Future droughts in Great Britain: from weather to Wensum

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

Nele Reyniers

April 2023

Supervisors:
Prof. Timothy Osborn, Dr. Nans Addor, Dr. Geoff Darch, Dr. Helen He

Registration number: 100295261

Abstract

Droughts can cause enormous damages and are changing under anthropogenic climate change. In this thesis, climate and hydrological models were used to investigate projected future droughts in Great Britain, emphasising the contributions of methodological and modelling choices to the resulting projections and their uncertainty. First, non-negligible biases were found in the regional climate model projections of UKCP18, the latest set of regional climate projections for the UK (UKCP18-RCM), and therefore, two bias adjustment methods were applied successfully. The ensemble projects wetter, warmer winters and hotter, drier summers, but with more complex changes for different temperature and precipitation indices, including different changes in the extreme ends of their ranges. These changes are well-preserved after bias-adjustment, but the change factor approach failed to capture changes in some precipitation metrics. From these data, projected changes in drought characteristics were derived using two contrasting drought indicators, the Standardised Precipitation Index (SPI) and the Standardised Precipitation Evapotranspiration Index (SPEI). Projected increases in drought frequency and extent, and the dominant contribution of future summer droughts to annual-scale deficits, are far greater based on the SPEI than based on the SPI, highlighting the important role of increased evaporative demand for future droughts. The largest changes were found for extreme droughts. Finally, an ensemble of hydroclimatic impact model chains was calibrated for the Wensum catchment. Detailed evaluation using multiple criteria strongly reduced the uncertainty in the streamflow drought frequency projections, especially the contribution from hydrological modelling choices. Increasing streamflow drought frequency was projected for the Wensum, with greater increases for more severe droughts and higher levels of global warming. This resulted from both a lengthening of individual drought events (especially for moderate droughts) and an increase in the number of events (primarily for more severe droughts). This work concludes with recommendations for future hydrological climate change impact studies.

Contents

List of Tables

List of Figures

Acknowledgements

First I would like to express my genuine thanks to my supervisors Tim Osborn, Nans Addor, Geoff Darch and Helen He, for all their support and guidance. Many thanks to Tim Osborn for his endless patience during the rather turbulent last years and his encouragement to widen my scientific horizons, to Nans Addor for never failing to come up with interesting food for thought, and to Geoff Darch for grounding my research in a high-stakes reality.

I would also like to thank Anglian Water and the University of East Anglia for providing the funding for this project, including a three-month extension related to the COVID-19 pandemic, and the Anglian Centre for Water Studies for making PhD projects like this possible. I would also like to acknowledge the people at Anglian Water and consultants at Mott MacDonald who provided me with some of the data and models used in this work. Finally, I would like to thank the UEA HPC team for their helpful answers to my ADA-questions.

An enormous thanks also goes out to my amazing fellow PhD students and friends in Norwich, who I am incredibly grateful to have gotten to know, and who I already miss dearly. Words cannot adequately express how grateful I am to my family and old friends in Belgium, for all the many different wonderful kinds of support they have shown throughout the past years and for bringing light and warmth into my life. And of course, to Arne, for being by my side, for believing in me, for tolerating my jokes, and simply for being him.

Finally, I would like to thank the UEA Symphony Orchestra and also Continuo – music has been a necessary relief.

For the 20's kids.

1

Introduction

This introduction provides the overarching motivation for this thesis, as well as key literature that gives common background for the upcoming chapters. Further literature, especially that which is relevant to a specific chapter, is reviewed in the discussion and introduction sections of the individual chapters.

1.1 Water management and its challenges in the UK

Climate change is an important threat to future water supply in the UK (Watts et al., 2015), however it needs to be seen in the wider context of multi-faceted pressures facing the UK water industry in upcoming decades. To improve the reliability and sustainability of water resources systems, water companies in England and Wales are required to submit a Water Resources Management Plan (WRMP) every five years, as per the Water Industry Act of 1991 (Environment Agency et al., 2022). In the WRMP, water companies forecast supply and demand for their water resources, and develop supply- and demand-side options to meet any potential forecast deficits. Water resources management plans of several UK water companies show that, without additional measures, future water supply may not be able to satisfy demand within the next few decades, a phenomenon known as the 'jaws of death' (Bevan, 2019). This is also the case for Anglian Water, the industrial partner of this doctoral project. As the water company with the largest area of coverage in England, Anglian Water is responsible for the water supply and water recycling for over 6 million customers. On average, abstraction for storage and treatment amounts to around 1100 Ml of water daily, which can climb up to about 1400 Ml/day during periods of high demand such as summer 2018. About half of their water supply comes from groundwater abstraction, with eight water reservoirs and eight direct supply river intakes making up the other half (Anglian Water, 2019).

The projected supply-demand imbalance is evidently a major concern to UK water companies, managers and consumers, and is attributed to four main categories of cause in their recent water resources management plans (the consultation phase is still ongoing for the 2024 water resources management plans, however, drafts are available on water companies' websites):

1. Demand increase. This is not limited to the effects of population growth (e.g. Thames Water, 2022; Anglian Water, 2022a; Severn Trent Water, 2022), but can also be influenced by e.g. new development plans and cultural shifts such as changed work-life habits after the COVID-19 lockdowns (e.g. Anglian Water, 2022a; Severn Trent Water, 2022).

2. Environmental and sustainability reductions in abstraction. Water companies are required to plan for improving, protecting, and especially not causing deterioration of, the environment (Environment Agency et al., 2022). For WRMP24, water companies are asked set out how to achieve sustainable abstractions by 2050, taking into account the effects of future demand and climate change (Environment Agency et al., 2022). Abstraction reductions required for avoiding or reversing damage to the environment arising from the exploitation of water resources pose a significant challenge for some water companies (e.g. Severn Trent Water, 2022), and makes up the largest portion of the projected baseline supply-demand imbalance in the draft WRMP24 of Anglian Water (e.g. Anglian Water, 2022a).

3. Climate change. Both historical and projected future impacts of climate change need to be taken into account, with emphasis on the treatment of the associated uncertainty (e.g. Anglian Water, 2022a). The uncertainty surrounding the future impact of climate change on water resources is multifold. First and foremost, there is the inherent uncertainty related to the trajectories of anthropogenic greenhouse gas emissions in the coming years and decades on which future climatic changes (and the consequences on water resources and drought impacts) depend. However, even for a fixed emissions scenario, there is uncertainty associated with the response of the global climate system and indeed local water resources impacts, due multiple factors including incomplete process understanding and limitations of climate and impact models and analysis methods (including those relied on by water companies for projected climate change impacts).

4. Drought resilience. Since WRMP 2019, the drought resilience aim set by the Environment Agency has changed from a 1/200 year drought (yearly probability of occurrence of 0.005) to a 1/500 year drought (yearly

probability of 0.002) (Environment Agency et al., 2022). In this context, drought resilience means that for droughts of these return periods, there should be no need for 'exceptional demand restrictions on customers' (for example standpipes) associated with a drought order (Environment Agency et al., 2022).

The uncertainties in the expected impacts of these factors are taken into account as *headroom*, which is added to the expected impacts taken into account (Environment Agency et al., 2022). Due to the inherently uncertain nature of important factors influencing future water supply (e.g. climate change) and demand (e.g. population growth), the WRMP guidelines recommend adaptive planning (Environment Agency et al., 2022). To complement the WRMP in tackling the challenge of drought resilience, water companies also have to submit a Drought Plan, which details the more short-term operational actions that the companies plans to take in the case of a drought in order to maintain water supply (Environment Agency, 2021). This covers a range of measures on the demand and supply side, from customer communications to (temporary) supply increases, and crucially needs to aim at minimising environmental drought impacts. As part of the drought plan, water companies also determine a set of drought triggers based on precipitation, soil moisture, river flow and/or groundwater metrics. These are used to define phases of drought development which are linked to specific sets of responses and actions in the drought response framework (e.g. Table 2.1 in Anglian Water, 2022b). Better understanding of the influence of climate change on the nature and characteristics of future droughts, at spatial and temporal scales relevant to water companies, and understanding and constraining the associated uncertainties, is invaluable to help support water companies in planning for future droughts.

1.2 What is drought? Definition, classification and quantification

No single satisfactory universally applicable definition for drought exists (Wilhite and Glantz, 1985; Mishra and Singh, 2010; Lloyd-Hughes, 2014; Ault, 2020). In the most general sense, a drought can be defined as a deficit of water compared to normal conditions. Depending on the context and the hydrological flux or reservoir in which this deficit takes place, the following conventional drought types can be defined (Mishra and Singh, 2010; Dai, 2011). A *meteorological drought* is a deficit of precipitation relative to values normally expected based on the local climatology. A *soil moisture drought*, defined by a

deficit in soil moisture, is also often called agricultural drought because of the importance of root zone soil moisture for crop growth. In Seneviratne et al. (2021), the consideration of agricultural and ecological droughts focuses on crop yield reductions and plant water stress, which are strongly linked to soil moisture but also allow for the effects of other meteorological and biophysical factors. Soil moisture droughts can then propagate further to *hydrological droughts*, which encompass anomalously low levels of streamflow, lake storage or groundwater levels (*groundwater drought*). A *socio-economic drought* occurs when one or more of these drought types leads to societal impacts (Mishra and Singh, 2010; Van Loon and Laaha, 2015).

Drought is distinct from *aridity* and *water scarcity* and should not be confused with these terms (e.g. Van Loon and Laaha, 2015). Aridity reflects the baseline climatic moisture availability of a region under normal circumstances (Mishra and Singh, 2010). Droughts can therefore occur in climates with different aridity levels, as they are temporary deviations relative to normal moisture availability. Water scarcity is also defined as a long-term condition and refers explicitly to imbalances between water demand and supply, i.e. it can also indicate over-exploitation of available resources (Loon and Van Lanen, 2013).

Many drought indicators (DI) have been developed over time to quantify these different types of droughts. A wide range of DIs can be found in the literature (e.g. Dracup et al., 1980; Keyantash and Dracup, 2002; Pedro-Monzonís et al., 2015) describing droughts of different types, from purely precipitation-based (e.g. McKee et al., 1993) to indicators considering the effects of human influences (Marcuello and Lallana, 2003). Keyantash and Dracup (2002) outline the following criteria for selecting a suitable drought indicator: robustness (is the DI useful in a range of conditions?), tractability (how easy is it to calculate the DI and acquire the data to do so?) and transparency (how understandable is the DI?), sophistication (which can present a trade-off with transparency), extendability (in time, depending on data availability) and dimensionality (units of the index values). Drought indicators are further discussed in the introduction, methods and discussion (sections 4.6.2 and 4.6.4) of Chapter 4, with a focus on the indicators used there and their role as a source of uncertainty.

Using one or more hydroclimatic variables or drought indicators, another key decision to make in drought research is to identify the *drought characteristics* of interest. Key characteristics which are often the focus of drought research are the *frequency* (how often do drought conditions occur?), *duration* (for how long does a drought event persist?), *spatial extent*, *peak intensity* (the maximum deviation from normal conditions as identified using the indicator of choice) and

total severity (typically defined by aggregating the drought intensity values over the duration of the drought event). As droughts are complex events evolving through space and time throughout their duration, methods have been developed to perform joint analyses of spatial and temporal drought characteristics (e.g. Sheffield et al., 2009; Shiau and Modarres, 2009; Halwatura et al., 2015) or even to fully characterise the evolution of individual drought events in space and time (Haslinger and Bloeschl, 2017; Vernieuwe et al., 2020). Other drought characteristics have also been derived from time series of drought indicators to investigate specific research questions, such as drought termination (Parry et al., 2016).

Furthermore, it is worth briefly noting the distinction between drought as a *hazard* and drought *risk*, as these terms can sometimes be conflated. In a climate change context, risks can result from the interactions between climate-related hazard(s), exposure to these hazards and the vulnerability of the systems exposed to these hazards (Ara Begum et al., 2022). When a drought occurs and people or ecosystems are exposed to it, the vulnerability of the affected agents regulates the associated levels of risk. Of these three factors constituting risk, the present work is primarily concerned the nature and changes of drought hazards and exposure of water resources to those hazards, and does not explicitly consider vulnerability. However, a major aim of the work is to contribute to reducing vulnerability to future droughts by providing information which can be helpful for adaptation planning.

1.3 Causes and consequences of drought

1.3.1 Drought development and propagation

Large-scale atmospheric conditions provide the starting point for developing a drought signal in terms of meteorological variables (precipitation, temperature, and variables contributing to atmospheric moisture demand). For the UK, these atmospheric conditions are strongly coupled to the surface energy budget over the Atlantic Ocean (Wilby et al., 1997). One of the most important important modes of climatic variability linked to UK weather and drought conditions, with substantial spatial variations in how it explains precipitation variability and hydrological droughts, is the North Atlantic Oscillation (NAO) (Wilby et al., 1997; Kingston et al., 2013; Rust et al., 2021). The NAO is defined by the variation in the meridional atmospheric pressure gradient across the North Atlantic, i.e. the difference in strength between the high pressure system over the Azores and the low pressure system over Iceland. Positive NAO winters are

associated with a strengthened westerly atmospheric flow into the UK, bringing moist Atlantic air masses and increased precipitation – especially over the western parts of the UK where it is enhanced orographically (Seager et al., 2020). Conversely, therefore, it is the negative NAO winters – where these anomalous conditions are approximately reversed – that are associated with drier than usual conditions over the UK (again, especially for the western regions). The NAO in other seasons also has some influence on rainfall in parts of the UK, though to a lesser extent than in winter. In the south east of GB, winter precipitation is strongly affected by the East Atlantic pattern (Hall and Hanna, 2018).

In addition to the NAO-related variations in westerly circulation, the atmospheric circulation patterns that give rise to precipitation deficits (especially in the east of the UK) are typically anticyclonic (high pressure) conditions, which deflect storm systems and cause low cloud formation and precipitation rates. Richardson et al. (2018) linked droughts in different parts of the UK to a novel set of 30 weather patterns developed by Neal et al. (2016), and found 6 patterns which occur more frequently during droughts in all parts of the UK plus some specific patterns which are only associated with droughts in particular regions - generally consistent with drought occurring under conditions of high atmospheric pressure (similar to the findings of Fleig et al., 2011), or in the downwind areas of the different weather patterns. In the analysis by Fleig et al. (2011), weather patterns with airflow from the south and south-east were linked to droughts in Great Britain, likely due to the low moisture content and higher temperatures of these continental winds. Atmospheric high-pressure blocking systems, which remain quasi-stationary for a period of time and generally divert storm tracks around centres of high atmospheric pressure, are often associated with droughts, drought-heatwaves, and other hydrometeorological extremes (Kautz et al., 2022). Under these blocking conditions, precipitation is largely reduced under the high pressure blocking centres (Sousa et al., 2017). Going in more detail, Sousa et al. (2017) analysed changes in the distribution of precipitation over the UK and Ireland for blocking circulation over the Atlantic, and showed a very strong reduction in the number of days with more precipitation than the median, paired with an increase in the number of light rainfall days (which do not contribute much to the total precipitation), together resulting in decreased mean precipitation.

Moving one stage further in the hydrological cycle, land-atmosphere interactions play an important role in the evolution of droughts. A lack of moisture on land can in turn feed back to precipitation decreases locally or downwind: when less water is available for evaporation, less moisture can be taken up by the atmosphere over land, which reduces the amount of

atmospheric moisture available for precipitation. At the same time, air temperature increases with the larger fraction of the available energy being converted to sensible heat due to the decrease in evaporation. Due to Clausius-Clapeyron, this then increases the saturation humidity level and thereby atmospheric moisture demand, leading to a self-intensifying feedback mechanism. The role of atmospheric evaporative demand in drought processes in GB is discussed in Chapter 4 (Reyniers et al., 2023).

The transformation of the drought signal from anomalous meteorological conditions through the terrestrial components of the hydrological cycle is termed drought propagation (Van Loon and Laaha, 2015). As precipitation deficits propagate to streamflow and groundwater levels, there is a general evolution from many short periods of deficit to fewer, longer, less intense periods of low streamflow and groundwater levels, which somewhat lag behind the rainfall deficits (Van Loon and Laaha, 2015). Drought propagation is heavily influenced by climate and catchment characteristics. Different processes govern hydrological drought development in arid regions, monsoon climates, warm seasonal climates, cold seasonal climates with a significant snow accumulation and snow-melt cycle, or climates with little seasonal variation (Van Loon and Laaha, 2015). As such, different climate types can be associated with different types of hydrological droughts, as classified by Van Loon and Van Lanen (2012). The time scale of propagation from precipitation deficits to streamflow drought is also significantly different between climate types, and is seasonally variable depending on the climate type (Gevaert et al., 2018; Ding et al., 2021). Furthermore, Apurv et al. (2017) was able to link three modes of drought propagation (based on the seasonal groundwater recharge cycle) to aridity, the degree of seasonal variation of precipitation, and most importantly the phase-shift between precipitation and potential evapotranspiration levels. Catchment storage (including aquifers, but also surface water bodies) has a crucial influence on the propagation of the meteorological drought signal to streamflow droughts (Van Loon and Laaha, 2015). Barker et al. (2016) showed for UK catchments that the influence of climate on drought propagation is regulated by catchment characteristics. System recovery after drought is also influenced strongly by aquifer hydrogeology and responsiveness (Stoelzle et al., 2014; Parry et al., 2016). Aquifer responsiveness also influences streamflow drought severity and duration (Van Lanen et al., 2013; Barker et al., 2016; Tijdeman et al., 2018). Even for given catchment characteristics and a given climate type, hydrological droughts tend to develop through multiple mechanisms: over 85% of the European catchments studied by Brunner et al. (2022) experienced 5 or 6 of the process-based drought types developed by Van Loon and Van Lanen (2012) and Van Loon et al. (2015). Moreover,

catchment characteristics might be a less important influence for abrupt drought terminations than for gradual, prolonged drought termination. Finally, Brunner et al. (2022) showed that the dominant drivers vary between moderate and more severe droughts.

Last but not least, human impacts on hydrological systems also shape hydrological drought development and propagation (Van Loon et al., 2016a), through abstraction, inter-basin transfers, reservoir construction and operation, sewage return and pipe leakage, and land use (Van Loon et al., 2022). In UK aquifers, for example, Wendt et al. (2020) identified two distinct typologies of ground water drought modifications caused by human water use, depending on the balance between groundwater abstraction and groundwater recharge. Tijdeman et al. (2018) investigated human influences on hydrological drought (propagation) characteristics in England and Wales, focusing primarily on abstractions and reservoir effects. They demonstrated a reduced correlation between precipitation and streamflow due to different human influences, and generally found varying (catchment-specific) degrees of impacts of human influences on drought characteristics. Land use changes have already affected streamflow and evaporation in various locations across the UK, not limited to urbanised areas (Teuling et al., 2019).

1.3.2 How does climate change affect droughts?

Drought is a major natural hazard linked to devastating socio-economic, health and ecological impacts, already getting worse in many parts of the world due to anthropogenic climate change (Seneviratne et al., 2021). Climate change affects droughts through different thermodynamic, dynamic and biophysical mechanisms taking place at multiple spatial and temporal scales (Douville et al., 2021; Seneviratne et al., 2021). Thermodynamic effects were identified as the main way in which human-induced global warming affects droughts with *high confidence* (Seneviratne et al., 2021). Following the Clausius-Clapeyron relation, a parcel of air at standard sea level pressure can hold about 7% more moisture for every temperature increase of 1 °C. While atmospheric water vapour content is expected to closely follow this, global mean precipitation increases at a much lower rate with increasing global temperatures (Douville et al., 2021). Beyond the global mean, changes to the water balance over land are region- and season- dependent, and are influenced by atmosphere dynamical changes and land surface interactions. These regionally variable changes over land are now understood to be too complex to be adequately described by the 'wet gets wetter / dry gets drier'-paradigm (Greve et al., 2014; Feng and Zhang,

2016; Yang et al., 2019; Xiong et al., 2022), though an amplification of the seasonal contrast between wet and dry periods is often found (Douville et al., 2021). Changes to dynamic circulation processes resulting from the increase in energy in the climate system tend to be more uncertain than direct thermodynamic effects (Shepherd, 2014).

The influence of climate change on mid-latitude precipitation results from a combination of a number of processes, whose relative contributions and interactions are still uncertain (Doblas-Reyes et al., 2021; Douville et al., 2021). Atmospheric evaporative demand is strongly affected by climate change. An increase in atmospheric evaporative demand leads to increasing evaporation from surface water, vegetation and soils, subject to moisture availability. Changes in transpiration from vegetation are crucial and complex, as they result from the combination of increasing atmospheric evaporative demand, increasing water use efficiency of plants due to increasing CO_2 concentration, increases in plant growth, and other plant physiological responses to temperature, vapour pressure deficit, soil moisture changes and CO_2 concentrations (Canadell et al., 2021). Land-atmosphere feedbacks play a key role in determining the impact of increasing temperatures on drought (Seneviratne et al., 2010; Miralles et al., 2019). Amplified evaporation rates directly lead to depletion of soil moisture, which can have a self-amplifying effect on atmospheric evaporative demand as the sensible heat flux increases when moisture availability becomes limiting for evaporation. Depending on the region, these feedbacks can also contribute significantly to precipitation deficits (e.g. Schumacher et al., 2022a).

While this work looks at impacts of climate change on the water supply side, it must be noted that climate change may also affect UK water demand, as households tend to consume more water as temperatures increase (Parker and Wilby, 2013; Xenochristou et al., 2020).

1.3.3 Impacts of drought

Droughts impacts can be observed in all sectors, including (aquatic) ecological harm, threats to domestic water supply, agricultural impacts, energy production disturbances, and economic losses in water-dependent industry sectors (Stahl et al., 2016). These different drought impacts can have complex interactions, which can be identified as co-occurring or cascading behaviour (de Brito, 2021). Importantly, the nature and severity of drought impacts depend not only on physical characteristics of the drought hazard, but are crucially dependent on vulnerability and depend on the local and global socioeconomic situation and governance (Savelli et al., 2021).

Public water supply is one of the most crucial sectors which can be impacted by streamflow and/or groundwater droughts (e.g. Stahl et al., 2016). In the UK, an extreme example of this is the implementation of standpipes and rota cuts during the extremely dry conditions of 1976 (Rodda and March, 2011). Potable water supply can be affected not only through a reduction in available water quantity in water resources, but also by water quality deterioration associated with drought or drought termination, such as algal blooms, increased pollutant concentrations, microbial contamination linked to waterborne diseases, or saline intrusions (Wright et al., 2013; Khan et al., 2015).

Hydrological droughts can also threaten energy production in the UK and globally, due to low flows disrupting the operation of hydro-power plants (Dallison et al., 2021) or thermal plants which rely on cooling water from rivers (Byers et al., 2020). Impacts on aquatic ecosystems arise not only due to reduced and slower flows, but also indirectly through the resulting higher water temperatures and deteriorated water quality (Lennox et al., 2019). Droughts lead to ecosystem and crop productivity reductions worldwide (Gampe et al., 2021), due to insufficient soil moisture availability (Anderegg et al., 2012), irrigation water shortages (Stahl et al., 2016) and/or high temperatures/potential evapotranspiration levels (Gampe et al., 2021) and especially high vapour pressure deficit (Schönbeck et al., 2022; Lu et al., 2022) (see also Section 4.6.3).

Droughts do not just affect ecosystems in the short term, but can also bring about long-lasting changes (Müller and Bahn, 2022). Moreover, while climate change affects the development and propagation of droughts, drought in turn reduces the effectiveness of the land carbon sink globally and regionally (Canadell et al., 2021, Cross-Chapter Box 5.1 and references therein). Globally, the effects of drought on crops have led to food insecurity and threats to the livelihoods of people that depend on agricultural production (Caretta et al., 2022), and in some parts of the world drought is a significant driver of migration (mostly short-distance from rural to urban areas; Cissé et al., 2022). In addition, the pressures associated with drought can significantly aggravate political tensions and conflict in different regions of the world, depending on the prevailing context and vulnerabilities (Kelley et al., 2015; von Uexkull et al., 2016). Drought can lead to a range of impacts on human health, which are heavily determined by vulnerability factors at the individual and population level. These include malnutrition (which can increase susceptibility to other mental and physical health problems), water-borne diseases, and death (Stanke et al., 2013). Finally, studies have shown that mental health impacts of drought and water scarcity exist and are complex, depending on many factors including source of income and its dependence on water (e.g.

farmers) and cultural aspects (e.g. Stanke et al., 2013; Bryan et al., 2020; Wutich et al., 2020).

Given the diverse and context-dependent range of potentially devastating drought impacts, understanding how droughts can be expected to change at the local scale depending on continued climate-relevant emissions, as well as the uncertainty associated with these projections is crucial.

1.4 Drought in the UK: past and future

1.4.1 Historic droughts

While the UK is not typically seen as a particularly drought-prone area, droughts have been a recurring concern in the past (e.g. Rodda and March, 2011; Kendon et al., 2013; Turner et al., 2021), including in very recent years (Schumacher et al., 2022b). The Drought Inventory (`https://www.ceh.ac.uk/our-science/projects/drought-inventory`, created by CEH as part of the Historic Droughts project) gives an overview of the major droughts and their impacts affecting the UK between 1890 and 2012, while Lister et al. (2018) provides an overview of the hydroclimatic context of 7 major historic droughts affecting East Anglia from 1920 to 2012. A number of studies have reconstructed historic droughts in (parts of) the UK, focusing on different hydroclimatic variables (e.g. Spraggs et al., 2015; Smith et al., 2018a; Murphy et al., 2020). A reconstruction by Murphy et al. (2020) revealed a number of extreme droughts in the 18th and 19th centuries, including an especially extreme event in 1765-1768. Spraggs et al. (2015) reconstructed historic streamflow and reservoir droughts during 1798–2010 specifically for the East Anglian region, with a specific focus on water resources. They showed that, in terms of reservoir droughts in the East Anglian region, the longest and most severe events since 1798 generally occurred after 1920, and the ranking of the worst droughts differs among East Anglian reservoirs due to natural and human factors. In what follows, an overview is given of notable historic droughts in light of their diversity and consequences for water supply, with emphasis on the East Anglian region (see Fig. 2.1.1 for a map of Anglian Water's water resources).

The 1920-1922 drought is one of the benchmark droughts of the 20th century and had severe impacts in among others the East Anglian region (Marsh et al., 2007). In the 1940s, multiple droughts took place across the UK, including groundwater droughts in different parts of the Chalk in the mid-1940s (UK Centre for Ecology and Hydrology, nd). A notable year was 1944, which is used as the reference

drought year for the River Nar in Anglian Water's WRMP19 (estimated return period greater than 200 years) (Anglian Water, 2019). In 1959, a short (February to September) but very intense drought led to severe impacts on water supply depending on gravity-fed reservoirs, while groundwater supply was only modestly affected even in regions with large precipitation deficits (Marsh and Turton, 1996; UK Centre for Ecology and Hydrology, nd).

The drought of 1975-1976, combined with the extremely hot summer of 1976, is used as a benchmark event in many places in the UK and elsewhere in western Europe due to its severity and widespread impacts (Rodda and March, 2011). While UK reservoirs and groundwater levels were generally in good condition before the start of the drought (early summer 1975), the combination of extremely low winter recharge and high water use and demand resulted in extremely low levels being reached throughout 1976. Many rivers reached record low flows, often with return periods over 100 years. In groundwater-dominated rivers, flow reductions started later, but recovery in autumn 1976 was also slower (Rodda and March, 2011). The drought had severe consequences for water supply and even led to failures and the use of standpipes for water supply, especially in the summer and autumn of 1976. The lasting impact of this extreme event on water management in the UK includes new legislation introduced in 1976 and 1977 (UK Centre for Ecology and Hydrology, nd).

In contrast to e.g. the short but intense 1959 drought, the 1988-1993 drought was marked by persistent, prolonged rainfall deficits and multiple weak groundwater recharge seasons, exacerbated by exceptionally high temperatures and evaporative losses. Low river flows were recorded in many catchments, including some of the lowest flows on record in eastern, southern and central English rivers by 1991-1992 (Marsh et al., 1994; UK Centre for Ecology and Hydrology, nd). This drought was particularly severe for the east English lowlands (Marsh et al., 1994), and 1988-1992 is also noted as the reference drought for the Wensum catchment in Anglian Water's WRMP19 (Anglian Water, 2019) with an estimated return period close to 200 years. The drought of 2010-2012 was characterised by two dry hydrological winters and very strong spatial contrasts in rainfall amounts, first between the east and west of GB and especially later between the north west and south east of GB (Kendon et al., 2013). In the spring of 2012, large declines in reservoir stocks across English and Wales led to the lowest stocks on record for some major reservoirs, including Anglian Water's Rutland reservoir. Hosepipe bans were issued in different parts of England during different phases of the drought (Kendon et al., 2013). This event has served as the observed baseline to explore storylines of factors that could have made this drought even more severe, assuming physical

plausibility of the modified factors (Chan et al., 2022).

The drought of 2018-2019, including the hot summer of 2018, had widespread impacts across the UK. After 2018, the drought was particularly significant in the south-east and east of England in groundwater-dominated catchments due to multi-year sustained precipitation deficits, with streamflow deficits persisting into late 2019 or even 2020 (Turner et al., 2021). At the time of writing, the 2022 drought is the most recent significant drought to affect the UK, and some areas (including parts of East Anglia) were still in drought status. In an attribution study focusing on the 2022 soil moisture drought in the Northern Hemisphere, Schumacher et al. (2022b) showed that climate change made the 2022 drought in western Europe far more likely, and that temperature increases made an important contribution to the soil moisture deficits.

1.4.2 Projections of future droughts in the UK

Many studies have been done on climate change impact on droughts in the UK. There is general agreement that climate change is expected to change drought frequency, duration and severity across the UK.

Multiple studies have used atmospheric variables simulated by different sets of climate models to assess projections of future droughts in the UK. Using change factors derived from the probabilistic (Strand 1) UKCP18 simulations based on a high emissions scenario (RCP8.5, see Section 2.4), Arnell and Freeman (2021) found projected increases in the frequency of 3-month aggregated precipitation deficits and 6-month aggregated climatic moisture balance deficits, primarily in England, which are largely avoided in an emissions scenario with stronger mitigation (RCP2.6). Hanlon et al. (2021) derived a range of extreme weather indicators from the regional UKCP18 projections (Strand 3 of the latest set of national climate projections for the UK; see Section 2.3), and found projected increases in 3- to 36-month aggregated rainfall deficits, with hot spots in the rain shadows of high elevation regions and the English lowlands (especially for longer aggregation levels). These projected changes are dependent on the level of global warming and are almost entirely avoided by limiting global warming to no more than 1.5 °C above pre-industrial levels.

Based on regional simulations from the previous generation of national climate projections for the UK (UKCP09), Rahiz and New (2013) found projected increases in drought frequency and intensity, as well as an evolution toward more spatially coherent droughts. An analysis of the PRUDENCE ensemble of regional climate projections shows changes in the frequency of 3-month

aggregated precipitation deficits over the UK, with ensemble agreement on increases in the south and south-east but decreases projected in the north and/or west in some ensemble members (Blenkinsop and Fowler, 2007). For 6-month deficits, their projections are more uncertain in sign, leading to decreases in drought frequency for most simulations for large parts of the UK (but most ensemble members project increases in drought frequency for the Anglian Water region). In most regions, they also find decreasing maximum severity of the 6-month deficits, although there is some disagreement in the sign of changes in all regions except along the south coast.

Moving beyond atmospheric variables, several hydrological impact modelling studies (see Section 1.5) for the UK have produced projections of possible future soil moisture availability and low flows. Kay et al. (2022) investigated changes in soil moisture droughts using the hydrological model Grid-to-Grid (G2G), driven by regional UKCP18 projections downscaled to a resolution of 1km. To mitigate the effect of errors in the regional UKCP18 projections (see Chapter 3), they chose to apply monthly mean correction factors on the precipitation simulations (using CEH-GEAR data as reference), prioritising simplicity over improving the representation of higher order moments of the precipitation distribution. Their estimates for potential evapotranspiration were not bias adjusted, and were modified with a representation of the effect of increased CO_2 on stomatal resistance (Kay et al., 2022). Given this approach, they found an increasing area for which the minimum monthly mean soil moisture content is lower than the estimated residual soil moisture content, a lengthening of the average duration of seasonal dry soil conditions, and decreasing average soil moisture content across Great Britain (especially in summer and autumn but with year-round reductions in some locations). This is broadly in agreement with earlier work by Rudd et al. (2019) using regional climate projections from the weather@home2 project to force G2G (after the same simple multiplicative bias adjustment of precipitation). They found increases in the severity of soil moisture droughts throughout Great Britain, alongside intensification of peak soil moisture deficits in the east of Scotland and south-central England (including East Anglia) and increases in the spatial extent of soil moisture drought.

Considering streamflow projections, Kay et al. (2021a) and Lane and Kay (2021) used the UKCP18 projections to assess projected changes in low flows with different return periods, simulated by the G2G model (also used by Kay et al. (2022) for soil moisture projections, see previous paragraph). However, while Lane and Kay (2021) forced the hydrological model with the downscaled climate projections themselves (after a simple bias adjustment consisting of

applying multiplicative correction factors based on monthly means to precipitation only), Kay et al. (2021a) applied mean change factors to historic time series of precipitation, potential evapotranspiration and temperature. The potential implications of this approach are discussed in Chapter 3. Both studies are in agreement on projected decreases in 7- and/or 30-day low flows with return periods between 1 and 25 years, and found larger decreases toward the south and east of Great Britain. Lane and Kay (2021) also found a delay of the 7-day minimum flow.

Cammalleri et al. (2020) used 11 different G/RCM combinations from the EURO-CORDEX ensemble to force (after bias adjustment with quantile mapping) the LISFLOOD model over Europe, and found projected increases in the severity of 10-year streamflow droughts with global warming for the UK (along with most of the Atlantic and Mediterranean regions). Kay et al. (2020) compared projections between the probabilistic projections of UKCP18 and its predecessor, UKCP09, using lumped and semi-distributed hydrological models (unavoidably using the change factor approach due to limits of the probabilistic projection data). They found a consistent decrease in low flows in the sample of UK catchments studied based on both generations of climate projections, and fairly similar results between UKCP18 and UKCP09, although the spread of the UKCP18-based projections tended to be larger and less dry than those based on UKCP09 (see also Prudhomme et al., 2012; Charlton and Arnell, 2014).

Rudd et al. (2019) and Kay et al. (2018) assessed projections of future droughts and low flows based on the weather@home2 simulations, both using the G2G hydrological model and applying the simple multiplicative mean bias adjustment to precipitation only. Kay et al. (2018) found projected reductions in 30- and 7-day low flows with return periods of 2 and 20 years, with greater decreases in the south and for the longer return period. They also showed that everywhere except along the west coast an adjustment of climate model-derived projected potential evapotranspiration for stomatal closure due to increasing CO_2 curbed the projected low flow reductions. This adjustment was also applied in Rudd et al. (2019), Lane and Kay (2021) and Kay et al. (2021a).

Rudd et al. (2019) focused on drought characteristics instead of low flows. They found projected increases in the total severity of drought events, with greater increases in the south and east Scotland, and the greatest range of projections in the south and east of the UK. The projected changes in peak drought intensity were spatially variable, while increasing drought durations were found in the southern and eastern regions, and finally the spatial extent (as a portion of GB in drought) of soil moisture and streamflow droughts was projected to increase.

1.5 Modelling hydrological impacts of climate change

Decision makers in different sectors need information on how climate change is most likely to affect their activities. Modelling the potential impacts of climate change resulting from a certain assumed emissions scenario is a key approach for providing such information. While impact modelling is done for a wide range of systems (see e.g. Rosenzweig et al., 2017), only modelling stages relevant within the scope of this thesis study are discussed below.

Climate models have been indispensable to further our understanding of the many ways in which greenhouse gas emissions and other forcing factors are affecting different components of the climate system, and they have made up an essential part of every IPCC report since the first one in 1990 for that reason (Cubasch and Cess, 1990; Chen et al., 2021). There exists an enormous variety of climate models covering a large range of complexity, from simple energy balance models to complex earth system models which represent a vast range of processes in the atmosphere, ocean, cryosphere, biosphere, carbon cycle and on land, on three-dimensional global grids (Flato et al., 2013). Climate models of different complexity (and indeed the hierarchy of climate model complexity itself, e.g. Vallis et al., 2018) can be used to answer different research questions or solve different problems. For example, the computationally cheap one-dimensional model used by Manabe and Wetherald (1967) would not get very far with projections of future droughts in the UK, but it was able to show that the vertical pattern of warming in the atmosphere corresponds to a greenhouse gas effect rather than to changes in solar irradiation, and managed to make an estimate of climate sensitivity that still falls within the very likely range (although on the lower side) of the climate sensitivity estimates given in the most recent IPCC assessment cycle (Forster et al., 2021).

While climate change is a global problem, its various impacts are generally more localised, and so detailed spatial information is required to support regional and local adaptation to the consequences of climate change. Global climate models (abbreviated here as GCMs - which is also used to abbreviate General Circulation Models, but technically those do not include the more complex class of earth system models) can provide important spatial information to some extent, but although great progress is regularly being made in refining their horizontal resolution, the vast majority still operates on a grid with a resolution of the order of 100 km, which falls short of the high resolution demanded for many scientific and practical purposes. Therefore, different downscaling techniques have been developed to provide more detailed spatial information based on global climate model simulations (e.g. Prein et al., 2015,

and references therein). A widely used downscaling strategy is to nest a higher resolution regional climate model (RCM) inside a coarser global climate model, with improved spatial detail and possibly more explicit representations of finer scale processes. With one-way nesting, the GCM provides boundary conditions for the regional climate model over a limited area of the globe, with a buffer zone along the boundary to smooth the transition from coarser boundary conditions to the desired resolution.

Potential future changes to hydrological variables such as streamflow are of crucial interest to many stakeholders. Typically, to investigate projected changes in hydrological hazards from climate change, precipitation and other meteorological variables simulated by a climate model are used as input for an off-line hydrological model, resulting in an *impact modelling chain* (Hakala et al., 2019). Despite continuous improvements, climate models still tend to contain non-negligible biases, which complicates (but does not necessarily inhibit) their direct or indirect use for climate change impact studies. Generally, climate models tend to perform better for simulations of temperature than for simulations of precipitation (Flato et al., 2013). Using the raw RCM simulations of precipitation and other variables such as temperature as input for hydrological models can lead to large errors in the resulting streamflow simulations. Bias correction tends to considerably improve the resulting streamflow simulations, and the choice of bias correction method can have a considerable (intended and unintended) impact on the resulting streamflow simulations (e.g. Rojas et al., 2011; Teutschbein and Seibert, 2012; Teng et al., 2015; Pastén-Zapata et al., 2020), as input biases are propagated through non-linear processes in the hydrological rainfall-runoff model (Wigley and Jones, 1985), so a careful consideration of the changes introduced by the bias-correction is necessary.

Note that statistical bias correction does not solve climate model problems at the source (i.e. imperfect process representations), and inherently assumes that biases are stationary in time, and that the simulations can be trusted to some degree, while needing correction of the mean and possibly several higher order moments. Bias correction is further discussed in Chapter 3.

Different types of hydrological models can be used to simulate streamflow (and groundwater, however this is not the focus of the present work) from meteorological observations or (bias adjusted) climate projections. A first distinction can be made between land surface models and hydrological models. Land surface models are generally physics-based, aim to realistically simulate different fluxes of energy and moisture on land and between atmosphere and land, and are used in particular to simulate the terrestrial component of the

water cycle in climate models. On the other hand, hydrological models are primarily developed to simulate streamflow, and the representation in these models of other water fluxes primarily serves this objective. Hydrological models can be classified based on their level of process and spatial complexity (Hrachowitz and Clark, 2017). In terms of spatial complexity, the main distinction made is between lumped and distributed models. Distributed models simulate streamflow over a large number of spatial hydrological units (such as the cells of a square grid or hydrological response units), which are associated with spatially variable parameters and are driven by spatially variable atmospheric forcing (Beven, 2012). In contrast, lumped hydrological models represent the catchment as a single homogeneous hydrological unit, they are run for a single value per parameter at a time and using atmospheric forcing averaged over the whole catchment. There exist hybrid configurations, often referred to as semi-distributed, which often consist of a lumped model relying on a number of elevation bands to simulate the influence of topography within the catchment. In terms of complexity, simple models that represent processes between rainfall and runoff by a combination of water buckets are typically called conceptual and rely on calibration for parameter estimation. At the other end of the complexity continuum lie models that leverage physics to represent water dynamics and whose parameters are typically estimated using geophysical datasets (e.g. high resolution soil product). The hydrological models used in this work (Section 2.5) to simulate the relatively flat Wensum catchment are lumped conceptual models.

In the last decades, modular modelling frameworks have been developed for better understanding structural uncertainty (e.g. Clark et al., 2008, 2015; Knoben et al., 2019; Craig et al., 2020; Dal Molin et al., 2021). What sets modular modelling frameworks apart from model inter-comparisons is that they allow to trace back structural uncertainty to the representation of specific processes or specific structural differences, as well as their ability to recombine structural elements into new models. Another aim for the development of modular frameworks is to select the model structure based on known physical characteristics of the catchments to model (Pomeroy et al., 2007). The work in Chapter 5 makes use of the widely used rainfall-runoff model GR6j (Pushpalatha et al., 2011) described in Section 2.5.1 and which is used by Anglian Water, combined with the modular modelling framework FUSE (Clark et al., 2008), described in Section 2.5.2.

When chaining multiple models (and, where necessary, the calibration of these models to estimate optimal parameter sets) and processing steps after each other, the uncertainty associated with each modelling choice propagates through

the next. This results in a cascade of uncertainty (Wilby and Dessai, 2010; Smith et al., 2018b). The literature on the contributions of different sources of uncertainty to the overall uncertainty in hydrological impact modelling studies (especially those focusing on droughts or low flows) is discussed in Chapter 5.

1.6 Thesis objectives and outline

This chapter provided the necessary background for understanding the broader context of the work presented here. The overarching aim of this thesis is to better characterise the projected impacts due to climate change on droughts over GB and East Anglia, with a specific focus on how methodological choices can affect the projected impacts and the associated uncertainty. Hereby, this thesis aims to develop valuable insights for future climate change impact studies, aside from generating the future drought projections themselves. The overall structure of this thesis follows consecutive steps along a hydroclimatic model chain, and moves from atmospheric variables over atmosphere-based drought indicators to streamflow drought characteristics. Within this narrative, the overall aim of this thesis can be split into the following specific objectives (and within each chapter, these objectives are refined further):

Chapter 3 evaluates and bias adjusts the state of the art climate model simulations used, and investigates projected trends in precipitation and temperature indices.

Chapter 4 investigates projected drought characteristics for the whole of GB based on two atmosphere-based drought indicators, and closely examines their differences.

Chapter 5 zooms in on the Wensum catchment to investigate projected changes in streamflow drought frequency, and explores model evaluation approaches to constrain the associated uncertainty.

Chapter 6 uses the ensemble of streamflow projections produced in Chapter 5 to explore novel indicators to characterise the distribution of durations of individual streamflow drought events, and investigates how the resulting metrics change over time for the Wensum.

Preceding this, **Chapter 2** provides an overarching introduction of the study areas, observation datasets, climate model simulations, scenario selection method and hydrological models used in the results Chapters 3, 4, 5 and 6. Finally, **Chapter 7** provides overarching conclusions and forms

recommendations to researchers and practitioners based on the work presented in the previous chapters, and outlines potential avenues for future research. The literature review is not concentrated into a single chapter, but instead distributed over the current Chapter, parts of Chapter 2, and the introduction and discussion sections of Chapters 3-6.

2

Study area, data and models

2.1 Study area

Great Britain (GB) is home to a variety in hydroclimatic regimes, land features, hydrogeology and human influences, which influence how anomalous meteorological conditions propagate to drought impacts (e.g. Barker et al., 2016). This is also influenced by, and influences in turn, water management and related infrastructure across the region. An overview of Great Britain's hydrology and gaps in the current state of its scientific understanding is provided by Wagener et al. (2021).

In the Köppen-Geiger climate classification system, Great Britain has a humid temperate oceanic climate (Peel et al., 2007). There is almost an order of magnitude difference in annual rainfall totals across GB, from the wet regions in the north and west to the less humid climate in the east and south east. The prevailing south-westerly winds bring in mild, humid air from the Atlantic, resulting in large rainfall amounts along the west coast. High elevation areas along the (north-)west of GB orographically enhance this rainfall, while projecting a rain shadow on the lowlands towards the east and south-east. Weather in GB is strongly influenced by variations in the position and strength of the polar jet stream, leading to wet, stormy winters (jet stream over UK) and warm, drier summers (jet stream north of UK) on average (Davies et al., 1997), but also plays an important role in extreme events, e.g. in blocking patterns.

The North Atlantic Oscillation (NAO) is an important metric of interannual variability of (especially winter) weather in GB (e.g. Wilby et al., 1997). During a positive phase of the NAO, defined by a stronger atmospheric pressure gradient between the Azores High and the Icelandic Low, the UK experiences wetter, milder winters. The magnitude (and even sign) of the influence of the NAO on precipitation varies spatially across GB, generally strongest in the north-west and weakest in the south-east (e.g. Wilby et al., 1997; Folland et al., 2015; West

et al., 2019), depending on the season (West et al., 2019). In fact, the East Atlantic Pattern has a stronger influence on winter precipitation in southeast GB (Hall and Hanna, 2018) and climatic variability has also been linked to the Scandinavian Pattern (Bueh and Nakamura, 2007; Comas-Bru and McDermott, 2014). Positive phases of the Summer North Atlantic Oscillation (SNAO) are correlated with higher temperatures and lower summer rainfall in GB, associated with anticyclonic conditions (Folland et al., 2009).

Moving from the atmosphere to the land surface, the difference in water availability across GB is also reflected in the spatial variability of evaporation controls across GB. Based on annual averages of MORECS data, Kay et al. (2013) show there is a north-west to south-east gradient from energy-limited to more moisture-limited evaporation. In the south-east, potential evaporation exceeds precipitation for several months from late spring to early autumn, while in parts of the north-west, monthly averaged potential evaporation doesn't exceed precipitation even in summer. Therefore, there are larger gaps between actual and potential summer evaporation in the south east compared to the north west (Kay et al., 2013). In an analysis of the correlation between annual evapotranspiration and incident solar radiation or precipitation for North America and Europe, Teuling et al. (2009) present a more complicated spatial pattern of evaporation controls over GB. In their analysis, which is limited to evaporation and radiation data on rain-free days in May-September, GB grid cells show any combinations of a weak to strong positive correlation with precipitation and a weak to strong positive correlation with radiation, showing GB in a transition region from energy-limited northern Europe to moisture-limited southern Europe.

Finally, the propagation of meteorological conditions to hydrological hazards in GB is strongly influenced by hydrogeology and by human-built water infrastructure. Groundwater and groundwater-influenced catchments respond to precipitation deficits aggregated over longer scales than other catchments, and these time scales show significant variation among individual catchments and sites (Bloomfield and Marchant, 2013; Barker et al., 2016). In addition, human interventions have modified the natural hydrological cycle in a continuous effort to fulfil the water-related needs of an evolving society, and in turn this (purposefully or not) modifies the propagation of anomalous meteorological conditions to drought impacts (Van Loon et al., 2016a). For example, the spatial dynamics of drought propagation over England are influenced by an expansive network of transfers and reservoirs (Dobson et al., 2020). In the region managed by Anglian Water, eight reservoirs of varying sizes store surplus water from wet periods for use during drier periods, five of which are connected

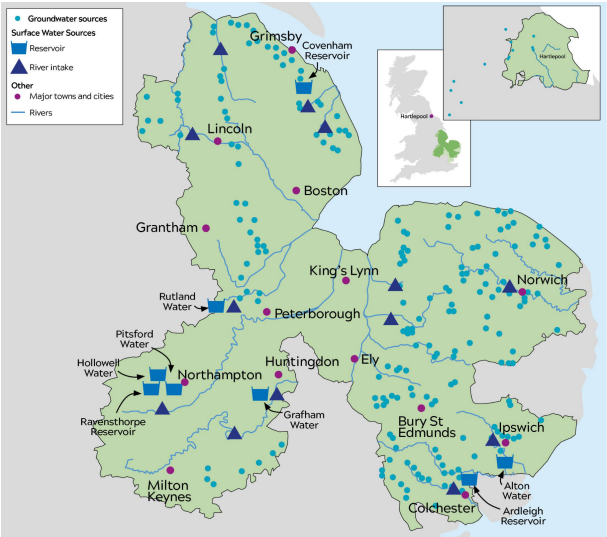

Figure 2.1.1: Map of Anglian Water's ground- and surface water resources. Reproduced from Fig. 1.3 in Anglian Water (2022b).

into one partially integrated system (Anglian Water, 2021).

The region managed by Anglian Water is located in the East of England, which is one of the driest (and warmer) regions of the UK characterised by relatively low precipitation (generally under 700 mm year^{-1}) and high atmospheric moisture demand. The locations of Anglian Water's water resources, including reservoirs and direct abstraction intakes (e.g. the Wensum, see below), are shown in Fig. 2.1.1 which was reproduced from Anglian Water's most recent draft drought plan (Anglian Water, 2022b). Agricultural land makes up three quarters of the region of East Anglia, described by the National Farmers Union as 'Britain's breadbasket' (National Farmers Union, 2016). In 2021, the East of England was responsible for about two thirds of England's sugar beat crop, one third of its potato crop, 28% of its wheat crops, 30% of its field vegetables and a quarter of its barley and oilseed rape crops (DEFRA, 2023). Aside from its crops, East Anglia is also a large supplier of eggs, chicken and pigs (National Farmers Union, 2016). As these agricultural activities result in a high water demand, it is important to note that AW generally does not supply water for agriculture.

The Wensum catchment, gauged at Costessey Mill (nrfa ID = 34004) spans 570.9 km^2 and is relatively flat, with altitudes between 5.3 and 96.1 mAOD. The catchment is a primarily agricultural area, with land cover consisting for 74.02% of arable and horticultural land, 13.59% of grassland, 8.98% of woodland and 3.06% of urban area (UKCEH, 2023). The River Wensum is groundwater-dominated, influenced by the chalk aquifer underlying the catchment. The River Wensum makes up Anglian Water's main water resource

for the public water supply of Norwich, a city with an estimated population of 144,000 people (2021 Census, Office for National Statistics, 2022). The bulk of the surface water abstraction took place at Costessey Mill from 1988 until 2019, when it was moved downstream to the Heigham intake in Norwich (Jones et al., 2004; Mott MacDonald, 2020, personal communication, 2021). Streamflow in the River Wensum is also significantly influenced by groundwater abstraction (e.g. Jones et al., 2004). These abstractions make the river particularly challenging for hydrological modelling studies (Jones et al., 2004). Similarly, in a benchmarking study, Lane et al. (2019) found that catchments in the south-east of England were generally challenging to model (as measured by the Nash Sutcliffe Efficiency, Nash and Sutcliffe, 1970). Low rainfall amounts, human influences and missing representation of groundwater processes in catchments underlain by the Chalk aquifer were all mentioned as possible explanations for this poor performance. Finally, the Wensum catchment is also of significant ecological importance, as recognised by its classification as a Site of Specific Scientific Interest and as a Special Area of Conservation (Wood et al., 2022). In fact, most of the world's chalk streams are in southern and eastern England and these catchments therefore have particular ecological significance (Rangeley-Wilson and CaBA CSRG Panel, 2021). These ecological aspects impact on management of, and abstraction from, rivers such as the Wensum (e.g. in the specification of 'hands-off' flows, below which abstraction of water can only take place under a specific temporary permit).

2.2 Gridded observation datasets

2.2.1 HadUK-Grid

The Met Office's HadUK-Grid dataset (Hollis et al., 2019) provided observations for daily precipitation used in Chapters 3, 4 and 5. It also provided the daily temperature observations used in Chapter 3. The in situ data interpolated to create this gridded dataset are mainly from the Met Office Integrated Data Archive System (MIDAS). Following the same gridding approach of HadUK-Grid's predecessors, Hollis et al. (2019) used inverse distance weighting to transform the in situ time series of climate variables to gridded datasets (reflecting point estimates for a 1 km-spaced regular grid, not grid cell averages). This was preceded by a transformation of the data to account for spatial variations. First, the variables were transformed to differences with (minimum and maximum temperature) or percentages of (rainfall) the long term average. Second, topography-dependent variations were

accounted for through regression with topographic variables (not applied for daily rainfall anomalies, but the influence of terrain elevation and aspect are included in the baseline climatology). Specifically, the topographic variables used to inform the transformation are elevation and aspect for rainfall and temperature variables, proximity to the coast for minimum and maximum temperatures, and proximity to urban areas for minimum temperatures (Hollis et al., 2019, Tables S3 and S4 in the Supplementary Information).

The base version of HadUK-Grid is provided on a 1 km grid, but it is also provided on coarser grids including a 12 km grid intended to facilitate comparison with UKCP18-RCM (Hollis et al., 2019). However, on inspection, a mismatch was found between the coordinates of the HadUK-Grid and UKCP18-RCM grid, with HadUK-Grid being offset by 1.5 km Easting and 2.5 km Northing of the OSGB36 projection relative to the UKCP18-RCM grid. The OSGB36 / British National Grid projection (Ordnance Survey of Great Britain 1936), a transverse Mercator type projection, is a coordinate reference system widely used over the British Isles which defines coordinates with eastings and northings. At the time the work in Chapter 3 was carried out, an updated version of HadUK-Grid was planned, with the 12 km grid to be modified to match the UKCP18-RCM grid (CEDA, personal communication, 25 March 2020), but timing was uncertain due to the COVID-19 pandemic (this update has since become available). As a workaround, a new 12 km dataset was constructed here from the 1 km HadUK-Grid data for daily precipitation, and minimum and maximum temperature to match the UKCP18-RCM 12 km grid. The 12 km dataset derived here was used throughout the thesis and is called HadUK-Grid. Consistent with Hollis et al. (2019), this was done by averaging all 1 km grid points that lay in each 12 km UKCP18-RCM grid cell. For the evaluation, bias adjustment and analysis of daily average temperature in Chapter 3, HadUK-Grid daily average temperature was derived by taking the average of daily minimum and daily maximum temperatures.

2.2.2 CHESS-PE

The CHESS-PE dataset provided by the Centre for Ecology and Hydrology (Robinson et al., 2020) was used as observational reference for PET in Chapters 3 and 4. Like for HadUK-Grid, this dataset was regridded from its native 1 km resolution to the 12 km resolution grid of UKCP18-RCM through averaging. The computation of PET relies on daily observations of air temperature (mean and range), specific humidity, wind speed, downward longwave and shortwave radiation, precipitation and air pressure. These variables were derived primarily from the MORECS dataset (horizontal resolution of 40 km), except for air

pressure and the daily temperature range which were derived from WFD, CRU TS 3.12 (both with a horizontal resolution of 0.5°) and CEH-GEAR (horizontal resolution of 1 km), respectively.

The method used to obtain PET in the production of CHESS-PE is an implementation of Penman-Monteith PET (Monteith, 1965) for a reference crop (Allen et al., 1998), which assumes a crop height of 0.12 m, a constant stomatal resistance of 70.0 s m^{-1}, an albedo of 0.23 and an emissivity of 0.92. For the calculation of vapour pressure deficit from temperature, Robinson et al. (2017) used a formulation based on Richards (1971). The equations and detailed explanations of the computation of PET for the CHESS-PE dataset can be found in Robinson et al. (2017).

2.3 The UKCP18 regional climate projections

UKCP18 is the most recent set of national climate projections for the UK. The research presented in this thesis makes use of its third strand, a perturbed physics ensemble (PPE) of regional climate model projections (UKCP18-RCM; Met Office Hadley Centre (2018)), available from the Centre for Environmental Data Analysis (CEDA). The RCM simulations were run over the EURO-CORDEX rotated pole grid with a horizontal resolution of 0.11°, which results in quasi-uniform 12 km spacing over the European domain (Murphy et al., 2018), and made available for the UK region with these coordinates as well as the OSGB36 projection which was used for the work presented in this thesis. The horizontal resolution of the RCM simulations over the UK region is 12 km, which for drought research presents a good trade-off between practicality and spatial detail compared to the UKCP18 high-resolution convective permitting simulations. Simulations of different variables are available from 1 December 1980 to 30 November 2080 on a daily time step (for practical reasons, December 1980 was left out of our analyses). The UKCP18-RCM simulations are a PPE, obtained by running the global climate model (GCM) HadGEM3-GC3.05 with perturbations in 47 parameters, which was then downscaled by one-way nesting with a regional configuration of this model using the same perturbed parameter sets as its driving GCM ensemble member.

The HadGEM3-GC3.05 parameter sets chosen for the global (Strand 2) and regional (Strand 3) UKCP18 PPE ensembles were selected in multiple stages based on different criteria, as summarised briefly below and explained in Murphy et al. (2018) and references therein. Following expert judgement, 47 parameters were selected to perturb in the convection, gravity wave drag,

boundary layer, cloud, large-scale precipitation, aerosols, and land surface interaction schemes (the ocean component was excluded from perturbations from the outset). From 2800 parameter sets sampled from their prior distributions, 442 survived a first selection based on 5-day hindcasts. These, plus an additional 115 parameter sets (so 557 in total), were further filtered based on 5-year atmosphere-only simulations, from which 39 were retained based on a wide range of performance metrics. Due to computational constraints, the 25 most diverse ensemble members were selected from these 39 to run coupled ocean-atmosphere projections. Based on historical runs, two ensemble members were excluded due to numerical instabilities, and three more were excluded because they simulated a spin-down of the Atlantic Meridional Overturning Circulation (AMOC) inconsistent with observations and large negative sea surface temperature biases in the North Atlantic.

The 20 remaining ensemble members were run to 2100. For the global projections (UKCP18 Strand 2), these 20 were reduced to the final 15 ensemble members after eliminating 5 more ensemble members due to unsatisfactory representations of the AMOC, the simulated historical trend in Northern Hemisphere surface temperatures, and European climatology. For the regional projections (UKCP18 Strand 3), 16 of the 20 parameter sets were selected first, aiming to maximise the spread in the 47-dimensional parameter space within the limitations of computational resources. These 16 ensemble members still contained 4 of the ensemble members eliminated for the global projections, which were therefore also eliminated, resulting in the final 12 ensemble members chosen to drive the regional projections (numbers 1, 4, 5, 6, 7, 8, 9, 10, 12, 13 and 15). These include the standard parameter set (number 1) as well as the weakest and strongest estimates of global aerosol forcing and climate feedback strength found in the 15 member global PPE, and maximises the spread of 47 parameter values (Murphy et al., 2018).

A limitation of this modelling setup relying on a single climate model structure is that it causes UKCP18-RCM to sample a narrower range of potential future outcomes than the UKCP18 global projections, which combine the 15 member global PPE with an ensemble of 13 other CMIP5 GCMs. However, complementing the analysis with extra ensemble members produced with different climate model structures (e.g. from the EURO-CORDEX projections Giorgi et al., 2009) was considered practically outside the feasible scope of this study. Furthermore, restricting this study to the use of UKCP18 makes the results more comparable to other work using UKCP18 for the UK (e.g. Arnell et al., 2021; Kay et al., 2021b; Hanlon et al., 2021), thereby resulting in a deeper understanding of the implications of these climate projections. It is also

important to note that in the end, the UKCP18 datasets were specifically developed for the UK, and that these sets of projections have been intended and recommended for use by decision makers in the UK to support adaptation to future climate change impacts (Met Office, 2019).

2.4 Warming scenarios

The work in Chapters 4 and 5 makes use of time slices representing warming levels of +2°C and +4°C above pre-industrial levels, consistent with the derived projections made available within the UKCP18 set of projections (Gohar et al., 2018). The rationale for this choice is explained below, as well as a framing of the realism of these two warming levels.

Importantly, HadGEM-GC3.1, to which HadGEM-GC3.05 is closely related (Lowe et al., 2018), has a high equilibrium climate sensitivity (ECS, the warming resulting from a doubling of CO_2) of 5.4 °C (Meehl et al., 2020). This lies substantially above the *likely* upper range of 4.5 °C from the IPCC's 6th Assessment Report (Arias et al., 2021), but cannot be ruled out. It samples the upper end of the UKCP18 probabilistic projections of global mean surface temperature change, and projects larger temperature increases than almost all of the CMIP5 models included in the UKCP18 global ensemble (see Fig. 2.13 in Lowe et al. (2018)). UKCP18-RCM presents a comparatively warm set of climate projections for the UK, notably tending toward a drier and hotter summer response than the UKCP18 selection of CMIP5 projections. UKCP18-RCM and its driving global PPE mostly cover the upper quartile of summer warming of the UKCP18 probabilistic projections, and the majority of the ensemble projects greater drying than the median summer precipitation change in the probabilistic projections. For winter, there is more overlap between the CMIP5 selection and the global and regional HadGEM-GC3.05 PPE, and a greater portion of the range of the probabilistic projections is sampled (see Fig. 5.2 in Murphy et al. (2018)).

