Estimating flood peaks and hydrographs for small catchments:

R4 – Estimating the median annual flood (QMED) in small catchments

FCERM Research & Development Programme

Research Report

Date: March 2024

Version: SC090031/R4

We are the Environment Agency. We protect and improve the environment.

We help people and wildlife adapt to climate change and reduce its impacts, including flooding, drought, sea level rise and coastal erosion.

We improve the quality of our water, land and air by tackling pollution. We work with businesses to help them comply with environmental regulations. A healthy and diverse environment enhances people's lives and contributes to economic growth.

We can't do this alone. We work as part of the Defra group (Department for Environment, Food & Rural Affairs), with the rest of government, local councils, businesses, civil society groups and local communities to create a better place for people and wildlife.

Published by:

Environment Agency
Horizon House, Deanery Road,
Bristol BS1 5AH

www.gov.uk/environment-agency

© Environment Agency 2024

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

Further copies of this report are available from our publications catalogue: Flood and Coastal Erosion Risk Management Research and Development Programme or our National Customer Contact Centre: 03708 506 506

Email: fcerm.evidence@environment-agency.gov.uk.

Author(s):
Gianni Vesuviano, Adam Griffin, Lisa Stewart, Andy Young

Keywords:
Flood frequency, United Kingdom, rural, urban, plot-scale, QMED, pooling, DDF, ReFH2, local data

Research contractor:
Centre for Ecology & Hydrology, Wallingford, Oxfordshire, OX10 8BB,
01491 838800.

Environment Agency's Project Manager:
Mark Whitling

Project number:
SC090031

Contents

Contents 3

Acknowledgements...5

Executive summary 6

Important Note: 8

1. Introduction 9

2. Study data...11

2.1 Uncertainty in QMED 23

3. Regression model 25

3.1 Sampling error covariance 25

3.2 Model error correlation 28

3.3 Estimating final regression parameters 29

4. Recalibrating existing FEH model 30

5. New QMED model for small catchments..40

5.1 Catchment descriptors 40

5.2 Model structure exploration..40

5.3 Proposed model...44

5.4 Verification on intermediate-size catchments...49

5.5 Verification on ‘extended’ data set 53

6. Using gauged baseflow index and flow statistics in estimating QMED 57

7. Donor transfer in small catchments..64

7.1 Revisiting donor transfer procedures 64

7.2 Recommendations for donor transfer..78

8. Conclusions and recommendations 80

8.1 Conclusions 80

8.2 Recommendations..80

References 82

List of abbreviations..84

Acknowledgements

Particular thanks are due to the Environment Agency's Project Board, members of which were responsible for developing the vision and scope of this research. The board has continued to advise and support the project team, providing valuable feedback, practical suggestions and lively discussions. The project team is also grateful to staff from the Environment Agency, Natural Resources Wales, the Scottish Environment Agency (SEPA) and the Rivers Agency for supplying data and advice on data quality, and to those who have acted as external reviewers.

Project Board

Anita Asadullah (Project Sponsor, Environment Agency)

Donna Wilson (Project Executive, Environment Agency)

Mark Whitling (Project Manager, Environment Agency)

Andy Eden (Environment Agency)

Tim Hunt (Environment Agency)

Sean Longfield (Environment Agency)

Peter Spencer (Environment Agency)

Rob Stroud (Environment Agency)

Glenda Tudor-Ward (Natural Resources Wales)

External reviewers

Thomas Kjeldsen (University of Bath)

Claire Samuel (Jacobs)

Project Team

Centre for Ecology & Hydrology (CEH)

Lisa Stewart (Project Manager),

Giuseppe Formetta

Adam Griffin

Ilaria Prosdocimi (now at University of Bath)

Gianni Vesuviano

Independent consultant

David MacDonald

JBA Consulting

Duncan Faulkner

Harriett Twohig-Howell

Maxine Zaidman

Lara Bentley

Alexander Siddaway

Wallingford HydroSolutions

Andy Young

Tracey Haxton

This project follows on from phase 1, which was jointly funded by the Environment Agency, the Centre for Ecology & Hydrology and JBA Consulting.

Executive summary

The current FEH statistical method as described by Kjeldsen and others (2008) is the UK's standard statistical method for estimating flood peaks of specified annual exceedance probabilities. It consists of an index flood, equal to the median annual flood, QMED, and a standardised growth curve. In ungauged catchments, the index flood is estimated by a regression equation using four catchment properties. This equation has been shown previously to have more uncertainty in small catchments than in larger catchments.

This report develops 3 general options for reducing uncertainty in estimating the index flood in small catchments:

Option 1 - recalibrating the index flood equation using a data set of small catchments only to provide a 'retuned' equation.

Option 2 - developing a new index flood regression equation, based on other catchment descriptors that are found to explain more variation in the observed QMED in small catchments.

Option 3 - developing a new index flood equation, based on a regression as above, but considering using gauged flow statistics in the regression. These flow statistics require some gauging but can be estimated using equipment that is accurate at flows much lower than QMED.

This study finds that both the recalibrated index flood equation (option 1) and new small catchments equation (option 2) offer improvements in terms of uncertainty in QMED. The new regression equation (option 2) is similar in form to option 1, relying on some measure of catchment area, rainfall and gridded soil properties. However, the improvement offered by option 2 over option 1 is considered too small to justify a complete overhaul of the model structure.

The gauged flow equation (option 3) provides the best model fit to observed median annual flows. However, the smaller number of catchments for which flow statistics were obtained (46) means that this regression equation cannot be generalised with confidence to all small catchments.

The results presented in this report suggested that option 1, published as Equation 16 in this report, can be used to estimate QMED in ungauged catchments smaller than 25 km^2. However, additional research, carried out on a further screened data set and reported in Section 3 of the 'Small catchments overview report' (R0) suggests that, while this new equation does reduce the spread of errors in QMED estimates, it also introduces a consistent bias. It is therefore recommended to continue using the existing (2008) FEH statistical QMED equation for estimating QMED in small catchments.

If reliable daily mean flow statistics are available, it is recommended that the 'QMED linking equation' included in WINFAP 4 (Wallingford HydroSolutions 2016a) should be used. A QMED estimate generated with this equation has less uncertainty than one generated via a catchment descriptor equation, but still contains too much uncertainty to be used as the donor value in a donor transfer procedure.

The potential benefits of donor transfer, whereby the estimated index flood is adjusted using a nearby long-term gauged record, has been re-examined for small catchments, although no clear conclusions can be drawn. For the existing all-catchment FEH equation used with one donor, the spread of errors is reduced compared to applying it with no donors, meaning that donor transfer is generally acting to reduce model errors at the sites where they are the largest. When used with the existing all-catchment FEH equation, donor transfer with two or six donors reduces residuals at more than half of the catchments tested but increases the spread of errors. In practice, this means that using the existing FEH equation with more than one donor at any arbitrary small catchment is slightly more likely to marginally improve the estimate of QMED than it is to make it a lot worse. However, the risk of degrading the final estimate is considered to be too high to recommend using more than one donor catchment.

Rejecting potential donors because they are too dissimilar to the catchment of interest in terms of AREA and SAAR is actively unhelpful: rejecting dissimilar donors can give small improvements for catchments where the QMED equation already performs well, but these improvements are more than outweighed by the risk of increasing the error in the estimate for catchments where the QMED equation does not perform well.

Using a single donor catchment is therefore recommended when the existing (2008) FEH QMED equation is applied to a small catchment. There is no advantage in this donor being similar to the catchment of interest in terms of AREA or SAAR.

For the retuned small catchment equation, donor transfer worsens the estimate of QMED more often than not, regardless of how many donors are used and whether they are screened for similarity in catchment descriptors. However, the mean and spread of errors given by the retuned small catchments equation without donor transfer is considerably less than that given by the original FEH equation with any number of donors, including zero (mean and spread of errors with one donor: −0.0210 and 0.3473 for FEH equation, −0.0099 and 0.3042 for retuned equation).

Donor transfer is unnecessary when the QMED linking equation (Equation 20) is used, as this method directly incorporates at-site flow data from the site of interest.

For larger catchments (that is, both target and potential donor > 25 km^2), donor selection should normally be based on centroid-centroid distance between the catchments. Further research is needed to determine whether it is justified to reject certain donors based on dissimilarities in any catchment descriptors.

This report does not consider how to reduce uncertainty or improve accuracy in estimation of the growth curve, as this topic is reported separately in SC090031/R5 – 'Pooling-group formation for small catchments'.

Important Note:

Work on Project SC090031 ‘Estimating flood peaks and hydrographs in small catchments (Phase 2)’ began in December 2013. Tasks carried out in the early stages of the project have already been documented in several project notes and reports, so it is possible that there may be inconsistencies, particularly in the various data sets and methods that have been applied at different points in time. This report provides a summary of the research carried out throughout the project, and we have detailed the data sets and methods used in each of the stages and tasks.

1. Introduction

Index flood methods are widely used for estimating flood frequency. These methods link the size of a specified flood (usually the median or mean of annual maxima) to a series of catchment properties via a regression equation. In the UK, the existing FEH statistical method (Kjeldsen and others, 2008) includes a 'QMED equation', used to estimate the median of annual maximum flows in an ungauged catchment. The equation estimates QMED as a function of four FEH catchment descriptors, corresponding to catchment area, mean annual rainfall, estimated baseflow index (via soil type), and the size and location of online water bodies (Equation 1).

Equation 1 – QMED equation

$$QMED = AREA^{0.8510} 0.1536^{\frac{1000}{SAAR}} FARL^{3.4451} 0.0460^{BFIHOST^2}$$

The calibration data set for this equation consisted of 602 rural catchments from 1.63 to 9,931 km^2 in area. Within this range it is unbiased, that is, it does not tend towards over or underestimation of QMED. However, the proportion of unexplained variance in QMED is higher for small catchments than it is for UK catchments as a whole (Vesuviano and others, 2016).

During this project, a total of 153 small catchments up to 40.9 km^2 in area and suitable for QMED estimation (according to National River Flow Archive criteria: nrfa.ceh.ac.uk), were identified. This is 80% more than the 85 catchments of equivalent size that were used, along with larger catchments, in calibrating the existing FEH statistical equation. Furthermore, the passage of time and additional data collection means that sampling uncertainty in these gauged records is lower than it was when developing the existing FEH statistical equation.

In this report, the feasibility of producing a unique QMED equation, calibrated using only small catchments, is explored. The calibration data set is discussed in Section 2 of this report. Section 3 outlines the design of a comprehensive regression model, accounting for covariance in sampling errors, which originate from the finite record lengths of the calibration QMED values, and correlation between modelling errors, which originates from spatial factors that cannot be easily related to catchment descriptors. Section 4 applies the regression model to re-fit the parameters of the existing FEH statistical QMED equation. In Section 5, forward stepwise regression is used to develop an alternative model form that is more likely (according to a maximum likelihood procedure) to explain variations in QMED in small catchments. Section 6 repeats the work of Section 5, but incorporates flow statistics, derived from points on the flow duration curve. These require some at-site gauging, though to a very limited degree when compared to gauging QMED directly. Section 7 repeats the work of Section 5, using a larger calibration data set that includes catchments that have not been assessed for QMED or pooling suitability by the measuring authorities. Section 8 lists the conclusions and recommendations of the previous sections.

For estimating flood peaks rarer than QMED, the user is directed to report SC090031/R5: ‘Pooling-group formation for small catchments’.

It is assumed that the reader has a detailed understanding of FEH methods, hydrological terminology, and catchment descriptors.

2. Study data

Developing a regression model for QMED in small catchments requires two types of data:

- gauged QMED data, to give target values for the regression
- catchment properties (descriptors) which can be used in the regression

Full annual maximum records, from which QMED is the median value, and values for 25 catchment descriptors are available for all 153 small, QMED-suitable catchments identified previously during this project. However, one of these 153 catchments (27032, Hebden Beck at Hebden), which was included in earlier phases of the project, is now excluded from this study, as dye tracing has shown that water entering Mossdale Caverns, which are inside the topographic catchment, is transferred out of the catchment (Faulkner, pers. comm.). A summary of this study data set compared with previous studies is shown in Table 1. Only directly comparable catchments (that is, up to 40.9 km^2 and marked as suitable for QMED estimation) are counted. The quantity of small-catchment data available for this study greatly exceeds that of previous UK index flood studies, both in number of catchments and in minimum, mean and maximum record length. Figure 1 shows a map of the gauging stations used in this project.

Table 1 - Summary of AMAX data sets for small catchments (< 40.9 km^2)

Dataset property	This study (2017)	Existing FEH (2008)	FEH (1999)
No. of gauges	152	85	117
Shortest record length	4	4	3
Longest record length	64	56	56
Mean record length	30.4	25.6	17.7
No. AMAX events	4627	2180	2071

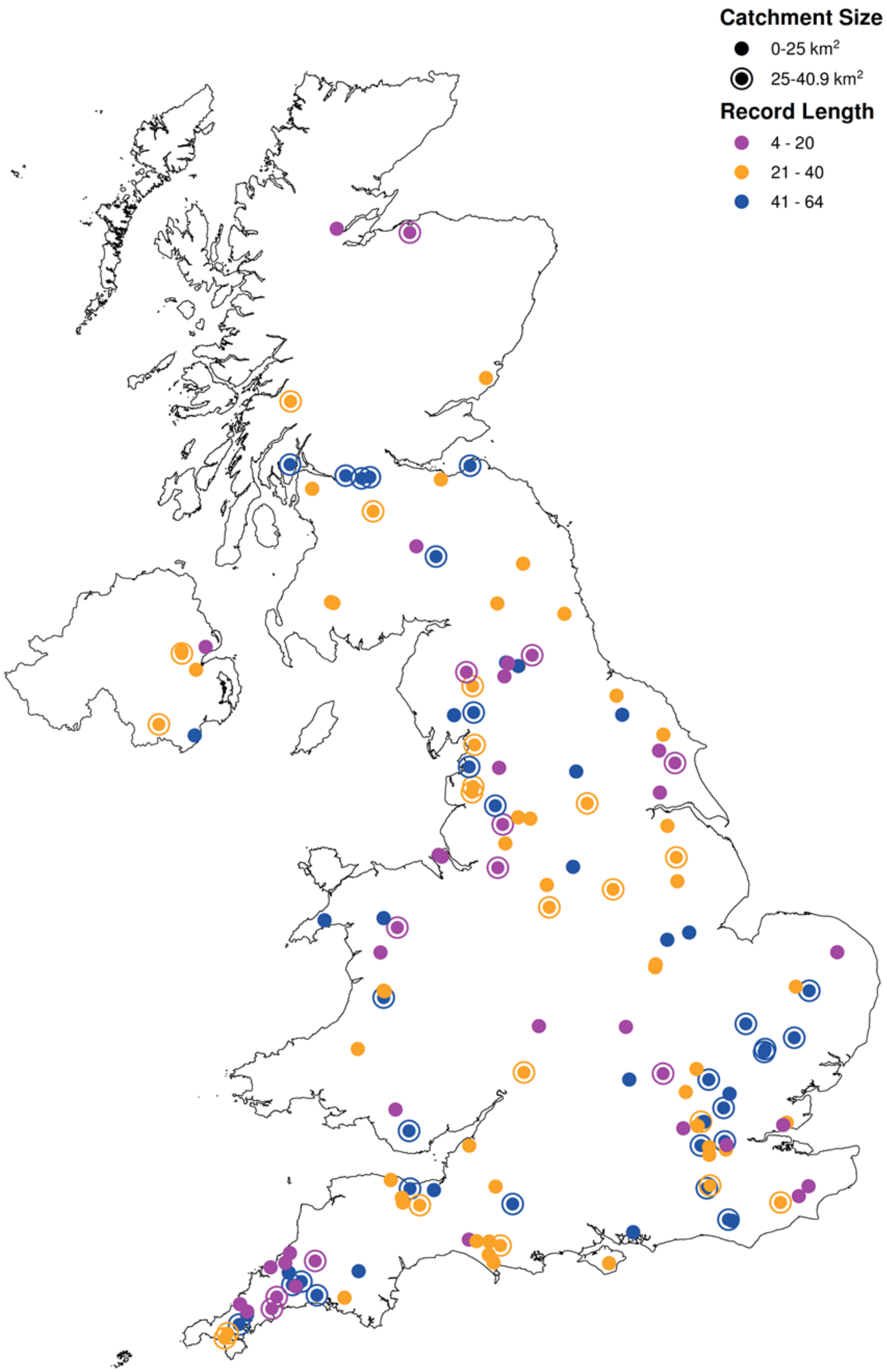

Figure 1 - Locations, areas, and AMAX record lengths of study catchments

Figures 2 and 3 show histograms of catchment descriptor values.

