

# The Spatial Coherence of European Droughts – Final Report

Science Report – [SC070079/SR3](#)

The Environment Agency is the leading public body protecting and improving the environment in England and Wales.

It's our job to make sure that air, land and water are looked after by everyone in today's society, so that tomorrow's generations inherit a cleaner, healthier world.

Our work includes tackling flooding and pollution incidents, reducing industry's impacts on the environment, cleaning up rivers, coastal waters and contaminated land, and improving wildlife habitats.

This report is the result of research commissioned and funded by the Environment Agency's Science Programme.

Published by:  
Environment Agency, Rio House, Waterside Drive,  
Aztec West, Almondsbury, Bristol, BS32 4UD  
Tel: 01454 624400 Fax: 01454 624409  
[www.environment-agency.gov.uk](http://www.environment-agency.gov.uk)

ISBN: XXXXXXXXXX

© Environment Agency –

All rights reserved. This document may be reproduced with prior permission of the Environment Agency.

The views and statements expressed in this report are those of the author alone. The views or statements expressed in this publication do not necessarily represent the views of the Environment Agency and the Environment Agency cannot accept any responsibility for such views or statements.

This report is printed on Cyclus Print, a 100% recycled stock, which is 100% post consumer waste and is totally chlorine free. Water used is treated and in most cases returned to source in better condition than removed.

Further copies of this report are available from:  
The Environment Agency's National Customer Contact Centre by emailing:  
[enquiries@environment-agency.gov.uk](mailto:enquiries@environment-agency.gov.uk)  
or by telephoning 08708 506506.

Author(s):  
Jamie Hannaford, Benjamin Lloyd-Hughes, Caroline Keef, Simon Parry, Christel Prudhomme

Dissemination Status:  
Publicly available

Keywords:  
Droughts, Europe, Low flow, Rainfall deficit, Streamflow deficit, UK, Water Management

Research Contractor:  
CEH, Maclean Building, Wallingford, OXON,  
OX108BB, Tel: 01491 692381

Environment Agency's Project Manager:  
Stuart Allen, Ipswich

Collaborator(s):  
Walker Institute  
JBA Consulting

Science Project Number:  
SC070079

Product Code:  
XXXXXXXXXXXX-E-P

# Science at the Environment Agency

Science underpins the work of the Environment Agency. It provides an up-to-date understanding of the world about us and helps us to develop monitoring tools and techniques to manage our environment as efficiently and effectively as possible.

The work of the Environment Agency's Science Group is a key ingredient in the partnership between research, policy and operations that enables the Environment Agency to protect and restore our environment.

The science programme focuses on five main areas of activity:

- **Setting the agenda**, by identifying where strategic science can inform our evidence-based policies, advisory and regulatory roles;
- **Funding science**, by supporting programmes, projects and people in response to long-term strategic needs, medium-term policy priorities and shorter-term operational requirements;
- **Managing science**, by ensuring that our programmes and projects are fit for purpose and executed according to international scientific standards;
- **Carrying out science**, by undertaking research – either by contracting it out to research organisations and consultancies or by doing it ourselves;
- **Delivering information, advice, tools and techniques**, by making appropriate products available to our policy and operations staff.

Steve Killeen

**Head of Science**

# Executive summary

Drought can cause serious problems across much of Europe. Many droughts are localised and short, but others are widespread and cause environmental and social effects that cross international boundaries. Some of the most important UK droughts were also significant droughts across much of Europe. Intuitively, it would seem that there may be considerable potential for developing improved drought monitoring and forecasting tools by examining the spatial coherence of droughts on a continental scale.

This project has considered the potential for developing new approaches to forecasting drought by asking the following research questions:

- Is there any systematic time lag between the onset and development of droughts in different parts of Europe?
- Can the onset and development of droughts in some parts of Europe provide an early warning for the development of droughts in other parts of Europe, and in particular, in the UK?
- Can these relationships be used to build reliable and robust operational tools for UK drought forecasting?

The method, which draws on a unique archive of flow and rainfall data from across much of Europe, involved the following steps.

1. Calculate a normalised deficiency index for each site – a measure of drought that allows comparison between locations with different climatological and hydrological regimes, and between different seasons
2. Group catchments with similar drought characteristics into regions
3. Develop standardised flow and rainfall deficiency indices for these regions
4. Analyse relationships between regions and develop statistical models to predict drought.

Twenty-four homogenous regions were identified across Europe; catchments within these groups frequently experience simultaneous streamflow deficiencies. Four distinct geographical regions emerged in the UK. A further group, comprising very slow-responding catchments (Base Flow Index > 0.8), was identified in southeast England.

For each of these regions, time series of regional streamflow and rainfall deficits were defined and a catalogue of regional drought severity developed, spanning 1901 – 2005 for meteorological droughts, and 1961 – 2005 for hydrological droughts. This enabled a characterisation of major drought periods, in terms of duration, seasonality and spatial coherence in the various regions. This drought catalogue is a major deliverable of this project, and will be of considerable practical utility for drought management and future research in the UK and in Europe.

For major post-1961 streamflow droughts, a comprehensive description of the extent and spatio-temporal development of the drought was provided. A standalone publication has been produced, which illustrates the evolution of streamflow and rainfall anomalies, along with climatic drivers and large-scale atmospheric circulation anomalies for major droughts (e.g. 1975 – 76; 1988 – 1992). From an appraisal of these events, it is clear that most droughts appear to have different characteristics, in terms of their duration, spatial coherence and seasonality. For example, a contrast was found between the 1976 drought, which was spatially consistent across much of Europe and was combined with a rainfall deficiency the preceding winter and a heat

wave in the summer, and the 1995-1997 drought, which was interspersed by wet episodes and had little long-lasting spatial coherence over Europe. In most historical events, the UK experienced drought simultaneously with other European regions, or earlier; there was little evidence of any systematic lag time which could be readily exploited in the development of early warning systems for the UK based on conditions in other parts of Europe.

An exploratory data analysis was then carried out, to determine whether there are relationships in the drought indicators which could be exploited to develop forecasting tools. Correlation analysis, multidimensional scaling and statistical modelling were applied to find relationships, which were generally fairly weak. Low correlations exist between regional drought deficiency time series of different regions, and the correlation patterns for hydrological and meteorological droughts are similar, albeit slightly higher for the latter. Correlations with the rest of Europe are stronger in winter than in summer for northern and western Britain, but are of similar magnitude all year round for southeast England. Although a relationship was identified between the length of a UK drought and the number of regions contemporaneously experiencing drought elsewhere in Europe, it was found that this relationship was not statistically significant.

Following these exploratory analyses, statistical models were built for each UK region, which predict the number of drought months that may occur in the next 6 months. Predictions are based on streamflow deficiencies in other European regions, so the models essentially predict 'drought from drought' – i.e. they use the spatial coherence of anomalies to derive forecasts for the UK based on deficiencies on the continent. The models forecast droughts in groundwater-dominated catchments in southeast England reasonably well. In northwest Britain, however, the predictive capability is poor.

Importantly, the models have some significant benefits when compared to previous seasonal forecasting studies – in particular, the approach is based on large regions, rather than being 'tuned' to particular catchments, and they enable forecasting of winter anomalies rather than just summer flows. Furthermore, the models perform reasonably well at forecasting the cessation of drought conditions. These attributes mean that the models could potentially be of high utility during long, multi-season drought events, to determine whether a drought is likely to intensify or to diminish. Whilst the predictive capacity is modest in some regions, the models clearly have potential for application in UK drought management, although there are also important practical considerations – in particular, the need for timely data supply from across Europe – which would need to be examined in further research before they could evolve into an operational tool.

Further analysis concentrated on attempting to explain observed patterns of spatial coherence, by linking drought indicators to large scale modes of atmospheric variability (e.g. the North Atlantic Oscillation and the East Atlantic-West Russia pattern). In some regions and some seasons, these predictors clearly play an important role in determining the spatial coherence of droughts. Whilst their predictive capability is relatively weak at present, there is undoubtedly scope for refining these relationships into tools for monitoring and providing indicative forecasts. An advantage of this approach is that some climatological indicators are routinely forecast (although the modest skill levels are a further obstacle to application at present).

The regional drought indicators are shown to be powerful tools for illustrating the dynamics of rainfall and streamflow deficiencies. They could therefore find application in UK and European drought monitoring systems. Again, there would be important practical limitations to consider, and further research would be needed to optimise the indicators for use in monitoring. However, they could potentially fill an important gap; existing monitoring European drought monitoring systems lack a streamflow component, whilst UK approaches (e.g. CEH's monthly Hydrological Summaries) consider runoff deficiencies but do not use any metrics tailored specifically to drought.

# Table of contents

<b>Science at the Environment Agency</b>	<b>iii</b>
<b>Executive summary</b>	<b>iv</b>
<b>Table of contents</b>	<b>vi</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Background	1
1.2 Objectives	2
1.3 Data	2
1.4 Methodology	2
1.5 Overview of the report	3
<b>2 Drought Indicators</b>	<b>4</b>
2.1 River Flow – Regional Deficiency Index (RDI)	4
2.1.1 Deriving at-site deficiency indices	4
2.1.2 Cluster Analyses	5
2.1.3 Regional Deficiency Index (RDI)	6
2.2 Rainfall – Regional Standardised Precipitation Index (RSPI)	7
2.2.1 Standardised Precipitation Index	7
2.2.2 Regional Standardised Precipitation Index	8
<b>3 Drought regions</b>	<b>9</b>
3.1 Indicators of regions	9
3.2 Modified regions	14
3.3 Final regions	15
<b>4 Drought Catalogues</b>	<b>17</b>
<b>5 Appraisal of Historical Droughts</b>	<b>20</b>
5.1 Rationale and approach	20
5.2 Spatio-temporal development of major European droughts	21
5.3 Drought evolution in North West Europe	22
5.4 Summaries of synoptic conditions	23
5.5 Meteorological Droughts before 1960	24
5.6 Observations on Spatial Coherence of Major droughts	24
<b>6 Exploratory Analysis</b>	<b>26</b>
6.1 Relationships in flow data	26
6.1.1 Raw RDI data	26
6.1.2 Drought periods	30
6.2 Precipitation	32
6.3 RDI Vs RSPI and lagged relationships	34
6.4 Temperature	36
<b>Science Report – The spatial Coherence of European Droughts – Final report</b>	<b>vi</b>

<b>7</b>	<b>Model Fitting</b>	<b>39</b>
7.1	Introduction	39
7.2	Linear modelling	39
7.3	Seasonal prediction models	40
7.4	Length of drought	44
7.4.1	Fitting distribution to length of drought	45
7.4.2	Best fitting models	46
<b>8</b>	<b>Linking Spatial Coherence to atmospheric Circulation</b>	<b>49</b>
8.1	Introduction	49
8.2	Correlation Analysis	49
<b>9</b>	<b>Evaluation of forecasting methods</b>	<b>54</b>
9.1	Evaluation of seasonal prediction models	54
9.2	Forecasting from large-scale circulation	57
9.2.1	Western Europe	57
9.2.2	North West UK	58
9.3	Comparison with previous forecasting work	58
<b>10</b>	<b>Potential operational utility of new methods</b>	<b>64</b>
10.1	Practical uses of the drought catalogues	64
10.2	Use of the RDI and RSPI in drought monitoring	64
10.3	Forecasting Methods	69
<b>11</b>	<b>Conclusions and suggestions for further research</b>	<b>71</b>
11.1	Summary of main findings	71
11.2	Data issues	73
11.3	Caveats and Limitations	74
11.4	Further research	74
	<b>References</b>	<b>78</b>

## Index of Figures

Figure 2-1	Definition of moving Q90 for a snow-driven catchment (left) and pluvial catchment (right)	5
Figure 2-2	The 22 original Drought European regions defined by (Demuth and Stahl, 2001) and (Prudhomme and Sauquet, 2007)	6
Figure 2-3	Example of RDI series, for the North East region of France	7
Figure 2-4	Standardised precipitations for Madrid and London June-July-August 1901-2005	8
Figure 3-1	Final European Drought regions	16
Figure 4-1	Example of Page 1 of the Drought Catalogue for South West Great Britain	18
Figure 4-2	Example of Page 2 of the Drought Catalogue for South West Great Britain	19
Figure 5-1:	Selected screenshots showing phases of the 2003 summer drought in the UK and Europe	21
Figure 5-2:	Matrix showing the evolution of RDI deficiencies in the 1975/76 drought	22
Figure 5-3:	Matrix showing the evolution of RDI deficiencies in the 1995 - 1997 drought	23
Figure 5-4:	Pressure anomalies associated with 1962 – 1965 drought, and temperature anomalies associated with the cold winter of 1962/1963	23
Figure 6-1:	Pairwise scatter plots of RDI for region pairs with highest correlation (left) and lowest correlation (right).	27
Figure 6-2:	Rank correlations with RDI for UK regions, left all data, middle, summer months, right, winter months	28
Figure 6-3:	Groupings for first three principal components for daily RDI data. Left PC1, middle PC2, right PC3. Approximate percentage variation explained for each PC – PC1 33%, PC2 11%, PC3 9%.	29
Figure 6-4:	Definition of drought periods	30
Figure 6-5:	Drought periods for all regions	31
Figure 6-6:	Correlations with RSPI_1 for UK regions, left all data, middle, summer months, right, winter months	33
Figure 6-7:	Groupings for first three principal components for monthly RSPI data. Left PC1, middle PC2, right PC3. Approximate percentage variation explained for each PC – PC1 38%, PC2 13%, PC3 10%.	34
Figure 6-8:	Lagged correlations with RDI for UK regions. Left plots are correlations using all 6 month seasons with previous season, middle plots are all summer seasons with the previous winter season; right plots are all winter seasons with the previous summer season.	35
Figure 6-9:	Time series plot of seasonal RDI and seasonal temperature anomalies for five UK regions. RDI plotted as solid black lines temperature anomalies plotted in grey, with summers highlighted in red and winters in blue.	38
Figure 7-1:	Scatter plots of monthly mean RDI and monthly maximum RDI for three selected UK regions.	40
Figure 7-2:	Regions where lagged correlation with any UK region is greater than 0.4	41
Figure 7-3:	Real and predicted number of future drought months for UK regions black lines real, red lines predicted.	43
Figure 7-4:	Scatter plots of drought maximum Vs drought length for UK regions	44
Figure 7-5:	Fitted distributions for each of the UK regions. Line indicates perfect fit, crosses indicate distribution function of data points (drought lengths), horizontal dashes indicate 95% confidence limits using the empirical data.	46
Figure 7-6:	QQ plots of fitted model drought length Vs real drought lengths	48
Figure 8-1	Rank correlations between RSPI1 and the North Atlantic Oscillation by month for the period 1951-2005.	51
Figure 8-2	Rank correlations between RSPI1 and the East Atlantic West Russia pattern by month for the period 1951-2005.	52
Figure 8-3	Rank correlations between RSPI1 and the Scandinavian Pattern by month for the period 1951-2005.	53
Figure 10-1	Comparison between the RDI Maps (with four UK regions, legend in centre) and Monthly Hydrological Summaries for the same months in 2003	65
Figure 10-2	Example of RDI and RSPI Metrics for April 1997	67
Figure 10-3	Comparison between RDI plots produced using three thresholds (l – r): Q70, Q90 and Q95. For SE Great Britain, non-Groundwater region	68

## Index of Tables

Table 3-1	Analysis of regions drought coherence	11
Table 3-2	Table of metrics for revised UK and Scandinavian regions	15
Table 6-1:	Highest 4 correlations for each region, correlation measured using Spearman's rho correlation coefficient. Different colours reflect different geographic regions; brown UK; blue France and Spain; green alps; pink Germany and further east and orange Scandinavia.	27
Table 6-2:	Correlations between length of drought in each UK region and number of other regions experiencing a drought (including other UK regions).	32
Table 6-3:	Lagged seasonal Spearman's rank correlations between seasonal RDI and previous seasonal RSPI_1 for UK regions.	36
Table 7-1:	Coefficient values for all fitted models positive values indicate a higher expected number of drought months, negative values indicate a lower expected number of drought months.	42
Table 7-2:	R <sup>2</sup> values for best fitting model	47
Table 8-1	Global circulation indices used in the correlation study with regionally coherent drought ( further details at the US Climate Prediction Center website ( <a href="http://www.cpc.noaa.gov/data/teledoc/telecontents.shtml">http://www.cpc.noaa.gov/data/teledoc/telecontents.shtml</a> ))	50
Table 9-1:	Performance of model in predicting long seasons with at least two drought months*.	55
Table 9-2	Skill scores for models to predict number of drought months in the next six.	57
Table 9-3	Contingency table for rule based prediction of winter drought over central western Europe 1961-2004.	57
Table 9-4	Contingency table for rule based prediction of winter drought over central north western UK 1961-2004.	58
Table 9-5	Review of previous seasonal forecasting literature	60



# 1 Introduction

## 1.1 Background

Drought can be defined as a “sustained and regionally extensive occurrence of below average water availability” (Tallaksen & Van Lanen, 2004), and generally begins with a protracted period of reduced (or nil) precipitation over a wide area that may subsequently propagate to the entire hydrological cycle. Different types of drought manifest at different temporal and spatial scales, depending on the physical processes involved and antecedent conditions (Zaidman *et al.*, 2002): this project deals primarily with meteorological droughts (rainfall deficiency); and hydrological droughts (river flow deficiency) which develop over a longer time scale and can occur much later than meteorological droughts depending on the storage characteristics of a catchment (Marsh *et al.*, 2007a).

Several major droughts experienced in the UK over the last 40 years (e.g. 1975-76, 1988-92, 1995-96, 2003, 2004-06) also affected large parts of mainland Europe. Such events have many wide-ranging impacts, including disruption to public water supply and the disposal of effluent, reductions in power generation, industrial, agricultural and forestry production, effects on fisheries, navigation and leisure amenities and the derogation of the environment, with reduced habitat for flora and fauna. Secondary impacts include increased risk, and incidence, of forest- and bush-fire. In financial terms, the cost of drought is substantial: droughts in Europe, for example, are estimated to have cost over €13 billion in 2003 (COPA COGECA, 2003), and at least €100 billion over the last 30 years (CEC, 2007). Anticipating the likelihood of a drought would allow better preparedness and mitigation of these impacts.

Previous research has examined droughts at a continental scale in Europe. The ARIDE (ENV4-CT97-0553) project (Demuth & Stahl, 2001) sought to characterise droughts in Europe, and their impact. ARIDE examined spatial coherence of droughts; for example, in developing tools to visualise spatio-temporal evolution of major drought events (Zaidman *et al.* 2002). The ARIDE project demonstrated that the spatial extent of droughts can be used as an indicator of drought severity (Stahl & Demuth, 1999; Stahl, 2001; Hisdal & Tallaksen, 2003). Some work was also directed towards attempting to predict droughts from Circulation Types (e.g. Stahl & Demuth, 1999), but met with a limited degree of success in many parts of Europe. Whilst this research led to an improved understanding of spatial coherence, the study did not attempt to develop forecasting or early warning systems. Furthermore, the study was carried out on data up to 1990 (thereby not including more recent events), and the detailed analyses of drought evolution focused on only two events (1975/6 and 1989 – 1990).

Other research has tended to examine variability in meteorological droughts, using indices such as the Palmer Drought Severity Index (PDSI) and Standardised Precipitation Index (SPI), e.g. Lloyd-Hughes & Saunders (2002); Van der Schrier *et al.* (2006); Briffa *et al.* (2009). Previous work has not examined both hydrological and meteorological droughts simultaneously, which is a weakness for hydrological studies as the propagation of drought from rainfall to river flow or groundwater deficits is complex (Tallaksen & Van Lanen, 2004).

There are therefore major gaps in our knowledge of spatial coherence of drought in Europe, particularly in terms of the space-time development of droughts, the large scale propagation of meteorological to hydrological droughts, and linkages of drought characteristics to large scale circulation.

## 1.2 Objectives

Droughts are hydrometeorological events that evolve to afflict large swathes of territory, typically for periods of several months or years. Because of their gradual emergence, droughts tend to show strong spatial coherence over wide regions. Such behaviour suggests improved understanding of how droughts develop in Europe could help with the management and mitigation of droughts in the UK. The project examined three main questions:

- Do historical European droughts show spatial coherence, and is there any systematic time lag between the onset and development of droughts in different parts of Europe?
- Can the onset and development of droughts in some parts of Europe provide an early warning for the development of droughts in other parts of Europe, and in particular, in the UK?
- Can these relationships be used to build reliable and robust operational tools for UK drought forecasting?

## 1.3 Data

The project benefited from access to the European Water Archive (EWA), the National River Flow Archive and the French Banque Hydro river flow data banks and the good quality river flow data they hold.

Through links with the EU-FP6 WATCH (Water and Global Change) project, it was possible to obtain a dataset of recently updated (to 2005) streamflow data for 10 countries across Europe, updated as part of a WATCH and UNESCO funded project (Stahl *et al.* 2008). This dataset is comprised of catchments with minimal anthropogenic disturbances on flow regimes, having gauging stations regarded to have good hydrometric performance and records extending back to 1962.

Added to this were additional data from the UK benchmark catchments, a set of undisturbed catchments previously used in UK trend analysis studies (Hannaford and Marsh, 2006), and data from Banque Hydro in France, which again were filtered so that only near-natural catchments were included, as employed by Prudhomme and Sauquet (2006).

The precipitation data are taken from the CRU 0.5° x 0.5° gridded analysis 1901-2006 (Mitchell and Jones, 2005). Gridded precipitation data are preferred over raw raingauge observations because they reduce biases arising from the irregular distribution of the raingauge network (Jones and Hulme, 1996, Dai *et al.*, 1997).

## 1.4 Methodology

Two drought indicators were used to examine the research question, representing hydrological and meteorological droughts. Both indicators estimate the proportion, within a delimited region, of the region experiencing a flow (rainfall) deficiency simultaneously: the Regional Drought Index (RDI) for hydrological droughts, and the Regional Standardized Precipitation Index (RSPI) for meteorological droughts. These indices are a measure of the spatial coherence of a drought within a region, a measure of drought severity. The methodology used to compute the indices allows for difference in the overall regional climates and in the seasonality of the river flow and rainfall regime to be incorporated in the indices. It was implemented consistently throughout Europe, resulting in daily time series of RDI and RSPI for 24 regions.

The RDI and RSPI are an objective measure to characterise droughts, easy to implement, and enabling consistent comparison of drought characteristics between regions. They form the backbone of the analysis, enabling:

- To identify, within a region, 'drought rich' (when there is a high degree of spatial coherence in the catchments experiencing very low levels river flow/rainfall for the time of the year) or 'drought poor' (when available data does not show very low river flow/rainfall observations) periods: the drought catalogues
- To identify, within a region, some major regional drought characteristics: duration of 'drought rich' periods, level of spatial coherence, seasonality in the 'drought rich' periods
- To examine the temporal and spatial evolution of major historical European droughts
- To identify patterns of spatial coherence common to major droughts, and to see whether they can be explained, for example using some large scale atmospheric indices
- To quantify any links between the RDI and RSPI time series of different regions
- To exploit any quantified links to develop tools for forecasting drought characteristics (for example start, duration or end)

## 1.5 Overview of the report

The methodology used to answer the key research questions in this project is described in Sections 2 and 3. The analysis of the historical droughts identified in Europe was developed into a Drought Catalogue (Section 4). Section 5 describes qualitative appraisals of large-scale European drought events (e.g. 1975 – 1976 and 2003), which were carried out to determine whether patterns of spatial coherence could be identified and exploited within the project. Section 6 describes an exploratory analysis of the drought indicators, focusing on developing statistical relationships which can underpin models. In Section 7, these analyses were built on with the development of statistical models designed to improve forecasting. Section 7 explores relationships between the drought indicators and large scale modes of atmospheric circulation. The tools developed in the preceding sections are evaluated in Section 9, using statistical methods to assess performance, whilst Section 10 considers their potential for operational applications. Finally, conclusions are presented in Section 11, which summarises the main findings, considers caveats and limitations, and ends with recommendations for further research.

As part of the project's objectives, a European-wide catalogue of droughts has been established, from 1961 for hydrological droughts, and as far back as 1901 for the meteorological droughts. Such a product is a powerful tool to examine the spatial coherence of droughts across different regions and Europe and their spatio-temporal development. An associated product generated in this project is a summary description of the spatio-temporal evolution of 5 major drought episodes that have affected Europe since 1961. These two products (Lloyd-Hughes *et al.* 2009 and Parry *et al.* 2009) are provided as standalone reports, complementary to this final project report.

# 2 Drought Indicators

This section introduces the drought indicators that were used to define historical droughts for the project. Two drought indicators – one describing hydrological drought, the other meteorological droughts – were derived. The former approach is derived from the Regional Deficiency Index (RDI), developed in the EU-funded project ARIDE (Demuth and Stahl, 2001). Meteorological droughts were defined following the same spatial coherence concept but are based on the Standardised Precipitation Index (McKee *et al.*, 1993). A description of the steps involved in deriving these indicators is given below.

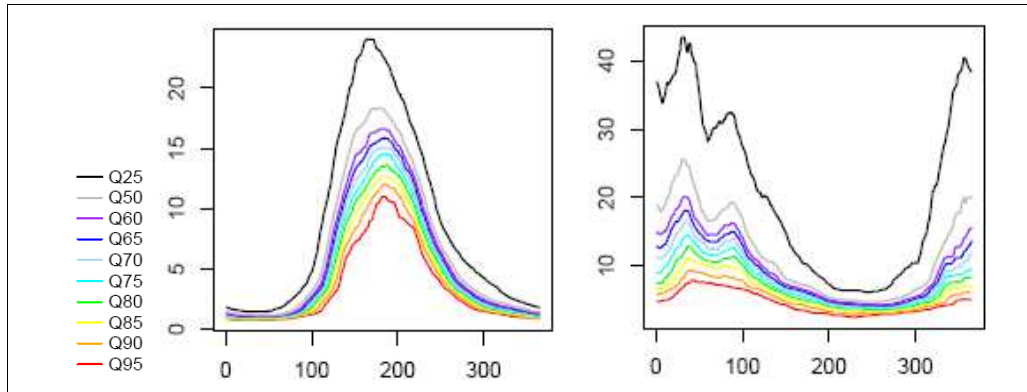
## 2.1 River Flow – Regional Deficiency Index (RDI)

The Regional Deficiency Index was developed by Demuth & Stahl (2001) as a way of characterising drought within homogeneous regions. Firstly, a time series of streamflow deficiencies are characterised in individual gauging station time series using a Deficiency Index. Secondly, homogenous regions are defined from these records, using a cluster analysis. Finally, for each homogeneous region, a Regional Deficiency Index is computed. The approach is described in detail in the following sections.

### 2.1.1 Deriving at-site deficiency indices

The first step requires the calculation of a Deficiency Index time series for each time series of daily flow data. In order to account for the inter-annual characteristics of river flow regimes, in which similar flow variations may be expected during typical years, the approach involves evaluation whether the flow on any given day falls below a daily varying (moving) low flow threshold – this is a flow deficiency. In this study, the flow exceeded 90% of the time (the Q90 flow) was used as the threshold, such that a different Q90 threshold value was calculated for every day of the year (i.e. 365 Q90 values). The daily “moving Q90” value is calculated by ranking all observed flows on the “day of interest” together with those 15 days either side of the day of interest. For example, the moving Q90 value for 16 January (the day of interest) would be calculated from all flow values recorded 1-15 January (15 days previous), 16 January (day of interest) and 17-31 January (15 days after) from every year of the data record. For a complete daily record of, say, 20 years’ duration, the moving Q90 for any day would thus be based on a sample of 620 (31x 20) observations.

