Griffin, Adam; Young, Andy; Stewart, Lisa. 2019. **Revising the BFIHOST catchment descriptor to improve UK flood frequency estimates**.

Revising the BFIHOST catchment descriptor to improve UK flood frequency estimates

Adam Griffin*[1], Andy Young[2], Lisa Stewart[1]

[1]Centre for Ecology & Hydrology, Benson Lane, Crowmarsh Gifford, Oxfordshire. UK.

[2]Wallingford HydroSolutions, Benson Lane, Crowmarsh Gifford, Oxfordshire. UK.

*Corresponding author: adagri@ceh.ac.uk

Short Title: Revising the BFIHOST catchment descriptor

ABSTRACT

The estimate of base flow index (BFI) based on the Hydrology of Soil Types (HOST) classification, BFIHOST, provides a measure of catchment responsiveness. BFIHOST is used with other variables to estimate the median annual maximum flood (QMED) in the UK standard Flood Estimation Handbook (FEH) statistical method and is also an explanatory variable in ReFH2, the FEH design hydrograph package.

The current estimates of BFIHOST are derived from a restricted linear model, and a number of issues in the catchment dataset have been identified since the original work in 1995. BFI calculated through base flow separation tends to be underestimated in clay-dominated catchments, and the calculation technique performs poorly in ephemeral catchments or those with missing data. The pragmatic bounding of BFI coefficients for permeable soils overlying aquifer outcrops is also problematic for small catchments.

This paper investigates alternative regression methods to improve base flow estimates using the HOST class data for 991 stations (compared to 575 in the original); beta regression was found to give the best performance. Combining multiple rare classes into single classes is also shown to improve

performance. The new version of BFIHOST was applied to the QMED equation, showing improved performance.

KEYWORDS

base flow, catchment hydrology, regression methods, statistical hydrology

INTRODUCTION

The Base Flow Index (BFI) is a widely applied, broad-scale measure used in the United Kingdom for measuring variation in the low flow regimes of gauged catchments. The index was originally proposed in the 1980s by the Institute of Hydrology (1980) and is a simple separation algorithm that disaggregates the daily mean flow record for a catchment into low and high frequency components. This is achieved by partitioning the daily mean flows into five-day blocks and identifying the minimum within each five-day series. The line connecting these minima represents the storage-driven component of the hydrograph. BFI is the ratio of the volume of this base flow to the total flow. This is shown for two catchments in Figure 1: a permeable, groundwater-dominated chalk catchment in the south of England and an impermeable catchment from the west of Scotland. The BFI is constrained to lie in the interval [0,1] and ranges from around 0.11 for perennial streams draining impermeable catchments to 0.98 for streams draining the most permeable catchments.

The HOST classification and BFIHOST

It is difficult to overstate the importance of soils and underlying geology in influencing the movement of water through the landscape at both the site and catchment scales. The Hydrology Of Soil Types (HOST) classification of the United Kingdom was developed in the mid-1990s to provide a hydrologically relevant classification of soils and parent geology to aid hydrological studies and analyses within the UK (Boorman *et al.*, 1995). The original objective was to replace the five-class Winter Rainfall Acceptance Potential (WRAP) mapping of UK soils, developed in the 1970s following the Soil Survey Field Handbook classification system (Hodgson, 1974).

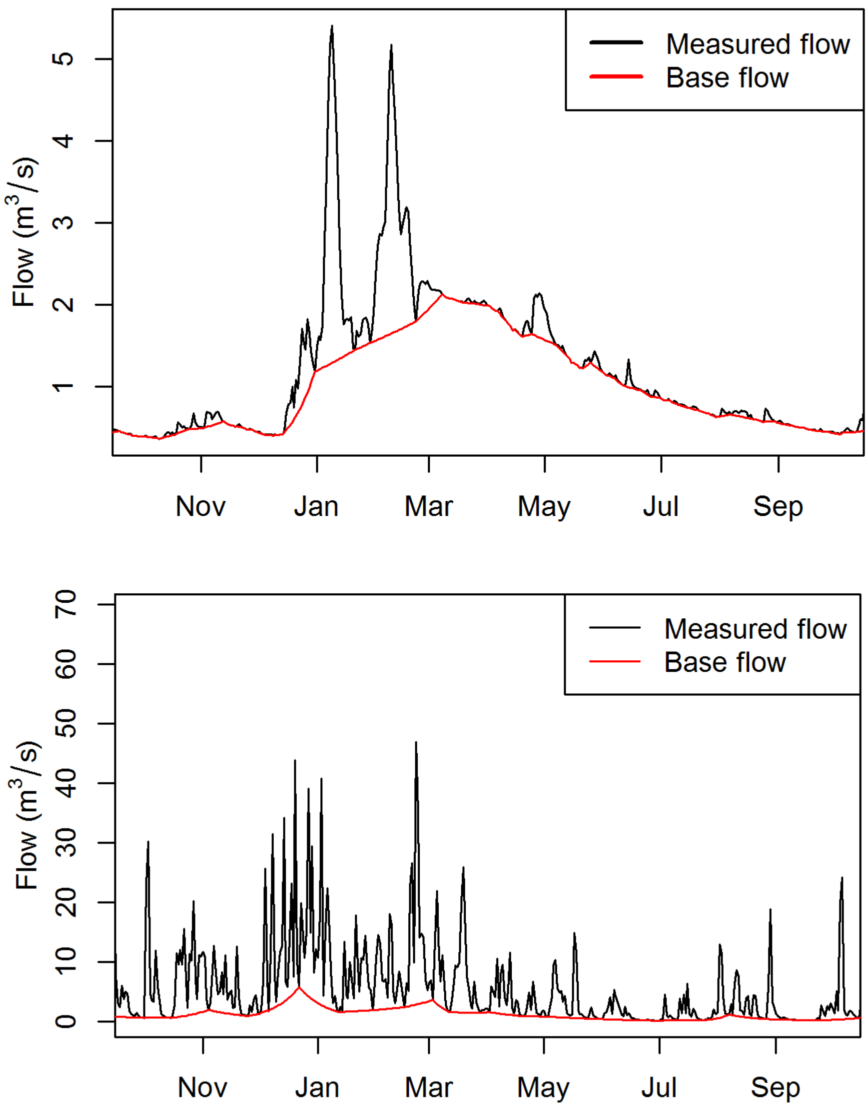

Figure 1. The base flow separation algorithm of Gustard *et al.* (1992) applied to two catchments for the water year 2013-2014. Top: Pang at Pangbourne (groundwater dominated). Bottom: Falloch at Glen Falloch (impermeable upland catchment)