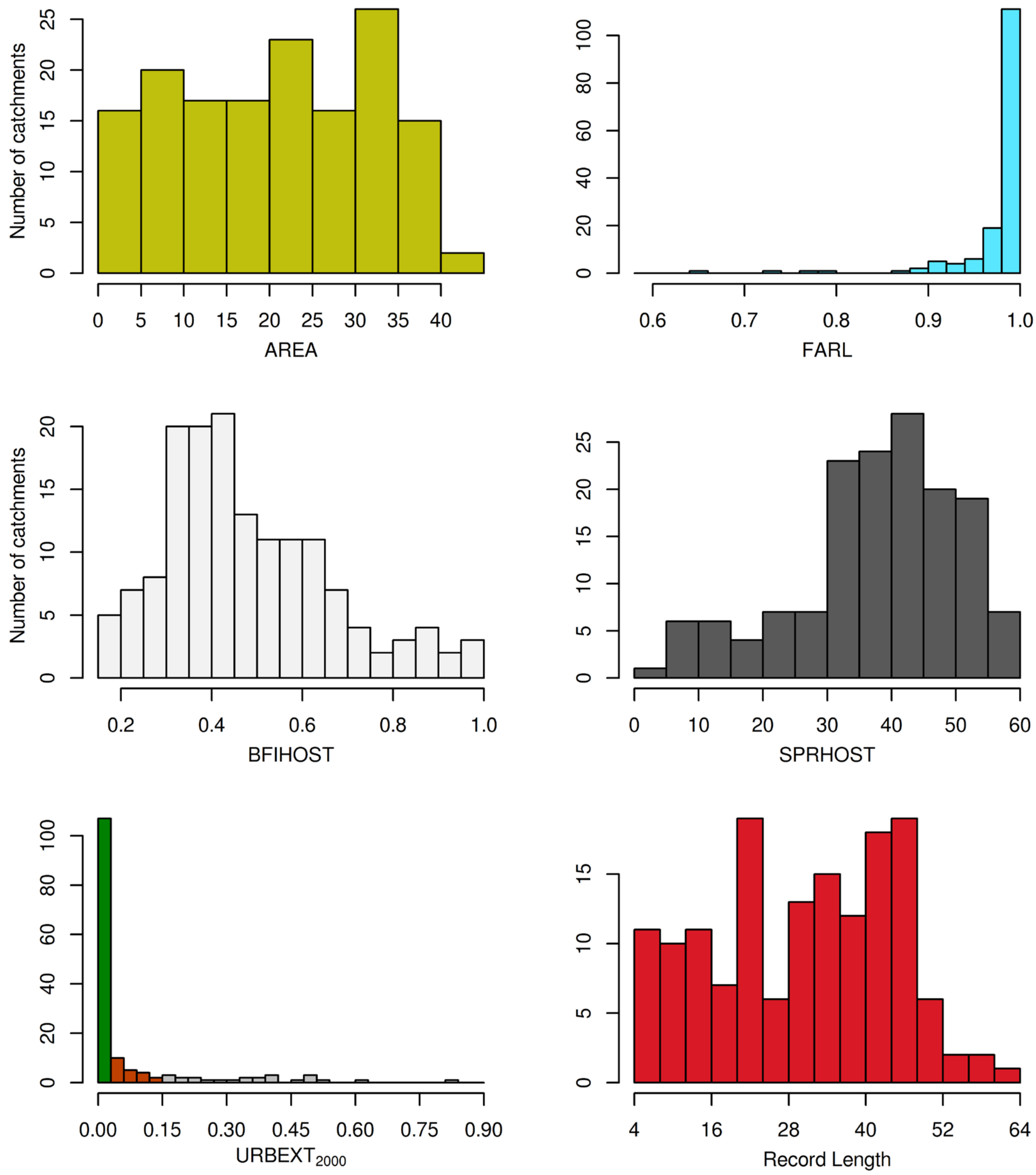

Figure 2 - Catchment descriptor histograms for study data set, part 1

The six histograms in Figure 2 plot the number of catchments (y-axis) by each descriptor value:

- top left histogram: AREA, km^2: catchment area derived from the Integrated Hydrological Digital Terrain Model (IHDTM) - for catchments over 0.5 km^2, IHDTM area is used rather than nominal area (estimated from Ordnance Survey mapping) as all other FEH methods are intended to be used with IHDTM area - for the four catchments under 0.5 km^2, including two with IHDTM area over 0.5 km^2, nominal areas are used instead - in practice, it is very important that practitioners check the accuracy of the IHDTM area using local information. See page 149 of the Flood Estimation Guidelines (LIT 11832: Environment Agency, 2022) for further details

- top right histogram: FARL: flood attenuation due to reservoirs and lakes - dependent on the fraction of catchment covered by on-line water bodies and their proximity to the catchment outlet
- middle left histogram: BFIHOST: baseflow index estimated from HOST soil data
- middle right histogram: SPRHOST: standard percentage runoff estimated from HOST soil data
- bottom left histogram: $URBEXT_{2000}$: weighted coverage of urban, suburban and inland bare ground as a proportion of the catchment, using the Land Cover Map 2000
- bottom right histogram: Record length, years: Number of valid AMAX values

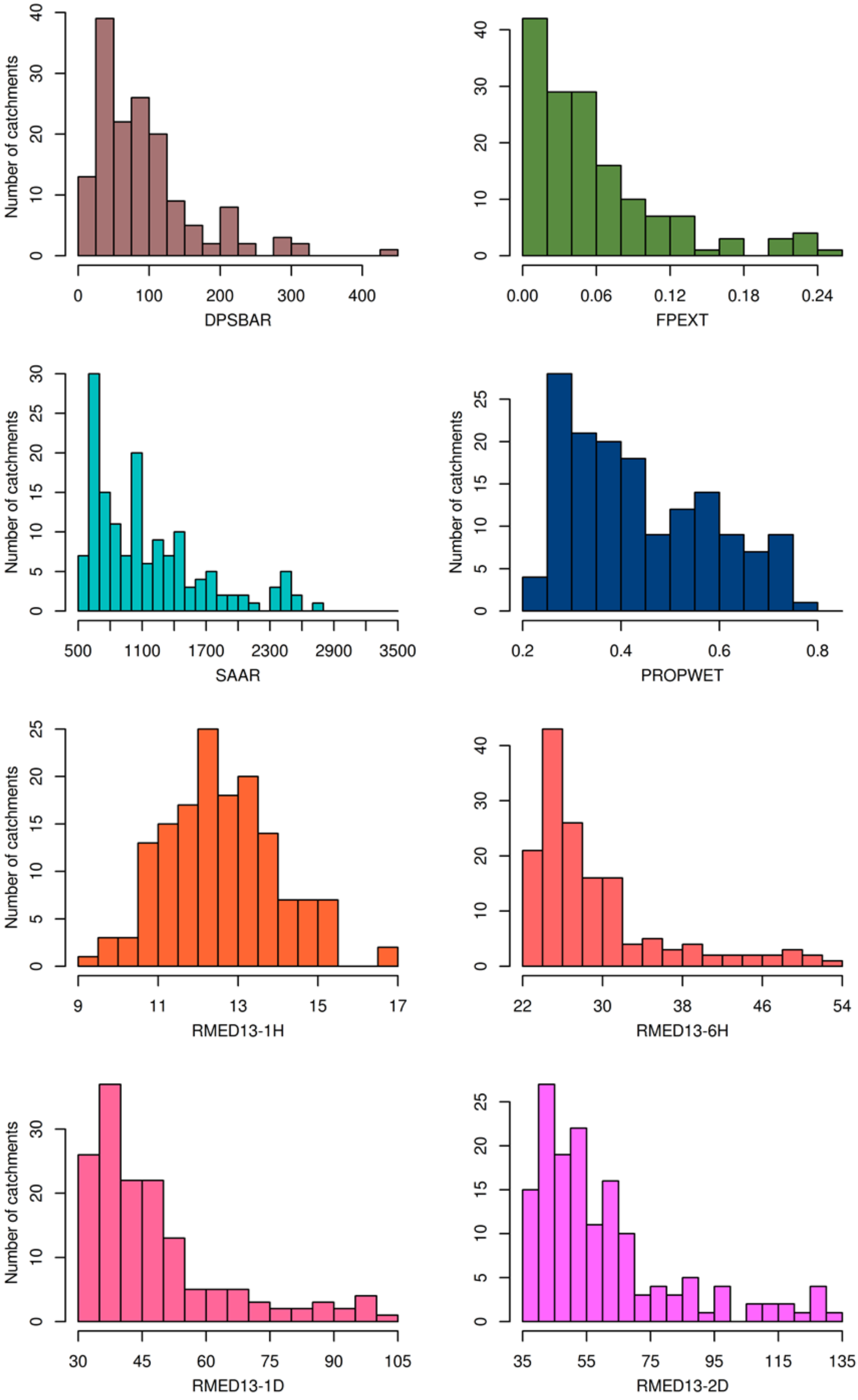

Figure 3 - Catchment descriptor histograms for study data set, part 2

The 8 histograms in Figure 3 plot the number of catchments (y-axis) by each descriptor value:

- top left histogram: DPSBAR, m km^{-1}: mean slope of drainage paths (defined on IHDTM)
- top right histogram: FPEXT: fraction of catchment inundated by 100-year flood.
- second row left histogram: SAAR, mm: mean annual rainfall depth recorded in catchment from 1961 to 1990
- second row right histogram: PROPWET: proportion of period from January 1961 to December 1990 when soil moisture deficit estimated via Met Office Rainfall and Evapotranspiration Calculation System (MORECS) is less than 6 mm
- third row left histogram: RMED13-1H, mm: one-hour, two-year rainfall depth from FEH13 rainfall model
- third row right histogram: RMED13-6H, mm: six-hour, two-year rainfall depth, from FEH13 rainfall model
- bottom row left histogram: RMED13-1D, mm: 24-hour, two-year rainfall depth, from FEH13 rainfall model
- bottom row right histogram: RMED13-2D, mm: 48-hour, two-year rainfall depth, from FEH13 rainfall model

Catchment descriptor histograms show a flat distribution of AREA from 0 to just over 40 km^2, as well as a wide spread of BFIHOST, PROPWET, DPSBAR, FPEXT, SAAR and other rainfall descriptors. Few catchments have significant lake or reservoir attenuation (FARL < 0.9) although this is also true of monitored UK catchments in general. Additionally, the proportion of catchments in this data set that are urbanised ($URBEXT_{2000}$ ≥ 0.03) is lower than in the data set of monitored UK catchments as a whole. However, this data set does contain two extremely heavily urbanised small catchments ($URBEXT_{2000}$ ≥ 0.6). While no urbanised catchments were used in developing either the original FEH (1999) or current ‘improved’ FEH (2008) statistical QMED equation, gauged QMED values obtained from urbanised catchments were used in this study to increase the calibration sample size. QMED estimates from urban gauges were deurbanised by applying the FEH urban adjustment procedures detailed by Wallingford HydroSolutions (2016b) in reverse. To avoid a discontinuity at $URBEXT_{2000}$ = 0.03, these procedures were applied to all catchments, although the impact on gauged QMED in essentially rural catchments was very small. Five-number summaries for all catchment descriptors are shown in Tables 2 and 3.

Table 2 - Five-number summary for study data set (0-40.9 km^2)

Descriptor	Min	25%	Med	75%	Max
AREA	0.04	10.83	21.74	30.66	40.83
FARL	0.645	0.979	0.997	1.000	1.000
FPEXT	0.000	0.0183	0.0438	0.0753	0.2498
FPDBAR	0.000	0.215	0.408	0.630	2.598
BFIHOST	0.172	0.345	0.438	0.594	0.985
SPRHOST	3.270	31.475	39.710	46.795	59.900
PROPWET	0.21	0.32	0.42	0.57	0.79
$URBEXT_{2000}$	0.0000	0.0000	0.0069	0.0397	0.8110
DPLBAR	0.170	3.440	5.230	6.585	11.010
DPSBAR	8.80	37.50	76.25	117.80	433.30
ALTBAR	21.0	82.0	148.5	289.0	656.0
SAAR	555	703	1030	1421	2766
Record length	4	20	32	42	64
RMED13-1H	9.5	11.8	12.5	13.5	16.8
RMED13-6H	22.1	25.1	26.65	30.9	52.7
RMED13-1D	31.0	37.05	42.5	52.1	101.0
RMED13-2D	36.6	44.7	53.7	67.1	131.2
QMED	0.066	2.877	6.448	13.068	69.466

Table 3 - Five-number summary for study data set (0-25 km^2)

Descriptor	Min	25%	Med	75%	Max
AREA	0.04	7.10	12.87	19.50	24.58
FARL	0.727	0.988	1.000	1.000	1.000
FPEXT	0.0000	0.0138	0.0375	0.0711	0.2373
FPDBAR	0.000	0.144	0.300	0.556	1.709
BFIHOST	0.172	0.320	0.430	0.597	0.985
SPRHOST	3.27	32.13	40.58	48.58	59.90
PROPWET	0.21	0.32	0.40	0.56	0.74
$URBEXT_{2000}$	0.0000	0.0000	0.0043	0.0369	0.8110
DPLBAR	0.17	2.63	3.74	5.25	9.28
DPSBAR	8.8	36.7	79.9	123.4	433.3
ALTBAR	21	81	161	341	656
SAAR	555	702	1006	1485	2531
Record length	4	17	29	38	64
RMED13-1H	9.6	11.8	12.6	13.4	16.8
RMED13-6H	22.6	25.3	26.5	31.5	50.4
RMED13-1D	32.2	37.1	42.4	53.5	97.2
RMED13-2D	37.4	44.6	53.3	69.2	127.1
QMED	0.066	1.818	4.852	9.124	38.510

The correlation matrices in Figures 4-6 show relationship between pairs of catchment descriptors. Each black dot in each sub-plot represents one catchment. A number below the diagonal indicates the correlation between the two catchment descriptors shown as a scatter plot above the diagonal. In each case, the two descriptors relevant to each scatter plot or correlation can be found by reading directly down/up and left/right from the scatter plot/correlation.

Considering pairs of catchment descriptors (Figure 4), different values of SAAR, BFIHOST, $URBEXT_{2000}$, ALTBAR and DPSBAR are distributed evenly across the full range of catchment areas. The lower-FARL (<0.9) catchments in this data set are, with one exception, larger than 10 km^2. There are a few strong correlations: between SAAR, ALTBAR and DPSBAR (i.e. wetter catchments have higher altitudes, and higher-altitude catchments are steeper). There are moderate negative correlations between URBEXT and the inter-related SAAR, ALTBAR and DPSBAR, due to the locations of major urban areas. There are also moderate negative correlations between BFIHOST and SAAR (wetter catchments tend to be less permeable) and BFIHOST and ALTBAR (due to the relationship between SAAR and ALTBAR), but not between BFIHOST and DPSBAR (since DPSBAR is not perfectly correlated with either SAAR or ALTBAR). FARL is not strongly correlated with any other descriptor, presumably as most catchment have similar FARL values near 1. AREA is also not strongly correlated with any other descriptor in this dataset, though it is noted that this small catchment dataset does not span a UK-representative range of areas.

The small catchments data set contains no catchments with both high SAAR and high BFIHOST values, though this is a characteristic of UK climate and soils rather than a potential void in the data set (Figure 5). Figure 5 also shows that the lack of more-urbanised, higher-SAAR catchments is consistent with version 4.1 of the NRFA peak flows data set (NRFA 2014), although this is also a consequence of UK climatic and topographical factors. However, Figure 5 also shows much smaller correlations between SAAR, BFIHOST and $URBEXT_{2000}$ than shown in Figure 4 for the small catchment data set, and a much greater proportion of urbanised catchments built on less permeable soils.

One objective of this study is to determine if descriptors like RMED13-1H, RMED13-6H, RMED13-1D or RMED13-2D have more predictive power over QMED in small catchments than SAAR. This is being investigated because the critical durations for small catchments are generally short: for example, if the critical duration in one catchment is six hours, then there should be an obvious link between the two-year flood and the six-hour, two-year storm. Testing this hypothesis, however, is complicated by the strong correlations between the two-year storms of different durations (0.67 to 1.00), and between these storms and descriptors acting on much longer timescales, namely SAAR and PROPWET (0.36 to 0.98, Figure 6).