The moving threshold approach is illustrated in Figure 2-1 for two catchments having contrasting river flow regimes: a snow-melt dominated regime (left); and a rainfall-dominated regime (right). The figure shows the discharge in cumecs (y-axis) associated with different flow exceedence percentiles (coloured lines) calculated for each day of the year (x-axis), from Q25 flow (black lines) to Q95 (red line). As can be seen, the orange (Q90) line does not remain constant throughout the year and is higher during high-flow seasons than in low flow seasons.



**Figure 2-1** Definition of moving Q90 for a snow-driven catchment (left) and pluvial catchment (right)

A single, constant, value clearly would not capture such intra-annual variation and would be misleading, as regards real periods of deficiency. The figures also illustrate the variability of the moving Q90 between one catchment and another.

The Deficiency Index of a particular streamflow daily time series is determined according to whether each daily streamflow value is equal to or lower than the corresponding daily threshold (Q90) value. If the flow is less than the threshold, then the conditions experienced during that day are amongst the 10% driest for that day and potentially represents a period of extreme low flow, or drought.

The *Deficiency Index* time series simply reduces the streamflow series to a binary time series populated with values of 0, when the flow is greater than the moving Q90 for that day, or 1, if lower, as given by the following equation:

$$DI(t) = 1 \quad \text{if } Q(t) \leq Q90(D)$$

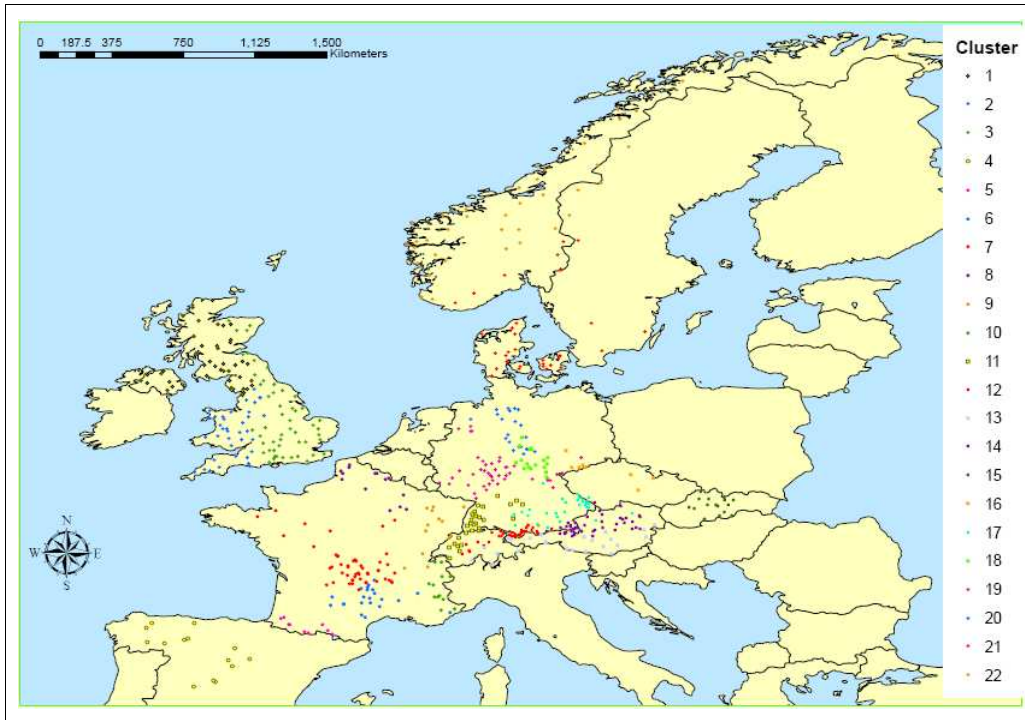
$$DI(t) = 0 \quad \text{if } Q(t) > Q90(D)$$

### 2.1.2 Cluster Analyses

The second step groups catchments that experience an abnormal low flow at the same time into homogeneous regions. The clustering is based on the premise that droughts are generated over large swathes of territory and, thus, affect many catchments simultaneously. In order to delineate regions that behave similarly according to a large-scale circulation conditions, and might experience a drought together, the Deficiency Index series of all catchments are compared. Clustering aims to put together all gauging stations (catchments) which have Deficiency Index equal to 1 on the same day, and to put catchments that have a Deficiency Index equal to 0 that day into different groups (see Stahl & Demuth, 1999 for further details).

There are many clustering methods, which are generally time-consuming to apply. Running a new cluster analysis would be time-consuming and would duplicate previous work. The decision was made, therefore, to build upon the original cluster analysis performed for Europe in the development of the RDI (Demuth and Stahl, 2001), which defined 19 clusters across Europe. In the aforementioned study, however, there was very limited data available for France; in the present study, a much denser dataset is available, and France was thought to be of particular importance in virtue of it being close to the UK. The decision was made, therefore, to introduce six homogenous regions in France, based on the cluster analysis performed by Prudhomme and Sauquet, 2007. When these 25 clusters were applied to the dataset used in the present study, a total of 22 homogeneous regions were available, using river flow data

from 10 countries, as shown in Figure 2-2. A review and redefinition of the original regions was conducted - this activity is described more fully in Section 3.



**Figure 2-2 22 original Drought European regions defined by Demuth and Stahl (2001) and Prudhomme and Sauquet (2007)**

### 2.1.3 Regional Deficiency Index (RDI)

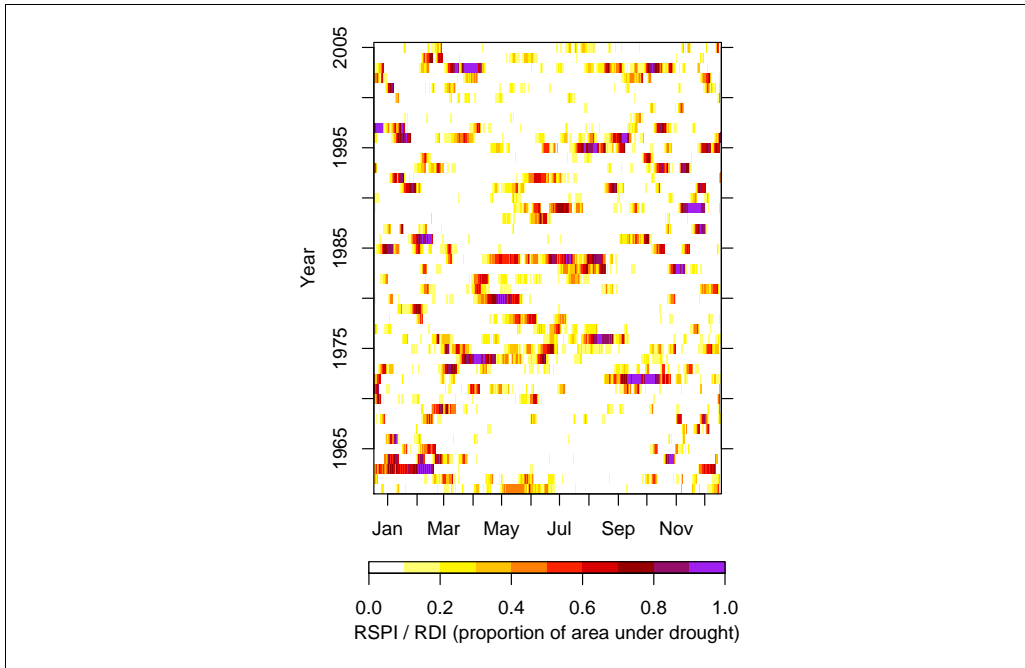
The final step in the establishment of the drought catalogue is to derive the Regional Deficiency Index (RDI). For each of the regions found to experience exceptional low flows at the same time (i.e. the clusters defined in 2.1.2), the proportion of the region experiencing such abnormal low discharge is derived by calculating, for each day of the catalogue, the arithmetic mean of the Deficiency Index series of all catchments within the region, as described in the following equation:

$$RDI(t) = \frac{1}{M} \sum_{i=1}^M DI_i(t)$$

with M the number of catchments with available data and derived DI series in the region. This represents the proportion of catchments in the region that experience abnormal low flows at that time.

By construction, the RDI series takes values between 0 and 1. A value of 0 reflects that none of the gauging stations (catchments) where data was available had exceptionally low flow, and thus, the region is not in a drought condition. In contrast, a value of 1 occurs when all catchments with data had extremely low discharge for the day (i.e. DI equals to 1). This situation is exceptional and defines a very severe drought. Values ranging between 0 and 1 define an event with very few catchments experiencing low flows conditions (RDI towards 0) to an event when the majority of the region experienced low flows (RDI towards 1). An RDI of 0.3 is recommended as a

minimum to define the existence of a severe drought, corresponding to 30% of the region showing a streamflow deficit (Stahl, 2001). An illustration of RDI series is given in Figure 2-3 for the region of North East France. Coloured bars show the periods where RDI is greater than 0.1. Contiguous bars of colour represent prolonged droughts, the darker the colour, the more severe the drought (higher RDI), and the greater the drought severity. Drought-poor periods are easily identified by the periods in white.

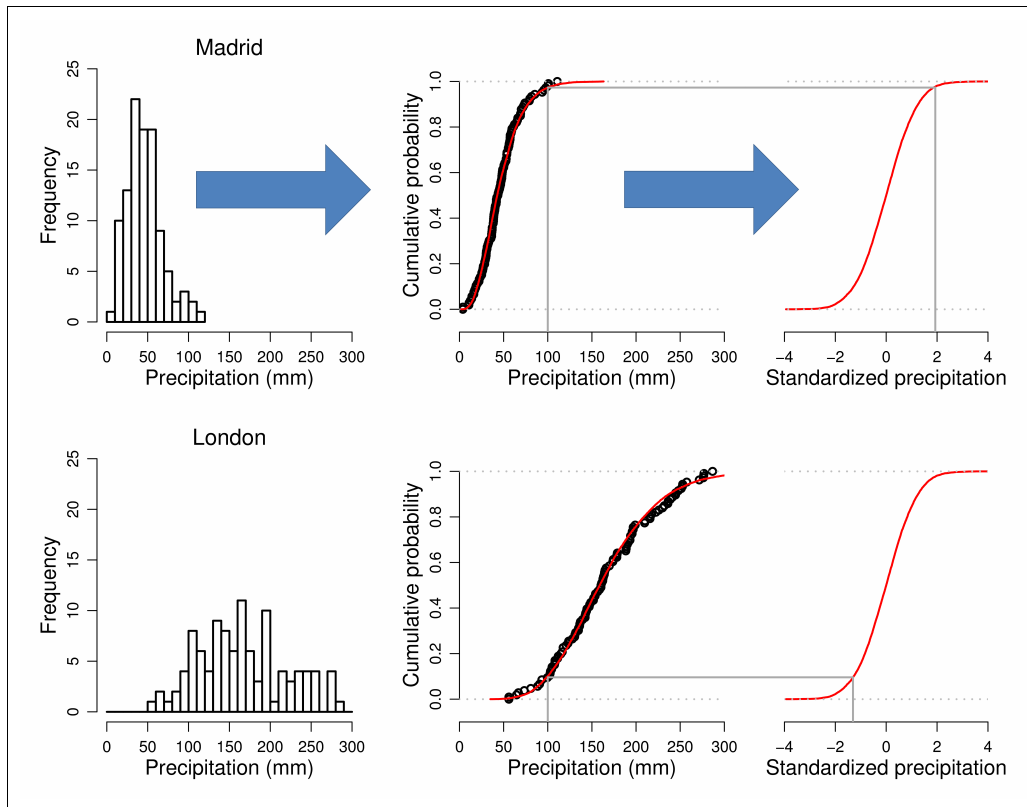


**Figure 2-3 Example of RDI series, for the North East region of France**

## 2.2 Rainfall – Regional Standardised Precipitation Index (RSPI)

### 2.2.1 Standardised Precipitation Index

In order to extend the spatial coverage and historical extent of the drought record beyond that possible from the hydrological gauge network, the Standardized Precipitation Index (SPI) (McKee *et al.*, 1993) was used as a meteorological proxy for drought. The SPI is the unit normal transformation of the time averaged precipitation time series climatologically appropriate to the particular location and time of year. This is illustrated in Figure 2-4 which shows the transformation of precipitation accumulated over a 3 month period (June-July-August) for two contrasting regions, namely, London and Madrid.



**Figure 2-4 Standardised precipitations for Madrid and London June-July-August 1901-2005**

The relative importance of a given amount of rainfall accumulated over a particular time period varies from place to place and from time to time in the year. The SPI value allows us to compare, say, 100mm over the summer in Madrid (+2 i.e. very wet) with 100mm over the summer in London (-1.3 i.e. quite dry). The example illustrates August SPI3. A similar comparison for precipitation accumulated in the six months from July to December would be described as December SPI6. The month always refers to the end of the accumulation period and SPI  $n$  to the  $n$ -month total that is being standardized.

Since the SPI is by definition normally distributed we can assign return periods to droughts of a given severity: Moderate < -1 (5 years), Severe < 1.5 (15 years), Extreme < -2 (40 years).

## 2.2.2 Regional Standardised Precipitation Index

Since our aim is to provide a direct mapping of the meteorological description to the hydrological, the Regional Standardized Precipitation Index (RSPI) follows directly from the homogeneous reporting units identified during the hydrological classification. It is defined as the proportion of the grid cells under the region boundary experiencing moderate drought i.e. with SPI < -1. A representative timescale for the SPI (e.g. 3, 6, or 12 months) was determined for each region by the maximum rank correlation between the index and the RDI. Storage effects were allowed for by the consideration of lagged correlations of up to 12 months.



# 3 Drought regions

In order to examine spatial coherence across Europe, the decision was made to use discrete hydrological regions as the framework for characterising drought. As described in Section 2.1.2, a set of 22 homogenous regions were defined on the basis of previous research projects. However, as the source river flow data used in the present study was different (and more up-to-date) to that used in the original research, a review was undertaken to ensure that the regions would be suitable for this project. A series of indicators were developed to test the existing regions (Section 3.1) and, on the basis of this testing, some regions were modified (Section 3.2).

Note that the construction of homogeneous regions using cluster analysis applies to hydrological droughts only. In this study, the meteorological droughts were then extracted using these same regions, but using a boundary derived from the gauging station points (discussed further in Section 3.3).

## 3.1 Indicators of regions

For each of the 22 original regions, metrics were derived that provide some summary of the characteristics of the region in terms of hydrology and droughts. These indicators are described in the next paragraphs, and given in Table 3-1. These metrics were used to evaluate whether the original regions, derived from two different studies and using different hydrometric records lengths, should be further refined. For each region, the following is provided:









- **Number of gauging stations.** The total number of streamflow records in a given region. A greater number of time series increases the significance of a larger RDI value for a given cluster, because more of the records have to be in deficit conditions to generate such a value;
- **Inter-gauge distance.** This is the median distance between all the gauges located in a region, given in km. Dense, small regions will have a smaller inter-gauge distance than larger areas;
- **Mean monthly hydrograph.** Derived for each station using the mean monthly discharge standardised by the mean annual discharge. The seasons of high and low flows provide a visual indicator of the dominant streamflow regime (e.g. rainfall-dominated, snow-melt dominated or mixed). The clusters have been derived using purely statistical techniques aiming to maximise the simultaneous occurrence of Drought Index. The streamflow regime is an independent measure of the hydrological homogeneity of the clusters;
- **Cluster Homogeneity.** Based on the definition of the Q90 and the derived DI series, the RDI of gauging stations within a “perfect” region (where a region is defined as a cluster of stations that experience drought at exactly the same time), will be 0 for 90% of the time, and equal to 1 for 10% of the time. The Cluster Homogeneity is the distance between the cumulative distribution function of the regional RDI and the ideal cumulative distribution. The Cluster Homogeneity measure can have values of 0 to 1 (or 0% to 100%), with a perfect region having a value of 0.
- **Relative proportion of severe droughts.** This measures how likely a small number of gauges in the region experience a drought: the larger the number of stations experiencing a drought at the same time, the more coherent the region is in terms of its response to drought conditions. This is calculated using:


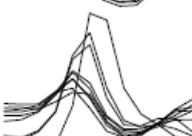








$$P_c = \frac{P(RDI > 0.5)}{P(RDI > 0)}.$$


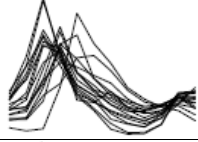
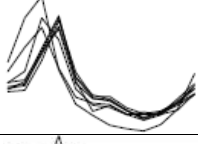

A totally coherent region (all stations have  $DI=1$  at the same time) has a  $P_c$  equal to one, while  $P_c$  equals zero if the probability of more than half the gauging stations in that cluster experiencing a drought at the same time is zero.

- **Correlation between RDI and RSPI** (on monthly series). This provides an indication of how well the regional meteorological drought series (RSPI) matches the hydrological drought series (RDI). This will be useful to extend the analysis prior to 1961, when most of the gauging stations records start.

**Table 3-1 Analysis of regions drought coherence**

	Region	Number of gauging stations	Inter-gauge distance (km)	Mean Hydrograph	Cluster homogeneity (%)	$P_c$	Correlation RDI and SPI
UK	1	43	197.7		7.5	0.146	0.13
	2	37	155.6		8.1	0.124	0.29
	3	57	207.1		9.1	0.079	0.14
Alps	11	28	119.5		7.7	0.152	0.62
	12	35	78.7		8.8	0.085	0.53
	13	25	215.3		11.4	0.021	0.50
Scandinavia	21	31	241.9		8.6	0.099	0.44
	22	39	385.8		10.9	0.024	0.05

	Region	Number of gauging stations	Inter-gauge distance (km)	Mean Hydrograph	Cluster homogeneity (%)	$P_c$	Correlation RDI and SPI
Western Europe	4	14	226.5		10.8	0.036	0.58
	5	9	90.5		8.1	0.160	0.33
	6	18	108.5		7.3	0.163	0.58
	7	48	132.7		7.5	0.128	0.65
	8	14	139.2		9.5	0.096	0.61
	9	12	134.5		7.0	0.214	0.59
	10	12	97.3		8.6	0.099	0.50
	18	16	72.3		7.9	0.109	0.50
	19	47	143.5		7.8	0.116	0.68
	20	19	111.3		8.5	0.111	0.58

	Region	Number of gauging stations	Inter-gauge distance (km)	Mean Hydrograph	Cluster homogeneity (%)	$P_C$	Correlation RDI and SPI
Eastern Europe	14	49	113.2		9.8	0.055	0.50
	15	17	81.6		9.7	0.081	0.55
	16	8	162.5		8.2	0.132	0.50
	17	30	115.8		9.3	0.086	0.61

In order to investigate possible relationships between drought occurrences of different regions, it is important that drought behaviour within individual regions is as homogeneous as possible, as it is unlikely that reliable correlations would be obtained between non-homogeneous regions. The reason for this is that if catchments within a region are not homogeneous, they are unlikely to simultaneously experience drought conditions. In RDI terms, such a region would never experience severe droughts and establishing statistically significant relationships with any other regions would be unlikely.

From Table 3-1, two areas were considered for cluster refinement: Region 3 (in the UK) and Region 22 (in Scandinavia):

Region 22, primarily located in Norway, is the least coherent in terms of  $P_C$ , the relative proportion of severe droughts, and has the smallest correlation between RDI and RSPI. It covers the largest geographical area and has the largest inter-gauge distance. It also contains gauges that measure flows for inland rivers with a predominantly snowmelt regime and those on the coast with a more rainfall-dominated regime. Re-arranging this region with the adjacent Region 21 could lead to more coherent sub-regions.

Region 3, covering most of the eastern side of Britain, although not the least coherent of all remaining regions, has the 4<sup>th</sup> lowest  $P_C$  value and is the least coherent British region. As one of the primary aims of Stage 2 of the project is to examine correlations between UK droughts and European droughts, a special effort was made into obtaining coherent UK regions. The large geographical area of Region 3, the high precipitation gradient from north east Scotland (where average annual rainfall can be >1000mm) and south east England (where average annual rainfall can be < 500mm), and its grouping of catchments having widely varying flow regimes, from slowly responding chalk catchments to quickly responding upland catchments, are considered to be the main reasons for the region's low coherence. Potential Evaporation (PE) is also likely

to be an important factor, with PE exceeding precipitation further south, and the reverse relationship further north.

## 3.2 Modified regions



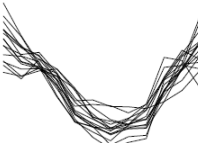
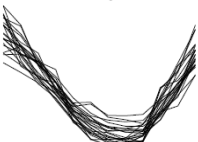
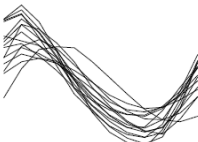
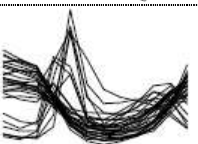

Investigation into the homogeneity of the Scandinavian clusters focused on attempting to revise the regions based on flow regimes. Table 3-1 shows that Region 21 and 22 have a mix of snowmelt- and rainfall-dominated regimes, so re-clustering attempted to separate out regions based on the two dominant processes. This did not yield any improvement in cluster homogeneity, so the original two regions were retained. However, some stations in the far north of Norway were removed because they are a long distance away from the remainder of the region, making the area very large, with by-far the highest inter-gauge distance, and covering an extensive climatological gradient.

Revision of the UK clusters focused on separating the large Eastern cluster. Firstly, this was separated into a North East and South East cluster, broadly following a distinction between previously-defined North East and Central homogeneous rainfall regions (a widely used scheme for classifying regional rainfall (e.g. Jones and Conway, 1997)). The two clusters thus derived were more homogeneous than the original cluster (both with a Cluster Homogeneity value of  $< 8\%$ ).

It was further decided to split the South Eastern Great Britain region into two subgroups on the basis of flow regimes. The rationale for this is that a high number of catchments in SE England are groundwater-dominated, and it likely these may respond differently to precipitation deficiencies, as compared to impermeable catchments. It would be expected that droughts would be slower-developing and longer. If the regimes are mixed, it is less likely that they will respond synchronously, resulting in lower RDI values and poorer cluster homogeneity. The region was therefore split into two based on the Base Flow Index (BFI, Gustard *et al.*, 1992) of catchments. An investigation was made into an appropriate BFI threshold, examining the distribution of BFI across the catchments. A threshold of 0.8 was chosen, to separate out those catchments which are generally dominated by baseflow; these catchments are all chalk catchments in southern and eastern England, so it is likely the regional response would be homogeneous. A lower BFI threshold would introduce catchments from other aquifers, and would also then require catchments from other clusters to be separated out. It should be noted that within the  $BFI < 0.8$  group, baseflow may still contribute a substantial component of the regime of any catchment – this should not be thought of as a group of ‘flashy’ catchments (average  $BFI = 0.43$ ).

The results in Table 3-2 show that the two South Eastern Great Britain (groundwater dominated and non-groundwater dominated) regions are homogeneous, with the homogeneity of the high baseflow region being particularly good (the lowest across all clusters). Although two distinct ‘sub-regions’ are defined, these should be thought of as an abstract classification rather than actual spatial entities. The BFI-based split cannot be used to create distinct geographical regions; hence, meteorological droughts are only studied for one South East Great Britain region.

**Table 3-2 Table of metrics for revised UK and Scandinavian regions**

	Region	Number of gauging stations	Inter-gauge distance (km)	Mean Hydrograph	Cluster homogeneity (%)	$P_c$	Correlation RDI and SPI
UK	NW Great Britain	33	159.8		7.3	0.162	0.52
	SW Great Britain	34	146.7		7.9	0.139	0.58
	NE Great Britain	16	151.9		7.6	0.157	0.58
	SE Great Britain non-groundwater dominated	23	132.3		7.7	0.132	0.63
	SE Great Britain groundwater dominated	14	130.4		6.8	0.230	0.69
Scandinavia	Southern Scandinavia	31	241.9		8.6	0.099	0.63
	NW Scandinavia	37	347.5		10.7	0.030	0.65

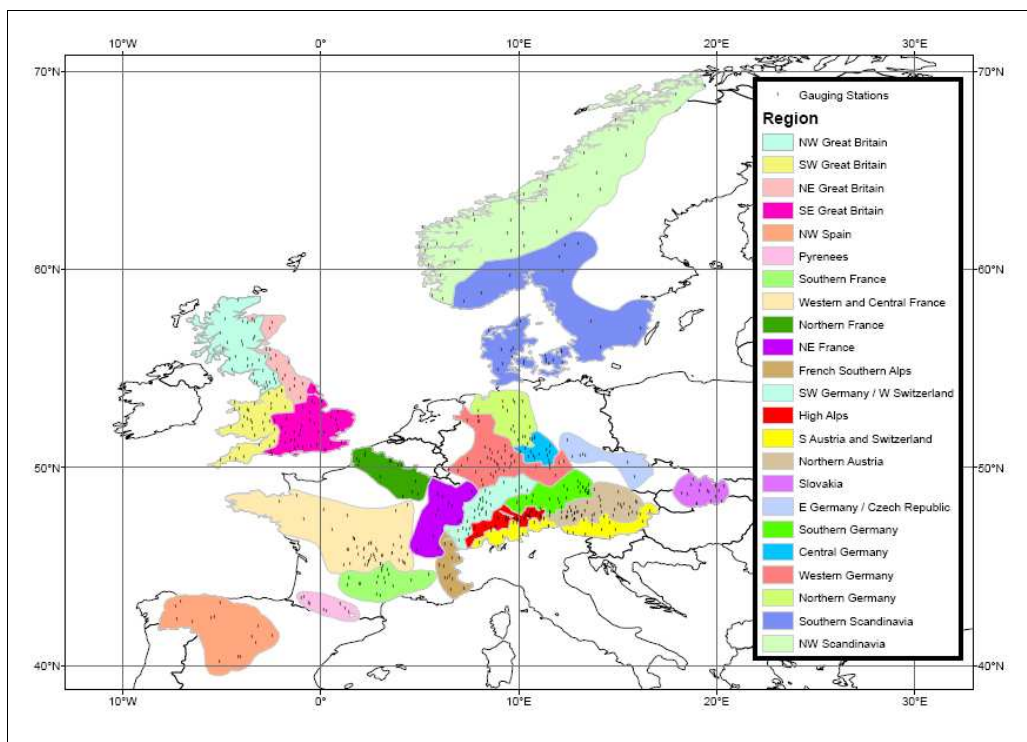
### 3.3 Final regions

Without a uniformly dense network of river gauges, establishing the boundaries of homogeneous Drought Regions is a difficult and somewhat arbitrary exercise. In a very strict sense, the RDI only provides information for the 579 European catchments considered, and not for any location outside these catchment boundaries. However, for any practical use, a wider definition of the regions is necessary. This procedure is explained in the next paragraphs.

The approach taken here is a compromise between objective criteria and manual correction. From the statistical clustering of the 579 DI series, 23 distinct groups

emerged (with an additional ‘groundwater-dominated’ sub-region, for one region of the UK). All regions show distinct geographical features that provided the basis for the final region boundaries.

Thiessen polygons were used to create boundaries between adjacent river gauges locations, and drawn around all 579 gauging stations. With no digital catchment boundaries available for the majority of the European river basins, this technique was considered to provide acceptable alternative to the use of hydrometric boundaries. Merging these polygons according to the homogeneous regions provided geographic extent of each drought region. A fixed buffer zone around each station ensured the majority of the regions showed unbroken areas. Where no data was available for some countries, it was decided to exclude that country from any of the final drought region. Despite providing artificial boundaries, following administrative limits rather than being-processed-based, this technique was preferred to avoid including some areas where the droughts regime could not be analysed within the framework of this project. Note that for some regions, the administrative boundaries also coincide with some hydrological boundaries (e.g. Southern limit of Region 14, South Austria and Switzerland). Final manual re-adjustment was necessary where the adjacent gauging stations were far apart.



