The predictability of UK drought using European weather patterns

Doug Richardson

School of Engineering

Newcastle University

A thesis presented for the degree of Doctor of Philosophy

February 2019

For Lem, Malcolm and John.

Abstract

This thesis explores the use of a 167-year daily weather pattern (WP) classification (MO-30) in UK meteorological drought prediction. As MO-30 was recently introduced, necessary analyses as a precursor to building a forecast model are conducted. First, an exploratory analysis of MO-30's fundamental characteristics and its relation to UK precipitation and drought climatology is carried out. Second, two novel methods to find weekly to seasonal persistence in MO-30 are used in order to assess if there is any inherent predictability within MO-30. Third, a statistical model based on historical analogues for predicting 30-day periods of WPs is constructed, from which precipitation forecasts are derived. Finally, a dynamical ensemble prediction system is applied to forecast WPs, with resultant precipitation estimated in the same way as for the statistical method.

MO-30 is shown to be suitable for precipitation-based analyses in the UK. Furthermore, intra-WP precipitation variability, defined by the interquartile range, is lower in MO-30 compared to another commonly used WP classification. Six WPs are associated with nationwide drought, with several other WPs linked to regional drought. Results from the persistence analysis show that there are multi-month periods when small sets of four to six WPs dominate, and some of these periods coincide with notable meteorological events, including droughts and storms. Some WPs also behave as 'attractors', showing increased probability of reoccurrence despite other WPs occurring in-between.

The statistical method for WP and precipitation forecasts is no more skilful than climatology, suggesting that the model did not adequately exploit the persistence identified previously. However, WPs are shown to be potentially useful for drought forecasting, as an idealised, perfect prognostic model (with WP observations as inputs rather than predictions) substantially improves skill, with a skill score of almost 0.5 (out of one) for north-eastern regions. Using a dynamical model to predict WPs, while keeping the precipitation estimation procedure the same as for the purely statistical method, yields overall higher skill compared to a benchmark statistical method for predicting droughts. The model also outperforms direct (modelled) dynamical precipitation forecasts for lead-times greater than 16 days during winter and autumn, with the greatest skill advantage for western regions. This is despite the relatively modest skill scores of all forecast models (rarely above 0.4). Again, high skill scores, of almost 0.8 on occasions, are achieved by the perfect prognostic model, demonstrating the potential for incorporating WPs into precipitation and drought forecast systems.

Acknowledgements

First I would like to thank my two supervisors, Hayley Fowler and Chris Kilsby, for their outstanding guidance and motivation over the past four-and-a-bit years. Aside from contributions to this thesis, their encouragement in everything from attending conferences to writing papers has helped me develop into the researcher I am today. I hope that we will get the chance to work together again in the future. I am also very grateful to Rob Neal and Rutger Dankers, whose excellent supervision during a placement at the Met Office provided some much-needed stimuli in the midst of the 'second-year blues'. Thanks must also go to those colleagues and friends in the School of Engineering, without whom this PhD experience would not have been so incredibly fun (most of the time).

Beyond the walls of the Cassie Building and the realms of scientific research, I owe a huge debt of gratitude to my circles of friends and family who have had to pretend, to varying degrees, to be interested in what I've been doing for four years. It doesn't matter what I say, to these people I will always be a 'weatherman'. In particular, thanks to my mum, whose unbounded enthusiasm means I am a small celebrity in the Peak Paddlers kayak club; to my dad, for the countless pieces of advice and Saturday trips to watch the Boro ("It's the hope that kills, kid"); and to my amazing brother, for those single-word conversations, long periods of silence and obscure Red Dwarf references.

Finally, a special mention to my two housemates, Alice and Rhi, who have provided me with all the support a PhD student could ever wish for. Thank you both, you wonderful girls, not least for putting up with my uninspiring rota of curries. Alice – I hope this thesis finally provides you with some evidence that yes, I did actually do some work.

Contents

List of publications

The following journal articles are products of this thesis and are referenced at the beginning of the relevant chapters. I took the lead role in all articles and was responsible for code development, figure drawing and writing. The co-authors contributed to research ideas and manuscript editing.

Richardson, D., Fowler, H. J., Kilsby, C. G. and Neal, R. (2018a), A new precipitation and drought climatology based on weather patterns. Int. J. Climatol, 38: 630–648. doi:10.1002/joc.5199

Richardson, D., Kilsby, C. G., Fowler, H. J. and Bárdossy, A. (2018b), Weekly to multi-month persistence in sets of daily weather patterns over Europe and the North Atlantic Ocean. Int. J. Climatol. doi:10.1002/joc.5932

List of figures

List of tables

Common abbreviations

BS	Brier Score
BSS	Brier Skill Score
CEE	Central and east England
DSI	Drought Severity Index
EC	Empirical counting method
ECDF	Empirical cumulative distribution function
ECMWF	European Centre for Medium-Range Weather Forecasts
EMULATE	European and North Atlantic daily to multi-decadal climate variability data set
ENSO	El Niño Southern Oscillation
EPS-P	Precipitation forecast model using ECMWF-EPS
EPS-WP	Precipitation forecast model using ECMWF-EPS WP
ES	East Scotland
ESP	Ensemble Streamflow Prediction
FM	Frequency method to predict WPs
GloSea5	Met Office global seasonal forecast system version 5
GpH	Geopotential height
GWL	Grosswetterlagen
HadUKP	Met Office Hadley centre UK precipitation data set
hPa	Hectopascal
IQR	Interquartile range
IS	Identical set
JSD	Jensen-Shannon divergence
KLD	Kullback-Leibler divergence
LWT	Lamb Weather Type
MC	Markov method for WP prediction
MM	Markov method for persistence
MO-8	Met Office classification of eight weather patterns
MO-30	Met Office classification of 30 weather patterns
MSE	Mean square error
MSLP	Mean sea-level pressure

NAO	North Atlantic Oscillation
NEE	Northeast England
NI	Northern Ireland
NS	North and northwest Scotland
NWE	Northwest England and north Wales
OD	One-different set
PA	Percentage anomaly
PC	Precipitation climatology
PDSI	Palmer Drought Severity Index
Perfect-WP	Precipitation forecast model derived from WP observations
PET	Potential evapotranspiration
RCM	Regional climate model
ROC	Relative operating characteristic
RPS	Ranked Probability Score
RPSS	Ranked Probability Skill Score
S2S	Sub-seasonal to seasonal
SEE	Southeast England
SLP	Sea-level pressure
SM	Similarity method
SPI	Standardised Precipitation Index
SPEI	Standardised Precipitation Evapotranspiration Index
SS	South and southwest Scotland
SST	Sea-surface temperature
SWE	Southwest England and south Wales
WP	Weather pattern

Chapter 1
Introduction

1.1 UK drought

The impacts of drought can be severe. It is thought to be the costliest form of extreme weather, and is responsible for huge numbers of deaths worldwide (World Meteorological Organization (WMO) and Global Water Partnership (GWP), 2017). In the UK, droughts are a recurrent feature of climate and events such as those in 1975-76, 1995 and 2010-12 had severe implications for many sectors, including agriculture, water resources and the economy, as well as for ecosystems and natural habitats (Marsh, 1995; Marsh *et al.*, 2007; Rodda and Marsh, 2011; Kendon *et al.*, 2013). The Met Office declared the summer of 2018 as the joint hottest on record along with 1976, 2003 and 2006 (Met Office, 2018), and the period May through July was extremely dry, rivalling that in 1976, which is typically regarded as the benchmark severe event. The 2018 drought had noticeable effects on reservoir and river levels (The Guardian, 2018c) and agriculture (The Guardian, 2018a). A remarkable effect of this drought was to dry the soil sufficiently to reveal the cropmarks of foundations of ancient settlements (The New York Times, 2018). Hydrological impacts were particularly severe in north-western regions. The Republic of Ireland and Northern Ireland saw restrictions in public water usage enforced (BBC, 2018a), while United Utilities, which serves North West England, announced a date for which a temporary use ban would come into effect (although this was later suspended after some timely precipitation) (United Utilities, 2018). By comparison, a very wet spring meant that groundwater aquifers remained well-stocked throughout the summer (Hannaford *et al.*, 2018), and so water resource systems were relatively unaffected in the southeast. This is one of the key differences between the summers of 1976 and 2018 – the former followed a dry winter, which is the season crucial for recharging groundwater aquifers (Rodda and Marsh, 2011).

In future, the effects of climate change are likely to result in more frequent, more severe and longer duration droughts for parts of the UK (Rahiz and New, 2013; Samaniego *et al.*, 2018; Spinoni *et al.*, 2018) and it is likely that further restrictions on water abstraction will be required (Water UK, 2016; Environment Agency, 2018). Furthermore, changing socio-economic factors, such as an increasing population, will exacerbate the physical impacts of drought and place higher demands on water resource systems. This is particularly relevant to large urban centres such as London, in which residents face the prospect of queuing for water during extremely hot summers (Water UK, 2016).

To mitigate the effects of drought, it is crucial that relevant sectors plan ahead, and drought forecasts have an important role in designing these strategies. UK droughts typically last between a few months and several years (Marsh *et al.*, 2007) and it is on these time-scales that forecasters should aim to provide predictions of drought. Different sectors require tailored forecasts, whether that be in the predicted variable (such as streamflow and reservoir levels for urban water suppliers), the forecast lead-time (for agricultural end-users this might be the crop-growing season) or spatial and temporal resolution. Even before considering methodological choices, this means there is a wide range of different requirements for drought prediction.

1.2 Drought forecasting

Drought forecasting, and hydro-meteorological prediction in general, may be classified into two broad categories, dynamical and statistical. Dynamical forecast models are physically-based, and use governing equations representing earth-system components, such as the atmosphere, oceans and land-surface, to propagate forecasts forward from a set of initial conditions (Smith *et al.*, 2012). Given the incredibly complex nature of these processes and the volume of data they require, dynamical models have only been feasible since the advent of high-performance computing. Even so, the technology is still sufficiently limited so as to make representing all known processes not yet feasible, meaning many must be parameterised rather than explicitly resolved (Smith *et al.*, 2012). These computational demands require dynamical models to be developed, run and hosted by institutions with the requisite facilities and financial resources, which are unsurprisingly few in number. In the UK, there are two main institutions, the Met Office and the European Centre for Medium-range Weather Forecasts (ECMWF). A consequence of this is that dynamical forecast end-users are reliant on these organisations making their data available, often at financial cost.

Statistical models, including methods often described as data-driven (such as machine learning), are based on historical data (observations). In prediction, they typically will relate a target variable (the predictand) with one or more temporally-lagged explanatory variables (predictors). For drought forecasting, the predictand might be precipitation, streamflow or a drought index, while the predictors may be exogenous (for example, atmospheric circulation patterns or sea-surface temperature) or be derived from antecedent conditions (such as using prior streamflow observations to predict future values). Statistical prediction models are typically conceptually simpler and less computationally demanding than their dynamical counterparts, making them more accessible to those without the resources to use dynamical models. Furthermore, in hydro-meteorology, they often perform adequately. For example, streamflow can be well-predicted using statistical models in the UK (Svensson, 2014) and

Australia (Wang *et al.*, 2009). Another area in which statistical models are useful is in benchmarking dynamical model performance. If the former is almost as skilful (or even equally, or more, skilful) as the latter, then an end-user may choose to use a statistical model in light of their relative conceptual simplicity. Even so, the majority of works in hydro-meteorological forecasting are focussed on dynamical models. Indeed, this thesis is part of a project named IMPETUS (Improving Predictions of Drought for User-Decision Making) (Shaffrey, 2014) and is the only component to be explicitly tasked with researching statistical models.

1.3 Weather pattern classifications

A typical drought forecast model will attempt to describe the relationships between a number of hydro-meteorological variables. Such data sets often exhibit high variability (noise) and so these relationships, if present, tend to be of insufficient strength to be useful in forecasting. There are some well-known exceptions. For example, the North Atlantic Oscillation is a key atmospheric teleconnection that influences UK weather and climate, although the strength of this relationship is highly dependent on the region and season (Hurrell, 1995; Wilby *et al.*, 1997; Mehta *et al.*, 2000; Qian *et al.*, 2000; Wilby, 2001; Fowler and Kilsby, 2002a; Hall and Hanna, 2018). In fact, Hall and Hanna (2018) argue that there is limited value in focussing solely on predicting this teleconnection in winter in order to forecast local-scale variables, as there are other modes of atmospheric circulation that are just as important, for example the East Atlantic pattern.

One way to avoid testing a plethora of potential predictors is to consider weather pattern (WP) classifications. A classification consists of a small number of individual WPs defined by some pressure variable, which together represent the key patterns of atmospheric circulation over a given region (Chapter 2 will describe WP classifications in detail). A noisy (continuous) time series of, say, daily sea level pressure (SLP) can be converted into a discrete time series of WPs where each day can only be one of a few states rather than one of (in principle) infinite values. This has made WP classifications an attractive prospect to many researchers over the last few decades, in fields such as historical climatology, climate change projections and, more recently, weather forecasting and dynamical forecast model verification (again, Chapter 2 will cover this in more detail, along with relevant references). For drought prediction, the use of WPs represents an opportunity. Medium- to long-range forecast skill of precipitation is low outside of the tropics (Kirtman and Pirani, 2009; Smith *et al.*, 2012; Saha *et al.*, 2014), yet forecast skill of atmospheric variables, and hence potentially WPs, can be much higher (Scaife *et al.*, 2014; Vitart, 2014; Lavers *et al.*, 2016a; Baker *et al.*, 2018; Ferranti *et al.*, 2018). Therefore, it is

possible that precipitation forecast skill could be increased by predicting WPs, and deriving predicted precipitation from these discrete time series.

The published literature on using WPs for drought forecasting is small. Only one study adopts a statistical approach to WP prediction (Fayos and Fayos, 2007) and none subsequently focus on drought. This is therefore the research gap that this thesis will attempt to address: what potential do WPs have in approaching the challenge of UK drought prediction? Can a statistical model produce skilful WP and precipitation forecasts, and how does this skill compare to that of a dynamical model? Only meteorological drought shall be considered, as precipitation deficits are also the primary cause of other drought types, and the relationship between air pressure (such as SLP) and precipitation is more direct and immediate than between air pressure and hydrological variables (Lavers *et al.*, 2010). Furthermore, this thesis shall focus on a specific WP classification introduced by the Met Office in 2016 (Neal *et al.*, 2016). This classification, called MO-30, offers, in principle, a significant upgrade on the most commonly used classification in UK WP analyses, the Lamb Weather Types (LWTs) (Lamb, 1972), due to a more objective definition process and a larger domain that captures large-scale atmospheric systems over the North Atlantic Ocean. However, outside of those in this thesis, there are no published studies on the relationship between MO-30 and UK precipitation. Indeed the only climatological study using it focusses on sea-surge and wave heights (Neal *et al.*, 2018). Therefore, research into this classification's hypothesised advantages compared to LWTs and its applicability to UK precipitation and meteorological drought analyses, including forecasting, is warranted.

1.4 Research aims and objectives

The overall aim of this thesis is to evaluate the potential of using WPs in statistical UK drought prediction. For WPs to have utility in this context, three key criteria must be satisfied:

1. That complex atmospheric variability can be satisfactorily encapsulated in a simple classification of WPs.
2. That the WPs are predictable.
3. That the WPs can be used to estimate precipitation and meteorological drought.

These criteria and the overall aim shall be addressed by conducting analyses assessing the suitability of MO-30 for drought studies; the potential of MO-30 for forecasting by exploring the persistence of the classification; and by developing and testing statistical drought prediction models against benchmark and dynamical methods.

This thesis will address the following research objectives:

- Conduct a thorough and comprehensive review of the literature surrounding drought, including its definition, its past and future occurrence in the UK and its relationship with potential predictors; statistical drought forecasting studies; dynamical models for drought prediction and WP classifications.

- Perform an exploratory analysis of MO-30 in terms of WP frequencies of occurrence, seasonality and comparisons with LWTs.

- Investigate the relationship between MO-30 (and a smaller set of clustered WPs) with UK daily precipitation climatology, again with reference to the LWTs.

- Analyse how MO-30 WP behaviour changes during UK droughts.

- Explore the persistence of MO-30 on weekly-to-seasonal time-scales and assess whether this persistence would be useful in a WP forecast model.

- Construct a statistical WP and drought forecast model based on the persistence previously identified, and compare against simpler methods.

- Develop a dynamical WP and drought forecast model and compare with statistical approaches. The dynamical model will expand upon the framework of existing Met Office operational forecast tools by using a more sophisticated statistical methodology for deriving precipitation estimates from the forecast WPs.

1.5 Thesis structure

A literature review of the topics described above will be presented in Chapter 2. In Chapter 3, MO-30 shall be formally introduced and its behaviour and relationship with UK precipitation and drought analysed, with comparisons made to the LWTs. The work in this chapter has been published in Richardson *et al.* (2018a). Chapter 4 will explore and quantify WP persistence on weekly-to-seasonal time-scales, and link this to possible physical mechanisms. The material in Chapter 4 has been published in Richardson *et al.* (2018b). A statistical forecast model will be developed in Chapter 5, where monthly WP occurrences shall be predicted based on historical analogues, and corresponding precipitation forecasts derived via a sampling procedure. In Chapter 6, output from the ECMWF ensemble prediction system will be used to create a WP reforecast data set, which shall then be utilised to assess the predictability of WPs, with precipitation forecasts estimated in the same way as in Chapter 5. This will enable a comparison between the developed statistical and dynamical approaches used in practice by the Met Office. Finally, conclusions shall be presented in Chapter 7, with a review of the results in this thesis,

their relevance for the wider field and some thoughts on future work and the operational potential of the described drought forecast models.

Chapter 2
Literature Review

2.1 Definition of drought

2.1.1 Drought classifications and characteristics

Drought is a worldwide phenomenon occurring in both high and low precipitation regions. It is a temporary period of dry weather relative to the local climatological norm and should not be confused with aridity, which is a permanent feature of climate only present in regions of low precipitation (Wilhite, 2009; Mishra and Singh, 2010). Despite this, arid areas are more prone to drought as they depend critically on a limited number of precipitation events (Sun *et al.*, 2006). A universal definition of drought proves elusive, with over 100 definitions in existence (Mishra *et al.*, 2015b). There are two main types of definition – conceptual and operational. Conceptual definitions are relative descriptions (e.g. moderate/extreme drought) used to aid the understanding of drought for those at risk of its effects. These effects vary across different social and economic sectors, and this has led to a plethora of definitions tailored to the situation at hand. Operational definitions are more specific and are used to measure the frequency, severity and duration of droughts for some return period (Mishra and Singh, 2010). The difficulty in defining drought lies in the time-scale of precipitation deficit accumulation, and in how these deficits result in reductions in different water sources (e.g. soil moisture, streamflow and reservoir storage) (McKee *et al.*, 1993). The lack of a single definition suitable for all purposes was originally thought of as a problem holding back drought research (Yevjevich, 1967), yet it is now well recognised that multiple definitions are necessary (e.g. Wilhite and Glantz, 1985; McKee *et al.*, 1993; McKee *et al.*, 1995; Panu and Sharma, 2002; Dai, 2011; Lloyd-Hughes, 2014). The American Meteorological Society (AMS; 2014) offers perhaps the most overarching definition:

> "[Drought is a] period of abnormally dry weather sufficiently long enough to cause a serious hydrological imbalance.
>
> Drought is a relative term, therefore any discussion in terms of precipitation deficit must refer to the particular precipitation-related activity that is under discussion."

The concept of drought as a term relative to the context in which it is being discussed is key, and in the literature there have arisen four distinct classifications of drought proposed by Wilhite and Glantz (1985) and adopted by the AMS (2013). The four classes are meteorological, agricultural, hydrological and socio-economic:

i. Meteorological or climatological drought is characterised primarily by a prolonged deficit in precipitation. Such an event might range from several months to years in duration and can develop and end suddenly (Heim, 2002). The propensity to form over short time-scales means that meteorological drought precedes and causes other types of drought (Dai, 2011).

ii. Agricultural drought is largely caused by dryness in the surface layers (root zone) of soil, affecting crop yields and plant growth. This results from below-average precipitation, intense but infrequent precipitation, low humidity or above-average evaporation (Heim, 2002; Dai, 2011). Depending on the antecedent moisture content of the soil surface layer, agricultural drought may lag meteorological drought (Heim, 2002). Agricultural drought has been studied far less than the meteorological or hydrological types, despite it possibly being the most important (Palmer, 1965). Panu and Sharma (2002) suggest that this is due to difficulties in measuring soil moisture content and the lack of sufficiently long historical records available. More pertinently, the study of agricultural drought is complex, and might be best analysed through a combination of variables (Panu and Sharma, 2002), leading into what Palmer (1965) called "the realms of soil physics, plant physiology, and agricultural economics." For example, different crops behave differently in drought situations, even in the same region, necessitating the consideration of a wide range of factors critical for crop water needs (see Food and Agriculture Organization of the United Nations, 1986; Mishra *et al.*, 2015a).

iii. Hydrological drought concerns the depletion of water sources such as reservoirs, groundwater and lakes to below-average levels that are not replenished due to a long-term deficit of precipitation (Heim, 2002; Dai, 2011). As there is a time lag between periods of below-average precipitation and the associated consequences for the hydrological system, hydrological drought initiation and termination generally lags that of meteorological and agricultural droughts.

iv. Socio-economic drought is conceptually different from the other three drought types as it is related to the failure of water resource systems to meet the demands of end-users. For example, economic impacts might be concerned with lost income from a reduced crop yield as a result of agricultural drought; a social impact could be the inability to supply enough water to urban communities, possibly leading to health problems (AMS, 2013).

What is notable from the different drought classifications is that they are functionally separated by their time-scales (McKee *et al.*, 1993). A short-term drought may have severe impacts for agriculture but only limited effects on the hydrologic community. Conversely, a season of normal precipitation that is preceded by seasons of below average precipitation could constitute a hydrological, but not an agricultural drought. This may be due to depleted water storage levels that are not fully replenished, whilst agriculture may be unaffected if the crop depends primarily on the current season's precipitation (Edwards and McKee, 1997).

It is also important to delineate certain characteristics of drought, namely duration, intensity, severity and spatial extent (Mishra and Singh, 2010). Unlike with other natural hazards, the initiation and termination points (and hence duration) of a drought are hard to define (Parry *et al.*, 2016); there is no sudden incident that marks the start of drought as there is with, say, earthquakes or flash floods. Does a drought begin with the first day in a sequence of no precipitation, or only when the effects are felt in water supply systems? Similarly, is the end to a drought marked by a return to normal precipitation amounts, or do reservoir storage levels have to be replenished before a drought can be considered over? Clearly this is related to the time-scales of the drought classifications described previously. Simpler to gauge is a drought's intensity, which is a measure of the moisture deficit, and severity (or magnitude) which is a function of the drought intensity and duration (AMS, 2013). An early example of characterising drought comes from Yevjevich (1967), who used the statistical theory of runs to describe droughts in terms of run-length (duration) and run-sum (cumulative deficit volume). Generally droughts are now characterised using a drought index; the time series of which can be used as a basis for evaluating parameters of interest (Mishra and Singh, 2010).

2.1.2 Drought indices

A drought index is a quantitative measure of the departure of some primary (moisture) variable from the average climatology computed from historical records (Dai, 2011). The primary variable depends on the type of drought. Precipitation is the determinant variable for meteorological drought, with secondary variables such as evapotranspiration also incorporated into some indices. River runoff/streamflow or reservoir/groundwater levels are the main variables for hydrological drought and soil moisture is often the primary variable for agricultural drought (Panu and Sharma, 2002). Although drought indices have been in existence at least since Munger's Index (Munger, 1916) in the early 1900s, the two most widely used indices today were developed in the latter half of the 20th Century, namely the Palmer Drought Severity Index (PDSI; Palmer, 1965) and the Standardised Precipitation Index (SPI; McKee *et al.*, 1993; McKee *et al.*, 1995).

The development of the PDSI was a milestone for drought study. The index is essentially a sum of the current moisture anomaly and a fraction of the previous index value (Lloyd-Hughes and Saunders, 2002), incorporating precipitation, evapotranspiration, runoff and soil moisture as variables. Negative (positive) PDSI values represent the degree of dryness (wetness), with more negative (positive) values indicating greater severity. Calculation of the PDSI is explained in various studies (e.g. Karl, 1983; Alley, 1984; Karl, 1986). Karl *et al.* (1985) suggested that in real-time the PDSI is more of a hydrological than meteorological index as it pertains to moisture inflow, outflow and storage; they named this the Palmer Hydrological Drought Index (PHDI). The PDSI has several limitations (e.g. Heim, 2002; Dai, 2011, and references therein). One problem is that the index is defined by a set of empirical relationships that were computed using observations from just nine US climate stations. Consequently, the standardisation procedure (used to ensure drought characteristics can be compared at different times and locations) may not result in the degree of spatiotemporal comparability that Palmer intended (Alley, 1984; Karl, 1986). Work by Wells *et al.* (2004) negated this issue with the development of a self-calibrating PDSI (sc-PDSI), which replaced the empirical constants in the PDSI with dynamically computed values based on the local climate. However, the principal shortcoming of the PDSI has not been resolved, namely its inherent time-scale of nine to twelve months (McKee *et al.*, 1993; McKee *et al.*, 1995; Guttman, 1998; Lloyd-Hughes and Saunders, 2002; Vicente-Serrano *et al.*, 2010; Jiang *et al.*, 2015). This means the PDSI (and sc-PDSI) should not be expected to provide useful information over all time-scales. Additionally, Guttman (1998) noted that the index has an autoregressive characteristic which sees PDSI values retain memory of up to four years.

A drought index intended for universal purpose should be applicable over a range of time-scales and locations. The SPI (McKee *et al.*, 1993) quantifies precipitation deficits on multiple time-scales. Its calculation involves fitting a probability density function to the precipitation data and transforming it to the standardised normal distribution. This gives the SPI the same benefits of standardisation as the PDSI. However, Guttman (1998) examined the spectral characteristics of SPI and PDSI time series, finding that only the former is spatially invariant, making the SPI more suitable when comparing drought in different regions. SPI values are interpreted the same way as PDSI values, with more negative (positive) values signifying a greater severity of dryness (wetness). The variable time-scale of the SPI is advantageous in that it allows the index to be used for different drought types, which are functionally separated by the time-scales over which moisture deficits accumulate (McKee *et al.*, 1993). Consequently, the SPI can be computed over short time-scales for meteorological and agricultural droughts and longer time-

scales for hydrological drought. Another advantage of the SPI is its simplicity. It is based solely on precipitation and requires the computation of just two parameters in comparison to the 68 needed in PDSI calculation (Lloyd-Hughes and Saunders, 2002). However, there are several criticisms of the SPI, including the assumption that a suitable theoretical distribution can be found to model the raw precipitation prior to standardisation (Lloyd-Hughes and Saunders, 2002). Several distributions have been suggested, for example the Gamma distribution (McKee *et al.*, 1993; Edwards and McKee, 1997), the Pearson Type-III distribution (Guttman, 1999) and the Weibull distribution (Sienz *et al.*, 2012). Linked to this, the SPI suffers from problems in fitting and interpretation in dry climates. The high frequency of zeros characterising datasets in such climates means the precipitation probability distribution is highly skewed. Then for short time-scales the resulting SPI distribution is non-normal and unable to indicate drought occurrence (Wu *et al.*, 2007). Furthermore, the length of record to which a distribution is fitted can have a large impact on the uncertainty of resulting SPI estimates; Carbone *et al.* (2018) found for several stations in Italy that parameters estimated on 60-year time series resulted in considerably less uncertain SPI series than the often-used 30-year records. By using a variable other than precipitation, the method used to calculate SPI can be applied to hydrological drought. For streamflow this is often known as the Standardised Streamflow Index (SSI; Vicente-Serrano *et al.*, 2012) and for groundwater this is called the Standardised Groundwater Index (Bloomfield and Marchant, 2013).

Although the simplicity of precipitation-based indices is an advantage in calculation, it may be regarded as a limitation in other respects. They do not consider other variables that can influence drought, such as temperature, wind speed and evaporation. Vicente-Serrano *et al.* (2010) argued that in regions where the variability of precipitation is lower than that of other variables, or regions where the temporal trends of other variables are not stationary, an exclusively precipitation-based index is not adequate. Furthermore, Manning *et al.* (2018) showed that, although precipitation dominates potential evapotranspiration (PET) in European soil moisture drought, the consideration of the latter improves the estimation of drought onset, persistence and severity at wet sites. Two indices similar in design to the SPI but with an additional component of PET have therefore been developed: the Standardised Precipitation Evapotranspiration Index (SPEI; Vicente-Serrano *et al.*, 2010) and the Reconnaissance Drought Index (RDI; Tsakiris *et al.*, 2007). Calculation of indices with a PET variable is more complex than for the SPI due to the need to compute the PET component, which includes a number of parameters. Some studies (Mavromatis, 2007; Jiang *et al.*, 2015) argue this complexity may be avoided by adopting the simple Thornthwaite method for PET calculation (Thornthwaite,

1948), as there is little difference in PDSI values calculated using this method compared to the more sophisticated Penman-Monteith equations (Monteith, 1965). However, work by Sheffield *et al.* (2012) shows that significantly different results are obtained when using each PET method in evaluating changes in global drought frequency and severity using the PDSI.

On the other hand, some studies recommend a simple index over a more complicated one, due to either precipitation being considered the primary variable controlling drought or the difficulties associated with obtaining evapotranspiration data (Oladipo, 1985; Lloyd-Hughes and Saunders, 2002; Blenkinsop and Fowler, 2007a). There are many more indices which are not discussed here, ranging from the theoretically simple Drought Severity Index (DSI; Phillips and McGregor, 1998), which is based on cumulative precipitation deficits; to more exotic reliability-resilience-vulnerability concepts (Maity *et al.*, 2013), which provide a tool to assess the frequency of, vulnerability to, and ability to recover from drought. A useful resource for drought indices and their application, including strengths and weaknesses, is provided by the World Meteorological Organization (WMO) and Global Water Partnership (GWP) (2016). The wide range of available drought indices allows the user to select one appropriately on a case-by-case basis. For analyses in humid regions where precipitation variability is greater than temperature variability, a precipitation-only drought index could be considered adequate. For regions with a nonstationary temperature trend, or for global warming scenarios, an index incorporating PET may be more suitable (Vicente-Serrano *et al.*, 2010).

2.2 Drought impacts and future projections

2.2.1 Drought effects, planning and policy

Drought affects more people than any other natural hazard (Mishra and Desai, 2005) and is among the most costly of natural disasters (World Meteorological Organization (WMO) and Global Water Partnership (GWP), 2017). It can have severe implications for natural habitats, ecosystems, and many social and economic sectors, such as agriculture and urban water supply (Heim, 2002). A growing population is placing increasingly higher demands on water availability and there are many places that experience water scarcity almost every year (Mishra and Singh, 2010). Climate change has exacerbated this problem. Dai (2011) showed that many regions have seen average temperature increases of 1-3°C and a decrease in precipitation in the period 1950-2008, and global drought recovery times have increased over the twentieth century (Schwalm *et al.*, 2017). Projections indicate that present-day aridity may become severe drought in a substantial part of the world later this century (Burke *et al.*, 2006; Dai, 2011; Cook *et al.*, 2018). This may be especially problematic for developing nations with limited resources

for disaster mitigation. In these less developed regions, the AMS (2013) lists the main impacts of drought as crop failures, clean drinking water shortages, famine, energy shortages, mass migration and political unrest. For developed countries the effects are different, mostly linked to socio-economic problems such as rising food costs or water restrictions.

In the UK, the limits of water restrictions that water companies must observe is set out in government legislation, namely the Water Industry Act 1991[1], as amended by the Environment Act 1995[2] and the Water Act 2014[3] (this piece of legislation will be collectively referred to as WIA 1991 from herein). The WIA 1991 details three measures which water companies may take in times of drought. The first (least severe) measure is a Temporary Use Ban (TUB; formerly a 'hosepipe ban') detailed in The Water Use (Temporary Bans) Order 2010, which can be implemented by water companies. A TUB restricts consumers' use of water for nonessential tasks like watering gardens or washing vehicles. Second is an Ordinary Drought Order (ODO), which allows water companies to impose stricter controls on water supply. Lastly, there is an Emergency Drought Order (EDO), which enables water companies to restrict water supplies further by introducing rota cuts and/or standpipes. While a TUB has been enforced as recently as summer 2018 (BBC, 2018a), there has not been an EDO put in place since 1976, and the government has made it clear that such measures should be avoided at all costs (Gavin *et al.*, 2013). Unlike TUBs, an EDO or ODO must be authorised by the Secretary of State.

Drought is a recurring phenomenon of UK climate with varying impacts (Marsh *et al.*, 2007). The summer 2018 heatwave was severe enough for water companies in Ireland to restrict usage (BBC, 2018a), while United Utilities (covering North West England) were only able to call off planned restrictions due to some timely rains (United Utilities, 2018). While the impact on the general public's (outside the Republic of Ireland and Northern Ireland) livelihoods was not significant, there were newsworthy incidents such as moorland fires (The Guardian, 2018b), mass fish die-offs (BBC, 2018b), agricultural stresses (The Guardian, 2018a) as well as the unusual effect of dried-out soils revealing the foundations of ancient settlements (The New York Times, 2018). Kendon *et al.* (2013) reviewed the 2010-12 drought, which had significant detrimental impacts for central, southern and eastern regions of England. Farmers struggled to grow crops, wildfires were a risk due to dry soils, while reduced river flow and dried-out

[1] Part III, chapter III of the Water Industry Act 1991 (https://www.legislation.gov.uk/ukpga/1991/56/contents)
[2] Paragraphs 139, 140 and 141 of Schedule 22 to the Environment Act 1995 (https://www.legislation.gov.uk/ukpga/1995/25/schedule/22)
[3] Sections 57-60 of the Water Act 2014 (https://www.legislation.gov.uk/ukpga/2014/21/part/2)

wetlands posed a threat to wildlife. The drought in 1975-76 is generally considered the most severe on record (e.g. Jones and Lister, 1998; Marsh *et al.*, 2007; Kendon *et al.*, 2013), with the UK experiencing the lowest 16-month precipitation totals since 1766 (Marsh *et al.*, 2007). In June 1975 there were five drought orders in place over the UK. One year later this number increased to 27 and by September 1976 there were 112 drought orders in force (Rodda and Marsh, 2011). Rodda and Marsh (2011) stated that from July 1976 emergency measures were implemented, including water pressure reductions, delivery of water by road tanker to isolated communities and restriction of spray irrigation. The water supply to consumers remained sufficient except in two Welsh regions where the supply was cut off at night, but by August cut-offs affected over one million people in Southeast Wales. In terms of percent of average precipitation, the most severe droughts after 1975-76 occurred in 1920-21 (62% of the average precipitation), 1943-44 and 1995-97 (both 70%), 1933-34 (72%), 2018 (72%) and 1990-92 (73%) (Kendon *et al.*, 2013; Met Office, 2018). However, this measure is not a reflection of the societal impacts of drought, as it does not take into account water resource levels.

Region	Groundwater (%)[4]
Yorkshire & North East	12.7
North West	7.7
Midlands	17.9
Anglian	38.5
South East	40.5
South West	18.4
Scotland	5.0
Northern Ireland	0.1

Table 2.1: Water supply extracted from groundwater (%) by EA region for England, plus Scotland and Northern Ireland.

Regions in the UK have different primary water sources. As shown in Table 2.1, Anglian and South East England extract a substantial portion of water from groundwater aquifers. Note that these are the rather large Environment Agency (EA) regions which do not represent individual water companies' jurisdiction. For example, Southern Water (located in the South East EA region) claim as much as 70% of their water comes from aquifers[5]; significantly more than the regional value of 40.5% in Table 2.1. On the other hand, northern England, Scotland, Wales and Northern Ireland depend mostly on surface water such as reservoirs and rivers. Droughts can therefore be present in one part of the UK and absent in others. Droughts in south and east England tend

[4] Figures for England EA regions calculated for the 2010-2013 average from https://www.gov.uk/government/statistical-data-sets/env15-water-abstraction-tables
Scotland figures are from 2011 at https://www.bgs.ac.uk/downloads/start.cfm?id=2429
Northern Ireland figures are from 2014 at http://www.niwater.com/siteFiles/resources/pdf/Reports/2014DrinkingWaterQualityAnnualReport.pdf
[5] https://www.southernwater.co.uk/groundwater

only to pose supply problems when flows have been low for at least 15 months, with insufficient precipitation during the critical winter recharge period failing to replenish groundwater levels. Conversely, in northern and western regions intense deficiencies in spring and summer precipitation can cause a hydrological drought on much shorter time-scales of four to nine months (Jones and Lister, 1998; Marsh *et al.*, 2007; Kendon *et al.*, 2013). For example, the Yorkshire region mostly uses single-season upland reservoirs which are recharged during winter and drawn-down the rest of the year, with little carry-over to the next year (Fowler and Kilsby, 2002b).

Serious water supply issues have been limited since 1976, with the exception of the Yorkshire drought in 1995 which required water to be brought in by tanker (Marsh *et al.*, 2007). This resilience in the UK water resources system is demonstrated by the 2003 and 2018 heatwaves. February to October 2003 was the driest run of these months since 1921 (Marsh, 2004), while May through July 2018 was the third driest stretch of these three months since records began in 1910 (CEH, 2018). Both were exceptionally warm with high evaporation and dry soils. However, both heatwaves only saw a modest impact on water supplies for regions predominantly fed by groundwater aquifers, as groundwater levels remained high prior to the onset of dry weather (Marsh *et al.*, 2007; Sefton *et al.*, 2018). Impacts on northern and western catchments were typically more severe. Spraggs *et al.* (2015) also highlighted the resilience in UK water resource systems in a comprehensive study over the period 1798-2010. The authors found that the most severe runoff deficiencies in the record did not necessarily lead to critical reservoir storage levels.

Despite this resilience, projected temperatures suggest the warmth of summers in 1976, 2003 and 2018 will be repeated and exceeded in the 21st Century (Rodda and Marsh, 2011). Therefore, regional forecasts of drought for a range of time-scales will become increasingly valuable to water companies and others concerned with implementing drought management and mitigation plans. The regional aspect is crucial, as UK droughts affect water supplies in different ways and over different time-scales depending on the type of water supply, as detailed above.

2.2.2 Projections of UK drought

There are some difficulties in comparing studies of drought projections for various reasons, such as the use of different future time-periods, data sets, drought metrics and models (Sheffield and Wood, 2008; Knutti *et al.*, 2010; Dai, 2011; Heinrich and Gobiet, 2012; Knutti and Sedláček, 2012; Trenberth *et al.*, 2013; Touma *et al.*, 2015). Nevertheless, this subsection

collates some results from key publications that attempted to quantify projected changes in drought characteristics. For brevity, and in keeping with the region of interest in this thesis, the discussed studies are restricted to those focussed on the UK.

Projections of future UK drought frequency agree that changes will not be spatially homogeneous. Most studies found a clear north-south divide, with northern (southern) regions facing a decrease (increase) in frequency, with the magnitude of these changes increasing for later periods (Blenkinsop and Fowler, 2007a; Vidal and Wade, 2009; Rahiz and New, 2013; Spinoni *et al.*, 2018). While annual-scale results from Spinoni *et al.* (2018) and Samaniego *et al.* (2018) also showed a north-south divide, droughts are expected to increase in frequency for the entire UK – the geographical split is related to the magnitude of the changes, with southern regions anticipated to see larger increases in frequency than northern regions. Spatial heterogeneity is also a characteristic of projected drought intensity, with Scotland facing less intense droughts and with the opposite effect for southern regions (Rahiz and New, 2013; Dai and Zhao, 2017; Zhao and Dai, 2017; Spinoni *et al.*, 2018). Rahiz and New (2013) used the Met Office Hadley Centre's regional climate model to assess future changes in the 2020s, 2050s and 2080s. For all periods, Scotland is projected to experience a large decrease in drought intensity during the dry season, with the opposite projected for Wales, southeast and southwest England. The intensity of wet season droughts is also anticipated to increase for some regions. Similarly, using the Coupled Model Inter-comparison Project Phase 3 (CMIP3) and Phase 5 (CMIP5) models, Dai and Zhao (2017) showed a north-south split in drought intensity changes between 2070 and 2099. By measuring projected changes in annual sc-PDSI, the authors showed that Scotland is expected to experience a decrease in drought intensity, whilst the rest of the British Isles may see increases in intensity, with the magnitude increasing further south. However, the entire British Isles (along with much of the world) is projected to spend more time in moderate and severe drought (Zhao and Dai, 2017). This implies that Scotland can expect longer duration droughts balanced by an overall wetter annual cycle, whilst the rest of the Isles can expect both longer duration droughts and drier year-round conditions.

Projected changes in precipitation are expected to propagate through to river flow and soil moisture. Summer soil moisture and streamflow is expected to decrease in the UK, and increases in winter flow are likely, subject to regional variations (Christierson *et al.*, 2012; Prudhomme *et al.*, 2012; Sanderson *et al.*, 2012; Ruosteenoja *et al.*, 2018). Furthermore, using seven global impact models driven by general circulation models, Prudhomme *et al.* (2014) showed that hydrological drought frequency and severity are likely to increase in western Europe over the next century (under the most severe emissions scenario, RCP8.5). Roudier *et*

al. (2016) obtained similar results across Europe using three hydrological models for different emissions scenarios and future periods. The greatest magnitude changes in both studies' simulations were driven by decreased precipitation and increased evaporation, while smaller changes were associated with increased precipitation offset by increased evaporation (Garner *et al.*, 2017). These changes in hydrological drought are in agreement with projected meteorological drought changes anticipated by Vidal and Wade (2009), Heinrich and Gobiet (2012) and Rahiz and New (2013) but contradict Blenkinsop and Fowler (2007b). The complexity of meteorological-hydrological-agricultural interactions in drought projections was highlighted by a recent EA report (Environment Agency, 2018), which suggests that many parts of the UK (particularly London and the southeast) will face severe water deficits by 2050. This warning comes as a result of research into linkages between socio-economic factors, such as an increasing population and inadequate water supply practices (for example unsustainable groundwater abstraction), together with the aforementioned projections of climate change and its effects on the hydro-meteorology of the UK.

2.3 Physical drivers of drought

2.3.1 Primary causes of drought

There has been much research in the last two decades into the atmospheric and oceanic processes that might drive drought occurrence. The direct cause of drought is usually an area of high pressure inhibiting cloud formation resulting in reduced precipitation. The US National Drought Mitigation Center (NDMC; 2015) states that the ability to forecast drought is dependent on the prediction of temperature and precipitation, both of which are modulated by a huge range of factors. Precipitation deficit in particular is the primary source of drought (Oladipo, 1985; Lloyd-Hughes and Saunders, 2002; Panu and Sharma, 2002; Hoerling and Kumar, 2003; Dai, 2011). Results from Hoerling and Kumar (2003) showed that winter and spring precipitation deficits were the main reason that large areas of the Northern Hemisphere mid-latitudes saw just 50% of the climatological annual average precipitation during 1998-2002, and drought projections by Dai (2011) indicated that the main reason for future drying in Central America, southwestern USA, the Mediterranean and southern Africa is a projected decrease in mean precipitation. This is despite the fact that, for many mid-latitude and subtropical regions, the main driver behind projected drought is increased evaporation (Dai, 2011). A model for drought forecasting on at least monthly time-scales, however, needs an easily measurable phenomenon related to precipitation at least one month in advance (Cordery and McCall, 2000). Focus should therefore be on long-memory processes such as large-scale atmospheric circulation patterns, oceanic systems, sea ice, snow cover, vegetation and soil

moisture content (Smith *et al.*, 2012; Orth and Seneviratne, 2013; Scaife *et al.*, 2014). What follows is a review of the main drivers of drought in the literature, with a focus on the UK and Europe.

2.3.2 El Niño Southern Oscillation

Globally, the El Niño Southern Oscillation (ENSO) is the largest source of seasonal forecast skill (Smith *et al.*, 2012), with significant statistical correlation found between ENSO indices and climate in many parts of the world (e.g. Ropelewski and Halpert, 1987; Halpert and Ropelewski, 1992; Piechota and Dracup, 1996; Mariotti *et al.*, 2002; Panu and Sharma, 2002; Moron and Plaut, 2003; Brönnimann, 2007; Vicente-Serrano Sergio *et al.*, 2011; Zhang *et al.*, 2015; Herceg-Bulić *et al.*, 2017). The ENSO cycle describes the fluctuations in temperature between the ocean and atmosphere in the tropical eastern Pacific Ocean, with anomalously warm (cool) phases known as El Niño (La Niña), generally lasting nine to 12 months and occurring on average every two to seven years (National Oceanic and Atmospheric Administration, 2015). The Northern Hemisphere drought in 1998-2002 coincided with a protracted La Niña event, with the Pacific Ocean exhibiting persistent negative sea-surface temperature (SST) anomalies in the Tropics. This resulted in a persistent tropospheric circulation pattern, which forced a belt of high pressure over the mid-latitudes (Hoerling and Kumar, 2003). This corresponds with results from Vicente-Serrano Sergio *et al.* (2011), who showed that La Niña events are related with drought over large parts of the Northern Hemisphere (plus some areas of South America and southern Africa). An almost opposite pattern was found for El Niño events, with the Southern Hemisphere more affected than during La Niña events, but Northern Hemisphere regions such as northern South America, the Indian subcontinent, Southeast Asia and parts of Canada and Alaska are also influenced. In all cases, the lag of the ENSO signal to drought effects varies by region. In the 1980s and early 1990s many papers questioned the strength or influence of the ENSO signal over climate in Europe (e.g. Ropelewski and Halpert, 1987; Halpert and Ropelewski, 1992), and hence its usefulness as a teleconnection for drought forecasting in this region (e.g. Pongrácz *et al.*, 2003). More recently, however, there have been many studies that found significant evidence of the influence of ENSO on various features of European and North Atlantic climate, including other atmospheric circulation patterns such as the North Atlantic Oscillation (NAO) (e.g. Moron and Plaut, 2003; Brönnimann, 2007; Scaife *et al.*, 2014; Herceg-Bulić *et al.*, 2017), mean sea-level pressure (MSLP; Folland *et al.*, 2015) and precipitation (e.g. Mariotti *et al.*, 2002; Ineson and Scaife, 2009; Shaman and Tziperman, 2010). These relationships have been found to vary in strength depending on where the main El Niño heating occurs (i.e. in the east or central Pacific

Ocean) and the temperature difference between the Pacific and Atlantic Oceans (López-Parages *et al.*, 2016).

Links between ENSO and precipitation have been found in various European regions. Mariotti *et al.* (2002) showed a significant influence of ENSO on the interannual variability of precipitation in the European-Mediterranean region, with the rainy season in the western Mediterranean arriving and retreating earlier than the climatological norm during El Niño events. A link between El Niño and Mediterranean precipitation was also found by Ineson and Scaife (2009); namely that El Niño precedes a decrease in precipitation over this region. Similarly, Shaman and Tziperman (2010) found a significant ENSO modulation effect on southwest Europe/western Mediterranean precipitation during August-November. Precipitation over central and northwestern Europe has been found to increase during El Niño events (Ineson and Scaife, 2009). It is possible that the North Atlantic region exhibits a more stable atmospheric response to La Niña than El Niño (Pozo-Vázquez *et al.*, 2005), and relationships between two different types of La Niña and their opposing effects on western Europe have been identified (see Zhang *et al.*, 2015, and references therein). For the UK winter half-year, Folland *et al.* (2015) found that La Niña events are associated with drier (wetter) than average conditions over much of England and Wales (northwest Scotland), while El Niño events yield largely opposite results.

2.3.3 The North Atlantic Oscillation and the Arctic Oscillation

The NAO (Walker and Bliss, 1932) describes the difference in sea-level atmospheric pressure between the Icelandic low and the Azores high. A stronger gradient between the two locations (called a positive phase of the NAO) results in an increased zonal flow, whereas a weaker gradient (a negative NAO) weakens the jet stream resulting in blocked conditions. The NAO is a key factor in modulating Atlantic seasonal climate and its effect on climate in the Northern Hemisphere has been widely studied (e.g. van Loon and Rogers, 1978; Wallace and Gutzler, 1981; Hurrell, 1995; Wilby *et al.*, 1997; Rodwell *et al.*, 1999; Qian *et al.*, 2000; Fowler and Kilsby, 2002a; Haylock and Goodess, 2004; Trigo *et al.*, 2004; Scaife *et al.*, 2008; Scaife *et al.*, 2014; Hidalgo-Muñoz *et al.*, 2015; Hall and Hanna, 2018; Tosunoglu *et al.*; Vazifehkhah and Kahya, 2018).

A positive phase NAO generally precedes higher winter European temperatures, drier conditions over most of central and south Europe and the Mediterranean region, and wetter conditions over the northwestern European coast (Hurrell, 1995; Rodwell *et al.*, 1999; Tosunoglu *et al.*; Vazifehkhah and Kahya, 2018). Results from Qian *et al.* (2000) and Haylock

and Goodess (2004) support the significance of the NAO's effect on European precipitation as the most important signal in winter, yet not as important as the North Sea pattern described in Qian *et al.* (2000) for other seasons. Over the Iberian peninsula, Trigo *et al.* (2004) found that the NAO has a strong influence on precipitation and flow in the three main rivers (the Douro, Tejo and Guadiana). Notably, their results showed that the one-month-ahead NAO index displays stronger correlation with the river flows in winter than the concurrent NAO index, whilst for autumn streamflow the NAO of the preceding winter may be an important predictor, possibly due to later spring and summer snowmelt contributions (Hidalgo-Muñoz *et al.*, 2015). This suggests there is potential for using the NAO as a predictor in streamflow forecasting with a lead-time of one month (at least for this region and season). Various studies have also identified the winter NAO as a potential predictor of UK summer streamflow (Wilby, 2001; Wilby *et al.*, 2004; Svensson and Prudhomme, 2005; Kingston *et al.*, 2013).

Over the UK, positive correlations between the winter NAO and precipitation have been found in northern and southern Scotland (but not eastern Scotland) and northwestern England, and negative correlations in southern, central and northeast England (Wilby *et al.*, 1997; Fowler and Kilsby, 2002a). However a more recent study did not show significant negative correlations for these regions (Hall and Hanna, 2018), implying that since around the year 2000 the relationship has weakened. According to Wilby *et al.* (1997), since before the turn of the 20th Century there have been four main phases of the NAO: pre-1900 (index close to zero), 1900-1930 (strong positive anomalies), 1930-1960s (negative anomalies) and 1960s onwards (strong positive phase). Correspondingly, Fowler and Kilsby (2002a) found that the proportion of dry days (PD) in Yorkshire in winter increased until the 1930s, decreased until the 1970s and increased again until the early 2000s, suggesting a low (high) NAO phase is concurrent with a decrease (increase) in PD (for this season and region). The NAO in summer is of lower amplitude than in winter but still has a significant effect on aspects of European climate (Folland *et al.*, 2009). For example, it exhibits positive correlation with UK land-surface temperatures and strong negative correlations with precipitation over the British Isles, among other effects (Folland *et al.*, 2009; Hall and Hanna, 2018). There is great potential for exploiting the relationship between the NAO and UK climate in a forecast model, as recent studies have shown the winter NAO can be skillfully forecast a month in advance using both dynamical (Scaife *et al.*, 2014; Dunstone *et al.*, 2016) and statistical (Hall *et al.*, 2017; Wang *et al.*, 2017) models. In each of these studies, the persistence of autumn sea ice is a key reason given for forecast skill.

The NAO is a regional manifestation of the hemispheric circulation pattern known as the Arctic Oscillation (AO; Thompson and Wallace, 1998). The AO is a large-scale climate pattern of

zonal flow resembling the NAO but centered over more of the Arctic (Thompson and Wallace, 1998). It influences wintertime climate not just over the Atlantic-European region but the Pacific as well (Thompson and Wallace, 2001). Broadly, a more positive AO index leads to increased temperatures across the US Midwest, central Canada and Europe and increased (decreased) precipitation over northern (southern and Mediterranean) Europe (Thompson and Wallace, 2001; Buermann *et al.*, 2003; Vazifehkhah and Kahya, 2018). Additionally, a positive index suggests higher variability in the storm-track between northeastern America and northern Europe while a negative index is associated with blocking over Alaska and the Atlantic (Thompson and Wallace, 2001). Recent progress has been made with regards to predicting the AO with a lead-time of up to two months, with hints that the winter index may be linked to the preceding October's snow cover extent in Eurasia (Cohen and Fletcher, 2007; Cohen and Jones, 2011; Riddle *et al.*, 2013) and the geopotential height of the 500 hPa surface (Kryjov and Min, 2016). However, the link between October snow cover in Eurasia and the AO only emerged in the 1970s, and the strength of the relationship has changed over time (Peings *et al.*, 2013), perhaps negating its usefulness in certain situations.

2.3.4 Oceanic processes

Oceanic and atmospheric processes are intimately linked. In the North Atlantic, Rodwell *et al.* (1999) argue much of the variability in the NAO can be reconstructed using SSTs, in particular those located off the east coast of the USA and south of Greenland. Scaife *et al.* (2014) found potential predictability of European winters using SSTs in this region: forecasts starting with cold (warm) states lead to winter predictions with a more negative (positive) NAO, though with low correlation. There is also evidence for a link between North Atlantic winter SST anomalies and atmospheric signals over the northwest Labrador Sea the following spring (Czaja and Frankignoul, 1999). A number of studies have also identified relationships between the aforementioned SST anomalies and UK summer streamflow (Colman and Davey, 1999; Wedgbrow *et al.*, 2002; Wilby *et al.*, 2004; Svensson and Prudhomme, 2005; Kingston *et al.*, 2013). Decadal changes in North Atlantic SSTs, referred to as Atlantic multidecadal variability (AMV, or the Atlantic Multidecadal Oscillation) (Sutton *et al.*, 2018), are thought to influence several features of European summer climate. These include temperature (Sutton and Hodson, 2005; Knight *et al.*, 2006; Sutton and Dong, 2012; Ghosh *et al.*, 2017; Ruprich-Robert *et al.*, 2017), mean sea-level pressure (Knight *et al.*, 2006; Ruprich-Robert *et al.*, 2017) and the position of the North Atlantic summer storm track (Dong *et al.*, 2013). Positive phase AMV appears to precede increased precipitation over western Europe (Sutton and Hodson, 2005; Dong *et al.*, 2013), although this feature is not as strong as for other variables such as

temperature (Ruprich-Robert *et al.*, 2017). Winter relationships are not as strong as during summer (Ruprich-Robert *et al.*, 2017). Recently, Kolstad and Årthun (2018) showed that winter European surface-air temperature is partly predictable by autumn SST anomalies in the Barents Sea and Norwegian Sea, with the latter in spring also influencing the following summer's temperatures. These relationships were not, however, stationary in time, with the link between autumn Barents Sea SST and winter European temperature since 1979 being stronger than at any point since 1900.

Arctic sea ice is also thought to exert an influence on climate around the world. A brief summary of links with various climatic phenomena over the Atlantic-European region is presented; see Budikova (2009) and references therein for a comprehensive review of global links. Retreating sea ice generally precedes lower surface air temperatures and precipitation increases throughout the mid-latitudes and a decrease in the speed of westerlies and storm intensity north of 45ºN. The atmospheric response to sea ice coverage bears resemblance to the NAO and AO: negative phases of the NAO/AO frequently coincide with reduced sea ice (Alexander *et al.*, 2004; Balmaseda *et al.*, 2010; Scaife *et al.*, 2014), with the potential for using November sea ice coverage over the Barents-Kara Seas region as a predictor for winter NAO (Scaife *et al.*, 2014; García-Serrano *et al.*, 2015). Hall *et al.* (2017) and Wang *et al.* (2017) have also had success using autumn sea ice, along with other variables, as a predictor in a statistical model used to forecast winter NAO. Furthermore, Caian *et al.* (2018) showed a robust link between winter sea ice concentration and the following winter NAO, which the authors noted may have important consequences for advances in multi-year forecasts.

2.3.5 Land-atmosphere feedbacks

Interaction between the land and the atmosphere can be crucial in drought development. European drought and heatwaves, such as during summer 2003 (García-Herrera *et al.*, 2010), have been linked to preceding precipitation deficits. Fischer *et al.* (2007) conducted RCM simulations to show that four major European heatwaves were preceded by anomalously low spring precipitation, while Vautard *et al.* (2007) showed winter and spring precipitation deficits played a role in 58 hot summers as a result of anomalously dry air transported northwards. Similarly, Hauser *et al.* (2016) estimated that dry soils over western Russia in 2010 increased the risk of a heatwave by a factor of six. There are numerous other studies showing a similar role of soil moisture in heat wave development (Mueller and Seneviratne, 2012; Quesada *et al.*, 2012; Miralles *et al.*, 2014). Soil moisture carries the memory of previous months' climate (Shukla and Mintz, 1982; Entin *et al.*, 2000) and appears to propagate through to streamflow and evapotranspiration (Koster *et al.*, 2010; Orth and Seneviratne, 2013). This may be useful

in predictive applications, as the memory of streamflow and evapotranspiration may be derived from the memory of soil moisture. In North America, soil moisture has been shown to be an important variable in forecasts of temperature and, to a lesser extent, precipitation (Beljaars *et al.*, 1996; Koster *et al.*, 2010).

2.4 Statistical drought forecasting: history and the state of the art

2.4.1 Early studies

Drought forecasting can be traced back as far as the middle of the 20th Century. Namias (1955) described weather forecasts at the time as being based primarily on trends in atmospheric patterns and empirical evidence recorded in previous studies. No detailed physical explanations have been presented for these observed trends simply because there was no such knowledge available. However, the study did recognise that the mechanisms driving drought are linked to atmospheric processes and that accurate predictions of these should be a key aim in attempts to improve drought forecasting capabilities. Yevjevich (1967) agreed, suggesting that a statistical method of drought prediction should involve looking at the time-lags between oceans and other potential sources of predictability that precede precipitation and runoff.

It was not until the development of popular drought indices that drought forecasting really became a prominent area of study, and even then, progress was relatively slow. Early works such as those by Davis and Rappoport (1974) and Rao and Padmanabhan (1984) applied exponential smoothing and Autoregressive Moving Average (ARMA) models to monthly Palmer Drought Index values, while a study by Karl *et al.* (1987) looked at the amount of precipitation required for terminating a drought and subsequently calculated the probability of receiving the necessary quantity using the gamma distribution. By enhancing the theory of runs, Moyé *et al.* (1988) enabled computation of the expected number and duration of droughts over a given number of years, although the study used only annual precipitation data and so the method may not be useful for forecasting with sub-annual lead-times. By deriving exact probability density functions of critical droughts in stationary second-order Markov chains, Şen (1990) predicted transitional probabilities for streamflow deficits and surpluses and the expected duration of droughts for any future period. The relatively small number of studies on this topic published in the three decades following Yevjevich (1967) is perhaps surprising, but since the mid-1990s significant advances have been made.

2.4.2 Modern techniques

Table 2.2 contains key details of every statistically-based drought forecasting study published since 1974 known to the author. The vast majority of studies focused on the USA, with China

also well represented. There are just three papers that attempted to forecast drought in the UK. For brevity, this review attempts only to discuss studies that have a specific focus on drought (or low-flow), and does not mention work on generic streamflow, precipitation etc. forecasting.

Modern statistical drought forecasting generally falls into one of two categories. The first of these is forecasting future values of some variable (e.g. a drought index, precipitation or streamflow), and has been performed in the context of time series analysis (Mishra and Desai, 2005; Modarres, 2007; Fernández *et al.*, 2009; Hwang and Gregory, 2009; Durdu, 2010; Han *et al.*, 2010; Chun *et al.*, 2013; AghaKouchak, 2014; Behrangi *et al.*, 2015; Myronidis *et al.*, 2018), regression models (Kumar and Panu, 1997; Cordery and McCall, 2000; Liu and Juárez, 2001; Chun *et al.*, 2013; Djerbouai and Souag-Gamane, 2016; Ahn *et al.*, 2017; Seibert *et al.*, 2017), machine learning methods (Wedgbrow *et al.*, 2005; Mishra and Desai, 2006; Morid *et al.*, 2007; Chen *et al.*, 2012; Dehghani *et al.*, 2014; Demirel *et al.*, 2015; Djerbouai and Souag-Gamane, 2016; Liu *et al.*, 2017; Seibert *et al.*, 2017) and hybrid models (Kim and Valdés, 2003; Mishra *et al.*, 2007; Bacanli *et al.*, 2009; Özger *et al.*, 2012; Awan and Bae, 2016; Liu *et al.*, 2017). The second category concerns forecasting transition probabilities from one drought class to another (Lohani and Loganathan, 1997; Rodó *et al.*, 1997; Steinemann, 2003; Carbone and Dow, 2005; Paulo *et al.*, 2005; Cancelliere *et al.*, 2007; Hannaford *et al.*, 2011; Madadgar and Moradkhani, 2013; Madadgar and Moradkhani, 2014; Bonaccorso *et al.*, 2015; Saghafian and Hamzekhani, 2015; Ahn and Palmer, 2016; Chen *et al.*, 2016; Hao *et al.*, 2016; Rezaeianzadeh *et al.*, 2016; Hao *et al.*, 2017; Meng *et al.*, 2017; Yan *et al.*, 2017). While the majority of studies detailed in this section use purely statistical models, there are a small number that make use of dynamical meteorological or hydrological forecast systems in conjunction with data-driven methods (Carbone and Dow, 2005; Hwang and Gregory, 2009; Madadgar and Moradkhani, 2014; Hao *et al.*, 2017; Yan *et al.*, 2017). The following subsections will look at this body of literature in more detail. Refer to Table 2.2 for the key details of every study mentioned in Section 2.4.

2.4.3 Time series analysis

By far the most common method of time series analysis uses Autoregressive Integrated Moving Average (ARIMA) models and its seasonal equivalent (SARIMA); see Box *et al.* (2008) for a full description. ARIMA models essentially consider the serial correlation between observations, while SARIMA models allow for nonstationarity within and across the seasonal cycle. They can be adapted to include exogenous time series as predictors (ARIMAX) and to account for heteroskedasticity (i.e. a non-constant variance and covariance function) in the Generalised Autoregressive Conditional Heteroskedastic model (Bollerslev, 1986). As

mentioned in Section 2.4.1, ARMA models (a special case of ARIMA where the time series of observations is considered stationary) were applied to the US Midwest by Davis and Rappoport (1974) and Rao and Padmanabhan (1984) to forecast PDSI values on monthly to annual time-scales. In a study in north-eastern India, Mishra and Desai (2005) used ARIMA and SARIMA models to forecast monthly SPI values with reasonable performance at a one or two-month lead-time, with a study in Turkey yielding similar results (Durdu, 2010). SARIMA models have also been used for streamflow drought forecasting. Modarres (2007) forecasted SSI values for a region in Iran one year ahead, and Fernández *et al.* (2009) used the same index to predict streamflow in Spain, finding that the inclusion of a variable relating precipitation and temperature improves the forecast. Han *et al.* (2010) predicted agricultural drought in northwest China using ARIMA models at lead-times of ten and 20 days using the Vegetation Temperature Condition Index (VTCI; Peng-Xin *et al.*, 2001). In the UK, Chun *et al.* (2013) applied a seasonal ARIMAX model to forecast monthly DSI values using SLP, temperature and the NAO index as predictors. The performance of ARIMA and SARIMA models in predicting northern Algeria SPI at lead-times up to two seasons was compared with two machine learning models by Djerbouai and Souag-Gamane (2016), who found the machine learning methods outperformed the stochastic models. More recently, Myronidis *et al.* (2018) tested the forecast performance of ten ARIMA models on monthly streamflow, with output post-processed to a drought index to reveal the likelihood of future drought states.

Other time series methods have been tested for drought forecasting. For South Carolina, Hwang and Gregory (2009) used an autoregressive model combined with seasonal forecast information from a dynamical system to predict PDSI one to three months in advance. To predict agricultural drought, AghaKouchak (2014) estimated future monthly soil moisture using a linear combination of antecedent soil moisture and resampled historical soil moisture, similar to the Ensemble Streamflow Prediction (ESP; Day, 1985) method commonly used in hydrology. AghaKouchak (2014) showed that the summer 2012 drought in the USA was predictable several months in advance. A similar methodology was applied to forecast SPI in the USA, using historical analogues of precipitation, surface air temperature and an ENSO-related index by Behrangi *et al.* (2015).

2.4.4 Regression analysis

The relationship between a dependent variable and a set of independent variables can be explored by regression analysis. In the context of drought forecasting, the dependent variable is that which is to be predicted (also called the predictand; typically a drought index), whilst the predictors are explanatory variables describing changes in the predictand (antecedent

conditions, teleconnection indices etc.). Kumar and Panu (1997) predicted agricultural droughts one season ahead in northwest India by regressing the yield of a particular crop against a variety of explanatory variables, including several based on precipitation and soil moisture. In Australia, Cordery and McCall (2000) examined the relationships between winter precipitation and two teleconnections: ENSO and the autumn geopotential height (GpH). By using partitioned data, the authors provided a framework where sufficiently low GpH prompts the calculation of the precipitation forecast. Using the Normalized Difference Vegetation Index (NDVI) as the predictand, Liu and Juárez (2001) forecasted drought onset in north-eastern Brazil in ENSO years, achieving moderate success at a four-month lead-time by constructing a model from September through December using only those years with distinct ENSO episodes. Chun *et al.* (2013) conducted a study in the UK using Generalised Linear Models (GLMs; Nelder and Wedderburn, 1972) to quantify the probability of occurrence, and amount, of daily precipitation in assessing drought severity and persistence. Ahn *et al.* (2017) used a hierarchical Bayesian model to forecast summer low-flows in north-eastern USA, utilising antecedent at-site and regional climate information. This kind of model allows for a parsimonious structure that accounts for spatial variation in the predictors across the target region. Indeed, the authors stated that the greatest benefits over the other tested models derived from the spatial heterogeneity of the climate teleconnections. In the Limpopo river basin, southern Africa, Seibert *et al.* (2017) found that multiple linear regression models outperformed machine learning methods in forecasting December to May SSI, despite the latter's non-linear capabilities. The models were tested over a range of catchments with forecast lead-times of up to 12 months. The smaller catchments' predictability was largely driven by antecedent streamflow whereas larger catchments relied more heavily on climate teleconnections for accurate prediction.

Ordinal regression can also be used to predict drought class probabilities. The United States Drought Monitor (USDM) classifies drought using SPI, SSI and a soil moisture variable and so predicting these classes is a way of forecasting meteorological, hydrological and agricultural drought simultaneously. Hao *et al.* (2016) assessed this method in Texas, southern USA, with satisfactory results. A similar log-linear model was applied to predict hydrological drought class transitions in northeast China, with antecedent SPI a better input predictor than SSI (Li *et al.*, 2016). Hao *et al.* (2017) extended the work of Hao *et al.* (2016) by predicting August USDM drought categories using ordinal regression, this time with seasonal climate forecasts from a dynamical model as input. A case study of the central USA 2012 summer drought showed encouraging results at a lead-time of one month, although not three months. The authors

suggested this was because the drought was a result of internal atmospheric processes and so the seasonal climate forecast failed to predict this. The persistence of drought in east China was assessed by Meng *et al.* (2017), who used logistic regression to model the probability of binary drought/non-drought occurrence one season ahead, showing that the Southern Oscillation Index bears influence on drought in this region. An interesting innovation by Lorenz *et al.* (2017) was to compute the (binary) probability of the drought state (using USDM) being more intense than the current observation, rather than the probability of transitioning to any other drought class. Furthermore, no attempt was made to quantify the magnitude of the intensification (i.e. how many drought categories the USDM increases by). The authors found the model is most skilful in forecasting high-amplitude flash droughts and at times when the USA is already experiencing intensifying drought.

