https://doi.org/10.5194/hess-2021-259
Preprint. Discussion started: 27 May 2021

Hydrology and Earth System Sciences
Discussions

How is Baseflow Index (BFI) impacted by water resource management practices?

John P. Bloomfield[1], Mengyi Gong[2,3], Benjamin P. Marchant[1], Gemma Coxon[4], Nans Addor[5]

[1]British Geological Survey, Wallingford, OX10 8BB, UK
[2]British Geological Survey, Keyworth, NG12 5GG, UK
[3]Lancaster University, Lancaster, LA1 4YR, UK
[4]University of Bristol, Bristol, BS8 1SS, UK
[5]University of Exeter, Exeter, EX4 4RJ, UK

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

Abstract. Water resource management (WRM) practices, such as abstractions and discharges, may impact baseflow. Here the CAMELS-GB large-sample hydrology dataset is used to assess the impacts of such practices on baseflow index (BFI) using statistical models of 429 catchments from Great Britain. Two complementary modelling schemes, multiple linear regression (LR) and machine learning (random forests, RF), are used to investigate the relationship between BFI and two sets of covariates (natural covariates only and a combined set of natural and WRM covariates). The LR and RF models show good agreement between explanatory covariates. In all models, the extent of fractured aquifers, clay soils, non-aquifers, and crop cover in catchments, catchment topography and aridity are significant or important natural covariates in explaining BFI. When WRM terms are included, groundwater abstraction is significant or the most important WRM covariate in both modelling schemes and discharge to rivers is also identified as significant or influential, although natural covariates still provide the main explanatory power of the models. Surface water abstraction is a significant covariate in the LR model but of only minor importance in the RF model. Reservoir storage covariates are not significant or are unimportant in both the LR and RF models for this large-sample analysis. Inclusion of WRM terms improves the performance of some models in specific catchments. The LR models of high BFI catchments with relatively high levels of groundwater abstraction show the greatest improvements, and there is some evidence of improvement in LR models of catchments with moderate to high discharges. However, there is no evidence that the inclusion of the WRM covariates improves the performance of LR models for catchments with high surface water abstraction or that they improve the performance of the RF models. These observations are used to formulate a conceptual framework for baseflow generation that incorporates WRM practices. It is recommended that information on WRM, particularly groundwater abstraction, should be included where possible in future large-sample hydrological data sets and in the analysis and prediction of BFI and other measures of baseflow.

1 Introduction

Baseflow, defined as streamflow fed from the deep subsurface and shallow subsurface storage between precipitation and/or snowmelt events (Tallaksen, 1995; Price, 2011), is a hydrological phenomenon that represents a whole catchment response to

https://doi.org/10.5194/hess-2021-259
Preprint. Discussion started: 27 May 2021

Hydrology and Earth System Sciences
Open Access
Discussions

meteorological and other environmental signals (Bloomfield et al., 2011). It integrates the outcomes of a wide range of natural and human-influenced surface and sub-surface catchment processes. These include geomorpohological controls related to surface topography and soils processes (Vivoni et al, 2007; Price et al., 2011; Singh et al., 2019) and geological controls on baseflow (Longobardi and Villani, 2008; Bloomfield et al., 2009; Kuentz et al., 2017; Carlier et al., 2018). Land use and land cover (LULC) change may also have profound effects on baseflow generation (Zhang and Schilling, 2006; Wang et al., 2014), including effects of changing forest cover and agriculture (Juckem et al., 2008; Zhang et al., 2017) and urbanization (Simmons and Reynolds, 1982; Chang 2007; Dow, 2007; McGrane 2015). Through these processes, the dynamics of baseflow generation is modulated by meteorological variability over a range of spatial and temporal scales (Van Loon and Laaha, 2015; Longobardi and Van Loon, 2018). There is growing evidence for the potential impact of climate change on baseflow across a variety of climate and catchment settings (Wang et al., 2014; Ficklin et al., 2016) and it has been proposed that this should be viewed in the context of increasing sensitivity of changes in droughts and low flows to wider anthropogenic influences (Van Loon et al., 2016; Sankarasubramanianet al., 2020).

In addition to climate and geophysical controls, there is evidence that baseflow may be impacted by water resource management (WRM) practices. Using a baseflow recession analysis, Wittenburg (2003) identified reduced baseflow resulting from abstraction for agricultural irrigation in northern Germany. Similarly, using an empirical analysis of baseflow recession, Wang and Cai (2009) modelled the impact of abstraction and returns on stream flow in a catchment in Illinois, USA and showed that they affected flow recession processes and increased low-flow magnitude but decreased low flow variability. In a statistical analysis of trends in baseflow in a catchment in Florida, USA, Webber and Perry (2006) demonstrated decline in baseflow due to over-abstraction of groundwater. Thomas et al. (2013) emphasised the importance of taking 'human interference' into account when estimating the baseflow recession constant after documenting higher baseflow recession constants associated with groundwater withdrawals from catchments in New Jersey, USA. A variety of modelling studies have simulated the impact of abstraction and other WRM practices on baseflow (Kirk and Herbert, 2002; Parkin et al., 2007; Sanz et al., 2011; de Graff et al., 2014). For example, de Graffe et al., (2014) calibrated the PCR-GLOBWB model with a dynamic allocation scheme to simulate surface water and groundwater abstractions and corresponding feedbacks. They found impacts of WRM were experienced during periods of low flows when the contribution of groundwater through baseflow is the largest and that return flows changed the timing and duration of the low flow periods, causing baseflow to be maintained for longer.

However, there have been no large-sample, data-led analyses of the impacts of WRM practices on baseflow. This is despite new opportunities being offered to investigate and quantify catchment processes through open access, large-sample hydrology datasets (Addor et al., 2020). Such datasets have been used to provide insights into catchment processes and functioning across multiple climate and catchment settings (Beck et al., 2013; Ochoa-Tocachi et al., 2016; Fouad et al., 2018; Gnann et al. 2019; Dudley et al., 2020). CAMELS-GB, a recently published large-sample hydrology dataset for Great Britain (GB) (Coxon et al., 2020a; 2020b), is unusual in that it contains quantitative information on WRM practices including surface water and groundwater abstractions, discharges, and reservoir numbers and capacities at the catchment scale. The aim of the present study is to use the CAMELS-GB large-sample dataset to identify which, if any, WRM activities influence baseflow;

https://doi.org/10.5194/hess-2021-259
Preprint. Discussion started: 27 May 2021

to assess the importance of these activities in the context of other factors known to influence baseflow, such as meteorology, catchment hydrogeology, catchment physiography, and LCLU (Price, 2011); and, to investigate if WRM factors are important in any particular catchment or management settings.

As Price (2011) has noted, there are four broad approaches to quantify baseflow, as follows: low flow event time series; flow-duration statistics; baseflow recession analysis; and, metrics of the proportion of baseflow to total flow, also known as baseflow indices. This study focusses on the latter and specifically on the two measures of Baseflow Index (BFI) reported in CAMELS-GB (Coxon et al., 2020a; 2020b). Two statistical models (multiple linear regression, LR, and machine learning using random forests, RF) are used here to investigate the relationships between BFI and WRM and other catchment covariates in the CAMELS-GB dataset. Note that in the present study the modelling is not designed to compare the respective efficacy of the models in estimating baseflow: this is not a model inter-comparison study (Refsgaard and Knudsen, 1996). Instead, the models are designed to provide complementary evidence for the nature and importance (or not) of WRM practices on influencing BFI.

2 Study area and data

2.1 Study area

This study focuses on 429 catchments across GB (Fig. 1) covering a wide range of climate-landscape-water management features (Fig. 2). Catchments in north and north-west of the study area tend to have higher mean elevations than those in the south and south-east (Coxon et al., 2020a). Meteorology tends to reflect the broad gradient in catchment physiography, with wet and cooler conditions typically prevalent in the north and west of the study area compared with relatively dry and warmer conditions in the south-east (Fig. 2a). The dominant land cover also reflects the prevailing physiographic and meteorological conditions with grass cover predominating in the north and west and crop cover in the south and east, with urban land-cover dominant in London and the other large cities of central and northern England (Fig. 2b). High productivity aquifers are found in the south-east and east (Fig. 2c; 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 that are present in central and eastern areas (Fig. 2d) (Bloomfield et al., 2009; Bricker and Bloomfield, 2014).

https://doi.org/10.5194/hess-2021-259
Preprint. Discussion started: 27 May 2021

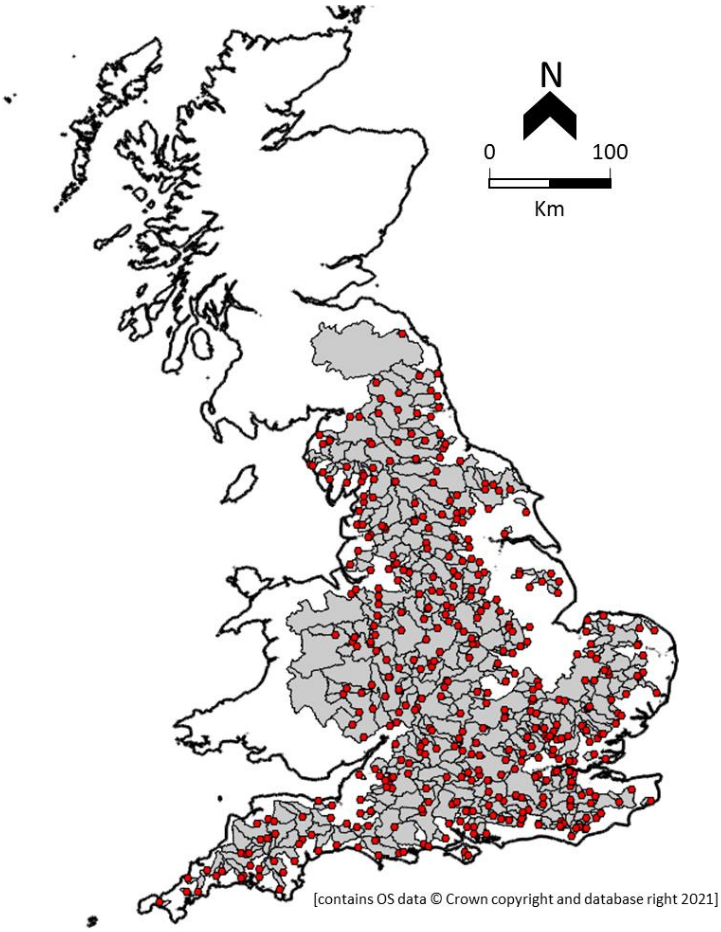


Figure 1. Location of catchments in the study area

Groundwater is used throughout England and forms on average about 30% of the public supply, as well as being used extensively for agricultural irrigation and industrial supplies (Ascott et al., 2017). For 2017 (the last year of published abstraction data) estimated actual abstractions from all sources (except tidal) in England totalled 10,395 million cubic metres (Mm^3), with 8,350 Mm^3 from surface waters and 2,044 Mm^3 from groundwater. Just over half of all these abstractions were used for public supply (5,332 Mm^3) (UK Government, 2020). Regionally groundwater use is more important in southern and eastern England where groundwater abstraction may consist of 100% of public supply (Ascott et al., 2020). Consequently, there is a tendency for more extensive surface water abstraction in the north and more groundwater abstraction in south-east (Fig. 2e and 2f) (Coxon et al., 2020b). Discharges are generally relatively high in catchments in and near major urban centres such as London, central England, and across parts of the north-west (Fig. 2b and 2g) while the highest reservoir capacity is generally associated with catchments in northern and western parts of the study region (Fig. 2h).

https://doi.org/10.5194/hess-2021-259
Preprint. Discussion started: 27 May 2021

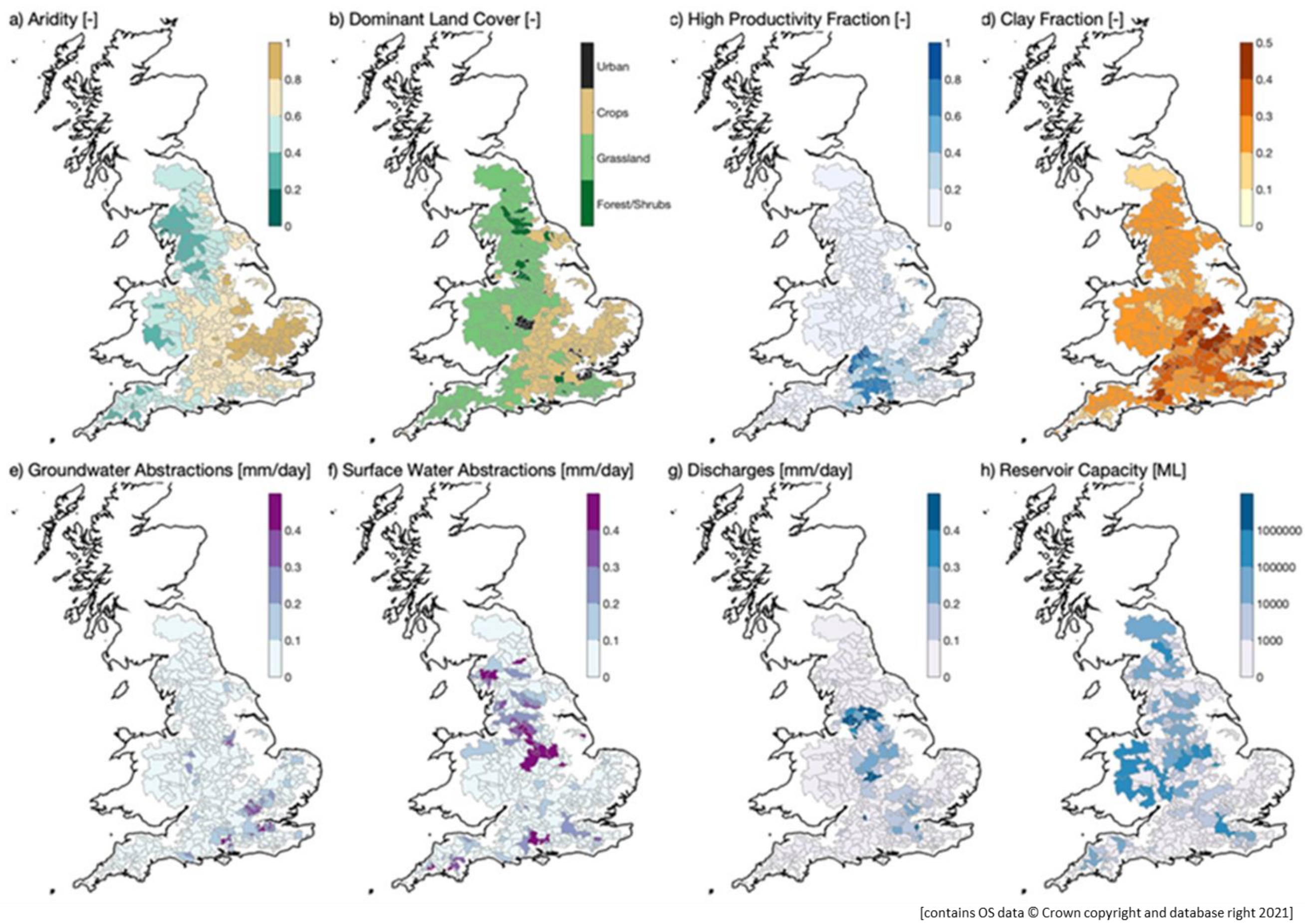


Figure 2. Selected catchment characteristics from CAMELS-GB.