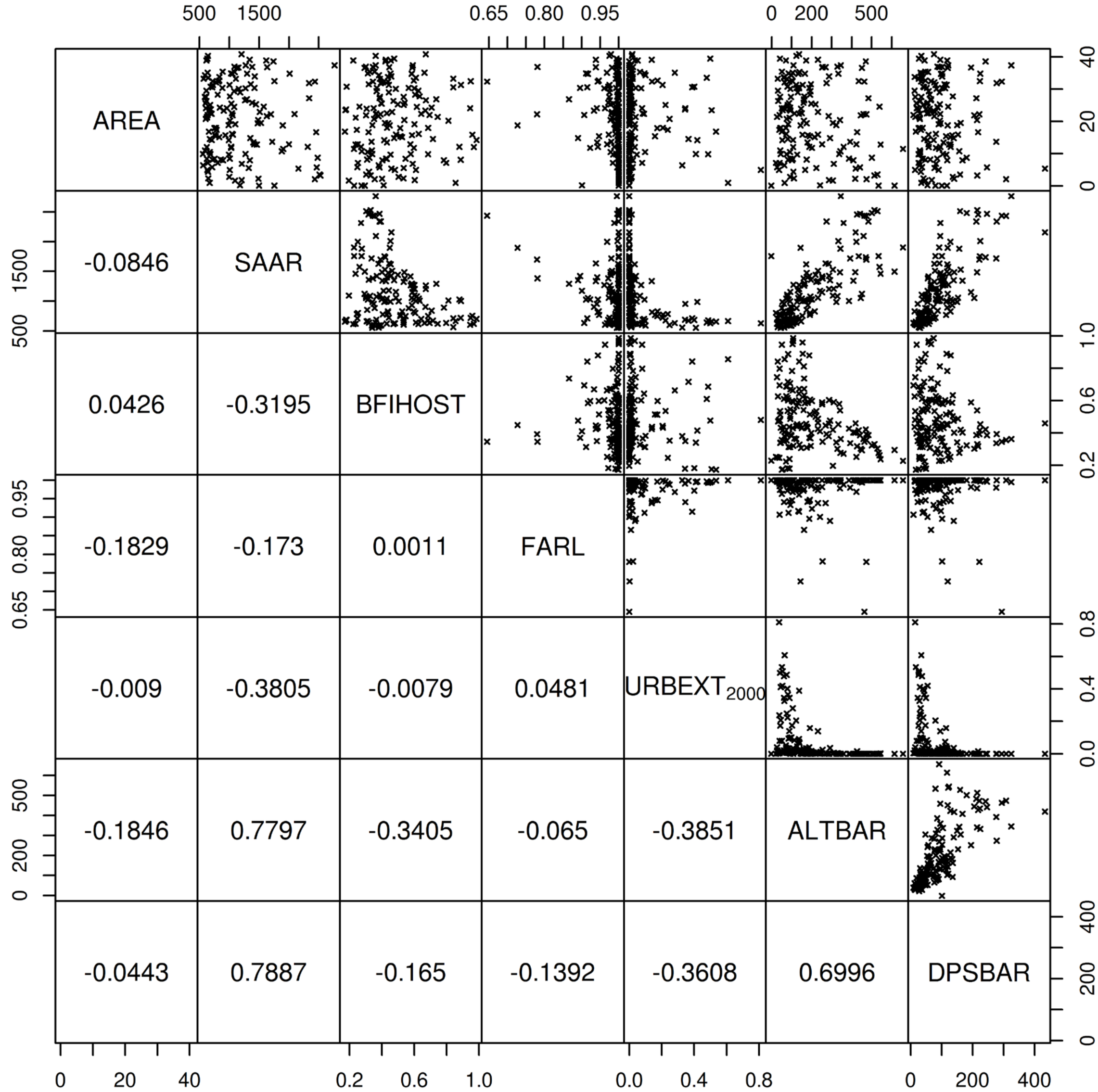

Figure 4 - Correlation matrix of catchment characteristics for small catchments data set

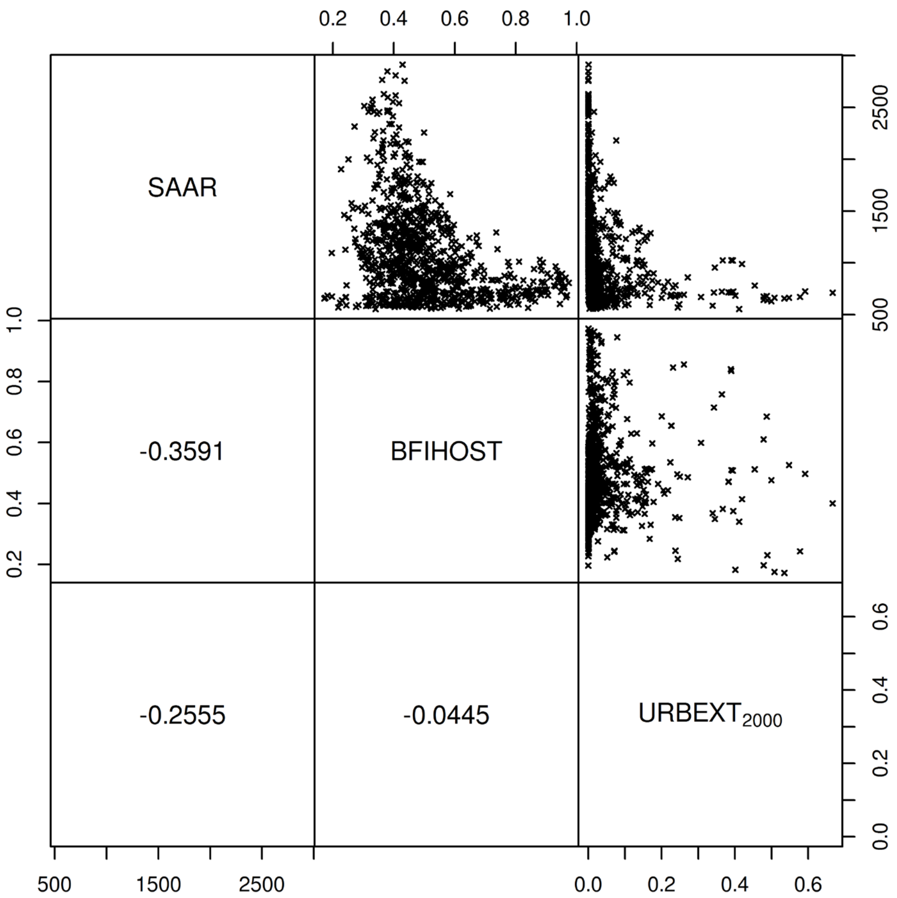

Figure 5 - Correlation matrix of catchment characteristics for NRFA peak flows data set v4.1

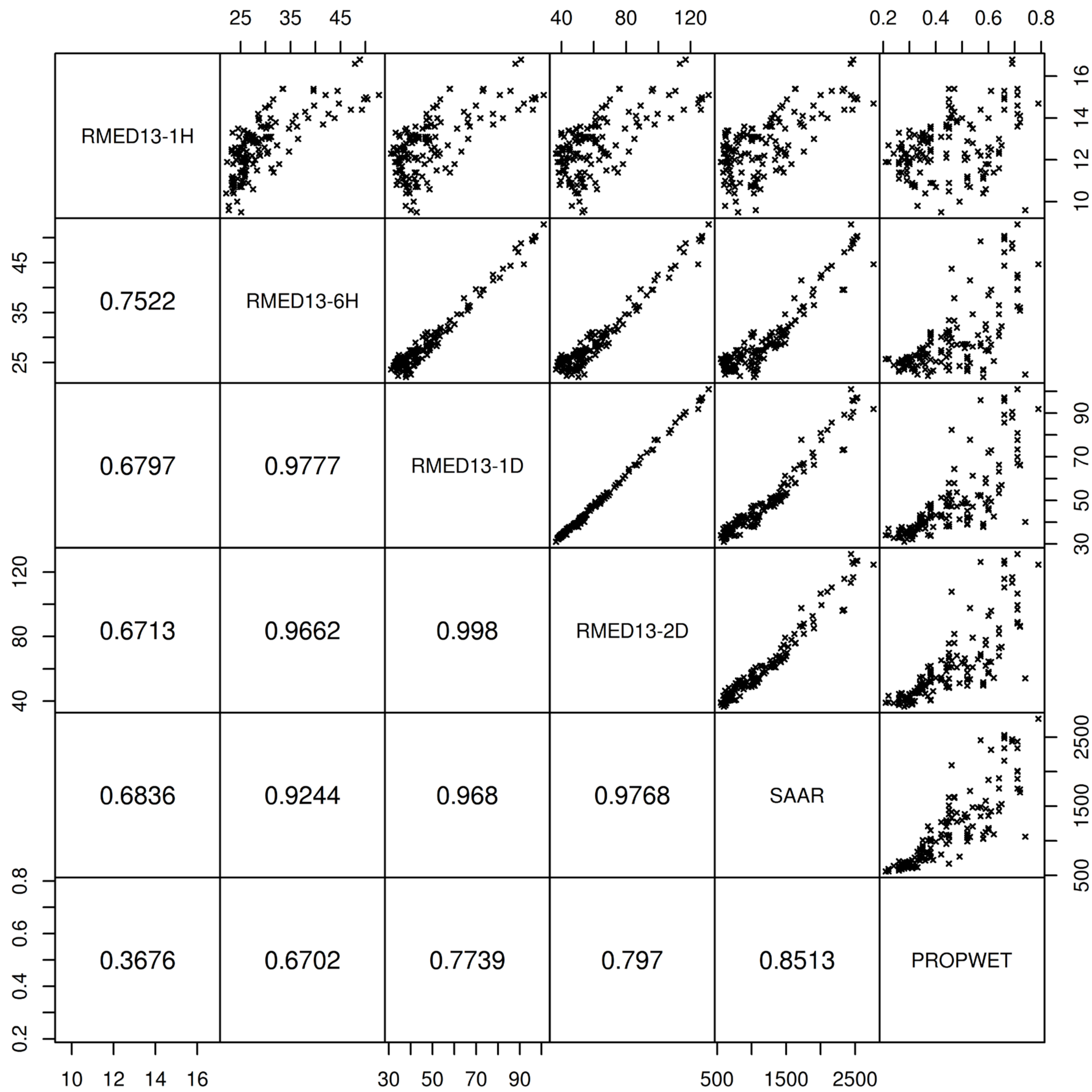

Figure 6 - Correlation matrix of rainfall-related descriptors for small catchments data set

Gauged QMED values at all stations were taken as the median of the annual maximum series, regardless of record length. This was either a single gauged value (for series of an odd length) or the arithmetic mid-point between two gauged values (for series of an even length). There is no lower record length threshold below which POT series are used instead of AMAX series; this is consistent with developing the existing 'improved' QMED equation. The justification given there applies equally here, namely that as the regression model accounts for the variance of and covariance between QMED observations, the higher uncertainty in QMED estimates made from short AMAX records (relative to equivalent POT records) is accounted for by lower weightings of those estimates. Short records are not adjusted for climatic variation, as the procedure does not allow the sampling variance of the adjusted records to be estimated. This makes it impossible to weight QMED observations according to how uncertain the values are. It is strongly recommended to use POT data, if available and of sufficient quality, and to perform climatic adjustment, if estimating QMED from a short record at a single gauging station, as concerns relating to QMED variance-covariance and overall model structural error are not relevant for that type of study.

In urbanised catchments, the gauged QMED value was reduced to an estimated as-rural value by applying the urbanisation relationships published by Wallingford HydroSolutions (2016b) in reverse (Equations 2 and 3). Using the existing urbanisation relationships means that any new QMED estimation equations developed in this study will be compatible with the existing urban adjustment procedures.

Equation 2 – QMED urbanisation equation

$$QMED_{urban} = QMED_{rural}(1 + IF \times URBAN)^{1.25} PRUAF^{1.33}$$

Equation 3 – PRUAF equation

$$PRUAF = 1 + IF \times URBAN\left(\frac{70}{69.366 - 65.686BFIHOST} - 1\right)$$

By default, $IF = 0.3$ and $URBAN = 1.567URBEXT_{2000}$ as shown in Equation 4:

Equation 4 – Relationship between IF, URBAN and URBEXT$_{2000}$

$$IF \times URBAN = 0.4701URBEXT_{2000}$$

WINFAP 4 specifies the use of IF in its urbanisation relationships so that, if the fraction of urban surfaces that is impermeable is known or suspected to be different from 30% in a particular catchment, the user can account for this.

This deurbanisation relationship presents its own uncertainty issues, including the absence of ‘as-rural’ QMED values to pair with gauged QMED values in urban catchments. The guidance in Wallingford HydroSolutions (2016b) does not report goodness-of-fit metrics such as R^2 or fse for the deurbanisation procedure. Moreover, 27 of the 93 study catchments have $URBEXT_{2000} \geq 0.03$, so any gains to be made from an investigation into only essentially rural catchments would be potentially obscured by the increased uncertainty resulting from reducing the data set size from 93 to 66 members.

As such, the urban adjustment procedure is applied unmodified, as fitting new relationships for small catchments is beyond the scope of this work, and, as mentioned previously, it is a pragmatic decision that preserves compatibility with existing techniques.

2.1 Uncertainty in QMED

An expression for the asymptotic sampling uncertainty of the p’th quantile of any continuous distribution was published by Mosteller (1946) and is repeated in Equation 5:

Equation 5 – Mosteller's (1946) expression for asymptotic sampling uncertainty

$$\sigma^2 = \frac{p(1-p)}{nf^2\left(F^{-1}(p)\right)}$$

Where n is the length of the sample (in this case, the number of AMAX values), f is the probability density function and F^{-1} is the value of the distribution at the p'th quantile. This expression calculates the theoretical variance of any quantile of the distribution, which can be used to estimate the theoretical uncertainty in any T-year flood (where p = 1 – 1/T). For the logarithm of the median of a generalised logistic (GLO) distribution, the expression simplifies greatly to give Equation 6 (Kjeldsen & Jones 2009):

Equation 6 – variance in QMED as a function of gauged record length and GLO distribution parameters

$$\sigma^2 = \frac{4\beta^2}{n}$$

Where σ^2 is the variance of ln(QMED) and β is equal to the scale parameter, α, of the distribution divided by the location parameter, ξ, of the distribution. α can be estimated from the AMAX series via L-moment methods (e.g. Hosking and Wallis 1997), while the median of the AMAX series is a reasonable estimate of ξ (Robson & Reed 1999, Kjeldsen and others, 2008).

3. Regression model

In this study, the regression model for QMED was developed using a maximum likelihood (ML) framework, identical to that used to develop the existing FEH statistical QMED equation. The key advantages of ML (and generalised least-squares regression) over ordinary least-squares regression are that observations can be weighted according to their variance and that covariance between observations can be accounted for. However, ML is preferred for this study as the relationship fitted to model residuals is less affected by sampling variations (Kjeldsen and Jones 2009). The general regression model is of the same form as the existing FEH statistical method (see Equation 7):

Equation 7 – general form of FEH QMED regression model

$$y_i = \boldsymbol{x}_i^T + \omega_i = \boldsymbol{x}_i^T \boldsymbol{\theta} + \varepsilon_i + \eta_i$$

In Equation 7, y_i is a log-transformed QMED estimate, $\mathbf{x}_i$ is a vector of catchment descriptors, $\boldsymbol{\theta}_i$ is a vector of regression model parameters and ω_i is total error, comprising sampling error, ε_i, and modelling error, η_i, for i = 1,…,M, where M is the number of gauged catchments. Both errors are assumed normally-distributed with zero mean values. Sampling and modelling errors at one site are assumed to be independent of each other. Diagonal elements of the modelling error covariance are assumed to be identical and equal to σ_η^2. Denoting the error covariance matrices by $\boldsymbol{\Sigma}_\omega$, $\boldsymbol{\Sigma}_\varepsilon$ and $\boldsymbol{\Sigma}_\eta$ respectively, the total error covariance matrix is given by Equation 8.

Equation 8 - The total error covariance matrix

$$\boldsymbol{\Sigma}_\omega = \boldsymbol{\Sigma}_\eta + \boldsymbol{\Sigma}_\varepsilon = \sigma_\eta^2\left(\boldsymbol{R}_\eta + \boldsymbol{\Sigma}_\varepsilon / \sigma_\eta^2\right) = \sigma_\eta^2 \boldsymbol{G}$$

In Equation 8, $\mathbf{R}_\eta$ is the model error correlation. The non-diagonal elements of $\mathbf{R}_\eta$ represent between-catchment model error correlation, which manifests as geographical clustering of over and underestimation, and occurs when there is a geographical effect unaccounted for by any catchment descriptor. Between-catchment sampling error correlation occurs because single rainfall events often fall across several nearby catchments, therefore one rainfall event could generate an AMAX flow event in several neighbouring catchments.