**Figure 3-1 Final European Drought regions**

Figure 3-1 shows the final regions and their spatial extent. Note that these geographical regions only impact on the RSPI series, derived by superposing the 0.5° grids of the monthly rainfall on these region boundaries. The RDI series remained defined from the 579 hydrological series.



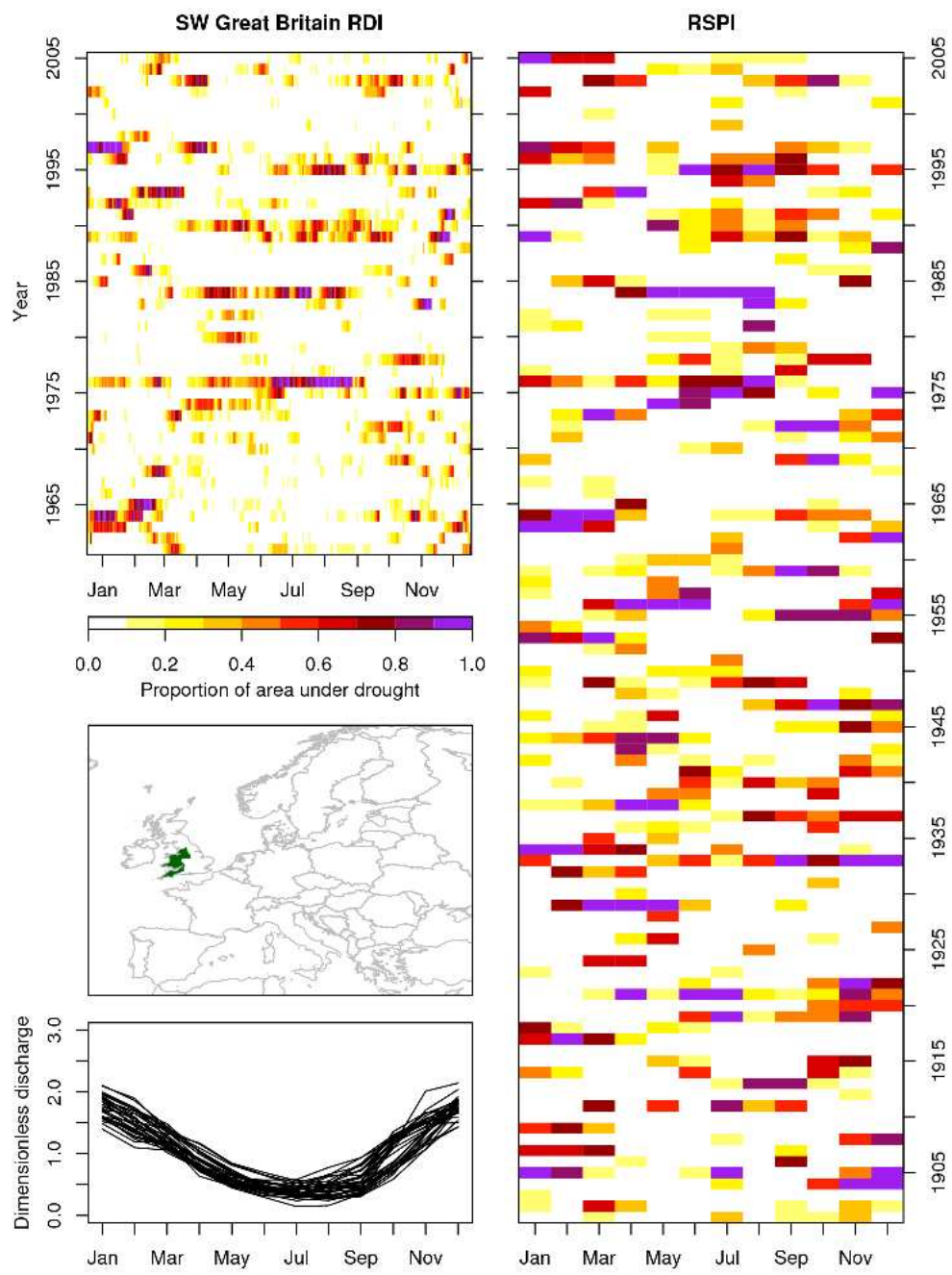
# 4 Drought Catalogues

Using the 24 regions described in Section 3, a Drought Catalogue was developed to present the key historical drought characteristics for each region. The catalogue is released as a standalone report (Lloyd-Hughes *et al.* 2009), so only a brief summary is presented here to illustrate the information in the catalogue pages. An example catalogue page is illustrated in Figure 4-1 and Figure 4-2, for South West Great Britain.

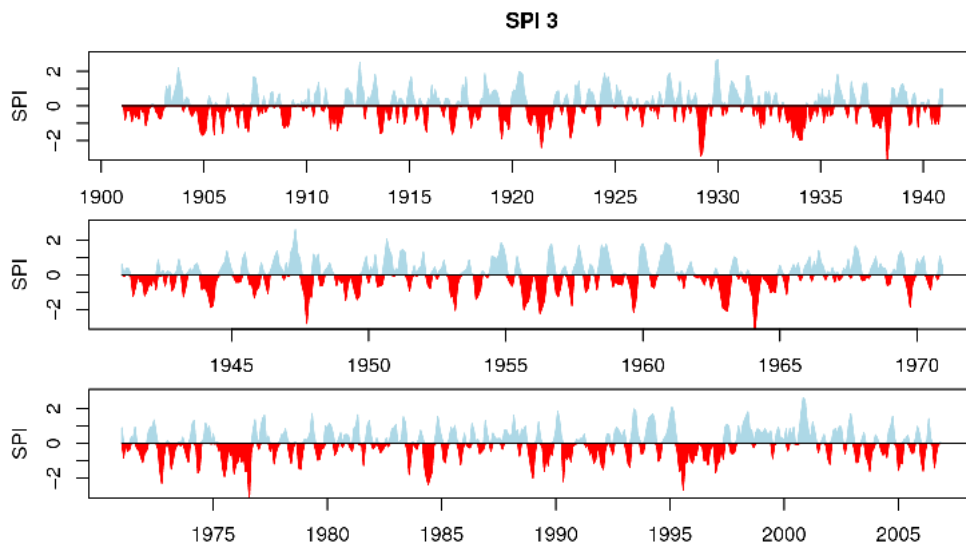
Data for each homogenous drought region are described over two pages in the catalogue. The first page provides a direct comparison of regional streamflow drought for the period 1961 to 2005 and the longer record meteorological indicator 1901-2005. Periods of coherent drought are easily picked out by blocks of colour. The darker the colour the more coherent is the drought across the region. Any potential lag between the hydrological response to the meteorological input is evident by a shift in the horizontal position of the coloured blocks. A map indicates the location of the region and a hydrograph plot (using a scaling by the mean to give dimensionless discharge) illustrates the flow regimes at each gauging station within the regional network.

The second page provides times series of the Standardized Precipitation Index. These compliment the regional plots by providing information on the severity of any spatially coherent events. Seasonality of the onset, duration, and spatial coherency of droughts within the region can be explored using circular plots at the bottom of the page. A drought beginning in February is represented as a line pointing to that month on the circle. The length of the radius indicates the coherency of the event where a radius of one, i.e. a line extending to the edge of the circle means that the whole of the region is affected. The colour of the line represents the duration of the event in months.

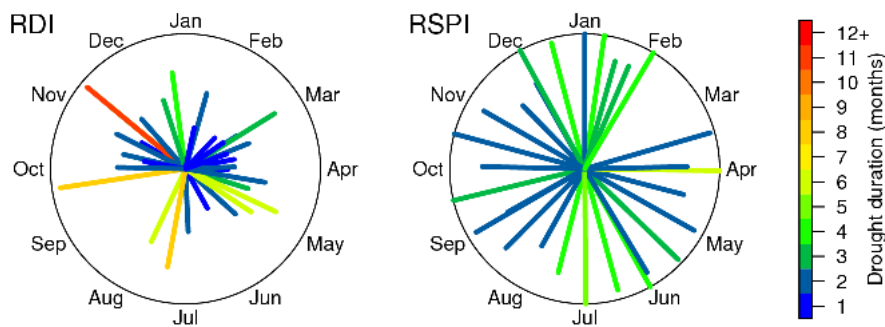
The catalogue entry is completed by a brief description of any special characteristics of this region such as particularly strong or weak correlations between stream flow and precipitation, suspected lack of homogeneity, or lack of data.



**Figure 4-1** Example of Page 1 of the Drought Catalogue for South West Great Britain



**Notes**  
 Regime: Precipitation-dominated, high flows in winter, low flows in summer.  
 The optimal meteorological proxy for streamflow drought is SPI 3 at a lag of 0 months with a rank correlation of 0.58.  
 Some long coherent hydrological droughts but the majority are less coherent and of shorter duration.  
 Multi-month hydrological droughts generally start from May to November.



Circular plots of drought onset and duration 1961–2005. Radii are proportional to the cluster area under drought. The unit circle is 100%.

**Figure 4-2 Example of Page 2 of the Drought Catalogue for South West Great Britain**

# 5 Appraisal of Historical Droughts

## 5.1 Rationale and approach

The aim of this component of the project was to describe the spatial coherence of rainfall and streamflow deficiencies during major historical droughts, to support the quantitative analysis carried out in the following sections – in particular, to explore whether there were any time-lags or patterns of antecedent conditions which may assist in the development of models predicting UK drought from those on the continent.

A particular objective of this work was to elucidate the onset and development of droughts on a European scale; this has rarely been done on a continental scale in Europe before, other than in the analyses carried out for two droughts by Zaidman *et al.* (2002). It was felt that this would therefore be a useful exercise, particularly if carried out for a number of major droughts in the recent past. By studying several events, recurring patterns could potentially be characterised; conversely, notable differences between the events may enable causative mechanisms to be isolated and used in the quantitative analyses.

There are clearly some patterns of ‘drought rich’ and ‘drought poor’ periods which can be seen across many regions (e.g. see Figure 6-5). From a qualitative assessment of those drought episodes which affected several European regions, as well as at least one UK region, the following drought periods were selected for detailed study:

**1962 – 1964; 1975 – 1976; 1988 – 1992; 1995 – 1997; 2003**

These generally correspond to episodes identified by previous authors (Marsh *et al.*, 2007a). The 2004 – 2006 drought, which had serious impacts in south eastern England, could not be studied as the streamflow data ended in 2005.

For each of the major events, a narrative of the spatial and temporal development of the drought is provided.

The following research questions were addressed:

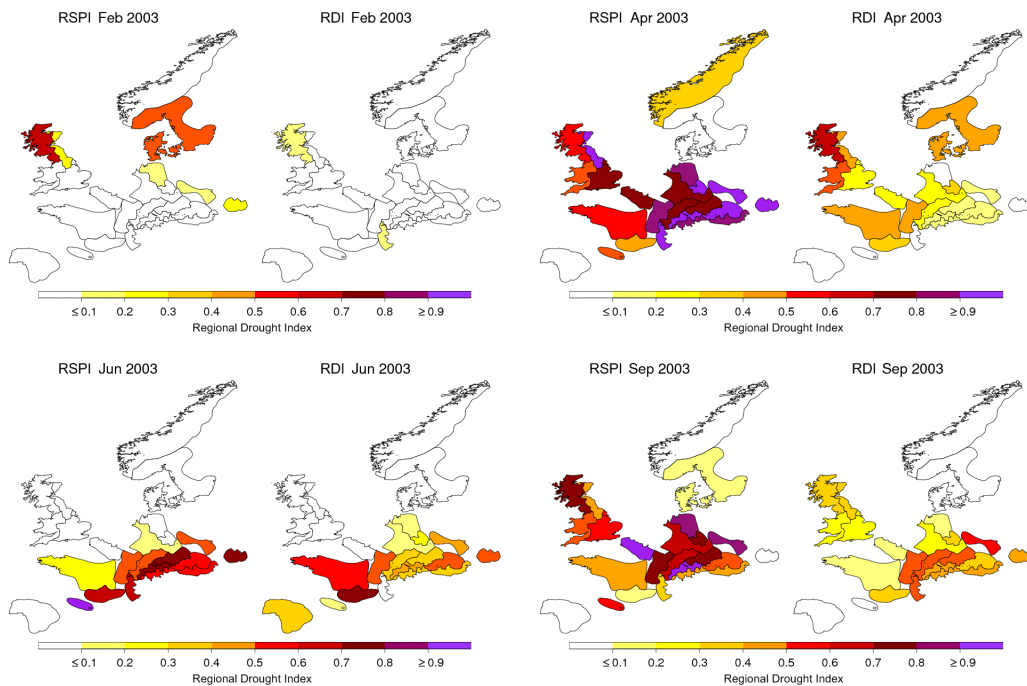
- How did the major droughts develop in time and space?
- Are there common patterns in drought onset and evolution, which occur between drought events?
- Is there any systematic time-lag between drought development in Europe and the UK, which could be used to inform the development of forecasting methods?

As a result of these analyses, a detailed 3-page drought summary was produced for each event. These are available as a separate publication (Parry *et al.* 2009), so only a brief overview is presented in this section. Sections (5.2 to 5.4) describe the methods used to elucidate drought development. Section 5.6 then reviews the main findings from the qualitative analyses.

## 5.2 Spatio-temporal development of major European droughts

To facilitate the description of the spatio-temporal evolution of major droughts, animations were built up of the RDI and RSPI developing through time. The animations were constructed using monthly data, so daily RDI had to be converted to a monthly index. These monthly indices were mapped for the 24 European regions, and can be iterated through sequentially to characterise temporal drought development.

This is a powerful tool for illustrating which regions tend to experience drought first, and the subsequent intensification, spread and decay of the drought. The maps also allow visualisation of the extent to which in the UK is coherent with the rest of Europe. For example, the 2003 drought is known to be a spatially extensive and severe drought across much of continental Europe. Whilst the summer drought was notable in GB, the animations reveal that the UK was only partly coherent with Europe in some phases of the drought; selected screen shots from the event are shown in Figure 5-1.



**Figure 5-1: Selected screenshots showing phases of the 2003 summer drought in the UK and Europe**

Parts of the UK experienced deficiencies first (see February map), and in April, GB was coherent with the rest of Europe. In May – July, the focus shifted to central Europe, and the UK experienced a comparatively damp interlude. Peak deficiencies in the UK occurred in Aug – Oct, with the European drought continuing. The cessation of the drought on the continent occurred in October, but deficiencies continued to persist in some UK regions until December.

### 5.3 Drought evolution in North West Europe

Whilst the pan-European animations allow the full extent of the study area to be considered, the amount of information presented is overwhelming. RDI and RSPI values are shown for 24 regions for each month, which is a vast amount of information over a four-year drought; this makes interpretation complex, and hampers the comparison between major drought episodes.

An alternative approach was developed, to summarise the evolution of a drought in a simple matrix showing the monthly RDI in each region (representing a row) for each month (representing a column). To simplify even further, this was carried out for the subset of regions closest to the UK, which exhibited the strongest statistical relationships in the model fitting (discussed in Section 7.3). Once the matrices are constructed, a simple shading scheme can be applied to highlight RDI values above pre-defined thresholds. Examples are shown in Figure 5-2 and Figure 5-3, for the 1975 – 1976 and 1995 – 1997 droughts.

In 1975/76, the drought started as a winter drought in the UK and Scandinavia and then spread to the continental interior; firstly to France before becoming persistent in Germany. Clearly, the drought is spatially coherent over Northern Europe, with high RDI values across the majority of NW Europe at the height of the summer drought, and long sequences of RDI deficits were observed in the summer for all regions except the Pyrenees. The drought ended abruptly in the UK in the autumn, but persisted in Europe, particularly Northern France and Germany.

In contrast, 1995 – 1997 is much more complex, with several distinct phases. The UK experienced an intense summer drought in 1995, which had no equivalent effect in NE Europe. In the winter of 1995/96, streamflow deficiencies were observed in the UK, but were not coherent; more coherent deficits were observed in Denmark and Germany. A summer drought developed in 1996, but this was generally not coherent. Following winter/spring deficiencies in the UK and France in 1996/1997, a summer drought then developed in the UK (particularly in SE groundwater catchments), but this was not observed elsewhere in Europe. Unlike in 1975/1976, there was no consistent pattern of evolution (with drought growing and then receding from various centres of action rather than following a west-east transition) and at no time were there coherent deficits across NW Europe. Nevertheless, the various phases had major impacts in the UK and other NW European regions, so whilst spatial coherence was less prominent, the drought as a whole was still of regional significance.

Month		1975												1976												1977
		APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	JAN			
UK	NORTH WEST			0.21	0.24																					
	SOUTH WEST			0.29	0.21																					
	NORTH EAST							0.21	0.29	0.42																
	SOUTH EAST								0.27	0.59									0.50	0.68	0.28					
	SE (Groundwater)								0.22	0.44									0.57	0.55	0.71	0.74	0.70	0.75	0.76	0.78
FRANCE	PYRENEES																									
	SOUTH																									
	WEST & CENTRAL																									
	NORTH																									
GERMANY	NORTH EAST																									
	SW GER & SWITZ																									
NORDIC	CENTRAL																									
	WEST																									
	NORTH																									
	S. SCANDINAVIA																									

Figure 5-2: Matrix showing the evolution of RDI deficiencies in the 1975/76 drought

		1995												1996												1997										
Month	Season	5	6	7	8	9	10	11	12	1	2	3	4	5	6	7	8	9	10	11	12	1	2	3	4	5	6	7	8	9	10	11				
UK	NORTH WEST				0.7	0.3		0.5	0.3	0.3	0.2	0.2						0.6														0.3	0.2			
	SOUTH WEST				0.3	0.7	0.3	0.2	0.3	0.4		0.2																								
	NORTH EAST				0.5	0.9			0.2								0.3	0.3	0.5	0.3													0.3	0.2		
	SOUTH EAST				0.2	0.4			0.2	0.2							0.3	0.4	0.2	0.4	0.4		0.2										0.2	0.3	0.3	
	SE (Groundwater)								0.2	0.2					0.2	0.2	0.2	0.2	0.3														0.5	0.5	0.4	
FRANCE	PYRENEES					0.2																												0.2		
	SOUTH																																	0.2		
	WEST & CENTRAL																																	0.2		
	NORTH								0.3	0.3																								0.2		
GERMANY	SW GER & SWITZ								0.3	0.3																								0.2		
	CENTRAL								0.3	0.3																								0.2		
	WEST								0.3	0.3																								0.2		
	NORTH								0.2	0.2																								0.2		
NORDIC	S. SCANDINAVIA					0.2		0.2	0.6																									0.3		

Figure 5-3: Matrix showing the evolution of RDI deficiencies in the 1995 - 1997 drought

## 5.4 Summaries of synoptic conditions

A further stage of the qualitative work was to examine the synoptic conditions associated with the major droughts. This was carried out to establish whether any commonalities could be observed in the synoptic conditions prevailing during (and before) the major droughts, which could potentially be of use in explaining spatial coherence (see also Section 8). Temperature and pressure anomalies were plotted for the major events, using gridded NCEP/NCAR reanalysis data for Europe. An example of a narrative of the synoptic conditions for 1962 – 1965 is provided below; Figure 5-4 shows the associated plots of temperature and pressure anomalies. Important modes of atmospheric circulation (e.g. the North Atlantic Oscillation) were also plotted for the historical drought period – more detailed analysis of the relationships between RDI, RSPI and teleconnections such as the NAO is given in Section 8.

*“This period began with the spectacularly cold winter of 1962/63 which is the 2nd coldest on record. Anomalously high pressure developed over the North Atlantic and persisted for much of the next two years. This pushed the NAO into a negative phase, blocking the northerly storm track, and directed rain bearing systems south in to the Mediterranean. The high pressure moved east in the early winter of 1964 further suppressing rainfall over central Europe. The East Atlantic / West Russia component of the atmospheric circulation remained on the negative side of neutral throughout the core years of this drought. This would have reduced the severity of winter droughts over southern Europe.”*

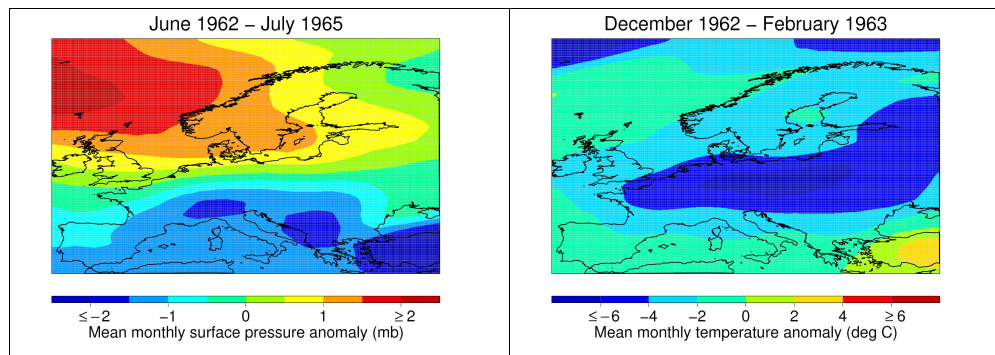


Figure 5-4: Pressure anomalies associated with 1962 – 1965 drought, and temperature anomalies associated with the cold winter of 1962/1963

## 5.5 Meteorological Droughts before 1960

The use of parallel hydrological and meteorological drought indices has allowed an extension of the analysis of major droughts in Europe back to the beginning of the twentieth century. Although the first half of the twentieth century can only be investigated through the meteorological index, the close agreement between the RDI and the RSPI from 1961 to 2005 suggests that the latter can be used independently with a fair degree of confidence.

The drought of 1920/21 is one example of a major event that has been detected using the RSPI meteorological index independently. An initial three-month period of deficient rainfall totals to end 1920, which predominantly affected northern and eastern regions of Europe severely, was immediately followed by a distinct abrupt shift in focus on the continent. The previously unaffected Iberian Peninsula and southern and western France emerged in January 1921 as the source region for a persistent year-long period of extensive and severe rainfall deficiencies. Throughout 1921, meteorological drought was very spatially coherent, exhibiting the most significant levels of deficits over a large proportion of mainland Europe. Following an initial phase featuring drought conditions concentrated on the south-western half of Europe during the first quarter of the year, precipitation deficiencies migrated gradually northwards and westwards across the UK, France, southern Germany and the Alps in a diagonal northwest-southeast band. The spatial extent of the most severe rainfall deficits decreased during late summer and autumn, but drought conditions were still present across a large part of Europe by year-end. Similar slow migrations have been observed during other major drought episodes, although they vary in their source area, pace and direction of evolution, and their termination location.

## 5.6 Observations on Spatial Coherence of Major droughts

Following these qualitative exploratory analyses, some remarks on the spatial coherence of major historical droughts can be made:

- Whilst there are some broad patterns between the major droughts, each tends to have a distinctive spatial signature; there are common elements associated with certain periods of each drought but, taken in entirety, there are few commonalities which recur in the major droughts.
- This feature of European droughts partly is likely to present challenges in identifying robust statistical relationships, particularly given the relatively short records involved (yielding a sample of only six major droughts).
- Often the UK is one of the first regions to experience drought, or experiences drought simultaneously with other parts of northwest Europe. This implies that there may be very little lead time between onset of droughts in Europe and the UK, which may confound the development of forecasting tools based on drought development elsewhere.
- The west – east migration seen in some major events is potentially a useful phenomenon for assisting in the development of improved monitoring and forecasting on a European scale. Whilst this is a feature seen in some phases of the major droughts, it is certainly not universal – so caution should be exercised in assuming that a drought developing in western Europe will spread



east (for example, 1984 was a severe summer streamflow drought in the UK, but did not express itself on the continent).

- Only the 1975/1976 drought is coherent on a pan-European scale for a persistent period. During other droughts, there are often short spatially coherent phases within droughts, the centres of action of which tend to shift in space.
- Some major UK droughts do not appear to have had any equivalent impact in Europe – for example, the summer droughts in 1984 and 1995. In contrast, there are some droughts which manifest themselves over a wide area in continental Europe, but are not expressed in the UK – for example in late 1971/1972, when a drought occurred across most of mainland Europe, but had very limited expression in the UK.
- During most long droughts there are distinct periods when the UK is in-phase and out-of-phase with Europe; even in intense summer droughts seen on a large scale over Europe, the UK is not necessarily in-phase with the continent.
- Some long droughts (1962 – 1964; 1995 – 1997; 1988 – 1992) result from a combination of both winter and summer deficiencies. The evolution of these events is very complex, in comparison with the major short-duration summer droughts, and is likely to escape classification by a simple index. Whilst there are often phases of intense summer drought, in these events the winter conditions are as important for dictating the overall deficiencies.
- Reviews of the synoptic situation associated with the drought reveal that they are all generally associated with major pressure anomalies, but that these differ substantially in terms of their intensity and location.
- All the droughts (except 2003) were associated with anomalies in the large-scale atmospheric circulation but, again, the characteristics of these anomalies varied significantly from event to event.

# 6 Exploratory Analysis

This section of the report provides a summary of the exploration of relationships within the data, which was carried out to identify patterns which may inform the development of forecasting methods.

Sections 6.1 to 6.3 focus on dependence between three different types of data sets; Regional Deficiency Index (RDI) data in different regions; Regional Standardised Precipitation Index (RSPI) data in different regions; and finally lagged relationships in both the RDI and RSPI in different regions. Section 6.4 provides a brief overview of relationships with temperature data, which was carried out as temperature was thought to be an important driver of drought which may confound simple relationships in the RDI and RSPI.

## 6.1 Relationships in flow data

### 6.1.1 Raw RDI data

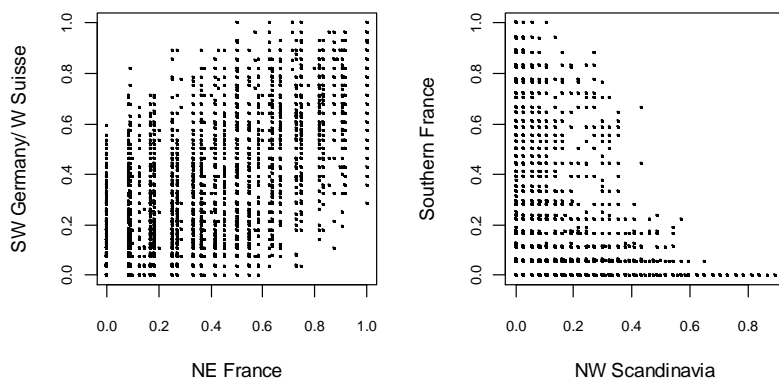
The first relationships examined were rank correlations between the raw RDI data for different regions. The analysis is based on rank correlations because the data do not follow a normal distribution. Table 6-1 gives the highest 4 correlations for each region.

As might be expected, in general, each region is most highly correlated with the other regions that are closest to it. The highest pairwise correlation is between regions 11 and 13 western France and south-west Germany/ western Switzerland and has a value of 0.70. In addition to the raw correlation measure, the pairwise scatter plots were also examined. Figure 6-1 shows two such plots. Even for the region pair with highest rank correlation (top left) the amount of scatter is very high. This shows that any model fitted to this relationship is likely to have a large amount of uncertainty associated with it. For this reason no statistical models were fitted to the raw data.