2.2 Data

The data used in this study have been taken from the CAMELS-GB large-sample hydrology data set for Great Britain (GB) (Coxon et al., 2020a, 2020b), itself part of the wider CAMELS (Catchment Attributes and MEteorology for Large-sample Studies) initiative (Newman et al., 2015; Addor et al., 2017; 2020; Alvarez-Garreton et al., 2018; Chagas et al., 2020). CAMELS-GB is unique in that it contains human-influence attributes for some catchments, and it is that sub-set of catchments which are used here. These initially consisted of 442 catchments for which there are 'human influence attributes' (Coxon et al., 2020a, Table2). However, these were further reduced to 429 catchments (Fig. 3) following a consideration of the estimates of BFI that are available for those catchments and the availability of data for the covariates of interest, as described below.

https://doi.org/10.5194/hess-2021-259
Preprint. Discussion started: 27 May 2021

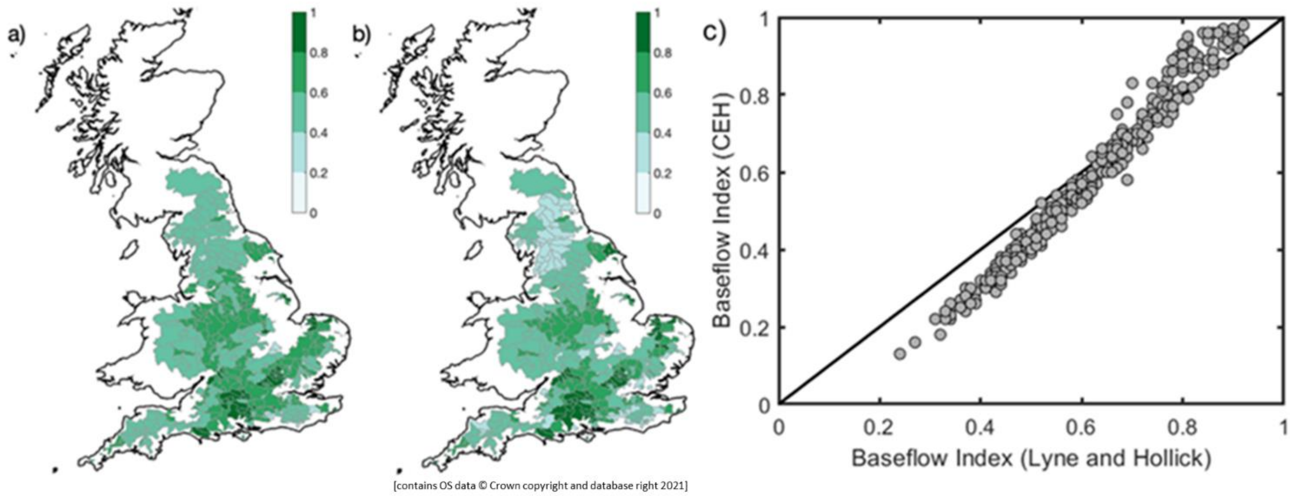


Figure 3. The distribution of a. BFI_CEH, b. BFI_LH, and c. relationship between the two measures of BFI with a 1:1 line for reference.

CAMELS-GB contains two estimates of baseflow. One index, 'baseflow_index_ceh' (BFI_CEH) (Fig3a.), is derived using a method developed by the UK Centre for Ecology & Hydrology and has been used in previous studies of baseflow and flow regimes in Great Britain (Gustard et al., 1992; World Meteorological Organization, 2008). The other, 'baseflow_index' (BFI_LH) (Fig. 3b), was estimated by baseflow separation using the Lyne and Hollick digital filter as implemented by Ladson et al., (2013). A comparison of the two CAMELS-GB baseflow indices (Fig. 3c) confirms the common observation that different techniques used for baseflow separation influence the estimated indices (Eckhardt, 2008; Beck et al., 2013, Addor et al., 2017). Given that the true BFI for any given catchment is unknown, catchments for analysis in this study have been selected where there is a reasonable agreement between the two baseflow indices. Ten catchments were removed where there is an absolute difference between BFI_CEH and BFI _LH of greater than 0.14, equivalent to the largest 2.5%tile of the absolute differences of the population. A further three catchments were removed due to missing covariate data leaving the 429 catchments for analysis (Figs. 1 and 3). Coxon et al. (2020b) note that the CAMELS-GB baseflow indices have been estimated for flow time series available during water years from 1st Oct 1970 to 30th Sept 2015, but that individual time series lengths and completeness may vary between catchments. On average, flow records for the 429 catchments are 91% complete with only 48 catchments with <75% complete. No sites have been omitted from the analysis based on the length of their flow records. Figure 3c shows that that there is a generally good linear agreement between the two estimated BFI indices but that for BFIs up to about 0.7 BFI_CEH is systematically lower than BFI_LH, and that at BFIs above about 0.7 BFI_CEH is systematically higher than BFI_LH. In addition, for sites above a BFI of about 0.7 the correlation between the two indices is reduced.

https://doi.org/10.5194/hess-2021-259
Preprint. Discussion started: 27 May 2021

21 of the CAMELS-GB catchment attributes (Coxon et al., 2020a) related to catchment physiography, climate, hydrogeology, land cover and soils as well as WRM practices have been selected as covariates for analysis (Table A1). The spatial distribution of selected covariates are provided in Fig. 2 and fully described in Coxon et al. (2020b). The 21 CAMELS-GB covariates used in this study have been selected to be representative of each of the major components in a new conceptual model of baseflow generation (Fig. 4). Price (2011) presented a conceptual model that illustrated how components of the terrestrial water cycle and specific catchment processes are related to baseflow based on stores and flows of water in catchments. It didn't, however, incorporate WRM concepts and how these might influence or modify baseflow. In addition, it didn't include aspects of catchment physiography as it focussed on catchment inputs, storage and losses. Fig. 4 is a revised conceptual diagram (building on Price et al, 2011) indicating conceptual relationships between baseflow, catchment compartments and processes that lead to baseflow generation, including aspects of WRM.

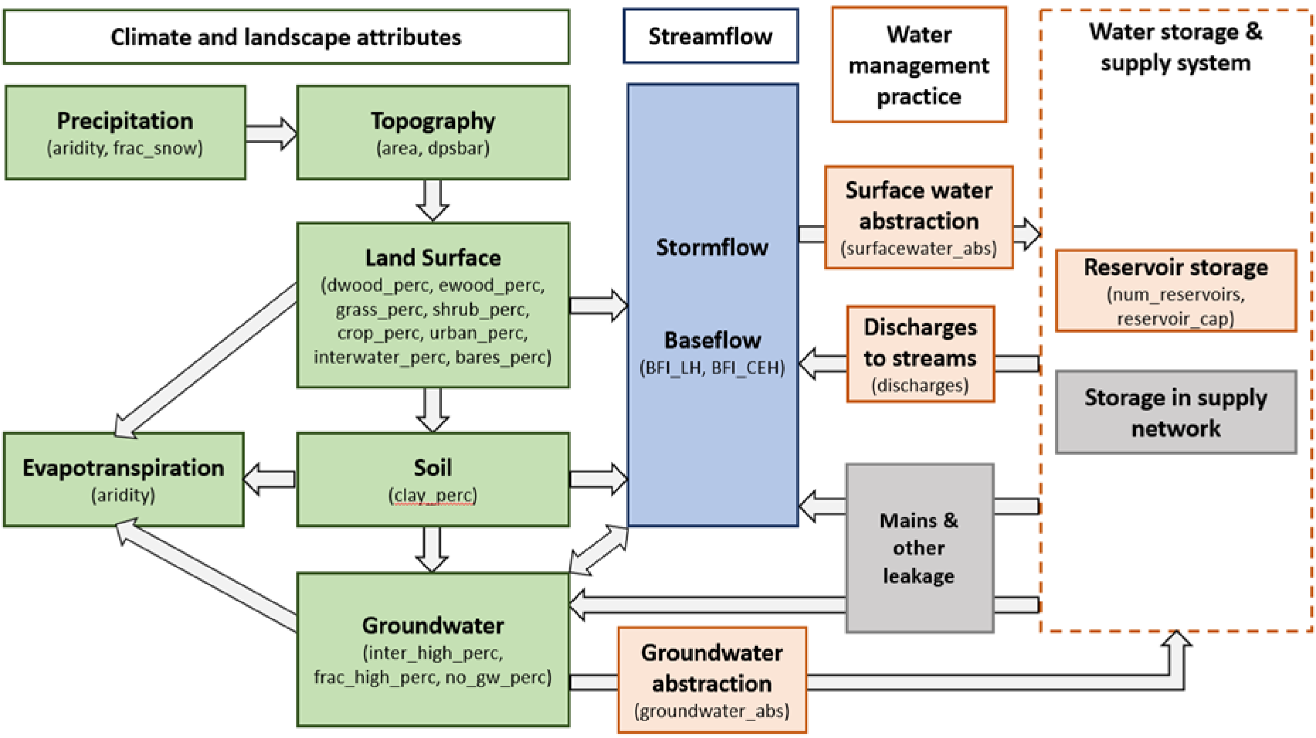


Figure 4. Conceptual model of the relationships between the major compartments of the terrestrial water cycle that exert an influence on baseflow. Baseflow and storm flow components are highlighted in blue, driving climatology, catchment characteristics and compartments are shaded in green, and human influences within the conceptual model are shaded in orange and grey (the latter outside the scope of this study). The 21 CAMELS-GB covariates and the two BFI parameters used in this study are listed against their respective compartments within the conceptual framework.

https://doi.org/10.5194/hess-2021-259
Preprint. Discussion started: 27 May 2021

Hydrology and Earth System Sciences
Open Access
Discussions

3. Modelling methods

Modelling is used in this study not for predictive purposes but to explore model structures and performance to assess the evidence for the relative importance (or not) of WRM practices in influencing BFI. Two complementary modelling schemes, a multiple linear regression (LR) scheme and a random forest scheme (RF), have been applied to two estimates of BFI (BFI_LH and BFI_CEH) using two sets of covariates (Set A and Set B). Set A consists of the 16 natural covariates and Set B consists of all 21 CAMELS-GB covariates, i.e. the combined natural and human influence covariates (Table A1). Consequently, eight models have been developed and evaluated. The LR and RF models are first calibrated for the Set A covariates, then a second separate calibration is undertaken using Set B covariates. The resulting model structures are investigated and their performance in estimating observed BFI compared without and with WRM covariates to understand the influence of WRM covariates on BFI.

The accuracy of the model estimates has been assessed using RMSE and by calculating Lin's concordance coefficient (Lin, 1989) for the predicted and measured values. Lin's coefficient indicates the degree of similarity between two variables. This is in contrast to a more standard Pearson correlation coefficient that is an indication of the explanatory power of a linear relationship between the two variables. Lin's concordance coefficient is calculated both to assess the accuracy of a given model at replicating the training data and in a 10-fold cross-validation procedure to explore the model accuracy at locations that were not used in calibration. If Lin's coefficient is substantially smaller upon cross-validation then this could be an indication that the model is overfitted.

3.1 Linear regression

Regression is commonly used to model the effect of a given set of covariates on a variable of primary interest (Fahrmier et al., 2013). Here generalised linear regression (Dobson, 2002) is used to investigate the relationship between BFI_LH and BEI_CEH and the 21 catchment covariates. Logit transformation was applied to the BFI data, as $y_i = \log(z_i/(1-z_i))$, where z_i is the BFI of catchment i. This is to ensure the fitted, back transformed BFI values are constrained between 0 and 1.
The model considered in this paper is a linear mixed model with the following form,

$$Y = X_1\beta_1 + \cdots + X_p\beta_p + \epsilon \qquad \epsilon \sim \mathcal{N}(0, \sigma^2 R) \qquad (1)$$

where $Y = (y_1, \dots\ y_n)'$ denotes the column vector of BFI values from n catchments, $X_j = (x_{j1}, \dots, x_{jn})'$, $j = 1, \dots p$, are the column vectors of the covariates (catchments attributes). The column vector ϵ represents the model residuals, which are assumed to follow a normal distribution, with covariance matrix $\sigma^2 R$, where R reflects the correlation between transformed BFI values. The linear sums of covariates in a linear mixed model are referred to as the fixed effects and the residual term as the random effects.