Secondly, the UKCP18-RCM simulations were only produced using the RCP8.5 emissions pathway, a very high emissions scenario characterised by high population growth and energy demand (Riahi et al., 2011). It is sometimes interpreted as the 'business as usual' scenario but this interpretation has been criticised (Hausfather and Peters, 2020; Riahi et al., 2022). Instead, RCP8.5 should be viewed as a high-end, high-risk scenario which has become considerably less likely due to developments in energy technologies and policy during the last decade (Riahi et al., 2022) but which cannot yet be completely

ruled out. Exploring such possible worst case scenarios is relevant, especially given the extremely high stakes of impacts of climate change on vital sectors such as water supply and food production. While all HadGEM-GC3.05 PPE simulations in UKCP18 follow the RCP8.5 emissions scenario, the exact CO_2 concentration pathways were designed to differ among the ensemble members to represent uncertainty in the carbon cycle processes converting carbon emissions to atmospheric CO_2 concentrations (Murphy et al., 2018). For UKCP18-RCM PPE, the regional downscaling HadGEM-GC3.05 configurations use the same CO_2 concentration pathways as their driving global simulations.

To reduce the influence of some of the assumptions described above and provide quasi-scenario-neutral projections of future drought, a time slice approach was used to represent different global warming levels (James et al., 2017) rather than using a specific future time period under a particular scenario. The two warming levels investigated, +2 and +4 °C above pre-industrial levels, are based on the same warming levels for which UKCP18 derived projections were produced and makes use of the years in which a 25-year running mean of global-mean temperature from the driving global model reaches the warming level of interest presented in Gohar et al. (2018). Warming levels are specified relative to pre-industrial conditions, defined as the 1850-1900 climate. As opposed to the CMIP5 part of the global UKCP18-ensemble, all 12 driving global PPE members reach +4 °C of global mean warming at some point (Gohar et al., 2018).

Importantly, the time slice approach would result in an accurate assessment of changes in GB drought projected at these warming levels if these changes would scale directly with global temperature increase (independent of the speed of change), and if the regional model has the same climate sensitivity as its driving global model. Neither of these requirements are likely to be fully met. UKCP18-RCM projects slightly weaker UK temperature responses towards the end of the simulated period than their driving global simulations (Fig. 5.2 in Murphy et al. (2018)). Additionally, midlatitude atmospheric circulation patterns in the selected time slices (which influence UK weather and therefore drought events) might be related more directly to the radiative forcing rather than to the global temperature increase projected under that radiative forcing by a particular GCM (Ceppi et al., 2018). Nevertheless, the applied time slice approach is a reasonable approximation and is frequently used for investigating impacts at different levels of global warming (e.g. James et al., 2017; Naumann et al., 2018; Samaniego et al., 2018).

As the time slices representing +2°C and +4°C of warming above pre-industrial levels are used for many of the results presented in this thesis, some context is

required for understanding their realism and implications. In the 2015 Paris agreement, 196 Parties (covering most of the world's nations, population and emissions) agreed to *'holding the increase in the global average temperature to well below 2 °C above pre-industrial levels and pursuing efforts to limit the temperature increase to 1.5 °C above pre-industrial levels, recognizing that this would significantly reduce the risks and impacts of climate change'* (United Nations, 2015). The latest IPCC Working Group III report on mitigation states that not strengthening national policies beyond those implemented by the end of 2020 would lead to a *very likely* projected global warming between 2.2 and 3.5 °C above pre-industrial levels by 2100 (IPCC, 2022). Including updated pledges and policy after COP27 in Sharm-el-Sheikh, the 2022 warming projections report by Climate Action Tracker gives a similar estimated warming range based on current policies and actions (2.2-3.4 °C) (Climate Action Tracker, 2022). They further report a substantial gap between projected 2100 warming based on 2030 NDC targets (1.9-2.9 °C, central estimate of 2.4 °C) and their 'best case' scenario where all announced targets are fully implemented (1.5-2.3 °C, central estimate of 1.8 °C). Indeed, multiple recent studies show that the optimistic scenario with full implementation of all pledges and announced targets could suffice to limit global warming to +2 °C, but would fall short for the +1.5 °C target (Meinshausen et al., 2022; United Nations Environment Programme, 2022; Climate Action Tracker, 2022).

The time slice representing two degrees of warming above pre-industrial levels thus represents an achievable future with improved policy and action compared to the current situation. The other time slice, with four degrees of warming, can be regarded as a possible higher-end warming scenario and falls within the very likely (5%-95%) ranges of high-emissions scenarios which can be considered either about as pessimistic as, or more pessimistic than, pathways based on current policies (Lee et al., 2021; Riahi et al., 2022, e.g. SSP3-7.0, SSP5-8.5 or C7-C8). Furthermore, multiple climate tipping points could be crossed which would result in irreversible state changes of elements in the climate system. Several of these could be reached even below +1.5 or +2 °C of warming above pre-industrial levels, and for some tipping points even at current global temperatures, although the number of possible tipping points and the risk of crossing them increases depending on global warming levels (Armstrong McKay et al., 2022). The high degree of uncertainty associated with these tipping points and their potential interactions (Ripple et al., 2023) further motivate the consideration of seemingly pessimistic scenarios or global mean warming levels for climate change impact studies.

2.5 Hydrological models

In this section, the hydrological rainfall-runoff models used in Chapter 5 are explained.

2.5.1 GR6j

GR6j (*modèle du Génie Rural à 6 paramètres Journalier*) is a widely used parsimonious lumped rainfall-runoff model with only 6 free parameters (see Table 2.1). The widely known four-parameter version (GR4j), proposed by Perrin et al. (2003) as an improvement of the GR3j model (Edijatno et al., 1999), was later modified to GR5j by Le Moine (2008) who implemented a groundwater exchange function in which the direction of the exchange is not fixed but depends on the water level in the routing store. This finally became GR6j with the addition of a new exponential routing store in parallel with the existing linear one, aimed at improving low-flow simulation without adversely impacting the simulation of high flows (Pushpalatha et al., 2011). This final addition was the result of a structural sensitivity analysis, in which different formulations were tested for the groundwater exchange term F, for an additional store in parallel or in series with the existing routing store, and for the parameterisation of the input flux split between the existing and new store. Based on this analysis, the existing groundwater exchange term from GR5j was kept, a fixed (instead of catchment-specific) split between the stores was deemed suitable, and an extra parallel store proved more advantageous than adding an extra store in series (Pushpalatha et al., 2011). Finally, an exponential store was found as the best option for this additional parallel store in GR6j (Pushpalatha et al., 2011). Besides its widespread usage, GR6j is primarily included in this study because it was recently adopted for operational (and strategic planning) use by Anglian Water. The model is used first to naturalise flows for FUSE calibration (section 5.5.1), and second to produce streamflow drought frequency (Chapter 5) and duration (Chapter 6) projections driven by UKCP18-RCM, which will be compared to a larger ensemble of projections generated with FUSE.

An overview of the GR6j model structure is given in Fig. 2.5.1 and its components are discussed briefly below. For more details and the equations, the reader is referred to Perrin et al. (2003). The soil architecture consists of three reservoirs: a soil moisture accounting production reservoir (S) and two parallel routing stores (R1 and R2). GR6j assumes that precipitation at a given time step will evaporate until either the PET is satisfied or no precipitation remains

Table 2.1: Free parameters in GR6j (adapted from Pushpalatha et al. (2011) and Perrin et al. (2003)).

Parameter	Units	Explanation
X1	mm	Maximum capacity of the production store
X2	mm	Groundwater exchange coefficient
X3	mm	Maximum capacity of linear routing store
X4	day	Time base of unit hydrograph
X5	-	Threshold for change in direction of groundwater exchange (introduced in version GR5j)
X6	mm	Maximum capacity of the exponential routing store (introduced in version GR6j)

(*'as if there were an interception storage of zero capacity'* (Perrin et al., 2003), which can be seen as a way to represent vegetation). From the remaining part of the precipitation (if any), the production store is recharged by an amount determined by a non-linear function of the production store state variable and the precipitation reduced by potential evapotranspiration. Only when potential evapotranspiration exceeds precipitation, evaporation from the production store is enabled, which is represented by a non-linear function of the production store state variable and potential evapotranspiration reduced by precipitation (analogous to the production store recharge function). Percolation is represented as a power function of the production store level, and while there is no division between free and tension storage implemented, the formulation implies only a (fixed) fraction of the production store level contributes to percolation. The percolated water, together with the precipitation remainder after production store recharge (i.e. surface runoff), is then split into two flow components for routing. The smallest portion is directly routed through a single hydrograph, while the larger portion is routed through a hydrograph toward two parallel routing stores with different non-linear behaviours. A catchment exchange term (F, depending on X2 and X5 (Table 2.1)) is applied to the linear routing store and the direct flow component.

2.5.2 FUSE

The modular rainfall-runoff modelling framework FUSE (Framework for Understanding Structural Errors) was developed by Clark et al. (2008) and has since been further developed by the community (Addor, 2020), for example with the addition of a snow module ((not needed in this study) Henn et al., 2015). The rationale behind FUSE is to better understand and diagnose differences between hydrological models of comparable complexity by systematically comparing model structures with controlled differences and similarities. FUSE is a framework enabling the user to set up and run different hydrological model

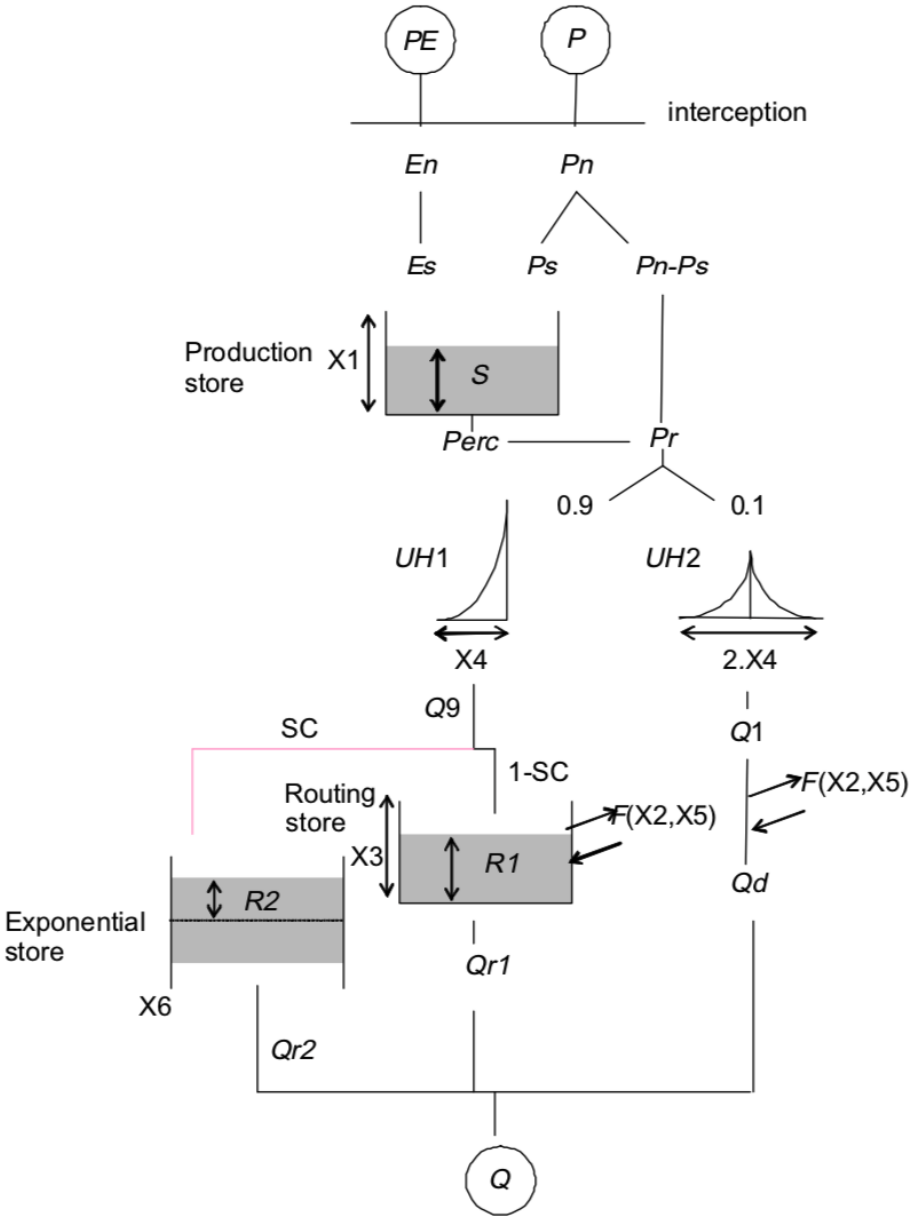

Figure 2.5.1: Wiring diagram of GR6j. Reproduced from Pushpalatha et al. (2011).

structures, i.e. to generate an ensemble of models without having to set up and run each model individually. This is made possible by the recombination of elements from four parent structures, which are based on existing hydrological models, and are referred to as follows in this thesis:

1. ARVI: based on versions of the Variable Infiltration Capacity Model (VIC) (Zhao, 1977, 1984; Wood et al., 1992; Liang et al., 1994) and borrowing from the ARNO model (Todini, 1996).

2. PRMS: based on the Precipitation-Runoff Modelling System (Leavesley

et al., 1983, 1996).

3. SACR: based on the SACRAMENTO model (Burnash et al., 1973; Burnash, 1995; Koren et al., 2004).

4. TOPM: based on TOPMODEL (Beven and Kirkby, 1979; Ambroise et al., 1996; Beven, 1997; Duan and Miller, 1997; Iorgulescu and Musy, 1997).

As illustrated in Fig. 2.5.2, the four parent structures follow a common architectural framework, with different formulations for each model component. The available choices and differences between the parent structures are briefly described below. The reader is referred to Clark et al. (2008) for more details and equations of the model choices. The soil is represented in two layers: an upper soil layer representing the unsaturated zone ($S1$) and a lower soil layer representing the saturated zone ($S2$). The unsaturated zone is implemented as a single reservoir in ARVI and TOPM. For SACR and PRMS, it is separated into one or two tension storage reservoirs and a free storage reservoir which is recharged when the tension storage is above capacity (i.e. when field capacity is exceeded), with a logistic function smoothing the transition around the capacity threshold (Kavetski and Kuczera, 2007). For these architectures, surface runoff and evaporation are generated from the (upper) tension reservoir, while percolation and interflow are generated from the free-flowing reservoir. The lower soil layer is implemented as a single reservoir in TOPM, PRMS, and ARVI, of which only ARVI has a fixed, finite size and allows evaporation to take place (see Fig. 2.5.2). In SACR, the architecture of the lower soil layer is more complex, represented as a tension storage reservoir (with evaporation) which overflows into two parallel free storage reservoirs (where baseflow is generated). FUSE also provides different options for the representation of some major fluxes from these reservoirs, which interact with the soil architecture choices.

The four options for base flow parameterisation are tied to the choice of lower zone architecture. Baseflow is represented as a linear function of the states of the infinite single store of PRMS and the two parallel free storages of the SACR lower zone, or as a non-linear function of the ARVI and TOPM single reservoirs. For TOPM, baseflow depends on the mean of the power-transformed topographic index, which is derived from a three-parameter Gamma distribution used to represent the distribution of the topographic index (which is thus not calculated directly from topographic data). Percolation from the unsaturated upper zone to the saturated lower zone is controlled by total upper zone storage (ARVI), upper zone free storage (PRMS, also used in the TOPM parent structure) or a combination of the upper zone free storage (linear) and total lower zone storage (non-linear) (SACR). For upper soil layer

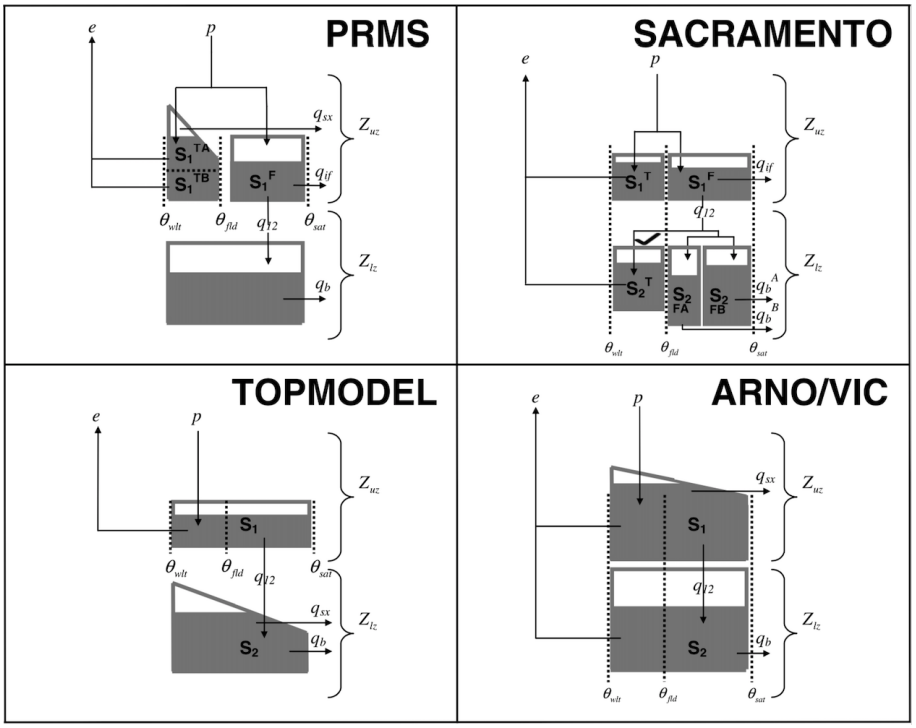

Figure 2.5.2: Wiring diagram of the four FUSE parent structures. Reproduced from Clark et al. (2008).

representations with a free storage reservoir (PRMS and SACR), interflow can be parameterised as a linear function of upper zone free storage or turned off (in this study, interflow is computed for PRMS and SACR and turned off for TOPM and ARVI).

Evaporation depends on PET and is limited by the state(s) of the source storage(s). There are two options provided for how to distribute the evaporation flux between two sources (which doesn't apply in the case of the TOPM soil architecture, where evaporation only comes from the single reservoir representing the upper soil layer). In the sequential option (parent structures PRMS, TOPM and SACR), evaporation from the lower soil layer (or lower tension reservoir of the upper soil layer for PRMS, see Fig. 2.5.2) only occurs when moisture availability of the upper soil layer is limiting, while in the root weighting option (parent structure ARVI) the relative contributions to evaporation of the upper and lower reservoirs depends on their relative root fractions.

Surface runoff is represented as a linear function of the saturated area and precipitation, with three available parameterisations of the saturated area: a linear function of the upper layer tension storage with a set maximum saturated area (PRMS, also used in SACR parent structure), a nonlinear function of the

saturation deficit of the upper soil layer (ARVI) and a function of the topographic index distribution beyond a critical topographic index which depends on the state of the total lower zone storage (TOPM). Infiltration excess runoff (also called Horton overland flow) is neglected in all structures. There is no explicit representation of vegetation stores or fluxes in the FUSE models, although (as for other missing processes) this is implicitly taken into account in the calibration. Finally, FUSE allows to represent the runoff time delay (routing) using a two-parameter Gamma distribution, which was enabled for all four structures used in this study.

The parameters for these different FUSE models are shown in Table 2.2. During the calibration with SCE (see Section 5.5.2), for some structures, some optimised parameter values closely approached the previously fixed boundaries for one or more structures. Therefore, for the parameters $S_{1,max}$, $S_{2,max}$, k_u, α, ψ, k_s, the feasible upper value limits from Clark et al. (2008) were increased, for λ the lower limit was decreased, and for the parameters ϕ_{tens}, v, v_A, v_B, χ and μ_τ the value range was expanded on both ends. It is possible that in some cases expanding the feasible range may lead to unrealistic parameter values in an attempt to solve fundamental issues with the model structures through calibration, however some adjustments can also be necessary to represent e.g. the long memory of the Wensum catchment.

Table 2.2: Free parameters in the FUSE models (adapted from Clark et al., 2008). Bracketed parent structure names give the source or inspiration for the expressions using the parameters.

Parameter	Units	Explanation
$S_{1,max}$	mm	Maximum storage in the upper layer (ALL)
$S_{2,max}$	mm	Maximum storage in the lower layer (ALL)
ϕ_{tens}	-	Fraction total storage as tension storage (PRMS, SACR)
ϕ_{rchr}	-	Fraction of tension storage in primary zone (upper layer; PRMS)
ϕ_{base}	-	Fraction of free storage in primary reservoir (lower layer; SACR)
r_1	-	Fraction of roots in the upper layer
k_u	mm day-1	Percolation rate (ARVI, PRMS)
c	-	Percolation exponent (ARVI, PRMS)
α	-	Percolation multiplier for the lower layer (SACR)
ψ	-	Percolation exponent for the lower layer (SACR)
κ	-	Fraction of percolation to tension storage in the lower layer (SACR)
k_i	mm day^{-1}	Interflow rate
k_s	mm day^{-1}	Base flow rate (ARVI, TOPM)
n	-	Base flow exponent (ARVI, TOPM)
v	day^{-1}	Base flow depletion rate for single reservoir (PRMS)
v_A	day^{-1}	Base flow depletion rate for primary reservoir (SACR)
v_B	day^{-1}	Base flow depletion rate for secondary reservoir (SACR)
$A_{c,max}$	-	Maximum saturated area (fraction; PRMS)
b	-	Exponent for saturated area (ARVI)
λ	m	Mean of the log-transformed topographic index distribution (TOPM)
χ	-	Shape parameter defining the topographic index distribution (TOPM)
μ_τ	days	Time delay in runoff (ALL)

3

Evaluation, analysis and bias correction of the UKCP18 regional climate projections.

Synopsis

The UKCP18 12km regional perturbed physics ensemble (UKCP18-RCM) is one of the three strands of the latest set of national climate projections produced for the UK by the Met Office, and is already widely being adopted in climate impact studies (the two other strands are global projections and local 2.2km projections). In this study, we report biases in the raw UKCP18-RCM simulations that are significant and are likely to deteriorate impact assessments if they are not adjusted. Two methods were used to bias correct UKCP18-RCM: quantile mapping and a variant developed for the 3rd phase of the Inter-Sectoral Impact Model Intercomparison Project (ISIMIP) designed to preserve the climate change signal. Specifically, daily temperature and precipitation simulations for 1981 to 2080 were adjusted for the 12 ensemble members. Potential evapotranspiration was also estimated over the same period using the Penman-Monteith formulation and then bias corrected. Both methods successfully corrected biases in a range of daily temperature, precipitation and potential evapotranspiration metrics, and reduced biases in multi-day precipitation metrics to a lesser degree. An exploratory analysis of the projected future changes confirms the expectation of wetter, warmer winters and hotter, drier summers, and shows uneven changes in different parts of the distributions of both temperature and precipitation. Both bias correction methods preserved the climate change signal almost equally well, as well as

the spread among the projected changes. The change factor method was used as a benchmark for precipitation, and we show that it fails to capture changes in a range of variables, making it inadequate for most impact assessments. By comparing the difference between the two bias correction methods with the differences between the 12 ensemble members, we show that the uncertainty in future precipitation and temperature changes stemming from the climate model parameterisation far outweighs the uncertainty introduced by selecting one of these two bias correction methods. We conclude by providing guidance on the use of the bias-corrected data sets. The data sets bias adjusted with ISIMIP3BA are publicly available in the following Zenodo repositories: `https://doi.org/10.5281/zenodo.6337381` (precipitation and temperature), `https://doi.org/10.5281/zenodo.6320707` (potential evapotranspiration).

Note: This chapter is based partly on a co-authored paper in preparation for submission to the journal Earth System Science Data (ESSD). That paper presents results only for precipitation and temperature, whereas this chapter substantially extends the paper by inclusion of the calculation, evaluation and bias correction of potential evapotranspiration. There are further significant differences between this chapter and the paper in the writing throughout, in terms of wording, emphasis, some restructuring, additional discussion and references.

Inclusion of research from the co-authored paper in this chapter of my thesis correctly represents my contributions to the overall work, namely co-design of the work; execution of data processing and analysis; evaluating and interpreting the results; making figures; and a large part of the writing and revising of the text. The work therefore clearly meets the requirements to be presented in this chapter as my original work, while noting and acknowledging contributions from the other authors of the manuscript: Qianyu Zha (quantile mapping, seasonal cycle (Fig.3.3.4) and bias and change maps (Figures 3.3.1, 3.3.2, 3.3.9 and 3.3.10)), Nicole Forstenhaeusler (bias and change maps (Figures 3.3.1, 3.3.2, 3.3.9 and 3.3.10)), Nans Addor (study design, writing, revision and supervision), Tim Osborn (writing, revision and supervision), Yi He (writing, revision and supervision), and Geoff Darch (revision, supervision).

3.1 Introduction

Climate model projections are essential to anticipate and adapt to future climate change impacts. Continuous efforts by the climate modelling community have led to major improvements in the realism of climate model simulations, yet significant discrepancies (biases) between observations and simulations of simulated variables remain (e.g. Kotlarski et al., 2014; Vautard et al., 2021). Biases in climate projections require particular attention when the projections are used to force impact models, for instance to assess future impacts on river streamflow, ecosystems or agricultural yields. Capturing the whole distribution of rainfall amounts is an essential prerequisite for hydrological modelling but is still notoriously challenging for many climate models. The response of impact models to forcing errors can be non-linear and amplify the bias's severity, hence biases are typically adjusted before the climate projections are used in impact models.

To this effect, a range of bias-correction (BC) methods have been developed and compared (Teutschbein and Seibert, 2012; Gutmann et al., 2014; Maraun et al., 2019). These methods essentially transform the simulations so that some of their statistical properties match those of the observations. This efficiently reduces biases and, as a result, can considerably improve impact simulations (Rojas et al., 2011; Hakala et al., 2018; Pastén-Zapata et al., 2020). We note, however, that residual biases remain after the correction, and can deteriorate the impact simulations (Teng et al., 2015). This highlights that these biases are not corrected and removed, but rather, adjusted - as such, the term 'bias adjustment' is more accurate and becoming more widely used, but here we use BC to match how these methods are more commonly described in the literature. In addition, the statistical nature of BC methods means they only address the symptoms and not the origin of model errors, i.e. they do not identify the causes of model biases nor account for them (Addor et al., 2016; Maraun et al., 2017).

Furthermore, the reliability of BC can be questioned because of its reliance on the assumption that climate model biases are stationary in time or under a changing climate state (Maraun, 2012; Ehret et al., 2012; Teutschbein and Seibert, 2012; Chen et al., 2015; Hui et al., 2020). A related issue is that BC can modify the simulated climate change trends. This poses an issue if the origin of model errors in daily variability in an evaluation period (to which BC is often calibrated) differs from the origin of potential errors in the model's climate change response (Maraun et al., 2017). These challenges are difficult to overcome. However, when following the alternative approach of using unadjusted climate model output, propagation of biases through impact models

can severely bias the resulting simulations and sometimes render them unusable (Hakala et al., 2020). As such, BC is currently necessary and widely used to assess climate change impacts, despite its imperfections.

In this chapter, the UKCP18 regional climate projections (UKCP18-RCM, Met Office Hadley Centre (2018), see Chapter 2) are evaluated and bias corrected, as these projections are used further in this thesis to investigate projected changes in atmospheric-based drought indicators (Chapter 4 and hydrological droughts (Chapter 5)). Specifically, the following questions are investigated:

1. How biased are the UKCP18-RCM projections, and does this require bias correction?
2. Can existing bias correction methods successfully correct errors in simple and more challenging metrics?
3. What climatic changes do the UKCP18-RCM ensemble members broadly project for the UK, and do the chosen bias correction techniques affect them?

An exploratory analysis of the changes projected by UKCP18-RCM is discussed in section 3.3.2, using metrics based on daily precipitation and daily average temperature.

3.2 Data and methods

3.2.1 Data

The UKCP18 regional climate projections

The UK Climate Projections 2018 (UKCP18) are the current generation of national climate projections for the UK, developed by the Met Office Hadley Centre as part of their Climate Programme (Met Office Hadley Centre, 2018) (see also Section 2.3). The UKCP18 regional projections (referred to as UKCP18-RCM here) were selected due to their unique suitability for impact modelling focusing on droughts in the UK. Droughts generally play out over a much longer time scale than floods, so the hourly resolution of the convection-permitting local projections is not required for investigating processes characterised by deficits accumulating over the span of weeks, months or years. The 12 km resolution over the UK of the UKCP18-RCM data provides

improved spatial detail compared to the global projections (60 km over the UK), while being more practical to handle than the larger files of the local projections (2.2 km over the UK). The UKCP18-RCM simulations were forced with the high-emissions RCP8.5 scenario (see also Section 2.4).

Potential evapotranspiration (PET) is not provided by UKCP18-RCM but it is a necessary input variable for some impact models (such as the rainfall-runoff models used in Chapter 5) and drought indicators (e.g. Chapter 4), so it was calculated off-line here. While rising temperatures lead to PET increases, changes in humidity, net radiation and wind speed can also play a significant role. Therefore, PET was calculated using the Penman-Monteith method, which includes the effect of all these variables and is recommended over simpler temperature-based methods (e.g. Dewes et al., 2017), although it is still subject to significant limitations (Milly and Dunne, 2016; Greve et al., 2019). The calculation of PET for the UKCP18-RCM used here relies on the same variant of the Penman-Monteith method used by Robinson et al. (2017), to ensure consistency with the CHESS-PE dataset. Specifically, the following variables simulated by the UKCP18-RCM ensemble were used: specific humidity, pressure at sea level, net downwelling longwave radiation, net downwelling shortwave radiation, wind speed at 10m and daily average surface air temperature. PET was set to zero wherever a calculated value was negative (which occurred for less than 1% of the values overall and, when split by ensemble member and month, also less than 1% for all cases except December in ensemble member 1 with 1.2% of negative values).

Observation data

As observational reference for the evaluation and bias correction of UKCP18-RCM precipitation and temperature, the Met Office's 1 km HadUK-Grid dataset (Hollis et al., 2019) was used after regridding to the UKCP18-RCM 12 km grid (consistent with Hollis et al. (2019), all 1 km grid points that lay in each UKCP18-RCM grid cell were averaged; see Section 2.2.1). For the bias correction and evaluation of PET, the CHESS-PE dataset provided by the Centre for Ecology and Hydrology was used (Robinson et al., 2020), also after regridding to the UKCP18-RCM 12 km grid (see Section 2.2.2). Daily data from 1981 to 2010 were used here.

3.2.2 Evaluation and trend analysis

Biases in UKCP18-RCM precipitation and temperature metrics were assessed over the reference period 1981-2010 (REF), using the 30-year temporal averages of a range of metrics computed for each grid cell. For analysing projected changes, the final 30-year period of the simulations was chosen as the future period (2051-2080; FUT) for each ensemble member.

Climate model errors are not necessarily homogeneous across the range of the precipitation and temperature distributions. Therefore, the initial model evaluation metrics shown consist of the errors in the mean, in a lower tail metric (Q05 for temperature and PET, dry day frequency for precipitation), and in Q95 as an upper tail metric. This is based on the results of a preliminary analysis.

Changes in temperature and precipitation extremes are generally of greater societal interest than changes in the mean (although extreme impacts do not always need extreme meteorological conditions to arise; see van der Wiel et al., 2020). Therefore, a set of moderate and extreme climate indicators was used to further evaluate the model error and analyse projected changes of simulated precipitation and temperature. These metrics were drawn from or inspired by the list of indices compiled by the Expert Team on Climate Change Detection and Indices (ETCCDI; `http://etccdi.pacificclimate.org/list_27_indices.shtml`), which have been extensively used in the literature, including in IPCC reports (IPCC, 2021). Daily mean temperature was used in this study, so the ETCCDI temperature indicators (which typically use daily minimum and maximum temperatures) were modified to use mean temperature. Table 3.1 gives an overview of all indices used and their definitions.

3.2.3 Bias correction

Comparison to observations revealed significant biases in the simulations of precipitation, temperature and PET, so these variables were statistically post-processed. Two closely related BC methods (quantile mapping and the ISIMIP3b approach, see below) were used to bias adjust UKCP18-RCM, to allow exploring the sensitivity of the results to the differences between the BC methods, including but not limited to whether or not the BC method explicitly aims to preserve the climate changes. Both BC methods were applied to each grid cell, ensemble member and calendar month combination separately, and were calibrated using simulated and observed data for 1981–2010. For

Table 3.1: Climatic indices used to evaluate and examine trends in precipitation, temperature and PET simulated by UKCP18 and derived PET.

Index	Description
Temperature	
yMEAN	Average annual (seasonal) mean daily-mean temperature (°C)
yMIN	Average annual (seasonal) minimum daily-mean temperature (°C)
yMAX	Average annual (seasonal) maximum daily-mean temperature (°C)
Qn	Daily-mean temperature exceeded on average n% of days in a year (season) (°C)
Precipitation	
prTOT	Average annual (seasonal) total precipitation (mm)
SDII	Simple precipitation intensity index: average annual (seasonal) mean wet-day precipitation intensity (mm)
DF	Average annual (seasonal) dry-day fraction (%)
Qn	Daily precipitation exceeded on average n% of days in a year (season) (mm)
Rx5day	Average annual (seasonal) maximum five-day total precipitation (mm)
CDD	Average annual (seasonal) maximum number of consecutive dry days (days)
CWD	Average annual (seasonal) maximum number of consecutive wet days (days)
Potential Evapotranspiration	
Qn	Daily PET exceeded on average n% of days in a year (season) (mm)
mean	Mean annual (seasonal) PET (mm)

precipitation, we also compared future projections from these BC methods with the change factor method, which is not a bias correction method but consists of applying projected changes in mean climate to observed time series. The change factor method was applied using the same period as the BC methods.

Quantile Mapping

*Method used by **Qianyu Zha**. A description of this method is included here as it shares some principles with the second bias correction method and to help interpret the differences between both BC methods later.*

Quantile mapping (QM; e.g. Piani et al., 2010) is a statistical transformation of the distribution of a modelled variable such that it matches the distribution of the observed variable. By construction, the resulting distributions of the simulations and observations match closely, removing deviations from the observed data in the mean, variance and higher order moments. The mapping may be applied to the quantiles of either the empirical cumulative distribution function (CDF) or to an assumed parametric CDF. Switanek et al. (2017) demonstrated that, where the parametric distribution is known to be a perfect fit to the observed and simulated data (because, e.g., it is synthetic data drawn from that distribution), then parametric QM reduces the influence of sample size-induced noise that degrades the empirical CDFs of observed and simulated

time series. However, Gudmundsson et al. (2012) found that empirical BC methods were more successful at bias correcting different precipitation quantiles than fitted distributions or simpler parametric methods.

Here, the choice was made to use QM with an empirical distribution, to complement the second BC method for which a parametric distribution is used. The resulting dataset will be refered to as BCQM. To implement QM, the R package qmap (Gudmundsson et al., 2012) was used, with QM applied to 1000 empirical quantiles. Note that using this large number of quantiles might lead to overfitting. However, our analysis shows that the differences between this non-parametric method and the parametric method introduced below are overall minor, implying that in this case the climate signal is not particularly sensitive to any overfitting that may have resulted from this, especially when compared to other decisions such as the climate model selection. QM was not used to bias adjust PET.

ISIMIP3b bias correction method

The second method used in this study is the univariate change preserving bias correction method developed for phase 3b of the Inter-Sectorial Impact Model Intercomparison Project (BCI3, Lange (2019)). A parametric quantile mapping method which approximately preserves the climate change signal in each quantile was applied to each variable independently, with a different distribution used for each variable. Here, we corrected daily average temperature, precipitation and potential evapotranspiration using the normal, gamma and Weibull distributions, respectively. The choices for temperature and precipitation here were motivated by their use in the literature and by Lange (2019), combined with verifying the results. For PET, the Weibull, gamma and beta distributions were fitted to sample grid cells representative of different climates across the UK. While all three would have been adequate, the Weibull distribution was used in the end due to slightly better performance in the sample cells, and, using the Kolmogorov-Smirnov test statistic, it showed a better fit than the beta distribution in most regions of the UK in most months.

For precipitation and potential evapotranspiration, the BCI3-option to apply a separate correction to the probability of occurrence of events beyond thresholds was used to adjust bias and preserve the projected change in the frequency of dry days / no PET (< 0.1 mm/day). For each combination of ensemble member, month and location, the change in the distributions of precipitation, PET and temperature between present-day (calibration) and future periods was preserved by computing the change for 50 quantiles and applying it to the

observed quantiles. This was done multiplicatively (but additively if there are large negative biases in the model data, to avoid obtaining unrealistically large values; Lange, 2019) for precipitation and PET, and purely additively for temperature. For PET and temperature , the transient trend was preserved by removing it before bias correction and adding it back afterwards, as described in Lange (2019).

The code used for this bias correction is version 2.4.1 of the Python code made available in a Zenodo repository by Lange (2020). Version 2.4.1 differs from the originally published method (version 1.0; Lange, 2019) in the equation used for the correction of the frequency of events beyond thresholds (in this case dry days) and code error fixes. The code was modified for this study in minor ways: to allow it to be used when simulations and observations have different calendars (360-day and Gregorian, respectively) and to suit the format of the data used.

The BCI3 method was applied to, and preserves changes between, two 30-year periods (the 1981-2010 reference period and future application periods). This can introduce artefacts such as step changes between adjacent future 30-year application periods and imperfect preservation of the trend within each 30-year application period. In order to produce a continuous, bias corrected 100-year sequence (1981-2080), the algorithm was applied to overlapping periods of 30 years (same length as the reference time period), with the starting point of the application period increasing in decadal steps. Then, the central 10 years (as well as the first and last 10 for the first and last 30-year period, respectively) of each run was extracted and concatenated to obtain the final semi-transient bias corrected time series. Note that the climate change signal in the very first and last decades of the resulting concatenated bias-corrected time series may be slightly stronger or weaker, respectively, because here the beginning and end of the application periods were used instead of the central decade. To estimate the magnitude of the effect of this strategy, the 10-year overlaps of pairs of 30-year bias corrected periods separated by 20 years were used to examine differences in the resulting distributions. The decades 2001-2010 and 2051-2060 (i.e. the earliest and latest of the six decades that were part of three 30-year chunks, respectively) were chosen for making scatter plots and QQ-plots for precipitation and temperature in two grid cells representing different UK climates. For each decade, the three sets of BC data are called ‘tail’, ‘middle’ and ‘head’, depending on whether they are the first, middle or last decade of the 30-year bias correction period, respectively.

Change factor method

The change factor method is commonly used in impact studies (e.g. Prudhomme et al., 2012; Kay et al., 2020). It is included here as a benchmark, for the purpose of demonstrating the limitations of basing impact studies on an observed time series perturbed only by a change in mean climate, and thus show the added value of assessing changes based instead on the RCM output itself. A multiplicative change factor (CF) was computed and applied to the simulated precipitation time series for each month m as follows:

$$CF_{m,p} = \frac{\bar{P}_{raw,FUT,m,p}}{\bar{P}_{raw,REF,m,p}}$$

where $\bar{P}$ is the mean precipitation from the raw UKCP18-RCM data over the future (FUT) or calibration (REF, i.e. reference) period for month m and ensemble member p. Each monthly CF is then applied to the observed calibration period time series to generate the CF precipitation values for each timestep i within month m:

$$P_{CF,i,p} = CF_{m,p} P_{obs,i}$$

3.3 Results and discussion

3.3.1 Evaluation

Bias of raw simulations

The maps in figure 3.3.1 show the ensemble mean errors of the raw UKCP18-RCM projections in the dry-day frequency, mean daily precipitation and the Q95 of precipitation in the reference period, expressed as a percentage of the observed value. In general, the frequency of dry days in UKCP18-RCM is too low (and therefore the wet-day frequency is too high), particularly in the winter and in regions of higher elevation. In summer, the dry-day frequency bias is very small for most of England. The precipitation mean and Q95 are strongly overestimated across the UK in winter, although in highly elevated areas this bias is smaller or even reversed in sign (especially for Q95). In summer, however, the mean and Q95 biases show a strong spatial variability, with underestimations toward the south and at high elevation levels, and a wet bias in the north of the UK.

These seasonal bias differences result in an annual bias of too few dry days almost everywhere, too wet mean precipitation in most regions, and more mixed wet and dry Q95 relative biases.

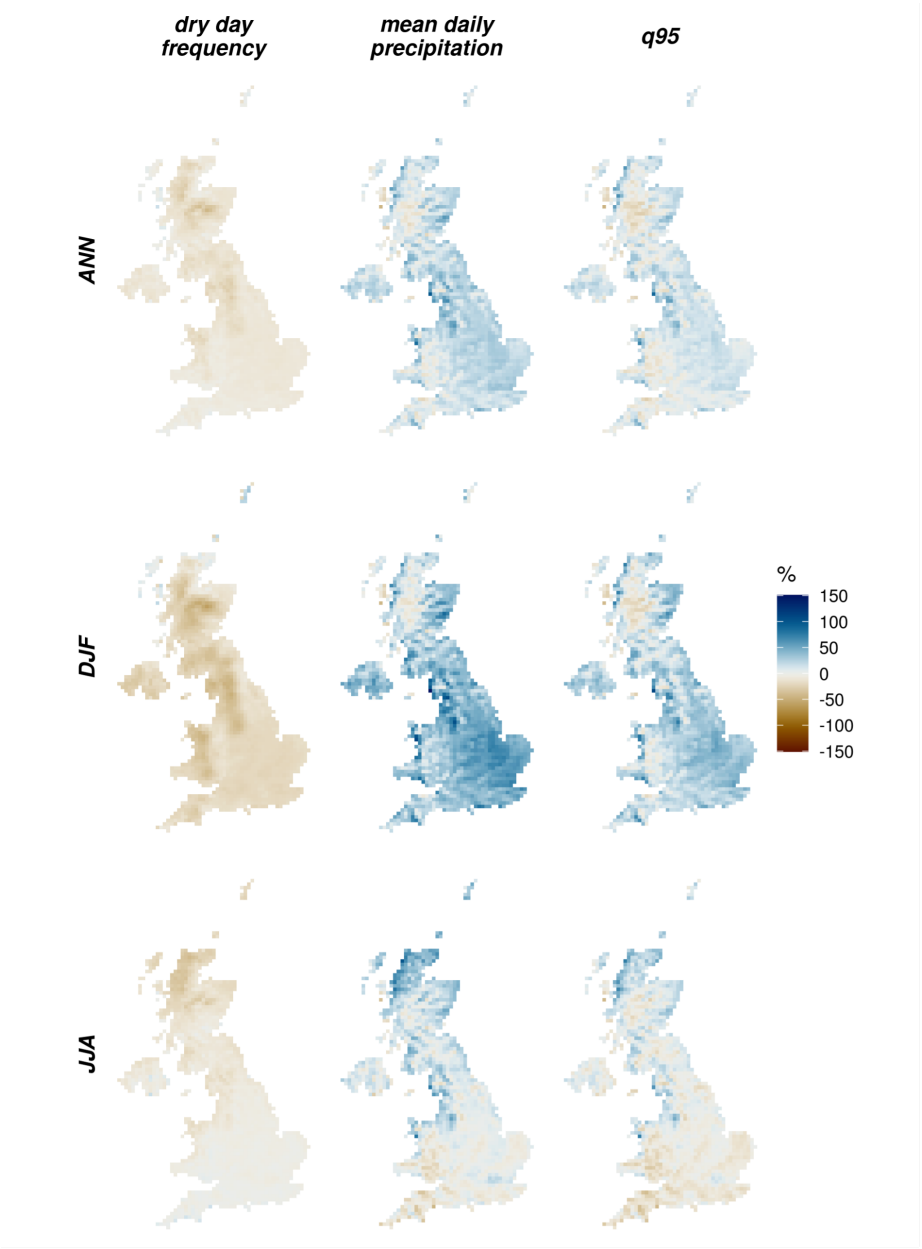

Figure 3.3.1: Precipitation biases in UKCP18-RCM for 1981-2010, expressed as a percentage of the observed values. The percentage bias for each ensemble member was computed and the mean across the ensemble is shown. Dry-day frequency is the percentage of days with P $<$ 1 mm; mean daily precipitation is the precipitation averaged over all days; q95 is the 0.95 quantile of precipitation across all days. Top, middle and bottom rows are for annual, DJF (December, January, February) and JJA (June, July, August), respectively.

The maps in figure 3.3.2 show the ensemble mean errors of the raw UKCP18-RCM projections in the mean daily temperature as well as the cold

(Q05) and hot (Q95) tails of the distribution. At the annual scale, temperature tends to be underestimated in UKCP18 across its distribution, as reflected by the mean, Q05 and Q95. Temperature biases in winter show a north-south gradient, with generally cold biases in the north and warm in the south. Along this gradient, biases are generally colder in the lower tail and warmer in the upper tail of the distribution, indicating an overestimation of temperature variability in winter. In the middle regions along the north-south axis, the biases in the cold and hot tails of the distribution have opposite signs. UKCP18-RCM underestimates temperature on cold days (Q05) in the north especially strongly. In summer, temperature is typically too low across the UK in all three indices considered. One particular deviation from this is the overestimation (or smaller cold bias) in (especially summer) Q95 in major built up areas (e.g. London in South East England). Biases in the representation of urban heat islands were previously documented by Lo et al. (2020), who found that UKCP18-RCM tends to overestimate the intensity of urban heat islands in summer, more so for nighttime than daytime.

The maps in Fig. 3.3.3 show the absolute UKCP18-derived PET biases in mm. Note that these maps do not show Northern Ireland, as this region is not included in the CHESS-PE dataset which served as the observational reference. The predominantly cool bias in the UKCP18-RCM ensemble mean (Fig. 3.3.2) contributes to the low PET bias in some regions and seasons (Fig. 3.3.3). Interestingly, however, the regional and seasonal variations in the PET biases do not closely follow the temperature bias patterns, implying that the bias in PET is not solely caused by errors in the daily average temperature. Note that the strong positive temperature biases shown by some grid cells along the coast in winter come from the use of the regridded 1km HadUK-Grid, which covers the land only, and contrary to UKCP18-RCM does not account for the warmer sea temperature in winter, leading to "observed" grid cell averages that are too low. The opposite but smaller effect occurs in summer, and a similar effect can be observed in the PET bias maps.

Figure 3.3.4 shows the UK average seasonal cycle of the raw and bias corrected precipitation and temperature. For precipitation, the raw UCKP18-RCM ensemble is too wet from November to June, and its range does not encompass the observations during these months. The seasonal timing also appears to be shifted: the driest and wettest months are delayed in the simulations (June and January) compared to the observations (May and October). The ensemble averaged daily mean temperature generally matches observations far more closely than precipitation (although individual members can contain biases of either sign). The largest mean temperature biases over the UK occur from

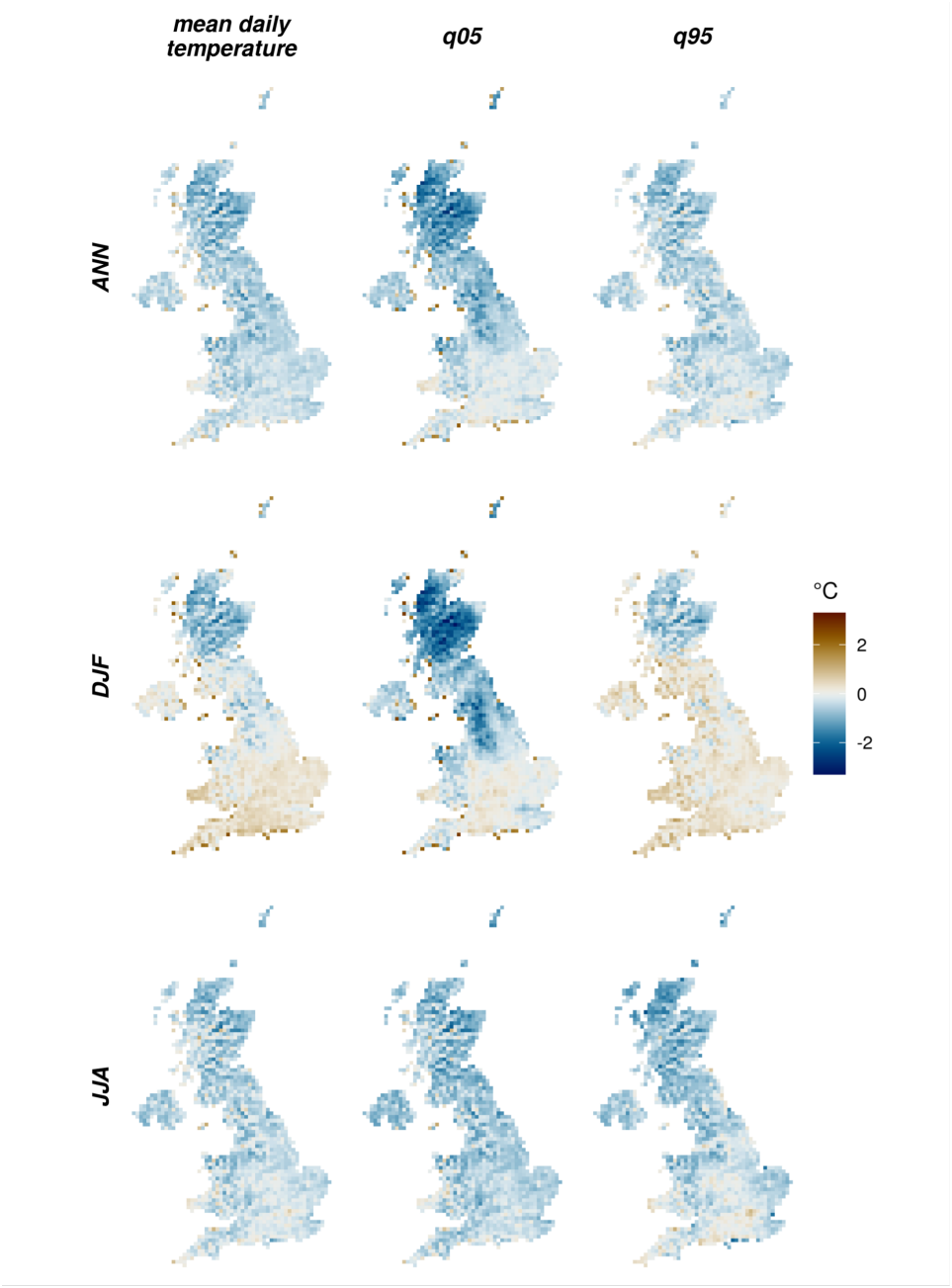

Figure 3.3.2: Temperature biases (°C) in UKCP18-RCM for 1981-2010. The bias for each ensemble member was computed and the mean across the ensemble is shown here. Q05 and Q95 are the 0.05 quantile and the 0.95 quantile across all days, respectively.

March to May, a period during which almost all members are too cold.

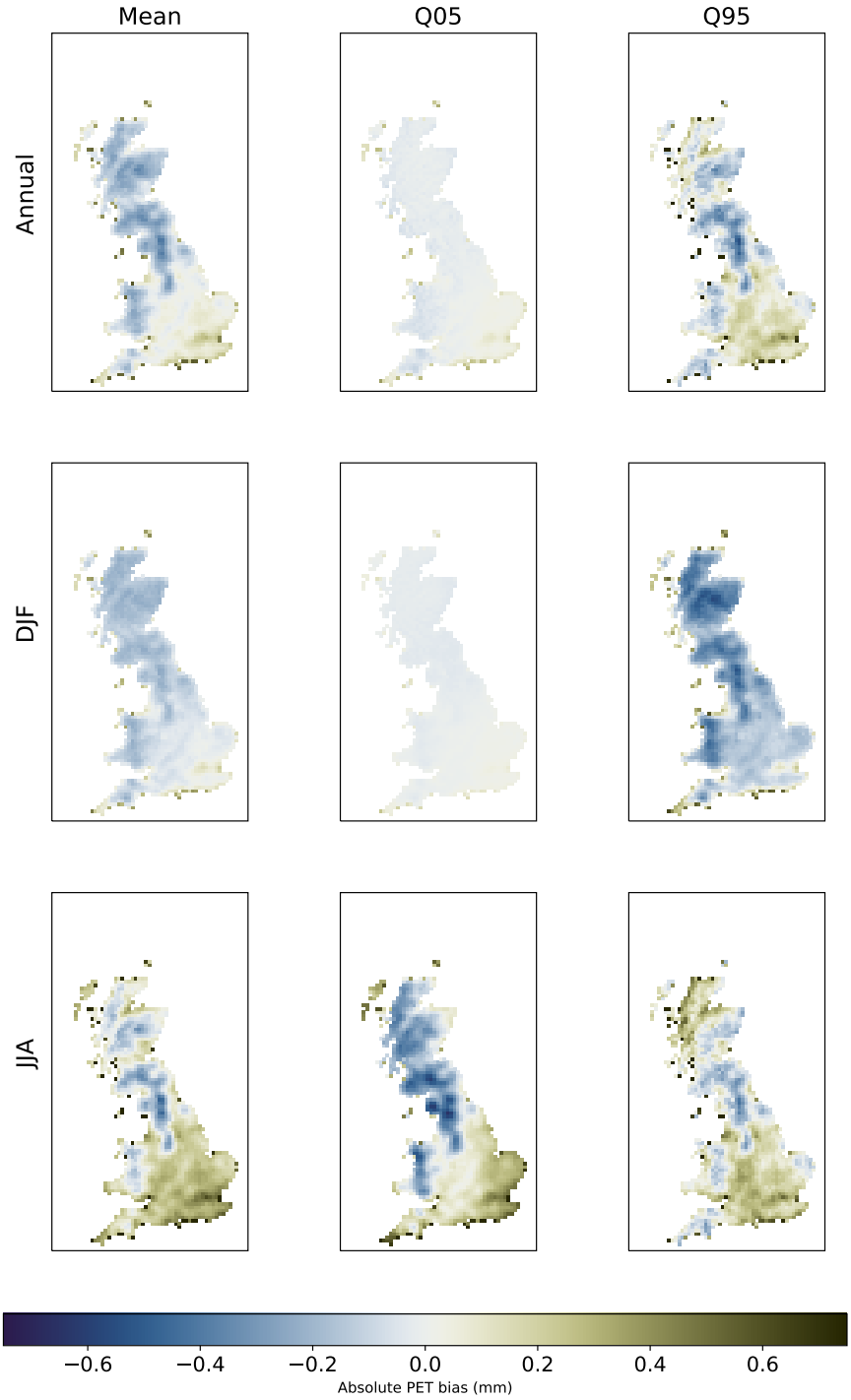

Figure 3.3.3: As Fig. 3.3.2 but for PET (mm/day).

Evaluation of bias correction

The biases of monthly mean temperature and precipitation are effectively corrected by both BCQM and BCI3, with observations very closely matching the processed ensemble members (Fig. 3.3.4). The remaining error in the statistics shown on the maps in Figures 3.3.1, 3.3.2 and 3.3.3 are small (and therefore not shown here, as they take up much space for little information).

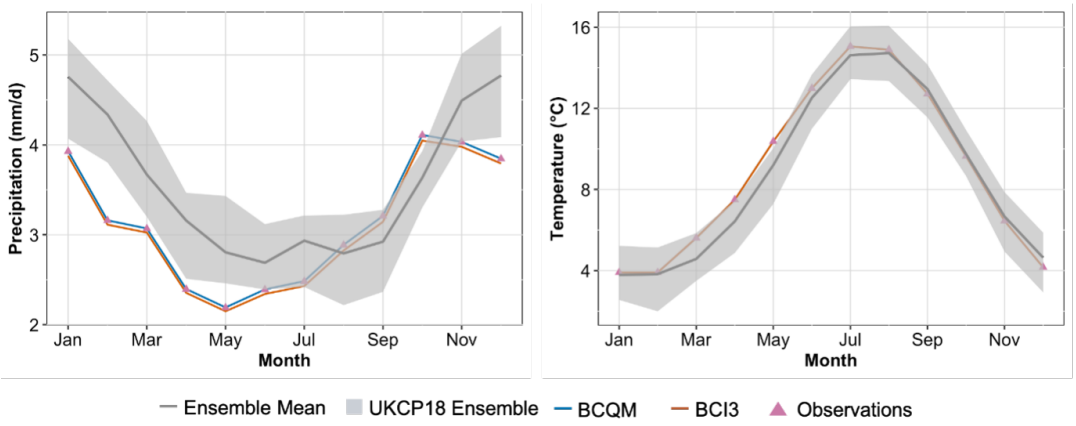

Figure 3.3.4: Comparison between observations and UKCP18-RCM simulations before and after bias correction for monthly precipitation (left, mm/day) and temperature (right, °C) averaged over the UK for 1980-2010. The grey shading shows the spread of the 12-member ensemble, prior to bias correction. The blue line lies underneath the red line in the temperature panel).

Figures 3.3.5 and 3.3.6 show heatmaps of the spatially averaged errors in precipitation and daily mean temperature respectively, before and after bias correction. The value of each metric was calculated for each RCM ensemble member and for the observations, at each grid cell for each year and then averaged over 1981-2010. The heatmaps show the spatial average of the absolute values of the difference between observation and simulation (as a percentage of the observed value for precipitation). For the metrics which are also shown on the maps in Figures 3.3.1 and 3.3.2, the UK-averaged errors are strongly reduced for the full year as well as winter and summer for all ensemble members, resulting in a large decrease in the ensemble mean errors. The standard deviation (sd, bottom rows) of the errors decreases largely as well, as the statistics of the ensemble members converge toward those of the observations for the reference period.

A close look at the figures in Fig. 3.3.5 and 3.3.6 reveals that biases may be smaller in BCQM than in BCI3, but this is not surprising, since empirical quantiles were fitted for BCQM while BCI3 relied on parametric distribution. This should not be interpreted as an advantage of BCQM over BCI3, as the slight edge conferred

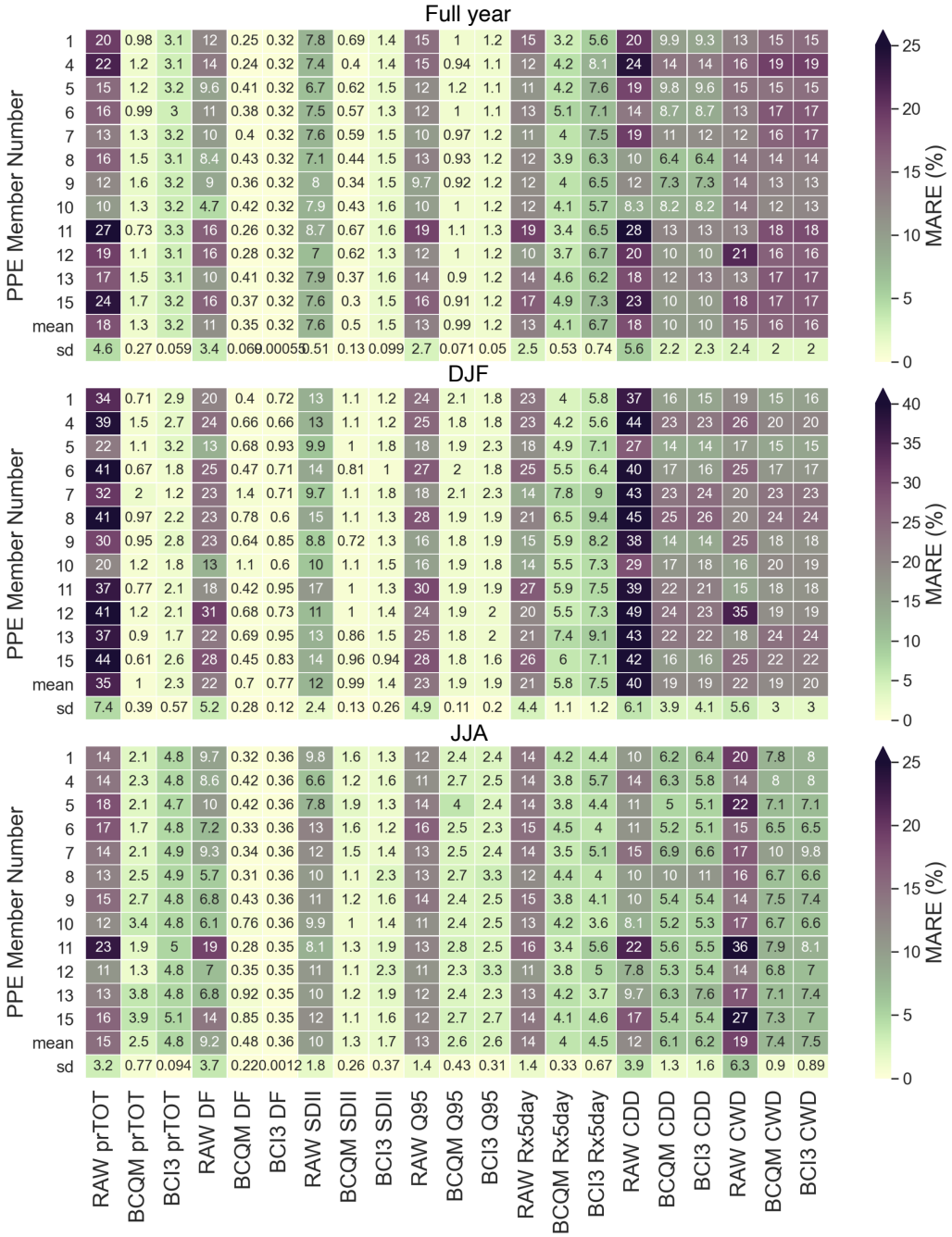

Figure 3.3.5: Mean absolute relative error (MARE, %) for calibration period precipitation, for the UKCP18 RCM simulations before (RAW) and after bias correction (BCQM and BCI3). Values shown are UK-wide averages for 1981-2010. prTOT = annual total precipitation, DF = fraction of days that are dry (P < 1 mm), SDII = mean wet-day precipitation, Q95 = 0.95 quantile of daily precipitation, Rx5day = maximum 5 consecutive-day precipitation, CDD (CWD) = maximum number of consecutive dry (wet) days. The mean and standard deviation across the 12-member RCM ensemble are included at the bottom of each panel.