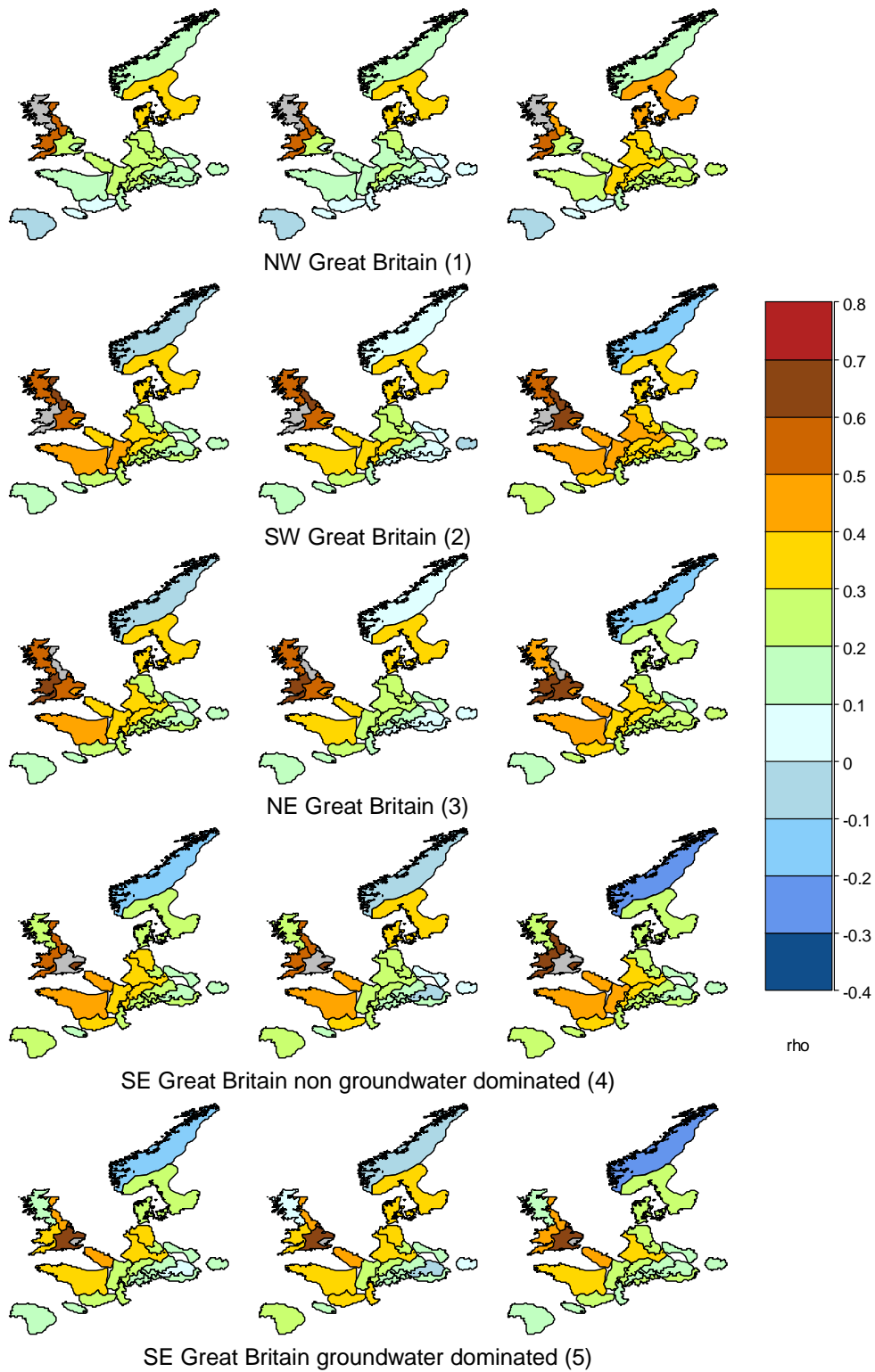
Figure 6-2 shows the correlations for all other regions with each individual UK region. In addition to showing the rank correlations for all data, the rank correlations are shown for just two seasons: summer (1<sup>st</sup> May-31<sup>st</sup> October) and winter (1<sup>st</sup> November to 30<sup>th</sup> April). These seasons correspond to the accepted recharge period for river flows (Marsh *et al.* 2007). Figure 6-2 seems to suggest that there is very little difference in correlation in the winter and the summer months. For Regions 1, 2, and 3 (NW, SW, and NE Great Britain) the correlations appear to be slightly higher in the winter months than in the summer months, however for Regions 4 and 5, the two SE Great Britain regions, the correlations seem to be almost identical with the exception of the correlations with Region 24, North West Scandinavia. The correlations with this region appear to be more strongly negative in the winter months than in the summer months. It is possible that this is due to correlations with the East Atlantic/West Russia circulation pattern (see Section 8).

**Table 6-1: Highest 4 correlations for each region, correlation measured using Spearman's  $\rho$  correlation coefficient. Different colours reflect different geographic regions; brown UK; blue France and Spain; green alps; pink Germany and further east and orange Scandinavia.**

Region	Highest		2 <sup>nd</sup> highest		3 <sup>rd</sup> highest		4 <sup>th</sup> highest	
	region	rho	region	rho	region	rho	region	rho
NW Great Britain (1)	2	0.55	3	0.51	23	0.40	4	0.29
SW Great Britain (2)	3	0.63	4	0.60	1	0.55	9	0.41
NE Great Britain (3)	2	0.63	4	0.58	1	0.51	5	0.44
SE GB non GW (4)	5	0.62	2	0.60	3	0.58	9	0.47
SE GB GW (5)	4	0.62	3	0.44	10	0.42	2	0.40
NW Spain (6)	8	0.44	7	0.35	9	0.31	4	0.20
Pyrenees (7)	8	0.37	6	0.35	9	0.34	12	0.31
S France (8)	9	0.63	6	0.44	7	0.37	4	0.34
W & C France (9)	8	0.63	11	0.54	13	0.50	4	0.47
N France (10)	21	0.61	11	0.54	13	0.51	22	0.50
NE France (11)	13	0.70	21	0.58	9	0.54	10	0.54
French S Alps (12)	15	0.36	13	0.35	9	0.35	11	0.34
SW Germany W Switzerland (13)	11	0.70	21	0.62	14	0.59	19	0.55
High Alps (14)	16	0.60	13	0.59	19	0.58	15	0.52
S Austria & Switzerland (15)	16	0.52	14	0.52	13	0.46	19	0.43
N Austria (16)	14	0.60	19	0.57	15	0.52	17	0.47
Slovakia (17)	16	0.47	15	0.41	19	0.40	18	0.38
E Germany & Czech Republic (18)	19	0.55	20	0.53	21	0.49	13	0.45
S Germany (19)	21	0.58	14	0.58	16	0.57	13	0.55
C Germany (20)	21	0.66	22	0.61	13	0.54	18	0.53
W Germany (21)	22	0.68	20	0.66	13	0.62	10	0.61
N Germany (22)	21	0.68	20	0.61	10	0.50	13	0.47
S Scandinavia (23)	22	0.44	21	0.41	1	0.40	10	0.39
NW Scandinavia (24)	23	0.28	1	0.19	22	0.06	18	0.02



**Figure 6-1: Pairwise scatter plots of RDI for region pairs with highest correlation (left) and lowest correlation (right).**



**Figure 6-2: Rank correlations with RDI for UK regions, left all data, middle, summer months, right, winter months**

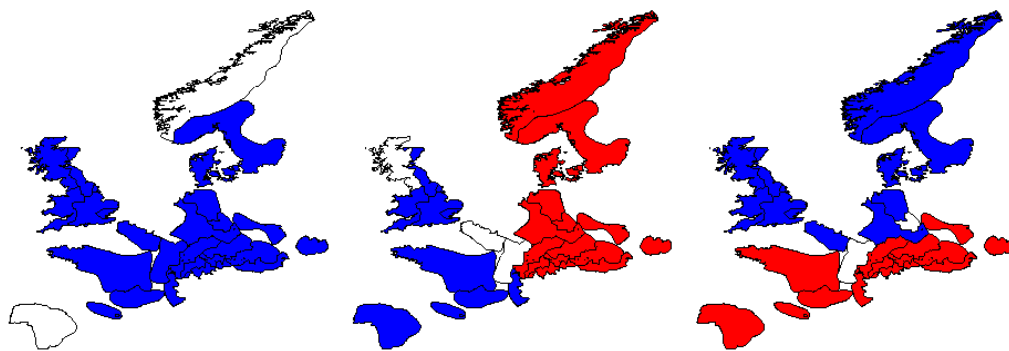
Although the correlations between the RDI values for different regions are relatively low, it may still be possible to use them to extract information about further groupings in the data. The technique used is a non-parametric technique called multi dimensional scaling (MDS). This is equivalent to a non parametric version of Principal Component Analysis (PCA). The only difference between MDS and PCA is that PCA uses Pearson's  $\rho$  as a measure of correlations between a pair of variables (here RDI for two regions) whereas MDS uses a rank correlation measure. Here Spearman's  $\rho$  is used.

The aim of both MDS and PCA is to find weighted sums of the original individual RDI variables that explain the maximum amount of variation in the data. So, here, each Principal Component (PC) is a weighted sum, which has the form

$$PC = W_1 \times RDI_1 + \dots + W_{24} \times RDI_{24}$$

where the weights,  $W_i$ , can be either positive or negative. If all the weights for a particular PC are positive or all negative then that PC shows a group of RDI variables that all act together – so tend to experience droughts at the same time. If some are positive and some negative then that PC identifies two groups of regions that tend to experience droughts at different times. It is possible to calculate how much of the total variation in the data each PC explains. For instance it is possible that a group of RDI regions may all be very highly correlated and so always experience a drought at the same time. In this situation there is likely to be a PC where all the weights are of the same sign that explains almost all the variation in the whole data set.

Strictly speaking PCA (and so MDS) is only valid for use with data that follow a multivariate normal distribution, which the RDI data does not. The techniques can still be useful when this assumption is invalid, but the results obtained should be used with caution. In particular the contrast groupings obtained are reasonably safe and informative, whereas the values of percentage variation in the data explained by each Principal Component should only be used as a very rough guide. Figure 6-3 shows groupings for the first three PCs. PCs 1-3 explain a sensible amount of variation in the data set, whereas the later PCs only explains a small amount of variation. The main grouping in the data comprises all regions apart from NW Spain and NW Scandinavia: this shows that most of Europe tends to experience some form of drought at the same time. The second grouping is a contrast between western Europe and eastern Europe and the third a contrast between northern and southern Europe.

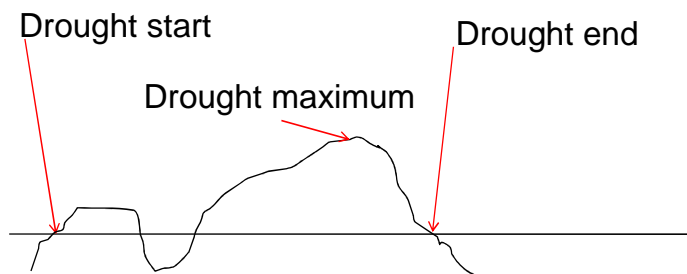


**Figure 6-3: Groupings for first three principal components for daily RDI data. Left PC1, middle PC2, right PC3. Approximate percentage variation explained for each PC – PC1 33%, PC2 11%, PC3 9%.**

### 6.1.2 Drought periods

Relationships between RDI in drought periods have been examined. To do this a drought period is defined and used to identify independent droughts for each region. The method used is the runs declustering method of Smith and Weissman (1994). In this context declustering refers to extracting independent 'clusters in time' of extreme high or low values. Independent droughts are defined as independent 'clusters in time' of high RDI values. Figure 6-4 illustrates this method. Droughts are defined to be independent if they are separated by a certain length of time (here 84 days, approximately 3 months). The start of a drought is defined as the first day the RDI is above the threshold, and the end of the drought is the last day it is above the threshold. Any time the RDI dips below the threshold for a shorter period of time than 84 days is still classed as a drought period.

The approach taken here is similar to the approach used to define independent floods when extracting a peaks-over-threshold (POT) record (NERC, 1975, Robson and Reed, 1999). The main difference between the method used here, and the method used in extracting POT data is that we do not include a lower threshold to identify independent droughts.



**Figure 6-4: Definition of drought periods**

Figure 6-5 shows times for which each region is in a drought period. Three aspects of the data can be seen clearly. First, there are periods of time (e.g. 1976) when most regions experience a drought at the same time. Secondly some regions are more prone to longer droughts than other. For instance compare Region 1, North West Britain and Region 5, groundwater dominated south East Britain. Region 1 has had a large number of drought which each have a short duration, whereas Region 5 has had many fewer droughts, but of longer duration. Finally there seems to be some clustering in drought occurrence, in particular the late 1960s and early 1980s seem to be particularly drought poor periods whereas the early-mid 1970s and 1990s seem to be drought rich periods.

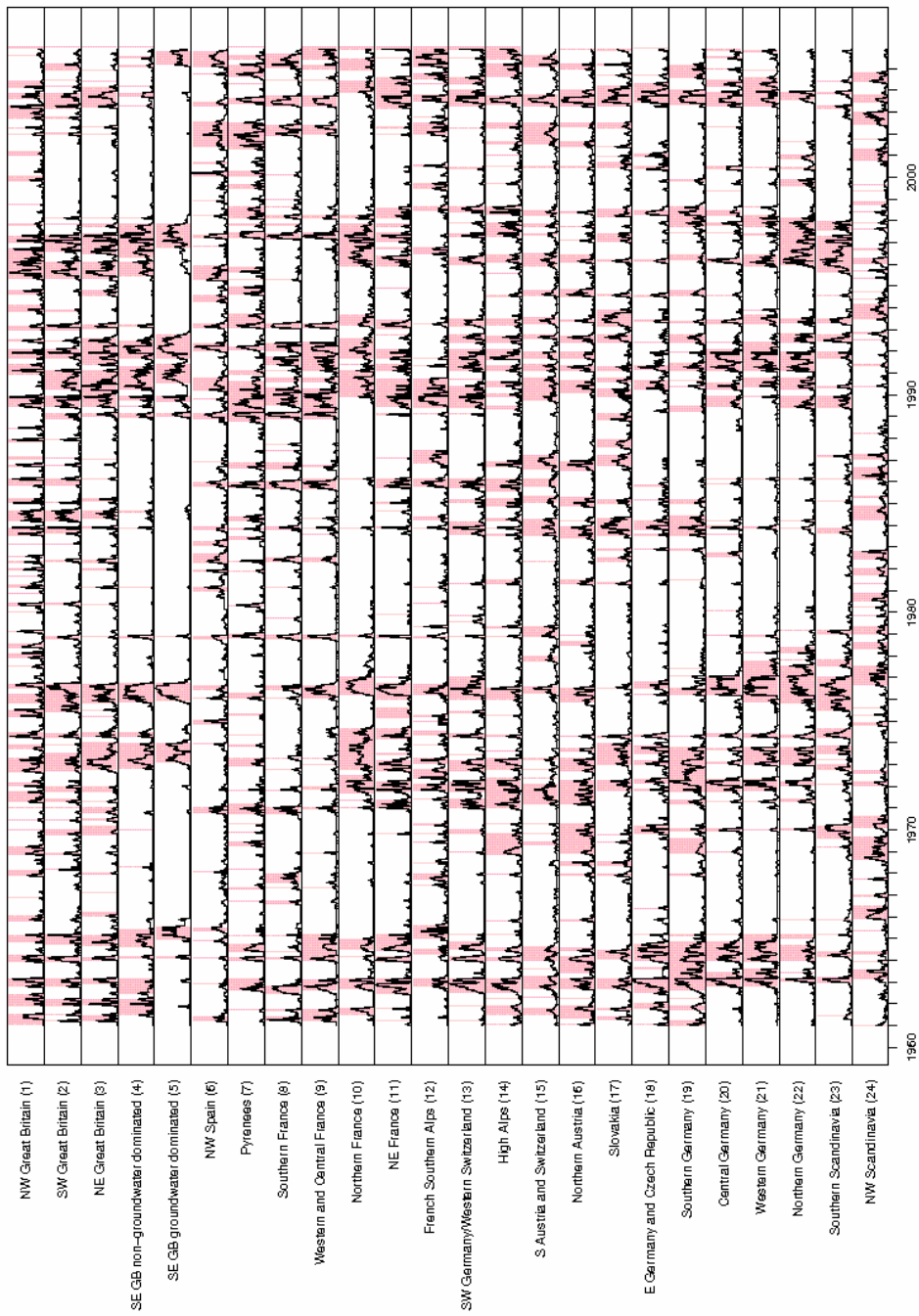


Figure 6-5: Drought periods for all regions

To see if the number of other regions experiencing a drought has any effect on the length of a drought in a particular UK region, two correlation indices were calculated:

- the correlations between the number of other regions experiencing a drought in the 28 days prior to the start of a drought, with the drought length for each UK region
- the correlations between the number of other regions experiencing a drought in the 56 days (i.e. the 28 days prior and 28 days after) surrounding the start day of each drought with the length of each drought for each UK region.

The correlations are all positive but fairly low (Table 6-2), but high enough to be included as covariate in a model to predict drought length.

**Table 6-2: Correlations between length of drought in each UK region and number of other regions experiencing a drought (including other UK regions).**

	Region 1 (NW GB)	Region 2 (SW GB)	Region 3 (NE GB)	Region 4 (SE GB non-GW)	Region 5 (SE GB GW)
28 days prior	0.146	0.156	0.083	0.104	0.087
56 days surrounding	0.195	0.14	0.172	0.132	0.037

## 6.2 Precipitation

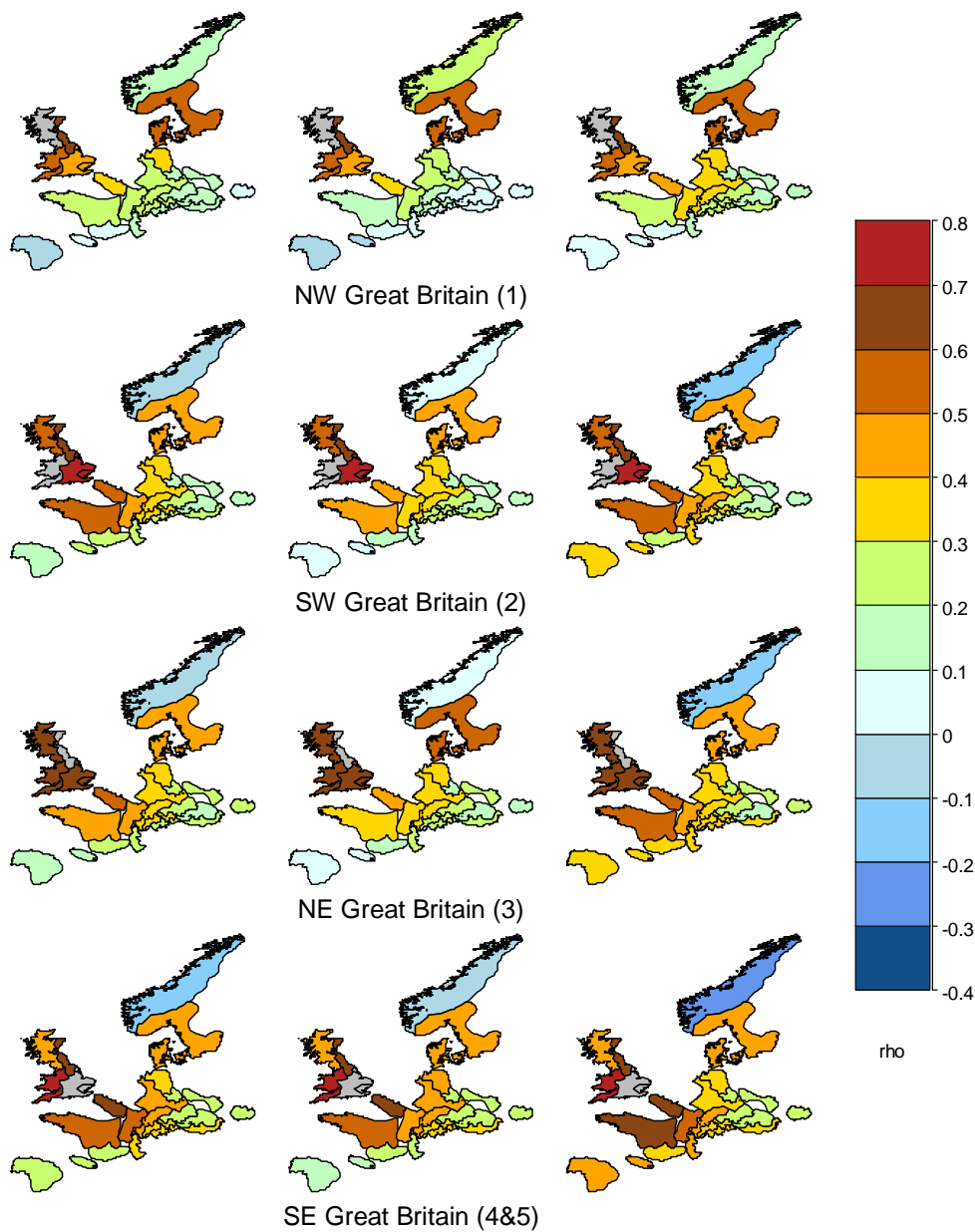
In this section the correlations between observed precipitation for the regions are examined, looking at the monthly RSPI values (RSPI\_1).

Figure 6-6 shows the correlations for all other regions with each individual UK precipitation region, and show slightly higher values than the correlation in streamflow (Figure 6-2), possibly because the percentage of rainwater reaching the rivers vary between regions due to factors such as seasonal soil moisture deficit (i.e. making the regional less coherent in terms of rainfall deficit).

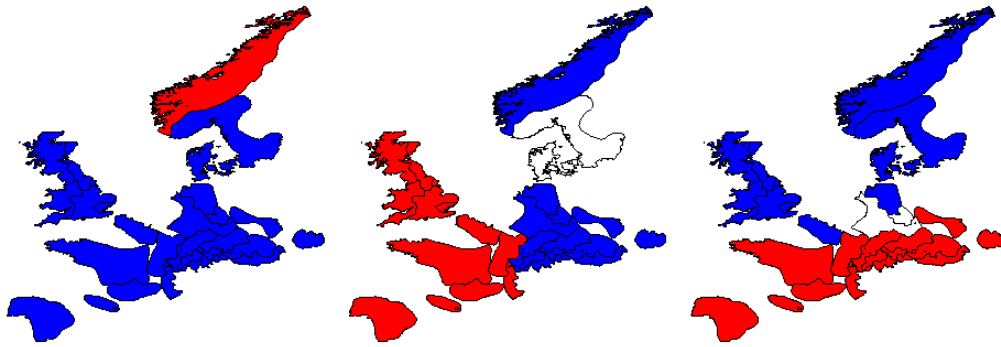
Summer and winter correlation are more similar than for streamflow, but the negative winter correlation between Region 4/5 and Region 24 still appears (Figure 6-6).

The Multi Dimensional Scaling procedure on the RSPI data showed very similar groupings for the first three PCs that found for RDI (Figure 6-7: RSPI; Figure 6-3). The main difference is that in PC1 Region 24, only NW Scandinavia appears to contrast with the rest of Europe (northern Spain was also contrasting with the rest of Europe for RDI).





**Figure 6-6: Correlations with RSPI\_1 for UK regions, left all data, middle, summer months, right, winter months**

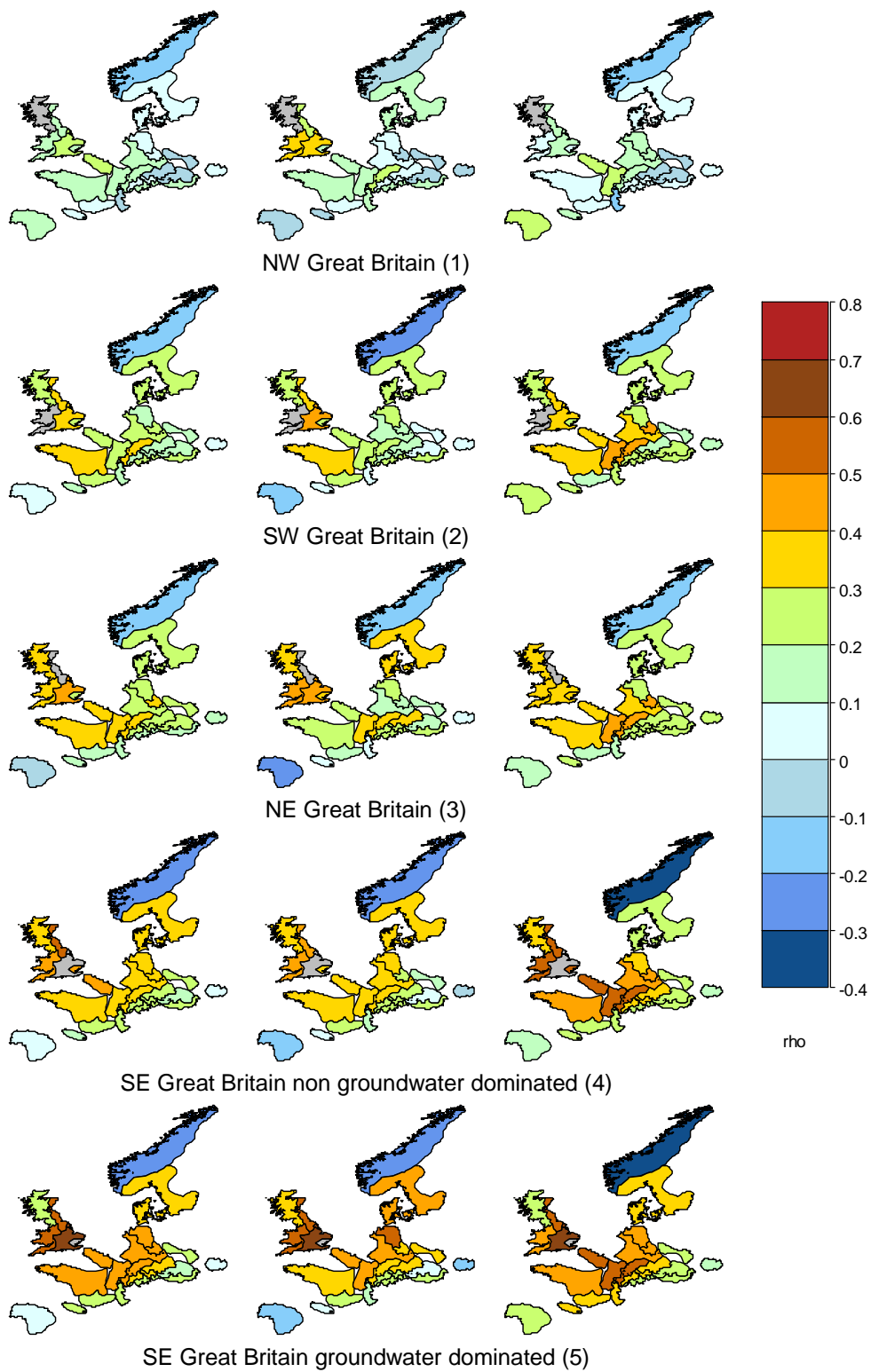


**Figure 6-7: Groupings for first three principal components for monthly RSPI data. Left PC1, middle PC2, right PC3. Approximate percentage variation explained for each PC – PC1 38%, PC2 13%, PC3 10%.**

### 6.3 RDI Vs RSPI and lagged relationships

In this section, the lagged correlations between RDI in different regions, and the lagged correlations between RDI and RSPI for each UK region, are described. In this analysis we examine the average RDI and average monthly RSPI for whole seasons.

