https://doi.org/10.5194/hess-2023-202
Preprint. Discussion started: 29 August 2023

Hydrology and Earth System Sciences
Discussions

Functional data analysis to quantify and investigate controls on and changes in baseflow seasonality

Kathryn A. Leeming[1], John P. Bloomfield[2], Gemma Coxon[3], Yanchen Zheng[3]

[1] British Geological Survey, Keyworth, NG12 5GG, UK
[2] British Geological Survey, Wallingford, OX10 8BB, UK
[3] School of Geographical Sciences, University of Bristol, Bristol, BS8 1SS, UK

Correspondence to: Kathryn A. Leeming (kle@bgs.ac.uk)

Abstract. Baseflow is the delayed component of streamflow from subsurface storage and is critical for sustaining ecological flows and ensuring water resource security. Understanding controls on and changes in baseflow, including the seasonality of baseflow, is therefore an important task. Baseflow seasonality has typically been investigated using pre-defined hydrological seasons. Instead, here, we investigate baseflow seasonality using data-led approaches that identify and cluster average annual baseflow hydrographs that exhibit early-, mid-, or late-seasonality. We apply a novel functional data analysis (FDA) approach and examine temporal changes in the timing of seasonal peaks in annual standardised baseflow hydrographs for 671 catchments across Great Britain (GB). We use data from the CAMELS-GB dataset for the period 1976 to 2015 split into two twenty-year time blocks (1976-1995 and 1996-2015). Functional clustering enables groups of catchments with similar distributions between time blocks to be identified. Changes in baseflow seasonality with time are investigated by identifying and characterising catchments that move between functional clusters and time blocks, while analysis of the timing of baseflow peaks provides additional temporal resolution to the early-, mid-, and late-season discretisation generated by the functional clustering. The analysis shows that baseflow seasonality has a spatio-temporally coherent structure across GB and catchment characteristics are a first order control on the form of seasonal baseflow clusters. Changes in climate are inferred to be the first order control on changes in baseflow seasonality between the two time blocks. A change to earlier seasonal baseflow in snow-melt influenced catchments in upland northern GB is associated with systematic warming across the two time blocks, and a move to earlier (later) baseflow seasonality across lowland southern, central and eastern (western, north-western and northern) catchments in GB is associated with earlier (later) seasonality in effective rainfall (defined as precipitation minus potential evapotranspiration). These changes in baseflow seasonality in non-snow-melt influenced catchments are consistent with the proposition that, in temperate environments, climate warming leading to vegetation phenology-mediated changes in evapotranspiration may be modifying the timing of hydrological cycles.

Keywords: baseflow seasonality; Functional Data Analysis (FDA); CAMELS-GB; effective rainfall; snow-melt

https://doi.org/10.5194/hess-2023-202
Preprint. Discussion started: 29 August 2023

Hydrology and Earth System Sciences
Discussions

1 Introduction

Baseflow is the delayed component of streamflow fed by subsurface storage between precipitation and/or snow-melt events (Tallaksen, 1995; Price, 2011; Zhang et al., 2017; Gnann et al., 2019; Singh et al., 2019). It is of interest for a number of reasons: baseflow can act to regulate the quality and temperature of stream flow (Jordan et al., 1997; Gomez-Velez et al., 2015; Hare et al., 2021), it supports ecological flows and ecosystem functioning (Poff et al., 1997; Boulton, 2003), and, importantly for water resources, it sustains surface flows at times when there is a deficit in precipitation and so baseflow is a significant component of streamflow during episodes of low flow and drought (Smakhtin, 2001; Miller et al., 2016). Consequently, there is a need to understand the controls on and changes in baseflow, including the seasonality of baseflow – the focus of the current study.

Baseflow is a hydrological phenomenon that represents an integrated, whole-catchment response to meteorological and other environmental change signals (Bloomfield et al., 2009; 2011; Price, 2011) as well as to water resource management practices, such as abstraction and discharge within catchments (Bloomfield et al., 2021). It typically exhibits catchment-specific responses over a wide range of spatio-temporal scales to variability or changes in driving climatology and longer-term climate change, and to changes in catchments, such as land-use and land-cover change. For example, in an analysis of trends in annual baseflow in 99 catchments in the Missouri River Basin, Ahiablame et al. (2017) described a strong positive correlation between increasing trends in precipitation and baseflow and a weak negative correlation with agricultural land use. Tan et al. (2020) recently investigated the impacts of climate change and human-induced changes in vegetation on baseflow at a global scale using data from 2,374 streamflow stations for the period 1970 to 2016. They used a combination of linear regression and the Mann-Kendall nonparametric trend test to identify the magnitude of changes in annual and seasonal baseflow. They found that changes were region-dependent due to regional differences in baseflow-generation processes, and that, at a global level, changes in precipitation, terrestrial water storage (TWS), normalized difference vegetation index (NDVI) and temperature contributed to the majority of changes in annual and seasonal baseflow. They also noted that seasonal baseflow changed, in part, due to the shift from snowfall to rainfall and warming effects on glacial retreat and the timing of snow-melt. This latter phenomenon has been well-documented in observational studies for a number of snow-dominated catchments, particularly in North America (Barnett et al. 2005; Barnett et al., 2008; Leppi et al., 2012; Kormos et al., 2016).

One of the motivations for the present study comes from work on the effects of climate change on vegetation phenology (Piao et al., 2019) and the implications for ecohydrology (Chen et al., 2022), and the concept that changing vegetation phenology may impact both the magnitude and the timing of key fluxes in the ecohydrological cycle. In addition to the effect of climate warming on changes in the timing of flows in snow-dominated catchments, anthropogenic climate warming may already be modifying the magnitude of evapotranspiration (ET) through earlier spring phenology (greenup) and delayed autumn

https://doi.org/10.5194/hess-2023-202
Preprint. Discussion started: 29 August 2023


Hydrology and Earth System Sciences
Discussions

phenology (senescence) so extending the growing season (Chen et al., 2022). For example, in a study of the phenological response of vegetation in eastern USA over the period 1992 to 2011, Kim et al (2018) found that phenological changes were increasing long-term changes in ET in addition to those driven by climatic warming (earlier greenup being more significant than delayed senescence) and that ecohydrological modelling of the phenological changes indicated reduced annual stream discharge. Similarly, Geng et al (2020) documented an extended growing season in the Luanhe Basin, China, over the period 1982–2015 (again, primarily due to earlier greenup) and demonstrated that although the main regulation of annual runoff was precipitation, changes in the length of growing season also played a key role in the changes of annual runoff. Unlike Tan et al (2020), where baseflow seasonality was investigated by analysing the characteristics of changes in baseflow within an analytical framework of pre-defined hydrological seasons, e.g. hydrological spring as defined by March-April-May (MAM), here, as with phenological studies (Kim et al., 2018; Geng et al., 2020; Chen et al., 2022), we are interested in quantifying changes in the timing of seasonal hydroecological signals, in our case baseflow, precipitation, temperature and effective rainfall.

Previous investigations of change in baseflow have typically used non-parametric methods such as the Mann-Kendall trend test, or slope-based methods (e.g. Ahiablame et al., 2017, Mohammed and Scholz, 2016, Bosch et al. 2017, Tan et al., 2020) to characterise annual and / or seasonal trends. Consistent changes across time can be identified using these methods. However, a drawback to using trend tests such as these is that only monotonic changes can be identified and the tests are not particularly suitable to characterise changes in seasonality with time other than by simply segmenting the data into pre-defined 'seasons' (as in Tan et al., 2020). In addition, care must be applied to the choice of test as autoregression and seasonal behaviour in a time series can affect the performance of trend tests (Hirsch and Slack, 1984). Previous studies of seasonality in precipitation (Liebermann et al., 2012) and modelled precipitation under future climate change (Dunning et al., 2018) have identified and analysed the timing of annual peaks and troughs in the hydrographs, however, this approach does not use all the information in the hydrograph to define the seasonality. An alternative approach is to use Functional Data Analysis (FDA) (Bouveyron et al., 2015b). FDA treats the data (for example seasonal hydrographs) as a curve, rather than as discrete, sequential temporal observations in a time series, and allows for general annual shapes in the data to be identified and compared. Functional clustering enables groups of catchments with similar seasonal patterns of baseflow across time blocks, i.e. that exhibit early-, mid-, or late-seasonality, to be identified. A number of other studies have applied FDA to hydrological data. For example, FDA has been applied to the analysis of fluvial flood events over time (Ternynck et al., 2016, Alaya et al., 2020), comparison of modelled versus observed discharge time series (Larabi et al., 2018), clustering of locations on a river network according to nitrate concentration (Haggarty et al., 2015) and, spatio-temporal analysis of precipitation and other climatic data (Suhaila and Yusop, 2017; Ghumman et al., 2020). However, to date, FDA has not yet been applied to the quantification of changes in hydrological seasonality.

To address the gap in observation-based understanding of changes in and controls on the seasonality of baseflow in catchments in temperate settings, a large-sample data set for Great Britain (GB), CAMELS-GB (Coxon et al., 2020a; 2020b) is analysed here using a novel application of FDA methods (Chebana et al., 2012; Ternynck et al., 2016). Data from the CAMELS-GB

https://doi.org/10.5194/hess-2023-202
Preprint. Discussion started: 29 August 2023


Hydrology and Earth System Sciences
Discussions

dataset catchments for the period 1976 to 2015 is split into two twenty-year time blocks (1976-95, 1996-2015) to characterise the distributions of average annual patterns of baseflow and how these vary over time. Following a description of the study area and analytical methods, functional clusters of early-, mid-, and late-season baseflow curves are identified and changes in the timing of seasonal peaks in the annual standardised baseflow hydrographs are evaluated and discussed in the context of potential catchment and climate controls on baseflow seasonality.