2.4.5 Machine learning methods

Machine learning has gained popularity since the advent of powerful computers that can process large amounts of information to identify patterns in data without the prior specification of a particular model. There are several examples of machine learning methods applied to drought forecasting. The earliest example is by Wedgbrow *et al.* (2005), who used "expert systems" to predict summer streamflow for the River Thames. The model classified streamflow according to winter MSLP, GpH and sea-ice predictors, successfully forecasting low-flow nearly 80% of the time between 1971 and 2001.

Artificial neural networks (ANNs; see ASCE Task Committee on Application of Artificial Neural Networks in Hydrology, 2000, for a detailed review) are models that can solve complex nonlinear problems by adaptively identifying patterns and relationships in the data. A typical three-layer feed-forward model consists of an input layer, a hidden layer and an output layer. The input layer receives the input variables (or nodes) which for drought forecasting may be lagged values of a drought index, or explanatory variables, or both. The hidden nodes represent some nonlinear transfer function used to process the information from the input nodes, whilst the output layer is the model output to be used for forecasting. One advantage of using ANNs is that the hidden layer mechanisms do not have to be explicitly defined, which is useful in drought forecasting as the physical processes may not be fully understood. On the other hand, this black-box effect leaves the modeller unable to learn more about the underlying processes.

ANN models been widely applied in hydrology (Dawson and Wilby, 2001; Abrahart *et al.*, 2012) and to drought forecasting in particular by several studies. Mishra and Desai (2006) used recursive and direct multi-step ANNs to forecast SPI values in north-eastern India, finding

improved performances compared to a SARIMA model, while the recursive multi-step ANN was useful for short-term forecasting and the direct multi-step ANN useful for long-term forecasting. In Iran, Morid *et al.* (2007) forecasted SPI and Effective Drought Index (EDI) on seasonal to annual time-scales, using different combinations of antecedent precipitation, SPI, EDI and two climate indices. More recently, an ANN was used to predict monthly SSI in Iran (Dehghani *et al.*, 2014; Rezaeianzadeh *et al.*, 2016) and southern Africa (Seibert *et al.*, 2017), and low-flow in the Moselle river basin (Demirel *et al.*, 2015).

Another machine learning method is random forest regression trees, which construct an ensemble of "decision trees" during training and output the mean of this ensemble. There are two cases of random forests used for drought prediction. Chen *et al.* (2012) forecasted SPI in northeast China at one- and six-month lead-times, finding that the random forest method outperformed ARIMA models. For the Limpopo basin in southern Africa, Seibert *et al.* (2017) found that random forest models did not perform as well as regression models in forecasting summer SSI. Support vector machines (SVMs) attempt to fit a function between two different classes, with every point of each class lying a specified distance from the function. Liu *et al.* (2017) used SVMs to evaluate the performance of at-site and remotely-sensed variables for short lead-time agricultural drought forecasting over the USA, finding the inclusion of remotely-sensed predictors extended the useful forecast horizon from two to four weeks.

2.4.6 Probability models

The other category of drought forecasting model involves calculating the probability of a drought class (or sometimes a drought index value) transitioning to another drought class. Şen (1990) applied a second order Markov chain to predict transition probabilities for streamflow deficits, while first order Markov chains were used in an early-warning system by Lohani and Loganathan (1997), where the full spectrum of drought progression possibilities was specified. Work has also been done using Markov chains to characterise the probabilities of drought persistence and duration as well as class transitions (Şen, 1990; Steinemann, 2003). Paulo *et al.* (2005) took this further by computing the average time in each drought class, the average time to reach a specific class for the first time and the most probable class up to three months ahead. Results in the aforementioned study were compared with those made using log-linear models, with the latter performing slightly better. Carbone and Dow (2005) combined resampled historical temperature and precipitation records with seasonal forecasts from a dynamical model, producing drought class probability predictions for South Carolina. Resampling methods were also used by Saghafian and Hamzekhani (2015) to generate precipitation thresholds that trigger hydrological drought classes in the following month. The authors applied

this model to a river in Iran, with only two of the 24 predictions being false alarms for two major drought years. The probability of drought severity, as well as occurrence, was predicted in Connecticut via frequency analysis of low-flow by Ahn and Palmer (2016), who made use of non-stationary copulas and generalised additive models for location, scale and shape parameters (GAMLSS; Rigby and Stasinopoulos, 2005) to construct bivariate joint return periods for the two drought characteristics. Similarly, Chen *et al.* (2016) used a copula-based probabilistic forecasting model to estimate the conditional distribution of drought in southern China one season ahead.

The applicability of Markov chains for drought forecasting was contested by Cancelliere *et al.* (2007), who questioned the general validity of the method due to insufficient observations of class transitions and percentage errors incurred by the Markov chain method. Instead, the authors proposed a fully analytical approach by deriving the exact auto-covariance matrix of SPI values and using this to compute drought class transition probabilities. Bonaccorso *et al.* (2015) used the same approach to calculate transition probabilities in two ways: from an SPI drought class to another drought class and from an SPI value to a drought class. As might be expected intuitively, the latter method yielded higher precision, as more information about the current drought condition is included in the model. In a comparison with ANNs, Rezaeianzadeh *et al.* (2016) found Markov chain models performed similarly well for hydrological drought class transitions in South Iran. Recently, Montaseri *et al.* (2018) assessed the long-term predictability of drought using Markov chains applied to an ensemble of synthetic (generated) SPI series, concluding that this approach yielded benefits over analysing only observed series.

Bayesian networks have also been found useful in drought prediction. They describe the dependencies between a set of variables via directed acyclic graphs and can be used to predict probability distributions. Madadgar and Moradkhani (2013) incorporated copula functions into this framework to estimate seasonal flow in Colorado, finding the forecast accuracy to be comparable with the commonly-used ESP, yet with more reliable forecast uncertainty. This approach was extended by setting up the model with runoff estimation from a hydrological model, allowing simultaneous hydrological drought forecasting across the whole basin (Madadgar and Moradkhani, 2014). Further innovation on this methodology was undertaken by Yan *et al.* (2017), who used a Monte-Carlo data assimilation technique to quantify the initial condition uncertainty from the hydrological model, rather than relying on a single deterministic input.

Hannaford *et al.* (2011) adopted a different approach to forecasting regional hydrological drought in the UK. By making use of the spatial coherence of drought across Europe,

predictions of the number of drought months in the next six months were made using current drought conditions from European regions exhibiting correlations with the UK. Although the model predicted some response for most droughts, the hit-rate was below 0.5 for all regions outside the southeast.

2.4.7 Hybrid models

By combining two or more models to form a hybrid, advantages from each separate model can be utilised, or disadvantages mitigated. Most drought forecasting hybrid models in the literature contain an ANN component. Kim and Valdés (2003) forecasted PDSI values on monthly to annual time-scales by combining an ANN model with dyadic wavelet transforms to try and capture nonlinear and nonstationary behaviour. The basic idea is that the wavelet transforms were used to decompose the PDSI signal and these sub-signals were forecasted and subsequently reconstructed to the original series using ANNs. Mishra *et al.* (2007) elected to combine nonlinear (ANN) and linear (ARIMA) models to forecast SPI values in northeast India up to six months ahead, with the ARIMA component included to capture the autocorrelation present in SPI time series. At a lead-time of one month, Bacanli *et al.* (2009) forecasted SPI series using the Adaptive Neuro-Fuzzy Inference System (ANFIS), which is a combination of ANNs and fuzzy logic (FL). An issue with FL modelling lies in defining the parameters and the design of fuzzy "if-then" rules, but the learning algorithm component of ANN attempts to mitigate this. ANFIS was also applied to forecasting SPI over the East Asian monsoon region using SST anomalies as predictors (Awan and Bae, 2016). The only example of a drought forecasting hybrid model not containing an ANN component is from Özger *et al.* (2012), who developed a wavelet and FL (WFL) model for long-lead forecasting of modified PDSI (PMDI) values using antecedent PMDI values and an ENSO index as predictors. Essentially the predictors and predictand were separated into their frequency bands and then the predictand series was reconstructed using the forecast bands. WFL was compared to a simple FL model, a pure ANN model and a combination wavelet-ANN model, with WFL exhibiting results of greater accuracy.

2.5 Dynamical models for drought prediction

Dynamical models are based on physical equations governing processes of the atmosphere, ocean, land and their interaction. Forecasts using these models are initialised at a particular time and run forward to the desired forecast horizon. Typically they are run as an ensemble (Leutbecher and Palmer, 2008), with different ensemble members initialised with varying initial conditions as a way to quantify uncertainty and to provide a probabilistic, rather than

deterministic, forecast. Due to the time-scales drought is associated with (i.e. at least monthly), it is seasonal forecast systems that are typically used for drought prediction (as opposed to weather forecasting systems that are used for hourly to weekly time-scales). There are a number of such systems, such as the European Centre for Medium Range Weather Forecasting (ECMWF) seasonal forecast system 4 (Molteni *et al.*, 2011) and the Met Office global seasonal forecasting system version 5 (GloSea5) (MacLachlan *et al.*, 2015).

For drought, the predictions of temperature and precipitation are key as they can be used to calculate drought indices (Yoon *et al.*, 2012; Yuan and Wood, 2013; Dutra *et al.*, 2014; Lavaysse *et al.*, 2015; Mo and Lyon, 2015; Caron *et al.*, 2018). However, seasonal forecast skill of these variables is low outside of the tropics due to the chaotic nature of the climate system (Kirtman and Pirani, 2009; Smith *et al.*, 2012; Saha *et al.*, 2014). Furthermore, several authors argued that predicting drought onset is essentially a stochastic problem at longer lead-times (Yuan and Wood, 2013; Dutra *et al.*, 2014), although these studies attempted to predict drought on a global scale. For European drought, Lavaysse *et al.* (2015) found that, using two ECMWF prediction systems, skill scores were improved when forecasting larger-scale droughts. Also using the ECMWF seasonal prediction system, Caron *et al.* (2018), predicted European summer drought (quantified using the SPEI) and compared it to a benchmark statistical model based on historical analogues. They found the dynamical model had slightly higher skill than the benchmark model in southern and western Europe and for longer lead-times. However, they concluded that overall skill was similar, and recommended the use of the benchmark as a computationally cheaper and faster method than the dynamical model. Xu *et al.* (2018) went a step further, finding that the majority of the statistical methods they tested in forecasting SPI in China had higher accuracy than a multi-model ensemble of dynamical models beyond a lead-time of one month. For a binary drought onset prediction, however, the dynamical models were more promising.

Skill in forecasting large-scale atmospheric circulation is often higher than for precipitation (Lavers *et al.*, 2014; Vitart, 2014; Lavers *et al.*, 2016a; Lavers *et al.*, 2016b; Lavers *et al.*, 2017; Baker *et al.*, 2018; Lavers *et al.*, 2018), and hence most drought indices. For example, forecasts of the winter NAO one season ahead are fairly skilful (Scaife *et al.*, 2014; Hall *et al.*, 2017), and these forecasts can in turn be used to predict precipitation. For example, in the UK, a hybrid dynamical-statistical approach that derived precipitation estimates from atmospheric pressure forecasts outperformed the direct dynamical system's precipitation forecast (Baker *et al.*, 2018). This is an example of downscaling, which aims to bridge the resolution gap between GCMs and the regional scale via statistical or dynamical techniques; see Fowler *et al.* (2007) for a

review. Hydrological drought forecasting requires different models to meteorological drought prediction. ESP uses historical analogues of streamflow as these drivers while climate model output is also often incorporated (Hao *et al.*, 2018). In the latter case, the forecast skill is dependent on both initial hydrologic conditions in the catchment and on the climate forecasts (Emerton *et al.*, 2016; Arnal *et al.*, 2017), while in general, as with most forecast systems, skill is dependent on region, season and lead-time (Shukla *et al.*, 2013; Harrigan *et al.*, 2018).

Dynamical models are advantageous compared to statistical models as they are able to capture nonlinear interactions of the earth system components, and as such are adaptable to a changing climate and capable of predicting unprecedented conditions (Hao *et al.*, 2018). Statistical models are reliant on historical relationships between predictors and the target variable, and these relationships are not guaranteed to continue in the future. Furthermore, statistical relationships are often too simplistic to properly represent the underlying physical processes, often resulting in unsatisfactory performance. However, the simplicity of statistical models (compared to dynamical) is also their primary advantage. They are easier and computationally cheaper to implement, and do not require the huge efforts in data assimilation, model calibration and parameterisation demanded by physical (dynamical) models (Hao *et al.*, 2018). Furthermore, dynamical models are subject to a variety of sources of uncertainty such as in the initial conditions, physical parameterisations of certain processes and internal variability. For example, GCMs tend to underestimate drought persistence on the global scale, with model and initial-condition uncertainty the main culprits (Moon *et al.*, 2018). Research in drought forecasting has not focussed on either statistical or dynamical modelling alone and hybrid methods are commonplace now. These methods come in various forms although generally involve calibrating a climate forecast from a dynamical model and merging this with one or more statistical forecasts and/or techniques (Lim *et al.*, 2011; Wang *et al.*, 2012; Madadgar and Moradkhani, 2013; Robertson *et al.*, 2013; Schepen *et al.*, 2014; Schepen and Wang, 2015; Madadgar *et al.*, 2016; Yan *et al.*, 2017; Lorenz *et al.*, 2018; Xu *et al.*, 2018).

2.6 Weather-pattern classifications

2.6.1 General concepts

A weather-pattern classification is a group of weather patterns (WPs; also called weather types, circulation patterns and circulation types), which typically characterise one or more meteorological variables over a particular domain and time-scale. Some studies have used the term “weather regimes”, which often describes the same thing as WPs, but in some cases they are fewer in number, defined over larger domains, are more persistent, or defined by their

highest probability of occurrence (Michelangeli *et al.*, 1995; Santos *et al.*, 2005; Boé, 2013; Ferranti *et al.*, 2015; Vigaud *et al.*, 2018). According to Huth *et al.* (2008), classifications were first employed several decades ago by meteorologists concerned with weather forecasting. Increases in computing power led to a change in methodology of these forecasts and classifications dropped out of use. The result of this was increased interest by climatologists, who are still the primary users of such classifications. Recently, however, WP classifications have made a resurgence in weather prediction. There are myriad studies describing and utilising WP classifications and this section does not attempt to discuss each one, but instead focuses on broad themes concerning WPs, with references to the key texts where appropriate. This shall be kept rather general, with specific studies or applications discussed in detail at the beginning of each of the following chapters. The review of WP classifications by Huth *et al.* (2008) was an important source of information in this section.

WPs are often defined by an air pressure variable, most commonly (M)SLP or a geopotential height in the lower- to mid-troposphere (up to 500 hPa). As such, they represent the broad-scale atmospheric circulation over a given region and time-scale, often daily (Huth *et al.*, 2008). WPs can then be used to make general estimates of climatic variables such as wind direction and speed, temperature and precipitation. They are a useful way of simplifying complex meteorological data sets into a few discrete states. For example, each day in a time series of daily MSLP can be classified (via some assignment procedure; see later) into one WP from the classification used. This achieves a reduction in data dimensionality by going from a continuous time series of many (theoretically infinite) MSLP values to a discrete time series of dimensionality equal to the number of WPs in the classification. Of course, this has consequences. Huth *et al.* (2008) warned: "Atmospheric circulation constitutes more a continuum than a system with several, clearly defined, well-separated states. Any classification of WPs should therefore be viewed as a purpose-made simplification of reality rather than a reality itself". Nevertheless, this is clearly considered a reasonable compromise by the many researchers that have used WPs.

2.6.2 Methodology: derivation and assignment of WP classifications

There are two methodological branches related to WP classifications:

1. The derivation method of the classification itself i.e. how the WPs are defined.
2. The assignment procedure used to classify each time step from the underlying data set used in (1) as one of the WPs.

Each of these branches can be approached using either subjective or objective methods, resulting in classifications that may be considered totally subjective, totally objective, or mixed. This last definition pertains to classifications with WPs that are defined subjectively, while the assignment procedure is carried out objectively. The earlier WP classifications are typically subjective. The two most well-known examples are the Grosswetterlagen (GWL) catalogue (Hess and Brezowsky, 1952) and the Lamb catalogue (Lamb, 1950; Lamb, 1972). The Lamb WPs (commonly abbreviated to LWTs) are daily, while the GWL are required to last for at least three consecutive days, a relatively rare requirement in classifications. Both classifications had their respective WPs defined prior to the assignment procedure, the former using expert knowledge and the latter using physical considerations such as airflow direction and degree of cyclonicity. Time series of these classifications were also subjectively derived, by visual attribution of daily pressure fields to the closest WP. In addition, both the GWL and Lamb catalogues exist as mixed classifications, where the assignment procedure is objective, based on pattern correlation for the GWL (James, 2007) and airflow direction, vorticity and intensity threshold criteria for the LWTs (Jenkinson and Collison, 1977).

Objective classifications are by far the most common, in no small part due to increases in computing power achieved over the last few decades making it easier for researchers to design classifications specific to their purpose. It is important to note that most objective classifications require some subjective input at some stage, for example when deciding on the number of desired WPs. Various methods have been used to define WP classifications, including correlation techniques (Lund, 1963; McKendry *et al.*, 1995; Saunders and Byrne, 1999; Brinkmann, 2000; Bischoff and Vargas, 2003; Schoof and Pryor, 2006), cluster analysis (Fernau and Samson, 1990; Wilson *et al.*, 1992; Cheng and Wallace, 1993; Corte-Real *et al.*, 1998; Esteban *et al.*, 2005; Santos *et al.*, 2005; Esteban *et al.*, 2006; Stahl *et al.*, 2006; Philipp *et al.*, 2007; Vrac *et al.*, 2007; Cheng *et al.*, 2010; Jiang, 2011; Boé, 2013; Ferranti *et al.*, 2015; Moron *et al.*, 2016; Neal *et al.*, 2016; Prein *et al.*, 2016; Rueda *et al.*, 2016; Vigaud *et al.*, 2018), principal components analysis (PCA) (Wilson *et al.*, 1992; Huth, 1996; Huth, 1997; Huth, 2000; Jacobeit *et al.*, 2003; Müller *et al.*, 2003; Almazroui *et al.*, 2015) and neural networks (Hewitson and Crane, 2002; Cassano *et al.*, 2006a; Cassano *et al.*, 2006b; Hope *et al.*, 2006; Lynch *et al.*, 2006; Michaelides *et al.*, 2007; Reusch *et al.*, 2007; Nigro *et al.*, 2011). A brief description of these methods follows; for more detail refer to the relevant references. For the purposes of this section, the word “field” refers the spatial distribution of the variable(s) used to define WPs in a classification, for example daily MSLP over Europe.

Correlation techniques derive WPs by identifying the field with the most correlations above a threshold with the other fields, and subsequently removing these fields from the data set. This procedure is repeated until either all days are classified or the size of the WPs (the number of fields assigned to a WP) drops below a pre-specified threshold. There is an analogous technique that replaces the correlation scores with sums-of-squares of differences. Correlation-based methods are adept at yielding adequate separation between resulting WPs, meaning the derived WPs are dissimilar to each other. A key disadvantage of this method is that there tends to be vastly unequal numbers of fields assigned to each WP, with the first identified WP often by far the largest. As a result, the derived WP frequencies of occurrence will be very different.

Cluster analysis is the most natural, and most common, approach to defining a WP classification. The classic clustering algorithm is k-means (Hartigan and Wong, 1979), which aims to partition the data into a number of clusters by minimising the within-cluster variance. This algorithm is non-hierarchical, meaning that the number of desired WPs must be specified prior to implementation. This introduces a degree of subjectivity into the derivation. Typically, the choice involves balancing the number of WPs with the within- and between-WP variability. Too few WPs results in within-WP variability being too large, while too many can yield a number of WPs that appear very similar to each other. To avoid this choice, some studies have employed hierarchical clustering algorithms, which begin with each field in its own cluster, and clusters are iteratively merged until some criteria is met (Kalkstein *et al.*, 1987). An improvement to the k-means clustering method is simulated annealing, used by Philipp *et al.* (2007) and Neal *et al.* (2016), which differs from k-means primarily in that it does not converge to a local optimum from which it cannot leave, but instead allows an object to leave a cluster. This allows all possible combinations of data points to be compared.

PCA is often used prior to cluster analysis in order to remove any collinearity between the input variables, which would give too much weight to the strongly correlated variables. PCA (with a different matrix configuration) is also used as a WP derivation tool. It is a natural method for classifying WPs, as it aims to transform a set of possibly correlated observations into a set of uncorrelated values (principal components), ordered such that the first principal component explains the largest amount of variance of the data.

Neural networks are the main nonlinear method used in creating WP classifications, in particular self-organising maps (SOM) which have featured in meteorological studies since the 1990s (Liu and Weisberg, 2011). SOMs are a method of projecting high-dimensional data onto low-dimensional space and as such are well-suited to creating WP classifications. A classification derived using SOM is a two-dimensional array of maps, with a key benefit being

easy interpretation of resulting WPs: the most dissimilar WPs are located at opposite ends of the array's diagonal.

2.6.3 Applications

The literature on applications of WP classifications is dominated by climatological studies. The most common classification property studied when analysing recent observed changes is WP frequencies of occurrence. This has been done in the case of investigating changes in the atmospheric circulation as a standalone feature (Lamb, 1950; Jenkinson and Collison, 1977; Bárdossy and Caspary, 1990; O'Hare and Sweeney, 1993; Stefanicki *et al.*, 1998; Goodess and Jones, 2002; Bartholy *et al.*, 2006; Kučerová *et al.*, 2017), although the variety of classifications used and lack of a common time period used make any conclusions about "true" physical changes difficult to justify. Some studies attempt to address this issue by analysing multiple classifications simultaneously, as this should reduce artefacts arising from individual data sets (Kučerová *et al.*, 2017; Stryhal and Huth, 2017; Stryhal and Huth, 2018). This issue was clearly highlighted by Philipp *et al.* (2010), who showed that characteristics of 22 European classifications varied enormously.

WP frequencies are also analysed in the context of hydro-meteorological phenomena. These studies might have been concerned with a general assessment of such phenomena, such as how a WP classification relates to precipitation climatology i.e. the entire precipitation distribution (Hay *et al.*, 1991; Bárdossy and Plate, 1992; Wilby, 1995; Goodess and Jones, 2002; Santos *et al.*, 2005; Hope *et al.*, 2006; Stahl *et al.*, 2006; Malby *et al.*, 2007; Trnka *et al.*, 2009; Lorenzo *et al.*, 2011; Fernández-González *et al.*, 2012; Almazroui *et al.*, 2015; Moron *et al.*, 2016; Prein *et al.*, 2016; Maheras *et al.*, 2018). More commonly, changes were analysed with respect to extreme events, such as extreme precipitation (Hay *et al.*, 1991; Wilby, 1998; Esteban *et al.*, 2005; Burt and Ferranti, 2012; Prein *et al.*, 2016; Maheras *et al.*, 2018), temperature extremes (Bischoff and Vargas, 2003; Müller *et al.*, 2003; Cassano *et al.*, 2006a; Blenkinsop *et al.*, 2009; van den Besselaar *et al.*, 2009), droughts (Hay *et al.*, 1991; Phillips and McGregor, 1998; Fowler and Kilsby, 2002b; Fleig *et al.*, 2010; Giuntoli *et al.*, 2013; Prein *et al.*, 2016), fluvial flooding (Wilby, 1993; Bárdossy and Filiz, 2005; Pattison and Lane, 2012; Wilby and Quinn, 2013) and coastal flooding (Ramos *et al.*, 2013; Rueda *et al.*, 2016; Neal *et al.*, 2018).

More detailed discussion of the literature surrounding WPs and UK precipitation and drought is in Chapter 3. The main purpose of these kind of studies is to identify which WPs coincide with extreme events. This provides insight into the common synoptic situations associated with such extremes. WP persistence is another characteristic that has been studied. Using the term

'singularities', Lamb (1950) identified periods of persistence for certain LWTs that tended to occur at the same times each year between 1898 and 1947. More commonly, WP persistence has been studied in relation to changes in atmospheric persistence (Werner *et al.*, 2000; Kyselý, 2002; Kyselý and Domonkos, 2006; Kyselý and Huth, 2006) and to European heatwaves (Kyselý, 2002; Domonkos *et al.*, 2003; Kyselý, 2007). A more in-depth review of the literature surrounding WP persistence is in Chapter 4.

WP classifications have also been used for future climate projections. They lend themselves to this kind of analysis as GCM output is complex, so reducing this complexity using classifications is an attractive proposition. Inference of future atmospheric circulation can be made by analysing the projected changes in WP frequencies between the current and future periods, with some studies then relating these changes to local-scale meteorological variables (Hay *et al.*, 1992; Wilby, 1994; Kidson and Watterson, 1995; Saunders and Byrne, 1996; Huth, 1997; Corte-Real *et al.*, 1999; Huth, 2000; Fowler *et al.*, 2005; Cassano *et al.*, 2006b; Hope, 2006; Lynch *et al.*, 2006; McKendry *et al.*, 2006; Schoof and Pryor, 2006; Demuzere *et al.*, 2009; Donat *et al.*, 2010; Lorenzo *et al.*, 2011; Brigode *et al.*, 2018; Fereday *et al.*, 2018).

More recently, WP classifications have found application in sub-seasonal to seasonal weather forecasting, in a shift back toward their original use. All forecasting applications that have used classifications have assigned output from a dynamical ensemble forecast system to WPs, which yielded probabilistic predictions of WPs. These predictions were then used to verify the underlying model skill (Nigro *et al.*, 2011; Ferranti *et al.*, 2015) or as a step toward forecasting some other variable (Vuillaume and Herath, 2017; Baker *et al.*, 2018; Neal *et al.*, 2018). Some studies did not explicitly forecast the WPs themselves, but assessed their utility in potential forecast models by considering 'perfect' WP forecasts (i.e. the observations) and whether this information would be beneficial in predicting climatic variables (Schiemann and Frei, 2010; Fernández-González *et al.*, 2012; Ramos *et al.*, 2013). A more comprehensive review of WP classifications in a forecasting context is given in Chapter 5.

2.7 Summary

This chapter has highlighted that there is a substantial amount of research on statistical drought forecasting, although only three studies are focussed on the UK. Of these, two are for hydrological drought (Wedgbrow *et al.*, 2005; Hannaford *et al.*, 2011) and the primary aim of the third was to compare droughts in the late 20th Century with projections for the 2080s (Chun *et al.*, 2013). This represents a clear opportunity for developing a more operationally-focussed meteorological drought prediction model, in the style of dynamical drought forecast systems

discussed in Section 2.5. This chapter has also reviewed the literature surrounding the links between UK precipitation (and associated drought indicators) and potential predictors such as atmospheric and oceanic variables. The general lack of strong, consistent, relationships (with the exception of some cases, such as the NAO with precipitation in winter) is motivation to consider WP classifications as a means to reduce the effective complexity (i.e. noise) of the analyses. The concept of WP classifications has been introduced, with the methodological choices involved in their derivation and how they have been applied, including the recent tendency towards forecasting. Not only is their use in dynamical forecasting an emerging field, there is only one study that attempts to predict WPs statistically (Fayos and Fayos, 2007), representing a well-defined research gap for researching statistical models for WP and a comparison with dynamical models. A further motivation for exploring the use of WPs in drought prediction is the higher skill that many dynamical forecast systems have in predicting atmospheric variables compared to precipitation (as discussed in Section 2.5).

Study	Drought Type	Area of study	Methodology	Lead time	Forecast month/season	Forecast variable
Davis and Rappoport (1974)	M	Midwest USA	Exponential smoothing, ARMA	1 month	All	PDSI
Rao and Padmanabhan (1984)	M	Midwest USA	ARMA	Monthly and yearly	All	PDSI
Karl *et al.* (1987)	M	USA	Calculate amount of precipitation required to terminate a drought then calculate the probability of receiving this amount by fitting the gamma distribution.	1, 2, 3 and 6 months	All	PDSI and PHDI
Moyé *et al.* (1988)	M	South USA	Enhanced Yevjevich (1967) theory of runs to enable calculation of expected number of droughts in *n* years and expected duration of a drought over next *n* years.	Return periods of *n* years	All	Precipitation
Şen (1990)	H	Switzerland, Romania and Midwest USA	Probabilistic second order Markov chain to predict transitional probabilities for streamflow deficits and surpluses and expected critical drought durations.	Any	All	Streamflow
Kumar and Panu (1997)	A	Northwest India	Linear regression predicting crop yield using as predictors the number of rainy days, rainfall amount, Soil Moisture Index (SMI) and a variable representing delay time in sowing.	The crop growing season. For this particular crop 13 weeks.	July – September	Grain yield
Lohani and Loganathan (1997)	M	South USA	Probabilistic early-warning system using first-order Markov chain to predict monthly drought class transitions.	Monthly	All	PDSI
Cordery and McCall (2000)	M	Australia	Partitioned data by identifying seasons and years with highest GpH and Southern Oscillation Index (SOI). Using these dates, precipitation was regressed against GpH and SOI.	Seasonal	All	Precipitation

Table 2.2: Studies on statistical drought forecasting detailing type of drought studied (M = meteorological, H = hydrological and A = agricultural), region of study, description of methods, forecast lead-time, forecast target month and forecast variable.

Study	Drought Type	Area of study	Methodology	Lead time	Forecast month/season	Forecast variable
Liu and Juárez (2001)	A	Northeast Brazil	Regressed NDVI against four different ENSO variables.	4 months	All	NDVI
Steinemann (2003)	M/H	Southeast USA	Used Markov chain model to characterise probabilities of drought class transitions, drought persistence and duration.	Monthly	All	PDSI, PDHI and SPI
Kim and Valdés (2003)	M	Northern Mexico	Hybrid of ANN combined with dyadic wavelet transforms.	1, 3, 6 and 12 months	All	PDSI
Carbone and Dow (2005)	M	South Carolina	Combined historical temperature and precipitation records with seasonal forecasts from a dynamical model.	Seasonal	All	Precipitation, temperature and PDSI
Mishra and Desai (2005)	M	Northeast India	ARIMA/SARIMA	1-6 months	All	SPI
Paulo *et al.* (2005)	M	Southern Portugal	Markov chain to estimate probabilities of different drought classes, expected time in each class, average time to reach a specific class for the first time from some other class and most probable class 1-3 months ahead. Also fitted several log-linear models to drought class transition matrices.	1-3 months	All	SPI
Wedgbrow *et al.* (2005)	H	River Thames, England	Used expert systems (similar to fuzzy logic and machine learning) with MSLP, GpH and sea ice predictors.	2 seasons	Summer and August	Streamflow
Mishra and Desai (2006)	M	Northeast India	Recursive and direct multi-step ANN.	1-6 months	All	SPI

Table 2.2 (cont.)

Study	Drought Type	Area of study	Methodology	Lead time	Forecast month/season	Forecast variable
Cancelliere *et al.* (2007)	M	Sicily, southern Italy	(i) Computing drought class transition probabilities by analytically deriving auto-covariance matrix of SPI values, and (ii) forecasting future SPI values based on conditional expectation of past precipitation values.	3, 6, 9 & 12 months	All	SPI
Mishra *et al.* (2007)	M	Northeast India	Hybrid combining ARIMA with two types of ANN: multi- and direct-step.	1-6 months	All	SPI
Modarres (2007)	H	Central Iran	ARIMA/SARIMA	12 months and return periods of 2-, 5-, 10- and 20-years.	All	SSI
Morid *et al.* (2007)	M	North-central Iran	ANN with NAO and ENSO indices as predictors.	3, 6, 9 and 12 months	All	SPI and EDI
Bacanli *et al.* (2009)	M	Central Turkey	ANFIS: hybrid of ANN and FL system.	1 month	All	SPI
Fernández *et al.* (2009)	H	Northwest Spain	SARIMA with precipitation and a temperature-precipitation relation variable as predictors.	12 months	All	SSI
Hwang and Gregory (2009)	M	South Carolina	Residual-resampled ARMA forecasts weighted with seasonal forecasts from a dynamical model.	1-3 months	All	PDSI and SPI
Han *et al.* (2010)	A	Northwest China	ARIMA	10 and 20 days	All	VTCI

Table 2.2 (cont.)

Study	Drought Type	Area of study	Methodology	Lead time	Forecast month/season	Forecast variable
Durdu (2010)	M	Southwest Turkey	ARIMA/SARIMA	1-6 months	All	SPI
Hannaford *et al.* (2011)	H	UK	GLM. Predicted number of "drought months" over next 6 months, using 6-monthly mean RDI from neighbouring countries.	1-6 months	All	Regional Deficiency Index
Chen *et al.* (2012)	H	Northeast China	ARIMA and random forest regression trees	1 and 6 months	All	SPI
Özger *et al.* (2012)	M	South USA	Wavelet and FL hybrid using an ENSO index as a covariate.	3, 6, 12 months	All	Modified PDSI
Chun *et al.* (2013)	M	UK	(i) SARIMA with SLP, temperature and the NAO index as exogenous variables, and (ii) GLM (both a rainfall amount and occurrence model) with SLP, temperature and relative humidity as covariates.	1-12 months	All	DSI
Madadgar and Moradkhani (2013)	H	Upper Colorado river basin, western USA	Bayesian networks with copula functions to link the marginal distributions of spring, winter and autumn flow.	One season	April-June	SSI
AghaKouchak (2014)	A	USA	6-month accumulations of soil moisture with resampling, similar to Ensemble Streamflow Prediction methods.	1-4 months	May-August	Standardised Soil Moisture Index
Bonaccorso *et al.* (2014)	M	Sicily, southern Italy	Models transition probabilities from drought class to drought class and from SPI values to drought class, with and without NAO as a predictor. The probabilities are calculated by deriving the auto-covariance matrix.	1-3 months	All	SPI

Table 2.2 (cont.)

Study	Drought Type	Area of study	Methodology	Lead time	Forecast month/season	Forecast variable
Dehghani *et al.* (2014)	H	Southwest Iran	Feed-forward ANN with principle components analysis applied to input predictors of precipitation and temperature. Implemented in a Monte-Carlo framework to determine the uncertainty.	1 month	All	SSI
Madadgar and Moradkhani (2014)	H	Upper Colorado river basin, western USA	Bayesian networks with copula functions to link the marginal distributions of spring, winter and autumn flow. The model is setup with runoff estimation from a hydrological model.	1-6 months	July-December	SSI
Behrangi *et al.* (2015)	M	USA	Construct probability density functions (PDFs) for SPI-6 using historical analogues with precipitation, SLP and ENSO data.	1, 2, 3 and 6 months	All	SPI
Demirel *et al.* (2015)	H	Moselle river basin (Germany, France and Luxembourg)	Two hydrological models and one ANN model.	1 season	All	Streamflow
Saghafian and Hamzekhani (2015)	H	Northwest Iran	Used precipitation thresholds to trigger hydrological drought conditions in the next month.	1 month	All	SSI
Ahn and Palmer (2016)	H	Connecticut river basin, Northeast USA	Nonstationary frequency analysis to low-flow occurrence and severity, using copulas and GAMLSS models.	Annual	All	Streamflow

Table 2.2 (cont.)

Study	Drought Type	Area of study	Methodology	Lead time	Forecast month/season	Forecast variable
Awan and Bae (2016)	M	East Asia monsoon region	ANFIS with antecedent SPI and SST as predictors.	3 month	All	SPI
Chen *et al.* (2016)	M	South China	Conditional distribution of drought given previous season's drought, using copulas.	1 season	All	SPEI
Djerbouai and Souag-Gamane (2016)	M	North Algeria	ANN compared with ARIMA/SARIMA	1-6 months	All	SPI
Hao *et al.* (2016)	M/H/A	Texas, southern USA	Ordinal regression	1 month	All	Drought category (using SPI, SSI and total column soil moisture percentile)
Li *et al.* (2016)	H	Northeast China	Log-linear models using antecedent meteorological and hydrological drought classes.	1 and 2 months	All	Drought class (based on SSI)
Rezaeianzadeh *et al.* (2016)	H	South Iran	Markov chain and ANN	1 month	All	SSI

Table 2.2 (cont.)

Study	Drought Type	Area of study	Methodology	Lead time	Forecast month/season	Forecast variable
Ahn *et al.* (2017)	H	Northeast USA	Bayesian hierarchical model conditioned on antecedent basin characteristics and climate teleconnections.	1 season	July-September	Streamflow
Hao *et al.* (2017)	M/H/A	USA	Ordinal regression with antecedent drought category and precipitation and temperature forecasts from a climate model as predictors.	1 and 3 months	August	Drought category
Meng *et al.* (2017)	M	East China	Logistic regression with previous season's SPI and SOI as covariates.	1 season	All	Probability of occurrence
Liu *et al.* (2017)	A	USA	Applies support vector machines and data assimilation methods with meteorological and remotely-sensed inputs.	1 to 4 weeks	October 2009 – September 2011 and April 2015 – March 2016	Soil water deficit index
Lorenz *et al.* (2017)	M/H/A	USA	Logistic regression with antecedent precipitation, evapotranspiration and soil moisture, plus current USDM index value as predictors.	2, 4 and 8 weeks	May - September	USDM
Seibert *et al.* (2017)	H	Southern Africa	Multiple linear models, ANN and random forest regression trees. Various teleconnections and SST indices tried as predictors.	1 to 12 months	December-May	SSI

Table 2.2 (cont.)

Study	Drought Type	Area of study	Methodology	Lead time	Forecast month/season	Forecast variable
Yan *et al.* (2017)	A	Columbia river basin, Northwest USA	Used a hydrological model with data-assimilation to model the land-surface states. This was used as input to a copula-based probabilistic forecast model.	1 and 2 seasons, and half-yearly	All	Drought category and probability of occurrence
Montaseri *et al.* (2018)	M	10 stations globally	Generated a synthetic ensemble of precipitation time series and applied a Markov model to monthly SPI classes to each.	Year multiples. Up to 26 years for drought inter-arrival times.	All	Inter-arrival time probabilities, drought duration, intensity.
Xu *et al.* (2018)	M	China	Compares several statistical models (mostly machine learning methods) with a dynamical multi-model ensemble system and hybrid models.	1 to 8 months.	All	SPI

Table 2.2 (cont.)

Chapter 3

A new precipitation and drought climatology based on weather patterns

The material in this chapter has been published in the following journal article:

Richardson, D., Fowler, H. J., Kilsby, C. G. and Neal, R. (2018a), A new precipitation and drought climatology based on weather patterns. Int. J. Climatol, 38: 630–648. doi:10.1002/joc.5199

Part of the Introduction of this journal article has been moved to Chapter 2 of this thesis. Other minor changes to the wording have been made to make the article more coherent in the context of this thesis.

3.1 Introduction

Most weather-pattern (WP) studies focussed on aspects of UK precipitation climatology have used Lamb Weather Types (LWTs) (Lamb, 1972), as these are defined on a region centred over the British Isles. Wilby *et al.* (1994) generated daily precipitation time series for several sites in England by relating precipitation to LWTs. These data were then input to a hydrological model to simulate daily flows. Fowler and Kilsby (2002a) linked LWTs to the winter and summer North Atlantic Oscillation (NAO; Walker and Bliss, 1932) and in turn to precipitation in Yorkshire, England. Also for Yorkshire, Fowler *et al.* (2005) simulated climate change scenarios using a stochastic precipitation model conditioned on LWTs, whilst Malby *et al.* (2007) found associations between Lake District orographic precipitation variations and changes in wet circulation types. LWTs have also been studied in the context of flooding and heavy precipitation in the UK. Results from a hydro-chemical model used by Wilby (1993) suggested flood frequencies in the East Midlands, England, are linked to changes in the occurrence of cyclonic and anticyclonic LWTs. Wilby (1998) successfully reproduced low-frequency heavy daily precipitation incidence by relating such events to LWTs for several sites in central and southern England. However, the model did not capture variations in mean wet day probabilities or persistence, which the author attributed to general deficiencies of the LWT classification system and the method used to group individual weather types together. Of the 27 objective LWTs, Pattison and Lane (2012) found that just five accounted for over 80% of recorded extreme floods in Carlisle in northwest England. Burt and Ferranti (2012) showed that an increase in winter heavy precipitation in upland northern England was linked to an increase in the amount of precipitation associated with westerly LWTs. The same authors also showed that a decrease in summer heavy precipitation in the same region was linked to a decrease in the amount of precipitation associated with cyclonic LWTs.

Studies investigating the relationship between WPs and drought in the UK are rarer and normally do not consider the UK as a whole, instead focussing on a small number of regions. Wilby (1993) showed that the occurrence of drought in the East Midlands is related to the frequency of anticyclonic LWTs. Phillips and McGregor (1998) analysed frequencies of LWTs during droughts in southwest England, finding that drought distribution in the region depends on the position of the controlling anticyclone. Using a similar methodology, Fowler and Kilsby (2002b) demonstrated that Yorkshire droughts between 1881 and 1998 were typically associated with changes in frequencies of cyclonic and anticyclonic LWTs, with further differences in easterly and westerly type occurrences for sub-regional droughts. Away from LWTs, Fleig *et al.* (2011) related objective Grosswetterlagen (GWL) (James, 2007) to a hydrological drought index in six regions covering Britain and Denmark. They found that of six WPs associated with drought, five feature a centre of high pressure to the north, although the governing pattern varied between regions and for drought events within each region.

A new set of WPs have been developed by the Met Office (Neal *et al.*, 2016). This classification consists of 30 WPs that are representative of the general atmospheric circulation over the UK and surrounding North Atlantic Ocean and European area, and will be referred to as MO-30. MO-30 has two main advantages over LWTs. First, the 30 patterns are derived objectively from first principles, without *a priori* categorisation of the resultant flow over the UK. By contrast, LWTs are preordained categories of flow direction (i.e. northerly, south-westerly etc.) to which daily pressure patterns are (objectively or subjectively) assigned. Second, the patterns are defined over a much larger area than that used for the LWTs. By including much of the North Atlantic Ocean, MO-30 better captures the large-scale atmospheric systems that drive weather in Europe. Also, the larger region size means that it can be applied to other European regions, whereas LWTs must be redefined for studies outside the UK (see Lorenzo *et al.*, 2008, for an example in Spain). A further set of eight WPs, named MO-8, has been defined by clustering types from MO-30. This is intended for use in long-range and seasonal forecasts.

MO-30 and MO-8 are used by the Met Office for several medium- to long-range probabilistic forecasting applications. Ensemble member forecast scenarios are objectively assigned to the closest matching WP definition, providing a probabilistic insight into the occurrence of different WPs throughout the forecast period. Once WP characteristics are understood, in terms of their climatologies or impact, it then becomes possible to interpret forecast output and describe likely consequences. Three operational WP forecast applications used at the Met Office: a tool for predicting airflow from Iceland that could potentially bring volcanic ash over

UK airspace, a model for potential UK coastal flood events and a model for assessing the likelihood of potential extreme precipitation events.

There are three objectives to this chapter. First, an exploratory analysis of several features of the new classification, such as WP frequencies and precipitation associated with each pattern. LWTs will be used as a comparison. Second, an investigation of how the new data set relates to the UK precipitation climatology, which on a monthly time-scale will be defined by the SPI (McKee *et al.*, 1993). Third, this chapter will expand upon previous work by linking WPs to drought in the UK as a whole, rather than on smaller regions. Drought will be quantified using the DSI after Phillips and McGregor (1998) and Fowler and Kilsby (2002b). Whilst the example provided is for the UK, the methodology is transferrable to other regions within the domain of the WPs.

This chapter is organised as follows. Section 3.2 describes the data sets used. The methodology for associating WPs with daily precipitation and linking their monthly frequencies to drought indices is detailed in Section 3.3. Results are shown in Sections 3.4 and 3.5, a discussion is in Section 3.6 and conclusions are presented in Section 3.7.

3.2 Data

The methodology for deriving MO-30 and MO-8 is described in Neal *et al.* (2016). Briefly, 154 years (1850 to 2003) of daily MSLP fields from the European and North Atlantic Daily to Multi-decadal Climate Variability data set (EMULATE; Ansell *et al.*, 2006) were grouped into 30 distinct clusters using a simulated annealing technique (Philipp *et al.*, 2007; Huth *et al.*, 2008). The data have a spatial resolution of 5° latitude and longitude; the domain used in the clustering was 30°W-20°E; 35°-70°N, covering most of Europe and the North Atlantic. Daily historic WP classifications are available from the EMULATE period (1850 to 2003), and have been extended from 2004 to the present using the European Centre for Medium-Range Weather Forecasts ERA-Interim data set (Dee *et al.*, 2011). Table 3.1 details a description of each WP and Figure 3.1 gives their definition according to MSLP anomalies. MO-8 was produced by repeatedly clustering patterns from MO-30 according to spatial correlation between the pairs of WPs. As a result, each WP in the smaller set comprises between one and seven WPs from the larger set. Table 3.2 contains descriptions of each WP in MO-8 and Figure 3.2 shows the MSLP anomaly maps.

Individual WPs will be referred to as WP*i*, with *i* indicating the WP number. Neal *et al.* (2016) ordered the WPs in both sets according to their annual historic occurrence in the period used for the clustering technique (1850 to 2003), with WP1 occurring most often and the last pattern

(WP30 or WP8) least often. The use of only MSLP anomalies in the clustering results in some seasonal grouping of WPs. This is particularly true for MO-30. Table 3.1 shows that lower-numbered WPs occur more often during the summer half-year (April through September; weak MSLP anomalies) and higher-numbered WPs occur more in winter (October through March; strong MSLP anomalies). Figure 3.3 shows the 11-year moving average frequency of each WP in MO-30. There is strong interannual variability for most WPs, and suggestions of a trend in several. Further discussion of WP frequencies and a comparison with long-term trends in LWT occurrences is in Section 3.6.

Each MO-30 WP is objectively classified as a LWT following the method in Jones *et al.* (1993); see Table 3.1 and Table 3.2. Using 16 grid points in the domain 20°W-10°E; 45°-65°N, values of wind flow and vorticity are calculated from daily MSLP data. LWTs are then defined using a set of rules related to the relative strength of these values. LWTs can be one of eight flow direction types (N, NE, E, SE, S, SW, W and NW) and/or one of anticyclonic (A) and cyclonic (C). Flow directions may be combined with cyclonic/anticyclonic types where appropriate. There is an additional 'unclassified' LWT to represent light indeterminate flow, denoted U, giving a total of 27 LWTs. Table 3.3 details the number of patterns from MO-30 assigned to each LWT. A discussion of this classification and comparisons with LWT frequencies is in Section 3.6. The LWT series between 1871 and 2015, derived using reanalysis products, is from Jones *et al.* (2013b). Historic occurrence of each LWT in this period is in Table 3.3.

UK daily and monthly precipitation data are from the Met Office Hadley Centre UK Precipitation data set (Alexander and Jones, 2000). The UK is split into nine regions (Figure 3.4) originally defined by Wigley *et al.* (1984): northeast England (NEE), central and east England (CEE), southeast England (SEE), southwest England and south Wales (SWE), northwest England and north Wales (NWE), east Scotland (ES), southwest and south Scotland (SS), northwest and north Scotland (NS) and Northern Ireland (NI). Regional precipitation series are available as daily totals from 1931 to the present and monthly totals from 1873 (for NEE, CEE, SEE, SWE and NWE) or 1931 (for the remaining regions). The data are estimates of the regional average calculated using at least seven evenly distributed precipitation gauging stations, see Alexander and Jones (2000) for details.

Weather pattern	Flow description	Historic occurrence (%)			LWT
		A	W	S	
1	Neutral north-westerly.	6.45	2.28	10.60	U
2	Cyclonic south-westerly.	5.63	3.16	8.09	SW
3	Anticyclonic south-westerly with a ridge over northern France.	5.26	2.59	7.92	ASW
4	Neutral westerly.	4.89	2.99	6.78	W
5	Neutral southerly with a centre of high pressure over Scandinavia.	4.86	2.85	6.86	S
6	Anticyclonic with a high pressure centre over the Azores.	4.88	3.22	6.53	A
7	Cyclonic south-westerly with a centre of low pressure west-north-west of Ireland.	4.92	2.70	7.13	U
8	Cyclonic westerly with a low pressure centre near the Shetland Islands.	4.62	3.20	6.04	C
9	Anticyclonic with high pressure over Iceland.	4.48	2.91	6.05	A
10	Anticyclonic westerly with high pressure over the Azores.	4.30	3.43	5.17	W
11	Cyclonic with a low pressure centre over southern Britain.	3.65	2.83	4.47	C
12	Anticyclonic south-south-westerly with high pressure over Poland.	3.60	4.23	2.97	SW
13	Anticyclonic north-westerly with high pressure south-west of Ireland.	3.49	4.03	2.95	NW
14	Cyclonic north-north-westerly with low pressure centred over southern Sweden.	3.07	3.87	2.27	NW
15	Neutral south-westerly. Very windy for northwest Britain.	3.02	4.39	1.65	SW
16	Anticyclonic south-south-easterly with high pressure near Denmark.	2.72	3.19	2.24	S
17	Anticyclonic east-south-easterly with a high pressure centre over Denmark.	2.57	3.99	1.16	AS
18	Anticyclonic south-westerly with high pressure over northern France.	2.59	4.38	0.81	ASW
19	Neutral northerly.	2.55	3.83	1.28	N
20	Cyclonic westerly with a deep centre of low pressure near Iceland. Very windy.	2.57	3.98	1.17	W
21	Cyclonic south-westerly with a deep low south of Iceland. Strong winds.	2.52	3.57	1.46	SW
22	Cyclonic southerly with low pressure west of Ireland.	2.28	3.15	1.42	S
23	Neutral westerly with high pressure north of Spain. Windy.	2.36	3.90	0.82	W
24	Cyclonic northerly with low pressure centred over the North Sea.	1.98	3.08	0.89	C
25	Anticyclonic northerly with high pressure over Northern Ireland.	2.08	3.37	0.80	A
26	Cyclonic north-westerly with low pressure centred near Norway. Very windy.	1.95	3.19	0.71	NW
27	Anticyclonic easterly with high pressure over the Norwegian Sea.	1.82	3.29	0.35	SE
28	Cyclonic south-easterly with low pressure southeast of the UK.	1.72	3.00	0.45	CSE
29	Cyclonic south-south-westerly with a deep low centre of pressure west of Ireland. Very windy.	1.64	2.73	0.55	C
30	Cyclonic west-north-westerly with deep low pressure southeast of Iceland. Very windy.	1.52	2.64	0.41	CW

Table 3.1 (previous page): For each weather pattern in MO-30, a description of the resultant flow over the UK, the historic occurrence (%) between 1850 and 2015 for all months (A), the winter half-year (W) and the summer half-year (S) and objectively assigned LWT class is listed.

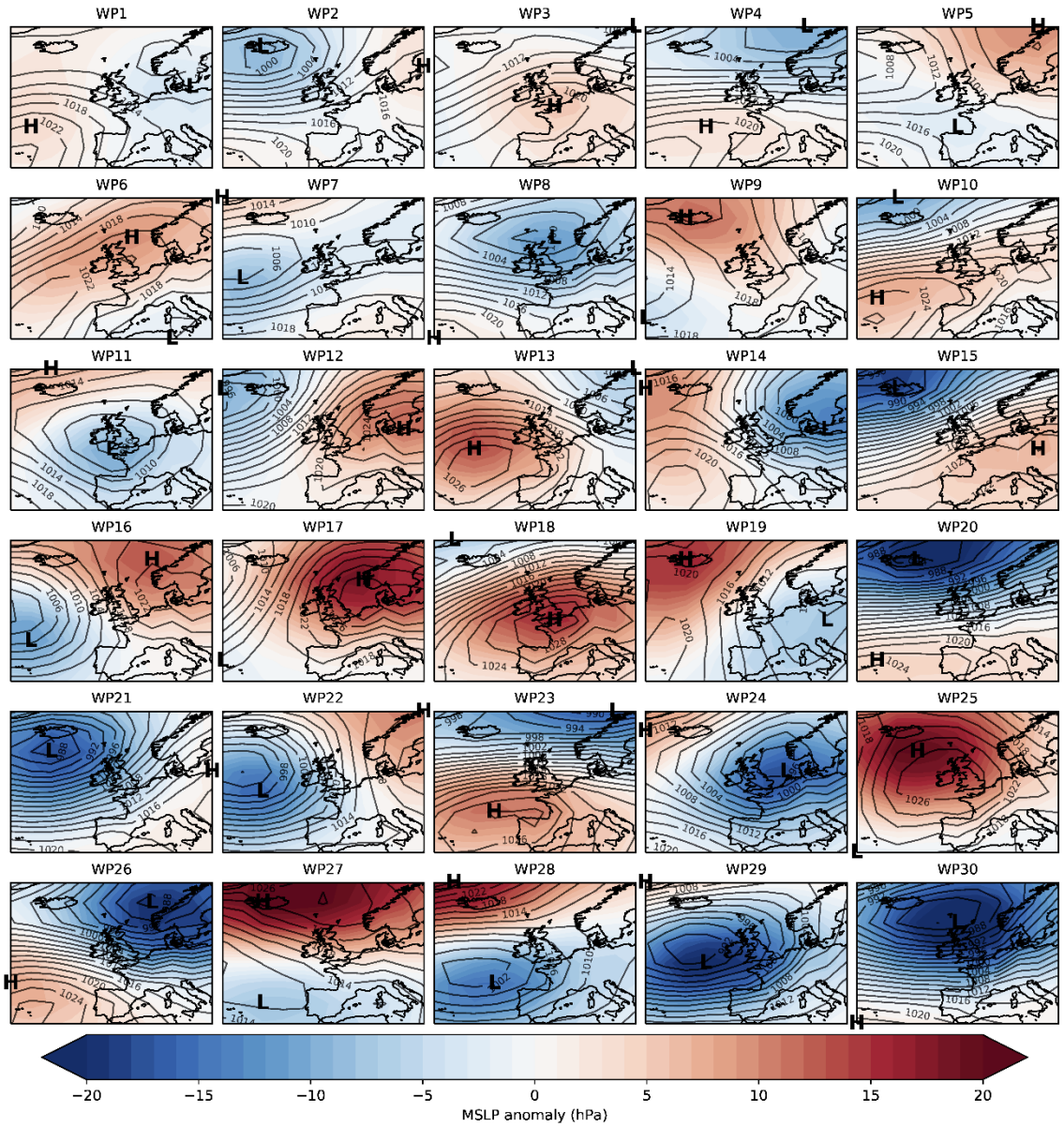

Figure 3.1: Definition of each weather pattern in MO-30. Red shading is for positive mean sea level pressure (MSLP) anomalies (hPa) and blue shading is for negative MSLP anomalies.

Weather pattern	Sub-patterns from MO-30	Flow description	Historic occurrence (%)			LWT
			A	W	S	
1	6, 9, 11, 19, 25, 27 and 28	Blocked, negative NAO pattern.	21.19	22.46	19.92	U
2	4, 8, 20, 23, 26 and 30	Zonal, positive NAO pattern.	17.92	19.91	15.93	W
3	1, 13, 14 and 24	Neutral north-westerly with low pressure northeast of the UK and high pressure to the southwest.	14.99	13.26	16.71	NW
4	2, 12, 15 and 21	Cyclonic south-westerly with low pressure centred near Iceland.	14.76	15.35	14.17	SW
5	5, 16, 17 and 22	Anticyclonic southerly with high pressure near Denmark and low pressure southwest of the UK.	12.43	13.18	11.69	S
6	3 and 18	Anticyclonic west-south-westerly with a centre of high pressure over northern France.	7.85	6.97	8.73	ASW
7	7 and 29	Cyclonic south-westerly with low pressure west of Ireland.	6.56	5.44	7.67	CSW
8	10	Anticyclonic westerly with high pressure over the Azores.	4.30	3.43	5.17	W

Table 3.2: As for Table 3.1, but for MO-8. Includes column detailing the patterns from MO-30 clustered into each MO-8 pattern.

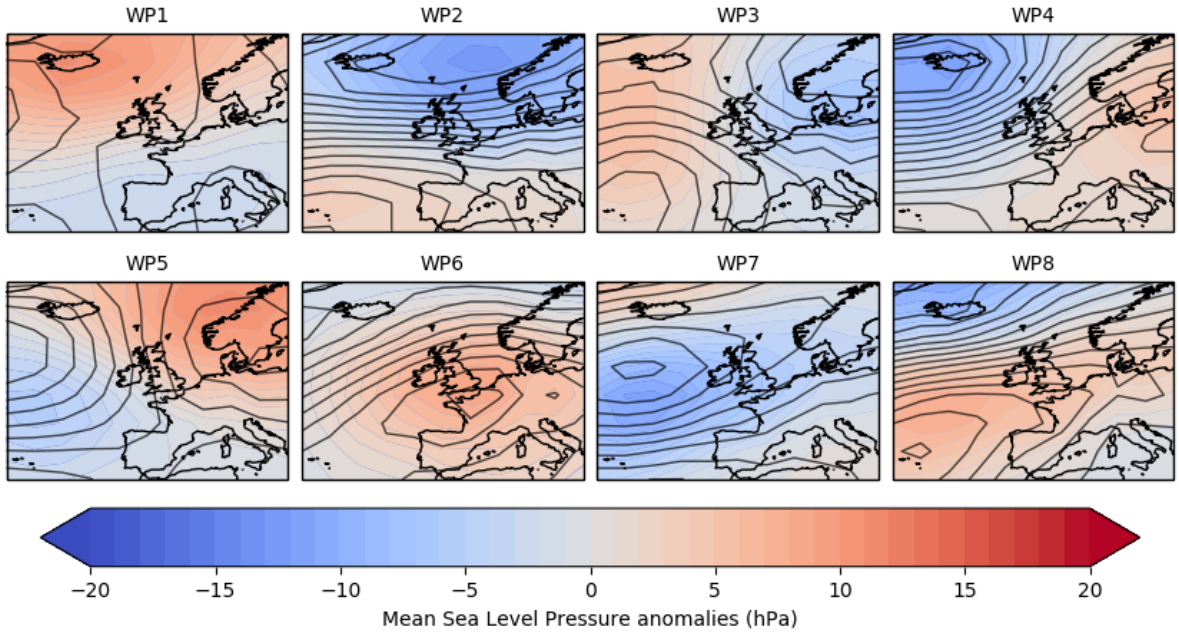

Figure 3.2: Definition of each weather pattern in MO-8. Sub-patterns from MO-30 are listed in parentheses. Red shading is for positive mean sea level pressure (MSLP) anomalies (hPa) and blue shading is for negative MSLP anomalies.

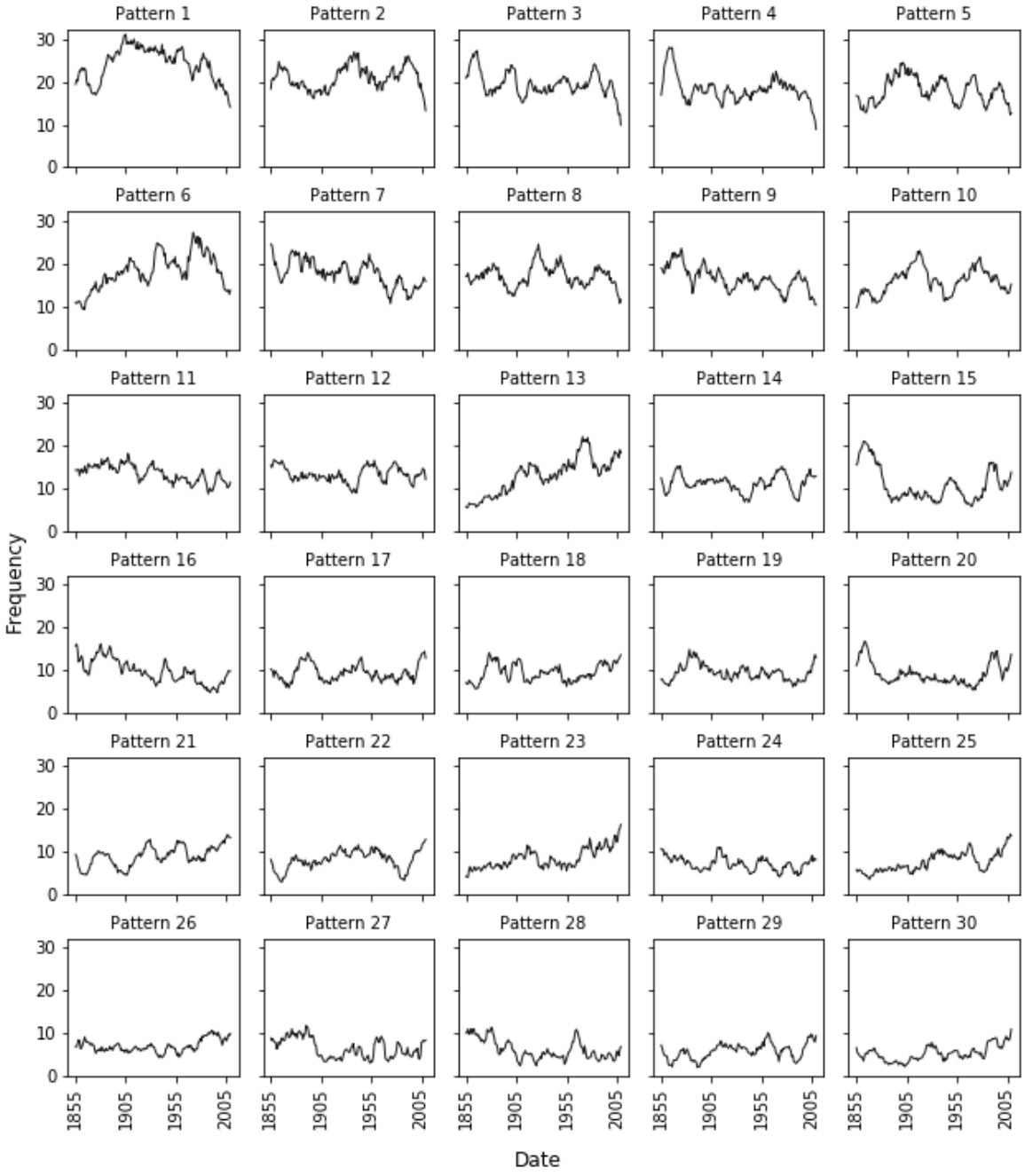

Figure 3.3: 11-year moving average frequencies of each weather pattern in MO-30. Dates represent the central year of each 11-year window.

LWT	Historic occurrence (%)	No. of patterns from MO-30
A	20.58	3
C	13.94	4
SW	9.45	4
W	9.02	4
S	5.64	3
NW	4.97	3
N	3.29	1
AW	3.20	0
SE	3.16	1
ASW	2.73	2
CSW	2.49	0
CW	2.24	1
CS	1.97	0
E	1.84	0
ANW	1.82	0
NE	1.74	0
AS	1.65	1
CNW	1.62	0
AN	1.19	0
ASE	1.17	0
CN	1.14	0
U	1.06	2
CSE	1.02	1
AE	1.00	0
ANE	0.80	0
CNE	0.67	0
CE	0.62	0

Table 3.3: LWT historic occurrence (%) between 1871 and 2015 and the number of weather patterns from MO-30 assigned to each LWT by the objective classification method.

3.3 Methodology

3.3.1 Associating daily precipitation with weather patterns and LWTs

Precipitation data between 1931 and 2015 are associated by WP, giving a distribution of daily precipitation totals for each WP. To allow comparison between regions, these data are divided by the regional mean daily precipitation. The result is a precipitation distribution for each WP expressed as a proportion of the regional average. The median and interquartile range (IQR) for each of these distributions are displayed in Figure 3.5 for MO-8 and Figure 3.6 for MO-30. Ideally, a WP classification would show distinct distributions for each WP. From Figure 3.5, it is clear that this is not the case for MO-8. Across all regions, half of the WPs (WP1, WP5, WP6 and WP8) have very similar median and IQR; WP2 and WP7 are also very similar. This is because the clustering process used to derive the WPs does not use precipitation but instead spatial correlation of MSLP anomalies. As a result, several WPs from MO-8 are composed of WPs from MO-30 that show very different precipitation distributions. For example, Table 3.2 shows that WP1 (a negative NAO-like pattern) contains a mixture of WPs featuring cyclonic and anticyclonic characteristics. Figure 3.6 shows that the WPs in MO-30 feature more distinct precipitation distributions, although there remain several subsets that appear similar (e.g. WP13 and WP3 or WP4 and WP2). The precipitation distributions of each WP in MO-8 are considered too similar to warrant further analysis in the context of precipitation climatology, so this set is excluded from the remainder of the analysis.

The same method is applied to the LWTs between 1871 and 2015, with median and IQR of daily precipitation associated with each LWT shown in Figure 3.7. For many of the wet LWTs, the IQR is much higher than for wet patterns in MO-30. In particular, the daily precipitation variability associated with cyclonic easterly LWTs (CE, CNE and CSE) in eastern England is large. The lower variability of precipitation associated with WPs in MO-30 compared to with LWTs suggests the former is better suited to precipitation analyses in the UK.

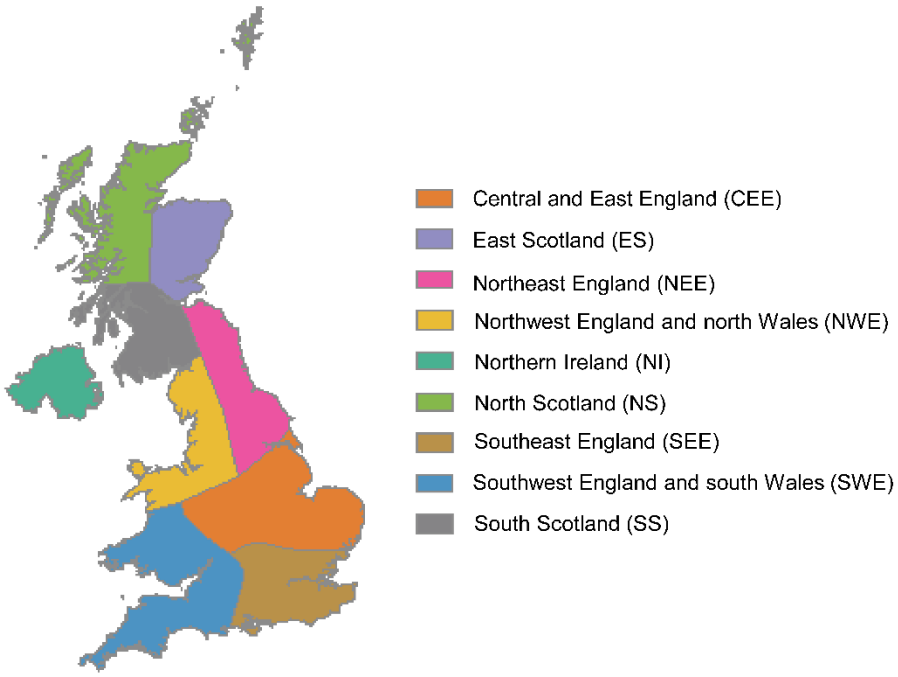

Figure 3.4: Regional boundaries of the HadUKP precipitation data set.