The HOST classification is based on a set of conceptual models of the hydrological processes taking place within the soils and, where relevant, the underlying geologies. The classification is based on soil series, and within these settings soils were differentiated through soil properties and wetness regimes, as indicated by the presence of gleying. This differentiation gave rise to 11 models which were subdivided into 29 classes based on either the geology of the substrate or other properties. Assignment of soil series to these classes was undertaken using a method based on K-means cluster analysis. Although HOST is based on soil series, for refinement of the classification it was applied as a national coverage using the 1:250,000 soil associations (of series) through reconnaissance mapping undertaken by the soil survey organisations across the UK at that time. The conceptual classification of soils was refined by assessing its ability to explain the variation in hydrological characteristics as expressed by BFI estimated from data for a sample of 575 gauged catchments across the UK. These catchments were selected on the basis of the quality of the hydrometry of the gauged record and the naturalness of the flow regime using the classification scheme of Gustard *et al.* (1992). A linear regression model of the relationship between BFI and the fractional extents of HOST classes across each catchment in this sample set was then developed and the resulting BFIHOST catchment descriptor (here referred to as $BFIHOST_{1995}$) was developed by applying this model across the UK to a 1km grid of the fractional extents of each of the 29 HOST classes, and taking an area-weighted average over each catchment.

Table 1 presents the HOST classification scheme, together with the model coefficients for BFIHOST and the percentage of the United Kingdom land surface that each class represents. In general, catchment permeability grades from left to right and top to bottom and on the whole this pattern is reflected in the BFIHOST model coefficients, with some notable exceptions. For example, the drained and eroded peats of classes 11 and 28 had higher BFIHOST than surrounding classes, not following this trend. To ensure model coefficients lay in an acceptable range, the regression was also bounded to prevent the coefficients taking values greater than 1 or less than 0.17. In the final model, additional constraints were placed on the sandstone and sand-dominated HOST classes 3 and 5 such that the coefficient value for both classes was constrained; class 3 restricted to having a value above 0.9, class 5 capped to having a value below 0.9.

Notwithstanding these criticisms, BFIHOST has been a well-used catchment descriptor. It is a key descriptor in the Flood Estimation Handbook (FEH) catchment descriptor equation for estimating the index variable QMED, the median of the annual maximum flood series, in ungauged catchments

(Kjeldsen *et al.*, 2008) and informs three out of the four model parameters in the design package for the ReFH2 event-based rainfall-runoff model (Wallingford Hydrosolutions, 2016a). It has also proved to be a key catchment descriptor for explaining the variation in parameter values for generalised rainfall-runoff models for estimating both daily mean flows (Young, 2006; Young *et al.*, 2006) and extreme flood events (Calver *et al.*, 1999).

The BFIHOST catchment descriptor has also found wide application in environmental management, for example, in the development of environmental standards for abstraction as part of the implementation of the Water Framework Directive in the UK (SNIFFER, 2006). Another example is its use in understanding the role of hydrology in explaining the interactions between dissolved organic carbon and nitrogen in rivers (Heppell *et al.*, 2017).

Operational use of research outputs often identifies limitations, and the BFIHOST model is no exception. In impermeable catchments (low BFI) such as those found in upland peat-dominated catchments, $BFIHOST_{1995}$ tends to overestimate at low gauged BFI values (Figure 2). This can be seen in Table 1 in the high coefficient for class 11 compared to the overall trend. On the other hand, HOST classes 23 and 25, which represent clay-based soils, show much lower values of BFIHOST than the overall trend would suggest. In such impermeable catchments, those with a BFI less than 0.2 are often associated with ephemerality and zero flow in summer months.

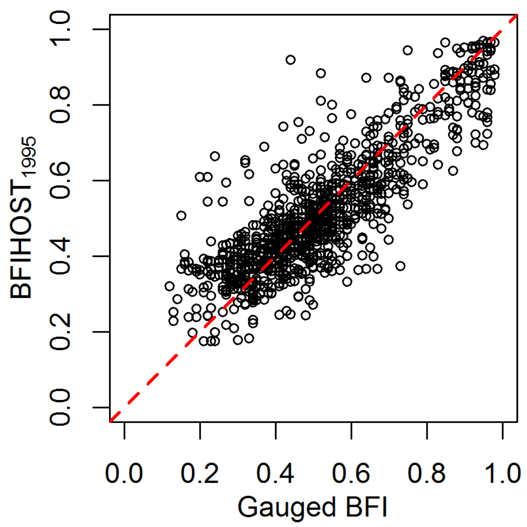

Figure 2: Comparison of gauged BFI estimates with $BFIHOST_{1995}$ (Boorman *et al.*, 1995) using present dataset of 991 catchments.

Substrate Hydrogeology	MINERAL SOILS						PEAT SOILS	
	Groundwater or aquifer	No impermeable or gleyed layer within 100cm	Impermeable layer within 100cm or gleyed layer at 40-100cm		Gleyed layer within 40cm		All	
Weakly consolidated, microporous, by pass flow uncommon (Chalk)	Normally present and at >2m	[1] 1.000 *(4.31)*	[13] 1.000 *(0.87)*		[14] 0.380 *(0.66)*		[15] 0.380 *(9.93)*	
Weakly consolidated, microporous, by pass flow uncommon (Limestone)		[2] 1.000 *(2.12)*						
Weakly consolidated, macroporous, by pass flow uncommon		[3] 0.900 *(1.58)*						
Strongly consolidated, non or slightly porous. By-pass flow common.		[4] 0.791 *(3.33)*						
Unconsolidated, macroporous, by-pass flow very uncommon.		[5] 0.900 *(5.07)*						
Unconsolidated, microporous, by-pass flow very common.		[6] 0.645 *(2.61)*						
Unconsolidated, macroporous, by-pass flow very uncommon.	Normally present and at ≤ 2m	[7] 0.792 *(1.01)*			IAC' <12.5	IAC' ≥12.5	Drained	Undrained
Unconsolidated, microporous, by-pass flow common.		[8] 0.560 *(1.62)*			[9] 0.734 *(3.68)*	[10] 0.520 *(2.21)*	[11] 0.927 *(0.55)*	[12] 0.170 *(2.94)*
Slowly permeable	No significant groundwater or aquifer	[16] 0.778 *(0.43)*	IAC' >7.5	IAC' ≤7.5	[24] 0.312 *(13.85)*		[26] 0.244 *(2.49)*	
			[18] 0.518 *(5.40)*	[21] 0.340 *(4.02)*				
Impermeable (hard)		[17] 0.609 *(9.28)*	[19] 0.469 *(2.16)*	[22] 0.315 *(1.10)*			[27] 0.259 *(0.83)*	
Impermeable (soft)			[20] 0.524 *(0.69)*	[23] 0.218 *(1.31)*	[25] 0.170 *(3.64)*			
Eroded Peat							[28] 0.581 *(0.58)*	
Raw Peat							[29] 0.226 *(5.73)*	