2 Study area, data and methods

2.1 Study area

The study area, GB, consists of England, Scotland and Wales (Fig. 1(a)) and includes a wide range of climate–landscape–water management features as described in Coxon et al. (2020b). Catchments in the north and north-west of the study area typically have higher mean elevations than those in the south and south-east of GB and the prevailing climatology reflects the broad gradient in catchment physiography. Wet and cooler conditions with reduced evapotranspiration are typically prevalent in the north and west of GB compared with relatively dry and warmer conditions in the south-east (Fig. 1(b)).

Annual mean precipitation over England shows no systematic trends with time since records began in 1766 and there has been no attribution of changes in annual mean precipitation to anthropogenic factors (Jenkins et al., 2008; Watts et al., 2015). Precipitation in the UK is, however, seasonal and variable over a range of spatio-temporal scales, with a tendency towards drier summers in the south-east and wetter winters in the north-west (Jenkins et al., 2008; Watts et al., 2015) and to showing significant inter-annual variations, including episodes of meteorological extremes (Bloomfield and Marchant, 2013). Although annual average precipitation has not changed significantly in the observational record, there is a tendency for increasing winter rainfall and with more winter rain falling during intense events (Jenkins et al., 2008; Burt and Ferranti, 2011; Jones et al., 2012). Air temperature has increased by about 1 °C between 1980 and 2015 (Jenkins et al., 2008; Watts et al., 2015) consistent with long-term global warming trends. However, there have been few studies of historical changes in associated evapotranspiration. Kay et al. (2013) documented increased potential- and actual-evapotranspiration (PET and AE) across GB between 1961 and 2012 and Watts et al. (2015) speculated that it is reasonable to hypothesize that PET has increased in line with decadal scale warming of air temperatures over GB, but a causal relationship has yet to be established.

High-productivity aquifers are largely found in the south-east and east of GB (Allen et al., 1997; Bloomfield et al., 2009; Marchant and Bloomfield, 2018), whereas less productive aquifers and non-aquifers are generally more extensive in the west and north-west. Catchments in which clay-dominated soils overlie mudrock and clay bedrock formations and catchments with extensive glacial till deposits are typically present in central and eastern areas (Bloomfield et al., 2009; Bricker and Bloomfield, 2014; Bloomfield et al., 2021). These regional variations in underlying hydrogeological conditions are reflected in spatial variation in Baseflow Index across GB (Fig. 1(c)) with the highest Baseflow Index associated with streams flowing over the unconfined Chalk (a fractured microporous limestone of Cretaceous age and the main aquifer in GB) of southern, south-east

https://doi.org/10.5194/hess-2023-202
Preprint. Discussion started: 29 August 2023


Hydrology and Earth System Sciences
Discussions

and eastern England (Fig. 1(a)), and the lowest baseflow in northern and western catchments where low permeability-low storage bedrock and superficial deposits predominate (Allen et al., 1997; Bloomfield et al., 2021).

In a previous analysis of baseflow data from CAMELS-GB using multiple linear regression, Bloomfield et al. (2021) showed that, even though natural covariates, such as topography, aridity, and fractional area of highly productive fractured aquifers, provide the main explanatory power for Baseflow Index (BFI) across the study area, BFI is also affected by groundwater abstraction and to a lesser extent discharges to rivers from sewage treatment works. Groundwater abstraction is focused on major aquifers, particularly the Chalk, in GB, however, discharges to rivers from sewage treatment works have no regional focus and are typically highest in catchments with large populations and high use of water (Bloomfield et al., 2021, Figure 2).

2.2 Data

Hydro-climatic timeseries for 671 catchments in Great Britain are used from the CAMELS-GB dataset (Coxon et al., 2020a; 2020b). These data are a combination of UK National River Flow Archive (NRFA) and meteorological time series, provided at a daily resolution for the period 1970 to 2015. The streamflow series were collected by agencies including the Environment Agency, Natural Resources Wales and the Scottish Environmental Protection Agency and then compiled and quality checked by the NRFA. The data feature a good spatial coverage of GB, and over 80 % of the locations have under 20 % missing flow data, converted to mm day^{-1}.

The daily precipitation data in CAMELS-GB are derived from the CEH-GEAR dataset (Keller et al., 2015). These data consist of observations from Met Office UK gauges, quality checked and converted to grid format using natural neighbour interpolation. Snow fraction is taken from the CAMELS-GB dataset (Coxon et al., 2020a) and given by fraction of precipitation falling as snow for days colder than 0 °C.

The CAMELS-GB temperature data used in this study are catchment daily averaged temperature from the CHESS-met dataset (Robinson et al., 2017) and CAMELS-GB PET data used in this study are catchment daily averaged PET for a well-watered grass based on the Penman–Monteith equation (Robinson et al., 2020).

The baseflow series are derived from the daily CAMELS-GB streamflow series using the Lyne-Hollick filter (Ladson et al., 2013). The Lyne-Hollick digital filtering approach is chosen as this enables separation of hundreds of baseflow series without requiring additional estimation of parameters. The monthly average baseflow values for each location are calculated, then the average seasonal shapes for each time block (1976-1995 and 1996-2015) are formed by taking the median over the years in the time block. Any locations and months with fewer than ten years of data within a twenty-year time block are removed from the dataset. This means that catchments may be present in one time block but not the other. These seasonal shapes are standardised to have mean zero and standard deviation of one, so that the FDA method considers different seasonal shapes rather than absolute levels or the magnitude of annual variation of baseflow.

BFI is taken from CAMELS-GB (Coxon et al., 2020a) and is estimated by the ratio of mean daily baseflow to daily discharge, where hydrograph separation has been performed using the Ladson et al. (2013) digital filter.

https://doi.org/10.5194/hess-2023-202
Preprint. Discussion started: 29 August 2023

Hydrology and Earth System Sciences
Discussions

To investigate changes over time in the hydro-climatic time series, the data are split into two twenty-year time blocks: hydrological years (October-September) 1976-1995, 1996-2015. Peaks and troughs in the standardised functional baseflow curves are identified as days from the start of the hydrological year.

2.3 Functional data analysis

Functional data consists of representations of functions $\{X_1, \dots, X_n\}$ observed on a domain $[t_1, t_2]$. In this work temporal domains are considered, but functional data can also be considered over other domains, such as spatial applications. The monthly data are considered to be observations from a curve on the domain $[0,12)$ months. The functional approach allows discretely sampled data to be considered as observations from a process acting on a continuous domain. Unlike the previously cited work using FDA in hydrological contexts (Haggarty et al., 2015, Ternynck et al., 2016, Suhaila and Yusop, 2017, Larabi et al., 2018, Alaya et al., 2020, Ghumman et al., 2020), the seasonal shapes are standardised to remove their mean and give unit variance for each curve. This allows the shapes defining the baseflow seasonality to be compared rather than the absolute values.

To identify similarities between the average annual baseflow curves at different locations and time blocks we apply the *funFEM* clustering method (Bouveyron et al., 2015b). The *funFEM* package runs the algorithm of the same name (Bouveyron et al., 2015a), which aims to find a discriminative functional subspace to separate the curves into different clusters. This involves starting with a matrix of coefficients containing a basis representation of each curve, then assuming a latent subspace exists such that the coefficients of the curves separate into distinct clusters. The clustering approach is different to testing methods such as Larabi et al. (2018) or Suhaila and Yusop (2017) as there is no explicit test of difference, instead this method allows for characterisation of the seasonal curves over space and time blocks.

Given original data from one curve, $X_i \in \{X_1, \dots, X_n\}$ and chosen basis with basis functions $\{\psi_j\}_{j=1}^{p}$ the curve is converted to basis coefficients using the representation:

$$X_i(t) = \sum_{j=1}^{p} \gamma_{i,j}\, \psi_j(t) \quad (1)$$

This is done using least squares to fit the basis coefficients γ by assuming that noisy versions of the curves are observed. These γ coefficients are then transformed to a subspace with dimension at most $K-1$, where K is the number of clusters. This projection and separation is performed by iterating over three steps detailed below: Function determination (F), Expectation (E) and Maximisation (M). This iteration is based upon the assumption that in the discriminant subspace, the coefficients of the curves are from a multivariate Gaussian with mean and covariance matrix dependent on the cluster membership. The basis coefficient vector for each curve:

$$\Gamma_i = (\gamma_{i,1}, \dots, \gamma_{i,p})' \quad (2)$$

is related to the corresponding subspace coefficients

$$\Lambda_i = (\lambda_{i,1}, \lambda_{i,d})' \quad (3)$$

via the noisy transformation

https://doi.org/10.5194/hess-2023-202
Preprint. Discussion started: 29 August 2023


Hydrology and Earth System Sciences
Discussions

$$\Gamma_i = U\Lambda_i + \epsilon_i. \quad (4)$$

Conditional on allocation to group k, Λ_i is assumed to be normally distributed with mean μ_k and covariance matrix Σ_k.

The FEM steps are required because the underlying cluster group allocation is unknown, so it is not possible to maximise the likelihood of the multivariate Gaussian directly. To find the most discriminative subspace both the parameters of the distribution and the transformation are iterated in the *funFEM* algorithm. Note the naming is derived from functional-EM (expectation maximisation), however the steps are set out in a different order (FME) in Bouveyron et al. (2015b), as follows:

F: Using the posterior probabilities of group membership (for the current iteration), this step aims to find the orientation matrix U that gives a subspace in which the clusters are best separated. This is defined as the subspace such that the variance between the groups is maximised, but the variance within the cluster is minimised.

M: Using the new orientation matrix U the log-likelihood is maximised and the parameters of the multivariate Gaussian distribution are updated.

E: Using the new parameters, update the posterior probabilities of group membership for each curve.

After repeated iteration of these steps the parameters of the Gaussian distribution and the transformation to the subspace are estimated. This latent information provides the cluster allocation and associated probabilities.

