Regional analysis of groundwater droughts using hydrograph classification

J.P. Bloomfield[1], B.P. Marchant[2], S.H. Bricker[2] and R.B. Morgan[3]

[1]{British Geological Survey, Wallingford, United Kingdom}

[2]{British Geological Survey, Keyworth, United Kingdom}

[3]{Environment Agency, Lincoln, United Kingdom}

Correspondence to: J. P. Bloomfield (jpb@bgs.ac.uk)

Abstract

Groundwater drought is a spatially and temporally variable phenomenon. Here we describe the development of a method to regionally analyse and quantify groundwater drought. The method uses a cluster analysis technique (non-hierarchical *k*-means) to classify standardised groundwater level hydrographs (the Standardised Groundwater level Index, SGI) prior to analysis of their groundwater drought characteristics, and has been tested using 74 groundwater level time series from Lincolnshire, UK. Using the test data set, six clusters of hydrographs have been identified. For each cluster a correlation can be established between the mean SGI and a mean Standardised Precipitation Index (SPI), where each cluster is associated with a different SPI accumulation period. Based on a comparison of SPI time series for each cluster and for the study area as a whole, it is inferred that the clusters are independent of the driving meteorology and are primarily a function of catchment and hydrogeological factors. This inference is supported by the observation that the majority of sites in each cluster are associated with one of the principal aquifers in the study region. The groundwater drought characteristics of the three largest clusters, that constitute ~80% of the sites, have been analyzed. There are differences in the distributions of drought duration, magnitude and intensity of groundwater drought events between the three clusters as a function of autocorrelation of the mean SGI time series for each cluster. In addition, there are differences between the clusters in their response to three major multi-annual droughts that occurred during the analysis period. For example, sites in the cluster with the longest SGI autocorrelation experience the greatest magnitude droughts and are the slowest to recover

from major droughts, with groundwater drought conditions typically persisting at least six months longer than at sites in the other clusters. Membership of the clusters is shown to be related to unsaturated zone thickness at individual boreholes. This last observation emphasises the importance of catchment and aquifer characteristics as (non-trivial) controls on groundwater drought hydrographs. The method of analysis is flexible and can be adapted to a wide range of hydrogeological settings while enabling a consistent approach to the quantification of regional differences in response of groundwater to meteorological drought.

1. Introduction

Groundwater drought is a type of hydrological drought characterised by sustained low groundwater levels, reduced base flow and reduced flows to springs and groundwater-fed rivers and wetlands (Van Lanen and Peters, 2000; Tallaksen and Van Lanen, 2004; Mishra and Singh, 2010; Van Loon, 2015). Like other hydrological aspects of drought, groundwater droughts are not a simple function of meteorological drivers. The impact of droughts on regional groundwater resources can vary in space and time. This is because the response of groundwater systems to meteorological droughts, through changes in groundwater levels and baseflow to groundwater supported rivers, is influenced by spatial variations in intrinsic catchment and aquifer characteristics and processes. These include highly non-linear unsaturated zone processes, recharge, and saturated groundwater storage, flow and discharge over a range of space and time scales (Tallaksen et al, 2009; Bloomfield and Marchant, 2013; Van Lanen et al., 2013; Van Loon and Laaha, 2015).

In order to improve the design and operation of groundwater drought monitoring networks, the analysis and interpretation of data from such networks, and, more generally, water resource management at the onset, during and after episodes of groundwater drought, there is a need for a much better understanding of the heterogeneous spatio-temporal response of aquifers to major meteorological droughts (Bloomfield and Marchant, 2013). This includes the need for robust methods to systematically characterise and quantify the heterogeneous response of groundwater to meteorological droughts at a regional scale prior to investigation and attribution of the causes of any heterogeneous response. Despite extensive work on the regional analysis of meteorological and other hydrological droughts, to date there has been no systematic investigation of heterogeneities in groundwater droughts at the regional scale. This paper describes the application of one such suite of methods to regionally analyse groundwater level hydrographs and to assess variations in the spatial response of groundwater to meteorological droughts using a case study from the UK.

1.1 Controls on spatial heterogeneity in groundwater drought

A few previous studies have presented evidence for the spatially heterogeneous response of groundwater to meteorological droughts. To help develop an optimal monitoring network for groundwater resources under drought conditions, Chang and Teoh (1995) described the heterogeneous response of groundwater levels at 13 observation boreholes to meteorological droughts across a basin in Ohio, USA, although they did not investigate the hydrogeological causes of the heterogeneity. Van Lanen (2005) and Van Lanen and Tallaksen (2007) observed that drought characteristics derived from groundwater levels have ‘spatial effects’, and noted that these spatial effects on groundwater drought are an important consideration when monitoring droughts using groundwater levels. Van Lanen and Tallaksen (2007) compared modelled groundwater recharge and discharge for a humid continental climate (Missouri, USA) and a tropical savannah climate (Guinea) for quick- and slow-responding catchments and showed that both climatology and the responsiveness of the catchment as defined by the aquifer characteristics have an influence on drought generation. Peters et al. (2006) investigated the propagation and spatial distribution of aspects of modelled groundwater drought, including recharge, groundwater level and groundwater discharge in the Pang catchment in the UK. They found that short droughts in groundwater levels were most severe near streams and were attenuated with distance from the streams; longer periods of below average recharge had more effect on suppressing groundwater levels on interfluves near groundwater divides, and that droughts in groundwater discharge are more attenuated upstream and less so downstream in the catchment. Tallaksen et al. (2009) also modelled the spatio-temporal response of the Pang catchment to drought events and found large differences between the spatio-temporal response of groundwater recharge, level and discharge and the driving meteorological droughts, where droughts in groundwater recharge and levels were found to cover relatively small areas, but last longer, than the meteorological droughts.

Mendicino et al. (2008) developed a groundwater resource index for drought monitoring and forecasting based on a simple distributed runoff/water balance model, and evaluated the use of the index in three catchments in southern Italy. They found that the groundwater resource index was highly spatially variable and related it to variations in hydraulic conductivity across the catchments. Using a newly developed groundwater drought index, the

Standardised Groundwater level Index (SGI), Bloomfield & Marchant (2013) also investigated hydrogeological controls on groundwater drought. Based on 14 observation boreholes in different catchments across England, UK, they showed that groundwater drought duration depended on the autocorrelation structure of SGI time series. This was in turn inferred to be both a function of spatially varying recharge processes and saturated flow processes within the local aquifer systems.

1.2 Regional analysis of groundwater drought

There has been significant work on the regional analysis of meteorological and other hydrological droughts. Cluster Analysis (CA), Principal Component Analysis (PCA) or some combination of both techniques have been used extensively by meteorologists and hydrologists to investigate the spatio-temporal distribution of hydrological variables, including drought indices (e.g. Klugman ,1978; Karl and Koscienly, 1982; Eder et al., 1987; Stahl and Demuth, 1999; 2001, Lana et al., 2001; Bonaccorso et al., 2003; Vincente-Serrano, 2006; Vicente-Serrano and Cuadrat-Prats, 2007; Raziel et al., 2008; Santos et al., 2010; Fleig et al., 2011; Hannaford et al., 2011; Lorenzo-Lacruz et al., 2013).

Although not previously applied to groundwater drought, CA and/or PCA techniques have been used to classify groundwater level hydrographs for a range of purposes. Winter et al. (2000) classified groundwater hydrographs from three small lake-dominated catchments to investigate groundwater recharge and differences in the hydrographs as a function of the geology of the catchments. Similarly, Moon et al. (2004) applied PCA to 66 groundwater level hydrographs from South Korea to characterise the spatial variability in groundwater recharge. Upton and Jackson (2011) used CA and PCA (following a methodology developed by Hannah et al., 2000) with 52 groundwater level hydrographs from the Pang and Lambourn catchments in the UK to produce regional or 'master' hydrographs for modelling the spatial distribution of groundwater flooding.

Here we present the first systematic regional analysis of groundwater droughts using a case study from Lincolnshire, UK. The case study consists of 74 groundwater hydrographs from an area of approximately 8,000 km^2 that includes three regionally important aquifers, the Lincolnshire Limestone, the Chalk and the Spilsby Sandstone aquifers, each with contrasting aquifer characteristics (section 2). The groundwater hydrographs have been normalised using the Standardised Groundwater level Index (SGI) technique of Bloomfield & Marchant (2013) and groups or clusters of similar groundwater hydrographs have been identified using CA, where hydrogeologically meaningful clusters are identified by explicitly searching for groups of hydrographs that can be explained by *a posteriori* knowledge of the groundwater system (section 4.2). The drought characteristics of the clusters have been quantified in terms of drought event duration, magnitude and intensity and the impact of the three major, multi-annual droughts on the SGI time series has been investigated (section 4.4). Controls on the groundwater drought response in each of the clusters have been explored and the results briefly discussed in terms of the implications for monitoring and managing groundwater droughts (section 5).

2. The case study

The case study area of Lincolnshire is situated in the east of England, UK. It is bounded by the North Sea to the east, the Wash estuary to the south and the Humber Estuary to the north (Fig. 1). The area is predominantly rural with highly productive agricultural and horticultural land, fens and estuarine wetlands. Lincoln, Boston and Scunthorpe are the principal small conurbations in the study area. The land is generally flat and low-lying, typically less than 30 m above sea level (m asl), apart from the Chalk of the Lincolnshire Wolds and the Lincolnshire Limestone outcrop which form northwest-southeast trending escarpments that reach elevations of approximately 150 m asl and 70 m asl respectively.

2.1 Hydrometeorology and drought history

As a first-order approximation, it is assumed that the broad meteorological drought history of the study area is spatially homogeneous. This assumption means that any relative differences in drought histories between sites or clusters need to be explained in terms of catchment or hydrogeological factors, rather than differences in the drought climatology. This assumption is tested as part of the analysis of correlations between precipitation and regional groundwater levels (see section 4.2). It is also supported by the observations that the whole study area is governed by the same broad climatic patterns, i.e. rain-bearing low pressure systems from the Atlantic and high pressure systems leading to a lack of rainfall, with only small variation in annual precipitation across the region (Marsh and Hannaford, 2008). The assumption is also consistent with the previously documented spatial coherence of major hydrological (surface water) droughts in the UK (Hannaford et al., 2011; Fleig et al., 2011; Folland et al., 2015) where the current study area falls within a homogeneous drought region (“region 4” of Hannaford et al., 2011, “region GB4” of Fleig et al., 2012; Kingston et al., 2013, and the "English Lowlands" of Folland et al., 2015) although it is noted that the effects of landscape processes can cause heterogeneous meteorological signals to become attenuated (Van Loon, 2015).

Mean annual rainfall varies across the study area from about 600 to 700 mm (Marsh and Hannaford, 2008). The groundwater hydrographs used in the study have been analysed from 1983 to 2012. During this period, three multi-annual episodes of drought have previously been documented by Marsh et al. (2007; 2013), Kendon (2013), Parry and Marsh (2013) and Folland et al. (2015) as follows: 1988 to 1992, 1995 to 1997 and 2010 to 2012. All are known to have been major drought events causing reduced surface flows and suppressed groundwater levels throughout large areas of central, eastern and southern UK as well as over parts of North West Europe (Lloyd-Hughes and Saunders, 2002; Lloyd-Hughes et al., 2010; Hannaford et al., 2011; Fleig et al., 2012 and Kingston et al., 2013).

2.2 Geology and hydrogeology

The study area consists of a sequence of Jurassic and Cretaceous aquifers separated by low permeability clay and shale units. The whole sequence generally dips gently eastwards and where each of the aquifer units passes under an overlying low permeability formation they typically become confined. The whole sequence is unconformably overlain by Quaternary superficial deposits. Figure 1 shows the distribution of the three main aquifers in the region: the Jurassic Lincolnshire Limestone; the Lower Cretaceous/Upper Jurassic Spilsby Sandstone, and the Upper Cretaceous Chalk, and includes a schematic cross-section of the hydrostratigraphy of the study area. These aquifers are hydrogeologically distinct from each other, and two of them, the Lincolnshire Limestone and the Chalk have previously documented spatially variability. Below we summarise these features as they inform the heuristic rules used in section 4.2 to guide the selection of clusters as part of the CA.

The Lincolnshire Limestone Formation is an oolitic limestone with fine-grained, micritic and peloidal units (Allen et al., 1997), and is up to 40 m thick at outcrop in the west. It dips and thins to the east where it becomes confined and eventually pinches out down-dip. Maximum unsaturated zone thickness is up to about 45m towards the southwest of the outcrop. Groundwater movement is almost entirely by fracture flow along well-developed bedding plane fractures and joints. Abstraction takes place mainly from the region immediately to the east of the outcrop. It has highly variable transmissivities and storage coefficients typical of a fractured limestone. Allen et al. (1997) have reported a wide range of transmissivity values for the Lincolnshire Limestone with an interquartile range of 260 to 2260 m^2 d^{-1} and a geometric mean of 660 m^2 d^{-1}, with slightly higher transmissivities being reported from the south of the region, and a very wide range of storage coefficients from $2x10^{-7}$ to 0.58.

The Spilsby Sandstone aquifer is up to about 30 m thick consisting of a variably, but often poorly cemented pebbly quartz sandstone with alternating thin clays and marls (Whitehead and Lawrence, 2006). It outcrops along the foot of the Wolds escarpment (Fig. 1) where it is associated with springs and maximum unsaturated zone thickness is about 30m. It dips to the east and away from outcrop and it is generally confined by clays above and below (Fig. 1). Jones et al., (2000) reported transmissivity values in the range 130 to 170 m^2 d^{-1}, and a geometric mean of 140 m^2 d^{-1} with storage coefficients ranging from $1x10^{-4}$ to $1x10^{-3}$ and with a geometric mean of $4x10^{-4}$.