Table 1: HOST classification (small upper numbers) with BFIHOST coefficients listed (percentage land cover in brackets). Based on Boorman *et al.*, 1995. IAC' = integrated air capacity.

Another problem associated with some HOST classes is linked to their relative rarity within the dataset. For example, HOST class 11 is only observed in 0.55% of the original 1995 dataset, and was also seen to be frequently co-located with permeable soils and so is likely poorly modelled by $BFIHOST_{1995}$ due to having some possible correlation with other soil types. It is hoped that using a larger more varied dataset may alleviate this.

Finally, the need to artificially change model coefficients (in the cases of classes 1,2,3,5 and 13) suggests that the model was not chosen carefully enough in the outset. An important part of model development is the choice of model form, be that linear regression, a more generalised additive model or even a more physically-based mathematical hydrological model. The parameter estimates in permeable soil classes represent a theoretical upper limit where the parameter values take a bounded value of 1. In practice, BFI estimates from gauged records may approach this theoretical limit but even the flow regimes of the most permeable catchments exhibit some direct response to rainfall and thus a value of 1 is not attained.

In this paper, the aim is to tackle these issues and produce a more robust model of base flow index. To achieve this, a more physically appropriate model will be investigated to avoid the need for capping coefficients. Secondly, methods will be investigated to improve the base flow estimates of the rarer HOST classes, by combining single HOST classes into groups. Thirdly, models focusing on the low-BFI catchments will be explored to improve the estimates in the peat and clay catchments. This will be then demonstrated through an application to estimation of QMED through the FEH catchment descriptor equation.

DATA

Station selection

The catchment dataset used in the 1995 development of the HOST classification was constrained to using catchments with flow records that were believed to be of good hydrometric quality and relatively free from artificial influence (Gustard *et al.*, 1992), and applied to all gauged flow records held by the National River Flow Archive (NRFA). However, information on water use and return was fragmented in England and Wales, and in Scotland and Northern Ireland there was no requirement to regulate abstraction until 2006. A necessarily conservative approach was taken to the selection of study catchments. Due to differences in the mapping and assignment of HOST classes within Northern Ireland compared to England, Scotland and Wales, catchments in Northern Ireland were not included. Instead, we feel that future work should be undertaken to properly develop a Northern Ireland-specific BFIHOST model.

A focus of the current study was to develop a much larger sample set of appropriate gauged records to inform the research. The NRFA holds flow records for UK gauged catchments judged by the measuring authorities as of a suitable hydrometric quality for public release. On the basis that BFI is relatively insensitive to hydrometric quality (pp. 24, Gustard et al., 1992), all stations were considered as being of a suitable hydrometric quality. The records for 1223 catchments with a minimum of 10 years of daily mean flow record were reviewed and 991 catchments selected according to the following criteria:

- the absence of significant upstream impounding reservoirs;
- less than 2% missing days in the record;
- generally perennial stream flow, determined on the basis that the flow that is equalled or exceeded for 95% of the time, Q95, is greater than zero.

Lakes were not chosen as a criteria for exclusion due to the very high percentage of catchments with some form of surface water storage. Instead, a HOST class (class 30) describing proportion of lake coverage is included; this is explained in more detail below. Previous research in Scotland (Gustard *et al.*, 1987) has shown that the estimation of BFI from gauged records is relatively insensitive to sample error and that, provided that extremely dry years are avoided, the error in BFI calculated from one year of record is typically approximately 5%. However, BFI values are sensitive to missing data within the period of calculation, and particularly within periods of low flow; a limit on missing data of 2%

was proposed by Gustard *et al.* (1992), and this motivated the use of this same limit in the present study.

Abstractions and discharges in the UK are, in the main, regulated to maintain the magnitude of influence to below a fraction of the lower flows observed within a catchment. When compared with the full range of variation in the daily mean flow records for a catchment (typically 3 to 4 orders of magnitude) these influences are relatively invariant. Errors at very high flows associated with out-of-bank flow are not as relevant as these occur in a small fraction of the higher-frequency components of the flow record. In contrast, impounding reservoirs have a significant impact upon the whole flow regime.

It was noted that some of the 991 gauged flow records showed non-negligible artificial influence compared to Q95. In particular, 332 catchments were seen to have net influence estimated to be greater than 20% of Q95. Fortunately, filtering out ephemeral catchments also screens out catchments with a very high net abstractive component. Sensitivity to this abstraction was investigated. The basic linear model was fitted with and without those 332 catchments with noticeable abstraction, and 10-fold cross-validation was performed in both cases (James et al., 2013). It was seen that the basic model performed similarly with and without the extra catchments, both in terms of model residual and cross-validation prediction errors. Therefore the rest of this paper will make use of all 991 flow records, as it is felt that a dataset that covers a broader range of catchments will lead to a model which is more representative of base flow across the whole UK. This does come at the slight cost of increased observation uncertainty, but as mentioned above (Gustard et al., 1992) shows that BFI is robust to this sort of data issue.

Figure 3 shows the locations of the catchments within the UK. The mean record length of stations within the sample is 42.3 years. Catchment areas range from 0.9 km^2 up to 9940 km^2, with a mean area of 343 km^2.