In this application to annual signals, the Fourier basis is chosen as it provides a periodic function space over the domain interval so the smoothed seasonal patterns can seamlessly repeat from year to year. The Fourier basis is constructed using a combination of sine and cosine functions defined over the year interval. Over the domain of one year, $[0,12)$ months, seven basis functions are used as a balance between flexibility and complexity. Although information criteria such as the Akaike Information Criterion (AIC) or the Bayesian Information Criterion (BIC) could be applied to choose the number of clusters, it is noted by Bouveyron et al. (2015b) that these can be less efficient in real data scenarios than with simulated test data. Here, the number of clusters is set to three for the seasonal baseflow clustering. This number is chosen to allow for clear comparisons to be made between the shapes of baseflow seasonality (here the clusters represent early-, mid- and late patterns) and for comparison of these groupings spatially and across the time blocks. The parameters of the latent subspace are such that the mean and variance of each cluster can be different, with diagonal covariance matrices Σ_k.

3 Results

In the following sections baseflow seasonality in GB is characterised by describing the spatio-temporal distribution of the functional seasonal (early-, mid- and late-season) baseflow clusters, and controls on those distributions are investigated by comparing functional seasonal baseflow cluster membership with BFI. Controls on changes in cluster membership and baseflow seasonality are then explored.

https://doi.org/10.5194/hess-2023-202
Preprint. Discussion started: 29 August 2023


Hydrology and Earth System Sciences
Discussions

3.1 Characterisation of baseflow seasonality

Figure 2(a) shows the standardised median seasonal baseflow curves (in grey) grouped by cluster (early-, mid-, and late-seasonality), with the cluster means overlaid (in bold colour). Figure 2(b) shows the spatial distribution of the catchments as a function of cluster memberships for each of the two time blocks. Figure 2(c) is a flow diagram showing how the cluster membership of individual catchments varies (or not) across the two time blocks, and Fig. 2(d) shows this information spatially. Table 1 gives the numbers of catchments assigned to each seasonal baseflow cluster and how they vary over time, and Table 2 contains the mean residual variance for each baseflow cluster and the timings of the peak and trough of the annual baseflow curves. Equivalent plots to Fig. 2 and Table 1 for precipitation, temperature and effective rainfall are given in the Supplementary Information (Figs. S1 to S3 and Tables S1 to S3).

The three functional seasonal baseflow clusters have broadly similar annual shapes, as shown in Fig. 2(a). All are slightly asymmetrical with relatively sharper peaks in baseflow during late winter and broader, less well-defined troughs in summer. However, the timing of the peaks (and troughs) in seasonal baseflow varies between the three clusters. Baseflow cluster 1 peaks (troughs) earliest in December (July), with cluster 2 peaking (troughing) around a month later, and cluster 3 peaking later still in February to March. In addition, there is a difference in the within cluster variation in baseflow seasonality with the smallest mean residual variance associated with cluster 2 and greatest associated with cluster 3 (Table 2). The functional clustering resulting in early-, mid- and late-season groups indicates that the main difference between the baseflow seasonality across catchments and time blocks is the timing of the peaks and troughs, rather than the seasonal shapes.

There are no spatial inputs to the clustering algorithm used to generate the clusters, the inputs are simply the average seasonal shapes for each location and time block. Consequently, any spatially coherent grouping of the clusters in Fig. 2(b) is a sign of similar seasonal behaviour in these catchments for a given time block. The maps of cluster membership (Fig. 2(b)) show consistent spatial relationships that persist across the two time blocks. Catchments in cluster 1 (associated with the earliest baseflow seasonality) are predominantly distributed throughout the west of GB across both time blocks with only a few isolated, outlying catchments in the second time block (1996-2015) in this cluster found in south-east England. Catchments in cluster 2 are predominantly distributed through a band running from eastern Scotland, down through central England to south-west England and a second, smaller spatially coherent region running from the easternmost area of England through south-east England to the south-east coast of England. Finally, cluster 3 catchments with the latest baseflow seasonality are predominantly situated in central, eastern and southern England (largely co-incident with the outcrop of the Chalk aquifer, Fig. 1(a)) with a small outlier of catchments distributed in the east Scottish Highlands. However, most of these latter catchments change to cluster 1 catchments by the second time block (Fig. 2(b)). As previously noted, cluster 3 has greatest within cluster variation in baseflow series (Fig. 2(a) and Table 2), which may result from the geographical diversity of regions in GB that contribute to this cluster (Fig. 2(b)) compared with clusters 1 and 2.

The majority of catchments have unchanged cluster membership between the two time blocks (Fig. 2(d)). Even though over time some catchments change cluster allocations, as shown by Table 1 and in the maps in Fig. 2(b) and 2(d), the overall spatial

https://doi.org/10.5194/hess-2023-202
Preprint. Discussion started: 29 August 2023


Hydrology and Earth System Sciences
Discussions

disposition of the seasonal baseflow clusters remains broadly similar between the two time blocks. The most noticeable changes over time are the increase in membership of cluster 1 (the cluster with the earliest peak in seasonal baseflow) from 233 catchments in 1976-95 to 328 catchments in 1996-2015, and the decrease in membership of cluster 3 (the cluster with the latest peak in seasonal baseflow) from 133 catchments to 88 catchments over the same time (Table 1). The increase in membership of cluster 1 is due to movement from clusters 2 and 3, and additionally some catchments that were not included in the first time block due to missing data. In addition, some catchments move from cluster 3 to cluster 2 between the time blocks. Of the catchments included in both time blocks, 97 catchments move to an earlier functional baseflow seasonality cluster in the last time block and 501 catchments do not change cluster. In comparison, only three move to a cluster with later baseflow seasonality.

3.1.1 Controls on baseflow seasonality

Given the strong spatial association between catchments with the latest baseflow seasonality (cluster 3, Fig. 2(b)) and the outcrop of the unconfined Chalk aquifer with associated high BFI (Fig. 1(a) and 1(c)), the association between the timings of baseflow seasonality and the baseflow index (BFI) of the catchments is considered here.

Figure 3 shows histograms of BFI for the catchments plotted by time block with different colours used for the different seasonal baseflow clusters. There is a broad alignment between the baseflow clusters and BFI. In general, the catchments in clusters with earlier (later) baseflow seasonality have lower (higher) BFI. In the later time block there is a wider spread of BFI in cluster 2, and a narrower spread of BFI in cluster 3, corresponding to the move of many of the cluster 3 catchments to cluster 2. Based on these observations, the first order control on baseflow seasonality is inferred to be catchment characteristics (that are in large part reflected by the BFI) rather than features of the driving climatology. This is supported by the observation that when the FDA methodology is applied to precipitation (six-month smoothed precipitation), temperature, and effective rainfall timeseries for the same set of CAMELS-GB catchments (Fig. S1, S2 and S3) there is no clear spatial correlation between the resulting clusters for the seasonality of the climatological variables and the equivalent baseflow seasonality clusters (Fig. 2(b)). For example, the annual patterns in precipitation (Fig. S1) appear to be moving earlier between the two time blocks, but spatially the eastern locations exhibit earlier peaks in baseflow. There are barely discernible differences between the three clusters of temperature seasonality (Fig. S2), indicating a process with shared annual distribution across this study area. For effective rainfall (Fig. S3), in the first time block relatively late effective rainfall seasonality dominates across much of GB (with the exception of the northwestern and northern GB that shows relatively early seasonality), but at the second time block both the early and late season effective rainfall cluster membership switches to the mid-seasonality cluster. Given that the majority of catchments across GB exhibited late seasonality in the first time block, the effect is that the majority of sites across GB move to earlier effective rainfall seasonality (with the exception of those in northwestern and northern GB that move to a slightly later seasonality).

https://doi.org/10.5194/hess-2023-202
Preprint. Discussion started: 29 August 2023


Hydrology and Earth System Sciences
Discussions

3.2 Controls on changes in baseflow seasonality

Although the study area is predominantly temperate in character, there are a few catchments in the north of GB, in the mountains of eastern Scotland, where snow accumulates during winter. Given previously documented links between the effects of warming on snow-melt influenced flow regimes (Barnett et al. 2005; Barnett et al., 2008; Leppi et al., 2012; Kormos et al., 2016; Tan et al., 2020), changes in baseflow seasonality in snow-influenced GB catchments are first explored. Then we explore the evidence for potential controls on changes in baseflow seasonality more generally across the study area.

3.2.1 Changes in baseflow seasonality associated with snow-melt influences catchments

One geographically distinct group of catchments that start in cluster 3 (late season baseflow) in the earlier time block (Fig. 4(a)) are 24 catchments in Scotland. The BFI of these catchments is shown in Fig. 4(b). These catchments have BFI values typically in the range 0.4 to 0.7, rather than the higher BFI associated with the majority of the groundwater-dominated catchments in cluster 3 located in the south east of England primarily on the Chalk (BFI typically >0.9). These Scottish catchments are mostly in locations with relatively high elevation (mean elevation greater than the 80th percentile of GB catchments, 315 m, for all but 4 of these catchments) and have a correspondingly high fraction of precipitation falling as snow (greater than or equal to the 80th percentile for GB catchments).

Of the 24 Scottish catchments, 20 move to cluster 1 (the earliest seasonal baseflow), 3 move to cluster 2 and 1 stays in cluster 3 in the later time block and it is inferred that the catchments that move from cluster 3 to earlier clusters are associated with earlier snow-melt associated with long-term warming in these high catchments. Figure 5 shows that warming January temperatures for the selected catchments over the two time blocks are consistent with the snow melting at an earlier time in the year: Fig. 5(a) shows an increase in median January temperature and Fig. 5(b) shows an increase in the proportion of days with temperature over 0 °C between the two time blocks. The only catchment to stay in cluster 3 (with the latest timed peak) in the later time block is the catchment with lowest temperature, highest elevation and highest proportion (17 %) of precipitation as snowfall. The movement from cluster 3 to cluster 1 for many of the selected catchments indicates a large change in the timing of the peak in annual baseflow (median change ~87 days), consistent with higher temperatures contributing to earlier melting of the snow-pack. This observation is consistent with similar observations from North America (Barnett et al. 2005; Barnett et al., 2008; Leppi et al., 2012; Kormos et al., 2016) and with the findings of Pohle et al (2019) who have shown that, over the period of analysis of this study, increased air temperature due to climate warming has led to earlier snowmelt across the Scottish Highlands.