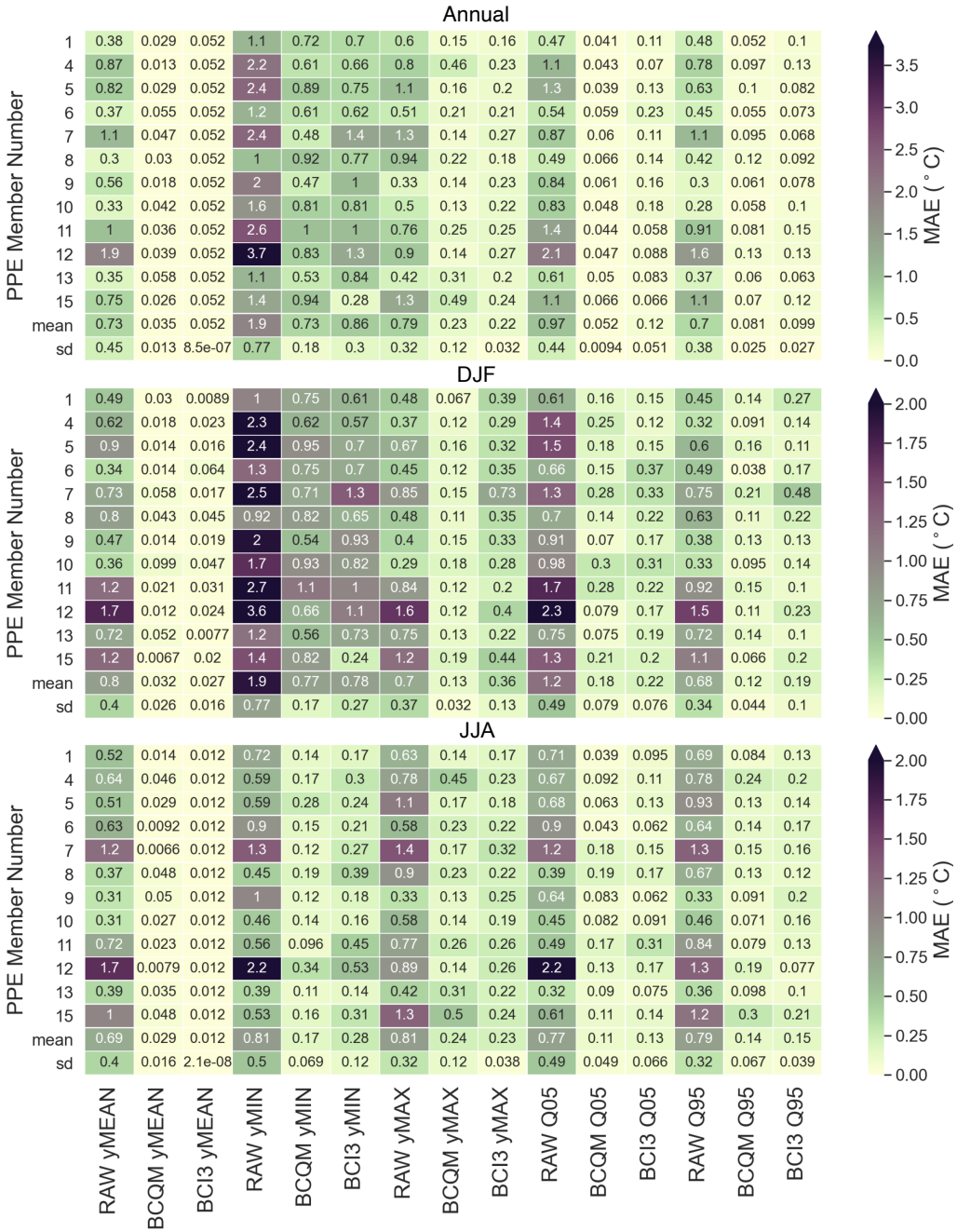

Figure 3.3.6: As Fig. 3.3.5 but for temperature, showing the bias (mean absolute error; MAE; °C) of the mean (MEAN), minimum (MIN), maximum (MAX), 0.05 quantile (Q05) and 0.95 quantile (Q95) of daily temperatures over each year (top), winter (middle) or summer (bottom).

by the non-parametric fit is likely to be lost when the bias-correction is applied to a period not used for calibration (Chen et al., 2019).

For precipitation, BCQM outperforms BCI3 for reducing the bias in the total precipitation (prTOT) (Fig. 3.3.5). Since both methods perform very well for correcting the DF and Q95 bias, the difference in skill for prTOT may be related to the wet end (> Q95) of the precipitation distribution: BCQM (with 1000 quantiles) adjusts the simulated extremes more precisely to fit the observed extremes, while the fit of the gamma distribution used in BCI3 can deviate more from the observations for those extremes. Note that this is not necessarily a flaw in BCI3: BCQM may be undesirably sensitive to the most extreme events occurring in the 30-year observations due to overfitting (Switanek et al., 2017). Due to the strongly skewed nature of precipitation distributions, its extremes (in this case, the tail beyond Q95) make a large contribution to the mean (Pendergrass and Knutti, 2018). Evaluation based on the mean may thus be influenced by how well these extreme quantiles are reproduced. In reverse, this reasoning implies that bias correction relying solely on errors and projected changes in the precipitation mean could be sensitive to the variability in observed and projected extreme events, including relatively rare events that might be included or excluded depending on the choice of calibration period.

Although biases in daily precipitation statistics (TOT, DF, SDII) are largely removed, biases in multi-day metrics (CWD, CDD and Rx5day) are reduced but persist. This is expected (Addor and Seibert, 2014), as both BC methods are designed to correct biases in daily values, but not the temporal structure of the time series. For instance, although the dry day frequency is well corrected by both BC methods, these methods fail to correct errors in the sequences of consecutive wet days in winter and on an annual basis. Figure A.0.1 reveals that this is due to an over-correction, from mostly overestimated toward the east to mostly underestimated toward the west, where the observed CWD is longer. On the contrary, both methods reduced the summer CWD and the CDD in each time period by about half on average.

For temperature, the biases in the mean are almost entirely removed by both methods (Fig. 3.3.6). The ensemble mean Q05 and Q95 biases are strongly reduced, by a factor 7 to 18 for the year and by a factor 3 to 7 for winter and summer. The most persistent biases are found in the winter (and annual) minimum daily temperature (computed for each season and then averaged over the whole period). The remaining winter maximum temperature bias is also markedly higher using BCI3 (about half the raw bias on average) than using BCQM (less than a quarter).

Importantly, across the metrics in Figures 3.3.5 and 3.3.6, the differences between BCQM and BCI3 are generally relatively modest during the reference period. Both methods efficiently reduce biases in single-day metrics, and they mostly struggle with the same metrics, and to some extent, with the same ensemble members.

To conclude this evaluation, Figures 3.3.7 and 3.3.8 assess whether errors were induced by concatenating the centre period of 30-year time series to make BCI3. These figures only show the results for ensemble member 1, as the results were similar across the ensemble. The 'middle' daily values (and quantiles for the QQ-plots) are plotted on the x-axis, and the corresponding 'tail' and 'head' values (quantiles) are plotted on the y-axis. For temperature, the uncertainty introduced by using different periods for change preservation is very small. For precipitation, the match between the quantiles is good (dark circles), although there is some uncertainty introduced for the most intense precipitation events (light circles). The precipitation range covered by these dots in the upper-right corner of the panels highlights the variability within the wettest quantiles, highlighting that when fitting the upper tail of a 30-year period distribution, it is unlikely to perfectly match the upper tail from another period, in particular in presence of climate change. Since some large differences in extreme rainfall can be introduced depending on the period used, the concatenated time series or the data from each 30y period can be used when the focus is on these events, in order to capture the influence of the bias correction. Overall, the differences appear to be small enough to justify concatenation.

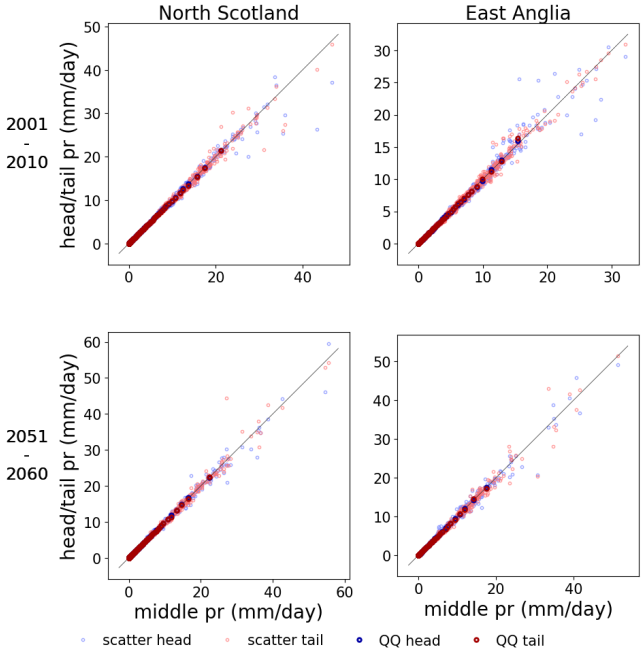

Figure 3.3.7: Quantile-quantile plot (dark colours) and scatter plot (light colours) of precipitation (pr) in two overlapping decades (top row: 2001-2010, bottom row: 2051-2060) for UKCP18-RCM ensemble member 1 for two example grid cells (columns). Tail, middle and head respectively refer to the first, middle and end 10 years of each 30-year period that overlaps in these decades. The line $y{=}x$ is shown for reference.

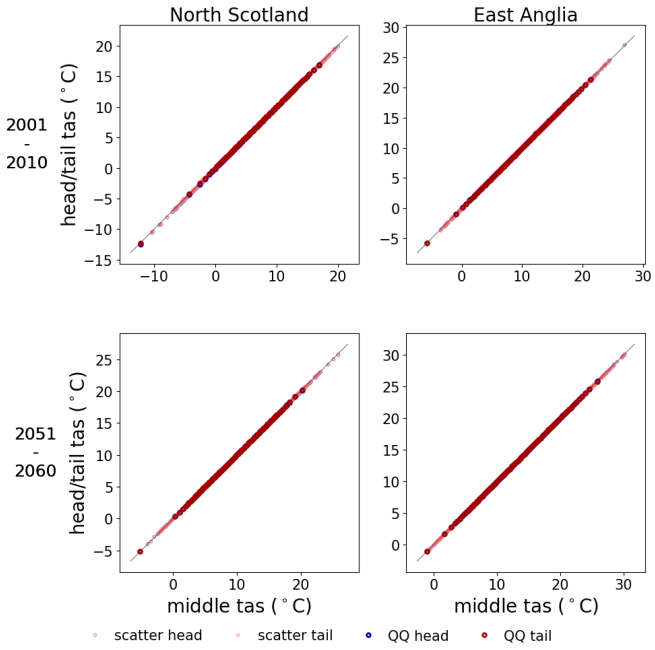

Figure 3.3.8: As Fig. 3.3.7 but for temperature.

3.3.2 Projected changes before and after bias correction

This section analyses the climate change signal (CCS) projected by UKCP18-RCM in the evaluation metrics used above (Section 3.3.1), and particularly considers whether BCQM and the change-preserving BCI3 alter the CSS compared to the raw simulations. Furthermore, for precipitation, these changes are compared to those simulated using the change factor approach. The CCS is defined as the relative or absolute difference between metrics computed over 1981-2010 and 2051-2080. Maps of projected changes in the mean precipitation and temperature (Figures 3.3.9 and 3.3.10) show the spatial variability of the CCS in UKCP18-RCM mean, while the CCS aggregated over the whole domain is summarised for each member and index in heat maps (Figures 3.3.11 and 3.3.12).

The CCS of mean precipitation and temperature in the raw projections (left columns in Figures 3.3.9 and 3.3.10) confirms the conclusions of the UKCP18 headline findings: warmer, wetter winters and hotter, drier summers. The precipitation CCS shows substantial regional variation in all seasons, influenced by topography. The summer drying signal is strongest toward South West England, while the winter wetting signal is weakest in the north-east of Scotland and strongest along the west and south coasts. The seasonally contrasting changes result in an overall precipitation decrease over most of the UK, except along the west coast of north England and Scotland. For temperature, the annual and winter projections are more spatially homogeneous than the summer projections, which show an increasing gradient from the north-west toward the south and south-east.

Overall, these regional variations in the projected changes in mean annual, winter and summer precipitation and temperature over the UK are well preserved in both BCQM and BCI3. In BCQM, the summer temperature increase contains more local variation and is somewhat exacerbated over the regions of higher elevation, and the winter precipitation increase is slightly greater than in the original projections. Spatial patterns in the local variations in the BCQM projected summer temperature changes resemble those in the summer temperature errors (Fig. 3.3.2). The raw changes are better preserved in BCI3, although the added value of this method in terms of change preservation is limited.

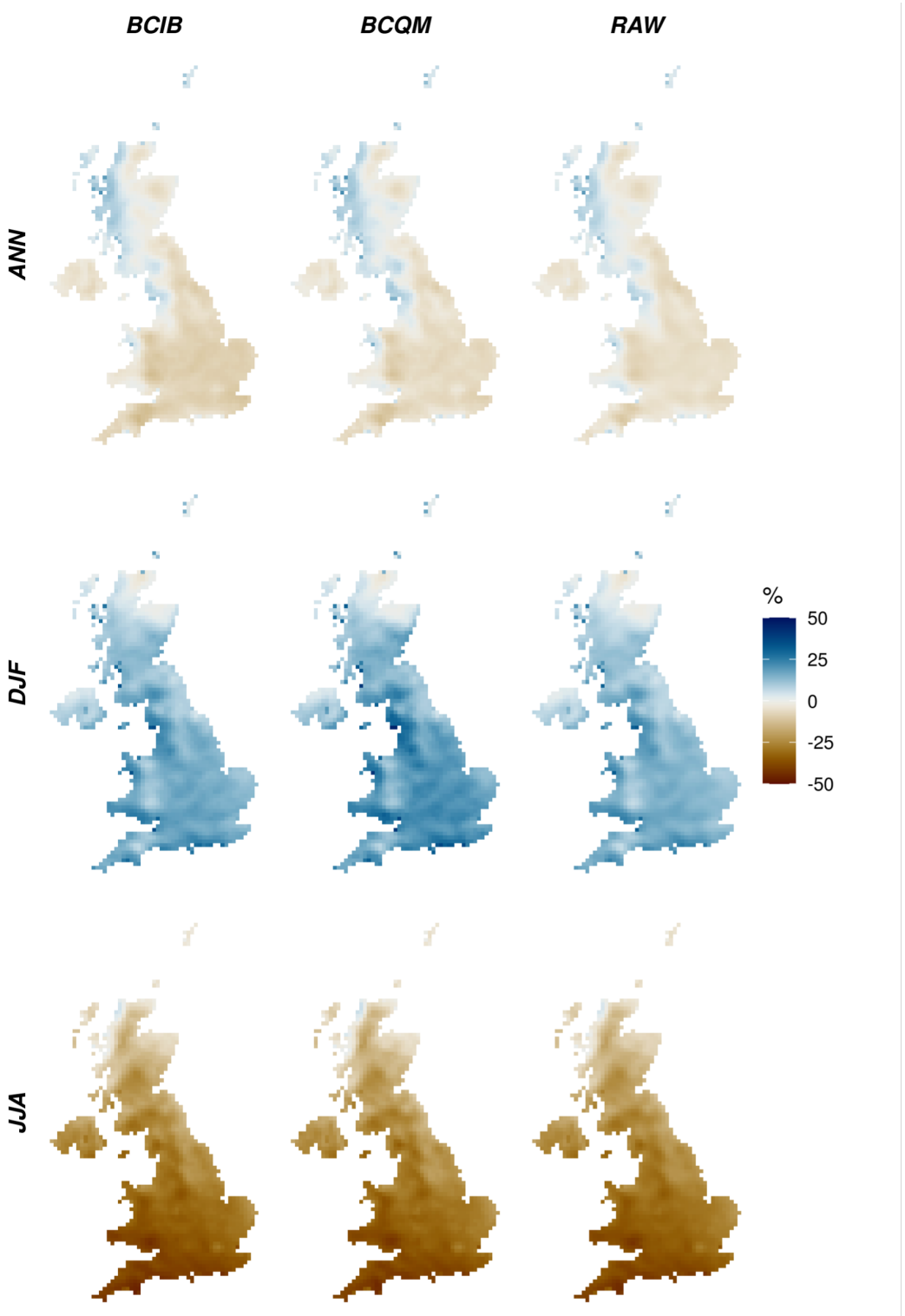

Figure 3.3.9: Projected changes in annual or seasonal mean precipitation from the ensemble mean of UKCP18-RCM simulations before (RAW) and after bias correction (BCQM and BCI3). Values shown are the percentage change from 1981-2010 to 2051-2080 under RCP8.5.

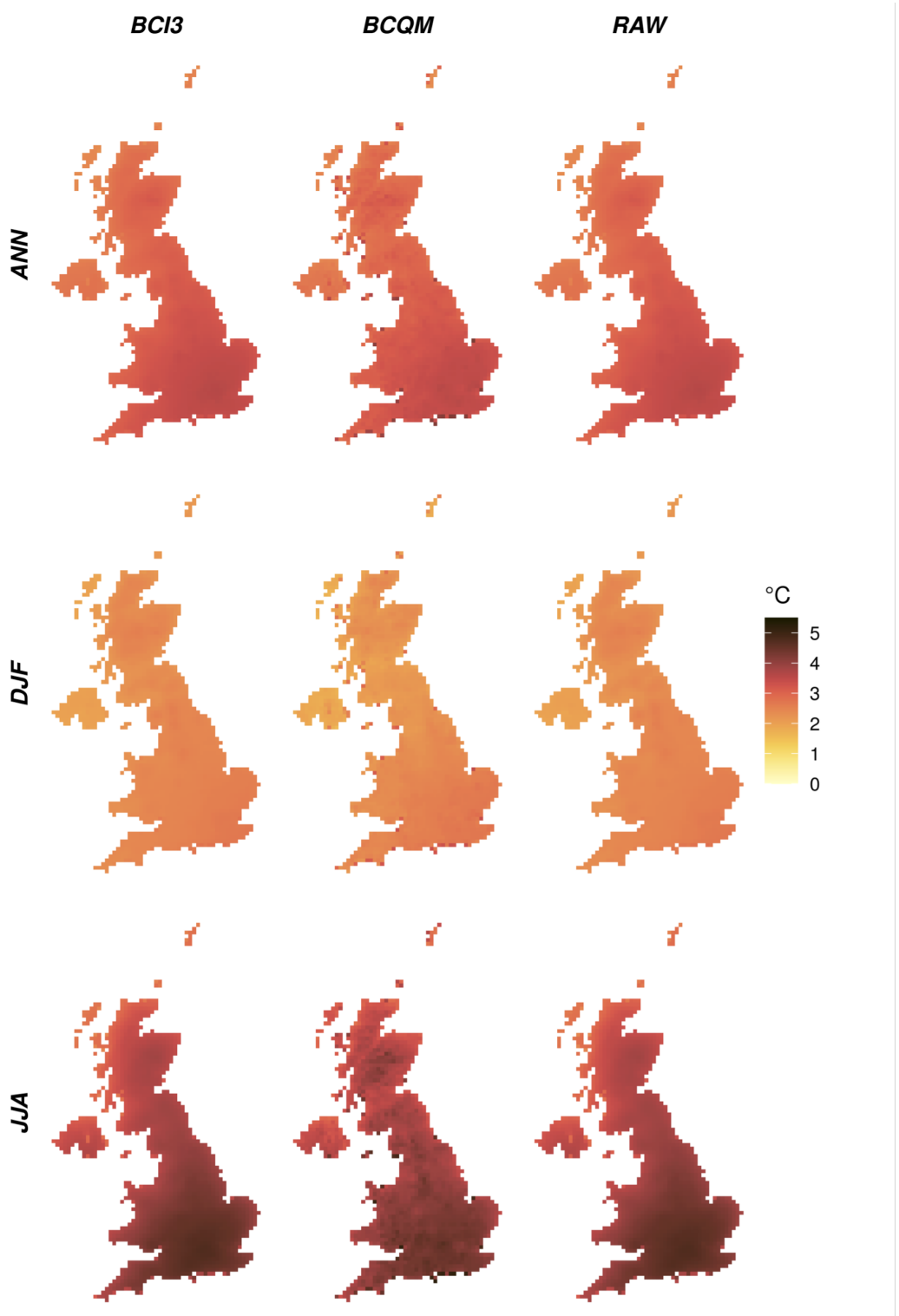

Figure 3.3.10: As Fig. 3.3.9, but for the change in temperature expressed in °C.

The heatmaps in Figures 3.3.11 and 3.3.12 summarize the UK-averaged projected changes in precipitation and temperature metrics, and allow to compare the contributions of i) the different ensemble members and ii) the two bias correction techniques to the uncertainty in the projected changes. Before proceeding to a discussion of the effects of the different bias correction techniques, the projected changes are briefly discussed below (based on the columns labelled 'RAW ...' in the heatmaps).

More so than the mean, precipitation variability has a profound influence on society, and how it is influenced by climate change is thus of great interest. In general, the 'wet' precipitation metrics show wetter (or less dry) projected changes than the total precipitation. The projected total precipitation decrease for summer is the combined effect of an increasing proportion of dry days (DF), and slightly decreasing precipitation falling on wet days, marked by greater relative decreases in the Q95 than in the SDII. The longest summer dry spell is projected to lengthen by a third on average (with wide variation across the ensemble). In winter, the relative increases in prTOT can be attributed to similar increases in SDII Q95 and Rx5day, rather than changes in the fraction of wet days. Interestingly, this is paired with a slight shortening of the longest wet day streak on average. Annually, these seasonal changes combine into a slight projected prTOT decrease and a modest increase in the fraction of dry days, combined with wetter wet days, including in the wet end of the precipitation distribution. The increases in intense 5-day maximum accumulated rainfall could lead to increasing river flooding. This is in agreement with increases in moderate- and high-impact 1-day rainfall threshold exceedances found by Hanlon et al. (2021).

The projected average temperature increases in the UK are greater in summer than in winter. Projected changes in the range of daily mean temperatures are opposite between winter and summer: in summer the Q95 and maximum daily mean temperatures increase more than the Q05 and minimum daily mean temperature, whereas in winter the colder end of the tail is projected to warm more than the warmer end of the tail. There exists a partial overlap between the ensemble members with high summer CDD increases (6, 9, 11 and 13), and the ensemble members with the highest maximum daily-mean temperature increases (5, 8, 9, 11, 13). This might point to limited soil moisture availability during dry spells in members 9, 11 and 13 exacerbating heatwaves during possible compound hot-dry events; however as only average statistics are shown here, further research would be needed to confirm this. In line with the projected increases in minimum temperatures, Hanlon et al. (2021) found a continued decreasing number of days with minimum and maximum

temperatures below 0°C (frost and icing days respectively). On the warm end of the distribution, the larger increases in the expected maximum mean daily temperatures match their findings of increasing days with maximum temperatures exceeding 25°C, and increasing nights with minimum temperatures over 20°C. Similarly, Arnell et al. (2021) found increases in the annual probability of experiencing at least one heat wave (based on regionally varying thresholds). Concerning adaption, the findings from these studies and ours thus imply a potential shift in health care and housing concerns from heating in winter toward cooling in summer (Hanlon et al., 2021; Arnell et al., 2021).

The differences between the rows for the different metrics and columns for bias correction methods in the heatmaps (Figures 3.3.11 and 3.3.12) show a wide range of projected changes among the ensemble members, which usually exceeds the differences from BC method. In other words, at least when aggregated over the UK, the differences between the ensemble members are much greater than between the two bias-correction methods. This is also summarised well by two last rows of each heat map panel for each metric: the standard deviation (summarised by the 'sd' row) across the ensemble is typically much larger than the difference between the BC methods in the 'mean' row. While over the reference period the BC significantly reduces the spread of the biases among the ensemble members (see row 'sd' in Figures 3.3.5 and 3.3.6), the spread in the projected changes is quite well retained by both BC methods (same row in Figures 3.3.11 and 3.3.12).

Although the differences in CCS between the two BC methods are relatively minor when compared to the differences between ensemble members for the same metric, there is a subtle added value of BCI3 over BCQM for preserving the changes, which is generally more visible looking at the individual ensemble members than in the ensemble means. It is worth noting that BCQM already largely preserves the projected changes, even though it was not explicitly designed to do so, which limits the potential benefit of even a perfect change preservation over this method. For temperature, the heatmaps in Fig. 3.3.12 show a limited added value of BCI3 compared to BCQM. This is in part due to the spatial averaging that took place in order to produce the heatmaps. Maps of the climate change signal for each month and ensemble member for different quantiles (not shown) and Fig. 3.3.10 show that BCQM modifies the projected temperature climate change signal differently in different regions of the UK, in particular over significant topography; this is not the case for BCI3, which better preserves the raw CCS spatial patterns. Finally, the differences in climate change signal between both BC methods are largest for several precipitation indices in winter, where BCI3 slightly improves the preservation of

the projected change. Interestingly, for many of the temperature indices and the winter precipitation indices, the spread of projected changes is also better preserved in BCI3 than in BCQM (compare the ensemble 'sd' values).

Lastly, the heatmaps in Fig. 3.3.11 highlight the pitfalls of the change factor (CF) approach for UKCP18-RCM. There are large differences between the CF approach and the raw (or bias corrected) projections when looking at the CCS in precipitation metrics beyond the seasonal or annual totals. The projected lengthening of the longest annual or summer dry day sequences and shortening of the longest annual, winter or summer wet day sequences in all ensemble members, which is well preserved by both BC methods, is largely or almost entirely disregarded using the CF approach. Most of the projected change in CDD and CWD can thus not simply be attributed to changes in the mean. The smaller projected decrease in the summer maximum 5-day accumulated precipitation (Rx5day) compared to the total precipitation decrease is not captured by the CF approach. Worse, at the annual scale the CF approach projects a decrease in the Q95 (similar to the prTOT changes), whereas the raw UKCP18-RCM projects an increase on average. At the annual scale, the CF approach is unable to capture the simulated change towards fewer (DF increase) but wetter (SDII increase) wet days, the combined effect of which is an overall decrease in total precipitation. This inability of the CF approach to account for changes in the variability of precipitation severely narrows its suitability for climate impact modelling studies on e.g. floods or droughts, for which the temporal variability and changes to the precipitation distribution are highly relevant.

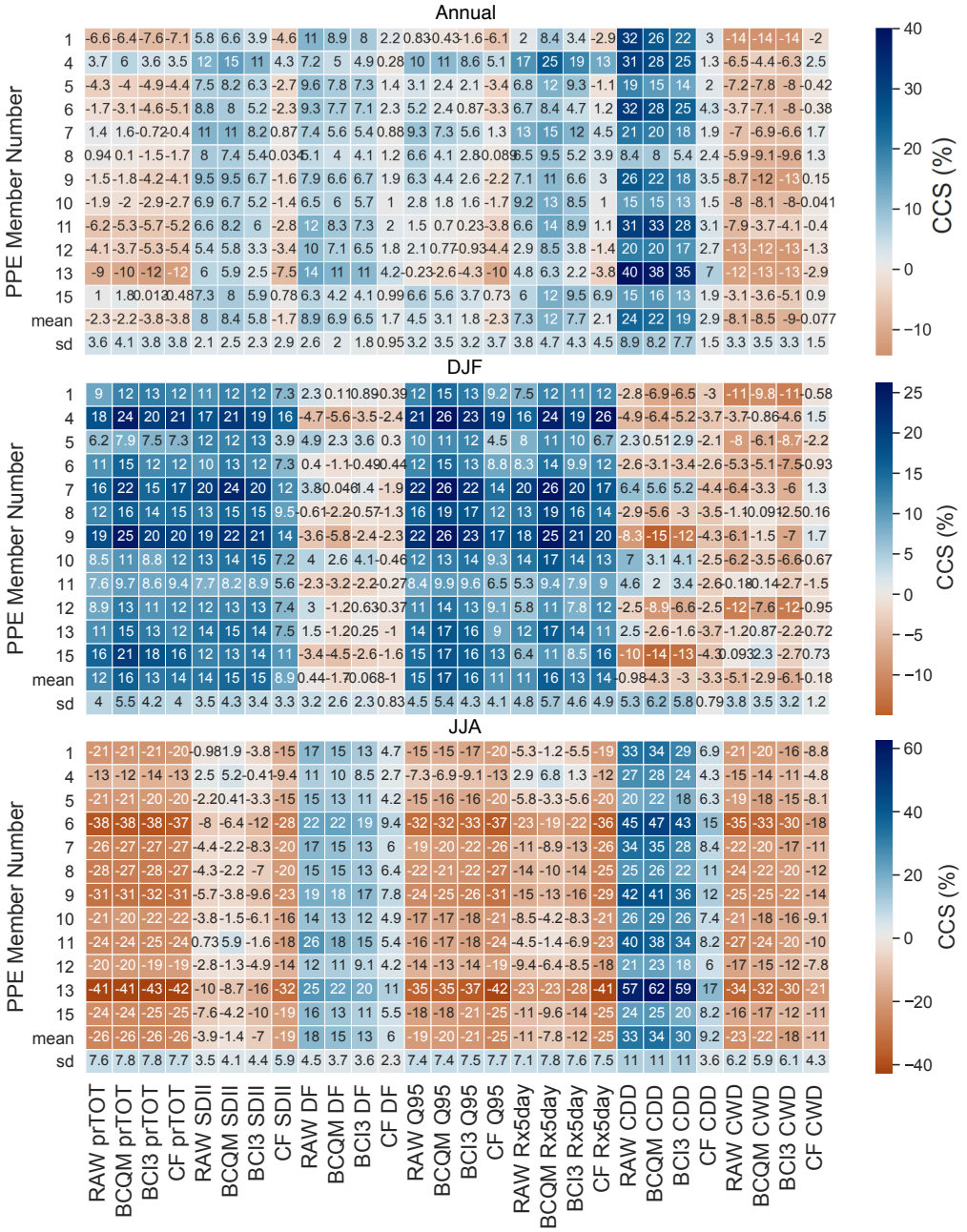

Figure 3.3.11: Projected changes (climate change signal; CCS) in precipitation characteristics in the ensemble of UKCP18-RCM simulations before (RAW) and after bias correction (BCQM and BCI3) and after applying a change factor (CF) to the observed time series. Each indicator shows the spatially average (UK-mean) of the changes by 2051-2080 compared to 1981-2010, expressed as a percentage of the observed values for 1981-2010. Statistics shown are the same as in Fig. 3.3.5.

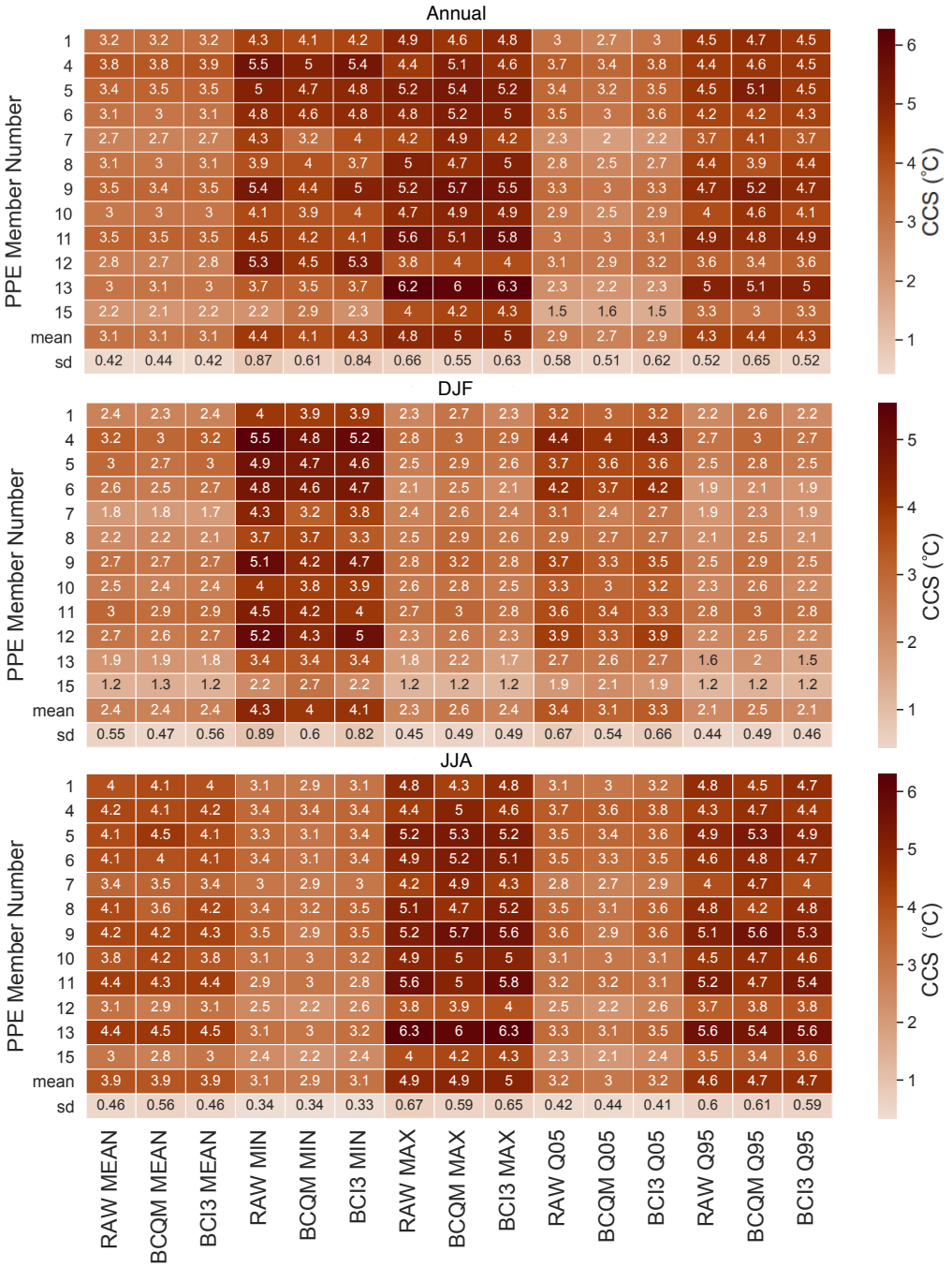

Figure 3.3.12: As Fig. 3.3.11 but for temperature characteristics and actual (°C) rather than percentage changes are shown. Statistics shown are the same as in Fig. 3.3.6.

3.4 Conclusions

3.4.1 General conclusions

Substantial biases in temperature, potential evapotranspiration and precipitation statistics were found in the UKCP-RCM simulations. To improve their usability for impact modelling, two bias correction methods were applied to all ensemble members of the UKCP-RCM perturbed physics ensemble: the widely used quantile mapping (BCQM) and the bias correction method developed for phase 3 of the ISIMIP project (BCI3), which is designed to preserve the climate change signal. Both methods successfully reduced errors in daily temperature and precipitation metrics, and reduced errors in multi-day precipitation metrics to a lesser degree. Both methods also satisfactorily conserved the climate change signal as well as the spread among the projections, with a minor improvement in BCI3 compared to BCQM. For potential evapotranspiration, similar analyses (figures not included here) show that BCI3 successfully reduced errors and conserved the climate change signal. Analysis of projected changes in temperature and precipitation metrics suggests a higher likelihood of extreme weather (hot, dry or wet), and confirms the headline findings of projected hotter, drier summers and warmer, wetter winters.

3.4.2 Recommendations for users of the bias corrected datasets

Potential users of these bias-corrected simulations are encouraged to consider the following points.

1. Both bias-corrected datasets may be of interest for users whose impact model is affected by the biases in the UKCP18-RCM simulations and whose application requires changes in the precipitation distribution and temporal variability to be captured as well as the mean.

2. Caution is needed with coastal grid cells, as the 1km data set used as reference for the bias correction only covers the land, which results in biases in grid cell averages along the coast (especially for temperature). This should not be an issue for users only interested in land temperature, but we recommend that users consider the fraction of each grid cell covered by land before using the data.

3. At the regional scale, the climate change signal is slightly better preserved by BCI3 (especially for temperature), so for this it is recommended above BCQM. Although the differences between BCI3 and BCQM are small, they

might grow or shrink after propagation in impact models, so it may be valuable to use both, as a limited way of sampling the uncertainty due to statistical postprocessing. However, the difference between BCQM and BCI3 is unlikely to be the greatest source of uncertainty.

4. Using the change factor method for precipitation is discouraged for applications where changes to the temporal variability and full distribution is important.

5. Potential users of these bias-corrected datasets are urged to consult the UKCP18 user guidance published by the Met Office, in particular guidance on bias correction (e.g. Fung et al., 2018). Since the uncertainty among the UCKP18-RCM ensemble members is large for changes in precipitation and temperature metrics studied here, it is recommended to employ multiple ensemble members to sample this uncertainty.

6. Users are encouraged to perform their own evaluation of these datasets to ensure that they are adequate for their planned use.

7. As discussed in the introduction, all bias-correction methods rely on strong assumptions about climate projections, which users of any bias-corrected climate projections should keep in mind when using the data and interpreting the resulting impacts.

4

Projected changes in droughts and extreme droughts in Great Britain strongly influenced by the choice of drought index

---

Synopsis

This work was published in HESS as a first-author paper (Reyniers et al., 2023).

Droughts cause enormous ecological, economical and societal damage, and are already undergoing changes due to anthropogenic climate change. The issue of defining and quantifying droughts has long been a substantial source of uncertainty in understanding observed and projected trends. Atmospheric-based drought indicators, such as the Standardised Precipitation Index (SPI) and the Standardised Precipitation Evapotranspiration Index (SPEI), are often used to quantify drought characteristics and their changes, sometimes as the sole metric representing drought. This study presents a detailed systematic analysis of SPI- and SPEI-based drought projections and their differences for Great Britain, derived from the most recent set of regional climate projections for the UK. We show that the choice of drought indicator has a decisive influence on the resulting projected changes in drought frequency, extent, duration and seasonality by 2 °C and 4 °C above pre-industrial levels. The increases projected in drought frequency and extent are far greater based on the SPEI than based on the SPI. Importantly, compared to droughts of all intensities, isolated extreme droughts are projected to increase far more in frequency and extent, and show more pronounced changes in the distribution of their event

durations. Further, projected intensification of the seasonal cycle is reflected in an increasing occurrence of years with (extremely) dry summers combined with wetter than average winters. Increasing summer droughts also form the main contribution to increases in annual droughts, especially using SPEI. These results show that the choice of atmospheric drought index strongly influences the drought characteristics inferred from climate change projections, comparable to the uncertainty from the climate model parameters or the warming level, and therefore potential users of these indices should carefully consider the importance of potential evapotranspiration in their intended context. The stark differences between SPI- and SPEI-based projections highlight the need to better understand the interplay between increasing atmospheric evaporative demand, moisture availability and drought impacts under a changing climate. The region-dependent projected changes in drought characteristics by two warming levels have important implications for adaptation efforts in GB, and further stress the need for rapid mitigation.

4.1 Introduction

Anthropogenic climate change is already affecting the frequency and intensity of droughts on all continents, through increases in atmospheric evaporative demand and in some regions, also through precipitation (Seneviratne et al., 2021). How much larger these changes become depends on current and future emissions, and understanding the impact of climate change on droughts is crucial importance given the serious ecological and socio-economic damage these events can inflict. However, quantitatively assessing changes to droughts is complicated by the difficulty of defining and quantifying droughts (Yevjevich, 1967). Distilled to its most simple form, a drought can be defined as a deficit of water relative to normal conditions (Sheffield et al., 2012). As this generalised definition is not very helpful for assessing drought hazards (Lloyd-Hughes, 2014), different types of drought are typically recognised, based on the context and the moisture quantity in which the deficit takes place (Wilhite and Glantz, 1985). A meteorological drought, indicating a period of below-normal precipitation, can develop into a soil moisture drought, also called agricultural drought due to its relevance for crop growth. These conditions can develop into low flows in rivers or low water levels in lakes, called hydrological drought (of which groundwater drought can be considered a sub-type).

Drought indices, of which a large number can be found in the literature (e.g.

Keyantash and Dracup, 2002), are frequently used to quantify different types of drought conditions. While indicators exist for variables relevant to different drought types, drought indices that only rely on atmospheric data are a popular choice due to (historical) data availability and due to their ease of use (they do not require the deployment of an impact model, such as a hydrological model). The Drought Severity Index (DSI; Phillips and McGregor, 1998), for example, uses precipitation only and has been used in previous studies on the impact of climate change on drought in the UK (e.g. Blenkinsop and Fowler, 2007; Rahiz and New, 2013; Hanlon et al., 2021). One of the most widely used drought indicators is the Standardised Precipitation Index or SPI (McKee et al., 1993), a precipitation-based index recommended by the World Meteorological Organisation (WMO and GWP, 2016). It is one of the indicators shown in the UK Water Resources Portal (https://eip.ceh.ac.uk/hydrology/water-resources), and has been used in earlier work on drought under climate change in the UK (e.g. Vidal and Wade, 2009; Arnell and Freeman, 2021). Since the introduction of the SPI, other standardised indicators have been developed that apply the standardisation principle of SPI to different (combinations of) drought-relevant variables. This includes the Standardised Precipitation Evapotranspiration Index (SPEI; Vicente-Serrano et al., 2009), which gives the anomaly in a simple climatic water balance, computed as the difference between precipitation and potential evapotranspiration (PET). This indicator was developed to be sensitive to the effect of global warming induced increases in atmospheric evaporative demand (AED), the potential of the atmosphere to evaporate water (depending on radiation, temperature, humidity and wind speed; Robinson et al., 2017). High AED can aggravate the effects of sustained precipitation deficits and accelerate drought development (e.g. Manning et al., 2018; Bloomfield et al., 2019; Pendergrass et al., 2020). Contrary to the SPI, SPEI is thus not purely an indicator of meteorological drought, but instead an atmospheric-based index that is "mostly related to the actual water balance in humid regions", reflects "an upper bound for overall water-balance deficits" during dry periods and in water-limited regions, and is also linked to vegetation stress (Seneviratne et al., 2021). These atmospheric-based indicators are widely used in climate change impact studies, although the consequences of their implicit assumptions with regards to evaporative stress are not always expressly considered. Along similar lines, a study by Satoh et al. (2021) found that, if the drought type is considered as a source of uncertainty for projections of future droughts, it constitutes a major one in many parts of the world.

This study focuses on Great Britain (GB) to compare projected drought changes as quantified using the SPI and SPEI. Despite not typically being thought of as a particularly drought-prone area, GB has experienced several droughts in the past

which lead to widespread impacts, including impacts on ecosystems (including algal blooms and fish kills), agriculture and domestic water supply (Rodda and March, 2011; Kendon et al., 2013; Turner et al., 2021). The impacts of climate change on future droughts in the UK is therefore a key concern for stakeholders including water managers and farmers (e.g. Watts et al., 2015). In this study, we aim to answer the following questions.

1. Based on atmospheric-based standardised drought indices, how are drought and extreme drought frequency, duration, extent and seasonal timing expected to change under different global warming levels?
2. What is the potential contribution of changes in PET and precipitation to the changes in these drought characteristics?
3. How sensitive are the quantified projected changes in drought characteristics to the choice of atmosphere-based drought indicator, and how does it compare to other sources of uncertainty?

To this end, we identify and characterise droughts and their projected changes in the most recent ensemble of regional climate projections for GB, using both SPI and SPEI (hereafter, SI for standardised indicators). We compare projected drought characteristics for both indices, to identify the potential role of changing PET. Although previous studies have compared historical and projected changes using these SI in different regions of the world (e.g. Stagge et al., 2017; Chiang et al., 2021), this study adds a new level of detail by an in-depth analysis of different drought characteristics and attention to within-GB regional differences, and is the first to use UKCP18 with these SI to assess projected changes in drought characteristics for GB. This helps further understand the potential future changes in the nature of GB droughts depending on global warming, and demonstrates the importance of the drought index choice for climate change impact studies and stakeholder usage.

4.2 Data

4.2.1 Observations

Datasets of PET and precipitation observations were needed for evaluation, bias correction of the UKCP18-RCM, calibration of SI and calculation of historical SI. The CHESS-PE (Robinson et al., 2020) and HadUK-Grid (Hollis et al., 2019) datasets were used for PET and precipitation respectively, using the following

time periods: 1961-2010 for the SI calibration (see Section 4.3.4), 1981-2010 for the bias correction, and 1981-2005 for comparison to the reference period UKCP18-data in this study. Both datasets were first regridded from their native 1km resolution to the 12km resolution grid of the UKCP18-RCM, by averaging of the 1km grid cells falling in each 12km cell. A land fraction was obtained based on the proportion of 1km grid cells with observations on land within each 12km grid cell, and used to exclude grid cells with a land fraction lower than 50% from the analysis. As no observation-based PET was available for Northern Ireland in CHESS-PE, this region was excluded from our study. The method used to obtain PET in the production of CHESS-PE is an implementation of Penman-Monteith PET for a reference grass crop (Allen et al., 1998), in which the calculation of vapour pressure deficit from temperature is based on Richards (1971) (Robinson et al., 2017).

4.2.2 UKCP18 regional climate projections

UKCP18 is the most recent set of national climate projections for the UK and has been produced by the Met Office Hadley Centre (Murphy et al., 2018). This study makes use of its third strand, produced with the aim of providing a range of storylines to support adaptation efforts in the UK: a perturbed physics ensemble (PPE) of regional climate projections (UKCP18-RCM; Met Office Hadley Centre, 2018), available from the Centre for Environmental Data Analysis. This ensemble of 12 simulations was constructed by dynamically downscaling global HadGEM3-GC3.05 simulations through one-way nesting with the same model at finer resolution. At both resolutions, HadGEM3-GC3.05 was perturbed in 47 parameters spread over model representations of convection, gravity wave drag, boundary layer, cloud, large-scale precipitation, aerosols, and land surface interactions (Murphy et al., 2018). The ensemble thus does not sample GCM-RCM structural uncertainty, only parameter uncertainty, and was designed to cover a range of possible futures. While multiple GCM-RCM structures would add another interesting dimension to the study, expanding the ensemble was outside the scope and capacity of the study. The horizontal resolution of the RCM simulations is 12km over GB (available on OSGB36 grid projection). As droughts tend to be more spread out in space and time, we judged that the 12km daily resolution of the UKCP18 RCM poses a better trade-off between practicality and spatiotemporal detail than the higher-resolution convective permitting simulations for this study. Simulations of different variables are available from 1 December 1980 to 30 November 2080 on a daily time step (for practical reasons, December 1980 was left out of our analysis).

4.3 Methods

4.3.1 Calculation of potential evapotranspiration

While AED increases with rising temperatures, changes in humidity, net radiation and wind speed can also play a significant role. Therefore, we represented AED by PET calculated using Penman-Monteith, which includes the effect of all these variables. This method leads to a more robust correlation between the resulting SPEI and soil moisture under a warming climate compared to using the temperature-only Thornthwaite method (Feng et al., 2017) and is recommended over simpler temperature-based methods (e.g. Dewes et al., 2017), however it is still subject to significant limitations (Milly and Dunne, 2016; Greve et al., 2019). The calculation of PET for the UKCP18-RCM follows the same variant of the Penman-Monteith method used by Robinson et al. (2017), to ensure consistency with CHESS-PE. It uses these variables simulated by the UKCP18-RCM ensemble: specific humidity, pressure at sea level, net downwelling longwave radiation, net downwelling shortwave radiation, wind speed at 10m and daily average surface air temperature. PET was set to zero wherever a calculated value was negative (which occurred for less than 1% of the values overall and, when split by ensemble member and month, also less than 1% for all cases except December in ensemble member 1 with 1.2% of negative values).

Detrending temperature

To investigate the influence of the projected temperature trend on changes in SPEI-based droughts and the deviation of SPEI from SPI, we also computed an alternate version of SPEI projections ($\text{SPEI}_{dtr-tas}$) using a detrended version of UKCP18-RCM temperature. For this, a linear trend was fitted to, and subsequently subtracted from, the simulated temperature time series for each grid cell and month separately. This detrended temperature dataset was used to compute PET as described above, resulting in a $\text{PET}_{dtr-tas}$ variable in which any trend left is due to trends in other variables (specific humidity, radiation, wind speed and pressure) or in interactions between variables. As these variables are closely intertwined in the climate models, this unavoidably introduces a physical discrepancy between temperature and the other variables used in the PET calculation. This is taken into account in the interpretation.

4.3.2 Bias adjustment

As comparison to observations revealed significant bias in the simulation of both precipitation and PET (see Figs. B.0.1 and B.0.2), these variables were statistically post-processed using the ISIMIP3b change preserving bias adjustment method (Lange, 2019) version 2.4.1 (Lange, 2020). The biases we observed for different quantiles were not equal to the biases observed in the mean, which is why we selected a bias adjustment method that took this into account. Similarly, biases also varied between months and locations, so the bias adjustment needed to be specific for each month and grid cell. The ISIMIP3b bias adjustment method is based on quantile mapping, but also preserves projected changes in the variables being corrected, and enables separate adjustment of the frequency of dry days – a desirable feature for drought research. For precipitation, the gamma distribution and mixed additive/multiplicative per-quantile change preservation were used. For PET and $PET_{dtr-tas}$, the Weibull distribution, detrending and mixed additive/multiplicative per-quantile change preservation were used. A dry threshold of 0.1 mm day^{-1} was selected below which there is considered to be no precipitation or PET. In what follows, UKCP18-RCM indicates the bias adjusted data.

4.3.3 Time slice selection

The UKCP18-RCM simulations used in this study are available for the RCP8.5 emissions scenario, and the models used have high global climate sensitivity compared to the CMIP5-ensemble and the probabilistic projections (Murphy et al., 2018). Therefore, to assess the impact of climate change on drought characteristics in scenarios with lower climate sensitivity and more mitigation (resulting in lower warming levels above pre-industrial times), a time slice approach was implemented to investigate changes at two specific global mean warming levels. A common fixed reference period (1981-2005) was used for all ensemble members to compare to these future time slices and observations. For each ensemble member, a time slice was selected from 12 years before to 12 years after the year in which the centred 25-year rolling mean global temperature exceeds + 2 °C and + 4 °C above pre-industrial levels (defined as 1850-1900) in the driving global model (see Table 2 in Gohar et al. (2018)). As opposed to the fixed reference period, the time periods used to represent different levels of warming are thus different for each ensemble member, depending on when their global driving models reach +2 and +4 °C above pre-industrial levels. Both warming levels are reached in all 12 ensemble

members, however for ensemble member 8 the time slice representing +4 °C is cut short 2 years by the end of the simulated period.

This approach would result in an accurate assessment of changes in GB drought projected at these warming levels if these changes would scale directly with global temperature increase (independent of the speed of change), and if the regional model has the same climate sensitivity as its driving global model. Neither of these requirements are likely to be fully met. UKCP18-RCM projects slightly weaker UK temperature responses towards the end of the simulated period than the driving global simulations (Fig. 5.2 in Murphy et al. (2018)). Also, midlatitude atmospheric circulation patterns in the selected time slices (which influence UK weather and therefore drought events) may respond to a higher level of radiative forcing than the global temperature increase levels used to select them (Ceppi et al., 2018). Nevertheless, the applied time slice approach is a reasonable approximation, and frequently used for investigating impacts at different levels of global warming.

4.3.4 Drought and aridity indicators

While drought refers to a period of below-normal water availability for a given context, aridity refers to the climatic average moisture availability (Dai, 2011). This is included in this study in order to help establish an understanding of the mean climatic changes projected for precipitation and PET in UKCP18-RCM, before proceeding to assessing projected changes in drought characteristics. To this end, the aridity index (AI) was calculated as the annual average ratio of precipitation to PET (e.g. United Nations Environment Programme, 1992; Feng and Fu, 2013; Greve et al., 2019), which is more intuitive to interpret than the standardised indicators. For time slices of 25 years, this gives:

$$AI = \frac{1}{25} \sum_{y=1}^{25} \frac{Precipitation_y}{PET_y}$$

The drought indices compared in this study are SPI and SPEI. Both are widely used in the literature to quantify droughts, and they imply contrasting assumptions of the surface water balance: for SPI, no evaporation takes place, while for SPEI, evaporation takes place and is not limited by moisture availability. Multi-scalar standardised climate indicators such as these allow for comparison of unusually dry (or wet) periods across locations with different climates. The SI are calculated as follows. First, the time series of a variable D (precipitation for SPI, precipitation minus PET for SPEI) is aggregated using a

specified accumulation period length of n months, such that the value for each month in the resulting time series is the average of that month and the n preceding months. Then, a suitable distribution F_{D} for that variable is fitted to the aggregated time series, for each month and location. The SI value for an accumulation period length n at a time step t is then defined as follows:

$$SIn_t = \phi^{-1}(F_D(D_{n,t}))$$

with $D_{n,t}$ indicating D accumulated over the n time steps preceding t (inclusive), and ϕ the standard normal distribution. Monthly values of SPI and SPEI are calculated using n of 3 to 24 months. Following recommendations provided by Stagge et al. (2015b), the two-parameter gamma distribution was used for calculating SPI and the generalised extreme value (GEV) distribution was used for calculating SPEI. For shorter SPI accumulation periods (1-3 months) and further into the future in the UKCP18-RCM simulations (with drying summers), there may be occurrences of zero accumulated precipitation for grid cells in drier regions. To take this possibility into account, the SPI values corresponding to the probability of zero accumulated precipitation were calculated separately following the method proposed by Stagge et al. (2015b), which avoids the mean SPI becoming larger than 0. A 50-year period (1961-2010) of observation-based data (regridded HadUK-Grid and CHESS-PE) was used to fit the distributions for the SPI and SPEI calculation. This observation-based calibration was also applied to the UKCP18-RCM data to allow a direct comparison of the results between climate model ensemble members and observations. This is appropriate because the bias adjustment brings the distributions of the reference period climate model data close to the observed distributions.

4.3.5 Drought characteristics

In order to compare changes in overall drought conditions to changes in more extremely dry conditions, we consider a category of "all/total drought" covering all SI of -1 and lower, and a category of "extreme" drought covering SI values of -2 and lower. These threshold values are a subset of the classification originally introduced by McKee et al. (1993), which has been extensively used in studies using standardised drought indicators. As in Stagge et al. (2015b), SI values were capped at -3 to limit the uncertainty induced by extrapolating into the very extreme tails of a distribution fitted to the relatively short time series available (see Section 4.3.4).

Given the importance of both space (e.g. extent, spatial connectivity, local

vulnerability) and time (e.g. seasonal timing, duration) for drought impacts, the spatiotemporal characterisation of droughts is an important element of any drought study. It is approached here in three ways. First, the frequency (fraction of the time in drought) of dry and extremely dry conditions was computed for each individual grid cell of GB separately, for each ensemble member and the observations. Second, the drought area extent was quantified as the fraction of the total GB area simultaneously in (extreme) drought. We then compute the frequency with which different drought extents are exceeded (fraction of time). Third, regionally averaged SI values were used to investigate drought seasonality and duration. For computing these regional averages, we used the UKCP18 administrative regions (ukcp18, 2021) shown in Fig. 4.3.1, as they represented a decent trade-off between the sizes of the regions, number of regions to compare and relevant differences in climatology, projected changes and societal relevance.

Figure 4.3.1: Map of administrative UKCP18 regions used for regional drought characterisation. Regions for which results are shown in the main text are highlighted in yellow.

For investigating the seasonal contributions to longer-term deficits (seasonality), we compared the 6-month aggregated regionally averaged SI (SI6) for March and September for each year to represent the winter and summer contributions to that year's overall dryness (SI12). Durations of individual drought events are defined as periods of continuously negative regionally averaged SI values reaching

a threshold value of -1 or lower, following the theory of runs (Yevjevich, 1967). Each event is then assigned to the time slice (reference period, +2 °C or +4 °C) that contains its central time step, and the number of occurrences of droughts with different duration categories is assessed. Extreme droughts are identified as events that have a peak (i.e. minimum) SI value below -2. To assess changes in drought duration and the occurrence of multi-year droughts, SI computed with an aggregation period of 6 months was used. This sub-yearly aggregation period is frequently used and linked to impacts (e.g. Stagge et al., 2017; Parsons et al., 2019), and ensures any resulting drought durations of a year or longer were sustained throughout all seasons. For seasonality, duration as well as the seasonal cycles of precipitation and PET, four regions are shown in the main text to represent the main results found across all regions, while results for the other regions are included in Appendix B.

4.4 Projected climatic changes

Regionally averaged seasonal cycles of precipitation (blue) and PET (yellow) are shown in Fig. 4.4.1 for North Scotland, North East England, South West England and East of England, and in Fig. B.0.3 for all regions. The four regions shown in Fig. 4.4.1 were selected to represent the spread of climatic regions and projected changes in climate and drought indicators of all 12 regions, and will be used throughout this chapter to discuss the spatial variability in projected changes. The observations plotted in the reference period column show a very close match with the UKCP18-RCM ensemble, which is the result of successful bias adjustment for each season. In all regions, existing seasonal patterns become more pronounced under a warming climate, and in most regions there is a shift in rainfall seasonality delaying the driest months (clearly visible in South West England). In summer, especially in the South and in the East, the combination of increasing PET and decreasing precipitation lead to an increasing gap between the two, and an increasing period where atmospheric demand for moisture exceeds supply (light yellow area). In some areas (e.g. North Scotland), the reference period precipitation exceeds PET year-round (light blue area), but a warming climate causes this gap to diminish or even crossed in late spring to summer. The ensemble spread in the simulated changes of precipitation, which is driven more by dynamical processes, is greater than that of PET. However, the ensemble broadly agrees on the pattern of projected changes.

Considering the annual average ratio of precipitation and PET, parts of GB are projected to become more arid in most ensemble members (Fig. 4.4.2). This is mostly the case in the (south-) east and the English Midlands, where in the

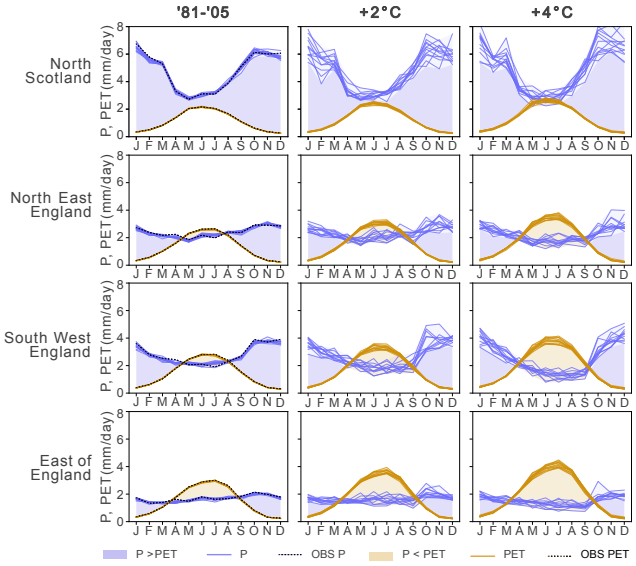

Figure 4.4.1: Seasonal cycle of precipitation (P; blue lines) and potential evapotranspiration (PET; orange lines) for the 12 bias-adjusted UKCP18-RCM ensemble members, for four selected regions. The different lines represent different ensemble members, while the observations are plotted in the first column in darker, dashed lines. See Fig. B.0.3 seasonal cycles for all 13 regions.

reference period the AI was already closer to 1 (annual PET roughly equal to annual precipitation), and PET starts to exceed precipitation on an annual basis under a +2 °C warming scenario. While the ensemble agrees very well on the spatial patterns of aridity changes, there is significant ensemble spread in the magnitude of change. In the +4 °C scenario, widespread AI decreases in the (south-) east and the Midlands are projected by all ensemble members, but only three ensemble members simulate small isolated areas in the South East crossing the threshold from humid to a dry sub-humid climate (aridity index < 0.65). The strong similarity of the reference period simulations to the observations (top row maps) is showing successful bias adjustment of daily precipitation and PET in the ratio of annual averages.