3.1 Sampling error covariance

In common with Equation 6, all elements of the sampling error covariance matrix are based on the asymptotic variance of the sample median (see Equation 9):

Equation 9 – general form of sampling error covariance between two gauged sites

$$\boldsymbol{\Sigma}_{\varepsilon,ij} = \begin{cases} 4\beta_i^2/n_i & i = j \\ 4\beta_i\beta_j \dfrac{n_{ij}}{n_i n_j} r_{\varepsilon,ij} & i \neq j \end{cases}$$

In Equation 9, n_{ij} denotes the number of recorded years shared by station i and station j, indicating that apparent covariance between flows can only be genuine if both flows relate to the same event. The term $r_{\varepsilon,ij}$ links the correlation between QMED values to the geographical distance between catchment centroids.

In common with the existing QMED equation it is assumed that the correlation between at-site sampling errors is the same as the correlation between QMED. Therefore, a bootstrap experiment was used to determine an expression for $r_{\varepsilon,ij}$. However, the expression derived here is subject to greater uncertainty as there are only 108 rural catchments (67 under 25 km^2) and just 354 pairs of gauges with a minimum 40-year record overlap (65 pairs where both catchments are under 25 km^2). Urban catchments were not used in this experiment, as each AMAX value would require a different level of adjustment to be deurbanised, depending on its estimated return period. The form of the expression for $r_{\varepsilon,ij}$ was copied from the existing FEH statistical method project report and the final equations are reported in Equation 10.

Equation 10 – Final equations for sampling error covariance

$$r_{\varepsilon,ij} = \begin{cases} 0.6729e^{(-0.1352d_{ij})} + (1 - 0.6729)e^{(-0.0043d_{ij})} & both\ catchments\ < \ 25\ km^2 \\ 0.7505e^{(-0.1472d_{ij})} + (1 - 0.7505)e^{(-0.0026d_{ij})} & both\ catchments\ < \ 40.9\ km^2 \end{cases}$$

The correlations between QMEDs are shown in Figure 7. As in Kjeldsen and others (2008) for the existing QMED equation, the correlation is vague but appears to be real. It is unclear if the model form re-used here is optimal, as there are few points to the left of the ‘knee’ in either relationship. The model requires correlation to rise to 1 at a centroid-centroid distance of exactly zero, and this model form allows that. However, it should be noted that it is possible in practice for catchments with centroids that are physically close to be very different in hydrological character (for example, catchment area).

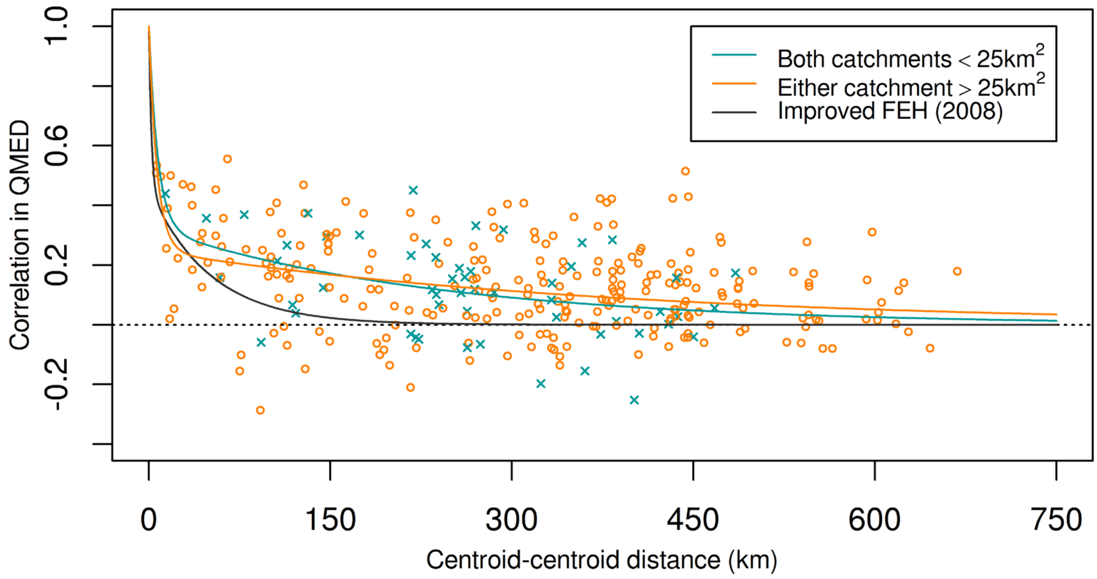

Figure 7 – Correlation between ln(QMED) values and centroid-centroid distance of catchment pairs

The scatter graph in Figure 7 plots the correlation in pairs of QMED values on the y-axis (from 0.2 to 1.0) against the centroid-centroid distance between pairs of catchments on the x-axis (from 0 to 750 km). The following elements are plotted:

- both catchments are <25km^2: blue line
- either catchment is >25km^2: orange line
- improved FEH (2008): black line

(The fitted equation is only shown for the existing FEH method.)

Equation 9 references both $r_{\varepsilon,ij}$ and β terms. In estimating the existing FEH equation for QMED, at-site values for β, derived from AMAX data, were not used. Instead, a regression model between β and catchment descriptors was formed (Equation 11), with the intention of reducing the sampling noise inherent in the data-derived values. Table 4 shows the summary statistics for Equation 11. This is able to model 50% of the variance in β with an fse of 1.412, comparing favourably with the 28% of variance modelled with an fse of 1.387 in the improved FEH study (Kjeldsen and others, 2008). A generalised regression equation for β would not necessarily be expected to fit the observed values very well, as a small spread of errors could indicate over-fitting to the observed sampling noise, particularly if achieved through a large number of parameters. However, this equation, which explains almost double the variance in β, is considered appropriate as it contains no more parameters than the equation used in the improved FEH study.

Equation 11 – Regression model for ln(β), used to estimate sampling error covariance

$$ln(\beta) = -1.1677 - 0.5769\frac{SAAR}{1000} - \frac{0.0387}{AREA} + \frac{0.5237}{1 + FPDBAR}$$

Table 4 - Summary statistics for regression model describing ln(*β*)

Variable	Coefficient	Std. err.	t-value	p-value
Intercept	−1.1677	0.1895	−6.164	0.0000
***SAAR*/1000**	−0.5769	0.0703	−8.211	0.0000
***AREA*$^{-1}$**	−0.0387	0.0096	−4.025	0.0001
(1 + *FPDBAR*)$^{-1}$	0.5237	0.2638	1.986	0.5015

$\sigma^2 = 0.1203$, df = 89, $R^2 = 0.513$

3.2 Model error correlation

Model error correlation is estimated as part of the ML procedure. Kjeldsen and others (2008) set the 'structure' of the correlation as equal to that for sampling error covariance (Equation 12):

Equation 12 – general form of model error correlation fitted to improved FEH QMED equation (Kjeldsen and others, 2008)

$$r_{\eta,j} = \phi_1 e^{(-\phi_2 d_{ij})} + (1 - \phi_1)e^{(-\phi_3 d_{ij})}$$

The regression constants $\boldsymbol{\varphi} = (\varphi_1, \varphi_2, \varphi_3)$ are unknown until the ML estimation procedure is complete. This is because model error cannot be estimated from the data set, as the true values of QMED are unknown. In this study, a simplified structure was used (Equation 13), as preliminary studies found that the term φ_1 was always fitted with a value near 1, making $(1 - \varphi_1)$ near zero and therefore redundant.

Equation 13 – general form of model error correlation fitted to small catchment QMED equations

$$r_{\eta,j} = e^{(-\phi d_{ij})}$$

Model error correlation describes the tendency of the model to over or underestimate QMED relative to gauged data. Therefore, when selecting a donor, it is more important to match model error than any other property (for example, catchment or hydrological similarity), as catchments with similar model error show similar ratios between gauged and estimated QMED values. Furthermore, an appropriate and successful regression model should show minimal residual bias against the terms used in that model, making those terms poor candidates to use in selecting a donor catchment.

3.3 Estimating final regression parameters

Given the error structures specified previously, and a chosen set of catchment descriptors, the regression constants are estimated through minimising the negative log-likelihood function. This measures the 'plausibility-of-fit' of the model, that is, how plausible it is that the estimates match the observations. Lower values of negative log-likelihood indicate more plausibility, given that values are only directly comparable for models fitted to the same data. The negative log-likelihood function is presented in Equation 14:

Equation 14 – Negative log-likelihood function

$$-l(\boldsymbol{\theta}, \sigma_n, \boldsymbol{\varphi} | \boldsymbol{X}, \boldsymbol{y}) = \frac{1}{2} ln[\det(\sigma_n^2 \boldsymbol{G})] + \frac{1}{2} (\boldsymbol{y} - \boldsymbol{X\theta})^T (\sigma_n^2 \boldsymbol{G})^{-1} (\boldsymbol{y} - \boldsymbol{X\theta})$$

Into this, it is noted that the value of **θ** that minimises the negative log-likelihood function can be expressed as shown in Equation 15:

Equation 15 – Value of θ (QMED model parameter vector) that minimises negative log-likelihood function

$$\widehat{\boldsymbol{\theta}} = (\boldsymbol{X}^T \boldsymbol{G}^{-1} \boldsymbol{X})^{-1} \boldsymbol{X}^T \boldsymbol{G}^{-1} \boldsymbol{y}$$

Therefore, the objective function can be minimised simply by searching over the model error correlation parameter and the model error variance simultaneously.

4. Recalibrating existing FEH model

The existing FEH statistical model for QMED estimation is presented in Equation 1. As this model was fitted to catchments up to almost 10,000 km^2, it is not expected to be optimal for small catchments specifically. Therefore, a quick and easy way to produce a new and improved QMED model for small catchments could be simply to re-apply the existing FEH statistical modelling framework to the project data set. In this section, the existing FEH statistical QMED equation is compared to 'retuned' versions of the same equation, with new regression constants, developed by applying the steps given in Section 3.

Equation 16 shows the 'retuned' regression equation, developed using 93 catchments – only those detailed in Section 2 that are under 25 km^2.

Equation 16 – The 'retuned' regression equation (based on 93 catchments)

$$QMED = 4.7260 AREA^{0.8335} 0.2763^{\left(\frac{1000}{SAAR}\right)} FARL^{2.6266} 0.0404^{BFIHOST^2}$$

- $\sigma_\eta^2 = 0.2823$
- $\phi = 3.0142$

Model error variance is solved as 0.2823 to give a model factorial standard error (fse) of 1.701, which is considerably higher than the values of 1.431 reported for the existing QMED equation and 1.549 reported for the first FEH statistical QMED equation (Robson and Reed 1999). However, for any given set of small catchments, the retuned model is more likely than either existing FEH equation to offer a sample fse, exp(s), close to the model fse, exp(σ_η).

Figure 8 shows regression diagnostic plots for the retuned FEH equation. These demonstrate that the model residuals are unrelated to the gauged or modelled value of QMED. However, most of the negative residuals do not follow a normal distribution.

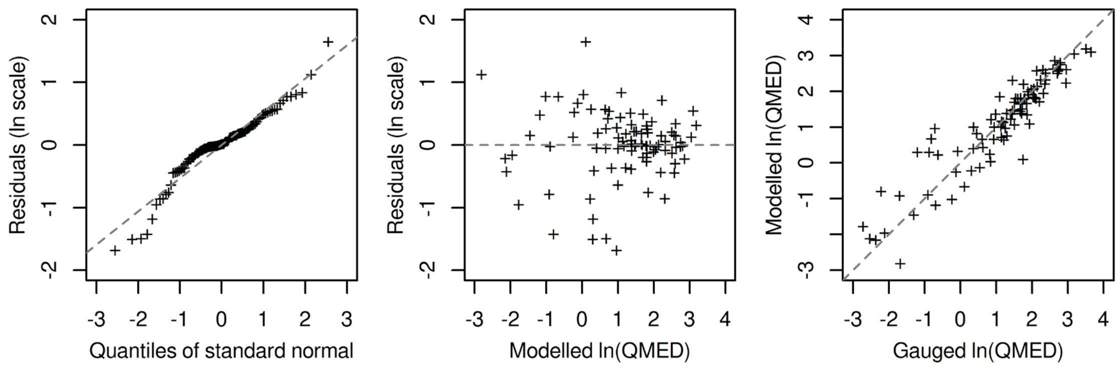

Figure 8 - Regression diagnostic plots for retuned FEH equation

The three scatter graphs in Figure 8 present regression diagnostic plots for the retuned FEH equation:

- left panel: residuals (ln scale) on the y-axis (from -2 to 2) and quantiles of standard normal on the x-axis (from -3 to 3)
- middle panel: residuals (ln scale) on the y-axis (from -2 to 2) and modelled ln(QMED) on the x-axis (from -3 to 4)
- right panel: modelled ln(QMED) on the y-axis (from -3 to 4) and gauged ln(QMED) on the x-axis (from -3 to 4)

Figure 9 shows the influence of each catchment descriptor in Equation 16 over the reasonable range of possible values (orange) and compares this with Equation 1 (grey).

For AREA and particularly FARL, it is shown that the retuned regression constants reduce the range of the catchment descriptor's influence, whereas for SAAR and BFIHOST, the range of influence is slightly extended. This does not indicate that FARL is less important in small catchments; it is due to the reduced range of FARL values in the new calibration data set (Figure 10), that is, it is a sampling issue. The value of BFIHOST has approximately the same effect on QMED, whether the calibration catchments range up to 25 km^2 or up to ~10,000 km^2.

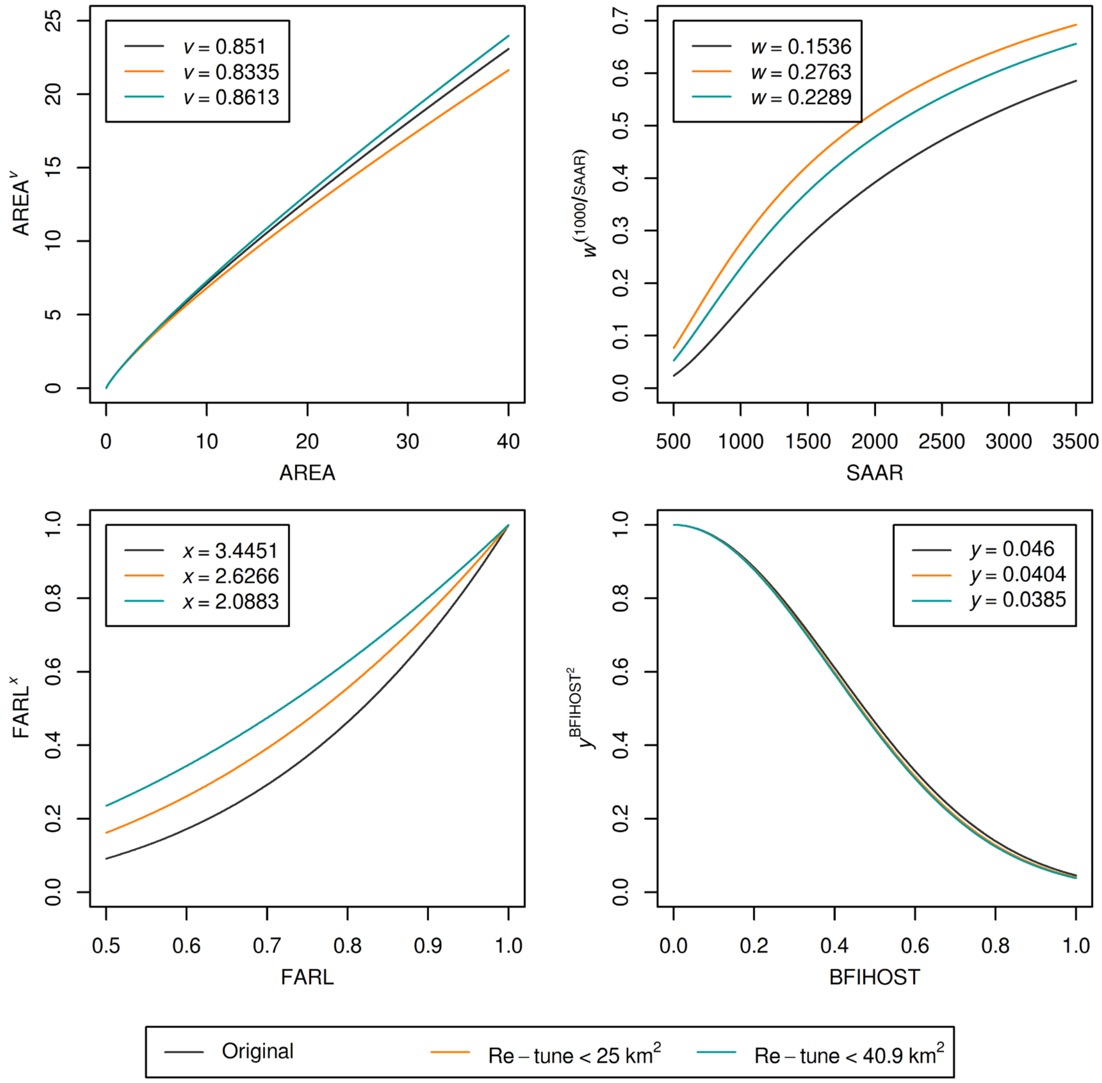

Figure 9 - Influence of catchment descriptors in original QMED equation (grey – Equation 1) and QMED equation retuned with catchments up to 25 km^2 (orange – Equation 16) and 40.9 km^2 (blue – Equation 17)

The four line graphs in Figure 9 show the influence of the four catchment descriptors over the reasonable range of possible values, compared with Equation 1:

- top-left graph: $AREA^v$ from 0 to 25 (y-axis) by AREA from 0-40 (x-axis)
- top-right graph: $w^{(1000/SAAR)}$ from 0.0 to 0.7 (y-axis) by SAAR from 500 to 3500 (x-axis)
- bottom-left graph: $FARL^x$ from 0.0 to 1.0 (y-axis) by FARL from 0.5 to 1.0 (x-axis)
- bottom-right graph: $y^{BFIHOST^2}$ from 0.0 to 1.0 (y-axis) by BFIHOST from 0.0 to 1.0 (x-axis)

On each graph the three lines show:

- original (grey line)
- re-tune <25 km^2 (orange line)
- re-tune <40.9 km^2 (blue line)

Equation 17 shows the 'retuned' regression equation when all 152 catchments discussed in Section 2 are used for calibration. The influence of these catchment descriptors is shown in Figure 9 using blue lines.