3.2.2 Other controls on changes in baseflow seasonality

In addition to the Scottish catchments described above, 73 catchments show a change to earlier baseflow seasonality between the time blocks (Fig. 2(d)) and, unlike the Scottish catchments, these catchments that move to earlier baseflow seasonality have a wide spatial distribution across large parts of GB. Two approaches have been used to investigate what may be

https://doi.org/10.5194/hess-2023-202
Preprint. Discussion started: 29 August 2023


Hydrology and Earth System Sciences
Discussions

controlling this systematic change to earlier baseflow seasonality in these catchments across GB: a paired catchment approach and an analysis of changes in the timing of peak seasonal baseflow.

Given the first order control of BFI and hence catchment characteristics on baseflow seasonality (Fig. 3), we first used three sets of paired catchments with different cluster allocations and contrasting changes (or not) in cluster allocation to investigate if there is evidence for catchment controls on the move to earlier baseflow seasonality (see Table S4 and Fig. S4 for details of the paired catchments and results). From the paired catchment study, there was no evidence of systematic control of catchment characteristics on cluster allocations and changes in cluster membership. Instead, we found that all catchments exhibit an earlier seasonality at the second time block regardless of whether a site changes cluster membership with time when details of the monthly baseflow curves are considered. Thus increased temporal granularity (beyond early-, mid-, and late-season functional clusters) in the seasonal analysis may be useful in exploring controls on changes in baseflow seasonality.

As the paired catchment analysis has indicated that changes in baseflow seasonality without a corresponding change in cluster allocation may be affected by factors other than catchment characteristics and to provide increased temporal granularity to the analysis, the timing of the peaks in the baseflow curves are identified for each location and time block and compared with comparable plots for precipitation, temperature, and effective rainfall to investigate if changes in the seasonality of climate variables may be influencing changes in baseflow seasonality. This is shown in Fig. 6 where the direction of change in the seasonal peak timing and magnitude for each of the variables is mapped across GB.

The majority of the symbols have a red outline in Fig. 6(d) indicating that most catchments across GB are exhibiting an earlier peak in baseflow in the second time block compared to the first by up to a month. Conversely, there are a few catchments, indicated by light blue triangles mostly situated in western, northwestern and northern GB, along with a few catchments near London that exhibit later baseflow peaks (by up to a month) in the second time block compared with the first. Most of these catchments with later peaks in baseflow in the second time block are in an area predominately allocated to cluster 1 in both time blocks. This could indicate that there is a limit to how early the peaks of the baseflow curves can be; as curves that are already early do not see a shift to an earlier peak in annual baseflow. Note that the Scottish catchments that have been inferred to be affected by snow-melt processes (section 3.2.1) show peak seasonal baseflow is more than one month earlier in the second time block compared with the first time block (Fig. 6(d)).

The predominant change in precipitation is to earlier peaks (Fig. 6(a)), except for locations in western Scotland where there is evidence for some later peaks in precipitation seasonality. This tendency to earlier peaks in precipitation is strongest (indicated by dark red filled triangles) in the east of GB. There is very little change shown in the timing of peaks in temperature seasonality (Fig. 6(b)), either the peaks are slightly later or there is no change. The map of change in effective rainfall (Fig. 6(c)) shows regions of later seasonal peaks in the western, northwestern and northern GB and in the far southeast of GB, with a band of earlier peaks in lowland southern, central and eastern England and up to the east of Scotland. There is a strong similarity in the overall patterns of changes in peak timings of the effective rainfall and baseflow seasonal curves in Fig. 6(c) and 6(d) (with the exception of a greater tendency for a change to later effective rainfall in the far southeast of England compared with generally earlier baseflow seasonality). Lin's concordance coefficient for the respective changes in peak timing of the baseflow

https://doi.org/10.5194/hess-2023-202
Preprint. Discussion started: 29 August 2023

compared with the effective rainfall is 0.40, whereas comparing the changes in peak timing of baseflow to precipitation and temperature yield coefficients of 0.00 and -0.01. In summary, unlike Tan et al. (2020), we have found no evidence for an association between changes in the seasonality of baseflow with changes in either the seasonality of precipitation or temperature. However, from Fig. 6(c) and 6(d), and given the Lin's concordance coefficients, there is some evidence that changes in the seasonality of effective rainfall may be associated with changes in baseflow seasonality across the study area.

4 Discussion

4.1 Controls on and changes in baseflow seasonality

The findings of this study are consistent with the results of previous analyses of the controls on BFI. Those studies have shown that catchment characteristics such as the fraction of highly productive aquifer, non-aquifers, clay soils, and crop cover within a catchment as well as other factors such as catchment topography are significant controls on BFI (Bloomfield et al., 2009; 2011; Price 2011; Bloomfield et al., 2021). Here, the seasonality of baseflow, i.e. whether catchments exhibit early-, mid- or late-seasonality, has been shown to be closely linked to BFI and hence to catchment characteristics. Consequently, because catchment characteristics vary spatially across the study area, see for example Figure 2 in Coxon et al (2020b), the spatial distribution of baseflow seasonality follows regional variations in catchment characteristics and BFI (Fig. 1(c)).

The findings of the present study also provide new insights into potential controls on changes in seasonal baseflow. The results of this study are consistent with and support the findings of previous studies that have described earlier snow-melt in snow-influenced catchments associated with warming (Barnett et al. 2005; Barnett et al., 2008; Leppi et al., 2012; Kormos et al., 2016; Pohle et al., 2019). However, importantly they provide the first evidence for systematic changes in baseflow seasonality in catchments in temperate settings and give an indication of at least one possible mechanism for those changes. The global-scale analysis of controls on changes in baseflow seasonality of Tan et al. (2020) found that changes in precipitation, terrestrial water storage (TWS), the normalized difference vegetation index (NDVI) and temperature contributed to the majority of changes in annual and seasonal baseflow. Here we have shown that neither change in the seasonality of precipitation nor temperature on their own appear to provide plausible explanations for change in baseflow seasonality. However, we have demonstrated some concordance or association between change in effective rainfall and baseflow seasonality. Given the absence of association between change in precipitation and baseflow seasonality, it is inferred that the change in effective rainfall seasonality is due to changes in the seasonality of PET, the latter presumably driven at least in part by long-term warming across the UK (Kay et al., 2013; Watts et al., 2015). This association is consistent with the vegetation phenology-mediated changes in PET proposed by Chen et al. (2022) under climate warming. An analysis of phenological change over the current study area linked to the characterisation of seasonal hydrological change would provide evidence to strengthen the case for a causal relationship between phenology and hydrological seasonality change in temperate regions, as would repeating the analysis over a wider study area to one that included both water- and energy-limited evaporation from groundwater storage (Condon and Maxwell, 2017; Condon et al., 2020).

https://doi.org/10.5194/hess-2023-202
Preprint. Discussion started: 29 August 2023

4.2 Application of the FDA methodology

This application of functional data analysis to seasonal data has shown differences over time in annual baseflow that may not have been identified using other seasonal approaches such as trend identification following seasonal averaging. The clustering part of the methodology allows for discrete categorisation of the time series and catchments with large changes in seasonal baseflow can be identified as those changing cluster. However, the discrete nature of the clustering analysis can mask smaller temporal changes that are present in the functional representation of the annual series.

The functional data analysis methodology is well-suited to annual patterns as the fitted curves can be defined as cyclic so that the end of the year continues seamlessly to the start of the next. Here the data are averaged over months before applying the functional method, however the daily data could also be used, as in Ternynck et al. (2016). The monthly averaging provides smoother starting curves for analysis. The months were also considered to be equally spaced through the year, which is an approximation. However, whilst the ability to compare the seasonal distributions of the data is a key factor of this work, standardising the curves also limits the information available to the distribution over the year and not the absolute values. This is related to the application of the FDA method to standardised shapes in this work and is not a feature of FDA approaches in general. Means and different variances can be included in FDA but would have distracted from the seasonal patterns in this work.

It is also noted that for this application, the cluster means are of similar shapes but different timing so the clusters are described as earlier and later versions of the annual pattern, but this will not always be the case for the resulting clusters. In other applications where the curves (seasonality or other series) are of different shapes the categorisation of clusters would focus on shape rather than timing.

4.3 Implications for future work – the need for better linked datasets

One of the key challenges we found when conducting this study was identifying clear links between changes in baseflow seasonality and climatic/geophysical/water management catchment attributes. Catchment attributes included in large sample datasets such as the CAMELS-GB dataset (Coxon et al., 2020a) typically represent a snapshot (i.e. specific year) or an average over time, rather than changes or trends. This hinders their application in large-sample studies when identifying controls on changes in time. For example, in an analysis of the effects of water resource management practices on BFI across GB, Bloomfield et al. (2021) demonstrated that both groundwater and surface water abstraction and effluent discharges to rivers have potentially had minor but systematic influences on BFI. They also noted that there have been a range of water resources management schemes and measures including conjunctive-use, low flow alleviation schemes, and hands-off flow measures all of which may modify BFI and potentially the seasonality of baseflow. However, systematic information on such schemes and how they have varied over time is unavailable in CAMELS-GB (Coxon et al., 2020a) and so their influence on the seasonality of baseflow cannot be investigated by the current study.

https://doi.org/10.5194/hess-2023-202
Preprint. Discussion started: 29 August 2023


Hydrology and Earth System Sciences
Discussions

Other changes to catchments over the analysis period that increased the responsiveness of catchments to precipitation, in central and central southern GB, such as changes in land cover (including, for example, increased urban coverage) or stream conveyance, may also have contributed to the switch to earlier seasonal peaking of baseflows and should also be the focus for future research.