4.5 Projected changes in drought characteristics

4.5.1 Drought frequency

Figure 4.5.1 shows the spatially averaged frequency of dry and extremely dry conditions based on SPI6 and SPEI6 for three time slices representing different warming levels. The scatter plots show the relationship between GB-averaged drought frequencies using SPI6 and SPEI6, as projected at different global mean warming level (GMWL) in the 12 ensemble members and the

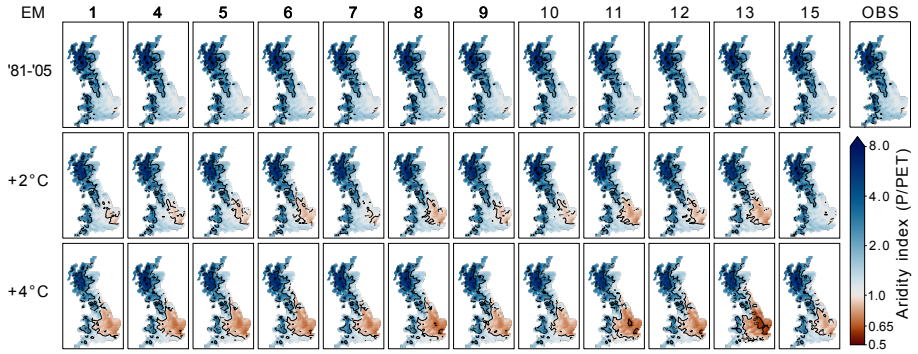

Figure 4.4.2: Aridity index (average annual P/PE) for the 12 bias-adjusted UKCP18-RCM ensemble members. The contours shown in black are powers of 2 and the level of 0.65 (below which a climate is classified as dry sub-humid).

observations. Considering a GB average, the UKCP18-RCM ensemble generally projects an increased frequency of moderate to extreme drought conditions with each warming level using both indicators. In the scatter plots, all points move upwards (more frequent SPEI6 events) with increasing global warming level, and most move to the right (more frequent SPI6 events), except for a few ensemble members for +2 °C. However, the GB-averaged drought frequency increases and future projections are, for each ensemble member and warming level, greater when quantified using SPEI6 than using SPI6. For those 3 ensemble members that project a slight decrease in total drought frequency based on SPI6 for +2 °C, including PET in the drought indicator (SPEI6) changes the sign of the projected change. To compare the differences between SI and GMWL, the SPEI6-based GB-average drought frequency projected at +2 °C is equal to or greater than the SPI6-based frequency projected at +4 °C for at least half the ensemble members in each drought category. For drought frequency quantified with SPI6, by +4 °C, the projected increases range from a few percent points to more than double the reference frequency, and between two- and eightfold for the extreme droughts.

At the same high warming level, the ensemble projects SPEI6-based drought almost half of the time (ensemble average: 46%), about half of which (ensemble average: 23%) are classed as extremely dry conditions. The ensemble spread (scatter) of future projections is substantial and grows with increasing warming level. Importantly, the projected relative increase for extreme drought frequency is far greater than for the total drought frequency. By +4 °C, the ensemble mean spatially averaged total drought frequency increases by a factor 1.7 for SPI6 and by a factor 3.1 for SPEI. For extreme droughts, however, these multiplication factors are 3.7 and 11.5, respectively. This disproportionate increase in the extreme drought category, which shows in projections based on both indicators, has potentially important implications for drought impacts,

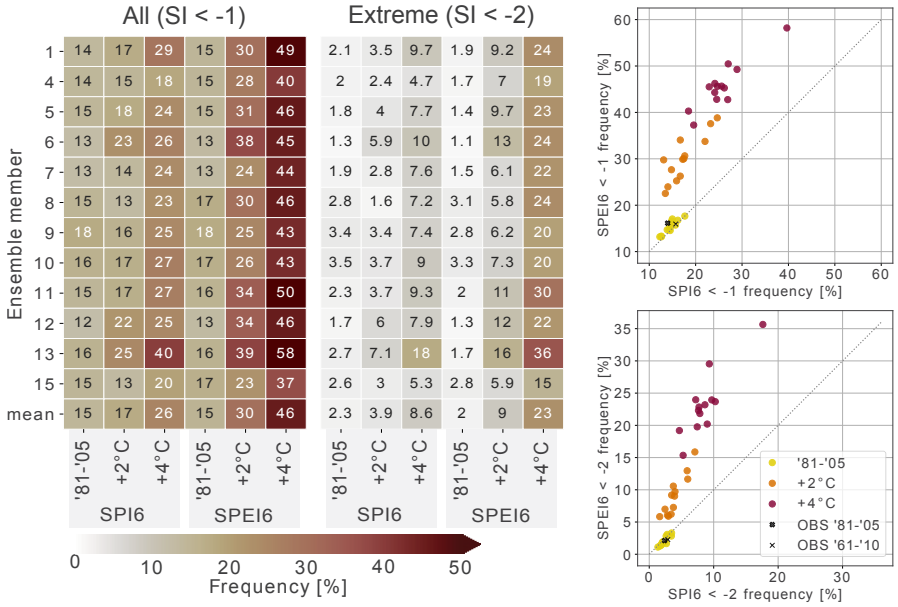

Figure 4.5.1: Spatially averaged projections of drought frequency, expressed as the fraction of months SI is below the threshold, for each ensemble member (rows) and the ensemble mean (bottom row), for three time slices (subcolumns) that correspond to the reference period and two different levels of global mean surface warming compared to pre-industrial levels, using both drought indices (columns). Frequencies are shown for all droughts (left), extreme droughts (middle), and as scatter plots (one point per ensemble member) comparing SPI6- and SPEI6-based frequencies of all droughts (top-right) and of extreme droughts (bottom-right). Spatial averages are across the whole of Great Britain.

such as stakeholders or ecosystems vulnerable only to extremely dry conditions (e.g. Parsons et al., 2019).

The maps in Fig. 4.5.2 show the spatial patterns of these drought frequency changes (for the ensemble average) and their differences between SPI6 and SPEI6. For the reference period, the ensemble-averaged GB mean total and extreme drought frequencies are 0.15 and 0.023 respectively, which are close to the theoretically expected values of 0.16 and 0.022. There is some variation around these values in space (Fig. 4.5.2) and among ensemble members (Fig. 4.5.1), which is not unexpected. Imperfections of the distribution fits in the calculation of SPI and SPEI, differences between the climates of the 1961-2010 and 1981-2005 periods (black markers in Fig. 4.5.1), any model errors remaining after bias adjustment and internal climate variability can all result in differences between the simulated drought frequency in the reference period and the theoretical frequencies that would be expected for the calibration data.

There is significant regional variability in projected drought frequency across GB inferred with either drought indicator, especially for extreme drought (Fig. 4.5.2).

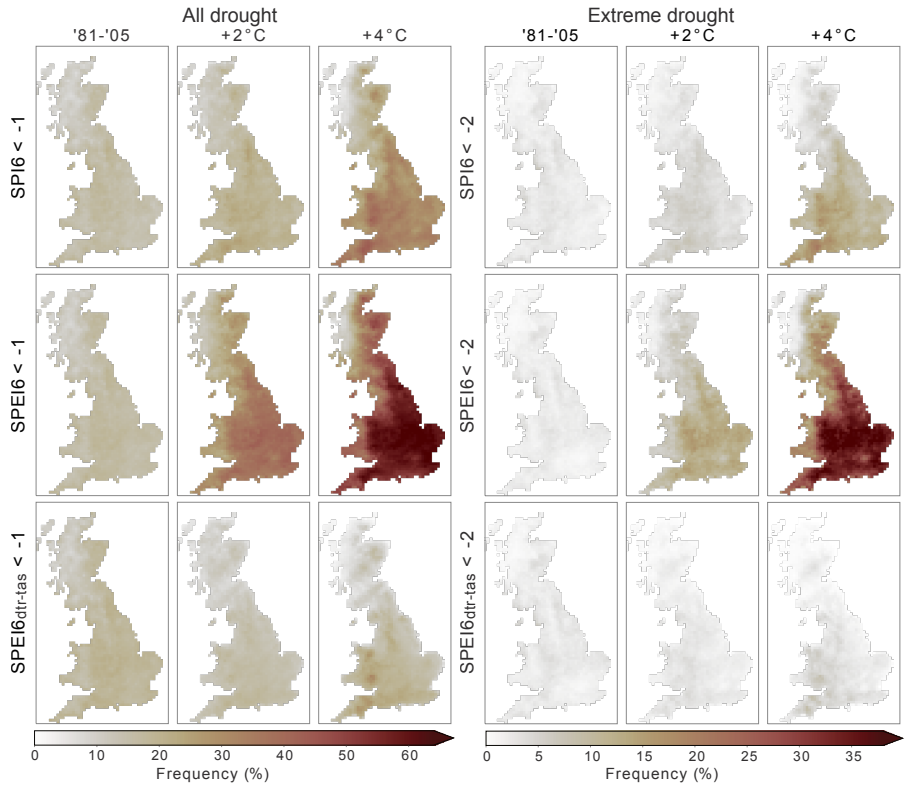

Figure 4.5.2: Ensemble averaged projected frequency of all (left) and extreme (right) dry conditions, expressed as the fraction of time SI falls below -1 or -2 respectively. Top: SPI6, middle: SPEI6 with projected temperature changes, bottom: SPEI6 with detrended temperature.

Both SI show a similar pattern, with the mildest increases or even decreases along the west coast, most notably in north-west Scotland. However, the areas projected to experience the greatest increase in frequency of dry conditions differ between the drought indices. In the SPI6-based projections, the greatest increases are projected in the rain shadows of highly elevated areas. For SPEI6, the largest increases are seen in these areas plus a larger area covering most of England (except near the west coast), including the East Midlands and East England where SPEI6-based drought conditions are projected around 60% of the time under +4 °C. These are already the least humid regions of GB (Fig. 4.4.2). For both indices, these regional patterns of change are amplified when looking at the higher warming level and when isolating extreme droughts.

The bottom row of Fig. 4.5.2 shows $\text{SPEI6}_{\text{dtr-tas}}$, which is the SPEI6 using PET calculated with detrended temperature simulations. With the projected temperature increase removed, SPEI6 shows only minor changes in drought frequency. Furthermore, at +4 °C, the projected drought frequencies using $\text{SPEI6}_{\text{dtr-tas}}$ are much less than those found for the precipitation-only SPI6. On the face of it, that suggests non-temperature influences may reduce PET

(offsetting some of the temperature-driven increase) and that PET calculation methods which only rely on temperature (e.g. Thornthwaite) may overestimate drought risk based on the UKCP18-RCM simulations. However, the effects of physically inter-dependent variables (especially temperature and humidity) cannot be truly separated. Crucially, here we use simulated specific, not relative, humidity to compute PET (Robinson et al., 2017). Whereas specific humidity is projected to increase over GB, relative humidity is projected to decrease as the saturated humidity increases faster with rising temperatures (not shown), contributing to the increased future PET. However, by detrending the temperature, the saturation humidity level computed in the PET calculation was reduced for future simulations, which, combined with the unadjusted specific humidity projections, resulted in artificially increased relative humidity and thus a decreased vapour pressure deficit term. The temperature effect shown by the SPEI6 - $\text{SPEI6}_{\text{dtr-tas}}$ difference (Fig. 4.5.2) therefore implies a far greater effect of temperature than if a PET formulation using the relative humidity projections would have been used (Fu and Feng, 2014; Robinson et al., 2017).

4.5.2 Spatial extent

Figure 4.5.3 and Fig. 4.5.4 show the observed and simulated extent-frequency curves of drought conditions for SPI and SPEI respectively, for different global warming levels (i.e. time slices) and using different aggregation levels. Moving upwards in this plot means an increase in the frequency of drought conditions with at least the spatial extent given by the horizontal axis (not necessarily in the same locations). Moving to the right in this plot means an increase in the spatial extent of drought conditions that is exceeded with a particular frequency (given by the vertical axis).

The relationship between frequency and drought extent for the reference period simulations generally match well with the observations. However, as the aggregation period increases, the frequencies of smaller drought extents are increasingly overestimated in the simulations, while the frequencies of larger drought extents are on average well represented (SPI) or become slightly underestimated (primarily SPEI), especially for the 12 and 24 month aggregation periods. The ensemble spread for the reference period simulations also increases with the aggregation level, enveloping the observations in all cases. The bias adjustment was done on the distributions of daily values of individual grid cells, not considering the spatial coherence in longer-term statistics, and without considering correlations between precipitation and PET.

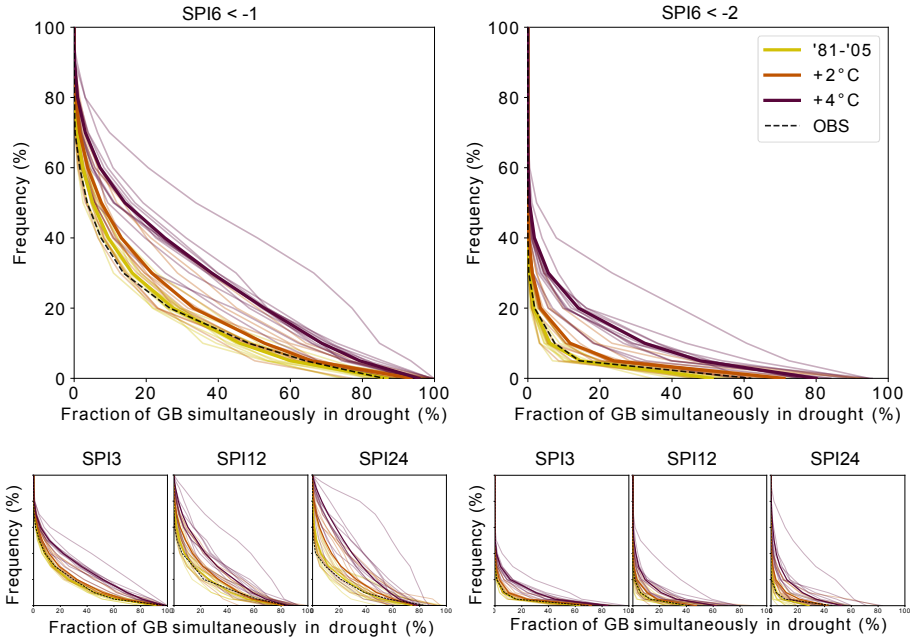

Figure 4.5.3: Extent-frequency curves for all (left) and extreme (right) drought extents based on SPI at different aggregation levels (subplots). The horizontal axis gives the drought extent (as fraction of GB area) that is reached or exceeded with a frequency given by the vertical axis.

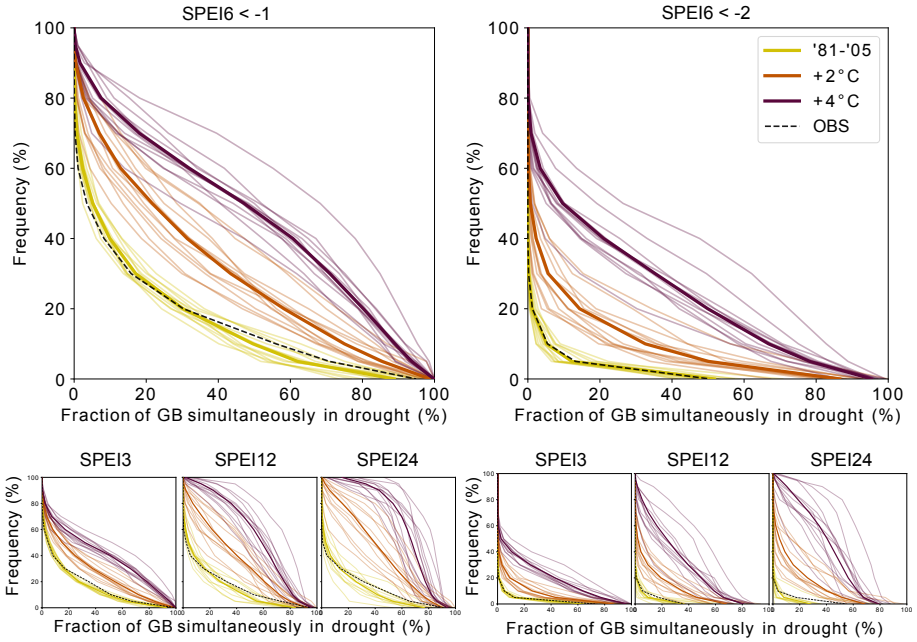

Figure 4.5.4: As Fig. 4.5.3 but for SPEI rather than SPI

These might be the reasons for these minor mismatches.

For a given drought extent, the relative change in frequency as global temperature increases is far greater for extreme droughts than for all droughts (for both SPI and SPEI). For instance, based on SPI6, the frequency that at least 20% of GB

simultaneously experiences a drought is 26% currently, and 44% with +4 °C of warming (mean of the ensemble). In contrast, SPI6-based extreme droughts covering 20% GB have a frequency of occurrence currently less than 4% of the time, but this frequency is projected to jump to 16% for a +4 °C warming (mean of the ensemble). Climate change-induced changes in the relationship between frequency and extent of droughts depend strongly on the drought metric used. SPI and SPEI both show increasing frequency of droughts of most extents, however the increase is much greater for SPEI. Moreover, different frequency changes are projected for droughts with different extents, e.g. greater changes for smaller drought extents using SPI. Using longer aggregation periods, the future projections move toward higher frequencies and extents, the ensemble spread increases, the difference between GMWL grows, and differences between drought indicators become more pronounced. For the extreme drought class, the maximum extent is projected to increase greatly with global warming, based on both SPI and SPEI. The ensemble mean maximum SPI6 area fraction in drought increases from just over 51.2% (an underestimation of observation-based maximum extent) to just over 71.1% by +2 °C and to 80.0% by +4 °C. For SPEI6, the ensemble-averaged simulated maximum extent and the overall frequency-extent relationship matches observations very closely, and the maximum extent is projected to increase from just over 51.8% to 86.5% by +2 °C, and to 95.4% (i.e. almost all of GB simultaneously in extreme drought) at +4 °C. The relative increase of maximum extreme drought extent projected due to global warming is greater for longer aggregation periods, for both indicators. Finally, the drought-free frequency, given by the difference between 100% and the intercept on the y-axis, is generally projected to decrease under climate change, again far more strongly for the extreme drought category and for SPEI-based drought.

4.5.3 Seasonal timing

Figure 4.5.5 shows the contributions that summer and winter deficits make to annual droughts according to SPI and SPEI for three global warming levels for the selected GB regions. Results for the other regions can be found in Figs. B.0.4 and B.0.5. The horizontal and vertical axis show SI6 for March and September respectively, indicating how dry or wet the hydrological winter and summer were in a given year. The September SI12, indicating the dryness of the corresponding hydrological year, is represented by the colours of the dots. For example, a grey dot with coordinates (1.1, -2.2) represents a normal annual value consisting of a wet winter and an extremely dry summer.

The increasing summer dryness is reflected by a general downwards shift of the

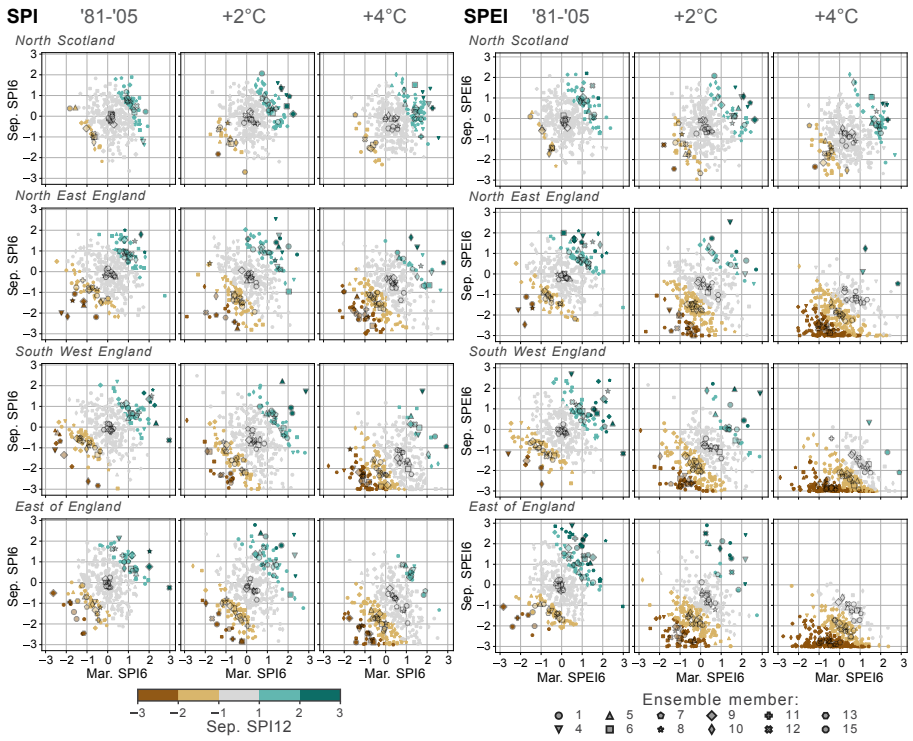

Figure 4.5.5: Values of September SI12 (hydrological year) plotted against the September SI6 (hydrological summer) and March SI6 (hydrological winter) from the same year used to compute SI12 for SPI (left) and SPEI (right). All years are shown for each time slice and ensemble member. SI6 values that exceed -3 or +3 are plotted at -3 or +3. The larger, transparent markers show the centroids of 5 SI12 classes: extremely dry, dry but not extremely dry, normal, wet but not etremely wet, extremely wet. See Figs. B.0.4 and B.0.5 for results for all 13 regions.

point cloud, while a rightward shift reflects wetter winters in some regions. Increases in the proportion of dry years are projected in most regions and can be attributed mainly to the summers of those dry years, especially for SPEI-based droughts. In several regions, such as the East of England, most summers in the ensemble are classified as dry by +4 °C, leading to their respective years to be classified as dry in about half (SPI +4 °C and SPEI +2 °) or almost all (SPEI +4 °C) cases. With increasing global warming levels, a growing number of years consist of a wet winter followed by a dry summer (bottom right corner beyond the (1, -1) coordinate), which is rare in the reference period simulations. In South West England, this even becomes the norm under +4 °C in these simulations (grey centroid dots). Using the SPEI, in all except the Scottish regions, an increasing number of these "contrasting" years is categorised as dry (or even extremely dry, in some regions under +4 °), which is not observed at all for the SPI. Using the SPI, in most of the western regions the number of "contrasting" years classified as wet based on their SI12 increases, which is observed to a lesser extent in SPEI. The implicit assumptions

on evapotranspiration in these indicators thus have a decisive influence on seasonal droughts and how they tip the annual water balance, demonstrating the importance of understanding the influence of these PET increases.

4.5.4 Duration

Figure 4.5.6 shows the number of simulated drought events within 6 drought duration categories (horizontal axis), based on SPI6 and SPEI6. Results for the other regions as well as droughts that reach extreme levels are shown in Figs. B.0.6 and B.0.7.

Overall, the drought indicator makes a large difference in the projected changes in the distribution of drought durations. The ensemble spread of the number of events in each drought category is large, for both indicators, and there is often a strong overlap between GMWL which is diminished when isolating droughts that reach extreme levels. In most regions, the SPI6-based projections show increases in droughts shorter than 6 months, while the SPEI6-based projections for this category are divided between decreases in the drier region (south and central to east), and increases or little change in the other regions. In almost all regions, however, there is an increase in 6-11 month droughts using both indicators. The decreases projected in the shortest SPEI6-drought category in half of the regions are generally paired with increases in longer droughts, suggesting that the larger projected drought frequency in these regions (see Fig. 4.5.2) is concentrated in fewer consecutive dry periods, with seasonal droughts getting pooled together into longer events.

Sustained multi-year droughts are a major concern for water managers (e.g. Marsh et al., 2007). They can also have less occurrences in a 25 year time slice by definition, and a larger share of the ensemble members contains zero multi-year events for a given time slice and region. Droughts lasting at least 2 years rarely occur more than once in a given time slice in our analysis, and never more than twice for a given duration bin. Therefore, for these events we discuss the total number of ensemble members that project at least 1 such event in any given time slice. Based on the SPI6, the number of ensemble members projecting at least one drought lasting from 2 to 3 years is not projected to change for most regions, although an increasing share of events reaching extreme levels is found in about half of the regions. Using the SPEI6, the number of ensemble members projecting at least one 2-3 year event increases with GMWL in the southern and central to eastern regions, and for events reaching extreme levels this increases in almost all regions. The number of these events simulated in a single time slice by a single ensemble member also

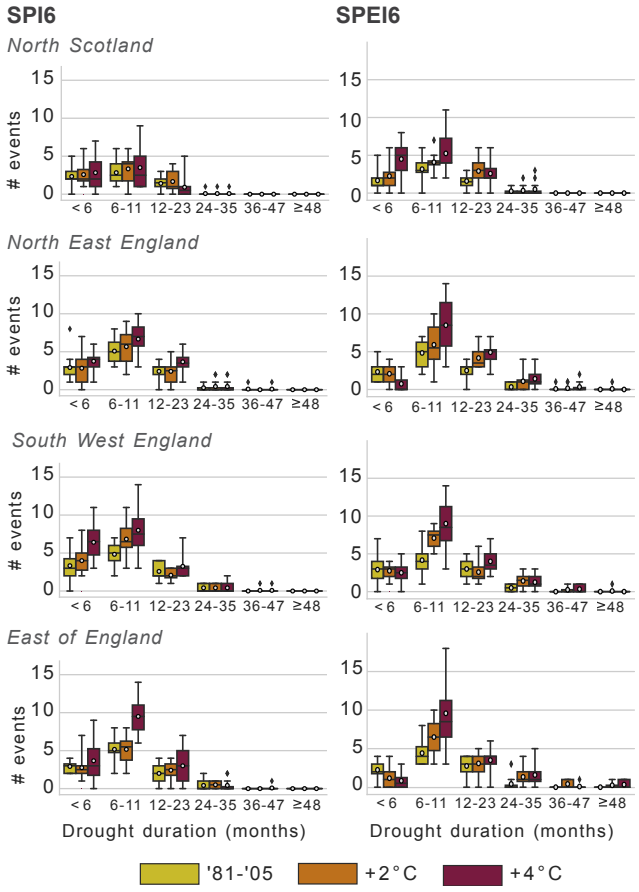

Figure 4.5.6: Number of drought events of all severities by duration for three 25-year periods corresponding to progressive warming scenarios in the selected regions, based on SPI6. White circles indicate the ensemble mean, boxes show the interquartile range, whiskers show the ensemble range except for members exceeding 5 x the interquartile range (diamonds). See Fig. B.0.6 for other regions.

increases in several regions using the SPEI6. Droughts lasting three years or longer in the reference period are simulated in either none or one of the ensemble members depending on the region, irrespective of the SI (with exception of the West Midlands for SPEI6: 2 ensemble members). A drought of four years or longer in the reference period is only simulated by one ensemble member in one region for each indicator. With increasing warming levels, more ensemble members simulated at least one +3 year SPEI6-drought event in the English regions and Wales, most of which reaching extreme levels at some point. This is in contrast with SPI6, where little change can be found in the number of ensemble members simulating such events (max. 2 ensemble members in any time slice and region). As increasing multi-year droughts across the ensemble are almost exclusively seen using SPEI, any indication toward a possible increased likelihood of these events depends on the influence of AED.

4.6 Discussion

4.6.1 Projected changes in atmospheric droughts

This section discusses the results presented above in the context of previous studies that have used meteorological and atmospheric based drought indices to investigate climate change impacts on droughts in the GB.

The spatial pattern of drought frequency changes in Fig. 4.5.2 is very broadly in agreement with the spatial patterns of drought intensity found by Hanlon et al. (2021) using UKCP18-RCM and the DSI, and increases in drought event occurrence found by Spinoni et al. (2018) using EURO-CORDEX and a combined atmosphere-based drought indicator. Nevertheless, the differences between our SPI6 drought frequency maps and the DSI6 drought severity maps in Hanlon et al. (2021) reveal how drought projections can be sensitive to the exact method used for drought quantification and characterisation, even considering the same variable (precipitation) and aggregation time scale (6 months). For example, the hot spots of the SPI6 drought frequency increase in Fig. 4.5.2 are further west than the DSI6 drought severity increases along the east of GB found by Hanlon et al. (2021). Moreover, by discriminating between all and extreme drought, we showed how the spatial patterns of drought frequency projections are similar but amplified in the extreme drought class. Furthermore, this study analysed for the first time the projected changes in drought extent as a function of its frequency. The difference in the shape of the observation-based extent-frequency curves between extreme and all drought conditions (Fig. 4.5.3 and Fig. 4.5.4) confirms the finding by Tanguy et al. (2021) that the most extreme droughts tend to be more localised than when all droughts are considered, and shows that this is also true using a drought definition including PET. Rahiz and New (2013) found a projected increase in drought spatial coherence using UKCP09 and the DSI6. Here, we not only showed increased drought extent and frequency, but notably a larger relative increase in the frequency of widespread extreme drought conditions, as well as strong increases in more localised extreme droughts. This is the case for SPI, but greatly amplified when including PET. Moreover, we showed that the observed and projected drought extent-frequency relationship is time scale-dependent. As discussed in Section 4.6.4, widespread dry and extremely dry conditions identified using a SI with one specific aggregation period, would likely lead to differential agricultural and water resources impacts depending on the relevant time scales in the affected regions.

Previous studies have often assessed changes in drought duration through the

mean and/or median duration or overall trends (e.g. Touma et al., 2015; García-Valdecasas Ojeda et al., 2021; Vicente-Serrano et al., 2021). Instead, here we isolated changes in events of different duration categories, which revealed a possibility of increasing multiyear droughts in some regions based on the SPEI6, but not the SPI6. Multi-year droughts were also assessed by Lehner et al. (2017), who found that for some studied regions (including Central Europe and the Mediterranean), progressive climate change is projected to increase the risk of 4 consecutive drought years (based on the PDSI, which is also sensitive to projected increases in PET). Rahiz and New (2013) considered changes in drought events lasting at least 3, 6, 10 and 12 months based on the DSI6 using the previous generation of UKCP regional projections (UKCP09), and found widespread increases in the number of events of at least 3 months, generally with stronger increases and ensemble agreement toward the south west. Seasonal timing and contributions of drought were assessed by investigating changes in the combination of March SI6, September SI6 and September SI12 for a given year. By visualising the relationship between annual conditions and the summer and winter half-years, this approach goes beyond assessing changes in seasonal and annual SI independently (e.g. Spinoni et al., 2018; Vicente-Serrano et al., 2021) in making use of the multiscalar property of these indices. In this way, it was shown that the dominant contribution to increasing deficits in the annual SI in many regions consists of increasing deficits in the hydrological summer SI6 (especially for SPEI), and that more years consisting of a dry summer preceded by a wet winter are projected in many regions. With an accumulation period of 3 months, Spinoni et al. (2018) found decreasing occurrence of drought events in winter and increasing occurrence in the other seasons, with the strongest increases in summer, with a spatial pattern dependent on the scenario and drought intensity considered. These and our results are in disagreement with Rahiz and New (2013), who found larger and more widespread drought frequency and intensity in the hydrological wet season, most likely due to a methodological difference in delineating seasons.

4.6.2 Differences between SPI and SPEI projections

We show that the magnitude of the difference between SPI- and SPEI-inferred projected changes is substantial for all considered drought characteristics. For drought frequency, using the 6 month aggregation period, it is comparable to the difference between +2 °C and +4 °C of warming above pre-industrial levels for the extent and frequency of drought and extreme drought. Within both warming scenarios, the difference in GB-averaged projected total drought frequency between SPI and SPEI is similar to the ensemble range, for either SI.

For extreme drought, the difference between SPI6 and SPEI6 is similar in size to the ensemble range according to SPI at +2 °C, and lies between the ensemble ranges of SPI and SPEI at +4 °C.

Previous studies found divergence in trends of drought characteristics between SPI and SPEI in observations (Stagge et al., 2017; Karimi et al., 2020; Ionita and Nagavciuc, 2021), historical climate simulations (Chiang et al., 2021) and future climate projections (e.g. Arnell and Freeman, 2021; García-Valdecasas Ojeda et al., 2021; Wang et al., 2021), with SPEI indicating increased drying compared to SPI. Increases in PET under a changing climate, combined with the high sensitivity of SPEI to PET changes, cause amplified projections of climatological drying and even a reversal of wetting trends in some parts of the world compared to when only changes in precipitation are considered (Cook et al., 2014). For the UK, Arnell and Freeman (2021) found that projected increases in drought frequency based on SPEI6 exceeded those based on SPI3, which is attributed to the inclusion of the effect of PET in SPEI, although the difference in aggregation period should also have contributed. Ionita and Nagavciuc (2021) found a divergence of observed SPI12 and SPEI12 trends over Europe for 1901-2019, with the strongest drying trends located over the Mediterranean and Central Europe regions. For GB, they found mostly non-significant SPEI12 trends from wetting in the north-west to drying in the south-east (mostly due to a summer drying trend), alongside (also mostly non-significant) wetting SPI12 trends, especially toward the north. For 1958-2014, Stagge et al. (2017) found a decreasing SPI6-based drought extent not being reflected in SPEI6-based drought extent trends over Europe. For GB, they found a non-significant difference between SPI6- and SPEI6-based drought frequency trends, with both SI6 showing significant wetting in the north. Through a detailed analysis, the present study showed substantial differences between SPI- and SPEI-based projections for drought frequency, the distribution of drought spatial extents (using different temporal aggregations), the distribution of drought event durations and the seasonal contributions to 12-month deficits. The stark differences between SI in projections of all these drought characteristics, combined with their emerging divergence in observations documented in the literature, invites more critical consideration before using one of these indicators in drought studies or monitoring, based on an understanding of the likely impacts of increasing PET.

4.6.3 The role of AED

The strong sensitivity to global warming of drought projections based on drought indicators relying on PET has been discussed before (e.g. Seneviratne, 2012; Dewes et al., 2017; Berg and Sheffield, 2018; Manning et al., 2018; Scheff et al., 2021), considering the following aspects. First, overly warming-sensitive PET formulations can lead to overestimating increases in drought. This is not only true for temperature-based methods such as Thornthwaite (Sheffield et al., 2012), but also for the FAO56 reference crop Penman-Monteith method used in this study and many others. Assuming a fixed stomatal resistance of the reference crop neglects the effects of increasing CO_2 on plant growth and stomatal conductivity, which has been identified as an important source (Milly and Dunne, 2016; Greve et al., 2019), but not the full explanation (Scheff et al., 2021), of off-line PET overestimation in climate change studies. The impact of the representation of influences of CO_2, temperature and vapour pressure deficit (Grossiord et al., 2020) on transpiration is likely highly relevant for the results presented in this study, as transpiration and bare soil evaporation respectively make up the largest and smallest fractions of total evapotranspiration in GB, with transpiration constituting the majority of AET in the English Lowlands (Blyth et al., 2019). Second, when looking at the variables standardised in SPI and SPEI as proxies for the surface water balance, the assumptions are respectively that no AET occurs, or that AET always occurs at its maximum rate (PET), neglecting possible limitations from moisture supply. In reality, the response of AET to increasing AED is complex, and the land-atmosphere interactions contributing to drought development and propagation, including the role of evapotranspiration under high AED, are active areas of ongoing research (e.g. Miralles et al., 2019; Vicente-Serrano et al., 2020; Gampe et al., 2021; Denissen et al., 2022; Massari et al., 2022; Zhao et al., 2022). Intuitively, enhanced AED leads to enhanced AET until moisture availability becomes limiting, after which the effect of AED on AET reduces. This implies a temporally variable response of AET to AED during drought development, evolution and recovery, dependent on moisture availability (e.g. Zhao et al., 2022), and different responses based on the regional climate (e.g. Vicente-Serrano et al., 2020). The sensitivity of SPEI to AED also varies between climates (Tomas-Burguera et al., 2020). Moreover, the behaviour of AET under drought crucially depends on land cover and plant physiology (e.g. Teuling et al., 2010; Grossiord et al., 2020), soil structure (e.g. Massari et al., 2022; Zhao et al., 2022), and geology (e.g. Bloomfield et al., 2019). Finally, due to equal aggregation periods used for precipitation and PET in SPEI, it is inherently implied that the drought development contributions of low precipitation and high PET anomalies are influential over the same time scales,

which is not always the case (Manning et al., 2018).

Interestingly, GB sits in the transition between humid, radiation-controlled Northern and Central Europe and more arid, precipitation-controlled Southern Europe (Teuling et al., 2009). Evaporation is generally more water-limited and negatively correlated with temperature in summer toward the south and east, and more energy-limited and positively correlated with temperature in summer in the north and west of GB on an annual basis (Seneviratne et al., 2006; Kay et al., 2013). This has important implications for the expected impacts of increasing AED on future droughts across GB, as the influence of AED varies between energy- and water-limited evaporation regimes, and the effect of AED increases can be more complicated in transitional regions (Vicente-Serrano et al., 2020). Indeed, Kay et al. (2013) found that observed trends in PET between 1961 and 2012 are greater than those for AET for England and Wales, while in energy-limited Scotland PET and AET trends are very similar. This is in contrast to Blyth et al. (2019), who found that modelled AET increased at a greater rate than PET in GB between 1961 and 2015, due to increases in precipitation and the large contribution of interception to total AET. Enhanced AED has already been shown to enhance streamflow droughts in GB, with a stronger effect in some regions than in others (Vicente-Serrano et al., 2019; Massari et al., 2022), as well as groundwater droughts in the major aquifer in south-east GB (Bloomfield et al., 2019).

The importance of the range of evaporation regimes for explaining drought propagation and drought impacts across GB has not received much attention in existing literature, but presents a valuable direction for further research. For example, the currently least humid areas of GB are projected to experience large increases in SPEI-drought, increases in aridity, and on average longer and more intense seasons where PET exceeds precipitation. The effect of extreme SPEI-drought conditions on soil moisture and streamflow droughts in these areas could be smaller than suggested by the magnitude of the PET contribution due to moisture availability becoming limited. In such conditions, vegetation may still be significantly impacted due to high AED and its components (Schönbeck et al., 2022). Understanding potential shifts in these evaporation regimes under climate change could help inform climate change adaptation strategies related to land and water use. To better understand the PET component of the projected SPEI-based drought projections for GB, we detrended temperature (which affected the vapour pressure deficit term and the slope of the Clausius-Clapeyron relation), after which no increases in $\mathrm{SPEI}_{\mathrm{dtr\text{-}tas}}$ drought frequency were projected in most regions of GB. Further research into projected changes for the different variables influencing PET (radiation,

temperature, relative and specific humidity, wind speed) is needed to better understand the strong contribution of PET to SPEI-based drought projections, and to help understand possible shifts in evaporative regimes over GB.

Based on the discussion above, and depending on the drought type or impact of interest, the SPEI-based results in this work may present a (conservative) upper limit of future drought risk, while using the SPI (and other precipitation-only indicators) is expected to underestimate these increases. Future changes in other drought types may thus end up in the range between SPI- or SPEI-based projections depending on the region (Touma et al., 2015; Lee et al., 2019). Thus, these results highlight the importance of understanding (changes in) the role AED plays in GB droughts and overall hydroclimatic changes under a changing climate.

4.6.4 From atmospheric indicators to impacts

Many studies have used a range of drought impact-related data to investigate the relationships of SI with different aggregation periods in GB and beyond (Stagge et al., 2015a; Folland et al., 2015; Barker et al., 2016; Bachmair et al., 2016, 2018; Haro-Monteagudo et al., 2018; Parsons et al., 2019; Gampe et al., 2021). This is not straightforward, as impact variables and reports of past droughts (based for instance on observed flow) are also influenced by water fluxes driven by the land surface (e.g. evaporation limited by soil moisture) and human actions (e.g. irrigation and water abstractions), which are not accounted for by SPI or SPEI. While studies linking SI to impacts agree in some aspects (e.g. longer SI aggregation periods for predicting streamflow drought in the south east than the north west; Bachmair et al., 2016; Barker et al., 2016), there is a lot of uncertainty left. In the UK, due to regional differences in climatology, hydrogeology and agricultural practice, the links between SI and various impacts are more meaningful at regional or local levels than at the national scale (Barker et al., 2016; Parsons et al., 2019). Socio-economic and physical vulnerability factors also influence the impacts resulting from droughts characterised by certain SPEI or SPI values (Blauhut et al., 2016). Additionally, previously established relationships between drought indicators and impacts may change under a changing climate (Feng et al., 2017). Therefore, despite established links between SI and impacts, it is difficult to quantitatively infer changes in agricultural, ecological and hydrological drought from drought projections based on SPI and SPEI alone. For example, agricultural drought impacts may be expected to increase due to the projected increase in summer drought frequency and intensity (Stagge et al., 2015a;

Haro-Monteagudo et al., 2018; Parsons et al., 2019), which is found for both indicators in most of GB, including in agriculturally important regions (Fig. 4.5.5). However, as the projections based on SPI are much milder than those based on SPEI, the magnitude of this increase depends on the importance of increasing AED and temperature for root zone accessible soil moisture and crop growth, as well as crop response to components of AED itself. The greater frequency and intensity of dry years (SI12), as well as the increasing extent and frequency of drought and extreme drought with longer aggregation periods, may indicate greatly increased risk of drought impacts on water resources in the southeast and east, and by extension irrigated agriculture in these regions. Smaller projected increases in drought frequency based on SI3 may indicate similarly smaller increases in streamflow drought in the northwest (Barker et al., 2016; Bachmair et al., 2016).

4.6.5 Study limitations

The set of regional climate projections in UKCP18, which this study relies upon, is not intended to represent a comprehensive, probabilistic view of possible changes, but rather to sample a broad range of possible futures and provide storylines suited for analysis of impacts (Murphy et al., 2018). The UKCP18-RCM projections were produced using the same GCM and RCM structure with perturbed parameter values, meaning that the climate model structural uncertainty has not been sampled. Finally, as opposed to an ensemble where only the initial conditions differ, the projections of such a perturbed physics ensemble cannot be combined in order to obtain longer time series for each level of global warming. This primarily limits our analysis of multi-year droughts. For those events, the length of the time slices used is also a limiting factor for investigating projected changes in the occurrence of such events.

The drought indices this study uses are among the most widely used ones. However, other indices exist that rely on precipitation or some combination of precipitation and AED. Choosing a different drought index that includes both moisture supply and demand, with a different degree of sensitivity to each component, could lead to slightly different results (Vicente-Serrano et al., 2015). Indeed, as emphasised in this study, the drought index choice itself can be a substantial source of uncertainty, due to the use of different variables representing different drought types (Satoh et al., 2021), but also between different drought quantifications based on the same variable (Sutanto and Van Lanen, 2021).

Aside from the vegetation assumptions in the PET calculation (see Section 4.6.3), vegetation assumptions in the UKCP18-RCM projections themselves present another potentially important limitation. In the UKCP18 "Soil Moisture and the Water Balance" fact sheet, Pirret et al. (2020a) write that "the models use prescribed vegetation, which means that the model does not represent how increasing atmospheric carbon or reduced soil moisture would affect vegetation, or any feedbacks that this may have on the atmosphere or land surface". This may lead to unrealistic changes in AET under a warming atmosphere with increasing CO_2, and thereby introduce errors in the simulated temperature and humidity, affecting PET.

Finally, using a different observation-based dataset for bias adjustment of PET such as the recently produced Hydro-PE dataset (Robinson et al., 2022), may also lead to quantitative differences in the results.

4.7 Conclusions

We used the regional climate model perturbed parameter ensemble from the latest set of national climate projections for the UK, UKCP18, to quantify projected changes in drought characteristics. For this, two related but contrasting atmospheric-based standardised drought indices were used and their results compared: the Standardised Precipitation Index (SPI) and the Standardised Precipitation Evapotranspiration Index (SPEI). The SPI gives the standardised anomaly of n-month aggregated precipitation. The SPEI is similar, except the variable being standardised is a climatological moisture balance given by precipitation minus potential evapotranspiration. When regarding these indicators as standardised proxies of the surface water balance, their implied assumptions are either that no evapotranspiration occurs (SPI) or that evapotranspiration is never limited by moisture availability (SPEI). We assess in detail the difference between these indices for investigating the impact of climate change on drought frequency, extent, seasonality and duration, for two categories of drought intensity. This is the first detailed systematic analysis of SPI- and SPEI-based drought projections and their differences for Great Britain.

Drought risk over Great Britain increases almost everywhere with increasing global mean surface temperature, including extreme drought risk. We find projected increases in drought frequency and extent with increasing global warming levels. These changes are far more pronounced for extremely dry conditions than for all drought conditions. The projected changes in drought

frequency, seasonality and duration show large regional differences across GB, with the greatest increases generally found in English regions and Wales, notably including some of the already least humid regions toward the south east, and little change (or even decreases) in drought in North and West Scotland. By assessing the relationship between drought spatial extents and their frequency in observations, reference period simulations and future projections, we showed that the reference period simulations capture the observed extent-frequency relationship quite well for both extreme and all droughts, and all (extreme) drought extents are projected to increase in frequency. Moreover, extreme droughts with extents greater than the most widespread drought in the reference period are projected to occur more often, depending on the warming level and especially for SPEI. Linking shorter-term contributions to longer-term deficits is an under-utilised possibility of (standardised) drought indicators that are applied over different time scales. By exploiting the multi-scalar nature of the standardised drought indicators, we found that increasing summer droughts are found to be the main contributor of increasing frequency of increasing longer-term dry conditions. Additionally, contrasting years, consisting of a wet winter combined with a dry summer, are also projected to increase in occurrence. However, the combined result of contrasting seasonal changes is a projected increase in dry years for most regions. Finally, the distribution of drought event durations is also projected to change. For both indicators (but especially for the SPI), the changes are far greater by +4 °C than by +2 °C, supporting the consensus that every additional degree translates into increasing extreme events.

The choice of atmosphere-based drought indicator can have a great impact on the derived drought characteristics, and thus great care should be taken when selecting a drought index for climate change studies. This study clearly showed this for drought frequency, the distribution of drought extents, drought event durations, and the seasonal character of annual deficits. The difference between the six-month aggregation period based SPI and SPEI is similar in magnitude to the UKCP18-RCM ensemble range of GB-averaged total and extreme drought frequency, and the +2 °C SPEI projections better resemble the SPI-based projections under +4 °C than under +2 °C for drought and extreme drought frequency, spatial extent and seasonality. The spatial pattern of simulated drought frequency is similar between the indicators, although there are differences in the regions with the strongest projected changes. Projected changes in the distribution of drought durations also differ between the indicators. Droughts shorter than 6 months are projected to increase in occurrence in most regions based on the SPI, but projected to decrease based on the SPEI in about half of these regions. On the other end, multi-year

droughts lasting over 3 years (based on 6-month aggregated indicators) only occur in some of the SPEI-based projections.

With the sizeable divide between projections based on both indicators, it becomes increasingly important to understand how atmospheric evaporative demand and temperature affect droughts and their propagation to impacts in GB. The large difference between SPI and SPEI in our results calls attention to the need to understand the influence of atmospheric evaporative demand changes on GB drought through land-atmosphere interactions, and its adequate representation in models. In particular, further research is needed to understand the effects of the contribution of PET to projected drought conditions across the range of climatological evaporation regimes in GB (from energy-limited to transitional and water-limited), and likely changes in these regimes. Moreover, analysing the contributions of changes in radiation, relative and specific humidity, temperature and wind speed can shed light on the PET component itself. Different modelling approaches can help understand how changes in atmospheric moisture demand and precipitation can affect future droughts. This can include making use of the simulated soil moisture, evaporation and runoff calculated in the UKCP18-RCM itself (Pirret et al., 2020a), as well as land surface modelling and hydrological modelling approaches which are valuable to shed light on projected changes in different components of the hydrological system (e.g. Lane and Kay, 2021; Kay et al., 2022). More generally, this work raises the question of how these changing drought characteristics translate into impacts for agriculture, water resources and ecosystems in GB. Under the current climate, according to the reviewed literature there is little difference between SPI and SPEI in their ability to predict different drought impacts. However, this is likely to change as SPI and SPEI diverge due to increasing PET. Understanding how the projected increases in atmospheric evaporative demand can be expected to affect different drought types through land-atmosphere interactions is therefore of paramount importance for understanding future drought risk in GB.

5

Evaluation and filtering approaches to increase confidence in streamflow drought projections from an ensemble of hydrological impact model-chains.

Synopsis

Despite continual improvement in our understanding of the consequences of anthropogenic greenhouse gas emissions on hydrological hazards such as droughts, decision makers are faced with much uncertainty for the future. This uncertainty is not only due to unknown future socio-economic developments affecting future greenhouse gas emissions, but also due to uncertainties in the climate response to those greenhouse gas emissions and interactions between catchments and those climatic changes. Hydroclimatic impact modelling chains, in which a hydrological model is driven by (post-processed) climate model simulations, are important tools to support adaptation efforts in essential sectors such as water supply. Therefore, it is important to understand how the individual links in these hydroclimatic impact modelling chains contribute to the uncertainty in projected impacts. In this study, an ensemble of model chains consisting of 12 UKCP18 regional climate model(RCM) simulations (same climate model structures with perturbed parameters) and 4 conceptual hydrological model structures with similar complexity, each with hundreds of parameter sets is used to simulate naturalised future streamflow drought frequency for the Wensum, a groundwater dominated catchment in one of the driest regions of Great Britain, for time slices representing

two levels of global mean warming: 2 °C and 4 °C above pre-industrial levels. The ensemble of model chains is evaluated based on six objective functions over a dry period, the robustness of the hydrological models' performance under a drying climate, and their drought frequency biases due to the hydrological model and the driving RCM. Finally, the sources of uncertainty and the distributions of projected drought frequency are compared before and after rejecting ensemble members based on all these criteria. In addition, the projections are compared to those produced using GR6j (with one parameter set), a hydrological model used by the company managing water for this region (Anglian Water). The detailed evaluation managed to reduce all sources of uncertainty, with the greatest decrease obtained for the contribution of the hydrological model structure and parameter set. The uncertainty due to the hydrological model structures was generally found to be smaller than that due to their parameter sets, yet this might be underestimated due to the relative similarity of the conceptual model structures used. Using the ensemble members that met all evaluation criteria, it was found that drought frequency in the Wensum catchment is projected to increase depending on future global warming, with far greater relative increases for more severe droughts. This could lead to important impacts due to the importance of the Wensum for ecology and as a water resource with direct abstraction. Importantly, this study shows that using an ensemble of hydrological model structures and/or parameter sets for climate change impact modelling can yield a wider spread of plausible future changes compared to when (as in many applications) only one hydrological model is used.

5.1 Introduction

Hydrological impact modelling is of immense value in informing climate change adaptation efforts, and is widely used and relied on by decision makers in government bodies and companies. With high stakes resting on their outcomes, it is especially relevant to understand the uncertainty in their simulations. A typical hydrological impact modelling study is constituted by different modelling stages, e.g. a hydrological model might be driven by a bias-adjusted regional climate model simulation, downscaled from a global simulation forced with a particular emissions scenario. Each link in such a hydrological impact

modelling 'chain' introduces a measure of uncertainty to the resulting projections, leading to a 'cascade of uncertainty' (Smith et al., 2018b).

Many studies have looked at attributing the total uncertainty of the projections to these different sources. The driving climate model is often understood to be the largest contributor of uncertainty in hydrological climate change impact studies, with a smaller but non-negligible contribution from the hydrological model (Wilby and Harris, 2006; Arnell, 2011; Gosling et al., 2011; Addor et al., 2014; Vetter et al., 2017; Krysanova and Hattermann, 2017; Samaniego et al., 2017). However, there is great variability in the relative importance of uncertainty sources between regions (Hagemann et al., 2013) or individual catchments (Arnell, 2011; Gosling et al., 2011; Addor et al., 2014; Vetter et al., 2017; Hattermann et al., 2018; Melsen et al., 2018; Visser-Quinn et al., 2019), seasons (Wilby, 2005; Arnell, 2011; Bae et al., 2011; Gosling et al., 2011; Addor et al., 2014; Bosshard et al., 2013; Hattermann et al., 2018), projected changes of precipitation, temperature and/or PET (Gosling et al., 2011; Melsen et al., 2018), streamflow regimes (Gosling et al., 2011; Bosshard et al., 2013; Vetter et al., 2017; Hattermann et al., 2018; De Niel et al., 2019; Visser-Quinn et al., 2019) and even hydrological hazard characteristics of interest (Visser-Quinn et al., 2019). Specifically, the hydrological model has previously been found to be more influential (or even dominant) for streamflow projections focusing on droughts or low-flows (Vansteenkiste et al., 2014; Vetter et al., 2017; Hattermann et al., 2018; Chegwidden et al., 2019; De Niel et al., 2019; Visser-Quinn et al., 2019), in groundwater-dominated (Visser-Quinn et al., 2019) or glacier-dominated (Addor et al., 2014) catchments, in more water-limited regions (Melsen et al., 2018; Chegwidden et al., 2019) (although Hagemann et al. (2013) concluded the contrary), and in regions with large projected increases in temperature (Melsen et al., 2018). This means that the hydrological model structure and/or parameter set can be expected to have a substantial influence on resulting projections of hydrological drought in groundwater-dominated catchments in the comparatively dry area of the East of England. Indeed, Lane et al. (2022) showed that hydrological model parameter uncertainty was associated with substantial uncertainty in this region compared to the rest of GB, and related this to difficulties in calibrating catchments in this region.

Calibrating and evaluating hydrological models and hydrological impact modelling chains comes with several key methodological considerations which can potentially affect the uncertainty in the resulting projections. First, climate model errors propagating through the (simulated) hydrological system can induce errors in projections of streamflow statistics (Teutschbein and Seibert,

2012). This is commonly addressed by post-processing climate model outputs to reduce errors in statistics of precipitation and other variables before using them to drive hydrological models, in a process called bias correction or bias adjustment (see Chapter 3). However, even if bias adjustment methods are able to almost fully remove biases in statistics of daily values, they can still leave substantial errors in multi-day precipitation metrics (Addor and Seibert, 2014). This is indicative of errors in the modelled temporal variability, which, after non-linear propagation through (simulated) hydrological processes, may result in simulated streamflow errors even if the distributions of the hydrological model input variables would perfectly match observations. In order to evaluate the quality of (bias-adjusted) climate model output for representing non-linear and multi-variate hydrological processes, climate models can be evaluated based on offline streamflow simulations (Hakala et al., 2018; Dakhlaoui and Djebbi, 2021). Hakala et al. (2018) showed that even though quantile mapping generally improved streamflow simulation in most cases, errors in some streamflow metrics may indeed persist, and low flows were relatively less improved than high flows.

Second, the limited transferability of hydrological models and their parameter values to different climatic conditions poses an important challenge to the calibration and evaluation of hydrological models for climate change impact studies. In particular, some previous studies indicate that an evolution toward substantially drier climatic conditions compared to the period used for model calibration could be particularly challenging for many lumped conceptual rainfall-runoff models (Vaze et al., 2010; Coron et al., 2012; Motavita et al., 2019; Bai et al., 2021). Moreover, the evolution and termination of multi-year droughts is difficult to represent well in some hydrological models (Trotter et al., 2022; Fowler et al., 2020), and Fowler et al. (2020) hypothesise that this might especially lead to problems for runoff projections in low-relief catchments with large storage capacities (which is the case for East English chalk catchments) and hydrogeology that enables a non-linear feedback between storage and runoff generation, such as saturation excess runoff.

Multiple studies have proposed ways to improve model calibration strategies to help tackle the issue of uncertainty arising from non-stationarities in catchment behaviour (and thus in optimal hydrological model parameter sets) as a result of climate change. This includes selecting calibration data (e.g. Motavita et al., 2019), choosing objective functions (e.g. Fowler et al., 2018) and developing targeted novel evaluation approaches. Indeed, several recent studies have proposed tests that hydrological models need to pass before applying them in climate change impact studies (e.g. Krysanova et al., 2018; Fowler et al., 2020;

Nicolle et al., 2021; Todorović et al., 2022). For example, the Robustness Assessment Test (RAT; Nicolle et al., 2021) aims to identify dependencies of model skill on interannual variations of climatic variables.

Third, model calibration strategies can also be informed by the specific variables of interest (Pushpalatha et al., 2012). For example, Smith et al. (2019) used a set of six objective functions selected to evaluate low- to high-flows for calibrating an ensemble of hydrological models, resulting in generally skilful reconstructions of streamflow drought for the UK.

This study brings together research on the partitioning of impact model chain uncertainty and hydrological model (HM) evaluation strategies for impact modelling. It quantifies the contributions of different modelling steps to the variability in ensembles whose members pass different evaluation criteria. Specifically, our work contributes to answering the following research questions:

1. Which modelling decisions (climate model parameterisation, hydrological model structure and hydrological model parameters) explain most of the uncertainty in the projected streamflow drought frequency?
2. Can the projection spread be reduced/constrained using a more stringent model evaluation?
3. How is streamflow drought frequency expected to change under climate change in a groundwater dominated catchment in a humid climate where annual evaporation rates are controlled by both moisture and energy availability?

5.2 Study area

The present study investigates drought frequency projections for the Wensum catchment in the East of England, United Kingdom (UK). The East of England is one of the driest regions of the UK, and water supply is facing increasing pressures from a changing climate, stricter requirements for avoiding harm to valuable ecological systems, as well as water demand increases (although these can be addressed through demand management measures) (Anglian Water, 2022a). The work presented in Chapter 4 showed that this region is projected to become more arid and face increasing meteorological drought frequency, particularly for extreme droughts, exacerbated by increasing PET. The relatively large contribution of PET and temperature in increasing atmospheric-based drought frequency in the East of England raises the

expectation that the hydrological model may contribute substantial uncertainty to hydrological impact projections (Gosling et al., 2011; Melsen et al., 2018). In a benchmarking study, Lane et al. (2019) found that catchments in the south-east of England were generally challenging to model (as measured by the NSE (Table 5.1)). Low rainfall amounts, human influences and missing representation of groundwater processes in catchments underlain by the chalk aquifer were all mentioned as possible explanations for this poor performance. The Wensum catchment is underlain by a highly productive chalk aquifer. The catchment gauged at Costessey Mill (national river flow archive (NRFA) ID: 34004) spans 570.9 km^2 and is relatively flat, with altitudes between 5.3 and 96.1 mAOD. A map showing the location of the catchment and its intermediary gauges is shown in Fig. 5.2.1. The Wensum is the main water resource for Norwich, an urban area with a population of 220 thousand (144 thousand in the City of Norwich; Office for National Statistics, 2022). The bulk of the surface water abstraction took place at Costessey Mill until 2019 and subsequently moved to the downstream Heigham intake. Wensum streamflow is also influenced by groundwater abstraction. The Wensum catchment is of significant ecological importance, and it is classified as a Site of Specific Scientific Interest and as a Special Area of Conservation.

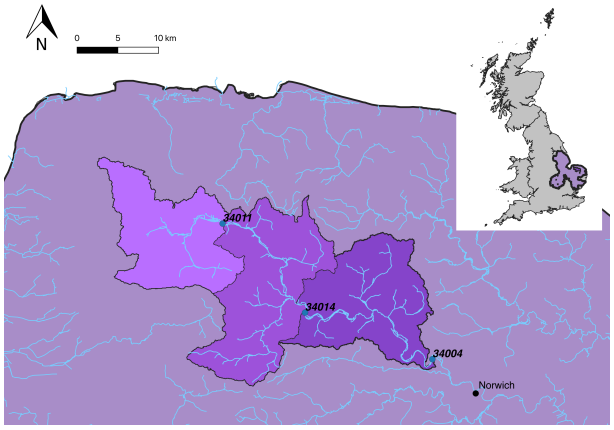

Figure 5.2.1: Map of the Wensum catchment to gauge 34004 (with its two intermediary gauges 34014 and 34011) and its situation in the Anglian Water region. Spatial data sources: National River Flow Archive (catchment boundaries, gauge locations), Anglian Water Ltd. (Anglian Water region), Office for National Statistics (GB boundary*) and Ordnance Survey (rivers*). *licensed under the Open Government Licence v.3.0. Contains Ordnance Survey data ©Crown copyright and database right 2023.