Equation 17 – The 'retuned' regression equation (based on 152 catchments)

$$QMED = 5.4800 AREA^{0.8613} 0.2289^{\left(\frac{1000}{SAAR}\right)} FARL^{2.0883} 0.0385^{BFIHOST^2}$$

- $\sigma_\eta^2 = 0.2630$
- $\phi = 3.4107$

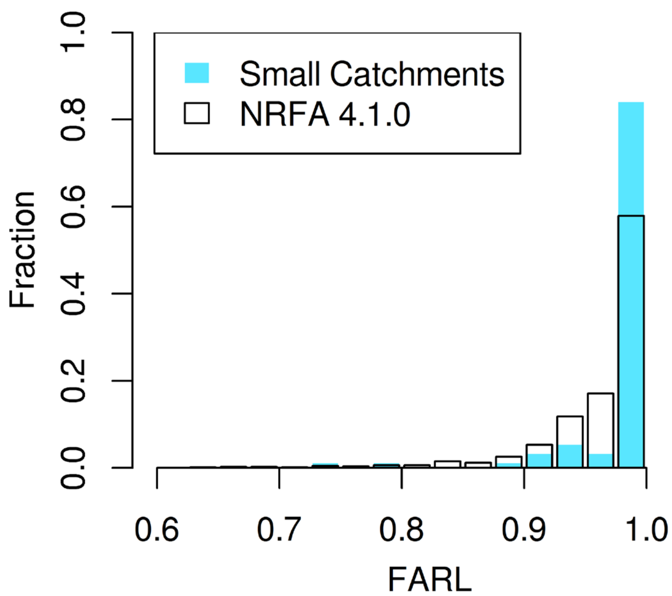

Figure 10 - Density plot of FARL

The bar chart in Figure 10 shows that the small catchment data set contains many more catchments with almost no reservoir and lake attenuation than the NRFA peak flow data set (the rightmost bar), though both are dominated by catchments with little attenuation (FARL > 0.9) Small catchments are plotted by blue bars. The white bars plot NRFA 4.1.0. The x-axis shows FARL (from 0.6 to 1.0) and the y-axis shows fraction (from 0.0 to 1.0).

Recalibration with the 152-catchment data set makes the relationship between QMED and AREA slightly more linear than in the FEH equation. It is plausible that the relationship between QMED and AREA may become more linear for smaller catchments, as rainfall and runoff processes become more homogeneous, and shorter flow paths and less flood plain area offer reduced opportunities for attenuation (Faulkner and others, 2012). Indeed, the principle of quoting greenfield runoff rates in l/s/ha presupposes a linear relationship between flow and area. It is worth noting, however, that all three equations have very similar relationships between QMED and AREA – indeed, Equation 16 has a slightly less linear relationship than Equation 1. This may be a result of the small range of areas used in calibration.

Calibration with 152 catchments up to 40.9 km^2 also reduces the influence of SAAR to a point midway between the two lines shown in Figure 9, but has little effect on the influence of either FARL or BFIHOST. Model error variance for this equation is 0.2630, giving a model fse of 1.670, the smaller value reflecting the increased quantity of calibration data.

Table 5 compares the performance statistics of the existing FEH statistical QMED equation in original and both retuned forms when applied to model the 93 catchments under 25 km^2.

Table 5 - Performance of existing FEH statistical QMED equation in three forms

Model variant	model fse $\exp(\sigma_\eta)$	R^2	RMSE*	sample fse $\exp(s)$
Existing ('improved')	1.431	0.8356	0.6116	1.843
Retuned, <25 km^2	1.701	0.8530	0.5596	1.744
Retuned, <40 km^2	1.670	0.8489	0.5671	1.757

Note: Sum of squares of residuals divided by number of degrees of freedom (93 – 4 parameters – 1 = 88).

Figure 11 plots QMED as estimated by the three parameterisations of the existing FEH statistical QMED equation. The positions of the crosses in the top-right of the plot suggest that the retuned models tend to underestimate the largest values of QMED. A similar observation was made by MacDonald and Fraser (2013) for their QMED equation, containing the same descriptors in the same form, minus FARL. The positions of the circles towards the left of the plot also suggest some overestimation in catchments with QMED values under 1 m^3/s. However, the retuned equations offer similar or smaller residuals than the original equation for these catchments, apart from catchments 29013 (Moor Beck at Clapgate Farm) and 33048 (Larling Brook at Stonebridge). Both of these catchments are exceptionally flat (DPSBAR = 25 and 8.8 respectively) and contain large floodplains (FPEXT = 0.0863 and 0.2331 respectively), so their flood peaks are likely to be influenced by factors other than the four descriptors contained in either the original or retuned FEH equations. Additionally, there is one catchment (39832, Wandle at Carshalton) for which all equations underestimate by a factor of five or more. As this catchment is more heavily urbanised than any of those used to estimate the relationship between urbanisation and QMED it is possible that the deurbanisation relationship, derived using mostly moderately-urbanised catchments, is too strong when applied to more heavily-urbanised small catchments. This is supported by the trend towards underestimating QMED in extremely-urbanised catchments shown by Vesuviano and others (2016).

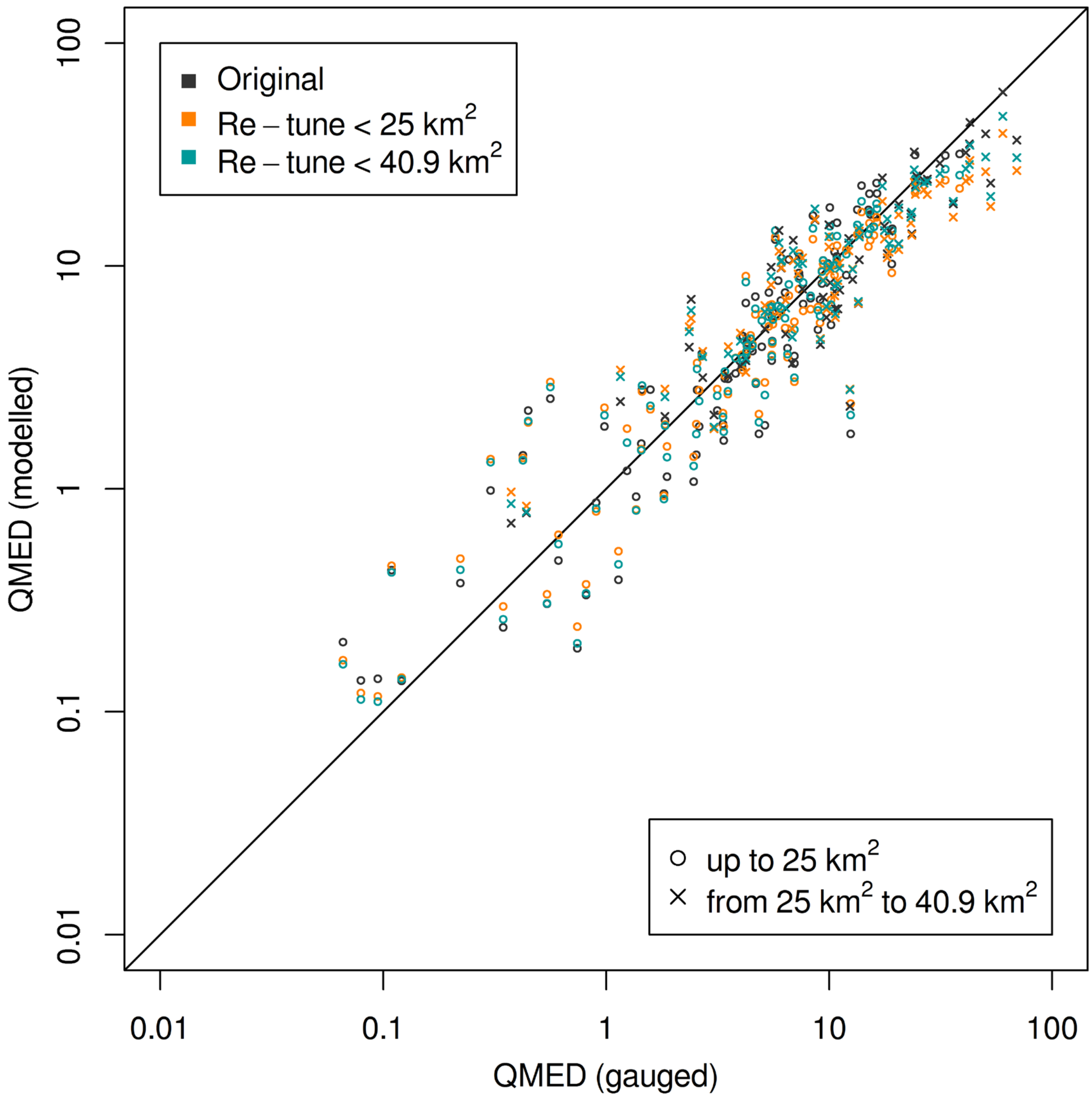

Figure 11 - Comparison of original and retuned FEH statistical models for QMED

The scatter graph in Figure 11 plots QMED (modelled) on the y-axis (from 0.01 to 100) against QMED (gauged) on the x-axis (from 0.01 to 100) for catchments up to 25km^2 (marked with circles) and catchments from 25km^2 to 40.9km^2 (marked with crosses) for the three parameterisations of the existing FEH statistical QMED equation:

- original (marked in black)
- re-tune <25km^2 (marked in orange)
- re-tune <40.9km^2 (marked in green)

Figure 12 compares estimates of QMED given by the FEH equation retuned on catchments under 25 km^2 (Equation 16) against model outputs from ReFH 2 (version 2.2). Good agreement is shown for most catchments, with QMED ranging from approximately 1 to 15 m^3/s. However, the retuned FEH equation shows smaller errors relative to gauged QMED for the very smallest values of QMED and ReFH 2.2 shows smaller errors for the very largest values of QMED.

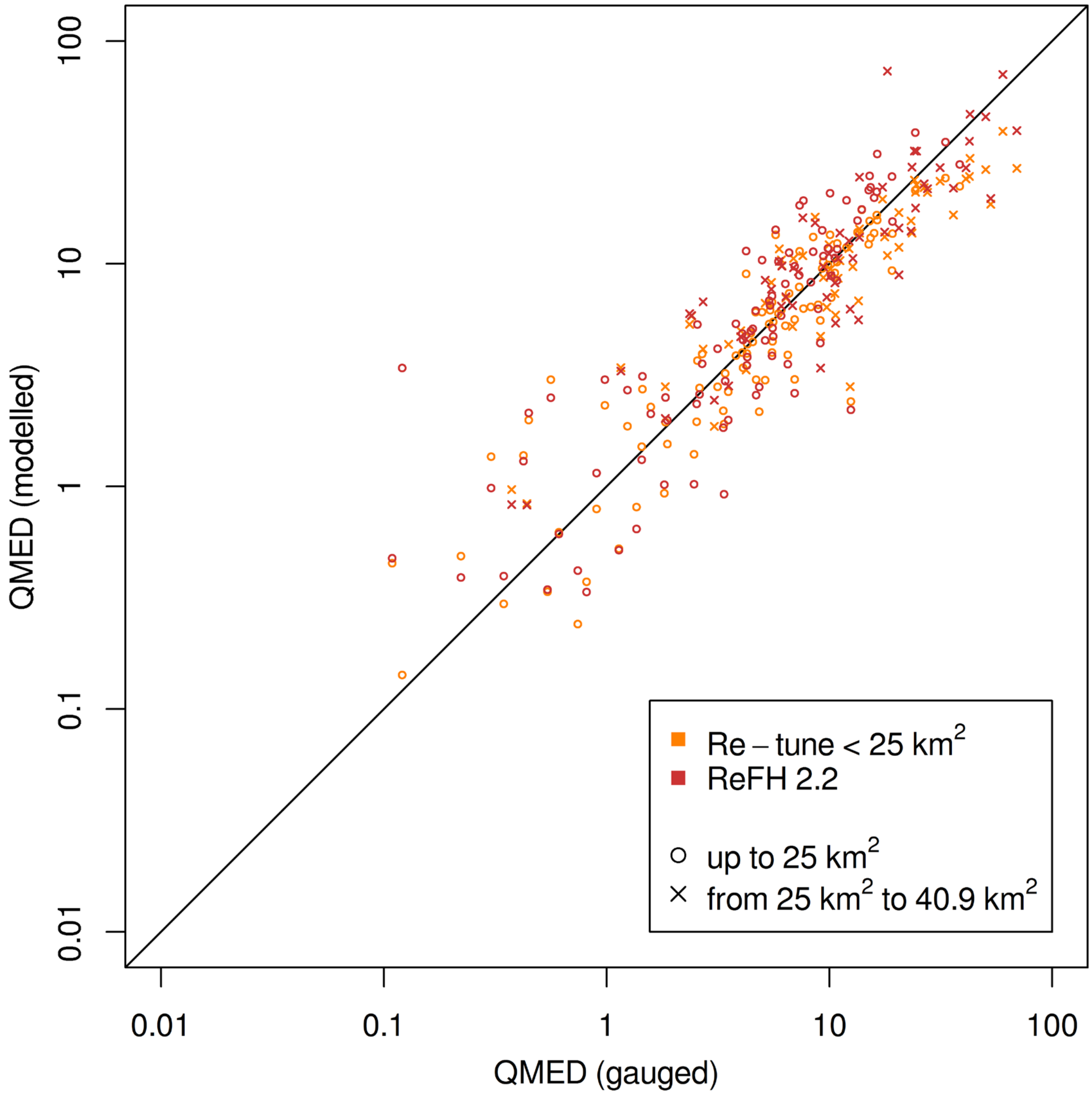

Figure 12 - Comparison of QMED estimated using retuned FEH equation and ReFH2.2

The scatter graph in Figure 12 plots QMED (modelled) on the y-axis (from 0.01 to 100) against QMED (gauged) on the x-axis (from 0.01 to 100) for catchments up to 25km^2 (marked with circles) and catchments from 25km^2 to 40.9km^2 (marked with crosses) for:

- re-tuned catchments <25km^2 (marked in orange)
- model outputs from ReFH 2 2.2 (marked in red)

Finally, Figure 13 shows the variation in retuned parameter values for each of the 93 cases where one catchment is excluded from the optimisation procedure on catchments under 25 km^2, as well as the corresponding sample fse. This shows that parameter values and sample fse are largely insensitive to the catchments selected – the largest bin in each sub-plot always corresponds to the parameter value given in Equation 16, and there is very little variation in values that do not fall into that bin. However, there are three outliers for the FARL parameter value: the two 'upper' outliers corresponding to excluding the catchment with the lowest and second-lowest FARL value (0.727 and 0.780) and the 'lower' outlier corresponding to excluding the catchment with the fifth-lowest FARL value (0.907). Therefore, in common with the existing FEH equation, it is recommended that the retuned FEH equation is not used as a main QMED estimation method in catchments with FARL < 0.9. Similarly, the three lowest and three highest outliers on the BFIHOST histogram correspond to excluding six of the nine most permeable catchments. While this does show that care should be taken in very permeable catchments (BFIHOST > 0.8), it does not indicate a bias in Equation 16 relative to BFIHOST, as the changes to the recalibrated BFIHOST parameter are evenly split between increasing and decreasing. Therefore, if a high-BFIHOST catchment is deliberately chosen as a donor for another, there is a risk that one showing a large positive residual is used to modify one with a large negative residual, or vice versa. This is discussed in more detail in Section 7.