More generally, however, to investigate changes in seasonality and how these cascade through the terrestrial water cycle in different catchment settings there is a need for better linked environmental data sets, including both hydrological and phenological data, and continued production of and investment in future large sample data sets such as the CAMELS family of data sets (Addor et al., 2017; Coxon et al., 2020a; Chagas et al., 2020), the related LamaH-CE large sample dataset for Central Europe (Klingler et al., 2021), and the Caravan meta-data set (Kratzert et al., 2023).

5 Conclusions

The FDA approach to clustering patterns in environmental data has been presented and applied to the seasonal distribution of baseflow in GB catchments. By splitting time series from the CAMELS-GB dataset into two twenty-year blocks, changes to earlier seasonal patterns have been seen in most of the baseflow series. In this work, the seasonal distribution rather than absolute values were considered, allowing for identification of shifts of pattern that may be missed by taking seasonal averages. The results from FDA have identified patterns and areas of similarity within the data. The changes to clusters are not formally tested hypotheses but are indicators of change that could be explored further.

The first order control on the membership of functional clusters of similar seasonal baseflow curves is inferred to be catchment characteristics due to their close association with BFI. However, changes in climate, and specifically warming, is the first order effect on changes in baseflow seasonality. For snow-influenced catchments in GB, there has been a shift towards earlier baseflow seasonality inferred to be due to earlier snow-melt associated with global warming. For other catchments, there is a geographical association between a shift to earlier seasonal effective rainfall and earlier seasonal baseflow. Here, again, it is proposed that warming has driven these changes, in this case possibly through a mechanism suggested by Chen et al (2022) namely by a change in vegetation phenology leading to changes in the timing of PET over the study area.

Our approach to identifying patterns of seasonal distribution of hydrological variables is not specific to the GB catchment data in CAMELS-GB. This method could also be applied to other hydrological or climate time series to identify similarities and changes over space and time. A natural extension to this work would be application to further CAMELS (Addor et al., 2017) and related datasets in other countries and regions. It has been shown that the FDA approach described here is a potentially powerful data-driven analytical tool to identify and quantify hydrological changes as a precursor to designing subsequent research and models to address specific process-based questions, such as elements of the causal cascade of process from changing phenology, through changing evapotranspiration and associated effective rainfall to changes in recharge, flow and discharge within catchments. An improvement in data availability, particularly for time-varying anthropogenic factors would aid further study.

https://doi.org/10.5194/hess-2023-202
Preprint. Discussion started: 29 August 2023

Code availability

R code available upon request from corresponding author.

Data availability

All figures contain/are derived from CAMELS-GB data (Coxon et al. 2020a). The CAMELS-GB dataset contains data supplied by the UK Centre for Ecology & Hydrology © UK Centre for Ecology & Hydrology; © University of Bristol; contains Environment Agency information © Environment Agency and/or database right; derived from UK bedrock hydrogeological mapping and UK superficial productivity mapping, BGS © UKRI.

Author contribution

KAL and JPB designed the study. KAL performed the analysis. All authors contributed to the investigation and discussion of results and preparation of the manuscript.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgements

KAL and JPB publish with the permission of the Executive Director of the British Geological Survey, part of UK Research and Innovation (UKRI).

KAL and JPB were funded on this study by the NERC Climate change in the Artic-North Atlantic region and impact on the UK (CANARI) project (NE/W004984/1). GC and YZ were supported by a UKRI Future Leaders Fellowship award (MR/V022857/1).

References

Addor, N., Newman, A. J., Mizukami, N., and Clark, M. P.: The CAMELS dataset: catchment attributes and meteorology for large-sample studies, Hydrol. Earth Syst. Sci., 21, 5293–5313, https://doi.org/10.5194/hess-21-5293-2017, 2017.

Ahiablame, L., Sheshukov, A. Y., Rahmini, V., and Moriasi, D.: Annual Baseflow Variations as Influenced by Climate Variability and Agricultural Land Use Change in the Missouri River Basin, J. Hydrol., 551, 188-202, https://doi.org/10.1016/j.jhydrol.2017.05.055, 2017.

https://doi.org/10.5194/hess-2023-202
Preprint. Discussion started: 29 August 2023

Alaya, M. A. B., Ternynck, C., Dabo-Niang, S., Chebana, F., and Ouarda, T. B. M. J.: Change Point Detection of Flood Events Using a Functional Data Framework, Adv. Water Resour., 137, 103522, https://doi.org/10.1016/j.advwatres.2020.103522, 2020.

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, WD/97/034, https://nora.nerc.ac.uk/id/eprint/13137/1/WD97034.pdf, 333pp, 1997.

Barnett, T. P., Adam, J. C., Lettenmaier, D. P.: Potential impacts of a warming climate on water availability in snow dominated regions, Nature, 438, 303–309, https://doi.org/10.1038/nature04141, 2005.

Barnett, T. P., Pierce, D. W., Hidalgo, H. G., Bonfils, C., Santer, B. D., Das, T., Bala, G., Wood, A. W., Nozawa, T., Mirin, A. A., Cayan, D. R., and Dettinger, M. D.: Human-Induced Changes in the Hydrology of the Western United States, Science, 319, 1080-1083, https://doi.org/10.1126/science.1152538, 2008.

Bloomfield, J. P., and Marchant, B. P.: Analysis of groundwater drought building on the standardised precipitation index approach, Hydrol. Earth Syst. Sci., 17, 4769-4787, https://doi.org/10.5194/hess-17-4769-2013, 2013.

Bloomfield, J. P., Allen, D. J., and Griffiths, K. J.: Examining geological controls on baseflow index (BFI) using regression analysis: An illustration from the Thames Basin, UK, J. Hydrol., 373, 164–176, https://doi.org/10.1016/j.jhydrol.2009.04.025, 2009.

Bloomfield, J. P., Bricker, S. H., and Newell, A. J.: Some relationships between lithology, basin form and hydrology: a case study from the Thames Basin, UK, Hydrol. Process., 25, 2518–2530, https://doi.org/10.1002/hyp.8024, 2011.

Bloomfield, J. P., Gong, M., Marchant, B. P., Coxon, G. and Addor, N.: How is Baseflow Index (BFI) impacted by water resource management practices?, Hydrol. Earth Syst. Sci., 25, 5355–5379, https://doi.org/10.5194/hess-25-5355-2021, 2021.

Bosch, D. B., Arnold, J. G., Allen, P. G., Lim, K.-J., and Park, Y. S.: Temporal Variations in Baseflow for the Little River Experimental Watershed in South Georgia, USA, J. Hydrol., 10, 110–21., https://doi.org/10.1016/j.ejrh.2017.02.002, 2017.

Boulton, A. J.: Parallels and contrasts in the effects of drought on stream macroinvertebrate assemblages, Freshw. Biol., 48, 1173–1185, https://doi.org/10.1046/j.1365-2427.2003.01084.x, 2003.

Bouveyron, C.: funFEM: Clustering in the Discriminative Functional Subspace, R package version 1.1., https://CRAN.R-project.org/package=funFEM, 2015a.

Bouveyron, C., Côme, E., and Jacques, J.: The Discriminative Functional Mixture Model for a Comparative Analysis of Bike Sharing Systems, Ann. Appl. Stat., 9, 1726–60, https://www.jstor.org/stable/43826443, 2015b.

Bricker, S. H. and Bloomfield, J. P.: Controls on the basin-scale distribution of hydraulic conductivity of superficial deposits: a case study from the Thames Basin, UK, Q. J. Eng. Geol. Hydrogeol., 47, 223-236, https://doi.org/10.1144/qjegh2013-072, 2014.

Burt, T.P., and Ferranti, E. J. S.: Changing patterns of heavy rainfall in upland areas: a case study from northern England. Int. J. Climatol., 32, 518–532, https://doi.org/10.1002/joc.2287, 2011.

Chagas, V. B. P., Chaffe, P. L. B., Addor, N., Fan, F. M., Fleischmann, A. S., Paiva, R. C. D., and Siqueira, V. A.: CAMELS-BR: hydrometeorological time series and landscape attributes for 897 catchments in Brazil, Earth Syst. Sci. Data, 12, 2075–2096, https://doi.org/10.5194/essd-12-2075-2020, 2020.

Chebana, F., Dabo-Niang, S., and Ouarda, T. B. M. J.: Exploratory functional flood frequency analysis and outlier detection, Water Resour. Res., 48, W04514, https://doi.org/10.1029/2011WR011040, 2012.

Chen, S., Fu, Y. H., Hao, F., Li, X., Zhou, S., Liu, C., and Tang, J.: Vegetation phenology and its ecohydrological implications from individual to global scales. Geogr. Sustain., 3, 334-338, https://doi.org/10.1016/j.geosus.2022.10.002, 2022.

Condon, L. E., and Maxwell, R. M.: Systematic shifts in Budyko relationships caused by groundwater storage changes. Hydrol. Earth Syst. Sci., 21, 1117–1135, https://doi.org/10.5194/hess-21-1117-2017, 2017.

Condon, L. E., Atchley, A. L., and Maxwell, R. M.: Evapotranspiration depletes groundwater under warming over the contiguous United States, Nature Communications, 11, 873, https://doi.org/10.1038/s41467-020-14688-0, 2020.

Coxon, G., Addor, N., Bloomfield, J. P., Freer, J., Fry, M., Hannaford, J., Howden N. J. K., Lane, R., Lewis, M., Robinson, E. L., Wagener, T., and Woods, R.: Catchment Attributes and Hydro-Meteorological Timeseries for 671 Catchments Across Great Britain (CAMELS-GB), NERC Environmental Information Data Centre, https://doi.org/10.5285/8344e4f3-d2ea-44f5-8afa-86d2987543a9, 2020a.