The Chalk is a microporous fractured limestone (Bloomfield et al, 1995). Storage and transmissivity are controlled by local sub-karstic development of the fracture network (Bloomfield, 1996; Maurice et al., 2006). The Chalk Group reaches a thickness of over 250 m. Groundwater flows from the recharge areas in the west eastward down dip towards and into the confined Chalk to the east. The Chalk bedrock surface was significantly altered during the Ipswichian interglacial of the Quaternary. As a result of glacial activity a cliff line and wavecut platform were eroded into the Chalk (Fig. 1). The Chalk to the east of the palaeo-cliff line is now buried beneath a covering of till, sand and gravel superficial deposits (Whitehead and Lawrence, 2006). Maximum unsaturated zone thickness occurs towards the northwest of the Chalk outcrop and is about 60m contrasting with the relatively thin unsaturated zone to the east of the palaeo-cliff line. Allen et al. (1997) and Whitehead and Lawrence (2006) have reported that transmissivity values differ between the northern and southern Chalk in Lincolnshire. In the northern part of the region transmissivity has an interquartile range of 1020 m^2 d^{-1} to 6070 m^2 d^{-1} with a geometric mean of 2350 m^2 d^{-1}, whereas in the southern area, in the region of the eroded Chalk, transmissivity is slightly reduced and has an interquartile range of 850 m^2 d^{-1} to 3010 m^2 d^{-1} with a geometric mean of 1380 m^2 d^{-1}. Similarly, Allen et al. (1997) report storage coefficients with an interquartile range of $3.5x10^{-5}$ to $1.5x10^{-3}$ and with a geometric mean of $2x10^{-4}$ for the northern Chalk and $6.1x10^{-5}$ to $2.7x10^{-3}$ and with a geometric mean of $1.5x10^{-3}$ for the southern Chalk.

The Quaternary superficial deposits in the study area comprise: glaciofluvial sand and gravels and tills; peat; tidal flat deposits; river terrace sands and gravels, and overlying alluvium.

The Lincolnshire Limestone Formation and the western part of the Chalk outcrop are largely absent of superficial cover.

3. Data and Methods

3.1 Data

Groundwater level data for the 74 observation boreholes (Fig. 1) has been provided by the Environment Agency from their groundwater level monitoring network database (Environment Agency, 2014). Prior to the study none of the sites were believed to be significantly impacted by abstraction although all three regional aquifers are used for public water supply, abstractions for agricultural irrigation and industrial use (Allen et al., 1997; Whitehead and Lawrence, 2006). Where observation boreholes penetrate both the Chalk and underlying Spilsby Sandstone aquifer, the boreholes are completed with screens so that they monitor water levels in only one of the two aquifers. Groundwater levels have been recorded over a range of frequencies, but typically at weekly to monthly time steps. Based on the raw groundwater level data, mean monthly groundwater levels have been estimated. If no observations were available for a given month then a linear interpolation was used to estimate the monthly groundwater levels following the method described by Bloomfield and Marchant (2013).

Precipitation data has been taken from the Centre for Ecology and Hydrology's Continuous Estimation of River Flows (CERF) 1km gridded precipitation dataset (Keller et al., 2005; Dore et al., 2012; Bloomfield and Marchant, 2013). CERF daily gridded precipitation data is generated from rain gauge data held in the UK Met Office national precipitation monitoring network. A triangular planes methodology is used to produce a daily 1km^2 grid based on a weighted average (inverse distance) of the three nearest rain gauges. Daily rainfall is then summed to give total monthly gridded rainfall. The precipitation data that is used with each groundwater level observation site is the monthly total for the CERF 1km^2 grid square that contains the given groundwater observation borehole.

3.2 Methods

3.2.1 Hydrograph normalisation using the SGI method

The groundwater level hydrographs have been normalised to the Standardised Groundwater level Index (SGI) of Bloomfield and Marchant (2013). This is a non-parametric normalization of data that assigns a value to the monthly groundwater levels based on their rank within groundwater levels for a given month from a given hydrograph. The normal scores transform is undertaken by applying the inverse normal cumulative distribution function to n equally spaced p_i values ranging from 1/(2*n*) to 1 - 1/(2*n*). The values that result are the SGI values. They are then re-ordered such that the largest SGI value is assigned to the i for which p_i is largest, the second largest SGI value is assigned to the i for which p_i is second largest and so on. In summary, for each of the 74 study sites, normalized indices are estimated from the groundwater level data for each calendar month using the normal scores transform. These normalized indices are then merged to form a continuous SGI. Precipitation records for each site have also been normalised. At each site a version of the Standardised Precipitation Index (SPI) after McKee et al. (1993) has been estimated for precipitation accumulation periods of 1, 2, ..., 36 months. For consistency between groundwater and precipitation indices, SPIs are estimated using the normal scores transform applied to accumulated precipitation data for each calendar month.

3.2.2 Cluster analysis

Cluster Analysis (CA) attempts to identify clusters of similar individuals amongst a multivariate dataset. In the context of this paper CA is used to form clusters of groundwater level hydrographs which exhibit similar fluctuations in their SGI time series. A wide range of CA algorithms exist. They are most coarsely distinguished according to whether or not they assume that the resultant clusters are hierarchical. Given the wide variety of algorithms it is difficult to decide upon the best approach to cluster a particular dataset. Webster and Oliver (1990) stress that this decision is rather subjective, although previous studies that have used CA to cluster hydrographs have typically justified their choice of algorithm by claiming that some produce more physically interpretable groupings. For example, Hannah et al. (2000) used the agglomerative hierarchical average linkage algorithm as they thought it was more interpretable than alternatives such as the centroid and Ward's clustering procedures. Webster and Oliver (1990) recommend that multiple clustering algorithms should be applied and expert knowledge of the system being investigated used to decide which set of clusters is most relevant. In this paper we adapt this approach by applying one hierarchical and one non-hierarchical method.

Hierarchical classifiers require a measure of the similarity (or dissimilarity) between each pair of individuals. Common examples include the Euclidean distance or the correlation between the measurements of the individuals. The pairwise similarities between s individuals are expressed in a $s \times s$ matrix **B**. A mathematical criterion is then used to allocate the individuals to different clusters in a manner that maximizes the similarity between the individuals within the groups whilst minimizing the similarity between individuals in different clusters. For our hierarchical clusters we measure the similarity between groundwater level hydrographs by the correlation matrix of their SGI time series and then apply the agglomerative hierarchical complete-linkage strategy (Webster and Oliver, 1990) to merge the boreholes into clusters.

We also apply the commonly used non-hierarchical *k*-means clustering algorithm. It is widely used in spatial analysis studies, for example, Santos et al. (2010), Raziei et al. (2012) and Sadri et al. (2014) have all used the *k*-means clustering algorithm to investigate the regional characteristics of droughts. The approach partitions the individuals into a specified number of clusters. A numerical optimization routine is used to select the partitioning which maximizes the similarity between each individual and the centroid of the cluster in which it is contained. Again there is flexibility in the choice of similarity measure and the manner in which the centroid of a cluster is calculated. We use the squared Euclidean distance between the vectors of time series observations from each site to assess similarity and define the centroid of a cluster as the multi-dimensional mean of the time series within the cluster.

Clustering methods do not produce a unique partitioning of a given data set on their own, and for both the hierarchical and non-hierarchical approaches there remains the issue of deciding upon the optimal number of clusters. This can be achieved by asking an expert on the system in question to compare the attributes of clusterings consisting of a different number of groups. Here we use a rule-based approach to help identify the number of clusters based on knowledge of the general hydrogeology of the study area. Bloomfield and Marchant (2013) have previously shown that groundwater drought characteristics are a function of unsaturated zone thickness in fractured aquifers such as the Lincolnshire Limestone and Chalk aquifers, and that when a broader range of aquifer types are considered groundwater drought characteristics are also a function of the hydraulic diffusivity of aquifers. Here we use these observations and knowledge of the spatial variation in these features across the three aquifers in the study area (section 2.2) to design rules to aid in the selection of clusters. The rules adopted for the current study are to identify the smallest number of clusters that: i.) broadly resolve the spatial distribution of the three aquifers across the study region, ii.) given the previously documented N-S variation in aquifer properties and unsaturated zone thickness across the Lincolnshire Limestone aquifer (Allen et al., 1997), that distinguish more than one

region of the Lincolnshire Limestone, and iii.) given variations in aquifer properties and unsaturated zone thickness across the Chalk aquifer both N-S and across the buried cliff line (Allen et al., 1997), that distinguish more than one region of the Chalk. Note that this set of rules is specific to the current study, however, for any given study area the target number of classes and hence the rules used can be adapted to reflect the regional hydrogeology and in particular any knowledge of heterogeneity in the aquifer systems under investigation.
However, mathematical criteria can also be used as a guide to clustering. We also calculate the RMSSD, the square root of sum of the squared Euclidean distance between each individual and the centroid of the group to which it is allocated. In combination with expert judgement related to the system under consideration, it is common practice to inform the choice of the number of clusters using plots of RMSSD versus cluster number. Since RMSSD decreases non-linearly as the number of clusters increases, a cluster number is selected associated with a decrease in the rate of RMSSD decline.

3.2.4 Autocorrelation structure of the SGI time series

Bloomfield and Marchant (2013) demonstrated the importance of the autocorrelation structure of SGI time series for groundwater drought studies by establishing a relationship between the range of significant autocorrelation in the SGI series, $m_{\max}$, and corresponding SPI. They showed that $m_{\max}$ scales linearly with $q_{\max}$, where $q_{\max}$ is the SPI accumulation period which leads to the strongest correlation between SGI and SPI. Both $m_{\max}$ and $q_{\max}$ are also used here to characterise and quantify groundwater droughts within each of the clusters of groundwater hydrographs and have been estimated as follows.

If the mean SGI for a borehole is denoted by $\overline{\mathrm{SGI}}$ then the kth sample autocovariance coefficient is defined to be

$$g_k = \frac{1}{n}\sum_{i=k+1}^{n}\{\mathrm{SGI}(i) - \overline{\mathrm{SGI}}\}\{\mathrm{SGI}(i-k) - \overline{\mathrm{SGI}}\} \qquad (1)$$

and the kth sample autocorrelation coefficient is

$$r_k = \frac{g_k}{g_0} \qquad (2)$$

where g_0 reduces to the population variance function (see Eqn. 1 when $k = 0$). The correlogram is a plot of r_k against k. If there is no correlation between the $\mathrm{SGI}(i)$ observed k months apart and if the SGI values are normally distributed then r_k is approximately normally distributed with mean zero and variance $1/n$. Therefore values of r_k with magnitude greater than $2/\sqrt{n}$ indicate significant correlation at approximately the 5 % level. We define the range of significant temporal correlation of a SGI time series to be the largest m, $m_{\max}$, for which $r_k > 2/\sqrt{n}$ for all $k \leq m$. Since all of our groundwater records are of $n = 355$ months the threshold on r_k is equal to 0.11.To estimate $q_{\max}$, Pearson correlation coefficients are calculated between SGI and SPI with accumulation periods of $q = 1, 2, \ldots, 36$ months and the accumulation period associated with the maximum correlation gives $q_{\max}$.

4. Results

4. 1 Identification of regional droughts from average SPI and SGI time series

Before undertaking the regional drought analysis, the correlation between mean SPI and SGI for the entire region, based on all 74 sites, has been investigated and the large-scale drought history of the study area has been defined.

Figure 2a is a heatmap showing the correlation coefficient between SPI for precipitation accumulation periods q = 1 to 36 months and SGI for lags between SPI and SGI of 0 to 5 months based on average values of SPI and SGI for all 74 sites. Dark blue denotes zero correlation and dark red a perfect correlation. Figure 2a shows that there is a good correlation between SPI and SGI. The strongest correlation (0.84, denoted by the closed black circle in Fig. 2a) is for a precipitation accumulation period ($q_{\max}$) of 12 months (SPI_{12}) with no lag between the SGI and SPI time series. This is consistent with the observations of Bloomfield and Marchant (2013) who previously reported $q_{\max}$ for a variety of groundwater hydrographs from the UK with an average of 13 months and Folland et al. (2015) who reported a $q_{\max}$ of 12 months for aggregated time series representing the English Lowlands. Figures 2b and 2c, the average SPI_{12} and SGI time series respectively, have similar features. For example, episodes of high groundwater levels in 1983, 1994, 2002, and 2008 correspond with high values of SPI_{12}. Three episodes of regionally significant groundwater drought associated with prolonged low groundwater levels from October 1988 to November 1993, May 1995 to February 1998, and from August 2010 to August 2012 correspond closely with episodes of meteorological drought in the SPI_{12} time series and are consistent with those identified by previous studies (Lloyd-Hughes and Saunders, 2002; Marsh et al., 2007; 2013; Kendon, 2013; Hannaford et al., 2011; Parry and Marsh, 2013; Folland et al., 2015). It is inferred from these observations that the large-scale drought history of the study area is represented well by the average SPI_{12} and SGI time series.