In this paper, the model parameters $\beta = (\beta_1, \dots, \beta_p)'$ are estimated using the generalized least squares estimator (Dobson, 2002):

$$\beta = (X'R^{-1}X)^{-1}X'R^{-1}Y \qquad (2)$$

https://doi.org/10.5194/hess-2021-259
Preprint. Discussion started: 27 May 2021

Hydrology and Earth System Sciences
Open Access
EGU
Discussions

These parameter values maximise the likelihood or probability that the data would have arisen from the estimated model. Standard linear regression requires the assumption that the residuals are independent and identically distributed (iid) and that the correlation matrix is equal to the identity matrix, I. Such an assumption can be inappropriate for landscape measurements, as they are not selected according to a randomised design, and are often correlated in space as a result of the underlying geology, climate, etc. In particular, the BFI measurements made from locations closer to each other are more likely to share some similarity than those a long distance apart. If this correlation is ignored then the significance of some model terms could be exaggerated.

A further issue is deciding which of the available covariates should be included. If too few covariates are included then some of the key drivers of BFI variation might be missed and the predictions that result might be imprecise. If too many covariates are included then the model might be overfitted. Some of the terms in an overfitted model replicate the random variation of the BFI values within the calibration data rather than generally applicable relationships between BFI and the covariates. Such a model can accurately predict the BFI for the sites used in calibration but performs less well on other data. The addition of a covariate to a model generally increases the maximised likelihood even in the absence of a true relationship between that covariate and the property of interest. The addition cannot decrease this likelihood because the original model can be achieved if $\beta_{p+1} = 0$. A statistical criterion must be used to decide whether the increase in likelihood upon the addition of a parameter is sufficient to justify the inclusion of that term.

The modelling procedure consists of three steps. In the first step, given the candidate covariates, variable selection is carried out using the stepwise selection routine based on the Akaike Information Criterion (AIC; Akaike, 1973). The AIC, which is twice the negative log-likelihood of the model minus 2 times the number of model parameters,

$$AIC = -2\log\{\mathcal{L}(\beta, \sigma^2; Y)\} - 2(p+1) \qquad (3)$$

The model with the lowest AIC is considered to be the best compromise between accuracy and complexity. The forwards selection routine starts with a model containing no covariates. Each candidate covariate is considered in turn and the AIC that results from its addition to the model is recorded. The covariate which leads to the largest decrease in AIC is added to the model. The iterative procedure continues until none of the remaining covariates lead to a decrease in AIC. This procedure is initially conducted assuming independent residuals (i.e., $R = I$) and is implemented using the “step” function from R package “stats”. In the second step, spatial correlation is assessed by calculating empirical variograms (Cressie, 1993) of the model residuals using the “variogram” function from R package “gstat”. The variogram indicates how the expected squared difference between a pair of residuals varies according to their distance apart. Finally, a model including spatial correlation in the residuals is estimated when inspection of the variogram indicates that this is necessary. Specifically, the spatial correlation is reflected by the non-zero off-diagonal elements in the correlation matrix, R which correspond to the values from an exponential correlation function (i.e., $r(d_{ij}) = \exp(-d_{ij}/\varphi)$, where d_{ij} is the Euclidean distance between two catchments i and j and φ is an estimated model parameter). The model with spatial correlation can be estimated by residual maximum likilihood (REML; Lark et al., 2006) using the “gls” function from R package “nlme”. The statistical significance of each covariate included in

https://doi.org/10.5194/hess-2021-259
Preprint. Discussion started: 27 May 2021

the model (i.e. whether the corresponding regression coefficient is significantly different to zero) is recorded for p values of 0.1, 0.05 and 0.001.

3.2 Machine learning

LR models require assumptions about the nature of baseflow variation that can restrict the patterns of variation which the model can represent. In the past few decades, machine learning methodologies have become increasingly popular for representing complex environmental variation (e.g. Hengl et al., 2018; Lange and Sippel, 2020; Nearing et al 2020). Machine learning algorithms lead to considerably more flexible relationships between environmental variables. For example, regression trees recursively partition observation locations according to a series of binary tests on their covariate values. Each location enters the tree at the initial decision node and then follows one of two branches according to the result of the initial test. Each branch leads to a network of further decision nodes and tests until the location is allocated to a terminal node. The predicted value of the environmental variable at an unobserved location is equal to the average of the training data that are allocated to the same terminal node. The tests at each node are optimised so that the total squared errors for a tree of a specified degree of complexity is minimised.

Regression trees can replicate complex nonlinear relationships that include interactions between different covariates but they are prone to overfitting. A regression tree can perfectly predict the variable of interest for some training data if the number of terminal nodes is equal to the number of training observations but it cannot be expected to perform exactly when predictions are made at other locations. Overfitting can be reduced by introducing stopping criteria to the trees (e.g. each terminal node must contain a specified proportion of the training data) or by using cross-validation to decide whether a particular decision node should be included in the tree. Overfitting might be further reduced by combining an ensemble of regression trees to form a random forest (Breiman, 2001). The trees within the ensemble differ because they are estimated for a different bootstrap sample of the available data and a different subset of the candidate covariates is considered at each decision node. The prediction of the variable of interest at a particular location is equal to the average prediction across all the trees. Addor et al. (2018) found that the inclusion of 500 trees in a random forest considerably stabilised predictions and smoothed relationships between their covariates and BFI measurements.

The random forest interprets the available data as if they were a random sample of the population of interest and does not account for spatial correlation amongst the observations. Also, the relationships implied by a random forest model cannot be stated in a simple parametric form such as Eq. (1) meaning that it can be a challenge to determine the drivers of variation. It is possible to assess the importance of each covariate by shuffling the values of that covariate amongst the observation locations and calculating the reduction in prediction accuracy. However, Schmidt et al. (2020) and Wadoux et al. (2020) advise caution when inferring causal relationships from random forest models. Wadoux et al. (2020) demonstrate that photographs of soil scientists projected across their study area can be utilised by a random forest to accurately map the soil carbon content. They suggest that knowledge discovery from machine learning models requires more than the recognition of patterns and successful prediction. Instead they recommend the pre-selection of relevant environmental covariates and the posterior interpretation and evaluation of the recognised patterns: this is the approach taken here with the selection of 21 covariates representative of the conceptual framework being analysed (Fig. 4).

Random forests are calibrated using the Matlab 'Treebagger' function with each forest containing 500 trees (consistent with Addor et al., 2018), the with-replacement bootstrap sample for each tree being of the same size as the set of available data and one third of the covariates are considered at each decision node. The 'Treebagger' function defines the importance of a covariate in a random forest to be equal to the increase in the mean squared error of all predictions averaged over all trees in the ensemble upon shuffling of the covariate values divided by the standard deviation of the predictions taken over the trees.

4. Results

4.1 Linear regression model structures

Regression models were developed for both BFI_LH and BFI_CEH, with covariates from Set A and from Set B. For all four models the variograms of the residuals indicated substantial spatial correlation. Therefore, the models were re-estimated by REML and included spatial correlation parameterised by an exponential function. Note that although the inclusion of the residual correlation structure does not alter signs of the estimated coefficients, the significance of the model covariates changed. Some covariates were no longer significant after accounting for the spatial correlations. This could imply that part of the variation in BFI that was previously explained by certain covariates in the iid model may have been a result of spatial correlation. The full LR models are listed in Table A2 and the distribution of residuals for the LR models are illustrated in Fig. A1.

Figure 5 shows the covariates identified as significant as well as the sign of the covariates. In this analysis, topography ("dpsbar"), climate ("aridity"), the spatial coverage of fractured aquifers ("frac_high_perc"), of crop coverage ("crop_perc") and of clay soils ("clay_perc") are highly significant in all four LR models, and the spatial coverage of areas with no active groundwater systems ("no_gw_perc") is also a significant covariate in all four models to different levels of significance (Fig. 5). In the LR models using Set B, surface and groundwater abstractions and discharges are all highly significant in explaining the variations in the BFI_LH and BFI_CEH although the number ("num_reservoirs") and capacity of reservoirs ("reservoir_cap") are not significant covariates. Urban land cover ("urban_perc"), previously noted as potentially influencing BFI in the Thames Basin in southern England (Bloomfield et al., 2009), is not a significant covariate in the LR models using covariate Set A once spatial correlation in the covariates has been accounted for, and is not significant at all when WRM covariates are include in the LR models.

In the LR models, the signs of the significant covariates in Fig. 5 are consistent with current process-based understanding of the generation of baseflow (Price et al., 2011) as represented in the revised conceptual model (Fig. 4) and with previous regression models of BFI in the study area (Bloomfield et al., 2009). For example, there is a significant inverse relationship between BFI and the fraction of clay soils within catchments, the fraction of catchments underlain by rocks with

https://doi.org/10.5194/hess-2021-259
Preprint. Discussion started: 27 May 2021

Hydrology and Earth System Sciences
Open Access
Discussions

essentially no groundwater, and the aridity of catchments. Conversely, all LR models indicate a significant positive correlation between BFI and the fraction of catchments underlain by fractured aquifers.

In all four LR models, the Lin's concordance coefficients between the fixed effects predictions and the observed BFI are similar upon training and validation indicating that the models are not overfitted (Table 1). The coefficients for the models using Set A to predict BFI_LH and BFI_CEH are 0.75 and 0.81 respectively. There are moderate negative correlations between the residuals from these models and the surface water and groundwater abstractions and discharges from Set B covariates (Table 2). There are negligible correlations between the residuals and the number and capacity of reservoirs covariates. When the WRM covariates are added to the model the Lin's concordance coefficients increase to 0.82 and 0.85 for BFI_LH and BFI_CEH respectively (Table 1).

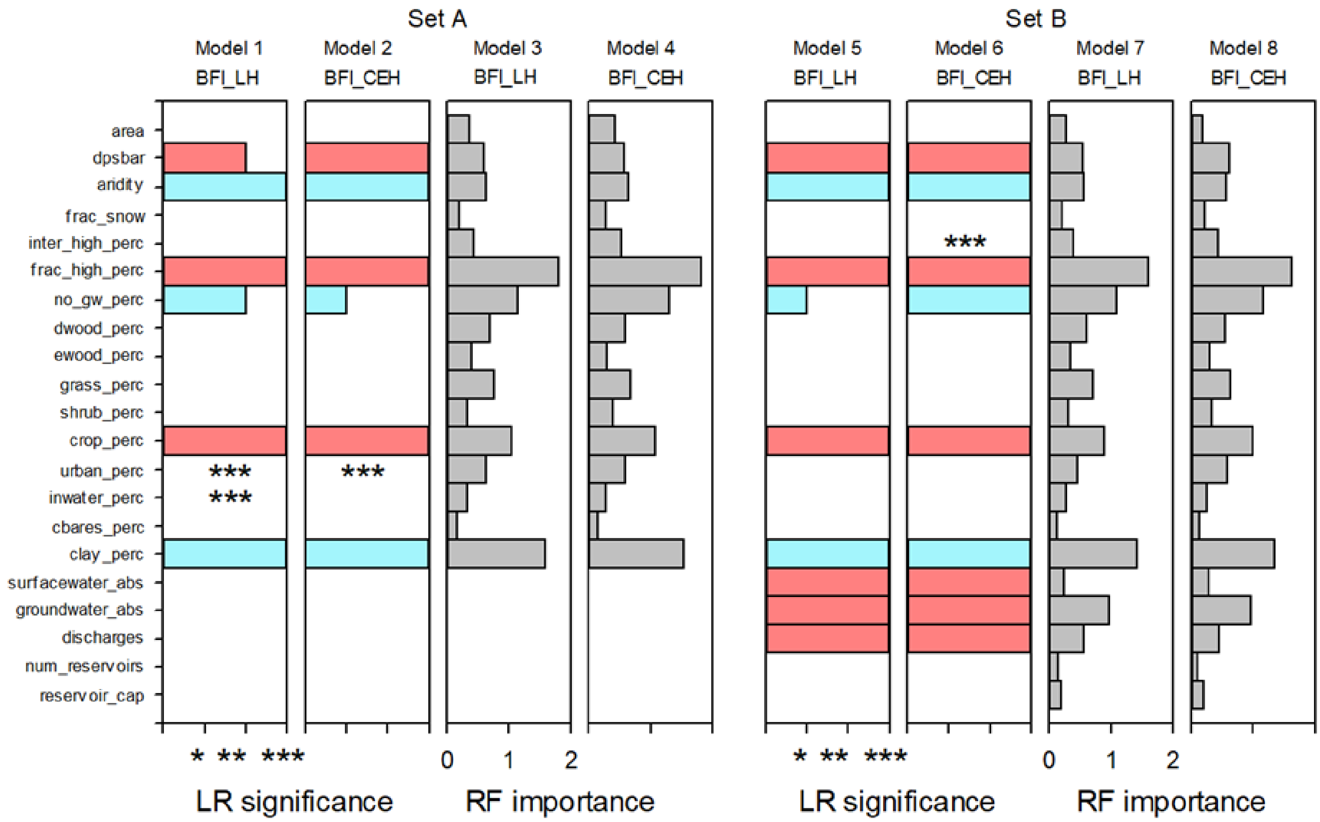


Figure 5. Signs and significance levels of the covariates in the LR models and the relative importance of covariates in the RF models. The signs of the significant covariates in the LR models are indicated using colour (pink for positive, blue for negative) and the corresponding significance levels of the covariates are indicated on the x-axis with asterisks (* for significance level between 0.05 and 0.1, ** for significance level between 0.01 and 0.05, *** for significance level below 0.001). Some covariates were only significant prior to accounting for the spatial correlations. These are marked with asterisks only in the figure at their respective level of significance. Table A1 gives full details of the regression coefficients. Relative RF importance ranges from zero to 2. Table 3 gives the scores for the relative importance of covariates in the four RF models.

https://doi.org/10.5194/hess-2021-259
Preprint. Discussion started: 27 May 2021

Hydrology and Earth System Sciences Discussions

Table 1. Lin's concordance coefficients between LR model predictions and the data.