Coxon, G., Addor, N., Bloomfield, J. P., Freer, J., Fry, M., Hannaford, J., Howden, N. J. K., Lane, R., Lewis, M., Robinson, E. L., Wagener, T., and Woods, R.: CAMELS-GB: Hydrometeorological time series and landscape attributes for 671 catchments in Great Britain, Earth Syst. Sci. Data, 12, 2459–2483, https://doi.org/10.5194/essd-12-2459-2020, 2020b.

Dunning, C., Black, E., and Allan, R. P.: Later wet seasons with more intense rainfall over Africa under future climate change, J. Climate, 31, 9719-9738, https://doi.org/10.1175/JCLI-D-18-0102.1, 2018.

Geng, X., Zhou, X., Yin, G., Hao, F., Zhang, X., Hao, Z., Singh, V. P., Fu, Y. H.: Extended growing season reduced river runoff in Luanhe River basin, J. Hydrol., 582, Article 124538, https://doi.org/10.1016/j.jhydrol.2019.124538, 2020.

Ghumman, A. R., Ateeq-ur-Rauf, Haider, H., and Shafiquzamman, Md.: Functional data analysis of models for predicting temperature and precipitation under climate change scenarios, J. Water Clim. Chang., 11, 1748-1765, https://doi.org/10.2166/wcc.2019.172, 2020.

Gnann, S. J., Woods, R. A., and Howden, N. J. K.: Is there a baseflow Budyko curve?, Water Resour. Res., 55, 2838–2855, https://doi.org/10.1029/2018WR024464, 2019.

Gomez-Velez, J. D., Harvey, J. W., Cardenas, M. B., and Kiel, B.: Denitrification in the Mississippi River network controlled by flow through river bedforms, Nat. Geosci., 8, 941–945, https://doi.org/10.1038/ngeo2567, 2015.

Haggarty, R. A., Miller, C. A., and Scott, E. M.: Spatially weighted functional clustering of river network data, J. R. Stat. Soc., C: Appl. Stat., 64, 491–506, https://doi.org/10.1111/rssc.12082, 2015.

Hare, D. K., Helton, A. M., Johnson, Z. C., Lane, J. W., and Briggs, M. A.: Continental-scale analysis of shallow and deep groundwater contributions to streams, Nat. Commun., 12, 1450, https://doi.org/10.1038/s41467-021-21651-0, 2021.

Hirsch, R. M., and Slack, J. R.: A Nonparametric Trend Test for Seasonal Data With Serial Dependence, Water Resour. Res., 20, 727-732, https://doi.org/10.1029/WR020i006p00727, 1984.

Jenkins, G. J., Perry, M. C., and Prior, M. J.: The climate of the United Kingdom and recent trends, Met Office Hadley Centre, Exeter, UK, 2008.

Jones, M. R., Fowler, H. J., Kilsby, C.G., Blenkinsop, S.: An assessment of changes in seasonal and annual extreme rainfall in the UK between 1961 and 2009, Int. J. Climatol., 33, 1178–1194, https://doi.org/10.1002/joc.3503, 2012.

Jordan, T. E., Correll, D. L., and Weller, D. E.: Relating nutrient discharges from watersheds to land use and streamflow variability, Water Resour., Res., 33, 2579–2590, https://doi.org/10.1029/97WR02005, 1997.

Kay, A. L., Bell, V. A., Blyth, E. M., Crooks, S. M., Davies, H. N., Reynard, N. S.: A hydrological perspective on evaporation: Historical trends and future projections in Britain. J. Water Clim. Change, 4, 193–208, https://doi.org/10.2166/wcc.2013.014, 2013.

Keller, V. D. J., Tanguy, M., Prosdocimi, I., Terry, J. A., Hitt, O., Cole, S. J., Fry, M., Morris, D. G., and Dixon, H.: CEH-GEAR: 1 km resolution daily and monthly areal rainfall estimates for the UK for hydrological and other applications, Earth Syst. Sci. Data, 7, 143–55, https://doi.org/10.5194/essd-7-143-2015, 2015.

Kim, J. H., Hwang, T., Yang, Y., Schaaf, C. L.,Boose, E., & Munger, J. W.: Warming-induced earlier greenup leads to reduced stream discharge in a temperate mixed forest catchment. Journal of Geophysical Research: Biogeosciences, 123, 1960–1975, https://doi.org/10.1029/2018JG004438, 2018.

Klingler, C., Schulz, K., and Herrnegger, M.: LamaH-CE. Large-sample data for Hydrology and environmental sciences for Central Europe, Earth Syst. Sci. Data, 13, 4529–4565, https://doi.org/10.5194/essd-13-4529-2021, 2021.

Kormos, P. R., Luce, C. H., Wenger, S. J., & Berghuijs, W. R.: Trends and sensitivities of low streamflow extremes to discharge timing and magnitude in Pacific Northwest mountain streams, Water Resour. Res., 52, 4990–5007, https://doi.org/10.1002/2015WR018125, 2016.

Kratzert, F., Nearing, G., Addor, N., Erickson, T., Gauch, M., Gilon, O., Gudmundsson, L., Hassidim, A., Klotz, D., Nevo, S., Shalev, G., and Matias, Y.: Caravan – A global community dataset for large sample hydrology, Sci. Data, 10, 61, https://doi.org/10.1038/s41597-023-01975-w, 2023.

Ladson, A. R., Brown, R., Neal, B., and Nathan, R.: A Standard Approach to Baseflow Separation Using the Lyne and Hollick Filter, Australas. J. Water Resour., 17, 173–8, https://doi.org/10.7158/13241583.2013.11465417, 2013.

Larabi, S., St-Hilaire, A., and Chebana, F.: A New Concept to Calibrate and Evaluate a Hydrological Model Based on Functional Data Analysis, J. Water Manag. Model., 26, C442, https://doi.org/10.14796/JWMM.C442, 2018.

Liebermann, B., Blade, I., Kildas, G. N., Carvalho, L. M. V., Senay, G. B., Allured, D., Leroux, S., and Funk, C.: Seasonality of African precipitation from 1996 to 2009, J. Climate, 25, 4304-4322, https://doi.org/10.1175/JCLI-D-11-00157.1, 2012.

Leppi, J., DeLuca, T., Harrar, S., Running, S.: Impacts of climate change on August stream discharge in the Central-Rocky Mountains, Clim. Change, 112, 997-1014, https://doi.org/10.1007/s10584-011-0235-1, 2012.

570 Marchant, B.P., and Bloomfield, J. P.: Spatio-temporal modelling of the status of groundwater droughts, J. Hydrol., 564, 397-413, https://doi.org/10.1016/j.jhydrol.2018.07.009, 2018.

Miller, M. P., Buto, S. G., Susong, D. D., and Rumsey, C. A. J.: The importance of base flow in sustaining surface water flow in the Upper Colorado River Basin. Water Resour. Res., 52, 3547–3562, https://doi.org/10.1002/2015WR017963, 2016.

Mohammed, R., and Scholz, M.: Impact of Climate Variability and Streamflow Alteration on Groundwater Contribution to
575 the Base Flow of the Lower Zab River (Iran and Iraq), Environ. Earth Sci., 75, 1392, https://doi.org/10.1007/s12665-016-6205-1, 2016.

Piao, S., Liu, Q., Chen, A., Janssens, I. A., Fu, Y., Dai, J., Liu, L., Lian, X., Shen, M., and Zhu, X.: Plant phenology and global climate change: Current progresses and challenges. Glob. Change Biol., 25 (6), 1922-1940, https://doi.org/10.1111/gcb.14619, 2019.

580 Poff, N. L., Allan, J. D., Bain, M. B., Karr, J. R., Prestegaard, K. L., Richter, B. D., Sparks, R. E., and Stromberg, J. C.: The natural flow regime, Bioscience, 47, 769–784, https://doi.org/10.2307/1313099, 1997.

Pohle, I., Helliwell, R., Aube, C., Gibbs, S., Spencer, M. and Spezia, L.: Citizen science evidence from the past century shows that Scottish rivers are warming, Sci. Total Environ., 659, 53-65, https://doi.org/10.1016/j.scitotenv.2018.12.325, 2019.

585 Price, K.: Effects of watershed topography, soils, land use, and climate on baseflow hydrology in humid regions: A review, Prog. Phys. Geogr., 35, 465–492, https://doi.org/10.1177/0309133311402714, 2011.

Robinson, E. L., Blyth, E., Clark, D. B., Comyn-Platt, E., Finch, J., and Rudd, A. C.: Climate hydrology and ecology research support system meteorology dataset for Great Britain (1961-2015) [CHESS-met] v1.2, NERC Environmental Information Data Centre, https://doi.org/10.5285/b745e7b1-626c-4ccc-ac27-56582e77b900, 2017.

590 Robinson, E. L., Blyth, E. M., Clark, D. B., Comyn-Platt, E., Rudd, A. C.: Climate hydrology and ecology research support system potential evapotranspiration dataset for Great Britain (1961-2017) [CHESS-PE], NERC Environmental Information Data Centre, https://doi.org/10.5285/9116e565-2c0a-455b-9c68-558fdd9179ad, 2020.

Singh, S. K., Pahlow, M., Booker, D. J., Shankar, U., and Chamorro, A.: Towards baseflow index characterisation at national scale in New Zealand, J. Hydrol., 568, 646-657, https://doi.org/10.1016/j.jhydrol.2018.11.025, 2019.

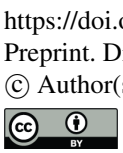
595 Smakhtin, V. U.: Low flow hydrology: A review, J. Hydrol., 240, 147–186, https://doi.org/10.1016/S0022-1694(00)00340-1, 2001.

Suhaila, J., and Yusop, Z.: Spatial and temporal variabilities of rainfall data using functional data analysis, Theor. Appl. Climatol., 129, 229–242, https://link.springer.com/article/10.1007/s00704-016-1778-x, 2017.

Tallaksen, L. M.: A review of baseflow recession analysis, J. Hydrol.. 165, 349–370, https://doi.org/10.1016/0022-
600 1694(94)02540-R, 1995.