METHODS

As mentioned above, $BFIHOST_{1995}$ was a linear regression model with bounds placed on the minimum (0.17) and maximum (1.0) values BFIHOST could take for any specific HOST class (Boorman *et al.*, 1995). In the first stage of the current study, the linear model was recalibrated using the new dataset and then directly compared with $BFIHOST_{1995}$. Simple linear regression models are easy to interpret, but have problems with extrapolation, application of them here possibly leads to estimated values of

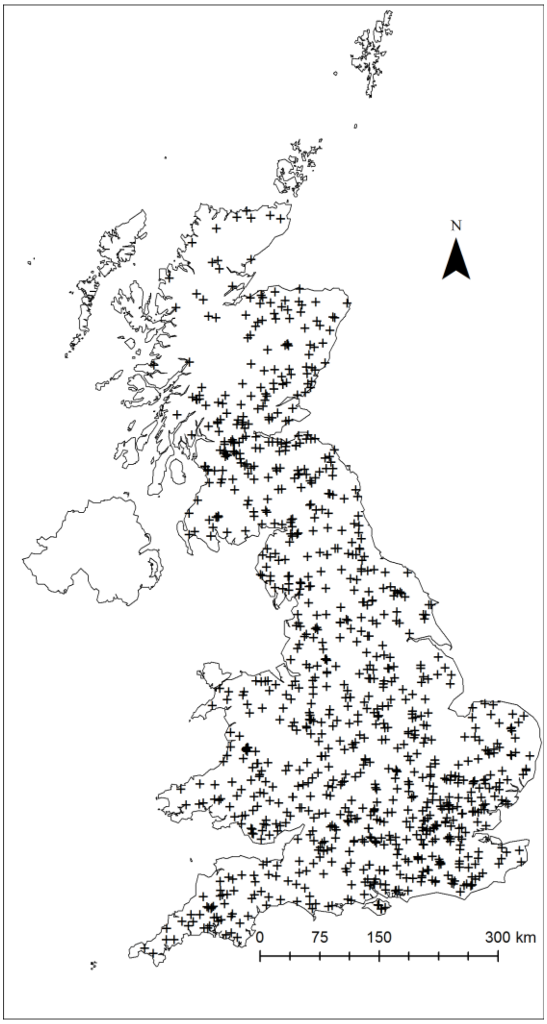

Figure 3: Locations of 991 catchments selected for new models of BFIHOST.

BFIHOST greater than 1 or less than 0. Catchment estimates are given by $BFIHOST = \sum_{i=1}^{29} \alpha_i h_i$ where α_i is the BFI coefficient for HOST class *i*, and h_i is the proportion of that catchment which is classified as belonging to HOST class *i*. This assumes that the error (BFI_{gauged} – $BFIHOST$) is normally distributed which may suggest that the true value could lie outside the permitted range even if the fitted model estimate lies within the range.

To more directly compare to $BFIHOST_{1995}$, the recalibrated linear model was also adjusted to cap coefficients to lie in the range $0 < \alpha < 1$. Within-range capping was not considered in this work. Linear models were implemented using the core *stats* R package (R Core Team, 2016).

Beta regression

As an alternative to the simple linear regression, a beta regression model was investigated. Rather than assuming that the residual follows a normal distribution as in linear regression, beta regression assumes that each observation comes from a beta distribution with mean fitted using a logit link function (Ferrari and Cribari-Neto, 2004). Here $\mathrm{logit}(x) = \log(x/(1-x))$, a function that turns the unit interval (0, 1) into the whole range from $-\infty$ to ∞; inverting the transformation naturally forces the resulting estimates of BFIHOST to lie strictly between zero and one. More specifically, the gauged values of BFI at site *i* are assumed to be distributed according to a beta distribution with mean μ_i and precision ϕ common to all sites. Here, μ_i is given by $\mathrm{logit}(\mu_i) = \sum_i \beta_i h_i$, where β_i are fitted model coefficients and h_i are as defined previously. The beta distribution is completely defined between 0 and 1 and has the probability density function given by:

$$f(x) = \frac{\Gamma(\phi)}{\Gamma(\phi\mu)\Gamma\big(\phi(1-\mu)\big)} x^{\mu\phi-1}(1-x)^{(1-\mu)\phi-1}, \qquad 0 < x < 1$$

where $\Gamma(x)$ is the gamma function. BFIHOST is then estimated by using

$$BFIHOST_{BETA} = \mathrm{logit}^{-1}\left(\sum_i \beta_i h_i\right)$$

with β_i and h_i as before. The beta regression was implemented using the *betareg* R package (Cribari-Neto and Zeileis, 2010). In order to better compare the coefficients directly, a set of “linear equivalent” values will be reported, corresponding to the value b_i of $BFIHOST_{BETA}$ in a catchment consisting solely of a single HOST class *i*, which can be calculated as $b_i = \mathrm{logit}^{-1}(\beta_i)$.

One benefit of using a beta regression is that it is designed for estimating proportions, and so the estimate plus the associated uncertainty (the error term) still lies within the interval of appropriate BFI values.

Regrouping variables

The 1995 BFIHOST model was a classification tool to aid the development of HOST whereas the objective of this study is to refine the estimation of BFI to give reliable estimates across all catchment scales and with HOST class representations ranging from a single dominant class to large catchments draining soils corresponding to many different HOST classes.

A 30-class model including a HOST class representing surface water extents (HOST class 30) is considered; nominally one would expect the BFI of a waterbody to be 1, as none of the water would flow over the top of a waterbody, rather through it. The effect of surface water, in the form of lakes and reservoirs (for example), is already well established in flood frequency estimation using catchment descriptor equations, where the extent to which such water bodies attenuate flow is significant in the QMED equation (Kjeldsen *et al.*, 2008). This 30-class model was compared to the 29-class model under the original 1995 coefficients, and under the recalibrated linear model.

Typically, the BFI for catchments which are dominated by rarer HOST classes is over- or under-estimated in $BFIHOST_{1995}$, since there are many fewer catchments in which those HOST classes are observed. For example, HOST class 11 only makes up 0.55% of the land cover in England, Wales and Scotland, and makes up less than 11.2% of any one catchment in the present dataset.

To analyse this, a series of models were tested where subsets of the 30 classes were combined replacing, for example, HOST classes 16 and 17 with a single (16+17) class (so $h_{(16+17)} = h_{16} + h_{17}$). Various combinations of HOST class groupings were investigated, as outlined in Table 3. Each of these groupings was chosen to combine similar soil classes (in terms of base flow) together to address poor representation of one or more classes within the group. This was then followed by considering combinations of the above single groups to develop the final model.