4.2 Regional analysis of the SGI hydrographs

CA has been used to analyse the heterogeneous response of groundwater to droughts across the study region. Clustering has been undertaken using both an agglomerative hierarchical complete-linkage algorithm and a non-hierarchical *k*-means clustering algorithm and the resulting clusters searched for those that are hydrogeologically meaningful and that can be explained by known features of the catchment and groundwater systems. Figure 3a is a dendrogram that fully illustrates the level of similarity between individuals within the clusters formed by the hierarchical clustering. The number of clusters is controlled through the threshold on the distance between groups. For example, a threshold of 0.62 leads to the six clusters shown in Fig. 3b. Figure 3c is an equivalent map showing the distribution of sites by clusters formed by *k*-means clustering for $k = 6$.

Figures 3b and 3c shows that the spatial distribution of sites as a function of the clusters formed by the hierarchical and non-hierarchical approaches are broadly similar, so the choice of clustering algorithm is based on a plot of RMSSD against number of clusters. Figure 4 shows that the RMSSD for the *k*-means clustering is systematically lower than that for the hierarchical clustering algorithm where there are three clusters or more, so we have chosen to use the non-hierarchical *k*-means clustering approach. Note also that both clustering algorithms are better than a clustering scheme based solely on the three classes of aquifer (e.g. Lincolnshire Limestone, Chalk and Spilsby Sandstone). However, an optimal number of *k*-mean clusters is not clearly evident in Fig. 4. After careful inspection of the clusters formed by a range of *k*-means clustering classes and a consideration of the study specific clustering rules described in section 3.2.2, $k = 6$ was selected. Based on *k*-means clustering where $k = 6$, Fig. 3c shows the distribution of sites between the six clusters (cluster 1 to cluster 6, or CL1, ... CL6).

It can be seen from Fig. 3c that the resulting *k*-means clusters have a degree of spatial coherency. We have previously assumed that such spatial correlations in the SGI time series are primarily a function of catchment and hydrogeological factors and not a consequence of heterogeneity in the driving meteorology. Here we test if this is the case, prior to further exploration of the features of each cluster, by investigating if precipitation associated with each cluster is substantially different from regional average precipitation. To do this, we first need to identify a representative accumulation period, $q_{\max}$, for precipitation for each cluster.

Figure 5 is a set of heatmaps, similar to Fig. 2a, showing the correlation between SPI for precipitation accumulation periods, q, 1 to 36 months, and SGI for lags between SPI and SGI time series of 0 to 5 months for each of the six clusters. Dark blue denotes zero correlation and dark red a perfect correlation with the strongest correlation for each cluster marked by the closed black circle. Table 1 gives q_{max} for each cluster and also gives the maximum associated correlation coefficient. In all cases, except CL2, the maximum correlation between SPI and SGI is found where there is no lag between the two time series. For CL2 it is found at a lag of one month. The highest correlations are for CL2, CL4 and CL1 at 0.86, 0.82 and 0.74 respectively. The correlations for CL3 and CL5 are moderate (0.36 and 0.53) and for CL6 there is effectively no correlation (0.09). This is consistent with the observations made in section 4.3 below that linear trends in CL3 and CL5 appear to affect the SGI time series and that the SGI hydrograph for CL6 appears to be anomalous, departing from the mean regional SGI and SPI signals. Values of q_{max} for CL1 to CL5 from Fig. 5 are 4, 16, 15, 9, and 17 months respectively. Based on these, Fig. 6 shows SPI time series for each cluster, where black lines are the mean SPI for the cluster and the red lines are average SPI across the study area based on the same cluster-specific q_{max}. Since Fig. 6 illustrates that the two SPI time series for each cluster are similar, we infer that heterogeneity in the driving meteorology across the study region, or at least between the clusters as defined here, does not play an important role in the clustering process and that membership of clusters is dominated by catchment or hydrogeological factors.

4.3 Characteristic features of the SGI hydrograph clusters

Figure 7 shows the mean SGI time series for each cluster. Two main qualitative observations can be made regarding the SGI hydrographs. Five of the six clusters have a similar overall form to the mean SGI hydrograph for the region (Fig. 2c) showing common patterns of low (and high) groundwater level stand. Whereas, CL6 appears to be an exception with a different overall form to the SGI hydrograph – it also exhibits an anomalous step change in SGI from drought to high groundwater level stand over an eight month period from May 1990 to December 1990. Secondly, two of the clusters, CL3 and CL5, appear to show declining linear trends in SGI making direct comparison of drought histories between these and other clusters problematic.

Bloomfield & Marchant (2013) have previously shown that m_{max}, a measure of the significant autocorrelation length of SGI time series, relates to features of groundwater drought. A similar analysis of autocorrelation structure of SGI time series for each cluster is presented here. Figure 8 shows autocorrelation plots for SGI hydrographs for each of the six clusters. In each figure the pale grey lines are autocorrelation plots for individual sites and the solid black line is the autocorrelation plot for the mean SGI time series for the cluster with the horizontal dashed line indicating the significant level of autocorrelation based on the record length. Based on these plots, values of m_{max} for the mean SGI time series for each cluster are given in Table 1. Values of m_{max} for CL3, CL5 and CL6 are anomalously large, consistent with the anomalous features of these SGI hydrographs described above. For the remaining clusters, Figure 8 and Table 1 show that CL1 has the shortest autocorrelation of 15 months. In comparison, CL2 has an autocorrelation of 23 months and CL4 is intermediate at 18 months.

These contrasting characteristics between the clusters can be seen clearly in Fig. 9a which illustrates SGI time series for all sites within each cluster, grouped in their respective clusters, and presented in the form of a heatmap where low values of SGI (associated with drought conditions) are in shades of green to red (increasing drought intensity) and episodes of high groundwater level stand are in shades of green to blue (increasing high groundwater levels). The three major episodes of drought can be seen clearly in the heatmaps for CL1, CL2 and CL4, but are obscured by the trends in CL3 and CL5 and absent in CL6. The degree of coherency of individual SGI time series within each cluster also appears to be consistent with differences in autocorrelation between the clusters. Figure 9b is a heatmap of the cross-correlation coefficients for all the individual SGI time series ordered as a function of the six clusters, where dark red denotes high correlations and dark blue denotes low correlations. Sites within CL1 and CL4, clusters with moderate or short autocorrelation, show relatively low levels of internal coherency compared with sites in CL2 with relatively long autocorrelation that are highly correlated.

Based on the above, the following is a summary of the features of each cluster:

- CL1 is dominated by sites from the northern parts of the Lincolnshire Limestone. The mean SGI time series of CL1has a relatively short autocorrelation (m_{max} of 15 months) and within the cluster SGI hydrographs are relatively variable.
- CL2 is dominated by sites from the northern part of the Chalk. The cluster has the longest mean SGI autocorrelation (m_{max} of 23 months) and hydrographs within CL2 are highly correlated indicating a high degree of coherency in groundwater levels across the northern part of the Chalk in the study area.

- CL3 is a relatively small cluster of six sites, four of which are from the confined Spilsby Sandstone and two from the Lincolnshire Limestone. The main feature of the cluster is a trend in decreasing SGI across the observational record. This trend is consistent with a previous water balance assessment for the Spilsby Sandstone (Whitehead and Lawrence, 2006) where annual groundwater deficits have been reported. The sites in this cluster are inferred to be possibly variably impacted by long-term abstraction. Given this inference and the small size of the cluster of sites, CL3 is not included in the subsequent analysis of groundwater droughts.
- CL4 is dominated by sites from the southern Lincolnshire Limestone and also includes five unconfined sites on the southern Chalk and one site located in the northern Lincolnshire Limestone. It has a moderate autocorrelation, m_{max} of 18 months. Individual SGI hydrographs within the cluster show a moderate degree of coherency.
- CL5 is a small cluster of five sites all from the southeastern Chalk to the east of the palaeo-wave cut platform and are the five sites closest to the coast. It has a moderately long autocorrelation, m_{max} of 28 months that may be affected by an apparent weak trend in declining SGI - there is only a weak correlation between SPI and SGI. Given the small size of the cluster and the apparent trend in mean SGI, CL5 is not included in the subsequent analysis of groundwater droughts.
- CL6 consists of three SGI hydrographs from the confined Spilsby Sandstone aquifer. The hydrographs are characterised by an anomalous step change in SGI from drought to high groundwater level stand over an eight month period from May 1990 to December 1990. The mean SGI hydrograph shows no correlation with the other five clusters and there is no correlation between SPI and SGI within the cluster. All three sites are within a radius of about 3 km of a public water supply borehole and it is inferred that groundwater levels may be influenced by abstraction. So, as with CL3 and CL5, this very small cluster is not included in the subsequent analysis of groundwater droughts.

4.4 Analysis of droughts using the hydrographs from CL1, 2 and 4

Clusters CL1, CL2 and CL4 consist of 61 of the 74 hydrographs analysed. Here the characteristics of groundwater droughts in these clusters are quantified and the response of the clusters to three major drought episodes is investigated.

The duration, magnitude and mean intensity of groundwater drought events have been investigated based on an analysis of the SGI hydrographs where, following the convention of

McKee et al. (1993), negative values of SGI denote drought conditions (note, however, that the current convention of the World Meteorological Organisation for SPI refers to drought conditions where SPI is continuously negative and reaches and intensity of -1.0 or less and that negative values between 0 and -1 are classified as near normal and simply indicate less than a median precipitation, World Meteorological Organisation, 2012). Groundwater drought duration, D, is taken to be the total number of consecutive months where SGI is negative. Groundwater drought magnitude, M, is taken to be the total cumulative value of monthly SGI for a given drought event, and mean drought intensity, I, is given by M/D. Summary drought statistics for CL1, CL2 and CL4 are given in Table 2.

Table 2 shows that there are differences in the character of the groundwater drought events in the SGI hydrographs for clusters CL1, CL2 and CL3. For example, CL1 has more than twice the number of drought episodes (39 episodes) than CL2 (15 episodes) and the average and maximum duration of droughts in CL1 (4.6 and 27 months respectively) are less than half those of CL2 (11.3 and 61 months). The mean drought event magnitude in CL1 (-2.9) is less than half that in CL2 (-7.9) and the mean drought event intensity in CL1 (-0.43) is almost twice that of CL2 (-0.28). In all cases, the drought event statistics for CL4 fall between those for CL1 and CL2. In summary, CL1 exhibits shorter, but generally more intense drought episodes compared with CL2, with CL4 drought events being of intermediate character. These relative drought phenomena are a consequence of the degree of autocorrelation in the respective SGI time series, where CL1 has a relatively short autocorrelation compared with relatively long autocorrelation for CL2. This observation is consistent with previous site specific and modelling studies that noted a similar relationship between the ‘flashiness’ or responsiveness of the groundwater system to meteorological divers and the number of droughts, where quickly responding groundwater systems typically experience more droughts than more slowly responding catchments (Peters et al. 2003; Van Loon and Van Lanen, 2012; Van Lanen et al. 2013).

There is a strong relationship between drought duration and magnitude for all three clusters, Fig. 10, where longer episodes of groundwater drought are associated with droughts of greater magnitude. However, there is no such regular or simple relationship between drought duration and intensity. Maximum drought intensity is similar for all three clusters, for CL1, CL2 and CL4 it is -1.10, -1.05 and -1.13 respectively (Table 2 and Fig. 11), and is associated with two of the major drought events, i.e. with the latter part of the 1988 to 1993 drought for CL2, and the 2010 to 2012 drought for CL1 and CL4. Figure 11 shows the empirical distribution of D, M and I for clusters CL1, CL2 and CL4. Drought duration (Fig. 11) in all three clusters is highly positively skewed with many short drought events and relatively few long drought events. As previously noted, the longest duration droughts are associated with CL2, the cluster with the longest autocorrelation in the SGI time series. These observations are consistent with those of Hisdal and Tallaksen (2003), Tallaksen et. al. (2009) and Fleig et al. (2011) who have also described strongly skewed distributions of hydrological drought durations.

Three major, multi-annual droughts have already been described from the regional (Fig. 2) and the cluster-specific (Figs. 7 and 9a) SGI time series. Table 3 summarises differences in the relationships between the driving meteorology and the drought characteristics of each cluster for the three major droughts. Each of the major drought episodes have been quantified using drought characteristics as applied to SPI_{12} and SGI for each of the clusters.

The 1988-1993 event was the longest of the three major droughts and consequently had the greatest drought magnitude. The groundwater and meteorological droughts start approximately contemporaneously in the winter of 1988. In CL2 the drought was continuous with negative SGI from November 1988 to November 1993, whereas in CL4 there were two short breaks in the drought and numerous breaks in the drought in CL1. In CL2 there was a gradual intensification in the drought magnitude across the event, peaking in June 1992 at an SGI of -1.85 (four months after the peak SPI_{12} meteorological drought). In contrast, not only were there short breaks in the drought in CL1 and CL4 but there were approximately annual cycles of drought intensification and decline over the four year period – these were particularly pronounced in CL4. This is seen in Fig. 9a where between 1988 and 1993 the drought status of CL4 is designated by the red tones in the heatmap, but that these tones show a series of approximately annual variations giving the appearance of vertical stripes during that period and within that cluster. However, the most pronounced differences in response to major droughts between clusters CL1, CL2 and CL4 is in the timing of the end of drought. Groundwater drought conditions ended in CL1 and CL4 in May 1993, seven months after the end of the meteorological drought, but this was still six months before the groundwater drought ended in CL2 (Fig. 9a).