Tan, X., Liu, B., and Tan, X.: Global changes in baseflow under the impacts of changing climate and vegetation, Water Resour. Res., 56, e2020WR027349, https://doi.org/10.1029/2020WR027349, 2020.

https://doi.org/10.5194/hess-2023-202
Preprint. Discussion started: 29 August 2023



Hydrology and Earth System Sciences
Discussions

Ternynck, C., Alaya, M. A. B., Chebana, F., Dabo-Niang, S., and Ouarda, T. B. M. J.: Streamflow Hydrograph Classification Using Functional Data Analysis, J. Hydrometeorol., 17, 327–344, https://doi.org/10.1175/JHM-D-14-0200.1, 2016.

Watts, G., Battarbee, R. W., Bloomfield, J. P., Crossman, J., Daccache, A., Durance, I., Elliott, J. A., Garner, G., Hannaford, J., Hannah, D. M., Hess, T., Jackson, C. R., Kay, A. L., Kernan, M., Knox, J., Mackay, J., Monteith, D. T., Ormerod, S. J., Rance, J., Stuart, M. A., Wade, A. J., Wade, S. D., Weatherhead, K., Whitehead, P. G., and Wilby, R. L.: Climate change and water in the UK – past changes and future prospects, Prog. Phys. Geog., 39, 6–28, https://doi.org/10.1177/0309133314542957, 2015.

Zhang, M., Liu, N., Harper, R., Li, Q., Liu, K., Wei, X., Ning, D., Hou, Y., and Liu, S.: A global review on hydrological responses to forest change across multiple spatial scales: Importance of scale, climate, forest type and hydrological regime, J. Hydrol., 546, 44–59, https://doi.org/10.1016/j.jhydrol.2016.12.040, 2017.

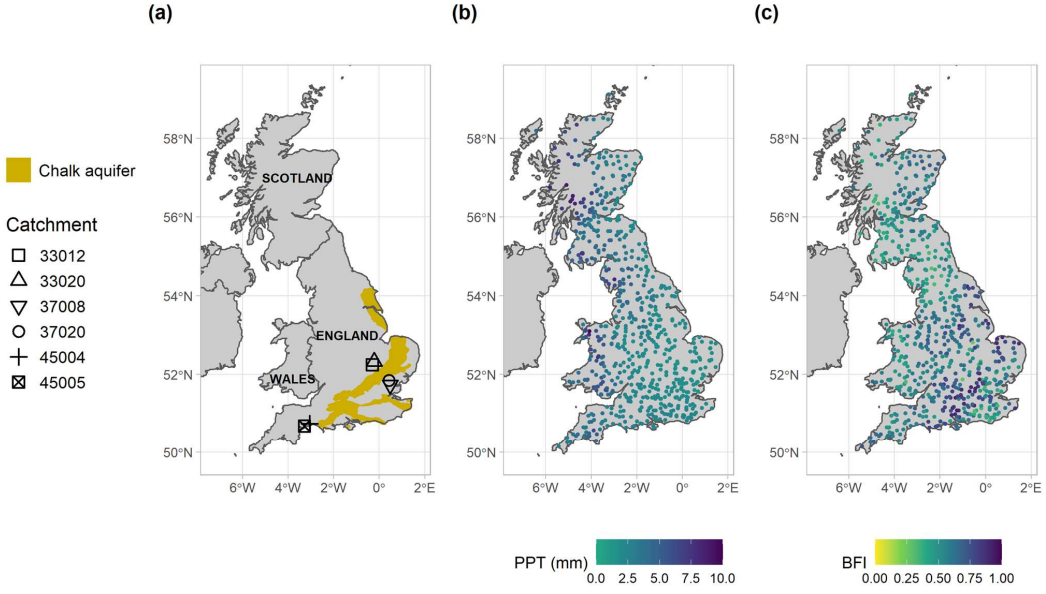
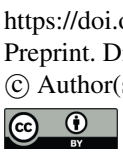


Figure 1: (a) Map of the study area with Chalk aquifer designation, country boundaries and paired catchments (described in Section 3.2.2). (b) Mean daily precipitation at CAMELS-GB catchments. (c) BFI at CAMELS-GB catchments.

https://doi.org/10.5194/hess-2023-202
Preprint. Discussion started: 29 August 2023



Hydrology and Earth System Sciences
Discussions
Open Access
EGU

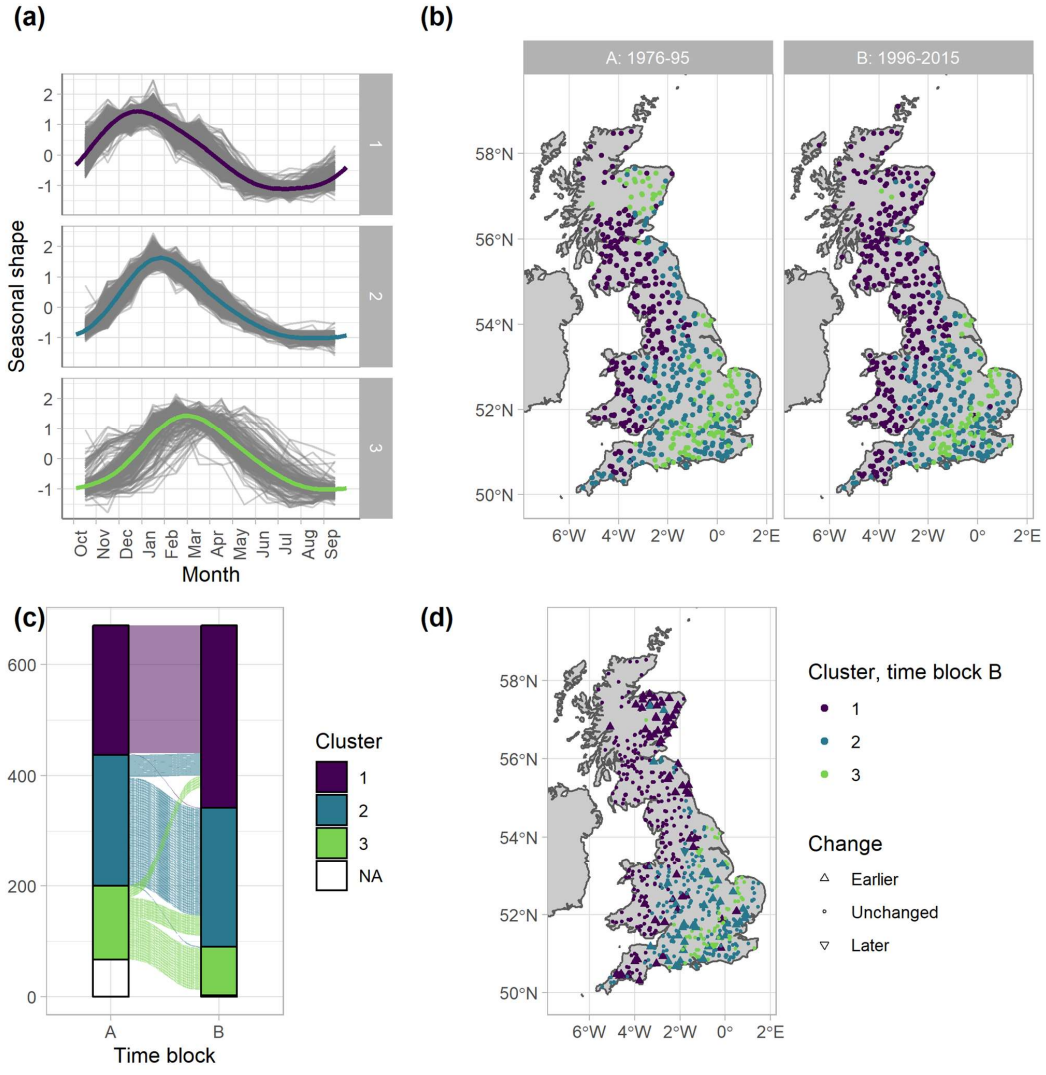


Figure 2: (a) Median annual baseflow for each location and time block plotted by cluster (grey lines), with the cluster means overlaid (in cluster colour and bold). (b) Cluster membership of each location within each time block. (c) Flow diagram of each location showing cluster membership over the time blocks. Each location is a thin line within the plot. (d) Map showing the cluster membership of each location in time block B (shown by colour), with triangular symbols denoting those locations in an earlier or later cluster compared with time block A.

https://doi.org/10.5194/hess-2023-202
Preprint. Discussion started: 29 August 2023

Hydrology and Earth System Sciences
Discussions

Cluster	Block A	Block B
1	233	328
2	237	252
3	133	88

Table 1: Number of locations assigned to each baseflow cluster within each time block.