Weighted models

The most successful models were tested using weighted regression models, where certain data points were given greater weighting in the regression method (either linear or beta). Catchments with a gauged BFI value of less than 0.4 were seen to perform particularly badly in the original 1995 model, and so a more focused approach to characterising the BFI in these catchments was felt to be important for use in prediction across the UK.

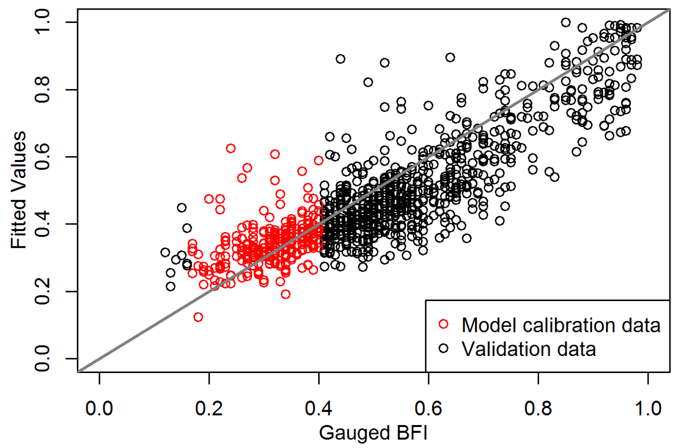

Figure 4: Plots showing accuracy of fit underweighting for linear 30-class model fitted with catchments with gauged BFI between 0.17-0.4 weighted six times as heavily as those outside the range.

To develop the weighted models, four methods were considered. The catchments with BFI less than 0.4 were weighted twice as much, three times as much, and six times as much as those with higher BFI. Also, a weighting inversely proportional to BFI was tried.

RESULTS AND DISCUSSION

The results in this section show model performance in terms of R^2 (a measure of variance of the data explained by the model), root mean squared error (RMSE, a measure of average accuracy of the model), fractional bias, and Akaike Information Criterion (AIC, Akaike, 1974) which is given by $AIC = 2k - 2\log(L)$, where k is the number of fitted parameters (the number of classes/groups selected), and L is the maximum value of the likelihood function of the fitted model given the dataset; note lower or more negative values show better models. One should note that the R^2 values are designed to assume normally distributed residuals about the model fitted values. This means that the values of R^2 associated with the beta regression may incorrectly claim under-performance of the model. Therefore, RMSE and AIC are more appropriate for comparing linear with beta regression models. Within type, R^2 can be more appropriately used as a comparison.

30-class model

First an investigation was performed to assess the original model, and the benefits of including HOST class 30 (surface water). Table 2 suggests that the old model much better explains the data than previously thought, with even further improvements in terms of R^2 and RMSE. The AIC of the original model cannot be determined due to lack of access to the original 1995 dataset; the RMSE is as reported in (Boorman *et al.*, 1995).

Model	**R^2**	**RMSE**	**AIC**	**BIAS**
1995 Model (as reported)	0.79	0.089	N/A	N/A
1995 Model (new dataset)	0.97*	0.101	N/A	+0.050
Recalibrated Linear 29-class	0.97	0.096	-1767	-0.052
Linear 30-class	0.97	0.095	-1797	-0.049
Capped 30-class	0.97	0.095	-1776	+0.037

Table 2: Comparison of various linear models. AIC and bias not documented for original 1995 model. R^2 for the 1995 model under the new data (starred) is only illustrative, not a true value.

However, one should note that this is not an ideal use of R^2 since the data used to calculate the model here was not the data used to calibrate the model in Boorman *et al.* (1995), and so the resultant value of R^2 should be treated with caution. The recalibrated 29-class model is a naïve linear regression, and therefore some of the coefficients extend beyond 0 and 1; HOST class 1 has a value of 1.02, class 11 has a value of 1.15, and class 13 has a value of 1.19. Including the HOST class 30 (surface water extent) does not improve this, with HOST class 1 taking a coefficient of 1.02, class 11 having a value of 1.17, class 12 having a value of -0.03, and class 27 taking a value of -0.13. This could be a significant problem for estimation in small catchments which predominantly consist of these HOST classes.

Reapplying the capping procedure of Boorman *et al.* (1995) on the new 29- and 30-class linear models only served to reduce the efficacy of the model (Table 2). Bias may have decreased as an artefact of restricting the model to the "correct" range. In this case, the model statistics were recomputed assuming that the error term had the same distribution in the capped and uncapped 30-class linear model.

Beta regression as an alternative

Table 3 shows the initial results of using the beta regression versus the linear model; the bias is slightly improved and the values of R^2 and RMSE are very similar. Recall that since linear regression coefficients and beta regression coefficients cannot be directly compared, Table 3 presents "linear equivalent" values. These correspond to the value b_i of $\mathrm{BFIHOST}_{\mathrm{BETA}}$ in a catchment consisting solely of a single HOST class *i*, which can be calculated as $b_i = \mathrm{logit}^{-1}(\beta_i)$, where β_i is the beta regression coefficient for HOST class *i*.

Table 5 shows that the beta regression model keeps all the "linear equivalent" coefficients between 0 and 1 without further modification, so that no catchment can have an estimated value of BFIHOST outside this range, which is desirable. Here the value of R^2 is computed as for a linear regression. In the *betareg* package, a "pseudo-R^2" specific to beta regression is also presented (Ferrari and Cribari-Neto, 2004). To more appropriately compare between linear and beta regression model, the linear regression formulation of R^2 is used.

Combined HOST classes

Table 3 shows the difference in including various combinations of the groupings outlined above. Here one notes that certain models perform slightly better (the (9+10) model is slightly better on all statistics), but upon examining the coefficients and the linear-model-equivalent values, it can be observed that they produce physically unrealistic values close to 0 or 1: HOST class 12 has a "linear-equivalent” coefficient of 0.074, and class 27 has a value of 0.068, much lower than reasonable. As a balance between physically reasonable and statistically “powerful” models must be reached, this led to the selection of a model using the grouped classes (7+8+9+10, 11+12+15, 16+18, 20+23, 26+27) with all other classes kept in isolation. In the rest of this paper, only the 30-class and (7+8+9+10, 11+12+15, 16+18, 20+23, 26+27) models will be considered; the latter will be referred to as the “22-class” model.