Model	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Scheme	LR	LR	LR	LR	RF	RF	RF	RF
	Set A	Set A	Set B	Set B	Set A	Set A	Set B	Set B
BFI data	BFI_LH	BFI_CEH	BFI_LH	BFI_CEH	BFI_LH	BFI_CEH	BFI_LH	BFI_CEH
Training	0.75	0.81	0.82	0.85	0.95	0.96	0.97	0.97
Validation	0.75	0.80	0.82	0.84	0.80	0.82	0.81	0.84

Table 2. Pearson correlation between Set A model residuals and Set B model covariates

Model Scheme	LR	LR	RF	RF
BFI data	BFI_LH	BFI_CEH	BFI_LH	BFI_CEH
surfacewater_abs	-0.16	-0.17	-0.21	-0.19
groundwater_abs	-0.36	-0.31	-0.27	-0.27
discharges	-0.27	-0.23	-0.16	-0.13
num_reservoirs	-0.02	-0.02	-0.03	-0.02
reservoir_cap	-0.01	-0.02	-0.04	-0.03

In summary, when spatial correlation effects are taken into account, the LR models do not appear to be overfitted, show a consistent though moderate improvement in explanatory power with the addition of the WRM covariates, and indicate that surface abstraction, groundwater abstraction, and discharges are all significant in explaining the variations in both the estimates of BFI.

4.2 Machine learning model structures

The relative importance of the covariates with respect to estimates of BFI are listed in Table 3 and illustrated in Fig. 5 for the RF models. Lin's concordance coefficients on training data are larger for the RF predictions than for the LR models (Table 1). However, upon cross-validation the RF coefficients decrease and are comparable to the LR model values. This could be an indication of overfitted RFs, perhaps because the spatial correlation previously identified amongst the data (see LR results above) has not been accounted for in the RF models. The most important covariates in the RF models using Set A covariates are consistent for both BFI_LH and BFI_CEH and are in descending order of importance: the fraction of catchments underlain by fractured aquifers ("frac_high_perc"), clay soils ("clay_perc"), extent of catchments underlain by rocks with essentially no groundwater ("no_gw_perc"), and crop and grass coverage ("crop_perc", "grass_perc") (Table 3 and Fig. 5).

https://doi.org/10.5194/hess-2021-259
Preprint. Discussion started: 27 May 2021

The residuals from the RF models are moderately and negatively correlated for the surface water and groundwater abstraction covariates (Table 2). The groundwater abstraction covariate has high importance in both RF models of Set B covariates (Table 3 and Fig. 5) and is in the top third of the covariates for both models. The discharges covariate has a moderate importance in the RF models, but the relative importance of the surface water abstraction covariate and the covariates for the number of reservoirs and for their total capacity is low (Table 3 and Fig. 5).

Table 3. Score of the relative importance of covariates in RF model

Covariate	Model 5	Model 6	Model 7	Model 8
area	0.36	0.43	0.28	0.18
dpsbar	0.59	0.58	0.54	0.62
aridity	0.63	0.64	0.55	0.57
frac_snow	0.20	0.28	0.21	0.22
inter_high_perc	0.43	0.52	0.39	0.44
frac_high_perc	1.81	1.82	1.59	1.62
no_gw_perc	1.14	1.3	1.09	1.16
dwood_perc	0.69	0.59	0.6	0.55
ewood_perc	0.39	0.30	0.34	0.31
grass_perc	0.76	0.67	0.70	0.64
shrub_perc	0.33	0.39	0.30	0.34
crop_perc	1.05	1.07	0.88	1.00
urban_perc	0.63	0.59	0.46	0.58
inwater_perc	0.34	0.27	0.27	0.26
cbares_perc	0.17	0.14	0.12	0.14
clay_perc	1.59	1.53	1.41	1.34
surfacewater_abs	0	0	0.24	0.28
groundwater_abs	0	0	0.96	0.96
discharges	0	0	0.55	0.45
num_reservoirs	0	0	0.15	0.10
reservoir_cap	0	0	0.19	0.21

In summary, RF models show that the majority of the power to explain variations in BFI is due to the natural covariates and when WRM covariates are included in the models, groundwater abstraction is the most important and discharges of moderate importance in explaining both estimates of BFI.

https://doi.org/10.5194/hess-2021-259
Preprint. Discussion started: 27 May 2021

Hydrology and Earth System Sciences
Discussions

4.3 Consistency between model structures

The structures of the LR and RF models (Fig. 5) are broadly insensitive to the BFI being modelled. Although this is reasonable given the correlation between BFI_LH and BFI_CEH (Fig. 3), this observation supports the inference that the models are robust. Importantly for the purposes of the present study, significant covariates in the LR models and covariates with relatively large importance in the RF models are consistent regardless of whether the models are developed using BFI_LH or BFI_CEH (Fig. 5).

There is a high level of agreement between the two modelling approaches regarding the significance or importance of the natural covariates in Set A. Both the LR and RF models indicate the primary importance of the presence of fractured aquifers in controlling BFI. This is consistent with the observation of Bloomfield et al (2009) where the percentage coverage of fractured aquifers in the Thames catchment in southern GB was found to be an important term in LR models of BFI. In the present study, in Models 1 to 4 the catchment fraction underlain by fractured aquifers is either a significant covariate or the covariate with the largest importance (Fig. 5), and catchment fraction of clay soils, those underlain by rocks with essentially no groundwater, and crop coverage are all significant in the LR models or have large importance in the RF models (Fig. 5). The two other catchment covariates identified as significant in the LR models (topography and aridity) also have high or moderate importance in the RF models.

The same natural covariates that are identified as significant or of high importance in the LR and RF models in Set A are also significant or important in models using the Set B covariates (Fig. 5) and the majority of the variation in BFI is described by the natural covariates (Table A2). From these observations, it is taken that WRM practices, rather than being the principle explanatory factor of variance in BFI, act to modify BFI controlled primarily by natural catchment processes. There are also similarities in the significant or importance of WRM covariates between the Set B models. In both cases groundwater abstraction is significant or important, discharges are significant or of moderate importance, and both reservoir numbers and capacities are either not significant or are of low importance. There is however a notable dissimilarity between the model structures with regard to surface water abstraction: it is a significant covariate in the LR models (Fig. 5, see Models 5 and 6) but is not important in the RF model (Table 3 and Fig. 5, see Models 7 and 8).

4.4 Evidence for the impact of water resources management practices

The observations relating to the effect of WRM on BFI have been investigated further by considering the extent to which particular catchment context and management settings influence the respective model performance. Figure 6 shows that, particularly for a number of relatively high BFI catchments in central England and SE England to the north of London (Fig. 6a), the LR model of BFI_LH using only natural covariates appears to underestimate BFI. Similar observations can be made with respect to estimates of BFI_CEH (Fig. A2a), with the additional observation that there are a few catchments in eastern England where the model appears to overestimate BFI. Inclusion of WRM covariates leads to some improvements in LR model estimates of BFI, with the largest improvements being in the high BFI catchments (Fig. 7a and A3a). These improvements are

https://doi.org/10.5194/hess-2021-259
Preprint. Discussion started: 27 May 2021

particularly seen in the relatively high BFI catchments immediately to the north of London (Fig. 6b). Note, however, that addition of WRM covariates to the models does not appear to improve the estimates of BFI_CEH in the catchments in eastern England, where the model still appears to overestimate BFI (Fig. A2b).

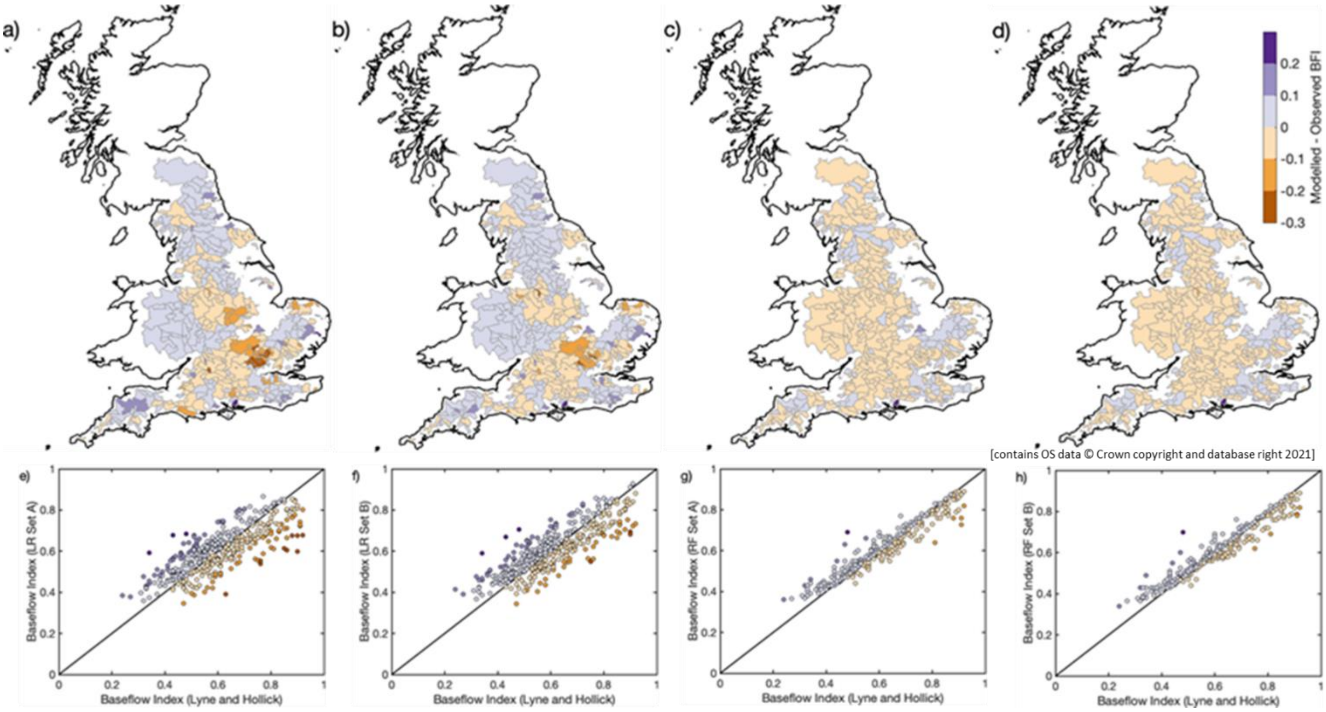


Figure 6. Maps of difference between modelled and observed BFI_LH (a to d) and corresponding scatter plots of BFI_LH against fitted BFI for covariate Sets A and B for LR and RF models (e to h)

To explore further which WRM covariates (groundwater abstraction, surface water abstraction, and discharges) may be contributing to the improvement of the LR models, the distribution of differences between model estimates and observed BFI as a function of the magnitude of the three WRM covariates have been plotted for BFI_LH (Fig. 8) and for BFI_CEH (Fig. A4). Figure 8 shows that for LR models using Set A, underestimation of BFI is greater in catchments with higher levels of groundwater abstraction and, to a lesser extent, with higher discharges. Whereas, there is no apparent systematic association between under- or overestimation of BFI_LH and levels of surface water abstraction. When the WRM covariates are included in the models (Set B), estimates of BFI_LH are noticeably improved in catchments with high levels of groundwater abstraction and to a lesser extent moderate to high discharges. Similar patterns are seen for models of BFI_CEH (Fig. A4). From this it is inferred that most of the improvement in the LR model performance when WRM covariates are included in the models is due to the groundwater abstraction covariate and, to a lesser extent, to the discharge covariate. Inclusion of the surface water abstraction covariate appears to have a negligible influence on estimates of BFI using LR models.

https://doi.org/10.5194/hess-2021-259
Preprint. Discussion started: 27 May 2021

385 Compared with the LR models, differences between estimates of BFI from the RF models and observed values of BFI_LH and BFI_CEH using Set A covariates are small and there are no clear regional patterns in model performance across the study area (Fig. 6 and A2), and, unlike the LR models, there is no evidence of preferential improvement in models as a function of catchment BFI (Fig. 7 and A3). Figure 8 shows that RF models using Set A covariates underestimate BFI in catchments with the highest levels of groundwater abstraction but there is no clear association between the performance of
390 these models and levels of surface water abstraction or discharges. Inclusion of WRM covariates in the RF models (Set B) does not appear to change these relationships: BFI is still underestimated in catchments with the highest levels of groundwater abstraction and there is still no clear association between model performance and levels of surface water abstraction or discharges. Similar relationships also hold for the RF models of BFI_CEH (Fig. A4). There is no noticeable improvement in the performance of the RF models with the inclusion of WRM covariates.