Cluster	Mean residual variance	Peak timing	Trough timing
1	0.083	85	283
2	0.050	116	319
3	0.182	152	338

Table 2: Mean residual variance per baseflow cluster (variance is calculated per time series and averaged over each cluster), approximate timing of cluster peak and trough (presented as days through the hydrological year)

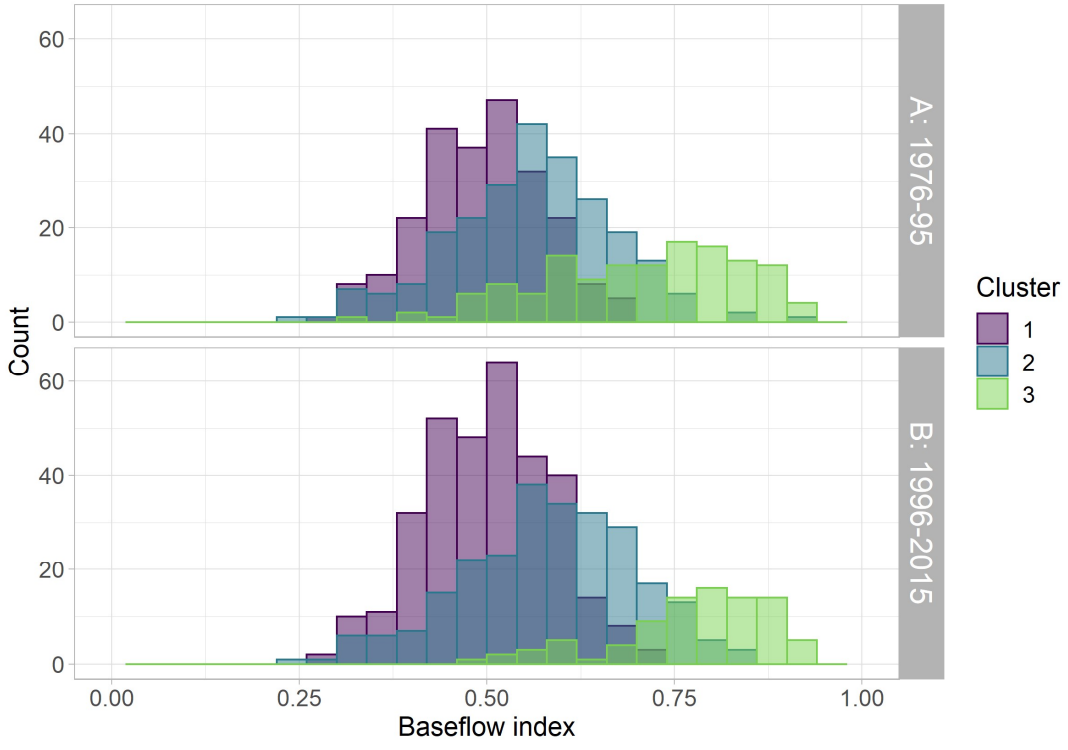
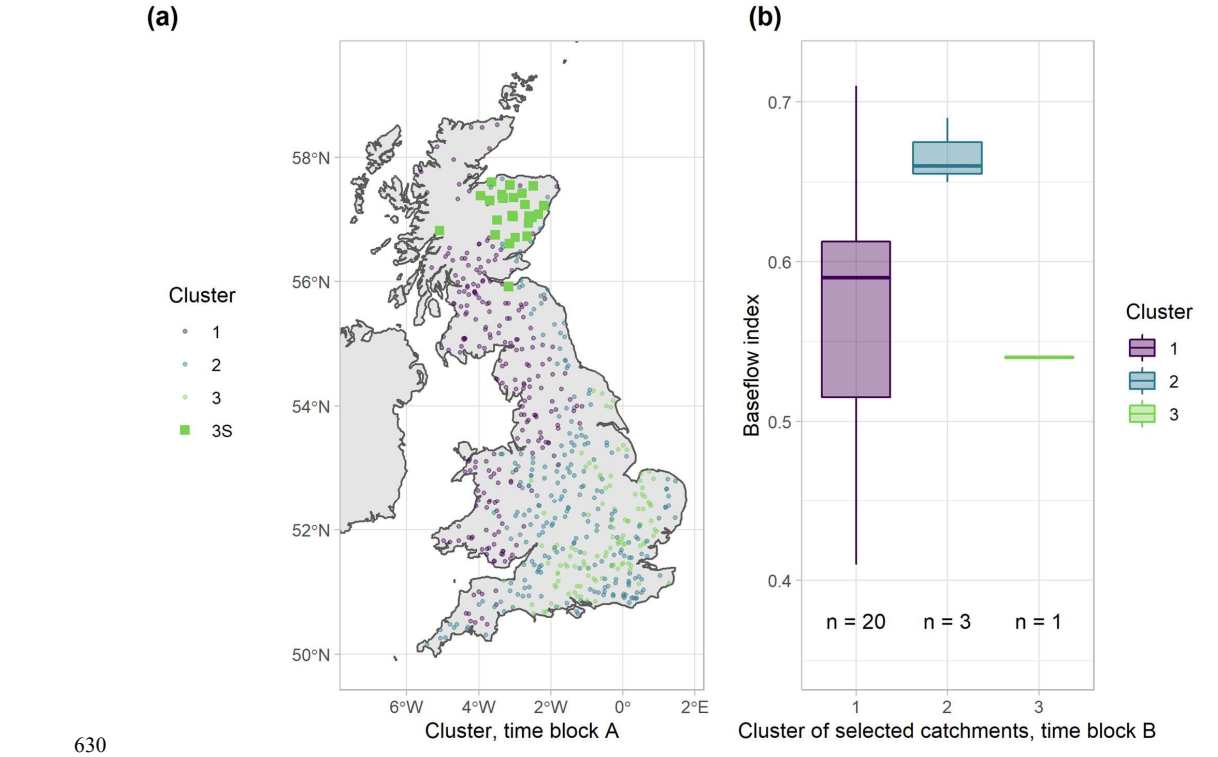


Figure 3: Histogram of baseflow index (BFI) for the cluster allocations in each time block (sub-plots). Histogram bars for different clusters are overlaid (not stacked).

https://doi.org/10.5194/hess-2023-202
Preprint. Discussion started: 29 August 2023



630

Figure 4: (a) Map denoting the selected Scottish catchments that start in cluster 3 in time block A (marked as Cluster 3S). (b) Boxplot of baseflow index (BFI) of the selected catchments, split according to their cluster allocation in time block B.

https://doi.org/10.5194/hess-2023-202
Preprint. Discussion started: 29 August 2023

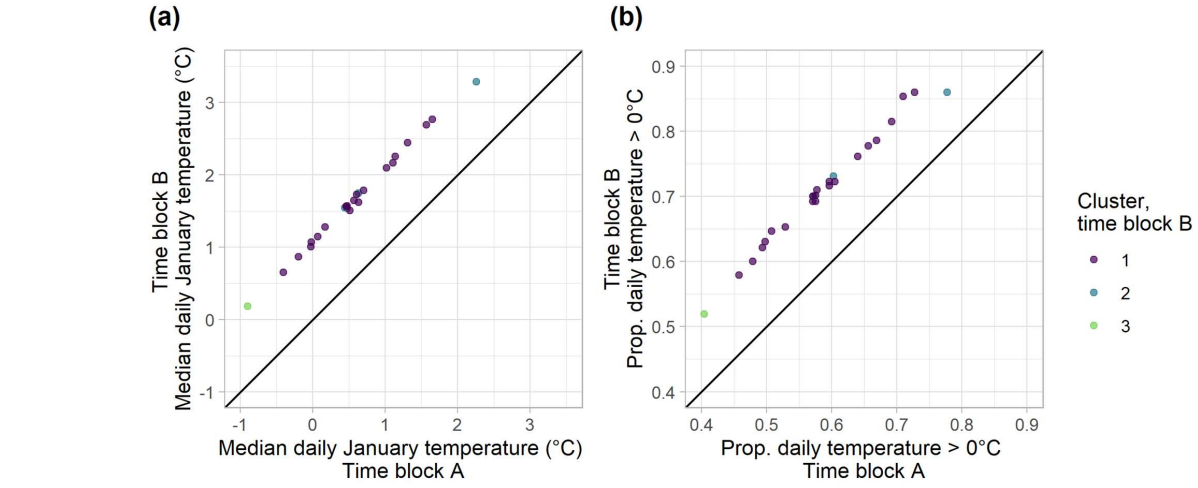


Figure 5: (a) Median January daily temperature for selected Scottish locations, calculated for each time block. (b) Proportion of January days with daily temperature greater than 0 °C.

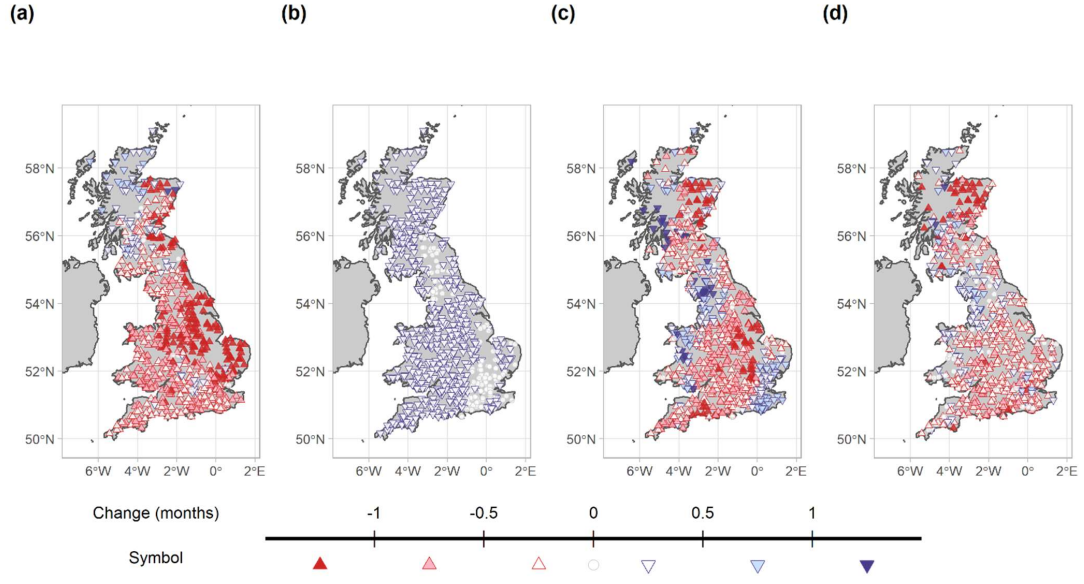


Figure 6: Changes in peak timing between the two time blocks (a) six-month smoothed precipitation, (b) temperature, (c) effective rainfall (PPT – PET), (d) baseflow. Red outlined up arrows indicate locations with earlier peaks, white circles indicate no change,

https://doi.org/10.5194/hess-2023-202
Preprint. Discussion started: 29 August 2023

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

and blue outlined down arrows indicate later seasonal peaks. The shapes are filled with colour according to the magnitude of change (in months).