3.3.2 Linking MO-30 frequencies with the SPI

UK monthly precipitation climatology in relation to MO-30 is now analysed using the SPI. Developed originally for drought applications, the SPI is equally valid for wet periods. The SPI is calculated by fitting a parametric probability density function to precipitation data and transforming it to the standard normal distribution (see Lloyd-Hughes and Saunders (2002) for full details). Standardisation means the SPI is comparable across different regions. Negative (positive) SPI values signify the degree of dryness (wetness). Typically, monthly precipitation data used in SPI calculation is first aggregated over some time-scale in months, k, resulting in an SPI series that represents the degree of dryness/wetness over a chosen period, denoted SPI-k. For example, SPI-3 might be used to describe meteorological or agricultural drought conditions, whilst SPI-12 is more suitable for describing long-term hydrological drought (e.g. McKee *et al.*, 1993; McKee *et al.*, 1995). There is ongoing debate about which statistical

distribution should be fitted to monthly precipitation data. McKee *et al.* (1993) originally used the gamma distribution; the Pearson type III (Guttman, 1999) and Weibull (Sienz *et al.*, 2012) have also been suggested. The choice of distribution is important, as it has a significant effect on the resultant SPI values (Sienz *et al.*, 2012; Guerreiro *et al.*, 2017).

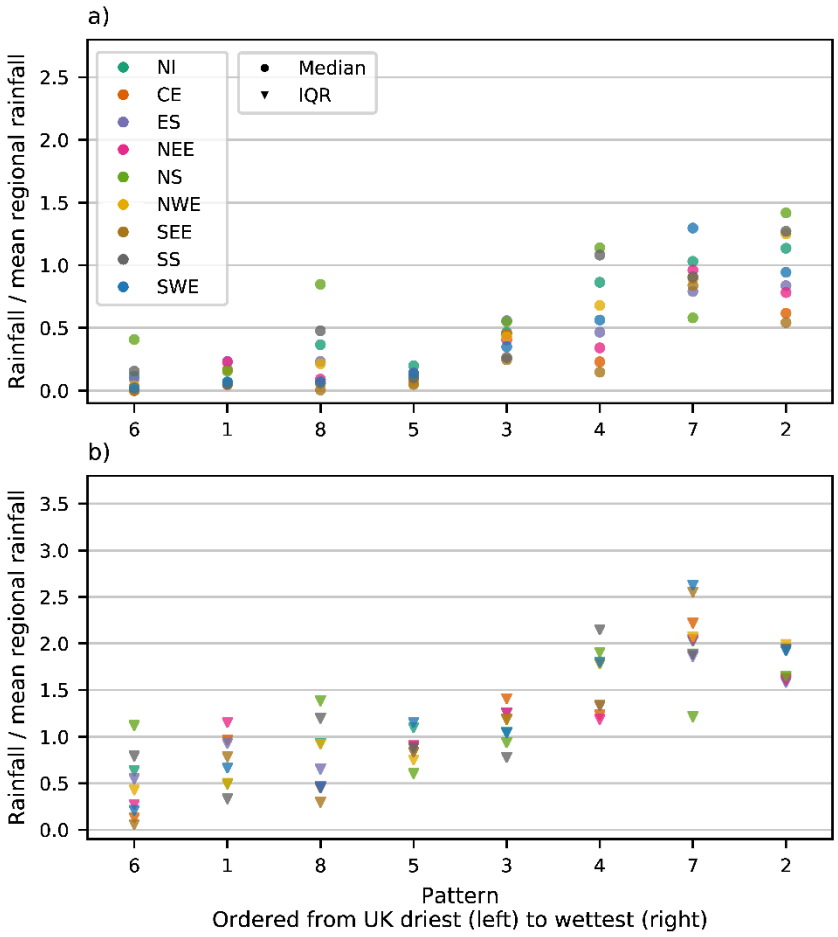

Figure 3.5: a) Median and b) interquartile range (IQR) of daily precipitation 1931-2015 for each region and each pattern in MO-8, expressed as the proportion of precipitation relative to the regional average. Weather patterns are ordered left to right from the lowest UK mean precipitation (i.e. averaged across all regions) to the highest.

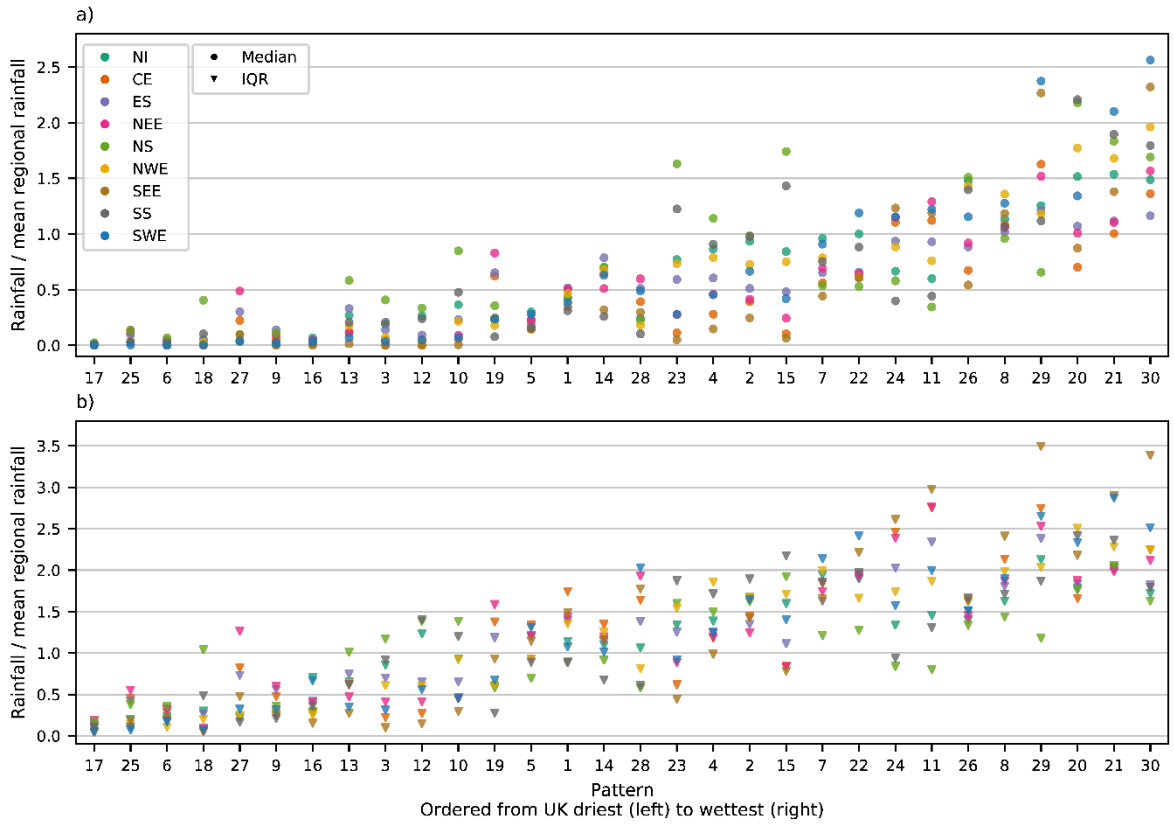

Figure 3.6: As Figure 3.5 but for each WP in MO-30.

SPI series for $k = 3, 6$ and 12 are calculated for each region. To maintain comparability across regions, the distributions are fitted using the common period, 1931 to 2015. The distributional fit of seven parametric distributions (the gamma, Weibull, exponential, Pearson type III, generalised extreme value, Gumbel and normal distributions) is assessed on each precipitation series using a modified Mean Square Error (MSE) metric defined by Papalexiou *et al.* (2013) and used for SPI distribution fitting by Guerreiro *et al.* (2017). This modified metric is advantageous compared to the classical MSE because it gives as much weight to values in the tails of the distribution as those in the middle. This is important for assessing extreme values. The modified MSE is defined as:

$$\text{MSE} = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{F(x_i)}{F_N(x_i)} - 1\right)^2,$$

Equation 3.1

where N is the sample size, $F(x)$ is the probability density function given any theoretical distribution and $F_N(x_i)$ is the probability of exceeding x_i using the empirical distribution

$$F_N(x_i) = \frac{r(x_i)}{N+1},$$

Equation 3.2

where $r(x_i)$ is the rank of x_i. The best-fitting distribution is that which yields the lowest MSE. For each aggregated precipitation series, the MSE is calculated monthly and then summed to give an estimate of the annual MSE. The lowest annual MSE scores for the majority of the series are obtained by fitting the Pearson type III distribution, with the gamma and Weibull distributions also represented. The chosen (series-specific) distributions are fitted to the precipitation series using maximum likelihood estimation and transformed to the standard normal distribution, yielding the SPI series.

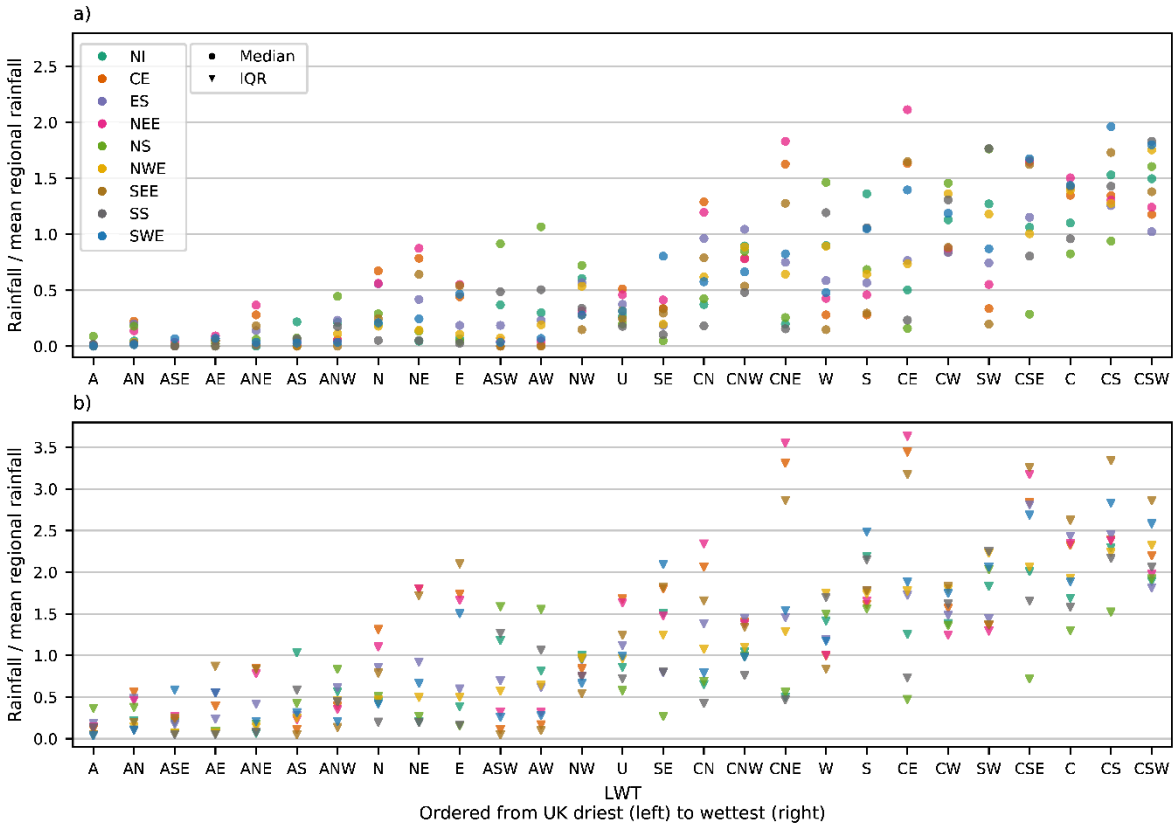

Figure 3.7: As Figure 3.5 but for LWTs.

The daily WP series is aggregated to monthly frequencies for each WP in MO-30. By examining WP frequencies during anomalously dry or wet periods it is possible to ascertain which are associated with these conditions. Here, "dry" and "wet" periods are defined as months where SPI-$k \leq -1$ and SPI-$k \geq 1$, respectively, corresponding to roughly 16% probability at each tail of the distribution. As SPI values are calculated based on precipitation aggregated over k months, WP frequencies are calculated over the same time-scales. Therefore, $k = 3, 6$ and 12 –monthly summed frequencies are calculated for each WP. The mean of these k-monthly frequencies during dry/wet periods is divided by the mean of the k-monthly frequencies over the entire record, giving the frequency anomalies of each WP during dry/wet periods. This is

then expressed as a percentage anomaly (PA). For a given WP and time-scale, PA is defined as:

$$PA = 100\left(\frac{\bar{y}}{\bar{x}} - 1\right),$$

Equation 3.3

where $\bar{x}$ is the mean frequency of the WP over the whole record and $\bar{y}$ is the mean frequency of the WP during dry/wet periods. For $PA = 0$, the average frequency of a particular WP during dry/wet periods is the same as the average over the entire record. A negative PA implies the WP occurs less frequently than normal, and vice versa for a positive PA. To enhance the readability of this chapter, PAs are referred to simply as "anomalies". The significance of the anomalies are assessed using Welch's *t*-test (Welch, 1947), with the null hypothesis that the means of the two samples (frequencies during dry/wet periods and frequencies over the remainder of the record) are the same. An assumption of this test is that the data follow a normal distribution. However, as the frequencies are count data exhibiting heavily skewed behaviour, its suitability must be assessed by checking whether nominal significance is preserved. This is done by repeatedly resampling the frequency data and performing the test on each sample with significance level $\alpha = 0.05$. The type I error rate (incorrect rejection of a true null hypothesis) is only slightly inflated above nominal significance (typically around 0.058) in a few cases; this is deemed adequate for the purpose of the study. The method described in this section is used for the winter and summer half-years (October through March and April through September, respectively) as well as annually.

For brevity, only selected results are presented. An anomaly for the dry/wet criterion is typically complemented by an anomaly of the opposite sign and roughly equal magnitude for the other criterion. As the focus is on droughts, SPI anomalies are presented for dry periods, with the corresponding wet period anomalies shown in the Appendix. Furthermore, results are not presented for SPI-6 and SPI-12. This is because they are very similar to SPI-3 results, although the anomaly magnitudes typically reduce as k increases. This is logical, as by extending the period over which monthly frequencies are considered, the likelihood of a larger range of WPs occurring increases. That is, a longer-term dry (wet) spell is more likely to contain a greater number of wet (dry) WP occurrences than for a shorter-term spell. A further reason for omitting SPI-6 and SPI-12 results is that, as winter and summer are considered as half-years, the WP frequencies for SPI-6 and SPI-12 include information from the preceding season. This is generally not the case for SPI-3, with much of the averaging taking place only in the considered season.

3.3.3 Linking MO-30 frequencies with the DSI

For drought-specific applications, the DSI is sometimes used. Based on cumulative precipitation deficits, rather than the k-monthly "snapshot" nature of the SPI, it is more appropriate for analysing the evolution and accumulated intensity of a drought. Furthermore, the DSI is non-parametric and does not suffer from the SPI's requirement of finding a suitable distribution to fit to the underlying data. Whilst the DSI is preferred here for drought analysis, the SPI is used in the previous section to assess wet, as well as dry periods, in which context it is more appropriate than the DSI. As with the SPI, the DSI can be calculated for different time-scales and is denoted DSI-k. The DSI calculation procedure is as follows:

1. For some time-scale, k, let the precipitation anomaly in month t be X_t. If $X_t < 0$ (i.e. precipitation is below the mean), and precipitation in the k-monthly period $X_t, X_{t-1}, \dots, X_{t-k+1}$ is also below its k-monthly mean then initiate a drought sequence in month t and set DSI-$k_t = -X_t$.

2. For the next month, $t+1$, the precipitation anomaly is X_{t+1}. Then DSI-$k_{t+1} = -X_{t+1} +$ DSI-k_t, if and only if the k-monthly mean of $X_{t+1}, X_t, \dots, X_{t-k+2}$ is not exceeded. If the k-monthly mean is exceeded then DSI-k_{t+1} is set to zero and the drought is terminated. This step is repeated for the entire precipitation series.

3. The DSI-k series is standardised by dividing the absolute deficit (in mm) by the mean annual precipitation and multiplying by 100. The index then expresses the cumulative precipitation deficit as a percentage of mean annual precipitation.

Note that the DSI can be negative when a precipitation surplus occurs over a time-scale shorter than k. The DSI has been used for assessing historical and projected drought conditions in the UK (Phillips and McGregor, 1998; Fowler and Kilsby, 2002b; Blenkinsop and Fowler, 2007a; Rahiz and New, 2012; Rahiz and New, 2013; Rahiz and New, 2014), Iberia (Guerreiro *et al.*, 2017) and Europe-wide (Blenkinsop and Fowler, 2007b).

For each region, DSI series are calculated using the same reference period as for the SPI (1931 to 2015). For each series, a threshold is selected such that roughly 5% of values are above the threshold. The months that these values correspond to are named "drought months". Some previous studies using DSI have selected arbitrary thresholds and defined a drought as when DSI exceeds this threshold over multiple locations (Phillips and McGregor, 1998; Fowler and Kilsby, 2002b). The regions used in this study, however, are large enough for a drought to occur in one region but not any other, justifying the use of different thresholds for different regions.

As for SPI dry and wet periods, WP *PA*s are calculated for drought months defined by the DSI series. However, Welch's *t*-test is not suitable in this case as nominal significance is not preserved. This is probably due to the sample size of drought months being far smaller than that of SPI wet/dry periods, leading to the test's assumptions being more easily violated. Other tests, such as the Mann-Whitney U-test (Mann and Whitney, 1947) and permutation *t*-test were tried with similar results. Therefore, DSI *PA*s are reported without testing for statistical significance. As before, the results are split into annual, winter and summer seasons. Results are presented for DSI-3, DSI-6 and DSI-12 to enable comparison of anomalies between droughts of different lengths. Winter and summer results are not shown, as they are qualitatively the same as for annual, particularly for DSI-6 and DSI-12.

3.4 Results: MO-30 frequency anomalies during SPI wet and dry periods

3.4.1 Annual

Figure 3.8 shows the anomalies of 3-monthly mean frequencies during annual SPI-3 dry periods for each region. The two WPs that occur statistically significantly more frequently during dry periods compared to normal across all regions are both anticyclonic LWT variants – WP6 and WP17. For wet periods, Figure A.1 shows that three cyclonic or westerly variants (WP8, WP21 and WP30) occur significantly more frequently than normal for all regions. Differences in anomalies between eastern and western regions are apparent. From Figure 3.8, western regions generally see a larger increase than eastern regions in the occurrence of WP9 and WP27 during dry periods compared to normal. WP27 is a south-easterly LWT so any precipitation associated with this pattern would mostly fall on East Britain. WP9 is anticyclonic, although it is hard to discern from Figure 3.1 why it is linked with more precipitation in western regions as there is no indication of flow direction. Dry periods in most eastern (northern and western) regions are associated with an increase (decrease) in the frequency of two WPs that would bring strong westerly winds over northern Britain - WP15 and WP23. From Figure A.1, the opposite effect of these regional differences is apparent for wet SPI periods. Furthermore, western regions are associated with larger decreases in the occurrence of three windy, westerly patterns (WP20, WP21 and WP26) during dry periods than eastern regions (Figure 3.8). Dry periods in NEE, CEE, SEE and SWE are associated with a statistically significant decrease in the occurrence of WP28, with significant increases for NS and SS.

Results in NS differ to the other regions (Figure 3.8). In general, the anomalies are greater in magnitude. Also, many WPs that occur more frequently than normal during dry/wet periods in NS occur less often than the average for other regions, and vice versa. For example, the

frequency of WP29 (a cyclonic LWT) is above average for NS yet is below average for all other regions except SS (for which it is near normal). The MSLP anomaly definition in Figure 3.1 implies strong south-south-westerly winds over the UK, so precipitation brought over the UK would fall heavily on other regions first, become moderate over SS before turning dry in NS. The differences between NS and other regions may partly be due to its location on the northern tip of the UK. It is exposed to both western and eastern coastlines, and so is on the front line of north-westerly, northerly and north-easterly winds.

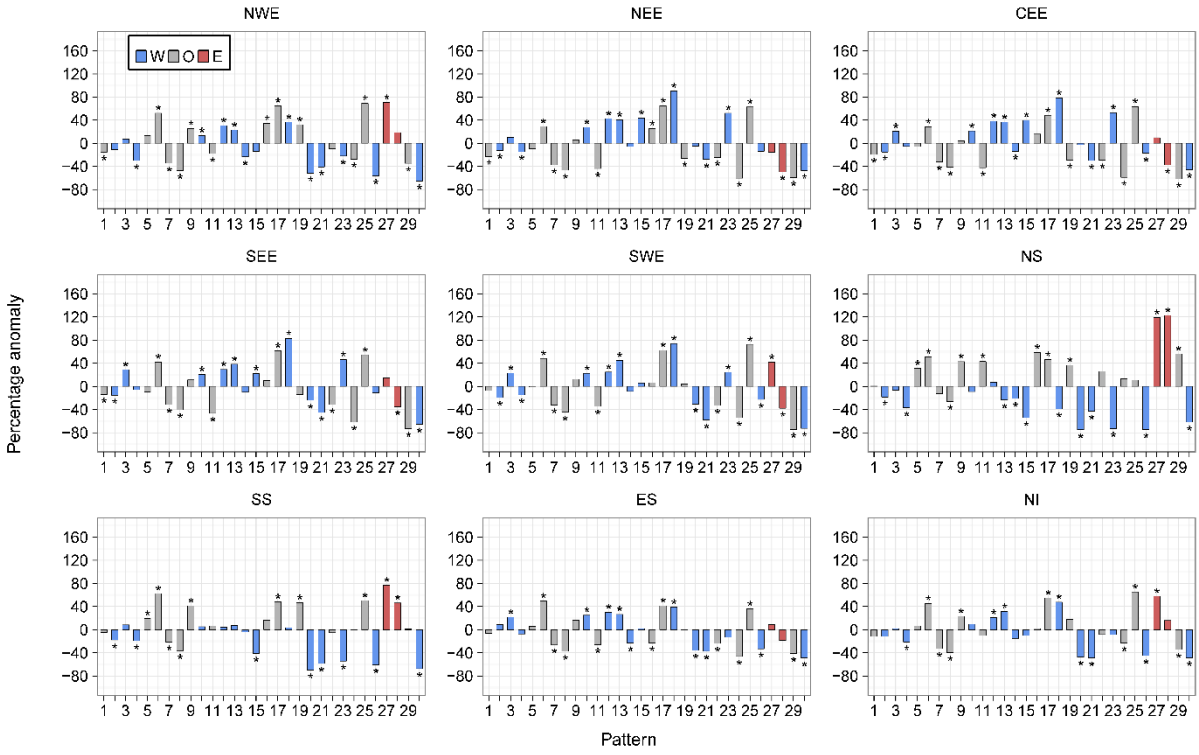

Figure 3.8: Annual (i.e. all months) three monthly mean frequency percentage anomalies of each weather pattern in MO-30 during dry periods defined by SPI-3 ≤-1. Blue and red bars indicate that the weather pattern contains a westerly (W) or easterly (E) component in its LWT equivalent, respectively. Grey bars represent all other types (O). An asterisk indicates statistical significance at the 95% level.

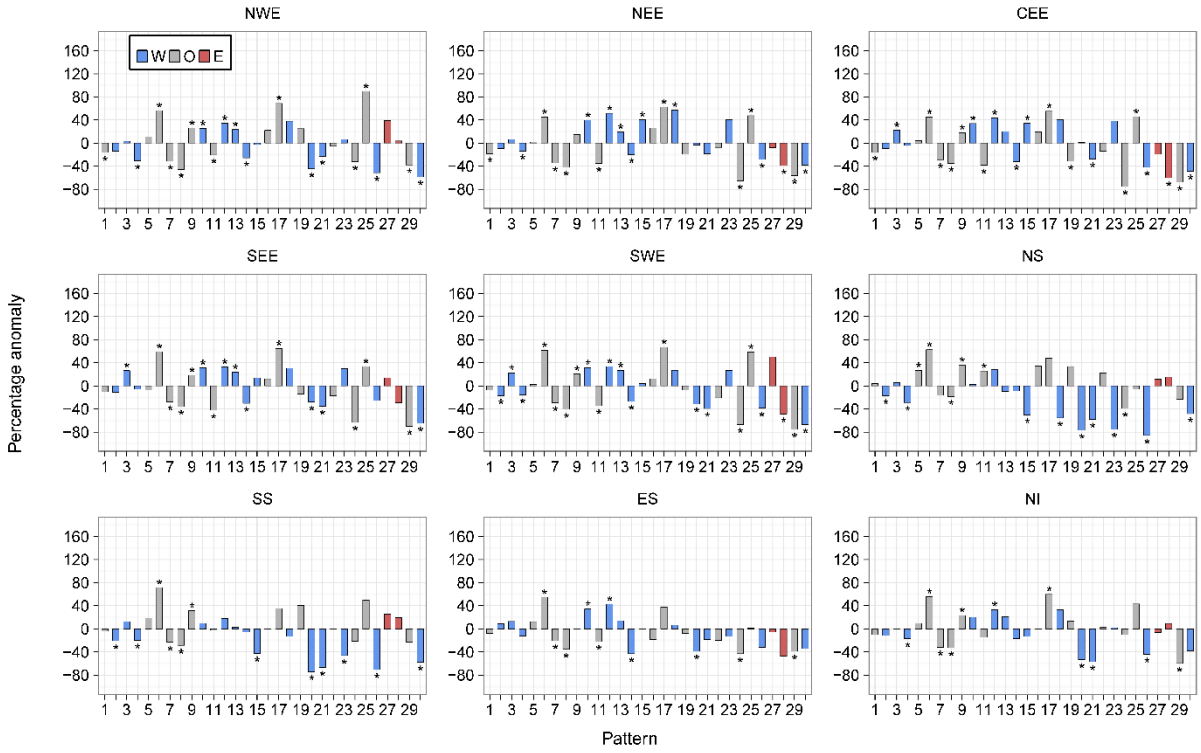

Figure 3.9: As Figure 3.8 but for summer.

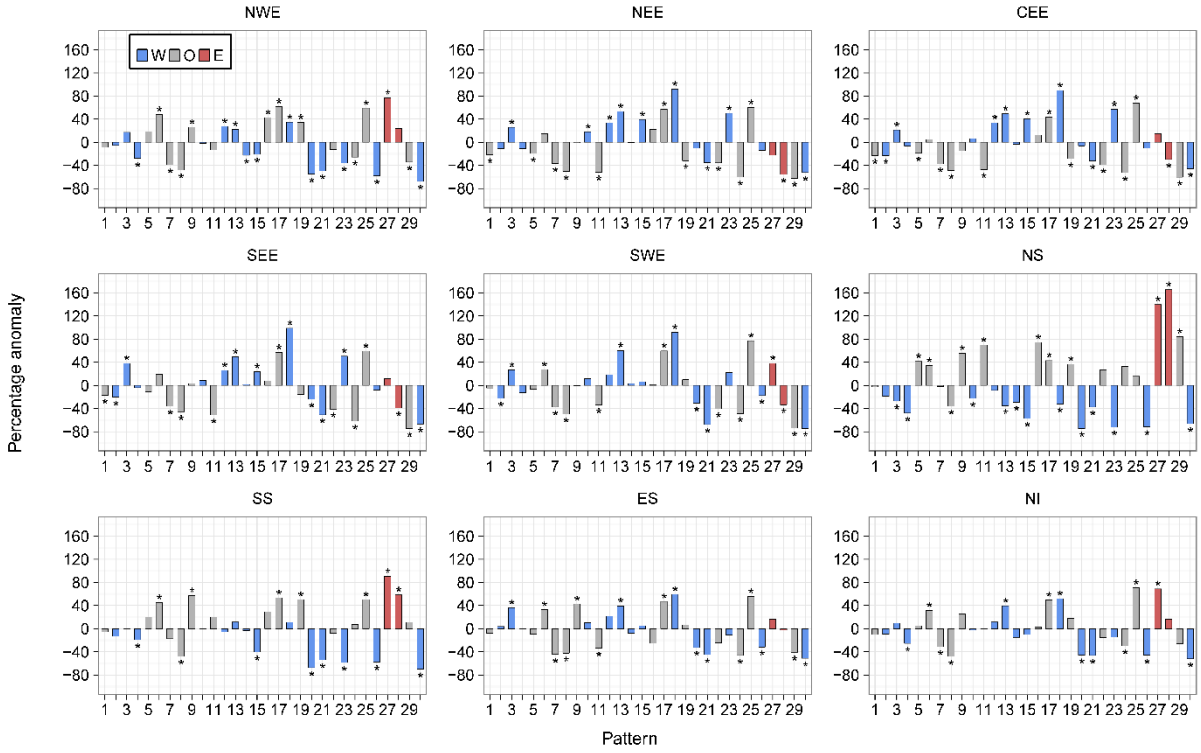

Figure 3.10: As Figure 3.8 but for winter.

3.4.2 Winter and summer

Neal *et al.* (2016) suggested that knowledge of the seasonal behaviour of the WPs could give an indication of an extreme weather event. Higher-numbered WPs occur more in winter than summer, so the occurrence of, say, WP30 in summer may indicate an extreme event on that day. This idea is extended here to monthly data by considering *PA*s calculated with respect to the seasonal average for winter and summer half-years.

Recall from Table 3.1 that lower- (higher-) numbered WPs are more associated with summer (winter). Figure 3.9 shows that for some regions (particularly those in Scotland, but not NWE), summer SPI-3 dry periods occur as a result of a decrease in the frequency of wet WPs that are associated with winter (for example WP20, 21 and 30), together with an increase in the occurrence of dry WPs that are associated with summer (particularly WP6). Winter SPI-3 dry periods generally see larger changes in the occurrence of WPs that are associated with winter compared to the WPs that are associated with summer. This is shown in Figure 3.10, with higher-numbered WPs tending to have higher-magnitude anomalies than lower-numbered WPs (except in ES). Wet SPI-3 periods in summer are characterised by an increase in the occurrence of wintery, wet WPs more than a decrease in the occurrence of dry WPs (Figure A.2). Figure A.3 shows that for wet periods in winter, a greater number of WPs feature strong anomalies than in summer. For example, in NWE, six WPs show non-significant anomalies in winter (Figure A.3) compared to 11 WPs in summer (Figure A.2).

3.5 Results: Defining drought months and MO-30 frequency anomalies

3.5.1 Identifying drought periods using DSI

Figure 3.11 displays the DSI-3, DSI-6 and DSI-12 series for NWE and NEE. As the time-scale increases, the threshold increases and drought months (black bars in Figure 3.11) become clustered together; droughts become more intense, less frequent and longer in duration. This is consistent with results from other studies (e.g. Phillips and McGregor, 1998; Fowler and Kilsby, 2002b). Results for other regions feature the same behaviour (not shown). Notable droughts can be identified from Figure 3.11. The 1995-96 drought is clearly visible in both regions. In NWE, for DSI-3 and DSI-6 this is the most intense drought in the record from 1873. Although still a major drought in the DSI-12 series, it is matched in intensity by several other episodes such as the 1975-76 drought and part of the long drought between 1890 and 1910 (see Marsh *et al.*, 2007). For DSI-3 in NEE there are many separate drought episodes of relatively low intensity. Other regions also feature a high frequency of droughts for this time-scale, although they are generally less intense for northern and eastern regions. A notable feature of Figure 3.11 is how

long and intense the mid-1970s drought was for NEE when considering DSI-12 (right column, third row). The DSI-12 value of almost 100 is unmatched in any other region or at any other time-scale and this drought accounts for the majority of the DSI values above the threshold (for this region and time-scale). The more recent 2010-12 drought is not evident in NWE or NEE but is for CEE, SEE and SWE regions (not shown). This is consistent with the exaggerated northwest-southeast precipitation gradient observed during this period (Kendon *et al.*, 2013).

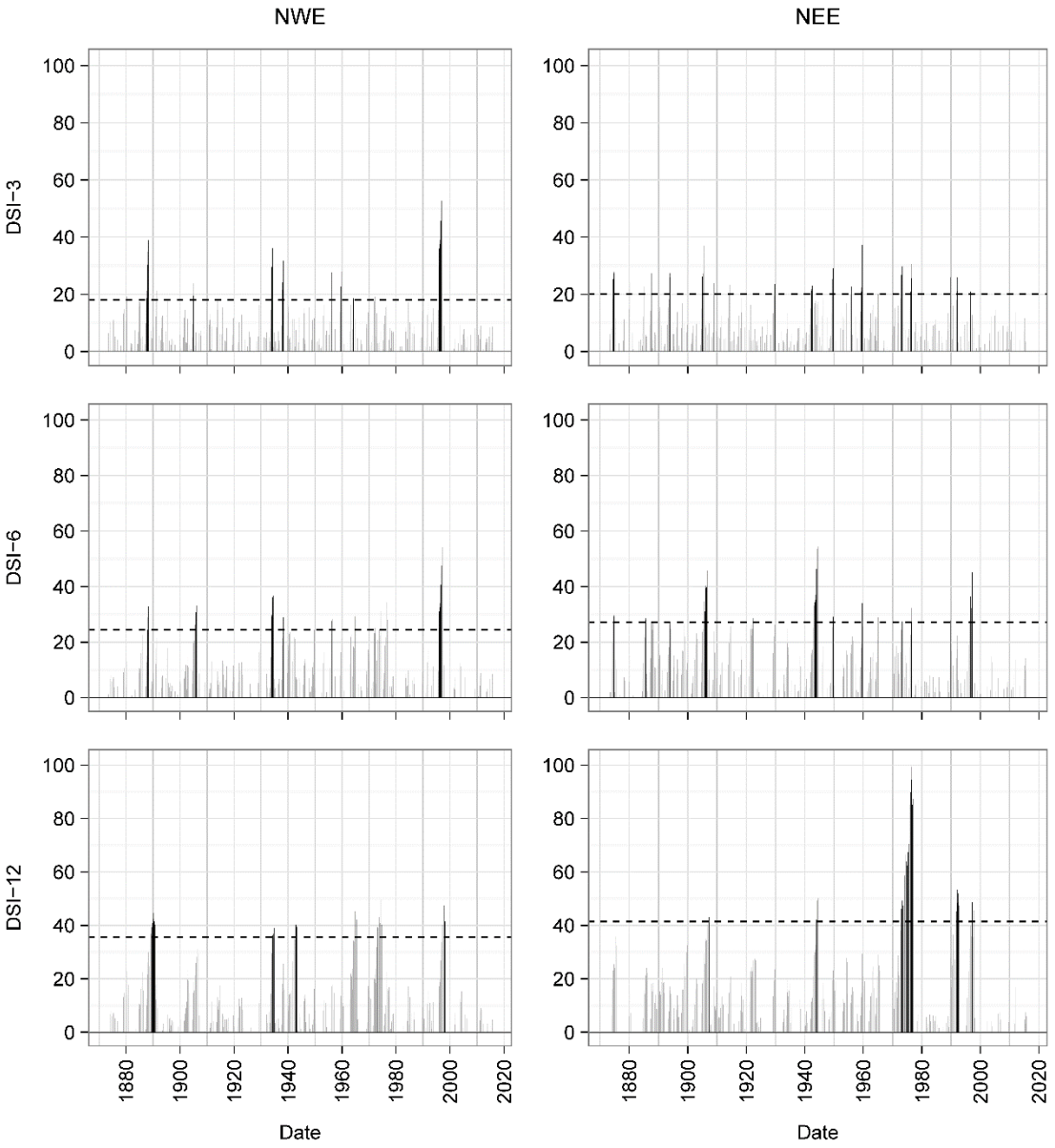

Figure 3.11: DSI series for NWE (left column) and NEE (right column) 1883-2015 indicated by grey bars. First row is DSI-3, second row is DSI-6 and third row is DSI-12. Drought months are represented by the black bars. The dashed horizontal line indicates the threshold for which DSI values are considered drought months.

3.5.2 MO-30 frequency anomalies during drought months

Figure 3.12 shows the results for annual three-monthly mean frequency anomalies for each region. WPs that are associated with dry (wet) conditions defined by SPI typically occur more (less) frequently during drought months than normal. DSI drought months represent extreme dryness better than SPI dry periods, as they account for 5% of each series, compared to 16% for SPI. The smaller-magnitude anomalies of lower-numbered WPs (e.g. WP6 through WP9 in Figure 3.12) during drought months compared to SPI wet/dry periods (Figure 3.8) implies that greater-intensity droughts are characterised by an increase or decrease in frequency of those WPs that occur less often annually (i.e. the higher-numbered WPs).

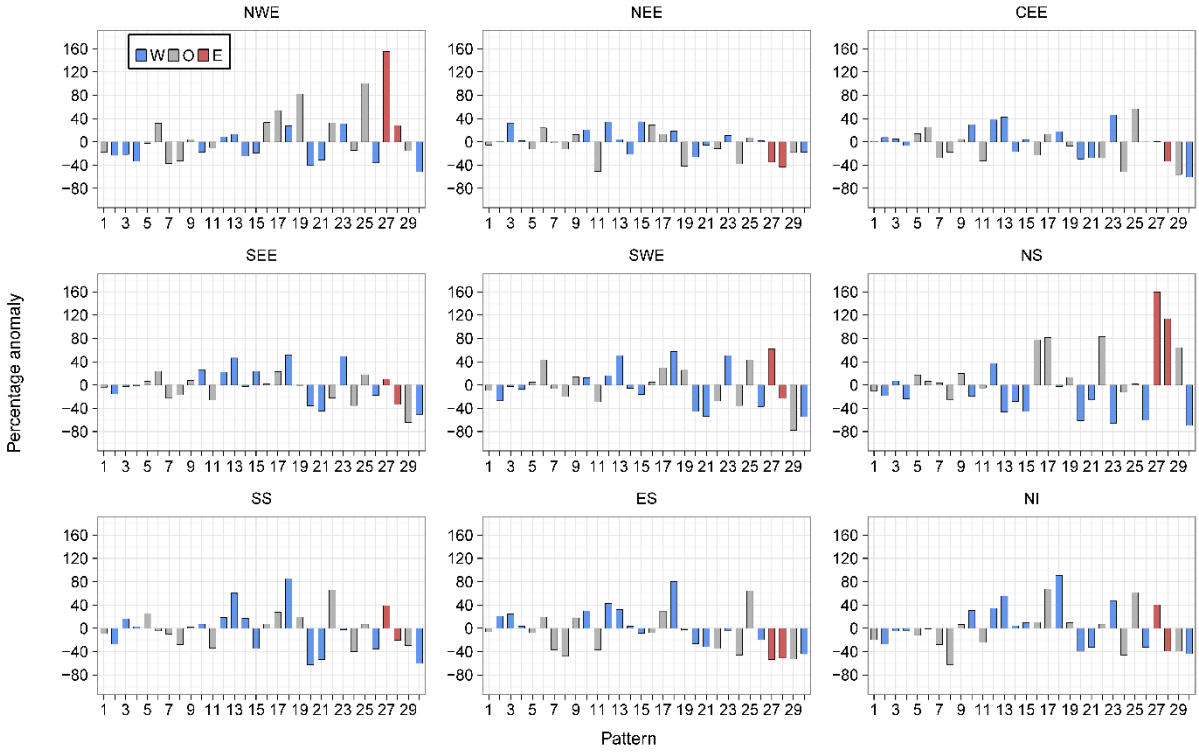

Figure 3.12: Annual (i.e. all months) three-monthly mean frequency percentage anomalies of each weather pattern in MO-30 during drought months defined by DSI-3. Blue and red bars indicate that the weather pattern contains a westerly (W) or easterly (E) component in its LWT equivalent, respectively. Grey bars represent all other types (O).

The results for WP17 (an AS LWT) are surprising. It has the lowest UK mean daily precipitation between 1931 and 2015 (Figure 3.6) and occurs significantly more frequently during SPI-3 dry periods than normal for all regions (Figure 3.8). However, it only occurs far more often than normal during DSI-3 drought months for NWE, NS and NI (Figure 3.12). The regional differences in anomalies for DSI-3 (Figure 3.12) correspond to those for SPI-3 wet and dry periods. WP27 occurs more (less) frequently than normal during drought months in western

(eastern) regions. This WP is particularly associated with droughts in NWE and NS, occurring around 160% more than normal. WP15 and WP23 (westerly variants) are not as strongly associated with DSI drought months as they are for SPI-3 dry periods in eastern regions. This is perhaps because neither WP is dominated by anticyclonic conditions; both would bring strong winds over the UK. Drought in NWE, NS and NI is characterised by larger increases in the occurrence of dry WP rather than decreases in the occurrence of wet WP. The opposite is true for NEE and SEE, where droughts appear more associated with a decrease in wet WPs (e.g. WP29 and WP30 for SEE and WP11 for NEE).

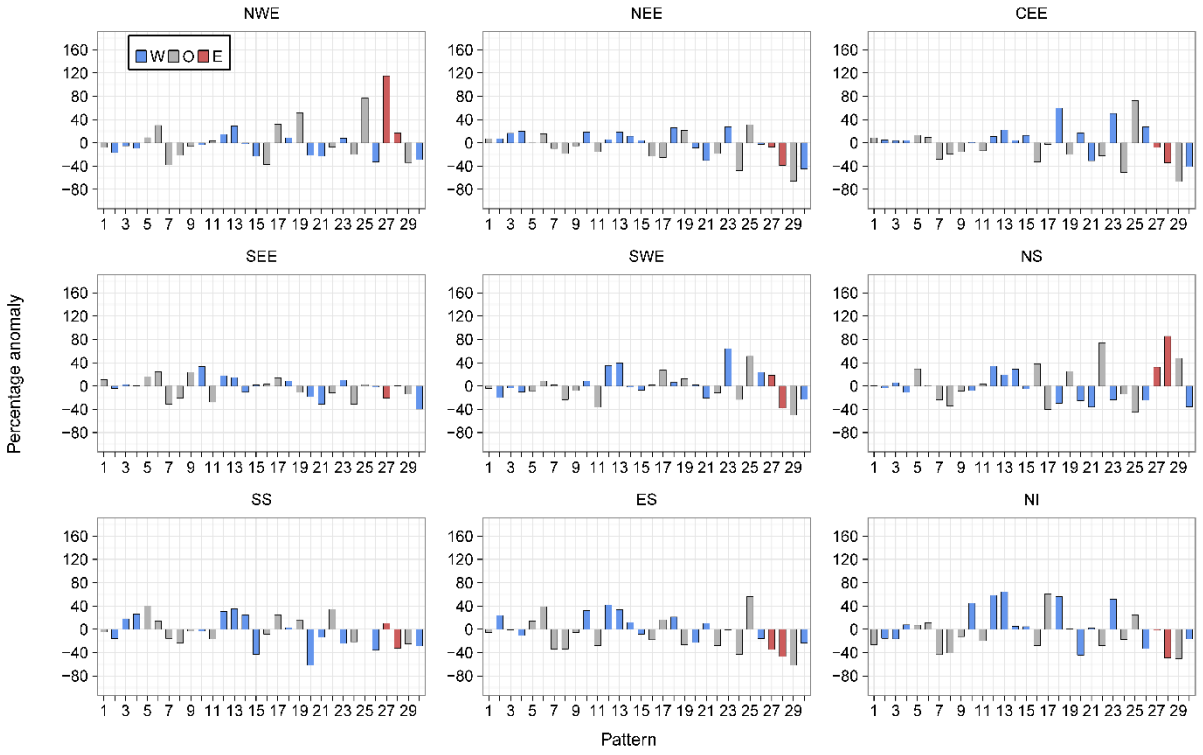

Figure 3.13: As Figure 3.12 but for DSI-6.

Unlike SPI, increasing the time-scale for DSI does not always yield lower-magnitude anomalies. NEE displays some of the lowest magnitude anomalies for DSI-3 (Figure 3.12), with similar results for DSI-6 (Figure 3.13). When considering 12-month droughts, however, anomaly magnitudes for this region increase to some of the greatest of all regions (Figure 3.14). Conversely, NWE features some of the largest positive anomalies for DSI-3 (Figure 3.12) and DSI-6 (Figure 3.13) compared to other regions, with droughts characterised by strong increases in the occurrence of WP19, WP25 and WP27. For DSI-12, however, it is less clear which WPs cause drought in NWE, except perhaps a strong decrease in the occurrence of WP29 (Figure 3.14). Also of note is how, in some regions, DSI-12 droughts are characterised by a smaller set of WPs than drought at smaller time-scales. For example, in CEE, Figures 3.12 and 3.13 show DSI-3 and DSI-6 droughts are attributed to increases and decreases in a number of patterns'

frequencies. By contrast, Figure 3.14 indicates DSI-12 droughts are typified by a strong increase in the occurrence of WP25, followed by the decrease in occurrence of WP24 and WP29.

3.6 Discussion

3.6.1 MO-30 and LWT frequencies

A significant feature of MO-30 frequencies of occurrences (Figure 3.3) is that, over the last 30 years of the record, days featuring stronger MSLP anomalies (WP12 through WP30 days) have become more frequent at the expense of days where MSLP anomalies are weaker (WP1 through WP11 days). These two groups of WPs are split seasonally, with WP1 through WP11 occurring more often during summer than winter and vice versa for the remaining patterns. This increase in the amount of "wintery" days featuring strong highs or deep lows might imply a rise in extreme weather events. Research suggests there has been an increase in precipitation intensity and floods in the UK since the 1960s (e.g. Osborn and Hulme, 2002; Fowler and Kilsby, 2003; Pattison and Lane, 2012; Jones *et al.*, 2013a; Jones *et al.*, 2014; Foulds and Macklin, 2016), so further work investigating these changes in relation to the changes in WP frequencies may be valuable.

Changes in WP frequencies are also comparable to changes in frequencies of LWTs derived from reanalysis products. Jones *et al.* (2013b) highlighted the most pronounced LWT trends as a decline in easterly LWTs between 1871 and the 1920s, and an increase in frequency of north-westerly LWTs between 1871 and 2010. Correspondingly, the only patterns in MO-30 with easterly flow components (according to their LWT classification), WP27 and WP28, also decrease in frequency until the 1920s. The north-westerly LWT changes are matched by a strong increase in the occurrence of WP13. The other north-westerly variants, WP14 and WP26, show less suggestion of an upward trend, implying the increase in NW LWTs corresponds almost totally to the increase in WP13. Additionally, Jones *et al.* (2013b) demonstrated that there was a slight increase in the frequency of westerly LWTs between 1871 and 2010. Of the four WPs in MO-30 assigned as westerly LWTs, two show similar behaviour (WP10 and WP23), one shows the opposite (WP4) and one shows no change (WP20) over the same period. It is important to note that changes in observation density in the original MSLP products (i.e. the reanalysis data used in the MO-30 and LWT derivations) may introduce artificial change points or trends in pattern frequencies. The EMULATE data set used for MO-30 consists of land station and ship observations in the period 1850-1881, with data beyond 1881 blended with a pre-existing data set. Prior to analysis, standard change point tests were performed on MO-8 and MO-30 frequencies, with no significant results found for this year. Therefore, trends

and changes points reported in this chapter are unlikely to be due to changes in the observation network in the original MSLP data.

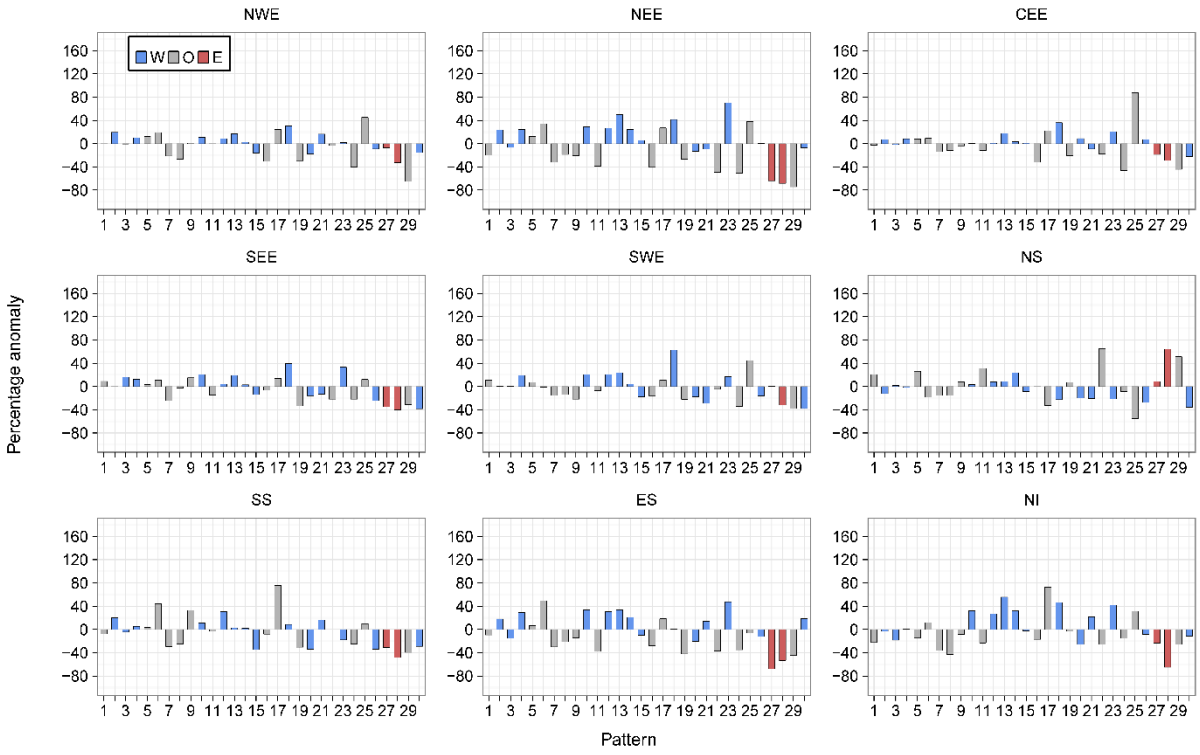

Figure 3.14: As Figure 3.12 but for DSI-12.

3.6.2 MO-30 comparison with LWTs

The six most frequent LWTs in the period 1871 to 2015 are the pure anticyclonic and cyclonic types plus four westerly and southerly variants (Table 3.3). Of the WPs in MO-30, three or four are mapped, via the objective method of Jenkinson and Collison (1977), to each of the six most frequent LWTs. Furthermore, the lack of easterly and north-easterly LWTs in MO-30 is reflected by their relative rarity, with E, AE, CE, NE, ANE and CNE LWTs accounting for a combined occurrence of just 6.67%. It is perhaps surprising that more of the WPs in MO-30 are not classified as pure anticyclonic LWTs, given the predominance of this type in the record (20.58%). Another interesting feature is that two WPs are assigned to the 'unclassified' LWT, which has a frequency of just 1.06%. Moreover, these two patterns occur relatively often (WP1 and WP7 account for a combined 11.37% of all pattern occurrences 1850-2015). MO-30 WP occurrences on the days of each LWT between 1871 and 2015 are counted (Figure A.4). The most frequent WP for each LWT does not always correspond to that pattern's average LWT classification. Often, however, the WP will feature similar behaviour to the more frequent WPs in terms of flow direction and cyclonicity. For example, on CW LWT days, WP30 is the fifth most frequent WP, yet its MSLP definition as a LWT is CW. This is explained partly by the

fact that WP definitions are composites of individual MSLP days and partly by WP30 occurring less often than other patterns. As with WP30, the four most frequent patterns on CW LWT days (WP8, WP4, WP1 and WP26) all feature westerly flow, with WP8 and WP26 additionally being cyclonic (Figure 3.1).

The distinction MO-30 makes between WPs that are defined as the same LWT is important. This distinction is linked to the larger region size used in the derivation of MO-30 compared to LWT. The inclusion of much more of the North Atlantic Ocean and Europe allows for a wider view of the dominant large-scale weather system for each pattern. For example, WP13 and WP14 are both defined as NW LWTs. Figure 3.1 shows that both WPs feature an anticyclone southwest of the UK and a cyclone to the northeast, causing a north-westerly flow over the region. However, in WP13 the anticyclone is closer to the UK and the cyclone further away than for WP14, resulting in WP13 being drier than WP14 overall (Figure 3.6). WP26 is also classed as a NW LWT, with the low and high in similar positions to WP14. The depth of the low-pressure anomaly is much greater however, and WP26 is the sixth wettest WP over the UK. Subtleties between other groups of WPs with the same LWT assignment is also evident (e.g. WP2, WP12, WP15 and WP21, which are all SW LWTs). In total, 18 of the 30 WPs in MO-30 feature some kind of westerly flow over the UK, compared to the nine possible westerly LWTs (W, SW and NW plus the A- and C-directional hybrids). By having a greater number of WPs representing the most common flow direction over the UK (i.e. westerly), more precise statements may be made about the precipitation expected. This could be useful in forecasting and in historical analyses inferring precipitation amounts from the WP on that day.

3.6.3 Suitability of MO-8 and MO-30 in UK-based precipitation analyses

To be useful in precipitation-based analyses, precipitation distributions should be distinct between WPs. Section 3.3.1 demonstrates that this is the case for MO-30, with most WPs exhibiting differences in median precipitation amount or variability. Furthermore, the distinction between WPs is generally greater than for LWTs, evidenced particularly by high variability in some of the wetter LWTs (Figure 3.7). MO-8 was produced by clustering WPs from MO-30 according to the spatial correlation of MSLP anomalies. This combines WPs from MO-30 with very different precipitation distributions and hence the resulting pooled precipitation distributions for some groups of WPs in MO-8 are very similar to each other. Therefore, although smaller sets of WPs may be preferable in some applications, it is recommended that for UK precipitation a different clustering method should be used to group WPs from MO-30. A WP classification with this design specification might include some atmospheric water component in the derivation. For example, for the contiguous USA, Prein *et*

al. (2016) classify WPs using sea level pressure, precipitable water and 700 hPa wind speed, as these variables are crucial in the physical processes driving precipitation (e.g. Doswell *et al.*, 1996).

Results for monthly frequency anomalies of WPs during dry/wet periods, defined by SPI thresholds, demonstrate the comparability of MO-30 and LWTs. In the original description of LWTs, Lamb (1972) describes anticyclonic and cyclonic types as dry and wet, respectively, across the UK, with regional differences in precipitation evident among the directional types. This general wet/dry behaviour mostly agrees with precipitation associated with MO-30 WPs: westerly variants typically occur more (less) often than the mean during wet (dry) periods in western regions and less (more) often than the mean during wet (dry) periods in eastern regions. The opposite is true for easterly types.

3.6.4 Weather patterns associated with drought

In general, the WPs associated with droughts are physically consistent with expected conditions (in terms of airflow direction and cyclonicity). WPs defined by flow from one direction tend to occur more often than normal during droughts in regions on the opposite side of the UK, and less often than normal during droughts in regions closer to the airflow direction source. WPs characterised by anticyclonic behaviour over a region are more likely to enhance drought and vice versa. This is in agreement with Phillips and McGregor (1998), who showed that several droughts in southwest England between 1962 and 1996 were mostly characterised by an increase in the occurrence of the N, NE, E and SE LWTs and their anticyclonic equivalents. This was coincident with a decrease in southerly and westerly LWTs. However, it is interesting to note that the WP showing the largest departure in frequency during six-month droughts in SWE is WP23 (Figure 3.13). This WP's LWT equivalent is W and so, intuitively, wet conditions would be expected, yet from the MSLP definition shown in Figure 3.1, SWE is close to the centre of high pressure and would therefore experience calmer, drier conditions than other regions. In Yorkshire (part of NEE), Fowler and Kilsby (2002b) found that droughts were associated with increases in the A, AE, ASE, AS and ASW LWTs, sometimes with concurrent decreases in westerly LWTs. Correspondingly, DSI-3 droughts in NEE coincide with increases in the frequency of several WPs from MO-30, such as WP3 (ASW), WP6 (A), WP12 (ASW), WP16 (AS), WP17 (ASE) and WP18 (ASW). There is also some agreement for WPs that decrease in occurrence during droughts. Notable exceptions are WP19 (N), WP24 (CN) and WP28 (CSE), which are in the "easterly" cluster in Fowler and Kilsby (2002b), and WP27 (AE). Disagreement between MO-30 and LWT frequencies such as this highlights how WP definitions in MO-30 are more subtle, and direct analogy to a particular LWT may be

unsuitable. In general there are difficulties in comparing different classifications, as often they have different numbers of weather patterns, are calculated for different domains and use different input data sets.

3.7 Conclusions

This chapter demonstrates the applicability of a new classification of 30 WPs to UK precipitation and meteorological drought analyses, and its advantages compared to LWTs. WPs in this classification mostly show more distinct differences in the daily precipitation distribution for nine UK regions than is the case for LWTs. A smaller set of eight WPs, however, as currently defined is not suitable for UK precipitation and meteorological drought analysis as the WPs show too much similarity in their precipitation amounts. Monthly frequency anomalies show which WPs occur more or less frequently during SPI-defined dry or wet periods across all regions, and any regional differences. It is demonstrated that, in general, the same WPs are responsible for dry or wet periods for time-scales of three, six and 12 months. The magnitude of anomalies associated with these WPs typically decreases as the time-scale is increased. WPs associated with dry (wet) conditions typically occur more (less) often than normal during drought months, which are defined using the DSI. Regional differences in the WPs associated with SPI dry conditions mostly hold for the DSI equivalent. WPs associated with drought can be summarised as follows.

- During droughts spanning the majority of UK regions, WP6, WP9, WP10, WP12, WP17 and WP25 occur more often than normal.
- Droughts in western (eastern) regions are generally accompanied by a rise (fall) in the frequency of WP27.
- Droughts in eastern (western) regions often see increases (decreases) in the number of WP15 days.
- WP23 occurs more often than normal during droughts in all regions except those in Scotland.

There are several opportunities for further research. Clustering WPs from MO-30 into a smaller set based on precipitation, in addition to spatial correlation of MSLP anomalies, might be useful for monthly or seasonal analyses where fewer patterns are desired. Furthermore, the methods presented in this chapter could be applied to other regions in Europe, to worldwide regions using another weather pattern data set, and to hydrological drought with a different drought index. Finally, the predictability of MO-30 can be investigated, with a focus on the high-risk

drought WPs highlighted in this chapter. As results here consider WP frequency anomalies on time-scales of at least three months, any forecast product arising from this research would require a minimum three-month lead-time. There are several global seasonal forecast models capable of providing probabilistic forecast output to at least this range (e.g. MacLachlan *et al.*, 2015), enabling a prediction of the spread of forecast frequency anomalies. This could form the basis for a probabilistic forecaster decision tool, highlighting periods with a higher likelihood of drought conditions months in advance.

Chapter 4

Weekly to multi-month persistence in sets of daily weather patterns over Europe and the North Atlantic Ocean

The material in this chapter is currently accepted (in press) for publication as:

Richardson, D., Kilsby, C. G., Fowler, H. J. and Bárdossy, A. (2018b), Weekly to multi-month persistence in sets of daily weather patterns over Europe and the North Atlantic Ocean. Int. J. Climatol.

Part of the Introduction of this journal article has been moved to Chapter 2 of this thesis, while some details from the Data section have been removed as they are covered in Chapter 3. Other minor changes to the wording have been made to make the article more coherent in the context of this thesis.

4.1 Introduction

Persistence in time series of daily WP classifications can provide useful information such as on the memory of broad-scale atmospheric circulation. Despite this, research of WP persistence has lagged behind that exploring their frequencies of occurrence and has focussed on changes and trends in mean persistence over time (Stefanicki *et al.*, 1998 ; Werner *et al.*, 2000; Kyselý, 2002; Kyselý and Domonkos, 2006; Kyselý and Huth, 2006; Kyselý, 2007; Blenkinsop *et al.*, 2009; Cahynová and Huth, 2009; Casado *et al.*, 2009; Kučerová *et al.*, 2017). Missing from the literature is analysis of the longest persistent sequences of WPs and the implications for atmospheric circulation and hydro-climatic variables. Furthermore, persistence has not, to the author's knowledge, been explored in a modelling context except by Fayos and Fayos (2007). Using the objective Lamb Weather Type classification (Lamb, 1972; Jenkinson and Collison, 1977), they developed a neural network forecast model to predict a WP based on the previous five days of WPs. Quantifying persistence in a statistical framework is a useful exercise as it could form the basis of a WP forecast model. Another motivation for this study is to explore a more relaxed definition for persistence. WP persistence has always been defined in previous works as uninterrupted sequences of a single WP (or WP cluster) and as a result these sequences are relatively short, typically with a mean of one week or less (Werner *et al.*, 2000; Kyselý, 2002; Kyselý and Domonkos, 2006; Kyselý and Huth, 2006; Cahynová and Huth, 2009; Casado *et al.*, 2009). An exception found by Kučerová *et al.* (2017) is for some classifications during the summer over the eastern Mediterranean, which show a mean persistence of up to 52 days, which the authors state reflects the classifications' unrealistic temporal behaviour. Here, a less

strict assumption is proposed, allowing for limited numbers of different WPs to break up occurrences of the WP(s) under consideration. This means much longer sequences that are broadly characterised, but not exclusively dominated, by one or more WPs could be identified, with little change in local climatic conditions.

In this chapter, the persistence of the 167-year MO-30 time series shall be explored using two methods. The first is an empirical counting method used to identify monthly to seasonal persistence in groups of WPs, which will be linked to UK precipitation and notable meteorological events. The second method focuses on shorter-term (up to three weeks) persistence in a novel Markov-model framework used to quantify the likelihood of WPs occurring given their occurrence up to 20 days prior. Results will be compared to a benchmark ensemble of WP time series. Section 4.2 outlines the WP and precipitation data and the two methods for quantifying WP persistence. Section 4.3 presents the results and discussion points and Section 4.4 provides conclusions.

4.2 Data and methodology

4.2.1 Data

The WP classification is MO-30 between 1850 and 2016, as described in Chapter 3. For daily precipitation, the Met Office HadUKP data set from 1931 to 2016 (Alexander and Jones, 2000) is used. The precipitation data is standardised for the nine regions by $\boldsymbol{Z} = \boldsymbol{X} - \tilde{x}$, where $\boldsymbol{X}$ is the distribution of daily precipitation for a particular month and region and $\tilde{x}$ is the median of $\boldsymbol{X}$.

4.2.2 An empirical counting method for identifying long-term weather pattern persistence

The first step in analysing WP persistence is to count on how many consecutive days each WP occurs. Annually, for 24 individual WPs the median persistence is just one day (Figure 4.1). The remaining six WPs have a median persistence of two days. The maximum persistence of any WP over the whole 167 years is 18 days for WP27, followed by 17 days for WP28. These are the only WPs characterised by easterly flow over the UK; this will be discussed in Section 4.3. Persistence is also linked to seasonality, with out-of-season WPs less likely to persist compared to the annual average (not shown).

The length of persistence for individual WPs might be suitable for linking with short-term events such as flooding, but not for events that evolve on longer time-scales, such as drought. To identify longer-term persistence, an empirical counting (EC) method is developed, considering sets of WPs, denoted S, that are treated as a single WP. For different sets of n WPs,

EC searches the MO-30 time series to find persistent periods of length at least v days containing WPs in the set. It is likely that long-duration dry and wet events will not be exclusively composed of whichever WPs are in the set under consideration - there will be days on which a different WP occurs that may not be dry or wet (and may indeed be the opposite to the majority of WPs during the event). However, if the number of these days is sufficiently low then the period under consideration may still be dry or wet overall. Therefore, EC permits a certain percentage, $a\%$, of days to be WPs outside the set. Sets of two and three WPs out of the 30 WPs in MO-30 have 435 and 4,060 unique, unordered combinations, respectively. It is feasible to compute persistence periods for each of these combinations. For higher-order WP combinations, the task becomes far more expensive in terms of time and computational burden: four- and five-WP sets have 27,405 and 142,506 unique combinations, respectively. Therefore, EC is extended to $n > 3$ by introducing a second set of n_f fixed WPs, S_f. For $n < 4$, $S_f = \emptyset$, the empty set. A persistence period is defined as a time period of at least v days where $(100 - a)\%$ of WPs are in S or S_f. Each WP in S and at least one WP in S_f must occur over this period. Using $S_f \neq \emptyset$ for $n > 3$ WPs then allows identification of much longer periods of persistence, as occurrences of WPs in S broken up by occurrences of WPs in S_f will not terminate a persistence period. Moreover, it enables the searching of hypothetically promising sets of more than three WPs while drastically reducing the total number of combinations to process.

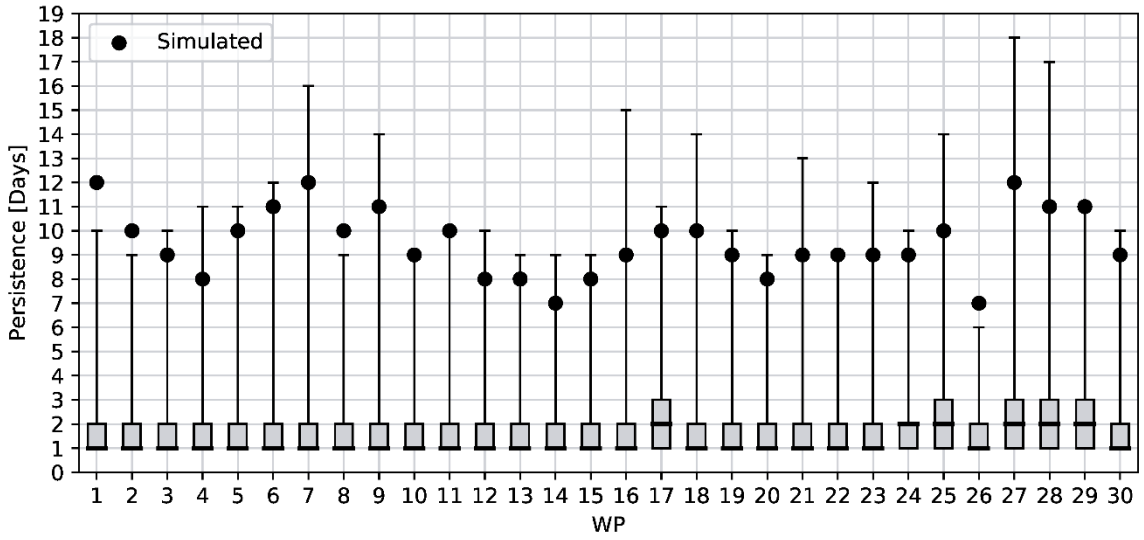

Figure 4.1: Box plots show the distribution of consecutive occurrences of each WP. Black circles represent the 90th percentile of the distribution of maximum consecutive occurrences from the 1000 simulated series. A simulated value less than the observed maximum indicates that 90% of the simulated series fail to capture the persistence of the observed series.