The 1995 to 1997 drought, although shorter than the 1988 to 1993 drought, followed a similar pattern with groundwater drought starting approximately contemporaneously with the meteorological drought. Although it was a continuous event for all three clusters (there were no breaks in the drought for CL1 and CL4), CL1 and CL4 again show approximately annual intensifications and declines in drought status during the episode. Such approximately annual changes in drought status are not seen in CL2. The 1995 to 1997 drought had the greatest magnitude in CL2 due to the prolonged end to the drought in this cluster, with groundwater drought in CL1 and CL4 finishing approximately contemporaneously with the meteorological drought but six months later in CL2. The 2011 to 2012 drought was much shorter than the other two multi-annual droughts, lasting just over a year starting relatively abruptly in early 2012 and finished abruptly in CL1 and CL4 in May 2012 in response to an unusual episode of spring recharge Parry et al. (2013). The groundwater drought in CL2 again finished relatively late, this time about three months later, in August 2012. The relatively short delay in the breaking of the groundwater drought in CL2 compared with CL1 and CL4 probably reflects the relatively smaller groundwater drought deficit accumulated due to the shorter duration and lower magnitude of the drought compared with the 1988 to 1993 and 1995 to 1998 drought episodes.

5. Discussion

The results of the regional analysis of droughts based on cluster analysis are consistent with current conceptualisations of the dynamics of drought in hydrological systems. Propagation of drought through catchments and in particular through the groundwater compartment is well documented (Peters et al., 2003; 2006; Tallaksen et al., 2006) and four components of drought propagation are recognised, i.e. pooling, attenuation, lag and lengthening, three of which (attenuation, lag and lengthening) are associated with modifications of drought signals in groundwater (Van Loon, 2015). Attenuation results in smoothing of the maximum drought anomaly, lag describes the delay in the onset of the drought signal as it passes through the hydrological cycle (for example, see Fig. 3a and Fig. 4 of Van Loon, 2015,), and lengthening extends the period of drought. Considering Table 3 that summarises the three multi-annual droughts and comparing event magnitude for SPI_{12}, CL1 CL2 and CL4 respectively, there is, as would be expected, evidence of a general attenuation of the SPI drought signal in the three clusters compared with SPI_{12}. Lagging of the multi-annual groundwater droughts behind meteorological droughts is not so easy to unambiguously quantify. Clearly the nature and degree of the lag is sensitive to the rainfall accumulation period used to define the meteorological drought index most closely correlated with SGI. In the present case, accumulation periods of 4, 16, and 9 months are required for CL1, 2 and 4 respectively to achieve optimal correlation between the SPI and SGI time series. Finally, the results of the present study strongly support the concept of lengthening of groundwater drought relative to meteorological drought (Van Loon, 2015). The results demonstrate that lengthening is most pronounced following longer and deeper groundwater droughts. They serve to emphasise that there can be significant differences in the lengthening response between different clusters, even within with the same aquifer. It also appears that the degree of lengthening may also be related to SGI autocorrelation (the greatest degree of lengthening is observed in cluster CL2 associated with the largest SGI autocorrelation, m_{max}).

The results of the regional analysis add to our current understanding of the controls on groundwater droughts. Bloomfield and Marchant (2013) investigated how unsaturated zone thickness and the hydraulic diffusivity of aquifers may relate to m_{max}. Using 14 SGI time series from four different aquifers around the UK (including one site from the Lincolnshire Limestone and nine sites on the Chalk, although none from the present study) they found that m_{max} was broadly an inverse function of log hydraulic diffusivity, $\log D_{diff}$ (where D_{diff} is given by T/S and where T is aquifer transmissivity and S is specific storage of the aquifer). Although they also noted that when fractured aquifers, such as the Lincolnshire Limestone and the Chalk that have similarly high hydraulic diffusivities, were specifically considered there is no clear relationship between m_{max} and $\log D_{diff}$. However, they did find a positive relationship between unsaturated zone thickness and m_{max} for fractured aquifers such as the Chalk and Lincolnshire Limestone. Based on this observation, they proposed that unsaturated zone drainage and recharge processes were an important contributory factor in determining autocorrelation or ‘memory’ in groundwater level hydrographs and by inference an influential factor on groundwater drought characteristics, particularly in fracture aquifer systems. Here we investigate if a similar relationship between m_{max} and unsaturated zone thickness holds for CL1, CL2 and CL4, clusters dominated by fractured aquifers.

Figure 12 shows box plots of unsaturated zone thickness for CL1, CL2 and CL4 as a function of m_{max} for each cluster (where unsaturated zone thickness is taken as the mean depth to groundwater recorded for sites in each cluster over the study period). In addition, corresponding observations for ten boreholes in fractured aquifers from Bloomfield and Marchant (2013) are also shown for reference. The results of the present study are consistent with those of Bloomfield and Marchant (2013, Fig. 13a) and show: increasing mean unsaturated zone thickness with increasing cluster m_{max}; increasing variability in unsaturated zone thickness with increasing cluster m_{max}; and increasing maximum unsaturated zone thickness with increasing cluster m_{max}. Bloomfield and Marchant (2013) previously noted that such observations are consistent with the findings of Peters et al. (2005), since unsaturated zone thickness is a function of distance to streams. However, in the present study area (Fig. 1) surface drainage is virtually absent from the northern Lincolnshire Limestone that dominates CL1 and is limited over both the Chalk (CL2) and the southern Lincolnshire Limestone (CL4). Instead we postulate that unsaturated zone thickness, and hence $m_{max,}$ is affected by more general catchment characteristics such as extent of outcrop, topography, intrinsic aquifer characteristics and aquifer thickness that all influence, through unsaturated zone drainage and saturated flow processes, the overall shape of the piezometric surface in the aquifers. For example, of the three aquifers in the study region the Chalk has the most extensive outcrop; it is the thickest aquifer, up to five times thicker than the Lincolnshire Limestone; and forms hills up to ~150 m asl compared to hills about 70 m asl across the southern Lincolnshire Limestone, while it is associated (CL2) with the largest m_{max} and the longest and highest magnitude droughts. As such, the relationships between unsaturated zone thickness, SGI autocorrelation and hence groundwater drought characteristics are not trivial and appear to reflect a number of fundamental catchment properties and processes that effect groundwater level dynamics and hence groundwater drought phenomena.

Although clustering of groundwater hydrographs is not novel in itself (Winter, 2000; Moon et al, 2004; Upton and Jackson, 2011) this is the first time these techniques have been systematically applied to investigate groundwater droughts. The approach described is generic and widely applicable and here we briefly highlight some of the methodological considerations, and implications for monitoring and prediction of groundwater droughts. The *k*-means clustering has been performed on the complete SGI hydrographs, including periods of relatively high groundwater level stand, even though the aim of the hydrograph classification has been to investigate regional variations in groundwater droughts. Yet the resulting clusters have been shown to effectively identify distinct regional groundwater drought responses across the study area. For example, they reflect the major drought history across the study region (Fig. 2 and Fig. 7), and identify spatially coherent hydrographs that are consistent with know hydrogeological differences across the study area (Fig. 3c and Fig. 9a). Eltahir and Yeh (1999) investigated the asymmetry of groundwater hydrographs to high and low groundwater level stands and noted that 'droughts leave a significantly more persistent signature on groundwater hydrology than floods'. They inferred that this phenomenon was because discharge of groundwater to streams is an efficient dissipation mechanism for wet anomalies and that this discharge is often strongly nonlinear. This may explain, at least in part, why the hydrograph classification scheme based on full hydrographs provides such a good basis for analysis of the heterogeneous response of groundwater to drought at the regional scale. However, there is potential for future work to investigate if the hydrograph classification can be improved by focussing on, or giving more weight to episodes of drought in the SGI time series.

In addition to identifying three clusters of SGI hydrographs, CL1, CL2 and CL4, that exhibit different characteristic responses to meteorological drivers, the *k*-means clustering also identified three relatively small clusters of SGI hydrographs, CL3, CL5 and CL6, where there were either: trends in the SGI time series; temporal anomalies expressed as anomalous phase relationships between cluster SGI and the regional SGI time series; or relatively poor coherency in SGI time series with a given cluster. In these three clusters it has been inferred that hydrographs may have been variably impacted by anthropogenic factors, such as groundwater abstraction. Although the CA was not specifically designed to identify anthropogenically impacted groundwater hydrographs the classification scheme could be used to that end since it can differentiate between clusters showing trends superimposed on the regional signals (e.g. CL3 and CL5) and clusters with anomalous phase relationships with the regional signal (e.g. CL6). The presence of a trend in a cluster of hydrographs may be indicative of an anthropogenic impact, for example from unsustainable abstraction (declining trend) or from groundwater rebound (rising trend). Where there is limited prior information regarding groundwater withdrawals across a region, a not uncommon situation in areas where abstraction is not highly regulated, cluster analysis could be used, either as it has been in the present study based on a set of heuristic rules to identify a suitable number of clusters, or in an exploratory manner. If it is used in a more exploratory manner, either hierarchical or non-hierarchical clustering could be undertaken and then clusters searched to identify spatially coherent clusters that show significant downward trends in hydrographs (where significance of trends in a cluster could be tested and quantified using standard tests, such as Mann-Kendall and Sen’s slope estimates). Any spatial coherence in clusters exhibiting downward trends may be taken as indicating the presence of potentially unsustainable abstraction. For the purposes of a study where the stationarity of the data is important, if trends in individual hydrographs are already known then either these hydrographs can be removed from an analysis or the trends could be identified and removed prior standardisation and clustering of the hydrographs.

It has been shown that there can be pronounced differences in the characteristics of multi-annual drought episodes between aquifers within a region (Fig. 9a). During multi-annual droughts some clusters temporarily go out of drought conditions while others will continually show deepening drought conditions over two or more years, and some clusters stay in groundwater drought for many months after groundwater (and meteorological) drought has ceased in other clusters. If observations such as these or similar can be made for a region they may have important implications for monitoring groundwater droughts and water resource management in multi-aquifer (cluster) systems. For example, at the end of a drought, sites in more quickly responding clusters may act as leading indicators of the end of groundwater drought at sites in more slowly responding clusters. In addition to the implications for groundwater monitoring particularly during long droughts, if there is sufficient understanding of regional variations in groundwater responses (i.e. relative differences in the timing and intensity of groundwater drought between different aquifers in a region or between sub-regions within an aquifer), then this understanding could be used to inform appropriate groundwater water resource management strategies and so may enable some of the worst impacts of the groundwater drought to be mitigated.

More generally we see a range of possible benefits to clustering groundwater hydrographs. For example, ‘sentinel’ boreholes within each cluster, those that are closest to the mean behaviour of a group, could be identified and used as indicative of the groundwater response of a wider area. Missing data is a common issue with groundwater hydrographs, and clustering techniques could potentially be used to identify suitable boreholes from which groundwater levels could be infilled. However, more importantly, clustering could be used in combination with groundwater models to aid the prediction of groundwater droughts. A range of techniques can be used to model groundwater hydrographs at a site, i.e. non-distributed groundwater models, including statistical models (Ahn 200; Bloomfield et al. 2003), artificial neural network models (Sreekanth et al. 2009) and ‘black box’ models (Mackay et al, 2014). The hydrograph cluster analysis could be used in combination with any of these techniques for groundwater drought forecasting. For example, forecasts of groundwater levels 1 to 3 months out are currently undertaken in the UK for selected sites using a black-box, lumped parameter model (Jackson et al. 2013; Mackay et al. 2014; Hydrological Outlooks, 2015) driven by probabilistic estimates of future rainfall. Regional inferences of future groundwater levels are then based on qualitative interpretations of the individual sites. Applying similar modelling systems to mean cluster hydrographs that are representative of spatially coherent regions of groundwater drought response instead of individual site specific hydrographs could enable more rigorous forecasts of the spatial distribution of groundwater drought.

6. Conclusions

Cluster analysis (CA) when applied to SGI time series of consistent length for multiple sites across a region has been shown to provide a robust approach to the regional analysis of groundwater droughts. In the present study an agglomerative hierarchical complete-linkage strategy and a *k*-means clustering strategy were tested. The *k*-means clustering was found to be most suitable. However, for any given case study a range of non-hierarchical algorithms and hierarchical classification schemes should be explored to see which is most appropriate.

A heuristic, rule-based approach was found useful in guiding the selection of the optimal number of clusters, where the rules applied prior knowledge of the hydrogeology of the study area including information related to spatial variations in catchment and aquifer characteristics. For the present case study, both non-hierarchical algorithms and hierarchical classification schemes provide better clustering of SGI time series than a simple three-fold classification simply based on geology alone, with the k-means clustering providing the best clustering. Membership of the resulting *k*-means clusters is shown to be dominated by hydrogeological factors and the effect of heterogeneity in precipitation over the study area on cluster composition is inferred to be negligible.

The clusters successfully discriminate different responses to groundwater drought both in terms of drought metrics for the complete time series and with respect to the detailed response of sites in each cluster during specific major episodes of multi-annual drought. Groundwater drought characteristics can be linked, through the autocorrelation structure of cluster hydrographs, to the distribution of unsaturated zone thickness. This reflects the role of a range of catchment and aquifer properties and processes that influence groundwater level dynamics, including topography, aquifer thickness and extent of outcrop, unsaturated zone drainage characteristics and saturated groundwater flow.

This approach to groundwater hydrograph clustering is flexible, can be applied in a wide range of hydrogeological settings where suitable hydrographs are available, and enables spatially variable responses of groundwater to drought to be quantified.

Acknowledgements

We would like to thank Henry Holbrook with help in preparation of the figures. The work described has been funded by the British Geological Survey (Natural Environment Research Council), and this paper is published with the permission of the Executive Director of the British Geological Survey (Natural Environment Research Council).