5.3 Data

5.3.1 Observations

Daily precipitation and temperature data were used from the HadUK-Grid dataset (Hollis et al., 2019) for 1966 to 2017. Daily average temperature was computed as the average of daily maximum and minimum temperatures. For observation-based PET, the CHESS-PE dataset was used (Robinson et al., 2020). For both the HadUK-Grid and CHESS-PE variables, the 1 km gridded datasets were resampled by averaging to a 12 km grid matching that of the UKCP18 regional climate projections (Section 5.3.2). Catchment averaged time series of daily precipitation, PET and temperature were then extracted from the resulting regridded datasets using the catchment area shape files available from NRFA.

The significant volumes of groundwater and surface water abstraction in the Wensum and other nearby catchments can make it more challenging to model. Time series of groundwater and surface water abstractions for the Wensum at 34004 and the two intermediary stations, 34011 (Fakenham) and 34014 (Swanton Morley Total) were provided by Anglian Water. These were used to produce the naturalised streamflow time series (see Section 5.5.1) which stood in as observations in this study for deriving observation-based streamflow drought characteristics and for calibrating and evaluating the hydrological models (see Section 5.4).

5.3.2 Climate projections

For producing streamflow projections, input data was used from the regional perturbed-parameter ensemble of the latest set of national climate projections for the UK (UKCP18-RCM; Met Office Hadley Centre, 2018). The 12 ensemble members have a spatial resolution of 12 km over the UK. The projections were produced by the Met Office Hadley Centre using the GCM HadGEM3-GC3.05, with perturbations in 47 different parameters, and then downscaling it with a regional configuration of the same model with matching parameter perturbations. As they were developed specifically for the UK, the UKCP18-RCM projections have already been used for a range of impact modelling studies focusing on the UK. The bias adjusted UKCP18-RCM-derived precipitation and potential evapotranspiration datasets produced in Chapter 3 were used here. Future projections were assessed for 22-year time slices representing two global mean warming levels (+2 °C and +4

°C), see Section 2.4.

5.4 Hydrological models

This study makes use of four models constructed with FUSE, a modular framework for rainfall-runoff modelling (Framework for Understanding Structural Errors; Clark et al., 2008). With FUSE, a large ensemble of hydrological model structures can be generated by recombining modular elements from four 'parent structures', which are each inspired by widely known existing hydrological models:

1. ARVI, based on versions of the Variable Infiltration Capacity Model (VIC) (Zhao, 1977, 1984; Wood et al., 1992; Liang et al., 1994) and borrowing from the ARNO model (Todini, 1996).
2. PRMS, based on the Precipitation-Runoff Modelling System (Leavesley et al., 1983, 1996).
3. SACR, based on the SACRAMENTO model (Burnash et al., 1973; Burnash, 1995; Koren et al., 2004).
4. TOPM, based on TOPMODEL (Beven and Kirkby, 1979; Ambroise et al., 1996; Beven, 1997; Duan and Miller, 1997; Iorgulescu and Musy, 1997).

See Chapter 2 and Clark et al. (2008) for a more detailed explanation of the FUSE models.

To constrain computational demands and to keep the members of our set of hydrological model structures as diverse as possible, this study only made use of these four parent models. These models were set up to run in lumped mode (i.e. the catchment is not divided into hydrological sub-units) for the Wensum catchment.

In addition, the lumped rainfall-runoff model GR6j (Perrin et al., 2003; Le Moine, 2008; Pushpalatha et al., 2011) was used to naturalize the Wensum streamflow time series and also to produce an additional set of streamflow projections (see Section 5.5.1 and Chapter 2). Since GR6j was included because it is used by Anglian Water (and their consultants), parameter sets for GR6j were obtained from Mott MacDonald through Anglian Water.

5.5 Methods

A large ensemble of hydrological simulations for the Wensum was generated using four hydrological model structures from the Framework for Understanding Structural Errors (FUSE; Clark et al., 2008), each with hundreds of possible parameter sets (Section 5.5.2), forced by each of the 12 UKCP18 RCM projections (Met Office Hadley Centre, 2018) for 2 global warming levels, with a time-slice approach based on the years when the centered 25-year mean reach +2 °C and +4 °C above pre-industrial levels (see Section 2.4). PET was derived using Penman-Monteith, and projections were bias adjusted using ISIMIP-3b change-preserving quantile mapping (Lange (2019), see Chapter 3). The ensemble of simulations was reduced by eliminating hydrological models and model chains based on four evaluation criteria (Section 5.5.3). Finally, drought frequency was derived from the simulated streamflow time series (Section 5.5.4), and the uncertainty in the projected drought frequency changes was assessed (Section 5.5.5).

5.5.1 Streamflow naturalisation

There is no standard way of naturalising streamflow, and little guidance is available (Terrier et al., 2021). However, the magnitude of streamflow and groundwater abstractions in the Wensum catchment warrants naturalisation (Jones et al., 2004). To account for the different effects of both groundwater and surface water abstractions and in different (upstream) subcatchments, the reconstitution method (Terrier et al., 2021) was used, in which a hydrological model is used to remove the abstractions (which requires the hydrological model to be able to represent (sub-)surface water abstraction fluxes, e.g. Rameshwaran et al., 2022). For this, the GR6j model was used, which has been calibrated for catchments in the region managed by Anglian Water, including the Wensum. As this model was calibrated in a semi-distributed setup with kinematic wave routing between subcatchments, the same setup and routing scheme was used, obtained from Anglian Water's consultants, Mott Macdonald. To simulate the naturalised streamflow, at each time step in the calibrated GR6j model, the surface water abstraction of that time step is removed from the simulated discharge, and the groundwater abstraction of that time step is removed from the lower soil reservoir. The resulting flow series is used as the naturalised flows.

Table 5.1: Definitions of the objective functions used for hydrological model evaluation and selection. Subscripts: R: observational reference flows, i.e. the naturalised flows; S: simulated flows. n is the number of time steps of the period over which the metrics are calculated. *:flows exceeded 95% of the time

Name (abbreviation)	Definition
Nash-Sutcliffe Efficiency (NSE; Nash and Sutcliffe, 1970)	$1 - \frac{\sum_{i=1}^{n}(S_i - R_i)^2}{\sum_{i=1}^{n}(R_i - \overline{R})^2}$
NSE of the logartihmically transformed flows (logNSE)	*as above but for log(S) and log(R)*
Kling-Gupta Efficiency (KGE; Gupta et al., 2009; Kling et al., 2012)	$1 - \sqrt{(r-1)^2 + (\beta-1)^2 + (\gamma-1)^2}$ with $\beta = \frac{\mu_S}{\mu_R}$ and $\gamma = \frac{CV_S}{CV_R} = \frac{\sigma_S/\mu_S}{\sigma_R/\mu_R}$
Volume Percent Error (VPE)	$\frac{\sum_{i=1}^{n} S_i - \sum_{i=1}^{n} R_i}{\sum_{i=1}^{n} R_i} * 100\%$
Mean Absolute Percent Error (MAPE)	$\frac{1}{n}\sum_{i=1}^{n} \frac{\lvert S_i - R_i \rvert}{R_i} * 100\%$
Absolute Percent Error of the Q95* (APE95)	$\frac{\lvert Q95(S) - Q95(R) \rvert}{Q95(R)} * 100\%$
Spearman correlation coefficient (r_s)	$\frac{cov(rank(S), rank(R))}{\sigma_{rank(S)}\sigma_{rank(R)}}$

5.5.2 Calibration approaches: ensemble generation

Two parameter set generation approaches were used and their resulting parameter sets pooled together: (1) calibration against naturalised flows using shuffled complex evolution (SCE; Duan et al., 1993); and (2) Monte Carlo runs with randomly sampled parameter sets. Performance metrics and their abbreviations used in various steps below are listed in Table 5.1.

First, an ensemble of parameters was derived for each structure using Shuffled Complex Evolution (SCE; Duan et al., 1993), a stochastic optimisation algorithm widely used for optimising hydrological (and other) models, thanks to its efficiency and robustness to local optima. Code for this optimisation method is integrated with the FUSE codebase, and as such SCE has been a frequent choice to calibrate FUSE models in previous studies (Clark et al., 2008; Staudinger et al., 2011; Konapala et al., 2020; Newman et al., 2021; Saavedra et al., 2022). SCE balances exploration and exploitation of the parameter space by letting multiple 'subpopulations' of points in the parameter space evolve towards better performance in their own neighbourhoods, before mixing them so that the information gained by each subpopulation is exchanged. Points in subpopulations can also be replaced by a randomly generated point in the allowed parameter space, potentially discovering promising unsampled areas of the parameter space (reflecting the role of mutation in natural evolution). The default objective function used in SCE implemented in the FUSE source code is the root mean square error (RMSE) between simulated and naturalised flows. However, as least squares optimisation without prior transformation can lead to

poor performance for simulating drying climates (Fowler et al., 2018), FUSE was modified to instead use the RMSE of the logarithmically transformed flows (logRMSE) as the objective function. The log-transformation increases the relative importance of reproducing low and medium flows compared to high flows, which is desirable when the ultimate application is to investigate streamflow droughts. The calibration period used is 1/10/1994 to 30/09/2012. This period includes drier (e.g. 1996 and 2010-2012) and wetter (e.g. 2007-2008) periods, starts after the multi-year drought of 1988-1992 used for evaluation (see Section 5.5.3), and is more than sufficiently long with 17 years (e.g. Motavita et al., 2019). For each FUSE model structure, 5 SCE calibrations were performed, each with a different random seed. In each case, the final optimal set of parameters was kept, plus 24 additional parameter sets sampled from the SCE iterations according to the following procedure:

1. Pre-select parameter set iterations passing all the following criteria for the calibration period: logNSE > 0.75, NSE > 0.5, VPE $< 15\%$ and error in coefficient of variation $< 15\%$.

2. Identify groups of similar parameter sets among these pre-selected parameter sets. This step was deemed necessary because parameter sets in later iterations become very similar as the optimal solution is approached, and a wider range of parameter values is desirable for quantifying parameter uncertainty. The BIRCH (Balanced Iterative Reducing and Clustering using Hierarchies; Zhang et al., 1997) cluster algorithm was applied to first create a tree-based summary of the parameter sets, resulting in a large number of clusters with highly uneven sizes. Although selecting parameter sets through clustering of the values passed in iteration is not standard, it has also been done by Mendoza et al. (2016) in a different setting, using iterations of the same SCE algorithm.

3. The parameter set with the highest logNSE was then drawn from each of the 25 largest clusters of the BIRCH data summary (everything was done separately for each of the SCE runs and for each structure). When these 25 parameters sets include the final result of the SCE run that has already been identified as the optimal solution (highest logNSE in the final cluster), this procedure generates only an additional 24 parameter sets.

Second, to capture a wider range of parameter sets in the uncertainty analysis, the SCE-derived ensemble was complemented with a fully random-generated ensemble. For this, 100,000 values for each adjustable parameter were drawn from an uniform distribution between plausible parameter value ranges (which were extended in range for the Wensum compared to the default limits

implemented; Clark et al., 2008). The FUSE structures were then run with each of these parameter sets. From the resulting simulations, parameter sets leading to logNSE < 0.4, NSE < 0.25 and VPE $> 35\%$ for the calibration period were discarded outright. For one structure (PRMSE), only 29 parameter sets remained after this step, so they were all used. For the other three structures, the 125 best parameter sets were selected from the remaining ones based on 5 objective functions. This was done by combining the top 25 according to each individual objective function, consecutively selecting from the parameter sets that were not yet in the top 25 according to a previously used objective function. This effectively increases the influence of metrics later in the order of selection. The top 25 selections were done in the following order: 1. NSE; 2. KGE; 3. MAPE; 4. APE95; 5. logNSE. The overall ensemble size thus amounts to 12 x (4 x 5 x (25 + 1) + (3 x 5 x 25 + 29)) or 11100 simulations.

5.5.3 Evaluation approaches: ensemble filtering

Four criteria ('filters') were used to evaluate and reduce the initial ensemble: (1) performance scores; (2) robustness; (3) drought frequency bias in the observation-driven flows; (4) drought frequency bias in the reference period RCM-driven flows. Good model performance during the historical period is valuable for improving the quality of impact modelling experiments (van Huijgevoort et al., 2014; Krysanova et al., 2018). First, a multi-objective evaluation was performed over the comparably dry period preceding the calibration period (KLEMEŠ, 1986). In this first filter ('Scores'), HM parameter sets are rejected or accepted based on their performance for observation-driven simulations during the period of 1 October 1988 to 30 September 1994, an intermittent long-term drought period which includes the lowest flow measured in the Wensum. Model performance was measured by a set of six objective functions: logNSE, NSE, KGE, r_s, MAPE and APE95 (see Table 5.1). Simple pass/fail thresholds were used for each score separately to keep this scores-based filter straightforward: logNSE > 0.80, NSE > 0.60, KGE > 0.75, $r_S > 0.85$, MAPE $< 25\%$ and APE95 $< 20\%$.

Second, a variant of the Robustness Assessment Test developed by Nicolle et al. (2021) was used to evaluate the applicability of each hydrological model parameter set in a drying climate. A HM ensemble member is eliminated if the HM shows a significantly decreased performance in drier years for any one of the metrics in the 'scores'. Specifically, for streamflow simulations from 1 October 1966 to 30 September 2017, the spearman rank correlation coefficient between the scores and precipitation minus PET for each hydrological year (so the

annual means for October - September) should not be significantly positive (for logNSE, NSE, KGE and r_s) or negative (for MAPE and APE95) (alpha = 0.05).

The final two filters are based on the absolute value of biases in drought frequency. Third, the difference in drought frequency in observation-driven streamflow simulations and the drought frequency in the streamflow observations is used to evaluate the ability of the hydrological model to capture the frequency of drought (given meteorological conditions of the real world). Fourth, the difference in drought frequency in observation-driven streamflow simulations and the drought frequency in UKCP18-driven streamflow simulations for the reference period is used. This evaluates the ability of the climate model with each of its 12 different parameterisations to simulate meteorology that results in the right frequency of resulting streamflow drought, after temporal propagation and aggregation through hydrological processes as given by a hydrological model structure and parameter set. For both types of drought frequency bias, the thresholds used to exclude a RCM-HM-parameter combination are if its frequency bias exceeds 2% for the Q99 drought, 3% for the Q95 drought, 4% for the Q90 drought and 5% for the Q80 drought. Finally, all evaluation criteria are combined so that all ensemble members need to pass all four evaluation criteria, resulting in the fully filtered ensemble.

5.5.4 Drought definition

As the choice of indicator can significantly influence results (Sutanto and Van Lanen, 2021), it is important to choose indicators and metrics that best reflect the motivation of a drought study. In this case, the drought definition was driven from a water supply perspective, with a seasonally invariant fixed threshold indicating whether or not abstraction is allowed. Thus, a fixed threshold method was used to quantify droughts defined by four different quantiles of the observed daily streamflow time series (1 October 1983 - 30 September 2005) used as drought thresholds: Q99, Q95, Q90, Q80, with Qn indicating the streamflow exceeded n% of the time. Using this definition of drought, the drought frequency was computed as the fraction of drought days in a given time slice.

5.5.5 Uncertainty partitioning

In this work, two time slices were employed representing warming levels and have max. 12 RCM parameterisations, max. 4 hydrological model structures, and max. hundreds of parameter sets for each hydrological model structure,

depending on the surviving members in each filter stage (Section 5.5.3). Each response (defined as the difference in drought frequency between the reference period and a warming level time slice) is thus the combined effect of the RCM parameterisation, HM structure and the specific parameter set used for that HM structure, and their interactions. The following model equation follows:

$$Y_{c,s,p|s} = Y_{o,o,o} + Y_{c,o,o} + Y_{o,s,o} + Y_{c,s,o} + Y_{o,o,p|s}$$

Let the variance components of the RCM parameters combined with internal climate variability, HM structure and HM parameters (conditioned on a given HM structure) be represented by C, S and P, respectively (and indicated by subscript positions 1, 2 and 3 in the equation above). Then, for a balanced design, the total variance V for drought frequency changes projected by a specific global warming level could be partitioned as follows:

$$V = C + S + CS + P|S$$

The parameter contribution is noted as $P|S$ because the number and meaning of hydrological model parameters depend on the hydrological model structure. Two of the filter steps described in Section 5.5.3 (RCM PPE-driven drought frequency bias and all filters combined) can lead to missing combinations of RCM PPE ensemble members and hydrological model structures, resulting in an unbalanced design in which the factors are not independent. As a result, in these cases a small part of the variance ($< 10\%$) cannot be attributed. To avoid wrongfully attributing this portion of the variance to one of the terms, Type I ANOVA was applied twice for each ensemble: once with RCM PPE as the first factor in the model, and once with the hydrological model structure as the first factor in the model. The sum of squares contributions of the RCM PPE and hydrological model structure main effects are then taken from the ANOVA where that factor was first in the model. The interaction terms and residuals (= parameter set contributions) are equal between these two ANOVA runs.

5.6 Results

5.6.1 Model chain evaluation and ensemble filtering

This section shows the evaluation of the RCM-HM modelling chain and subsequently gives an overview of the ensemble composition after applying each filter.

The first evaluation is to assess the performance of the hydrological model ensemble (structure and parameters) at reproducing the naturalised 'observed' streamflow and whether this performance deteriorates significantly during dry years. Figure 5.6.1 shows the *dry period performance* of the hydrological model ensemble according to the six objective functions, combined with their robustness to drier years. The relationship between objective function values and their correlation with dryness is different between the hydrological model structures and parameter generation strategies, although non-robustness to drying is found for a number of models for each objective function. The SCE-sampled parameter sets can be recognised as they are generally more clustered together, and tend towards the better-performing end of the ranges, although especially for TOPM this includes a range of non-robust parameter sets in logNSE, KGE and r_s. However, the randomly generated parameter sets also add a number of well-performing and robust parameter sets. The link between the objective function values and its robustness is generally weak, but varies between the individual metrics. Correlation coefficients between the objective function and the absolute value of its r_s with P-PET range from -0.65 (logNSE) to 0.15 (MAPE), and the weakest correlation was found for the low-flow-focused metric, APE95 (0.10). The correlation coefficients for NSE, KGE and r_s were -0.20, -0.32 and -0.24 respectively. While the signs of these correlation coefficients consistently indicates more robust models for better objective function values, the magnitudes mostly indicate a weak relationship. In other words, a model that performs well according to a standard objective function is not necessarily reliable in a changing climate. For the logNSE, the non-robust parameter sets tend to score worse than the robust ones, whereas for the NSE and the r_s spread of the scores of the non-robust parameter sets is similar to that of the robust parameter sets. For the KGE, the parameter sets that are not robust to drier years tend to have good KGE values, that is, KGE is of limited value to discriminate between hydrological models that perform well and poorly under climate change. Note that models that perform (significantly) worse in wetter years are not eliminated, as only robustness to drier years is deemed important here.

Figure 5.6.2 shows two sources of bias in the drought frequency simulations by hydrological model and RCM. For each combination of RCM parameter and hydrological model (structure and parameters), the x-coordinate is the absolute drought frequency bias in the observation-driven simulation, attributable to the hydrological model, while the y-coordinate is the drought frequency bias of the reference period RCM-driven simulations relative to the observation-driven simulation with the same hydrological model. The latter gives the streamflow drought frequency error component attributable to the bias-adjusted RCM, as

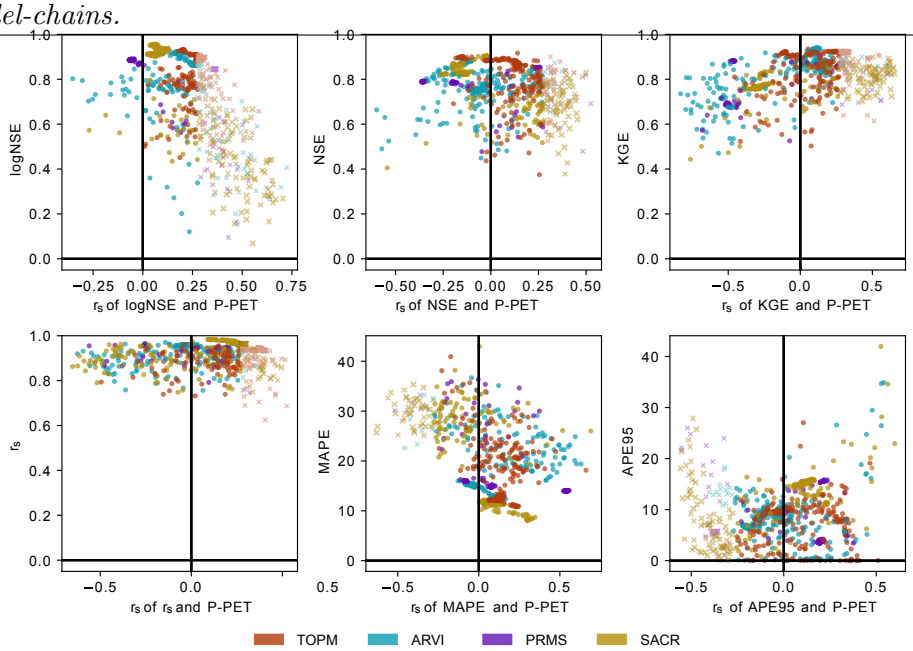

Figure 5.6.1: Objective functions (y axis) and their dependence on climatological dryness (x axis) which represents robustness, for each member of the hydrological model ensemble (driven by observations). Colours indicate the FUSE model structure. Robustness is shown by the Spearman correlation coefficient r_s between the objective function value and P-PE of the different hydrological years. $r_s = 0$ is the most robust, while deviations from zero indicate a lack of robustness. Semi-transparent 'x' markers indicate statistically significant drops in performance for drier years ($\alpha = 0.05$).

mediated through the hydrological processes of a specific hydrological model. The wide hydrological model drought frequency bias spread varies with the quantile threshold used for defining droughts, with a minimum value of minus the observed frequency itself (when the simulated drought frequency is zero). For the Q80, Q90 and Q95 droughts, SACR tends more towards an overestimation of the observed drought frequency, while most substantial underestimations are found in the ARVI and TOPM simulations. The evaluation of RCM-driven streamflow drought frequency reveals errors in the representation of hydrological drought for different RCM PPE members sets. In particular, ensemble member 15 (which projected the weakest increases in SPI- and SPEI-based drought frequency in Chapter 4) consistently overestimates present-day streamflow drought compared to the observation-driven streamflow drought frequency simulations, for most hydrological model structures and parameter sets. In the rest of the ensemble, underestimation of the drought frequency in the RCM-driven flows (compared to the observation-driven flows simulated by the same hydrological model) is more common.

As biases in the distributions of daily precipitation and potential evaporation

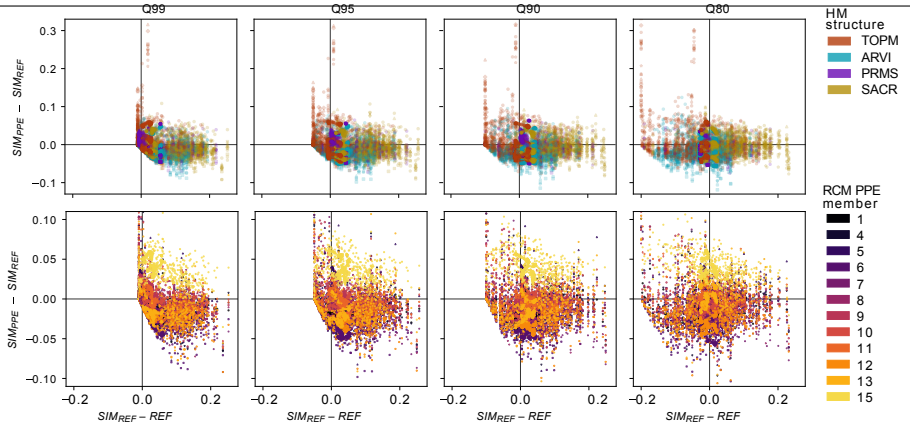

Figure 5.6.2: Comparison of drought frequency biases for all drought thresholds (columns), highlighted by HM structure (top row) and RCM PPE ensemble member with a reduced y axis range (bottom row). Drought frequency bias of the observation-driven streamflow simulations relative to the observed streamflow is plotted on the x axis. Drought frequency bias of the reference period PPE-driven simulations relative to the observation-driven evaluation runs (with the same hydrological model) is plotted on the y axis. The y axis on the bottom row is limited compared to the top row to show more detail on the bulk of the points, excluding outliers.

were successfully adjusted for different months and locations (see Chapter 3), the RCM component of the drought frequency bias can point to errors in aspects which were not (fully) corrected, including the temporal sequence of the precipitation and PET time series, and the relationship between precipitation and PET. However, the RCM contribution to the frequency bias of a given ensemble member can also differ depending on the hydrological model used to transform the RCM data into streamflow time series. This can be seen by variations in the changing vertical spread of the RCM ensemble members along the horizontal axis and among the hydrological model structures. In the second and fourth quadrants of Fig. 5.6.2, the drought frequency error contributions of the hydrological model and the RCM are opposite and therefore partially cancel each other out. For our ensemble applied to the Wensum, this is the case for many ensemble members, primarily with drought frequency underestimations due to the RCM being partly cancelled out by over-estimations due to the HM (fourth quadrant) especially for the more extreme droughts (Q95 and Q99). If the drought frequency bias would be evaluated by comparing the RCM-driven simulations directly with the observations, a number of models would be retained where errors in the hydrological model cancel out errors in the driving RCM data for the reference period, which could lead to biased future projections if the error sources no longer cancelled out under a changing climate. This means it can be important to evaluate the bias contributions of the hydrological model and the RCM separately to reveal opposing model

errors.

Ensemble members are eliminated based on applying different model evaluation criteria ('filters'). Figure 5.6.3 shows how many ensemble members remain after applying each filter individually (performance scores, robustness to dry years, drought-frequency of the HM when driven by observations, or when driven by each RCM ensemble member) or applying all filters together.

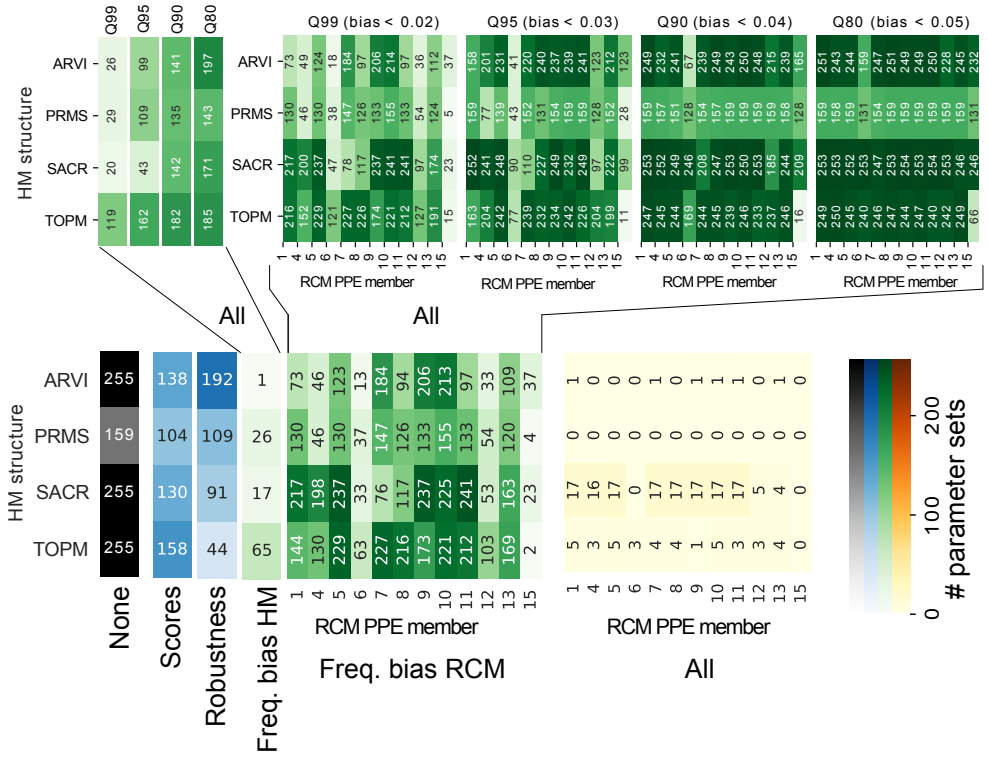

Figure 5.6.3: Overview of the number of ensemble members passing each filter stage. For the two filters based on drought frequency bias, the top row displays the number of passing ensemble members for each drought severity threshold separately, and the bottom row displays the resulting filter requiring that the bias is sufficiently small for all thresholds.

The first two filters (blue in Fig. 5.6.3) respectively show the number of ensemble members with acceptable values ('Scores') which do not deteriorate significantly in a drying climate ('Robustness') for all 6 performance criteria (Table 5.1). The number of remaining parameter sets per hydrological model again shows the added value of evaluating not only the model performance, but also the robustness of model performance to a drying climate. In particular, for the TOPM HM structure, most ensemble members perform significantly worse in drier years ('Robustness') even though a slight majority of the ensemble members shows good values of all 6 objective functions in the dry evaluation period ('Scores'). The opposite is true for the ARVI structure. For the drought-frequency criteria, ensemble counts are given separately for each

drought severity, as well as the number of members that meet the criteria for all four severities. For the two drought frequency-based filters, as the quantile-based thresholds for the more severe droughts (Q99, Q95) become more sensitive to the time period from which they are derived, the acceptability thresholds were chosen to be more lenient (relative to the optimal value) for more severe drought categories. The majority of the hydrological models lead to an unsatisfactory reproduction of the drought frequency in the naturalised observations, despite many more ensemble members performing well over the dry evaluation period as measured by the 6 objective functions ('Scores'). The breakdown of ensemble sizes after filtering for hydrological model drought frequency biases based on the individual drought frequency thresholds show that, for all structures, the most severe drought threshold rejected the most parameter sets. For ARVI and TOPM ensemble members were rejected based on multiple drought thresholds, while for SACR and PRMS most ensemble members with good Q99-drought representation were retained (few SACR models met the Q99 criteria in the first place). Generally, more ensemble members are retained based on the RCM component compared to the HM component of the drought frequency bias. Similar to the HM component, the requirement that the Q99 drought frequency should not deviate from 1% of the time by more than 2% of the time (i.e. Q99 or lower flows should occur between 0 and 3% of the time) eliminates the most ensemble members. The number of ensemble members eliminated for a given RCM PPE member can differ strongly among the hydrological model structures, although for ensemble members 15, 6 and 12 the majority of the parameter sets are ruled out for each hydrological model structure. For RCM PPE member 15, all of the remaining simulations are eliminated based on the other hydrological model evaluation criteria. This ensemble member interacts particularly poorly with TOPM for representing drought frequency, not just for Q99 but also for the less extreme drought thresholds. RCM PPE member 10 retains more than 83% of each HM structure's ensemble, possibly indicating a more realistic representation of precipitation and PET time series leading to hydrological drought. Note that for the two drought frequency bias filters, good performance in all four drought severity thresholds was required for consistency, i.e. to make sure that the same ensemble of model chains was retained for projections of drought frequency for each severity level. For other applications, a more targeted filtering focusing on a specific threshold (e.g. requiring a good representation of Q80 droughts but not necessarily Q99) may be prioritised.

The ensemble remaining after applying all filters (scores, robustness, and both frequency bias sources) consists mostly of SACR simulations, with a smaller number of TOPM ensemble members (most likely due to less robust ensemble

members) and one ARVI parameter set forced by 6 out of 12 RCM PPE members.

5.6.2 Drought frequency projections

This section discusses the projections of drought frequency change using the remaining model chains before and after filtering the ensemble based on different criteria. In particular, we discuss how the projections (best estimate and uncertainty range) are affected by the application of the model evaluation filters to the ensemble. Figure 5.6.4 shows the distributions of drought frequency projections resulting from ensembles passing various filters, while Table 5.2 shows the minimum, median and maximum drought frequency change of the fully filtered ensemble ('All', gold violin plots). The lumpy shapes of the distributions reflect similar, related projections based on the same driving RCM PPE members or groups of similar hydrological model parameters (from the SCE-sampling). Generally, most individual filters do not substantially change either the ensemble median drought frequency change or the quartiles of the distribution of outcomes. Only the filter evaluating observation-driven hydrological model drought frequency bias tends to slightly lower the median and quartiles of the drought frequency change projections, primarily for +4°C and the more severe droughts. Applying all filters combined leads to a slightly larger reduction in the median (and quartiles) of the drought frequency change projections, with the largest difference found in the Q99 projections by +4 °C, from a median increase of 0.16 to one of 0.10 after applying all filters. Both extreme tails of the drought frequency change projections are reduced by the combination of multiple evaluation criteria. Of course, the degree to which the various filters affect the number of remaining parameter sets and the range of the projections depends on the thresholds chosen to eliminate poorly behaving parameter sets in each filter.

Figure 5.6.5 explores further the unfiltered ensemble (filters: 'None', grey) and the fully filtered ensemble (filters: 'All', gold) violin plots of Fig. 5.6.4 for each hydrological model structure separately and in comparison with the GR6j-based projections. The distributions of the drought-frequency projections by different hydrological models mostly overlap. Nevertheless, there are some potentially important differences, which are generally larger (in relative terms) for the more extreme drought thresholds and for the higher warming scenario. Among the FUSE ensembles, the ARVI models generally project the greatest, and the TOPM models the smallest, increases in drought frequency. Notably, GR6j (which was used to produce the naturalised flows serving as the observational reference)

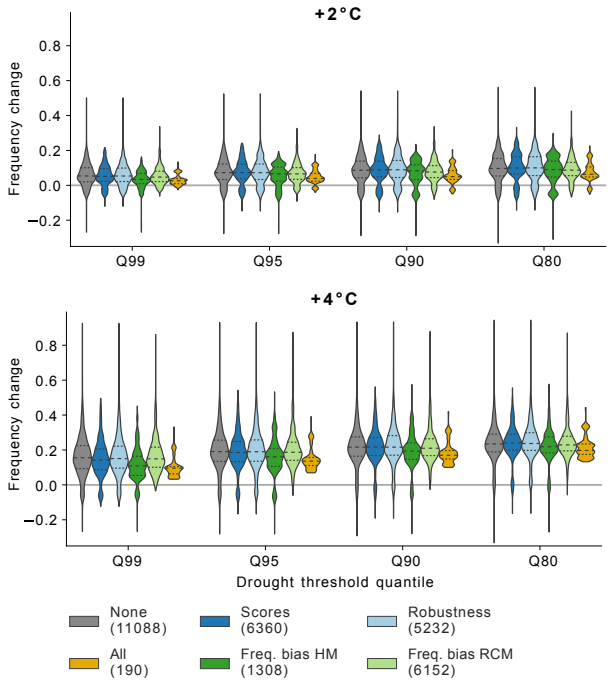

Figure 5.6.4: Violin plots of the change in drought frequency in the different filtered ensembles, for two global mean warming levels and four drought frequency thresholds. The median is given by the dashed lines, quartiles by the dotted lines, and the violin plots are capped at the range of the data shown, without extrapolating. The violin plots are scaled to have the same width to improve readability, while the number of simulations making up each violin is shown between brackets in the legend.

Table 5.2: Minimum, median and maximum of the changes projected by the fully filtered ensemble of model chains (FUSE hydrological models only) for drought frequency (expressed as a fraction of the days in drought) by two global mean warming levels, for streamflow droughts of four severity levels.

	+2 °C			+4 °C		
Drought severity	min	median	max	min	median	max
Q80	-0.051	0.063	0.23	0.13	0.20	0.44
Q90	-0.048	0.050	0.20	0.10	0.17	0.42
Q95	-0.040	0.039	0.18	0.069	0.14	0.39
Q99	-0.025	0.027	0.13	0.032	0.10	0.33

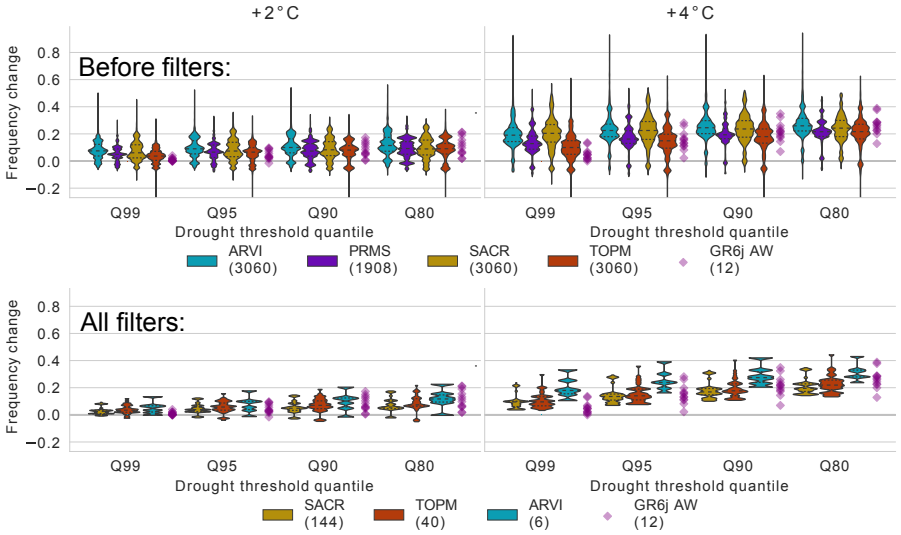

Figure 5.6.5: Violin plots per hydrological model structure for the two warming levels, for the initial full ensemble (top row) and the most strictly filtered ensemble (bottom row). The median is given by the dashed lines, quartiles by the dotted lines, and the violin plots are capped at the range of the data shown, without extrapolating. The violin plots are scaled to have the same width to improve readability, while the number of simulations making up each violin is shown between brackets in the legend.

generally projects relatively small changes, compared to the full FUSE ensemble, in the frequency of the most extreme drought threshold.

In Fig. 5.6.6, the projected drought frequency based on the fully filtered ensemble is shown in more detail, for the reference period and both warming scenarios. They are ranked by Q99 drought frequency so that indicators depicting HM structure and RCM PPE member can show any systematic dependence of drought frequency on those modelling chain choices. Due to their (by definition) more rare occurrence in the reference period, the relative increases in the frequency of more extreme streamflow droughts are far greater than the relative changes in the more moderate droughts 5.6.6. They are also associated with a greater spread in the projections of relative changes. For example, for +2°C, the relative changes in the Q99 drought frequencies range between a 64% decrease and a 6200% increase, and the relative changes in the Q95% drought frequency range between a 61% decrease and an 541% increase.

Figure 5.6.6 further demonstrates the value of using multiple HM structures or parameter sets for projections of hydrological climate change impacts. While the simulations with GR6j are clustered at the low end of the drought frequency simulations for the reference period, they are more spread out through the ensemble for the future time slices (see also Figs. 6.3.4-6.3.7 in Chapter 6), with

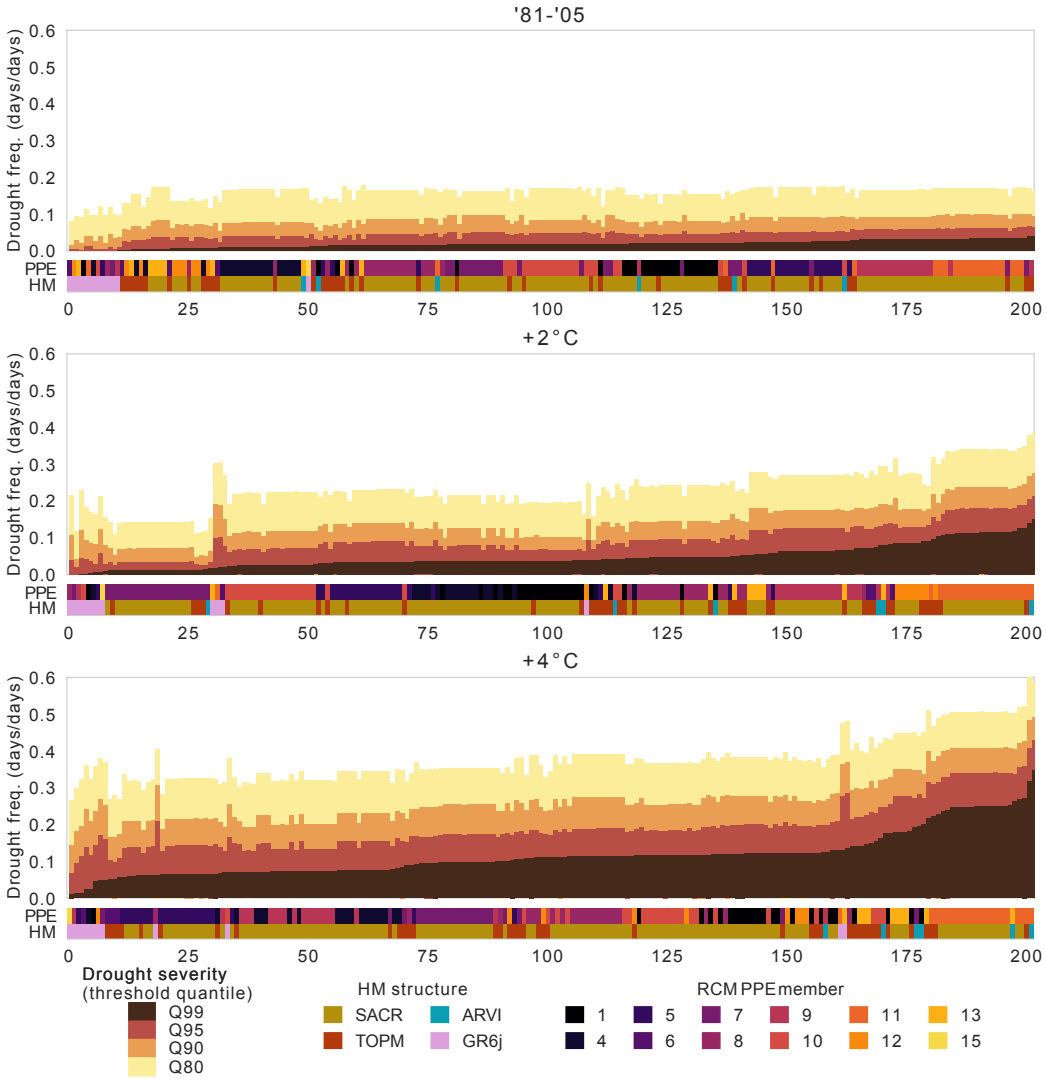

Figure 5.6.6: Projected drought frequency (fraction of days experiencing drought) in the fully filtered ensemble of streamflow drought projections, plus the GR6j-based projections. The projections are ranked from low to high simulated Q99 drought frequency for each scenario. The RCM PPE member and hydrological model structure used for each projection is given by the colour coded stripes below the projections bar plot.

the notable exception of the most severe drought threshold, for which GR6j gives weaker projected increases in drought frequency compared to the rest of the ensemble and compared to other simulations based on the same driving RCM PPE members (11 and 13). For the one surviving ARVI parameterset, all 6 remaining simulations are towards the high end of the projected drought drought frequencies and of the projections driven by a given RCM PPE member, especially for +4 °C and consistently for the different drought thresholds. The simulations with SACR and TOPM, both of which have multiple remaining parameter sets and driving RCM PPE members, both span

about the full range of drought frequency projections (see also Figs. 6.3.4-6.3.7 in Chapter 6). In the Q99 ranking, RCM PPE number 11 projects the strongest drought frequency levels for all time slices. Finally, the reference period simulations on the top row reveal that, while the FUSE ensemble was evaluated for the hydrological model structure and RCM parameter contributions separately, in the combined modelling chain some of the reference drought frequencies deviate substantially from their theoretical frequencies due to the combination of remaining errors in both sources of the same sign.

5.6.3 Quantifying contributions to projection uncertainty

Figure 5.6.7 shows the quantified contributions of the different uncertainty sources to the total variance for unfiltered and filtered ensembles, drought thresholds and both global mean warming levels. Note that the blue bars represent not only the uncertainty from the RCM parameters but also internal variability, as these two sources cannot be separated using the available RCM ensemble. The combination of the regional climate model parameter set and internal variability explains the largest part of the variance for each time slice, drought threshold and filtering stage. The variability in the projected drought frequency changes increases drastically between +2 °C and +4 °C. The combination of hydrological model structure and parameter set generally explains a substantial portion of the total variance in the projected drought frequency changes. Importantly, their relative influence is greater for more extreme drought thresholds and for the stronger warming level. The interaction term of hydrological model structure and the RCM parameter set tends to be small. When eliminating ensemble members with unsatisfactory performance based on the 'scores' filter, their robustness, drought frequency bias sources or all combined, the total variance and the magnitude of the contributions of the individual uncertainty sources generally change, but some filters have a larger effect than others (e.g. the robustness filter has little influence on the variance or individual contributions despite removing a large number of ensemble members). For all drought thresholds, the scores filter led to the greatest reduction in the hydrological model parameter uncertainty contributions, whereas the robustness filter barely affected the relative uncertainty contributions. Eliminating hydrological models with unacceptable drought frequency biases in the observation-driven runs also led to a reduction in the contribution of the hydrological model choices, except for a possible increase in the model structure contribution for the Q99 and Q95 drought frequencies. Evaluating the RCM-driven drought frequency bias led to reductions in the total variance for both warming levels, and in the relative contributions of the

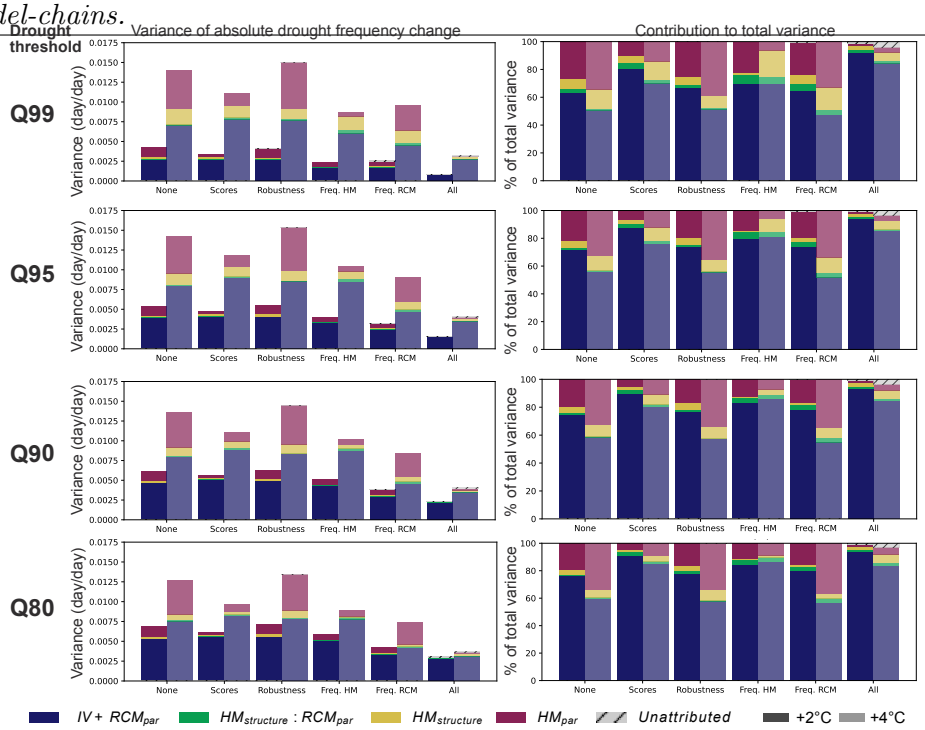

Figure 5.6.7: Quantitative variance partitioning for all ensembles of projected drought frequency changes. Left: absolute variance of the change in drought frequency (day/day). Right: relative contributions of each uncertainty source (i.e. left plots divided by total variance). Each pair of bars gives results for +2°C (dark) or +4°C (pale).

RCM parameters for Q99 and possibly Q95 in the highest warming level. For the 'All' filter, there is a small portion of the variance that cannot be attributed to a single source or interaction due to dependencies introduced by the unbalanced dataset. The 'All' filter, which leaves only ensemble members that did not get eliminated based on any of the four filters, strongly shrinks the relative contribution of the hydrological model components, especially the contribution from HM parameter uncertainties.

5.7 Discussion

5.7.1 Drought projections

This section compares our results for the fixed threshold-based drought frequency projections to previous work on drought and low-flow projections for the East of England and in GB. Previous studies have projected a GB-wide decrease in low flows (Kay et al., 2018, 2021b; Lane and Kay, 2021) or increases in hydrological drought event occurrence (Arnell et al., 2021), peak intensity or severity (Rudd

et al., 2019), with the Anglian region and the south east among the areas with the largest projected drying (Rudd et al., 2019; Kay et al., 2021b; Lane and Kay, 2021; Arnell et al., 2021). Rudd et al. (2019) also found that the Anglian region and the south east were characterised by a greater range of projected drought event peak intensity, total severity and duration relative to the rest of GB. Moreover, Arnell et al. (2021) showed that the projections based on the global HadGEM3.05 PPE resulted in stronger increases in drought frequency and far more uncertainty compared to the ensemble members sampled from CMIP5, for most GB regions including the East of England. Since they used fixed time slices to represent future changes (instead of the warming-based time slices used in this study, see Section 2.4), this is likely due to a combination of stronger warming in the HadGEM3.05 PPE and a stronger drought response for a given warming level. The UKCP18 regional projections used in this study were produced using 12 of these 15 global HadGEM3.05 PPE members, and thus can be expected to simulate similarly strong drying relative to the CMIP5 models. Although their study used only one hydrological model while here multiple hydrological models are used to produce a range of drought frequency projections from each driving RCM PPE member, the behaviour of the driving RCM PPE means the hydrological drought frequency projections produced in this study can also be expected to be relatively dry (though still plausible) compared to if e.g. CMIP5 was used to drive the hydrological models.

The present study focused on one groundwater-dominated catchment in the Anglian region, the Wensum, and found that, after rigorous evaluation, the ensemble of RCM and hydrological model combinations projects drought frequency changes of -0.025 to +0.13 by +2 °C and +0.032 to +0.33 by +4 °C for the Q99 drought threshold. Larger absolute (but smaller relative) changes in drought frequency were found for more moderate drought thresholds. The uncertainty of the magnitude of these changes is large, but this study shows that it can be partially constrained by multi-criteria evaluation of the hydrological models and the RCM-driven streamflow simulations. The large range in plausible projections of future drought frequency (see Figs. 5.6.4 and 5.6.6), especially for the Q99 droughts, highlights the need for robust adaptation resilient to a range of possible (drier) futures – although multi-objective model chain filtering reduced this range substantially, especially discarding the most extreme models due to lack of skill in one or more areas. Finally, the relatively small increases in drought frequency projected with the GR6j model structure compared to the FUSE models is notable. This could be specific to the single parameter set used for GR6j – possibly, a different parameter set might lead to different changes. Another possible contributing factor might be the different treatment of evaporation between GR6j and (some of) the FUSE models, given

that large increases in PET are projected (in SACR and ARVI, evaporation can take place from the lower soil layer, which is not possible in GR6j; and in the FUSE models evaporation is limited by water contents of the soil stores, while in GR6j evaporation is first controlled by concurrent precipitation (representing interception) and then by the content of the production reservoir). However, further research would be needed to understand the causes of the weaker drying projected by GR6j, for example looking at the realism of internal stores and whether these could represent more extreme droughts than the reference period (Fowler et al., 2020, , see also Chapter 7), and to sample GR6j parameter uncertainty (which could be substantial).

5.7.2 Model chain uncertainty quantification: context from the literature

The consistently larger total uncertainty under +4 °C compared to the +2 °C scenario is expected and in agreement with previous studies (e.g. Wilby, 2005; Bosshard et al., 2013; Vetter et al., 2017). While for each time slice the contribution from hydrological model decisions was non-negligible, the combination of climate model uncertainty and internal variability contributed the largest portion of the total uncertainty. Internal variability has previously been shown to make up a potentially substantial source of uncertainty in hydrological impact modelling studies (Vidal et al., 2016; Chawla and Mujumdar, 2018; Chegwidden et al., 2019). The influence of internal variability on projected changes is also expected to be more important for near-future time slices, relative to the strength of the forced climate change (Hawkins and Sutton, 2009; Gao et al., 2020). As such, our finding that the choice of driving RCM ensemble member explains a larger majority of the variance for the lower warming level than for the higher warming level is likely associated with this relatively larger contribution of internal variability rather than due to the GCM-RCM parameter set alone.

In this study, the importance of the hydrological model was found to be generally greater for changes in more extreme droughts. However, for a very different catchment, Bosshard et al. (2013) found that over the full streamflow range, the relationship between contributions of uncertainty sources and streamflow quantiles is non-monotonic, depending on the underlying hydrological processes linked to streamflow quantile ranges. The larger contribution of hydrological model structure and parameter uncertainty for the +4°C scenario is in agreement with previous work showing that the low flow changes become more dependent on the hydrological model under stronger

warming scenarios (Bosshard et al., 2013) and further in the future (Bae et al., 2011; Addor et al., 2014; Vidal et al., 2016).

In the present study, the uncertainty contribution from the hydrological model parameter set was generally larger than that from the hydrological model structure, but was also more substantially reduced after filtering the ensemble. The contribution of the hydrological model parameter choice to the resulting uncertainty is influenced by the methods used for generating the parameter sets and defining which parameter sets behave satisfactorily. In agreement with our results, the few studies which attempted to study separately the uncertainty sources of the hydrological model structure and the parameter set found that both decisions contributed to the spread in resulting projections (Wilby and Harris, 2006; Bastola et al., 2011; Vansteenkiste et al., 2014; Mendoza et al., 2016), although in the setup of De Niel et al. (2019) the choice of parameter set made little difference. Although our results could suggest sampling hydrological model parameter uncertainty might be more important but also easier to reduce than using different hydrological models, generalisations based on one catchment or streamflow metric should be treated with great care (Mendoza et al., 2016). Moreover, the FUSE models used in this study are relatively similar in structure compared to the wider spectrum of existing hydrological models and land surface models – the FUSE models are all lumped conceptual rainfall-runoff models with architectures of comparable complexity following a common framework, which for example doesn't include an explicit representation of vegetation. As such, the contribution found here of the hydrological model structure may be conservative compared to an ensemble of more diverse models (Vansteenkiste et al., 2014). It is worth noting however that even among these fairly similar model structures, large differences in their ability to simulate low flows exist, as illustrated by the much larger number of SACR than PRMS models (144 vs 0) that passed all the tests.

Finally, the interaction terms between the hydrological model structure and the RCM parameter set in our study are consistently small. However, the two- and three-way interactions involving the hydrological parameter sets were not separated from the main parameter effect in this study. Given the substantial contributions of interaction terms found for some catchments in previous studies (Bosshard et al., 2013; Addor et al., 2014; Vetter et al., 2017; Chawla and Mujumdar, 2018), it is likely that the bulk parameter contribution reported here includes interactions between parameter sets (for a given model structure) and the driving RCM PPE member.

5.7.3 Constraining model chain uncertainty sources

This study illustrated that hydrological impact model chain evaluation is an essential tool for constraining uncertainty of projected changes in drought frequency under a changing climate. After requiring that all models have (1) good performance in six objective functions for a dry historical period, (2) are robust in these six objective functions, and have acceptable (3) hydrological model and (4) climate model contributions to the drought frequency bias in the reference period simulations, the contribution of the hydrological model ensemble to the spread in drought frequency change projections becomes minor, even for the highest global mean warming level and for the most severe drought frequency. The reduction in the uncertainty contribution of the hydrological model structure and parameter sets is relatively greater compared to that of the combination of internal variability and climate model parameter sets. This is partly expected due to the fundamentally different approach in dealing with biases in the consecutive modelling stages: while the simulations of atmospheric variables from the regional climate models were bias adjusted (Chapter 3), the biases in streamflow simulations were not corrected but rather served as a basis for ensemble member elimination. The regional climate models also already resulted from an extensive selection process (Section 2.3, Murphy et al. (2018)). Moreover, the filter approach is unlikely to strongly constrain the uncertainty contribution of natural variability. For the TOPM and PRMS structures, different model chains were eliminated based on different evaluation criteria. For the SACR structure, the Q99 drought frequency bias due to the hydrological model component was the most stringent requirement: most model chains passing this evaluation criterion also performed well on all other tests and when driven by 8 out of 12 driving UKCP18 RCM PPE members. Similarly, for ARVI, the single parameter set with an acceptable hydrological model component of the drought frequency bias at all four thresholds also passes the scores and robustness filters, and produces satisfactory drought frequency when forced by half of the RCM PPE members. Furthermore, this study demonstrated that both the hydrological model performance and its robustness to drier (or wetter, depending on projected changes) years need to be evaluated and can serve as a basis to exclude ensemble members. While for the purpose of this study it was considered that model chains failing one or more of the evaluation criteria should not be used, and show their exclusion reduces uncertainty in drought frequency projections (van Huijgevoort et al., 2014), it is possible that some remaining models still contain important errors or that some eliminated models would have added valuable information to the range of projections.

However, investigating the process representation reasons behind model failure points and to the ranges of drought frequency projections was beyond the scope of this study. Previous work has identified baseflow processes (Vansteenkiste et al., 2014; Zheng et al., 2021), evapotranspiration (Vidal et al., 2016) and, where relevant, snow processes (Vidal et al., 2016) as crucial model components leading to differences in projections of low-flows and drought by different hydrological models. Many model chains were rejected in this study due to biases in their simulated Q99 streamflow frequency, a criterion which likely (implicitly) rejected those hydrological models with poor baseflow simulation and RCMs whose meteorological time series, after temporal aggregation, do not lead to the right storage conditions. As this statistic applies to the tail of a distribution of relatively strongly autocorrelated daily streamflow values, deviations from the observed Q99 drought frequency are also more dependent on natural climate variability (which is why the threshold for satisfactorily performing model chains was chosen to be relatively forgiving). The FUSE structure SACR is the FUSE parent structure with the most complex baseflow reservoir representation, consisting of a tension reservoir and two parallel free-flowing reservoirs as opposed to a single state variable (see also Section 2.5.2). This might have been beneficial for simulating streamflow over the baseflow-dominated Wensum catchment, however, several (more different) parameter sets also remained for the simpler TOPM structure (and one for ARVI for some RCM PPE members). The importance of representing evaporation from the lower soil layer is unclear: neither of PRMS and TOPM represent this, and the former was entirely eliminated while the latter was not. To identify the importance of individual model components (such as the architecture of the upper and lower storage reservoirs) for simulating droughts in the Wensum and other catchments, future research could use FUSE to perform controlled experiments where only one of the model components is varied (see also Chapter 7). Moreover, other studies have proposed promising approaches to evaluate and select hydrological models based explicitly on realistic internal model fluxes and states (e.g. Fowler et al., 2020; Saavedra et al., 2022). Finally, although Dakhlaoui et al. (2017) concluded that the hydrological model makes little difference for the evaluation of climate simulations for a range of streamflow metrics, our work shows that using different hydrological models forced with the same bias-adjusted UKCP18-RCM PPE-member can lead to differences in streamflow drought frequency bias, due to the propagation of remaining errors in the time series of precipitation and PET through different simulated hydrological processes.

5.7.4 Study limitations and potential developments to address them

For evaluating model robustness, streamflow simulation skill was compared between years based on the overall precipitation minus potential evapotranspiration of those years. However, for a relatively long-memory catchment such as the Wensum, the performance of a non-robust model might vary with the atmospheric dryness aggregated over more than one year, which is neglected in our approach. Future applications of the robustness assessment test could consider adapting the length of the aggregation period for the hydrometeorological variable which is being compared to a performance variable to the responsiveness of the catchment to that variable.

To constrain the ensemble of streamflow projections, a threshold-based cutoff was used to either keep or eliminate each ensemble member based on each criterion. This method is intentionally kept simple, but has the limitation that there is no consideration of the model performance beyond a pass/fail classification, and so this information is lost. Along similar lines, our simple multi-objective evaluation criterion of simply passing all threshold-based tests does not further take into account the strengths and weaknesses of different ensemble members. Future work could modify the approach to treat differently those ensemble members that only marginally fail on one criterion from those with substantial shortcomings across multiple criteria.