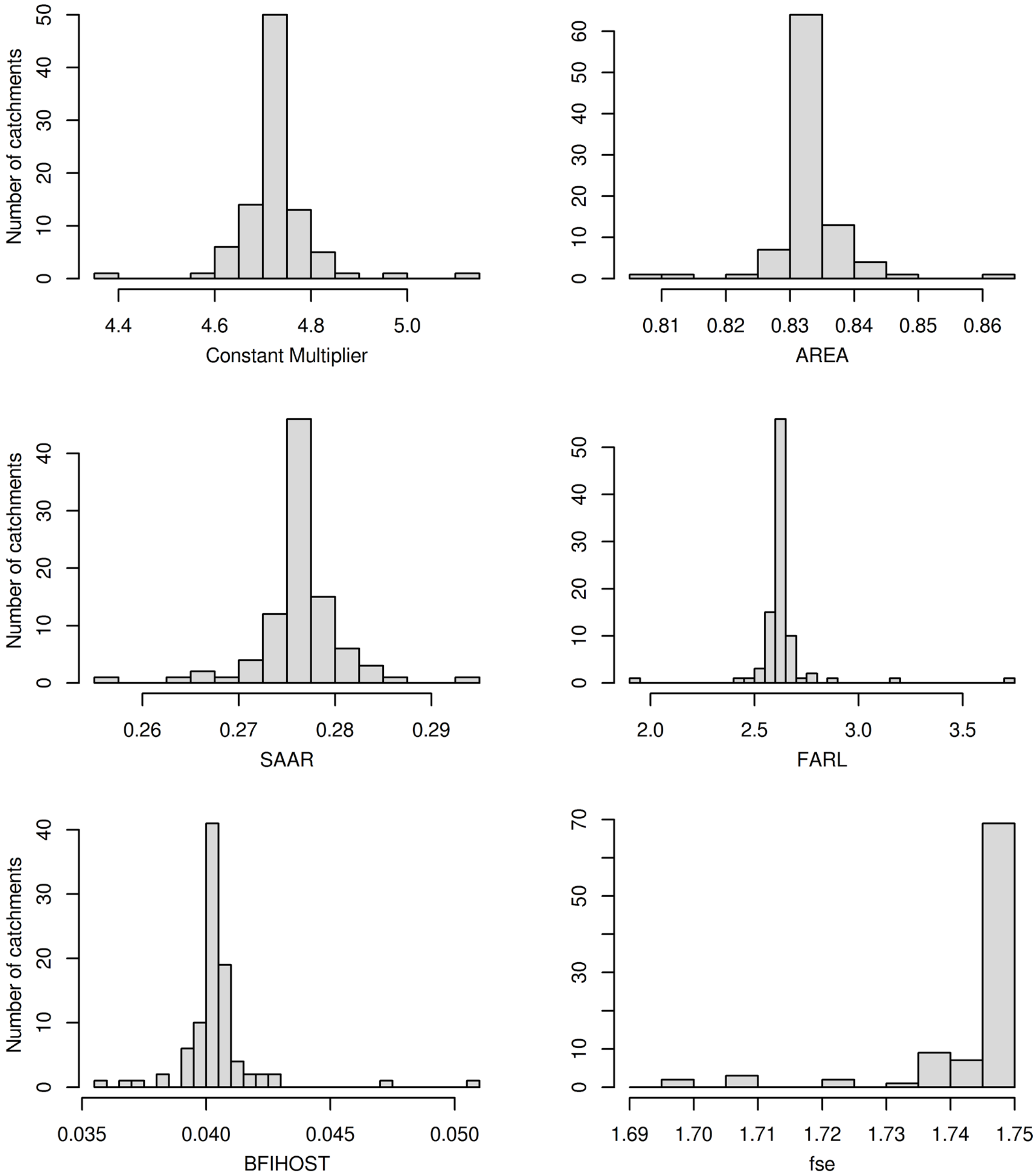

Figure 13 - Retuned (< 25 km^2) FEH parameter values when excluding each catchment in turn

The six bar charts in Figure 13 show the variation in retuned parameter values when one catchment is excluded:

- Top-left bar chart: Constant Multiplier on the x-axis (from 4.4 to 5.0). The y-axis shows the number of models with that value of constant initial multiplier, from 0 to 50, when the FEH QMED model is retuned in 93 cases, where each case excludes one of the 93 catchments from the calibration data set.
- Top-right bar chart: AREA on the on the x-axis (from 0.81 to 0.86). The y-axis shows the number of models with that value of exponent applied to AREA, from 0 to 60, when the FEH QMED model is retuned in 93 cases, where each case excludes one of the 93 catchments from the calibration data set.

- Middle-left bar chart: SAAR on the x-axis (from 0.26 to 0.29). The y-axis shows the number of models with that value applied to SAAR,from 0 to 40, when the FEH QMED model is retuned in 93 cases, where each case excludes one of the 93 catchments from the calibration data set.
- Middle-right bar chart: FARL on the x-axis (from 2.0 to 3.5). The y-axis shows the number of model with that value of exponent applied to FARL,from 0 to 50, when the FEH QMED model is retuned in 93 cases, where each case excludes one of the 93 catchments from the calibration data set.
- Bottom-left bar chart: BFIHOST on the x-axis (from 4.4 to 5.0). The y-axis shows the number of model with that value applied to BFIHOST,from 0 to 40, when the FEH QMED model is retuned in 93 cases, where each case excludes one of the 93 catchments from the calibration data set.
- Bottom-right bar chart: fse on the x-axis (from 1.69 to 1.75). The y-axis shows the number of model with that value of fse,from 0 to 70, when the FEH QMED model is retuned in 93 cases, where each case excludes one of the 93 catchments from the calibration data set.

5. New QMED model for small catchments

Although retuning the FEH equation has improved QMED estimation in small catchments, it is possible that a model structure with different catchment descriptors may offer further benefits. Here, the procedure of building a new model is described, using the 93 catchments up to 25 km^2 for calibration. This new regression model for ln(QMED) was built up using forward stepwise regression, one catchment descriptor at a time, until adding any new catchment descriptors could not improve the model further.

5.1 Catchment descriptors

The existing FEH statistical model for QMED tested only a limited selection of catchment descriptors during model development. In this study, all 25 standard catchment descriptors, plus one derived specifically for this study (RMED13-6H, six-hour median annual rainfall from FEH13 rainfall model), were considered initially. From these, FPLOC was excluded, as it is incalculable for catchments with limited floodplain cover. Catchment descriptors related to urbanisation were excluded, as these are incalculable for rural catchments. In addition, all RMED descriptors were updated to their FEH13 values for this model building procedure. To avoid confusion with the existing catchment descriptors, these are referred to as RMED13 descriptors.

Certain variables in the FEH model were transformed in order to reduce patterns in the model residuals against that variable. In this study, all catchment descriptors were made available in six forms: x, ln(x), e^x, 1/x, x^2 and $\sqrt{x}$. Before being transformed, some catchment descriptors were scaled, specifically:

- all SAAR and DPSBAR values were divided by 1,000 to allow exponentials to be taken (necessitated by data limitations)
- all SPRHOST values were divided by 100, that is, scaled to the same [0,1] range as BFIHOST, FPEXT and other common catchment descriptors
- all FPEXT and FPDBAR values were increased by 1 to allow logarithms to be taken
- all ASPBAR values were divided by 360, that is, scaled to the range [0,1], and then increased by one to allow logarithms to be taken
- all ALTBAR values were divided by 100, positioning all values in the approximate range 0.2 to 6.6, to allow exponentials to be taken
- RMED13-1H, RMED13-6H, RMED13-1D and RMED13-2D were divided by 10, 50, 100 and 100 respectively, positioning all values in the range 0.3 to 1.7, to allow exponentials to be taken

5.2 Model structure exploration

As stated, the regression model for ln(QMED) was built in stages, with stage 0 being just a constant. At each subsequent stage n, the ML estimation was run with the models from the previous stage, adding each catchment descriptor in turn to those models. The catchment descriptors that minimised the log-likelihood function by the most over the stage n – 1 model were then added to that model to give stage n models. This process continued until adding a new model stage either had negligible effect on the log-likelihood function, resulted in an increase in model structural error or resulted in collinearity between a transformed parameter and either another transformed parameter or a constant value.

Table 6 summarises model development, where the column ℓ reports the negative log-likelihood function. The branch notation is used to relate models with more parameters to the models with fewer parameters on which they were built. For example, branches AA, AB and AC are built by adding a second parameter to the one-parameter branch A; for each branch, the last letter is assigned alphabetically in order of negative log-likelihood function.

Table 6 - Development of small-catchment QMED regression model form

Stage	Branch	Variable(s)	ℓ	fse
1	A	ln(AREA)	50.32	3.474
1	B	√(SPRHOST/100)	57.20	3.096
1	C	$BFIHOST^2$	58.04	3.121
2	AA	ln(AREA), √(SPRHOST/100)	8.64	1.923
2	AB	ln(AREA), $BFIHOST^2$	13.59	2.026
2	AC	ln(AREA), 1/RMED13-1D	24.58	2.431
2	BA	√(SPRHOST/100), ln(AREA)	8.64	1.923
2	BB	√(SPRHOST/100), ln(LDP)	27.01	2.242
2	BC	√(SPRHOST/100), ln(DPLBAR)	27.92	2.258
2	CA	$BFIHOST^2$, ln(AREA)	13.59	2.026
2	CB	$BFIHOST^2$, ln(DPLBAR)	31.54	2.365
2	CC	$BFIHOST^2$, ln(LDP)	32.34	2.380
3	AAA	ln(AREA), √(SPRHOST/100), (1000/SAAR)	−7.41	1.731
3	AAB	ln(AREA), √(SPRHOST/100), (100/RMED13-2D)	−6.56	1.722
3	AAC	ln(AREA), √(SPRHOST/100), (100/RMED13-1D)	−6.47	1.732
3	ABA	ln(AREA), $BFIHOST^2$, (100/RMED13-2D)	−11.06	1.685
3	ABB	ln(AREA), $BFIHOST^2$, (100/RMED13-1D)	−9.94	1.696
3	ABC	ln(AREA), $BFIHOST^2$, (1000/SAAR)	−8.54	1.711

Stage	Branch	Variable(s)	ℓ	fse
4	ABAA	ln(AREA), $BFIHOST^2$, (100/RMED13-2D), $(1+ASPBAR/360)^2$	−14.56	1.651
4	ABAB	ln(AREA), $BFIHOST^2$, (100/RMED13-2D)), $FARL^2$	−13.97	1.657
4	ABAC	ln(AREA), $BFIHOST^2$, (100/RMED13-2D), (1/DPLBAR)	−12.54	1.656
4	ABBA	ln(AREA), $BFIHOST^2$, (100/RMED13-1D), $(1+ASPBAR/360)^2$	−14.11	1.669
4	ABBB	ln(AREA), $BFIHOST^2$, (100/RMED13-1D), $FARL^2$	−12.61	1.683
4	ABBC	ln(AREA), $BFIHOST^2$, (100/RMED13-1D), (1000/DPSBAR)	−11.68	1.658
4	ABCA	ln(AREA), $BFIHOST^2$, (1000/SAAR), $FARL^2$	−11.24	1.683
4	ABCB	ln(AREA), $BFIHOST^2$, (1000/SAAR), (100/RMED13-2D)	−11.10	1.685
4	ABCC	ln(AREA), $BFIHOST^2$, (1000/SAAR), $(1+ASPBAR/360)^2$	−10.58	1.703

Note: although SAAR and all four RMED descriptors are all correlated (Figure 5) and arguably ‘similar’, one key question for this research is to discover if rainfall descriptors based on shorter timescales have more influence on QMED in small catchments than SAAR does. Therefore, branches differing only in the selected rainfall descriptor were not immediately discontinued, as was the case with branches containing DPLBAR (mean drainage path length) or LDP (longest drainage path length).

The following branches were selected for further investigation:

- Stage 1 (A, B, C).
- Stage 2 (AA, AB).
- Stage 3 (ABA, ABB and ABC).

As would be expected, AREA is shown here to be the most likely single descriptor controlling QMED, followed by HOST descriptors. By stage 3, all models are of a form containing ln(AREA), one HOST descriptor and one rainfall descriptor. By stage 4, all models contain the exact terms ln(AREA) and $BFIHOST^2$, as used in the existing FEH model, as well as the inverse of a rainfall descriptor and, in three cases, FARL. The fact that different rainfall descriptors can be selected to result in almost equally good equations demonstrates that all of them are strongly correlated and suggests that there is no clear 'best' among them. As will be demonstrated in Section 5.5, it also suggests that the order in which rainfall descriptors are selected, from most to least likely, may be influenced by the selection of calibration catchments.

SPRHOST is shown here to be more likely than BFIHOST to describe variations in QMED, unless used together with a rainfall descriptor. However, SPRHOST was deliberately excluded when developing the existing FEH model because BFIHOST was found to be a more efficient descriptor of hydrological soil properties in earlier work (Kjeldsen and others, 2005). More specifically, the regression equation between BFI and HOST class is based on 575 catchments, while the regression between SPR and HOST class is based on 170 catchments. However, even BFIHOST is matched inconsistently to gauged BFI values, and both BFIHOST and SPRHOST are based on coarse mapped grids that distribute HOST class fractions uniformly over 1 km grid cells – comparable in size to the smallest catchments in the calibration data set.

5.3 Proposed model

Model development was stopped at stage 4, where the improvement in fse over the best stage 3 model was marginal, and the transformed fourth descriptor acted over only a small range of values in the calibration data set.

Model 3ABA (black line).

The six line graphs in Figure 14 evaluate potential new QMED models for small catchments:

- Model 3ABA (dark grey line)
- Model 3ABB (dark green line).
- Model 4ABAB (light grey line).
- Model 4ABBC (light green line).

Each graph represents how the QMED value generated by each potential QMED model varies with:

- Top-left graph: AREA.
- Top-right graph: BFIHOST.
- Middle-left graph: RMED13-2D.
- Middle-right graph: RMED13-1D.
- Bottom-left graph: FARL.
- Bottom-right graph: DPSBAR.

Each grey vertical line on each graph indicates one calibration catchment.

Figure 14 shows that AREA and BFIHOST have similar influence over QMED in each model, therefore suggesting that these contribute the most information on estimated QMED. RMED13-2D has similar influence in models 3ABA and 4ABAB, and the presence or absence of FARL[2] does not change any of the other fitted coefficients, showing that FARL is not collinear with AREA, BFIHOST or RMED13-2D. However, the influence of RMED13-1D changes from model 3ABB to 4 ABBC, suggesting some collinearity between RMED13-1D and DPSBAR.