References

Ahn, H.: Modelling groundwater heads based onsecond-order difference time series models. Journal of Hydrology, 234, 82-94, 2000.

Allen, D. J., Brewerton, L. J., Coleby, L. M., Gibbs, B. R., Lewis, M. A., MacDonald, A. M., Wagstaff, S. J., and Williams, A. T.: The physical properties of major aquifers in England and Wales, British Geological Survey Research Report WD/97/34, Keyworth, UK, 1997.

Bloomfield, J. P.: Characterisation of hydrogeologically significant fracture distributions in the Chalk: An example from the Upper Chalk of southern England, Journal of Hydrology, 184, 335-379, 1996.

Bloomfield, J. P., and Marchant, B. P.: Analysis of groundwater drought building on the Standardised Precipitation Index approach, Hydrology and Earth System Sciences, 17, 4769-4787, 2013.

Bloomfield, J. P., Brewerton, L.J, and Allen, D.J.: Regional trends in matrix porosity and dry density of the chalk of England, Quarterly Journal of Engineering Geology, 28, S131-S142, 1995.

Bloomfield, J.P., Gaus, I., and Wade, S.D.: A method for investigating potential impacts of climate change scenarios on annual minimum groundwater levels. Water and Environment Journal, 17, 86-91, 2003.

Bonaccorso, B., Bordi, A., Cancelliere, A., Rossi,G., and Sutera, A.: Spatial variability of drought: An analysis of the SPI in Sicily, Water Resource Management, 17, 273-296, 2003.

Chang, T. J. and Teoh, C. B.: Use of the kriging method for studying characteristics of ground water droughts, Journal of the American Water Resources Association, 31, 1001-1007, 1995.

Dore, A. J., Kryza, M., Hall, J. R., Hallsworth, S., Keller, V. J. D., Vieno, M., and Sutton, M. A.: The influence of model grid resolution on estimation of national scale nitrogen deposition and exceedance of critical loads, Biogeosciences, 9, 1597-1609, 2012

Eder, B. K., Davis, J. M., and Monahan, J. F.: Spatial and temporal analysis of the Palmer drought severity index over south-eastern United States, Journal of Climatology, 7, 31-51, 1987.

Eltahir, E. A. B., and Yeh, P. J-F.: On the asymmetric response of aquifer water level to floods and droughts in Illinois, Water Resources Research, 35, 1199-1217, 1999.

Environment Agency. 2014. National groundwater level database for England, Environment Agency. http://data.gov.uk/data last retrieved 26 August 2014.

Fleig, A. K., Tallaksen, L. M., Hisdal, H., and Hannah, D. M.: Regional hydrological drought in north-western Europe: Linking a new regional drought area index with weather types, Hydrological Processes, 25, 1163-1179, 2011.

Folland C.K., Hannaford, J., Bloomfield J.P., Kendon, M., Svensonn, C., Marchant, B.P., Prior, J., and Wallace, E. Multi-annual droughts in the English Lowlands: a review of their characteristics and climate drivers in the winter half year, Hydrology and Earth Systems Sciences, 19, 2353-2375, 2015.

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

Hannah, D. M., Smith, B. P. G., Gurnell, A. M., and McGregor, G. R.: An approach to hydrograph classification, Hydrological Processes, 14, 317-338, 2000.

Hisdal, H., and Tallaken, L. M.: Estimation of regional meteorological and hydrological drought characteristics, Journal of Hydrology, 281, 230-247, 2003.

Hydrological Outlooks. 2015. http://www.hydoutuk.net last downloaded 25th August 2015

Jackson, C. R., Pachocka, M., and Mackay, J. D: Hydrological outlook: outlook based on modelled groundwater levels and associated climate forecasts. British Geological Survey Report OR/13/046, Nottingham, UK. 2013. http://nora.nerc.ac.uk/503966/ last downloaded 25 August 2015

Jones, H. K., Morris, B. L., Cheney, C. S., Brewerton, L. J., Merrin, P. D., Lewis, M. A., MacDonald, A. M., Coleby, L. M., Talbot, J. C., McKenzie, A. A., Bird, M. J., Cunningham, J. and Robinson, V. K.: The physical properties of minor aquifers in England and Wales, British Geological Survey Technical Report, WD/00/4, Keyworth, UK, 2000.

Karl, R. T., and Koscienly, A. J.: Drought in the United States:1895-1981, Journal of Climatology, 2, 313-329, 1982.

Keller, V., Young, A. R., Morris, D., and Davies, H.: Continuous Estimation of River Flows (CERF), Technical Report: Estimation of Precipitation Inputs, Environment Agency R&D Project Report WD-101, Centre for Ecology and Hydrology, Wallingford, 2005.

Kendon, M., Marsh, T. J., and Parry, S.: The 2010-2012 drought in England and Wales, Weather, 68, 88-95, 2013

Kingston, D. C., Flieg, A. K., Tallaksen L. M., and Hannah, D. M.: Ocean-atmosphere forcing of summer streamflow droughts in Great Britain, American Meteorological Society, 14, 331-344, 2013

Klugman, M. R.: Drought in the Upper Midwest, 1931-1969, Journal of Applied Meteorology, 17, 1425-1431, 1978.

Lana, X., Serra, C., and Burgueno, A.: Patterns of monthly rainfall shortage and excess in terms of the standardised precipitation index for Catalonia (NE Spain). International Journal of Climatology, 21, 1669-1691, 2001.

Lloyd-Hughes, B., and Saunders, M. A.: A drought climatology for Europe, International Journal of Climatology, 22, 1571-1592, 2002.

Lloyd-Hughes, B., Prudhomme, C., Hannaford, J., Parry, S., Keef, C., and Rees, H. G.: Drought catalogues for UK and Europe. Environment Agency Science Report, SC070079/SR, Environment Agency, Bristol, 2010.

Lorenzo-Lacruz, J., Moran-Tejeda, E., Vincente-Serrano, S. M., and Lopez-Moreno, J. I.: Streamflow droughts in the Iberian Peninsula between 1945 and 2005: spatial and temporal patterns, Hydrology and Earth System Sciences, 17, 119-134, 2013.

Mackay, J.D., Jackson, C.R., and Wang, L. A lumped parameter conceptual model to similate groundwater level time series. Environmental Modelling software, 61, 229-245. 2014

Marsh, T. J., and Hannaford, J.: UK Hydrometric Register, Hydrological data UK series, Centre for Ecology and Hydrology, Wallingford, 2008.

Marsh, T. J., Cole, G., and Wilby, R.: Major droughts in England and Wales. 1800-2006, Weather, 62, 87-93, 2007.

Marsh, T. J., Simon, P., Kendon, M., and Hannaford, J.: The 2010-12 drought and subsequent extensive flooding: a remarkable hydrological transformation, Centre for Ecology & Hydrology, Wallingford, 54pp, 2013.

Maurice, L. D., Atkinson, T. C., Barker, J. A., Bloomfield, J. P., Farrant, A. R., and Williams, A. T.: Karstic behaviour of groundwater in the English Chalk, Journal of Hydrology, 330, 63-70, 2006.

McKee, T. B., Doesken, N. J., and Leist, J.: The relationship of drought frequency and duration time scales, 8th Conference on Applied Climatology, Anaheim, California, 17 to 22 January 1993, 179-184, 1993.

Mendicino, G., Senatore, A., and Versace, P.: A Groundwater Resource Index (GRI) from drought monitoring and forecasting in a Mediterranean climate, Journal of Hydrology, 357, 282-302, 2008.

Mishra, A. K., and Singh, V. P.: A review of drought concepts, Journal of Hydrology, 391, 202-216, 2010.

Moon, S. K., Woo, N. C., and Lee, K. S.: Statistical analysis of hydrographs and water-table fluctuation to estimate groundwater recharge, Journal of Hydrology, 292, 1989-209, 2004.

Parry, P., Marsh, T., Kendon, M.: 2012: from drought to floods in England and Wales, Weather, 68, 268-274. 10.1002/wea.2152, 2013

Peters, E., Torfs, P. J. J. F., Van Lanen, H. A. J., and Bier, G.: Propagation of drought through groundwater – a new approach using linear reservoir theory, Hydrological Processes, 17, 3023-3040, 2003.

Peters, E., Van Lanen, H. A. J., Torfs, P. J. J. F., and Bier, G.: Drought in groundwater – drought distribution and performance indicators, Journal of Hydrology, 306, 302-317, 2005.

Peters, E., Bier, G., Van Lanen, H. A. J., and Torfs, P. J. J. F.: Propagation and spatial distribution of drought in a groundwater catchment, Journal of Hydrology, 321, 257-275, 2006.

Raziel, T., Bordi, I., and Pereira, L. S.: A precipitation-based regionalisation for Western Iran and regional drought variability, Hydrology and Earth System Sciences, 12, 1309-1321, 2008.

Sadri, S and Burn, D. D.: Copula-based pooled frequency analysis of droughts in the Canadian Praries, Journal of Hydrological Engineering, 19, 277-289, 2014.

Santos, J. F., Pulido-Calvo, I., and Portela, M. M.: Spatial and temporal variability of droughts in Portugal, Water Resources Research, 46, W03503, 2010.

Stahl, K., and Demuth, S.: Linking streamflow drought to the occurrence of atmospheric circulation patterns, Hydrological Sciences Journal, 44, 467-482, 1999.

Stahl, K., and Demuth, S.: Regional classification. In: Demuth, S. Stahl, K. (Eds.) ARIDE – Assessment of the Regional Impact of Drought in Europe, Institute of Hydrology, University of Freiburg, Frieburg, Germany, pp. 98-105, 2001.

Sreekanth, P.D., Geethanjali, N, Sreedevi, P.D., Ahmed, S., Ravi Kumar, N., and Kamala Jayanthi, P.D.: Forecasting groundwater levels using neural networks, Current Science., 96, 933-939, 2009.

Tallaksen, L. M., and Van Lanen, H. A. J.: Hydrological drought. Processes and estimation methods for streamflow and groundwater, Developments in Water Sciences 48, Elsevier, Netherlands, 2004.

Tallaksen, L. M., Hisdal, H. And Van Lanen, H. A. J.: Propagation of drought in a groundwater fed catchment, the Pang in UK. In Demuth, S. (Ed.), FRIEND 2006 – Water Resources Variability: Processes, Analyses and Impacts, IAHS Publication 308, 128-133, 2006.

Tallaksen, L. M., Hisdal, H., and Van Lanen, H. A. J.: Space-time modelling of catchment scale drought characteristics, Journal of Hydrology, 375, 363-372, 2009.

Upton, K. A. and Jackson, C. R.: Simulation of the spatio-temporal extent of groundwater flooding using statistical methods of hydrograph classification and lumped parameter models, Hydrological Processes, 25, 1949-1963, 2011.

Van Lanen, H.A.J.: On the definition of groundwater drought. Geophysical Research Abstract, 7, 10867. 2005.

Van Lanen, H.A.J., and Peters, E.: Definition, effects and assessment of groundwater droughts. In: Vogt, J.V. and Somma, F. (Eds.), Drought and Drought Mitigation in Europe, Kulwer, Dordrecht, 49-61, 2000.

Van Lanen, H.A.J. and Tallaksen, L.M.. 2007. Hydrological drought, climate variability and change, in: Climate and Water, edited by: Heinonen,M., Proceedings of the third International Conference on Climate and Water, Helsinki, 3–6 September 2007, 488–493, 2007.

Van Lanen, H. A. J., Wanders, N., Tallaksen, L. M., and Van Loon, A. F.: Hydrological drought across the world: impact of climate and physical catchment structure, Hydrol. Earth Syst. Sci. 17: 1715–1732, doi:10.5194/hess-17-1715-2013, 2013.

Van Loon, A. F.: Hydrological drought explained, WIREs Water, doi:10.1002/wat2.1085, 2015.

Van Loon, A. F. and Laaha G.: Hydrological drought severity explained by climate and catchment characteristics, Journal of Hydrology, 526, 3-14, 2015.

Van Loon, A. F. and Van Lanen, H. A. J.: A process-based typology of hydrological drought, Hydrol. Earth Syst. Sci. 16: 1915-1946, doi:10.5194/hess-16-1915-2012, 2012.

Vicente-Serrano, S.M.: Spatial and temporal analysis of droughts in the Iberian Peninsula [1910-2000]. Hydrological Sciences Journal, 51, 83-97, 2006.

Vicente-Serrano, S.M., and Cuadrat-Prats, J.M.: Trends in drought intensity and variability in the middle Ebro valley (NE of the Iberian peninsula) during the second half of the twentieth century. Theoretical Applied Climatology, 88, 247-258, 2007.

Webster, R. and Oliver, M.A.: *Statistical Methods in Soil and Land Resource Survey*. Oxford University Press, U.K. 1990

Whitehead, E.J. and Lawrence, A.R.: The Chalk aquifer system of Lincolnshire. British Geological Survey Research Report, RR/06/03, 2006. http://nora.nerc.ac.uk/3699/

World Meteorological Organization. Standardized Precipitation Index User Guide, WMO-No. 1090, Geneva http://www.wamis.org/agm/pubs/SPI/WMO_1090_EN.pdf Last downloaded, 13th August 2015

Table 1. Summary of features of the six *k*-means clusters.