Model	**R^2**	**RMSE**	**AIC**	**BIAS**
Linear 30-class model	0.970	0.095	-1797	-0.049
Beta 30-class model	0.970	0.098	-1819	+0.059
(7+8)	0.970	0.095	-1821	+0.052
(9+10)	0.970	0.095	-1822	+0.052
(11+12+15)	0.969	0.096	-1800	+0.053
(13+14)	0.970	0.095	-1816	+0.052
(16+18)	0.970	0.095	-1812	+0.052
(20+23)	0.970	0.095	-1822	+0.052
(26+27)	0.970	0.095	-1815	+0.052
(28+29)	0.970	0.095	-1819	+0.052
(7+8+9+10)	0.970	0.095	-1825	+0.052
(7+8+9+10, 26+27)	0.970	0.095	-1820	+0.052
(7+8+9+10, 11+12+15, 26+27)	0.969	0.096	-1800	+0.054
(7+8+9+10, 11+12+15, 16+18, 20+23, 26+27)	0.968	0.096	-1794	+0.054

Table 3: Comparison of beta regression models under different groupings of HOST classes. Bracketed groupings replace the individual constituent classes, all other unlisted classes are kept individually.

Weighted models

Finally, the selected 22-class model and the 30-class model under linear and beta regression were investigated using the weighting scheme outlined in the Methods section. Small relative weightings were tried but showed little difference compared to the unweighted model (not shown), and very high relative weightings instead produced a poor fit for high-BFI stations (Figure 4). Finally, a weighting inversely proportional to BFI was also tried, but this also gave far too much weight to the low BFI catchments and led to a poorly fitting model (not shown). Hence the four candidate models were compared to similar models weighting catchments with gauged BFI between 0.17 and 0.4 three times as heavily as other catchments. Although R^2 and RMSE do not greatly change, continued improvement can be observed in Table 4 in terms of AIC. The final weighted 22-class beta regression model is denoted $BFIHOST_{2019}$

Model	R^2	RMSE	AIC	BIAS
Weighted Linear 30-class model	0.962	0.098	-1687	+0.039
Weighted Beta 30-class	0.968	0.0975	-2923	+0.040
Weighted Linear 22-class	0.960	0.100	-1657	-0.001
Weighted Beta 22-class	0.967	0.097	-1798	+0.006

Table 4: Comparison of various weighted regression models under linear and beta formulations, and under 22- and 30-class configurations.

To summarise the findings, Table 5 and Figure 5 show the coefficients from the linear regression models and the "linear-equivalent" values from the beta regression models. For the less frequently observed HOST classes, a more reasonable value of BFIHOST is obtained when combined with a more abundant HOST class; classes 11 and 12 are a clear example of this (with land cover of 0.55% and 2.94%, respectively). Figure 6 shows a residual plot as described in Chien (2011), which performs in a similar way to a standard residual plot as used for linear regression models. There is no obvious trend or pattern to the residuals as a function of $logit(BFIHOST_{2019})$, which suggests that the model type is appropriate for this data.

Example application to flood estimation

As a simple application of $BFIHOST_{2019}$, a standard use of BFIHOST will be investigated: the estimation of QMED. To do this, a subset of the above investigated dataset has been used: 605 stations which are rural and determined by the NRFA to be suitable for QMED estimation. Here, $BFIHOST_{1995}$ and the new $BFIHOST_{2019}$ were both used in the QMED catchment descriptor equation without recalibration:

$$QMED = 8.3062\, AREA^{0.8510} 0.1536^{\frac{1000}{SAAR}} FARL^{3.4451} 0.046^{(BFIHOST^2)}$$

HOST	$BFIHOST_{1995}$	Linear 30- class	Beta 30-class	Linear 22-class + weighted	Beta 22-class + weighted	HOST	$BFIHOST_{1995}$	Linear 30- class	Beta 30-class	Linear 22-class + weighted	Beta 22-class + weighted
1	1.000	1.031	0.956	1.035	0.949	16	0.778	0.617	0.559	0.491	0.492
2	1.000	1.007	0.927	0.938	0.881	17	0.609	0.564	0.564	0.514	0.514
3	0.900	0.690	0.671	0.709	0.704	18	0.518	0.485	0.480	0.491	0.492
4	0.791	0.819	0.793	0.773	0.756	19	0.469	0.456	0.467	0.214	0.234
5	0.900	0.974	0.888	0.982	0.878	20	0.524	0.390	0.233	0.319	0.302
6	0.645	0.463	0.486	0.367	0.402	21	0.340	0.354	0.333	0.328	0.323
7	0.792	0.830	0.756	0.718	0.682	22	0.315	0.398	0.397	0.369	0.374
8	0.560	0.756	0.825			23	0.218	0.137	0.119	0.319	0.302
9	0.734	0.884	0.705			24	0.312	0.317	0.326	0.325	0.333
10	0.520	0.634	0.561			25	0.170	0.247	0.258	0.190	0.209
11	0.927	1.608	0.984	0.266	0.271	26	0.244	0.272	0.287	0.222	0.249
12	0.170	0.010	0.090			27	0.259	-0.184	0.055		
13	1.000	1.124	0.933	1.224	0.955	28	0.582	0.455	0.474	0.424	0.447
14	0.380	0.660	0.671	0.664	0.702	29	0.226	0.253	0.257	0.246	0.254
15	0.380	0.296	0.304	0.266	0.271	30	N/A	2.263	1.000	1.583	0.991

Table 5: HOST coefficients for linear models and "linear-equivalent" values for beta regression models. Shaded regions correspond to combined classes.