The algorithm for EC is as follows, using $n = 2$ and $n_f = 3$ as an example:

1. Define a set of three fixed WPs $S_f = \{W_1 = i, W_2 = j, W_3 = l\}$ and a set of two WPs $S = \{W_4 = m, W_5 = o\}$ for $i, j, l, m, o = 1, \ldots, 30$, with $i \neq j \neq l \neq m \neq o$.

2. Set time $t = 0$, the length of the moving window in days, v, and the percentage of days per moving window allowed to be WPs not in S or S_f as $a\%$.

3. Take the $t, \ldots, t + v$ days of the WP time series. If $(100 - a)\%$ of WPs in this window are in S or S_f, each WP in S occurs at least once and at least one WP from S_f occurs, then define a new persistence period of length v days and proceed to Step 4. If not satisfied, set $t = t + 1$ and repeat.

4. Set $t = t + 1$ and take the $t, \ldots, t + v$ days of the WP time series. If $(100 - a)\%$ of WPs in this window are also in S or S_f, each WP in S occurs at least once and at least one WP from S_f occurs, then the length of the persistence period becomes $v^* = v + 1$ days. Repeat this step until less than $(100 - a)\%$ of WPs in S and S_f are in the window, at least one WP from S does not occur or no WPs in S_f occur.

5. The persistence period is terminated. Trim the persistence period at either end such that the period must begin and end with a WP in either S or S_f.

6. Set $t = t + 1$ and return to Step 3 until the entire time series has been searched.

7. Choose a new $S^* \neq S$ and repeat from Step 3 until every combination of WPs has been processed.

A persistence period of v^* days will not necessarily contain at least $v^*(100 - a)\%$ days featuring the WPs in S or S_f, but each of the v-day windows within the period will have at least $v^*(100 - a)\%$ days of these WPs. EC effectively treats all WPs in the sets as the same WP, and therefore the identification of multi-WP persistence periods does not necessarily mean that the individual WPs persist within that period. It is possible that a period could consist of many single-day occurrences of WPs in S and/or S_f.

For each identified persistence period, the observed precipitation anomalies, Z_o, and expected precipitation anomalies given the WPs, $\boldsymbol{Z}_e$, are calculated. The expected precipitation anomalies for each v-day persistence period are calculated as follows.

1. Let $\boldsymbol{Y}$ be the time series of precipitation anomalies for a particular region.

2. Set the month, M, as that corresponding to the middle day of the persistence period, and let the season, s, comprise M and the two months either side.

3. For each day in the persistence period, $T = t, \dots, t + v$, generate $Q = 10{,}000$ draws with replacement, $\boldsymbol{y_T} = (y_{1,T}, \dots, y_{Q,T})$, from the conditional distribution of precipitation anomalies given the WP occurring in s, $\boldsymbol{Y}_{s,WP}$,

4. Over all days, sum the generated samples to obtain a distribution of expected precipitation anomalies, $\boldsymbol{Z_e} = \sum_{T=t}^{t+v} \boldsymbol{y_T}$.

The observed persistence period precipitation is the sum of the daily anomalies, Y_T, over the period, given by $Z_o = \sum_{T=t}^{t+v} Y_T$. While Z_e can be estimated for the same period that MO-30 is available for (1850 to 2016), Z_o is only available from 1931.

Two sets of parameters are considered for EC. First, to test for monthly persistence amongst small sets of WPs, the parameters are chosen as follows: $n = 3$ WPs, $v = 30$ days and $a = 10\%$. As $n < 4$ no fixed WPs are required, so $S_f = \emptyset$. This parameter combination is denoted EC30. These are strict criteria: to be satisfied, three WPs will have to dominate 30-day periods with only three days permitted to be other WPs. The second parameter combination is chosen to seek multi-month persistence. After trying various numbers of WPs in S and S_f, a suitable compromise between number of WPs and length of persistence is identified as follows: $n = 3$ WPs, $n_f = 3$ WPs, $v = 60$ days and $a = 20\%$ (the same ratio of v and a as for EC30). This parameter combination is named EC60. Choosing to investigate the link between persistence and drought, the three driest WPs according to UK mean daily precipitation (Figure 3.6) are selected and so S_f = {WP6, WP17, WP25}. This choice reflects both frequently occurring summer WPs (WP6) and less frequent, wintry WPs (WP17 and WP25). As the three driest WPs, they are unsurprisingly characterised by anticyclonic conditions over the UK (Figure 3.1).

4.2.3 A Markov model for quantifying persistence

As EC is not a statistical model, it is difficult to use results arising from its application in further applications such as forecasting. Therefore, another approach to quantifying persistence in MO-30 is taken, underpinned more formally by a statistical model. A common way of modelling discrete data is using Markov processes (Norris, 1997) to quantify the probability of state (in this case WP) transitions. As before, let W_t represent a particular WP*i* on day *t*. Therefore, the first-order Markov assumption is:

$$\Pr(W_t|W_{t-1}, W_{t-2}, \dots, W_1) = \Pr(W_t|W_{t-1}),$$

Equation 4.1

i.e. the probability of a WP occurring is only dependent on the WP of the previous day. The transition probabilities are given by the transition matrix, **P**, with each element calculated as $p_{i,j} = \Pr(W_t = i|W_{t-1} = j)$, for $i, j = 1, \dots, 30$. The most likely transition for any WP is the same-state transition and the seasonality is also evident, with the lower-numbered, summer WPs more likely to transition between themselves and vice-versa for the winter WPs (Figure 4.2). To increase the number of days influencing each transition, higher-order Markov chains could be considered. However, the number of parameters needed to estimate Markov models increases exponentially with the order, and the sample size of some multi-day transitions could be very small, making their transition probability estimates highly uncertain. Fortunately, for persistence, consideration of all possible transitions to a high order is not necessary; only the return of the WP system to its initial state is of interest. For example, in the second-order case, rather than considering $\Pr(W_t = i|W_{t-1} = j, W_{t-2} = l)\,, i, j, l = 1, \dots, 30$, it is sufficient to consider $\Pr(W_t = i|W_{t-1} = j, W_{t-2} = i)$. Instead of 30 permutations given by the values l could take, only one is necessary, that given by i.

This idea is generalised to higher orders by considering $u \geq 1$ occurrences of $W_t = i$ in the k-day period prior to the final transition of W_{t-1} to W_t. This conditional probability can be expressed as

$$\Pr\left(W_t = i|W_{t-1} = j, \sum_{T=t-k-1}^{t-2} \mathrm{I}_{W_T}(i) \geq u\right), i, j = 1, \dots, 30,$$

Equation 4.2

where $\mathrm{I}_{W_t}(x)$ is the indicator function defined as

$$\mathrm{I}_{W_t}(x) \coloneqq \begin{cases} 1 \; if \; x = W_t, \\ 0 \text{ otherwise.} \end{cases}$$

Equation 4.3

That is, Equation 4.2 calculates the probability of a WP*i* occurring given a different WP*j* on the previous day and u occurrences of WP*i* in the k days prior to the final transition. For notational simplicity Equation 4.2 is re-written as

$$\Pr\left(W_t = i \middle| W_{t-1} = j, W^{(k,u)} = i\right).$$

Equation 4.4

The phrase "k-day period" will always refer to the time period prior to the final transition, $t-k-1,\ldots,t-2$. Equation 4.4 reveals whether the occurrence of a particular WP in a given window is likely to be followed by a repeat of this WP, despite other WPs occurring during the k-day period and between the k-day period and the final W_t. Furthermore, these probabilities can be compared with the equivalent first-order transition probabilities in **P**:

$$D_{i,j}^{(k,u)} = \Pr(W_t = i|W_{t-1} = j) - \Pr\left(W_t = i\middle|W_{t-1} = j, W^{(k,u)} = i\right).$$

Equation 4.5

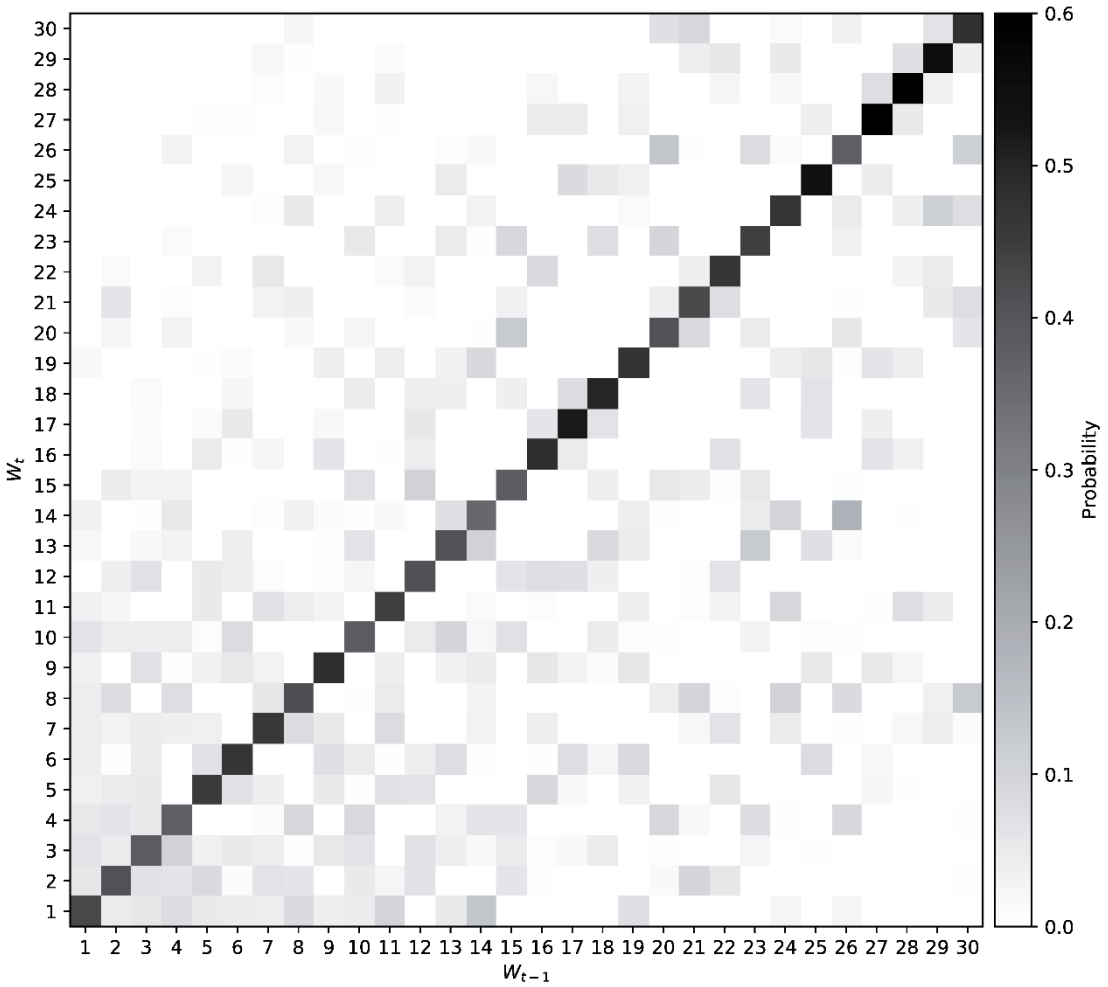

Figure 4.2: First-order transition probabilities for MO-30 WPs.

Negative $D_{i,j}^{(k,u)}$ occur when $\Pr\left(W_t = i\middle|W_{t-1} = j, W^{(k,u)} = i\right) > \Pr(W_t = i|W_{t-1} = j)$ i.e. when WPi has already occurred at least u times in the k-day window, and the resulting probability of $W_{t-1} = j$ transitioning to $W_t = i$ is greater than if the sequence was considered independently. These transitions are interesting as they indicate persistence. This Markov method will be referred to as MM. The choices of k and u used for MM are shown in Table 4.1. Higher values of k are chosen to investigate whether there is long-range dependence in

time, and the range of values chosen for u are to assess the differences between relaxed and strict persistence rules for each k.

k	u
5	1, 2, 3, 4
10	2, 3, 4, 5, 6, 7
15	3, 5, 7, 9, 10, 11, 12
20	3, 5, 7, 10, 12, 15

Table 4.1: Choices of window length, k, and strictness of persistence, u, for the Markov model (MM).

4.2.4 Comparing with a benchmark ensemble of synthetic weather pattern series

A simulation study is conducted to test whether a time-inhomogeneous first-order Markov model can reproduce observed persistence in the MO-30 data. Inhomogeneity in time is a necessary imposition, as ignoring the seasonality in MO-30 would result in the under-estimation of persistence. For example, as summer is dominated by WP1 through WP11, using a non-seasonal transition matrix would result in higher numbers of the other WPs and therefore fewer summer WPs, and hence a decreased likelihood of persistence. An ensemble of 1000 WP series is simulated using the following procedure:

1. Calculate the 12 monthly transition matrices, $\mathbf{P^m}, m = 1, \dots, 12$, of MO-30: $$p_{i,j}^m = \Pr(W_t = i | W_{t-1} = j), \text{for } i, j = 1, \dots, 30.$$
2. Generate an initial W_1 according to its frequency of occurrence and set time $t = 1$.
3. Generate a random number p^* from the standard uniform distribution $U(0,1)$.
4. Find the index q such that: $$\sum_{j=1}^{q-1} p_{i,j}^m < p^* < \sum_{j=1}^{q} p_{i,j}^m$$ where $i = W_t$.
5. Set $W_{t+1} = q$ and $t = t + 1$.
6. Repeat steps 3 to 5 until $t = T$, the length of the series to be generated.

The benchmark ensemble has reasonable skill at replicating the observed WP frequencies of occurrence (Figure B.1). For most WPs and months, the observed frequencies fall within 90% of the simulated series. There are some significant exceptions. In particular, the simulation methodology underrepresents the occurrences of the two easterly WPs, WP27 and WP28, between January and March. Winter appears to be the least-well simulated season, with WP4, WP8, WP10, WP22, WP23, WP25, WP29 and WP30 all featuring at least one month over- or

under-simulated, implying that the first-order transition probabilities are least suited to modelling this season. The simulated series also fail to capture the observed maximum persistence of individual WPs, with 90% of the ensemble underestimating this statistic for 22 WPs (Figure 4.1).

4.3 Results

4.3.1 Empirical counting method results

Despite the transient nature of individual WPs, there are multi-month periods where small numbers of WPs dominate. Furthermore, within these periods, the persistence of individual WPs tends be above average. EC30 identifies 67 persistence periods from 54 sets with a median persistence of 31 days (Figure 4.3a), equivalent to roughly one persistence period every 2.5 years. The longest persistence period was 48 days over winter 1962/63, which was characterised by anticyclonic, often easterly, flow over the UK, as can be seen from the high frequencies of WP25, WP27 and WP28 (Figure 4.4a). Another example is for 30 days over winter 1990, with WPs in this period suggesting very windy and often stormy conditions, culminating in a 10-day sequence of WP30, which is the longest amount of time this WP has persisted over the whole 167 years (Figure 4.4b).

EC60 identifies 79 persistence periods from 58 sets, with almost half lasting for at least nine weeks (Figure 4.3b). The two longest examples are almost a full season in length, over summer 1968 and, again, in winter 1962/63 (Figures 4.4c and d). The former features a near month-long stretch of anticyclonic WPs (WP6, WP9, WP17 and WP25) between 16th July and 10th August. The latter shows how the inclusion of three WPs in S_f can extend a persistence period by almost a month (Figures 4.4a and d). However, the key aspect of this persistence period is the dominance of easterly WPs between 26th December and 9th February, which is completely captured by the shorter persistence period with $S_f = \emptyset$ (Figure 4.4a).

All cases in Figure 4.4 highlight how, within each period, individual WP persistence is greater than average, with many cases of persistence greater than two days (the highest average persistence for any WP). This was particularly stark during the winter of 1995/96 (Figure 4.4e), with nine-, 11- and 12-day stretches of WP17, WP28 and WP27, respectively. That the behaviour of these daily, typically ephemeral WPs is sometimes reduced down to small sets occurring over such long time-scales is remarkable and suggests that the WPs, and possibly atmospheric circulation in general, can be more predictable than usual for months at a time. A dominant easterly flow might be the most likely circumstance under which WP predictability is enhanced, as WP27 and WP28 feature strongly in EC30 and (less so) EC60 persistence

periods (Figures 4.3c and 4.3d). However, any enhanced predictability during such periods is expected, as there are just two easterly WPs but many more featuring westerly flow (over the UK). Further discussion of the consequences of the small number of easterly WPs is presented later.

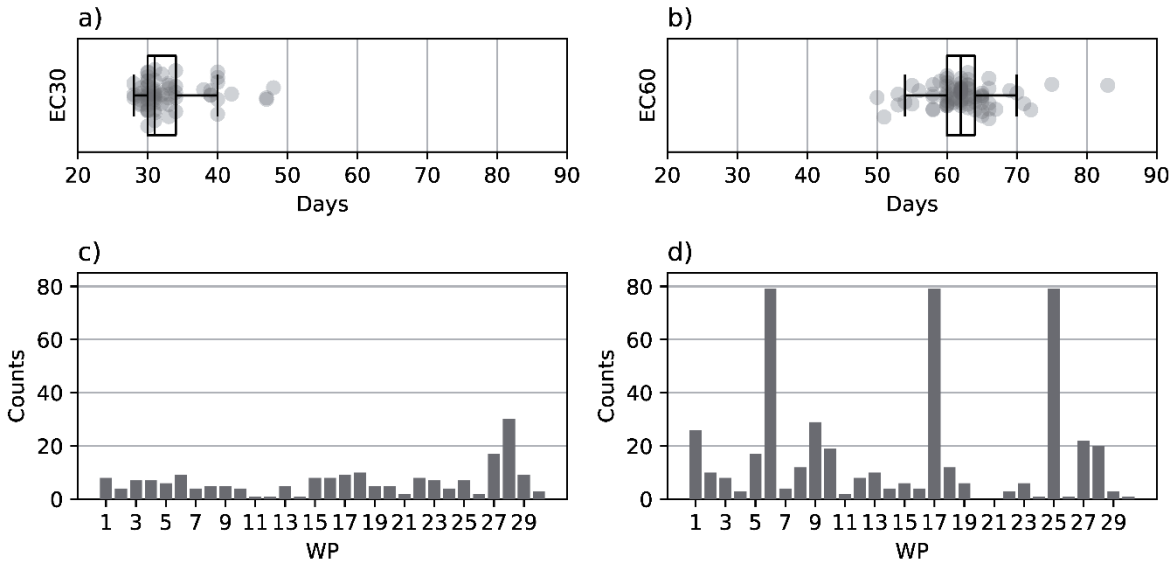

Figure 4.3: a) and b) show box plots and the underlying data for the number of days in each persistence period for EC30 and EC60. Whiskers are 1.5 times the interquartile range beyond the 25th and 75th percentiles. c) and d) are counts of how often each WP appears in the persistence period sets.

EC30 persistence periods are most numerous in winter (December through February; 43 out of 67), while those identified by EC60 are more common in summer (June through August; 37 out of 79), with winter also well-represented (25 out of 79); refer to Table 4.2. The reason that there are many summer persistence periods of at least 60 days in length, but fewer that are of at least 30 days is due to the percentage of days allowed to be WPs not in S or S_f (10% for EC30 and 20% for EC60). Although the ratio of this percentage and the minimum persistence period length is the same for both EC30 and EC60, the latter allows for longer sequences of WPs outside S or S_f (six consecutive days compared to three days for EC30). Therefore, summer persistence periods identified by EC60 have stretches of WPs not in S or S_f of length greater than three days, whereas those identified by EC30 do not.

Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
EC30	25	14	4	2	3	3	9	3	0	0	0	4
EC60	9	0	9	0	1	22	14	1	0	3	4	16

Table 4.2: Number of persistence periods beginning in each month.

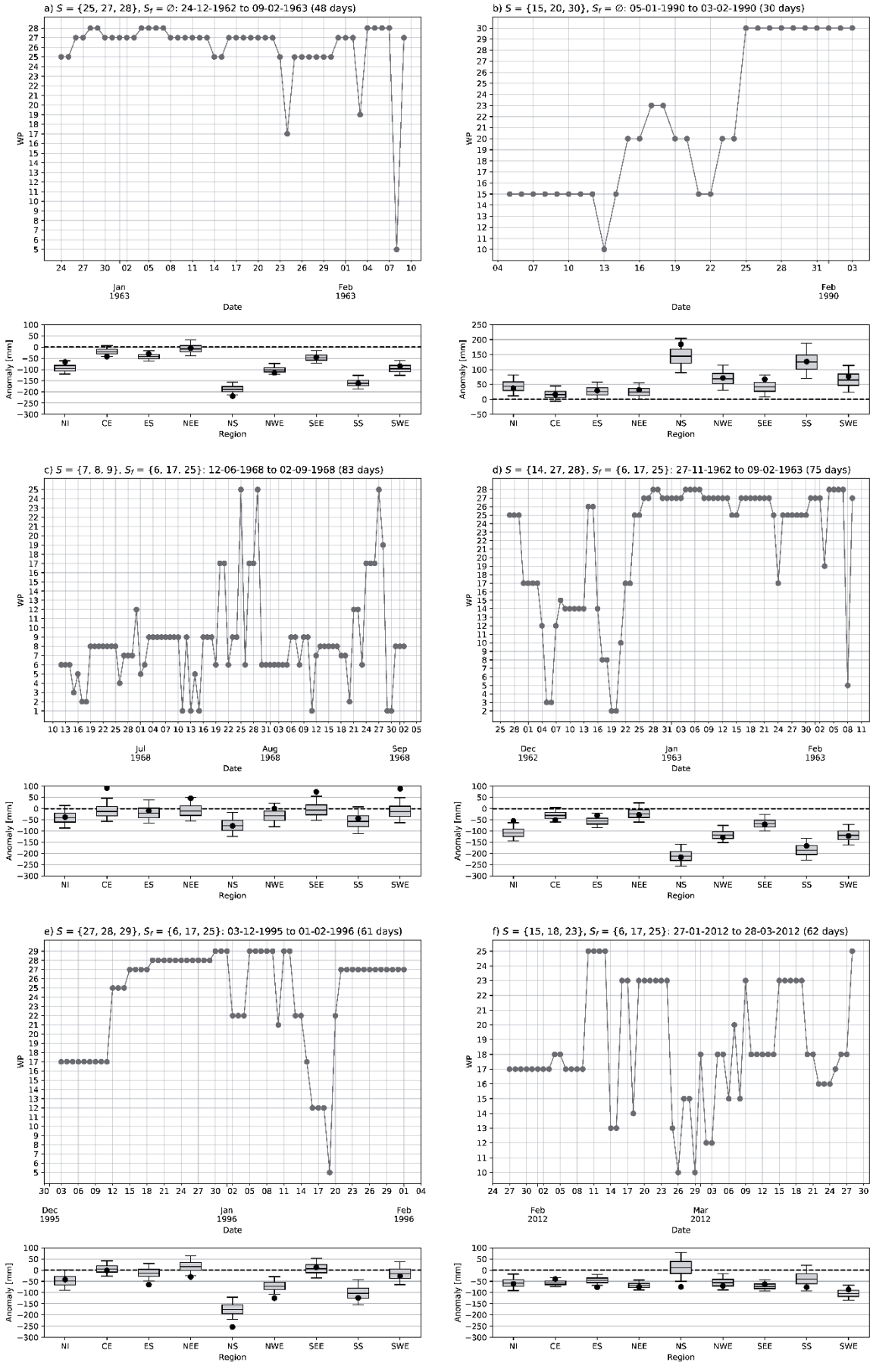
a) S = {25, 27, 28}, S_f = ∅: 24-12-1962 to 09-02-1963 (48 days)
b) S = {15, 20, 30}, S_f = ∅: 05-01-1990 to 03-02-1990 (30 days)
c) S = {7, 8, 9}, S_f = {6, 17, 25}: 12-06-1968 to 02-09-1968 (83 days)
d) S = {14, 27, 28}, S_f = {6, 17, 25}: 27-11-1962 to 09-02-1963 (75 days)
e) S = {27, 28, 29}, S_f = {6, 17, 25}: 03-12-1995 to 01-02-1996 (61 days)
f) S = {15, 18, 23}, S_f = {6, 17, 25}: 27-01-2012 to 28-03-2012 (62 days)
WP
Date
Anomaly [mm]
Region
NI
CE
ES
NEE
NS
NWE
SEE
SS
SWE

Figure 4.4 (previous page): Top figure in each sub-plot shows the WP time series. The bottom figure in each sub-plot shows the expected precipitation anomalies derived from WPs (boxplots) and observed precipitation anomalies (circles). Plotted for two EC30 persistence periods during a) winter 1962/63 and b) winter 1990 and four EC60 persistence periods during c) summer 1968, d) winter 1962/63, e) winter 1995/96 and f) winter/early spring 2012.

The other key feature highlighted by the seasonal distribution of persistence periods is how the number of winter events decreases for EC60 compared to EC30. This is due to the different numbers of WPs associated with each season. In winter, there are more WPs that tend to occur, and so it is less likely there will be persistence of sufficient length to satisfy the conditions of EC60. This is even more the case for the spring and autumn. In these seasons the WP behaviour is transitioning between winter and summer and the frequencies of occurrence are more evenly spread between all 30 WPs.

Overall, persistence periods are associated with drier conditions for western and northern regions than for eastern and southern regions. These persistence periods sometimes coincide with notable meteorological events. Given the WPs that occurred during EC30 persistence periods, the majority of precipitation simulations are below average in NI, NWE, NS, SS and SWE and average for the other regions (Figure 4.5a). This east-west divide is probably due to the predominance of easterly WPs WP27 and WP28 in EC30 periods (Figure 4.3c), resulting in wetter conditions for the eastern UK. Precipitation for individual persistence periods, on the other hand, is likely to be determined largely by the flow direction. Expected and observed anomalies have very similar distributions for both EC30 and EC60, with the main difference being greater 95th percentile anomalies for observed EC30 anomalies (Figure 4.5). Including the three driest WPs in S_f for EC60 does not change the median and lower-tail distribution statistics by very much, with the main effect being a reduction in the upper-tail anomalies, particularly in the observations (Figure 4.5d). The eastern regions remain wetter than those in the west, again likely due to WP27 and WP28 being amongst the most common WPs in EC60 persistence periods and also to the higher precipitation variability associated with WP25 in eastern regions (Figure 3.6).

Examples of easterly-flow persistence are winter 1962/63 (Figures 4.4a and d) and winter 1995/96 (Figure 4.4e), with observed and expected precipitation greater for eastern regions. These correspond to notable meteorological events, with the former known to be a particularly cold and dry winter (Prior and Kendon, 2011) and the latter part of the 1995/96 Yorkshire drought (Fowler and Kilsby, 2002b). Another example coincident with drought conditions was towards the end of the 2010-12 drought (Figure 4.4f), which was particularly severe for central, southern and eastern England in meteorological, hydrological and agricultural contexts

(Kendon *et al.*, 2013). The persistence period occurred over February and March of 2012, with that March being the driest on record since 1953 (Kendon *et al.*, 2013). EC30 also yielded a particularly notable storm: an extremely wet January 1990 (Figure 4.4b), which culminated in the Burns' Day Storm (McCallum, 1990) over southern Scotland on the 25th and 26th January. However, results imply that it is difficult to say anything about precipitation based on WP persistence alone. While there does appear to be a geographical difference (drier conditions in western regions), this may be an artefact caused by the lack of variety of easterly WPs in MO-30; this will be discussed more later.

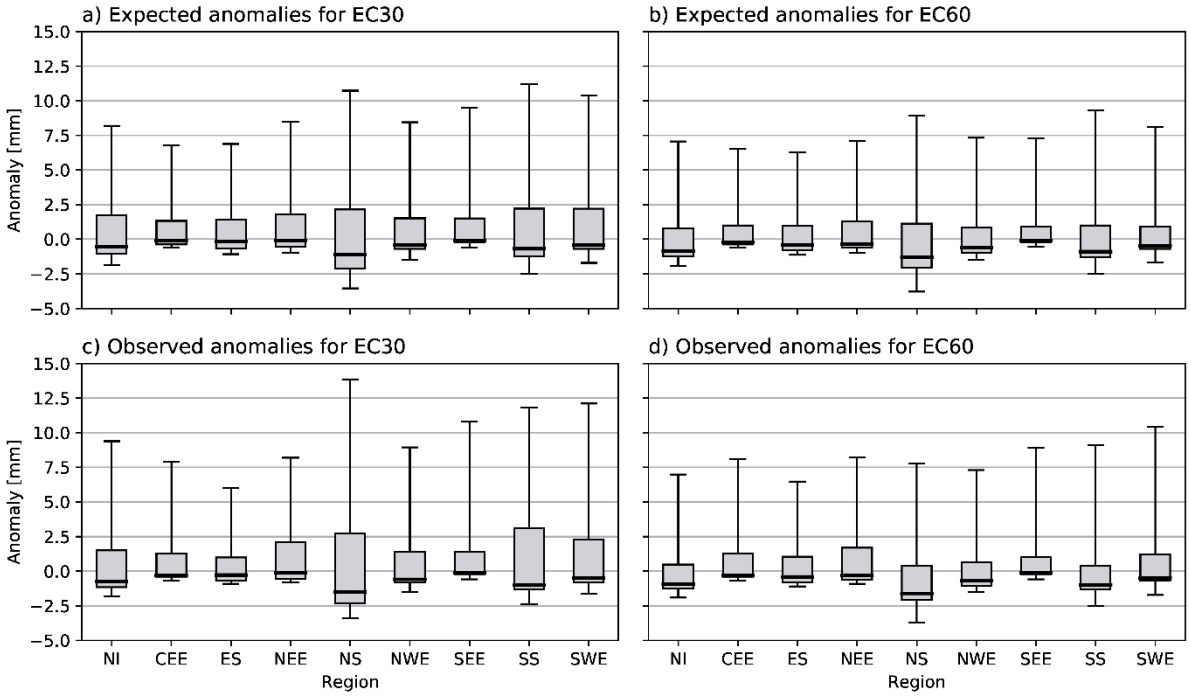

Figure 4.5: Boxplots of a) and b) expected, and c) and d) observed precipitation anomalies during persistence periods for a) and c) EC30, and b) and d) EC60. Expected anomalies are sampled from the distributions of precipitation anomalies for WPs occurring in each persistence period. Whiskers represent the 5th and 95th percentiles.

4.3.2 Markov model results

For some WPs, their probability of occurrence is increased if they have occurred previously, including situations in which other WPs occur in-between. This holds for different lengths of the k-day window for up to almost three weeks. Over all WPs, the more often a WP occurs in this window, the higher probability it has of reoccurring after the final transition. Results are shown for MM with $k = 5$, with stricter persistence rules (i.e. higher u) yielding more probability in the lower tail of the $D_{i,j}^{(5,u)}$ distributions (Figure 4.6a to d). This is also true for other k (not shown). This suggests that when a particular pattern of atmospheric circulation has

been dominating over recent days, it is more likely to persist into the future (dependent on the actual synoptic situation represented by the WPs).

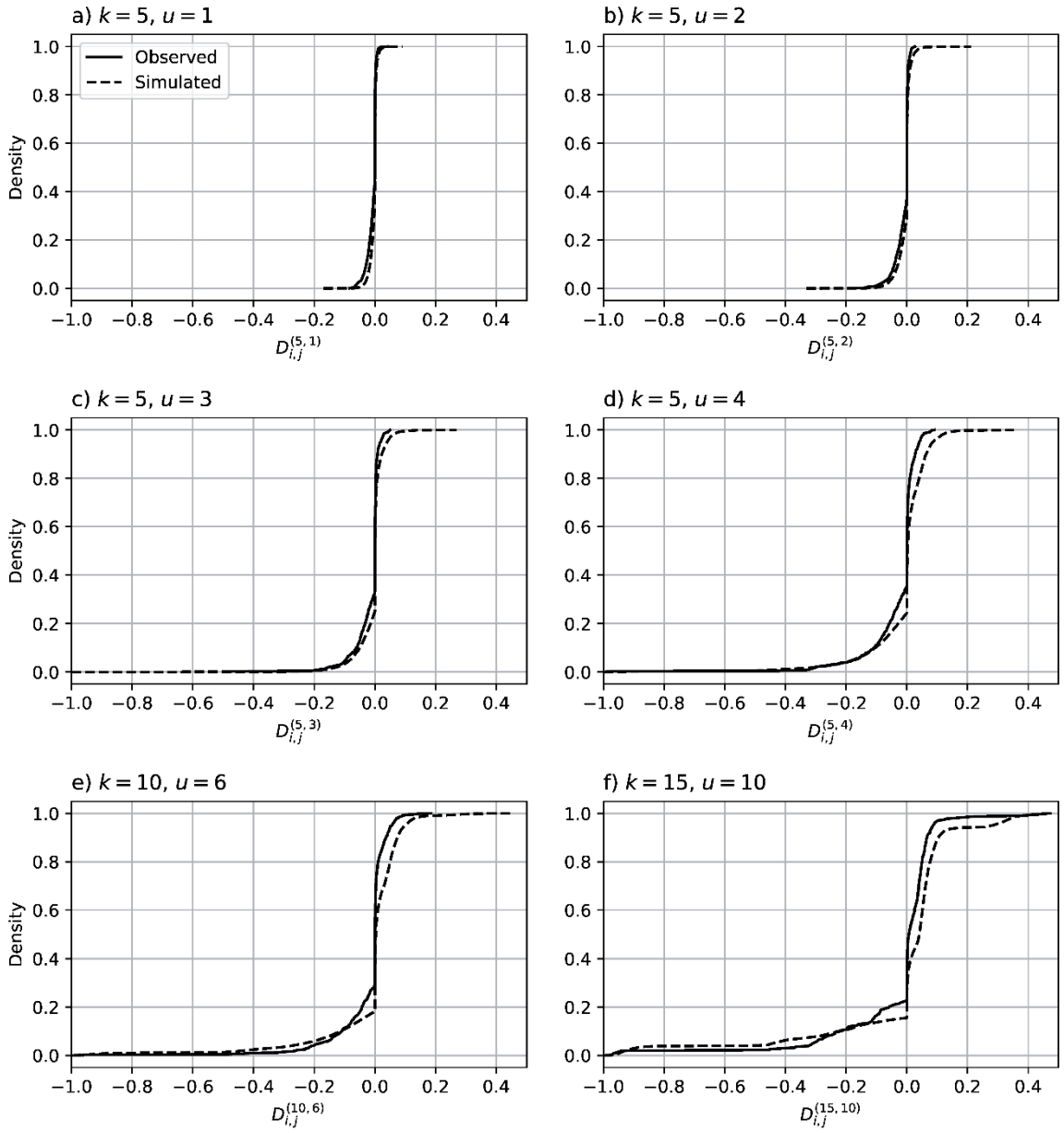

Figure 4.6: ECDFs for $D_{i,j}^{(5,u)}$ with a) $u = 1$, b) $u = 2$, c) $u = 3$, d) $u = 4$, and for e) $D_{i,j}^{(10,6)}$ and f) $D_{i,j}^{(15,10)}$ for observed (solid line) and simulated (dashed line) series.

To identify the WP transitions that are influenced by this type of persistence, the number of times each WP transition is below the 20th percentile of the $D_{i,j}^{(k,u)}$ distributions is counted (for all 23 combinations of k and u). Persistence of anticyclonic WPs appear to influence each other. For example, WP6 shows higher probabilities of reoccurring if the previous WP is WP6, WP17, WP18 or WP25. Similarly, some cyclonic WPs are more likely to persist if the penultimate WP is also cyclonic, such as WP7 preceded by WP7, WP8, WP11, WP22, WP24 or WP28 (Figure

4.7). These results are collated to identify the most persistent WPs by considering the transitions independent of the penultimate WP, equivalent to summing the counts given by the rows of Figure 4.7. It appears as though the summer WPs are more persistent than the winter WPs, with WP1 and WP7 the two most persistent (Figure 4.8a). However, this seasonal divide is also present in the simulated series. When taking the difference between the observed and simulated medians, this contrast is less stark (Figure 4.8b). Accounting for seasonality, the most persistent WPs are WP28, WP16, WP7, WP23, WP18 and WP25. Apart from WP7, these occur more often in winter than summer, although the following three most persistent WPs are all associated with summer (WP2, WP4 and WP6). There does not appear to be a link between WP persistence and the corresponding WP's MSLP definition, with high persistence from WPs that differ by their MSLP anomalies (e.g. cyclonic WP7 and WP28 versus anticyclonic WP18 and WP25) and direction of flow over the UK (e.g. westerly or south-westerly WP7, WP18, WP23; southerly WP16 and south-easterly WP28).

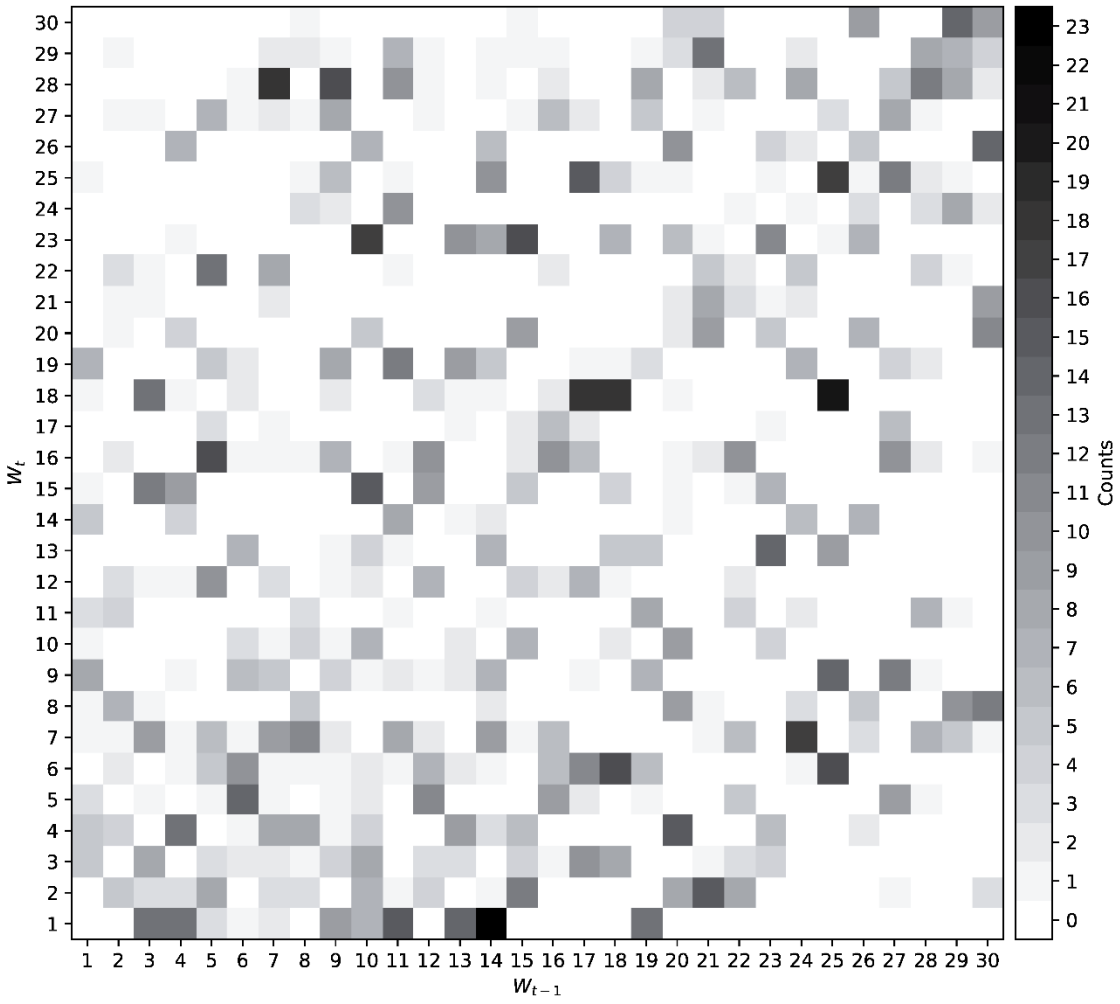

Figure 4.7: Counts of how often each individual value (i.e. WP transition) in the $D_{i,j}^{(k,u)}$ distribution is below the 20th percentile, over all k and u.

The persistence characteristics outlined previously are not replicated by the first-order transition probabilities. Those WP transitions that are more likely when conditioned on persistence (i.e. when $D_{i,j}^{(k,u)} < 0$) are more numerous and have higher magnitude transition probabilities for the observations than for the benchmark ensemble. This is shown by the differences between the lower tails of the observed and simulated $D_{i,j}^{(k,u)}$ ECDFs (Figure 4.6). An exception is how, for higher k and u, there is more probability in the lowest part of the distribution for the benchmark, not observed, WP series. This can be likened to 'repelling' behaviour: for longer k-day windows with stricter persistence rules, the occurrence of certain WPs in the prior window reduces the chance of these WPs persisting. However, this is likely due to these strict transitions occurring infrequently; there is a much smaller sample size in the observed series than over the 1000 simulated series. Two examples of this behaviour are for $D_{i,j}^{10,6}$ and $D_{i,j}^{15,10}$ (Figures 4.6e and f). The simulated ensemble also does not capture the variation in persistence amongst WPs, with the only differences attributable to seasonality: summer WPs are more persistent than winter WPs (boxplots in Figure 4.8a), whereas the reality is more complex (circles in Figure 4.8a). For computational reasons EC is not run for the synthetic series. However, upper-tail statistics of uninterrupted persistence are, for the majority of individual WPs, not reproduced in the ensemble (Figure 4.1). As discussed earlier, persistence periods are characterised by above-average persistence of individual WPs and therefore it may be inferred that the first-order Markov principle is insufficient to replicate persistence periods.

4.3.3 Weather patterns with an easterly flow

The only two WPs with a distinct easterly flow over the UK, WP27 and WP28, stand out throughout this study for several reasons. They are the WPs with the two longest uninterrupted sequences over the entire record and have the highest or joint-highest mean and median persistence (Figure 4.1). They are also the most common WPs in EC30 persistence periods (Figure 4.3c) and are amongst the most common for those identified by EC60 (Figure 4.3d). Furthermore, the first-order Markov assumption is least suited to these two WPs, as they show the greatest difference between observed and simulated maximum persistence (Figure 4.1), and their monthly frequencies are particularly under-simulated in winter (Figure B.1). As there are just two easterly WPs, one might expect most daily MSLP fields with an easterly flow to be assigned to them and hence for these WPs to have high within-WP variability. To check this, the spatial correlation between the WPs and their underlying (concurrent) MSLP fields is calculated. These results show that, in fact, WP27 and WP28 have amongst the lowest within-WP variability (Figure B.2). As such, the fact that these WPs stand out cannot be attributed to

an unsatisfactory assignment process and instead suggests that long periods of atmospheric persistence in the European-North Atlantic domain are characterised by easterly flow patterns closely resembling WP27 and WP28.

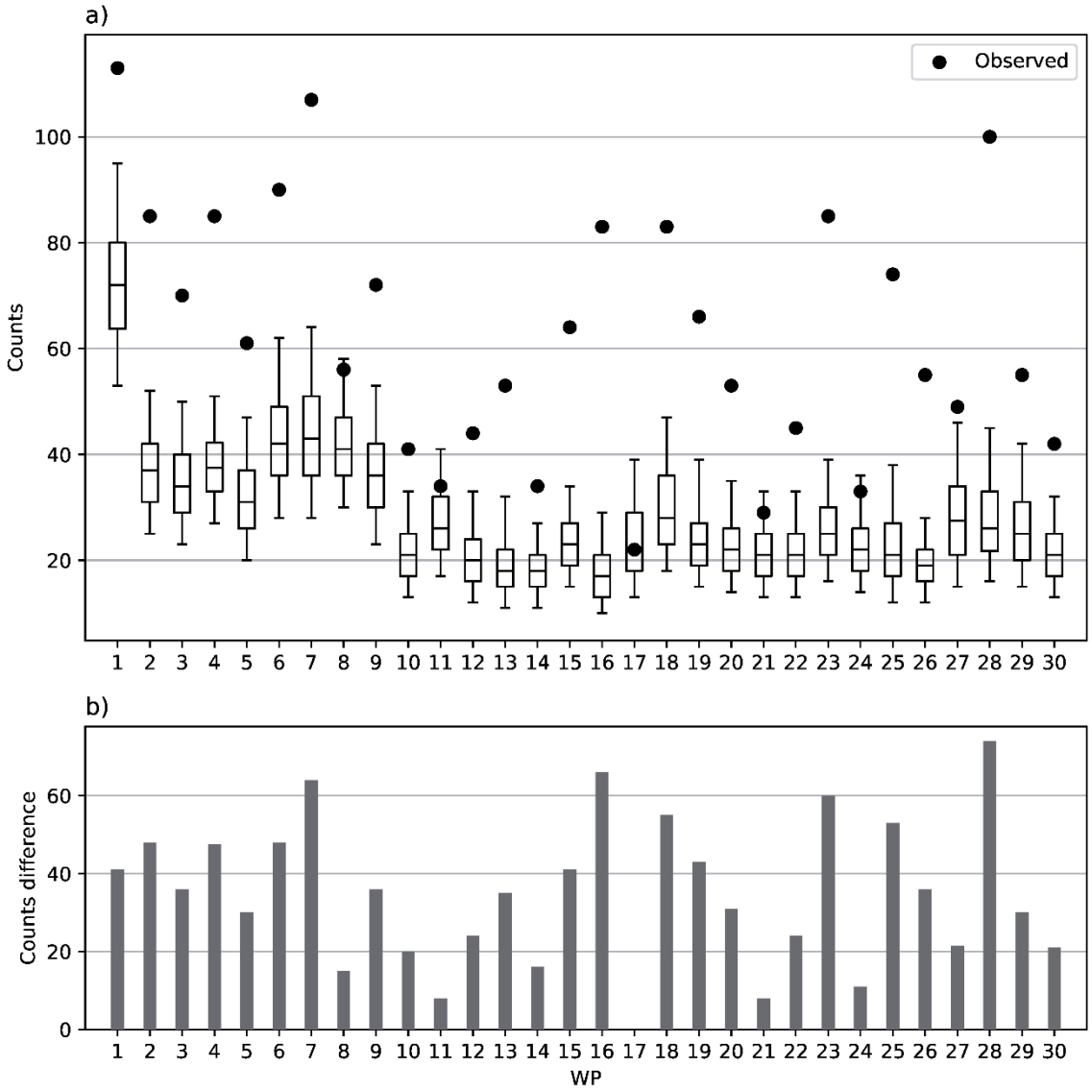

Figure 4.8: a) Independent of the penultimate WPj, how often each value (i.e. WP transition) in the $D_{i,j}^{(k,u)}$ distribution is below the 20th percentile, over all k and u. Observed series represented by the black circles and 1000 simulated series by the boxplots. Whiskers are the 5th and 95th percentiles. b) Observed counts minus median of the simulated counts.

4.4 Conclusions

The persistence properties of a time series of 30 WPs defined by MSLP over the North Atlantic Ocean and Europe have been investigated. By developing an empirical counting method that accounts for more relaxed persistence rules than currently in the literature, it was shown that there are multi-month periods when normally transient WP behaviour can be reduced down to many occurrences of small sets of WPs. These periods generally result in drier than average

conditions for western UK regions and have coincided with notable meteorological events, including droughts. This geographical divide is a result of the predominance of easterly WPs in the identified persistence periods. Persistence properties were also analysed using a novel Markov model, which quantified the difference between first-order WP transition probabilities and those conditioned on prior occurrences, without the need for estimating the large numbers of parameters usually required by high-order Markov chains. Some WPs are more likely to reoccur than others, although there does not appear to be a link between their persistence and the MSLP definition.

WP frequencies of occurrence are likely modulated by lower-frequency (slowly varying) physical processes, such as large-scale atmospheric teleconnection patterns and land-atmosphere interactions. For example, studies have associated variations in the North Atlantic Oscillation (NAO) with changes and trends in the frequencies of WPs from a range of classifications (e.g. Stefanicki *et al.*, 1998; Casado *et al.*, 2009; Kučerová *et al.*, 2017). In particular, during the 1990s the NAO entered a more positive phase (Ostermeier and Wallace, 2003), which means an increased pressure gradient between the teleconnection's high and low pressure centres near the Azores and Iceland, respectively, resulting in a strengthening of the prevailing westerly flow. Coinciding with this, Kučerová *et al.* (2017) show that WPs featuring zonal (meridional) flow became more (less) frequent over northern and central Europe. Similarly, more frequent northerly WPs over the central and eastern Mediterranean corresponded to an eastward shift of the NAO pressure centres during the same decade (Jung *et al.*, 2003; Beranová and Huth, 2008). To the author's knowledge, there are no studies formally linking changes in WP persistence to changes in large-scale atmospheric circulation (or other physical mechanisms), offering an opportunity for further research. WP persistence, explored as both mean persistence over time (see references in Section 4.1) and longer-term persistence (as analysed in this study) could be linked to variations in atmospheric and oceanic processes thought to influence MSLP in the relevant domains. As highlighted by Kučerová *et al.* (2017), these kind of climatological studies should be done using multiple, and indeed as many WP classification as possible, in order to avoid the over-interpretation of results arising from use of a single classification. This merits investigation, as previous research suggest that there are certain conditions that foster atmospheric persistence. The monthly persistence of low-frequency atmospheric circulation patterns, such as the NAO and other teleconnections, is well known (van den Dool and Livezey, 1984; Barnston and Livezey, 1987; Perlwitz and Graf, 2001). It is thought that this persistence is partly modulated by stratospheric processes, as anomalies in the stratosphere tend to propagate into the troposphere at lead-times of one to two

weeks (Baldwin and Dunkerton, 1998; Thompson *et al.*, 2002). Of particular interest for Europe is how stratospheric anomalies in the northern hemisphere have a similar surface signature to some teleconnection patterns, including the NAO (Baldwin and Dunkerton, 2001).

Land-atmosphere feedbacks may also contribute to modifying WP frequencies of occurrence and persistence. For example, the importance of soil moisture in contributing to heatwaves has been well documented, with major European events, such as that over western Russia in 2010, being preceded by anomalously dry soils (Fischer *et al.*, 2007; Vautard *et al.*, 2007; García-Herrera *et al.*, 2010; Mueller and Seneviratne, 2012; Quesada *et al.*, 2012; Miralles *et al.*, 2014; Hauser *et al.*, 2016). Relatedly, several studies have linked changes in the mean persistence of WPs from the Grosswetterlagen classification (Hess and Brezowsky, 1952) between 1988 and 1997 to severe central European heatwaves in the 1990s (Kyselý, 2002; Kyselý and Domonkos, 2006; Kyselý, 2007). Interestingly, these studies show that the persistence of all WPs increase during this period, not just those typically associated with heatwaves (i.e. anticyclonic WPs), although as Kyselý and Domonkos (2006) note, persistent atmospheric anomalies can support climate anomalies in both directions. However, Cahynová and Huth (2009) questioned the credibility of changes in WP persistence of this classification in the 1980s, as they were not reflected in analyses of other data sets. This again reinforces the point made by Kučerová *et al.* (2017) that studies of this nature should utilise multiple classifications.

This study may be useful for further research in several ways. For example, the methods used to identify and quantify persistence are applicable to any WP classification, making them useful for any region in the world that has such a classification. Furthermore, this chapter demonstrated the potential to build a statistical forecast model for WP occurrences by showing that their transition probabilities change when conditioned on persistence. While physically-based WP forecast models are in use operationally (Neal *et al.*, 2016), there is no data-driven approach to this problem, except that by Fayos and Fayos (2007). Harnessing the persistence shown here may be a viable option for building this model, which may be of value as a way of benchmarking the dynamical systems.

Chapter 5

A statistical approach to weather pattern and precipitation forecasting

5.1 Introduction

WP classifications are potentially useful in weather forecasting as they offer a way of reducing huge and complex information output from dynamical ensemble forecast systems. By classifying predictions of atmospheric circulation into one of a set of discrete types (i.e. the WPs of a chosen classification), the likelihood of occurrence of WPs for each day in the forecast lead-time can be determined. Then, features like the probable date of a major change in circulation (Huth *et al.*, 2008) can be estimated. In fact, according to Huth *et al.* (2008), until a few decades ago weather forecasting was the primary application of WP classifications, but advances in computing capabilities led to changes in methodology and a waning interest in WPs. However, the use of WP classifications in this field has undergone recent resurgence, although published works are still scarce, especially for data-driven methods. The author can find only one study predicting WPs using a purely data-driven approach, that by Fayos and Fayos (2007). That study used self-organising maps (a type of artificial neural network) to classify the similarity between all five-day sequences of LWTs between 1947 and 1994. This allowed identification of the most likely five-day WP sequences that preceded each LWT, and enabled one-day-ahead predictions based on the previous five days. A few other studies, while not explicitly forecasting WPs, have explored the usefulness of WPs in prediction of climatic variables by considering "perfect prognostic" methods (Fernández-González *et al.*, 2012; Ramos *et al.*, 2013). In these works, the WP observations were used rather than forecasts, and model skill is evaluated by how well these perfect forecasts translate to predictions of other variables.

In this chapter, the potential of a data-driven model to predict WPs, and subsequently precipitation, over 30-day periods is explored. It is possible that the persistence of the WPs in MO-30 identified in Chapter 4 can provide skill in a statistical forecast model. This chapter builds on the results found in Chapter 4 through the development of a novel approach to WP forecasting based on sampling from historical analogues. Firstly, the model's ability to predict WP occurrences is assessed and compared with the performance of two other methods. Following this, the models are extended with a second component to predict regional UK precipitation, with skill evaluated separately to the WP forecast component.

This chapter is organised as follows. Section 5.2 describes the data sets used, the forecast methods for both WP and precipitation prediction as well as the verification procedure. In

Section 5.3, the results of the forecast models are presented and Section 5.4 investigates the sources of model skill. Finally, Section 5.5 provides some concluding remarks and recommendations for future work.

5.2 Data and methodology

5.2.1 Data

The MO-30 WP data set for 1850-2016 and the Met Office HadUKP daily precipitation data set for 1931-2016 used are the same as described in previous chapters.

5.2.2 Weather pattern forecast methods

Chapter 4 showed that the occurrence of WPs over a daily to weekly window can influence the future transition probabilities of the WP system. A prior window of this time-scale is built into a forecast model based on historical analogues. Historical analogues form the basis of Ensemble Streamflow Prediction (Day, 1985), where hydrological models are driven by past observations that may be weighted based on different criteria such as climate indices or GCM forecast output (Hamlet and Lettenmaier, 1999; Carpenter and Georgakakos, 2001; Yao and Georgakakos, 2001; Werner *et al.*, 2004; Wang *et al.*, 2011). The concept of analogues for the atmosphere has also been studied, with the hypothesis that similar circulation patterns should yield similar local (weather) conditions (Lorenz, 1969; Obled *et al.*, 2002). This method has been used to predict precipitation for a network of European river catchments between 1953 and 1998, using seasonal analogues of 500 hPa geopotential height over Europe from which an empirical conditional precipitation distribution is used for the forecast (Obled *et al.*, 2002). Also, Caron *et al.* (2018) used a historical analogue approach to predict European summer drought for 1981-2010 as a benchmark for forecasts by the ECMWF seasonal forecast system 4 (Molteni *et al.*, 2011), finding little difference in skill.

Here, the set of WPs that occurred in the k-day window prior to the forecast initialisation date is considered. Then the historical time series (up to the date of the forecast) is searched for analogues i.e. windows of length k days in which the set of WPs is the same as for the target window. The WP that follows each of these analogue periods then forms the forecast sample from which the predicted WP can be randomly selected. Taking the set of WPs ignores the frequencies and the order of the WPs but has the advantage of providing a larger sample size of analogues from which to forecast. Using an ensemble approach by randomly selecting a desired number of WPs from the forecast sample provides a probabilistic forecast. Should there not be any identical analogues, then the WP series is instead searched for periods where the set

of WPs only differs to the target set by one WP. The former type of analogue is termed as an 'identical set' (IS) and the latter as a 'one-different set' (OD). Should there be no IS or OD then the next WP is predicted using the non-homogeneous first-order Markov chain (as described in Chapter 4.2.4). This method as a whole is called "the similarity method" (SMk, where k is the size of the prior window) and can be written as follows:

1. Set time $t = t_0$, the forecast initialisation date and let a WP at time t be denoted as $W(t) = i$, for $i = 1, \dots, 30$. Define the number of ensemble members, N.

2. For the latest k days $t - k + 1, \dots, t$, let $H(t) = \big(H_1(t), \dots, H_{30}(t)\big)$, where the $H_i(t)$ are the frequencies of each WP in this window (therefore $\sum H_i(t) = k$).

3. Find the set of WPs that occurred in this window $B(t) = \big(B_1(t), \dots, B_{30}(t)\big)$, with
$$B_i(t) \coloneqq \begin{cases} 1 \; if \; H_i(t) > 0, \\ 0 \text{ otherwise.} \end{cases}$$

4. For all $\tau < t - k + 1$ find $B(\tau)$. If there is at least one instance where $B(\tau) = B(t)$ (i.e. IS), randomly select one of these sets (i.e. one of these τ), set $W(t + 1) = W(\tau + 1)$ and then go to step 7.

5. If there is no $B(\tau) = B(t)$, establish whether there is a period satisfying $|B(\tau) - B(t)| = 1$ (i.e. OD). If so, randomly select one of these sets, set $W(t + 1) = W(\tau + 1)$ and go to step 7.

6. If there is no $|B(\tau) - B(t)| = 1$, generate $W(t + 1)$ using the first-order Markov transition probabilities as described in Chapter 4.2.3.

7. Repeat steps 2 to 6 N times.

8. Set $t = t + 1$ and return to step 2 until the desired forecast horizon is reached.

Forecasts of past observations (reforecasts, sometimes referred to as hindcasts) using SM5, SM10 and SM15 are compared with two other methods. The first is using the non-homogeneous (i.e. month-dependent) first-order Markov transitions (MC; Chapter 4.2.4). Note that in any cases when there are no IS or OD analogues for SM to use, it defaults to Markov transitions and so is equivalent to MC. The second method is a benchmark (or naïve) forecast that randomly chooses a WP based on the observed frequencies of each WP for the month of the reforecast (FM); this model may be thought of as the WP 'frequencies of occurrence' climatology. As WP frequencies and transition probabilities are implicitly linked, there will likely be numerous

occasions when FM and MC perform to a similar level of skill. SM, FM and MC are all performed as an ensemble of 1000 members.

As the focus is on droughts, reforecast initialisation dates are chosen to be those at the start of particularly dry 30-day periods. For each region and season the two 30-day periods with the lowest precipitation are selected for a total of 72 reforecasts. The reforecast periods are chosen such that they do not overlap. Should two regions have the same such reforecast period then the next lowest precipitation period from one of the regions is taken. To ensure that data during the reforecast period is not used to train the model, only data available up until the reforecast initialisation date is used. To guarantee that at least 30 years of precipitation data is available for verification purposes, only dates from 1961 onwards are made available for selection.

5.2.3 Precipitation forecast method

The daily precipitation data are discretised into m bins with historical probabilities p_b for $b = 1, \dots, m$. Dry days (zero precipitation) form one bin and bin intervals increase for higher precipitation values, see Table 5.1. This gives a discrete distribution of precipitation interval frequencies, $D(z)$, with conditional distributions for each WP given by $D(z|W = i), i = 1, \dots, 30$. Additionally, w 30-day summed precipitation intervals s_c for $c = 1, \dots, w$ are defined (Table 5.1). Forecast probabilities of these 30-day intervals are derived from the WP forecast models as follows:

1. Set the ensemble member $e \in (e_1, \dots, e_N)$, where N is the number of ensemble members; time $t = t_0$, the first day of the forecast, and then the predicted WP by ensemble member e at time t is $W_e(t) = i$ for $i = 1, \dots 30$.

2. Set $p_0 = 0$, calculate the probabilities $p_1, \dots p_m$ of each of the m daily precipitation bins from the discrete precipitation distribution that is conditional on $W_e(t)$ and the current season, S, denoted $D(z|W = i, S = S^*)$. The seasons are defined as winter (December through February), spring (March through May), summer (June through August) and autumn (September through November).

3. Define the maximum value of each bin as $l_{p_b}, b = 1, \dots m$, with $l_{p_0} = 0$. Note that $l_{p_0} = l_{p_1} = 0$, ensuring zero precipitation days can be simulated.

4. Generate n random variables $p_k^* \sim U(0,1)$ for $k = 1, \dots, n$.

5. For each p_k^*, find the index q such that

$$\sum_{j=0}^{q} p_j < p_k^* < \sum_{j=0}^{q+1} p_j \,.$$

Set $P_q = \sum_{j=0}^{q-1} p_j$ and $P_{q+1} = \sum_{j=0}^{q} p_j$, the cumulative probabilities of the bins adjacent to p_k^*.

6. Define the difference between the adjacent bins as $\alpha = P_{q+1} - P_q$ and the difference between the random number and the lower cumulative probability as $\beta = p_k^* - P_q$.

7. Estimate the precipitation value for each p_k^* as $x_k(t) = l_{p_q} + \frac{\beta}{\alpha}\left(l_{p_{q+1}} - l_{p_q}\right)$. There is now n predicted daily precipitation values at time t, $\boldsymbol{x}(t) = (x_1(t), \dots, x_n(t))$.

8. Set $t = t + 1$ and repeat steps 3 to 6 until the final day of the forecast, $t_{\max}$, is processed.

9. Sum the daily precipitation vectors and divide by the random-sample size $(\sum_\tau \boldsymbol{x}(t))/n$ for $\tau = t_o, \dots, t_{\max}$ and discretise according to the w summed precipitation bins $s_1, \dots s_w$ to obtain a distribution of relative frequencies for this ensemble member $\boldsymbol{f_e} = (f_1, \dots, f_w)$.

10. Set a new ensemble member $e^* \in (e_1, \dots e_N), e^* \neq e$ and repeat steps 2 to 9 until every ensemble member has been processed.

11. Sum each ensemble member's distribution of summed precipitation relative frequencies and divide by the number of ensemble members to obtain a final forecast probability distribution:

$$\boldsymbol{F} = \left(\sum_e \boldsymbol{f_e}\right)/N.$$

The number of ensemble members used in the WP forecast models is $N = 1000$ and also $n = 10{,}000$. The precipitation climatology (PC; expressed as frequencies of 30-day summed precipitation bins) of the month corresponding to the reforecast initialisation date is used as a benchmark method for comparison with SM, MC and FM. As with WP forecasts, only data up until the forecast initialisation date is used.

5.2.4 Forecast verification

Forecast verification is done separately for the WP and precipitation forecast components, as the inability to predict WPs well would not necessarily lead to a poor precipitation forecast - an incorrectly forecast WP might have a similar precipitation distribution to the observed WP. In

this case, the WP model would have poor skill but the precipitation model may perform better. This analysis will focus on the predictive skill over a reforecast horizon of 30 days, as the objective is to provide monthly WP and precipitation outlooks that are more applicable to general water resources management and drought planning than flood prediction.

Daily precipitation		30-day precipitation sums	
p_b	Range of precipitation, x (mm)	s_c	Range of summed precipitation, y (mm)
p_1	0	s_1	$0 \leq y < 10$
p_2	$0 < x \leq 1$	s_2	$10 < y \leq 20$
…	Intervals of 1mm	…	Intervals of 10mm
p_{11}	$9 < x \leq 10$	s_{25}	$240 < y \leq 250$
p_{12}	$10 < x \leq 15$	s_{26}	$250 < y \leq 300$
p_{13}	$15 < x \leq 20$	…	Intervals of 50mm
p_{14}	$20 < x \leq 30$		
…	Intervals of 10mm		

Table 5.1: Range of daily precipitation, x, for each bin p_b and of 30-day precipitations sums, y, for each bin s_c.

The Jensen-Shannon divergence (JSD), suitable for measuring the distance between two probability distributions (Lin, 1991), is used to evaluate WP forecast skill. It is based on information entropy, which is used to measure uncertainty. An information-theoretic approach to verification is not widespread, although there is some published research on the topic (Leung and North, 1990; Kleeman, 2002; Roulston and Smith, 2002; Ahrens and Walser, 2008; Weijs *et al.*, 2010; Weijs and Giesen, 2011). The JSD measures the forecast performance by quantifying the distance between distributions of the observed and forecast WP frequencies. The JSD is based on the Kullback-Leibler divergence (KLD) (Kullback and Leibler, 1951). Let P and Q be two discrete probability distributions. The KLD from Q to P is given by:

$$D_{KL}(P||Q) = -\sum_{i=1}^{W} P_i \log_2 \frac{Q_i}{P_i},$$

Equation 5.1

measured in bits (i.e. a binary unit of information). $D_{KL}(P||Q)$ is only defined if $Q_i = 0 \Rightarrow P_i = 0$. Furthermore, as $\lim_{x \to 0} x \log x = 0$, whenever $P_i = 0$ the i^{th} term of $D_{KL}(P||Q)$ is taken as zero.

$D_{KL}(P||Q) \geq 0$ and $D_{KL}(P||Q) = 0 \Rightarrow P = Q$. In this application $W = 30$, the number of WPs and $P = (p_{f,1}, \dots, p_{f,30})$ and $Q = (q_{f,1}, \dots, q_{f,30})$ are the vectors of observed and forecast WP relative frequencies, respectively. (Because these are relative frequencies, $\sum P = 1$ and $\sum Q = 1$.) As there would inevitably be some cases where the model forecasts no occurrences of some WPs (i.e. when Q contains zeros), $D_{KL}(P||Q)$ will be undefined at times. Using the JSD avoids this problem; it is defined as:

$$D_{JSD}(P||Q) = \frac{1}{2} D_{KL}(P||M) + \frac{1}{2} D_{KL}(Q||M),$$

Equation 5.2

where $M = (P + Q)/2$. Unlike the KLD, the JSD is symmetric i.e. $D_{JSD}(P||Q) \equiv D_{JSD}(Q||P)$. Also, $0 \leq D_{JSD}(P||Q) \leq 1$, with a score of zero indicating P and Q are the same (a perfect forecast).

A measure suitable for evaluating precipitation forecast performance is the ranked probability score (RPS) (Epstein, 1969; Murphy, 1971), which is an extension of the Brier Score for multiple event categories (Wilks, 2011). As the predictand (summed precipitation bins) is ordinal, RPS is appropriate due to its sensitivity to distance (Staël von Holstein, 1970). For J event categories, define a vector of forecast probabilities $\boldsymbol{p} = (p_1, \dots, p_J)$ and a binary vector $\boldsymbol{o} = (o_1, \dots, o_J)$ indicating which event category occurred. Then the cumulative forecasts and observations are given by $P_e = \sum_{i=1}^{e} p_i$ and $O_e = \sum_{i=1}^{e} o_i$, for $e = 1, \dots, J$, respectively. For a single forecast:

$$RPS^* = \sum_{e=1}^{J} (P_e - O_e)^2$$

Equation 5.3

which is represented as a skill score, $RPSS^*$, by

$$RPSS^* = 1 - \frac{RPS^*}{RPS^*_{ref}},$$

Equation 5.4

where RPS^*_{ref} is the score of the reference forecast using the climatological event category frequencies (PC). To jointly evaluate N forecasts:

$$RPS = \frac{1}{N}\sum_{k=1}^{N} RPS_k^*,$$

Equation 5.5

and

$$RPSS = 1 - \frac{RPS}{RPS_{ref}}.$$

Equation 5.6

A perfect score is achieved when $RPSS = 1$, which is also the upper limit. Negative (positive) values indicate the forecast is performing worse (better) than the reference forecast.