395

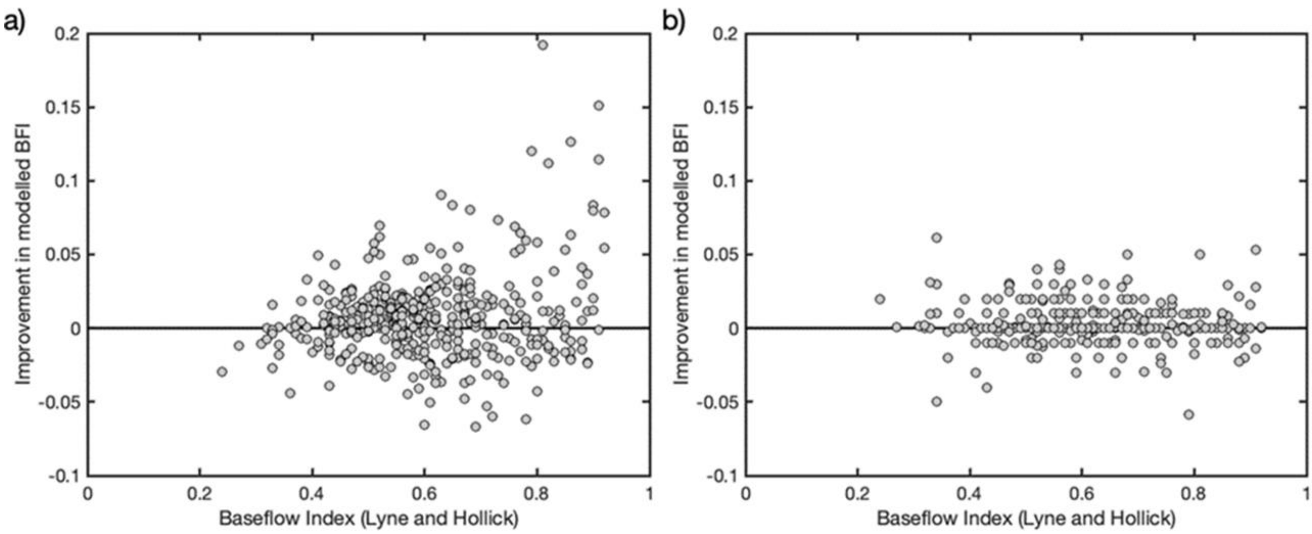
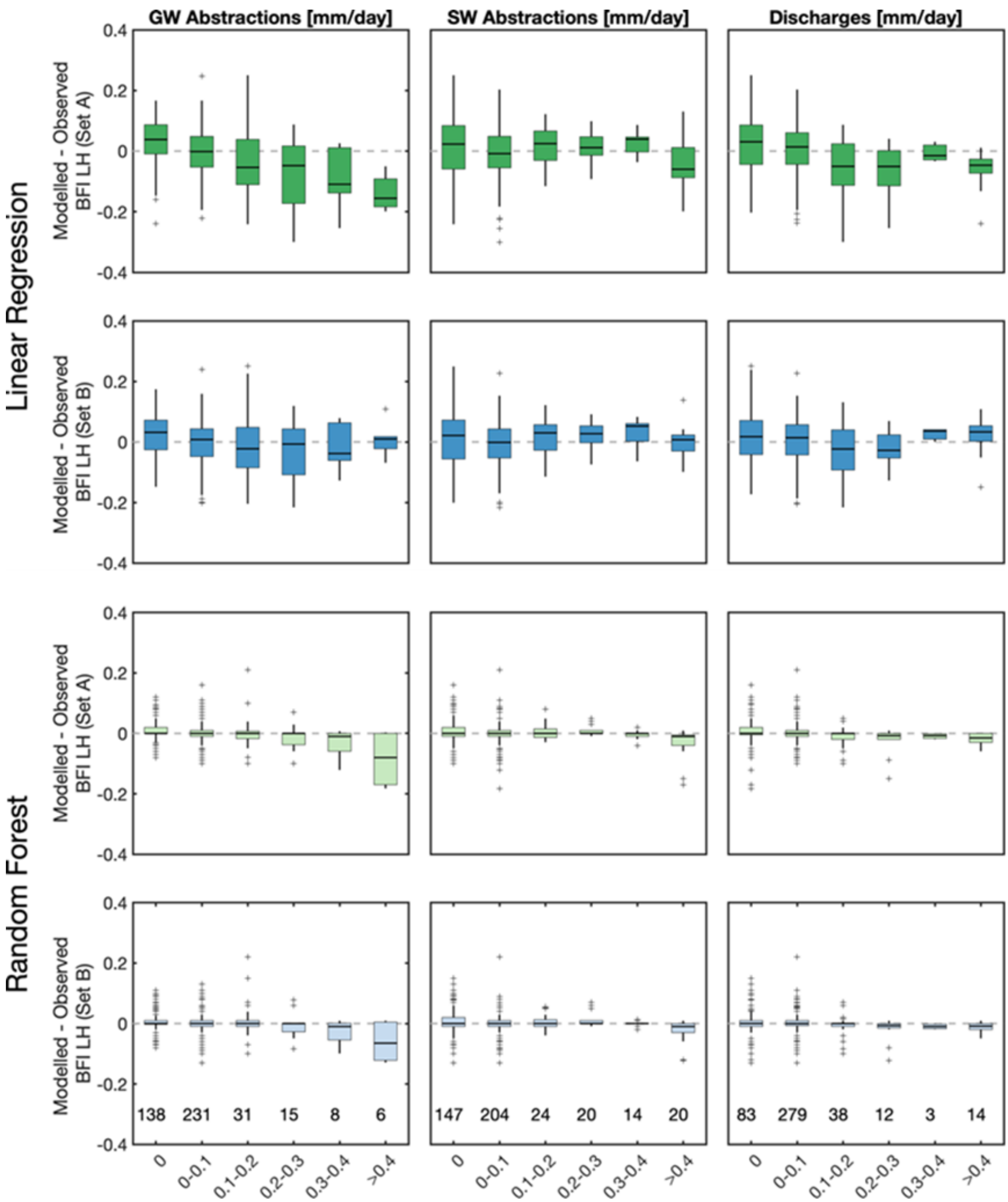


Figure 7. Scatter plots of improvement in modelled BFI as a function of observed BFI_LH for a) LR and b) RF models

https://doi.org/10.5194/hess-2021-259
Preprint. Discussion started: 27 May 2021



400 **Figure 8.** Comparison of observed and modelled BFI_LH for covariate Sets A and B, for LR and RF models and as a function of different human management categories.

https://doi.org/10.5194/hess-2021-259
Preprint. Discussion started: 27 May 2021

Hydrology and Earth System Sciences
Open Access
Discussions

5. Discussion

Given that this is not a model inter-comparison study, the following discussion focusses on insights to be gained from considering both the LR and RF model structures and the performance of the models with the inclusion of WRM covariates. The two contrasting modelling approaches, one relatively simple and tractable (LR modelling) and the other considerably more flexible but potentially harder to interpret (RF modelling), have resulted in remarkably similar model structures with high levels of consistency between both natural and WRM covariates being identified as either significant (LR models) or important (RF models). All results found are plausible in terms of the new conceptual model of baseflow generation (Fig. 4). In addition, analysis of the performance of the models appears to show that the LR models may offer some additional (model dependent) insights into which WRM covariates may be important in which catchment settings.

The results of the models are, however, subject to standard caveats for such types of analysis. Inclusion of spatial correlation in the LR models was necessary and led to some otherwise significant covariates being removed, and the LR models were unable to represent non-linear relationships between the covariates. The RF models did not take into account spatial correlation identified in the LR analysis and there was some evidence of overfitting of the RF models, but they are able to represent any non-linearites that are present between the covariates that could not be included in the LR models. More generally, the CAMELS-GB data on which the models are based are measurements made in the landscape and are not the result of designed experiments (Coxon et al., 2020a; 2020b). Consequently, they have not been sampled at random and confounding variation cannot be controlled in the manner that can be achieved by a statistical experimental design. Therefore, the models must include some potentially additional sources of variation rather than those just found within the covariates of interest (e.g. natural and water management covariates) leading to potentially more complex model structures and potentially to some mismatches between model structure and reality.

Notwithstanding these caveats, both modelling approaches point to the same natural covariates contributing to the majority of variation in BFI, including hydrogeology attributes such as the role of highly productive aquifers and non-aquifers influencing BFI. Given the importance of these covariates for BFI estimates, also previously reported by Bloomfield et al. (2009), it is interesting to note that the hydrogeology attributes in CAMELS-GB are "heuristic, based on hydrogeological inference that is based on mapped lithologies rather than on statistical analysis of borehole yields" (Coxon et al., 2020) (similar heuristic hydrogeological attributes are found in other CAMELS datasets, Addor et al., 2017). As a consequence, although uncertainties in hydrogeology attributes are particularly difficult to constrain they may be particularly influential in the overall performance of the models described here.

Both modelling approaches are largely consistent in identifying the most influential WRM covariates, namely: the importance of groundwater abstraction, the modest effect of discharges to streams, and the unimportance of reservoirs in influencing BFI. What is less clear is why surface water abstraction was significant in the LR model but unimportant in the RF model even though conceptually it is expected to be important in effecting BFI (Fig. 4). Differences between the two modelling approaches may amplify the effects of linear and non-linear relationships between groundwater and surface water

https://doi.org/10.5194/hess-2021-259
Preprint. Discussion started: 27 May 2021

abstraction and other covariates, including other WRM covariates. Water resources in England are currently well-regulated within the context of the European Water Framework Directive and daughter Directives (European Commission, 2000), however, a wide range of WRM practices are used to manage low flow and drought in England including: low flow alleviation schemes, 'hands-off' flow measures, and conjunctive use schemes (Wendt et al., 2020; 2021). Low flow alleviation schemes and 'hands-off' flow measures constrain the amount of water that is abstracted from rivers with abstractions being reduced or stopped at a given low flow trigger level. Conjunctive use schemes, such as the Shropshire Groundwater Scheme (Shepley et al., 2009), use combined management of groundwater and surface water abstractions to maintain ecological flows in catchments where there is an important groundwater resource in the catchment. These different WRM schemes have operated in different catchments in GB in response to perceived local regulatory and management needs, and in any given catchment WRM practices have evolved over the period which the CAMELS-GB baseflow indices have been estimated. All of these factors will affect the relationships between surface water abstraction and the other covariates and hence the respective sensitivities of the models to surface water abstraction effects.

CAMELS-GB reflects the dominant WRM practices for GB only. Consequently, once systematic WRM practice information becomes available for other regions it is recommended that the present study should be extended to test additional WRM attributes and the applicability of the findings in other WRM regimes. For example, CAMELS-GB does not explicitly include information about WRM practices associated with hydropower schemes or seasonal changes in abstraction (e.g. for irrigation), so the effect of such schemes on BFI has not been assessed. In addition, CAMELS-GB does not include any information on within and between catchment water transfers (note the absence of these WRM terms from the conceptual model, Fig. 4). In addition, the approach to assessing the effect of WRM practices on BFI could also be applied and tested for relevance in other climate settings such semi-arid environments (Mwakalila et al, 2002), or where snowmelt is an important component of baseflow generation (Miller et al., 2014; Barnhart et al., 2016) once systematic information on WRM practices is available in those settings.

Finally, there is an active debate on the comparative merits of process-based hydrological modelling and machine learning in hydrological forecasting. Specifically, questions have been asked related to the extent to which hydrological processes and our understanding of the uniqueness of place as encapsulated in our conceptual models of the terrestrial water cycle (Wagner et al., in review) has a role in hydrological prediction in the 'age of machine learning' (ML) (Bevan 2020; Nearing et al 2020). Although the purpose of the present modelling was not to develop models capable of predicting BFI, or of undertaking a model intercomparison between two different stochastic modelling approaches, it is interesting to note that there have been clear benefits in applying both simple statistical models (LR models) and more flexible machine learning approaches (RF models) to the same parameter space to identify common model structures and covariates of interest. Typically ML might have previously been applied to a much larger parameter space – over 100 catchment parameters and variables are available in CAMELS-GB – however limiting the analysis to 21 covariates has enabled direct comparison of LR and RF model structures. Even though the models described here are not process-based and have not been developed to predict BFI in

https://doi.org/10.5194/hess-2021-259
Preprint. Discussion started: 27 May 2021

Hydrology and Earth System Sciences Discussions
Open Access EGU

ungauged catchments, the results provide evidence to extend our current process understanding of baseflow based beyond individual LR (Bloomfield at al., 2009; Carlier et al., 2018; Zhang et al 2020) and RF (Addor, et al., 2018) studies.

6. Conclusions

By exploring model structures from complementary statistical modelling schemes applied to BFI data for GB, the aim of the present study was to identify which, if any, WRM activities influence baseflow; to assess the influence of these activities relative to other natural factors known to effect BFI; and, to investigate if WRM factors are important in any particular catchment settings or management context. The following concluding observations can be made:

- Modelling shows that variation in BFI is predominantly controlled by natural (meteorological and catchment) characteristics, the preeminent of which is the extent of high productivity fractured aquifers within a catchment.
- Although not the major control on variation in BFI, there is evidence that WRM practices systematically modify BFI in GB. Groundwater abstraction appears to be the most influential of these practices. Reservoir storage appears to be unimportant in modifying BFI, at least in the current study area and for the specific reservoir attributes provided in CAMELS-GB.
- The effects of WRM practices on BFI are most evident in groundwater-dominated catchments where there are the highest levels of groundwater abstraction.
- WRM practices can (Fig. 4) and should where appropriate be incorporated in future conceptual models of BFI and baseflow generation, and consequently data and information about WRM practices should be included in future large-sample catchment datasets and in future investigations of baseflow.