Cluster	Number of sites				Statistic		
	Total	Lincolnshire Limestone	Spilsby Sandstone	Chalk	SPI/SGI maximum correlation	Representative accumulation period, q_{max} (Months)	Autocorrelation range, m_{max} (Months)
CL1	13	13	0	0	0.74	4	15
CL2	23	2	0	21	0.86	16	23
CL3	6	2	4	0	0.36	15	60
CL4	24	19	0	5	0.82	9	18
CL5	5	0	0	5	0.53	17	28
CL6	3	0	3	0	0.09	-	-
Total	74	36	7	31			

Table 2. Summary of drought event statistics for clusters C1, C2 and C4.

	CL1	CL2	CL4
Number of Drought events	39	15	18
Mean duration (months)	4.6	11.3	9.1
Maximum duration (months)	27	61	49
Mean event magnitude	-2.9	-7.9	-6.6
Mean event intensity	-0.43	-0.28	-0.4
Maximum event intensity	-1.1	-1.05	-1.13
No. events where $I < -1$	3	2	2

Table 3. Summary of the 1988-93, 1995-98 and 2011-12 drought events for clusters CL1, CL2 and CL4 (where D_{event}, M_{event} and I_{event} denote indices for drought event duration, magnitude and intensity respectively).

Drought episode	Drought index	Regional SPI_{12}	Mean SGI CL1	Mean SGI CL2	Mean SGI CL4
1988 to 1993	Start date	Dec-88	Oct-88	Nov-88	Oct-88
	End date	Oct-92	May-93	Nov-93	May-93
	D_{event}	47	56	61	56
	M_{event}	-56.8	-37	-63.6	-41.6
	I_{event}	-1.2	-0.7	-1.0	-0.7
1995 to 1998	Start date	May-95	May-95	Aug-95	Jul-95
	End date	Oct-97	Jul-97	Feb-98	Aug-97
	D_{event}	30	27	31	26
	M_{event}	-34.3	-18.7	-32.4	-29.3
	I_{event}	-1.1	-0.7	-1.0	-1.1
2010 to 2012	Start date	Jan-11	May-11	Jan-11	Jul-10
	End date	Apr-12	May-12	Aug-12	May-12
	D_{event}	16	13	20	23
	M_{event}	-16.1	-13.9	-11.7	-21
	I_{event}	-1.0	-1.1	-0.6	-0.9

Figure captions

Figure 1. Case study area (left) and simplified geology map (right) showing locations of the observation boreholes. Cross-section (bottom) illustrating the stratigraphic/depth relationships between the three major aquifers in the study region: the Lincolnshire Limestone, the Spilsby Sandstone and the Chalk.

Figure 2. a. SPI/SGI correlation as a heatmap, b. mean SPI_{12} time series and c. mean SGI time series for all 74 hydrographs.

Figure 3. a. cluster dendrogram for hierarchical classification (k=6) of SGI time series, b. map showing the distribution of sites by clusters based on hierarchical classification (k=6), and c. map showing the distribution of sites by clusters formed by k-means clustering (k = 6).

Figure 4. RMSSD as a function of the number of clusters for the hierarchical and non-hierarchical k-means clustering algorithms and for a three-fold classification based on geology alone.

Figure 5. Heatmaps of Pearson correlation between SGI and SPI for q = 1 to 36 months and for lags up to 5 months. Maximum correlation is denoted by the closed black circles.

Figure 6. Mean SPI times series for each of the k-means clusters based on the accumulation period q_{max} for each cluster. Where the black line is SPI based on gridded precipitation series for sites in a given cluster and the red line is SPI for the mean rainfall across the whole study area based on the different aggregation periods, q_{max}, for each cluster.

Figure 7. Mean SGI time series for each of the six k-means clusters.

Figure 8. Correlograms for each of the mean SGI time series (bold) and individual site time series (grey) for each of the six k-means clusters showing variation in the autocorrelation function (ACF) for lags up to 60 months.

Figure 9. Heatmaps showing a.) SGI varying with time for all 74 sites as function of the six k-means clusters (left), and b.) correlations between all pairs of sites sorted as a function of the six k-means clusters (right).

Figure 10. Drought magnitude versus drought duration for sites in clusters CL1, CL2 and CL4.

Figure 11. Empirical distributions of a. drought duration, b. drought magnitude, and c. drought intensity for clusters CL1, CL2 and CL4.

Figure 12. SGI autocorrelation (m_{max}) as a function of unsaturated zone thickness.

Figure 1.

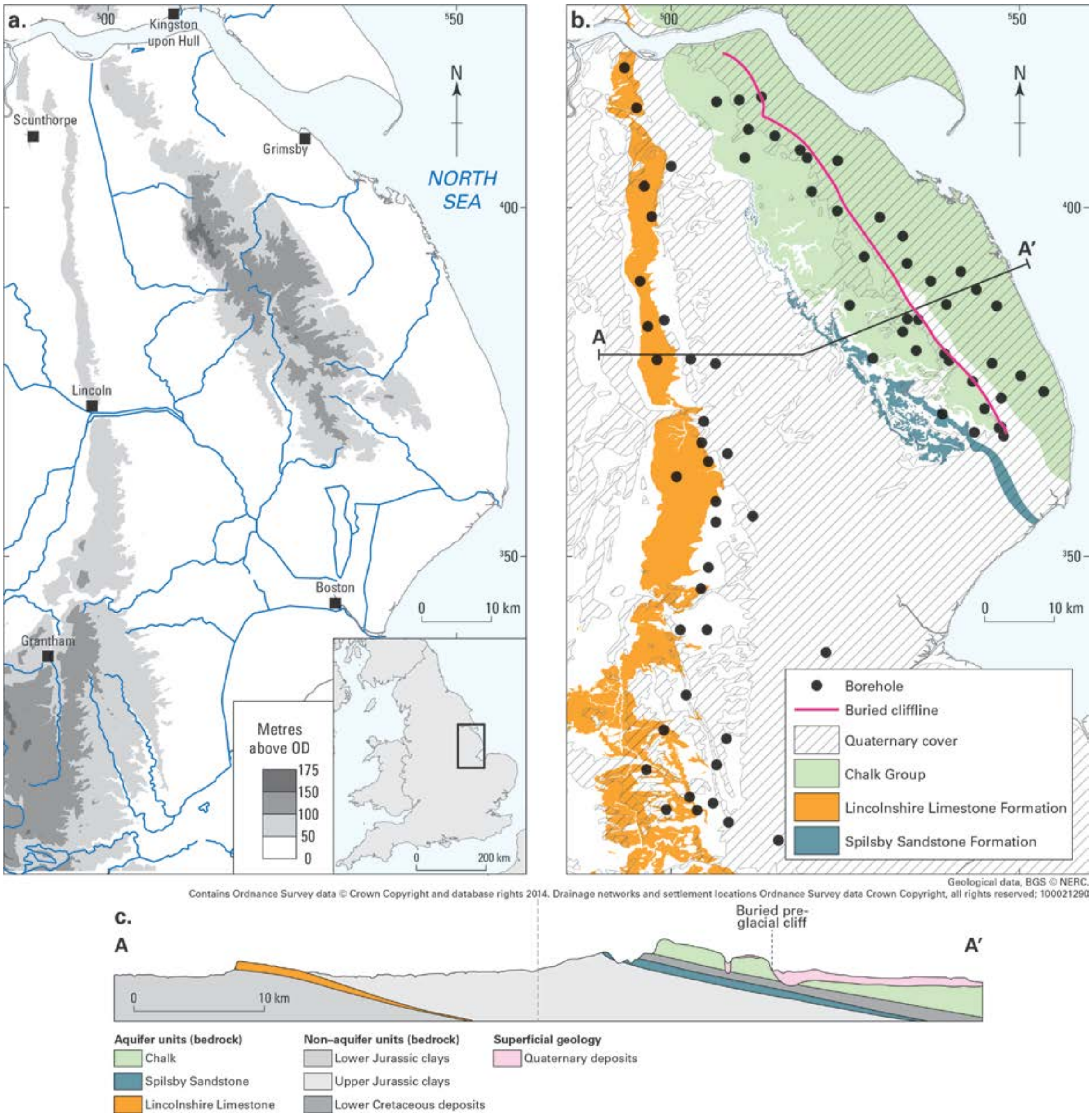

a.
Kingston upon Hull
Scunthorpe
Grimsby
NORTH SEA
Lincoln
Boston
Grantham
N
0 10 km
Metres above OD
175
150
100
50
0
0 200 km
b.
A
A′
N
0 10 km
Borehole
Buried cliffline
Quaternary cover
Chalk Group
Lincolnshire Limestone Formation
Spilsby Sandstone Formation
Geological data, BGS © NERC.
Contains Ordnance Survey data © Crown Copyright and database rights 2014. Drainage networks and settlement locations Ordnance Survey data Crown Copyright, all rights reserved; 100021290
c.
A
A′
Buried pre-glacial cliff
0 10 km
Aquifer units (bedrock)
Chalk
Spilsby Sandstone
Lincolnshire Limestone
Non-aquifer units (bedrock)
Lower Jurassic clays
Upper Jurassic clays
Lower Cretaceous deposits
Superficial geology
Quaternary deposits

Figure 2.

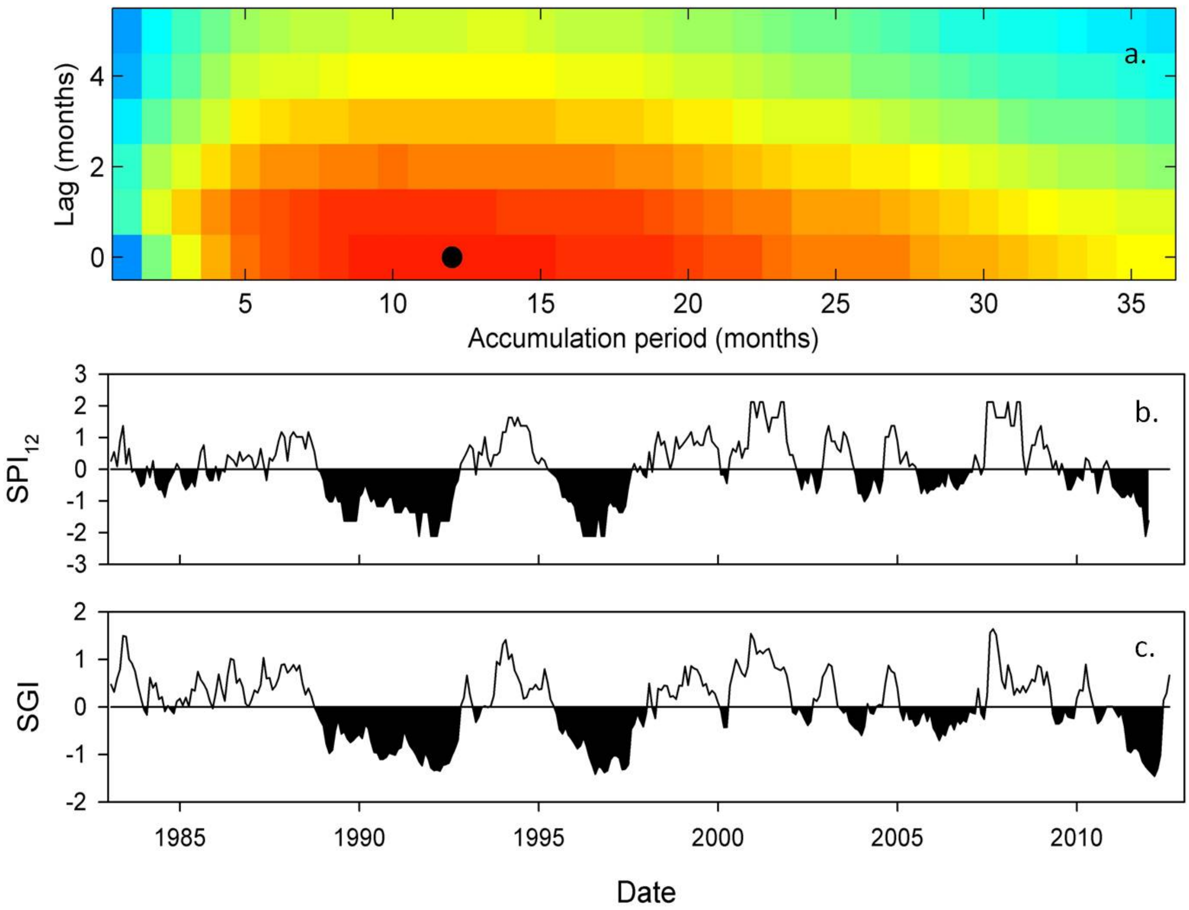
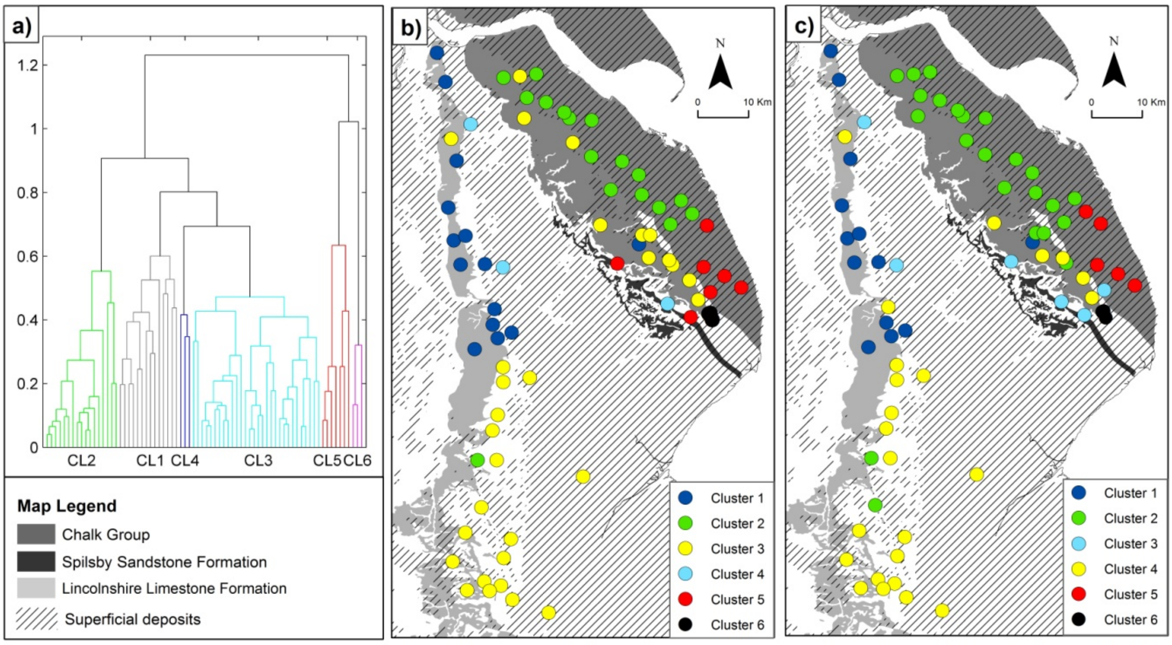

a.
Lag (months)
0
2
4
5
10
15
20
25
30
35
Accumulation period (months)
b.
SPI12
3
2
1
0
-1
-2
-3
c.
SGI
2
1
0
-1
-2
1985
1990
1995
2000
2005
2010
Date

Figure 3.