Figure 6-8 shows lagged RDI-RDI correlations. Region 1 has a much lower general level of lagged correlation than the other UK regions, possibly illustrating that the catchments in this region respond much quicker to rainfall than those in the other UK regions (Keef *et al.*, 2009). Apart from Region 5, the lagged correlations are generally lower than the simultaneous correlations (Figure 6-8; cf. Figure 6-2). The difference in correlation between summer and winter is relatively small, suggesting that it will be unnecessary to separate any further analysis into winter and summer.



**Figure 6-8: Lagged correlations with RDI for UK regions. Left plots are correlations using all 6 month seasons with previous season, middle plots are all**

**summer seasons with the previous winter season; right plots are all winter seasons with the previous summer season.**

Table 6-3 shows the rank correlations between seasonal RDI and seasonal RSPI\_1 for the five UK regions. The differences between summer and winter correlations are relatively small. Moreover, the differences in correlation at lags 1 and 2 seem to reflect the differences in catchment response time, faster responding catchments have little memory and so the impact of the previous 6 months precipitation deficit is lower. The regions with the slower responding catchments (4&5) have higher levels of lagged correlation. The region with the quickest responding catchments also has the lowest level of correlation at lag 0. This could be due to the very short length of droughts in this region. By averaging RDI over a 6 month period we are likely to be averaging drought periods and non drought periods. This averaging will reduce the signal of the seasonal RDI.

**Table 6-3: Lagged seasonal Spearman’s rank correlations between seasonal RDI and previous seasonal RSPI\_1 for UK regions.**

		Lag 0	Lag 1	Lag 2
Region 1	All	0.604	0.150	0.144
	Summer	0.657	0.086	0.202
	Winter	0.622	0.212	0.095
Region 2	All	0.727	0.228	0.205
	Summer	0.766	0.210	0.308
	Winter	0.730	0.225	0.101
Region 3	All	0.705	0.355	0.247
	Summer	0.735	0.308	0.361
	Winter	0.714	0.345	0.172
Region 4	All	0.785	0.561	0.362
	Summer	0.809	0.665	0.337
	Winter	0.762	0.478	0.418
Region 5	All	0.750	0.613	0.503
	Summer	0.700	0.564	0.484
	Winter	0.818	0.664	0.530

## 6.4 Temperature

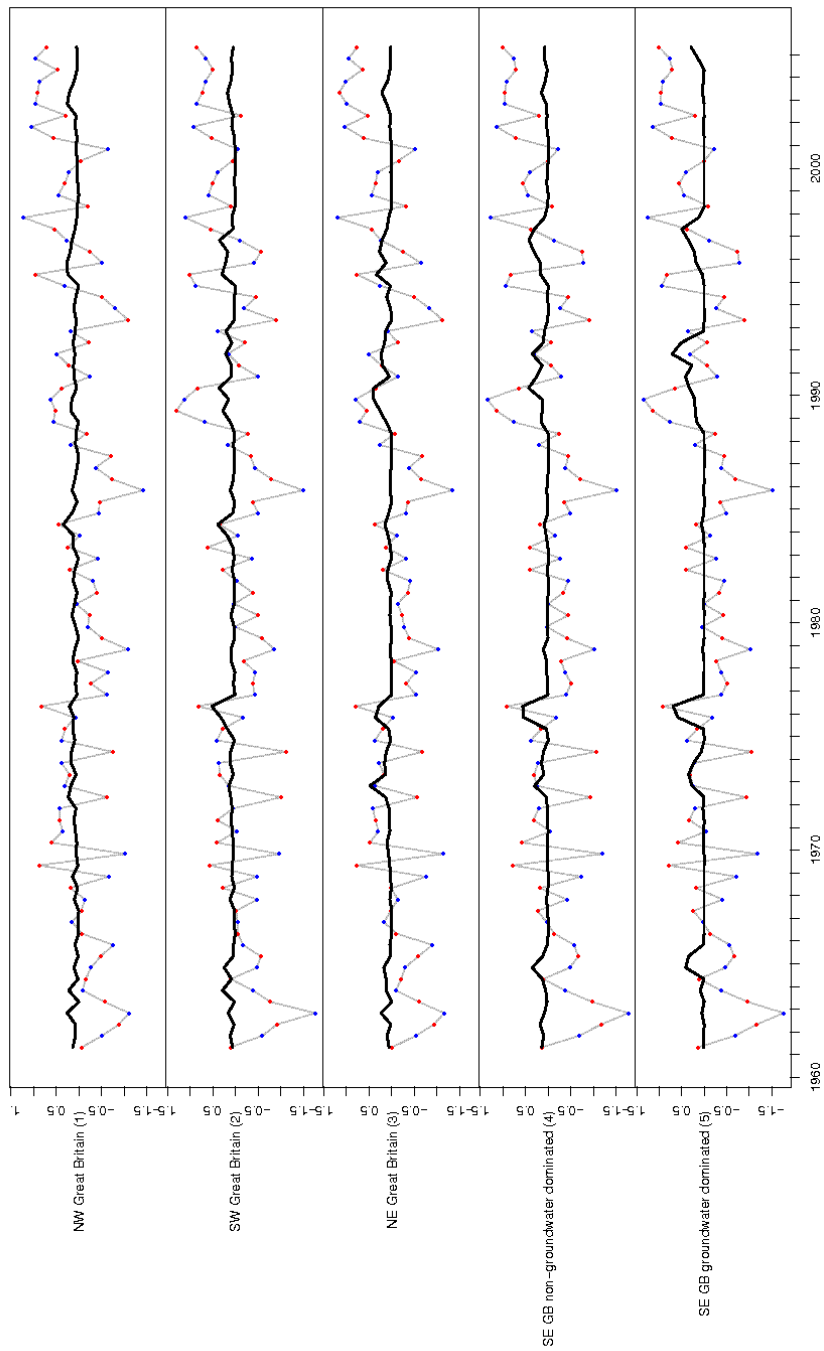
The development of hydrological droughts can be accelerated by high temperatures, which increase the amount of rainfall that is lost through evaporation from the soil. In addition, this loss must be replenished before rainfall can generate river flows, lengthening the drought. So a simple improvement to any statistical models that attempt to predict the length of droughts may be to include a covariate that describes the temperature.

The daily mean temperature data used is the NCAR/NCEP reanalysis data (freely downloadable from <http://www.cdc.noaa.gov>, at a 2.5 degree grid). The period of record used is 1<sup>st</sup> January 1961 to 31<sup>st</sup> December 2005 to match the period of record of RDI. For each drought region, the daily regional temperature time series is, for each day, the average of the temperature for each grid point of the region. Seasonality was accounted for using the moving average technique, similarly to RDI, using a time window of +/- 10 days around the day of interest. Daily regional anomaly temperature

time series were derived as the difference between the daily regional temperature and the moving average temperature for that day.

Spearman's rank correlation coefficient was used to evaluate if including temperature information could improve drought predictability. This has two advantages over using Pearson's correlation coefficient; firstly, it is valid for data that is not normally distributed; secondly, because it uses the ranks of the data it is less sensitive to possible measurement or averaging errors in the data. The rank correlations were calculated between seasonal average temperature anomalies and seasonal average RDI for each region. The analysis was carried out separately for summers and winters because the weather systems associated with low rainfall (anti-cyclonic conditions) are associated with colder winters and hotter summers. Results showed that generally hot summers were associated with high RDI and cold winters were associated with high RDI particularly in Scandinavia.

The correlations between temperature and RDI suggest that including temperature data could improve a statistical drought forecasting model, but the non-stationarity in temperature due to climate change (and already apparent in measured temperature, as seen in Figure 6-9 and as reported in the Central England Temperature, e.g. Karoly & Stott, 2006) would make it difficult to implement. This has not considered further in the project.



**Figure 6-9: Time series plot of seasonal RDI and seasonal temperature anomalies for five UK regions. RDI plotted as solid black lines temperature anomalies plotted in grey, with summers highlighted in red and winters in blue.**

# 7 Model Fitting

## 7.1 Introduction

This section introduces the statistical concepts behind linear modelling (Section 7.2) and presents the results of the statistical model fitting, designed to predict:

- the number of drought months in the next 6 months, a useful planning period for water resources (Section 7.3)
- the length of a drought, important indicator of drought severity (Section 7.4).

The focus is on predicting droughts in the UK, and models were constructed to each UK region. It would be possible to fit similar models to other European regions.

## 7.2 Linear modelling

The techniques used to model the data are standard linear modelling techniques, rather than multilevel models which are likely to add complexity to the models, without adding any real value.

Two types of linear modelling are commonly used:

- General linear modelling which assumes (1) that the response variable (the variable for which prediction is required) has a Normal distribution; (2) that the relationship between the response variable and the explanatory variables is linear, i.e. that a fixed change in an explanatory variable relates to the same change in response variable for the whole range of data. A way of thinking about this is that if a straight line can be fitted to points on a scatter plot then the relationship is linear, whereas if the line should be curved then the relationship is non-linear.
- Generalised linear modelling, where the assumption of normality is relaxed. Instead the response variable is assumed to come from a particular 'family' of distributions, the exponential family. The exponential family of distributions contains most commonly used distribution functions including the Normal, Binomial, Poisson, Exponential and Gamma distributions.

A general linear model with one explanatory variable is of the form:

$$Y = \beta_0 + \beta_1 X + \varepsilon,$$

where  $Y$  is the response variable,  $X$  the explanatory variable and  $\varepsilon$  the error term, or residual. The residuals are equal to the difference between the predicted response value and the observed response variable.

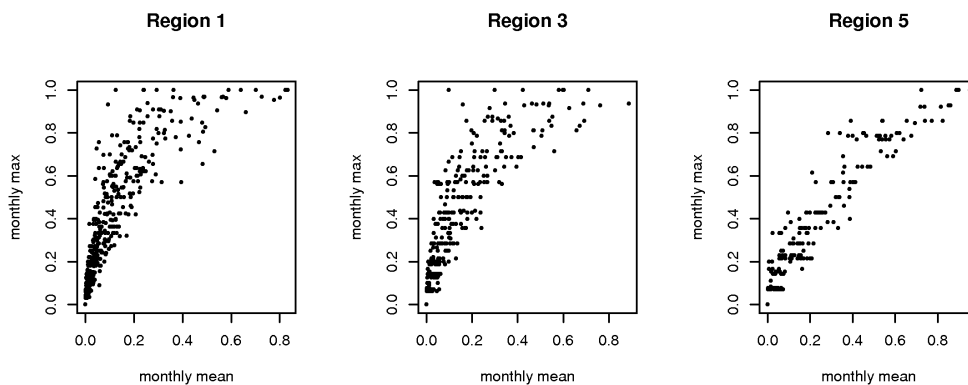
In generalised linear modelling the fitted model is of the form

$$Y = h(\beta_0 + \beta_1 X) + \varepsilon,$$

where the response variable  $Y$  has a distribution function from the exponential family. The function  $h(\cdot)$  is termed a link function. It is this link function that allows the relaxation of the assumption of normality. Each distribution in the exponential family has a particular link function, termed the canonical link function, which makes the maths of the model work best. For the Normal distribution the canonical link is the identity function, so  $h(\theta) = \theta$ , where  $\theta = \beta_0 + \beta_1 X$ . For the Binomial distribution function the canonical link is the logit function so  $h(\theta) = e^\theta / (1 + e^\theta)$ .

## 7.3 Seasonal prediction models

The seasonal prediction models developed in this study aim to predict the number of 'drought months' in the next 6 months using information about the RDI and RSPI in the previous 6 months. A 'drought month' for each region is defined as one for which the monthly average RDI exceeds the 0.9 quantile for the monthly mean average RDI for that region. The reason the monthly average RDI is used to identify drought months is that this will identify months when flows in a particular region are low for a significant amount of time. The relationship between the maximum RDI value in a particular month and the mean for that month was examined (Figure 7-1), and it was found that when the monthly mean value was high then the monthly maximum value was always high. However having a high maximum RDI did not guarantee a high mean RDI. This is especially true for Region 1 which tends to have very short droughts. By using monthly mean RDI rather than monthly maximum RDI, it is likely that only the driest months are included as drought months.



**Figure 7-1: Scatter plots of monthly mean RDI and monthly maximum RDI for three selected UK regions.**

In building these statistical models, seasonality was not taken into account, because the difference in correlation (shown in Figure 6-2 and Figure 6-8) between summer and winter seasons is relatively small.

The data used as covariates in these models are the 6 month mean RDI and mean monthly RSPI. For example, a prediction of the number of months in January-June 2010 that are drought months would use the mean RDI values for July-December 2009 and also the mean monthly RSPI values for July-December 2009. Furthermore, there is a possibility that there may be some interaction between RSPI and RDI for the region of interest. This may take into account some of the differences in the relationship between RDI and RSPI; for instance hot summers may have a different level of correlation than colder summers, but because temperature data is not included in the model, this cannot be assessed directly.

Figure 7-2 shows European regions that have a significant seasonal correlation (greater than 0.4) with any UK region for RDI at seasonal lag 1. In building the seasonal prediction models, attention was therefore focused on these regions.





**Figure 7-2: Regions where lagged correlation with any UK region is greater than 0.4**

The statistical model fitted to the data is a generalised linear model with the assumption that the number of drought months in the next 6 follows a Binomial distribution. This assumption would be safe if each month in the next 6 had an independent probability of being a drought month. It is equivalent to saying that if one month is a drought month, then the likelihood of the next drought month being a drought month is unchanged. Because droughts typically last for longer than one month this assumption is likely to be false. The effect this will have on the fitted model is that it will tend to underestimate the probability of obtaining high numbers of expected drought months.

Table 7-1 gives coefficient ( $\beta$ ) values for all five fitted models. With the exception of RSPi\*RDI (a combination of average RSPi and RDI) the range of values for each covariate is broadly similar and so the actual numbers can be compared directly. However, the actual numbers for RSPi\*RDI are much smaller, so the actual size of the multiplier of these covariates is slightly misleading. For all five models a high average RSPi in the previous 6 months raises the expected number of drought months in the future 6 months whereas the combination of RSPi and RDI lowers the expected number of drought months. This suggests that if the previous 6 months have very little rainfall, and the flows are very low then expected number of drought months in the next 6 months will be lower.

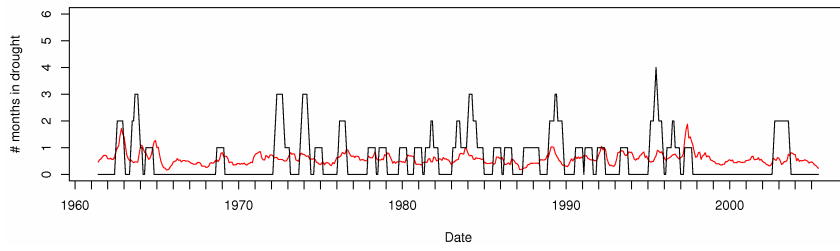
The other coefficient values are a mixture of positive and negative values, a function of the fact that certain patterns of European drought that lead to droughts in the UK while others don't, and because of the correlation in the European drought regions.

**Table 7-1: Coefficient values for all fitted models positive values indicate a higher expected number of drought months, negative values indicate a lower expected number of drought months.**

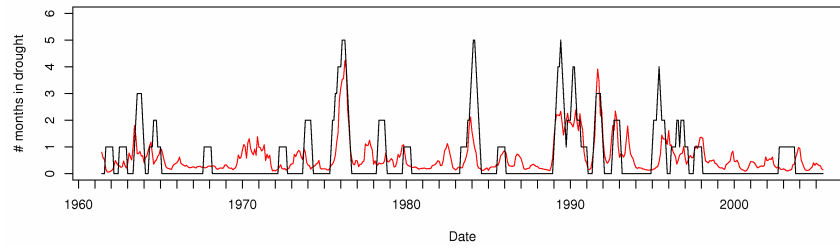
	Region 1	Region 2	Region 3	Region 4	Region 5
Intercept	-2.7	-3.36	-4.03	-4.3	-4.34
mean RSPI	3.58	5.32	5.64	4.89	6.17
Region 1		-3.08			-4.98
Region 2			-5.75	-3.9	-4.03
Region 3		6.38	7.04	4.29	9.25
Region 4	2.56			8.5	7.34
Region 5	-1.18		0.92	-1.83	
Region 7	1.63				
Region 8		4.38	2.47	-3.48	-2.62
Region 9				4.42	4.65
Region 10	1.34	-7.03	-3.96	-4.62	-1.87
Region 11		-8.45	2.37		-3.24
Region 12	-3.16	-3.11		-1.67	-2.58
Region 13	1.16	8.11		2.48	
Region 20		-8.9	2.62	3.53	3.79
Region 21		7.51	-3.26	-5.17	-2.84
Region 22		6.51	3.43	3.25	4.3
Region 23	-1.23	3.01	3.11	4.16	
RSPI*RDI	-7.56	-13.52	-12.57	-12.89	-9.62

Figure 7-3 shows the predicted and observed number of drought months in each 6 month period. The poorest performing model is that for Region 1, which is the region with the lowest lagged correlations. With the exception of this region, the models tend to predict something for most droughts. Interestingly the models predict a drought for 1970 for Regions 2-5 when no drought occurred in the UK. This may be due to the fact that there were some minor (but extensive) anomalies in France and Germany at the time, which were not observed in the UK. Generally, these models appear to be promising for improving seasonal forecasting, but a fuller appraisal of their skill level is needed. The performance of these models is considered further in Section 9, an evaluation of the new methods.

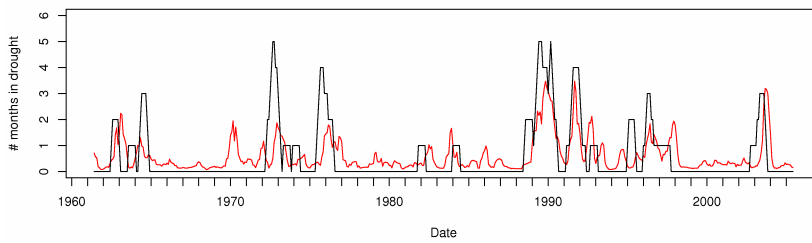
The models predict the ends of droughts reasonably well. In most of the major droughts the number of predicted drought months falls sharply at the end of the drought. This suggests that the mixture of positive and negative parameters in the fitted model may reflect real spatial patterns of European droughts.



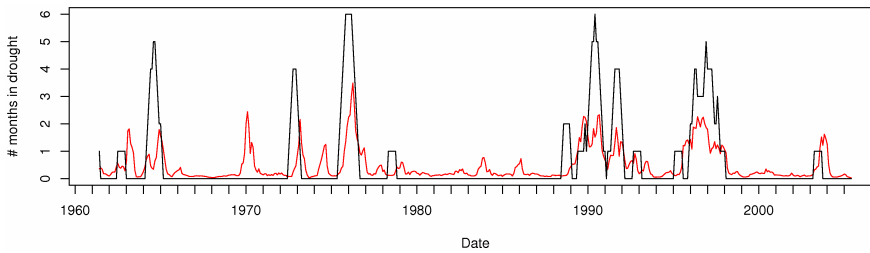
NW Great Britain (1)



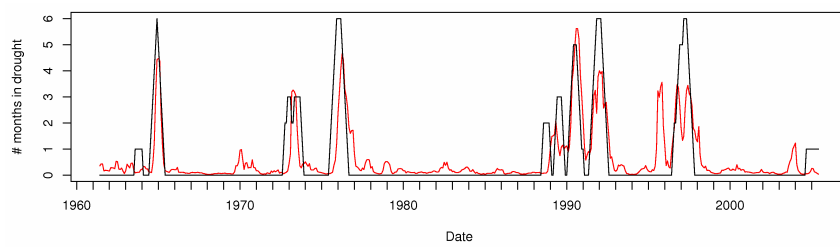
SW Great Britain (2)



NE Great Britain (3)



SE Great Britain non – groundwater dominated (4)

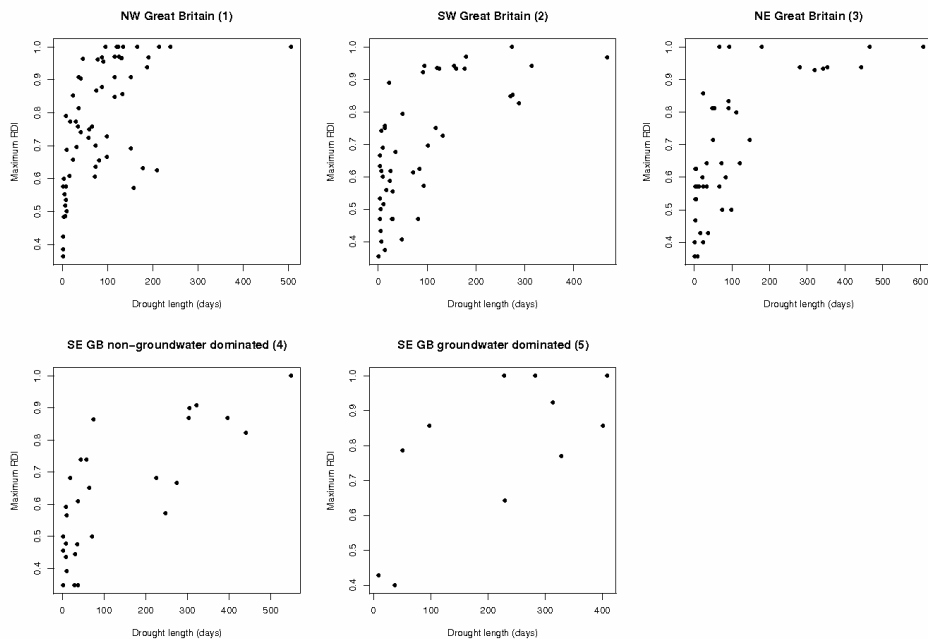


SE Great Britain groundwater dominated (5)

**Figure 7-3: Real and predicted number of future drought months for UK regions  
black lines real, red lines predicted.**

## 7.4 Length of drought

In addition to the seasonal forecast models an attempt was also made to fit a model to predict the length of a drought, given that a region is already in a drought period. Drought length is defined as the length of time between the drought start and drought end, as identified in Figure 6-4. For the UK, drought length and drought maximum are highly correlated (Figure 7-4) so drought length should be predictable from maximum drought intensity. The models fitted attempt to predict length of a drought in each UK region, conditional on the value of RDI in all other regions. An examination is also made of whether or not the following factors could also be used as predictors: 1) a simple indicator function of drought or no drought in other regions, 2) the total number of other regions and 3) whether or not the drought started in summer or winter.



**Figure 7-4: Scatter plots of drought maximum Vs drought length for UK regions**

To ensure that the same data twice is not used twice in each model, three separate sets of model were fitted. The models are of the form (~ means 'is proportional to'):

- Drought length in UK region ~ mean RDI in each region in previous 28 days to start of drought + whether or not drought started in summer
- Drought length in UK region ~ whether or not a drought in other regions in previous 28 days to start of drought + whether or not drought started in summer + total number of other regions that experience a drought in previous 28 days to start of drought
- Drought length in UK region ~ mean RDI in each region in previous 28 days to start of drought or first 28 days of start of drought + whether or not drought started in summer + total number of other regions that experience a drought in previous 28 days to start of drought or first 28 days of start of drought

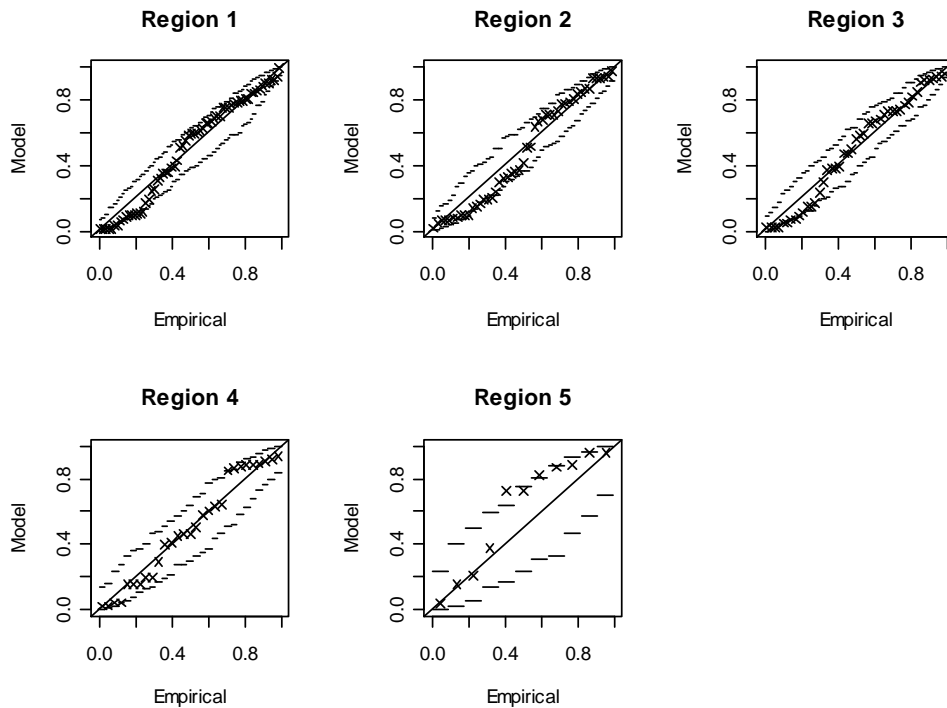
### 7.4.1 Fitting distribution to length of drought

One of the initial findings of the exploratory analysis is that the distribution of drought length is skewed, and seems to have a heavy tail. This is equivalent to saying that most droughts last a short length of time, but that a few droughts last a very long time. The plots in Figure 7-4 show that for most UK regions most droughts have lasted less than 100 days but that a few droughts have lasted longer than 300 days.

This data can be modelled in two ways: 1) transform the drought lengths for each region so that they follow a normal distribution, and then use general linear modelling; 2) select a suitable distribution from the exponential family and use this in generalised linear modelling. Which is the best option depends on the distribution of the data. In this situation the distribution of drought length has such a heavy tail that the first option is best.

The distribution chosen to fit to the data is the Generalised Pareto Distribution (GPD). This is a flexible distribution that arises as the limiting distribution of peaks-over-threshold data. For all UK regions this is a reasonable distribution (Figure 7-5, with data points indicate the probability of observing a particular drought length obtained from the fitted model and the empirical data points show the probability of observing a particular drought length obtained using the empirical distribution function). For Region 5 (SE GB, groundwater dominated), GPD does not provide a good fit, possibly because only 11 droughts were identified, and no model for drought length was fitted.

These models fit better in the main body of the data: these fitted distributions describe the average-length drought better than they describe the extremes. Because these models are used to transform the data to normal scale, it is unlikely to affect the performance of the fitted linear models on the normal scale. However, when transforming back to the original scale the actual values of the longest droughts are likely to be unrealistic.



**Figure 7-5: Fitted distributions for each of the UK regions. Line indicates perfect fit, crosses indicate distribution function of data points (drought lengths), horizontal dashes indicate 95% confidence limits using the empirical data.**

The drought lengths are transformed so that they follow a normal distribution using the probability integral transform. This is exactly the same procedure as is used in calculating the standardised precipitation index.

#### 7.4.2 Best fitting models

The method of model selection was as follows. For each UK region (apart from Region 5, SE GB groundwater dominated) each set of models were fitted independently. The best fitting model to drought length using mean RDI in other regions in the 28 days prior to the start of drought as explanatory variables was then selected. Then the same procedure was carried out using whether or not each other region experienced a drought in the 56 days surrounding the start day of the drought. This model selection was done using Akaike's Information Criterion (AIC). To choose the best fitting model overall the  $R^2$  of each model was used.