7. Data availability

The CAMELS-GB dataset used in this study is available from the Environmental Information Data Centre (EIDC) at https://doi.org/10.5285/8344e4f3-d2ea-44f5-8afa-86d2987543a9 and the supporting paper describing the data (Coxon et al., 2020b) is available from Earth System Science Data at https://essd.copernicus.org/articles/12/2459/2020/

8. Author contributions

JPB designed the study and undertook the literature review. MG and BPM performed the modelling and all authors contributed to the analysis of the results. JPB prepared the manuscript and GC prepared the figures (except Fig. 4 and 5 which were produced by JPB). MG, BPM and NA reviewed and edited the manuscript.

https://doi.org/10.5194/hess-2021-259
Preprint. Discussion started: 27 May 2021

Hydrology and Earth System Sciences
Discussions

9. Competing interests

The authors declare that they have no conflict of interest.

10. Acknowledgements

JPB and BPM publish with the permission of the Director, British Geological Survey (Natural Environment Research Council, UKRI).

11. References

Addor, N., Newman, A. J., Mizukami, N., and Clark, M.P.: The CAMELS data set: 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.

Addor, N., Nearing, G., Prieto, C., Newman, A. J., Le Vine, N. and Clark, M. P.: A ranking of hydrological signatures based on their predictability in space, Water Resour. Res., 54(11), 8792–8812, doi:10.1029/2018WR022606, 2018.

Addor, N., Do, H. X., Alvarez-Garreton, C., Coxon, G., Fowler, K., and Mendoza, P. A.: Large-sample hydrology: recent progress, guidelines for new datasets and grand challenges, Hydrological Sciences Journal, 65, 712-725, https://doi.org/10.1080/02626667.2019.1683182, 2020

Akaike, H.: Information theory and an extension of the maximum likelihood principle. In: Second International Symposium on Information Theory (eds. BN Petrov and F Csaki), pp 267-81, Akadémiai Kiadó, Budapest, 1973.

Alvarez-Garreton, C., Mendoza, P. A., Boisier, J.P., Addor, N., Galleguillos, M., Zambrano-Bigiarini, M., Lara, A., Puelma, C., Cortes, G., Garreaud, R., McPhee, J., and Ayala, A.: The CAMELS-CL dataset: catchment attributes and meteorology for large sample studies - Chile dataset. Hydrol. Earth Syst. Sci., 22, 5817-5846, https://doi.org/10.5194/hess-22-5817-2018, 2018.

Asano, Y., and Uchida. T.: Flow path depth is the main controller of mean baseflow transit times in a mountainous catchment, Water Resour. Res., 48, W03512, https://doi.org/10.1029/2011WR010906, 2012.

Ascott, M. J.: Groundwater resources in the UK, http://www.bgs.ac.uk/research/groundwater/waterResources/GroundwaterInUK/home.html (last access: 8 December 2020), 2017.

Ascott, M. J., Bloomfield, J. P., Karapanos, I., Jackson, C., Ward, R., McBride, A., Dobson, B., Kieboom, N., Holman, I., and Van Loon, A.: Managing public groundwater supplies during drought: Managing groundwater supplies subject to drought: perspectives on current status and future priorities from England (UK), Hydrogeology Journal, https://doi.org/10.1007/s10040-020-02249-0, 2020.

Barnhart, T.B., Molotch, N. P., Livneh, B., Harpold, A. A., Knowles, J.F., and Schneider, D.: Snowmelt rate dictates streamflow, Geophys. Res. Lett., 43, 8006-8016, https://doi.org/10.1002/2016GL069690, 2016.

Bayliss, A.: Catchment descriptors. In Flood estimation handbook (Vol. 5). Wallingford, England: Institute of Hydrology, 1999.

Beck, H. E., van Dijk, A. I. J. M., Miralles, D.G., de Jeu, R. A. M., Bruijnzeel, L. A. S., McVicar, T. R., and Schellekens, J.: Global patterns in base flow index and recession based on streamflow observations from 3394 catchments. Water Resources Res., 49, 7843-7863, https://doi.org/10.1002/2013WR013918, 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.

Breiman, L.: Random forests. Machine Learning, 45, 5-32, https://doi.org/10.1023/A:1010933404324, 2001.

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. and Hydrogeol., 47, 223-236, https://doi.org/10.1144/qjegh2013-072, 2014.

British Geological Survey.: BGS hydrogeology 625k mapping, https://www.bgs.ac.uk/products/hydrogeology/maps.html (last access: 8 April 2021), 2019.

Carlier, C., Wirth, S. B., Cochand, F., Hunkeler, D., and Brunner, P.: Geology controls streamflow dynamics, J. Hydrol., 566, 756-769, https://doi.org/10.1016/j.jhydrol.2018.08.069, 2018.

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.

Chang, H.: Comparative streamflow characteristics in urbanizing basins in the Portland Metropolitan Area, Oregon, USA, Hydrol. Process., 21, 211-222, https://doi.org/10.1002/hyp.6233, 2007.

Clayton, H. J., Morris, S. E., McIntyre, N. R., and Greaves, M.: The hydrological impact of low-flow alleviation measures, Proc. Inst. Civ. Eng. Water Man., 161, 171-180, https://doi.org/10.1680/wama.2008.161.4.171, 2008.

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., Woods, R.: Catchment attributes and hydro-meteorological timeseries for 671 catchments across Great Britain (CAMELS-GB). NERC Environmental Information Data Centre, [data set], available at: https://doi.org/10.5285/8344e4f3-d2ea-44f5-8afa-86d2987543a9 (last access: 8 April 2021), 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.

Cressie, N.: Statistics for Spatial Data Revised Edition. Wiley, 1993.

https://doi.org/10.5194/hess-2021-259
Preprint. Discussion started: 27 May 2021

Hydrology and Earth System Sciences
Discussions

de Graaf, I. E. M., van Beek, L. P. H., Wada, Y., and Bierkens, M. F. P.: Dynamic attribution of global water demand to surface water and groundwater resources: Effects of abstractions and return flows on river discharges, Advances Wat. Res., 64, 21-33, https://doi.org/10.1016/j.advwatres.2013.12.002, 2104

Dobson, A. J.: An introduction to generalized linear models. Chapman & Hall/CRC Boca Raton, 2002.

Dow, C. L.: Assessing regional land-use/cover influences on New Jersey Pinelands streamflow through hydrograph analysis, Hydrol. Process., 21, 185-197, https://doi.org/10.1002/hyp.6232, 2007.

Dudley, R. W., Hirsch, R. M., Archfield, S. A., Blum, A.G., and Renard, B.: Low streamflow trends at human-impacted and reference basins in the United States, J. Hydrol., 580, 124254, https://doi.org/10.1016/j.jhydrol.2019.124254, 2020.

Durant, M. J., and Counsell, C. J.: Inventory of reservoirs amounting to 90% of total UK storage. NERC Environmental Information Data Centre, [data set], available at: https://doi.org/10.5285/f5a7d56c-cea0-4f00-b159-c3788a3b2b38 (last access: 8 April 2021), 2018.

Eckhardt, K.: A comparison of baseflow indices, which were calculated with seven different baseflow separation methods, J. Hydrol., 352, 168-173, https://doi.org/10.1016/j.jhydrol.2008.01.005, 2008.

European Commission.: Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy. https://ec.europa.eu/environment/water/water-framework/index_en.html (last access: 9 April 2021), 2002.

Fahrmier, L., Kneib, T., Lang, S., and Marx, B.: Regression Models, Methods and Applications. Springer, Berlin, Heidelberg, 2013.

Ficklin, D. L., Robeson, S. M., and Knouft, J. H.: Impacts of recent climate change on trends in baseflow and stormflow in United States watersheds, Geophys. Res. Lett., 43, 5079-508, https://doi.org/10.1002/2016GL069121, 2016.

Fouad, G., Skupin, A., and Tague, C. L.: Regional regression models of percentile flows for the contiguous United States: Expert versus data-driven independent variable selection, J. Hydrol.: Reg. Studies, 17, 64-82, https://doi.org/10.1016/j.ejrh.2018.04.002, 2018.

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

Gustard, A., Bullock, A. and Dixon, J. M.: Low flow estimation in the United Kingdom, [online] Available from: http://nora.nerc.ac.uk/id/eprint/6050/ (last access: 8 April 2021), 1992.

Hengl, T., Nussbaum, M., Wright, M.N., Heuvelink, G. B. M., and Gräler, B.: Random forest as a generic framework for predictive modeling of spatial and spatio-temporal variables, PeerJ, 6, Article e5518, https://doi.org/10.7717/peerj.5518, 2018.

Hiederer, R.: Mapping soil typologies: spatial decision support applied to the European Soil Database., Publications Office, Luxembourg. [online] Available from: http://dx.publications.europa.eu/10.2788/87286 (Accessed August 2020), 2013.

Joint Research Centre. 2019. World Atlas of Desertification. https://wad.jrc.ec.europa.eu/patternsaridity , last downloaded August, 2020

590 Juckem, P. F., Hunt, R. J., Anderson, M. P., and Robertson, D. M.: Effects of climate and land management change on streamflow in the driftless area of Wisconsin, J. Hydrol., 355, 123-130, https://doi.org/10.1016/j.jhydrol.2008.03.010, 2008.

Kirk, S., and Herbert, A. W.: Assessing the impact of groundwater abstractions on river flows, Geol. Soc. Spec. Pub., 193, 211-233, https://doi.org/10.1144/GSL.SP.2002.193.01.16, 2002.

Kuentz, A., Arheimer, B., Hundecha, Y., and Wagener, T.: Understanding hydrologic variability across Europe through catchment classification, Hydrol. Earth Syst. Sci., 21, 2863-2879, https://doi.org/10.5194/hess-21-2863-2017, 2017.

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, 25-34, 2013.

Lange, H., and Sippel, S.: Machine Learning Applications in Hydrology. In: Levia, D. F., Carlyle-Moses, D. E.,, Iida, S., Michalzik, B., Nanko, K., and Tischer, A. (eds.) Forest-Water Interactions, Ecological Studies (Analysis and Synthesis), vol. 240. Springer, Cham. 2020.

Lark, R. M., Cullis, B. R., and Welham, S. J. On spatial prediction of soil properties in the presence of a spatial trend: the empirical best linear unbiased predictor (E-BLUP) with REML, European J. Soil Science 57, 787–799, https://doi.org/10.1111/j.1365-2389.2005.00768.x, 2006.

Lin, L. I-K.: A concordance correlation coefficient to evaluate reproducibility. Biometrics, 45, 255-268, https://doi.org/10.2307/2532051, 1989.

Longobardi, A., and Van Loon, A. F.: Assessing baseflow index vulnerability to variation in dry spell length for a range of catchment and climate properties, Hydrological Process., 32, 2496–2509, https://doi.org/10.1002/hyp.13147, 2018.

Longobardi, A., and Villani, P.: 2008. Base flow index regionalization analysis in a Mediterranean area and data scarcity context: role of the catchment permeability index, J. Hydrol., 355, 63-75, https://doi.org/10.1016/j.jhydrol.2008.03.011, 2008.

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.

McGlynn, B., McDonnell, J., Stewart, M., and Seibert, J.: On the relationships between catchment scale and streamwater mean residence time, Hydrol. Process., 17, 175-181, https://doi.org/10.1002/hyp.5085, 2003.

McGrane, S. J.: Impacts of urbanisation on hydrological and water quality dynamics, and urban water management: a review, Hydrol. Sciences Journal, 61, 2295-2311, https://doi.org/10.1080/02626667.2015.1128084, 2016.

Miller, M. P., Susong, D. D., Shope, C.L., Heilweil, V. M., and Stolp, B.J.: Continuous estimation of baseflow in snow melt-dominated streams and rivers in the Upper Colorado River Basin: A chemical hydrograph separation approach, Water Resour. Res., 50, 6986-6999, https://doi.org/10.1002/2013WR014939, 2014.

Morris, D. G., and Flavin, R. W.: A digital terrain model for hydrology, In: Proc. 4th International Symposium on Spatial Data Handling, vol. 1, pp. 250-262, Zürich, 1990.

Muñoz-Villers, L. E., Geissert, D. R., Holwerda, F., and McDonnell, J. J.: Factors influencing stream baseflow transit times in tropical montane watersheds, Hydrol. Earth Syst. Sci., 20, 1621-1635, https://doi.org/10.5194/hess-20-1621-2016, 2016.

https://doi.org/10.5194/hess-2021-259
Preprint. Discussion started: 27 May 2021

Hydrology and Earth System Sciences Discussions
Open Access
EGU

Mwakalila, S., Feyen, J., and Wyesure, G.: 2002. The influence of physical catchment properties on baseflow in semi-arid environments, Journal of Arid Environments, 52, 245-258, https://doi.org/10.1006/jare.2001.0947, 2002.

Newman, A. J., Clark, M.P., Sampson, K., Wood, A., Hay, L. E., Bock, A., Viger, R. J., Blogett, D., Brekke, L., Arnold, J. R., Hopson, T., and Duan, Q.: Development of a large-sample watershed-scale hydrometeorological dataset for the contiguous USA: dataset characteristics and assessment of regional variability in hydrologic model performance, Hydrol. Earth Syst. Sci., 19, 209-223, https://doi.org/10.5194/hess-19-209-2015, 2015.

Ochoa-Tocachi, B. F., Buytaert, W., and De Bievre, B.: Regionalization of land-use impacts on streamflow using a network of paired catchments, Water Resour. Res., 52, 6710-6729, https://doi.org/10.1002/2016WR018596, 2016.

Parkin, G., Birkinshaw, S. J., Younger, P.L., Rao, Z., and Kirk, S.: A numerical modelling and neural network approach to estimate the impact of groundwater abstractions on river flows, J. Hyrol., 339, 15-28, https://doi.org/10.1016/j.jhydrol.2007.01.041, 2007.