In our analysis of the contributions of different links in the hydrological model chain to the variance of projected changes, some but not all sources of uncertainty were comprehensively considered. First, climate model parameter uncertainty was sampled by using a perturbed-physics ensemble, but did not sample climate model structural uncertainty as only one climate model was used (HadGEM3-GC3.05) (see also Section 5.7.1). Second, as discussed in Section 5.7.2, the influence of internal variability was not quantitatively separated from the influence of the RCM parameter, and instead only looked at their combined effect in the 'RCM_{par} + IV' term. Third, for other steps in the impact modelling chain only one method was used, and thus the associated uncertainty was not sampled. In particular, only one post-processing or bias-adjustment method was used (ISIMIP3-BA), as well as only one representation of potential evapotranspiration (Penman-Monteith for FAO56 reference crop). The bias adjustment step is not expected to be a dominant source of uncertainty, although its uncertainty contribution may be significant (Addor et al., 2014). Studies looking at uncertainty from different PET formulations used in hydrological climate change impact modelling chains found

it can be a significant source of uncertainty (Kay and Davies, 2008; Bae et al., 2011; Seiller and Anctil, 2016; Thompson et al., 2014). Despite not investigating the influence of these other methodological choices, this study provides improved insight into the variability stemming from hydrological model structure and parameter sets compared to the RCM parameter set, and how this is affected by constraining the ensemble based on multiple criteria.

This study suffers from the usual drawbacks of offline hydrological impact modelling studies where there is no dynamic feedback between the hydrological model and the climate model providing meteorological inputs. Assumptions related to land-atmosphere interactions form another limitation of this study. Namely, the stomatal resistance coefficient used to compute PET in this study was not adjusted for CO_2 fertilising effects or physiological responses to high temperatures or vapour pressure deficit. Future work could including the effect of increasing atmospheric CO_2 concentrations in the PET calculation and this may reduce the projected increase of drought frequency, especially for the higher warming level (Kay et al., 2018).

Finally it is important to stress that the simulations and drought frequency analyses produced in this study are based on naturalised flows, in order to avoid making assumptions about future water use and to work with the limitations of process representation in the FUSE models. However, changes in ground- and surface-water abstraction, such as due to population increase or water saving measures, are also expected to have a major influence on future drought risk.

5.8 Conclusions

Model evaluation approaches were investigated to improve confidence in projected streamflow drought changes for the Wensum catchment in the east of England. To this end, the UKCP18 RCM data were used to drive an ensemble of parameterisations for the four parent hydrological models of the modular modelling framework FUSE.

With increasing global warming, a strong increase in streamflow drought frequency is projected for the Wensum. Increases in the most severe drought class considered, defined by the flow exceeded 99% of the time in the observations, are responsible for a dominant part of the increases, especially for the +4°C scenario. The variability from the RCM PPE plus internal variability is responsible for the largest portion of uncertainty in the projections by +2°C, and remains important or even dominant by +4°C. The hydrological model

structure, parameter and interactions are a larger source of uncertainty for more extreme droughts, and their combined contribution is not negligible in either scenario or for any drought severity class.

The ensemble of projections was constrained based on a set of 6 objective functions applied to measure model performance specifically during dry years, on their robustness to a drying climate and on the representation of streamflow drought frequency by the hydrological model and driving RCM. The ensemble resulting from filtering by all of these criteria projects a reduced range of plausible drought frequency changes. Filtering by 6 objective function values eliminated outlying projected changes of both signs, and substantially decreased the contribution of the hydrological model parameter sets to the total variance of the projections. Filtering on the robustness of these objective functions for a drying climate eliminated a larger number of ensemble members, while leaving the proportional contributions to the variance of future projections relatively unaffected. Evaluating the quality of the drought frequency representation due to the hydrological model changed the distribution shape of the drought frequency projections and changed the relative variance contribution of the hydrological model choices, primarily reducing the parameter set contribution. Finally, the combinations of RCM and hydrological model were evaluated for their ability to represent streamflow drought frequency, i.e. realistic sequences of precipitation and PET inputs that result in streamflow drought through non-linear hydrological propagation of moisture deficits. This revealed likely shortcomings in some individual ensemble members for accurately representing streamflow drought, in some cases only by some hydrological models. Importantly, separately evaluating the quality of the drought frequency representation by the hydrological model and the regional climate model (as temporally aggregated into streamflow by a particular hydrological model) prevented the inclusion of model chains in which one of these error sources (partly) cancel out the other.

For future studies and applications using similar hydrological impact model chains, this study recommends a rigorous ensemble evaluation approach considering multiple aspects of streamflow simulation and projection quality. Despite recent progress, more research is needed to understand which process representations (model components) are the main reasons behind differences between projections by different hydrological models, and how this knowledge can help cope with uncertainty in projections of hydrological hazards. Previous work has identified snow accumulation/snowmelt, evaporation and baseflow processes as important for determining hydrological model uncertainty contributions. Modular modelling framework such as FUSE present

opportunities to systematically quantify the effects of e.g. different evaporation or saturated layer and baseflow representations on the range of low flow or drought frequency projections.

6

Quantifying streamflow drought durations and their projected changes

---

Synopsis

Drought duration is one of the most important characteristics of streamflow drought and as such is often quantified in drought research. In this chapter, alternative approaches are explored for better understanding the temporal structure (alongside the total occurrence of drought days), and applied to the rigorously-filtered ensemble of streamflow projections for the River Wensum resulting from Chapter 5. The analysis shows that both a lengthening of drought events and an increase in the number of events contribute to the projected increase in drought frequency, although lengthening is relatively more important for the more moderate drought threshold considered. Furthermore, the contributions of droughts with different fixed categories of duration are assessed. While this study uses a fixed threshold on a daily time scale for defining drought conditions, the proposed approach can also be applied to other drought indicators (including with other time scales).

6.1 Introduction

The drought duration is one of the main hydrological drought characteristics and is commonly the focus of research (e.g. Van Loon and Laaha, 2015; Rudd et al., 2017; Cammalleri et al., 2020; Baran-Gurgul, 2022; Gebrechorkos et al., 2022). The duration of an event has significant relevance for drought impacts and their management. For example, given the same frequency of drought days, a single

long sustained drought presents a different challenge than if the same drought frequency would be made up of multiple short and intermittent droughts. For example, Wu et al. (2022) showed that dissolved organic carbon (an important measure of water quality) in rivers is strongly correlated with drought duration, and might depend more strongly on event duration than severity. In a changing climate, it is therefore relevant to not only have information on projected changes in drought frequency, but also on projected changes in the temporal structure of droughts.

Different approaches exist for describing the duration distribution of droughts (given a certain total number of drought days). In previous work looking at this characteristic, drought duration has mainly been defined as the length of a consecutive period of drought conditions (Yevjevich, 1967), based on monthly or daily time scales (Sutanto and Van Lanen, 2021). Studies using the Standardised Streamflow Index (SSI) as an indicator for streamflow drought tend to define duration on a monthly time scale since SSI is typically calculated on monthly flow data (e.g. Barker et al., 2016; Wang et al., 2020), whereas previous studies using variable (e.g. Van Lanen et al., 2013; Van Loon et al., 2022) or fixed (e.g. Zelenhasić and Salvai, 1987; Fleig et al., 2006) flow threshold approaches have computed drought duration on either a monthly or daily basis. Streamflow drought duration and its changes have been described in the literature by using different statistics, often looking at the mean duration (e.g. van Huijgevoort et al., 2014; Wanders and Wada, 2015; Tijdeman et al., 2018; Konapala and Mishra, 2019; Rudd et al., 2019; Sutanto and Van Lanen, 2021; Gebrechorkos et al., 2022; Van Loon et al., 2022), but also considering the longest duration (e.g. Barker et al., 2016; Baran-Gurgul, 2022) and/or the distribution and variability of drought durations (e.g. Van Lanen et al., 2013; Gebrechorkos et al., 2022). A targeted closer investigation of the drought duration distribution can yield additional interesting insights. For example, Sutanto and Van Lanen (2021) showed that, for large parts of Europe, the majority of streamflow droughts have a duration of under 30 days (based on a variable threshold method with daily resolution).

Hydrometeorological extremes related to either too much or too little water (and multi-hazards involving both, e.g. Matanó et al., 2022) can have both enormous socio-economic and ecological impacts. While there are large differences in the processes, impacts and spatiotemporal scales involved with droughts compared to precipitation extremes or floods (Van Loon and Laaha, 2015), it can be enriching for drought research to draw parallels and connect with methods and concepts from research on floods and precipitation extremes (e.g. Van Loon and Van Lanen, 2012; Van Loon et al., 2016b; Brunner and

Stahl, 2023). For example, in a study illustrating the severe unevenness of precipitation distributions globally, Pendergrass and Knutti (2018) calculated the number of wettest days during which half of the annual precipitation falls. Although the distribution of drought durations in the Wensum is far less uneven, the present study takes inspiration from their analysis, replacing precipitation intensity with drought duration to investigate for the first time the relative contribution of the most persistent streamflow droughts to the total streamflow drought frequency.

This study follows on the work presented in Chapter 5, and has the following objectives:

1. To develop useful indicators of the distribution of drought durations and of how drought days are dispersed over time.
2. To explore the relative contribution of the longest individual drought events to the overall number of drought days.
3. To explore the influence of the hydrological model structure on these metrics.
4. To better understand the contributions to the projected increase in drought frequency for the Wensum simulated in Chapter 5.

6.2 Methods and Data

6.2.1 Streamflow drought simulations

This chapter uses the filtered ensemble of UKCP18 RCM-driven streamflow simulations from Chapter 5.

6.2.2 Drought duration quantification

As in Chapter 5, droughts are defined using the fixed threshold approach based on four different quantiles (Q99, Q95, Q90, Q80) of the naturalised flows, see Section 5.5.4).

Drought event duration is defined as the length of a period of consecutive drought days. In some cases, droughts can have minor interruptions of only one or a few days where streamflow values lie above the threshold, leading to consecutive dependent droughts which can be considered as part of the same

event for practical purposes (Zelenhasić and Salvai, 1987). Therefore, for droughts with an inter-event time shorter than a critical inter-event time (Zelenhasić and Salvai, 1987), we consider them as the same event and pool their durations together. Sensitivity analyses by Fleig et al. (2006) and Tallaksen et al. (1997) testing critical inter-event time values between zero and infinity in multiple catchments showed little change to the mean drought duration if this critical inter-event time is increased beyond 10 days. Although they recommend a value of 5 days, here, we apply a critical inter-event time of 10 days, as in Van Loon and Van Lanen (2012) and Loon and Van Lanen (2013). As in Zelenhasić and Salvai (1987) but unlike Tallaksen et al. (1997), the inter-event time is not added to the drought duration, so that the sum of drought durations does not become inflated compared to the drought frequency analysed in Chapter 5. Individual minor droughts are not neglected (Zelenhasić and Salvai, 1987; Tallaksen et al., 1997; Fleig et al., 2006), as this would also lead to a mismatch between the total sum of drought durations and the drought frequency (i.e. total number of drought days) in Chapter 5, so the contributions from droughts with any duration are investigated.

6.3 Results

6.3.1 Developing metrics of temporal clustering and expected duration of drought conditions

Figure 6.3.1 shows the distribution of durations of individual streamflow drought events for 1983-2005, ranked from longest to shortest events, using the four different drought severity thresholds. The brighter colours shows the longest drought events that together make up at least half of the total number of drought days (i.e. the frequency multiplied by the length of the time series), while the muted colours show the shortest drought events making up the other half. We class (arbitrarily, but inspired by Pendergrass and Knutti, 2018) the first group as the 'long duration' events, and the minimum duration of these events is annotated on Fig. 6.3.1. The proportion of events with durations exceeding this minimum long duration is larger for more severe drought thresholds. As expected, the number of events is smaller for more severe drought thresholds, and for Q95 and Q99 there are fewer events than the number of years in the analysed time slice, i.e. there are fewer than one event per year on average. Because the Wensum streamflow has an average seasonal cycle with miminum flows in summer, recharge in autumn and maximum flows in winter and a fixed threshold is used, the moderate droughts can be assumed

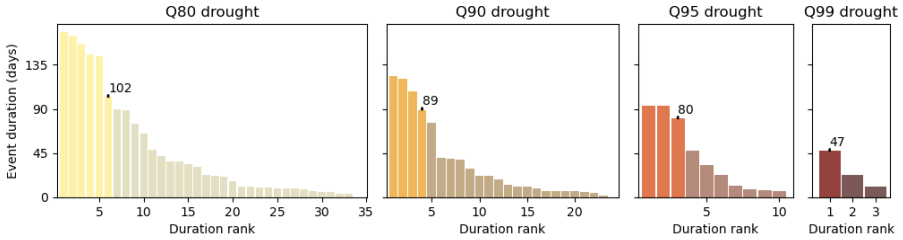

Figure 6.3.1: Distribution of drought durations for the naturalised flows (1/10/1983 - 30/9/2005). The colours indicate the half of the drought days made up by the longest droughts (brighter colours) vs. the half of the drought frequency made up by the shortest droughts (muted colours). The shortest drought that lies in the longer-duration group is annotated with its length in days.

to coincide mostly with summer seasons.

As a measure of expected drought duration, the 'minimum long duration' is defined as the length of the shortest drought in the category of longest events that together make up at least half of the total drought number of drought days (i.e. the rightmost brightly coloured bars in Fig. 6.3.1). This is the smallest duration for which, on a given drought day, there is at least a 50% chance of being in a drought lasting at least this long. Equivalently, if each drought day is assigned the value of the duration of the event to which it belongs, the minimum long duration is the median of those values (within rounding error). For example, at least half of the Q80 total drought frequency is covered by droughts longer than 102 days (almost 3.5 months), which make up less than a fifth of the events. For the Q99 drought, only three separate drought events were found (consisting of consecutive drought days interrupted by fewer than 10 days with streamflow greater than Q99), the longest of which lasted longer than the others combined with a duration of 47 days. The number of droughts shorter (muted colours) and longer (bright colours) than the minimum long duration will also be considered further.

6.3.2 Projected changes in drought duration and temporal clustering of drought days

Having defined our metric for drought duration in the previous section, we can now apply it to the future projections. Figure 6.3.2 shows the 'minimum long duration' for the projections using the fully filtered FUSE-ensemble plus the GR6j-based projections, plotted against the total duration for each of the three time slices. The total drought duration (i.e. drought days) divided by the total days in the time slice would give the drought frequency. For a given total drought duration value, there is a large range in the minimum long drought

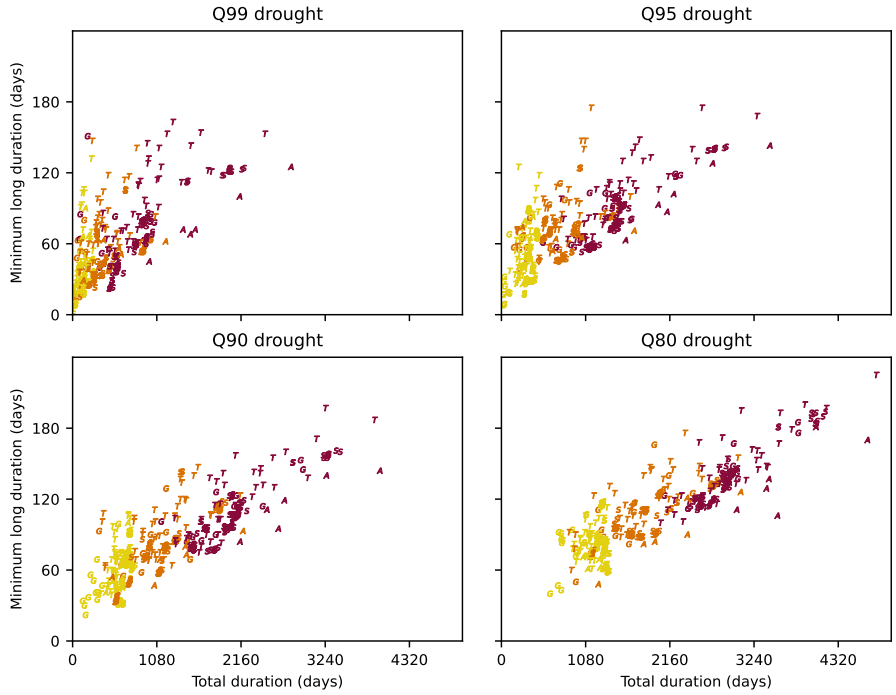

Figure 6.3.2: The minimum duration of the longest droughts that make up at least 50% of the total drought duration (y axis) plotted against the total drought duration (x axis), for the simulations using the fully filtered ensemble of projections plus the GR6j-based projections. Markers give the first character of the name of the hydrological model structures (ARVI, SACR, TOPM and GR6j). Colours and sub-plots indicate global mean warming levels (yellow: 1983-2005; orange: +2 °C; red: +4 °C) and drought severity thresholds, respectively.

duration. This means that different ensemble members projecting the same drought day frequency simulate these drought days as more or less aggregated in time, i.e. in fewer longer events or in more shorter events. This is the case for all time slices, although the spread of potential outcomes increases slightly with global warming level. As the total number of drought days per time slice increases with global mean warming level, the minimum long duration generally also lengthens, indicating an increased risk of long drought events occurring. The increase in minimum long drought duration is more pronounced for less severe drought thresholds. This can likely be linked to prolonged seasonal summer low flows.

The temporal aggregation of the projected drought frequency is further investigated in Fig. 6.3.3, where for each ensemble member the number of individual events longer and shorter than the 'minimum long duration' is plotted against the value of the minimum long drought duration. The increasing number of both shorter and longer drought events, alongside the increasing minimum long duration, indicates that the projected drought

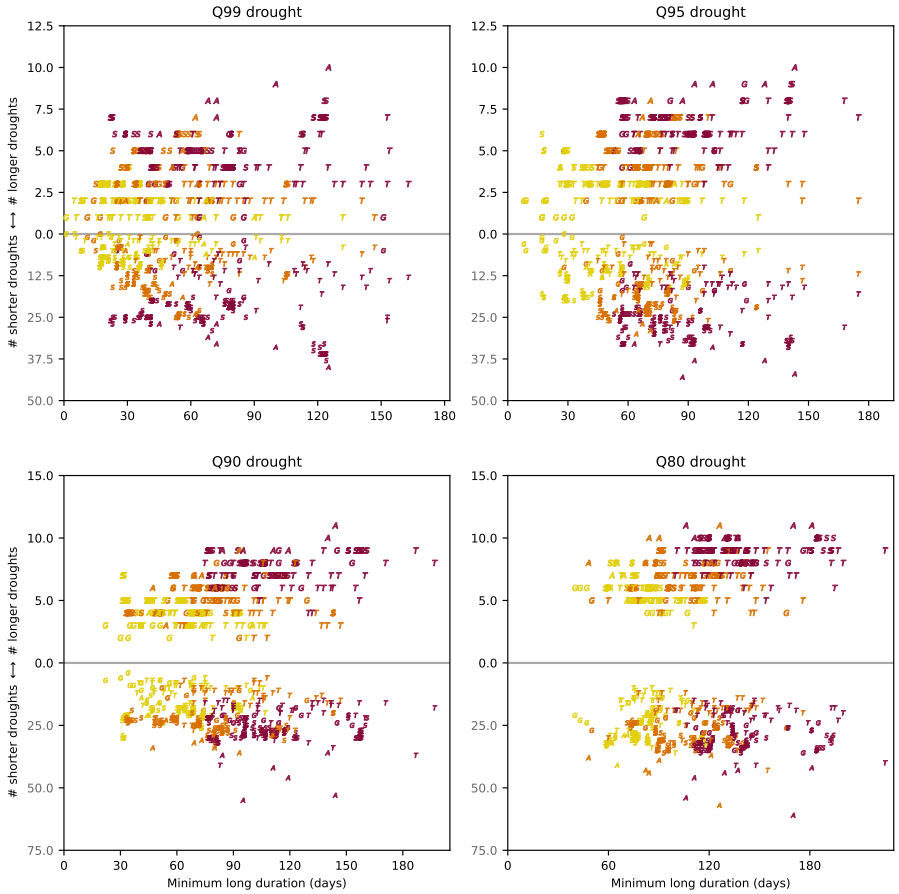

Figure 6.3.3: The number of drought events longer (positive black y axis) and shorter (negative grey y axis) than the minimum duration of the longest droughts that make up at least 50% of the total drought duration, i.e. the number of shortest and longest events contributing to half of the total drought duration. Markers give the first character of the name of the hydrological model structures (ARVI, SACR, TOPM and GR6j). Colours and sub-plots indicate global mean warming levels (yellow: 1983-2005; orange: +2 °C; red: +4 °C) and drought severity thresholds, respectively. Two outliers for Q99 and one for Q95 lie outside the axis ranges.

frequency is driven by a combination of a higher number of droughts and lengthening of the drought events. For the more severe drought thresholds (Q95 and especially Q99), the increasing number of events seems to play a greater role than the lengthening of the events, which plays a greater role for Q80. This can be expected: with longer Q80-based droughts, more of these events can be expected to reach Q99-drought levels at some point, thereby increasing the number of Q99 events. For Q80 (and Q90 to a lesser degree), the relative increase projected in the number of droughts longer than the minimum long drought duration is smaller than the relative increase in the number of shorter droughts, indicating the distribution becoming more symmetrical as these moderate-severity droughts lengthen with global warming.

Looking at the hydrological model structures, there is a stratification in the number of drought events, which is clearer for the number of shorter droughts than the number of longer droughts. For the reference period, TOPM simulates fewer drought events than SACR, with the ARVI simulations falling in between, especially for the shorter droughts, and the GR6j-simulations tend toward fewer and shorter droughts. Under the +2 °C and +4 °C global warming levels, the number of drought events simulated by ARVI (driven by 6 out of 12 UKCP18-RCM PPE members) increases more than for the other model structures, although the events are not necessarily shorter (especially for +4 °C), thereby resulting in a higher total number of drought days (see also Fig. 5.6.5). In the +2 °C scenario and for the severe drought thresholds in the +4 °C scenario, the TOPM-ensemble generally simulates fewer, but longer droughts than SACR (and ARVI for +2 °C). GR6j mostly falls between TOPM and SACR for Q90 and Q80, but simulates fewer and shorter severe droughts (Q99 and Q95).

Up to this point, the projected changes in the drought durations and temporal clustering of droughts were investigated without focusing on the occurrence of droughts with specific durations. However, assessing changes in droughts longer than a set period of time can also be informative. In Figures 6.3.4 to 6.3.7, the drought frequencies shown in 5.6.6 are split into the contributions of drought events of different durations, for the three warming scenarios and four different drought thresholds. The ensemble of model chains is ranked by the total drought duration (and thus by the drought frequency) for each severity level, and the hydrological structure and climate model parameter set are given by the double colour-coded x-axis. For each combination of UKCP18-RCM PPE ensemble member, hydrological model structure and hydrological model parameter set, the combined height of the stacked bars shows the total number of drought days (black) and frequency of drought days (grey) within each time slice. The colours in the bar chart display how many of these drought days are grouped into runs of different lengths, i.e. the contribution of droughts with different durations to the total drought duration or frequency.

Some model chains give very similar projections, as seen by the horizontal lines formed by (almost) equal-height bar plots. These generally occur for model chains using the SACR model structure combined with the same UKCP18-RCM PPE member, and reflect the similarity of the SACR hydrological model parameter sets leading to very similar drought projections. Sudden jumps in the bar heights reflecting different durations, even for similar total durations, can be related on one hand to different hydrological models tending toward more persistent or more numerous droughts, and on the other hand to the internal variability

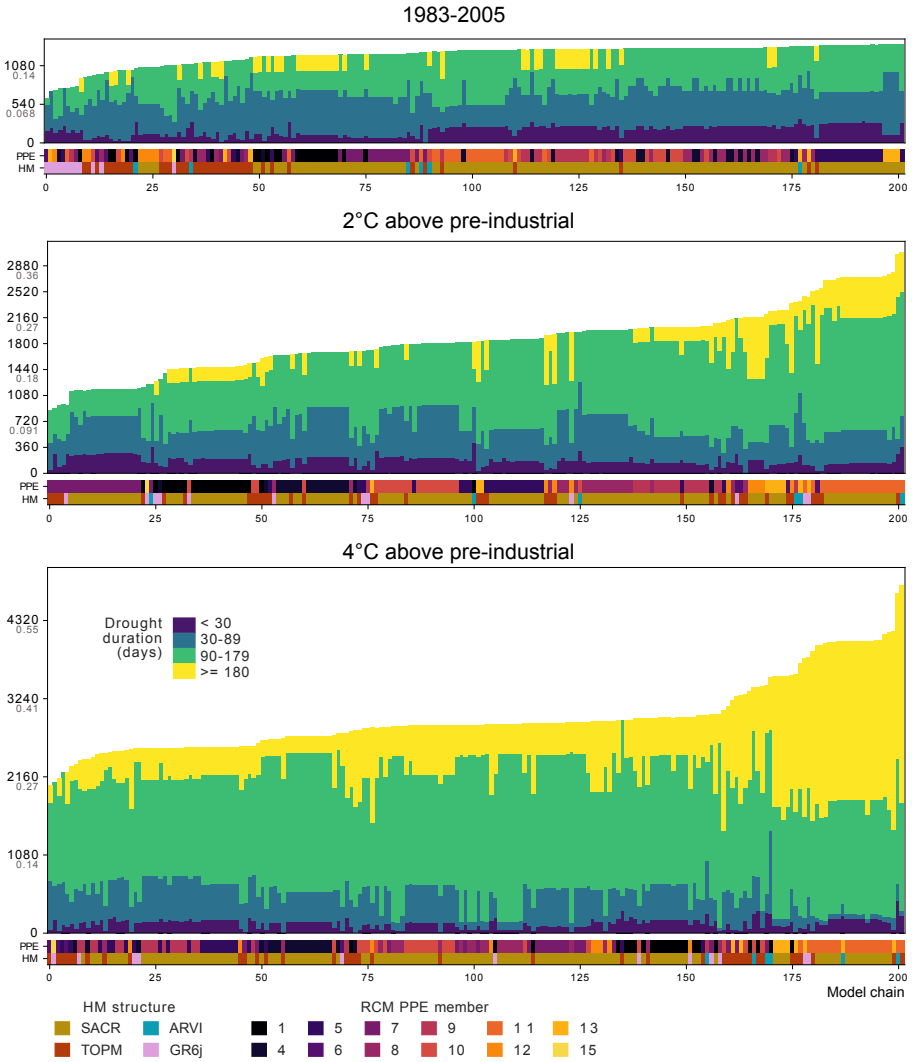

Figure 6.3.4: Projected Q80 drought event durations in the fully filtered ensemble of streamflow drought projections, plus the GR6j-based projections. The stacked bar plot colours indicate the duration of the drought events to which that proportion of drought days belongs: shorter than 30 days, 30 to 90 days, 90 to 180 days or over 180 days. The projections are ranked from low to high simulated total Q80 drought duration for each scenario. The RCM PPE member and hydrological model structure used for each projection are given by the colour coded stripes below the projections bar plot. The figure heights are scaled so that the y-axes are visually comparable. The grey secondary labels of the y-axis ticks show the frequency-equivalent of the total drought duration value, i.e. the total drought duration divided by 22 * 360 days.

and climate model parameter uncertainty reflected by the UKCP18-RCM PPE member, where durations of individual drought events within a given time slice vary significantly. Two initial observations can be made looking at the two ends of the model chain ranking. First, the GR6j-simulations are concentrated toward the left of these figures (projecting lower total drought frequency). This is the case for the reference period, and for more severe drought thresholds also for the future

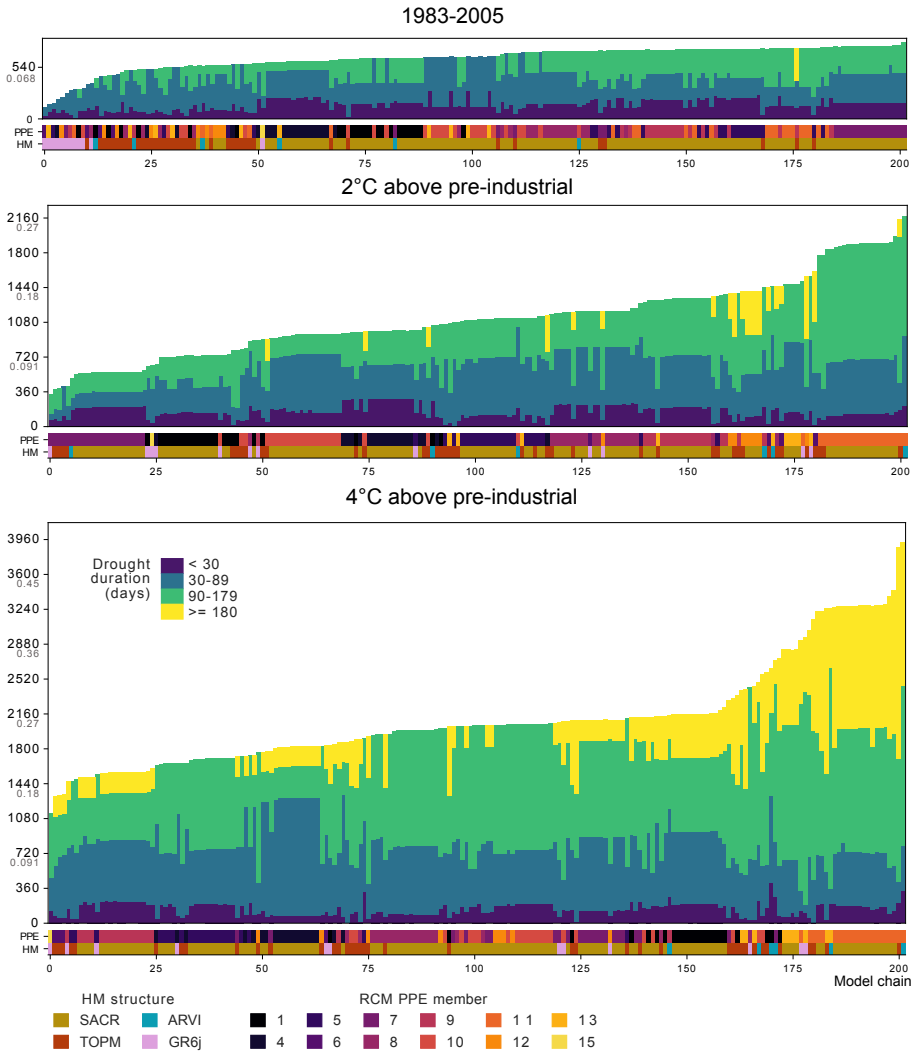

Figure 6.3.5: As Fig. 6.3.4 but for Q90 droughts.

time slices. The implications of this are discussed further below. Second, the 'cliffs' found towards the right of these figures show model chains which project the highest drought frequencies, deviating more substantially from most of the ensemble of projections and the next-in-rank. These cliffs are most obvious for the +4 °C-scenario, and notably involve UKCP18-RCM PPE member 11 driving SACR (with a parameter set leading to drier projections), TOPM and ARVI (with its only remaining parameter set). These model chains passed all filter steps in Chapter 5, and provide valuable information about the range of possible future drought scenarios.

In the reference period for Q90, Q95 and Q99, only one ensemble member (RCM PPE 9 + TOPM) simulates a streamflow drought longer than 6 months (one yellow bar in the top panel of each figure), whereas a wider range of ensemble members do so for Q80 droughts (mostly with SACR, though all

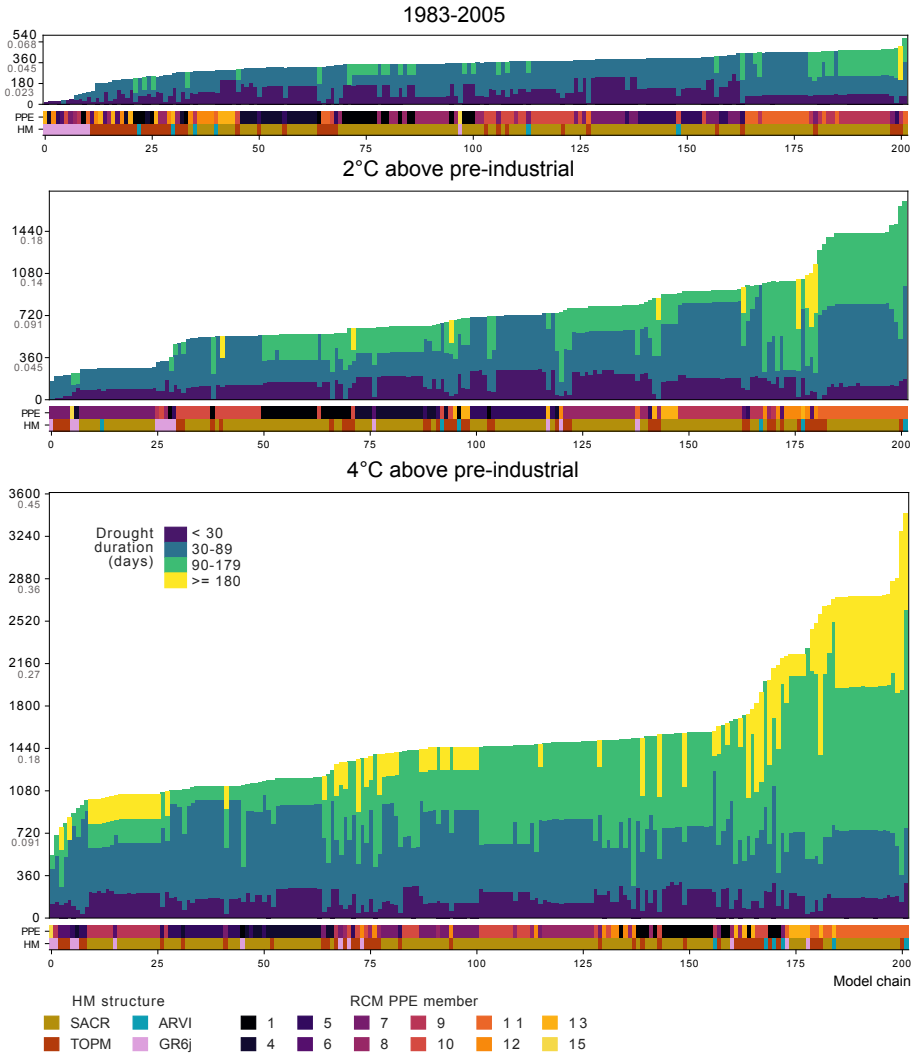

Figure 6.3.6: As Fig. 6.3.4 but for Q95 droughts.

hydrological model structures simulate some droughts of this length). For the more moderate drought thresholds, the increase in drought frequency is driven by an increasing contribution of long sustained dry periods (over 3 and 6 months), while the contribution of shorter droughts is steady or decreases. For Q99, there is an important increasing contribution of droughts sustained for 30-90 days (< 30 days) under progressively increasing global warming levels, combined with increases in droughts longer than 3 months for some ensemble members by +2 °C and for almost all ensemble members by +4 °C. For the high +4 °C warming scenario, streamflow droughts sustained for longer than 6 months are projected to occur in almost all model chains using the Q80 threshold, and even occur for the Q99 threshold in about one quarter of the ensemble, driven by UKCP18-RCM PPE members 7 and 8 (only with TOPM), or 1, 9, 10 and 11 (with ARVI, SACR or TOPM). These long-lasting streamflow droughts are found far less in the +2°C scenario, with minority contributions in

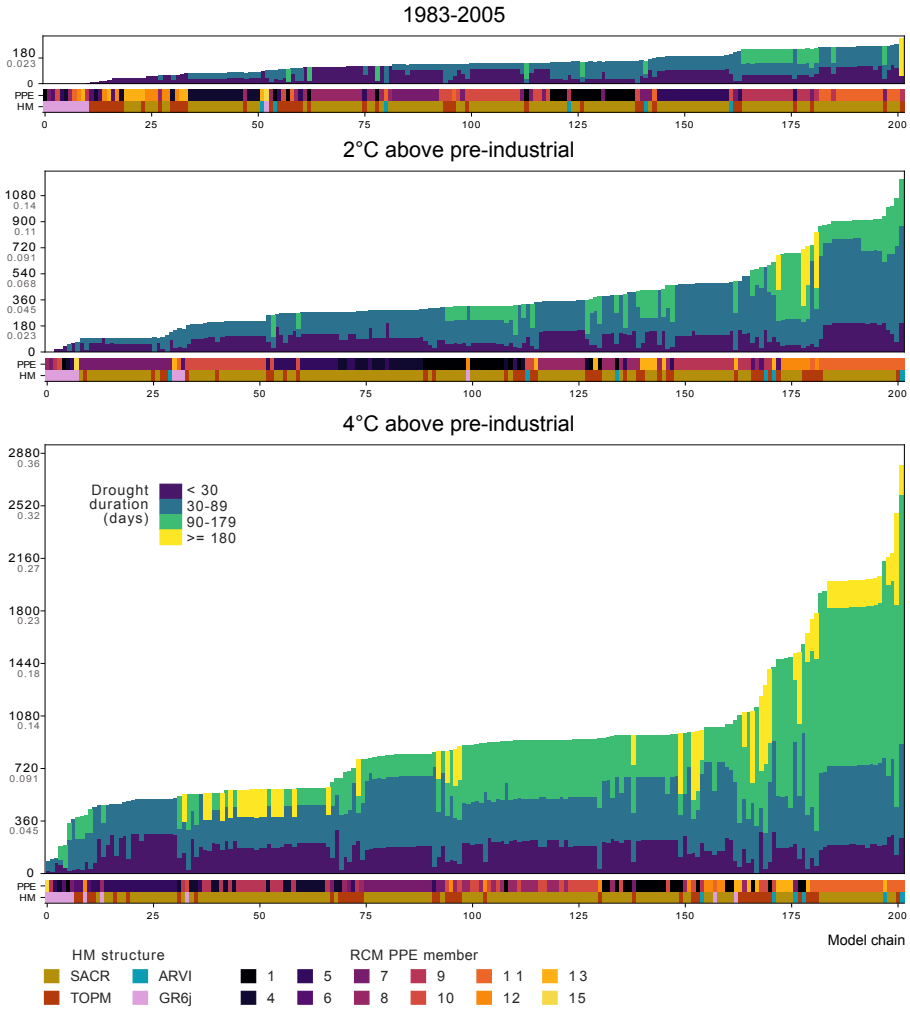

Figure 6.3.7: As Fig. 6.3.4 but for Q99 droughts.

about half of the Q80 ensemble members and occurrences in only a few ensemble members for the more extreme thresholds (UKCP18-RCM PPE members 6 and 12 with TOPM). Periods with stream flow below the Q95 or Q99 that last for at least 3 months (green or yellow in these figures) are quite commonly simulated even with only +2°C of warming.

Depending on the hydrological model and the meteorological forcing from the driving PPE, similar levels of total drought durations can occur but that are more distributed or more aggregated in time (i.e. larger contributions from shorter or longer droughts), as was also reflected in the range in minimum long drought duration for ensemble members with similar drought frequencies (Fig. 6.3.2). Interestingly, the model chains with GR6j consistently underestimate the total number of drought days in the reference period, whereas the rest of the ensemble (which has been filtered to require an acceptably small bias in drought frequency for each hydrological model and

regional climate model combination) simulates higher drought frequency levels, closer to the theoretically expected numbers of drought days matching the drought frequencies of 0.2, 0.1, 0.05 and 0.01 (but still mostly underestimating moderate droughts). Whereas the rest of the ensemble was selected specifically to have a low drought frequency bias (relative to GR6j-derived naturalised flows) from both the hydrological model when driven by observational data and the combination of regional climate model and hydrological model, the GR6j-driven projections were not, and used the same hydrological model as was used for the naturalised 'observed' flows from which the four quantile-based drought thresholds were derived. The low reference period drought frequency simulated by GR6j could possibly point to problems in the UKCP18-RCM ensemble, but not necessarily so: as shown in Fig. 5.6.2, the drought frequency bias in RCM-driven streamflow simulations can vary significantly with the hydrological model used. For the +2 °C and +4 °C global warming levels, the GR6j simulations are more spread throughout the ensemble of hydrological model chains, although for Q99 (and less so for Q95) they are still mostly clustered at the low end of the range (except for the simulations forced by UKCP18-RCM PPE members 11 and 13).

6.4 Discussion

6.4.1 Projected changes for the Wensum: comparison to the literature

The filtered ensemble simulates events with a wide range of durations, depending on hydrological model and regional climate model. The occurrence of Q99 droughts lasting at least 3 months, which for the reference period are present in only a minority of the ensemble members, greatly increases under +4 °C of warming above pre-industrial levels, and even contributes over half of the total number of drought days in ensemble members with stronger projected increases in Q99 drought frequency. The ensemble also simulates very notable increases in what is defined here as the 'minimum long duration' (i.e. the minimum duration of the longest droughts that make up 50% of the total drought days, or equivalently, there is at least a 50% chance of a drought day being part of a drought of this duration), with e.g. 68 days (ensemble median; full ensemble range: 22 to 163 days) under +4 °C of warming for the Q99 drought conditions compared with 30 days (full ensemble range: 5 to 265) in the UKCP18 RCM-driven simulation of present-day conditions 6.3.2.

The projected changes in the temporal aggregation of streamflow drought days

for the Wensum has not been assessed in previous studies. Although Rudd et al. (2019) did not find a projected lengthening of drought event durations for the Anglian region (using an alternative, monthly drought definition), the uncertainty range of the simulated drought durations was projected to increase, and a potential lengthening of drought events was found for the southern English regions. The present study found clear lengthening of projected streamflow drought events under climate change for the Wensum basin, for different seasonally fixed thresholds using a daily resolution. Because seasonally fixed streamflow thresholds were used for identifying drought days, the simulated droughts (i.e. periods with low flow) generally occur outside the recharge season. However, the seasonal timing of these events might also shift: Lane and Kay (2021) found projected delays in the annual 7-day minimum flows across GB, which were significant for the Wensum for some UKCP18 RCM PPE members.

6.4.2 Quantification of drought and its temporal structure

In the work presented here and in Chapter 5, the fixed threshold method is used to identify droughts. This choice was made to reflect the seasonally invariant nature of the environmental flow limits (or 'Hands-Off Flow') below which UK water companies are not allowed to abstract (e.g. Anglian Water, 2022b). Although recent decades saw a paradigm shift towards quantifying droughts as anomalies relative to seasonally varying thresholds, these may be further removed from impacts, and practitioners tend to rely on fixed thresholds to trigger drought responses (Stahl et al., 2020) (although seasonally variable thresholds are also used in practice, e.g. Anglian Water (2022b)). The results for projected changes in the distribution of drought durations presented here for the Wensum would likely be different if a different drought indicator were chosen, for example deficits relative to a seasonally variable threshold (Sutanto and Van Lanen, 2021). However, the proposed approach for assessing the distribution of streamflow drought duration and how it is projected to change can be applied to droughts identified using any indicator, as long as the indicator can make a sharp distinction between drought and non-drought status for a streamflow time series.

By comparing the total drought duration (which is equivalent to the drought frequency (Chapter 5) multiplied by the time slice length) to the minimum long-event duration, this study revealed substantial variability across the ensemble in the temporal connectedness of a given simulated number of drought days. This could be expected to lead to differences in drought impacts and management

implications, although research is needed to establish links between the temporal structure of drought events (given the same long-term drought frequency) and their impacts. With increasing global warming, the 'minimum long duration' generally also lengthens. This means the length of the drought events that will be experienced in at least 50% of the drought days is projected to increase, which is much more interpretable than the often-used mean value of the duration. Furthermore, we showed that the increase in the frequency of more severe (Q95 and Q99) drought days can be attributed relatively more to an increasing number of events than to a lengthening of events, although increasing frequencies for all drought severity levels were due to a combination of both factors.

6.5 Conclusion

Drought duration is an important characteristic with relevance to drought impacts and management. Here, different quantification approaches were proposed to gain further insights into the expected duration of droughts and the temporal structure of a given frequency of drought days. This analysis revealed novel insights for the Wensum case study catchment. The projected increasing frequency of (fixed threshold-based) droughts of different severities (see Chapter 5) can be attributed to both an increase in the number of events and the lengthening of drought events. For the more moderate droughts, lengthening of drought events was more important, while for the most severe drought threshold, the increasing number of events had an important contribution to the increasing number of drought days, alongside lengthening of the drought events. Furthermore, the contributions of streamflow droughts of different set durations to the total number of drought days was also analysed and a remarkable contrast was found between the outcomes from GR6j with a single parameter set and those from the FUSE-based ensemble of model structures and parameter sets. The proposed approach for analysing the temporal connectedness of a given drought day frequency and the distribution of drought durations could be applied to streamflow droughts identified based on different indicators (e.g. variable versus fixed threshold).

7

Conclusions and future research

7.1 Summary

This thesis investigated projected climate change impacts on droughts and explored the sources of uncertainty involved, including modelling choices and drought indicator choices.

First, the UKCP18-RCM simulations of precipitation, temperature and potential evapotranspiration used in this work were evaluated and bias adjusted. This resulted in a substantial improvement in most statistics considered, although biases remained in metrics related to the temporal structure of precipitation time series. The bias-adjusted simulations were used to derive projected changes in atmospheric-based drought characteristics using the SPI and SPEI, two complementary drought indicators. This work showed regionally varying projected increases in drought frequency and extent, which were greater for the more extreme drought category. It also showed projected changes in the distribution of drought durations, which are region- and indicator-dependent but generally include a lengthening of expected drought durations, and for SPEI include a possible increase in multi-year droughts across the RCM ensemble. Moreover, in many GB regions, projected increases in annual-scale SPI and SPEI droughts were driven by projected increasing hydrological summer deficits. Importantly, the projected changes depend strongly on the drought indicator choice, as increases in drought conditions were far greater for SPEI-based projections than SPI-based projections, leading to significant uncertainty in the projections depending on the inclusion or exclusion of PET. Therefore, this work emphasised the importance of improving our understanding of the role of atmospheric moisture demand and evaporation in the propagation of droughts across the diverse range of climatic regimes and hydrogeology found in GB.

Going further, the 12-member ensemble of bias adjusted UKCP-18 projections

was combined with an ensemble of hydrological models to project changes in drought for the Wensum, an East England catchment with great importance as a water resource as well as a site of ecological interest. The ensemble of streamflow simulations was evaluated and filtered based on multiple skill and robustness metrics, which enabled us to constrain (i.e. reduce the uncertainty in) the projected changes in streamflow drought frequency at different severity levels (based on a fixed rather than seasonally-varying drought threshold). The resulting ensemble agreed well on projected increases of streamflow drought frequency, again with far greater increases projected in the frequency of the most extreme category of drought conditions considered. However, there was still a large spread in the magnitude of projected changes, primarily from the combined contribution of internal variability and climate model parameter uncertainty. Finally, the filtered ensemble of streamflow projections was used to explore changes in the temporal structure of these projected drought frequencies. To this end, a novel approach to assess the distribution of drought durations was developed based on an analysis of drought duration distributions in the naturalised historic flows. This revealed that both an increase in the number of drought events and the lengthening of the most often experienced drought events contributed to the projected increase in drought frequency.

7.2 Recommendations

The influence of methodological choices for the simulation, quantification and analysis of droughts was a major focus of this thesis, and as such, several recommendations can be made for both researchers and practitioners.

1. Significant biases were found in UKCP18-RCM simulations of precipitation, temperature and derived potential evapotranspiration. These biases were especially large for precipitation. Because projected changes were found to differ between the lower and higher parts of the distributions, it is recommended to use a bias adjustment method that accounts for different changes along the precipitation and temperature distributions. This seems especially important for applications with a focus on extremes. Following from this work, it is recommended to avoid using the change factor approach if a good representation of meteorological extremes is important, and to preferably make use of the climate simulations themselves.

2. The choice of (atmospheric-based) drought indicator can have a decisive influence on the resulting conclusions regarding projections of future

drought. Researchers (and other data analysts) should, depending on data availability, resources, spatio-temporal resolution and relevant impact driver information, select the drought quantification method best suited to the underlying goal pertaining to impact modelling or process understanding. In particular, given the large role of atmospheric evaporative demand in how climate change influences droughts, take into account that drought projections based on metrics calculated from precipitation alone would likely underestimate projected changes in the frequency and intensity in many impact-related drought variables, whereas SPEI might overestimate future impacts.

3. If sufficient time and resources are available, hydrological modellers working on drought projections should consider using more than one hydrological model structure and parameter set, so that they can sample from a wider range of plausible modelled futures. A targeted evaluation of (calibrated) hydrological models based on their performance under a changing climate is necessary for impact modelling studies, as good performance under the reference climate does not necessarily imply good performance under, for example, a drier climate. In general, a rigorous ensemble evaluation tailored to the hydrological hazards of interest can be undertaken to constrain and improve confidence in an ensemble of hydroclimatic climate change impact projections.

4. As duration can be an important drought characteristic for impacts and management of drought, and this thesis has shown that the distribution of drought durations can change in multi-faceted ways, the concluding recommendations are to (a) consider assessing changes to specific relevant aspects of the temporal distribution of droughts, and (b) use metrics that capture more features than simply the mean event duration. The 'minimum long drought duration' was suggested here as an alternative metric to provide complementary insights into projected changes of drought frequency. Practitioners can consider potential implications of different changes in the length versus the number of droughts contributing to increasing drought frequency depending on their purposes.

7.3 Further remarks

7.3.1 Implications of the implementation of vegetation in the UKCP18-RCM models

The JULES land surface model which provided the land surface component in HadGEM-GC3.05 does not simulate vegetation interactively, and instead makes use of a time-varying prescribed vegetation distribution (Murphy et al., 2018). According to (Pirret et al., 2020b), this prescribed vegetation does not represent increases in LAI and stomatal closure (with their counteracting effects on transpiration) due to increasing CO_2 levels. The prescribed changes in vegetation composition are based on the land use scenarios developed by (Hurtt et al., 2011), implemented as changes in C3- and C4-grass cover compared to broadleaf and needleleaf tree and shrub cover, with constant proportions of plant functional types within both vegetation cover categories (Murphy et al., 2018). There is thus no feedback mechanism in this model configuration through which changes in climate affect vegetation cover and physiological responses, which in turn influences climatic changes.

In reality, vegetation structure and physiology both play an important role in determining land-atmosphere energy and moisture fluxes, and therefore (spatial patterns of) temperature (e.g. Alo and Anagnostou, 2017; Arora et al., 2020) as well as precipitation (e.g. Alo and Anagnostou, 2017; Skinner et al., 2017; Cui et al., 2022). Increasing atmospheric CO_2 concentrations lead to increased water use efficiency and thus lower transpiration by plants due to decreased stomatal conductance (e.g. Field et al., 1995), which is counteracted or tempered by increases in transpiration due to increasing leaf area (e.g. Mankin et al., 2017). The net effect of plant physiological responses to changes in CO_2 is uncertain and varies depending on the region, vegetation type and climate (variability) (e.g. Mankin et al., 2017, 2019; Liu et al., 2020, ; see also Section 4.6.3). Moreover, in addition to the influence of direct anthropogenic land use changes, observed and expected future changes in vegetation cover (Franklin et al., 2016) can further influence regional climatic changes. Of course, vegetation also influences climate change through carbon storage and the uptake and release of CO_2. However, this aspect of vegetation-climate interactions is not further discussed here, because the HadGEM-GC3.05 configuration used to produce the UKCP18-G/RCM simulations did not feature an interactive carbon cycle, and instead the carbon cycle uncertainty was represented by a range of CO_2 concentration scenarios based on those used in the UKCP18 probabilistic projections (Murphy et al., 2018).

Shifts in vegetation properties taking into account these vegetation-CO_2 and vegetation-climate interactions may thus differ from the prescribed vegetation scenario, and therefore lead to differences in the resulting climate projections and derived impact projections, such as streamflow and drought projections for the Wensum. The net effects of the lack of dynamic representation of vegetation on low flow simulations depends on projected changes in vegetation structure, the balance of transpiration changes due to the (opposing) plant physiological responses to CO_2, and vegetation-climate feedback effects emerging from plant physiological responses to climatic changes influenced by vegetation. For future modelling studies, it would be beneficial to include dynamic vegetation in order to better estimate the regional effects of vegetation-climate interactions (Yu et al., 2016).

7.3.2 Bias adjustment

In this work, the climate model simulations were statistically bias-adjusted before they were used for analysis of projected meteorological droughts and driving an ensemble of hydrological models. The use of bias correction is motivated by the aim to improve the usability of the resulting projections, through reducing error in the simulated drought-related variables. This rests on the assumption that, even though the climate models are an imperfect representation of the physical, biological and chemical processes of the climate system and can therefore produce biased simulations, the representation of the climate system is nevertheless of sufficient quality to provide enormously valuable information about potential future developments of this very complex system when forced by different scenarios. Previous studies have shown that bias adjusting climate model output can significantly improve the resulting projections of derived hydrological impact indicators (e.g. Teutschbein and Seibert, 2012; Hakala et al., 2018; Pastén-Zapata et al., 2020). However, no statistical bias adjustment technique is able to correct all error arising from imperfections in complex physics-based models. Consequentially, challenges for bias adjustment include the robust correction of multi-variate relations (Zscheischler et al., 2019) and the temporal structure of simulated variables (Addor and Seibert, 2014). In recent years, novel approaches have been developed to account for the correction (or avoiding deterioration) of, among others, multivariate statistics (e.g. Cannon, 2016) and temporal structures (e.g. Nguyen et al., 2016). There is still discussion concerning the (case study-specific) added value of more complex bias adjustment methods (e.g. François et al., 2020; Tootoonchi et al., 2022), and evidently none of these techniques correct the underlying climate model errors giving rise to the

statistical discrepancies between simulations and observations which are being corrected. Additionally, impact modellers can also bias adjust the climate model-driven simulated impact variables (e.g. streamflow), which is not usually done at present. This would put the (final) statistical correction step closer to the research questions at hand, and may alleviate the effect of propagating (remaining) biases between (bias adjusted) climate model simulation and impact variable of interest. However, this should not replace critical model evaluation, calibration and selection. Further research is needed to assess the potential benefits and pitfalls of impact model bias adjustment approaches.

7.3.3 Naturalised flows

The hydrological modelling work in this thesis was based on naturalised flow time series. This does not fully represent the reality of the Wensum catchment, which is strongly influenced by human influences such as groundwater and surface water abstraction. Basing the projections on naturalised flows was nevertheless identified as an appropriate approach for several reasons. First, given the large influence of water demand on future projections of drought water scarcity, the impact of any water demand assumption was expected to have a substantial influence on the results of the work. The use of naturalised flows neglects the influence of water demand on the hydrological system, rather than assuming a specific constant or changing water demand profile. The substantial uncertainty surrounding future development of water demand depends on socio-economic factors and technological developments, and developing appropriate demand development scenarios was considered outside the scope of this work. Rather, the focus was put on the uncertainty due to model representations of physical processes in the hydroclimatic system, which is the main value of this chapter for informing decision-making. Second, the set of hydrological models used in this work contains simplified representations of physical processes in the hydrological systems, but would require modifications in order to adequately represent the influence of (changing) water demand. Calibrating the FUSE models to non-naturalised flows would lead to an undesirable compensation effect, where the values of parameters describing a (simplified) physical system would also have to account for the effect of human water abstractions. While this approach may still lead to acceptable objective function values, it is very unlikely that the resulting parameter sets would have been suitable for application under a changing climate. However, the influence of (changes in) water demand and management actions is important, and definitely deserves closer consideration in future (interdisciplinary) hydrological impact projection studies.

7.4 Potential directions for future research

This thesis contributed to a better understanding of projected changes in drought characteristics over GB, associated uncertainty sources related to different methodological, modelling and analysis decisions, and the quality of the UKCP18 projections, but many (new and old) questions remain at the end of this work. Below, some outstanding directions for potential future research are outlined.

To start back in the atmosphere, one potential avenue for future work could be to establish links between the atmospheric-based and hydrological drought projections found in this study and the larger-scale atmospheric drivers giving rise to these conditions. This could potentially help understand the nature of the substantial climate model uncertainty found for both meteorological, climatic water balance-based and hydrological droughts (see e.g. Figures 4.5.1 and 5.6.7), as well as help improve streamflow drought forecasting if linked to shorter-term weather forecasts. The responses of thermodynamically controlled variables to climate change are typically more robust than indirect changes in dynamic processes (Shepherd, 2014). Previous studies have established regionally varying relationships between 30 fixed weather types and meteorological drought risk for the UK (Richardson et al., 2018), and gave an overview of biases and projected changes in the occurrence of these weather types in the UKCP18 projections (Pope et al., 2022). In Fig. C.0.1, a preliminary exploratory analysis is shown of the relative occurrence of drought in months with each of these weather types (for all 30 plus the 8 'summary' weather types), for the first UKCP18 RCM ensemble member. For many weather types, the relative occurrence of drought conditions changes between the time slices representing different warming levels, and for some of the 30 weather types there are noticeable differences between SPI and SPEI in this (e.g. WT7, WT27). Future work could look at the changes in weather types plus changes in the (concurrent or lagged) relationship between these weather types and (atmospheric, soil moisture or hydrological) drought variables of interest, possibly after aggregating temporally over a moving window or seasonally, and split by GB regions (see e.g. Richardson et al., 2018; Pope et al., 2022), as well as the implications of these biases for impact studies relying on these climate projections (e.g. Addor et al., 2016). Isolating components of projected changes by thermodynamics or dynamical processes in the projected drought changes may help contextualise, understand and reduce the associated uncertainty.

The findings in Chapter 4 emphasise the importance of understanding how

increasing temperature and atmospheric evaporative demand impact drought risk across GB. Further work on drought propagation processes across the range of hydrogeological properties and climatic regimes of GB would be valuable to better understand how climate change can be expected to impact water resources and other drought-affected sectors in GB. For example, an outstanding research question with high relevance for drought forecasting and water management in GB is how the relevant response time scales of hydrological drought development might change under a changing climate with higher potential evapotranspiration rates. Another key research area not explicitly addressed in the present study is the potential changes in vegetation controls on the propagation of meteorological deficits to streamflow and groundwater droughts, such as the plant physiological effects of rising CO_2 concentrations and of increasingly high levels of vapour pressure deficit on transpiration fluxes (Vicente-Serrano et al., 2022). In future modelling studies addressing these questions, the effects of climatic changes could be considered in tandem with the effects of (historic, planned or hypothetical) human-caused land use changes on the surface water balance. Moreover, observation-based studies are invaluable for understanding drought processes (e.g. Bloomfield et al., 2019), and can help verify and improve models and modelling approaches relied on by decision makers. Crucially, given the urgency of adaptation efforts, translation into practical improvements to the (modelling) tools and guidance relied on by decision makers should be prioritised (for example, by investigating the implications of improved process understanding of drought propagation changes under increasing AED for parameter uncertainty in hydrological models relied on by practitioners, or by assessing how potential propagation response time changes affect water resource systems).

In this work, only the four FUSE parent models were used, and significant differences between these structures were found in Chapters 5 and 6, including for evaluating RCM-driven streamflow drought frequency simulations and the temporal structure of simulated drought days as quantified through drought duration statistics. A next step could dig deeper into understanding the reasons behind these differences by exchanging individual model structural components in a controlled way, thereby fully exploiting the potential of FUSE. Examining not only the resulting statistics but also hydrological model state variables and fluxes could further illuminate the causes of differences in model behaviours (e.g. Fowler et al., 2020; Saavedra et al., 2022). Given the strong projected rise in PET, the sensitivity of projected changes in streamflow (and groundwater) droughts to the adequate representation of its effect in hydrological models, as well as the formulation of potential evapotranspiration (e.g. related to assumptions about vegetation responses to increased CO_2 concentrations and

warming Kay et al., 2018), deserves specific attention. FUSE has already proven an interesting tool in this work and other studies (e.g. Saavedra et al., 2022). However, investing in its software development and documentation (including a publicly shared system of FUSE model naming conventions) would have the potential to significantly benefit the hydrological modelling community in academia and in practice. One possible option could be to write a wrapper for FUSE in higher-level languages such as Python or R, which are generally more widely known by likely potential users than FORTRAN (in which FUSE is written).

In Chapter 5 streamflow drought projections were made for only one catchment, chosen due to its importance as one of Anglian Water's water resources with direct river abstraction. Future work could expand the ensemble modelling and evaluation approach to other catchments, including catchments feeding into reservoirs (e.g. the Nene, Bedford Ouse and Welland). A next step could then be to drive water resources simulations with industry models such as AQUATOR or PyWR, and investigate the propagation of uncertainty sources into projections of drought impacts on reservoir levels.

The simulations from Chapter 5 present an ensemble of specific plausible futures. In a next step, extreme events and streamflow time series could be extracted from this rigorously evaluated set of streamflow simulations, to represent a range of events that would challenge water resources systems. Working with such derived narratives or 'storylines' to crystallise (Smith et al., 2018b) the potential changes can help develop adaptation strategies by developing tangible 'what if' scenarios (Shepherd et al., 2018; Shepherd and Lloyd, 2021). Although storylines can be composed by perturbing historic events via applying changes to some climatic variables (e.g. Chan et al., 2022), selecting events produced from simulations by a full hydroclimatic impact modelling chain could potentially lend increased plausibility to the resulting scenarios, and ensures that important information on changing variability provided by the climate models is taken into account.

8

Bibliography

Addor, N. (2020). Framework for Understanding Structural Errors 2. GitHub Repository available at `https://github.com/naddor/fuse`. [Accessed: 2020-05-11].