Geological data, BGS © NERC. Contains Ordance Survey data © Crown copyright and database right (2014)

Figure 4.

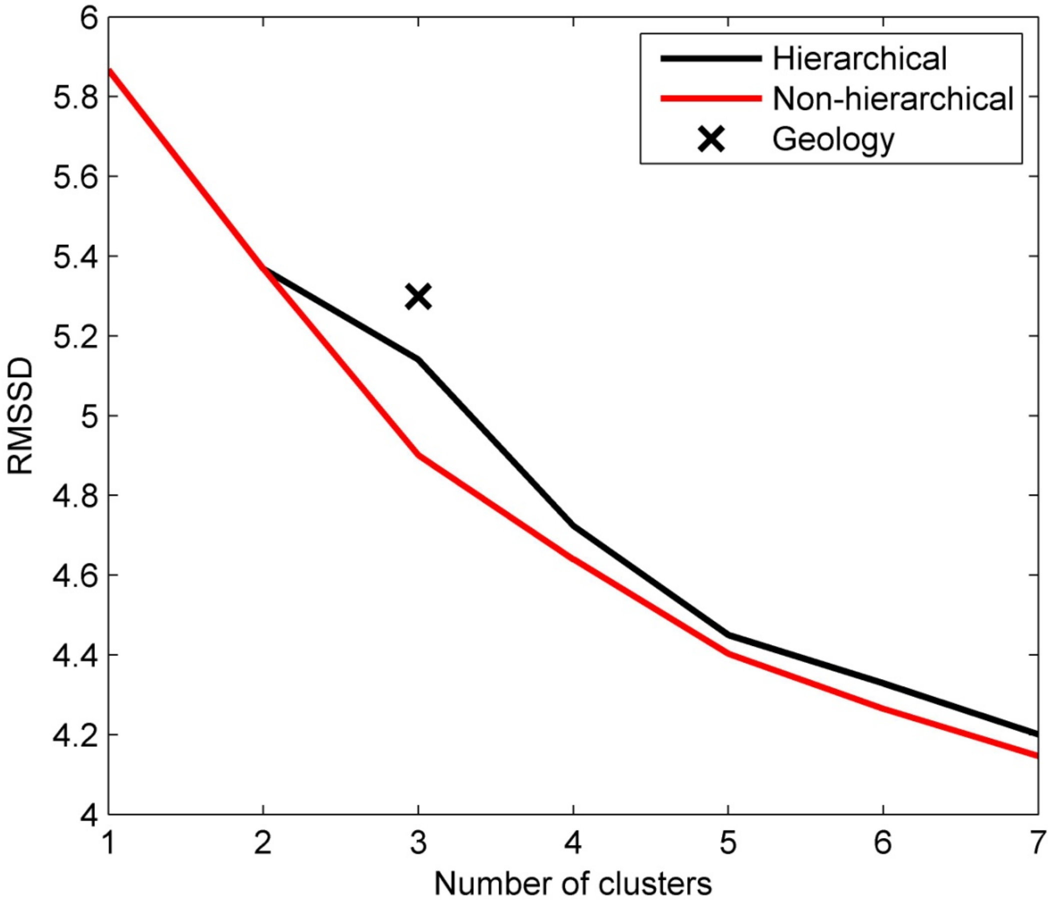

Hierarchical
Non-hierarchical
Geology
RMSSD
Number of clusters
6
5.8
5.6
5.4
5.2
5
4.8
4.6
4.4
4.2
4
1
2
3
4
5
6
7

Figure 5.

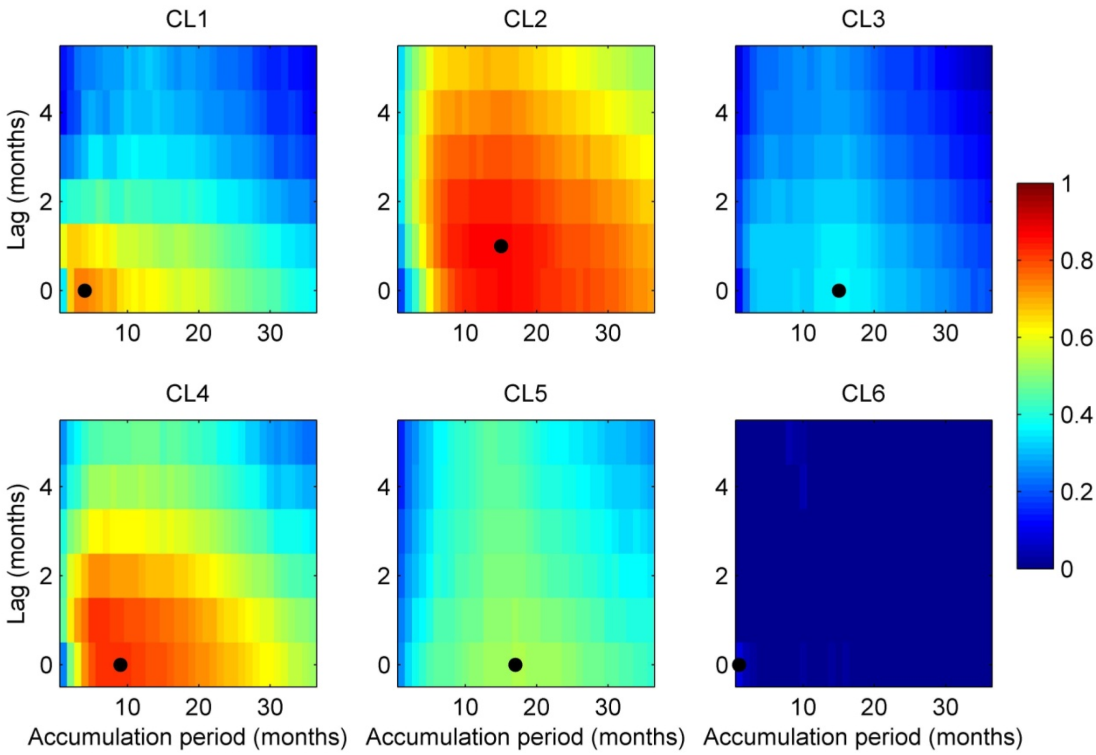

CL1
CL2
CL3
CL4
CL5
CL6
Lag (months)
Lag (months)
Accumulation period (months)
Accumulation period (months)
Accumulation period (months)
0
2
4
10
20
30
0
0.2
0.4
0.6
0.8
1

Figure 6.

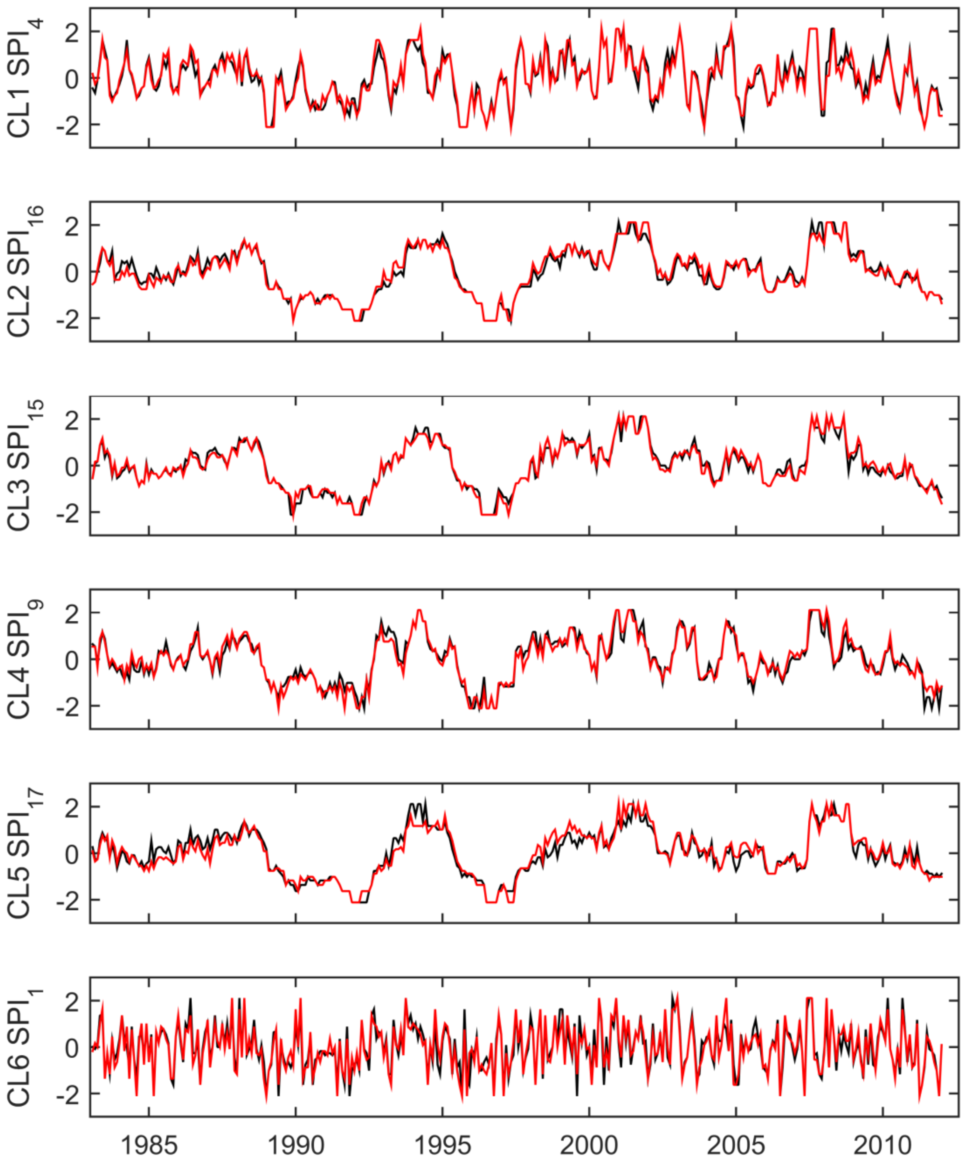

CL1 SPI4
CL2 SPI16
CL3 SPI15
CL4 SPI9
CL5 SPI17
CL6 SPI1
2
0
-2
1985
1990
1995
2000
2005
2010

Figure 7.