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

Price, K., Jackson, C. R., Parker, A. J., Reitan, T., Dowd, J., and Cyterski, M.: Effects of watershed land use and geomorphology on stream low flows during severe drought conditions in the southern Blue Ridge Mountains, Georgia and North Carolina, United States, Water Resour. Res., 47, W02516, https://doi.org/10.1029/2010WR009340, 2011.

Refsgaard, J. C., and Knudsen, J.: Operational validation and intercomparison of different types of hydrological models, Water Resources Res., 32, 2189-2202, https://doi.org/10.1029/96WR00896, 1996.

Rowland, C. S., Morton, R. D., Carrasco, L., McShane, G., O'Neil, A. W., and Wood, C. M.: Land Cover Map 2015 (1km percentage target class, GB). NERC Environmental Information Data Centre, [data set], available at: https://doi.org/10.5285/505d1e0c-ab60-4a60-b448-68c5bbae403e (last access: 8 April 2021), 2017.

Sankarasubramanian, A., Wang, D., Archfield, S., Reitz, M., Vogel, R. M., Mazrooei, A., and Mukhopadhyay, S.: Beyond the long-term water balance: evolving Budyko's supply–demand framework for the Anthropocene towards a global synthesis of land-surface fluxes under natural and human-altered watersheds. Hydrol. Earth Syst. Sci., 24, 1975-1984, https://doi.org/10.5194/hess-24-1975-2020, 2020.

Sanz, D., Castano, S., Cassiraga, E., Sahuquillo, A., Gómez-Alday, J. J., Peña, S., and Calera, A.: Modeling aquifer-river interactions under the influence of groundwater abstraction in the Mancha Oriental System (SE Spain). Hydrogeol. J., 19, 475-487, https://doi.org/10.1007/s10040-010-0694-x, 2011.

Schmidt, L., Hesse, F., Attinger, S., and Kumar, R.: Challenges in applying machine learning models for hydrological inference: a case study of flooding events across Germany, Water Resources Res., 56, e2019WR025924, https://doi.org/10.1029/2019WR025924, 2020.

Shepley, M. G., Streetley, M., Voyce, K., and Bamford, F.: Management of stream compensation for a large conjunctive use scheme, Shropshire, UK, Water and Environment Journal, 23, 263-271, https://doi.org/10.1111/j.1747-6593.2008.00158.x, 2009.

Simmons, D. L., and Reynolds, R. J.: Effects of urbanization on base-flow of selected south-shore streams, Long Island, New York, Water Resources Bulletin, 18, 797–805, https://doi.org/10.1111/j.1752-1688.1982.tb00075.x, 1982.

Singh, S., Pahlowb, M., Bookera, D. J., Shankara, U., and Chamorroc, 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.

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

Thomas, B. F., Vogel, R. M., Kroll, C. N., and Famiglietti, J. S.: Estimation of the base flow recession constant under human interference, Water Resources Res., 49, 7366-7379, https://doi.org/10.1002/wrcr.20532, 2013

UK Government.: Estimated licensed and actual abstractions from all surface and groundwater sources by purpose: 2000 to 2017, [data set], available at: https://www.gov.uk/government/statistical-data-sets/env15-water-abstraction-tables last downloaded Oct 2020 (last access 9 April 2021), 2020

Van Loon, A. F., and Laaha, G.: Hydrological drought severity explained by climate and catchment characteristics, J. Hydrol., 526, 3-14, https://doi.org/10.1016/j.jhydrol.2014.10.059, 2015.

Van Loon, A. F., Gleeson, T., Clark, J., Van Dijk, A. I. J. M., Stahl, K., Hannaford, J., Di Baldassarre, G., Teuling, A. J., Tallaksen, L. M., Uijlenhoet, R., Hannah, D. M., Sheffield, J., Svoboda, M., Verbeiren, B., Wagener, T., Rangecroft, S., Wanders, N., and Van Lanen, H. A. J.: Drought in the Anthropocene, Nature Geoscience, 9, 89-91, https://doi.org/10.1038/ngeo2646, 2016.

Vivoni, E. R., Entekhabi, D., Bras, R. L., and Ivanov, V.Y.: Controls on runoff generation and scale-dependence in a distributed hydrologic model, Hydrol. and Earth Syst. Sci., 11, 1683-1701, https://doi.org/10.5194/hess-11-1683-2007, 2007.

Wadoux, A. M. J-C., Samuel-Rosa, A., Poggio, L., Mulder, V. L.: A note on knowledge discovery and machine learning in digital soil mapping, European J. Soil Sci., 71, 133-136, https://doi.org/10.1111/ejss.12909, 2020.

Wang, D., and Cai, X.: Detecting human interferences to low flows through base flow recession analysis, Water Resources Research, 45, W07426, https://doi.org/10.1029/2009WR007819, 2009.

Wang, R., Kalin, L., Kuang, W., and Tian, H.: Individual and combined effects of land use/cover and climate change on Wolf Bay watershed streamflow in southern Alabama, Hydrol. Process. 28, 5530-5546, https://doi.org/10.1002/hyp.10057, 2014.

Weber, K. A., and Perry, R. G.: Groundwater abstraction impacts on spring flow and base flow in the Hillsborough River Basin, Florida, USA, Hyrdogeol. J., 14, 1253-1264 , https://doi.org/10.1007/s10040-006-0040-5, 2006.

Wendt, D. E., Van Loon, A. F., Bloomfield, J. P., and Hannah, D. M.: Asymmetric impact of groundwater use on groundwater droughts, Hydrol. Earth Syst. Sci., 24, 4853-4868, https://doi.org/10.5194/hess-24-4853-2020, 2020.

Wendt, D. E., Bloomfield, J. P., Van Loon, A. F., Garcia, M., Heudofrer, B., Larson, J., and Hannah, D. M.: Demonstrating the impact of integrated drought policies on hydrological droughts, Nat. Haz. Earth Syst. Sci., https://doi.org/10.5194/nhess-2021-129 , 2021

Wittenberg, H.: 2003. Effects of season and man-made changes on baseflow and flow recession: case studies, Hydrol. Proc., 17, 2113-2123, https://doi.org/10.1002/hyp.1324, 2003.

World Meteorological Organization.: Manual on low-flow estimation and prediction, WMO, Geneva. http://www.wmo.int/pages/prog/hwrp/publications/low-flow_estimation_prediction/WMO%201029%20en.pdf (last access 9 April 2021), 2008.

Zhang, Y-K., and Schilling, K. E.: Increasing streamflow and baseflow in Mississippi River since the 1940 s: Effect of land use change, J. Hydrol., 324, 412-422, https://doi.org/10.1016/j.jhydrol.2005.09.033, 2006.

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.

12. Appendix

Table A1. Description of the CAMESL-GB covariates used in the modelling and analysis.

Covariate class	CAMELS-GB covariate	Details of CAMELS-GB covariate	Context
Catchment physiography	area	Catchment area (km^2) based on date from UKCEH's Integrated Hydrological DTM (Morris and Flavin, 1990).	Catchment area is commonly identified as an important factor in explaining variability in low flows (Price et al., 2011). However, it less important with respect to mean residence and transit times where topographic relief appears to be more important (McGlynn et al, 2003; Asano and Uchida, 2012; Munoz-Villiers et al., 2016).
	dpsbar	Catchment mean drainage slope path (m km^{-1}).	Mean drainage path slope (Bayliss, 1999) is an index of catchment steepness and is estimated as the mean of all inter-nodal slopes from UKCEH's Integrated Hydrological DTM for a given catchment (Morris and Flavin, 1990).
Climate indices	aridity	Aridity (-). Aridity in CAMELS-GB, as with the other CAMELS data sets, is calculated as the ratio of mean daily potential evapotranspiration to mean daily precipitation (Addor et al., 2017, Coxon et al., 2020b). In the present study	The primary input to the catchment water balance and hence to baseflow generation is precipitation minus evapotranspiration (Price 2011, Fig.1).

https://doi.org/10.5194/hess-2021-259
Preprint. Discussion started: 27 May 2021

Hydrology and Earth System Sciences
Open Access
EGU
Discussions

		it has been reformulated as usually estimated (Joint Research Centre, 2019).	
	frac_snow	Fraction of precipitation falling as snow (for days colder than 0°C) was estimated by Coxon et al., (2020b).	Barnhart et al (2016) demonstrated a strong correlation between snowmelt rate and baseflow efficiency for catchments from western USA.
Hydrogeology classes	inter_high_perc	Percentage of catchment designated as being underlain by rock with intergranular flow & high productivity (%) (Hydrogeological attributes for each catchment were derived from the UK bedrock hydrogeological maps, British Geological Survey, 2019).	As Price (2011) notes, catchment geology is a primary control on baseflow-generating process. Three of the nine CAMELS-GB hydrogeological attributes have been selected as covariates, these include the two high groundwater productivity attributes and the attribute that denotes essentially no groundwater. Bloomfield et al., (2009) had previously explained 97% of the variance in BFI for 44 catchments in the Thames Basin, UK, using a model that regressed four hydrogeological classes including two high productivity and two low productivity classes on BFI.
	fract_high_perc	Percentage of catchment designated as being underlain by rock with flow through fractures & high productivity (%).	
	no_gw_perc	Percentage of catchment designated as being underlain by rocks with essentially no groundwater (%).	
Land cover	dwood_perc	Percentage of catchment designated as deciduous woodland coverage (%) (Attributes for each catchment were derived from the UK Land Cover Map 2015 produced by UKCEH, Rowland et al., 2017).	Processes associated with the transformation of the hydrological inputs, in forests and shrubby vegetation, such as interception, throughflow and stem flow, at the ground surface, such as ponding and infiltration, and in the soil, such as deep drainage and recharge (Price, 2011) depend on the nature of land use and land cover.
	ewood_perc	Percentage of catchment designated as evergreen woodland coverage (%).	
	grass_perc	Percentage of catchment designated as grass and pasture coverage (%).	
	shrub_perc	Percentage of catchment designated as medium scale vegetation (shrubs) coverage (%).	
	crop_perc	Percentage of catchment designated as crops coverage (%).	

	urban_perc	Percentage of catchment designated as suburban and urban coverage (%).	
	interwater_perc	Percentage of catchment designated as inland water coverage (%).	
	bares_perc	Percentage of catchment designated as bare soil and rocks coverage (%).	
Soil	clay_perc	Percentage clay content of soil (%). Soil attributes for each catchment were based on the European Soil Database Derived Data product (Hiederer, 2013).	Using data from over 600 catchments in the CAMELS-US dataset, Addor et al., (2018) used ML to compare the influence of catchment attributes on a variety of hydraulic signatures including BFI_LH. Soil clay fraction was the most negatively correlated attribute with BFI_LH (Addor et al., 2012, Fig. 4).
Water resource management	surfacewater_abs	Mean surface water abstraction (mm day^{-1}). Mean surface water and groundwater abstraction and discharge data were estimated by Coxon et al., (2020) based on monthly actual abstractions and returns for the period January 1999 – December 2014.	Wittenburg (2003), Wang and Cai (2009), Webber and Perry (2006) and Tomas et al. (2013) have all previously identified changes in features of baseflow in catchments subject to groundwater abstraction or due to returns flows.
	groundwater_abs	Mean groundwater abstraction (mm day^{-1}).	
	discharges	Mean discharges (mm day^{-1}). Discharge data consists of daily discharges into water courses from water companies and other discharge permit holders reported to the Environment Agency from 1 January 2005 to 31 December 2015.	
	num_reservoirs	Number of reservoirs in the Catchment (-). Reservoir attributes were taken from an open source UK reservoir inventory (Durant and Counsell, 2018).	
	reservoir_cap	Total storage capacity of reservoirs in the catchment (ML).	

https://doi.org/10.5194/hess-2021-259
Preprint. Discussion started: 27 May 2021

Hydrology and Earth System Sciences
Open Access
EGU
Discussions

Table A2. Coefficients of the four LR models, and associated spatial structural parameters and summary statistics for the models

	Model 1 Set A BFI_LH	Model 2 Set B BFI_CEH	Model 3 Set A BFI_LH	Model 4 Set B BFI_CEH
intercept	1.3652	1.4068	1.1372	1.2137
dpsbar	0.0029**	0.0056***	0.0034***	0.0063***
aridity	-0.2182***	-0.3220***	0.238***	-0.4002
inter_high_perc				(-0.0031***)
frac_high_perc	0.0107***	0.0194***	0.0105***	0.0031***
no_gw_perc	-0.0028**	-0.0035*	-0.0018*	-0.0021***
crop_perc	0.0096***	0.0157***	0.0089***	0.0149***
urban_perc	(0.0027***)	(0.0032***)		
inwater_perc	(0.0850***)			
clay_perc	-0.0412***	-0.0538***	-0.0350***	-0.0476***
surfacewater_abs			0.3278***	0.5239***
groundwater_abs			1.3861***	1.8737***
discharges			0.7099***	0.7285***
Spatial structure parameters				
Range	0.504	0.446	0.473	0.426
Nugget	0.408	0.383	0.496	0.387
Summary of models				
MSPE	0.117	0.360	0.138	0.289
Residual std.	0.435	0.642	0.388	0.581
R^2 (iid model)	0.627	0.669	0.703	0.732