None of the models fitted particularly well. The  $R^2$  values for the best fitting model are given in Table 7-2. Figure 7-6 shows plots of fitted drought lengths obtained from the model against the actual drought lengths. Part of the reason for the poor fit of the models is the poor fit of the fitted GPD distributions to the tail of the data. However even on the scale on which the linear models were fitted the QQ plots have a very large amount of scatter and are relatively poor fitting models (not shown).

**Table 7-2: R<sup>2</sup> values for best fitting model**

	Region 1	Region 2	Region 3	Region 4
R <sup>2</sup>	0.161	0.296	0.298	0.400

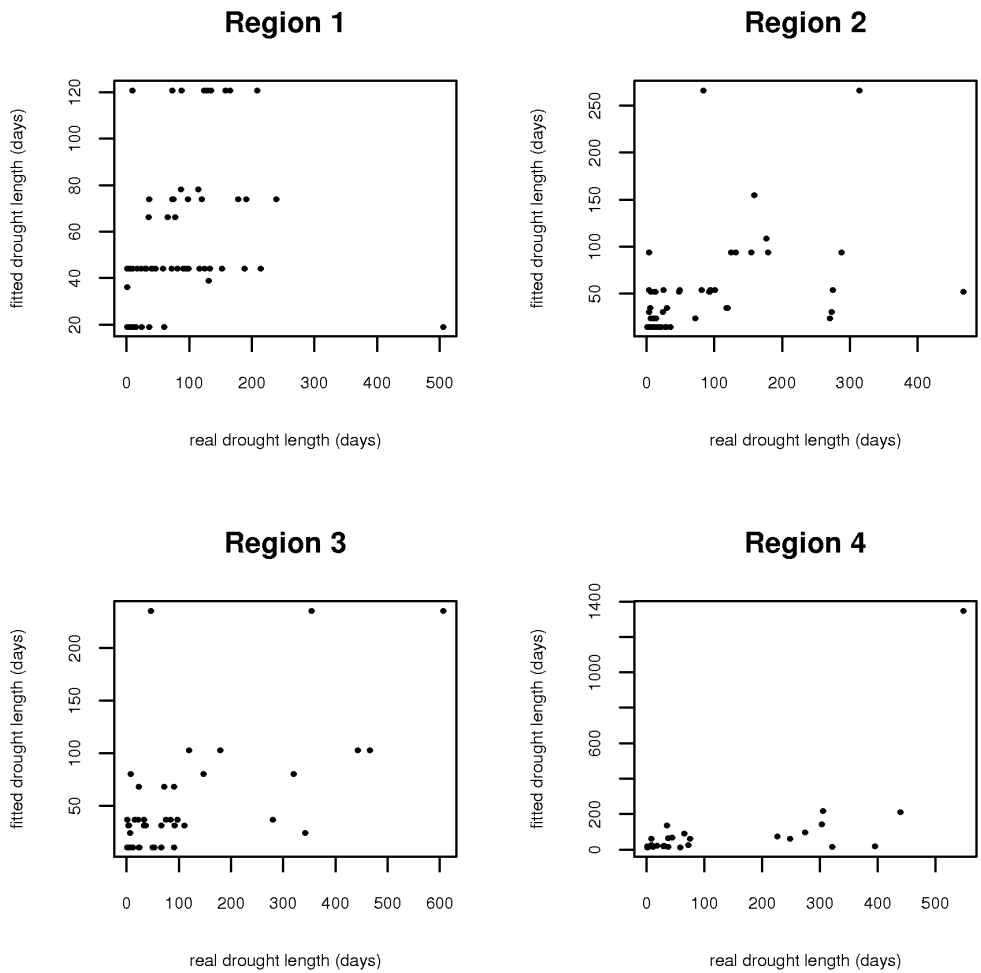
For Region 1 (NW GB) the best fitting model includes indicator variables for drought/no drought in the previous 28 days to the start of the drought. Having a drought in NW Spain seemed to shorten the length of the drought, whereas having droughts in E Germany and Czech Republic and Southern Germany seems to increase the length of the drought. Interestingly none of the other UK regions appear in this best fitting model.

For Region 2 (SW GB) the best fitting model includes indicator variables for drought/no drought in the 28 days either side of the start day of the drought. Surprisingly having a drought in NW Great Britain seems to shorten the length of the drought, whereas having a drought in the Northern France, NE France and French Southern Alps regions seem to increase the length of the drought.

For Region 3 (NE GB) the best fitting model includes indicator variables for drought/no drought in the 28 days either side of the start day of the drought. Having a drought in the Southern Germany and Southern Scandinavia regions seems to increase the length of the drought. Also droughts that start in summer seem to last longer than droughts that start in winter.

For Region 4 (SE GB, non-groundwater dominated) the best fitting model includes the average values for the previous 28 days to the start day of the drought. Higher RDI values in Northern France seem to indicate longer droughts, and droughts that start in summer seem to last longer than droughts that start in winter.

These models were not investigated further because none of them provide a good fit to the data.



**Figure 7-6: QQ plots of fitted model drought length Vs real drought lengths**



# 8 Linking Spatial Coherence to atmospheric Circulation

## 8.1 Introduction

Day to day local weather conditions can be highly variable and reliable forecasts are limited to a few days ahead. The same is not true for the global atmospheric circulation. Large scale patterns are known to occur and persist for weeks or months and it is often possible to forecast these several seasons ahead. It seems reasonable to expect that spatially coherent droughts might be related to the large scale atmospheric circulation. Relationships between low streamflow anomalies and the North Atlantic Oscillation have been found previously in Europe (Shorthouse & Arnell, 1999).

In this section, a summary is given of exploratory analyses which attempt to relate the drought indicators used in this study with large-scale modes of atmospheric variability. Any links between spatial coherence and climatological indicators would enable better explanation of any observed patterns, and may also facilitate more practical operational use, as large-scale atmospheric circulation can be forecasted.

## 8.2 Correlation Analysis

An assessment of the links between coherent drought and climate was made by considering correlations between a range of well known atmospheric patterns such as the North Atlantic Oscillation (NAO) and the El Niño Southern Oscillation (ENSO) with the RSPI. Specifically, rank correlations were computed between RDI and the climate indices listed in Table 8-1 for the period 1950-2005.

Seasonality in the linkages was investigated by performing the correlations individually for each month. Potential predictive relationships were investigated by computing lagged correlations between the particular regional index and antecedent values of the indices from 1 to 6 months.

No significant lagged correlations were detected between the regional RDI indices and the climatic indices. In particular, the phase of ENSO appears to have a negligible impact on the distribution of European droughts. These negative results tend to rule out the possibility of building predictive models for regional European drought from antecedent climatic conditions.

However, a number of interesting contemporaneous correlations have been identified and long range predictions of the large scale atmospheric circulation, where these exist, may prove to be useful in predicting the onset and development of coherent drought.

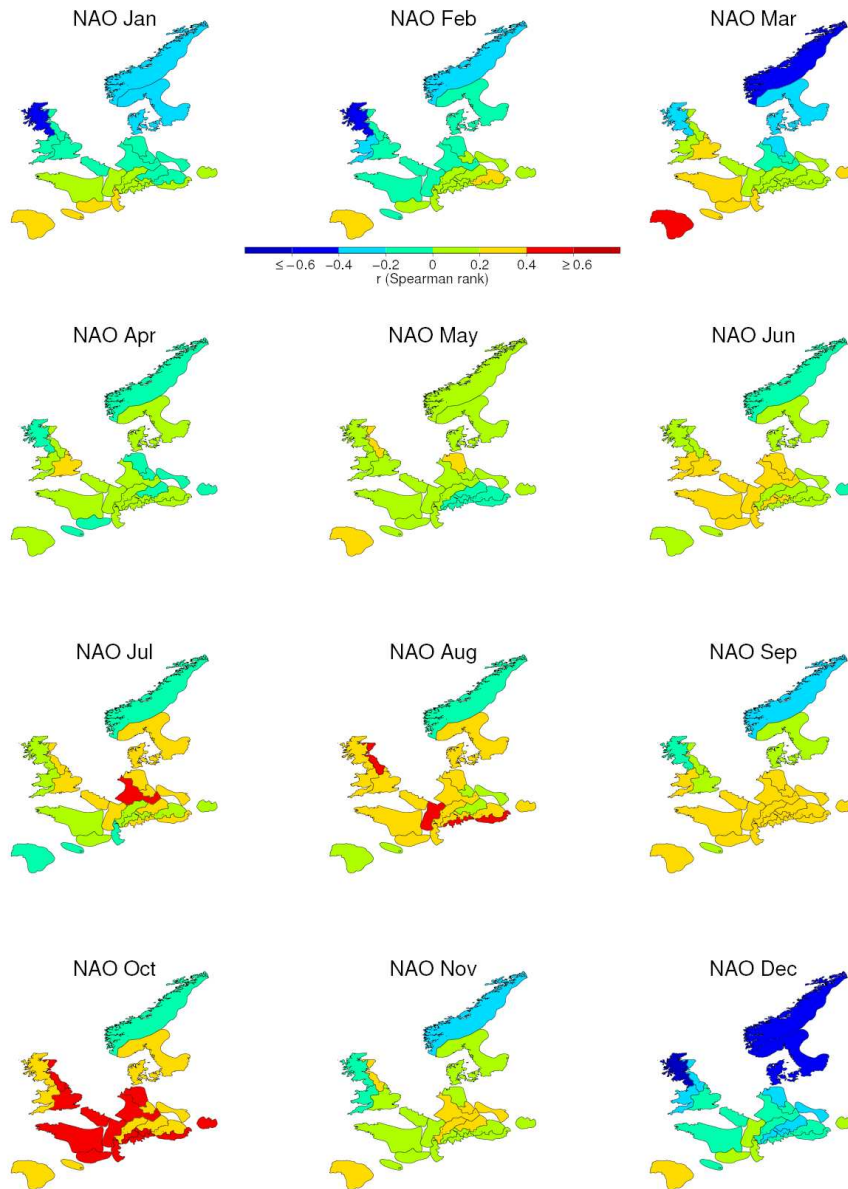
The North Atlantic Oscillation (Figure 8-1) is seen to modulate winter drought in a dipole like manner (i.e. showing a gradient from positive to negative from one end of the domain to another). Positive NAO tends to mitigate drought conditions over Northern latitudes whilst exacerbating drought conditions to the South. Positive NAO during the late summer / early autumn is linked to drought over central western Europe. A similar response is seen to the East Atlantic / West Russian pattern in January/February (Figure 8-2).

Scandinavian droughts appear to be driven by the Scandinavian pattern (Figure 8-3).

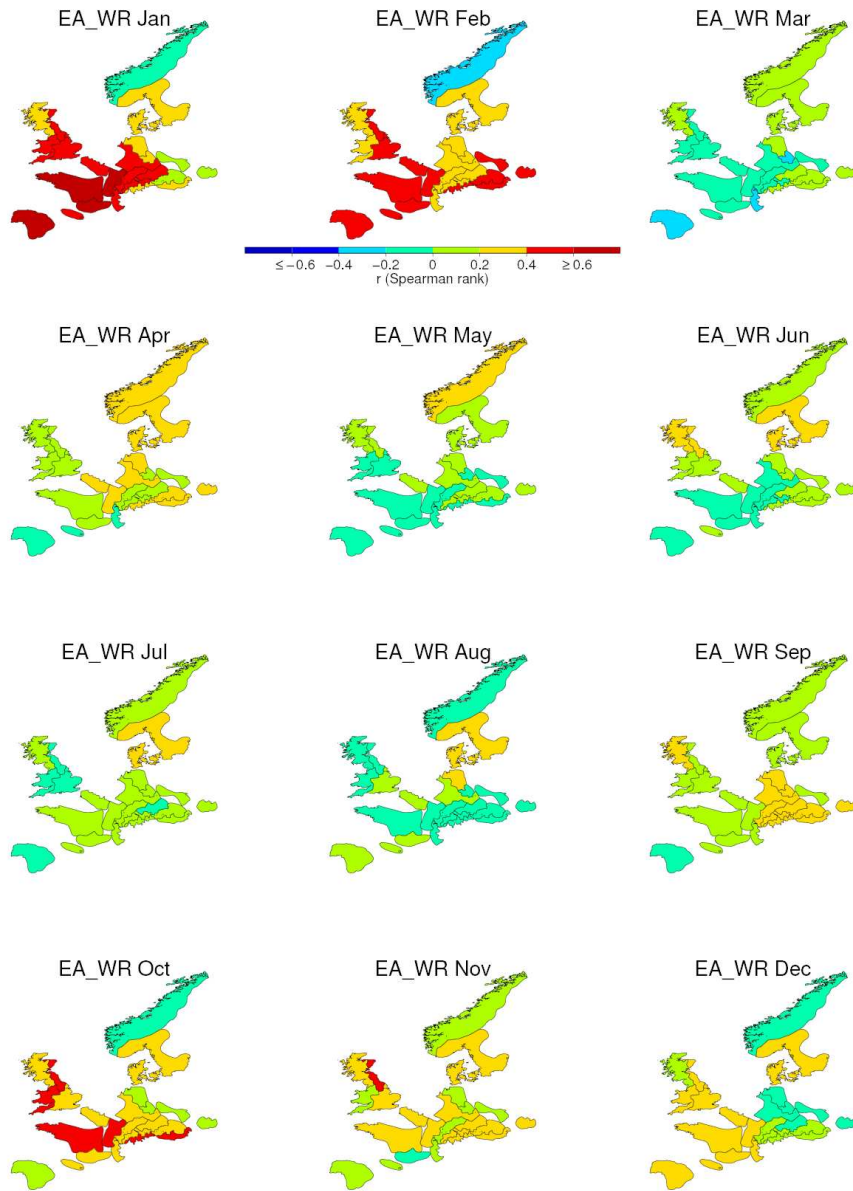
The patterns of correlation show striking similarities to two out of the three major groupings of drought (principal components) seen in Figure 6-3 in Section 6.1. Principal Component 1 appears to be driven by a combination of the Scandinavian and EA / West Russian patterns. PC3 is almost certainly a response to the NAO. We conclude that three large scale patterns (EA/WR, NAO, and SCA) are the dominant atmospheric drivers for coherent drought.

**Table 8-1 Global circulation indices used in the correlation study with regionally coherent drought (further details at the US Climate Prediction Center website (<http://www.cpc.noaa.gov/data/teledoc/telecontents.shtml>)).**

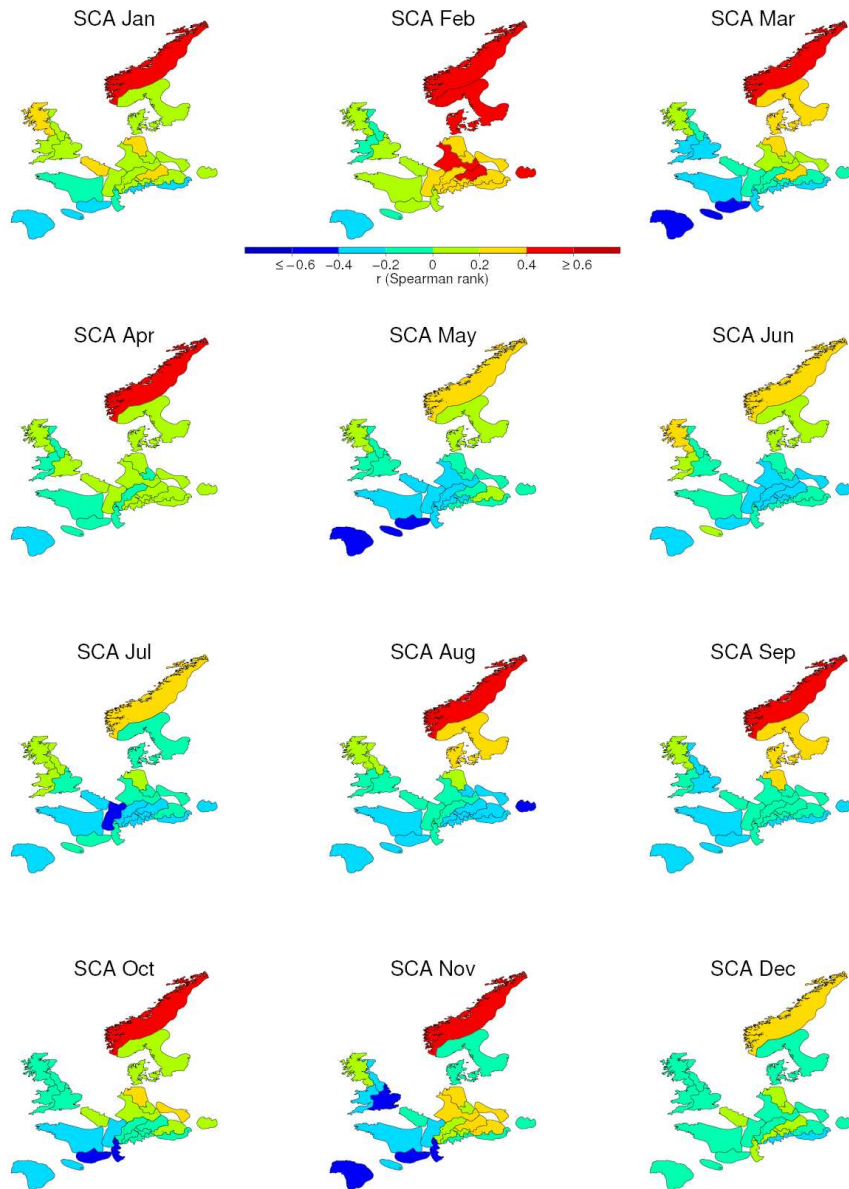
<b>Index</b>	<b>Name</b>
NAO	North Atlantic Oscillation
ENSO (N34)	El Niño Southern Oscillation
WP	West Pacific Pattern
EA	East Atlantic Pattern
EA_WR	East Atlantic/West Russia Pattern
EP_NP	EastPacific/ North Pacific Pattern
PNA	Pacific/ North American Pattern
POL	Polar/ Eurasia Pattern
PT	Pacific Transition Pattern
SCA	Scandinavia Pattern
TNH	Tropical/ Northern Hemisphere Pattern



**Figure 8-1 Rank correlations between RDI and the North Atlantic Oscillation by month for the period 1951-2005.**



**Figure 8-2 Rank correlations between RDI and the East Atlantic West Russia pattern by month for the period 1951-2005.**



**Figure 8-3 Rank correlations between RDI and the Scandinavian Pattern by month for the period 1951-2005.**

# 9 Evaluation of forecasting methods

In this section, the performance of the new forecasting methods is evaluated. In Section 9.1, the seasonal prediction models first described in Section 7.3 are evaluated using statistical tests. In Section 9.2, the relationships with large-scale atmospheric circulation are evaluated, in the context of developing a forecasting tool. Section 9.3 then reviews previous research on seasonal forecasting in the UK, to provide some background context against which the methods can be compared.

## 9.1 Evaluation of seasonal prediction models

Table 3-1 summarises the performance of the models when predicting two or more drought months from:

- the number of successful predictions
- the number of false positives (i.e. when a drought was forecasted but not observed)
- the number of false negatives (i.e. when a drought was not forecasted but observed).
- the percentage of 6 month periods in which there actually were more than two months where the model correctly predicted more than two drought months
- the percentage of 6 month periods where the model predicted more than two drought months but there were fewer than 2 drought months observed

For a model with good predictive capability the first of these percentages should be high and the second low.

**Table 9-1: Performance of model in predicting long seasons with at least two drought months\*.**

	Region 1 (NW GB)	Region 2 (SW GB)	Region 3 (NE GB)	Region 4 (SE GB non GW)	Region 5 (SE GB GW)
Observed and predicted	0 (0%)	17 (3%)	16 (3%)	12 (2%)	37 (7%)
Predicted and not observed	0 (0%)	5 (1%)	7 (1%)	6 (1%)	16 (3%)
Observed and not predicted	20 (4%)	20 (4%)	30 (6%)	42 (8%)	20 (4%)
Not predicted not observed	509 (96%)	487 (92%)	476 (90%)	469 (89%)	456 (86%)
Proportion of observed events correctly predicted (hit rate)	NA	0.46	0.35	0.22	0.65
Proportion of false positives	NA	0.23	0.30	0.33	0.30

\* Each cell gives number of six-month periods where at least two months in the six were predicted and/or observed to have a drought month. Percentages are percentages of total number of six-month periods for each region.

The worst model is for Region 1, probably because this region has a much lower level of lagged correlation than the other regions. For Regions 2-5 the proportion of forecasted false positives is fairly low; when the model predicts that more than two out of the next 6 months are likely to be drought months, then it is more than likely that this is going to occur. However, with the exception of the model for Region 5 none of the models had a hit rate of greater than 0.5, so fewer than half of the observed occurrences of more than 2 drought months were correctly predicted.

Three criteria were used to evaluate the performance of our models in predicting UK drought.

- The first criterion measures how well the models are able to predict the actual number of drought months in the next six.
- Model performance when predicting more than a certain number of drought months in the next six. One difference between these two criteria in practice is that the second more heavily penalises models that consistently under predict.
- How well the models predict whether or not the drought will ease in the next six months.

To measure this last criterion, an examination was made of the differences in predicted numbers of droughts months from month to month, and differences in actual numbers of drought months from month to month. If the difference between two months is negative (so both fewer than last month) for both predicted and actual, then the

prediction is a success. This third criterion assesses how well the models predict when a drought is going to end.

To assess predictive skill objectively the skill of the forecasting models must be compared with what would be predicted with no knowledge apart from that of the long term climate. For the first criterion skill is measured against what would be predicted using the long term climate, measured by the average number of drought months in any 6 month block ( $\bar{x}$ ). In measuring skill this the standard skill score ( $SS$ ) is used. It is calculated using the following formula

$$SS = 1 - \frac{MSE_f}{MSE_c},$$

where  $MSE_f = \frac{1}{N} \sum_{i=1}^N (p_i - o_i)^2$ , and  $MSE_c = \frac{1}{n} \sum_{i=1}^n N(\bar{x} - o_i)^2$ ,  $p_i$  denotes the predicted number of drought months (obtained from the model) and  $o_i$  denotes the observed number of drought months (obtained from the real data).

For the second criterion an assessment is made of how well the model predicts more than  $x$  drought months, when  $x = 0, 1, \dots, 6$ . Model performance is assessed using the ROC skill score, which measures the hit rate proportional to the false alarm rate. The hit rate is the proportion of occurrences (more than  $x$  drought months) that are correctly predicted. The false alarm rate is the proportion of predicted occurrences that were false as they do not correspond to real occurrences, so predicting false events. If a model has no predictive skill then we may expect that the hit rate would be equal to the false alarm rate.

The ROC skill score is also used to assess criterion three. Here we measure how well the model predicts fewer drought months when the real number of drought months is equal to 1, ..., 6. In this way, model performance is assessed over all the possible values.

The ROC score is obtained by plotting the hit rate against the false alarm rate, and calculating the area underneath the curve. Every ROC curve starts at the point (0,0), where no occurrences are observed or predicted, and ends at the point (1,1), where occurrences are always observed and predicted. If a forecasting tool has perfect skill, then the hit rate will be one, and the false alarm rate zero, and the area underneath the curve will be 1. If a model has no skill then the hit rate will be equal to the false alarm rate, and the curve will be a diagonal line, and the area underneath the curve will be 0.5. The ROC skill score is the difference between the ROC curve we observe and what we would expect under no skill (i.e. a ROC skill score of zero indicates no skill).

Table 9-2 shows the three skill scores calculated for each of the five models. For all of the models, all three skill scores are positive, indicating that our models do better than we would expect under no skill. All models show real forecasting skill, in particular to predict the end of droughts (i.e. fewer drought months in one month than were predicted in the last month). For Region 1, the models show very little skill.



**Table 9-2 Skill scores for models to predict number of drought months in the next six.**

	Standard Skill score	ROC Skill score	ROC skill score – drought easing
Region 1 (NW GB)	0.059	0.068	0.031
Region 2 (SW GB)	0.403	0.308	0.338
Region 3 (NE GB)	0.331	0.341	0.448
Region 4 (SE GB non-groundwater)	0.364	0.454	0.588
Region 5 (SE GB groundwater)	0.561	0.578	0.704

## 9.2 Forecasting from large-scale circulation

The relatively low strength of the correlations between the drought indicators and the climatic indices, and the variability between seasons and regions, limit the potential for building quantitative drought models from circulation indices. However, it is likely that long range predictions of circulation pattern will provide useful additional qualitative information. We tested this hypothesis by tabulating the strength of the EA/WR, NAO, and SCA patterns ‘as if’ we had perfect forecasts of the atmospheric circulation.

### 9.2.1 Western Europe

Qualitatively, from the correlation maps, it may be expected that large scale coherent drought would develop over Western Europe given:

1. Positive NAO in late summer / autumn.

and/or

2. Positive EA/WR in January / February.

If a coherent western European drought is defined as a mean RDI of more than 10% over France, western Germany, and southeast UK then it is possible to tabulate the reliability of the above:

**Table 9-3 Contingency table for rule-based prediction of winter drought over central Western Europe 1961-2004.**

		Observed	
		Drought	No drought
Predicted	Drought	6	13
	No drought	5	20

Assuming that the observed rate of the occurrence of drought remains constant at 11 out of 44 years, an estimate can be made of the likelihood that a set of wild guesses will do better than the rule based forecasts. The rule is found to be 30% more reliable than a guess at predicting the occurrence of drought but offers no improvement in the forecast of no drought. The statistical significance of the association of predicted and observed drought / no drought can be computed using the p-value from Fisher's exact test. This is the probability of seeing a more extreme table (one with more apparent association) by chance alone. The probability of seeing an association as strong or stronger than that shown in by chance is 0.30. This is quite probable and it can be concluded that the statistical significance of the results is weak.

### 9.2.2 North West UK

A similar rule based forecast can be constructed for winter drought over the north west UK can be constructed from the anti-correlation with winter time NAO over this region:

**Table 9-4 Contingency table for rule-based prediction of winter drought over central north western UK 1961-2004.**

		Observed	
		Drought	No drought
Predicted	Drought	10	6
	No drought	12	16

This rule is found to be 25% more reliable than a guess at predicting the occurrence of drought and 15% better at forecasting no drought. The significance of this result is much greater and the probability of obtaining the same result or better by chance is 0.17.

## 9.3 Comparison with previous forecasting work

Table 9-5 summarises the key findings of some previously published papers on seasonal forecasting in the UK and in Europe.

Globally, there is an extensive and growing body of research into seasonal forecasting. The techniques employed can broadly be subdivided into statistical methods, where relationships are derived between one or more predictors and a target variable of interest (e.g. rainfall or precipitation), and dynamical methods, which rely on the integration of General Circulation Models (GCMs) or other modelling techniques. A majority of previous studies have employed statistical methods, which have generally been shown to be more successful than dynamical approaches (e.g. see the review of Easey *et al.* 2006).

Within the UK, a number of studies have employed statistical methods to forecast rainfall or river flows. A majority have focused on forecasting summer river flows from predictors for the previous winter (e.g. Sea Surface Temperatures, SSTs, or large scale atmospheric circulation patterns such as the North Atlantic Oscillation) using linear

regression techniques. The focus on summer flows is because this is generally the time of greatest water resources stress in UK catchments.