5.3 Results

The reforecast periods are referred to as Hd, for $d = 1, \ldots, 72$, ordered first by season and then by time. H1 through H18 correspond to winter, H19 through H36 to spring, H37 through H54 to summer and H55 through H72 to autumn. See Table 5.2 for details.

5.3.1 Features of the similarity method

Unsurprisingly, longer prior windows (larger k) mean there are fewer occasions when SM can use the IS or OD components as there is a lower likelihood of there being an analogue in the historical record. Even when there are some analogues available, the samples from which WP predictions are randomly selected are smaller in size. A hypothesised corollary is that as k is increased, SM becomes more similar in skill to MC, as it will increasingly use the Markovian component. Table 5.3 shows the mean and standard deviation, over all Hd, of the number of times the SM models use IS, OD or Markov components over the 30,000 individual daily reforecasts (1000 ensemble members over a 30-day reforecast period). For five-day windows, almost 80% of SM runs utilise IS, diminishing to around 30% for 15-day windows. For this largest window (SM15), the Markov component is used most frequently (on average across the reforecasts); SM15 should therefore be closer in forecast skill to MC than SM5 or SM10. This is not reflected in the reforecast results, as differences between the SM variants and MC are small in magnitude (Figure 5.1).

The sample sizes of IS and OD decrease as k increases (Figure 5.2), with a much larger difference between $k = 5$ and $k = 10$ than between $k = 10$ and $k = 15$. On average, half of IS sample sizes are below 17 for SM5, nine for SM10 and four for SM15. For OD, half of the sample sizes are below 28.5 for SM5, 12 for SM10 and seven for SM5. While the use of IS,

OD and Markov components clearly differs between SM variants, it does not seem to have much effect on model performance in spring or autumn. For winter (summer) the Markov component predicts different WPs to those following the analogues at lower (higher) skill (Figure 5.3).

Season	Reforecast ID	Initialisation date
Winter	H1	27/02/1961
	H2	20/02/1962
	H3	30/12/1962
	H4	02/02/1963
	H5	27/01/1965
	H6	09/02/1968
	H7	26/02/1969
	H8	14/12/1972
	H9	25/02/1973
	H10	22/02/1984
	H11	31/01/1986
	H12	10/01/1992
	H13	19/02/1993
	H14	20/12/1996
	H15	21/01/1998
	H16	10/02/2004
	H17	28/02/2011
	H18	28/02/2012
Spring	H19	08/05/1961
	H20	30/05/1962
	H21	18/05/1968
	H22	23/05/1970
	H23	28/03/1974
	H24	01/04/1976
	H25	05/05/1978
	H26	19/04/1980
	H27	28/05/1982
	H28	19/04/1984
	H29	28/04/1990
	H30	04/05/1991
	H31	24/03/1997
	H32	26/03/2002
	H33	22/03/2003
	H34	22/03/2007
	H35	02/05/2008
	H36	05/04/2011
Summer	H37	16/07/1968
	H38	09/08/1972
	H39	07/06/1975
	H40	08/08/1976
	H41	15/06/1977
	H42	27/06/1979
	H43	11/06/1981
	H44	10/08/1981
	H45	03/07/1983
	H46	27/06/1984
	H47	19/06/1986
	H48	09/07/1989
	H49	13/07/1990
	H50	11/06/1995
	H51	23/07/1995
	H52	27/08/1996
	H53	24/08/2007
	H54	31/08/2014
Autumn	H55	24/11/1963
	H56	12/10/1964
	H57	19/09/1969
	H58	05/09/1971
	H59	08/09/1972
	H60	01/10/1978
	H61	25/10/1983
	H62	08/10/1985
	H63	06/09/1986
	H64	19/11/1987
	H65	02/09/1989
	H66	11/11/1989
	H67	30/11/1995
	H68	10/09/2002
	H69	26/09/2007
	H70	05/09/2009
	H71	26/11/2010
	H72	12/11/2013

Table 5.2: ID number, season and initialisation date of the 72 reforecasts.

SM component	SM5	SM10	SM15
Identical Set (IS)	23497.12 (1041.47)	15576.24 (3841.83)	9174.03 (6052.46)
One-Different Set (OD)	6253.32 (999.61)	10667.75 (2427.33)	10029.97 (2877.8)
Markov	249.56 (56.64)	3756.01 (2410.07)	10796.0 (6927.11)

Table 5.3: Mean and standard deviation (in brackets) of number of times identical set (IS), one-different set (OD) and Markov components of SM methods (for k = 5, 10 and 15) are used in reforecasts.

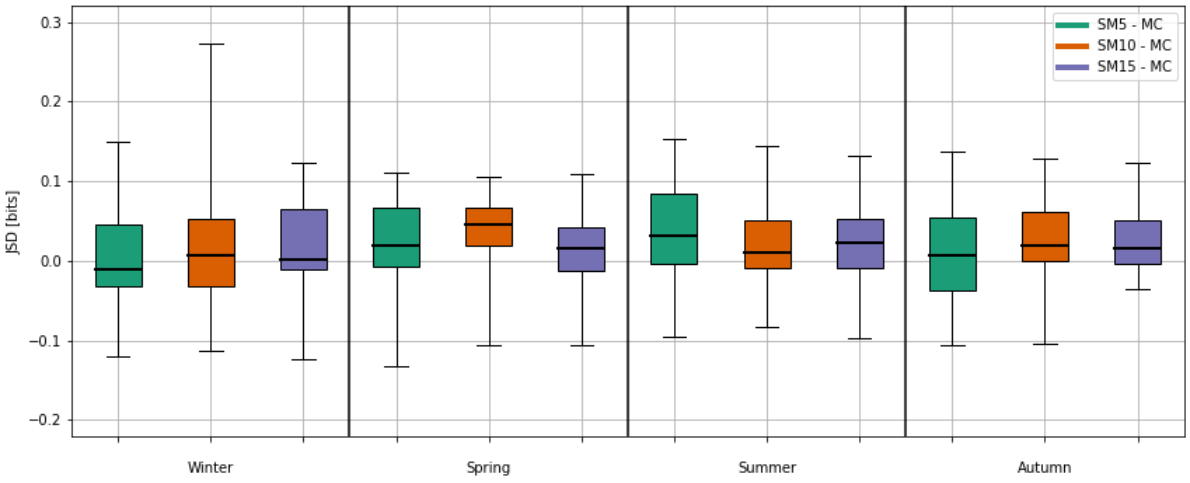

Figure 5.1: Boxplots of differences in Jensen-Shannon Divergence (JSD) between each SMk and MC. Whiskers are the range of the data.

Changing the size of the *k*-day window for SM does not have much effect on reforecast performance except in winter. JSD results for SM5, SM10 and SM15 are similar in spring and autumn, while SM15 is marginally the most skilful in summer and SM5 is the most skilful in winter (Figure 5.3). When there are large differences in skill between the three variants, SM5 tends to outperform at least one of the other models in winter and autumn but is typically the least skilful in summer (not shown). Figure 5.4a shows that for H68, none of the models predicted the standout feature – the dominance of WP9. The added skill of SM5 over the other two models appears to come from its better prediction of the number of times WP16, and to a lesser extent WP27, occurred. A large number of different WPs occurred in the 15-day period prior to this reforecast (Figure 5.4b). As such it is likely that the weaker performance of SM10 and SM15 is due to fewer IS and OD analogues available in the training period (the five days prior to H68 contains four different WPs, increasing to seven WPs for 10 days prior and eight

WPs for the full 15 days). This results in behaviour more similar to MC, which correspondingly performs worse for this reforecast than all SM models (not shown).

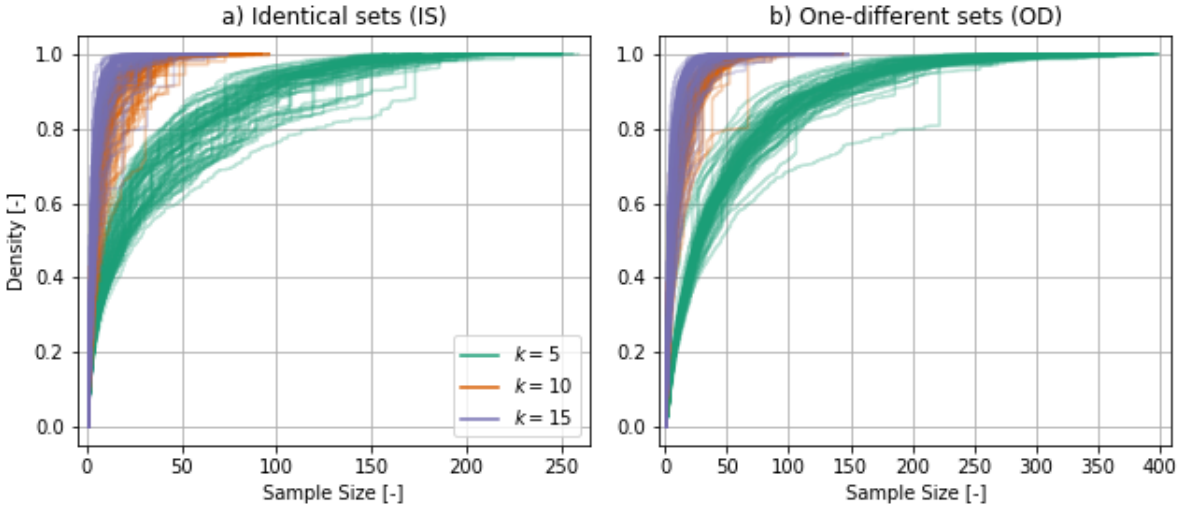

Figure 5.2: ECDF of sample sizes used in each a) identical set (IS) and b) one-different set (OD) component of SMk for k = 5, 10 and 15.

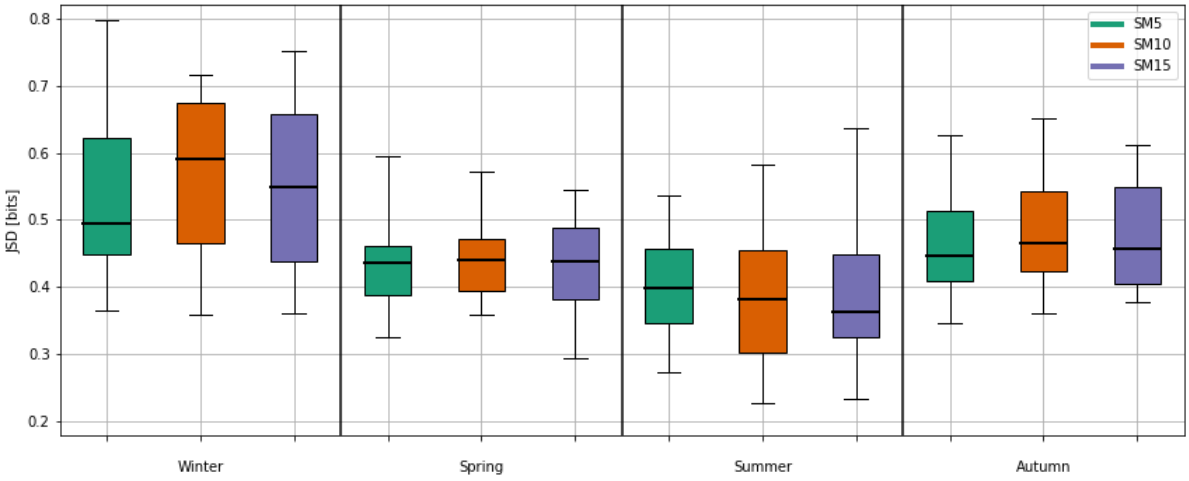

Figure 5.3: Boxplots of SM5, SM10 and SM15 Jensen-Shannon Divergence (JSD) for the reforecast periods. Whiskers are the range of the data.

On the other hand, higher *k* improves reforecast skill for H40 (Figure 5.6a). The high frequency of WP6 is best predicted by SM15, implying that the inclusion of WP13 in the prior window precipitated higher likelihood of a WP6 forecast. SM15 over-predicted WP9 compared to the other models, surely a result of the prior window including WP9 on more occasions that the other models (WP9 would be in the prior window for at least the first six reforecast days for SM15 compared to once and no times for SM10 and SM5, respectively). Note, however, that the influence of WPs in the prior window can only definitively be discussed for the first *k* days of the reforecast period, as the models update daily according to their own predictions, not the

observations. This point regarding the number of model runs for which WPs remain in prior windows shall be returned to later.

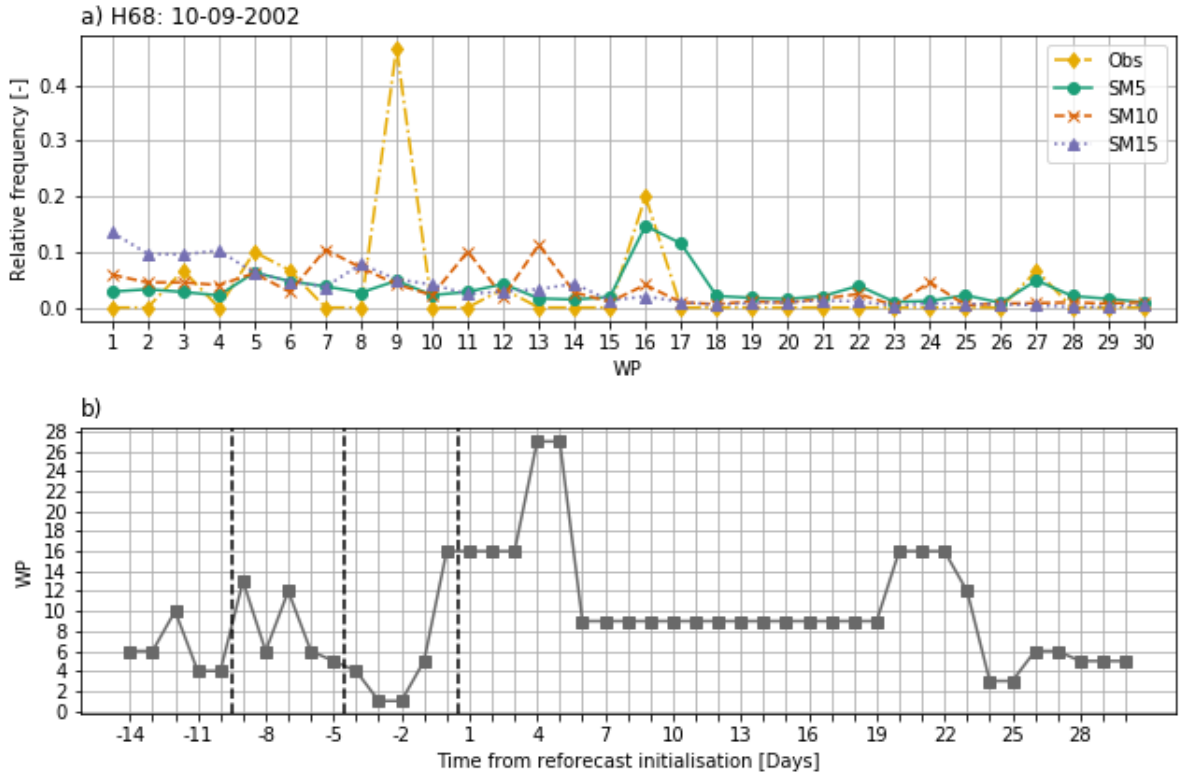

Figure 5.4: a) Relative frequencies of WPs for H68 for observations and SM5, SM10 and SM15 reforecasts; b) observed WP time series for H68 and preceding 15 days. Vertical dashed lines delineate the prior windows for each SMk.

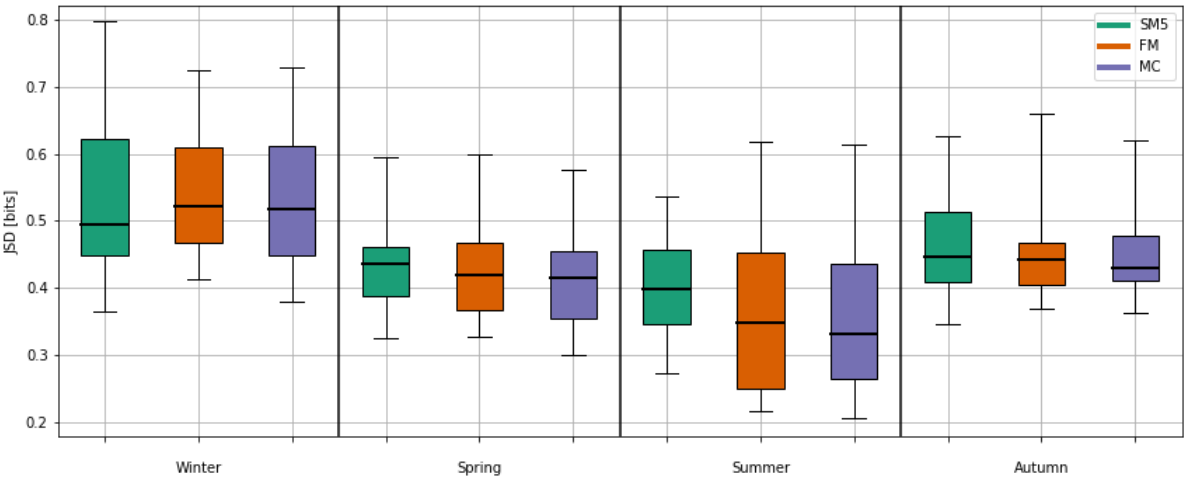

Figure 5.5: Boxplots of SM5, MC and FM Jensen-Shannon Divergence (JSD) for the reforecast periods. Whiskers are the range of the data.

5.3.2 Weather pattern reforecasts

For brevity, this chapter will focus on SM5 and only relay the key points regarding SM10 and SM15. Overall, SM5, FM and MC yield similar levels of reforecast skill across all seasons except summer (Figure 5.5). In summer, MC has marginally better skill than the benchmark, FM, with SM5 clearly the worst. SM10 and SM15 have skill scores closer to FM and MC in this season, but are still not as skilful (not shown). On occasions when SM5 has considerable skill over FM and MC, skill for all models tends to be lower compared to other reforecasts in that season e.g. H3, H30, H53 and H71 (not shown). For example, none of the models accurately predict the high frequency of WP13 and WP25 for H53 (Figure 5.7a). For MC and FM this is because these two WPs do not occur in summer often. The added skill of SM5 over the other models in this case is mostly due to the comparably lower reforecast frequencies for the majority of typical summer WPs (WP1 through WP11). As such, SM5 provides a more skilful reforecast because of the WPs it does not predict, rather than the WPs it does predict.

This reforecast is the exception rather than the rule; when SM5 outperforms MC and FM it is mostly also due to correctly predicting higher frequencies of WPs, such as WP6, WP10 and WP13 for H65 (Figure 5.8). For this reforecast, SM5 did not predict the increased occurrence of WP21 and WP25. This is likely a result of these WPs occurring (in the observed series) toward the end of the reforecast period. At this stage, the daily WP reforecasts are updated with information from the reforecasts themselves, not observations, and are therefore subject to a wider range of WPs in the prior windows and hence more variable predictions.

H45 (Figure 5.9) is a good example of why SM5 tends to be the least skilful model in the summer, with lower (higher) predicted frequencies of WPs that mostly occur in summer (winter). SM5 better predicts the frequencies of only nine WPs, and is particularly poor at reforecasting the predominance of WP1. The simpler models are more skilful than SM in this season probably because there are fewer WPs that tend to occur than in winter and so climatological and transition probability-based forecasts are less likely to diverge significantly from the observations.

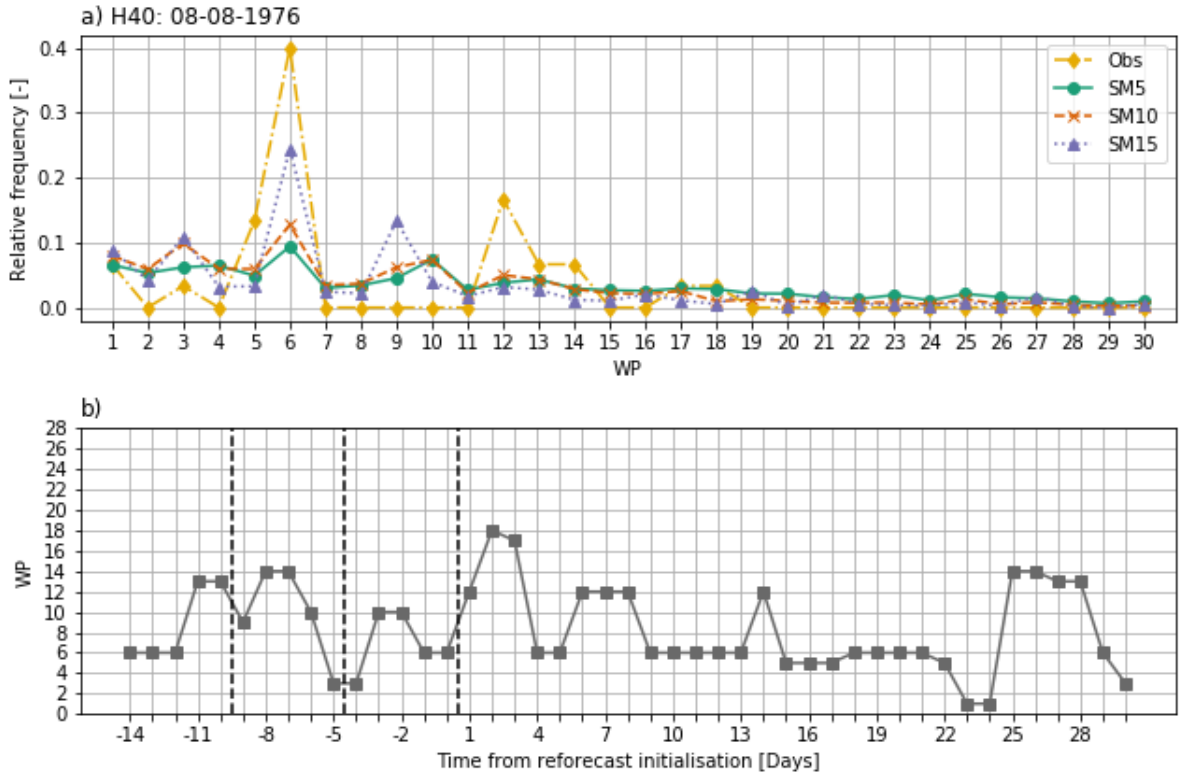

Figure 5.6: As for Figure 5.4 but for H40.

SM predominantly predicts the next day's WP as one of the WPs in the previous k days. That is, the observed WPs following the historical analogues used in the SM method are generally one of the WPs in the analogue. Consider H53 (Figure 5.7) as an example. The prior window for the first reforecast day features WP1, WP6 and WP19 and as a result, roughly 60% of the ensemble members predict one of these WPs for the next day (Figure 5.7c). The reforecast frequency of WP1 drops sharply for the second day, as it would no longer be present in prior windows of those ensemble members that did not predict WP1 for the first reforecast day (Figure 5.7b). Likewise, the reforecast frequencies for WP19 are much smaller on the fourth day. On the other hand, as WP6 is in the prior window for at least the first five days of reforecasts (more for those ensemble members that also predicted WP6), the predicted frequencies remain high. SM5 therefore accounts for "interrupted persistence" of WPs. That is, rather than only predicting continued persistence via uninterrupted sequences of a single WP, the use of unordered sets in SM enables the model to predict WPs that are not necessarily the most recently occurring prior to the reforecast date.

This is conceptually similar to a k^{th}-order Markov model, for which there would be, in some cases, higher transition probabilities to those WPs that had been observed in the previous k days. Furthermore, the order of the WPs in the prior window is implicitly accounted for. From the first forecast day the ensemble begins to diverge, but the first k model runs include at least

one observed WP in the prior window, and all ensemble members will include these WPs. Therefore, the most recent WP observations are guaranteed to be in the prior window for longer than the preceding WPs, meaning these WPs are more likely to be predicted in the early stages of the forecast period.

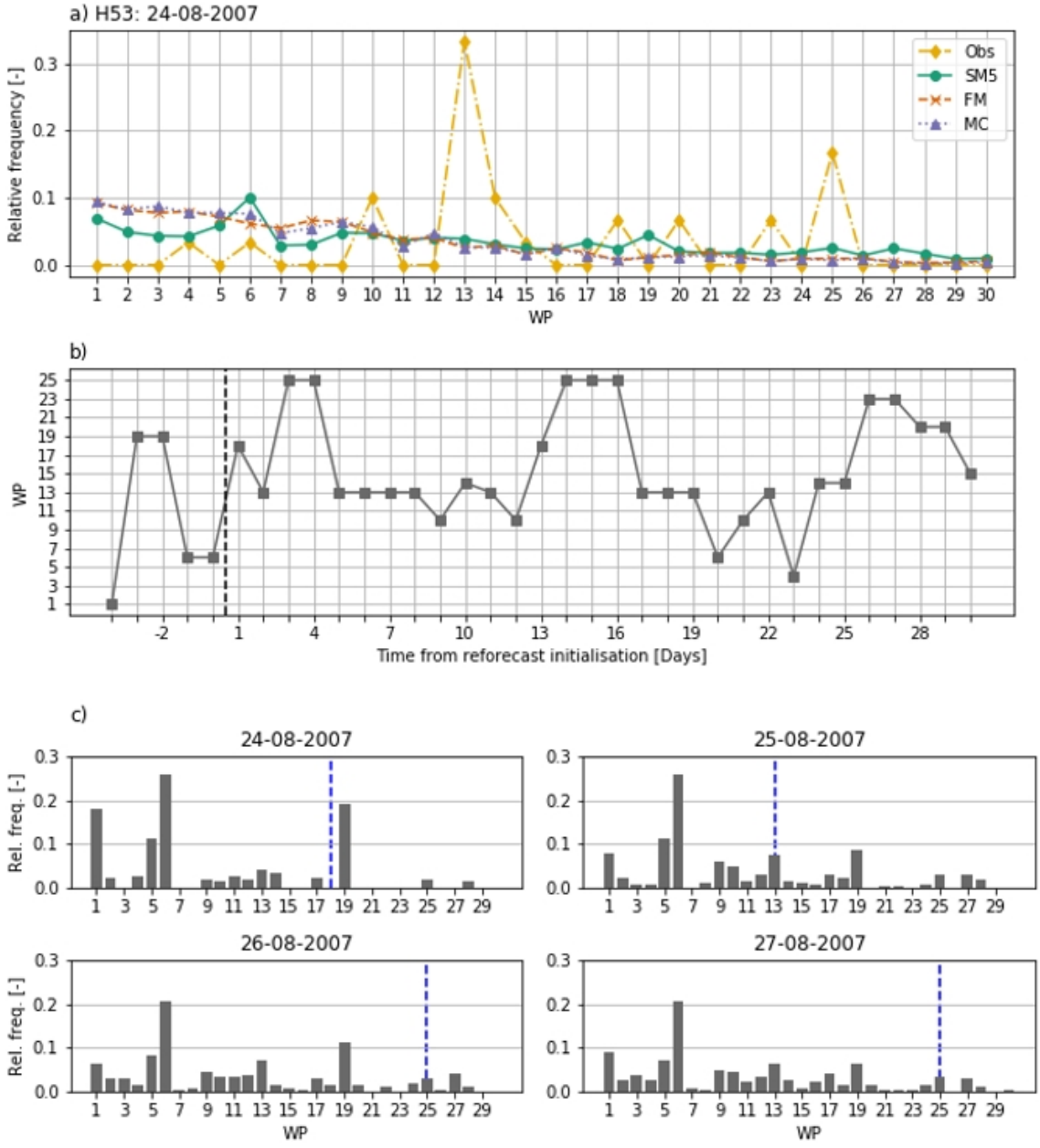

Figure 5.7: a) Relative frequencies of WPs for H53 for observations and SM5, FM and MC reforecasts; b) observed WP time series for H53 and preceding 5 days, with vertical dashed lines delineating the prior window and reforecast period; and c) SM5 reforecast probabilities of each WP for each of the first four days in the reforecast period. Dashed blue vertical lines indicate the observed WP.

SM is capable of producing unusual WP frequencies in the forecast, whereas FM and MC predict the climatological average. This is related to the concept of forecast *sharpness* (Allan, 1993), which is a quality reflecting a forecast system's ability to predict events that deviate

from climatology. Forecast sharpness is clearly seen in some of the reforecast results. Reforecast frequencies for H65 (Figure 5.8) show this feature, with SM5 better predicting the spikes in frequency of WP6, WP10 and WP13. FM and MC, on the other hand follow the clear climatological pattern of lower frequencies for higher-numbered WPs. Similarly, SM5 correctly reforecasts increased frequencies of WP10 in H45 (Figure 5.9), although this model is less skilful than MC and FM. H53 (Figure 5.7) shows increased sharpness is not always advantageous, with the climatological models' predictions of WP6 frequency more accurate than from SM5, despite the latter identifying the slight spike in frequency. The sharpness of SM is tied to its habit of predicting WPs from the prior window, as described in the previous paragraph: the WPs with higher reforecast frequencies are typically those in the prior window.

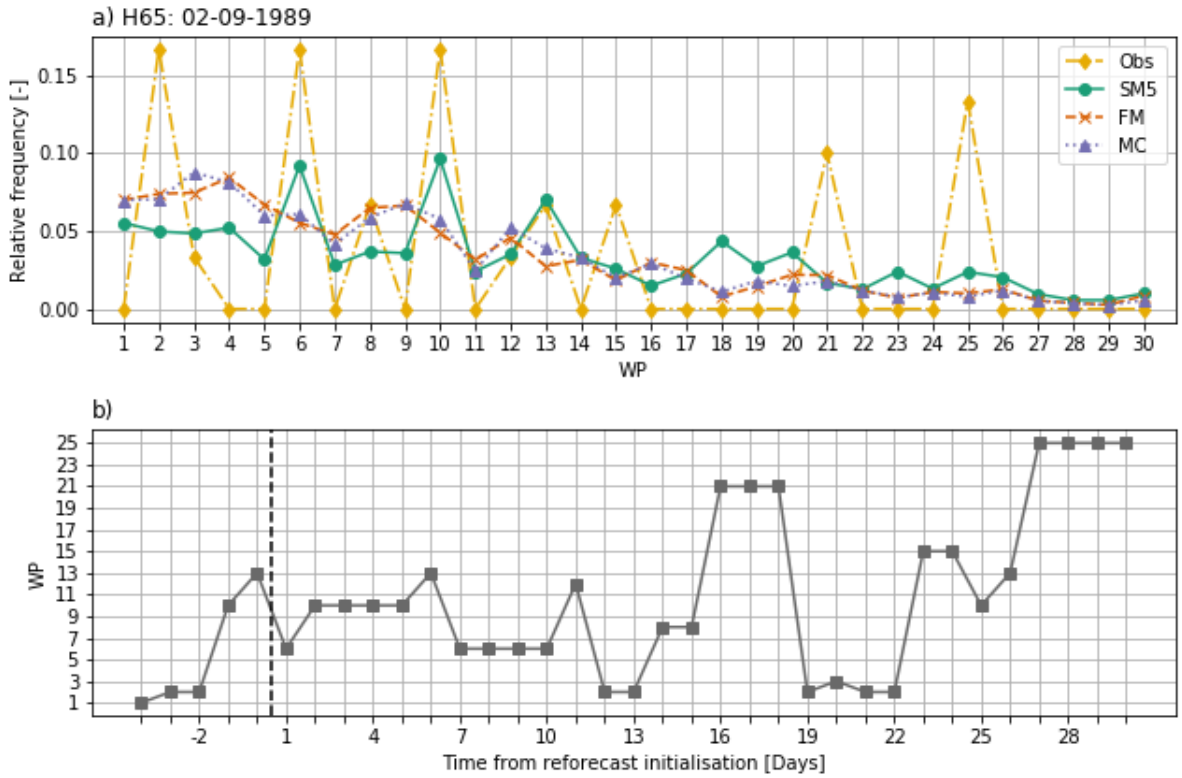

Figure 5.8: As Figure 5.7 (a) and (b) but for H65.

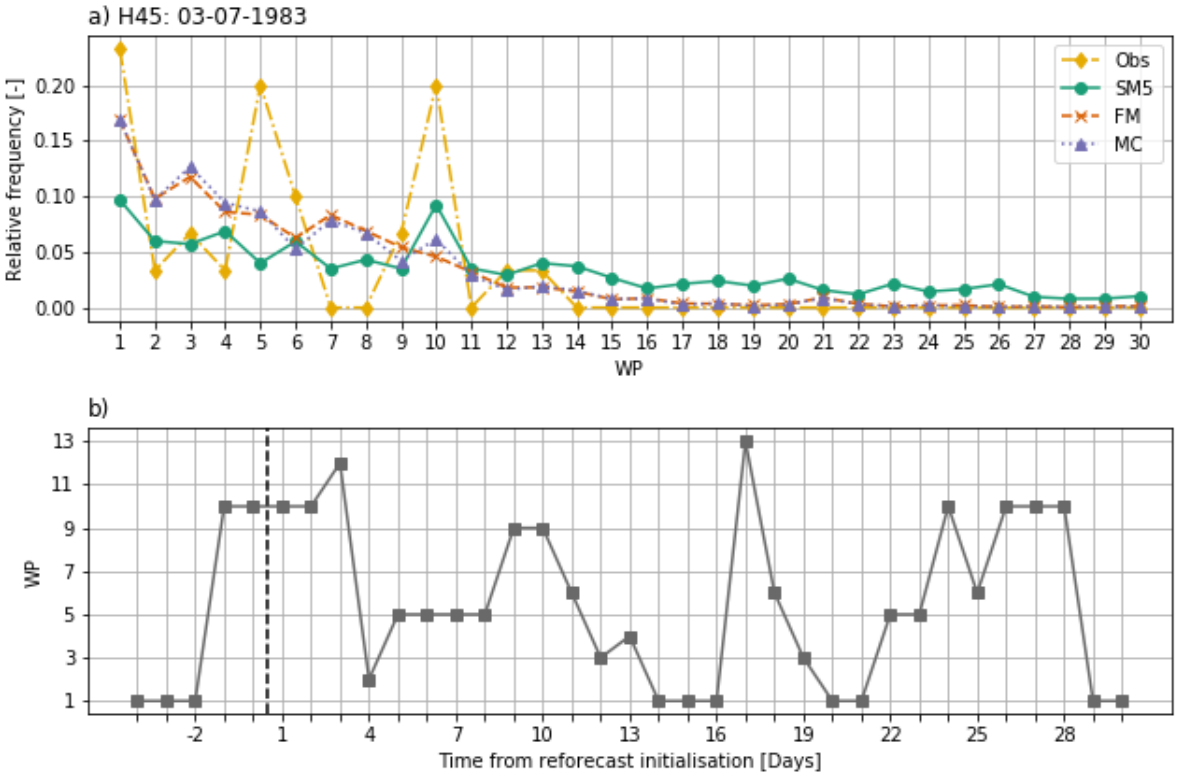

Figure 5.9: As Figure 5.7 (a) and (b) but for H45.

5.3.3 Precipitation reforecasts

Precipitation climatology is a more skilful prediction method than any of the forecasts derived from the WP forecast models. No model has $RPSS > 0$ for any region (Figure 5.10), indicating that, on average, PC outperforms SM, MC and FM. There are some cases when the skill score for an individual reforecast, $RPSS^*$, is greater than zero for SM5 (Table 5.4). When SM5 outperforms PC, it follows that the WPs predicted are associated with precipitation amounts closer to the observed than climatology. This is not to say the WP predictions are necessarily skilful themselves – the forecast WPs might be different to the observations but have similar precipitation distributions. For example, the least skilful summer WP reforecast, according to JSD, is H53, for which SM5 is unable to capture the high frequency of WP13 (Figure 5.7a). However, for the precipitation forecasts of this period, SM5 outperforms PC in over half of all regions (Table 5.4). As such, in terms of resulting precipitation, the underestimation of the frequency of WP13 appears to be compensated for by overestimation of other WPs' frequencies of occurrence, with these overestimated WPs being associated with precipitation distributions that closely match the observed precipitation (in these regions for this reforecast period). The inability of the WP forecast models to estimate precipitation with higher skill than climatology is particularly clear when taking into account the fact that the reforecasts were selected to be

the driest periods on record i.e. they are necessarily dissimilar to climatology. Therefore, PC itself is unlikely to be skilful and the WP models even less so.

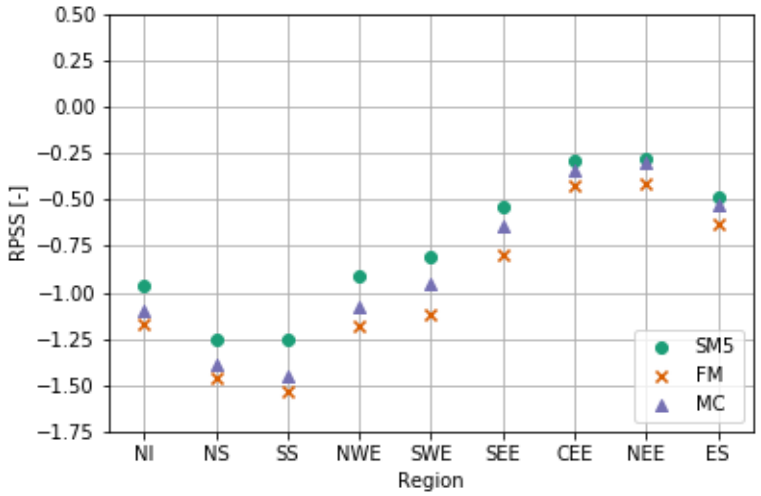

Figure 5.10: Ranked Probability Skill Score (RPSS) by region for SM5, FM and MC for H1 to H72.

Region	Winter	Spring	Summer	Autumn	Total
NI	H3, H5, H15	H27		H67, H68, H71	7
NS	H3, H5, H14	H20, H31	H42, H53	H57, H60, H61, H67, H68, H71	13
SS	H3, H5, H14, H15			H57, H60, H61, H67, H68, H71	10
NWE	H3, H5, H15	H27	H53	H64, H67, H68, H71	9
SWE	H3, H5, H15	H27, H35		H64, H65, H67, H71	9
SEE	H7	H27, H35	H37, H45	H55, H64, H69, H71	9
CEE	H7, H18	H35	H40, H44, H45, H52, H53	H55, H58, H59, H61, H64, H65, H67, H68, H69, H71, H72	19
NEE	H2, H5, H7, H11, H12, H15	H27	H37, H38, H40, H44, H45, H53, H54	H55, H58, H60, H61, H64, H67, H71, H72	22
ES	H3, H5, H11, H12, H15, H17		H37, H40, H44, H45, H48, H49, H53, H54	H57, H60, H61, H64, H65, H67, H68, H71, H72	23

Table 5.4: SM5 reforecasts for which $RPSS^* > 0$ by season, with total number of such cases.

Forecast skill is not spatially or temporally homogeneous. Models exhibit, on average, superior performance in eastern regions (Figure 5.10). For SM5, this is driven by more cases of the model outperforming precipitation climatology ($RPSS^* > 0$) for eastern regions (except SEE

and CEE in winter) than for western regions (Table 5.4). In addition, scores are generally lower overall in spring and summer for western regions. The situations where SM5 is more skilful than PC are most common in autumn and least common in spring (Table 5.4). Given that model skill of the WP forecasts is not vastly different between these two seasons (Figure 5.5), it is likely that the WPs predicted in autumn are more similar to the observations than those in spring, but in terms of their conditional precipitation distributions rather than the actual WPs themselves. On the other hand, some of the highest-skill WP reforecasts occur during summer (Figure 5.5), subject to the aforementioned east-west divide (Table 5.4). This suggests that the WPs can provide a superior representation of low-precipitation periods in the east compared to climatology, but this is not true in the west. Note, however, that these reforecasts with $RPSS^* > 0$ are in the minority; precipitation climatology is still the preferred method for all regions on average.

Over all regions, winter has the second highest number of cases with $RPSS^* > 0$. A possible reason for WP models outperforming precipitation climatology in this season and autumn is the above-average occurrences of easterly WPs (WP27 and WP28) during reforecast periods (Figures 5.11a and d). Many of the reforecast periods dominated by these WPs correspond with $RPSS^* > 0$: H3, H7, H11, H14, H55, H64, H67 and H71 satisfy this for at least two regions, while H2, H4, H6 and H66 satisfy this for one or no regions. It seems likely that WP models have higher skill than precipitation climatology for easterly-dominated reforecast periods because phases characterised by easterly flow are rarer (correspondingly, WP27 and WP28 are the fourth and third least common WPs annually, respectively). Precipitation generally follows a decreasing gradient from northwest to southeast and so on occasions when this does not hold, PC will be less appropriate, whereas SM5 forecasts that predict some easterly WPs will better estimate observed conditions.

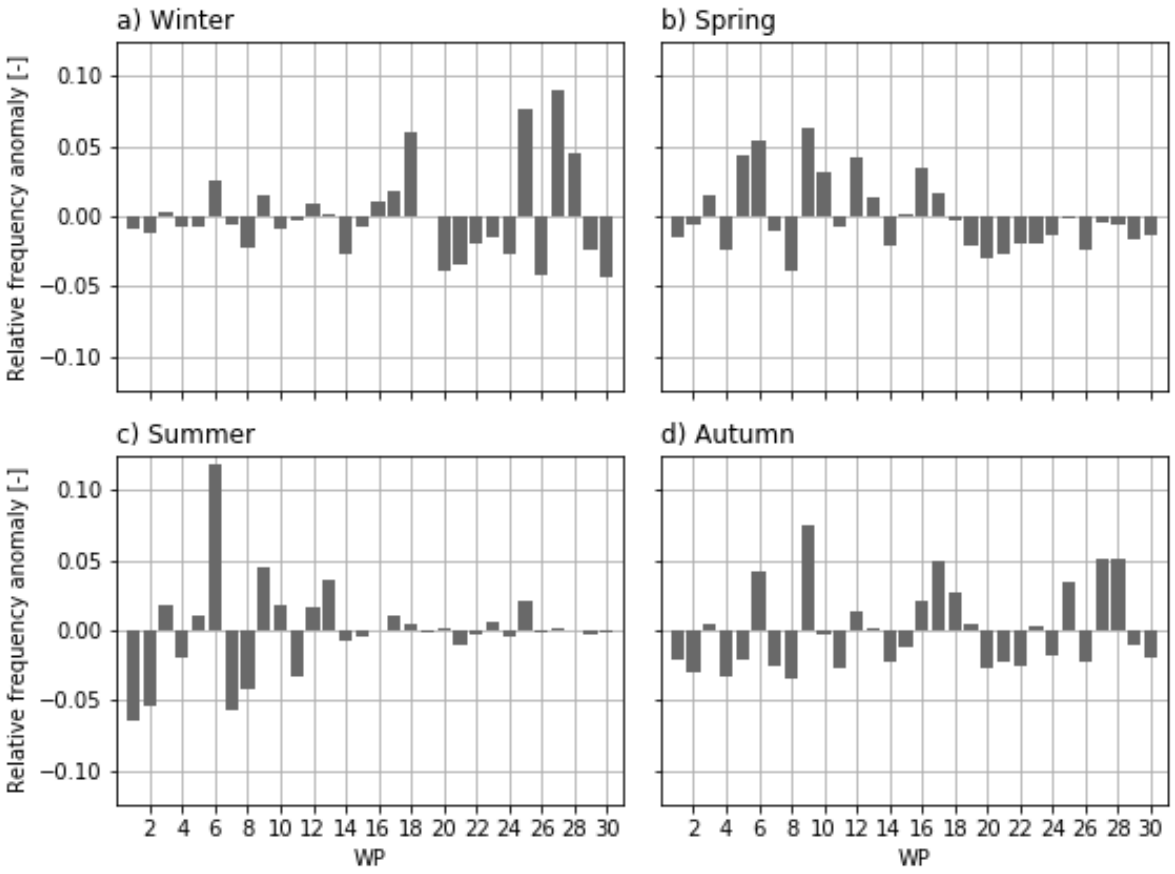

Figure 5.11: Seasonally aggregated relative frequency anomalies of observed WPs during the reforecast periods H1 to H72.

5.4 Investigating model deficiencies

This section attempts to identify reasons behind the overall poor skill of SM's precipitation forecasts. First, model performance may be hindered by the choice of dry reforecast periods, and so the models are run on a new set of 72 reforecast periods, H73 through H144, which observed average precipitation. For each season and each region, two 30-day periods that featured observed precipitation between the 35th and 65th percentiles of the distribution are randomly selected. Half of the reforecast dates are chosen from the first half of the record (1961 through 1989), with the remainder from the second half (1989 through 2016). The methods to predict WPs and precipitation are exactly as described previously for the dry periods.

WP forecast skill from SM5, MC and FM is very similar for average precipitation periods as for dry periods (not shown). As before, there is little to separate the models in terms of overall skill, and the same seasonal differences are evident (not shown). Precipitation forecast skill is higher when forecasting periods of average precipitation for all models in every region except NS (less skill) and SS (very similar skill) (Figure 5.12). This might be surprising, as for these dates, it would be reasonable to expect the benchmark precipitation forecast, PC, to be more

skilful than for drought dates, and resulting model skill scores to be lower (as RPSS scales model skill against PC). As this is not the case, the logical conclusion is that both model (SM5, FM and MC) and benchmark (PC) skill improves when predicting periods of average precipitation, but the magnitude of improvement is greater for the former. Despite this improvement, model skill is still poor overall, with only those for SEE and CEE showing an improvement upon climatology.

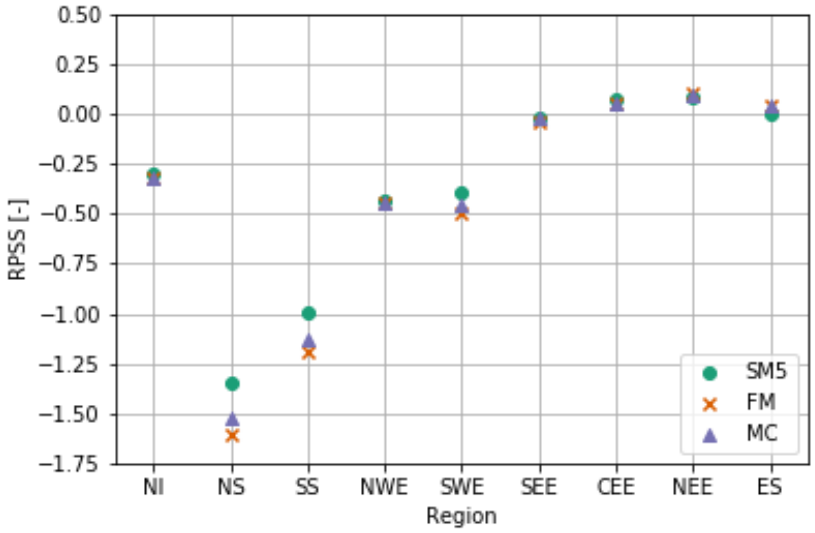

Figure 5.12: Ranked Probability Skill Score (RPSS) for SM5, FM and MC for average precipitation reforecast periods H73 to H144.

Precipitation forecasts from SM5 are derived using the two model components of WP predictions and subsequent precipitation estimation. An attempt to isolate the source of model skill is made by considering a perfect prognostic model. Instead of deriving the precipitation estimates from WPs predicted via historical analogues or otherwise, the observed WPs are used as input. This model, Perfect-WP, is an idealised model rather than one with operational potential. Testing Perfect-WP may offer information on whether it is the lack of skill in predicting WPs that propagates through to poor precipitation forecast skill, or if the procedure to estimate precipitation from the predicted WPs is the issue.

As expected, forecast skill is higher using Perfect-WP for both dry and climatologically average periods (Figure 5.13). For the latter, model skill is lower than precipitation climatology for all but one region, SEE. This suggests that, if forecasting average precipitation is the aim, directly using precipitation climatology is a superior method to utilising WP forecasts, even if the WPs could be predicted exactly. For predicting dry periods, on the other hand, Perfect-WP is more skilful than PC for all regions except NS (less skill) and SS (same skill) (Figure 5.13). The implication is that, when forecasting during periods of non-climatological precipitation,

knowing the WPs that will occur in advance would be of benefit. By extension, at least part of the low model skill from SM5 (and MC and FM) comes from inadequate forecasting of the WPs themselves, and improvements to this model component may provide more skilful precipitation forecasts.

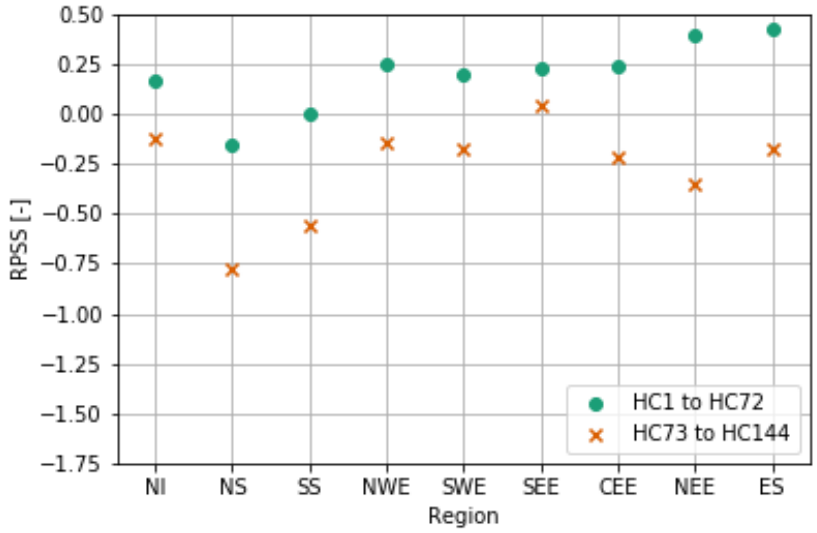

Figure 5.13: Ranked Probability Skill Score (RPSS) for Perfect-WP for dry reforecast periods H1 to H72 (green circles) and average precipitation reforecast periods H73 to H144 (orange crosses).

5.5 Conclusions

In this chapter, a WP and precipitation forecast model (SM) based on historical analogues identified from the previous five, 10 or 15 days has been developed. SM was compared to two models, one based on the first-order Markov WP transition probabilities (MC) and the other on the climatological WP frequencies (FM). Precipitation forecasts were then derived by weighted sampling from the conditional distributions of precipitation given the predicted WPs. The models were tested on 144 30-day periods between 1961 and 2016, 72 of which observed very low precipitation and the remainder observed average precipitation. Forecast skill of the WPs was assessed using the Jensen-Shannon divergence (JSD), with results indicating little difference in performance between the three models. Summer is an exception, when the benchmark models perform slightly better than SM. On the other hand, SM shows promise of higher forecast sharpness than FM or MC, meaning it is potentially capable of predicting unusual WP frequencies rather than consistently predicting the average. Forecast skill for precipitation was verified using the Ranked Probability Skill Score (RPSS), with precipitation climatology (PC) used as the benchmark. It was shown that, in general, PC is a superior prediction method to SM, MC or FM, even for periods of very low precipitation. However,

results also showed that a perfect prognostic model, that uses WP observations as the 'forecast', to derive precipitation estimates is more skilful than PC during low-precipitation periods. This suggests that given requisite improvement of the WP forecasts, useful precipitation predictions could be derived.

There are several reasons that may explain why the historical analogue approach to predicting WPs here is not skilful. First, the previous chapter confirmed that some WPs do display the tendency to persist on time-scales of up to three weeks. This was shown using a Markov model by higher probabilities of certain WPs occurring given their prior occurrence (of varying frequencies) in a preceding window (or varying length). SM is similar in that it uses preceding windows from which to make a prediction, and is sampling from historical analogues, which is comparable to deriving transition probabilities. However, SM may have failed to capture the persistence due to the selection of analogues based on the set of WPs that occurred in the prior window, rather than also taking into account the WP frequencies and their time-ordering. This links to the conclusion of the previous chapter, which discussed the relationship between WP persistence and physical mechanisms such as low-frequency atmospheric circulation patterns and land-atmosphere feedbacks. If the order of the WPs is crucial in determining the following WPs, then altering SM to account for this may be of benefit.

A second reason for SM's lack of skill is that the calibration period used may be unsuitable. SM makes use of the entire WP series, from 1850, prior to the date of the forecast. The behaviour of MO-30 WPs has changed over time, as illustrated by changes in WP frequencies of occurrence (Figure 3.3). Therefore, some historical analogues used by SM might not be representative of the behaviour of MO-30 at the time of the forecast. Instead of using the entire time series, it might be more appropriate to select analogues from a subsample of the series. These subsamples should be selected according to some exogenous process that has some influence on WP behaviour. The NAO cannot be considered exogenous – its index indicates the relative strength of low and high pressure over Iceland and the Azores, respectively, and the WPs are themselves defined using pressure. Therefore, it would not be surprising to see the behaviour of the WPs change for different NAO states. Indeed, previous authors have confirmed this for the Lamb catalogue, finding positive (negative) correlations between the NAO index and westerly (anticyclonic and northerly) LWTs (Sweeney and O'Hare, 1992; Wilby *et al.*, 1997; Fowler and Kilsby, 2002a).

This dependence between WPs and the NAO offers an opportunity for extending the potential forcing mechanisms of the NAO to WP classifications. There is evidence that the state of the NAO is influenced on short (~10-day) time-scales by the phase of the Madden-Julian

Oscillation (MJO); see Hannachi *et al.* (2017) and references therein. For example, Cassou (2008) showed that positive (negative) phases of the NAO tended to follow MJO-related convection over the Indian (western Pacific) Ocean. Investigating whether the same mechanism holds for WP frequencies of occurrences and persistence would be an interesting line of analysis. On longer time-scales, research suggests that tropical Atlantic and Pacific sea-surface temperatures influence atmospheric circulation over Europe and the North Atlantic (Terray and Cassou, 2002; Cassou *et al.*, 2004; Fereday *et al.*, 2008; Hannachi *et al.*, 2017). A decrease (increase) in the occurrence of the positive (negative) phases of the NAO have been linked to a warm tropical Atlantic, for example (Terray and Cassou, 2002; Cassou *et al.*, 2004). For the historical analogue WP forecast model presented here, selecting subsamples based on preceding states of the MJO and/or tropical sea-surface temperatures might be a good start, although this would need to be justified by some analysis exploring any changes in behaviour related to these variables.

Although results using the perfect prognostic model suggest that the method used to derive precipitation estimates from WPs can be skilful for dry periods (Figure 5.13), a potential improvement to this model component is to further condition the WP-precipitation conditional distribution prior to sampling. As with conditioning the selection of analogues, WP precipitation could be conditioned by a physical process, the NAO for example. It is well known that the NAO affects precipitation amounts in the UK (see references in Chapter 2), but this does not mean the precipitation associated with WPs also changes with the NAO, as it may only be the WP frequencies of occurrence that change (with their precipitation distributions unchanged). Numerous studies have shown that UK precipitation associated with the LWTs varies on decadal time-scales, and that these variations have coincided with shifts in the NAO (Sweeney and O'Hare, 1992; Wilby, 1994; Wilby, 1997; Fowler and Kilsby, 2002a). However, these analyses are somewhat confounding due to the aforementioned dependence between the NAO and WPs - it is difficult to separate the NAO's influence on the WPs' precipitation directly from its influence on WP frequencies of occurrence. An investigation exploring a more formal, causative relationship between the NAO and WP-precipitation would be interesting, and to the author's knowledge, there are no published studies on this.

A key objective of this thesis is to develop statistical and dynamical models for drought prediction using WPs. In this context, the key result from this chapter is that adequate improvement in the WP forecasting component can yield skilful precipitation forecasts; the focus in the following chapter shall be on improving this factor. While it would be possible to tweak the existing SM model according to suggestions outlined earlier, a different approach

will be pursued: a hybrid dynamical-statistical forecast model, where the WPs are predicted using output from ensemble seasonal forecast systems (the dynamical component) and precipitation forecasts are derived from these in the same manner as in this chapter (the statistical component).

Chapter 6

Drought and precipitation forecasting using dynamical model weather-pattern predictions

6.1 Introduction

A key result from Chapter 5 is that for an idealised model in which the WPs were predicted perfectly, resulting precipitation forecasts were substantially improved compared with those using statistically-predicted WPs. As such, this chapter will explore whether using a dynamical forecast model to predict WPs yields higher skill in both forecasts of the WPs themselves and of precipitation. This will allow a comparison between WP-based statistical and dynamical models and also with direct (modelled) precipitation forecasts from the dynamical system. As such systems tend to be more skilful in predicting atmospheric variables than precipitation in the medium to long range (Kirtman and Pirani, 2009; Lavers *et al.*, 2014; Saha *et al.*, 2014; Vitart, 2014; Lavers *et al.*, 2016a), a hypothesis for this chapter is that using pressure forecasts to predict WP occurrences, and then estimating precipitation and drought from these, will provide a gain in skill compared to forecasting precipitation directly from dynamical models.

Until recently, the capability of dynamical models to predict WP occurrences had been little researched. Ferranti *et al.* (2015) evaluated the forecast skill of the medium-range ECMWF ensemble using WPs. They objectively defined four WPs according to 500 hPa geopotential heights (Z500) over the North Atlantic and European sector. Model forecasts of daily Z500 for October through April between 2007 and 2012 were then assigned to the closest matching WP using the root-mean-square difference. Verification scores indicated that there was superior skill for predictions initialised during negative phases of the North Atlantic Oscillation (NAO). Similarly, WPs were used to evaluate the skill of the Antarctic Mesoscale Prediction System by Nigro *et al.* (2011). The Met Office predict MO-8 and MO-30 WPs using a post-processing system named "Decider". Using an ensemble prediction system, such as GloSea5 (MacLachlan *et al.*, 2015) or the ECMWF Ensemble Prediction System extended range forecast model (henceforth EMCWF-EPS) (Buizza *et al.*, 2007; Vitart *et al.*, 2008), forecast SLP fields over Europe and the North Atlantic Ocean are assigned to the best-matching WP according to the sum-of-squared differences between the forecast SLP and WP MSLP fields. Decider therefore produces a probabilistic prediction of WP occurrences for each day in the forecast lead-time; see Neal *et al.* (2016) for an example that used ECMWF-EPS to forecast MO-8 WPs. Furthermore, Decider has various operational applications: predicting the possibility of flow transporting volcanic ash originating from Iceland into UK airspace, highlighting potential

periods of coastal flood risk in the UK (Neal *et al.*, 2018) and as an early-forecast system for UK fluvial flooding. For Japan, Vuillaume and Herath (2017) defined a set of WPs according to MSLP. These WPs were used to refine bias-correction procedures, via regression modelling, of precipitation from two global ensemble forecast systems. The authors found that improvements from the bias-correction method using WPs was strongly dependent on the WP, but overall superior to the global (non-WP) method.

The aforementioned studies have all considered daily WPs. Other studies have addressed the seasonal time-scale. Baker *et al.* (2018) reforecast UK regional winter precipitation between the winters of 1992/93 and 2011/12 using GloSea5, which has little raw skill in forecasting this variable (MacLachlan *et al.*, 2015). GloSea5 has, however, been shown to skilfully forecast the winter NAO (Scaife *et al.*, 2014). The authors exploited this by constructing two winter MSLP indices over Europe and the North Atlantic, and reforecasts of these indices were derived from the raw MSLP fields. A simple regression model then related these indices to regional precipitation and produced more skilful forecasts than the raw model output.

This chapter investigates to what extent replacing the statistical component of WP prediction from the previous chapter with a dynamical model improves forecast skill. This is done by creating a 20-year daily WP reforecast data set, on which WP and precipitation forecast skill are assessed at a range of lead-times. The data and methods used are in Section 6.2, results in Section 6.3 and concluding remarks in Section 6.4.

6.2 Data and methodology

6.2.1 Data

For this analysis a 20-year daily WP probabilistic reforecast data set is created using the sub-seasonal to seasonal (S2S) project (Vitart *et al.*, 2017) data archive, which, through ECMWF, hosts reforecast data for a multitude of variables and by a range of models from around the globe. In particular, ECMWF-EPS, which is a coupled atmosphere-ocean-sea ice model with a lead-time of 46 days, is used. This model is run at a resolution of 0.25° up to day 15 and 0.5° beyond this, is run at 00Z, twice weekly (Mondays and Thursdays) and has 11 ensemble members for the reforecasts (compared to 51 members for the real-time forecasts). For further details, refer to the model webpage (ECMWF, 2017).

Daily reforecasts of MSLP between 02/01/1997 and 28/12/2016, inclusive, are used. The domain is 30°W-20°E, 35°-70°N with a resolution of 5° (the same as MO-30). These fields are converted to anomalies by removing a smoothed climatology and subsequently assigned to the closest matching MO-30 WP via minimising the sum-of-squared differences. Both the MSLP

climatology and the WP definitions are the same as those used by Neal *et al.* (2016) to ensure consistency, i.e. they are derived from the EMULATE MSLP data set (Ansell *et al.*, 2006) between 1850 and 2003. This is compared against an observed WP time series to measure forecast skill, but this series is slightly different to that described in previous chapters. WP assignment prior to 2004 was previously done using MSLP data, and from 2004 onwards using 12Z SLP fields from ERA-Interim reanalysis data (Dee *et al.*, 2011). The ECMWF-EPS reforecasts, however, are issued for 00Z of each day. To ensure consistency across the forecast and observation data sets, the observed WP series is recalculated using ERA-Interim 00Z fields from 1979 to 2017 (inclusive) using the same assignment procedure as in the original series. This does not qualitatively change key statistics of MO-30, such as WP frequencies of occurrence (not shown).

The regional daily precipitation data set, HadUKP, is the same as described in previous chapters, with data from 1979 through 2017 used. Precipitation forecasts are also compared with precipitation predicted directly by ECMWF-EPS. For the same dates as the WP reforecast data set, modelled precipitation is extracted at a resolution of 0.5° from the S2S archive. The grid cells are assigned to whichever of the nine HadUKP regions the cell centres lie in. Taking the daily mean of all cells over each region produces a probabilistic reforecast data set of precipitation for each of the HadUKP regions.

6.2.2 Weather pattern forecast models and verification procedure

Two models for the WP forecasts are compared. The first is ECMWF-EPS, which shall be referred to as EPS-WP (in practice this is the WP reforecast data set discussed in the previous subsection). The second model is a 1000-member, first-order, nonhomogeneous Markov chain, the same as that used in Chapter 5. This model is used as the statistical comparison with EPS-WP for two reasons. First, the historical analogue model described in the previous chapter is very data-intensive and requires significant computing time (hence the relatively small forecast verification sample used in that chapter). Second, the Markov model overall performed comparably with the analogue approach, making it a suitable and easily implementable proxy for how the analogue model might have performed.

The JSD, as described in Chapter 5 (Equation 5.2), is used to assess model performance of WP prediction. As the forecast sample is substantially bigger than that used in Chapter 5, skill is analysed at the monthly, not seasonal scale. The middle date of each forecast period is used to assign the month. Forecast skill is calculated for lead-times of 16, 21, 26, 31, 36, 41 and 46 days. Results suggest that the JSD is unsuitable for comparing skill at different lead-times. This

is because the JSD measures the distance between probability distributions. At the shorter lead-times, the forecast relative frequency distribution tends to be much noisier than that of the observations i.e. a greater number of different WPs are predicted than observed. As the lead-time is increased, the observations become noisier and as a result the JSD tends to score the differences between these distributions as more similar (a smaller difference). Interestingly, experimentation shows that at very short lead-times (up to five days), the JSD score for ECMWF-EPS may be lower than at some higher lead-times. This is because model forecasts at these time-scales tend to exhibit low variability (the number of different WPs predicted by ensemble members) and high accuracy, yielding a small JSD. Further investigation of the behaviour of information theoretic skill scores used to compare across lead-times would be interesting, but is beyond the scope of this thesis. The JSD is used to compare WP forecast skill of EPS-WP and the Markov model, considering each lead-time separately.

6.2.3 Precipitation and drought forecast models and verification procedure

Four precipitation forecast models are compared (Table 6.1), at the same lead-times as the WP predictions. Three models are driven first by a WP component: the EPS-WP and Markov models described above and Perfect-WP, a perfect prognostic model that uses WP observations as input (as described in Chapter 5.4). The procedure by which precipitation is estimated from the WP forecasts is the same as described in Chapter 5.2.3 except for the following changes:

- The conditional distributions of precipitation given the WPs are calculated using the new MO-30 time series described earlier. Instead of computing these distributions according to the season, they are computed on a rolling 91-day climatology. This should provide a more realistic simulation of sampled precipitation.
- Owing to a shorter record compared to Chapter 5, the cross-validation procedure used is different. Instead of using all data up to the forecast initialisation date, and adding in new data as it becomes available, a leave-one-year-out technique is used. This enables maximum use of observations in validation.

The fourth model is the direct ECMWF-EPS precipitation forecasts (EPS-P), processed to provide probabilistic predictions of regional precipitation as described earlier.

Precipitation and drought forecast verification is done separately. First, a general assessment of precipitation forecast skill is measured using the RPSS described in Chapter 5 (Equation 5.6). As with the WP verification, the data are split by month. Second, model performance in predicting dichotomous drought/non-drought events is evaluated for two classes of drought

severity. The first class, D_{mild}, corresponds to mild drought: precipitation sums (over the length of the considered lead-time) being below the 30.9th percentile of the summed precipitation distribution. The second class is moderate drought, D_{mod}, with such sums being below the 15.9th percentile. These percentiles are calculated using for each region and month using the whole data set from 1979 through 2017, and are chosen as they correspond to SPI values of -0.5 and -1.

Model	WP component	Precipitation component
Markov	Predicted using a first-order Markov chain	Estimated by sampling from conditional distributions of precipitation given the WPs.
EPS-WP	Predicted by assignment of forecast SLP fields from ECMWF-EPS	
Perfect-WP	Observed WPs	
EPS-P	-	Forecast by ECMWF-EPS

Table 6.1: Details of precipitation and drought forecast models.

The Brier Skill Score

Three verification techniques are used to assess skill in predicting droughts. The first is the Brier Skill Score (BSS). The BSS is based on the Brier Score (BS) (Brier, 1950), which measures the mean-square error of probability forecasts for a dichotomous event, in this case the occurrence or non-occurrence of drought. BS is defined as

$$BS = \frac{1}{n}\sum_{i=1}^{n}(y_i - o_i)^2,$$

Equation 6.1

where y_i and o_i are the forecast probabilities and binary event outcomes, respectively, and n is the number of forecasts. The BS is converted to a relative measure, or skill score, by setting

$$BSS = 1 - \frac{BS}{BS_{ref}},$$

Equation 6.2

where BS_{ref} is the score of a reference forecast given by the quantiles associated with each drought threshold, 0.309 for D_{mild} and 0.159 for D_{mod}. As with the RPSS, a perfect score is achieved when $BSS = 1$ and negative (positive) values indicate the forecast is performing worse (better) than BS_{ref}.