705

https://doi.org/10.5194/hess-2021-259
Preprint. Discussion started: 27 May 2021

Hydrology and Earth System Sciences
Open Access
Discussions

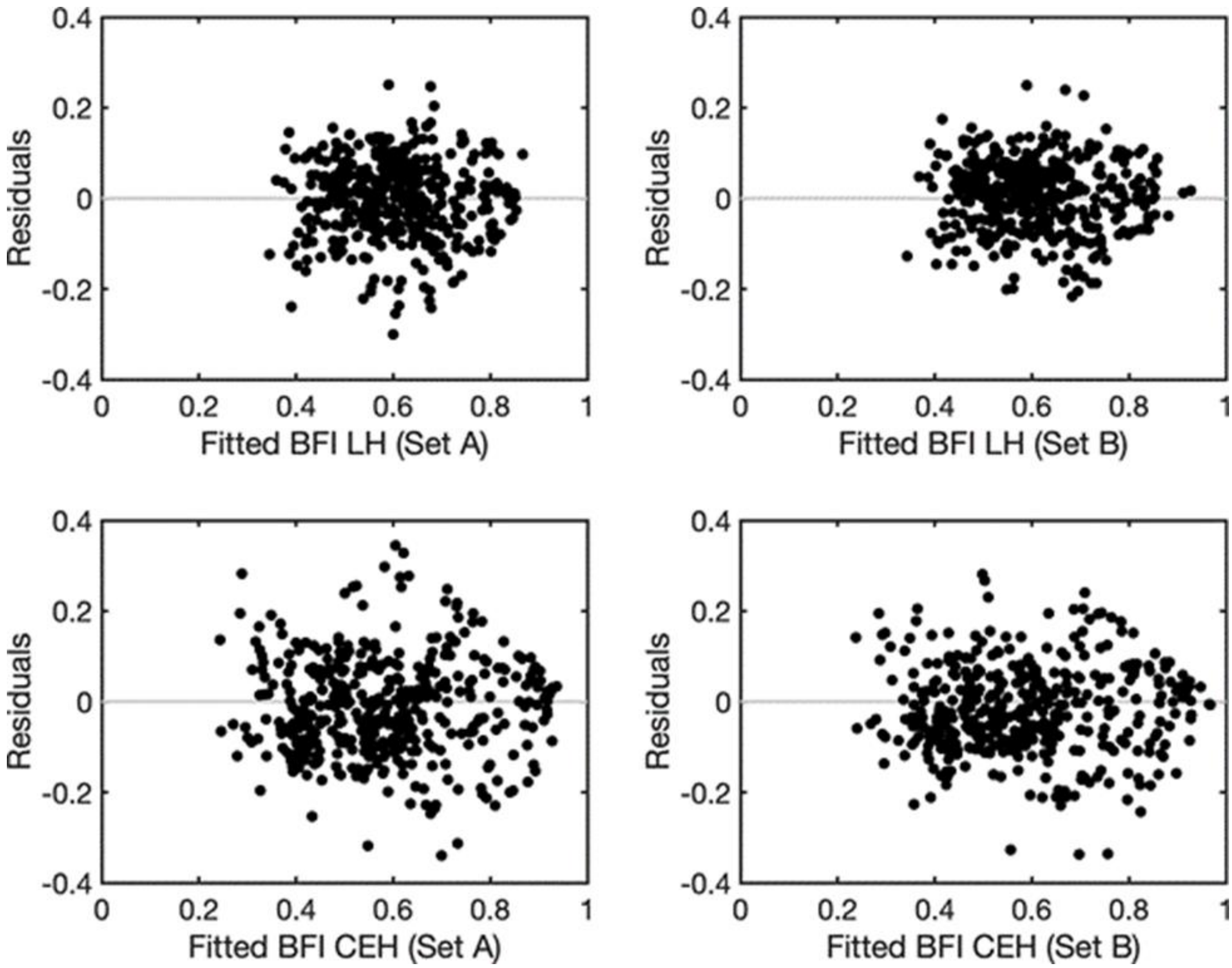


Figure A1. Distribution of residuals for LR models (Models 1 to 4)

710

https://doi.org/10.5194/hess-2021-259
Preprint. Discussion started: 27 May 2021

Hydrology and Earth System Sciences
Open Access
Discussions

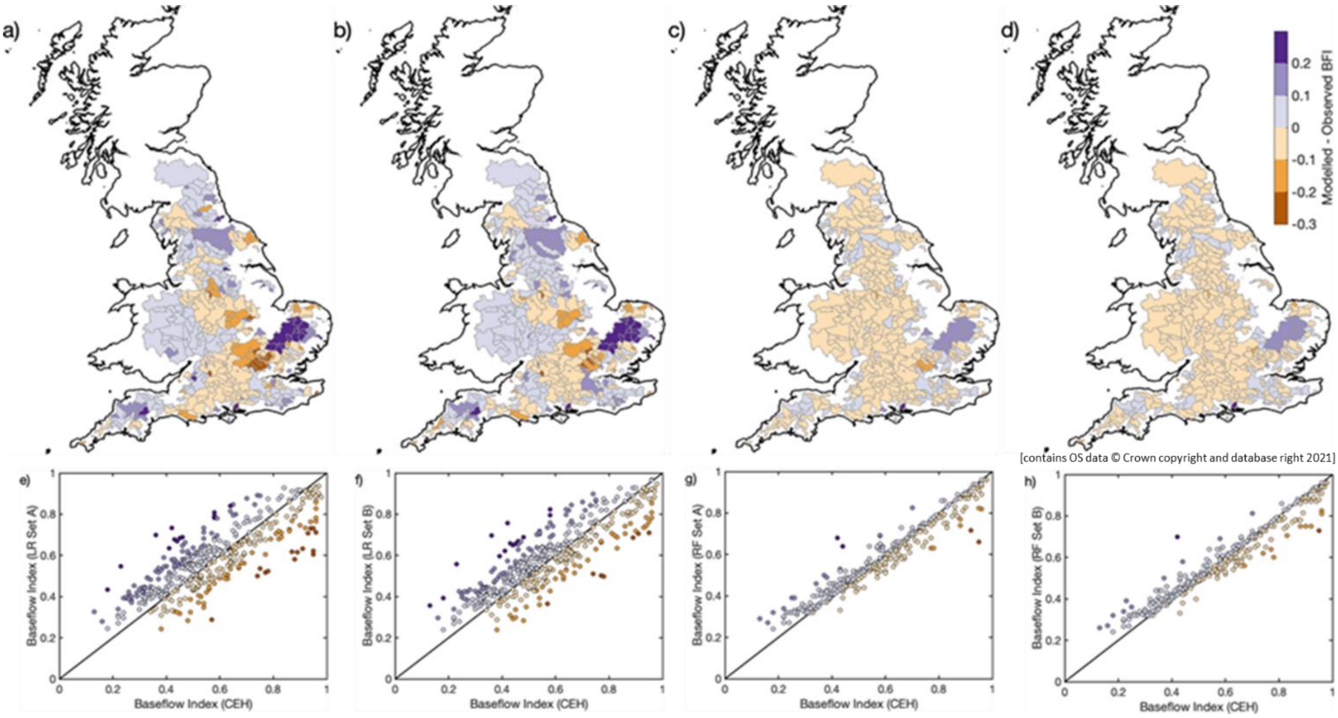


Figure A2. Maps of difference between modelled and observed BFI_CEH (a to d) and corresponding scatter plots of BFI_CEH against modelled BFI for covariate Sets A and B for LR and RF models (e to h) [Note this is the same as Fig. 6, but for BFI_CEH].

715

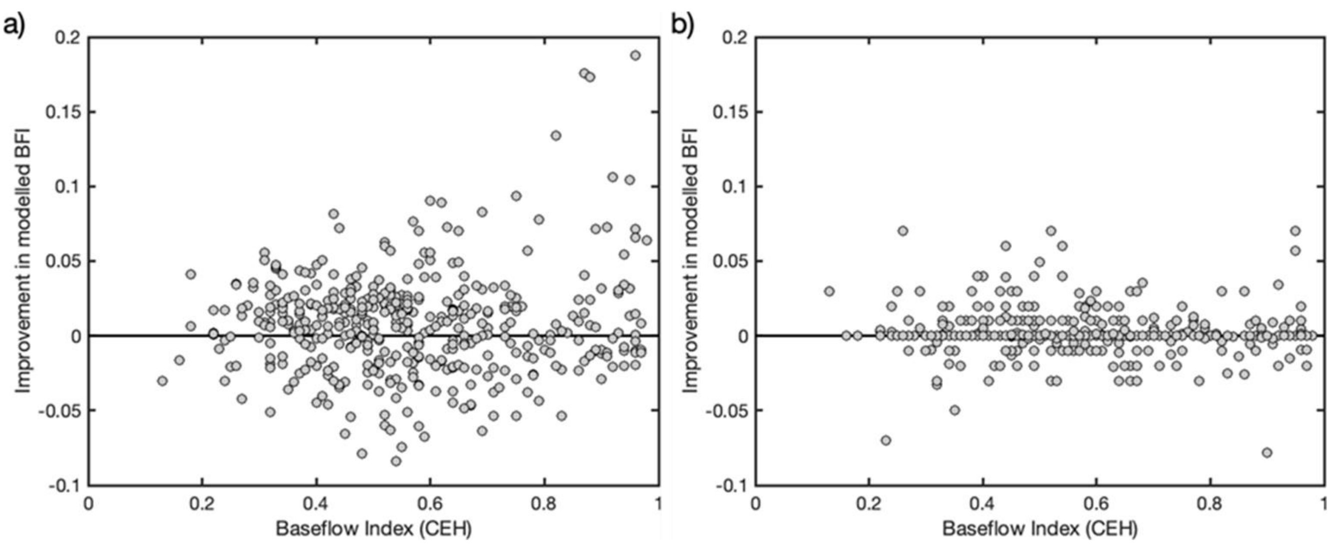


Figure A3. Scatter plots of improvement in modelled BFI as a function of observed BFI_CEH for a) LR and b) RF models. [Note this is the same as Fig. 7, but for BFI_CEH].

https://doi.org/10.5194/hess-2021-259
Preprint. Discussion started: 27 May 2021

720

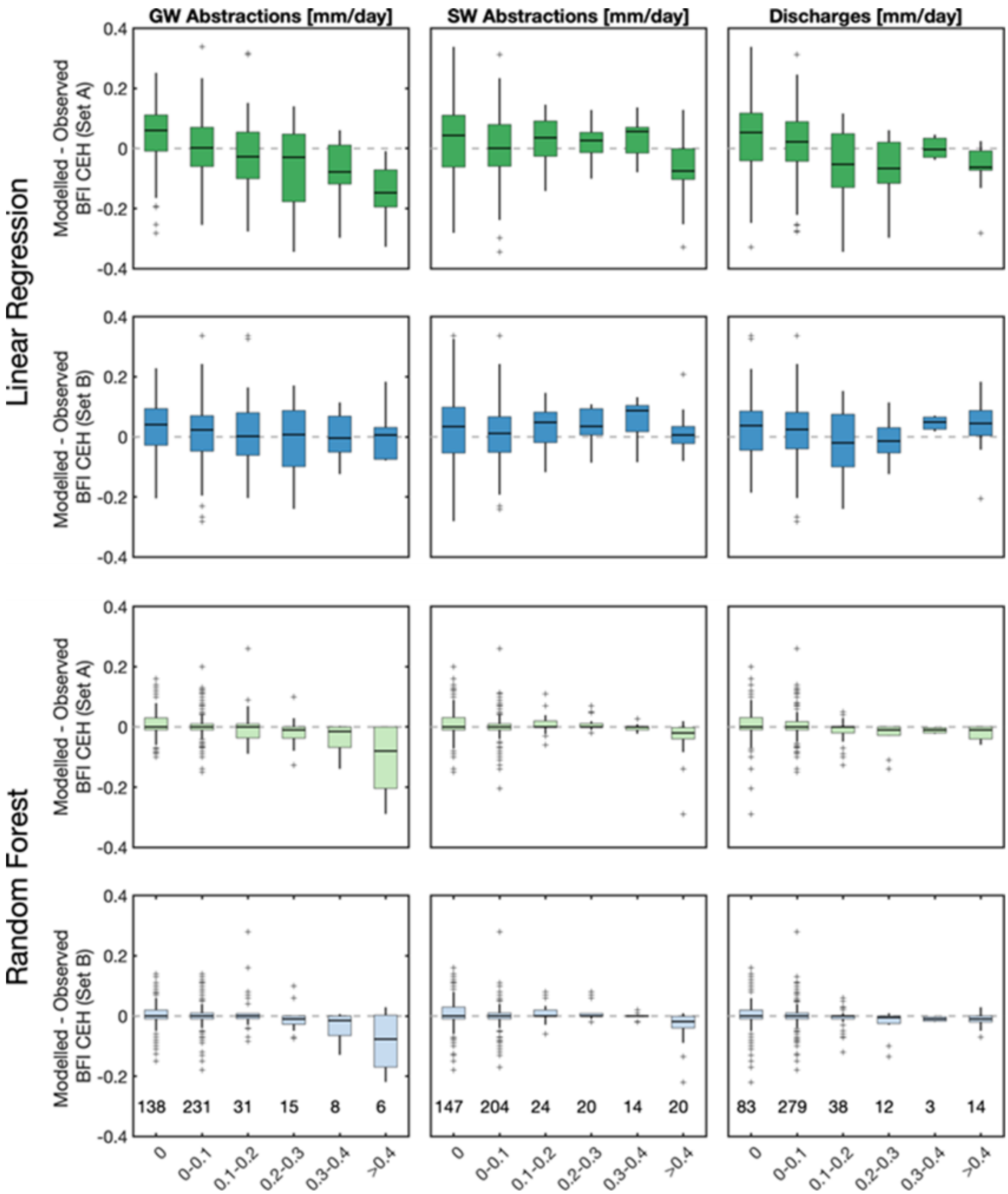


Figure A4. Comparison of observed and modelled BFI_CEH for covariate Sets A and B, for LR and RF models and as a function of different human management categories. [Note this is the same as Fig. 8, but for BFI_CEH].