The overall skill level of most of these techniques is relatively low, if quantified in terms of the amount of variance explained by the regression relationships. However, these studies have met with some degree of success at providing indicative forecasts. In general the techniques show potential for providing indications of higher or lower than average summer flows given certain conditions in the predictors; for example, the Expert System method of Wedgbrow *et al.* (2005) found limited explained variance, but performed significantly better when forecasting low flows only and was able to correctly model the signs of June – August anomalies in a significant majority of cases, including major drought years. The relatively limited skill level achieved in previous forecasting studies in the UK, along with much of Europe, is partly due to the lower climatic variability in the mid-latitudes. In comparison, seasonal forecasting is generally more successful in areas such as the sub-tropics, such as Australia, where climates are strongly associated with large-scale circulation patterns such as the El Niño/Southern Oscillation (ENSO, e.g. Ropelewski & Folland, 2001).

Dynamical models show little skill at predicting weather for the UK at the seasonal scale (Stockdale, 2007). However, physical relationships between weather and the state of the oceans have been found (Colman & Davey, 1999) but current models are not yet capable of representing these particular processes. Undeterred, the UK Met Office use these to issue seasonal outlooks based upon a mixed statistical and physical model forecast process. At a lead time of 3 months these are typically around 30% better than a wild guess at predicting the correct tercile (above average, average, below average) for temperature and 10% better for precipitation.

**Table 9-5 Review of previous seasonal forecasting literature**

Reference	Domain	Predictors	Predictands	Methods	Results and Skill/Performance
Wilby, 2001	12 catchments (around UK)	Winter (JF): <ul style="list-style-type: none"> <li>o NAO</li> <li>o SSTs (North Atlantic)</li> <li>o CET</li> </ul>	Summer: Monthly Q (JJA). 1961 - 1999.	<ul style="list-style-type: none"> <li>o Correlation</li> <li>o Regression</li> </ul>	<ul style="list-style-type: none"> <li>o NAO explained most variance in August (c.40% on the r. Stour in SE England)</li> </ul> <p><b>Stour ‘best case’ example:</b></p> <ul style="list-style-type: none"> <li>o Forecasts capture general behaviour, but are poor for individual seasons</li> <li>o Number of correct predictions was statistically significant</li> </ul>
Wedgbrow <i>et al.</i> 2002	14 catchments (around UK)  [reconstructed river flows]	Winter (monthly): <ul style="list-style-type: none"> <li>o POL</li> <li>o NAO</li> <li>o SSTs (T1)</li> </ul>	Summer (monthly and average): <ul style="list-style-type: none"> <li>o Q</li> <li>o PDSI</li> </ul>	<ul style="list-style-type: none"> <li>o Correlation</li> <li>o Split-sample analysis</li> <li>o ‘Index of Forecasting Potential’</li> </ul>	<ul style="list-style-type: none"> <li>o Negative relation between winter POL, NAO, SSTs and following summer PDSI and Q. Highest in the East (<math>r &lt; 0.5</math>), for POL.</li> <li>o Coherence tests show +ve SSTs and POL precede below average Q in 70 – 100% of summers (NW and SW)</li> <li>o Negative NAO may also precede low Autumn flows in the SE</li> </ul>
Wilby <i>et al.</i> 2004	Thames	Winter (JF) <ul style="list-style-type: none"> <li>o Sea Ice extent</li> <li>o MSLP</li> <li>o NAO</li> <li>o AO</li> <li>o POL</li> <li>o SOI</li> </ul>	Summer (JJA, A) <ul style="list-style-type: none"> <li>o Q</li> <li>o P</li> <li>o PDSI</li> <li>o CET</li> </ul>	<ul style="list-style-type: none"> <li>o Stepwise linear regression (+ lognormal)</li> <li>o Jackknife testing</li> </ul>	<ul style="list-style-type: none"> <li>o Hindcasts were consistently better than climatology</li> <li>o Most significant gains were with PDSI, then Q</li> <li>o Signs of anomalies generally captured, levels of explained variance were modest.</li> </ul>

		<ul style="list-style-type: none"> <li>○ Geopotential heights</li> </ul>				
McGregor & Phillips, 2004	Southwest peninsula of England	<ul style="list-style-type: none"> <li>○ SSTs (N Atlantic)</li> <li>○ Mean monthly sea level pressure</li> <li>○ Vorticity</li> <li>○ Atmospheric water vapour flux</li> </ul>	Rainfall, at monthly and seasonal timescales	<ul style="list-style-type: none"> <li>○ Multiple linear regression models</li> <li>○ Multivariate specification and prediction models</li> </ul>	<ul style="list-style-type: none"> <li>○ Predictability of rainfall at monthly and seasonal timescales is more dependent on SSTs than on atmospheric circulation variability</li> <li>○ Whilst SSTs account for the majority of the variance in monthly and seasonal rainfall, the inclusion of the atmospheric water vapour flux improves rainfall prediction</li> </ul>	
Wedgbrow <i>et al.</i> 2005	Thames	<ul style="list-style-type: none"> <li>Winter (generally JF) <ul style="list-style-type: none"> <li>○ Sea Ice extent</li> <li>○ MSLP</li> <li>○ NAO</li> <li>○ AO</li> <li>○ POL</li> <li>○ SOI</li> </ul> </li> <li>Geopotential heights</li> </ul>	<ul style="list-style-type: none"> <li>Summer (JJA, A) <ul style="list-style-type: none"> <li>○ Q</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>○ Expert Systems</li> </ul>	<ul style="list-style-type: none"> <li>○ When forecasting low flows only, correctly forecasts 79% of Aug and 89% of summers in the lowest quartile</li> <li>○ Amount of variance explained is low, but signs are correct for 71% of JJA anomalies</li> <li>○ Models correctly predicted below-average flows for major droughts (1976, 1984, 1995, 1997)</li> </ul>	
Svensson & Prudhomme, 2005	20 catchments (around UK), grouped into NW and SE clusters	<ul style="list-style-type: none"> <li>Winter (DJF) <ul style="list-style-type: none"> <li>○ River flow</li> <li>○ SSTs (N Atlantic)</li> <li>○ N America land air temperature</li> <li>○ NAO</li> <li>○ EA</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>Summer (JJA) <ul style="list-style-type: none"> <li>○ Q</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>○ Correlation</li> <li>○ Linear regression</li> </ul>	<ul style="list-style-type: none"> <li>○ Equations fitted to all data explain 55% and 61% of the variance for the NW and SE clusters, respectively</li> <li>○ Cross-validation correlations between predicted and observed series are 0.54 and 0.62</li> <li>○ Models are fairly skilful at avoiding false positive or</li> </ul>	

					negative prediction of extremes	
Fletcher & Saunders, 2006	Northern hemisphere	Atlantic SST NH snow cover	Winter (DJF) NAO.	Regression Hindcasting with block elimination.	6 to 9% improvement over a climatological forecast.	
Stockdale, 2007	Global 2.5x2.5 degree resolution	N/a	Seasonal averages of meteorological fields e.g. temperature, precipitation, and surface pressure.	Ensemble of coupled ocean-atmosphere dynamical models.	Correlation skill for November forecasts of DJF winter precipitation and temperatures over western Europe ranges between zero over Spain to 0.6 over Denmark.	

The new seasonal prediction method developed in this project differs substantially from previous statistical approaches. The new method attempts to forecast 'drought from drought', i.e. using spatial coherence of drought to facilitate some degree of early warnings in the UK given certain conditions of drought indicators for regions elsewhere on the continent. The new method therefore differs in its configuration, which confers several advantages compared to previous approaches:

- Being based on the RDI, the new seasonal prediction approach enables forecasting for large-scale regions of the UK, whereas previous approaches are generally fine-tuned to particular catchments.
- Previous approaches focused on river flows generally examine monthly flows, whereas this method examines the probability of the number of months in drought, according to a pre-defined RDI criterion. The new method therefore provides a more explicit assessment of the likelihood of being in drought (although this obviously is sensitive to the definition of what constitutes a drought) rather than experiencing a lower-than-average flow.
- Similarly, the new method can enable ongoing prediction of drought development once a drought has started. It performs reasonably well at forecasting the cessation of drought conditions. This would be useful during a drought – judging the likelihood of a drought ending could be of high utility for supporting water management decisions – but the termination or continuation of drought conditions has previously been very challenging to predict. In this context, the tool may have potential for application as a drought monitoring tool, for giving some indication of the likelihood of drought conditions intensifying or ameliorating in the next season.
- The new prediction indexes deficiencies at any time of year. Whereas previous approaches generally focus on summer conditions, the new method can also enable prediction of winter droughts. Similarly, rather than just focusing on a short-term anomaly, the method is well suited to examining longer multi-season, multi-year droughts. This is particularly advantageous, as a majority of severe historical droughts are long droughts which result from decreased winter replenishment of water resources (Marsh *et al.* 2007a; Parry *et al.* 2009), which are often then exacerbated with a summer water resource drought.

As the new method is very different from previous approaches, a direct comparison of skill levels is not appropriate. However, it is notable that the new forecasting model generally has a fairly low 'false alarm' rate, but this may be because few droughts are predicted.

The climatology-based methodology outlined in Section 9.2 has some parallels with previous approaches, in that it attempts to forecast conditions based on atmospheric circulation. However, most previous approaches exploit lagged relationships between winter predictors and summer river flows. The new approach has the benefit of attempting to predict winter droughts directly through the winter predictors. However, the low skill level of the qualitative models suggests that they are of limited utility at present, other than for indicative purposes.

# 10 Potential operational utility of new methods

## 10.1 Practical uses of the drought catalogues

The drought catalogues (Lloyd-Hughes *et al.* 2009) are potentially of very high utility for a range of purposes, both operationally and in terms of future research.

The catalogues provide an accessible and clear picture of the major historical droughts in a particular region of interest. They would therefore be useful to practitioners and researchers examining the severity, spatial coherence, duration and seasonality of historical drought episodes in Europe, as well as the historical sequencing of wet and dry periods. In the UK, the catalogues and the individual drought summaries provide further quantitative and qualitative information to supplement reports which have been produced following major drought episodes, such as 1988 – 1992 (Marsh *et al.* 1994) and 2003 (Marsh, 2004). Similarly, the catalogues compliment previous studies which have attempted to identify major historical droughts and to assess their impacts (Marsh *et al.* 2007a), particularly given that the assessment of meteorological drought extends back to the beginning of the twentieth century.

The drought catalogues and summaries will undoubtedly have wider resonance within the drought research and drought management communities in Europe. The catalogues and drought summaries represent one of the most comprehensive compilations of drought characteristics and spatio-temporal evolution of major European droughts. They have more information and are more up-to-date than previous pan-European drought datasets gathered for the ARIDE project (Demuth & Stahl, 2001).

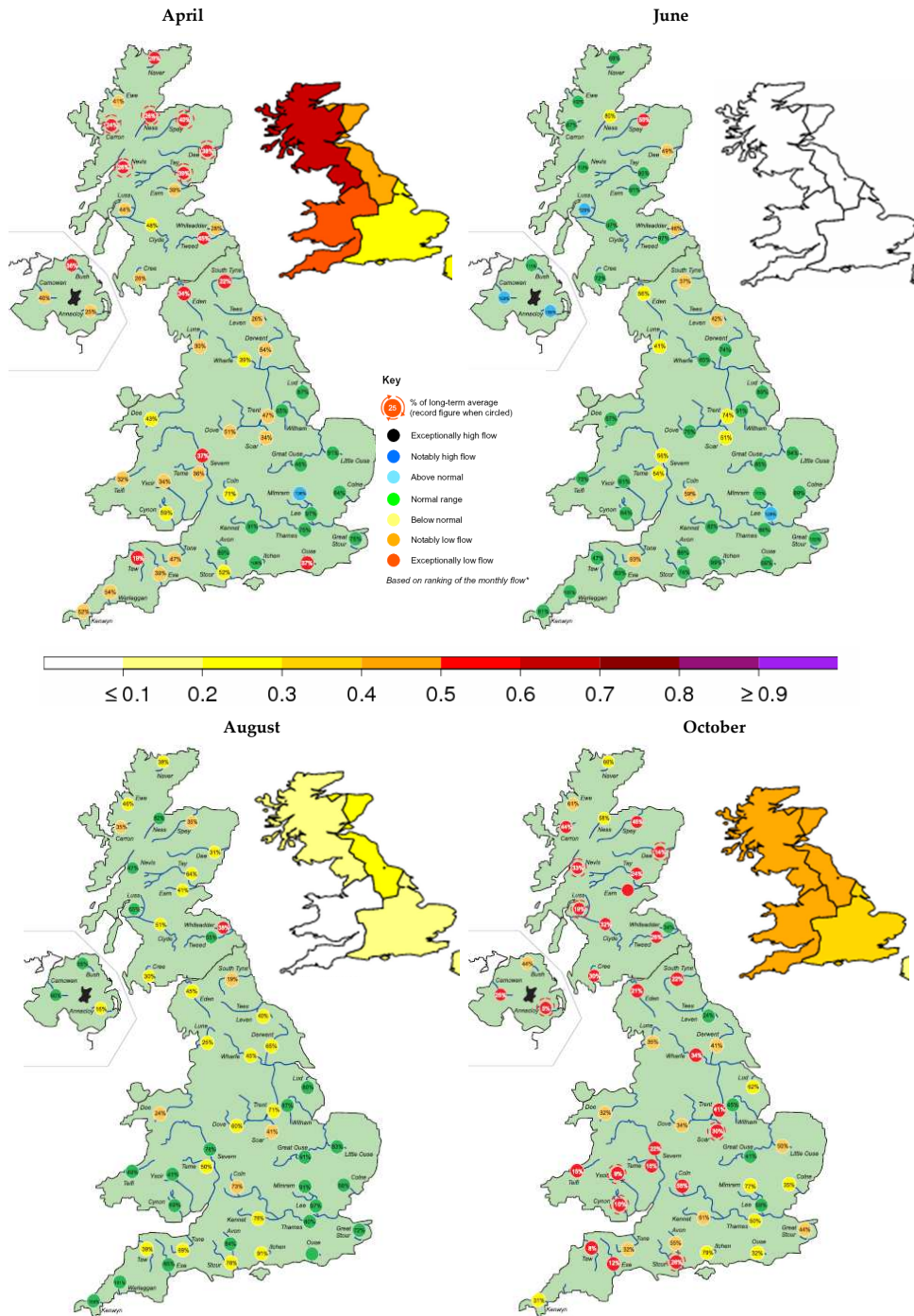
A recommendation of this project is therefore that the catalogues and drought summaries are disseminated and publicised as widely as possible within the European drought community. The catalogues and associated outputs are being used as part of the EU-FP6 WATCH (Water and Global Change) project which represents one way of disseminating the catalogues to the European community, most likely through links with the European Drought Centre (EDC, <http://www.geo.uio.no/edc/>) and related projects such as XEROCHORE, an EU Support Action which aims to develop a future EU drought policy (<http://www.feem-project.net/xerochore/>).

## 10.2 Use of the RDI and RSPI in drought monitoring

Alongside their use as a foundation for developing forecasting tools within this project, the RDI and RSPI datasets have potential for wider operational application. One key potential benefit is that the indicators could add a new dimension to efforts to monitor the development of drought conditions in the UK and Europe. For example, in the UK, the indicators could be integrated into the EA Water Situation Reports (Environment Agency, 2009) or the CEH Monthly Hydrological Summaries, released as part of the National Hydrological Monitoring Programme (NHMP: CEH, 2009). At present, these feature maps of monthly rainfall and river flow accumulations compared to long-term averages, and hydrographs of recent river flows. Whilst the NHMP and WSRs are aimed at monitoring the contemporary water resource situation, at present, there are no specifically drought-orientated metrics featured within either publication. Some drought indices could therefore be useful extensions. To compare the RDI indices with



the CEH hydrological summary, snapshots at two-month intervals during the 2003 summer drought are shown in Figure 10-1.

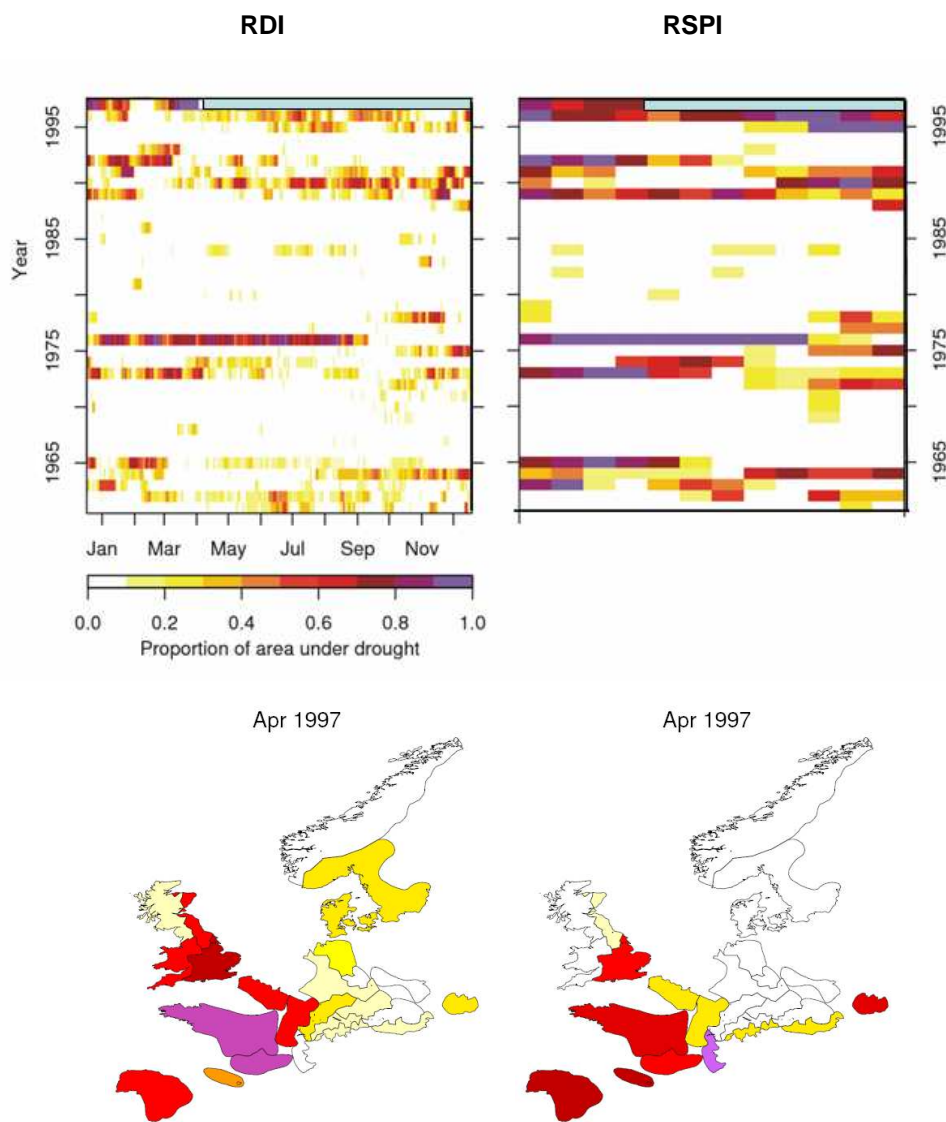


**Figure 10-1 Comparison between the RDI Maps (with four UK regions, legend in centre) and Monthly Hydrological Summaries for the same months in 2003**

Qualitatively, there is a very good degree of agreement between the regional RDI plots and the monthly summaries. The monthly summaries have a much higher degree of detail by allowing individual catchments to be differentiated. Some differences are clearly apparent, but this is unsurprising given the fact that different sets of catchments are used, and the Hydrological Summary focuses on at-site monthly runoff (as a percentage of the mean) as opposed to the proportion of sites under Q90. Some differences are also likely to be due to the different catchments used, especially as the RDI map shown features the non-groundwater dominated catchments only for the South East Great Britain region, whereas the Hydrological Summary mixes groundwater and non-groundwater dominated catchments. Generally, however, the agreement is encouraging, and suggests that despite their different formulation, the techniques would be complementary, but would focus on different facets of the flow response.

A distinct advantage of the RDI and RSPI is that they would be representative of a region rather than particular catchments, and therefore provide an indication of drought as an areal phenomenon, integrating as they do a measure of drought severity in terms of the prevalence of streamflow or rainfall deficiencies in a region. The contemporary RDI and RSPI values for any region could be assessed and compared with values from previous droughts, to give an indication of the severity and spatial coherence of a developing drought event. Furthermore, the indicators can be generated for 24 European regions (with scope for expansion given data availability), so potentially could be used to examine the spatio-temporal development of future large scale European droughts. The RDI and RSPI have been shown to be powerful tools for examining historical droughts, and the same techniques (e.g. drought animations) could be used for future droughts. Such monitoring would further be informed by lessons from historical droughts which are encapsulated in the historical drought summaries (Parry *et al.* 2009), and by linking the current RDI and RSPI values in a given accumulation period with climatic indicators (see 10.3 for further consideration of the potential for forecasting).

As an example of the potential use of the indicators for monitoring, a case study is presented in Figure 10-2, simulating the indicators as they would have been applied in April 1997 for South East England (non-groundwater). The RDI map also shows the European picture – in this case, a large scale deficiency persisted over the whole of the UK, France and Spain. Similarly, the time series plot in Figure 10-2 shows a much longer view than is typically featured in the CEH monthly Hydrological Summaries. In the monthly summary, runoff is normally mapped for the month and over one long-term accumulation period. The RDI and RSPI time series enables the deficiencies to be appraised over much longer periods – in this case, rainfall and runoff deficiencies were developing since mid-1995. Whilst similar information is tabulated in the monthly summary (e.g. 3-, 6- and 9- month accumulations), the RDI/RSPI approach enables a more direct visual appraisal. Similarly, the RDI plot clearly shows that the April 1997 deficit is the most pronounced April flow deficit in the available record (since 1961).

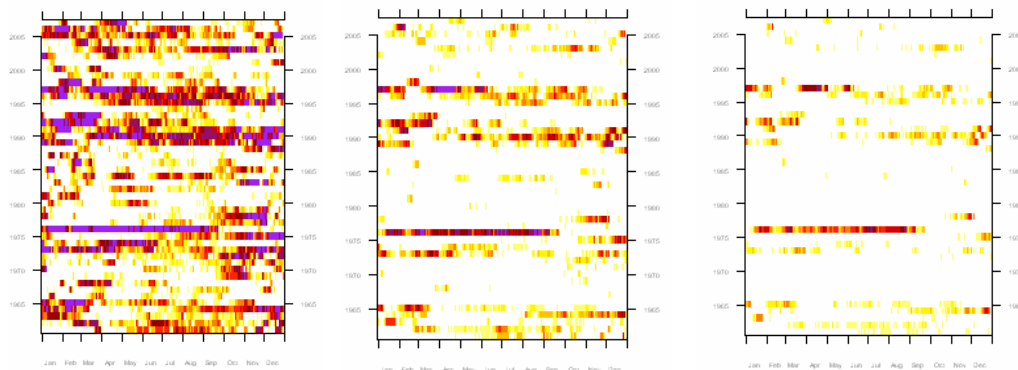


**Figure 10-2 Example of RDI and RSPI Metrics for April 1997**

Before the methods could be used operationally, a number of important questions would need to be addressed. The following questions illustrate some caveats and associated constraints on the use of the indicators at present – these are further followed up in the suggestions for further research (Section 11.4).

- Is the choice of regions (in the UK and in Europe) appropriate? Within the UK, the current regions may not conform to operational or administrative boundaries – further work would be needed to examine the suitability of the regions for operation purposes (in particular, whether the groundwater catchments should be treated separately or integrated), and this also applies to Europe.

- Is this the right set of catchments for the UK? Currently, 132 benchmark catchments are used, a deliberate policy to ensure that drought responses approximate a natural, climate-driven drought response. Many are relatively small upland catchments, which are not necessarily representative of a wider region for monitoring purposes. The catchments used in the National Hydrological Monitoring Programme or Water Situation Reporting datasets (which tend to be larger rivers) may be more appropriate.
- Is the formulation of the indicators adequate, and appropriate for operational use? For example, Q90 was used as a threshold for calculating whether a flow is under deficiency. Q95 is also a widely used drought indicator (arguably with more resonance within the low flow and hydroecology communities). However, there is inevitably a trade-off between making the threshold severe enough to be relevant, and capturing enough droughts. This is illustrated in Figure 10-3, which compares three thresholds for one region. With Q95 as a threshold, only the most severe droughts are captured (it is much rarer for a high number of catchments to be simultaneously in drought). In contrast, when Q70 is used, most years contain a severe deficit and short droughts concatenate into long sequences of deficiency. Similarly, the RSPI is defined (Section 2.2.2 on the basis of a proportion of grid cells under moderate drought ( $SPI < -1$ )), and the RSPI can potentially be calculated for a range of durations; tuning would be needed to ensure the indicator is appropriate for monitoring purposes.



**Figure 10-3 Comparison between RDI plots produced using three thresholds (left – right): Q70, Q90 and Q95. For SE Great Britain, non-Groundwater region**

In addition to these more theoretical considerations, some practical issues would need addressing before using the RDI and RSPI for monitoring. Data constraints, principally timeliness of provision, are likely to be especially challenging on a European scale (see Section 11.2)

Finally, the RDI and RSPI should be appraised for their suitability against other methods used for drought monitoring. As well as the current approaches used by WSRs and Hydrological Summaries, there are many alternative indicators actively used in monitoring. Extensive operational drought monitoring programmes are already used in the USA (National Drought Mitigation Centre, NDMC: <http://drought.unl.edu/>) and in a prototype European Drought Observatory (EDO: <http://edo.jrc.ec.europa.eu/php/index.php>) developed at the Joint Research Centre (JRC) for the European Commission. Similarly, a global drought monitor is produced on a monthly basis by the Benfield Hazard Research Centre at University College

London (<http://drought.mssl.ucl.ac.uk/>). These systems use a range of indicators, including the SPI and the widely-used Palmer Drought Severity Index (PDSI) which are spatially distributed and thus give a comprehensive picture of spatio-temporal development of drought. There is an extensive reliance on modelled and remotely sensed data in these systems, however, and they generally focus on meteorological and agricultural droughts; there is currently limited use of hydrological data. It is therefore very likely that the RDI (or some equivalent measure of regionalised streamflow) would have a place in drought monitoring on a European scale, if the practical issues are resolved. Similarly, on a UK scale the RDI and RSPI have potential for detailed monitoring based on observational data, to complement the much lower-spatial resolution approaches employed by the EDO.

## 10.3 Forecasting Methods

The seasonal prediction models developed in this study have the potential to be applied in drought management in the UK. As discussed in Section 9, their formulation confers distinct advantages compared to previous forecasting tools. The seasonal prediction models allow an assessment of the likelihood of a drought intensifying or diminishing. This property is potentially of very high operational utility. Application of the method would enable forecasting on a regional scale, of the likelihood of UK drought given the current situation on the continent. Whilst the forecast would be tempered with ‘health warnings’ due to the moderate predictive skill of the models (particularly in some regions, such as the North West GB), it would enhance the current capabilities for early warning – for both summer and winter – of developing droughts, and similarly for drought decay.

This is certainly an improvement on the current situation, where the outlook is generally driven by monitoring of rainfall or runoff accumulations, compared to historical precedents, with no real capacity for a ‘forward look’ (other than the use of long range meteorological forecasts such as those provided by the Met Office). Application of the new method would enable the likelihood of regional streamflow drought to be assessed. This intelligence could potentially be of use for drought early warning and then ongoing monitoring during droughts for the Environment Agency, Defra and other water managers and policymakers (such as the water companies).