The BS can be decomposed into three terms (Murphy, 1973):

$$BS = \frac{1}{n}\sum_{i=1}^{I} N_i(y_i - \bar{o}_i)^2 - \frac{1}{n}\sum_{i=1}^{I} N_i(\bar{o}_i - \bar{o})^2 + \bar{o}(1 - \bar{o}).$$

Equation 6.3

I can be thought of as the specified number of forecast probability subsamples (bins) that the forecasts are split into. Then, N_i are the number of times each forecast is used in each subsample and so $n = \sum_{i=1}^{I} N_i$, the total number of forecasts, as in Equation 6.1. The conditional (subsample) event relative frequency is defined as:

$$\bar{o}_i = p(o_1|y_i) = \frac{1}{N_i}\sum_{k \in N_i} o_k,$$

Equation 6.4

where $o_k = 1$ if the event occurs for the kth forecast-event pair and $o_k = 0$ if it does not. Note that the summation is only for those k that correspond to the particular forecasts in subsample $i = 1, \ldots, I$. The unconditional (overall) event relative frequency, or sample climatology, is given by:

$$\bar{o} = \frac{1}{n}\sum_{k=1}^{n} o_k = \frac{1}{n}\sum_{i=1}^{I} N_i \bar{o}_i .$$

Equation 6.5

The three terms in Equation 6.3 are known as the *reliability*, *resolution* and uncertainty, and so this equation can be written as:

$$BS = \text{reliability} - \text{resolution} + \text{uncertainty}.$$

Equation 6.6

The *reliability* term is a weighted average of the squared differences between the forecast probabilities and the observed relative event frequencies in each subsample. For a reliable forecast model, the forecast probabilities in each subsample would be close to the corresponding event relative frequencies, with small event relative frequencies for low-probability forecasts and high event relative frequencies for high-probability forecasts.

Resolution is a way of assessing the forecast's ability to discriminate between subsample forecast periods with event relative frequencies that differ from each other. Although the forecast probabilities y_i do not appear explicitly in this term (Equation 6.3), it still depends on

the forecasts because the events are ranked when the entire sample is split into the subsamples (Equation 6.4). As with reliability, resolution is a weighted average of squared differences, this time between the subsample event relative frequencies and the sample climatological relative frequency. Therefore, if the forecast subsamples have very different relative frequencies to the sample climatology, the resolution term will be large.

Finally, the uncertainty depends only on the variability of the observations and cannot be influenced by the forecaster (Wilks, 2011). A corresponding equation to Equation 6.6 for the BSS is obtained by dividing the reliability and resolution terms by the uncertainty term (Kharin and Zwiers, 2003):

$$BSS = \frac{\text{resolution} - \text{reliability}}{\text{uncertainty}}.$$

Equation 6.7

The uncertainty term is always positive and so to maximise accuracy a forecaster would look to maximise the resolution while minimising the reliability terms.

Reliability diagrams – forecast reliability, resolution and sharpness

Reliability diagrams offer a convenient way to assess the reliability and resolution of a forecast model (Wilks, 2011). These diagrams consist of two parts, which together show the full joint distribution of forecasts and observations. The first element is the calibration function, $p(o_1|y_i)$, for $i = 1, \ldots n$, where o_1 indicates the event (here, a drought) occurring. The calibration function is visualised by plotting the event relative frequencies against the forecast probabilities and indicates how well calibrated the forecasts are. Forecast probabilities are split into $I = 10$ bins (subsamples) of 10% probability and the mean of all forecast probabilities in each bin is the value plotted on the diagrams (Bröcker and Smith, 2007). Points along the 1:1 line represent a well-calibrated, *reliable*, forecast, as event probabilities are equal to the forecast probabilities and suggest that users can interpret the forecasts at 'face value'. Points to the right (left) of the diagonal indicate over-forecasting (under-forecasting) of the number of drought events.

The forecast *resolution* can also be deduced from the calibration function. For a forecast with poor resolution, the event relative frequencies $p(o_1|y_i)$ only weakly depend on the forecast probabilities. This is reflected by a smaller difference between the calibration function and the horizontal line of the climatological event frequencies and suggests that the forecast is unable to resolve when a drought is more or less likely to occur than the climatological probability. Good resolution, on the other hand, means that the forecasts are able to distinguish different

subsets of forecast occasions for which the subsequent event outcomes are different to each other.

The second element of reliability diagrams is the refinement distribution, $p(y_i)$. This expresses how confident the forecast models are by counting the number of times a forecast is issued in each probability bin. This feature is also called *sharpness*. A low-sharpness model would overwhelmingly predict drought at the climatological frequency, while a high-sharpness model would forecast drought at extreme high and low probabilities, reflecting its level of certainty with which a drought will or will not occur, independent of whether a drought actually does subsequently occur or not.

Relative operating characteristics

A final diagnostic used is the relative operating characteristics (ROC) curve (Mason, 1982; Wilks, 2011), which visualises a model's ability to discriminate between events and non-events. Being conditioned on the observations, the ROC curve may be considered a measure of potential usefulness – it essentially asks what the forecast is, given that a drought has occurred. The ROC curve plots the hit rate (when the model forecasts a drought and a drought subsequently occurs) against the false alarm rate (when the model forecasts a drought but a drought does not then occur). The hit rate and false alarm rate are calculated for cumulative probabilities between 0% and 100% at intervals of 10%. A skilful forecast model will have a hit rate greater than a false alarm rate, and the ROC curve would therefore bow towards the top-left corner of the plot. The ROC curve of a forecast system with no skill would lie along the diagonal, as the hit rate and false alarm rate would be equal, meaning the forecast is no better than a random guess. The area under the ROC curve (AUC) is a useful scalar summary. AUC ranges between zero and one, with higher scores indicating greater skill.

6.3 Results

To reduce information overload, the discussion of results is limited to the three lead-times of 16, 31 and 46 days. Furthermore, results are not shown for every combination of lead-time and drought class as this would be a large amount of figures. Key results not shown will be conveyed via the text. Precipitation and drought forecast results are aggregated from monthly to three-month seasons for visual clarity and regional results are combined for the ROC and reliability diagrams for the same reason.

6.3.1 Weather pattern forecasts

EPS-WP is more skilful than the Markov model for every month and every lead-time, although the difference in skill between the two models decreases as the lead-time increases. As discussed in the previous section, JSD is not used to compare skill across lead-times; the focus is instead on the difference in skill between the two models. Note how JSD scores for both models, especially for Markov, decrease as the lead-time increases (Figure 6.1), which is the opposite of the expected (and usual) effect, hence the argument as to the unsuitability of the JSD in lead-time comparisons. The skill difference between models is much larger for a lead-time of 16 days compared to a lead-time of 46 days. For a 46-day lead-time, the difference in skill is negligible for May through October; in fact, these months have the smallest differences in JSD for all lead-times. This is presumably because the summer months are associated with fewer WPs compared to winter (Table 3.1), resulting in a more skilful Markov model due to higher transition probabilities.

Figure 6.1: Jensen-Shannon Divergence scores for EPS-WP and Markov models for lead-times of 16, 31 and 46 days.

6.3.2 Precipitation forecasts

Unsurprisingly, Perfect-WP is uniformly the most skilful precipitation 'forecast' model for all regions, seasons and lead-times, except for some regions and seasons with a 16-day lead-time. At this shortest lead-time, Perfect-WP is the most skilful in all cases during winter (Figure 6.2a) and in all cases except NS during spring (Figure 6.2b) and NEE during autumn (Figure 6.2d), for which EPS-P is the most skilful. The only season in which Perfect-WP does not have the highest skill for most regions is during summer, when EPS-P is the most skilful (Figure 6.2c). For lead-times of 31 and 46 days, perfectly predicting WPs would enable by far the most skilful precipitation estimates of any model, for all regions and seasons (Figures 6.3 and 6.4). This model is obviously not practical, but the results serve to show that WPs are a potentially useful tool in medium-range precipitation forecasting. For readability, the phrase "forecast models" is

used to refer to EPS-WP, EPS-P and Markov from herein, to distinguish them from Perfect-WP, which cannot be used for forecasting.

During summer and spring, all three forecast models are well matched, although for a 16-day lead-time Markov is the least skilful. For this lead-time, EPS-P mostly scores similarly to EPS-WP, although it has higher skill for some regions (Figures 6.2b and 6.2c) and even outperforms Perfect-WP for several regions in summer (Figure 6.2c). At lead-times of 31 and 46 days, there is little difference in forecast model skill during spring and summer, although in summer NI and SWE appear to benefit from dynamical WP predictions (i.e. EPS-WP), as do the four eastern regions from any kind of WP forecast (EPS-WP and Markov; Figures 6.3 and 6.4). On the other hand, using WP predictions is to the detriment of precipitation forecast skill in spring for SEE, as shown by the superior performance of EPS-P (Figures 6.3b and 6.4b). This split between the east and west is also found by Lavaysse *et al.* (2015), who used ECMWF-EPS to predict meteorological drought with a one month lead-time.

For winter and autumn, EPS-WP is the most skilful forecast model except when considering a 16-day lead-time, for which EPS-P is often the best performer. Scotland benefits most from the use of EPS-WP, as even at the shortest lead-time this model is superior (Figures 6.2a and 6.2d). Note that the skill of the WP forecasts matters, as Markov is associated with poor precipitation skill at this lead-time, which corresponds to its low skill in forecasting the WPs compared to EPS-WP (Figure 6.1). EPS-WP is the most skilful forecast model for 31- and 46-day lead-times; EPS-P and Markov score fairly evenly overall for a 31-day lead-time, with the former model the least skilful for a 46-day lead-time (Figures 6.3 and 6.4). The difference in skill between EPS-WP and Markov is much larger for northern and western regions, particularly in winter. Therefore the improvement in skill by predicting the WPs with a dynamical model, rather than Markov (Figure 6.1), translates to a spatially non-uniform gain in skill for precipitation, with western and northern regions the principal beneficiaries. However, it is difficult to say why this is the case, as from the JSD scores alone it is not clear whether EPS-WP is better at predicting all WPs, or just some of them. Similarly, RPSS scores are for all forecast dates, which masks whether the improvement in precipitation forecast skill comes from an improvement over all periods or if the gain is made for predictions of dry or wet periods.

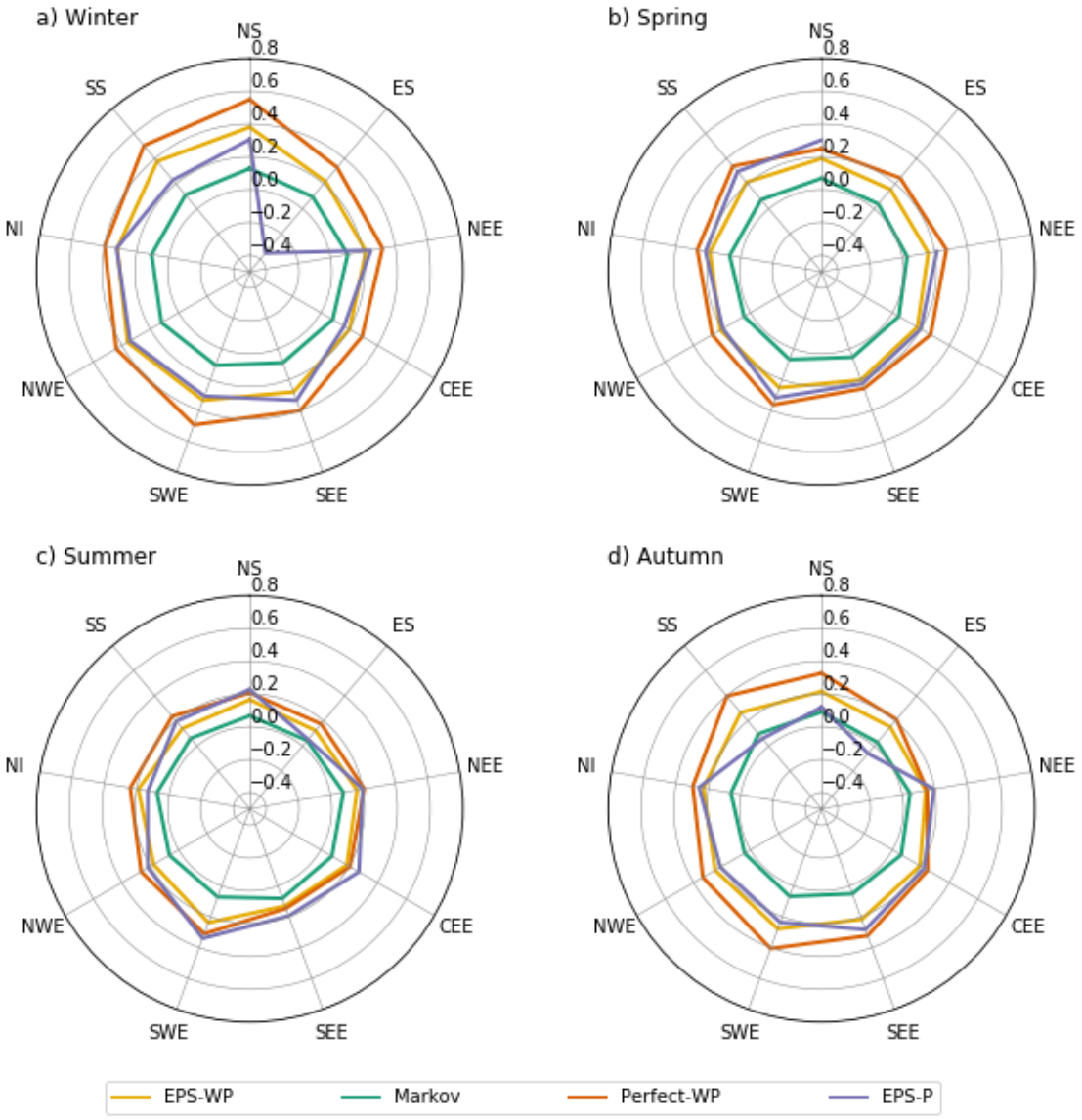

Figure 6.2: Regional Ranked Probability Skill Scores for each model and season with a lead-time of 16 days. Scores lower than -0.5 are omitted for visual clarity. The omitted scores are for EPS-P in ES during spring (-0.54).

The key conclusions from this subsection are that, for winter and autumn, precipitation forecasts are notably more skilful when derived from dynamical predictions of WPs compared to either simple statistical WP predictions or direct precipitation forecasts from a dynamical model. Furthermore, the relative gain in skill is greater for longer lead-times, mainly as a result of a notable drop in skill for EPS-P when comparing a 31-day to a 46-day lead-time, whereas other forecast models' score changes are less severe. For spring and summer, EPS-P is marginally the most skilful model at a 16-day lead-time, with little to choose between all three forecast models at the longer lead-times. A potential reason for the lower skill of WP-based models compared to EPS-P in summer is that the WPs associated with this season tend to be less clear-

cut in terms of being associated with dry or wet conditions (Figure 3.6), possibly as a result of their higher intra-WP variability compared to winter WPs (Figure B.2). Only WP6, WP8 and WP9 are distinctly dry or wet and so precipitation estimates from summer WPs may not be appropriate for periods of non-normal precipitation.

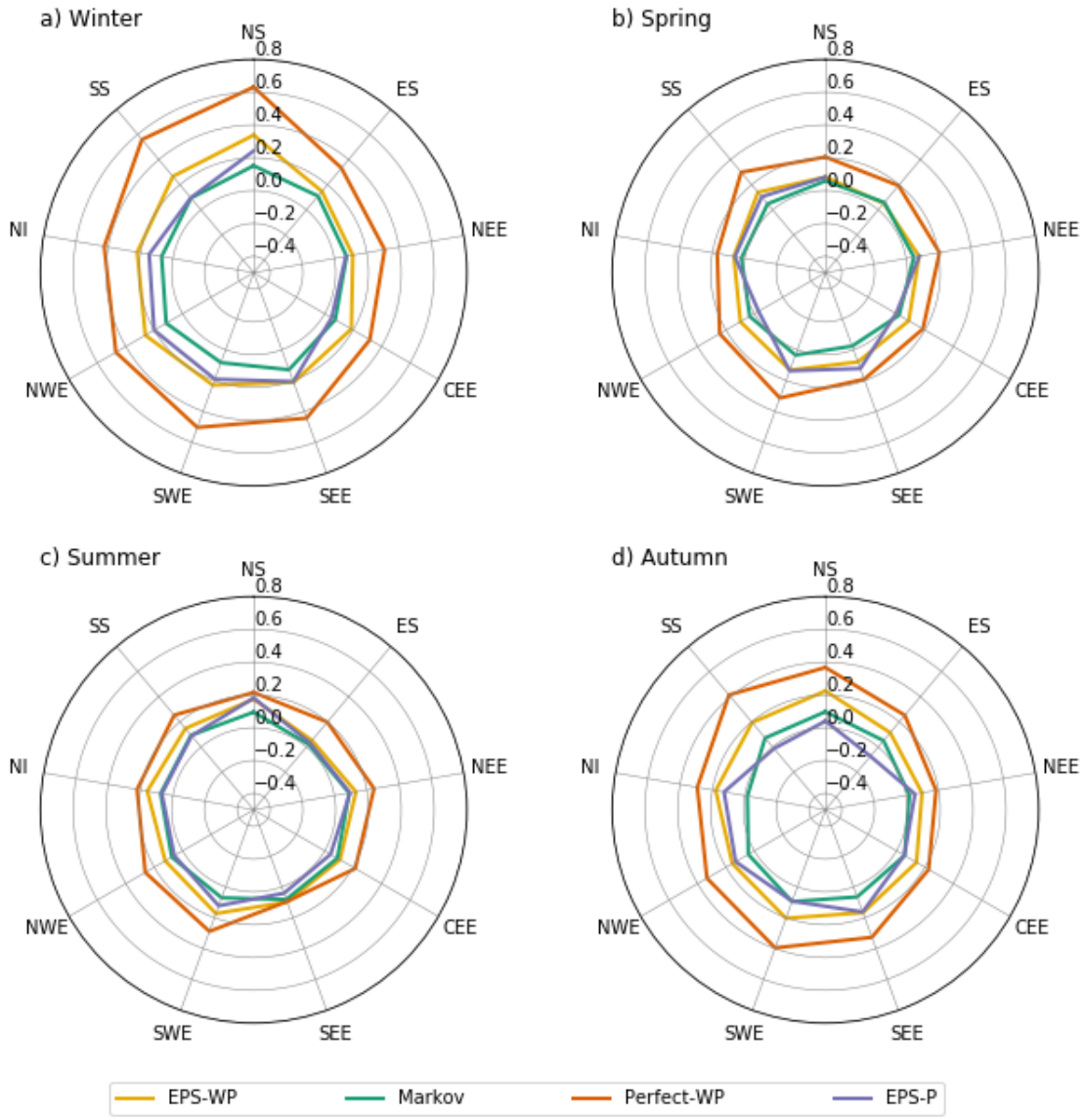

Figure 6.3: As for Figure 6.2 but with a lead-time of 31 days. The omitted scores are for EPS-P in ES during winter (-0.78) and spring (-0.91).

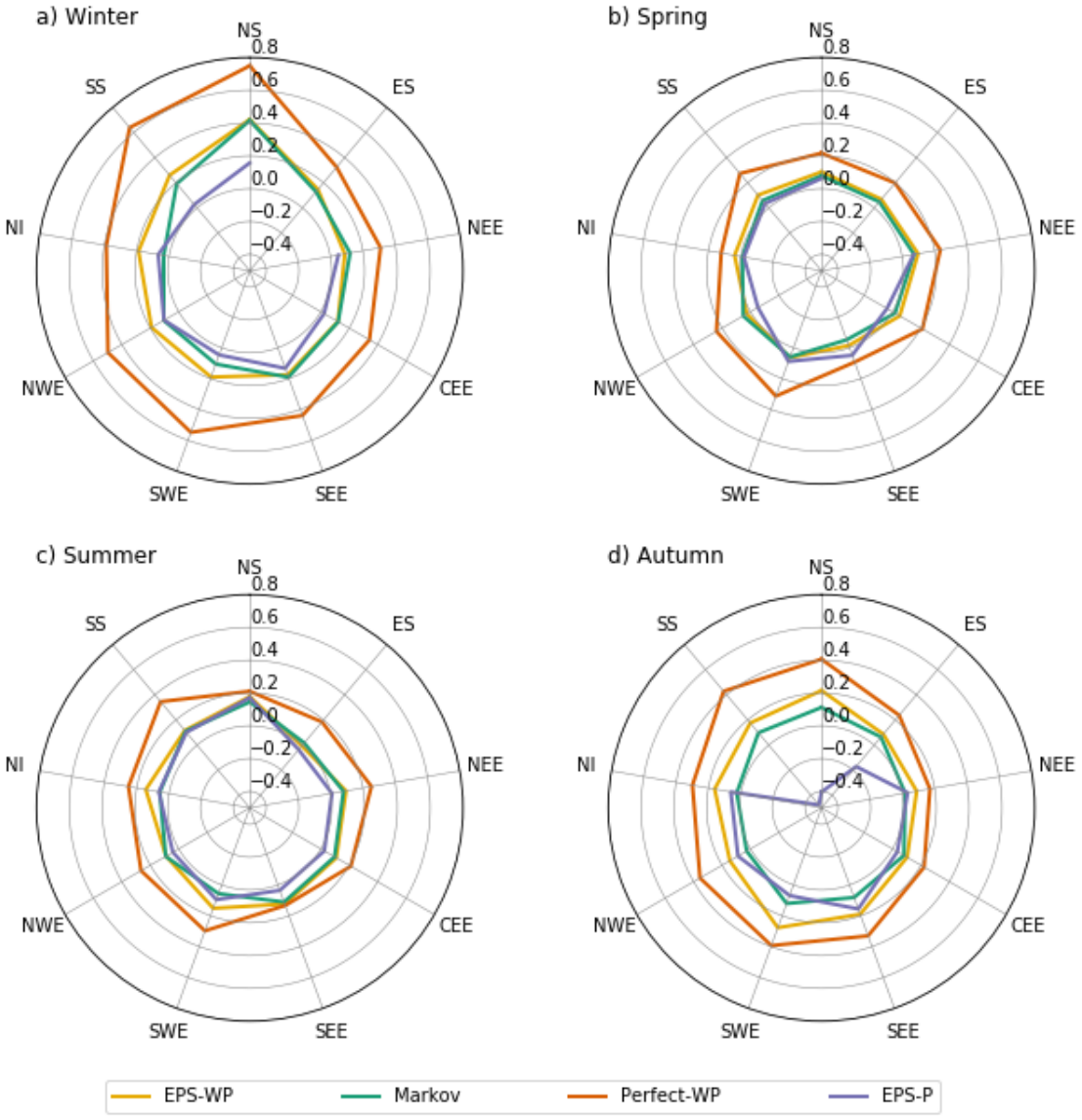

Figure 6.4: As for Figure 6.2 but with a lead-time of 46 days. The omitted scores are for EPS-P in ES during winter (-1.15) and spring (-1.37).

6.3.3 Drought forecasts

Forecast accuracy

Forecast accuracy for mild and moderate drought is qualitatively similar to that of general precipitation in terms of regional and lead-time differences. EPS-WP is overall the most skilful model, although this is less the case for a 16-day lead-time, for the three regions in East England and for most regions in spring and summer. To reduce the amount of figures, those for D_{mild} are omitted as the results are more similar to the RPSS results than those of D_{mod}. Also omitted is the figure for the 31-day lead-time as the BSS scores are qualitatively similar to the 46-day lead-time. For the shortest lead-time, EPS-P has the highest accuracy for predicting winter and

autumn drought of both classes, except in Scotland, for which EPS-WP has the highest skill (Figure 6.5). Indeed this model has the highest skill for the other lead-times during these seasons (Figure 6.6). A key difference is that eastern England droughts are at least as accurately predicted by EPS-P as by EPS-WP for the two longer lead-times (Figure 6.6), whereas for precipitation forecasting the latter tended to be more accurate (Figures 6.3 and 6.4). Difference in model skill is lower for spring and summer drought forecasts, particularly for moderate drought (Figure 6.6). In fact, for this class, there is very little or no gain in skill by using WPs at 31- and 46-day lead-times for spring and summer compared to EPS-P (Figure 6.6). Furthermore, at these lead-times both models are less skilful than issuing climatological drought probabilities (shown by a BSS of zero) except for spring predictions of eastern and southern droughts. This suggests that, during spring and summer, deriving precipitation from predicted WPs may be useful if forecasting mild drought, but not for more severe droughts.

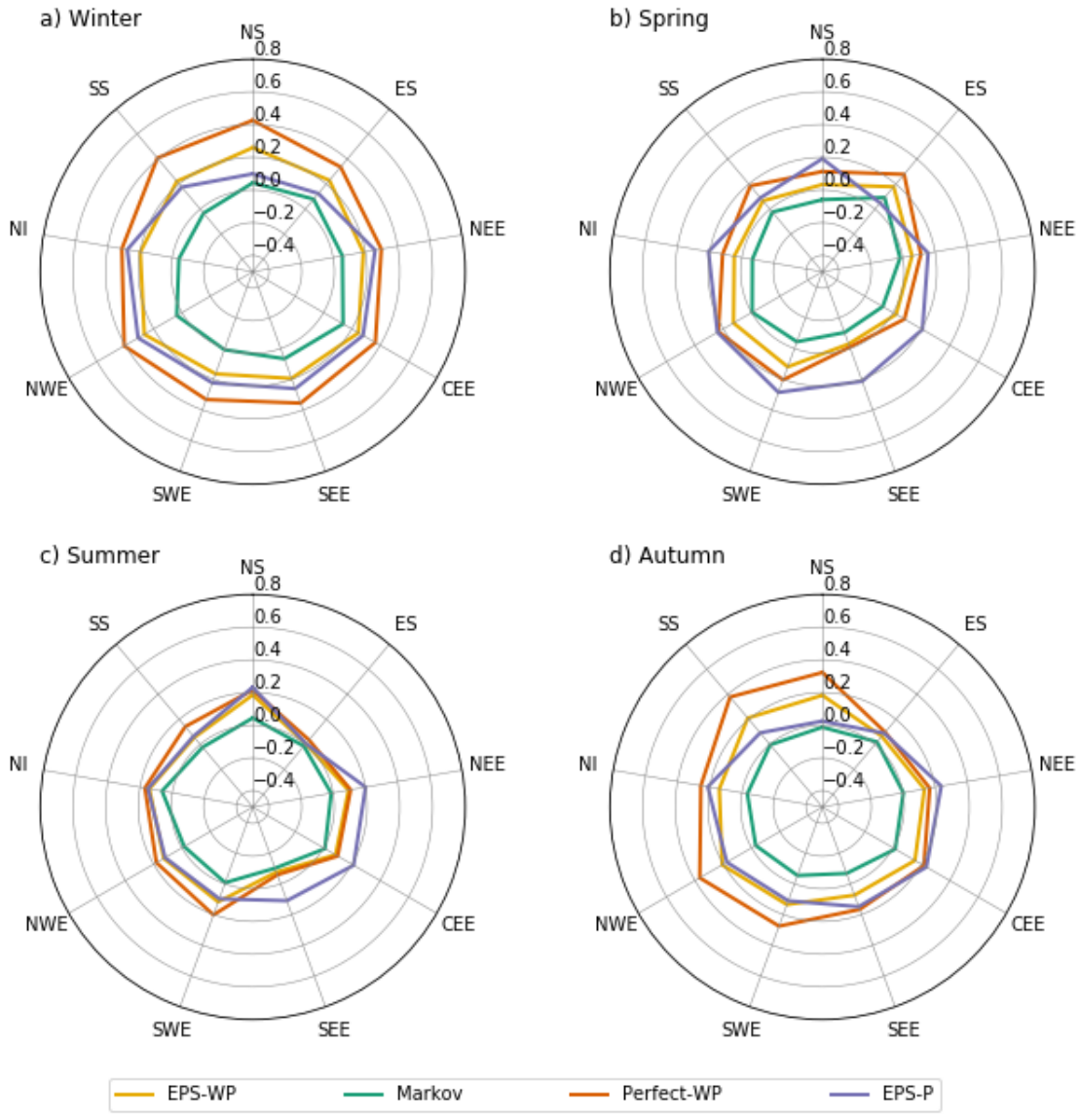

Figure 6.5 (previous page): Brier Skill Scores for predicting moderate drought at a 16-day lead-time.

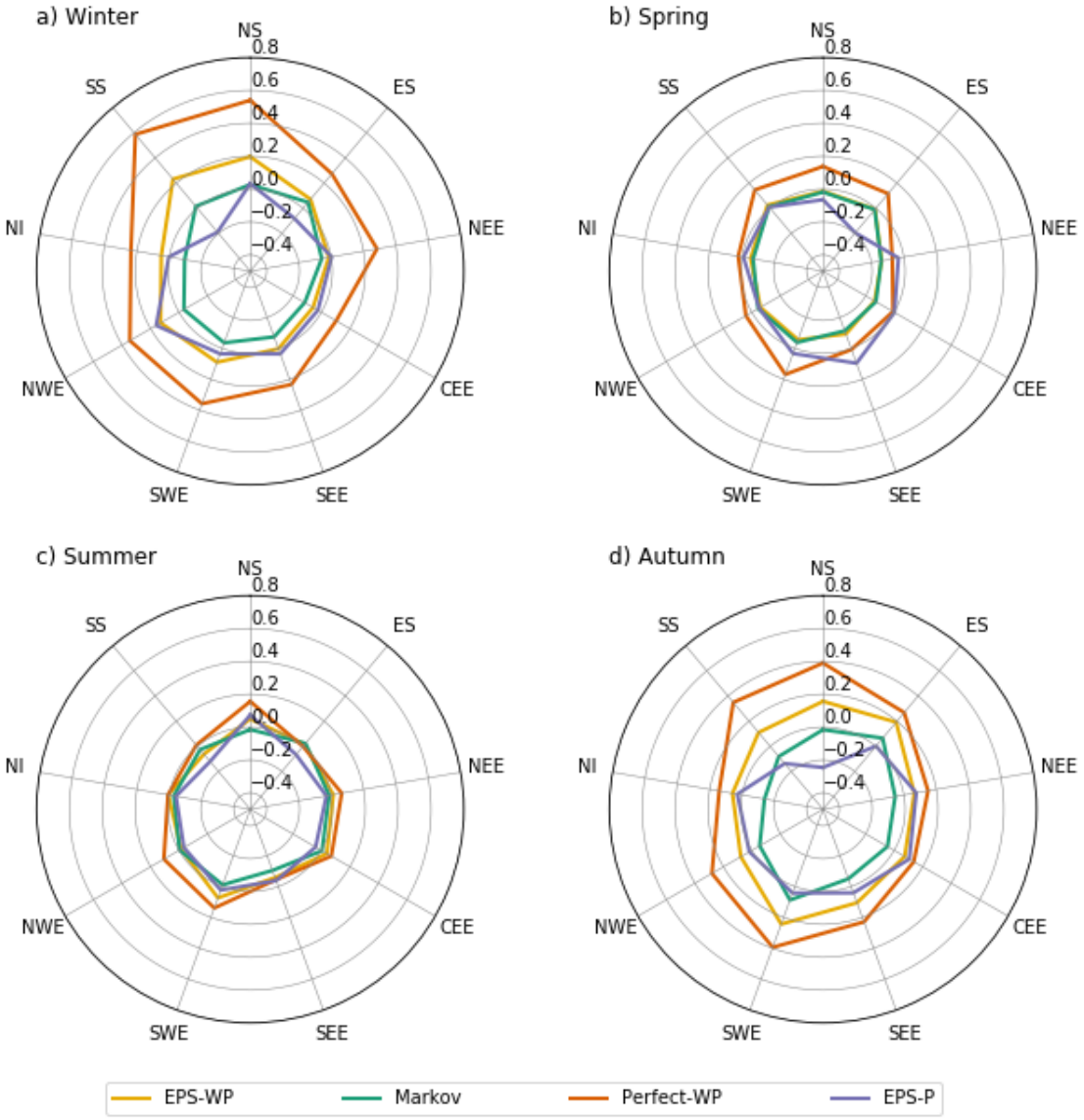

Figure 6.6: As Figure 6.5 but at a 46-day lead-time.

Relative operating characteristics

All models are better able to discriminate between drought and non-drought events than random chance, with Perfect-WP the most able and Markov the least able, subject to similar caveats regarding lead-time and season as for the BSS and RPSS results. During summer and spring, EPS-P has the highest AUC of any of the three forecast models (Figures 6.7, 6.8 and 6.9), and for a 16-day lead-time scores similarly to Perfect-WP (not shown). On the other hand, EPS-WP has the highest skill during winter and autumn at the other lead-times, particularly for mild

drought. Markov is consistently the least suitable model for predicting drought although still represents a better method of doing so than random chance.

A use of the ROC curve is to provide end-users with information on how to apply the considered forecast models. As the plotted points on each curve indicate the hit rate and false alarm rate associated with predicting droughts at each probability interval, they can be used to make an informed decision in selecting a probability threshold for issuing a drought forecast. For example, should a forecaster choose to issue a mild drought warning in winter at a 20% probability level and 46-day lead-time (Figure 6.8), then they would expect EPS-WP to achieve a hit rate roughly double that of the false alarm rate (60% and 30%, respectively). EPS-P, meanwhile, has a slightly higher hit rate but at the expense of a higher false alarm rate (65% and 40%). Knowing the WPs in advance (Perfect-WP) would allow for an outstanding score – roughly a 75% hit rate compared to a 10% false alarm rate. For mild droughts, a 20% probability threshold for EPS-WP and EPS-P achieves at least 60% hit rates at all lead-times, whereas for moderate droughts, this threshold will only achieve such rates at a 16-day lead-time and during autumn for all lead-times. In general, it appears that these low probability thresholds yield the best compromise between hits and false alarms, although in practice, the costs (e.g. financial) associated with false alarms and missed events will determine how responders use these probabilities.

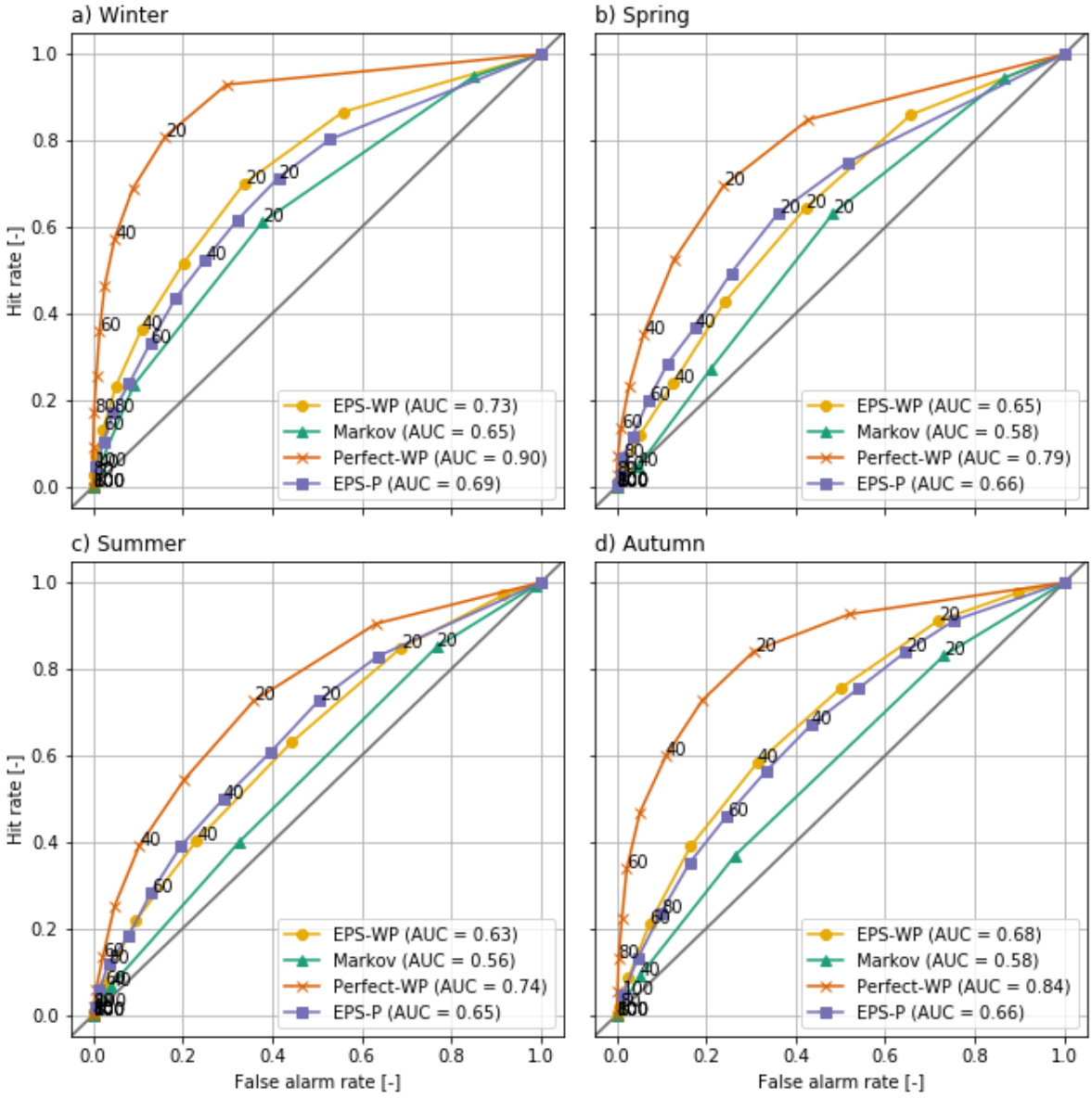

Figure 6.7: Relative operating characteristics (ROC) curves and area under ROC curve (AUC) for mild drought with a 31-day lead-time. Annotated values indicate drought forecast probability thresholds.

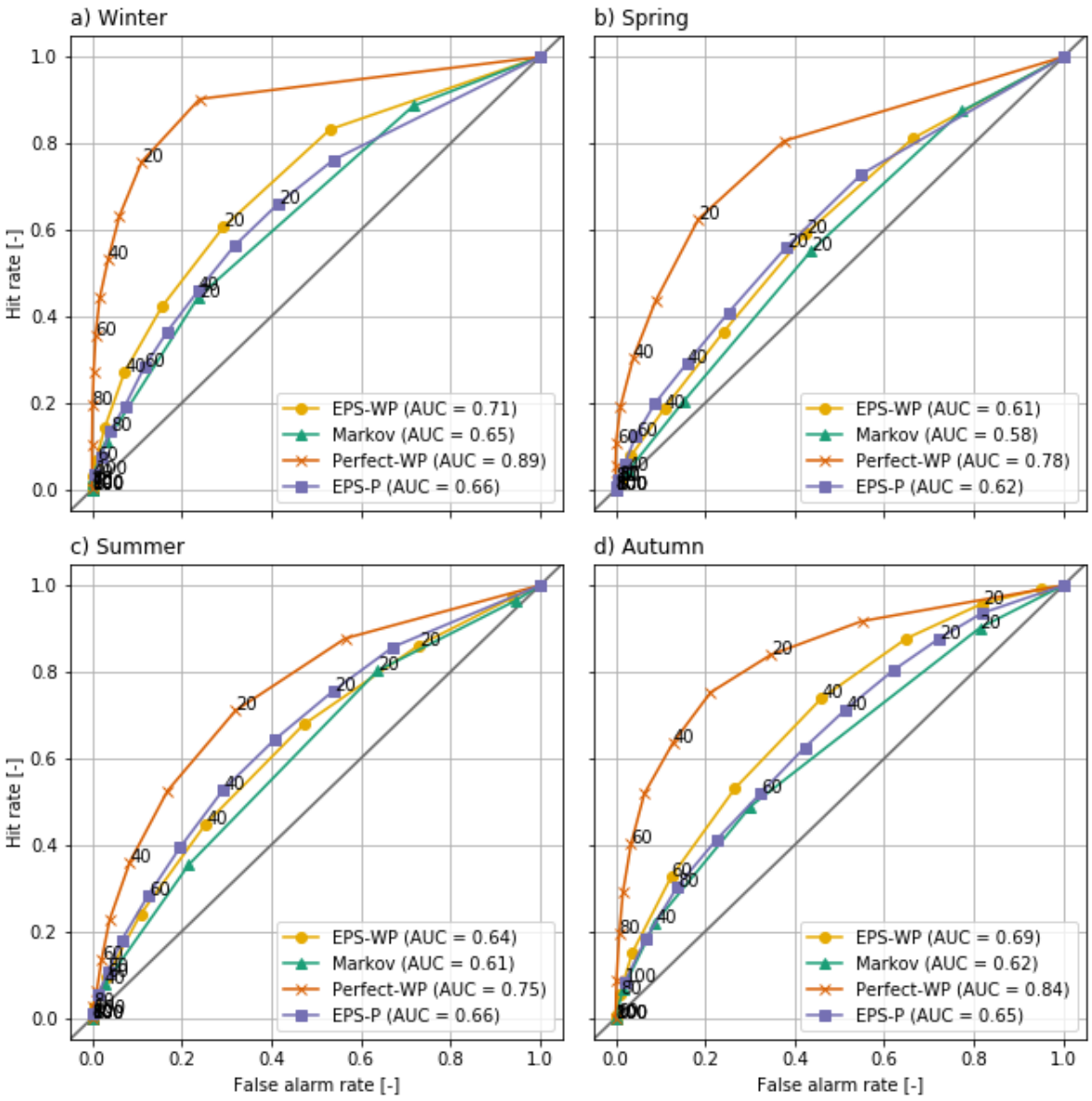

Figure 6.8: As Figure 6.7 but for a 46-day lead-time.

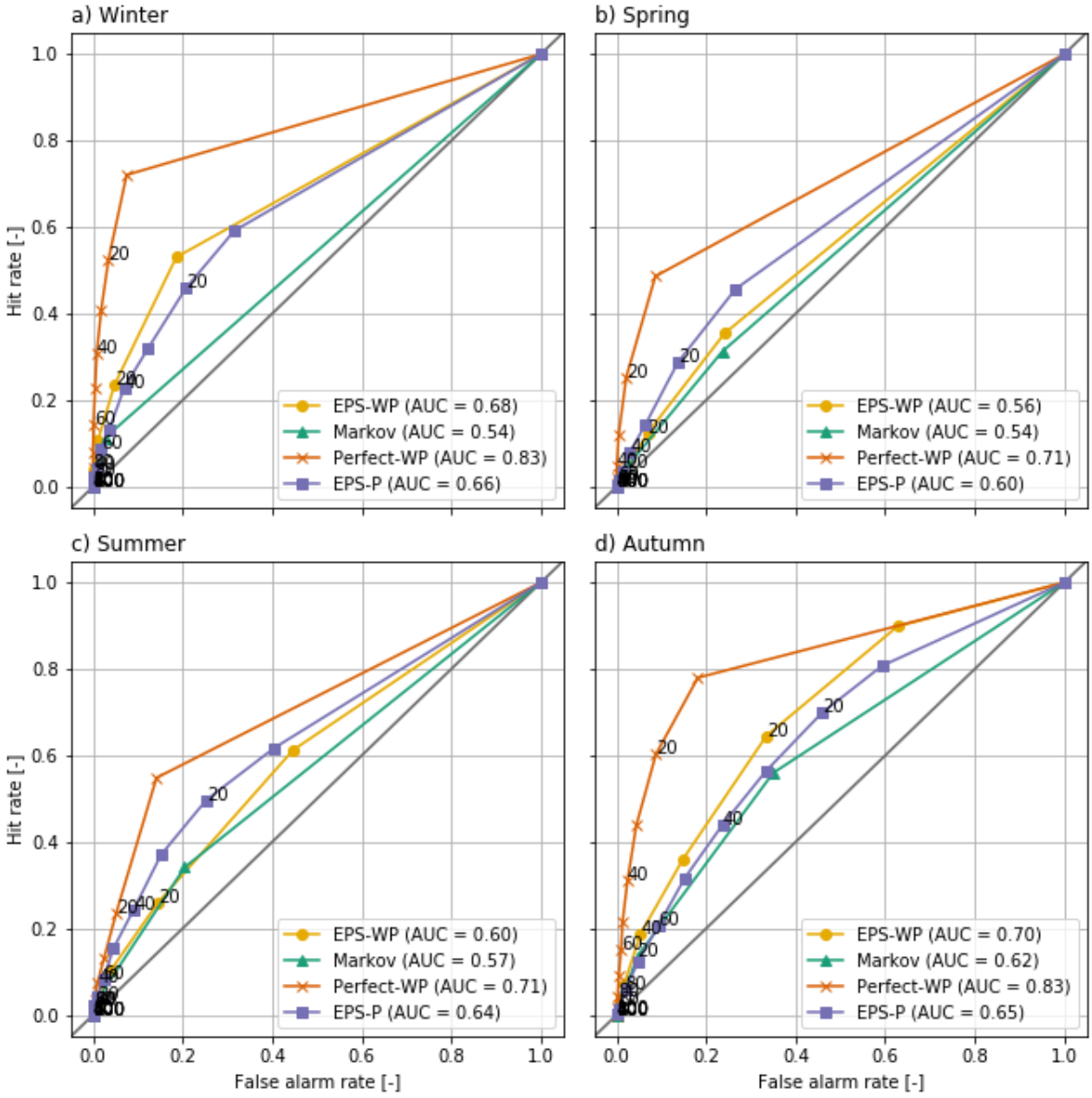

Figure 6.9: As Figure 6.7 but for moderate drought with a 46-day lead-time.

Forecast reliability, resolution and sharpness

EPS-WP is the most reliable forecast model, and while all three WP-driven forecast models tend to under-forecast droughts, EPS-P only does so for lower probability thresholds, with the higher thresholds resulting in this model over-forecasting. This is particularly true for shorter lead-times and during winter, although is still clear for 31-day lead-times in some seasons (Figures 6.10 and 6.11). Sometimes EPS-WP follows the same pattern as EPS-P and over-forecasts drought occurrence for higher predicted probabilities (e.g. Figures 6.10c, e, g and 6.11c). However, the total number of forecasts issued in these intervals is generally smaller than for EPS-P, as the refinement distributions show most clearly for mild drought (Figure 6.10). This means the corresponding points of the calibration function are less reliable for EPS-WP (and Markov) due to smaller sample sizes (Bröcker and Smith, 2007). In fact, all three WP-

based models have occasions when there are no issued forecasts with certain probabilities. These are high probabilities for Perfect-WP and EPS-WP (Figures 6.11c and e) but can be as low as between 30% and 40% for Markov (Figures 6.11e and g). As such, although EPS-WP appears the most reliable model from looking only at the calibration function, there is less certainty of this fact for moderate drought and for higher forecast probabilities. This erratic behaviour of the conditional event relative frequencies is most obvious in Figure 6.11c and may be explained by the very low sample sizes of forecasts issued with anything but a small probability (Figure 6.11d) (Wilks, 1995). An interesting result is that forecasts from EPS-WP are more reliable than from Perfect-WP (Figures 6.10 and 6.11), despite having lower accuracy (Figures 6.5 and 6.6). As a more skilful BSS is composed of smaller reliability and larger resolution terms (Equation 6.7), it follows that the resolution of Perfect-WP is sufficiently large to overcome the larger reliability term compared to EPS-WP and yield an overall more accurate forecast model. These under- or over-forecasting biases must be taken into account by an operational forecaster using these models.

A key difference apparent from the calibration function relates to the ability of the models to identify subsets of forecast situations where the subsequent event relative frequencies are different, i.e. the forecast resolution. An almost completely consistent feature across all lead-times and drought classes is the poorer resolution of EPS-P, particularly obvious in autumn (Figures 6.10g and 6.11g), with the conditional event relative frequencies quite clearly closer to the climatological average compared to the other models. This should be considered in conjunction with the sharpness of the forecast, which is relatively high for this model as shown by the numbers of issued extreme probabilities, particularly those in the upper-tail (Figures 6.10h and 6.11h). This combination of poor resolution and high sharpness can be thought of as "overconfidence" (Wilks, 2011) – on the occasions that EPS-P issues a forecast indicating the likelihood of a drought is very high, the actual likelihood of a drought subsequently occurring is lower. To compensate for this overconfidence, a user could adjust the probabilities to be less extreme to make the forecasts more reliable.

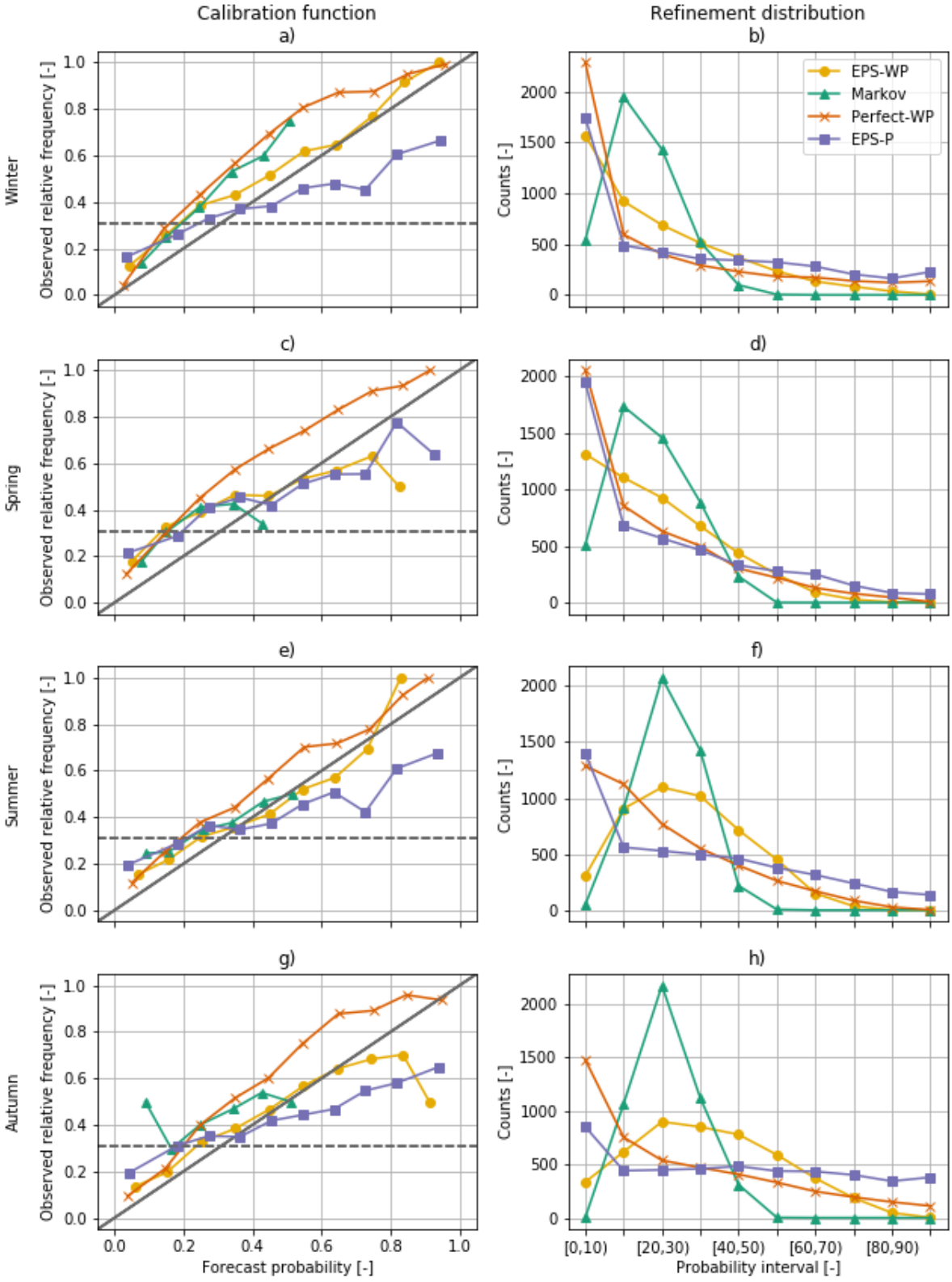

Figure 6.10: Calibration functions (first column) and refinement distributions (second column) for mild drought with a 31-day lead-time. For the calibration function diagrams, the solid diagonal line indicates perfect reliability and the dashed horizontal line the event relative frequency for mild drought (0.309).

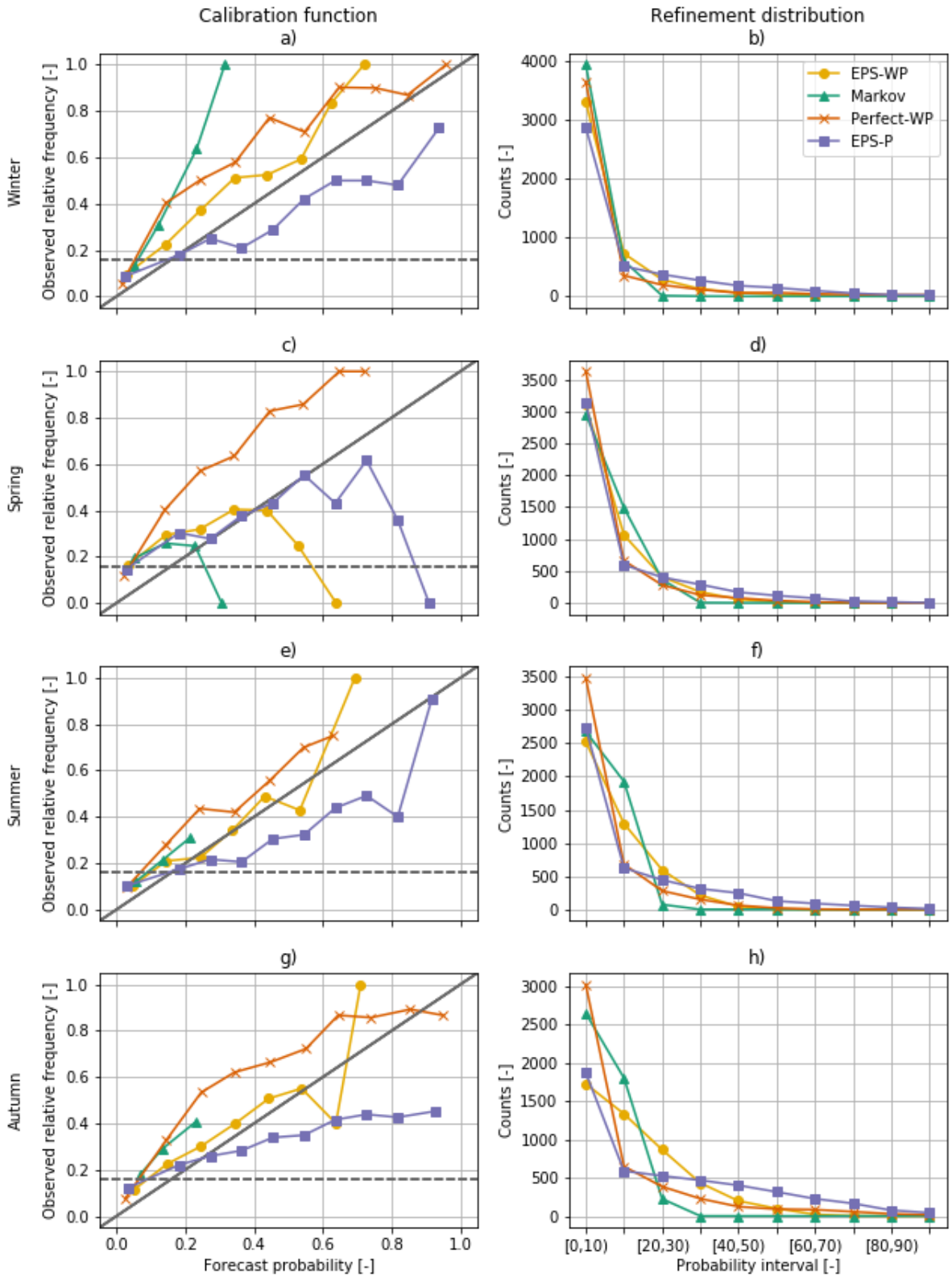

Figure 6.11: As Figure 6.10 but for moderate drought (event relative frequency of 0.159).

These refinement distributions can be compared to those of the Markov model, which exhibits low sharpness, overwhelmingly predicting droughts at the climatological frequency (second column of Figures 6.10 and 6.11). This means that the Markov model is not a useful operational tool in these situations, as similar forecasts could be obtained simply by using the climatological drought frequency. The refinement distributions for EPS-WP show that for mild drought in winter and spring and for moderate drought during all seasons, the model predicts droughts with low probabilities the majority of the time (Figures 6.10b, d and 6.11b, d, f, h). For mild droughts in summer and autumn, however, this model mostly issues forecasts close to the climatological frequency, although not nearly as regularly as the Markov model (Figure 6.10f, h). As with adjusting for bias, a forecaster can use model resolution and sharpness when assessing drought forecast probabilities output by a model.

6.4 Conclusions

This chapter has compared the performance of a dynamical forecast system (EPS-WP) and a first-order Markov model in predicting WP occurrences over a range of lead-times, showing that the dynamical model is always more skilful, although the difference in skill reduces with lead-time. Precipitation forecasts were derived from these WP predictions, and compared to direct precipitation predictions from the dynamical system (EPS-P). EPS-P has the highest overall skill in precipitation and drought forecasts for a 16-day lead-time, whereas EPS-WP predictions provided the greatest skill for longer 31- and 46-day lead-times. The potential in improving WP forecasts was demonstrated further by showing that an idealised, perfect prognosis model (Perfect-WP) would provide much more skilful precipitation and drought forecasts, with high hit rates and low false alarm rates.

Reliability diagrams showed that WP-based models only issue binary drought forecasts with either very low probabilities or probabilities close to the climatological average. In particular, there is little to gain in using the Markov model in mild drought prediction over the climatological frequency, as it tends to issue drought forecasts with this probability anyway. EPS-P has the highest sharpness, predicting drought occurrence with a wide range of probabilities. In particular, it issues greater numbers of high-probability drought forecasts compared to WP-based methods. However, this model also has poor resolution, indicating it is an overconfident forecast model. Overall, drought forecasts issued by EPS-WP are the most reliable, i.e. the forecast probabilities are most similar to the subsequent event probabilities (they “mean what they say”) (Wilks, 2011). Perfect-WP tends to under-forecast the number of drought events, while EPS-P over-forecasts drought events, particularly moderate droughts. Reliability diagrams are useful to aid users in adjusting for an over- or under-forecasting bias.

Given the results presented here, EPS-WP is recommended for the following drought forecast situations.

- Winter and autumn 31- and 46-day forecasts.
- Winter and autumn 16-day forecasts for Scotland (ES, NS and SS).
- Spring and summer 16-day forecasts for ES.
- Summer 31- and 46-day forecasts of mild drought for eastern and southern regions.

EPS-P is recommended for:

- Winter and autumn 16-day forecasts for all regions except those in Scotland.
- Spring and summer 16-day forecasts for all regions except ES
- Spring 31- and 46-day forecasts for all regions except those in Scotland.

Otherwise, the use of climatological drought frequencies represents the most parsimonious (in terms of skill versus model complexity) choice for:

- Summer 31- and 46-day forecasts for mild drought in northern and western regions and moderate drought in all regions.
- Spring 31- and 46-day forecasts for Scotland.

Focussing on the 31- and 46-day lead-times (that are more useful for drought prediction than 16-day forecasts), winter and autumn are clear-cut, with EPS-WP recommended for every region. Summer is more complex. Less severe droughts (D_{mild}) are best predicted by EPS-WP for the eastern and southern regions, but drought climatological frequencies are suggested over the forecast models for western and northern regions and more severe droughts (D_{mod}) for all regions. In spring, climatology is also recommended for Scotland, with the use of EPS-P for the remaining regions.

The higher skill of EPS-WP during winter (and possibly autumn) is probably due to the typically higher skill that medium- to long-range dynamical forecast systems have in predicting atmospheric variables in this season compared to other seasons (Scaife *et al.*, 2014; MacLachlan *et al.*, 2015). Furthermore, the relationship between the NAO (which is the primary mode of North Atlantic/European atmospheric circulation) and precipitation is stronger in this season (Hurrell and Deser, 2009; Lavers *et al.*, 2010; Svensson *et al.*, 2015). This is particularly true for western regions (Jones *et al.*, 2013a; Svensson *et al.*, 2015; van Oldenborgh *et al.*, 2015;

Hall and Hanna, 2018), which potentially explains the greater difference in precipitation and drought forecast skill between EPS-WP and EPS-P in these seasons. The skill of precipitation forecasting using observed WPs (Perfect-WP) is also lower for eastern regions than western regions in winter, implying that MO-30 is not as suited for deriving precipitation in the east. This is potentially because the WPs are more closely related to the NAO in this season, which is not always strongly related to eastern precipitation, than other modes of atmospheric variability that are, for example the East Atlantic pattern (Hall and Hanna, 2018). However, in general forecast skill is lower for eastern regions independent of the model.

By analysing the skill of a perfect prognosis model, this chapter has demonstrated the potential for using WP forecasts to derive precipitation and drought predictions. Currently, dynamical models such as the ECMWF system used here represent the best method of predicting WPs. A useful piece of further research would be to assess the forecast skill of other models, and even multi-model ensembles, at predicting MO-30 WPs or other WP classification systems. Another potential method to improve precipitation and drought forecast skill would be to alter the process by which precipitation is derived from the WPs. Here the entire conditional distribution of precipitation given the WP and season was used in the sampling procedure, but this may not be the optimal way of estimation. It is possible that other factors influence the precipitation from WPs, such as slowly-varying atmospheric and oceanic processes. For example, it would be interesting to see if conditioning the distributions further on the state of the NAO index, or some North Atlantic SST index, and sampling precipitation from these, would improve forecast skill. This is potentially most useful in predicting more severe forms of drought (D_{mod} in this study), for which skill is lower than for mild drought.

Chapter 7
Conclusions

7.1 Summary of results

This thesis has explored the utility of European and North Atlantic weather patterns (WPs) in statistical methods for UK drought prediction. The research aims and objectives set out in Chapter 1 were justified in Chapter 2 by conducting a thorough review of the literature surrounding statistical methods for drought prediction. By reporting on the myriad potential atmospheric and oceanic predictors of UK drought, the literature review showed that their potential in forecasting is highly dependent on lead-time, season and region. WP classifications, therefore, represented an opportunity to analyse drought without the need for testing numerous potential predictors. Chapter 2 also highlighted how just a few published drought prediction studies focussed on the UK, none of which utilised WPs in their methodologies.

Chapter 3 introduced a Met Office WP classification, MO-30, which has been used throughout this thesis. The chapter conducted some exploratory analyses of MO-30, cataloguing WP descriptions (in terms of cyclonicity and flow direction over the UK), frequencies of occurrences (including changes over time) and their seasonal behaviour. Comparisons with the most commonly used, UK-focussed WP classification, the Lamb catalogue, reinforced previous studies' conclusions on changes in European circulation, such as a decrease in easterly flow from the 1870s until the 1920s (Jones *et al.*, 2013b), and on recent increases in precipitation intensity and flooding (Pattison and Lane, 2012; Jones *et al.*, 2013a; Jones *et al.*, 2014). In particular, with reference to the first of the three "key criteria" of Chapter 1.4 (that the WPs must satisfactorily encapsulate atmospheric variability), the two WP classifications' similarities in circulation changes suggests that MO-30 is adequately able to summarise atmospheric variability, at very least to the same capability as the Lamb catalogue. Work in this chapter justified the use of MO-30 in UK precipitation-based analyses by showing that it is more suited to representing regional UK daily precipitation than the Lamb catalogue. In particular, precipitation variability for cyclonic easterly Lamb weather types (LWTs) was large, more so than for any WP in MO-30. However, a smaller set of eight WPs, derived from MO-30 and named MO-8, was shown to be unsuitable for the analyses in this thesis. The eight WPs were derived by repeatedly clustering the 30 WPs from MO-30 on the spatial correlation of their MSLP patterns. This meant WPs with very different precipitation distributions were grouped together and, as a result, WPs in MO-8 had precipitation distributions too similar to each other to be useful.

For the period 1878 to 2015 (1931 to 2015 for Scotland and Northern Ireland), it was shown which MO-30 WPs change during anomalously wet and dry periods by examining changes in their frequencies of occurrence during positive and negative Standardised Precipitation Index (SPI) periods, respectively. The Drought Severity Index (DSI) was used to link changes in these frequencies with drought explicitly, with the index calculated over a range of time-scales in order to represent short- and long-term droughts. Unsurprisingly, WPs which are typically dry for a region tend to occur more frequently than normal during droughts, and vice versa for wet WPs. Furthermore, the magnitude of these changes decrease for longer droughts. Overall, six WPs that are associated with drought over the majority of the UK were identified: WP6, WP9, WP10, WP12, WP17 and WP25. By showing that the relationships between the WPs and precipitation and meteorological drought align with *a priori* expectations and other studies' results (Phillips and McGregor, 1998; Fowler and Kilsby, 2002b), the second "key criterion" from Chapter 1.4 (that there is a relationship between the WPs and precipitation/drought) has been satisfied. In the context of the overall aim of using WPs in drought prediction, Chapter 3 may be thought of as an exploratory analysis, necessary to justify the use of MO-30 in further analyses, and because there is no published literature on such fundamental characteristics outside of this thesis. The work in Chapter 3 has been published as a journal article (Richardson *et al.*, 2018a).

Chapter 4 explored the potential of using MO-30 in predictive applications by investigating whether the classification's time series from 1850 to 2016 exhibited evidence of inherent predictability (persistence). Two novel methods for identifying persistence in the time series were developed. The first was an empirical counting method that searched for minimum 30- or 60-day periods dominated by three to six WPs. By allowing for a small number of days to be WPs other than those of interest, the method accounted for the inevitable small changes in atmospheric conditions that can occur over such long periods. The second method was a Markov model that enabled an assessment as to whether WP transition probabilities change when conditioned on persistence up to 20 days prior. This model was developed without the need for estimating the large number of parameters usually required for high-order Markov chains. Results showed that there were multi-month periods (up to 83 days) that were characterised by the dominance of groups of three to six WPs. Furthermore, some of these periods corresponded to notable meteorological events, such as the Burn's Day Storm in January 1990 and the mid-1990s Yorkshire drought. Finding such long periods of persistence in daily WPs was remarkable, and illustrated how typically variable atmospheric conditions can be relatively stable on monthly time-scales, and that WPs can be used to represent this. The

chapter also demonstrated that some WPs behave as 'attractors', showing increased probability of reoccurrence despite other WPs occurring in-between. However, there was no evident link between the persistence statistics of each WP and their flow characteristics, except for those featuring an easterly flow over the UK, which are amongst the most persistent. The work in this chapter has been published as a journal article and represents the first paper investigating WP persistence in this way (Richardson *et al.*, 2018b).