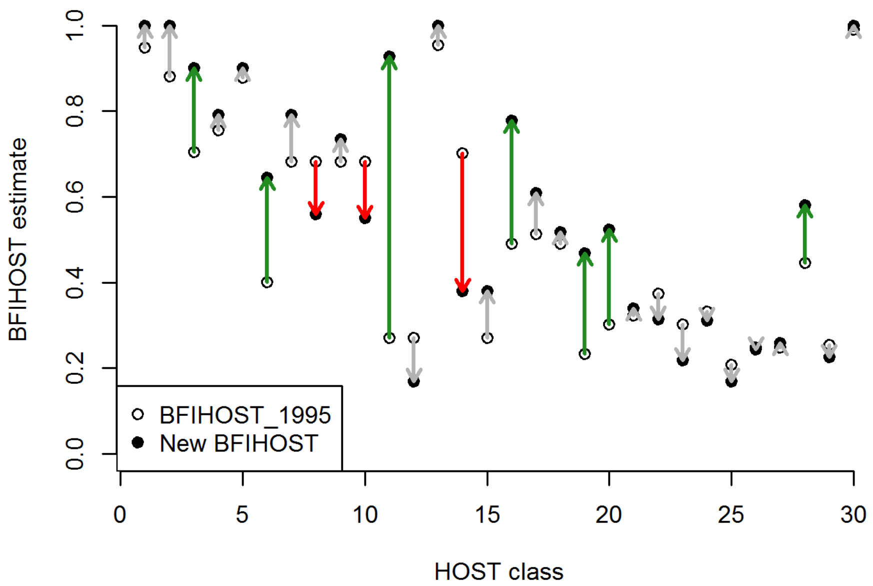

Figure 5: Comparison of BFIHOST estimates of single-class catchments, with strong differences highlighted with upward arrows (New BFIHOST much larger) and downward arrows (New BFIHOST much smaller).

where AREA is the area of the catchment in km^2, SAAR is the standardised annual average rainfall in mm (based on data from 1961-1990), FARL is a coefficient describing attenuation due to lakes and reservoirs, and QMED is measured in m^{3}s^{-1} (Kjeldsen *et al.*, 2008). Figure 7 shows the values estimated for QMED under the two methods of deriving BFIHOST. Here it can be seen that the value for QMED under the new BFIHOST$_{2019}$ performs slightly better, particularly for catchments with smaller values of QMED (comprising smaller catchments and more permeable catchments), where the QMED model typically performs less well (Vesuviano *et al.*, 2016).

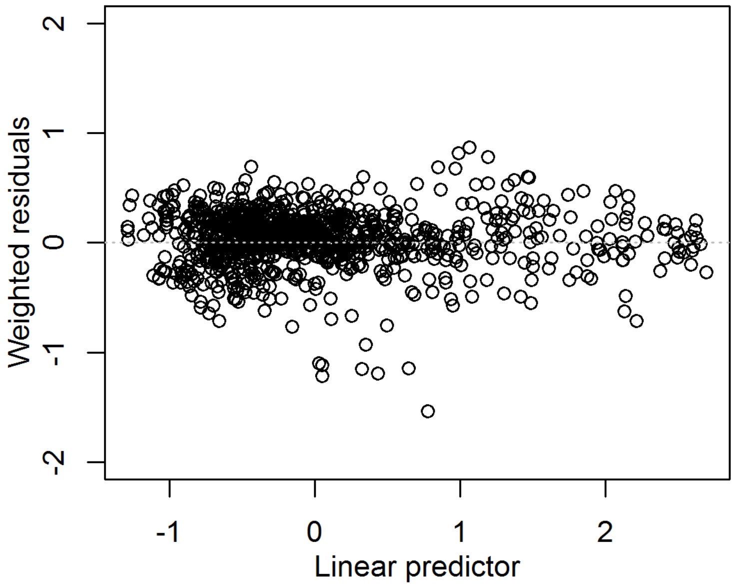

Figure 6: Beta regression residual plot showing scaled residuals against estimates (linear predictors).

Table 6 shows that estimation of QMED and BFIHOST in impermeable catchments (BFI < 0.4) is improved (in terms of factorial standard error) by the new model at all catchment sizes but performance is very similar across sizes of catchments. Hence, $BFIHOST_{2019}$ can address the concerns about parameter estimates for specific scarce HOST classes without a loss of estimation performance when used in the context of generally larger catchments. It should be noted that the equation was calibrated using the $BFIHOST_{1995}$ estimates and thus further model improvement may be gained by recalibrating the equation using $BFIHOST_{2019}$.

CONCLUSIONS

This paper has investigated the potential for an updated method of estimating base flow at ungauged locations, improving on $BFIHOST_{1995}$ by applying a beta regression model instead of the original capped linear regression model, in addition to using a new larger dataset of over 900 gauged catchments.

Choosing a beta regression allowed a model to be fitted which naturally gives estimates strictly between zero and one, avoiding hydrologically unrealistic estimates for BFI. Some HOST classes do not occur in any great quantities at any location, or are highly concentrated in an extremely small number of locations. In locations where these are present, base flow is often poorly estimated by $BFIHOST_{1995}$.

AREA	BFI	QMED fse BFIHOST$_{1995}$	QMED fse BFIHOST$_{2019}$	BFI fse BFIHOST$_{1995}$	BFI fse BFIHOST$_{2019}$
All	All	1.535549	1.531486	1.102568	1.103955
< 40 km^2	All	1.743654	1.755493	1.123471	1.128958
> 40 km^2	All	1.499231	1.492109	1.098892	1.099484
All	< 0.4	1.479287	1.450971	1.103977	1.083578
All	> 0.4	1.555649	1.559788	1.102043	1.11066
< 40 km^2	< 0.4	1.692531	1.658577	1.121898	1.102398
< 40 km^2	> 0.4	1.792398	1.846157	1.12499	1.151069
> 40 km^2	< 0.4	1.397002	1.369949	1.097331	1.076356
> 40 km^2	> 0.4	1.52839	1.526364	1.099365	1.105668

Table 6: Description of factorial standard error of QMED and BFI under BFIHOST$_{1995}$ and BFIHOST$_{2019}$ for small/large and permeable/impermeable catchments.

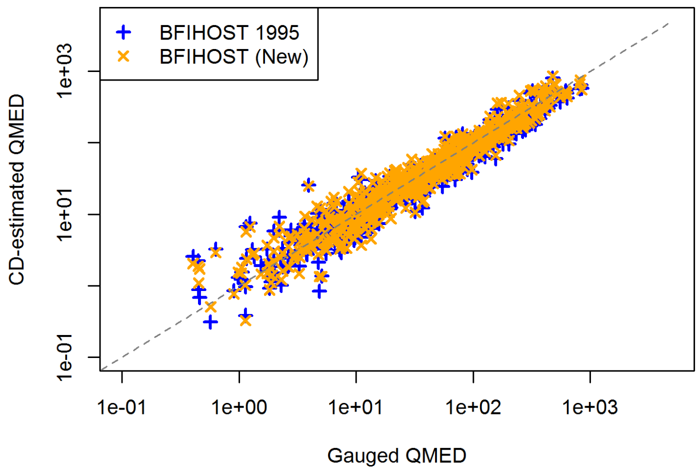

Figure 7: Comparison of fit of Catchment Descriptor (CD) QMED equation under the old BFIHOST from Boorman et al 1995, and the new BFIHOST model.