Addor, N., Rohrer, M., Furrer, R., and Seibert, J. (2016). Propagation of biases in climate models from the synoptic to the regional scale: Implications for bias adjustment: Circulation Biases in RCM Simulations. *Journal of Geophysical Research: Atmospheres*, 121(5):2075–2089.

Addor, N., Rössler, O., Köplin, N., Huss, M., Weingartner, R., and Seibert, J. (2014). Robust changes and sources of uncertainty in the projected hydrological regimes of Swiss catchments. *Water Resources Research*, 50(10):7541–7562.

Addor, N. and Seibert, J. (2014). Bias correction for hydrological impact studies - beyond the daily perspective: INVITED COMMENTARY. *Hydrological Processes*, 28(17):4823–4828.

Allen, R. G., Pereira, L. S., Raes, D., and Smith, M., editors (1998). *Crop Evapotranspiration: Guidelines for Computing Crop Water Requirements*. Number 56 in FAO Irrigation and Drainage Paper. Food and Agriculture Organization of the United Nations, Rome.

Alo, C. A. and Anagnostou, E. N. (2017). A sensitivity study of the impact of dynamic vegetation on simulated future climate change over southern europe and the mediterranean. *International Journal of Climatology*, 37(4):2037–2050.

Ambroise, B., Beven, K., and Freer, J. (1996). Toward a Generalization of the TOPMODEL Concepts: Topographic Indices of Hydrological Similarity. *Water Resources Research*, 32(7):2135–2145.

Anderegg, W. R. L., Berry, J. A., Smith, D. D., Sperry, J. S., Anderegg, L. D. L., and Field, C. B. (2012). The roles of hydraulic and carbon stress in a widespread climate-induced forest die-off. *Proceedings of the National Academy of Sciences*, 109(1):233–237.

Anglian Water (2019). WRMP 2019 Technical Document: Supply Forecast. `https://www.anglianwater.co.uk/siteassets/household/about-us/supply-forecast.pdf`.

Anglian Water (2021). Drought Plan 2022 Appendix 4: Drought management for Anglian Water reservoirs. `https://www.anglianwater.co.uk/siteassets/household/about-us/aws-drought-plan-2022---appendix-4-reservoirs.pdf`.

Anglian Water (2022a). Draft Water Resources Management Plan 2024: Main Report. Technical report.

Anglian Water (2022b). Drought Plan 2022. Technical report.

Apurv, T., Sivapalan, M., and Cai, X. (2017). Understanding the Role of Climate Characteristics in Drought Propagation. *Water Resources Research*, 53(11):9304–9329.

Ara Begum, R., Lempert, R., Ali, E., Benjaminsen, T., Bernauer, T., Cramer, W., Cui, X., Mach, K., Nagy, G., Stenseth, N., Sukumar, R., and Wester, P. (2022). Point of Departure and Key Concepts. In:. In *Climate Change 2022: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [H.-O. Pörtner, D.C. Roberts, M. Tignor, E.S. Poloczanska, K. Mintenbeck, A. Alegría, M. Craig, S. Langsdorf, S. Löschke, V. Möller, A. Okem, B. Rama (Eds.)]*., pages 121–196. Cambridge University Press, Cambridge, UK and New York, NY, USA.

Arias, P. A., Bellouin, N., Jones, R. G., Naik, V., Plattner, G.-K., Rogelj, J., Sillmann, J., Storelvmo, T., Thorne, P. W., Trewin, B., Rao, K. A., Adhikary, B., Allan, R. P., Armour, K., Barimalala, R., Canadell, J. G., Cassou, C., Cherchi, A., Collins, W., Corti, S., Cruz, F., Dentener, F. J., Dereczynski, C., Luca, A. D., Diongue, A., Doblas-Reyes, F. J., Dosio, A., Douville, H., Engelbrecht, F., Fyfe, J. C., Gillett, N. P., and Goldfarb, L. (2021). Technical Summary. In Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou (eds.). *Cambridge University Press, Cambridge, United Kingdom and New York, NY, US*, pages 33–144.

Armstrong McKay, D. I., Staal, A., Abrams, J. F., Winkelmann, R., Sakschewski, B., Loriani, S., Fetzer, I., Cornell, S. E., Rockström, J., and Lenton, T. M.

(2022). Exceeding 1.5°C global warming could trigger multiple climate tipping points. *Science*, 377(6611):eabn7950.

Arnell, N. W. (2011). Uncertainty in the relationship between climate forcing and hydrological response in UK catchments. *Hydrology and Earth System Sciences*, 15(3):897–912.

Arnell, N. W. and Freeman, A. (2021). The effect of climate change on agro-climatic indicators in the UK. *Climatic Change*, 165(1-2):40.

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*, 31:100265.

Arora, V. K., Katavouta, A., Williams, R. G., Jones, C. D., Brovkin, V., Friedlingstein, P., Schwinger, J., Bopp, L., Boucher, O., Cadule, P., et al. (2020). Carbon–concentration and carbon–climate feedbacks in cmip6 models and their comparison to cmip5 models. *Biogeosciences*, 17(16):4173–4222.

Ault, T. R. (2020). On the essentials of drought in a changing climate. *Science*, 368(6488):256–260.

Bachmair, S., Svensson, C., Hannaford, J., Barker, L. J., and Stahl, K. (2016). A quantitative analysis to objectively appraise drought indicators and modeldrought impacts. *Hydrology and Earth System Sciences*, 20(7):2589–2609.

Bachmair, S., Tanguy, M., Hannaford, J., and Stahl, K. (2018). How well do meteorological indicators represent agricultural and forest drought across Europe? *Environmental Research Letters*, 13(3):034042.

Bae, D.-H., Jung, I.-W., and Lettenmaier, D. P. (2011). Hydrologic uncertainties in climate change from IPCC AR4 GCM simulations of the Chungju Basin, Korea. *Journal of Hydrology*, 401(1):90–105.

Bai, P., Liu, X., and Xie, J. (2021). Simulating runoff under changing climatic conditions: A comparison of the long short-term memory network with two conceptual hydrologic models. *Journal of Hydrology*, 592:125779.

Baran-Gurgul, K. (2022). The spatial and temporal variability of hydrological drought in the Polish Carpathians. *Journal of Hydrology and Hydromechanics*, 70(2):156–169.

Barker, L. J., Hannaford, J., Chiverton, A., and Svensson, C. (2016). From meteorological to hydrological drought using standardised indicators. *Hydrology and Earth System Sciences*, 20(6):2483–2505.

Bastola, S., Murphy, C., and Sweeney, J. (2011). The role of hydrological modelling uncertainties in climate change impact assessments of Irish river catchments. *Advances in Water Resources*, 34(5):562–576.

Berg, A. and Sheffield, J. (2018). Climate Change and Drought: The Soil Moisture Perspective. *Current Climate Change Reports*, 4(2):180–191.

Bevan, J. (2019). Escaping the jaws of death: Ensuring enough water in 2050. Waterwise Conference, Environment Agency. Transcript available at `https://www.gov.uk/government/speeches/escaping-the-jaws-of-death-ensuring-enough-water-in-2050`.

Beven, K. (1997). TOPMODEL: A critique. *Hydrological Processes*, 11(9):1069–1085.

Beven, K. J. (2012). *Rainfall-Runoff Modelling: The Primer*. Wiley-Blackwell, Chichester, West Sussex ; Hoboken, NJ, 2nd ed edition.

Beven, K. J. and Kirkby, M. J. (1979). A physically based, variable contributing area model of basin hydrology / Un modèle à base physique de zone d'appel variable de l'hydrologie du bassin versant. *Hydrological Sciences Bulletin*, 24(1):43–69.

Blauhut, V., Stahl, K., Stagge, J. H., Tallaksen, L. M., De Stefano, L., and Vogt, J. (2016). Estimating drought risk across Europe from reported drought impacts, drought indices, and vulnerability factors. *Hydrology and Earth System Sciences*, 20(7):2779–2800.

Blenkinsop, S. and Fowler, H. J. (2007). Changes in European drought characteristics projected by the PRUDENCE regional climate models. *International Journal of Climatology*, 27(12):1595–1610.

Bloomfield, J. P. and Marchant, B. P. (2013). Analysis of groundwater drought building on the standardised precipitation index approach. *Hydrology and Earth System Sciences*, 17:4769–4787.

Bloomfield, J. P., Marchant, B. P., and McKenzie, A. A. (2019). Changes in groundwater drought associated with anthropogenic warming. *Hydrology and Earth System Sciences*, 23(3):1393–1408.

Blyth, E. M., Martínez-de la Torre, A., and Robinson, E. L. (2019). Trends in evapotranspiration and its drivers in Great Britain: 1961 to 2015. *Progress in Physical Geography: Earth and Environment*, 43(5):666–693.

Bosshard, T., Carambia, M., Goergen, K., Kotlarski, S., Krahe, P., Zappa, M., and Schär, C. (2013). Quantifying uncertainty sources in an ensemble

of hydrological climate-impact projections. *Water Resources Research*, 49(3):1523–1536.

Brunner, M. I. and Stahl, K. (2023). Temporal hydrological drought clustering varies with climate and land-surface processes. *Environmental Research Letters*, 18:034011.

Brunner, M. I., Van Loon, A. F., and Stahl, K. (2022). Moderate and severe hydrological droughts in Europe differ in their hydro-meteorological drivers. *Water Resources Research*, n/a(n/a):e2022WR032871.

Bryan, K., Ward, S., Roberts, L., White, M. P., Landeg, O., Taylor, T., and McEwen, L. (2020). The health and well-being effects of drought: Assessing multi-stakeholder perspectives through narratives from the UK. *Climatic Change*, 163(4):2073–2095.

Bueh, C. and Nakamura, H. (2007). Scandinavian pattern and its climatic impact. *Quarterly Journal of the Royal Meteorological Society*, 133(629):2117–2131.

Burnash, R. J. C. (1995). The NWS River Forecast System - catchment modeling. *Computer models of watershed hydrology*, pages 311–366.

Burnash, R. J. C., Ferral, R. L., and McGuire, R. A. (1973). *A Generalized Streamflow Simulation System: Conceptual Modeling for Digital Computers*. U.S. Department of Commerce, National Weather Service, and State of California, Department of Water Resources.

Byers, E. A., Coxon, G., Freer, J., and Hall, J. W. (2020). Drought and climate change impacts on cooling water shortages and electricity prices in Great Britain. *Nature Communications*, 11(1):2239.

Cammalleri, C., Naumann, G., Mentaschi, L., Bisselink, B., Gelati, E., De Roo, A., and Feyen, L. (2020). Diverging hydrological drought traits over Europe with global warming. *Hydrology and Earth System Sciences*, 24(12):5919–5935.

Canadell, J. G., Monteiro, P., Costa, M., Cotrim da Cunha, L., Cox, P., Eliseey, A., Henson, S., Ishii, M., Jaccard, S., Koven, C., Lohila, A., Patra, P., Piao, S., Rogelj, J., Syampungani, S., Zaehle, S., and Zickfeld, K. (2021). Global Carbon and other Biogeochemical Cycles and Feedbacks. In *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou (Eds.)]*., pages 673–816. Cambridge University Press,, Cambridge, United Kingdom and New York, NY, USA.

Cannon, A. J. (2016). Multivariate Bias Correction of Climate Model Output: Matching Marginal Distributions and Intervariable Dependence Structure. *Journal of Climate*, 29(19):7045–7064.

Caretta, M., Mukherji, A., Arfanuzzaman, M., Betts, R. A., Gelfan, A., Hirabayashi, Y., Lissner, T., Liu, J., Lopez-Gunn, E., Morgan, R., Mwanga, S., and Supratid, S. (2022). Water. In *Climate Change 2022: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [H.-O. Pörtner, D.C. Roberts, M. Tignor, E.S. Poloczanska, K. Mintenbeck, A. Alegría, M. Craig, S. Langsdorf, S. Löschke, V. Möller, A. Okem, B. Rama (Eds.)]*., pages 551–712. Cambridge University Press, Cambridge, UK and New York, NY, USA,.

Ceppi, P., Zappa, G., Shepherd, T. G., and Gregory, J. M. (2018). Fast and Slow Components of the Extratropical Atmospheric Circulation Response to CO2 Forcing. *Journal of Climate*, 31(3):1091–1105.

Chan, W. C., Shepherd, T. G., Facer-Childs, K., Darch, G., and Arnell, N. W. (2022). Storylines of uk drought based on the 2010–2012 event. *Hydrology and Earth System Sciences*, 26(7):1755–1777.

Charlton, M. B. and Arnell, N. W. (2014). Assessing the impacts of climate change on river flows in England using the UKCP09 climate change projections. *Journal of Hydrology*, 519:1723–1738.

Chawla, I. and Mujumdar, P. (2018). Partitioning uncertainty in streamflow projections under nonstationary model conditions. *Advances in Water Resources*, 112:266–282.

Chegwidden, O. S., Nijssen, B., Rupp, D. E., Arnold, J. R., Clark, M. P., Hamman, J. J., Kao, S. C., Mao, Y., Mizukami, N., Mote, P. W., Pan, M., Pytlak, E., and Xiao, M. (2019). How Do Modeling Decisions Affect the Spread Among Hydrologic Climate Change Projections? Exploring a Large Ensemble of Simulations Across a Diversity of Hydroclimates. *Earth's Future*, 7(6):623–637.

Chen, D., Rojas, M., Samset, B., Cobb, K., Diongue Niang, A., Edwards, P., Emori, S., Faria, S., Hawkins, E., Hope, P., Huybrechts, P., Meinshausen, M., Mustafa, S., Plattner, G.-K., and Tréguier, A.-M. (2021). In Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy,

J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou (eds.)]. *Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA*, pages 147–286.

Chen, J., Brissette, F. P., and Lucas-Picher, P. (2015). Assessing the limits of bias-correcting climate model outputs for climate change impact studies. *Journal of Geophysical Research: Atmospheres*, 120(3):1123–1136.

Chen, J., Brissette, F. P., Zhang, X. J., Chen, H., Guo, S., and Zhao, Y. (2019). Bias correcting climate model multi-member ensembles to assess climate change impacts on hydrology. *Climatic Change*, 153(3):361–377.

Chiang, F., Mazdiyasni, O., and AghaKouchak, A. (2021). Evidence of anthropogenic impacts on global drought frequency, duration, and intensity. *Nature Communications*, 12(1):2754.

Cissé, G., McLeman, R., Adams, H., Aldunce, P., Bowen, K., Campbell-Lendrum, D., Clayton, S., Ebi, K., Hess, J., Huang, C., Liu, Q., McGregor, G., Semenza, J., and Tirado, M. (2022). Health, Wellbeing and the Changing Structure of Communities. In *Climate Change 2022: Impacts,Adaptation and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [H.-O. Pörtner, D.C. Roberts, M. Tignor, E.S. Poloczanska, K. Mintenbeck, A. Alegría, M. Craig, S. Langsdorf, S. Löschke, V. Möller, A. Okem, B. Rama (Eds.)]*., pages 1041–1170. Cambridge University Press.

Clark, M. P., Nijssen, B., Lundquist, J. D., Kavetski, D., Rupp, D. E., Woods, R. A., Freer, J. E., Gutmann, E. D., Wood, A. W., Brekke, L. D., Arnold, J. R., Gochis, D. J., and Rasmussen, R. M. (2015). A unified approach for process-based hydrologic modeling: 1. Modeling concept. *Water Resources Research*, 51(4):2498–2514.

Clark, M. P., Slater, A. G., Rupp, D. E., Woods, R. A., Vrugt, J. A., Gupta, H. V., Wagener, T., and Hay, L. E. (2008). Framework for Understanding Structural Errors (FUSE): A modular framework to diagnose differences between hydrological models. *Water Resources Research*, 44(12).

Climate Action Tracker (2022). Warming Projections Global Update. Technical report.

Comas-Bru, L. and McDermott, F. (2014). Impacts of the EA and SCA patterns on the European twentieth century NAO–winter climate relationship. *Quarterly Journal of the Royal Meteorological Society*, 140(679):354–363.

Cook, B. I., Smerdon, J. E., Seager, R., and Coats, S. (2014). Global warming and 21st century drying. *Climate Dynamics*, 43(9):2607–2627.

Coron, L., Andréassian, V., Perrin, C., Lerat, J., Vaze, J., Bourqui, M., and Hendrickx, F. (2012). Crash testing hydrological models in contrasted climate conditions: An experiment on 216 Australian catchments. *Water Resources Research*, 48(5).

Craig, J. R., Brown, G., Chlumsky, R., Jenkinson, R. W., Jost, G., Lee, K., Mai, J., Serrer, M., Sgro, N., Shafii, M., Snowdon, A. P., and Tolson, B. A. (2020). Flexible watershed simulation with the Raven hydrological modelling framework. *Environmental Modelling & Software*, 129:104728.

Cubasch, U. and Cess, R. (1990). Chapter 3: Processes and Modelling. In *Climate Change. The IPCC Scientific Assessment.* Cambridge University Press, Cambridge.

Cui, J., Lian, X., Huntingford, C., Gimeno, L., Wang, T., Ding, J., He, M., Xu, H., Chen, A., Gentine, P., et al. (2022). Global water availability boosted by vegetation-driven changes in atmospheric moisture transport. *Nature Geoscience*, 15(12):982–988.

Dai, A. (2011). Drought under global warming: A review. *Wiley Interdisciplinary Reviews-Climate Change*, 2(1):45–65.

Dakhlaoui, H. and Djebbi, K. (2021). Evaluating the impact of rainfall–runoff model structural uncertainty on the hydrological rating of regional climate model simulations. *Journal of Water and Climate Change*, 12(8):3820–3838.

Dakhlaoui, H., Ruelland, D., Tramblay, Y., and Bargaoui, Z. (2017). Evaluating the robustness of conceptual rainfall-runoff models under climate variability in northern Tunisia. *Journal of Hydrology*, 550:201–217.

Dal Molin, M., Kavetski, D., and Fenicia, F. (2021). SuperflexPy 1.3.0: An open-source Python framework for building, testing, and improving conceptual hydrological models. *Geoscientific Model Development*, 14(11):7047–7072.

Dallison, R. J. H., Patil, S. D., and Williams, A. P. (2021). Impacts of climate change on future water availability for hydropower and public water supply in Wales, UK. *Journal of Hydrology: Regional Studies*, 36:100866.

Davies, T., Kelly, M., and Osborn, T. J. (1997). Explaining the climate ofthe British Isles. In *Climates of the British Isles: Present, Past and Future (Ed. M. Hulme and E. Barrow)*, pages 11–32. Routeledge, London.

de Brito, M. M. (2021). Compound and cascading drought impacts do not happen by chance: A proposal to quantify their relationships. *Science of The Total Environment*, 778:146236.

De Niel, J., Van Uytven, E., and Willems, P. (2019). Uncertainty Analysis of Climate Change Impact on River Flow Extremes Based on a Large Multi-Model Ensemble. *Water Resources Management*, 33(12):4319–4333.

DEFRA (2023). Agricultural facts: East of England Region. Official Statistics.

Denissen, J., Teuling, A., Pitman, A., Koirala, S., Migliavacca, M., Li, W., Reichstein, M., Winkler, A., Zhan, C., and Orth, R. (2022). Widespread shift from ecosystem energy to water limitation with climate change. *Nature Climate Change*, 12.

Dewes, C. F., Rangwala, I., Barsugli, J. J., Hobbins, M. T., and Kumar, S. (2017). Drought risk assessment under climate change is sensitive to methodological choices for the estimation of evaporative demand. *PLOS ONE*, 12(3):e0174045.

Ding, Y., Xu, J., Wang, X., Cai, H., Zhou, Z., Sun, Y., and Shi, H. (2021). Propagation of meteorological to hydrological drought for different climate regions in China. *Journal of Environmental Management*, 283:111980.

Doblas-Reyes, F., Sörensson, A., Almazroui, M., Dosio, A., Gutowski, W., Haarsma, R., Hamdi, R., Hewitson, B., Kwon, W.-T., Lamptey, B., Maraun, D., Stephenson, T., Takayabu, I., Terray, L., Turner, A., and Zuo, Z. (2021). Linking Global to Regional Climate Change. In Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou (eds.)]. Technical report, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Dobson, B., Coxon, G., Freer, J., Gavin, H., Mortazavi-Naeini, M., and Hall, J. W. (2020). The Spatial Dynamics of Droughts and Water Scarcity in England and Wales. *Water Resources Research*, 56(9):e2020WR027187.

Douville, H., Raghavan, K., Renwick, J., Allan, R., Arias, P., Barlow, M., Cerezo-Mota, R., Cherchi, A., Gan, T., Gergis, J., Jiang, D., Khan, W., Pokam Mba, W., Rosenfeld, D., Tierney, J., and Zolina, O. (2021). Water Cycle Changes. In Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou (eds.)]. Technical report, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA,.

Dracup, J. A., Lee, K. S., and Paulson, E. G. (1980). On the definition of droughts. *Water Resources Research*, 16(2):297–302.

Duan, J. and Miller, N. L. (1997). A generalized power function for the subsurface transmissivity profile in TOPMODEL. *Water Resources Research*, 33(11):2559–2562.

Duan, Q. Y., Gupta, V. K., and Sorooshian, S. (1993). Shuffled complex evolution approach for effective and efficient global minimization. *Journal of Optimization Theory and Applications*, 76(3):501–521.

Edijatno, De Oliveira Nascimento, N., Yang, X., MAKHLOUF, Z., and Michel, C. (1999). GR3J: A daily watershed model with three free parameters. *Hydrological Sciences Journal*, 44(2):263–277.

Ehret, U., Zehe, E., Wulfmeyer, V., Warrach-Sagi, K., and Liebert, J. (2012). "Should we apply bias correction to global and regional climate model data?". *Hydrology and Earth System Sciences*, 16(9):3391–3404.

Environment Agency (2021). Drought: How water companies plan for dry weather and drought. Available at: `https://www.gov.uk/government/publications/drought-managing-water-supply/drought-how-water-companies-plan-for-dry-weather-and-drought`. [Accessed: 2023-04-09].

Environment Agency, Natural Resources Wales, and The Water Services Regulation Authority (2022). Water resources planning guideline. `https://www.gov.uk/government/publications/water-resources-planning-guideline/water-resources-planning-guideline`. [Accessed: 2023-01-19].

Feng, H. and Zhang, M. (2016). Global land moisture trends: Drier in dry and wetter in wet over land. *Scientific Reports*, 5(1):18018.

Feng, S. and Fu, Q. (2013). Expansion of global drylands under a warming climate. *Atmospheric Chemistry and Physics*, 13(19):10081–10094.

Feng, S., Trnka, M., Hayes, M., and Zhang, Y. (2017). Why Do Different Drought Indices Show Distinct Future Drought Risk Outcomes in the U.S. Great Plains? *Journal of Climate*, 30(1):265–278.

Field, C. B., Jackson, R. B., and Mooney, H. A. (1995). Stomatal responses to increased co2: implications from the plant to the global scale. *Plant, Cell & Environment*, 18(10):1214–1225.

Flato, G., Marotzke, J., Abiodun, B., Braconnot, P., Chou, S., Collins, W., Cox, P., Driouech, F., Emori, S., Eyring, V., Forest, C., Gleckler, P., Guilyardi, E., Jakob, C., Kattsov, V., Reason, C., and Rummukainen, M. (2013). Evaluation of Climate Models. In *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assess- Ment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (Eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Fleig, A. K., Tallaksen, L. M., Hisdal, H., and Demuth, S. (2006). A global evaluation of streamflow drought characteristics. *Hydrology and Earth System Sciences*, 10(4):535–552.

Fleig, A. K., Tallaksen, L. M., Hisdal, H., and Hannah, D. M. (2011). Regional hydrological drought in north-western Europe: Linking a new Regional Drought Area Index with weather types. *Hydrological Processes*, 25(7):1163–1179.

Folland, C. K., Hannaford, J., Bloomfield, J. P., Kendon, M., Svensson, C., Marchant, B. P., Prior, J., and Wallace, E. (2015). Multi-annual droughts in the English Lowlands: A review of their characteristics and climate drivers in the winter half-year. *Hydrology and Earth System Sciences*, 19(5):2353–2375.

Folland, C. K., Knight, J., Linderholm, H. W., Fereday, D., Ineson, S., and Hurrell, J. W. (2009). The Summer North Atlantic Oscillation: Past, Present, and Future. *Journal of Climate*, 22(5):1082–1103.

Forster, P., Storelvmo, T., Armour, K., Collins, W., Dufresne, J.-L., Frame, D., Lunt, D., Mauritsen, T., Palmer, M., Watanabe, M., Wild, M., and Zhang, H. (2021). *The Earth's Energy Budget, Climate Feedbacks and Climate Sensitivity. In Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou (Eds.)].* Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Fowler, K., Knoben, W. J. M., Peel, M. C., Peterson, T. J., Ryu, D., Saft, M., Seo, K.-W., and Western, A. (2020). Many Commonly Used Rainfall-Runoff Models Lack Long, Slow Dynamics: Implications for Runoff Projections. *Water Resources Research*, 56(5):e2019WR025286.

Fowler, K., Peel, M., Western, A., and Zhang, L. (2018). Improved Rainfall-Runoff Calibration for Drying Climate: Choice of Objective Function. *Water Resources Research*, 54(5):3392–3408.

François, B., Vrac, M., Cannon, A. J., Robin, Y., and Allard, D. (2020). Multivariate bias corrections of climate simulations: Which benefits for which losses? *Earth System Dynamics Discussions*, pages 1–41.

Franklin, J., Serra-Diaz, J. M., Syphard, A. D., and Regan, H. M. (2016). Global change and terrestrial plant community dynamics. *Proceedings of the National Academy of Sciences*, 113(14):3725–3734.

Fu, Q. and Feng, S. (2014). Responses of terrestrial aridity to global warming. *Journal of Geophysical Research: Atmospheres*, 119(13):7863–7875.

Fung, F., Lowe, J., Mitchell, JFB., Murphy, J., Bernie, D., Gohar, L., Harris, G., Howard, T., Kendon, E., Maisey, P., Palmer, M., and Sexton, D. (2018). UKCP18 Guidance: How to use the UKCP18 Land Projections. *Met Office Hadley Centre, Exeter*.

Gampe, D., Zscheischler, J., Reichstein, M., O'Sullivan, M., Smith, W. K., Sitch, S., and Buermann, W. (2021). Increasing impact of warm droughts on northern ecosystem productivity over recent decades. *Nature Climate Change*, pages 1–8.

Gao, C., Booij, M. J., and Xu, Y.-P. (2020). Assessment of extreme flows and uncertainty under climate change: Disentangling the uncertainty contribution of representative concentration pathways, global climate models and internal climate variability. *Hydrology and Earth System Sciences*, 24(6):3251–3269.

García-Valdecasas Ojeda, M., Gámiz-Fortis, S. R., Romero-Jiménez, E., Rosa-Cánovas, J. J., Yeste, P., Castro-Díez, Y., and Esteban-Parra, M. J. (2021). Projected changes in the Iberian Peninsula drought characteristics. *Science of The Total Environment*, 757:143702.

Gebrechorkos, S. H., Pan, M., Lin, P., Anghileri, D., Forsythe, N., Pritchard, D. M. W., Fowler, H. J., Obuobie, E., Darko, D., and Sheffield, J. (2022). Variability and changes in hydrological drought in the Volta Basin, West Africa. *Journal of Hydrology: Regional Studies*, 42:101143.

Gevaert, A. I., Veldkamp, T. I. E., and Ward, P. J. (2018). The effect of climate type on timescales of drought propagation in an ensemble of global hydrological models. *Hydrology and Earth System Sciences*, 22(9):4649–4665.

Giorgi, F., Jones, C., and Asrar, G. R. (2009). Addressing climate information needs at the regional level: The CORDEX framework. *WMO Bulletin*, 58(3):175–183.

Gohar, L., Bernie, D., Good, P., and Lowe, J. A. (2018). UKCP18 Derived Projections of Future Climate over the UK. *Met Office, Exeter*, page 47.

Gosling, S. N., Taylor, R. G., Arnell, N. W., and Todd, M. C. (2011). A comparative analysis of projected impacts of climate change on river runoff from global and catchment-scale hydrological models. *Hydrology and Earth System Sciences*, 15(1):279–294.

Greve, P., Orlowsky, B., Mueller, B., Sheffield, J., Reichstein, M., and Seneviratne, S. I. (2014). Global assessment of trends in wetting and drying over land. *Nature Geoscience*, 7(10):716–721.

Greve, P., Roderick, M. L., Ukkola, A. M., and Wada, Y. (2019). The aridity Index under global warming. *Environmental Research Letters*, 14(12):124006.

Grossiord, C., Buckley, T. N., Cernusak, L. A., Novick, K. A., Poulter, B., Siegwolf, R. T. W., Sperry, J. S., and McDowell, N. G. (2020). Plant responses to rising vapor pressure deficit. *New Phytologist*, 226(6):1550–1566.

Gudmundsson, L., Wagener, T., Tallaksen, L. M., and Engeland, K. (2012). Evaluation of nine large-scale hydrological models with respect to the seasonal runoff climatology in Europe. *Water Resources Research*, 48(11).

Gupta, H. V., Kling, H., Yilmaz, K. K., and Martinez, G. F. (2009). Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology*, 377(1):80–91.

Gutmann, E., Pruitt, T., Clark, M. P., Brekke, L., Arnold, J. R., Raff, D. A., and Rasmussen, R. M. (2014). An intercomparison of statistical downscaling methods used for water resource assessments in the United States. *Water Resources Research*, 50(9):7167–7186.

Hagemann, S., Chen, C., Clark, D. B., Folwell, S., Gosling, S. N., Haddeland, I., Hanasaki, N., Heinke, J., Ludwig, F., Voss, F., and Wiltshire, A. J. (2013). Climate change impact on available water resources obtained using multiple global climate and hydrology models. *Earth System Dynamics*, 4(1):129–144.

Hakala, K., Addor, N., Gobbe, T., Ruffieux, J., and Seibert, J. (2020). Risks and opportunities for a Swiss hydroelectricity company in a changing climate. *Hydrology and Earth System Sciences*, 24(7):3815–3833.

Hakala, K., Addor, N., and Seibert, J. (2018). Hydrological Modeling to Evaluate Climate Model Simulations and Their Bias Correction. *Journal of Hydrometeorology*, 19(8):1321–1337.

Hakala, K., Addor, N., Teutschbein, C., Vis, M., Dakhlaoui, H., and Seibert, J. (2019). Hydrological Modeling of Climate Change Impacts. In *Encyclopedia of Water*, pages 1–20. John Wiley and Sons.

Hall, R. J. and Hanna, E. (2018). North Atlantic circulation indices: Links with summer and winter UK temperature and precipitation and implications for seasonal forecasting. *International Journal of Climatology*, 38(S1):e660–e677.

Halwatura, D., Lechner, A. M., and Arnold, S. (2015). Drought severity–duration–frequency curves: A foundation for risk assessment and planning tool for ecosystem establishment in post-mining landscapes. *Hydrol. Earth Syst. Sci.*, page 23.

Hanlon, H. M., Bernie, D., Carigi, G., and Lowe, J. A. (2021). Future changes to high impact weather in the UK. *Climatic Change*, 166(3):50.

Haro-Monteagudo, D., Daccache, A., and Knox, J. (2018). Exploring the utility of drought indicators to assess climate risks to agricultural productivity in a humid climate. *Hydrology Research*, 49(2):539–551.

Haslinger, K. and Bloeschl, G. (2017). Space-Time Patterns of Meteorological Drought Events in the European Greater Alpine Region Over the Past 210 Years. *Water Resources Research*, 53(11):9807–9823.

Hattermann, F. F., Vetter, T., Breuer, L., Su, B., Daggupati, P., Donnelly, C., Fekete, B., Flörke, F., Gosling, S. N., Hoffmann, P., Liersch, S., Masaki, Y., Motovilov, Y., Müller, C., Samaniego, L., Stacke, T., Wada, Y., Yang, T., and Krysnaova, V. (2018). Sources of uncertainty in hydrological climate impact assessment: A cross-scale study. *Environmental Research Letters*, 13(1):015006.

Hausfather, Z. and Peters, G. P. (2020). Emissions – the 'business as usual' story is misleading. *Nature*, 577(7792):618–620.

Hawkins, E. and Sutton, R. (2009). The Potential to Narrow Uncertainty in Regional Climate Predictions. *Bulletin of the American Meteorological Society*, 90(8):1095–1108.

Henn, B., Clark, M. P., Kavetski, D., and Lundquist, J. D. (2015). Estimating mountain basin-mean precipitation from streamflow using Bayesian inference. *Water Resources Research*, 51(10):8012–8033.

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):151–159.

Hrachowitz, M. and Clark, M. P. (2017). HESS Opinions: The complementary merits of competing modelling philosophies in hydrology. *Hydrology and Earth System Sciences*, 21(8):3953–3973.

Hui, Y., Xu, Y., Chen, J., Xu, C.-Y., and Chen, H. (2020). Impacts of bias nonstationarity of climate model outputs on hydrological simulations. *Hydrology Research*, 51(5):925–941.

Hurtt, G. C., Chini, L. P., Frolking, S., Betts, R., Feddema, J., Fischer, G., Fisk, J., Hibbard, K., Houghton, R., Janetos, A., et al. (2011). Harmonization of land-use scenarios for the period 1500–2100: 600 years of global gridded annual land-use transitions, wood harvest, and resulting secondary lands. *Climatic change*, 109:117–161.

Ionita, M. and Nagavciuc, V. (2021). Changes in drought features at the European level over the last 120 years. *Natural Hazards and Earth System Sciences*, 21(5):1685–1701.

Iorgulescu, I. and Musy, A. (1997). Generalization of TOPMODEL for a power law transmissivity profile. *Hydrological Processes*, 11(9):1353–1355.

IPCC (2021). Annex VI: Climatic Impact-driver and Extreme Indices [Gutiérrez J.M., R. Ranasinghe, A.C. Ruane, R. Vautard (eds.)]. In Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou (eds.)]. *Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA*, pages 2205–2214.

IPCC (2022). *Summary for Policymakers. In: Climate Change 2022: Mitigation of Climate Change. Contribution of Working Group III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [P.R. Shukla, J. Skea, R. Slade, A. Al Khourdajie, R. van Diemen, D. McCollum, M. Pathak, S. Some, P. Vyas, R. Fradera, M. Belkacemi, A. Hasija, G. Lisboa, S. Luz, J. Malley, (Eds.)].* Cambridge University Press, Cambridge, UK and New York, NY, USA.

James, R., Washington, R., Schleussner, C.-F., Rogelj, J., and Conway, D. (2017). Characterizing half-a-degree difference: A review of methods for identifying regional climate responses to global warming targets. *WIREs Climate Change*, 8(2):e457.

Jones, PD., Lister, DH., and Kostopoulou, E. (2004). Reconstructed river Flow Series from 1860s to Present: Updating Previously Reconstructed Series to 2002. Environment Agency Research Report SC030242. Science Report SC040052/SR. Environment Agency.

Karimi, M., Vicente-Serrano, S. M., Reig, F., Shahedi, K., Raziei, T., and Miryaghoubzadeh, M. (2020). Recent trends in atmospheric evaporative demand in Southwest Iran: Implications for change in drought severity. *Theoretical and Applied Climatology*, 142(3-4):945–958.

Kautz, L.-A., Martius, O., Pfahl, S., Pinto, J. G., Ramos, A. M., Sousa, P. M., and Woollings, T. (2022). Atmospheric blocking and weather extremes over the Euro-Atlantic sector – a review. *Weather and Climate Dynamics*, 3(1):305–336.

Kavetski, D. and Kuczera, G. (2007). Model smoothing strategies to remove microscale discontinuities and spurious secondary optima in objective functions in hydrological calibration. *Water Resources Research*, 43(3).

Kay, A. and Davies, H. (2008). Calculating potential evaporation from climate model data: A source of uncertainty for hydrological climate change impacts. *Journal of Hydrology*, 358(3-4):221–239.

Kay, A. L., Bell, V. A., Blyth, E. M., Crooks, S. M., Davies, H. N., and Reynard, N. S. (2013). A hydrological perspective on evaporation: Historical trends and future projections in Britain. *Journal of Water and Climate Change*, 4(3):193–208.

Kay, A. L., Bell, V. A., Guillod, B. P., Jones, R. G., and Rudd, A. C. (2018). National-scale analysis of low flow frequency: Historical trends and potential future changes. *Climatic Change*, 147(3):585–599.

Kay, A. L., Griffin, A., Rudd, A. C., Chapman, R. M., Bell, V. A., and Arnell, N. W. (2021a). Climate change effects on indicators of high and low river flow across Great Britain. *Advances in Water Resources*, 151:103909.

Kay, A. L., Lane, R. A., and Bell, V. A. (2022). Grid-based simulation of soil moisture in the UK: Future changes in extremes and wetting and drying dates. *Environmental Research Letters*, 17(7):074029.

Kay, A. L., Rudd, A. C., Fry, M., Nash, G., and Allen, S. (2021b). Climate change impacts on peak river flows: Combining national-scale hydrological modelling and probabilistic projections. *Climate Risk Management*, 31:100263.

Kay, A. L., Watts, G., Wells, S. C., and Allen, S. (2020). The impact of climate change on U. K. river flows: A preliminary comparison of two generations of probabilistic climate projections. *Hydrological Processes*, 34(4):1081–1088.

Kelley, C. P., Mohtadi, S., Cane, M. A., Seager, R., and Kushnir, Y. (2015). Climate change in the Fertile Crescent and implications of the recent Syrian drought. *Proceedings of the National Academy of Sciences*, 112(11):3241–3246.

Kendon, M., Marsh, T., and Parry, S. (2013). The 2010–2012 drought in England and Wales. *Weather*, 68(4):88–95.

Keyantash, J. and Dracup, J. A. (2002). The Quantification of Drought: An Evaluation of Drought Indices. *Bulletin of the American Meteorological Society*, 83(8):1167–1180.

Khan, S. J., Deere, D., Leusch, F. D. L., Humpage, A., Jenkins, M., and Cunliffe, D. (2015). Extreme weather events: Should drinking water quality management systems adapt to changing risk profiles? *Water Research*, 85:124–136.

Kingston, D. G., Fleig, A. K., Tallaksen, L. M., and Hannah, D. M. (2013). Ocean–Atmosphere Forcing of Summer Streamflow Drought in Great Britain. *Journal of Hydrometeorology*, 14(1):331–344.

KLEMEŠ, V. (1986). Operational testing of hydrological simulation models. *Hydrological Sciences Journal*, 31(1):13–24.

Kling, H., Fuchs, M., and Paulin, M. (2012). Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios. *Journal of Hydrology*, 424–425:264–277.

Knoben, W. J. M., Freer, J. E., Fowler, K. J. A., Peel, M. C., and Woods, R. A. (2019). Modular Assessment of Rainfall–Runoff Models Toolbox (MARRMoT) v1.2: An open-source, extendable framework providing implementations of 46 conceptual hydrologic models as continuous state-space formulations. *Geoscientific Model Development*, 12(6):2463–2480.

Konapala, G., Kao, S.-C., and Addor, N. (2020). Exploring Hydrologic Model Process Connectivity at the Continental Scale Through an Information Theory Approach. *Water Resources Research*, 56(10):e2020WR027340.

Konapala, G. and Mishra, A. (2019). Quantifying climate and catchment control on hydrological drought in continental United States. *Water Resources Research*, 56:e2018WR024620.

Koren, V., Reed, S., Smith, M., Zhang, Z., and Seo, D.-J. (2004). Hydrology laboratory research modeling system (HL-RMS) of the US national weather service. *Journal of Hydrology*, 291(3-4):297–318.

Kotlarski, S., Keuler, K., Christensen, O. B., Colette, A., Déqué, M., Gobiet, A., Goergen, K., Jacob, D., Lüthi, D., van Meijgaard, E., Nikulin, G., Schär,

C., Teichmann, C., Vautard, R., Warrach-Sagi, K., and Wulfmeyer, V. (2014). Regional climate modeling on European scales: A joint standard evaluation of the EURO-CORDEX RCM ensemble. *Geoscientific Model Development*, 7(4):1297–1333.

Krysanova, V., Donnelly, C., Gelfan, A., Gerten, D., Arheimer, B., Hattermann, F., and Kundzewicz, Z. W. (2018). How the performance of hydrological models relates to credibility of projections under climate change. *Hydrological Sciences Journal*, 63(5):696–720.

Krysanova, V. and Hattermann, F. F. (2017). Intercomparison of climate change impacts in 12 large river basins: Overview of methods and summary of results. *Climatic Change*, 141(3):363–379.

Lane, R. A., Coxon, G., Freer, J., Seibert, J., and Wagener, T. (2022). A large-sample investigation into uncertain climate change impacts on high flows across Great Britain. *Hydrology and Earth System Sciences*, 26(21):5535–5554.

Lane, R. A., Coxon, G., Freer, J. E., Wagener, T., Johnes, P. J., Bloomfield, J. P., Greene, S., Macleod, C. J. A., and Reaney, S. M. (2019). Benchmarking the predictive capability of hydrological models for river flow and flood peak predictions across over 1000 catchments in Great Britain. *Hydrology and Earth System Sciences*, 23(10):4011–4032.

Lane, R. A. and Kay, A. L. (2021). Climate Change Impact on the Magnitude and Timing of Hydrological Extremes Across Great Britain. *Frontiers in Water*, 3.

Lange, S. (2019). Trend-preserving bias adjustment and statistical downscaling with ISIMIP3BASD (v1.0). *Geoscientific Model Development*, 12(7):3055–3070.

Lange, S. (2020). ISIMIP3BASD 2.4.1. Zenodo repository. Available at `https://zenodo.org/record/3898426`.

Le Moine, N. (2008). *Le bassin versant de surface vu par le souterrain: une voie d'amélioration des performances et du réalisme des modèles pluie-débit? Sciences de l'environnement.* Doctorat Géosciences et Ressources Naturelles, Université Pierre et Marie Curie, Paris VI.

Leavesley, G., Restrepo, P. J., Markstrom, S., Dixon, M., and Stannard, L. (1996). The Modular Modeling System (MMS): User's Manual. USGS Numbered Series 96-151, Geological Survey (U.S.).

Leavesley, G. H., Lichty, R. W., Troutman, B. M., and Saindon, L. G. (1983). Precipitation-Runoff Modeling System: User's Manual. Technical report.

Lee, J.-Y., Marotzke, J., Bala, G., Cao, L., Corti, S., Dunne, J., Engelbrecht, F., Fischer, E., Fyfe, J., Jones, C., Maycock, A., Mutemi, J., Ndiaye, O., Panickal, S., and Zhou, T. (2021). *Future Global Climate: Scenario-Based Projections and Near- Term Information. In Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou (Eds.)].* Cambridge University Press, Cambridge, UK and New York, NY, USA.

Lee, M.-H., Im, E.-S., and Bae, D.-H. (2019). A comparative assessment of climate change impacts on drought over Korea based on multiple climate projections and multiple drought indices. *Climate Dynamics*, 53(1-2):389–404.

Lehner, F., Coats, S., Stocker, T. F., Pendergrass, A. G., Sanderson, B. M., Raible, C. C., and Smerdon, J. E. (2017). Projected drought risk in 1.5°C and 2°C warmer climates. *Geophysical Research Letters*, 44(14):7419–7428.

Lennox, R. J., Crook, D. A., Moyle, P. B., Struthers, D. P., and Cooke, S. J. (2019). Toward a better understanding of freshwater fish responses to an increasingly drought-stricken world. *Reviews in Fish Biology and Fisheries*, 29(1):71–92.

Liang, X., Lettenmaier, D. P., Wood, E. F., and Burges, S. J. (1994). A simple hydrologically based model of land surface water and energy fluxes for general circulation models. *Journal of Geophysical Research: Atmospheres*, 99(D7):14415–14428.

Lister, D., Osborn, T., Jones, P., and Darch, G. (2018). Observed droughts in the greater Anglian region since 1920. CRU Research Publication 22, http://www.cru.uea.ac.uk/publications/crurp.

Liu, N., Kala, J., Liu, S., Haverd, V., Dell, B., Smettem, K. R., and Harper, R. J. (2020). Drought can offset potential water use efficiency of forest ecosystems from rising atmospheric co2. *Journal of Environmental Sciences*, 90:262–274.

Lloyd-Hughes, B. (2014). The impracticality of a universal drought definition. *Theoretical and Applied Climatology*, 117(3):607–611.

Lo, Y. T. E., Mitchell, D. M., Bohnenstengel, S. I., Collins, M., Hawkins, E., Hegerl, G. C., Joshi, M., and Stott, P. A. (2020). U.K. Climate Projections: Summer Daytime and Nighttime Urban Heat Island Changes in England's Major Cities. *Journal of Climate*, 33(20):9015–9030.

Loon, V. and Van Lanen, H. A. J. (2013). Making the distinction between water scarcity and drought using an observation-modeling framework. *Water Resources Research*, 49(3):1483–1502.

Lowe, J. A., Bernie, D., Bett, P., Bricheno, L., Brown, S., Calvert, D., Clark, R., Edwards, T., Fosser, G., Fung, F., Gohar, L., Good, P., Gregory, J., Harris, G., Howard, T., Kaye, N., Kendon, E., Krijnen, J., Maisey, P., McDonald, R., McInnes, R., McSweeney, C., Mitchell, J. F. B., Murphy, J., Palmer, M., Roberts, C., Rostron, J., Thornton, H., Tinker, J., Tucker, S., Yamazaki, K., and Belcher, S. (2018). UKCP18 Science Overview Report. Met Office, Exeter.

Lu, H., Qin, Z., Lin, S., Chen, X., Chen, B., He, B., Wei, J., and Yuan, W. (2022). Large influence of atmospheric vapor pressure deficit on ecosystem production efficiency. *Nature Communications*, 13(1):1653.

Manabe, S. and Wetherald, R. T. (1967). Thermal Equilibrium of the Atmosphere with a Given Distribution of Relative Humidity. *Journal of the Atmospheric Sciences*, 24(3):241–259.

Mankin, J. S., Seager, R., Smerdon, J. E., Cook, B. I., and Williams, A. P. (2019). Mid-latitude freshwater availability reduced by projected vegetation responses to climate change. *Nature Geoscience*, 12(12):983–988.

Mankin, J. S., Smerdon, J. E., Cook, B. I., Williams, A. P., and Seager, R. (2017). The curious case of projected twenty-first-century drying but greening in the american west. *Journal of Climate*, 30(21):8689–8710.

Manning, C., Widmann, M., Bevacqua, E., Loon, A. F. V., Maraun, D., and Vrac, M. (2018). Soil Moisture Drought in Europe: A Compound Event of Precipitation and Potential Evapotranspiration on Multiple Time Scales. *Journal of Hydrometeorology*, 19(8):1255–1271.

Maraun, D. (2012). Nonstationarities of regional climate model biases in European seasonal mean temperature and precipitation sums. *Geophysical Research Letters*, 39(6).

Maraun, D., Shepherd, T. G., Widmann, M., Zappa, G., Walton, D., Gutierrez, J. M., Hagemann, S., Richter, I., Soares, P. M. M., Hall, A., and Mearns, L. O. (2017). Towards process-informed bias correction of climate change simulations. *Nature Climate Change*, 7(11):764–773.

Maraun, D., Widmann, M., and Gutiérrez, J. M. (2019). Statistical downscaling skill under present climate conditions: A synthesis of the VALUE perfect predictor experiment. *International Journal of Climatology*, 39(9):3692–3703.

Marcuello, C. and Lallana, C. (2003). Indicator Fact Sheet: Water Exploitation Index. European Environment Agency.

Marsh, T., Cole, G., and Wilby, R. (2007). Major droughts in England and Wales, 1800–2006. *Weather*, 62(4):87–93.

Marsh, T., Monkhouse, R., Arnell, N. W., Lees, M., and Reynard, N. (1994). *The 1988-92 Drought.* Institute of Hydrology, Wallingford.

Marsh, T. J. and Turton, P. S. (1996). The 1995 drought — a water resources perspective. *Weather*, 51(2):46–53.

Massari, C., Avanzi, F., Bruno, G., Gabellani, S., Penna, D., and Camici, S. (2022). Evaporation enhancement drives the European water-budget deficit during multi-year droughts. *Hydrology and Earth System Sciences*, 26(6):1527–1543.

Matanó, A., de Ruiter, M. C., Koehler, J., Ward, P. J., and Van Loon, A. F. (2022). Caught Between Extremes: Understanding Human-Water Interactions During Drought-To-Flood Events in the Horn of Africa. *Earth's Future*, 10(9):e2022EF002747.

McKee, T. B., Doesken, N. J., and Kleist, J. (1993). The Relationship of Drought Frequency and Duration to Time Scales. Eighth Conference on Applied Climatology, 17-22 January 1993, Anaheim, California.

Meehl, G. A., Senior, C. A., Eyring, V., Flato, G., Lamarque, J.-F., Stouffer, R. J., Taylor, K. E., and Schlund, M. (2020). Context for interpreting equilibrium climate sensitivity and transient climate response from the CMIP6 Earth system models. *Science Advances*, 6(26):eaba1981.

Meinshausen, M., Lewis, J., McGlade, C., Gütschow, J., Nicholls, Z., Burdon, R., Cozzi, L., and Hackmann, B. (2022). Realization of Paris Agreement pledges may limit warming just below 2 °C. *Nature*, 604(7905):304–309.

Melsen, L. A., Addor, N., Mizukami, N., Newman, A. J., Torfs, P. J. J. F., Clark, M. P., Uijlenhoet, R., and Teuling, A. J. (2018). Mapping (dis)agreement in hydrologic projections. *Hydrology and Earth System Sciences*, 22(3):1775–1791.

Mendoza, P. A., Clark, M. P., Mizukami, N., Gutmann, E. D., Arnold, J. R., Brekke, L. D., and Rajagopalan, B. (2016). How do hydrologic modeling decisions affect the portrayal of climate change impacts? *Hydrological Processes*, 30(7):1071–1095.

Met Office (2019). What is UKCP? `https://www.metoffice.gov.uk/research/approach/collaboration/ukcp/about/what-is-ukcp`.

Met Office Hadley Centre (2018). UKCP18 Regional Projections on a 12km grid over the UK for 1980-2080. `https://catalogue.ceda.ac.uk/uuid/589211abeb844070a95d061c8cc7f604`.

Milly, P. C. D. and Dunne, K. A. (2016). Potential evapotranspiration and continental drying. *Nature Climate Change*, 6(10):946–949.

Miralles, D. G., Gentine, P., Seneviratne, S. I., and Teuling, A. J. (2019). Land-atmospheric feedbacks during droughts and heatwaves: State of the science and current challenges: Land feedbacks during droughts and heatwaves. *Annals of the New York Academy of Sciences*, 1436(1):19–35.

Mishra, A. K. and Singh, V. P. (2010). A review of drought concepts. *Journal of Hydrology*, 391(1-2):202–216.

Monteith, J. L. (1965). Evaporation and environment. *Symposia of the Society for Experimental Biology*, 19:205–234.

Motavita, D. F., Chow, R., Guthke, A., and Nowak, W. (2019). The comprehensive differential split-sample test: A stress-test for hydrological model robustness under climate variability. *Journal of Hydrology*, 573:501–515.

Mott MacDonald (2020). River Wensum: Costessey Boreholes Permit. HRA - Stage 2 Appropriate Assessment. Technical report.

Müller, L. M. and Bahn, M. (2022). Drought legacies and ecosystem responses to subsequent drought. *Global Change Biology*, 28(17):5086–5103.

Murphy, C., Wilby, R. L., Matthews, T., Horvath, C., Crampsie, A., Ludlow, F., Noone, S., Brannigan, J., Hannaford, J., MacLeman, R., and Jobbova, E. (2020). The forgotten drought of 1765–1768: Reconstructing and re-evaluating historical droughts in the British and Irish Isles. *International Journal of Climatology*, n/a(n/a).

Murphy, J. M., Harris, G. R., Sexton, D. M. H., Kendon, E. J., Bett, P. E., Clark, R. T., Eagle, K. E., Fosser, G., Fung, F., Lowe, J. A., McDonald, R. E., McInnes, R. N., McSweeney, C. F., Mitchell, J. F. B., Rostron, J. W., Thornton, H. E., Tucker, S., and Yamazaki, K. (2018). UKCP18 Land Projections: Science Report. Technical report, Met Office, Exeter.

Nash, J. E. and Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part I — A discussion of principles. *Journal of Hydrology*, 10(3):282–290.

National Farmers Union (2016). Farming in East Anglia. `https://www.nfuonline.com:443/about-us/our-offices/east-anglia/east-anglia-key-content/farming-in-east-anglia/`.

Naumann, G., Alfieri, L., Wyser, K., Mentaschi, L., Betts, R. A., Carrao, H., Spinoni, J., Vogt, J., and Feyen, L. (2018). Global Changes in Drought Conditions Under Different Levels of Warming. *Geophysical Research Letters*, 45(7):3285–3296.

Neal, R., Fereday, D., Crocker, R., and Comer, R. E. (2016). A flexible approach to defining weather patterns and their application in weather forecasting over Europe. *Meteorological Applications*, 23(3):389–400.

Newman, A. J., Stone, A. G., Saharia, M., Holman, K. D., Addor, N., and Clark, M. P. (2021). Identifying sensitivities in flood frequency analyses using a stochastic hydrologic modeling system. *Hydrology and Earth System Sciences*, 25(10):5603–5621.

Nguyen, H., Mehrotra, R., and Sharma, A. (2016). Correcting for systematic biases in GCM simulations in the frequency domain. *Journal of Hydrology*, 538:117–126.

Nicolle, P., Andréassian, V., Royer-Gaspard, P., Perrin, C., Thirel, G., Coron, L., and Santos, L. (2021). Technical note: RAT – a robustness assessment test for calibrated and uncalibrated hydrological models. *Hydrology and Earth System Sciences*, 25(9):5013–5027.

Office for National Statistics (2022). Population and household estimates, England and Wales: Census 2021 - Office for National Statistics.

Parker, J. M. and Wilby, R. L. (2013). Quantifying Household Water Demand: A Review of Theory and Practice in the UK. *Water Resources Management*, 27(4):981–1011.

Parry, S., Wilby, R. L., Prudhomme, C., and Wood, P. J. (2016). A systematic assessment of drought termination in the United Kingdom. *Hydrology and Earth System Sciences*, 20(10):4265–4281.

Parsons, D. J., Rey, D., Tanguy, M., and Holman, I. P. (2019). Regional variations in the link between drought indices and reported agricultural impacts of drought. *Agricultural Systems*, 173:119–129.

Pastén-Zapata, E., Jones, J. M., Moggridge, H., and Widmann, M. (2020). Evaluation of the performance of Euro-CORDEX Regional Climate Models for assessing hydrological climate change impacts in Great Britain: A comparison

of different spatial resolutions and quantile mapping bias correction methods. *Journal of Hydrology*, 584:124653.

Pedro-Monzonís, M., Solera, A., Ferrer, J., Estrela, T., and Paredes-Arquiola, J. (2015). A review of water scarcity and drought indexes in water resources planning and management. *Journal of Hydrology*, 527:482–493.

Peel, M. C., Finlayson, B. L., and McMahon, T. A. (2007). Updated world map of the Köppen-Geiger climate classification. *Hydrology and Earth System Sciences*, 11(5):1633–1644.

Pendergrass, A. G. and Knutti, R. (2018). The Uneven Nature of Daily Precipitation and Its Change. *Geophysical Research Letters*, 45(21):11,980–11,988.

Pendergrass, A. G., Meehl, G. A., Pulwarty, R., Hobbins, M., Hoell, A., AghaKouchak, A., Bonfils, C. J. W., Gallant, A. J. E., Hoerling, M., Hoffmann, D., Kaatz, L., Lehner, F., Llewellyn, D., Mote, P., Neale, R. B., Overpeck, J. T., Sheffield, A., Stahl, K., Svoboda, M., Wheeler, M. C., Wood, A. W., and Woodhouse, C. A. (2020). Flash droughts present a new challenge for subseasonal-to-seasonal prediction. *Nature Climate Change*, 10(3):191–199.

Perrin, C., Michel, C., and Andréassian, V. (2003). Improvement of a parsimonious model for streamflow simulation. *Journal of Hydrology*, 279(1):275–289.

Phillips, I. D. and McGregor, a. G. R. (1998). The utility of a drought index for assessing the drought hazard in Devon and Cornwall, South West England. *Meteorological Applications*, 5(4):359–372.

Piani, C., Haerter, J. O., and Coppola, E. (2010). Statistical bias correction for daily precipitation in regional climate models over Europe. *Theoretical and Applied Climatology*, 99(1):187–192.

Pirret, J., Fung, F., Lowe, J., McInnes, R., Mitchell, J., and Murphy, J. (2020a). UKCP Factsheet: Soil Moisture. Met Office, Exeter.

Pirret, J., Fung, F., Mitchell, J. F. B., and McInnes, R. (2020b). Soil moisture and the water cycle in the UK Climate Projections. *EGU General Assembly Conference Abstracts*, page 8852.

Pomeroy, J. W., Gray, D. M., Brown, T., Hedstrom, N. R., Quinton, W. L., Granger, R. J., and Carey, S. K. (2007). The cold regions hydrological model: A platform for basing process representation and model structure on physical evidence. *Hydrological Processes*, 21(19):2650–2667.

Pope, J. O., Brown, K., Fung, F., Hanlon, H. M., Neal, R., Palin, E. J., and Reid, A. (2022). Investigation of future climate change over the British Isles using weather patterns. *Climate Dynamics*, 58(9-10):2405–2419.

Prein, A. F., Langhans, W., Fosser, G., Ferrone, A., Ban, N., Goergen, K., Keller, M., Tölle, M., Gutjahr, O., Feser, F., Brisson, E., Kollet, S., Schmidli, J., Lipzig, N. P. M., and Leung, R. (2015). A review on regional convection-permitting climate modeling: Demonstrations, prospects, and challenges. *Reviews of Geophysics*, 53(2):323–361.

Prudhomme, C., Young, A., Watts, G., Haxton, T., Crooks, S., Williamson, J., Davies, H., Dadson, S., and Allen, S. (2012). The drying up of Britain? A national estimate of changes in seasonal river flows from 11 Regional Climate Model simulations. *Hydrological Processes*, 26(7):1115–1118.

Pushpalatha, R., Perrin, C., Le Moine, N., Mathevet, T., and Andréassian, V. (2011). A downward structural sensitivity analysis of hydrological models to improve low-flow simulation. *Journal of Hydrology*, 411(1):66–76.

Pushpalatha, R., Perrin, C., Moine, N. L., and Andréassian, V. (2012). A review of efficiency criteria suitable for evaluating low-flow simulations. *Journal of Hydrology*, 420–421:171–182.

Rahiz, M. and New, M. (2013). 21st Century Drought Scenarios for the UK. *Water Resources Management*, 27(4):1039–1061.

Rameshwaran, P., Bell, V. A., Brown, M. J., Davies, H. N., Kay, A. L., Rudd, A. C., and Sefton, C. (2022). Use of Abstraction and Discharge Data to Improve the Performance of a National-Scale Hydrological Model. *Water Resources Research*, 58(1):e2021WR029787.

Rangeley-Wilson, C. and CaBA CSRG Panel (2021). Chalk Stream Restoration Strategy 2021 Main Report. CaBA Chalk Stream Restoration Group. Available at `https://catchmentbasedapproach.org/wp-content/uploads/2021/10/CaBA-CSRG-Strategy-MAIN-REPORT-FINAL-12.10.21-Low-Res.pdf`.

Reyniers, N., Osborn, T. J., Addor, N., and Darch, G. (2023). Projected changes in droughts and extreme droughts in Great Britain strongly influenced by the choice of drought index. *Hydrology and Earth System Sciences*, 27(5):1151–1171.

Riahi, K., Rao, S., Krey, V., Cho, C., Chirkov, V., Fischer, G., Kindermann, G., Nakicenovic, N., and Rafaj, P. (2011). RCP 8.5—A scenario of comparatively high greenhouse gas emissions. *Climatic Change*, 109(1):33.

Riahi, K., Schaeffer, R., Arango, J., Calvin, K., Hasegawa, T., Jiang, K., Kriegler, E., Matthews, R., Robertson, S., Sebbit, A. M., Steinberger, J., Khourdajie, A. A., Brutschin, E., Byers, E., Carleton, T., Clarke, L., Chaturvedi, V., Chen, W., and Martínez, J. T. (2022). Mitigation Pathways Compatible with Long-term Goals. In *IPCC, 2022: Climate Change 2022: Mitigation of Climate Change. Contribution of Working Group III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [P.R. Shukla, J. Skea, R. Slade, A. Al Khourdajie, R. van Diemen, D. McCollum, M. Pathak, S. Some, P. Vyas, R. Fradera, M. Belkacemi, A. Hasija, G. Lisboa, S. Luz, J. Malley, (Eds.)]*. Cambridge University Press, Cambridge, UK and New York, NY, USA.