There are two main aspects to consider in using these models operationally to predict drought in Europe. The first is data availability. The data used are RDI and RSPI values, which are based on monthly river flow and precipitation from a large number of stations over Europe. To employ this method as an operational tool, these data must first be collected and processed within a month, and then it must be made available to the Environment Agency. There are potentially a number of operational and political barriers to overcome before this is possible (see also Section 11.2). However, one key finding of this project is that the method demonstrates the feasibility of using spatial coherence of droughts on the continent to forecast drought in the UK. If the availability of streamflow data is an issue, future research could investigate the potential for linking UK RDI values to just the RSPI or other readily available meteorological datasets from Europe; this analogue approach could possibly exploit the concept delivered in this study, but may circumvent some of the practical limitations.

The second aspect to consider before using these models operationally is the sensitivity to the definition of what constitutes a drought. In this study, drought is defined in terms of how likely a certain level of dryness is to be observed. The study has not attempted to assess what level of dryness would be relevant as a drought definition for operational forecasting. If a different definition of dryness is used, then the models would need to be re-fitted. These re-fitted models will almost certainly not

be identical to the models fitted in this project, although it is likely that the general performance of these models would be similar.

The large scale circulation methods discussed in Section 9.2 could also potentially be used to infer the likelihood of drought development. One of the advantages of this approach is that it is relatively simple, requiring only monthly values of the NAO or EA/WR to enable some degree of early warning. However, the modest skill of the approach at present suggests that further work would be required before a similar system could be developed into an operational early warning tool. Furthermore, methods based on climate indices are contingent on perfect knowledge of the signs of the NAO and EA/WR patterns ahead of the upcoming winter. In reality, an operational scheme will be reliant on less than perfect seasonal forecasts of these quantities. The best available seasonal forecasts of the winter NAO issued in late summer / early autumn provide improvements over chance of between 6 and 9% (Fletcher & Saunders 2006). The EA/WR is not routinely forecast and no reliability figures are available. However, its nature is similar to the NAO and comparable reliability is expected. Combining the uncertainties of the rule based drought models with the uncertainty in the forecasts of NAO and EA/WR, we estimate that the reliability of an operational model issuing forecasts of drought / no drought in the forthcoming winter will be no better than 3% more reliable than chance. Improvements in forecasting such as those expected with EUROSIP (<http://www.ecmwf.int/products/catalogue/pseth.html>) will increase this figure.

Some research suggest that in the UK, the centres of highest correlation between rainfall (river flow) and the mean sea level pressure in the North Atlantic moves from month to month, and is always consistently greater than the correlation with the NAO (Lavers *et al.*, submitted). This would suggest that climate indices might not be the right tools to forecast droughts, and that more research is needed to identify the most indicators with the largest forecasting capacity in Europe.

In general, the forecasting tools delivered in this project show some potential for application for water management in the UK, and possibly Europe. At present, there are major issues to consider in terms of practical applicability, and the skill levels are relatively low. However, if efforts are invested in improving the predictability through further research, and increasing data availability, the methods could be developed into operational tools in future.



# 11 Conclusions and suggestions for further research

The new methods developed in this study clearly show potential for improving the UK's capabilities for early warning and forecasting of droughts, as well as monitoring conditions during a drought. Section 11.1 summarises the key findings of the project. There are some important caveats associated with the methods and their development which should be considered. Section 11.2 addresses the importance of data issues. Section 11.3 focuses on caveats and limitations associated with the methodologies employed. These are then further expanded on in the suggestions for further research, in Section 11.4.

## 11.1 Summary of main findings

- This study has employed a comprehensive and up-to-date dataset of streamflow and precipitation data from a number of European countries to examine the spatial coherence of droughts at a continental scale; unlike most previous large-scale drought research, this study combines assessments of both hydrological and meteorological droughts in a consistent framework across Europe.
- The Regional Deficiency Index (RDI) and Regional Standardised Precipitation Index (RSPI) have been shown to be powerful tools for assessing the spatial coherence of drought for large regions. Twenty four regions were identified across Europe; catchments within these groups frequently experience simultaneous streamflow deficiencies.
- For these 24 European regions, time series of regional streamflow and rainfall deficits were produced, and catalogues of regional drought severity were developed, spanning 1901 – 2005 for meteorological droughts, and 1961 – 2005 for hydrological droughts. These catalogues provide a comprehensive and clear picture of drought characteristics in each of the regions, capturing: drought duration, seasonality and spatial coherence. This 'Drought Catalogue' is a major deliverable of this project, and is published as a separate report (Lloyd-Hughes *et al.* 2009). The catalogue is a powerful tool for regional drought visualisation, and will be of practical utility to for drought management, policy-making, and future research in the UK and Europe.
- A detailed analysis was conducted of the spatial and temporal evolution of a number of historical droughts, using the RDI and RSPI datasets; the aim of this was to elucidate spatio-temporal dynamics of historical droughts, which may prove fruitful for informing the development of forecasting models. However, these analyses were felt to be of wide practical utility, so a comprehensive description of the spatio-temporal evolution of all major post-1961 European droughts was produced. These 'historical drought summaries' are another major deliverable of this project, released in a standalone report (Parry *et al.* 2009). These summaries illustrate the temporal evolution of streamflow and rainfall deficits, along with information on climatic drivers and atmospheric circulation anomalies, and also provide some narrative on the drought evolution, and major impacts of the events.
- From an appraisal of these events, it is clear that most droughts appear to have different characteristics, in terms of their duration, spatial coherence and

seasonality. For example, a contrast was found between the 1976 drought, which was spatially consistent across much of Europe and was combined with a rainfall deficiency the preceding winter and a heat wave in the summer, and the 1995-1997 drought, which was interspersed by wet episodes and had little long-lasting spatial coherence over Europe. In most historical events, the UK experienced drought simultaneously with other European regions, or earlier; there was little evidence of any systematic lag time which could be readily exploited in the development of early warning systems for the UK based on conditions in other parts of Europe.

- An exploratory data analysis was then carried out, to determine whether there are relationships in the drought indicators which could be exploited to develop forecasting tools. Correlation analysis, multidimensional scaling and statistical modelling were applied to find relationships, which were generally fairly weak.
- Low correlations exist between regional drought deficiency time series of different regions, and the correlation patterns for hydrological and meteorological droughts are similar, albeit slightly higher for the latter. Correlations with the rest of Europe are stronger in winter than in summer for northern and western Britain, but are of similar magnitude all year round for southeast England.
- Although a relationship was identified between the length of a UK drought and the number of regions contemporaneously experiencing drought elsewhere in Europe, it was found that this relationship was not statistically significant.
- Following these exploratory analyses, statistical models were built for each UK region, which predict the number of drought months that may occur in the next 6 months. Predictions are based on streamflow deficiencies in other European regions, so the models essentially predict 'drought from drought' – i.e. they use the spatial coherence of anomalies to derive forecasts for the UK based on deficiencies on the continent. The models forecast droughts in most regions relatively well, particularly in groundwater-dominated catchments in southeast England. In northwest Britain, however, the predictive capability is poor.
- Importantly, the models have some significant benefits when compared to previous seasonal forecasting studies – in particular, the approach is based on large regions, rather than being 'tuned' to particular catchments, and they enable forecasting of winter anomalies rather than just summer flows. Furthermore, the models perform reasonably well at forecasting the cessation of drought conditions. These attributes mean that the models could potentially be of high utility during long, multi-season drought events, to determine whether a drought is likely to intensify or to diminish.
- Whilst the predictive capacity is modest in some regions, the models clearly have potential for application in UK drought management, although there are also important practical considerations – in particular, the need for timely data supply from across Europe – which would need to be examined in further research before they could evolve into an operational tool. Nevertheless, the results prove the concept; that in some regions, it is possible to forecast 'drought from drought', using spatial coherence to predict drought without recourse to other sets of predictors.
- Further analysis concentrated on attempting to explain observed patterns of spatial coherence, by linking drought indicators to large scale modes of atmospheric variability (e.g. the North Atlantic Oscillation and the East Atlantic-West Russia pattern). In some regions and some seasons, these predictors clearly play an important role in determining the spatial coherence of droughts. Whilst their predictive capability is relatively weak at present, there is



undoubtedly scope for refining these relationships into tools for monitoring and providing indicative forecasts. An advantage of this approach is that some climatological indicators are routinely forecast (although the modest skill levels are a further obstacle to application at present).

- The regional drought indicators are shown to be powerful tools for illustrating the dynamics of rainfall and streamflow deficiencies. They could therefore find application in UK and European drought monitoring systems. Again, there would be important practical limitations to consider, and further research would be needed to optimise the indicators for use in monitoring. However, they could potentially fill an important gap; existing monitoring European drought monitoring systems lack a streamflow component, whilst UK approaches (e.g. CEH's monthly Hydrological Summaries) consider runoff deficiencies but do not use any metrics tailored specifically to drought.

## 11.2 Data issues

The river flow dataset used in this study is undoubtedly one of the largest, most up-to-date streamflow datasets for Europe. However, obtaining a dataset of this size for Europe was only made possible through previous research projects (Stahl *et al.* 2008 and Prudhomme & Sauquet, 2006), the former of which had dedicated resources available to collate the data from a multitude of European partner institutions – a major logistical challenge.

At present, not all of these data are available on the European Water Archive (EWA). Even when they are eventually archived centrally, they are only routinely updated on an annual basis. Accessing these data in a timely manner would be a major obstacle to using the methods developed in this study for monitoring and forecasting. Nevertheless, data sharing is increasingly recognised as an important priority for European Water management, and is one of the key recommendations of the XEROCHORE project (<http://www.feem-project.net/xerochore/>). The following data related issues should be considered when appraising the results of this study:

- It was not possible to make any appraisal of data quality from other countries in Europe. However, the various organisations responsible for submitting the data were asked for undisturbed, good quality sites (Stahl *et al.* 2008), so it is assumed that the majority of the data are reliable.
- The distribution of stations over Europe (see Figure 3-1) is somewhat irregular, with high densities of stations in some areas (e.g. Germany). There are very limited data in Southern Europe or Eastern Europe, so the study domain is essentially north western and central Europe. Similarly, no data was available in some key areas which, *a priori*, would be thought to be influential for early warning in the UK. In particular, the distribution is patchy over Northern France (due to heavy human disturbances) and there were no data from the Benelux countries. Republic of Ireland data was not available at the time, and this may also be of utility in early warning (particularly given the west-east migration of some major drought events).
- In order to get a standard time period across the region, the RDI was only computed to 2005, as the majority of European data was only updated to this point. The study therefore missed out the 2004 – 6 drought, which was significant in the UK (Marsh *et al.* 2007b).

## 11.3 Caveats and Limitations

In addition to the data issues, the following points should be considered, regarding the choice of methodology used in this project:

- The decision was made to use discrete regions, which are artificial constructs defined by statistical analyses. Whilst they provide a framework for the spatial coherence project, in reality the variability of hydrological and meteorological conditions is much more complex.
- The regions used in this study were inherited from previous research (a necessary response to the timescale and resources available). Whilst some effort was made to ensure the regions are homogeneous in terms of drought response, it is likely that they are not optimal for characterising droughts, particularly given the size of some regions. However, any attempt to regionalise is dependent on the distribution of the data; given the irregular network available to this study, it is unlikely that results would be substantially different given alternative regions.
- The RDI and RSPI are both regional measures of the proportion of a region that is under drought. It is therefore important to consider that drought in this study has been defined in terms of regional coherence, which is only one way of indexing drought – although approaches which integrate drought areal extent as a feature of severity are widely used (e.g. Severity-Area-Frequency analysis, Hisdal & Tallaksen, 2003; Severity-Area-Duration analysis, e.g. Sheffield *et al.*, 2009). Nevertheless, the fact remains that drought characterisation is sensitive to the homogeneity of regions (see above bullet point), and it is possible that in some regions historical droughts are less well represented owing to heterogeneity in catchment response rather than a true lack of drought conditions.
- The RDI and RSPI are both similar in construction, but use different thresholds to define drought. Correlation analyses have shown the RDI and RSPI to be strongly associated, but they were not compared systematically. Caution would therefore need to be exercised in extrapolating to hydrological droughts from meteorological droughts (for example, when comparing the RDI and RSPI in the regional catalogues, pre-1961 when only meteorological data was available). The complex propagation of meteorological to hydrological droughts (Tallaksen and VanLanen, 2004) further underlines the need for caution in interpreting the catalogues.
- Non-stationarity in the statistical relationships used to develop prototype forecasting tools was not considered. This is not thought to be especially problematic, but should be considered in future if the tools were to be extended to future droughts.

## 11.4 Further research

The following recommendations are suggested as important priorities for further research:

- The drought catalogue and historical drought summaries are a powerful way of visualising drought characteristics at a regional scale in Europe. Future work could build on the catalogues by corroborating the data in the catalogues with information on drought impacts. Evidence of environmental and socio-economic impacts would bolster the catalogue, by providing an independent assessment of drought severity, to compare with the relative

- magnitude of drought events assessed by the objective metrics and the observed data. The catalogues should be considered a 'dynamic' document, which can be populated with additional information (particularly drought impacts) from previous research or gathered from experts in the various European countries for which catalogue data is available.
- The relationships between RDI and RSPI should be examined more fully, to explore whether meteorological drought can be used as a reasonable surrogate for hydrological drought. Relationships in spatial coherence using RSPI and/or other meteorological data should be examined, to see whether an analogue to the seasonal prediction model could be developed, which may offer similar benefits but would avoid the reliance on European streamflow data (for example, employing raingauge data if available, or using reanalysis data as a surrogate).
  - Further work should be undertaken to examine the regions used in this study, and to determine whether they are suitable for a range of monitoring and forecasting purposes. If more data becomes available for a wider part of Europe, and for areas of sparse data coverage, a new regionalisation could be attempted, which may improve the homogeneity and coverage of drought regions.
  - Alternative approaches to examining spatial coherence on a European scale should be considered – for example, to address the question as to whether regions are needed at all. Some form of Severity-Area-Duration analysis (e.g. Sheffield *et al.* 2009) may be suitable for examining spatio-temporal development of large scale European droughts, using gridded rainfall and gridded runoff datasets.
  - The present study has focused primarily on the period for which streamflow is available, i.e. the early 1960s – present. Further work could expand the 'sample' of historical droughts by concentrating on meteorological data and examining the spatio-temporal development of historical droughts on a European scale (e.g. 1921 – 1922; the 1940s). The regional meteorological droughts could be corroborated using at-site data from selected long streamflow records from Europe (Stahl *et al.* 2008).
  - Research should be conducted to examine the potential for developing the prototype tools delivered in this study into operational monitoring and forecasting tools. A scoping study could be undertaken to investigate whether practical issues (e.g. data availability) could be overcome. Similarly, more extensive work should be undertaken to test the utility and suitability of the RDI and RSPI for operational purposes – for example, to address the questions raised in Section 10.2, on the sensitivity of the RDI to flow thresholds, choice of catchments etc.
  - The scope of this study has primarily been focused on forecasting UK drought from conditions in Europe. The datasets and analyses produced herein could potentially be used for prediction in other European regions, if further work was conducted to assess the feasibility of this approach for other 'target' regions. The west-east migration of some events suggests early warning in France, southern Scandinavia and Germany may benefit from monitoring of drought development on the Atlantic margin, but further research would be needed to investigate such an approach in detail.
  - Although a model for predicting length of drought (Section 7.4) proved elusive in the present study, such a tool would be of considerable utility. One reason that this was unsuccessful is that the fit of the distribution to length of drought was poor. Future work may benefit from breaking this

- approach into smaller steps, starting by attempting to fit a good distribution to drought length.
- In this study, some preliminary work explored the possibility of using temperature as an explanatory variable in drought prediction. This was found to be of limited utility within the scope of this study, but temperature is undoubtedly an important factor in some major droughts (particularly when associated with heatwave conditions, e.g. in 1976, 1995 and 2003). Furthermore, there are likely to be interactions between temperature and rainfall/streamflow (e.g. in enhancing evapotranspiration, thereby exacerbating streamflow deficiencies in responsive catchments; or in affecting snowmelt-in alpine Europe). Further work could attempt to integrate temperature into the predictive framework, perhaps using a joint-probability approach.
  - The relationships with large-scale atmospheric circulation patterns should be studied in more detail, to determine whether climatological indicators could successfully be used to explain spatial coherence and thus improve predictability. For example, to explore correlation between RDI and atmospheric variables, such as MSLP over North Atlantic, rather than fixed climate indices such as the NAO, as strongest correlation centres between catchment rainfall and mslp (for example) have been shown to move seasonally (Lavers et al., submitted)
  - This study examined relationships with large-scale climatology, but fruitful relationships may also be found between drought indicators and Weather Types. Previous research has exploited this for rainfall data (Fowler & Kilsby, 2002). At present, research is underway to investigate links between regional streamflow drought in the UK and weather types (Fleig, *et al.* 2009). Links between the present study and Fleig *et al* 2009 should be explored in detail.
  - The development of climate forecasting tools is currently a very active area of research, with new products emerging for various forecasting time lead, from medium range (up to 10 days) and seasonal forecast (up to 12 months) to decadal and multi-decadal projections (up to 100 years). However, their ability of reproduce the climate conditions and patterns leading to droughts have not yet been fully researched (e.g. blocking patterns, weather types, combined rainfall and temperature spatial patterns etc). In addition, for these products to become real operational tools, it is fundamental that research is done to understand how to exploit their potential, for example in developing some bias correction techniques, and in identifying and quantifying the different forecasting skills associated with different atmospheric conditions. This would help improving short term drought forecasting as well as future drought risk in Europe.



# References

- BRIFFA, K. R., Van der SCHRIER, G. & Jones, P.D. Wet and dry summers in Europe since 1750: evidence of increasing drought. *International Journal of Climatology*, *in press*
- COPA COGECA, 2003. Assessment of the impact of the heat wave and drought of the summer 2004 on agriculture and forestry.
- CEC, 2007. Addressing the challenge of water scarcity and droughts in the European Union. Communication from the Commission to the European Parliament and the Council. Commission of the European Communities. Brussels, 18.7.2007, COM(2007) 414 final.
- CEH, 2009. National Hydrological Monitoring Programme, Water Watch website. [http://www.ceh.ac.uk/data/nrfa/water\\_watch.html](http://www.ceh.ac.uk/data/nrfa/water_watch.html) Page accessed: 6th July 2009
- COLMAN, A. & DAVEY, M. 1999. Prediction of summer rainfall, temperature and pressure in Europe from preceding winter North Atlantic Ocean temperature. *International Journal of Climatology*, 19, 513 – 136
- DAI, A., FUNG, I. & DEL GENIO, A., 1997. Surface observed global land precipitation variations during 1900-1988. *Journal of Climate*, 11, 2943-2962.
- DEMUTH, S. & STAHL, K. (Eds.), 2001. ARIDE - Assessment of the Regional Impact of Droughts in Europe. Institute of Hydrology, University of Freiburg.
- EASEY, J., PRUDHOMME, C. & HANNAH, D. 2006. Seasonal forecasting of river flows: a review of the state-of-the-art. In Demuth, S., Gustard, A., Planos, E., Scatena, F. And Servat, E. *Climate Variability and Change – Hydrological Impacts*. Proceedings of the fifth FRIEND conference, Havana. IAHS Publication 308, pps 158 – 162
- ENVIRONMENT AGENCY, 2009. Water Situation for England and Wales. <http://www.environment-agency.gov.uk/research/library/publications/33995.aspx> Page Accessed: 6th July 2009
- FLEIG, A.K., TALLAKSEN, I.M., HISDAL, H. & HANNAH, D.M. 2009. Regional hydrological droughts and weather types in North West Europe. *Geophysical Research Abstracts*. Vol. 11, EGU2009-1053, 2009
- FLETCHER, C. G & SAUNDERS, M.A., 2006. Winter North Atlantic Oscillation hindcast skill: 1900-2001, *Journal of Climate*, 19, 5762-5776.
- FOWLER, H.J. & KILSBY, C.G. 2002. A weather-type approach to analyzing water resource drought in the Yorkshire region from 1881 to 1998. *Journal of Hydrology*, 262, 177-192
- GUSTARD, A., BULLOCK, A. & DIXON, J. M., 1992. Low flow estimation in the United Kingdom. IN *HYDROLOGY*, I. O. (Ed. Wallingford: Institute of Hydrology Report 108.
- HANNAFORD, J., MARSH, T.J. 2006. An assessment of trends in UK runoff and low flows using a network of undisturbed catchments. *International Journal of Climatology* 26: 1237 – 1253
- HISDAL, H., TALLAKSEN, L. M., 2003. Estimation of regional meteorological and hydrological drought characteristics. *Journal of Hydrology*, 281(3), 230-247
- JONES, P. D. & CONWAY, D., 1997. Precipitation in the British Isles: An analysis of area-average data updated to 1995. *International Journal of Climatology*, 17 (4), 427-438.
- JONES, P. D. & HULME, M., 1996. Calculating regional climatic time series for temperature and precipitation: methods and illustrations. *International Journal of Climatology*, 16, 361-377.
- KAROLY, D.J., & STOTT, P.A. 2006. Anthropogenic warming of central England temperature. *Atmospheric Science Letters*, 7, 81-85.

- KEEF, C. SVENSSON, C. TAWN, J.A., 2009 Spatial dependence in extreme river flows and precipitation for Great Britain, submitted to Journal of Hydrology
- LAVERS, D., PRUDHOMME, C. & HANNAH, D.M., submitted. Large scale climate, precipitation and British river flow: identifying hydroclimatological connections, dynamics and mechanisms. Submitted to Journal of Hydrology
- LLOYD-HUGHES, B., SAUNDERS, M.A., 2002. A drought climatology for Europe. International Journal of Climatology, International Journal of Climatology, 22, 1571 - 1592
- LLOYD-HUGHES, B., PRUDHOMME, C., HANNAFORD, J., PARRY, S., KEEF, C., REES, G. 2009. Drought Catalogues for UK and Europe.
- MARSH, T.J. 2004. The UK drought of 2003: a hydrological review. Weather, 59, 224 - 230
- MARSH, T.J., MONKHOUSE, R.A., ARNELL, N.W., LEES, M.L. & REYNARD, N.S. 1994. The 1988 – 1992 drought. Hydrological Data UK Series. Institute of Hydrology and British Geological Survey. Wallingford.
- MARSH, T., COLE, G. & WILBY, R., 2007a. Major droughts in England and Wales, 1800-2006. Weather, 62 (4), 87-93.
- MARSH, T.J., BOOKER, D. & FRY, M. 2007b. The 2004 – 2006 Drought. Hydrological Data UK series. Centre for Ecology and Hydrology, Wallingford.
- McGREGOR, G.R. & PHILLIPS, I.D. 2004. Specification and prediction of monthly and seasonal rainfall over the south west peninsula of England. Quarterly Journal of the Royal Meteorological Society. 130, 193 – 210.
- MCKEE, T. B., DOESKEN, N. J. & KLIEST, J., 1993. The relationship of drought frequency and duration to time scales. Proceedings of the 8th Conference on Applied Climatology, 17-22 January. Anaheim, CA.: American Meteorological Society.
- MITCHELL, T. D. & JONES, P. D., 2005. An improved method of constructing a database of monthly climate observations and associated high-resolution grids. International Journal of Climatology, 25 (6), 693-712.
- NERC, 1975, Flood Studies Report. Natural Environmental Research Council
- PARRY, S., LLOYD-HUGHES, B., HANNAFORD, J., PRUDHOMME, C, KEEF, C. 2009. Drought Summaries of Spatio-temporal evolution of major European Droughts since 1960.
- PRUDHOMME, C. & SAUQUET, E., 2006. Modelling a regional drought index in France. Wallingford: CEH - Cemagref.
- ROBSON, A. and REED. D. W. , 1999 Flood Estimation Handbook Vol 3: Statistical procedures for flood frequency estimation. Wallingford: Institute of Hydrology.
- SHORTHOUSE, C. A. & ARNELL, N.W. 1997. Spatial and temporal variability on European river flows and the North Atlantic Oscillation. *FRIEND '97-Regional Hydrology: Concepts and Models for Sustainable Water Resource Management* (Proceedings of the Psotojna, Slovenia, Conference, Sep-Oct 1997), IAHS. Publ. no.246.
- SHEFFIELD, J., ANDREADIS, K.M., WOOD, E.F., & LETTENMAIER, D.P. 2009. Global and continental drought in the second half of the twentieth century: Severity-Area-Duration analysis and temporal variability of large-scale events. Journal of Climate, 22, 1962 – 1981.
- SMITH, R.L. and WEISSMAN, I. 1994 Estimating the extremal index. J.R. Statist. Soc.B 56, 515-528
- STAHL, K., 2001. Hydrological drought - a study across Europe. Geowissenschaftlichen Fakultät. Freiburg: Albert-Ludwigs Universität Freiburg.
- STAHL, K. & DEMUTH, S. 1999. linking streamflow drought to the occurrence of atmospheric circulation patterns. Hydrological Sciences Journal, 44, 467 - 482
- STAHL, K. HISDAL, H, TALLAKSEN, L.M. Van LANEN, H.A.J, HANNAFORD, J & SAUQUET, E. 2008. Trends in low flows and streamflow drought across Europe. Report to UNESCO, Fee contract 450040322. 45pp.

- STOCKDALE, T., 2007. EUROSIP: Seasonal forecasting with ocean-atmosphere models, COSMOS General Assembly, Mainz 23-25 May 2007.
- SVENSSON, C. & PRUDHOMME, C. 2005. Prediction of British summer flows using winter predictors. *Theoretical and Applied Climatology*, 82, 1 - 15
- TALLAKSEN, L. M. & VAN LANEN, H. A. J., 2004. Hydrological drought - Processes and estimation methods for streamflow and groundwater. Elsevier.
- VAN DER SCHRIER, G., BRIFFA, K.R., JONES, P.D. & OSBORN, T.J. 2006. Summer moisture variability across Europe. *Journal of Climate*, 19, 2818 - 2834
- | WEDGBROW, C.S., WILBY, R.L., FOX, H.R. & O'HARE, G.O. 2002. Prospects for seasonal forecasting of summer drought and low river flow anomalies in England and Wales. *International Journal of Climatology*, 22, 219 – 236 Deleted: B
- | WEDGBROW, C.S., WILBY, R.L., & FOX, H.R. 2005. Experimental seasonal forecasts of low summer flows in the River Thames, UK, using expert systems. *Climate Research*, 28, 133 – 141. Deleted: B
- WILBY, R. 2001. Seasonal forecasting of river flows in the British Isles using North Atlantic pressure patterns. *Journal of the Chartered Institute of Water and Environmental Management*. 15, 56 – 63
- | WILBY, R., WEDGBROW, C.S., & FOX, H.R. 2004. Seasonal predictability of the summer hydrometeorology of the river Thames, UK. *Journal of Hydrology*, 295, 1 – 16. Deleted: B
- ZAIDMAN, M. D., REES, H.G. & YOUNG, A. R., 2002. Spatio-temporal development of streamflow droughts in north-west Europe. *Hydrology and Earth System Sciences*. 5 (4), 733 - 751



We are The Environment Agency. It's our job to look after your environment and make it a **better place** – for you, and for future generations.

Your environment is the air you breathe, the water you drink and the ground you walk on. Working with business, Government and society as a whole, we are making your environment cleaner and healthier.

The Environment Agency. Out there, making your environment a better place.

Published by:

Environment Agency  
Rio House  
Waterside Drive, Aztec West  
Almondsbury, Bristol BS32 4UD  
Tel: 0870 8506506  
Email: [enquiries@environment-agency.gov.uk](mailto:enquiries@environment-agency.gov.uk)  
[www.environment-agency.gov.uk](http://www.environment-agency.gov.uk)

© Environment Agency

All rights reserved. This document may be reproduced with prior permission of the Environment Agency.

