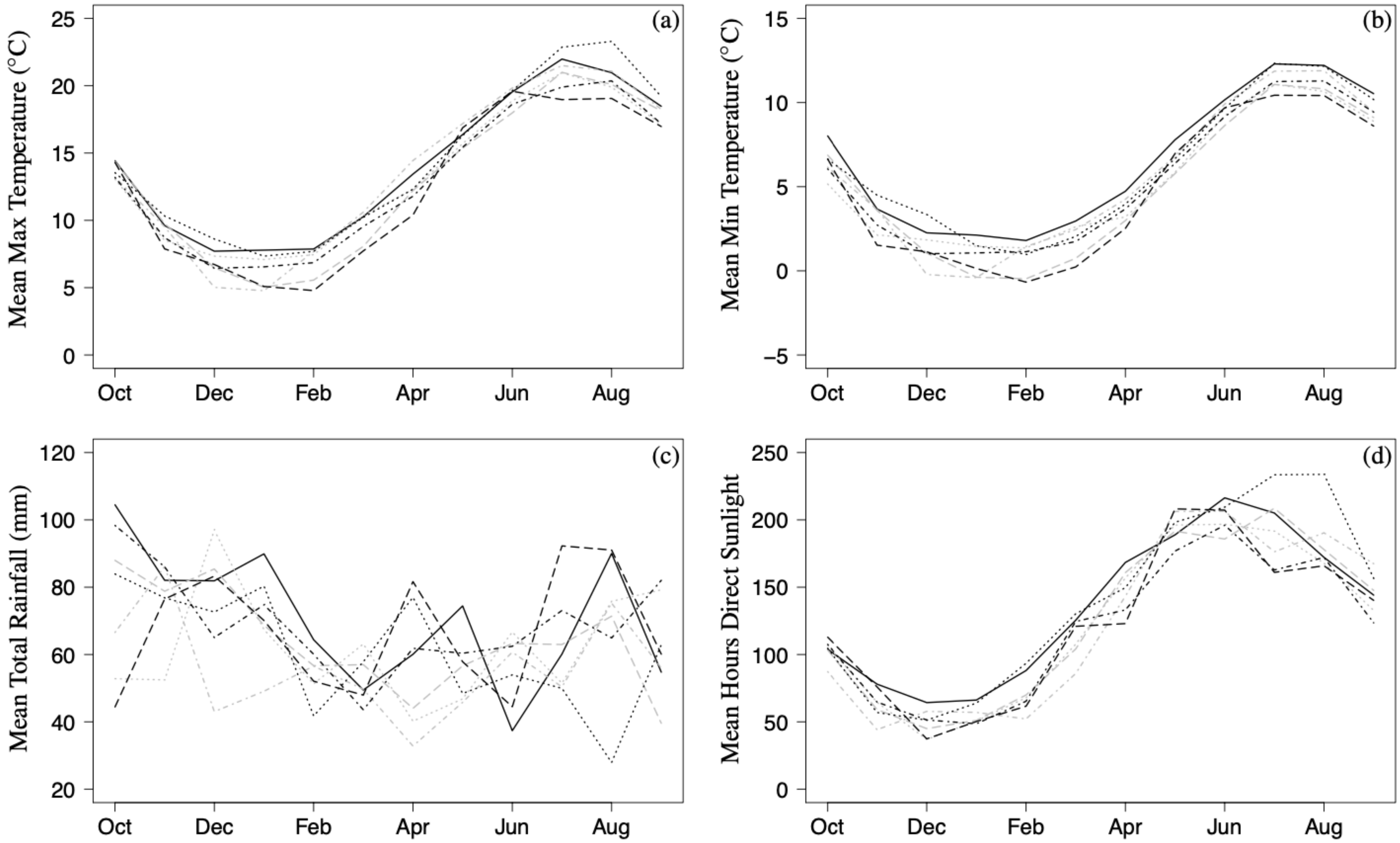
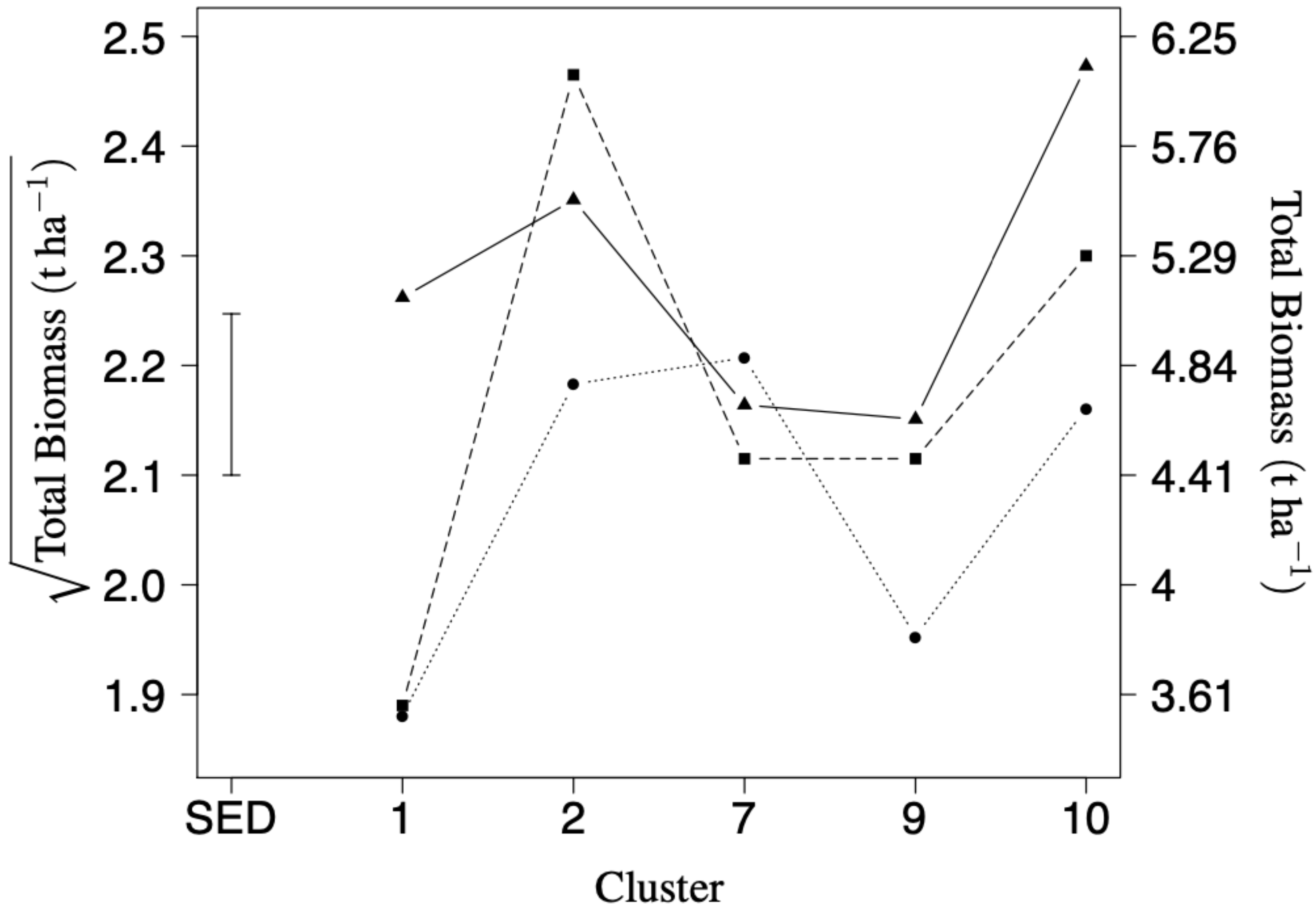

(a)
Mean Max Temperature (°C)
25
20
15
10
5
0
Oct
Dec
Feb
Apr
Jun
Aug
(b)
Mean Min Temperature (°C)
15
10
5
0
−5
Oct
Dec
Feb
Apr
Jun
Aug
(c)
Mean Total Rainfall (mm)
120
100
80
60
40
20
Oct
Dec
Feb
Apr
Jun
Aug
(d)
Mean Hours Direct Sunlight
250
200
150
100
50
0
Oct
Dec
Feb
Apr
Jun
Aug

Figure 5



2.5
2.4
2.3
2.2
2.1
2.0
1.9
√Total Biomass (t ha⁻¹)
6.25
5.76
5.29
4.84
4.41
4
3.61
Total Biomass (t ha⁻¹)
SED
1
2
7
9
10
Cluster