This issue was still present in the original beta regression model, giving incredibly high/low values for these classes, due to insufficient information to fit accurately. To this end, HOST classes were combined, grouping rare classes with HOST classes that are more abundant and have similar physical and hydrological properties (Table 2). This led to a 22-class model which gave a good fit without a need to impose artificial constraints on the model parameterisation.

To demonstrate its applicability and validity, this new $BFIHOST_{2019}$ was used in the existing QMED catchment descriptor equation. The resultant estimates are an improvement over the use of the equation with the original BFIHOST estimates, despite the equation having been fitted using those original BFIHOST estimates. This is in addition to the core objective of resolving unrealistic estimation in small catchments dominated by single HOST class values that were poorly represented in the dataset used in the original model development. To extend this work, it would be fruitful to recalibrate the QMED equation using the generalised linear model developed in Kjeldsen *et al.*, (2008), and also to recalibrate the parameters for the ReFH2 model (Wallingford Hydrosolutions, 2016b), namely C_{max}, BL and BR.

REFERENCES

Akaike, H. 1974 *A new look at the statistical model identification*, IEEE Transactions on Automatic Control, 19(6), 716-723. doi: 10.1109/TAC.1974.1100705

Boorman, D., Hollis, J. M., and Lilly, A. 1995 *Hydrology of soil types: a hydrologically-based classification of the soils of United Kingdom*, Institute of Hydrology, Wallingford, UK. doi:10.1029/98GL02804

Calver, A., Lamb, R., and Morris, S. E. 1999 *Flood frequency estimation using continuous rainfall-runoff modelling*. Proceedings of the Institute of Civil Engineers - Water and Maritime Engineering, **136**, 225–234. doi:10.1016/S0079-1946(96)00010-9

Chien, L. C. 2011 *Diagnostic plots in beta-regression models*. Journal of Applied Statistics **38**(8), 1607–1622. doi:10.1080/02664763.2010.515677

Cribari-Neto, F., and Zeileis, A. 2010 *Beta Regression in R*. Journal of Statistical Software **34**(2), 1–24. Retrieved from http://www.jstatsoft.org/v34/i02

Ferrari, S., and Cribari-Neto, F. 2004 *Beta Regression for Modelling Rates and Proportions*. Journal of Applied Statistics **31**(7), 799–815. doi:10.1080/0266476042000214501

Gustard, A., Bullock, A., and Dixon, J. M. 1992 *Low flow estimation in the United Kingdom*, Institute of Hydrology, Wallingford, UK.

Gustard, A., Marshall, D. C. W., and Sutcliffe, M. F. 1987 *Low Flow Estimation in Scotland*, Institute of Hydrology, Wallingford, UK.

Heppell, C. M., Binley, A., Trimmer, M., Darch, T., Jones, A., Malone, E., Collins, A. L., Johnes, P. J., Freer, J. E. & Lloyd, C. E. M. 2017 *Hydrological controls on DOC nitrate resource stoichiometry in a lowland, agricultural catchment, southern UK*. Hydrology and Earth System Sciences **21**(9), 4785–4802. doi:10.5194/hess-21-4785-2017

Hodgson, J. M. 1974 *Soil Survey Field Handbook: Technical Monograph No. 5.* Rothamsted Experimental Station, Harpenden, UK.

Institute of Hydrology 1980 *Low flow studies: Report No. 1 - Research Report*. Institute of Hydrology, Wallingford, UK.

James, G., Witten, D., Hastie, T., and Tibshirani, R. 2013 *An Introduction to Statistical Learning with Applications in R*. Springer-Verlag, New York, USA. doi:10.1007/978-1-4614-7138-7

Kjeldsen, T. R., Jones, D. A., and Bayliss, A. C. 2008 *Improving the FEH statistical procedures for flood frequency estimation.* Joint Defra/Environment Agency Flood and Coastal Erosion Risk Management R&D Programme, Science Report SC050050. London, UK.

R Core Team 2016 R: A Language and Environment for Statistical Computing. Vienna. Retrieved from https://www.r-project.org/

SNIFFER. 2006 *Development of Environmental Standards (Water Resources)*. WFD48 Stage 3: Environmental Standards, Edinburgh, UK. Retrieved from https://www.sniffer.org.uk/wfd48-stage-3-report-pdf

Vesuviano, G., Stewart, L., Haxton, T., Young, A., Hunt, T., Spencer, P., and Whitling, M. 2016 *Reducing uncertainty in small-catchment flood peak estimation*. E3S Web of Conferences **7**, 01008. doi:10.1051/e3sconf/20160701008.

Wallingford Hydrosolutions. 2016a *The Revitalised Flood Hydrograph Model ReFH2 : Technical Guidance*, Wallingford Hydrosolutions Ltd., Wallingford, UK.

Wallingford Hydrosolutions. 2016b *WINFAP 4 QMED Linking equation.* Wallingford Hydrosolutions Ltd., Wallingford, UK.

Young, A. R. 2006 *Stream flow simulation within UK ungauged catchments using a daily rainfall-runoff model*. Journal of Hydrology, **320**(1-2), 155-172. doi:10.1016/j.jhydrol.2005.07.017

Young, A. R., Keller, V., and Griffiths, J. 2006 Predicting low flows in ungauged basins: a hydrological response unit approach to continuous simulation. In *Climate Variability and Change - Hydrological Impacts*, Demuth, S., Gustard, A., Planos, E., Scatena, F. & Servat, E. (*ed.*)International Association of Hydrological Sciences Press, Wallingford, UK, pp. 134–138. doi:10.1063/1.478943

ACKNOWLEDGEMENTS

The authors would like to thank the British Hydrological Society for the opportunity to publish this work. The authors would also like to thank the editors and reviewers for their helpful comments.

None of the contributing authors are IAW members.