Richards, J. M. (1971). A simple expression for the saturation vapour pressure of water in the range -50 to 140°c. *Journal of Physics D: Applied Physics*, 4(4):L15–L18.

Richardson, D., Fowler, H. J., Kilsby, C. G., and Neal, R. (2018). A new precipitation and drought climatology based on weather patterns. *International Journal of Climatology*, 38(2):630–648.

Ripple, W. J., Wolf, C., Lenton, T. M., Gregg, J. W., Natali, S. M., Duffy, P. B., Rockström, J., and Schellnhuber, H. J. (2023). Many risky feedback loops amplify the need for climate action. *One Earth*, 6(2):86–91.

Robinson, E., Blyth, E., Clark, D., Comyn-Platt, E., and Rudd, A. (2020). Climate hydrology and ecology research support system potential evapotranspiration dataset for Great Britain (1961-2017) [CHESS-PE]. *NERC Environmental Information Data Centre*.

Robinson, E. L., Blyth, E. M., Clark, D. B., Finch, J., and Rudd, A. C. (2017). Trends in atmospheric evaporative demand in Great Britain using high-resolution meteorological data. *Hydrology and Earth System Sciences*, 21(2):1189–1224.

Robinson, E. L., Brown, M. J., Kay, A. L., Lane, R. A., Chapman, R., Bell, V. A., and Blyth, E. M. (2022). Hydro-PE: Gridded datasets of historical and future Penman-Monteith potential evaporation for the United Kingdom. *Earth System Science Data Discussions*, pages 1–44.

Rodda, J. and March, T. (2011). The 1975/76 Drought – a contemporary and retrospective view. Centre for Ecology & Hydrology. Available at `http://nora.nerc.ac.uk/id/eprint/15011/1/CEH_1975-76_Drought_Report_Rodda_and_Marsh.pdf`.

Rojas, R., Feyen, L., Dosio, A., and Bavera, D. (2011). Improving pan-European hydrological simulation of extreme events through statistical bias correction of RCM-driven climate simulations. *Hydrology and Earth System Sciences*, 15(8):2599–2620.

Rosenzweig, C., Arnell, N. W., Ebi, K. L., Lotze-Campen, H., Raes, F., Rapley, C., Smith, M. S., Cramer, W., Frieler, K., Reyer, C. P. O., Schewe, J., van Vuuren, D., and Warszawski, L. (2017). Assessing inter-sectoral climate change risks: The role of ISIMIP. *Environmental Research Letters*, 12(1):010301.

Rudd, A. C., Bell, V. A., and Kay, A. L. (2017). National-scale analysis of simulated hydrological droughts (1891-2015). *Journal of Hydrology*, 550:368–385.

Rudd, A. C., Kay, A. L., and Bell, V. A. (2019). National-scale analysis of future river flow and soil moisture droughts: Potential changes in drought characteristics. *Climatic Change*, 156(3):323–340.

Rust, W., Cuthbert, M., Bloomfield, J., Corstanje, R., Howden, N., and Holman, I. (2021). Exploring the role of hydrological pathways in modulating multi-annual climate teleconnection periodicities from UK rainfall to streamflow. *Hydrology and Earth System Sciences*, 25(4):2223–2237.

Saavedra, D., Mendoza, P. A., Addor, N., Llauca, H., and Vargas, X. (2022). A multi-objective approach to select hydrological models and constrain structural uncertainties for climate impact assessments. *Hydrological Processes*, 36(1):e14446.

Samaniego, L., Kumar, R., Breuer, L., Chamorro, A., Flörke, M., Pechlivanidis, I. G., Schäfer, D., Shah, H., Vetter, T., Wortmann, M., and Zeng, X. (2017). Propagation of forcing and model uncertainties on to hydrological drought characteristics in a multi-model century-long experiment in large river basins. *Climatic Change*, 141(3):435–449.

Samaniego, L., Thober, S., Kumar, R., Wanders, N., Rakovec, O., Pan, M., Zink, M., Sheffield, J., Wood, E. F., and Marx, A. (2018). Anthropogenic warming exacerbates European soil moisture droughts. *Nature Climate Change*, 8(5):421–426.

Satoh, Y., Shiogama, H., Hanasaki, N., Pokhrel, Y., Boulange, J. E. S., Burek, P., Gosling, S. N., Grillakis, M., Koutroulis, A., Müller Schmied, H., Thiery, W., and Yokohata, T. (2021). A quantitative evaluation of the issue of drought definition: A source of disagreement in future drought assessments. *Environmental Research Letters*, 16(10):104001.

Savelli, E., Rusca, M., Cloke, H., and Di Baldassarre, G. (2021). Drought and society: Scientific progress, blind spots, and future prospects. *WIREs Climate Change*, n/a(n/a):e761.

Scheff, J., Mankin, J. S., Coats, S., and Liu, H. (2021). CO2-plant effects do not account for the gap between dryness indices and projected dryness impacts in CMIP6 or CMIP5. *Environmental Research Letters*, 16(3):034018.

Schönbeck, L. C., Schuler, P., Lehmann, M. M., Mas, E., Mekarni, L., Pivovaroff, A. L., Turberg, P., and Grossiord, C. (2022). Increasing temperature and vapour pressure deficit lead to hydraulic damages in the absence of soil drought. *Plant, Cell & Environment*, 45(11):3275–3289.

Schumacher, D. L., Keune, J., Dirmeyer, P., and Miralles, D. G. (2022a). Drought self-propagation in drylands due to land–atmosphere feedbacks. *Nature Geoscience*, pages 1–7.

Schumacher, D. L., Zachariah, M., Otto, F., Barnes, C., Philip, S., Kew, S., Vahlberg, M., Singh, R., Heinrich, D., Arrighi, J., van Aalst, M., Thalheimer, L., Raju, E., Hauser, M., Hirschi, M., Gudmundsson, L., Rodell, M., Li, S., Yang, W., Vautard, R., Harrington, L. J., and Seneviratne, S. I. (2022b). High temperatures exacerbated by climate change made 2022 Northern Hemisphere soil moisture droughts more likely. World Weather Attribution.

Seager, R., Liu, H., Kushnir, Y., Osborn, T. J., Simpson, I. R., Kelley, C. R., and Nakamura, J. (2020). Mechanisms of Winter Precipitation Variability in the European–Mediterranean Region Associated with the North Atlantic Oscillation. *Journal of Climate*, 33(16):7179–7196.

Seiller, G. and Anctil, F. (2016). How do potential evapotranspiration formulas influence hydrological projections? *Hydrological Sciences Journal*, 61(12):2249–2266.

Seneviratne, S. I. (2012). Historical drought trends revisited. *Nature*, 491(7424):338–339.

Seneviratne, S. I., Corti, T., Davin, E. L., Hirschi, M., Jaeger, E. B., Lehner, I., Orlowsky, B., and Teuling, A. J. (2010). Investigating soil moisture–climate interactions in a changing climate: A review. *Earth-Science Reviews*, 99(3):125–161.

Seneviratne, S. I., Lüthi, D., Litschi, M., and Schär, C. (2006). Land–atmosphere coupling and climate change in Europe. *Nature*, 443(7108):205–209.

Seneviratne, S. I., Zhang, X., Adnan, M., Badi, W., Dereczynski, C., Di Luca, A., Ghosh, S., Iskandar, I., Kossin, J., Lewis, S., Otto, F., Pinto, I., Satoh,

M., Vicente-Serrano, S. M., Wehner, M., and Zhou, B. (2021). Weather and Climate Extreme Events in a Changing Climate. In *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou (Eds.)].*, pages 1513–1766. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Severn Trent Water (2022). Draft Water Resources Management Plan 2024 Main Narrative. Severn Trent Water, Coventry.

Sheffield, J., Andreadis, K. M., Wood, E. F., and Lettenmaier, D. P. (2009). Global and Continental Drought in the Second Half of the Twentieth Century: Severity–Area–Duration Analysis and Temporal Variability of Large-Scale Events. *Journal of Climate*, 22(8):1962–1981.

Sheffield, J., Wood, E. F., and Roderick, M. L. (2012). Little change in global drought over the past 60 years. *Nature*, 491(7424):435–438.

Shepherd, T. G. (2014). Atmospheric circulation as a source of uncertainty in climate change projections. *Nature Geoscience*, 7(10):703–708.

Shepherd, T. G., Boyd, E., Calel, R. A., Chapman, S. C., Dessai, S., Dima-West, I. M., Fowler, H. J., James, R., Maraun, D., Martius, O., Senior, C. A., Sobel, A. H., Stainforth, D. A., Tett, S. F. B., Trenberth, K. E., van den Hurk, B. J. J. M., Watkins, N. W., Wilby, R. L., and Zenghelis, D. A. (2018). Storylines: An alternative approach to representing uncertainty in physical aspects of climate change. *Climatic Change*, 151(3-4):555–571.

Shepherd, T. G. and Lloyd, E. A. (2021). Meaningful climate science. *Climatic Change*, 169(1):17.

Shiau, J. T. and Modarres, R. (2009). Copula-based drought severity-duration-frequency analysis in Iran. *Meteorological Applications*, 16(4):481–489.

Skinner, C. B., Poulsen, C. J., Chadwick, R., Diffenbaugh, N. S., and Fiorella, R. P. (2017). The role of plant co2 physiological forcing in shaping future daily-scale precipitation. *Journal of Climate*, 30(7):2319–2340.

Smith, K., Tanguy, M., Hannaford, J., and Prudhomme, C. (2018a). Historic reconstructions of daily river flow for 303 UK catchments (1891-2015).

Smith, K. A., Barker, L. J., Tanguy, M., Parry, S., Harrigan, S., Legg, T. P., Prudhomme, C., and Hannaford, J. (2019). A multi-objective ensemble

approach to hydrological modelling in the UK: An application to historic drought reconstruction. *Hydrology and Earth System Sciences*, 23(8):3247–3268.

Smith, K. A., Wilby, R. L., Broderick, C., Prudhomme, C., Matthews, T., Harrigan, S., and Murphy, C. (2018b). Navigating Cascades of Uncertainty — As Easy as ABC? Not Quite.... *Journal of Extreme Events*, 05(01):1850007.

Sousa, P. M., Trigo, R. M., Barriopedro, D., Soares, P. M. M., Ramos, A. M., and Liberato, M. L. R. (2017). Responses of European precipitation distributions and regimes to different blocking locations. *Climate Dynamics*, 48(3):1141–1160.

Spinoni, J., Vogt, J. V., Naumann, G., Barbosa, P., and Dosio, A. (2018). Will drought events become more frequent and severe in Europe? *International Journal of Climatology*, 38(4):1718–1736.

Spraggs, G., Peaver, L., Jones, P., and Ede, P. (2015). Re-construction of historic drought in the Anglian Region (UK) over the period 1798–2010 and the implications for water resources and drought management. *Journal of Hydrology*, 526:231–252.

Stagge, J. H., Kingston, D. G., Tallaksen, L. M., and Hannah, D. M. (2017). Observed drought indices show increasing divergence across Europe. *Scientific Reports*, 7(1):1–10.

Stagge, J. H., Kohn, I., Tallaksen, L. M., and Stahl, K. (2015a). Modeling drought impact occurrence based on meteorological drought indices in Europe. *Journal of Hydrology*, 530:37–50.

Stagge, J. H., Tallaksen, L. M., Gudmundsson, L., Loon, A. F. V., and Stahl, K. (2015b). Candidate Distributions for Climatological Drought Indices (SPI and SPEI). *International Journal of Climatology*, 35(13):4027–4040.

Stahl, K., Kohn, I., Blauhut, V., Urquijo, J., De Stefano, L., Acácio, V., Dias, S., Stagge, J. H., Tallaksen, L. M., Kampragou, E., Van Loon, A. F., Barker, L. J., Melsen, L. A., Bifulco, C., Musolino, D., de Carli, A., Massarutto, A., Assimacopoulos, D., and Van Lanen, H. A. J. (2016). Impacts of European drought events: Insights from an international database of text-based reports. *Natural Hazards and Earth System Sciences*, 16(3):801–819.

Stahl, K., Vidal, J.-P., Hannaford, J., Tijdeman, E., Laaha, G., Gauster, T., and Tallaksen, L. M. (2020). The challenges of hydrological drought definition, quantification and communication: An interdisciplinary perspective. In *Proceedings of the International Association of Hydrological Sciences*, volume 383, pages 291–295. Copernicus GmbH.

Stanke, C., Kerac, M., Prudhomme, C., Medlock, J., and Murray, V. (2013). Health Effects of Drought: A Systematic Review of the Evidence. *PLoS Currents*, 5.

Staudinger, M., Stahl, K., Seibert, J., Clark, M. P., and Tallaksen, L. M. (2011). Comparison of hydrological model structures based on recession and low flow simulations. *Hydrology and Earth System Sciences*, 15(11):3447–3459.

Stoelzle, M., Stahl, K., Morhard, A., and Weiler, M. (2014). Streamflow sensitivity to drought scenarios in catchments with different geology. *Geophysical Research Letters*, 41(17):6174–6183.

Sutanto, S. J. and Van Lanen, H. A. J. (2021). Streamflow drought: Implication of drought definitions and its application for drought forecasting. *Hydrology and Earth System Sciences Discussions*, pages 1–29.

Switanek, M. B., Troch, P. A., Castro, C. L., Leuprecht, A., Chang, H.-I., Mukherjee, R., and Demaria, E. M. C. (2017). Scaled distribution mapping: A bias correction method that preserves raw climate model projected changes. *Hydrology and Earth System Sciences*, 21(6):2649–2666.

Tallaksen, L. M., Madsen, H., and Clausen, B. (1997). On the definition and modelling of streamflow drought duration and deficit volume. *Hydrological Sciences Journal*, 42(1):15–33.

Tanguy, M., Haslinger, K., Svensson, C., Parry, S., Barker, L. J., Hannaford, J., and Prudhomme, C. (2021). Regional differences in spatiotemporal drought characteristics in Great Britain. *Frontiers in Environmental Science*, 9.

Teng, J., Potter, N. J., Chiew, F. H. S., Zhang, L., Wang, B., Vaze, J., and Evans, J. P. (2015). How does bias correction of regional climate model precipitation affect modelled runoff? *Hydrology and Earth System Sciences*, 19(2):711–728.

Terrier, M., Perrin, C., de Lavenne, A., Andréassian, V., Lerat, J., and Vaze, J. (2021). Streamflow naturalization methods: A review. *Hydrological Sciences Journal*, 66(1):12–36.

Teuling, A. J., de Badts, E. A. G., Jansen, F. A., Fuchs, R., Buitink, J., Hoek van Dijke, A. J., and Sterling, S. M. (2019). Climate change, reforestation/afforestation, and urbanization impacts on evapotranspiration and streamflow in Europe. *Hydrology and Earth System Sciences*, 23(9):3631–3652.

Teuling, A. J., Hirschi, M., Ohmura, A., Wild, M., Reichstein, M., Ciais, P., Buchmann, N., Ammann, C., Montagnani, L., Richardson, A. D., Wohlfahrt,

G., and Seneviratne, S. I. (2009). A regional perspective on trends in continental evaporation. *Geophysical Research Letters*, 36(2).

Teuling, A. J., Seneviratne, S. I., Stöckli, R., Reichstein, M., Moors, E., Ciais, P., Luyssaert, S., van den Hurk, B., Ammann, C., Bernhofer, C., Dellwik, E., Gianelle, D., Gielen, B., Grünwald, T., Klumpp, K., Montagnani, L., Moureaux, C., Sottocornola, M., and Wohlfahrt, G. (2010). Contrasting response of European forest and grassland energy exchange to heatwaves. *Nature Geoscience*, 3(10):722–727.

Teutschbein, C. and Seibert, J. (2012). Bias correction of regional climate model simulations for hydrological climate-change impact studies: Review and evaluation of different methods. *Journal of Hydrology*, 456–457:12–29.

Thames Water (2022). Draft Water Resources Management Plan 2024 - Section 6: Allowing for Risk and Uncertainty, and Baseline Supply-Demand Balance. Technical report.

Thompson, J. R., Green, A. J., and Kingston, D. G. (2014). Potential evapotranspiration-related uncertainty in climate change impacts on river flow: An assessment for the Mekong River basin. *Journal of Hydrology*, 510:259–279.

Tijdeman, E., Hannaford, J., and Stahl, K. (2018). Human influences on streamflow drought characteristics in England and Wales. *Hydrology and Earth System Sciences*, 22(2):1051–1064.

Todini, E. (1996). The ARNO rainfall—runoff model. *Journal of Hydrology*, 175:339–382.

Todorović, A., Grabs, T., and Teutschbein, C. (2022). Advancing traditional strategies for testing hydrological model fitness in a changing climate. *Hydrological Sciences Journal*, 67(12):1790–1811.

Tomas-Burguera, M., Vicente-Serrano, S. M., Peña-Angulo, D., Domínguez-Castro, F., Noguera, I., and Kenawy, A. E. (2020). Global Characterization of the Varying Responses of the Standardized Precipitation Evapotranspiration Index to Atmospheric Evaporative Demand. *Journal of Geophysical Research: Atmospheres*, 125(17):e2020JD033017.

Tootoonchi, F., Haerter, J. O., Todorović, A., Räty, O., Grabs, T., and Teutschbein, C. (2022). Uni-and multivariate bias adjustment methods in nordic catchments: Complexity and performance in a changing climate. *Science of the Total Environment*, 853:158615.

Touma, D., Ashfaq, M., Nayak, M. A., Kao, S.-C., and Diffenbaugh, N. S. (2015). A multi-model and multi-index evaluation of drought characteristics in the 21st century. *Journal of Hydrology*, 526:196–207.

Trotter, L., Saft, M., Peel, M. C., and Fowler, K. J. A. (2022). Symptoms of performance degradation during multi-annual drought: A large-sample, multi-model study. *Earth and Space Science Open Archive*.

Turner, S., Barker, L. J., Hannaford, J., Muchan, K., Parry, S., and Sefton, C. (2021). The 2018/2019 drought in the UK: A hydrological appraisal. *Weather*, 99(99).

UK Centre for Ecology and Hydrology (n.d.). Drought inventory. https://www.ceh.ac.uk/our-science/projects/drought-inventory. [Accessed: 2023-03-16].

UKCEH (2023). Catchment Info for 34004 - Wensum at Costessey Mill. `https://nrfa.ceh.ac.uk/data/station/spatial/34004`. [Accessed: 2023-03-07].

ukcp18 (2021). UKCP18 Spatial Files, GitHub [data set]. Data set. Available at `https://github.com/ukcp-data/ukcp-spatial-files`. [Accessed: 2021-10-12].

United Nations (2015). Paris agreement to the united nations framework convention on climate change. T.I.A.S. No. 16-1104.

United Nations Environment Programme (1992). *World Atlas of Desertification*. 0-340-55512-2.

United Nations Environment Programme (2022). Emissions Gap Report 2022: The Closing Window — Climate crisis calls for rapid transformation of societies. Technical report, Nairobi.

Vallis, G. K., Colyer, G., Geen, R., Gerber, E., Jucker, M., Maher, P., Paterson, A., Pietschnig, M., Penn, J., and Thomson, S. I. (2018). Isca, v1.0: A framework for the global modelling of the atmospheres of Earth and other planets at varying levels of complexity. *Geoscientific Model Development*, 11(3):843–859.

van der Wiel, K., Selten, F. M., Bintanja, R., Blackport, R., and Screen, J. A. (2020). Ensemble climate-impact modelling: Extreme impacts from moderate meteorological conditions. *Environmental Research Letters*, 15(3):034050.

van Huijgevoort, M. H. J., van Lanen, H. A. J., Teuling, A. J., and Uijlenhoet, R. (2014). Identification of changes in hydrological drought characteristics from a multi-GCM driven ensemble constrained by observed discharge. *Journal of Hydrology*, 512:421–434.

Van Lanen, H. a. J., Wanders, N., Tallaksen, L. M., and Van Loon, A. F. (2013). Hydrological drought across the world: Impact of climate and physical catchment structure. *Hydrology and Earth System Sciences*, 17(5):1715–1732.

Van Loon, A. F., Gleeson, T., Clark, J., Van Dijk, A. I. J. M., Stahl, K., Hannaford, J., Di Baldassarre, G., Teuling, A. J., Tallaksen, L. M., Uijlenhoet, R., Hannah, D. M., Sheffield, J., Svoboda, M., Verbeiren, B., Wagener, T., Rangecroft, S., Wanders, N., and Van Lanen, H. A. J. (2016a). Drought in the Anthropocene. *Nature Geoscience*, 9(2):89–91.

Van Loon, A. F. and Laaha, G. (2015). Hydrological drought severity explained by climate and catchment characteristics. *Journal of Hydrology*, 526:3–14.

Van Loon, A. F., Ploum, S. W., Parajka, J., Fleig, A. K., Garnier, E., Laaha, G., and Van Lanen, H. a. J. (2015). Hydrological drought types in cold climates: Quantitative analysis of causing factors and qualitative survey of impacts. *Hydrology and Earth System Sciences*, 19(4):1993–2016.

Van Loon, A. F., Stahl, K., Di Baldassarre, G., Clark, J., Rangecroft, S., Wanders, N., Gleeson, T., Van Dijk, A. I. J. M., Tallaksen, L. M., Hannaford, J., Uijlenhoet, R., Teuling, A. J., Hannah, D. M., Sheffield, J., Svoboda, M., Verbeiren, B., Wagener, T., and Van Lanen, H. A. J. (2016b). Drought in a human-modified world: Reframing drought definitions,understanding, and analysis approaches. *Hydrology and Earth System Sciences*, 20(9):3631–3650.

Van Loon, A. F. and Van Lanen, H. A. J. (2012). A process-based typology of hydrological drought. *Hydrology and Earth System Sciences*, 16(7):1915–1946.

Van Loon, A. F. V., Rangecroft, S., Coxon, G., Werner, M., Wanders, N., Baldassarre, G. D., Tijdeman, E., Bosman, M., Gleeson, T., Nauditt, A., Aghakouchak, A., Breña-Naranjo, J. A., Cenobio-Cruz, O., Costa, A. C., Fendekova, M., Jewitt, G., Kingston, D. G., Loft, J., Mager, S. M., Mallakpour, I., Masih, I., Maureira-Cortés, H., Toth, E., Oel, P. V., Ogtrop, F. V., Verbist, K., Vidal, J.-P., Wen, L., Yu, M., Yuan, X., Zhang, M., and Lanen, H. A. J. V. (2022). Streamflow droughts aggravated by human activities despite management. *Environmental Research Letters*, 17(4):044059.

Vansteenkiste, T., Tavakoli, M., Ntegeka, V., De Smedt, F., Batelaan, O., Pereira, F., and Willems, P. (2014). Intercomparison of hydrological model structures and calibration approaches in climate scenario impact projections. *Journal of Hydrology*, 519:743–755.

Vautard, R., Kadygrov, N., Iles, C., Boberg, F., Buonomo, E., Bülow, K., Coppola, E., Corre, L., van Meijgaard, E., Nogherotto, R., Sandstad, M., Schwingshackl, C., Somot, S., Aalbers, E., Christensen, O. B., Ciarlo, J. M.,

Demory, M.-E., Giorgi, F., Jacob, D., Jones, R. G., Keuler, K., Kjellström, E., Lenderink, G., Levavasseur, G., Nikulin, G., Sillmann, J., Solidoro, C., Sørland, S. L., Steger, C., Teichmann, C., Warrach-Sagi, K., and Wulfmeyer, V. (2021). Evaluation of the Large EURO-CORDEX Regional Climate Model Ensemble. *Journal of Geophysical Research: Atmospheres*, 126(17):e2019JD032344.

Vaze, J., Post, D. A., Chiew, F. H. S., Perraud, J. M., Viney, N. R., and Teng, J. (2010). Climate non-stationarity – Validity of calibrated rainfall–runoff models for use in climate change studies. *Journal of Hydrology*, 394(3):447–457.

Vernieuwe, H., Baets, B. D., and Verhoest, N. E. C. (2020). A mathematical morphology approach for a qualitative exploration of drought events in space and time. *International Journal of Climatology*, 40(1):530–543.

Vetter, T., Reinhardt, J., Flörke, M., van Griensven, A., Hattermann, F., Huang, S., Koch, H., Pechlivanidis, I. G., Plötner, S., Seidou, O., Su, B., Vervoort, R. W., and Krysanova, V. (2017). Evaluation of sources of uncertainty in projected hydrological changes under climate change in 12 large-scale river basins. *Climatic Change*, 141(3):419–433.

Vicente-Serrano, S. M., Beguería, S., and López-Moreno, J. I. (2009). A Multiscalar Drought Index Sensitive to Global Warming: The Standardized Precipitation Evapotranspiration Index. *Journal of Climate*, 23(7):1696–1718.

Vicente-Serrano, S. M., Domínguez-Castro, F., Murphy, C., Hannaford, J., Reig, F., Peña-Angulo, D., Tramblay, Y., Trigo, R. M., Mac Donald, N., Luna, M. Y., Mc Carthy, M., Van der Schrier, G., Turco, M., Camuffo, D., Noguera, I., García-Herrera, R., Becherini, F., Della Valle, A., Tomas-Burguera, M., and El Kenawy, A. (2021). Long-term variability and trends in meteorological droughts in Western Europe (1851–2018). *International Journal of Climatology*, 41(S1):E690–E717.

Vicente-Serrano, S. M., McVicar, T. R., Miralles, D. G., Yang, Y., and Tomas-Burguera, M. (2020). Unraveling the influence of atmospheric evaporative demand on drought and its response to climate change. *WIREs Climate Change*, 11:e632.

Vicente-Serrano, S. M., Miralles, D. G., McDowell, N., Brodribb, T., Domínguez-Castro, F., Leung, R., and Koppa, A. (2022). The uncertain role of rising atmospheric CO2 on global plant transpiration. *Earth-Science Reviews*, 230:104055.

Vicente-Serrano, S. M., Peña-Gallardo, M., Hannaford, J., Murphy, C., Lorenzo-Lacruz, J., Dominguez-Castro, F., López-Moreno, J. I., Beguería, S., Noguera, I., Harrigan, S., and Vidal, J.-P. (2019). Climate, Irrigation, and Land Cover

Change Explain Streamflow Trends in Countries Bordering the Northeast Atlantic. *Geophysical Research Letters*, 46(19):10821–10833.

Vicente-Serrano, S. M., Van der Schrier, G., Beguería, S., Azorin-Molina, C., and Lopez-Moreno, J.-I. (2015). Contribution of precipitation and reference evapotranspiration to drought indices under different climates. *Journal of Hydrology*, 526:42–54.

Vidal, J. P., Hingray, B., Magand, C., Sauquet, E., and Ducharne, A. (2016). Hierarchy of climate and hydrological uncertainties in transient low-flow projections. *Hydrology and Earth System Sciences*, 20(9):3651–3672.

Vidal, J.-P. and Wade, S. (2009). A multimodel assessment of future climatological droughts in the United Kingdom. *International Journal of Climatology*, 29(14):2056–2071.

Visser-Quinn, A., Beevers, L., Collet, L., Formetta, G., Smith, K., Wanders, N., Thober, S., Pan, M., and Kumar, R. (2019). Spatio-temporal analysis of compound hydro-hazard extremes across the UK. *Advances in Water Resources*, 130:77–90.

von Uexkull, N., Croicu, M., Fjelde, H., and Buhaug, H. (2016). Civil conflict sensitivity to growing-season drought. *Proceedings of the National Academy of Sciences*, 113(44):12391–12396.

Wagener, T., Dadson, S. J., Hannah, D. M., Coxon, G., Beven, K., Bloomfield, J. P., Buytaert, W., Cloke, H., Bates, P., Holden, J., Parry, L., Lamb, R., Chappell, N. A., Fry, M., and Old, G. (2021). Knowledge gaps in our perceptual model of Great Britain's hydrology. *Hydrological Processes*, 35(7):e14288.

Wanders, N. and Wada, Y. (2015). Human and climate impacts on the 21st century hydrological drought. *Journal of Hydrology*, 526:208–220.

Wang, F., Wang, Z., Yang, H., Di, D., Zhao, Y., Liang, Q., and Hussain, Z. (2020). Comprehensive evaluation of hydrological drought and its relationships with meteorological drought in the Yellow River basin, China. *Journal of Hydrology*, 584:124751.

Wang, T., Tu, X., Singh, V. P., Chen, X., and Lin, K. (2021). Global data assessment and analysis of drought characteristics based on CMIP6. *Journal of Hydrology*, 596:126091.

Watts, G., Battarbee, R. W., Bloomfield, J. P., Crossman, J., Daccache, A., Durance, I., Elliott, J. A., Garner, G., Hannaford, J., Hannah, D. M., Hess, T., Jackson, C. R., Kay, A. L., Kernan, M., Knox, J., Mackay, J., Monteith, D. T., Ormerod, S. J., Rance, J., Stuart, M. E., Wade, A. J., Wade, S. D.,

Weatherhead, K., Whitehead, P. G., and Wilby, R. L. (2015). Climate change and water in the UK – past changes and future prospects:. *Progress in Physical Geography*.

Wendt, D. E., Van Loon, A. F., Bloomfield, J. P., and Hannah, D. M. (2020). Asymmetric impact of groundwater use on groundwater droughts. *Hydrology and Earth System Sciences*, 24(10):4853–4868.

West, H., Quinn, N., and Horswell, M. (2019). Regional rainfall response to the North Atlantic Oscillation (NAO) across Great Britain. *Hydrology Research*, 50(6):1549–1563.

Wigley, T. M. L. and Jones, P. D. (1985). Influences of precipitation changes and direct CO 2 effects on streamflow. *Nature*, 314(6007):149–152.

Wilby, R. L. (2005). Uncertainty in water resource model parameters used for climate change impact assessment. *Hydrological Processes*, 19(16):3201–3219.

Wilby, R. L. and Dessai, S. (2010). Robust adaptation to climate change. *Weather*, 65(7):180–185.

Wilby, R. L. and Harris, I. (2006). A framework for assessing uncertainties in climate change impacts: Low-flow scenarios for the River Thames, UK. *Water Resources Research*, 42(2).

Wilby, R. L., O'Hare, G., and Barnsley, N. (1997). The North Atlantic Oscillation and British Isles climate variability, 1865–1996. *Weather*, 52(9):266–276.

Wilhite, D. A. and Glantz, M. H. (1985). Understanding the Drought Phenomenon: The Role of Definitions. *Drought Mitigation Center Faculty Publications*, 20.

WMO and GWP (2016). *Handbook of Drought Indicators and Indices (M. Svoboda and B.A. Fuchs)*. Integrated Drought Management Programme, Integrated Drought Management Tools and Guidelines Series 2. Geneva.

Wood, A., Wake, H., and McKendrick-Smith, K. (2022). River Wensum Special Area of Conservation – Evidence Pack. Natural England Technical Information Note. TIN201. Technical report, Natural England.

Wood, E. F., Lettenmaier, D. P., and Zartarian, V. G. (1992). A land-surface hydrology parameterization with subgrid variability for general circulation models. *Journal of Geophysical Research*, 97(D3):2717.

Wright, B., Stanford, B. D., Reinert, A., Routt, J. C., Khan, S. J., and Debroux, J. F. (2013). Managing water quality impacts from drought on drinking water

supplies. *Journal of Water Supply: Research and Technology-Aqua*, 63(3):179–188.

Wu, J., Yao, H., Yuan, X., and Lin, B. (2022). Dissolved organic carbon response to hydrological drought characteristics: Based on long-term measurements of headwater streams. *Water Research*, 215:118252.

Wutich, A., Brewis, A., and Tsai, A. (2020). Water and mental health. *WIREs Water*, 7(5):e1461.

Xenochristou, M., Kapelan, Z., and Hutton, C. (2020). Using Smart Demand-Metering Data and Customer Characteristics to Investigate Influence of Weather on Water Consumption in the UK. *Journal of Water Resources Planning and Management*, 146(2):04019073.

Xiong, J., Guo, S., Abhishek, Chen, J., and Yin, J. (2022). Global evaluation of the "dry gets drier, and wet gets wetter" paradigm from a terrestrial water storage change perspective. *Hydrology and Earth System Sciences*, 26(24):6457–6476.

Yang, T., Ding, J., Liu, D., Wang, X., and Wang, T. (2019). Combined Use of Multiple Drought Indices for Global Assessment of Dry Gets Drier and Wet Gets Wetter Paradigm. *Journal of Climate*, 32(3):737–748.

Yevjevich, V. (1967). An objective approach to definitions and investigations of continental hydrologic droughts. *Journal of Hydrology*, 7(3):353.

Yu, M., Wang, G., and Chen, H. (2016). Quantifying the impacts of land surface schemes and dynamic vegetation on the model dependency of projected changes in surface energy and water budgets. *Journal of Advances in Modeling Earth Systems*, 8(1):370–386.

Zelenhasić, E. and Salvai, A. (1987). A method of streamflow drought analysis. *Water Resources Research*, 23(1):156–168.

Zhang, T., Ramakrishnan, R., and Livny, M. (1997). BIRCH: A New Data Clustering Algorithm and Its Applications. *Data Mining and Knowledge Discovery*, 1:141–182.

Zhao, M., A, G., Liu, Y., and Konings, A. G. (2022). Evapotranspiration frequently increases during droughts. *Nature Climate Change*, pages 1–7.

Zhao, R. J. (1977). Flood forecasting method for humid regions of China. *East China College of Hydraulic Engineering, Nanjing, China.*

Zhao, R. J. (1984). Watershed hydrological modelling. *Water Resources and Electric Power Press, Beijing*, 1984.

Zheng, W., Lamačová, A., Yu, X., Krám, P., Hruška, J., Zahradníček, P., Štěpánek, P., and Farda, A. (2021). Assess hydrological responses to a warming climate at the Lysina Critical Zone Observatory in Central Europe. *Hydrological Processes*, 35(9):e14281.

Zscheischler, J., Fischer, E. M., and Lange, S. (2019). The effect of univariate bias adjustment on multivariate hazard estimates. *Earth System Dynamics*, 10(1):31–43.

A

Chapter 3: Evaluation, analysis and bias correction of the UKCP!8 regional climate projections

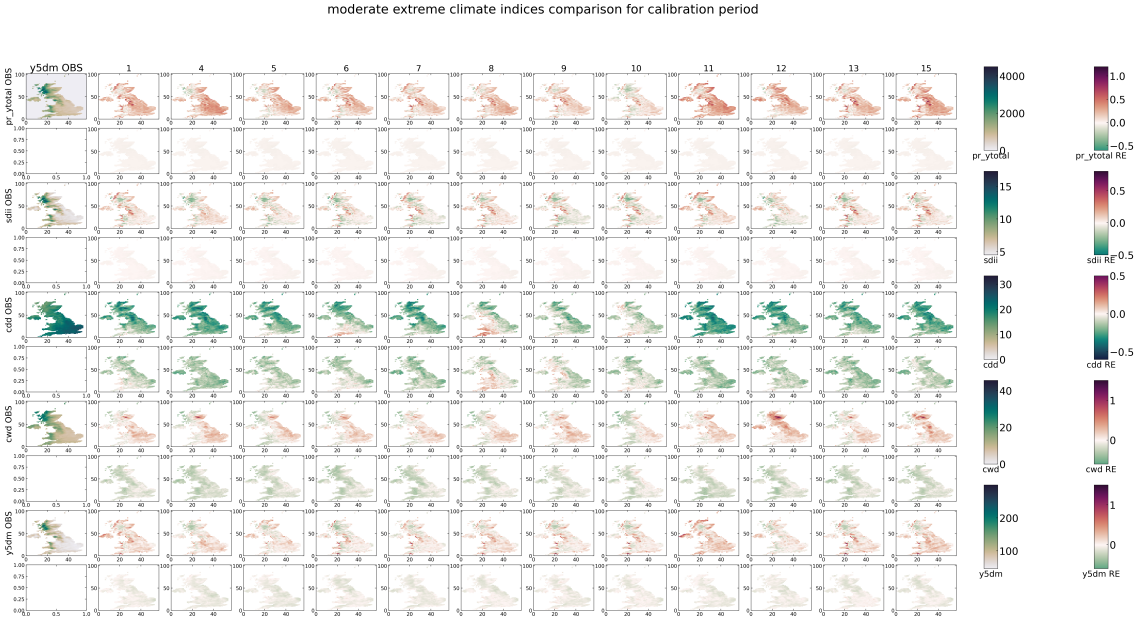

Figure A.0.1: Maps of observed values of precipitation metrics (left column) and their biases in UKCP18-RCM (all other columns, headed by ensemble member number) for the reference period (1981-2010). Odd rows show the observed values and the bias before bias adjustment, even rows show the remaining bias after bias adjustment by BCI3.

B

Chapter 4: Projected changes in droughts and extreme droughts in Great Britain strongly invluenced by the choice of drought index

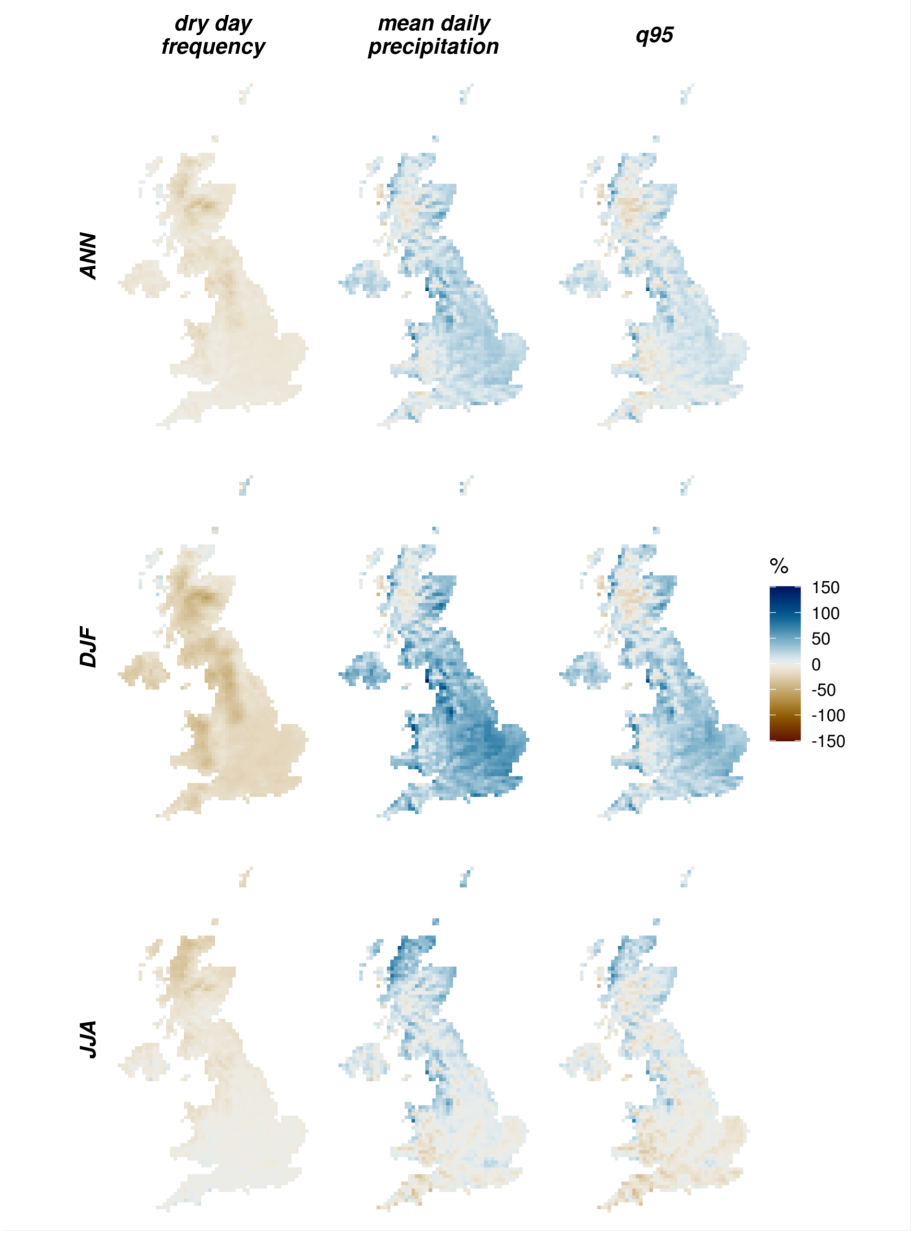

Figure B.0.1: Mean precipitation biases in UKCP18-RCM for 1981-2010, expressed as a percentage of the observed values. The bias for each ensemble member was computed and the mean across the ensemble is shown here. Dry-day frequency is the percentage of days with P ¡ 1 mm; q95 is the 0.95 quantile of precipitation. Created by Nicole Forstenhäusler.

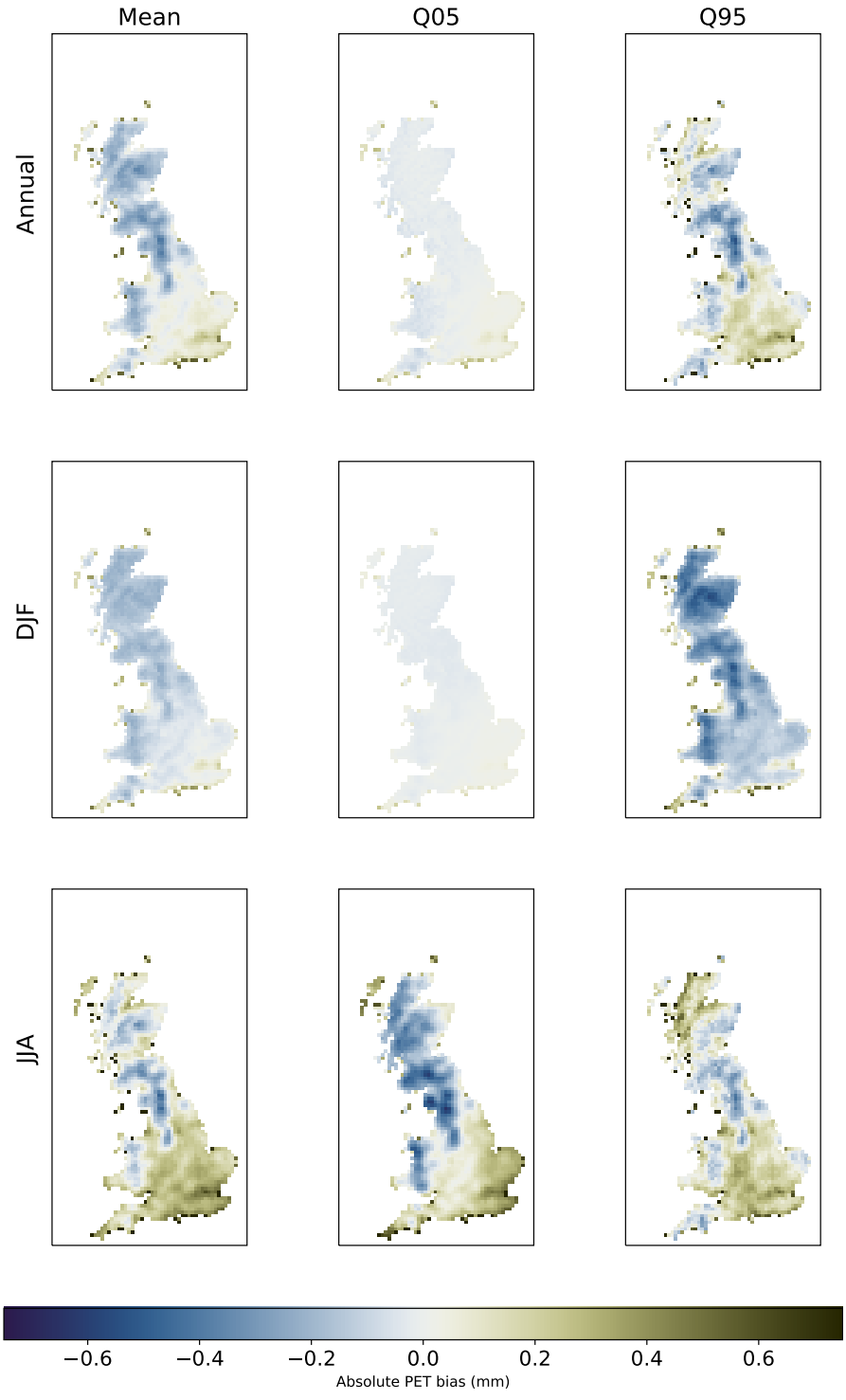

Figure B.0.2: Mean PET biases (mm) in UKCP18-RCM for 1981-2010. The bias for each ensemble member was computed and the mean across the ensemble is shown here. Q05 and Q95 are the biases in the 0.05 and 0.95 quantiles respectively.

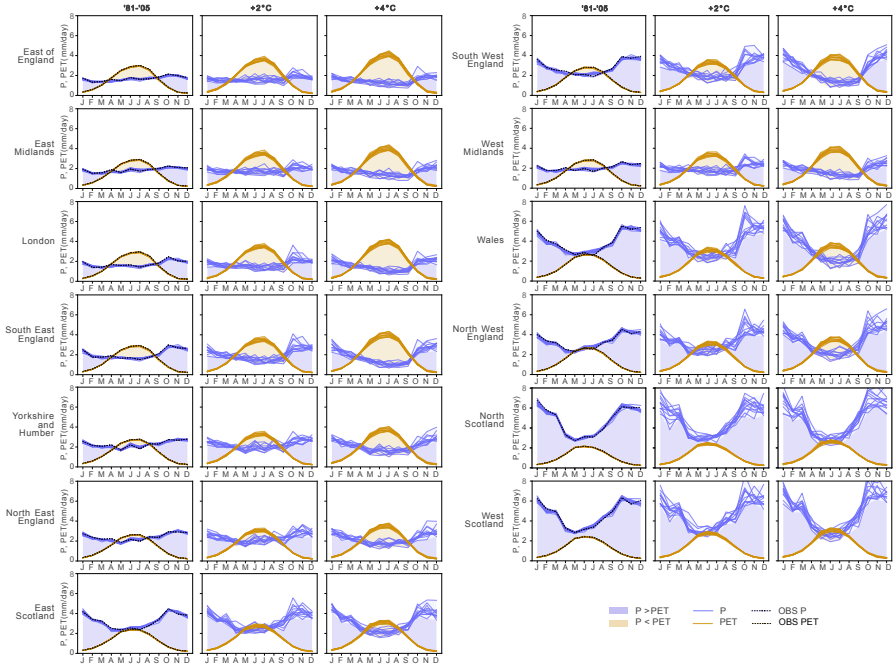

Figure B.0.3: Seasonal cycle of precipitation (P; blue lines) and potential evapotranspiration (PET; orange lines) for the 12 bias-adjusted UKCP18-RCM ensemble members, for all UKCP18 administrative regions. The different lines represent different ensemble members. Observations are shown in darker, dashed lines.

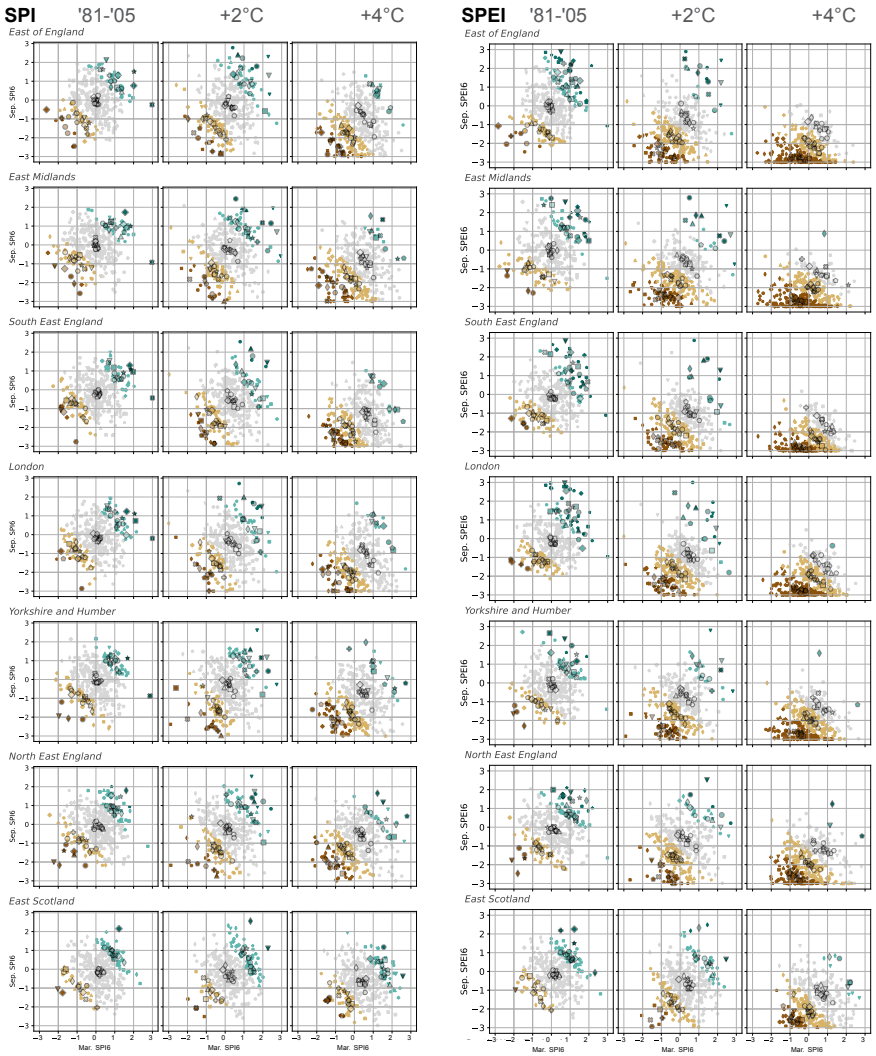

Figure B.0.4: As Figure 4.5.5 but for all GB regions (continued in Fig. B.0.5).

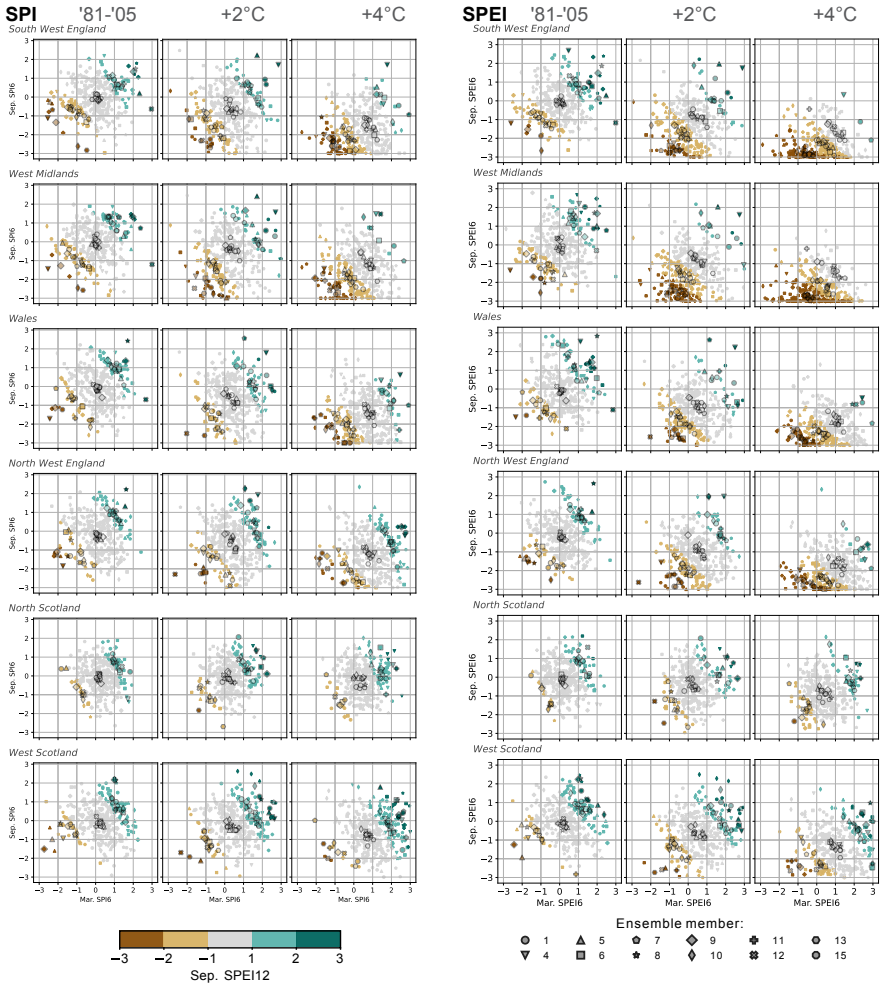

Figure B.0.5: As Figure 4.5.5 but for all GB regions (continuation of Fig. S4).

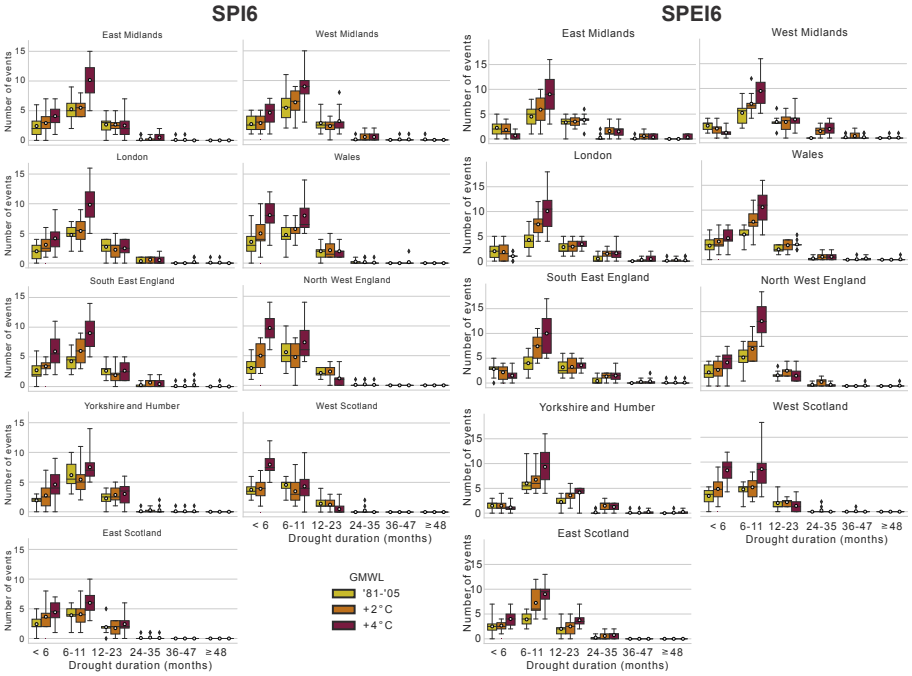

Figure B.0.6: As Figure 4.5.6 but for the other regions.

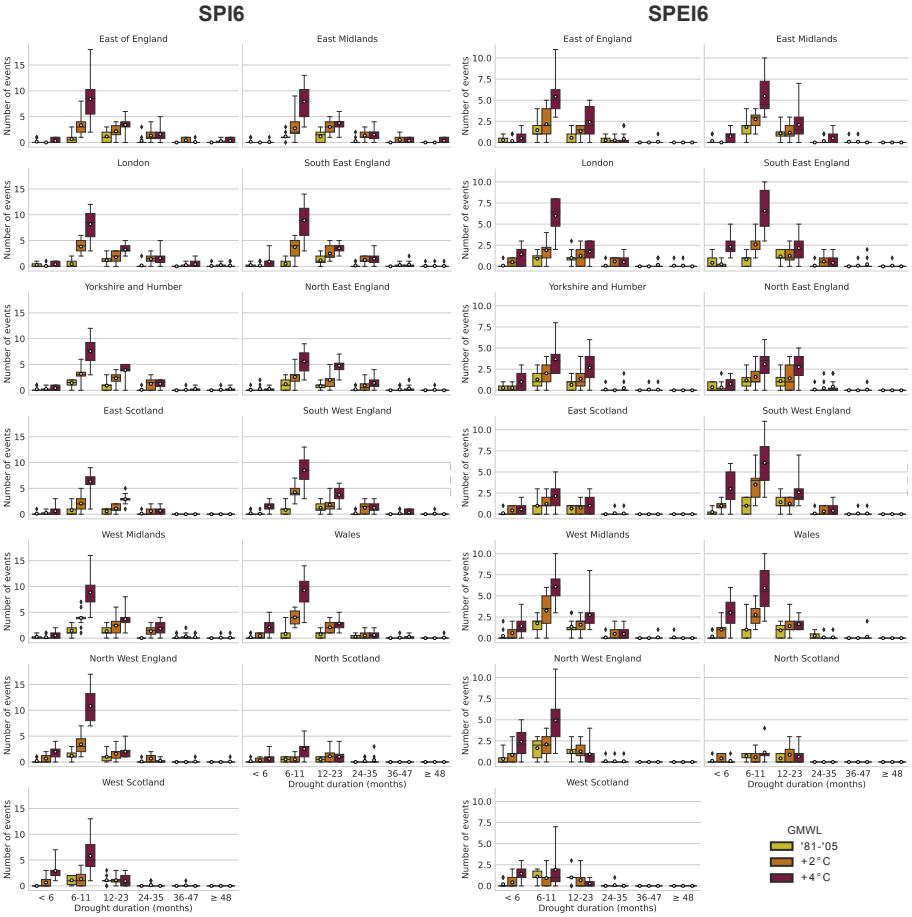

Figure B.0.7: As Figure 4.5.6 but for all regions and isolating drought events that reach extreme levels at some point.

C

Chapter 7: Conclusions and future research

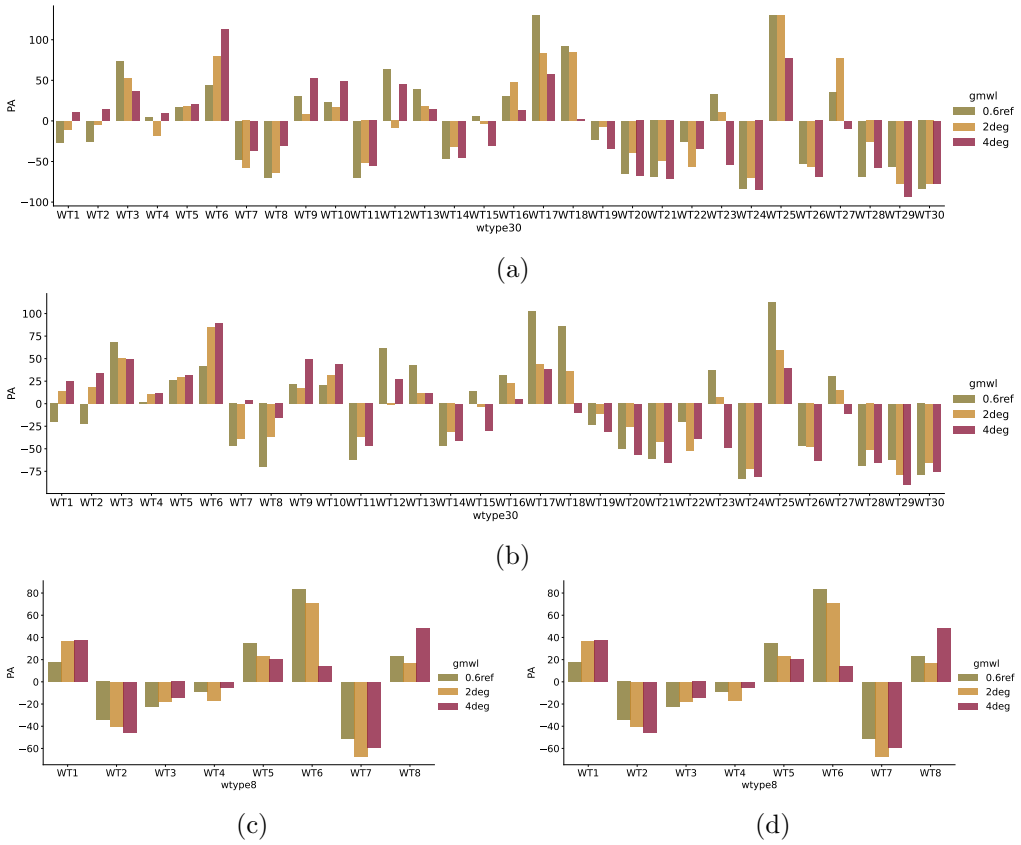

Figure C.0.1: The relative occurrence (percent anomaly; PA) of weather patterns during droughts in the East of England as defined by (a), (c) SPI1; (b), (d) SPEI1, for UKCP18-RCM ensemble member 1. (a),(b): all 30 weather patterns; (c),(d): 8 summary weather patterns As in Chapter 4, the drought indicators are computed based on bias adjusted data, whereas the weather patterns are derived from raw simulations of atmospheric pressure patterns.