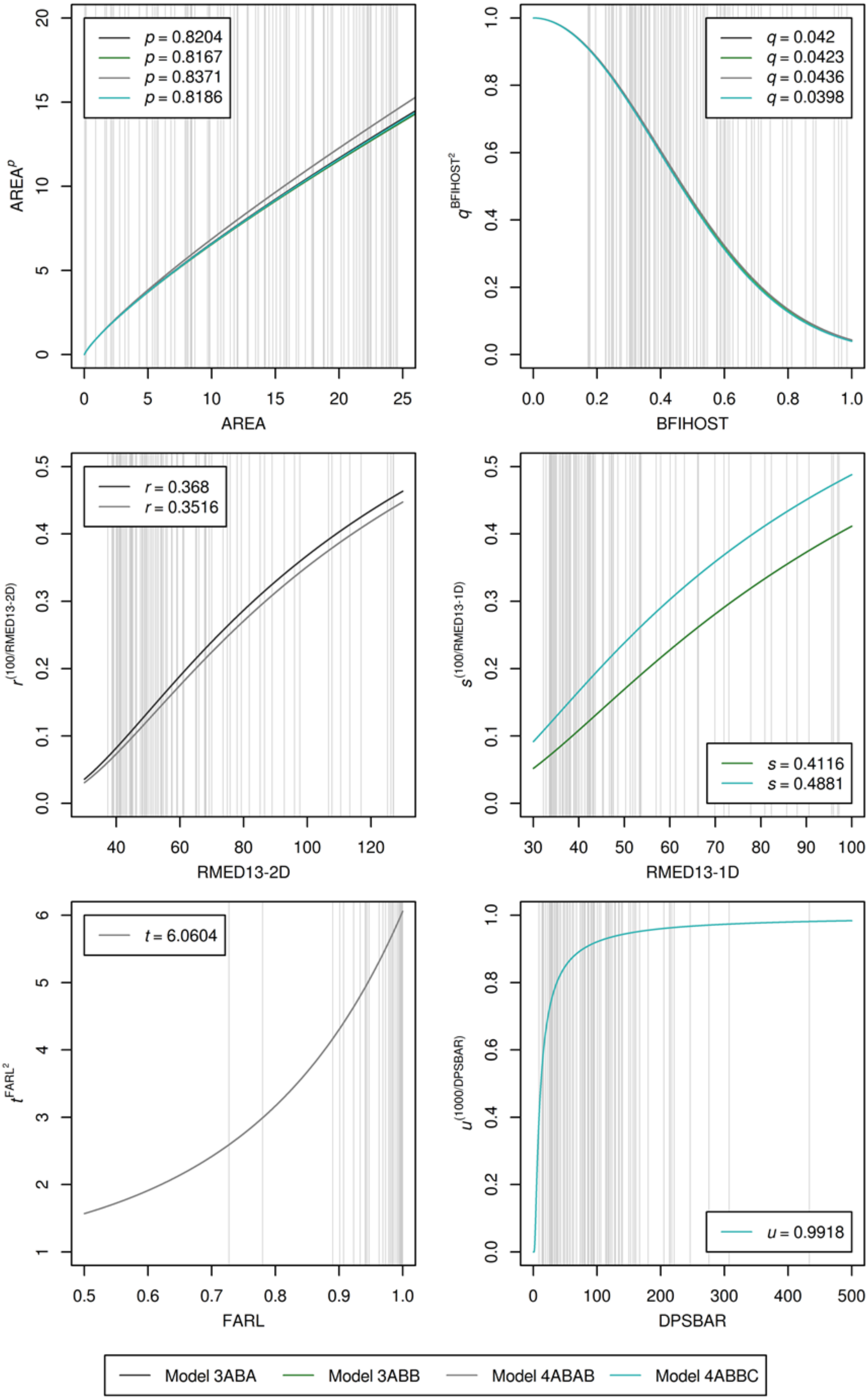

Figure 14 - Influence of catchment descriptors in proposed QMED models

The eight scatter graphs in Figure 15 show the residual plots of the four models against two catchment descriptors: FARL and DPSBAR. Only Model 4ABAB includes FARL, and it is the only model whose residuals are unbiased with respect to FARL. This suggests that the descriptor FARL is required to model QMED properly in small catchments. However, this conclusion is complicated by the compressed range of FARL values in the calibration data set and the resulting high leverage of the few calibration catchments with lower FARL values. Nevertheless, based on the evidence available, FARL is shown to exert influence over QMED – this conclusion is also physically sensible. As FARL is not strongly correlated with other descriptors, incorporating it in the QMED model should help in QMED modelling in catchments with FARL up to ~0.9 without having a negative effect in catchments with FARL above 0.9.However, all model residuals are unbiased with respect to DPSBAR, despite only Model 4ABBC including DPSBAR, suggesting that DPSBAR is not a necessary descriptor to include in a small-catchment QMED model.

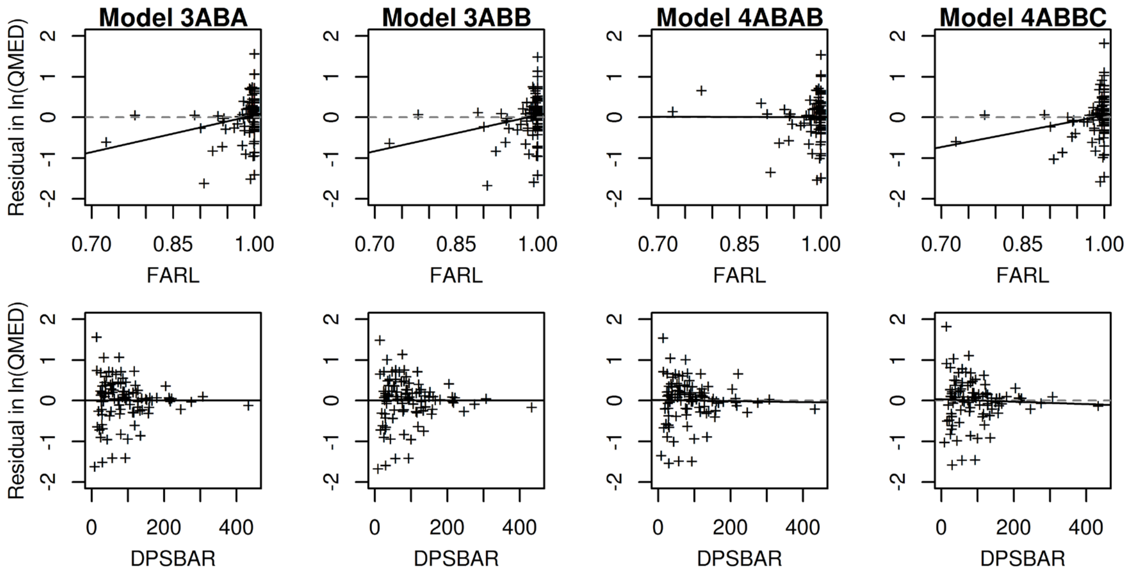

Figure 15 - Residual (ln(QMED$_{OBS}$) – ln(QMED$_{MOD}$)) plots for candidate QMED models

Equation 18 shows model 4ABAB, which is selected for modelling QMED in small catchments. As reported in Table 6, model negative log-likelihood is –13.97 and model factorial standard error is 1.657, which is lower than that found for either retuned FEH QMED equation developed in Section 4.

Equation 18 – Model 4ABAB

$$QMED = 1.3274 AREA^{0.8371} 0.3516^{\frac{100}{RMED_{13-2D}}} 6.0604^{FARL^2} 0.0436^{BFIHOST^2}$$

- $\sigma_\eta^2 = 0.2549$
- $\phi = 3.0930$

Equation 18 is notable for containing three of the four descriptors used in the existing FEH statistical equation (AREA, BFIHOST and FARL), two in exactly the same form. The dissimilar descriptor (RMED13-2D) is highly correlated with SAAR, and is used with the same transformation. The structure of both this new model and MacDonald and Fraser's (2013) model implies that the processes governing floods in catchments of ~10 km^2 are not too different from those governing floods in larger catchments. Regression diagnostic plots for model 4ABAB are shown in Figure 16.

- Left graph: ordered model residuals against normally distributed values (Q-Q plot)
- Middle graph: model residuals against modelled ln(QMED)
- Right graph: modelled ln(QMED) against observed ln(QMED)

These demonstrate that the model residuals are approximately normally distributed, when the limited sample size is considered, and that the model residuals are unrelated to either the gauged or modelled value of ln(QMED).

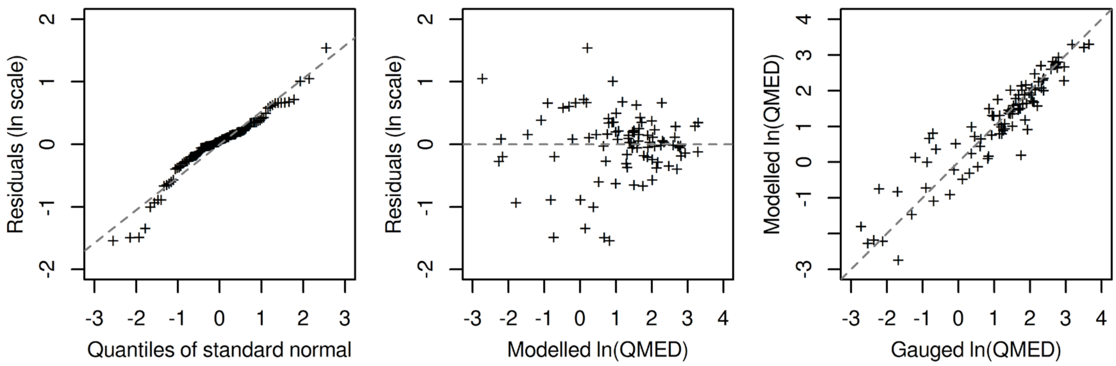

Figure 16 - Regression diagnostic plots for QMED model 4ABAB

A larger set of model residual plots is shown in Figure 17 for model 4ABAB only, plotting the residuals from model 4ABAB against nine catchment descriptors. Figure 17 demonstrates that the model residuals are neither strongly correlated with any of the catchment descriptors in the model (AREA, RMED13-2D, BFIHOST, FARL), nor with several other descriptors (DPSBAR, FPEXT, SPRHOST, PROPWET, SAAR). Although the limited number of permeable catchments (17) meant that they could not be considered separately, Figure 17 shows that model 4ABAB is unbiased with respect to BFIHOST.

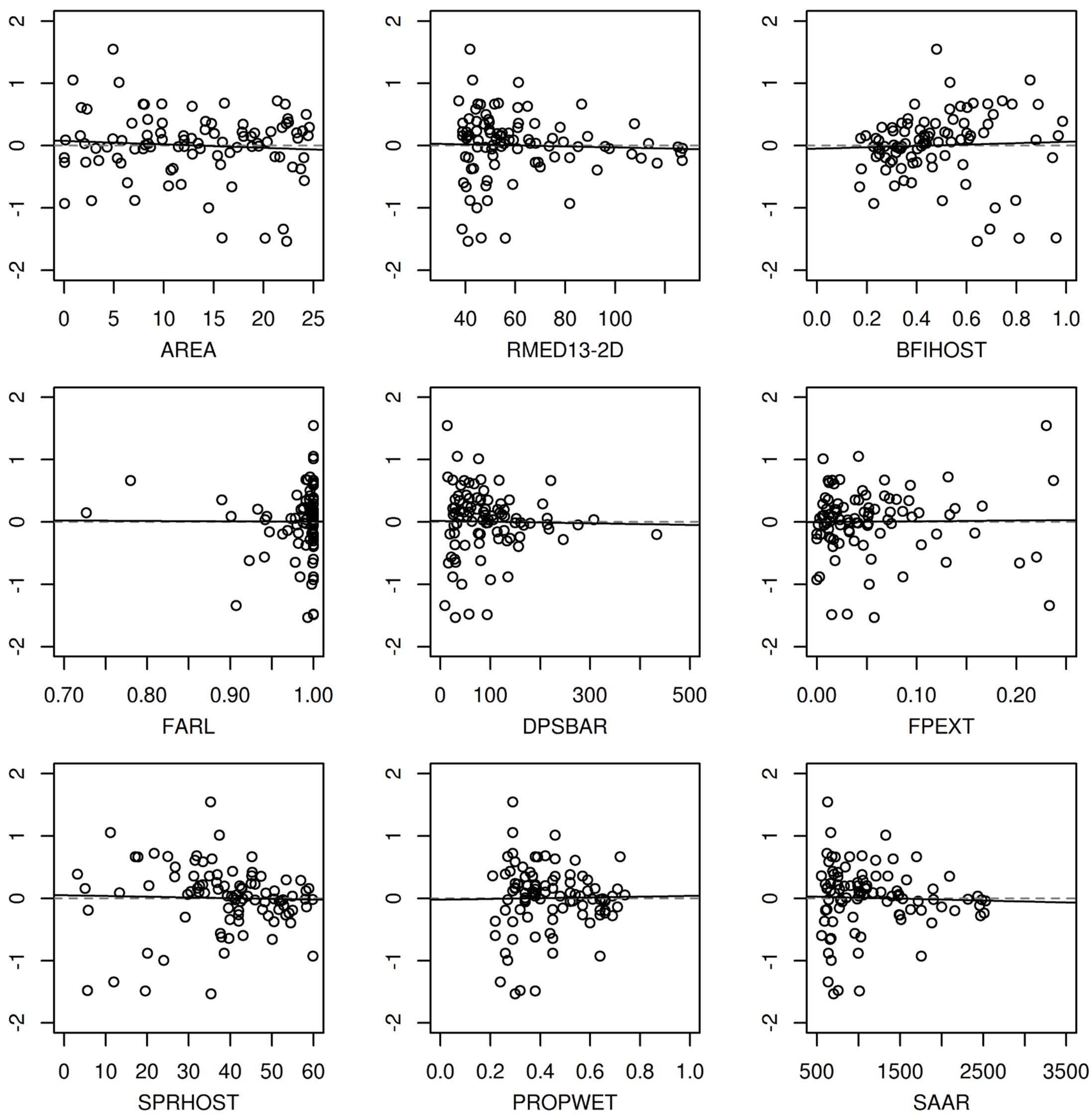

Figure 17 - Residual ($\ln(QMED_{OBS}) – \ln(QMED_{MOD})$) plots for QMED model 4ABAB

5.4 Verification on intermediate-size catchments

The proposed model (4ABAB) was calibrated to 93 catchments under 25 km^2. As the data set described in Section 2 actually contains 152 catchments up to 40.9 km^2, there are 59 additional catchments available to check how Equation 18 performs on catchments slightly above the intended upper limit on area. Figure 18 demonstrates that the model selected for small catchments is also valid for those up to 40.9 km^2.

In the left panel of Figure 18, the mean residual and 95% confidence interval is plotted in eight bins, where each bin contains one-eighth of the 152 study catchments after sorting by AREA. This shows that the model tends to slightly underestimate QMED on average in catchments between 25 and 40.9 km^2: the three highest bins show positive residuals (model underestimation), while the five lowest bins show mixed over and underestimation. The right panel of Figure 18 compares ln(QMED) as modelled by the existing FEH equation and model 4ABAB against the observed value of ln(QMED) in the 152 catchments. This does not appear to show that either one model is more accurate than the other in terms of residual. Table 7 compares the performance statistics of model 4ABAB, the existing FEH statistical QMED equation and the FEH statistical equation retuned to catchments under 25 km^2, on the extended data set of catchments up to 40.9 km^2. While this shows model 4ABAB to be slightly better at fitting the data, the difference is considered too marginal to justify a new model structure, particularly one that includes a model output (RMED13-1D). The range of FARL values included in the small catchments data set is also considered too compressed to justify a change to the transformation applied to FARL, especially as that transformation was found suitable for a larger data set with many more low-FARL catchments.

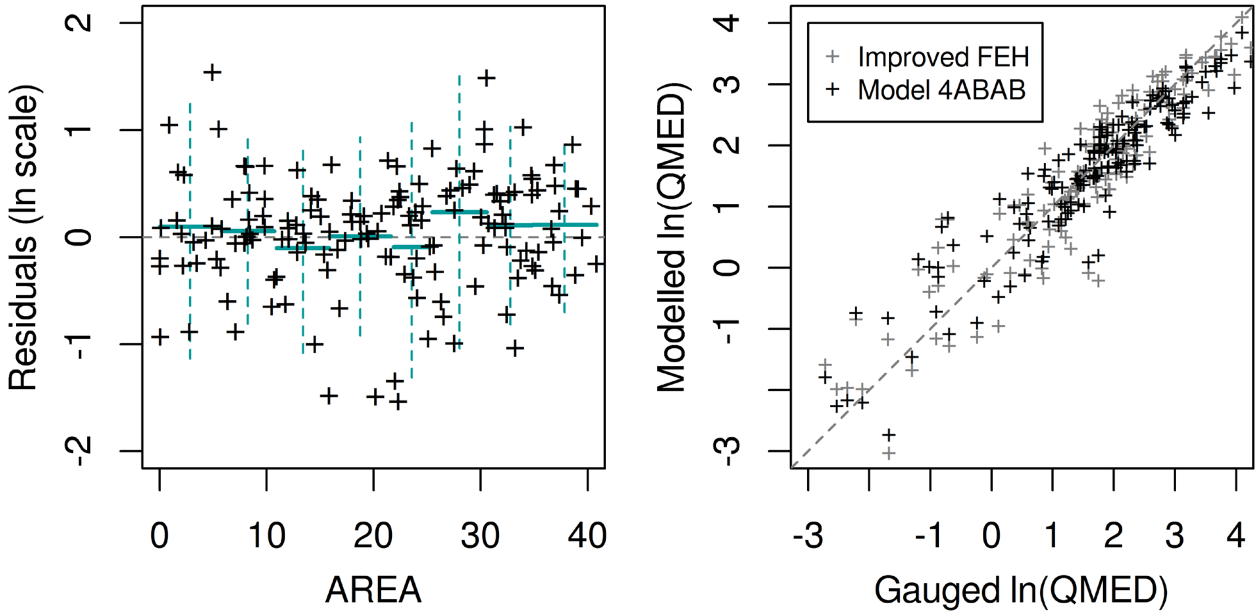

Figure 18 - Verification plots for model 4ABAB on catchments up to 40.9 km^2

Table 7 - Performance of new and existing models for QMED, tested on 152 catchments up to 40.9 km^2

Model	model fse $\exp(\sigma_\eta)$	R^2	RMSE*	sample fse $\exp(s)$
4ABAB	1.657	0.8604	0.5401	1.716
Existing FEH	1.431	0.8480	0.5635	1.757
Existing FEH Retuned, <25 km^2	1.701	0.8574	0.5457	1.726

Note: Sum of squares of residuals divided by number of degrees of freedom (152 – 4 parameters – 1 = 147).