These results indicated that persistence is a phenomenon of atmospheric circulation over Europe and the North Atlantic, which can be characterised by MO-30. Chapter 5 attempted to harness these persistence properties in a forecast model. The model, SM, was based on historical analogues, and operated by identifying past situations when the atmospheric conditions were similar. Specifically, the model identified periods in the WP time series that featured identical, or very similar, sets of WPs as the period prior to the forecast date. Probabilistic reforecasts of WPs over 144 30-day periods were generated by sampling from these analogues. From these WP reforecasts, precipitation estimates were derived by sampling from the conditional distributions of precipitation given each reforecast WP. This model was compared against a first-order nonhomogeneous Markov chain and a benchmark method based on climatological WP frequencies of occurrence. The results showed that SM performed very similarly to the simpler methods for WP prediction. Furthermore, all models were less skilful in reforecasting precipitation than climatology, with Ranked Probability Skill Scores (RPSS) no higher than -0.25. However, an idealised, perfect prognostic model, where the observed WPs were used in place of reforecast WPs, achieved a substantial improvement in drought reforecast skill, which for most regions was higher than climatology, with RPSS reaching almost 0.5 in north-eastern regions.

Findings from Chapter 5 suggested that an improvement in drought prediction skill could be achieved with more skilful WP forecasts. To this end, Chapter 6 explored the utility of a dynamical model to predict WPs, and subsequently precipitation and drought. The capability of the ECMWF extended-range ensemble system (ECMWF-EPS) to predict WPs (EPS-WP) was tested at a range of lead-times (from 16 to 46 days). This model was benchmarked against an idealised model using WP observations rather than forecasts (Perfect-WP), a model using first-order nonhomogeneous Markov chain WP predictions and direct precipitation forecasts from the dynamical model (EPS-P). For WP predictions, EPS-WP always outperformed the Markov approach, although the difference in skill decreased as the lead-time increased. Therefore, the third "key criterion" of Chapter 1.4 (that the WPs are predictable) is satisfied, although not by the statistical methods presented here. For drought prediction, EPS-P had the

highest skill for shorter lead-times (up to a Brier Skill Score (BSS) of 0.3 for moderate drought), while EPS-WP was often the most skilful for longer lead-times (BSS up to 0.2). This aligns with the documented low skill of precipitation forecasts from dynamical models beyond the medium-range (Cuo *et al.*, 2011), and supports previous research that exploiting atmospheric predictability is a useful precursor to forecasting local-scale variables at these lead-times (Lavers *et al.*, 2016a; Lavers *et al.*, 2016b). Again, the skilful performance of the idealised model, Perfect-WP, highlighted the potential usefulness of WPs in drought prediction, with high BSS scores of 0.6 on occasions, plus high hit-rates (80%) and low false-alarm rates (20%) achievable.

7.2 Results in context of the existing literature

This thesis has presented analyses that are important steps in the fields of both WPs and the predictability of UK drought. Relatively recent works have demonstrated the potential of forecasting atmospheric variables as a precursor to predicting local-scale variables such as precipitation beyond the short-range forecasts associated with weather prediction (Lavers *et al.*, 2014; Lavers *et al.*, 2016a; Lavers *et al.*, 2016b; Baker *et al.*, 2018). Furthermore, the prediction of WPs is itself an active, and relatively new, area of research (Vuillaume and Herath, 2017; Neal *et al.*, 2018). Chapter 6 showed the benefits of utilising WP forecasts as an intermediate step in precipitation and drought prediction by the applying the ECMWF-EPS system: forecast skill was often higher when WPs predicted by the dynamical model were used to estimate future precipitation and drought compared to direct precipitation forecasts. This supports the results of the aforementioned studies, providing further evidence that this is a valid research area that could be of benefit to the medium- to long-range weather forecast community.

While the historical analogue and Markov WP forecast models used in Chapter 5 did not yield particularly skilful results, there is remarkably little published research in using statistical methods for predicting WPs. This is perhaps understandable due to fairly widespread access to dynamical model output (at least in the UK and similarly wealthy nations) and hence the ability to predict WPs using these models, which tend to be considered superior *a priori*. However, research into statistical techniques is valuable for those who may require less data-intensive or financially expensive methods. In addition, statistical models have a role to play in benchmarking the performance of dynamical models. None of the studies utilising the latter for WP prediction have compared their performance to a statistical technique (Vuillaume and Herath, 2017; Baker *et al.*, 2018; Neal *et al.*, 2018). This is a missed opportunity, as the simpler, less expensive statistical methods may be of (at least) similar skill in certain situations, which could benefit those end-users who do not have access to dynamical model output. Indeed,

results in Chapter 6 showed that a simple Markov model produced precipitation and drought forecasts of similar skill to the ECMWF model in spring and summer at a 46-day lead-time.

Chapter 4 approached the topic of WP persistence in a novel way. While previous studies focussed on changes in the average persistence of WPs over time (Stefanicki *et al.*, 1998; Kyselý and Domonkos, 2006; Kučerová *et al.*, 2017), this chapter explicitly attempted to identify the longest periods of persistence. The objectives for this analysis were new to the field and, as such, the results and methods described are of benefit to the research community. The methods introduced are easily applied to other WP classifications, opening up their application to regions beyond Europe and the North Atlantic. Furthermore, the key result that the usually ephemeral WPs may, for extended periods, be far more persistent than usual could be valuable in the atmospheric sciences beyond pure WP analyses, perhaps in justifying the use of WP classifications to identify other aspects of atmospheric behaviour. The other significance of Chapter 4 was to show that the behaviour of the WPs resembled that described by atmospheric chaos theory (Lorenz, 1963). In particular, similarities were found between WP behaviour and strange attractors, a feature of some chaotic systems, including the famous Lorenz system. The presence of a (chaotic) strange attractor forces the system to evolve toward a set of numerical values, but with sensitive dependence on the initial conditions (Grebogi *et al.*, 1987; Boeing, 2016). Analogously, Chapter 4 showed, using a Markov model, that some WPs exhibit attractor-like behaviour, with their occurrence increasing the probability that the WP system will return to the same WP (i.e. the same state) in the future.

That the WPs appear to behave similarly to these highly nonlinear systems should not be a surprise, as they are a simplified representation of such systems. There must be some physical reason that the WPs have preferred sequences (i.e. those represented by their seasonal behaviour, transition probabilities and persistence statistics), suggesting that the poor skill of this thesis' statistical models for WP prediction, relative to the dynamical model, is due to an inability to properly capture the physical mechanisms driving WP occurrences. As stated by Hannachi *et al.* (2017), "…it is necessary to consider nonlinear scale interactions in order to fully understand and accurately predict atmospheric low-frequency behaviour". While the WPs are not exactly the low-frequency systems that Hannachi *et al.* (2017) were referring to, they are analogous. Therefore, it is likely that the difference in skill is at least partly attributable to the dynamical model's nonlinear capabilities. Using the same ECMWF model as used in this thesis, Dawson *et al.* (2012) demonstrated the importance of representing small-scale atmospheric processes when simulating large-scale circulation, including the NAO. They found that the model integrated with a "climate" horizontal spatial resolution (125km grid spacing at

the equator) was inferior compared to the model integrated with a "weather forecast" resolution (16km grid spacing), probably due to less realistic orography and a less realistic representation of Rossby waves (Hannachi *et al.*, 2017). Due to the likely dependence between the NAO and WPs (see the discussion in Chapter 5.5), it is possible that WPs cannot be accurately forecast without proper representation of small-scale processes, which would be challenging for a statistical model.

Finally, Chapter 3 expanded on the current literature linking WP classifications to UK climatology by relating MO-30 to meteorological drought, which has only been studied in a WP context in a few cases (Phillips and McGregor, 1998; Fowler and Kilsby, 2002b), and never on a nationwide basis. This work therefore provides a useful resource for those interested in the common patterns of broad-scale atmospheric circulation that are associated with drought for different UK regions and over different time-scales. In addition, the chapter outlined the applicability of MO-30 in UK precipitation-based analyses, including highlighting advantages compared to the Lamb catalogue, which has long been the de facto classification for such applications.

7.3 Future work

Recommendations for future work have been made at the end of each chapter. These were typically specific ideas for a direct continuation of the chapters' analyses e.g. modifications to a particular model. This section shall present some more general suggestions pertaining to the thesis as a whole.

7.3.1 The choice of WP classifications

WP classifications are essentially simplifications of atmospheric circulation and their characteristics are sensitive to methodological choices made in their development. For example, a consequence of the simulated annealing procedure used to define the WPs in MO-30 is that westerly WPs significantly outnumber easterly WPs. Similarly, the subjective method of defining LWTs results in equal numbers of easterly and westerly LWTs, but the latter occur far less frequently (Table 3.3). This would be problematic when, for example, constructing a Markov model for this classification, as the small sample sizes of these LWTs would make the transition probability estimates highly uncertain. Furthermore, Kučerová *et al.* (2017) showed that trend statistics of WP frequencies of occurrence and persistence are highly variable between a large number of different classifications. They showed that even WPs with similar flow configurations exhibit differences in the direction of trend, leading the authors to conclude that we should be wary of generalising climatological conclusions from a single, or limited

number of, WP classifications. Therefore, use of another classification alongside MO-30 might help reinforce the inference of some of the analyses in this thesis. In particular, it would be interesting to see if some of the easterly LWTs exhibited similar persistence statistics to WP27 and WP28 of MO-30. Results in Chapter 4 suggested that flow patterns very similar to these two WPs have persisted for multiple months, but following the recommendations by Kučerová *et al.* (2017), the analysis should be done with further WP classifications before drawing general conclusions about atmospheric persistence.

Using additional WP classifications for the analyses in this thesis would present an opportunity for testing a selection with a range of different features. Both MO-30 and the Lamb catalogue have been used for a variety of applications, but many studies chose to derive their own WP classifications for their specific task (see references in Chapter 2.6). By doing so, users are able to tailor the classification to suit their requirements. For example, to analyse precipitation changes in the US Southwest, Prein *et al.* (2016) incorporated wind speed and precipitable water variables (together with the traditional choice of SLP) into their WPs, as these are crucial in the physical processes leading to precipitation (Doswell *et al.*, 1996). Custom classifications may also be derived by clustering existing classifications. MO-8, presented in Chapter 3, is an example of this. Similarly, LWTs have previously been clustered together according to their northern England precipitation characteristics (Fowler *et al.*, 2000; Fowler and Kilsby, 2002b; Fowler and Kilsby, 2002a; Fowler *et al.*, 2005). Fewer WPs may be advantageous in a forecasting context, as the sample sizes of WP occurrences would be greater than those of a classification with more WPs and hence transition probability estimates would be more reliable. This is a trade-off, however, as clustering WPs would generally result in greater intra-WP variability, as seen in the precipitation distributions of the MO-8 WPs (Figure 3.5). Furthermore, the aforementioned studies focussed on the relatively small region of Yorkshire in northern England. To adopt a similar approach in this thesis would require clustering WPs separately for each region, resulting in nine separate classifications.

Clearly, using more classifications would increase the complexity of the analysis. In particular, multiple classifications might result in 'information overload' for end-users of operational models, as they may need to be familiar with WPs from each classification. In fact, the Flood Forecasting Centre (FFC; an EA and Met Office partnership) have two WP forecast tools for different purposes (discussed later), both using the MO-30 classification. This means the operational team can quickly assess the predicted WPs and the potential hydro-meteorological consequences.

7.3.2 Operational potential

There is no system in place specifically to provide guidance of upcoming UK drought probabilities. Perhaps the closest services are provided by the Centre for Ecology and Hydrology (CEH) - the Hydrological Outlook[6] and the UK Drought Portal[7] - together with national authorities' monthly summaries of water resource conditions e.g. the EA Water Situation Reports[8] for England. The Drought Portal is not a forecast product; it provides near real-time updates on the current meteorological drought status to aid monitoring and early-warning of drought. It is a web-based tool that allows users to quickly assess the current short- and long-term precipitation deficits/surpluses for hydrological catchments across the UK. However, as acknowledged by CEH on the portal website, it is not recommended for operational use due to uncertainties in the 5km gridded precipitation estimates, which are derived from a sparse gauge network. The Hydrological Outlook, on the other hand, is a forecast tool; it provides insight into the likely hydrological conditions for the next three months, updated monthly. The Met Office provides predictions of precipitation and temperature that are used by CEH as inputs to derive the hydrological components of the outlook (river flow and groundwater levels). As the name implies, the hydrology is the primary focus, but it is the precipitation predictions that are most analogous to forecasts of meteorological drought. The Met Office uses GloSea5 (MacLachlan *et al.*, 2015) to forecast UK precipitation. This information is disseminated to readers of the Hydrological Outlook by comparing probability density function plots of the climatology and the forecast, together with concise textual passages relaying the key information using phrases such as "the probabilities of above- and below-average rainfall do not differ significantly from normal[9]". Importantly, this information is presented for the UK as a whole, not regionally, which is surely too coarse for practical use by many potential end-users.

Therefore, there is arguably clear space for a forecast tool that provides probabilistic predictions of meteorological drought at the regional scale. Relevant examples of operational forecast tools are found in the FFC, a specialised hydro-meteorological service that provides county-level river, surface water, coastal/tidal and groundwater flood-risk guidance for core responders in England and Wales. The service aims to issue flood warnings with a lead-time of 2 or 3 days (Flood Forecasting Centre, 2017). For longer-range river and coastal flood guidance, the FFC

[6] https://www.hydoutuk.net
[7] https://eip.ceh.ac.uk/droughts
[8] https://www.gov.uk/government/collections/water-situation-reports-for-england
[9] http://www.hydoutuk.net/files/5215/4169/2237/A42018_Forecast-precip-NDJ-v1.pdf

use two tools that are driven by WP predictions from Decider. As described in Chapter 6.1, Decider is a Met Office post-processing system for producing probabilistic predictions of daily WP occurrences for MO-30. Driven by an ensemble forecast model, such as ECMWF-EPS as used in Chapter 6, forecast fields of daily SLP are assigned to the closest-matching WPs via the sum-of-squared differences. Coastal Decider and Fluvial Decider are FFC medium- to long-range forecast guidance tools underpinned by Decider's WP predictions. Coastal Decider flags high-risk periods during which UK coastal sites may be at increased risk of flooding (Neal *et al.*, 2018), while Fluvial Decider provides regional UK warnings of potential extreme precipitation, and hence possibly flood, events. Separate from the FFC, there is a third WP-based operational forecast tool used within the Met Office that uses Decider's WP forecasts to assess the probability of flow into UK airspace originating in Iceland. This is important during volcanic episodes due to the potential for ash dispersion to affect aviation.

In a similar manner to these models, there is operational potential in the drought forecast model presented in Chapter 6, EPS-WP. This model has the same structure, with probabilistic WP predictions (functionally identical to Decider) used to estimate some future meteorological variable: in this case regional drought probabilities. It would be straightforward to implement an operational version of EPS-WP by institutes who produce, or have access to, probabilistic daily SLP (or MSLP) forecasts over the European and North Atlantic domain, as it is relatively trivial and computationally inexpensive to convert these fields to WP predictions and subsequently estimate drought likelihoods. However, the interpretation of "straightforward to implement" is limited to the technical side i.e. producing the forecast data. An operational product would need further work on designing effective visualisations and other requirements of an operational team.

EPS-WP has only been tested with a maximum lead-time of 46 days, and at two- to three-week lead-times the direct ECMWF-EPS precipitation forecasts (i.e. EPS-P) tended to be more accurate. Therefore, the use of EPS-WP is advocated for operational drought prediction at 31- to 46-day lead-times, i.e. the monthly to sub-seasonal range. However, in certain seasons either EPS-P or drought climatological probabilities were more appropriate (Chapter 6.4). Consequently, an operational product may benefit from a seamless combination of models, with the forecasts only produced or disseminated by the model that exhibited highest skill for that region, season and lead-time. Alternatively, the product could run a suite of models in parallel, with forecasts presented for each model, leaving it to the expertise of the operational team to base their forecasts on their knowledge of model differences.

However, it is recognised that EPS-WP was not developed in discussion with potential end-users. Therefore there may be some gap between the existing capabilities of EPS-WP and the requirements of clients, whether they are an operational team or recipients of a "drought outlook" similar to the Hydrological Outlook. Assuming that probabilistic predictions of meteorological drought for monthly to sub-seasonal lead-times are attractive to some, there are still various features that could be easily modified or replaced depending on user requests. For example, drought classes in Chapter 6 were based on relative precipitation thresholds, but it would be simple to alter the definition to be based on a drought index. As the models were tested only in their capability to predict precipitation, this index would have to be derived from this variable alone. The obvious choice would be the SPI, which would also align this proposed tool with the indicators used in the Drought Portal. A further modification might be the regions used. While those used throughout this thesis are of a suitable size for drought studies, end-users might have different requirements, for example those concerned with river catchment management might require catchment boundaries. Switching from the HadUKP data set to a gridded precipitation product would allow different boundaries to be drawn depending on the application. Of course, suitable verification analyses would need to be completed to justify these modifications. Finally, a crucial component of a forecast tool or client-facing product is the communication strategy, i.e. the text and visualisations. Development of such aspects must necessarily follow the structural model requirements outlined above and as such, discussion of these features is considered to be beyond the scope of this thesis.

List of references

Abrahart, R.J., Anctil, F., Coulibaly, P., Dawson, C.W., Mount, N.J., See, L.M., Shamseldin, A.Y., Solomatine, D.P., Toth, E. and Wilby, R.L. (2012) 'Two decades of anarchy? Emerging themes and outstanding challenges for neural network river forecasting', *Progress in Physical Geography: Earth and Environment*, 36(4), pp. 480-513.

AghaKouchak, A. (2014) 'A baseline probabilistic drought forecasting framework using standardized soil moisture index: application to the 2012 United States drought', *Hydrol. Earth Syst. Sci.*, 18(7), pp. 2485-2492.

Ahn, K.-H., Palmer, R. and Steinschneider, S. (2017) 'A hierarchical Bayesian model for regionalized seasonal forecasts: Application to low flows in the northeastern United States', *Water Resources Research*, 53(1), pp. 503-521.

Ahn, K.-H. and Palmer, R.N. (2016) 'Use of a nonstationary copula to predict future bivariate low flow frequency in the Connecticut river basin', *Hydrological Processes*, 30(19), pp. 3518-3532.

Ahrens, B. and Walser, A. (2008) 'Information-Based Skill Scores for Probabilistic Forecasts', *Monthly Weather Review*, 136(1), pp. 352-363.

Alexander, L.V. and Jones, P.D. (2000) 'Updated Precipitation Series for the U.K. and Discussion of Recent Extremes', *Atmospheric Science Letters*, 1(2), pp. 142-150.

Alexander, M.A., Bhatt, U.S., Walsh, J.E., Timlin, M.S., Miller, J.S. and Scott, J.D. (2004) 'The Atmospheric Response to Realistic Arctic Sea Ice Anomalies in an AGCM during Winter', *Journal of Climate*, 17(5), pp. 890-905.

Allan, H.M. (1993) 'What Is a Good Forecast? An Essay on the Nature of Goodness in Weather Forecasting', *Weather and Forecasting*, 8(2), pp. 281-293.

Alley, W.M. (1984) 'Palmer Drought Severity Index: Limitations and Assumptions', *Journal of climate and applied meteorology*, 23(7), pp. 1100-1109.

Almazroui, M., Dambul, R., Islam, M.N. and Jones, P.D. (2015) 'Atmospheric circulation patterns in the Arab region and its relationships with Saudi Arabian surface climate: A preliminary assessment', *Atmospheric Research*, 161-162, pp. 36-51.

American Meteorological Society (AMS) (2013) *Drought: An Information Statement of the American Meteorological Society [Online]*. Available at: https://www.ametsoc.org/policy/2013drought_amsstatement.html (Accessed: 11 May 2015).

American Meteorological Society (AMS) (2014) *Drought. Glossary of Meteorology [Online]*. Available at: http://glossary.ametsoc.org/wiki/Drought (Accessed: 11 May 2015).

Ansell, T.J., Jones, P.D., Allan, R.J., Lister, D., Parker, D.E., Brunet, M., Moberg, A., Jacobeit, J., Brohan, P., Rayner, N.A., Aguilar, E., Alexandersson, H., Barriendos, M., Brandsma, T., Cox, N.J., Della-Marta, P.M., Drebs, A., Founda, D., Gerstengarbe, F., Hickey, K., Jónsson, T., Luterbacher, J., Ø, N., Oesterle, H., Petrakis, M., Philipp, A., Rodwell, M.J., Saladie, O., Sigro, J., Slonosky, V., Srnec, L., Swail, V., García-Suárez, A.M., Tuomenvirta, H., Wang, X., Wanner, H., Werner, P., Wheeler, D. and Xoplaki, E. (2006) 'Daily Mean Sea Level Pressure Reconstructions for the European–North Atlantic Region for the Period 1850–2003', *Journal of Climate*, 19(12), pp. 2717-2742.

Arnal, L., Wood, A.W., Stephens, E., Cloke, H.L. and Pappenberger, F. (2017) 'An Efficient Approach for Estimating Streamflow Forecast Skill Elasticity', *Journal of Hydrometeorology*, 18(6), pp. 1715-1729.

ASCE Task Committee on Application of Artificial Neural Networks in Hydrology (2000) 'Artificial Neural Networks in Hydrology. I: Preliminary Concepts', *Journal of Hydrologic Engineering*, 5(2), pp. 115-123.

Awan, J.A. and Bae, D.-H. (2016) 'Drought prediction over the East Asian monsoon region using the adaptive neuro-fuzzy inference system and the global sea surface temperature anomalies', *International Journal of Climatology*, 36(15), pp. 4767-4777.

Bacanli, U., Firat, M. and Dikbas, F. (2009) 'Adaptive Neuro-Fuzzy Inference System for drought forecasting', *Stochastic Environmental Research and Risk Assessment*, 23(8), pp. 1143-1154.

Baker, L.H., Shaffrey, L.C. and Scaife, A.A. (2018) 'Improved seasonal prediction of UK regional precipitation using atmospheric circulation', *International Journal of Climatology*, 38, pp. 437-453.

Baldwin, M.P. and Dunkerton, T.J. (1998) 'Quasi-biennial modulation of the southern hemisphere stratospheric polar vortex', *Geophysical Research Letters*, 25(17), pp. 3343-3346.

Baldwin, M.P. and Dunkerton, T.J. (2001) 'Stratospheric Harbingers of Anomalous Weather Regimes', *Science*, 294(5542), pp. 581-584.

Balmaseda, M.A., Ferranti, L., Molteni, F. and Palmer, T.N. (2010) 'Impact of 2007 and 2008 Arctic ice anomalies on the atmospheric circulation: Implications for long-range predictions', *Quarterly Journal of the Royal Meteorological Society*, 136(652), pp. 1655-1664.

Bárdossy, A. and Caspary, H.J. (1990) 'Detection of climate change in Europe by analyzing European atmospheric circulation patterns from 1881 to 1989', *Theoretical and Applied Climatology*, 42(3), pp. 155-167.

Bárdossy, A. and Filiz, F. (2005) 'Identification of flood producing atmospheric circulation patterns', *Journal of Hydrology*, 313(1–2), pp. 48-57.

Bárdossy, A. and Plate, E.J. (1992) 'Space-time model for daily rainfall using atmospheric circulation patterns', *Water Resources Research*, 28(5), pp. 1247-1259.

Barnston, A.G. and Livezey, R.E. (1987) 'Classification, Seasonality and Persistence of Low-Frequency Atmospheric Circulation Patterns', *Monthly Weather Review*, 115(6), pp. 1083-1126.

Bartholy, J., Pongrácz, R. and Pattantyús-ábrahám, M. (2006) 'European cyclone track analysis based on ECMWF ERA-40 data sets', *International Journal of Climatology*, 26(11), pp. 1517-1527.

BBC (2018a) *Irish hosepipe ban to continue [Online]*. Available at: https://www.bbc.co.uk/news/world-europe-44979323 (Accessed: 18/09/2018).

BBC (2018b) *Ladygrove Lakes closed after discovery of dead fish [Online]*. Available at: https://www.bbc.com/news/uk-england-oxfordshire-44616683 (Accessed: 20/09/2018).

Behrangi, A., Nguyen, H. and Granger, S. (2015) 'Probabilistic Seasonal Prediction of Meteorological Drought Using the Bootstrap and Multivariate Information', *Journal of Applied Meteorology and Climatology*, 54(7), pp. 1510-1522.

Beljaars, A., C. M., Viterbo, P., Miller, M., J. and Betts, A., K. (1996) 'The Anomalous Rainfall over the United States during July 1993: Sensitivity to Land Surface Parameterization and Soil Moisture Anomalies', *Monthly Weather Review*, 124(3), pp. 362-383.

Beranová, R. and Huth, R. (2008) 'Time variations of the effects of circulation variability modes on European temperature and precipitation in winter', *International Journal of Climatology*, 28(2), pp. 139-158.

Bischoff, S.A. and Vargas, W.M. (2003) 'The 500 and 1000 hPa weather type circulations and their relationship with some extreme climatic conditions over southern South America', *International Journal of Climatology*, 23(5), pp. 541-556.

Blenkinsop, S. and Fowler, H.J. (2007a) 'Changes in drought frequency, severity and duration for the British Isles projected by the PRUDENCE regional climate models', *Journal of Hydrology*, 342(1–2), pp. 50-71.

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

Blenkinsop, S., Jones, P.D., Dorling, S.R. and Osborn, T.J. (2009) 'Observed and modelled influence of atmospheric circulation on central England temperature extremes', *International Journal of Climatology*, 29(11), pp. 1642-1660.

Bloomfield, J.P. and Marchant, B.P. (2013) 'Analysis of groundwater drought building on the standardised precipitation index approach', *Hydrol. Earth Syst. Sci.*, 17(12), pp. 4769-4787.

Boé, J. (2013) 'Modulation of soil moisture–precipitation interactions over France by large scale circulation', *Climate Dynamics*, 40(3), pp. 875-892.

Boeing, G. (2016) 'Visual Analysis of Nonlinear Dynamical Systems: Chaos, Fractals, Self-Similarity and the Limits of Prediction', *Systems*, 4(4), p. 37.

Bollerslev, T. (1986) 'Generalized autoregressive conditional heteroskedasticity', *Journal of Econometrics*, 31(3), pp. 307-327.

Bonaccorso, B., Cancelliere, A. and Rossi, G. (2015) 'Probabilistic forecasting of drought class transitions in Sicily (Italy) using Standardized Precipitation Index and North Atlantic Oscillation Index', *Journal of Hydrology*, 526, pp. 136-150.

Box, G.E.P., Jenkins, G.M. and Reinsel, G.C. (2008) *Time series analysis: forecasting and control*. 4th edn. Hoboken, N.J: John Wiley.

Brier, G.W. (1950) 'Verification of forecasts expressed in terms of probability', *Monthly Weather Review*, 78(1), pp. 1-3.

Brigode, P., Gérardin, M., Bernardara, P., Gailhard, J. and Ribstein, P. (2018) 'Changes in French weather pattern seasonal frequencies projected by a CMIP5 ensemble', *International Journal of Climatology*, 38(10), pp. 3991-4006.

Brinkmann, W.A.R. (2000) 'Modification of a correlation-based circulation pattern classification to reduce within-type variability of temperature and precipitation', *International Journal of Climatology*, 20(8), pp. 839-852.

Bröcker, J. and Smith, L., A. (2007) 'Increasing the Reliability of Reliability Diagrams', *Weather and Forecasting*, 22(3), pp. 651-661.

Brönnimann, S. (2007) 'Impact of El Niño-Southern Oscillation on European climate', *Reviews of Geophysics*, 45(3).

Budikova, D. (2009) 'Role of Arctic sea ice in global atmospheric circulation: A review', *Global and Planetary Change*, 68(3), pp. 149-163.

Buermann, W., Anderson, B., Tucker, C.J., Dickinson, R.E., Lucht, W., Potter, C.S. and Myneni, R.B. (2003) 'Interannual covariability in Northern Hemisphere air temperatures and greenness associated with El Niño-Southern Oscillation and the Arctic Oscillation', *Journal of Geophysical Research: Atmospheres*, 108(D13).

Buizza, R., Bidlot, J.-R., Wedi, N., Fuentes, M., Hamrud, M., Holt, G. and Vitart, F. (2007) 'The new ECMWF VAREPS (Variable Resolution Ensemble Prediction System)', *Quarterly Journal of the Royal Meteorological Society*, 133(624), pp. 681-695.

Burke, E.J., Brown, S.J. and Christidis, N. (2006) 'Modelling the recent evolution of global drought and projections for the twenty-first century with the Hadley Centre climate model', *Journal of Hydrometeorology*, 7(5), pp. 1113-1125.

Burt, T.P. and Ferranti, E.J.S. (2012) 'Changing patterns of heavy rainfall in upland areas: a case study from northern England', *International Journal of Climatology*, 32(4), pp. 518-532.

Cahynová, M. and Huth, R. (2009) 'Enhanced lifetime of atmospheric circulation types over Europe: fact or fiction?', *Tellus A*, 61(3), pp. 407-416.

Caian, M., Koenigk, T., Döscher, R. and Devasthale, A. (2018) 'An interannual link between Arctic sea-ice cover and the North Atlantic Oscillation', *Climate Dynamics*, 50(1), pp. 423-441.

Cancelliere, A., Mauro, G.D., Bonaccorso, B. and Rossi, G. (2007) 'Drought forecasting using the Standardized Precipitation Index', *Water Resources Management*, 21(5), pp. 801-819.

Carbone, G., J, Lu, J. and Brunetti, M. (2018) 'Estimating uncertainty associated with the standardized precipitation index', *International Journal of Climatology*, 38(S1), pp. e607-e616.

Carbone, G.J. and Dow, K. (2005) 'Water resource management and drought forecasts in South Carolina', *JAWRA Journal of the American Water Resources Association*, 41(1), pp. 145-155.

Caron, L., Doblas-Reyes, F.J., Guemas, V., Marcos, R., Massonnet, F., Mishra, N., Prodhomme, C., Turco, M., Ardilouze, C., Ceglar, A. and Toreti, A. (2018) *Improving predictions and management of hydrological extremes (IMPREX). D3.4 Hydro-meteorological indices from successively improving seasonal prediction systems, for sectoral impact assessment*. [Online]. Available at: http://imprex.eu/system/files/generated/files/resource/deliverable3-4-imprex-v1-0.pdf (Accessed: 18/05/2018).

Carpenter, T.M. and Georgakakos, K.P. (2001) 'Assessment of Folsom lake response to historical and potential future climate scenarios: 1. Forecasting', *Journal of Hydrology*, 249(1), pp. 148-175.

Casado, M.J., Pastor, M.A. and Doblas-Reyes, F.J. (2009) 'Euro-Atlantic circulation types and modes of variability in winter', *Theoretical and Applied Climatology*, 96(1), pp. 17-29.

Cassano, E., N., Lynch, A., H., Cassano, J., J. and Koslow, M., R. (2006a) 'Classification of synoptic patterns in the western Arctic associated with extreme events at Barrow, Alaska, USA', *Climate Research*, 30(2), pp. 83-97.

Cassano, J.J., Uotila, P. and Lynch, A. (2006b) 'Changes in synoptic weather patterns in the polar regions in the twentieth and twenty-first centuries, part 1: Arctic', *International Journal of Climatology*, 26(8), pp. 1027-1049.

Cassou, C. (2008) 'Intraseasonal interaction between the Madden-Julian Oscillation and the North Atlantic Oscillation', *Nature*, 455(7212), pp. 523-527.

Cassou, C., Terray, L., Hurrell, J.W. and Deser, C. (2004) 'North Atlantic winter climate regimes: spatial asymmetry, stationarity with time, and oceanic forcing', *Journal of Climate*, 17(5), pp. 1055-1068.

CEH (2018) *UK Hydrological Status Update - early August 2018 [Online]*. Available at: https://www.ceh.ac.uk/news-and-media/blogs/uk-hydrological-status-update-early-august-2018 (Accessed: 20/09/2018).

Chen, J., Li, M. and Wang, W. (2012) 'Statistical Uncertainty Estimation Using Random Forests and Its Application to Drought Forecast', *Mathematical Problems in Engineering*, 2012, p. 12.

Chen, Y.D., Zhang, Q., Xiao, M., Singh, V.P. and Zhang, S. (2016) 'Probabilistic forecasting of seasonal droughts in the Pearl River basin, China', *Stochastic Environmental Research and Risk Assessment*, 30(7), pp. 2031-2040.

Cheng, C.S., Li, G., Li, Q. and Auld, H. (2010) 'A Synoptic Weather Typing Approach to Simulate Daily Rainfall and Extremes in Ontario, Canada: Potential for Climate Change Projections', *Journal of Applied Meteorology and Climatology*, 49(5), pp. 845-866.

Cheng, X. and Wallace, J.M. (1993) 'Cluster Analysis of the Northern Hemisphere Wintertime 500-hPa Height Field: Spatial Patterns', *Journal of the Atmospheric Sciences*, 50(16), pp. 2674-2696.

Christierson, B.v., Vidal, J.-P. and Wade, S.D. (2012) 'Using UKCP09 probabilistic climate information for UK water resource planning', *Journal of Hydrology*, 424, pp. 48-67.

Chun, K.P., Wheater, H. and Onof, C. (2013) 'Prediction of the impact of climate change on drought: an evaluation of six UK catchments using two stochastic approaches', *Hydrological Processes*, 27(11), pp. 1600-1614.

Cohen, J. and Fletcher, C. (2007) 'Improved Skill of Northern Hemisphere Winter Surface Temperature Predictions Based on Land–Atmosphere Fall Anomalies', *Journal of Climate*, 20(16), pp. 4118-4132.

Cohen, J. and Jones, J. (2011) 'A new index for more accurate winter predictions', *Geophysical Research Letters*, 38(21).

Colman, A. and Davey, M. (1999) 'Prediction of summer temperature, rainfall and pressure in Europe from preceding winter North Atlantic Ocean temperature', *International Journal of Climatology*, 19(5), pp. 513-536.

Cook, B.I., Mankin, J.S. and Anchukaitis, K.J. (2018) 'Climate Change and Drought: From Past to Future', *Current Climate Change Reports*, 4(2), pp. 164-179.

Cordery, I. and McCall, M. (2000) 'A model for forecasting drought from teleconnections', *Water Resources Research*, 36(3), pp. 763-768.

Corte-Real, J., Qian, B. and Xu, H. (1998) 'Regional climate change in Portugal: precipitation variability associated with large-scale atmospheric circulation', *International Journal of Climatology*, 18(6), pp. 619-635.

Corte-Real, J., Quian, B. and Xu, H. (1999) 'Circulation patterns, daily precipitation in Portugal and implications for climate change simulated by the second Hadley Centre GCM', *Climate Dynamics*, 15(12), pp. 921-935.

Cuo, L., Pagano, T.C. and Wang, Q.J. (2011) 'A Review of Quantitative Precipitation Forecasts and Their Use in Short- to Medium-Range Streamflow Forecasting', *Journal of Hydrometeorology*, 12(5), pp. 713-728.

Czaja, A. and Frankignoul, C. (1999) 'Influence of the North Atlantic SST on the atmospheric circulation', *Geophysical Research Letters*, 26(19), pp. 2969-2972.

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

Dai, A. and Zhao, T. (2017) 'Uncertainties in historical changes and future projections of drought. Part I: estimates of historical drought changes', *Climatic Change*, 144(3), pp. 519-533.

Davis, J.M. and Rappoport, P.N. (1974) 'The Use of Time Series Analysis Techniques in Forecasting Meteorological Drought', *Monthly Weather Review*, 102(2), pp. 176-180.

Dawson, A., Palmer, T.N. and Corti, S. (2012) 'Simulating regime structures in weather and climate prediction models', *Geophysical Research Letters*, 39(21).

Dawson, C.W. and Wilby, R.L. (2001) 'Hydrological modelling using artificial neural networks', *Progress in Physical Geography: Earth and Environment*, 25(1), pp. 80-108.

Day, G.N. (1985) 'Extended Streamflow Forecasting Using NWSRFS', *Journal of Water Resources Planning and Management*, 111(2), pp. 157-170.

Dee, D.P., Uppala, S.M., Simmons, A.J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M.A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A.C.M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A.J., Haimberger, L., Healy, S.B., Hersbach, H., Hólm, E.V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A.P., Monge-Sanz, B.M., Morcrette, J.J., Park, B.K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.N. and Vitart, F. (2011) 'The ERA-Interim reanalysis: configuration and performance of the data assimilation system', *Quarterly Journal of the Royal Meteorological Society*, 137(656), pp. 553-597.

Dehghani, M., Saghafian, B., Nasiri Saleh, F., Farokhnia, A. and Noori, R. (2014) 'Uncertainty analysis of streamflow drought forecast using artificial neural networks and Monte-Carlo simulation', *International Journal of Climatology*, 34(4), pp. 1169-1180.

Demirel, M.C., Booij, M.J. and Hoekstra, A.Y. (2015) 'The skill of seasonal ensemble low-flow forecasts in the Moselle River for three different hydrological models', *Hydrol. Earth Syst. Sci.*, 19(1), pp. 275-291.

Demuzere, M., Werner, M., van Lipzig, N.P.M. and Roeckner, E. (2009) 'An analysis of present and future ECHAM5 pressure fields using a classification of circulation patterns', *International Journal of Climatology*, 29(12), pp. 1796-1810.

Djerbouai, S. and Souag-Gamane, D. (2016) 'Drought Forecasting Using Neural Networks, Wavelet Neural Networks, and Stochastic Models: Case of the Algerois Basin in North Algeria', *Water Resources Management*, 30(7), pp. 2445-2464.

Domonkos, P., Kyselý, J., Piotrowicz, K., Petrovic, P. and Likso, T. (2003) 'Variability of extreme temperature events in south–central Europe during the 20th century and its relationship with large-scale circulation', *International Journal of Climatology*, 23(9), pp. 987-1010.

Donat, M.G., Leckebusch, G.C., Pinto, J.G. and Ulbrich, U. (2010) 'European storminess and associated circulation weather types: future changes deduced from a multi-model ensemble of GCM simulations', *Climate Research*, 42(1), pp. 27-43.

Dong, B., Sutton, R.T., Woolings, T. and Hodges, K. (2013) 'Variability of the North Atlantic summer storm track: mechanisms and impacts on European climate', *Environmental Research Letters*, 8(3), p. 034037.

Doswell, C.A., Brooks, H.E. and Maddox, R.A. (1996) 'Flash Flood Forecasting: An Ingredients-Based Methodology', *Weather and Forecasting*, 11(4), pp. 560-581.

Dunstone, N., Smith, D., Scaife, A., Hermanson, L., Eade, R., Robinson, N., Andrews, M. and Knight, J. (2016) 'Skilful predictions of the winter North Atlantic Oscillation one year ahead', *Nature Geosci*, 9(11), pp. 809-814.

Durdu, Ö. (2010) 'Application of linear stochastic models for drought forecasting in the Büyük Menderes river basin, western Turkey', *Stochastic Environmental Research and Risk Assessment*, 24(8), pp. 1145-1162.

Dutra, E., Pozzi, W., Wetterhall, F., Di Giuseppe, F., Magnusson, L., Naumann, G., Barbosa, P., Vogt, J. and Pappenberger, F. (2014) 'Global meteorological drought – Part 2: Seasonal forecasts', *Hydrol. Earth Syst. Sci.*, 18(7), pp. 2669-2678.

ECMWF (2017) *ECMWF Model Description CY43R1 [Online]*. Available at: https://confluence.ecmwf.int/display/S2S/ECMWF+Model+Description+CY43R1 (Accessed: 03/06/2018).

Edwards, D.C. and McKee, T.B. (1997) *Characteristics of 20th century drought in the United States at multiple time scales* (Climatology Rep. 97–2). Colorado State University, Ft. Collins, CO. [Online]. Available at: http://ccc.atmos.colostate.edu/edwards.pdf.

Emerton, R.E., Stephens, E.M., Pappenberger, F., Pagano, T.C., Weerts, A.H., Wood, A.W., Salamon, P., Brown, J.D., Hjerdt, N., Donnelly, C., Baugh, C.A. and Cloke, H.L. (2016) 'Continental and global scale flood forecasting systems', *Wiley Interdisciplinary Reviews: Water*, 3(3), pp. 391-418.

Entin, J.K., Robock, A., Vinnikov, K.Y., Hollinger, S.E., Liu, S. and Namkhai, A. (2000) 'Temporal and spatial scales of observed soil moisture variations in the extratropics', *Journal of Geophysical Research: Atmospheres*, 105(D9), pp. 11865-11877.

Environment Agency (2018) *State of the environment: water resources*. [Online]. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/709924/State_of_the_environment_water_resources_report.pdf (Accessed: 27/06/2018).

Epstein, E., S. (1969) 'A Scoring System for Probability Forecasts of Ranked Categories', *Journal of Applied Meteorology*, 8(6), pp. 985-987.

Esteban, P., Jones, P.D., Martín-Vide, J. and Mases, M. (2005) 'Atmospheric circulation patterns related to heavy snowfall days in Andorra, Pyrenees', *International Journal of Climatology*, 25(3), pp. 319-329.

Esteban, P., Martin-Vide, J. and Mases, M. (2006) 'Daily atmospheric circulation catalogue for western Europe using multivariate techniques', *International Journal of Climatology*, 26(11), pp. 1501-1515.

Fayos, J. and Fayos, C. (2007) 'Wind Data Mining by Kohonen Neural Networks', *PLOS ONE*, 2(2), p. e210.

Fereday, D., Chadwick, R., Knight, J. and Scaife, A. (2018) 'Atmospheric Dynamics is the Largest Source of Uncertainty in Future Winter European Rainfall', *Journal of Climate*, 31(3), pp. 963-977.

Fereday, D.R., Knight, J.R., Scaife, A.A., Folland, C.K. and Philipp, A. (2008) 'Cluster Analysis of North Atlantic–European Circulation Types and Links with Tropical Pacific Sea Surface Temperatures', *Journal of Climate*, 21(15), pp. 3687-3703.

Fernández-González, S., del Río, S., Castro, A., Penas, A., Fernández-Raga, M., Calvo, A.I. and Fraile, R. (2012) 'Connection between NAO, weather types and precipitation in León, Spain (1948–2008)', *International Journal of Climatology*, 32(14), pp. 2181-2196.

Fernández, C., Vega, J., Fonturbel, T. and Jiménez, E. (2009) 'Streamflow drought time series forecasting: a case study in a small watershed in North West Spain', *Stochastic Environmental Research and Risk Assessment*, 23(8), pp. 1063-1070.

Fernau, M.E. and Samson, P.J. (1990) 'Use of Cluster Analysis to Define Periods of Similar Meteorology and Precipitation Chemistry in Eastern North America. Part I: Transport Patterns', *Journal of Applied Meteorology*, 29(8), pp. 735-750.

Ferranti, L., Corti, S. and Janousek, M. (2015) 'Flow-dependent verification of the ECMWF ensemble over the Euro-Atlantic sector', *Quarterly Journal of the Royal Meteorological Society*, 141(688), pp. 916-924.

Ferranti, L., Magnusson, L., Vitart, F. and Richardson, D.S. (2018) 'How far in advance can we predict changes in large-scale flow leading to severe cold conditions over Europe?', *Quarterly Journal of the Royal Meteorological Society*, 144(715), pp. 1788-1802.

Fischer, E.M., Seneviratne, S.I., Lüthi, D. and Schär, C. (2007) 'Contribution of land-atmosphere coupling to recent European summer heat waves', *Geophysical Research Letters*, 34(6).

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), pp. 1163-1179.

Fleig, A.K., Tallaksen, L.M., Hisdal, H., Stahl, K. and Hannah, D.M. (2010) 'Inter-comparison of weather and circulation type classifications for hydrological drought development', *Physics and Chemistry of the Earth, Parts A/B/C*, 35(9–12), pp. 507-515.

Flood Forecasting Centre (2017) *Annual Review 2016/17*. [Online]. Available at: http://www.ffc-environment-agency.metoffice.gov.uk/media/pdf/FFC-Annual-Review-2016-17.pdf (Accessed: 09/11/2018).

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', *Hydrol. Earth Syst. Sci.*, 19(5), pp. 2353-2375.

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

Food and Agriculture Organization of the United Nations (1986) *Chapter 3: Crop water needs [Online]*. Available at: http://www.fao.org/docrep/s2022e/s2022e07.htm (Accessed: 12 May 2015).

Foulds, S.A. and Macklin, M.G. (2016) 'A hydrogeomorphic assessment of twenty-first century floods in the UK', *Earth Surface Processes and Landforms*, 41(2), pp. 256-270.

Fowler, H.J., Blenkinsop, S. and Tebaldi, C. (2007) 'Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling', *International Journal of Climatology*, 27(12), pp. 1547-1578.

Fowler, H.J. and Kilsby, C.G. (2002a) 'Precipitation and the North Atlantic Oscillation: a study of climatic variability in northern England', *International Journal of Climatology*, 22(7), pp. 843-866.

Fowler, H.J. and Kilsby, C.G. (2002b) 'A weather-type approach to analysing water resource drought in the Yorkshire region from 1881 to 1998', *Journal of Hydrology*, 262(1–4), pp. 177-192.

Fowler, H.J. and Kilsby, C.G. (2003) 'A regional frequency analysis of United Kingdom extreme rainfall from 1961 to 2000', *International Journal of Climatology*, 23(11), pp. 1313-1334.

Fowler, H.J., Kilsby, C.G., O'Connell, P.E. and Burton, A. (2005) 'A weather-type conditioned multi-site stochastic rainfall model for the generation of scenarios of climatic variability and change', *Journal of Hydrology*, 308(1–4), pp. 50-66.

Fowler, H.J., Kilsby, C.G. and O'Connell, P.E. (2000) 'A stochastic rainfall model for the assessment of regional water resource systems under changed climatic condition', *Hydrol. Earth Syst. Sci.*, 4(2), pp. 263-281.

García-Herrera, R., Díaz, J., Trigo, R.M., Luterbacher, J. and Fischer, E.M. (2010) 'A Review of the European Summer Heat Wave of 2003', *Critical Reviews in Environmental Science and Technology*, 40(4), pp. 267-306.

García-Serrano, J., Frankignoul, C., Gastineau, G. and de la Cámara, A. (2015) 'On the Predictability of the Winter Euro-Atlantic Climate: Lagged Influence of Autumn Arctic Sea Ice', *Journal of Climate*, 28(13), pp. 5195-5216.

Garner, G., Hannah, D.M. and Watts, G. (2017) 'Climate change and water in the UK: Recent scientific evidence for past and future change', *Progress in Physical Geography*, 41(2), pp. 154-170.

Gavin, H., Hammond, C. and Piper, B. (2013) *Managing Through Drought: Code of Practice and Guidance for Water Companies on Water use Restrictions - 2013 [incorporating lessons from the 2011-12 drought]*. [Online]. Available at: http://water.wuk1.emsystem.co.uk/home/policy/publications/archive/industry-guidance/managing-through-drought-cop/managing-through-drought-code-of-practice.pdf.

Ghosh, R., Müller, W.A., Baehr, J. and Bader, J. (2017) 'Impact of observed North Atlantic multidecadal variations to European summer climate: a linear baroclinic response to surface heating', *Climate Dynamics*, 48(11), pp. 3547-3563.

Giuntoli, I., Renard, B., Vidal, J.P. and Bard, A. (2013) 'Low flows in France and their relationship to large-scale climate indices', *Journal of Hydrology*, 482, pp. 105-118.

Goodess, C.M. and Jones, P.D. (2002) 'Links between circulation and changes in the characteristics of Iberian rainfall', *International Journal of Climatology*, 22(13), pp. 1593-1615.

Grebogi, C., Ott, E. and Yorke, J.A. (1987) 'Chaos, Strange Attractors, and Fractal Basin Boundaries in Nonlinear Dynamics', *Science*, 238(4827), pp. 632-638.

Guerreiro, S.B., Kilsby, C. and Fowler, H.J. (2017) 'Assessing the threat of future megadrought in Iberia', *International Journal of Climatology*, 37(15), pp. 5024-5034.

Guttman, N.B. (1998) 'Comparing the Palmer Drought Index and the Standardized Precipitation Index', *JAWRA Journal of the American Water Resources Association*, 34(1), pp. 113-121.

Guttman, N.B. (1999) 'Accepting the Standardized Precipitation Index: a Calculation Algorithm', *JAWRA Journal of the American Water Resources Association*, 35(2), pp. 311-322.

Hall, R., J., Scaife, A., A., Hanna, E., Jones, J., M. and Erdélyi, R. (2017) 'Simple Statistical Probabilistic Forecasts of the Winter NAO', *Weather and Forecasting*, 32(4), pp. 1585-1601.

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), pp. e660-e677.

Halpert, M.S. and Ropelewski, C.F. (1992) 'Surface Temperature Patterns Associated with the Southern Oscillation', *Journal of Climate*, 5(6), pp. 577-593.

Hamlet, A.F. and Lettenmaier, D.P. (1999) 'Columbia River Streamflow Forecasting Based on ENSO and PDO Climate Signals', *Journal of Water Resources Planning and Management*, 125(6), pp. 333-341.

Han, P., Wang, P.X., Zhang, S.Y. and Zhu, D.H. (2010) 'Drought forecasting based on the remote sensing data using ARIMA models', *Mathematical and Computer Modelling*, 51(11–12), pp. 1398-1403.

Hannachi, A., Straus, D.M., Franzke, C.L.E., Corti, S. and Woollings, T. (2017) 'Low-frequency nonlinearity and regime behavior in the Northern Hemisphere extratropical atmosphere', *Reviews of Geophysics*, 55(1), pp. 199-234.

Hannaford, J., Lloyd-Hughes, B., Keef, C., Parry, S. and Prudhomme, C. (2011) 'Examining the large-scale spatial coherence of European drought using regional indicators of precipitation and streamflow deficit', *Hydrological Processes*, 25(7), pp. 1146-1162.

Hannaford, J., Turner, S., Lewis, M. and Clemas, S. (2018) *Hydrological summary for the United Kingdom: August 2018*. Wallingford, UK. [Online]. Available at: http://nora.nerc.ac.uk/id/eprint/520977/.

Hao, Z., Hong, Y., Xia, Y., Singh, V.P., Hao, F. and Cheng, H. (2016) 'Probabilistic drought characterization in the categorical form using ordinal regression', *Journal of Hydrology*, 535, pp. 331-339.

Hao, Z., Singh, V.P. and Xia, Y. (2018) 'Seasonal Drought Prediction: Advances, Challenges, and Future Prospects', *Reviews of Geophysics*, 56(1), pp. 108-141.

Hao, Z., Xia, Y., Luo, L., Singh, V.P., Ouyang, W. and Hao, F. (2017) 'Toward a categorical drought prediction system based on U.S. Drought Monitor (USDM) and climate forecast', *Journal of Hydrology*, 551, pp. 300-305.

Harrigan, S., Prudhomme, C., Parry, S., Smith, K. and Tanguy, M. (2018) 'Benchmarking ensemble streamflow prediction skill in the UK', *Hydrol. Earth Syst. Sci.*, 22(3), pp. 2023-2039.

Hartigan, J.A. and Wong, M.A. (1979) 'Algorithm AS 136: A K-Means Clustering Algorithm', *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 28(1), pp. 100-108.

Hauser, M., Orth, R. and Seneviratne Sonia, I. (2016) 'Role of soil moisture versus recent climate change for the 2010 heat wave in western Russia', *Geophysical Research Letters*, 43(6), pp. 2819-2826.

Hay, L.E., McCabe, G.J., Wolock, D.M. and Ayers, M.A. (1991) 'Simulation of precipitation by weather type analysis', *Water Resources Research*, 27(4), pp. 493-501.

Hay, L.E., McCabe, G.J., Wolock, D.M. and Ayers, M.A. (1992) 'Use of weather types to disaggregate general circulation model predictions', *Journal of Geophysical Research: Atmospheres*, 97(D3), pp. 2781-2790.

Haylock, M.R. and Goodess, C.M. (2004) 'Interannual variability of European extreme winter rainfall and links with mean large-scale circulation', *International Journal of Climatology*, 24(6), pp. 759-776.

Heim, R., R. Jr. (2002) 'A Review of Twentieth-Century Drought Indices Used in the United States', *Bulletin of the American Meteorological Society*, 83(8), pp. 1149-1166.

Heinrich, G. and Gobiet, A. (2012) 'The future of dry and wet spells in Europe: a comprehensive study based on the ENSEMBLES regional climate models', *International Journal of Climatology*, 32(13), pp. 1951-1970.

Herceg-Bulić, I., Mezzina, B., Kucharski, F., Ruggieri, P. and King, M.P. (2017) 'Wintertime ENSO influence on late spring European climate: the stratospheric response and the role of North Atlantic SST', *International Journal of Climatology*, 37, pp. 87-108.

Hess, P. and Brezowsky, H. (1952) 'Katalog der Grosswetterlagen Europas [*Catalogue of large scale weather conditions over Europe*]', *Berichte des Deustchen Wetterdienstes in der US-zone*, 33.

Hewitson, B.C. and Crane, R.G. (2002) 'Self-organizing maps: applications to synoptic climatology', *Climate Research*, 22(1), pp. 13-26.

Hidalgo-Muñoz, J.M., Gámiz-Fortis, S.R., Castro-Díez, Y., Argüeso, D. and Esteban-Parra, M.J. (2015) 'Long-range seasonal streamflow forecasting over the Iberian Peninsula using large-scale atmospheric and oceanic information', *Water Resources Research*, 51(5), pp. 3543-3567.

Hoerling, M. and Kumar, A. (2003) 'The perfect ocean for drought', *Science*, 299(5607), pp. 691-694.

Hope, P.K. (2006) 'Projected future changes in synoptic systems influencing southwest Western Australia', *Climate Dynamics*, 26(7), pp. 765-780.

Hope, P.K., Drosdowsky, W. and Nicholls, N. (2006) 'Shifts in the synoptic systems influencing southwest Western Australia', *Climate Dynamics*, 26(7), pp. 751-764.

Hurrell, J.W. (1995) 'Decadal Trends in the North Atlantic Oscillation: Regional Temperatures and Precipitation', *Science*, 269(5224), pp. 676-679.

Hurrell, J.W. and Deser, C. (2009) 'North Atlantic climate variability: The role of the North Atlantic Oscillation', *Journal of Marine Systems*, 78(1), pp. 28-41.

Huth, R. (1996) 'Properties of the circulation classification scheme based on the rotated principal component analysis', *Meteorology and Atmospheric Physics*, 59(3), pp. 217-233.

Huth, R. (1997) 'Continental-Scale Circulation in the UKHI GCM', *Journal of Climate*, 10(7), pp. 1545-1561.

Huth, R. (2000) 'A circulation classification scheme applicable in GCM studies', *Theoretical and Applied Climatology*, 67(1), pp. 1-18.

Huth, R., Beck, C., Philipp, A., Demuzere, M., Ustrnul, Z., Cahynová, M., Kyselý, J. and Tveito, O.E. (2008) 'Classifications of Atmospheric Circulation Patterns', *Annals of the New York Academy of Sciences*, 1146(1), pp. 105-152.

Hwang, Y. and Gregory, J.C. (2009) 'Ensemble Forecasts of Drought Indices Using a Conditional Residual Resampling Technique', *Journal of Applied Meteorology and Climatology*, 48(7), pp. 1289-1301.

Ineson, S. and Scaife, A.A. (2009) 'The role of the stratosphere in the European climate response to El Nĩo', *Nature Geoscience*, 2(1), pp. 32-36.

Jacobeit, J., Wanner, H., Luterbacher, J., Beck, C., Philipp, A. and Sturm, K. (2003) 'Atmospheric circulation variability in the North-Atlantic-European area since the mid-seventeenth century', *Climate Dynamics*, 20(4), pp. 341-352.

James, P.M. (2007) 'An objective classification method for Hess and Brezowsky Grosswetterlagen over Europe', *Theoretical and Applied Climatology*, 88(1), pp. 17-42.

Jenkinson, A.F. and Collison, F.P. (1977) 'An initial climatology of gales over the North Sea', *Synoptic climatology branch memorandum*, 62, p. 18.

Jiang, N. (2011) 'A new objective procedure for classifying New Zealand synoptic weather types during 1958–2008', *International Journal of Climatology*, 31(6), pp. 863-879.

Jiang, R., Xie, J., He, H., Luo, J. and Zhu, J. (2015) 'Use of four drought indices for evaluating drought characteristics under climate change in Shaanxi, China: 1951–2012', *Natural Hazards*, 75(3), pp. 2885-2903.

Jones, M.R., Blenkinsop, S., Fowler, H.J. and Kilsby, C.G. (2014) 'Objective classification of extreme rainfall regions for the UK and updated estimates of trends in regional extreme rainfall', *International Journal of Climatology*, 34(3), pp. 751-765.

Jones, M.R., Fowler, H.J., Kilsby, C.G. and Blenkinsop, S. (2013a) 'An assessment of changes in seasonal and annual extreme rainfall in the UK between 1961 and 2009', *International Journal of Climatology*, 33(5), pp. 1178-1194.

Jones, P.D., Harpham, C. and Briffa, K.R. (2013b) 'Lamb weather types derived from reanalysis products', *International Journal of Climatology*, 33(5), pp. 1129-1139.

Jones, P.D., Hulme, M. and Briffa, K.R. (1993) 'A comparison of Lamb circulation types with an objective classification scheme', *International Journal of Climatology*, 13(6), pp. 655-663.

Jones, P.D. and Lister, D.H. (1998) 'Riverflow reconstructions for 15 catchments over England and Wales and an assessment of hydrologic drought since 1865', *International Journal of Climatology*, 18(9), pp. 999-1013.

Jung, T., Hilmer, M., Ruprecht, E., Kleppek, S., Gulev, S., K. and Zolina, O. (2003) 'Characteristics of the Recent Eastward Shift of Interannual NAO Variability', *Journal of Climate*, 16(20), pp. 3371-3382.

Kalkstein, L., S., Tan, G. and Skindlov, J., A. (1987) 'An Evaluation of Three Clustering Procedures for Use in Synoptic Climatological Classification', *Journal of Climate and Applied Meteorology*, 26(6), pp. 717-730.

Karl, T., Knight, R.W. and National Climatic Data, C. (1985) *Atlas of monthly Palmer hydrological drought indices (1931-1983) for the contiguous United States*. Asheville, N.C.: National Climatic Data Center.

Karl, T., Quinlan, F. and Ezell, D.S. (1987) 'Drought Termination and Amelioration: Its Climatological Probability', *Journal of Climate and Applied Meteorology*, 26(9), pp. 1198-1209.

Karl, T.R. (1983) 'Some Spatial Characteristics of Drought Duration in the United States', *Journal of Climate and Applied Meteorology*, 22(8), pp. 1356-1366.

Karl, T.R. (1986) 'The Sensitivity of the Palmer Drought Severity Index and Palmer's Z-Index to their Calibration Coefficients Including Potential Evapotranspiration', *Journal of Climate and Applied Meteorology*, 25(1), pp. 77-86.

Kendon, M., Marsh, T. and Parry, S. (2013) 'The 2010–2012 drought in England and Wales', *Weather*, 68(4), pp. 88-95.

Kharin, V.V. and Zwiers, F.W. (2003) 'Improved Seasonal Probability Forecasts', *Journal of Climate*, 16(11), pp. 1684-1701.

Kidson, J.W. and Watterson, I.G. (1995) 'A synoptic climatological evaluation of the changes in the CSIRO nine-level model with doubled CO2 in the New Zealand region', *International Journal of Climatology*, 15(11), pp. 1179-1194.

Kim, T. and Valdés, J. (2003) 'Nonlinear Model for Drought Forecasting Based on a Conjunction of Wavelet Transforms and Neural Networks', *Journal of Hydrologic Engineering*, 8(6), pp. 319-328.

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), pp. 331-344.

Kirtman, B. and Pirani, A. (2009) 'The State of the Art of Seasonal Prediction: Outcomes and Recommendations from the First World Climate Research Program Workshop on Seasonal Prediction', *Bulletin of the American Meteorological Society*, 90(4), pp. 455-458.

Kleeman, R. (2002) 'Measuring Dynamical Prediction Utility Using Relative Entropy', *Journal of the Atmospheric Sciences*, 59(13), pp. 2057-2072.

Knight, J.R., Folland, C.K. and Scaife, A.A. (2006) 'Climate impacts of the Atlantic multidecadal oscillation', *Geophysical Research Letters*, 33(17).

Knutti, R., Furrer, R., Tebaldi, C., Cermak, J. and MeehlGerald, A. (2010) 'Challenges in Combining Projections from Multiple Climate Models', *Journal of Climate*, 23(10), pp. 2739-2758.

Knutti, R. and Sedláček, J. (2012) 'Robustness and uncertainties in the new CMIP5 climate model projections', *Nature Climate Change*, 3, p. 369.

Kolstad, E.W. and Årthun, M. (2018) 'Seasonal Prediction from Arctic Sea Surface Temperatures: Opportunities and Pitfalls', *Journal of Climate*, 31(20), pp. 8197-8210.

Koster, R.D., Mahanama, S.P.P., Yamada, T.J., Balsamo, G., Berg, A.A., Boisserie, M., Dirmeyer, P.A., Doblas-Reyes, F.J., Drewitt, G., Gordon, C.T., Guo, Z., Jeong, J.H., Lawrence, D.M., Lee, W.S., Li, Z., Luo, L., Malyshev, S., Merryfield, W.J., Seneviratne, S.I., Stanelle, T., van den Hurk, B.J.J.M., Vitart, F. and Wood, E.F. (2010) 'Contribution of land surface initialization to subseasonal forecast skill: First results from a multi-model experiment', *Geophysical Research Letters*, 37(2).

Kryjov, V.N. and Min, Y.-M. (2016) 'Predictability of the wintertime Arctic Oscillation based on autumn circulation', *International Journal of Climatology*, 36(12), pp. 4181-4186.

Kučerová, M., Beck, C., Philipp, A. and Huth, R. (2017) 'Trends in frequency and persistence of atmospheric circulation types over Europe derived from a multitude of classifications', *International Journal of Climatology*, 37(5), pp. 2502-2521.

Kullback, S. and Leibler, R.A. (1951) 'On Information and Sufficiency', *Ann. Math. Statist.*, 22(1), pp. 79-86.

Kumar, V. and Panu, U. (1997) 'Predictive Assessment of Severity of Agricultural Droughts based on Agro-climatic Factors', *JAWRA Journal of the American Water Resources Association*, 33(6), pp. 1255-1264.

Kyselý, J. (2002) 'Temporal fluctuations in heat waves at Prague–Klementinum, the Czech Republic, from 1901–97, and their relationships to atmospheric circulation', *International Journal of Climatology*, 22(1), pp. 33-50.

Kyselý, J. (2007) 'Implications of enhanced persistence of atmospheric circulation for the occurrence and severity of temperature extremes', *International Journal of Climatology*, 27(5), pp. 689-695.

Kyselý, J. and Domonkos, P. (2006) 'Recent increase in persistence of atmospheric circulation over Europe: comparison with long-term variations since 1881', *International Journal of Climatology*, 26(4), pp. 461-483.

Kyselý, J. and Huth, R. (2006) 'Changes in atmospheric circulation over Europe detected by objective and subjective methods', *Theoretical and Applied Climatology*, 85(1), pp. 19-36.

Lamb, H. (1972) 'British Isles weather types and a register of the daily sequence of circulation patterns 1861-1971', *Geophysical Memoirs*, 116, p. 85.

Lamb, H.H. (1950) 'Types and spells of weather around the year in the British Isles: Annual trends, seasonal structure of the year, singularities', *Quarterly Journal of the Royal Meteorological Society*, 76(330), pp. 393-429.

Lavaysse, C., Vogt, J. and Pappenberger, F. (2015) 'Early warning of drought in Europe using the monthly ensemble system from ECMWF', *Hydrol. Earth Syst. Sci.*, 19(7), pp. 3273-3286.

Lavers, D., Prudhomme, C. and Hannah, D.M. (2010) 'Large-scale climate, precipitation and British river flows: Identifying hydroclimatological connections and dynamics', *Journal of Hydrology*, 395(3), pp. 242-255.

Lavers, D., Zsoter, E., Richardson, D. and Pappenberger, F. (2017) 'An Assessment of the ECMWF Extreme Forecast Index for Water Vapor Transport during Boreal Winter', *Weather and Forecasting*, 32(4), pp. 1667-1674.

Lavers, D.A., Pappenberger, F., Richardson, D.S. and Zsoter, E. (2016a) 'ECMWF Extreme Forecast Index for water vapor transport: A forecast tool for atmospheric rivers and extreme precipitation', *Geophysical Research Letters*, 43(22), pp. 11,852-11,858.

Lavers, D.A., Pappenberger, F. and Zsoter, E. (2014) 'Extending medium-range predictability of extreme hydrological events in Europe', *Nature Communications*, 5, p. 5382.

Lavers, D.A., Richardson, D.S., Ramos, A.M., Zsoter, E., Pappenberger, F. and Trigo, R.M. (2018) 'Earlier awareness of extreme winter precipitation across the western Iberian Peninsula', *Meteorological Applications*, 25(4), pp. 622-628.

Lavers, D.A., Waliser, D.E., Ralph, F.M. and Dettinger, M.D. (2016b) 'Predictability of horizontal water vapor transport relative to precipitation: Enhancing situational awareness for forecasting western U.S. extreme precipitation and flooding', *Geophysical Research Letters*, 43(5), pp. 2275-2282.

Leung, L.-Y. and North, G., R. (1990) 'Information Theory and Climate Prediction', *Journal of Climate*, 3(1), pp. 5-14.

Leutbecher, M. and Palmer, T.N. (2008) 'Ensemble forecasting', *Journal of Computational Physics*, 227(7), pp. 3515-3539.

Li, J., Zhou, S. and Hu, R. (2016) 'Hydrological Drought Class Transition Using SPI and SRI Time Series by Loglinear Regression', *Water Resources Management*, 30(2), pp. 669-684.

Lim, E.-P., Harry, H.H., David, L.T.A., Andrew, C. and Oscar, A. (2011) 'Dynamical, Statistical–Dynamical, and Multimodel Ensemble Forecasts of Australian Spring Season Rainfall', *Monthly Weather Review*, 139(3), pp. 958-975.

Lin, J. (1991) 'Divergence measures based on the Shannon entropy', *IEEE Transactions on Information Theory*, 37(1), pp. 145-151.

Liu, D., Mishra, A.K., Yu, Z., Yang, C., Konapala, G. and Vu, T. (2017) 'Performance of SMAP, AMSR-E and LAI for weekly agricultural drought forecasting over continental United States', *Journal of Hydrology*, 553(Supplement C), pp. 88-104.

Liu, W.T. and Juárez, R.I.N. (2001) 'ENSO drought onset prediction in northeast Brazil using NDVI', *International Journal of Remote Sensing*, 22(17), pp. 3483-3501.

Liu, Y. and Weisberg, R.H. (2011) 'A review of self-organizing map applications in meteorology and oceanography', in Mwasiagi, J.I. (ed.) *Self-Organizing Maps Applications and Novel Algorithm Design*. Rijeka, Croatia: InTech.