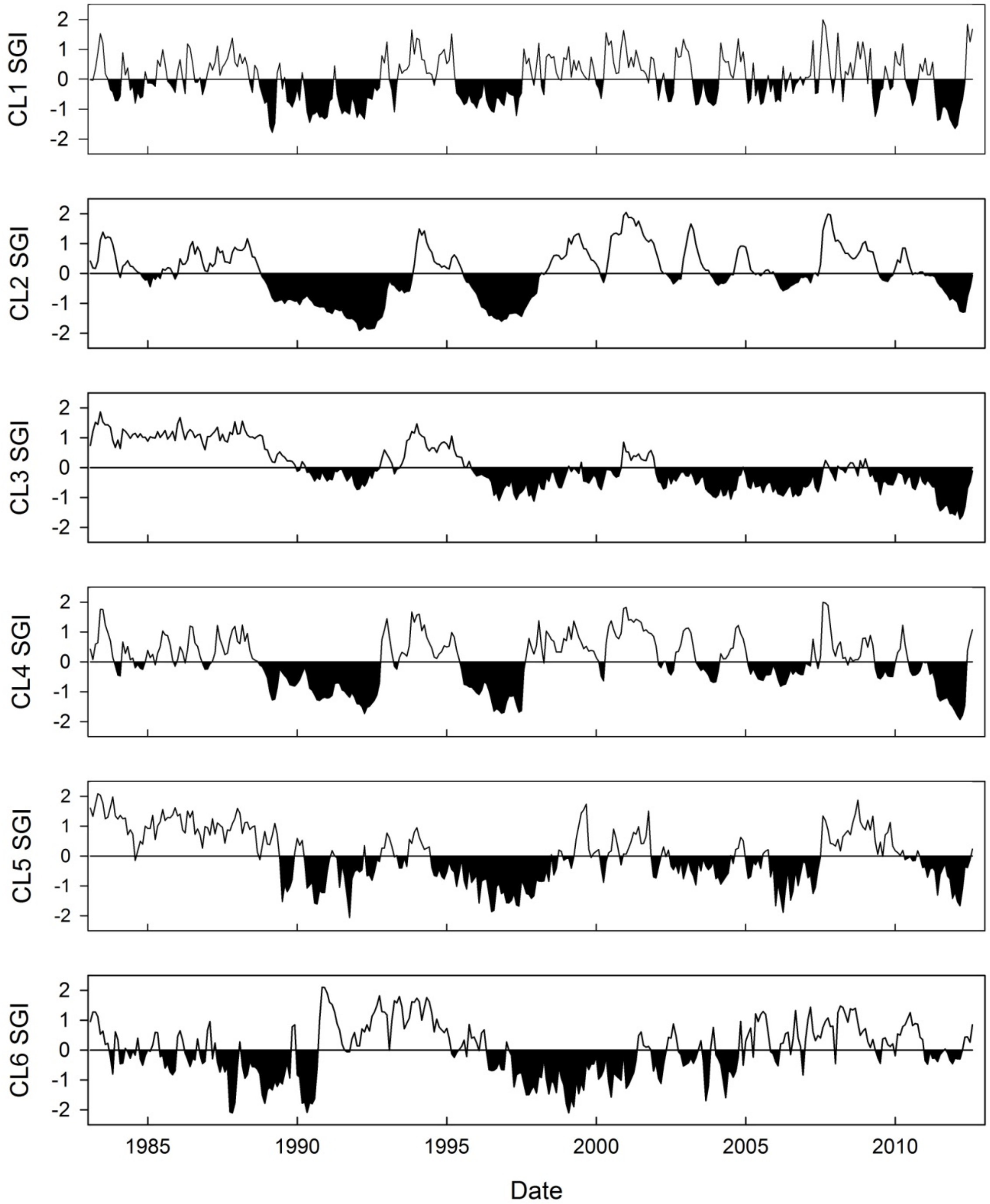

CL1 SGI
CL2 SGI
CL3 SGI
CL4 SGI
CL5 SGI
CL6 SGI
2
1
0
-1
-2
1985
1990
1995
2000
2005
2010
Date

Figure 8.

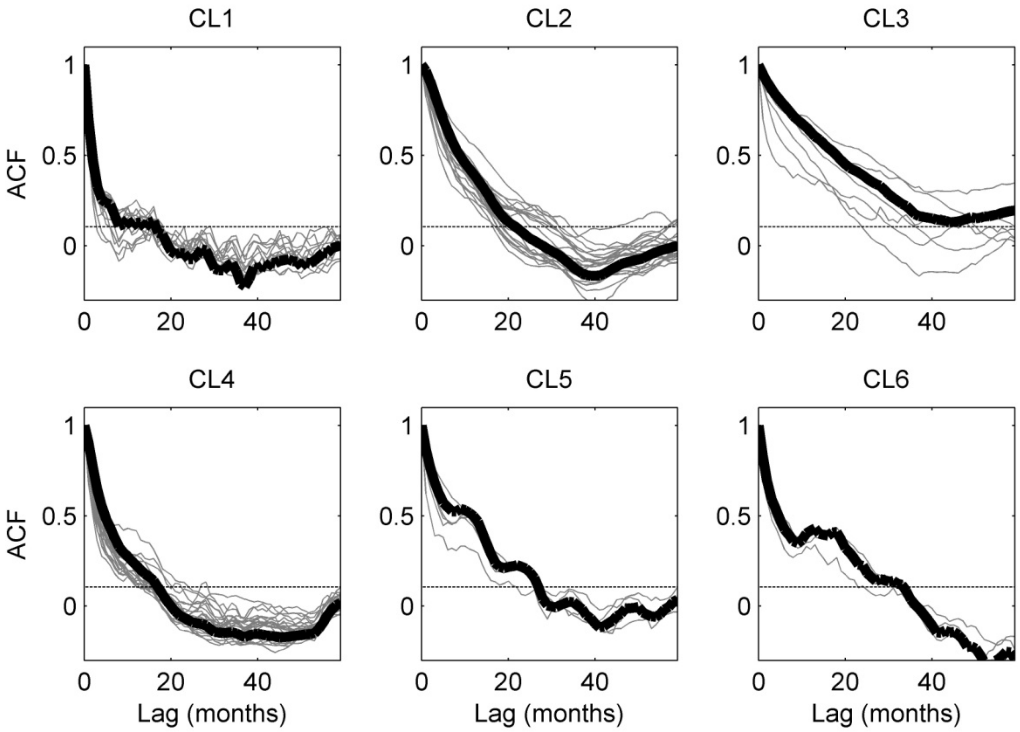
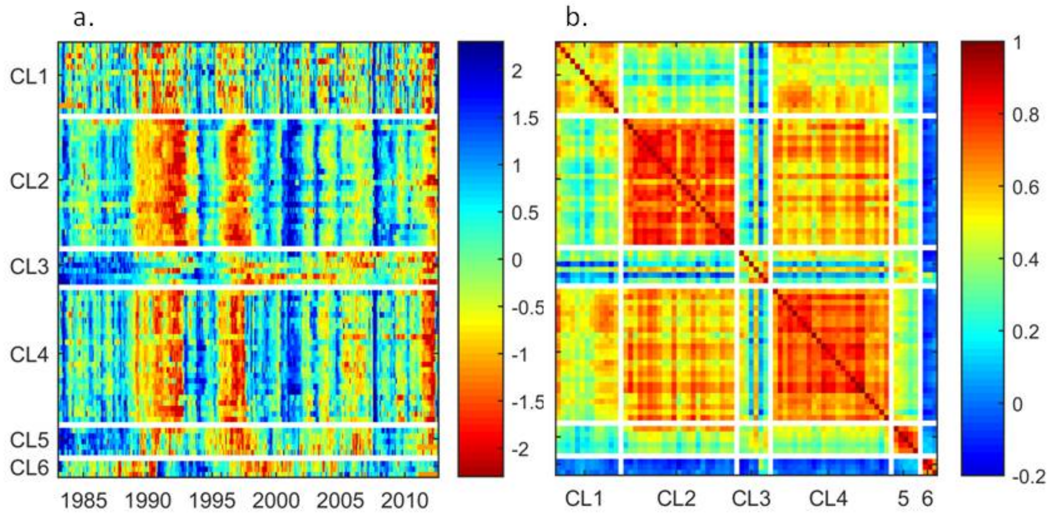
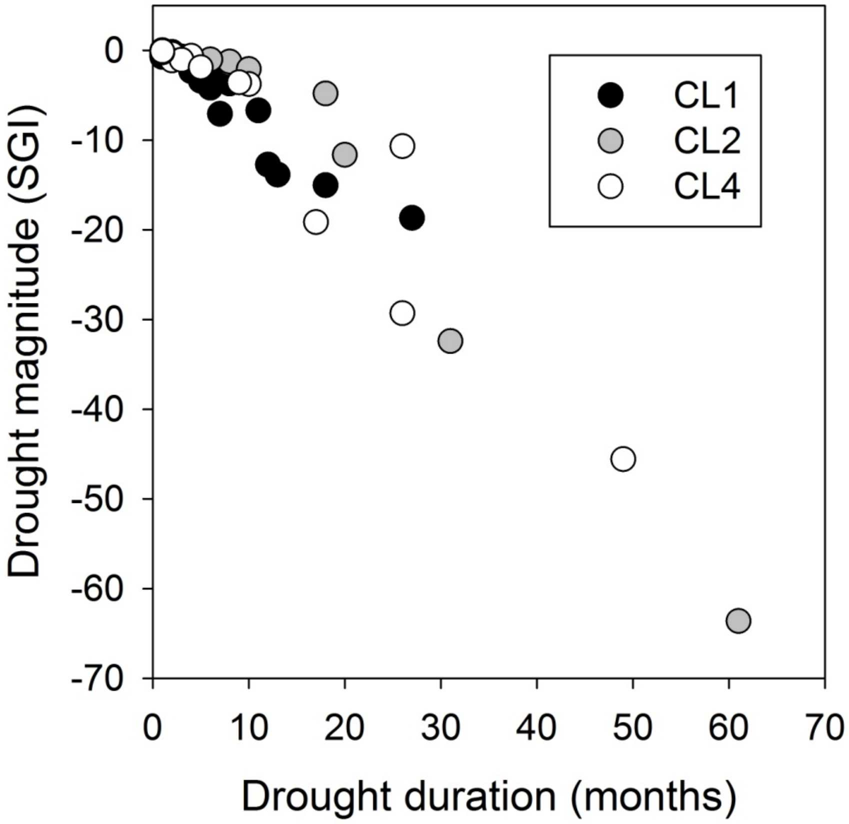

CL1
CL2
CL3
CL4
CL5
CL6
ACF
ACF
1
0.5
0
0
20
40
Lag (months)
Lag (months)
Lag (months)

Figure 9.



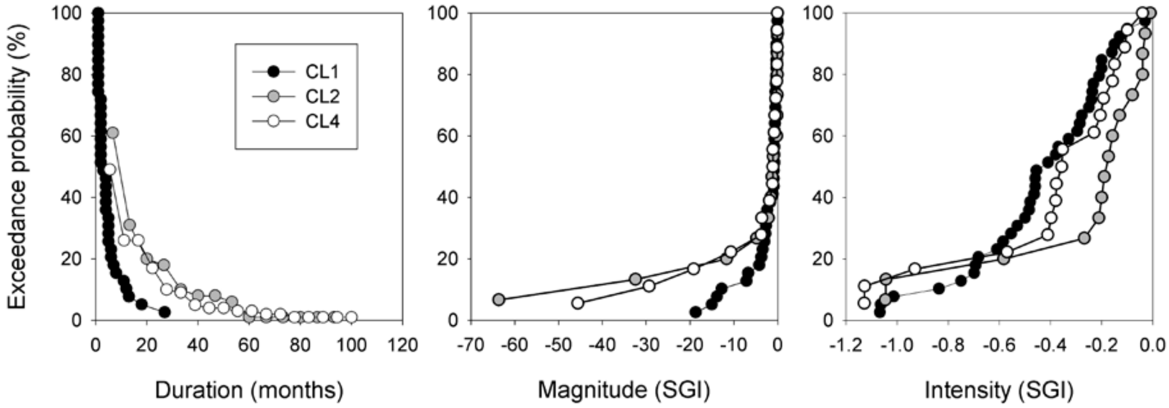
a.
b.
CL1
CL2
CL3
CL4
CL5
CL6
1985 1990 1995 2000 2005 2010
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
CL1 CL2 CL3 CL4 5 6
1
0.8
0.6
0.4
0.2
0
-0.2

Figure 10.



CL1
CL2
CL4
Drought magnitude (SGI)
0
-10
-20
-30
-40
-50
-60
-70
0
10
20
30
40
50
60
70
Drought duration (months)

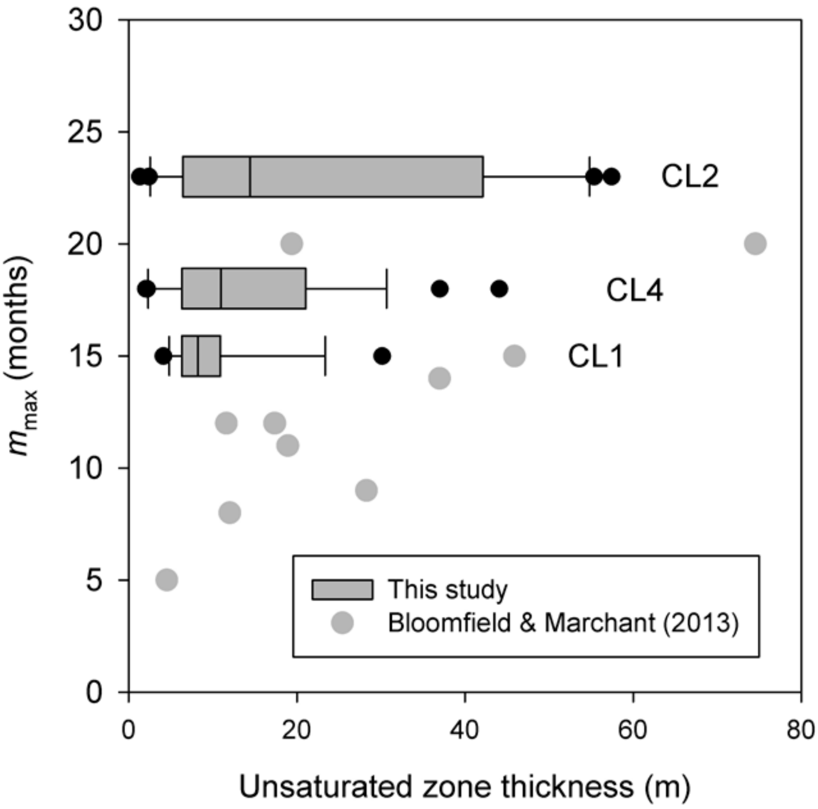
1039 Figure 11.



1040

1041

Figure 12.



CL2
CL4
CL1
This study
Bloomfield & Marchant (2013)
m_{max} (months)
Unsaturated zone thickness (m)
0
5
10
15
20
25
30
0
20
40
60
80