Lloyd-Hughes, B. (2014) 'The impracticality of a universal drought definition', *Theoretical and Applied Climatology*, 117(3-4), pp. 607-611.

Lloyd-Hughes, B. and Saunders, M.A. (2002) 'A drought climatology for Europe', *International Journal of Climatology*, 22(13), pp. 1571-1592.

Lohani, V.K. and Loganathan, G.V. (1997) 'An early warning system for drought management using the Palmer Drought Index', *JAWRA Journal of the American Water Resources Association*, 33(6), pp. 1375-1386.

López-Parages, J., Rodríguez-Fonseca, B., Dommenget, D. and Frauen, C. (2016) 'ENSO influence on the North Atlantic European climate: a non-linear and non-stationary approach', *Climate Dynamics*, 47(7), pp. 2071-2084.

Lorenz, D., J. , Otkin, J., A. , Svoboda, M., Hain, C., R. , Anderson, M., C. and Zhong, Y. (2017) 'Predicting the U.S. Drought Monitor Using Precipitation, Soil Moisture, and Evapotranspiration Anomalies. Part II: Intraseasonal Drought Intensification Forecasts', *Journal of Hydrometeorology*, 18(7), pp. 1963-1982.

Lorenz, D.J., Otkin, J.A., Svoboda, M., Hain, C.R. and Zhong, Y. (2018) 'Forecasting Rapid Drought Intensification Using the Climate Forecast System (CFS)', *Journal of Geophysical Research: Atmospheres*, 123(16), pp. 8365-8373.

Lorenz, E., N. (1969) 'Atmospheric Predictability as Revealed by Naturally Occurring Analogues', *Journal of the Atmospheric Sciences*, 26(4), pp. 636-646.

Lorenz, E.N. (1963) 'Deterministic Nonperiodic Flow', *Journal of the Atmospheric Sciences*, 20(2), pp. 130-141.

Lorenzo, M.N., Ramos, A.M., Taboada, J.J. and Gimeno, L. (2011) 'Changes in Present and Future Circulation Types Frequency in Northwest Iberian Peninsula', *PLOS ONE*, 6(1), p. e16201.

Lorenzo, M.N., Taboada, J.J. and Gimeno, L. (2008) 'Links between circulation weather types and teleconnection patterns and their influence on precipitation patterns in Galicia (NW Spain)', *International Journal of Climatology*, 28(11), pp. 1493-1505.

Lund, I.A. (1963) 'Map-Pattern Classification by Statistical Methods', *Journal of Applied Meteorology*, 2(1), pp. 56-65.

Lynch, A., Uotila, P. and Cassano, J.J. (2006) 'Changes in synoptic weather patterns in the polar regions in the twentieth and twenty-first centuries, part 2: Antarctic', *International Journal of Climatology*, 26(9), pp. 1181-1199.

MacLachlan, C., Arribas, A., Peterson, K.A., Maidens, A., Fereday, D., Scaife, A.A., Gordon, M., Vellinga, M., Williams, A., Comer, R.E., Camp, J., Xavier, P. and Madec, G. (2015) 'Global Seasonal forecast system version 5 (GloSea5): a high-resolution seasonal

forecast system', *Quarterly Journal of the Royal Meteorological Society*, 141(689), pp. 1072-1084.

Madadgar, S., AghaKouchak, A., Shukla, S., Wood, A.W., Cheng, L., Hsu, K.-L. and Svoboda, M. (2016) 'A hybrid statistical-dynamical framework for meteorological drought prediction: Application to the southwestern United States', *Water Resources Research*, 52(7), pp. 5095-5110.

Madadgar, S. and Moradkhani, H. (2013) 'A Bayesian Framework for Probabilistic Seasonal Drought Forecasting', *Journal of Hydrometeorology*, 14(6), pp. 1685-1705.

Madadgar, S. and Moradkhani, H. (2014) 'Spatio-temporal drought forecasting within Bayesian networks', *Journal of Hydrology*, 512, pp. 134-146.

Maheras, P., Tolika, K., Anagnostopoulou, C., Makra, L., Szpirosz, K. and Károssy, C. (2018) 'Relationship between mean and extreme precipitation and circulation types over Hungary', *International Journal of Climatology*, 38(12), pp. 4518-4532.

Maity, R., Sharma, A., Nagesh Kumar, D. and Chanda, K. (2013) 'Characterizing Drought Using the Reliability-Resilience-Vulnerability Concept', *Journal of Hydrologic Engineering*, 18(7), pp. 859-869.

Malby, A.R., Whyatt, J.D., Timmis, R.J., Wilby, R.L. and Orr, H.G. (2007) 'Long-term variations in orographic rainfall: analysis and implications for upland catchments', *Hydrological Sciences Journal*, 52(2), pp. 276-291.

Mann, H.B. and Whitney, D.R. (1947) 'On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other', pp. 50-60.

Manning, C., Widmann, M., Bevacqua, E., Van Loon, A., F., 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), pp. 1255-1271.

Mariotti, A., Zeng, N. and Lau, K.M. (2002) 'Euro-Mediterranean rainfall and ENSO—a seasonally varying relationship', *Geophysical Research Letters*, 29(12), pp. 59-1-59-4.

Marsh, T., Cole, G. and Wilby, R. (2007) 'Major droughts in England and Wales, 1800–2006', *Weather*, 62(4), pp. 87-93.

Marsh, T.J. (1995) 'The 1995 drought - a water resources review in the context of the recent hydrological instability', *LTA*, 155(47), p. 149.

Marsh, T.J. (2004) 'The UK drought of 2003: A hydrological review', *Weather*, 59(8), pp. 224-230.

Mason, I. (1982) 'A model for assessment of weather forecasts', *Australian Meteorological Magazine*, 30(4), pp. 291-303.

Mavromatis, T. (2007) 'Drought index evaluation for assessing future wheat production in Greece', *International Journal of Climatology*, 27(7), pp. 911-924.

McCallum, E. (1990) 'The Burns' Day Storm, 25 January 1990', *Weather*, 45(5), pp. 166-173.

McKee, T.B., Doesken, N.J. and Kleist, J. (1993) 'The relationship of drought frequency and duration to time scales', *Proceedings of the 8th Conference on Applied Climatology*. American Meteorological Society Boston, MA. Available at: http://clima1.cptec.inpe.br/~rclima1/pdf/paper_spi.pdf.

McKee, T.B., Doesken, N.J. and Kleist, J. (1995) 'Drought monitoring with multiple time scales', *Ninth Conference on Applied Climatology. American Meteorological Society, Boston*. Available at: http://ccc.atmos.colostate.edu/droughtmonitoring.pdf.

McKendry, I.G., Stahl, K. and Moore, R.D. (2006) 'Synoptic sea-level pressure patterns generated by a general circulation model: comparison with types derived from NCEP/NCAR re-analysis and implications for downscaling', *International Journal of Climatology*, 26(12), pp. 1727-1736.

McKendry, I.G., Steyn, D.G. and McBean, G. (1995) 'Validation of synoptic circulation patterns simulated by the Canadian climate centre general circulation model for western north America: Research note', *Atmosphere-Ocean*, 33(4), pp. 809-825.

Mehta, V.M., Suarez, M.J., Manganello, J.V. and Delworth, T.L. (2000) 'Oceanic influence on the North Atlantic Oscillation and associated Northern Hemisphere climate variations: 1959-1993', *Geophysical Research Letters*, 27(1), pp. 121-124.

Meng, L., Ford, T. and Guo, Y. (2017) 'Logistic regression analysis of drought persistence in East China', *International Journal of Climatology*, 37(3), pp. 1444-1455.

Met Office (2018) *Was summer 2018 the hottest on record? [Online]*. Available at: https://www.metoffice.gov.uk/news/releases/2018/end-of-summer-stats (Accessed: 18/09/2018).

Michaelides, S.C., Liassidou, F. and Schizas, C.N. (2007) 'Synoptic Classification and Establishment of Analogues with Artificial Neural Networks', *Pure and Applied Geophysics*, 164(6), pp. 1347-1364.

Michelangeli, P.-A., Vautard, R. and Legras, B. (1995) 'Weather Regimes: Recurrence and Quasi Stationarity', *Journal of the Atmospheric Sciences*, 52(8), pp. 1237-1256.

Miralles, D.G., Teuling, A.J., van Heerwaarden, C.C. and Vilà-Guerau de Arellano, J. (2014) 'Mega-heatwave temperatures due to combined soil desiccation and atmospheric heat accumulation', *Nature Geoscience*, 7, p. 345.

Mishra, A.K. and Desai, V.R. (2005) 'Drought forecasting using stochastic models', *Stochastic Environmental Research and Risk Assessment*, 19(5), pp. 326-339.

Mishra, A.K. and Desai, V.R. (2006) 'Drought forecasting using feed-forward recursive neural network', *Ecological Modelling*, 198(1–2), pp. 127-138.

Mishra, A.K., Desai, V.R. and Singh, V.P. (2007) 'Drought Forecasting Using a Hybrid Stochastic and Neural Network Model', *Journal of Hydrologic Engineering*, 12(6), pp. 626-638.

Mishra, A.K., Ines, A.V.M., Das, N.N., Prakash Khedun, C., Singh, V.P., Sivakumar, B. and Hansen, J.W. (2015a) 'Anatomy of a local-scale drought: Application of assimilated remote sensing products, crop model, and statistical methods to an agricultural drought study', *Journal of Hydrology*, 526(0), pp. 15-29.

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

Mishra, A.K., Sivakumar, B. and Singh, V.P. (2015b) 'Drought processes, modeling, and mitigation', *Journal of Hydrology*, 526(0), pp. 1-2.

Mo, K.C. and Lyon, B. (2015) 'Global Meteorological Drought Prediction Using the North American Multi-Model Ensemble', *Journal of Hydrometeorology*, 16(3), pp. 1409-1424.

Modarres, R. (2007) 'Streamflow drought time series forecasting', *Stochastic Environmental Research and Risk Assessment*, 21(3), pp. 223-233.

Molteni, F., Stockdale, T., Balmaseda, M.A., Balsamo, G., Buizza, R., Ferranti, L., Magnusson, L., Mogensen, K., Palmer, T.N. and Vitart, F. (2011) 'The new ECMWF seasonal forecast system (System 4)' *Technical Memorandum* [Technical Memorandum]. November. ECMWF, p. 49.

Montaseri, M., Amirataee, B. and Yasi, M. (2018) 'Long-term probability of drought characteristics based on Monte Carlo simulation approach', *International Journal of Climatology*.

Monteith, J.L. (1965) 'Evaporation and environment', *Symp. Soc. Exp. Biol.*, 19, pp. 205-234.

Moon, H., Gudmundsson, L. and Seneviratne, S.I. (2018) 'Drought Persistence Errors in Global Climate Models', *Journal of Geophysical Research: Atmospheres*, 123(7), pp. 3483-3496.

Morid, S., Smakhtin, V. and Bagherzadeh, K. (2007) 'Drought forecasting using artificial neural networks and time series of drought indices', *International Journal of Climatology*, 27(15), pp. 2103-2111.

Moron, V., Gouirand, I. and Taylor, M. (2016) 'Weather types across the Caribbean basin and their relationship with rainfall and sea surface temperature', *Climate Dynamics*, 47(1), pp. 601-621.

Moron, V. and Plaut, G. (2003) 'The impact of El Niño–southern oscillation upon weather regimes over Europe and the North Atlantic during boreal winter', *International Journal of Climatology*, 23(4), pp. 363-379.

Moyé, L.A., Kapadia, A.S., Cech, I.M. and Hardy, R.J. (1988) 'The theory of runs with applications to drought prediction', *Journal of Hydrology*, 103(1–2), pp. 127-137.

Mueller, B. and Seneviratne, S.I. (2012) 'Hot days induced by precipitation deficits at the global scale', *Proceedings of the National Academy of Sciences*, 109(31), pp. 12398-12403.

Müller, G.V., Compagnucci, R., Nuñez, M.N. and Salles, A. (2003) 'Surface circulation associated with frost in the wet Pampas', *International Journal of Climatology*, 23(8), pp. 943-961.

Munger, T.T. (1916) 'Graphic method of representing and comparing drought intensities', *Monthly Weather Review*, 44(11), pp. 642-643.

Murphy, A., H. (1973) 'A New Vector Partition of the Probability Score', *Journal of Applied Meteorology*, 12(4), pp. 595-600.

Murphy, A., H. (1971) 'A Note on the Ranked Probability Score', *Journal of Applied Meteorology*, 10(1), pp. 155-156.

Myronidis, D., Ioannou, K., Fotakis, D. and Dörflinger, G. (2018) 'Streamflow and Hydrological Drought Trend Analysis and Forecasting in Cyprus', *Water Resources Management*, 32(5), pp. 1759-1776.

Namias, J. (1955) 'Some Meteorological Aspects of Drought', *Monthly Weather Review*, 83(9), pp. 199-205.

National Drought Mitigation Center (NDMC) (2015) *Predicting Drought [Online]*. Available at: http://drought.unl.edu/DroughtBasics/PredictingDrought.aspx (Accessed: 15 July 2015).

National Oceanic and Atmospheric Administration (2015) *What are El Niño and La Niña ? [Online]*. Available at: http://oceanservice.noaa.gov/facts/ninonina.html.

Neal, R., Dankers, R., Saulter, A., Lane, A., Millard, J., Robbins, G. and Price, D. (2018) 'Use of probabilistic medium- to long-range weather-pattern forecasts for identifying periods with an increased likelihood of coastal flooding around the UK', *Meteorological Applications*, 25(4), pp. 534-547.

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), pp. 389-400.

Nelder, J.A. and Wedderburn, R.W.M. (1972) 'Generalized Linear Models', *Journal of the Royal Statistical Society. Series A (General)*, 135(3), pp. 370-384.

Nigro, M., A. , Cassano, J., J. and Seefeldt, M., W. (2011) 'A Weather-Pattern-Based Approach to Evaluate the Antarctic Mesoscale Prediction System (AMPS) Forecasts: Comparison to Automatic Weather Station Observations', *Weather and Forecasting*, 26(2), pp. 184-198.

Norris, J.R. (1997) *Markov Chains*. Cambridge: Cambridge University Press.

O'Hare, G. and Sweeney, J. (1993) 'Lamb's Circulation Types and British Weather: An Evaluation', *Geography*, 78(1), pp. 43-60.

Obled, C., Bontron, G. and Garçon, R. (2002) 'Quantitative precipitation forecasts: a statistical adaptation of model outputs through an analogues sorting approach', *Atmospheric Research*, 63(3), pp. 303-324.

Oladipo, E.O. (1985) 'A comparative performance analysis of three meteorological drought indices (Nebraska)', *Journal of Climatology*, 5(6), pp. 655-664.

Orth, R. and Seneviratne, S.I. (2013) 'Propagation of soil moisture memory to streamflow and evapotranspiration in Europe', *Hydrol. Earth Syst. Sci.*, 17(10), pp. 3895-3911.

Osborn, T.J. and Hulme, M. (2002) 'Evidence for trends in heavy rainfall events over the UK', *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 360(1796), pp. 1313-1325.

Ostermeier, G.M. and Wallace, J.M. (2003) 'Trends in the North Atlantic Oscillation–Northern Hemisphere Annular Mode during the Twentieth Century', *Journal of Climate*, 16(2), pp. 336-341.

Özger, M., Mishra, A.K. and Singh, V.P. (2012) 'Long Lead Time Drought Forecasting Using a Wavelet and Fuzzy Logic Combination Model: A Case Study in Texas', *Journal of Hydrometeorology*, 13(1), pp. 284-297.

Palmer, W.C. (1965) *Meteorological drought*. US Department of Commerce, Weather Bureau Washington, DC, USA.

Panu, U.S. and Sharma, T.C. (2002) 'Challenges in drought research: some perspectives and future directions', *Hydrological Sciences Journal*, 47(sup1), pp. S19-S30.

Papalexiou, S.M., Koutsoyiannis, D. and Makropoulos, C. (2013) 'How extreme is extreme? An assessment of daily rainfall distribution tails', *Hydrol. Earth Syst. Sci.*, 17(2), pp. 851-862.

Parry, S., Prudhomme, C., Wilby, R.L. and Wood, P.J. (2016) 'Drought termination: Concept and characterisation', *Progress in Physical Geography: Earth and Environment*, 40(6), pp. 743-767.

Pattison, I. and Lane, S.N. (2012) 'The relationship between Lamb weather types and long-term changes in flood frequency, River Eden, UK', *International Journal of Climatology*, 32(13), pp. 1971-1989.

Paulo, A.A., Ferreira, E., Coelho, C. and Pereira, L.S. (2005) 'Drought class transition analysis through Markov and Loglinear models, an approach to early warning', *Agricultural Water Management*, 77(1–3), pp. 59-81.

Peings, Y., Brun, E., Mauvais, V. and Douville, H. (2013) 'How stationary is the relationship between Siberian snow and Arctic Oscillation over the 20th century?', *Geophysical Research Letters*, 40(1), pp. 183-188.

Peng-Xin, W., Xiao-Wen, L., Jian-ya, G. and Conghe, S. (2001) 'Vegetation temperature condition index and its application for drought monitoring', *Geoscience and Remote Sensing Symposium, 2001. IGARSS '01. IEEE 2001 International*. 2001. Available at: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=976083&tag=1.

Perlwitz, J. and Graf, H.F. (2001) 'The variability of the horizontal circulation in the troposphere and stratosphere – a comparison', *Theoretical and Applied Climatology*, 69(3), pp. 149-161.

Philipp, A., Bartholy, J., Beck, C., Erpicum, M., Esteban, P., Fettweis, X., Huth, R., James, P., Jourdain, S., Kreienkamp, F., Krennert, T., Lykoudis, S., Michalides, S.C., Pianko-Kluczynska, K., Post, P., Álvarez, D.R., Schiemann, R., Spekat, A. and Tymvios, F.S. (2010) 'Cost733cat – A database of weather and circulation type classifications', *Physics and Chemistry of the Earth, Parts A/B/C*, 35(9–12), pp. 360-373.

Philipp, A., Della-Marta, P.M., Jacobeit, J., Fereday, D.R., Jones, P.D., Moberg, A. and Wanner, H. (2007) 'Long-Term Variability of Daily North Atlantic–European Pressure Patterns since 1850 Classified by Simulated Annealing Clustering', *Journal of Climate*, 20(16), pp. 4065-4095.

Phillips, I.D. and McGregor, 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), pp. 359-372.

Piechota, T.C. and Dracup, J.A. (1996) 'Drought and Regional Hydrologic Variation in the United States: Associations with the El Niño-Southern Oscillation', *Water Resources Research*, 32(5), pp. 1359-1373.

Pongrácz, R., Bogardi, I. and Duckstein, L. (2003) 'Climatic forcing of droughts: a Central European example', *Hydrological Sciences Journal*, 48(1), pp. 39-50.

Pozo-Vázquez, D., Gámiz-Fortis, S.R., Tovar-Pescador, J., Esteban-Parra, M.J. and Castro-DÍez, Y. (2005) 'North Atlantic Winter SLP Anomalies Based on the Autumn ENSO State', *Journal of Climate*, 18(1), pp. 97-103.

Prein, A.F., Holland, G.J., Rasmussen, R.M., Clark, M.P. and Tye, M.R. (2016) 'Running dry: The U.S. Southwest's drift into a drier climate state', *Geophysical Research Letters*, 43(3), pp. 1272-1279.

Prior, J. and Kendon, M. (2011) 'The UK winter of 2009/2010 compared with severe winters of the last 100 years', *Weather*, 66(1), pp. 4-10.

Prudhomme, C., Giuntoli, I., Robinson, E.L., Clark, D.B., Arnell, N.W., Dankers, R., Fekete, B.M., Franssen, W., Gerten, D., Gosling, S.N., Hagemann, S., Hannah, D.M., Kim, H., Masaki, Y., Satoh, Y., Stacke, T., Wada, Y. and Wisser, D. (2014) 'Hydrological droughts in the 21st century, hotspots and uncertainties from a global multimodel ensemble experiment', *Proceedings of the National Academy of Sciences*, 111(9), pp. 3262-3267.

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), pp. 1115-1118.

Qian, B., Corte-Real, J. and Xu, H. (2000) 'Is the North Atlantic Oscillation the most important atmospheric pattern for precipitation in Europe?', *Journal of Geophysical Research: Atmospheres*, 105(D9), pp. 11901-11910.

Quesada, B., Vautard, R., Yiou, P., Hirschi, M. and Seneviratne, S.I. (2012) 'Asymmetric European summer heat predictability from wet and dry southern winters and springs', *Nature Climate Change*, 2, p. 736.

Rahiz, M. and New, M. (2012) 'Spatial coherence of meteorological droughts in the UK since 1914', *Area*, 44(4), pp. 400-410.

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

Rahiz, M. and New, M. (2014) 'Does a rainfall-based drought index simulate hydrological droughts?', *International Journal of Climatology*, 34(9), pp. 2853-2871.

Ramos, A.M., Pires, A.C., Sousa, P.M. and Trigo, R.M. (2013) 'The use of circulation weather types to predict upwelling activity along the western Iberian Peninsula coast', *Continental Shelf Research*, 69, pp. 38-51.

Rao, A.R. and Padmanabhan, G. (1984) 'Analysis and modeling of Palmer's drought index series', *Journal of Hydrology*, 68(1–4), pp. 211-229.

Reusch, D.B., Alley, R.B. and Hewitson, B.C. (2007) 'North Atlantic climate variability from a self-organizing map perspective', *Journal of Geophysical Research: Atmospheres*, 112(D2).

Rezaeianzadeh, M., Stein, A. and Cox, J.P. (2016) 'Drought Forecasting using Markov Chain Model and Artificial Neural Networks', *Water Resources Management*, 30(7), pp. 2245-2259.

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

Richardson, D., Kilsby, C.G., Fowler, H.J. and Bárdossy, A. (2018b) 'Weekly to multi-month persistence in sets of daily weather patterns over Europe and the North Atlantic Ocean', *International Journal of Climatology*.

Riddle, E., Butler, A., Furtado, J., Cohen, J. and Kumar, A. (2013) 'CFSv2 ensemble prediction of the wintertime Arctic Oscillation', *Climate Dynamics*, 41(3-4), pp. 1099-1116.

Rigby, R.A. and Stasinopoulos, D.M. (2005) 'Generalized additive models for location, scale and shape', *Journal of the Royal Statistical Society Series C-Applied Statistics*, 54, pp. 507-544.

Robertson, D.E., Pokhrel, P. and Wang, Q.J. (2013) 'Improving statistical forecasts of seasonal streamflows using hydrological model output', *Hydrology and Earth System Sciences*, 17(2), pp. 579-593.

Rodda, J.C. and Marsh, T.J. (2011) *The 1975-76 Drought - a contemporary and retrospective review*. [Online]. Available at: http://www.ceh.ac.uk/data/nrfa/nhmp/other_reports/CEH_1975-76_Drought_Report_Rodda_and_Marsh.pdf.

Rodó, X., Baert, E. and Comin, F.A. (1997) 'Variations in seasonal rainfall in Southern Europe during the present century: Relationships with the North Atlantic Oscillation and the El Niño-Southern Oscillation', *Climate Dynamics*, 13(4), pp. 275-284.

Rodwell, M.J., Rowell, D.P. and Folland, C.K. (1999) 'Oceanic forcing of the wintertime North Atlantic Oscillation and European climate', *Nature*, 398(6725), pp. 320-323.

Ropelewski, C.F. and Halpert, M.S. (1987) 'Global and Regional Scale Precipitation Patterns Associated with the El Niño/Southern Oscillation', *Monthly Weather Review*, 115(8), pp. 1606-1626.

Roudier, P., Andersson, J.C.M., Donnelly, C., Feyen, L., Greuell, W. and Ludwig, F. (2016) 'Projections of future floods and hydrological droughts in Europe under a +2°C global warming', *Climatic Change*, 135(2), pp. 341-355.

Roulston, M., S. and Smith, L., A. (2002) 'Evaluating Probabilistic Forecasts Using Information Theory', *Monthly Weather Review*, 130(6), pp. 1653-1660.

Rueda, A., Camus, P., Tomás, A., Vitousek, S. and Méndez, F.J. (2016) 'A multivariate extreme wave and storm surge climate emulator based on weather patterns', *Ocean Modelling*, 104, pp. 242-251.

Ruosteenoja, K., Markkanen, T., Venäläinen, A., Räisänen, P. and Peltola, H. (2018) 'Seasonal soil moisture and drought occurrence in Europe in CMIP5 projections for the 21st century', *Climate Dynamics*, 50(3), pp. 1177-1192.

Ruprich-Robert, Y., Rym, M., Frederic, C., Stephen, Y., Tom, D. and Gokhan, D. (2017) 'Assessing the Climate Impacts of the Observed Atlantic Multidecadal Variability Using the GFDL CM2.1 and NCAR CESM1 Global Coupled Models', *Journal of Climate*, 30(8), pp. 2785-2810.

Saghafian, B. and Hamzekhani, F.G. (2015) 'Hydrological drought early warning based on rainfall threshold', *Natural Hazards*, 79(2), pp. 815-832.

Saha, S., Shrinivas, M., Xingren, W., Jiande, W., Sudhir, N., Patrick, T., David, B., Yu-Tai, H., Hui-ya, C., Mark, I., Michael, E., Jesse, M., Rongqian, Y., Malaquías Peña, M., Huug van den, D., Qin, Z., Wanqiu, W., Mingyue, C. and Emily, B. (2014) 'The NCEP Climate Forecast System Version 2', *Journal of Climate*, 27(6), pp. 2185-2208.

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), pp. 421-426.

Sanderson, M.G., Wiltshire, A.J. and Betts, R.A. (2012) 'Projected changes in water availability in the United Kingdom', *Water Resources Research*, 48(8).

Santos, J.A., Corte-Real, J. and Leite, S.M. (2005) 'Weather regimes and their connection to the winter rainfall in Portugal', *International Journal of Climatology*, 25(1), pp. 33-50.

Saunders, I.R. and Byrne, J.M. (1996) 'Generating regional precipitation from observed and GCM synoptic-scale pressure fields, southern Alberta, Canada', *Climate Research*, 06(3), pp. 237-249.

Saunders, I.R. and Byrne, J.M. (1999) 'Using synoptic surface and geopotential height fields for generating grid-scale precipitation', *International Journal of Climatology*, 19(11), pp. 1165-1176.

Scaife, A.A., Arribas, A., Blockley, E., Brookshaw, A., Clark, R.T., Dunstone, N., Eade, R., Fereday, D., Folland, C.K., Gordon, M., Hermanson, L., Knight, J.R., Lea, D.J., MacLachlan, C., Maidens, A., Martin, M., Peterson, A.K., Smith, D., Vellinga, M., Wallace, E., Waters, J. and Williams, A. (2014) 'Skillful long-range prediction of

European and North American winters', *Geophysical Research Letters*, 41(7), pp. 2514-2519.

Scaife, A.A., Folland, C.K., Alexander, L.V., Moberg, A. and Knight, J.R. (2008) 'European Climate Extremes and the North Atlantic Oscillation', *Journal of Climate*, 21(1), pp. 72-83.

Schepen, A. and Wang, Q.J. (2015) 'Model averaging methods to merge operational statistical and dynamic seasonal streamflow forecasts in Australia', *Water Resources Research*, 51(3), pp. 1797-1812.

Schepen, A., Wang, Q.J. and David, E.R. (2014) 'Seasonal Forecasts of Australian Rainfall through Calibration and Bridging of Coupled GCM Outputs', *Monthly Weather Review*, 142(5), pp. 1758-1770.

Schiemann, R. and Frei, C. (2010) 'How to quantify the resolution of surface climate by circulation types: An example for Alpine precipitation', *Physics and Chemistry of the Earth, Parts A/B/C*, 35(9), pp. 403-410.

Schoof, J.T. and Pryor, S.C. (2006) 'An evaluation of two GCMs: simulation of North American teleconnection indices and synoptic phenomena', *International Journal of Climatology*, 26(2), pp. 267-282.

Schwalm, C.R., Anderegg, W.R.L., Michalak, A.M., Fisher, J.B., Biondi, F., Koch, G., Litvak, M., Ogle, K., Shaw, J.D., Wolf, A., Huntzinger, D.N., Schaefer, K., Cook, R., Wei, Y., Fang, Y., Hayes, D., Huang, M., Jain, A. and Tian, H. (2017) 'Global patterns of drought recovery', *Nature*, 548(7666), pp. 202-205.

Sefton, C., Turner, S., Lewis, M. and Clemas, S. (2018) *Hydrological summary for the United Kingdom: April 2018*. Wallingford, UK. [Online]. Available at: http://nora.nerc.ac.uk/id/eprint/520084/1/HS_201804.pdf.

Seibert, M., Merz, B. and Apel, H. (2017) 'Seasonal forecasting of hydrological drought in the Limpopo Basin: a comparison of statistical methods', *Hydrol. Earth Syst. Sci.*, 21(3), pp. 1611-1629.

Şen, Z. (1990) 'Critical drought analysis by second-order Markov chain', *Journal of Hydrology*, 120(1), pp. 183-202.

Shaffrey, L.C. (2014) *IMPETUS: Improving Predictions of Drought for User Decision-Making. [Online]*. *[Online]*. Available at: http://gtr.rcuk.ac.uk/project/B543AD83-6491-42F8-BA47-FF205AC2864C (Accessed: 7 November 2014).

Shaman, J. and Tziperman, E. (2010) 'An Atmospheric Teleconnection Linking ENSO and Southwestern European Precipitation', *Journal of Climate*, 24(1), pp. 124-139.

Sheffield, J. and Wood, E.F. (2008) 'Projected changes in drought occurrence under future global warming from multi-model, multi-scenario, IPCC AR4 simulations', *Climate Dynamics*, 31(1), pp. 79-105.

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

Shukla, J. and Mintz, Y. (1982) 'Influence of Land-Surface Evapotranspiration on the Earth's Climate', *Science*, 215(4539), pp. 1498-1501.

Shukla, S., Sheffield, J., Wood, E.F. and Lettenmaier, D.P. (2013) 'On the sources of global land surface hydrologic predictability', *Hydrol. Earth Syst. Sci.*, 17(7), pp. 2781-2796.

Sienz, F., Bothe, O. and Fraedrich, K. (2012) 'Monitoring and quantifying future climate projections of dryness and wetness extremes: SPI bias', *Hydrology and Earth System Sciences*, 16(7), pp. 2143-2157.

Smith, D.M., Scaife, A.A. and Kirtman, B.P. (2012) 'What is the current state of scientific knowledge with regard to seasonal and decadal forecasting?', *Environmental Research Letters*, 7(1).

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), pp. 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(0), pp. 231-252.

Staël von Holstein, C.-A., S. (1970) 'A Family of Strictly Proper Scoring Rules Which Are Sensitive to Distance', *Journal of Applied Meteorology*, 9(3), pp. 360-364.

Stahl, K., Moore, R.D. and Mckendry, I.G. (2006) 'The role of synoptic-scale circulation in the linkage between large-scale ocean–atmosphere indices and winter surface climate in British Columbia, Canada', *International Journal of Climatology*, 26(4), pp. 541-560.

Stefanicki, G., Talkner, P. and Weber, R.O. (1998) 'Frequency Changes of Weather Types in the Alpine Region since 1945', *Theoretical and Applied Climatology*, 60(1), pp. 47-61.

Steinemann, A. (2003) 'Drought Indicators and Triggers: a Stochastic Approach to Evaluation', *JAWRA Journal of the American Water Resources Association*, 39(5), pp. 1217-1233.

Stryhal, J. and Huth, R. (2017) 'Classifications of Winter Euro-Atlantic Circulation Patterns: An Intercomparison of Five Atmospheric Reanalyses', *Journal of Climate*, 30(19), pp. 7847-7861.

Stryhal, J. and Huth, R. (2018) 'Trends in winter circulation over the British Isles and central Europe in twenty-first century projections by 25 CMIP5 GCMs', *Climate Dynamics*.

Sun, Y., Solomon, S., Dai, A. and Portmann, R.W. (2006) 'How Often Does It Rain?', *Journal of Climate*, 19(6), pp. 916-934.

Sutton, R.T. and Dong, B. (2012) 'Atlantic Ocean influence on a shift in European climate in the 1990s', *Nature Geosci*, 5(11), pp. 788-792.

Sutton, R.T. and Hodson, D.L.R. (2005) 'Atlantic Ocean Forcing of North American and European Summer Climate', *Science*, 309(5731), pp. 115-118.

Sutton, R.T., McCarthy, G.D., Robson, J., Sinha, B., Archibald, A.T. and Gray, L.J. (2018) 'Atlantic Multidecadal Variability and the U.K. ACSIS Program', *Bulletin of the American Meteorological Society*, 99(2), pp. 415-425.

Svensson, C. (2014) 'Seasonal river flow forecasts for the United Kingdom using persistence and historical analogues', *Hydrological Sciences Journal*, pp. 1-17.

Svensson, C., Brookshaw, A., Scaife, A.A., Bell, V.A., Mackay, J.D., Jackson, C.R., Hannaford, J., Davies, H.N., Arribas, A. and Stanley, S. (2015) 'Long-range forecasts of UK winter hydrology', *Environmental Research Letters*, 10(6), p. 064006.

Svensson, C. and Prudhomme, C. (2005) 'Prediction of British summer river flows using winter predictors', *Theoretical and Applied Climatology*, 82(1-2), pp. 1-15.

Sweeney, J.C. and O'Hare, G.P. (1992) 'Geographical Variations in Precipitation Yields and Circulation Types in Britain and Ireland', *Transactions of the Institute of British Geographers*, 17(4), pp. 448-463.

Terray, L. and Cassou, C. (2002) 'Tropical Atlantic Sea Surface Temperature Forcing of Quasi-Decadal Climate Variability over the North Atlantic–European Region', *Journal of Climate*, 15(22), pp. 3170-3187.

The Guardian (2018a) *Farmers in drought summit amid fears of food supply crisis [Online]*. Available at: https://www.theguardian.com/environment/2018/jul/28/farmers-drought-summit-food-supply-fears-crisis-heatwave (Accessed: 18/09/2018).

The Guardian (2018b) *Saddleworth Moor fire declared major incident as residents evacuated [Online]*. Available at: https://www.theguardian.com/uk-news/2018/jun/27/manchester-moorland-fire-declared-a-major-incident-by-police (Accessed: 20/09/2018).

The Guardian (2018c) *UK water firm asks for more water from Lake District to fight drought [Online]*. Available at: https://www.theguardian.com/uk-news/2018/jul/20/uk-water-firm-asks-for-more-water-from-lake-district-to-fight-drought (Accessed: 18/09/2018).

The New York Times (2018) *The Mysterious Landscapes of Heat-Scorched Britain [Online]*. Available at: https://www.nytimes.com/2018/08/15/opinion/britain-drought-cropmarks-summer.html (Accessed: 18/09/2018).

Thompson, D.W.J., Baldwin, M.P. and Wallace, J.M. (2002) 'Stratospheric Connection to Northern Hemisphere Wintertime Weather: Implications for Prediction', *Journal of Climate*, 15(12), pp. 1421-1428.

Thompson, D.W.J. and Wallace, J.M. (1998) 'The Arctic oscillation signature in the wintertime geopotential height and temperature fields', *Geophysical Research Letters*, 25(9), pp. 1297-1300.

Thompson, D.W.J. and Wallace, J.M. (2001) 'Regional Climate Impacts of the Northern Hemisphere Annular Mode', *Science*, 293(5527), pp. 85-89.

Thornthwaite, C.W. (1948) 'An Approach toward a Rational Classification of Climate', *Geographical Review*, 38(1), pp. 55-94.

Tosunoglu, F., Can, I. and Kahya, E. (2018) 'Evaluation of spatial and temporal relationships between large-scale atmospheric oscillations and meteorological drought indexes in Turkey', *International Journal of Climatology*, 38(12), pp. 4579-4596.

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(0), pp. 196-207.

Trenberth, K.E., Dai, A., van der Schrier, G., Jones, P.D., Barichivich, J., Briffa, K.R. and Sheffield, J. (2013) 'Global warming and changes in drought', *Nature Climate Change*, 4, p. 17.

Trigo, R.M., Pozo-Vázquez, D., Osborn, T.J., Castro-Díez, Y., Gámiz-Fortis, S. and Esteban-Parra, M.J. (2004) 'North Atlantic oscillation influence on precipitation, river flow and water resources in the Iberian Peninsula', *International Journal of Climatology*, 24(8), pp. 925-944.

Trnka, M., Kyselý, J., Možný, M. and Dubrovský, M. (2009) 'Changes in Central-European soil-moisture availability and circulation patterns in 1881–2005', *International Journal of Climatology*, 29(5), pp. 655-672.

Tsakiris, G., Pangalou, D. and Vangelis, H. (2007) 'Regional Drought Assessment Based on the Reconnaissance Drought Index (RDI)', *Water Resources Management*, 21(5), pp. 821-833.

United Utilities (2018) *Hosepipe ban in North West called off [Online]*. Available at: https://www.unitedutilities.com/corporate/newsroom/latest-news/hosepipe-ban-in-north-west-called-off/ (Accessed: 18/09/2018).

van den Besselaar, E.J.M., Klein Tank, A.M.G. and van der Schrier, G. (2009) 'Influence of circulation types on temperature extremes in Europe', *Theoretical and Applied Climatology*, 99(3), p. 431.

van den Dool, H.M. and Livezey, R.E. (1984) 'Geographical Distribution and Seasonality of Month-to-Month Correlation of Monthly Mean 700 mb Heights', *Monthly Weather Review*, 112(3), pp. 610-615.

van Loon, H. and Rogers, J.C. (1978) 'The Seesaw in Winter Temperatures between Greenland and Northern Europe. Part I: General Description', *Monthly Weather Review*, 106(3), pp. 296-310.

van Oldenborgh, G.J., Stephenson, D.B., Sterl, A., Vautard, R., Yiou, P., Drijfhout, S.S., von Storch, H. and van den Dool, H. (2015) 'Drivers of the 2013/14 winter floods in the UK', *Nature Climate Change*, 5, p. 490.

Vautard, R., Yiou, P., D'Andrea, F., de Noblet, N., Viovy, N., Cassou, C., Polcher, J., Ciais, P., Kageyama, M. and Fan, Y. (2007) 'Summertime European heat and drought waves induced by wintertime Mediterranean rainfall deficit', *Geophysical Research Letters*, 34(7).

Vazifehkhah, S. and Kahya, E. (2018) 'Hydrological drought associations with extreme phases of the North Atlantic and Arctic Oscillations over Turkey and northern Iran', *International Journal of Climatology*, 38(12), pp. 4459-4475.

Vicente-Serrano, S., López-Moreno, J., Beguería, S., Lorenzo-Lacruz, J., Azorin-Molina, C. and Morán-Tejeda, E. (2012) 'Accurate Computation of a Streamflow Drought Index', *Journal of Hydrologic Engineering*, 17(2), pp. 318-332.

Vicente-Serrano Sergio, M., López-Moreno Juan, I., Gimeno, L., Nieto, R., Morán-Tejeda, E., Lorenzo-Lacruz, J., Beguería, S. and Azorin-Molina, C. (2011) 'A multiscalar global evaluation of the impact of ENSO on droughts', *Journal of Geophysical Research: Atmospheres*, 116(D20).

Vicente-Serrano, S.M., Beguería, S. and López-Moreno, J.I. (2010) 'A multiscalar drought index sensitive to global warming: The standardized precipitation evapotranspiration index', *Journal of Climate*, 23(7), pp. 1696-1718.

Vidal, J.-P. and Wade, S. (2009) 'A multimodel assessment of future climatological droughts in the United Kingdom', *International Journal of Climatology*, 29(14), pp. 2056-2071.

Vigaud, N., Robertson, A.W. and Tippett, M.K. (2018) 'Predictability of Recurrent Weather Regimes over North America during Winter from Submonthly Reforecasts', *Monthly Weather Review*, 146(8), pp. 2559-2577.

Vitart, F. (2014) 'Evolution of ECMWF sub-seasonal forecast skill scores', *Quarterly Journal of the Royal Meteorological Society*, 140(683), pp. 1889-1899.

Vitart, F., Ardilouze, C., Bonet, A., Brookshaw, A., Chen, M., Codorean, C., Déqué, M., Ferranti, L., Fucile, E., Fuentes, M., Hendon, H., Hodgson, J., Kang, H.-S., Kumar, A.,

Lin, H., Liu, G., Liu, X., Malguzzi, P., Mallas, I., Manoussakis, M., Mastrangelo, D., MacLachlan, C., McLean, P., Minami, A., Mladek, R., Nakazawa, T., Najm, S., Nie, Y., Rixen, M., Robertson, A.W., Ruti, P., Sun, C., Takaya, Y., Tolstykh, M., Venuti, F., Waliser, D., Woolnough, S., Wu, T., Won, D.-J., Xiao, H., Zaripov, R. and Zhang, L. (2017) 'The Subseasonal to Seasonal (S2S) Prediction Project Database', *Bulletin of the American Meteorological Society*, 98(1), pp. 163-173.

Vitart, F., Buizza, R., Alonso Balmaseda, M., Balsamo, G., Bidlot, J.-R., Bonet, A., Fuentes, M., Hofstadler, A., Molteni, F. and Palmer, T.N. (2008) 'The new VarEPS-monthly forecasting system: A first step towards seamless prediction', *Quarterly Journal of the Royal Meteorological Society*, 134(636), pp. 1789-1799.

Vrac, M., Hayhoe, K. and Stein, M. (2007) 'Identification and intermodel comparison of seasonal circulation patterns over North America', *International Journal of Climatology*, 27(5), pp. 603-620.

Vuillaume, J.-F. and Herath, S. (2017) 'Improving global rainfall forecasting with a weather type approach in Japan', *Hydrological Sciences Journal*, 62(2), pp. 167-181.

Walker, G.T. and Bliss, E.W. (1932) 'World Weather V', *Memoirs of the Royal Meteorological Society*, 4(36), pp. 53-84.

Wallace, J.M. and Gutzler, D.S. (1981) 'Teleconnections in the Geopotential Height Field during the Northern Hemisphere Winter', *Monthly Weather Review*, 109(4), pp. 784-812.

Wang, E., Zhang, Y., Luo, J., Chiew, F.H.S. and Wang, Q.J. (2011) 'Monthly and seasonal streamflow forecasts using rainfall-runoff modeling and historical weather data', *Water Resources Research*, 47(5).

Wang, L., Ting, M. and Kushner, P.J. (2017) 'A robust empirical seasonal prediction of winter NAO and surface climate', *Scientific Reports*, 7(1), p. 279.

Wang, Q.J., Andrew, S. and David, E.R. (2012) 'Merging Seasonal Rainfall Forecasts from Multiple Statistical Models through Bayesian Model Averaging', *Journal of Climate*, 25(16), pp. 5524-5537.

Wang, Q.J., Robertson, D.E. and Chiew, F.H.S. (2009) 'A Bayesian joint probability modeling approach for seasonal forecasting of streamflows at multiple sites', *Water Resources Research*, 45(5).

Water UK (2016) *Water resources long-term planning framework*. [Online]. Available at: https://www.water.org.uk/news-water-uk/latest-news/research-shows-more-action-needed-protect-against-growing-drought-risk (Accessed: 19/09/2018).

Wedgbrow, C.S., Wilby, R. and Fox, H.R. (2005) 'Experimental seasonal forecasts of low summer flows in the River Thames, UK, using Expert Systems', *Climate Research*, 28(2), pp. 133-141.

Wedgbrow, C.S., Wilby, R.L., Fox, H.R. and O'Hare, G. (2002) 'Prospects for seasonal forecasting of summer drought and low river flow anomalies in England and Wales', *International Journal of Climatology*, 22(2), pp. 219-236.

Weijs, S., V. and Giesen, N.v.d. (2011) 'Accounting for Observational Uncertainty in Forecast Verification: An Information-Theoretical View on Forecasts, Observations, and Truth', *Monthly Weather Review*, 139(7), pp. 2156-2162.

Weijs, S., V., Nooijen, R.v. and Giesen, N.v.d. (2010) 'Kullback–Leibler Divergence as a Forecast Skill Score with Classic Reliability–Resolution–Uncertainty Decomposition', *Monthly Weather Review*, 138(9), pp. 3387-3399.

Welch, B.L. (1947) 'The Generalization of 'Student's' Problem when Several Different Population Variances are Involved', *Biometrika*, 34(1-2), pp. 28-35.

Wells, N., Goddard, S. and Hayes, M., J. (2004) 'A Self-Calibrating Palmer Drought Severity Index', *Journal of Climate*, 17(12), pp. 2335-2351.

Werner, K., David, B., Martyn, C. and Subhrendu, G. (2004) 'Climate Index Weighting Schemes for NWS ESP-Based Seasonal Volume Forecasts', *Journal of Hydrometeorology*, 5(6), pp. 1076-1090.

Werner, P.C., Gerstengarbe, F.W., Fraedrich, K. and Oesterle, H. (2000) 'Recent climate change in the North Atlantic/European sector', *International Journal of Climatology*, 20(5), pp. 463-471.

Wigley, T.M.L., Lough, J.M. and Jones, P.D. (1984) 'Spatial patterns of precipitation in England and Wales and a revised, homogeneous England and Wales precipitation series', *Journal of Climatology*, 4(1), pp. 1-25.

Wilby, R. (1995) 'Simulation of precipitation by weather pattern and frontal analysis', *Journal of Hydrology*, 173(1), pp. 91-109.

Wilby, R.L. (1993) 'The influence of variable weather patterns on river water quantity and quality regimes', *International Journal of Climatology*, 13(4), pp. 447-459.

Wilby, R.L. (1994) 'Stochastic weather type simulation for regional climate change impact assessment', *Water Resources Research*, 30(12), pp. 3395-3403.

Wilby, R.L. (1997) 'Non-stationarity in daily precipitation series: implications for GCM down-scaling using atmospheric circulation indices', *International Journal of Climatology*, 17(4), pp. 439-454.

Wilby, R.L. (1998) 'Modelling low-frequency rainfall events using airflow indices, weather patterns and frontal frequencies', *Journal of Hydrology*, 212–213, pp. 380-392.

Wilby, R.L. (2001) 'Seasonal Forecasting of River Flows in the British Isles Using North Atlantic Pressure Patterns', *Water and Environment Journal*, 15(1), pp. 56-63.

Wilby, R.L., Greenfield, B. and Glenny, C. (1994) 'A coupled synoptic-hydrological model for climate change impact assessment', *Journal of Hydrology*, 153(1), pp. 265-290.

Wilby, R.L., O'Hare, G. and Barnsley, N. (1997) 'The North Atlantic Oscillation and British Isles climate variability, 1865–1996', *Weather*, 52(9), pp. 266-276.

Wilby, R.L. and Quinn, N.W. (2013) 'Reconstructing multi-decadal variations in fluvial flood risk using atmospheric circulation patterns', *Journal of Hydrology*, 487, pp. 109-121.

Wilby, R.L., Wedgbrow, C.S. and Fox, H.R. (2004) 'Seasonal predictability of the summer hydrometeorology of the River Thames, UK', *Journal of Hydrology*, 295(1–4), pp. 1-16.

Wilhite, D. (2009) 'Drought Monitoring as a Component of Drought Preparedness Planning', in Iglesias, A., Cancelliere, A., Wilhite, D., Garrote, L. and Cubillo, F. (eds.) *Coping with Drought Risk in Agriculture and Water Supply Systems*. Springer Netherlands, pp. 3-19.

Wilhite, D.A. and Glantz, M.H. (1985) 'Understanding: the Drought Phenomenon: The Role of Definitions', *Water International*, 10(3), pp. 111-120.

Wilks, D.S. (1995) 'Chapter 7 Forecast verification', in Wilks, D.S. (ed.) *International Geophysics*. Academic Press, pp. 233-283.

Wilks, D.S. (2011) 'Chapter 8 - Forecast Verification', in Wilks, D.S. (ed.) *International Geophysics*. Academic Press, pp. 301-394.

Wilson, L.L., Lettenmaier, D.P. and Skyllingstad, E. (1992) 'A hierarchical stochastic model of large-scale atmospheric circulation patterns and multiple station daily precipitation', *Journal of Geophysical Research: Atmospheres*, 97(D3), pp. 2791-2809.

World Meteorological Organization (WMO) and Global Water Partnership (GWP) (2017) *Benefits of action and costs of inaction: Drought mitigation and preparedness - a literature review (N. Gerber and A. Mirzabaev).*

World Meteorological Organization (WMO) and Global Water Partnership (GWP) (2016) *Handbook of Drought Indicators and Indices*.

Wu, H., Svoboda, M.D., Hayes, M.J., Wilhite, D.A. and Wen, F. (2007) 'Appropriate application of the Standardized Precipitation Index in arid locations and dry seasons', *International Journal of Climatology*, 27(1), pp. 65-79.

Xu, L., Chen, N., Zhang, X. and Chen, Z. (2018) 'An evaluation of statistical, NMME and hybrid models for drought prediction in China', *Journal of Hydrology*, 566, pp. 235-249.

Yan, H., Moradkhani, H. and Zarekarizi, M. (2017) 'A probabilistic drought forecasting framework: A combined dynamical and statistical approach', *Journal of Hydrology*, 548, pp. 291-304.

Yao, H. and Georgakakos, A. (2001) 'Assessment of Folsom Lake response to historical and potential future climate scenarios: 2. Reservoir management', *Journal of Hydrology*, 249(1), pp. 176-196.

Yevjevich, V. (1967) 'An objective approach to definitions and investigations of continental hydrologic droughts', *Hydrology Papers Colorado State University, Fort Collins*.

Yoon, J.-H., Mo, K. and Wood, E.F. (2012) 'Dynamic-Model-Based Seasonal Prediction of Meteorological Drought over the Contiguous United States', *Journal of Hydrometeorology*, 13(2), pp. 463-482.

Yuan, X. and Wood, E.F. (2013) 'Multimodel seasonal forecasting of global drought onset', *Geophysical Research Letters*, 40(18), pp. 4900-4905.

Zhang, W., Wang, L., Xiang, B., Qi, L. and He, J. (2015) 'Impacts of two types of La Niña on the NAO during boreal winter', *Climate Dynamics*, 44(5-6), pp. 1351-1366.

Zhao, T. and Dai, A. (2017) 'Uncertainties in historical changes and future projections of drought. Part II: model-simulated historical and future drought changes', *Climatic Change*, 144(3), pp. 535-548.

Appendices

Appendix A: Supporting information for Chapter 3

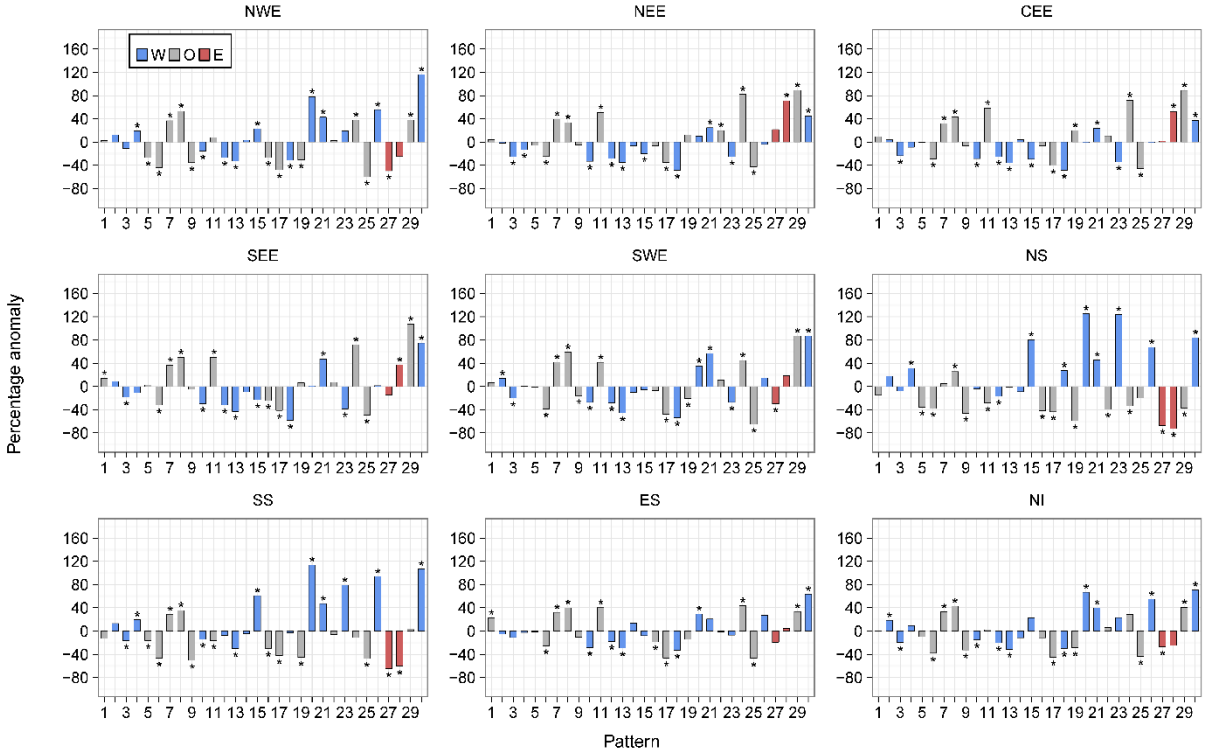

Figure A.1: Annual (i.e. all months) three monthly mean frequency percentage anomalies of each weather pattern in MO-30 during wet periods defined by SPI-3 ≥ 1. Blue and red bars indicate that the weather pattern contains a westerly (W) or easterly (E) component in its Lamb weather type equivalent, respectively. Grey bars represent all other types (O). An asterisk indicates statistical significance at the 95% level.

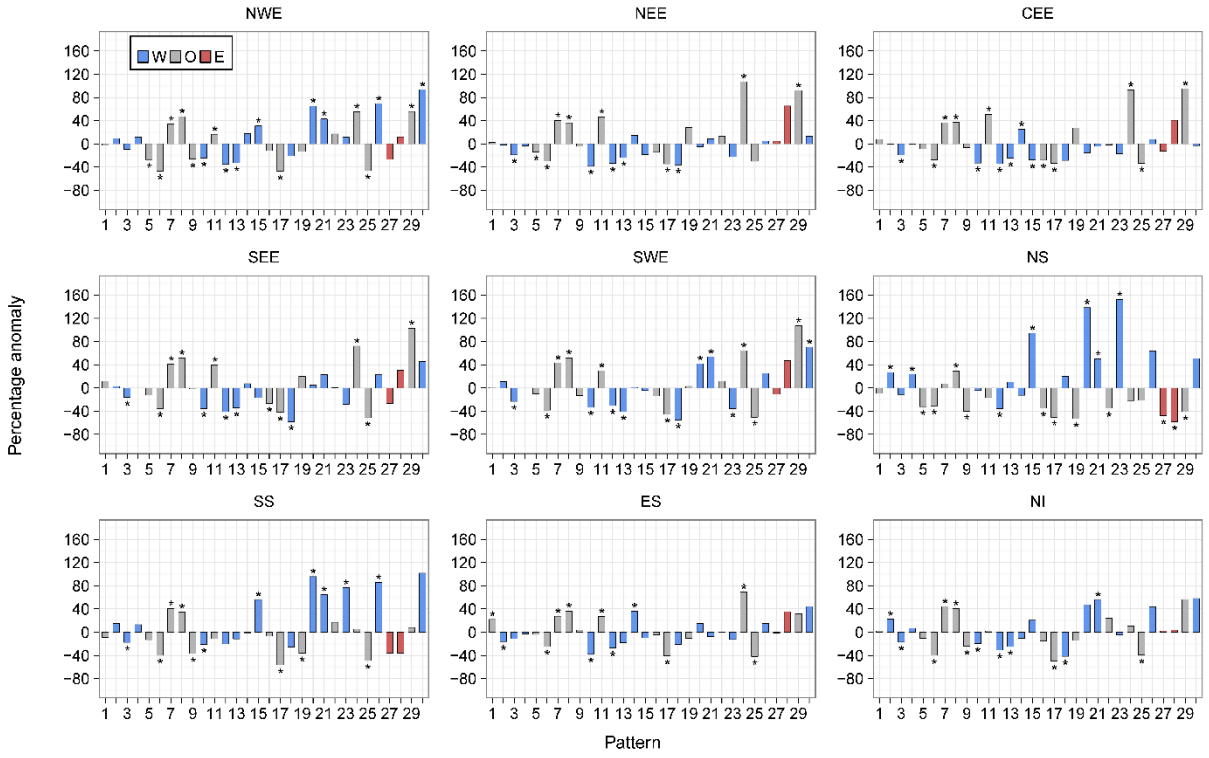

Figure A.2: As Figure A.1, but for summer.

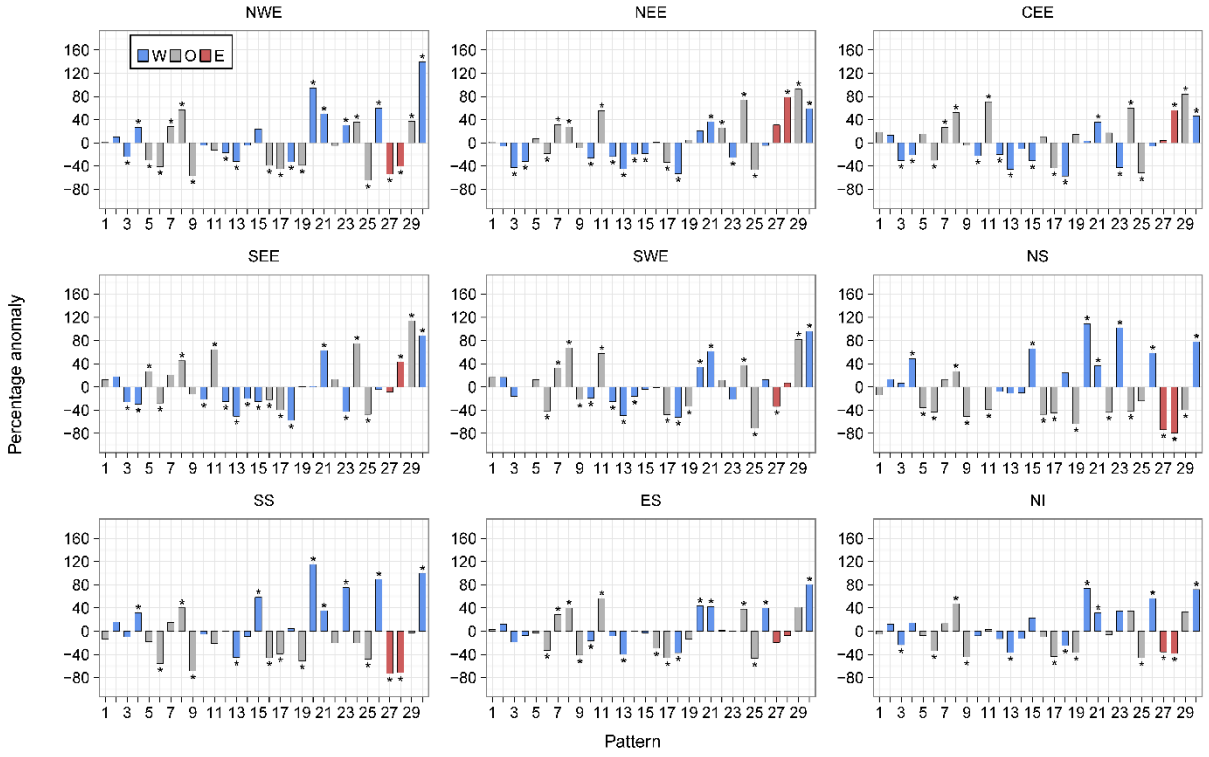

Figure A.3 As Figure A.1, but for winter.

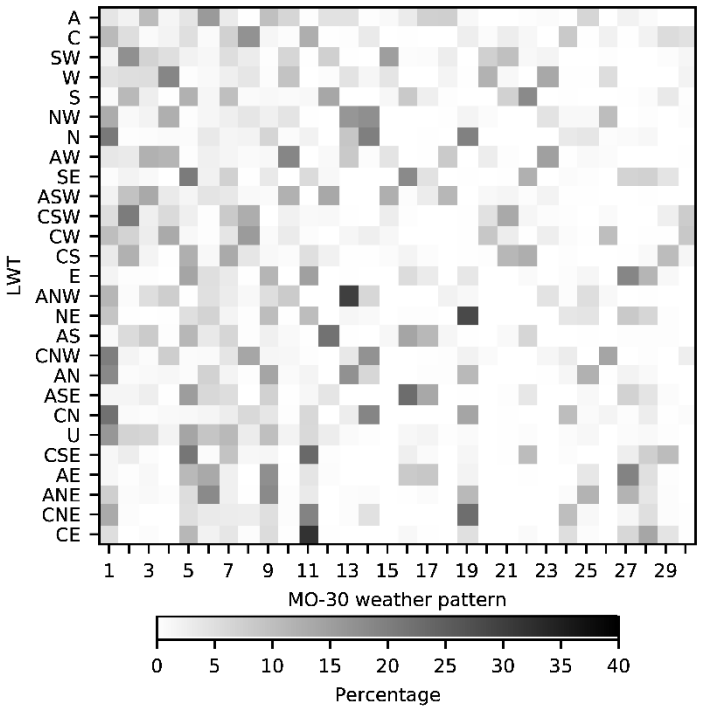

Figure A.4: Percentage occurrence of each weather pattern in MO-30 for each Lamb weather type (LWT) day between 1871 and 2015. Rows sum to 100%.

Appendix B: Supporting information for Chapter 4

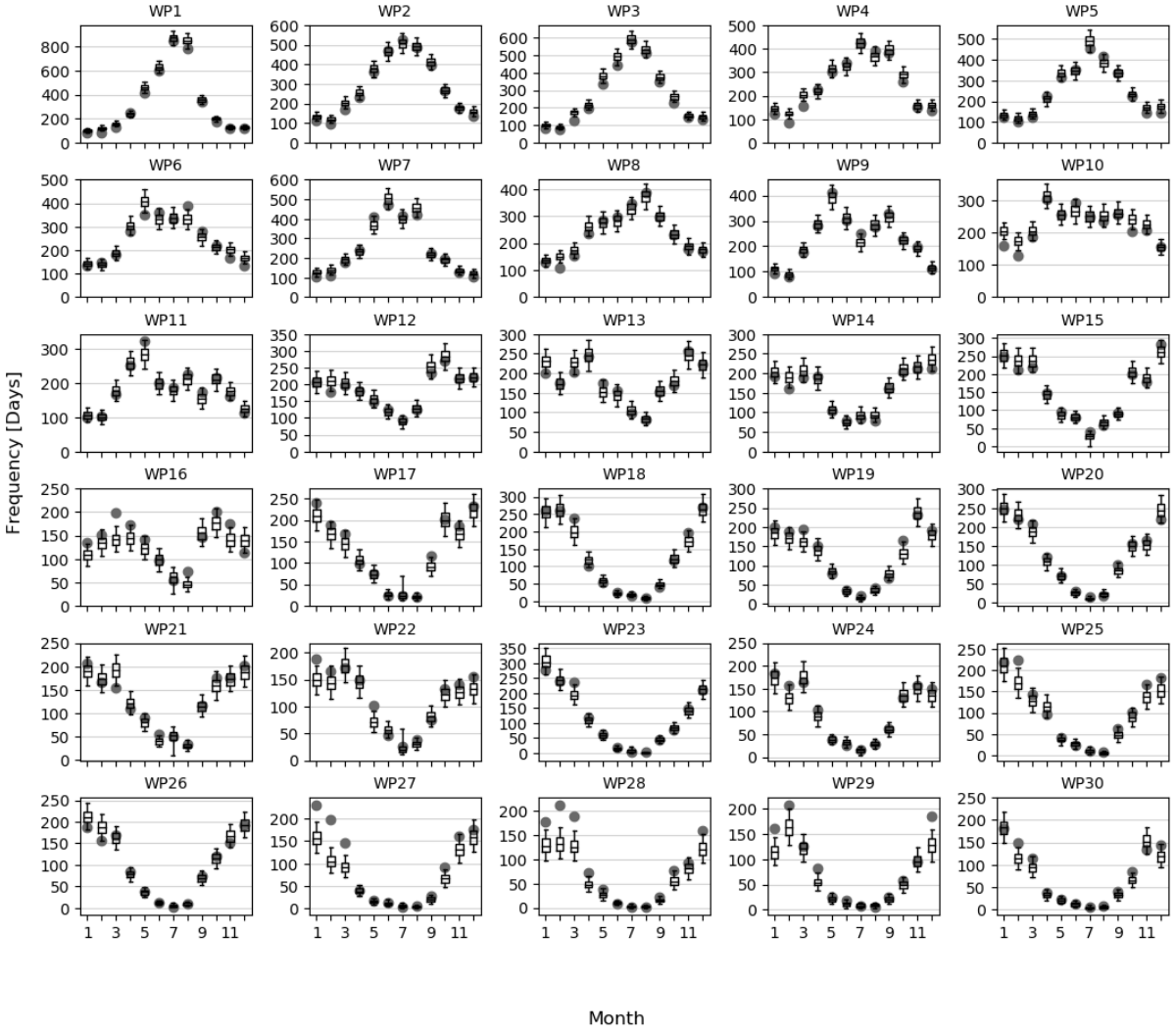

Figure B.1: Frequencies of MO-30 weather patterns by month for the observed series (grey circles) and 1000 simulated series (boxplots). Whiskers are the 5th and 95th percentiles.

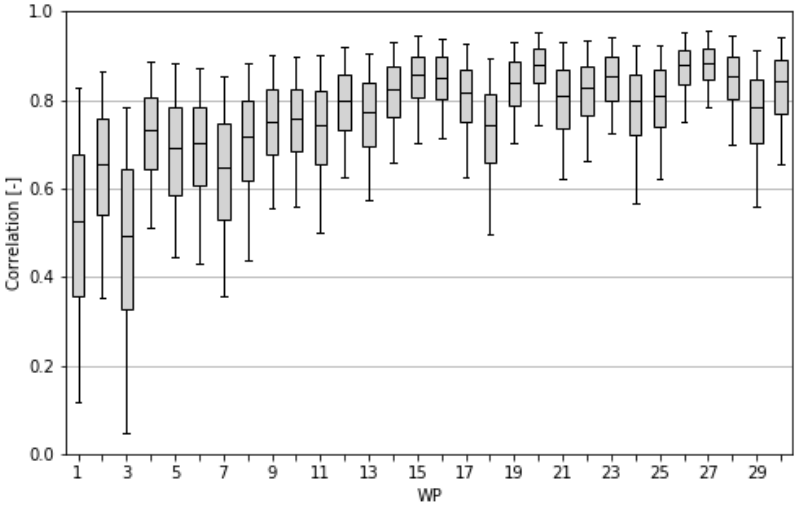

Figure B.2: Boxplots showing correlation between each weather pattern and the concurrent underlying SLP anomaly fields 1850-2016. Whiskers are the 5th and 95th percentiles.