Evidence from Table 7 and Figure 18 suggests that the performance of model 4ABAB reduces as catchment size increases, and that the model is unlikely to be suitable for much larger catchments.

Based on the trade-off between model performance and model structure, it is recommended using the retuned FEH equation for small catchments under 25 km^2 and the existing FEH equation for catchments over 40.9 km^2. For estimating QMED in catchments 25 and 40.9 km^2, a linear interpolation method, detailed in Appendix A, is intended to provide a smooth transition from the retuned FEH equation for small catchments to the existing FEH equation for larger catchments.

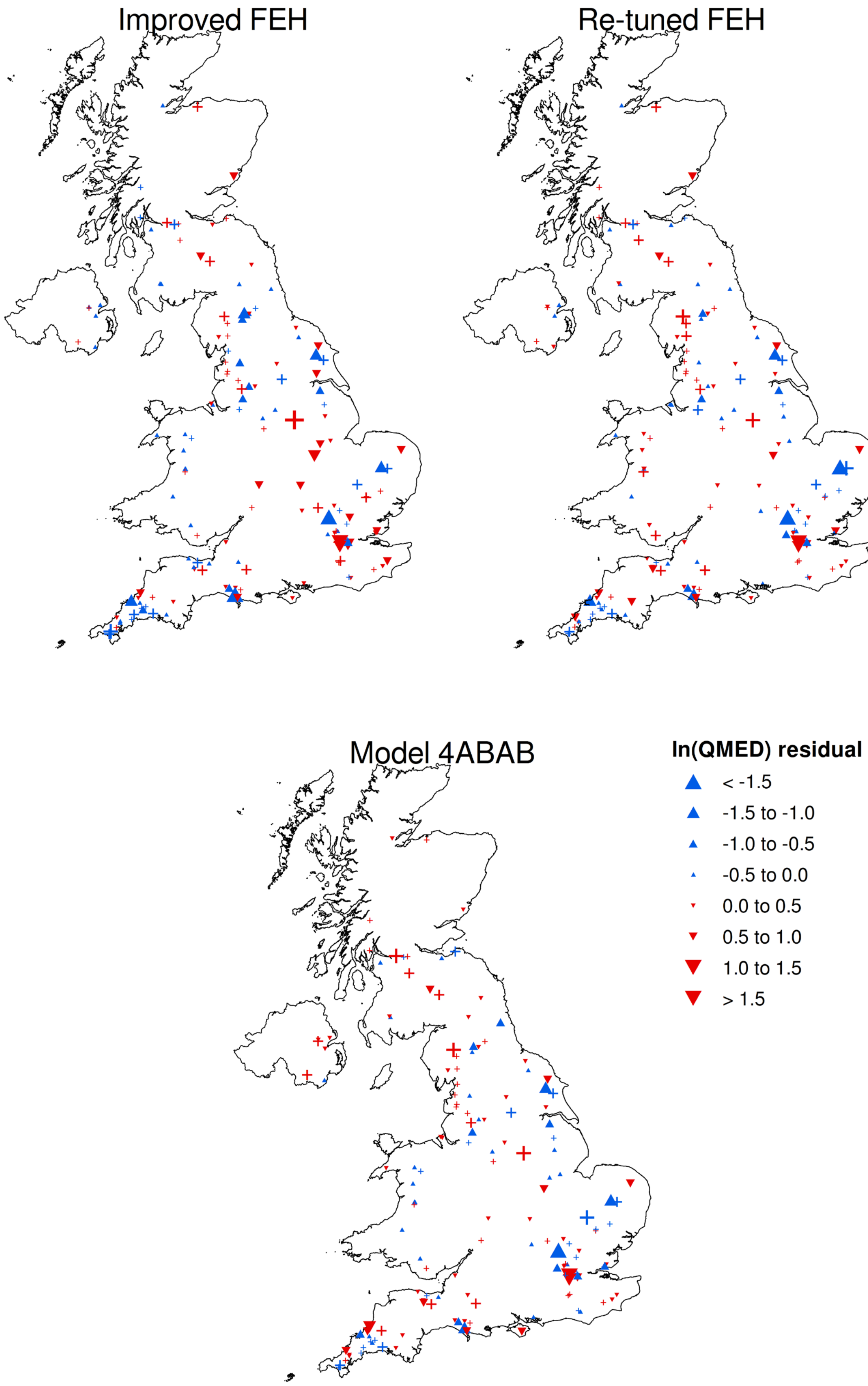

Figure 19 - Model residuals ($\ln(QMED_{OBS}) - \ln(QMED_{MOD})$) across UK for existing FEH model, retuned (25 km^2) FEH model and model 4ABAB

The three maps in Figure 19 show the spatial distribution of model residuals ($\ln(QMED_{OBS})$ – $\ln(QMED_{MOD})$) across the UK:

- Top-left map: improved FEH.
- Top-right map: re-tuned FEH.
- Bottom map: model 4ABAB.

The legend shows:

- <-1.5 (very large blue triangles).
- -1.5 to -1.0 (large blue triangles).
- -1.0 to -0.5 (medium blue triangles).
- -0.5 to 0.0 (small blue triangles).
- 0.0 to 0.5 (small red triangles).
- 0.5 to 1.0 (medium red triangles).
- 1.0 to 1.5 (large red triangles).
- >1.5 (very large red triangles).

5.5 Verification on ‘extended’ data set

MacDonald and Fraser (2013), when developing an ordinary least-squares regression model for QMED, used a total of 135 catchments up to 25 km^2. 63 of these catchments are shared with the calibration data set of 93 that was used to develop the new small-catchment QMED model presented in Equation 18. These have been formally classified as suitable for estimating QMED. The remaining 72 have either not been classified by the measuring authority or have been classified as unsuitable for estimating QMED. Lack of a rating, however, does not necessarily indicate low data quality. In this subsection, model development is explored for a data set of 132 catchments up to 25 km^2: the 93 used in developing Equation 18 plus a further 39 used by MacDonald and Fraser. Not all of MacDonald and Fraser’s data set can be used, as not all of their calibration catchments have associated AMAX records available, which are necessary to estimate sampling error. In addition, eight of MacDonald and Fraser’s catchments were explicitly rejected from this study. Table 8 summarises model development with these catchments, similarly to Table 6. Identical procedures and model selection criteria were used.

Table 8 - Development of small-catchment QMED regression model form using 'extended' data set

Stage	Branch	Variable(s)	ℓ	fse
1	A	ln(AREA)	60.11	3.501
1	B	√(SPRHOST/100)	78.26	2.995
1	C	$BFIHOST^2$	78.69	3.004
2	AA	ln(AREA), √(SPRHOST/100)	7.73	1.899
2	AB	ln(AREA), $BFIHOST^2$	11.98	1.983
2	AC	ln(AREA), (100/RMED13-2D)	34.64	2.293
2	BA	√(SPRHOST/100), ln(AREA)	7.73	1.899
2	BB	√(SPRHOST/100), ln(DPLBAR)	34.83	2.190
2	BC	√(SPRHOST/100) ln(LDP)	35.23	2.183
2	CA	$BFIHOST^2$, ln(AREA)	11.98	1.983
2	CB	$BFIHOST^2$, ln(DPLBAR)	40.63	2.026
2	CC	$BFIHOST^2$, ln(LDP)	41.86	1.969
3	AAA	ln(AREA), √(SPRHOST/100), 1000/SAAR	−14.63	1.695
3	AAB	ln(AREA), √(SPRHOST/100), 100/RMED13-2D	−14.59	1.696
3	AAC	ln(AREA), √(SPRHOST/100), 100/RMED13-1D	−13.57	1.703
3	ABA	ln(AREA), $BFIHOST^2$, 100/RMED13-2D	−23.00	1.640
3	ABB	ln(AREA), $BFIHOST^2$, 100/RMED13-1D	−22.06	1.645

Stage	Branch	Variable(s)	ℓ	fse
3	ABC	ln(AREA), $BFIHOST^2$, ln(SAAR/1000)	−21.22	1.649
4	ABAA	ln(AREA), $BFIHOST^2$, 100/RMED13-2D, exp(ASPBAR)	−26.04	1.620
4	ABAB	ln(AREA), $BFIHOST^2$, 100/RMED13-2D, $ASPVAR^2$	−25.42	1.622
4	ABAC	ln(AREA), $BFIHOST^2$, 100/RMED13-2D, $FARL^2$	−24.83	1.628
4	ABBA	ln(AREA), $BFIHOST^2$, 100/RMED13-1D, $ASPBAR^2$	−25.41	1.623
4	ABBB	ln(AREA), $BFIHOST^2$, 100/RMED13-1D, $ASPVAR^2$	−24.27	1.629
4	ABBC	ln(AREA), $BFIHOST^2$, 100/RMED13-1D, $FARL^2$	−23.68	1.635
4	ABCA	ln(AREA), $BFIHOST^2$, ln(SAAR/1000), (1000/DPSBAR)	−23.69	1.634
4	ABCB	ln(AREA), $BFIHOST^2$, ln(SAAR/1000), $ASPVAR^2$	−23.51	1.633
4	ABCC	ln(AREA), $BFIHOST^2$, ln(SAAR/1000), $ASPBAR^2$	−23.46	1.634

The following branches were selected for further investigation:

- Stage 1 (A, B, C).
- Stage 2 (AA, AB).
- Stage 3 (ABA, ABB and ABC).

With this ‘extended’ data set of small catchments, the best performing equations describing QMED are largely identical to those found using the 93-catchment ‘high-quality’ data set. Equations containing exact terms in common with the existing FEH equation, that is, ln(AREA) and $BFIHOST^2$, very quickly establish themselves as highly successful descriptive forms. By stage 3, all equations comprise ln(AREA), one HOST term and one rainfall term. In common with the earlier analysis in Section 5.2, performance of the stage

3 models is only slightly affected by the choice of rainfall descriptor. This reinforces the suggestion that there is no clear 'best' among the rainfall descriptors and that, given another different set of calibration catchments, the ranking of the five rainfall descriptors, from most to least likely, could be different.

6. Using gauged baseflow index and flow statistics in estimating QMED

It is stated in Section 5 that BFIHOST and SPRHOST are based on coarse mapped grids. These are the HOST (Hydrology of soil types) grids, which have a cell size of 1 km^2. Within each cell, the percentage of soil matching each of 29 different classes is stored to the nearest integer. Each soil class is assigned one BFIHOST and one SPRHOST value, and the cell-average BFIHOST and SPRHOST values are calculated, based on the percentage cover of each class in each 1 km grid cell. As a result, the BFIHOST and SPRHOST values are assumed uniform over each square kilometre. While this approximation has limited effect in large catchments, problems may arise in small catchments, particularly if their shapes do not align neatly with the grids.

Gauged estimates of BFI can be derived from daily mean streamflow records using the method of Gustard and others (1992), as can estimates of the flow duration curve. Although gauged BFI values derived from less than five years of flow are regarded as provisional, gauged BFI is more stable than other low flow metrics and, by definition, gives a more accurate indication of a river's baseflow than the BFIHOST value.

The National River Flow Archive (NRFA) holds gauged values of BFI and flow duration curve points *Q*95, *Q*70, *Q*50 and *Q*10 (daily mean flow exceeded on 95%, 70%, 50% and 10% of days) for 61 of the 93 catchments that were used for model calibration in Section 5. These were last updated on 30 September 2015, for stations that were still open then. A total of 46 catchments have $URBEXT_{2000}$ < 0.03 and are therefore essentially rural.

It is therefore suggested that some of these statistics may be useful in estimating QMED. It should be noted that these statistics were deliberately neglected in Section 5 because they are only available for approximately two-thirds of the catchments in this study data set – requiring these statistics would considerably reduce the quantity of calibration data. Additionally, there is no FEH procedure for deurbanising the flow duration curve equivalent to that for deurbanising gauged QMED, so the influence of urbanisation is inseparable from the flow duration curve. Furthermore, gauged daily flows are required to be natural or naturalised, so that catchment-specific abstractions and discharges are excluded from general models. The naturalness of the gauged flows obtained for this study was not checked. Therefore, this section and any results should be regarded as a preliminary proof-of-concept rather than a concrete recommendation.

Table 9 shows the first three stages of model development for 46 essentially rural small catchments when gauged BFI, Q95, Q70, Q50 and Q10 are made available for selection in addition to all previous catchment descriptors. As some values of Q95 and Q70 are zero, a constant offset of 1 was added to all values of Q95 and Q70 before starting model development, to allow logarithms to be taken if necessary.

Equation 19 presents model 3AAA, which has the lowest model structural error. It is noted, however, that three stage 3 models are almost identical in both negative log-likelihood function and model structural error.

Equation 19 – Model 3AAA

$$QMED = 9.9812DPLBAR^{0.3196}Q_{10}^{0.6725}0.0503^{BFIHOST^2}$$

- $\sigma_\eta = 0.0693$
- $\phi = 0.6577$

Despite its low model structural error, Equation 19 is not recommended for using in small catchments due to the small size of the calibration data set.

Table 9 - Development of small-catchment QMED regression model form with additional gauged flow descriptors available

Stage	Branch	Variable(s)	ℓ	fse
1	A	ln(Q10)	−8.97	1.866
1	B	ln(Q50)	7.73	2.463
1	C	ln(SPRHOST/100)	9.76	2.100
2	AA	ln(Q10), $BFIHOST^2$	−28.76	1.388
2	AB	ln(Q10), BFI^2	−25.47	1.437
2	AC	ln(Q10), ln(SPRHOST/100)	−20.98	1.488
2	BA	ln(Q50), BFI^2	−18.65	1.493
2	BB	ln(Q50), $BFIHOST^2$	−17.19	1.509
2	BC	ln(Q50), ln(Q10)	−12.04	1.725
2	CA	ln(SPRHOST/100),	−19.95	1.488
2	CB	ln(SPRHOST/100),	−14.62	1.560
2	CC	ln(SPRHOST/100),	−12.69	1.577
3	AAA	ln(Q10), $BFIHOST^2$, ln(DPLBAR)	−35.97	1.301
3	AAB	ln(Q10), $BFIHOST^2$, √AREA	−35.94	1.306
3	AAC	ln(Q10), $BFIHOST^2$, ln(LDP)	−35.41	1.304
3	ABA	ln(Q10), BFI^2, $PROPWET^2$	−36.04	1.304
3	ABB	ln(Q10), BFI^2, 1/(1+Q95)	−33.60	1.332
3	ABC	ln(Q10), BFI^2, DPLBAR	−32.16	1.339

Stage	Branch	Variable(s)	ℓ	fse
3	ACA	ln(Q10), ln(SPRHOST/100), √AREA	−29.86	1.361
3	ACB	ln(Q10), ln(SPRHOST/100), $BFIHOST^2$	−29.16	1.384
3	ACC	ln(Q10), ln(SPRHOST/100), ln(DPLBAR)	−27.79	1.379

The following branches were selected for further investigation:

- Stage 1 (A, B, C).
- Stage 2 (AA, AB, AC).

Table 9 shows that, for the 46 essentially rural catchments with gauged flow statistics, the most effective predictors of QMED are descriptors corresponding to points on the flow duration curve – highest first. This is not surprising, as QMED corresponds to one of the very highest instants on the flow duration curve; the FEH (volume 3, page 8) states that QMED might be only the 10^{th} or 11^{th} highest flow ever reached in a 13-year record. While AREA is the catchment descriptor that explains the most variation in QMED in the absence of gauged data, it is somewhat correlated with both QMED and Q10 (Figure 20). However, high flow levels, such as Q10, are directly related to the flow regime, whereas AREA isn't. Nevertheless, many stage 3 models feature AREA, DPLBAR or LDP as well as ln(Q10), thereby incorporating further, explicit information on catchment size or flow path length, in order to identify and model differences in QMED that cannot be explained through Q10 alone.

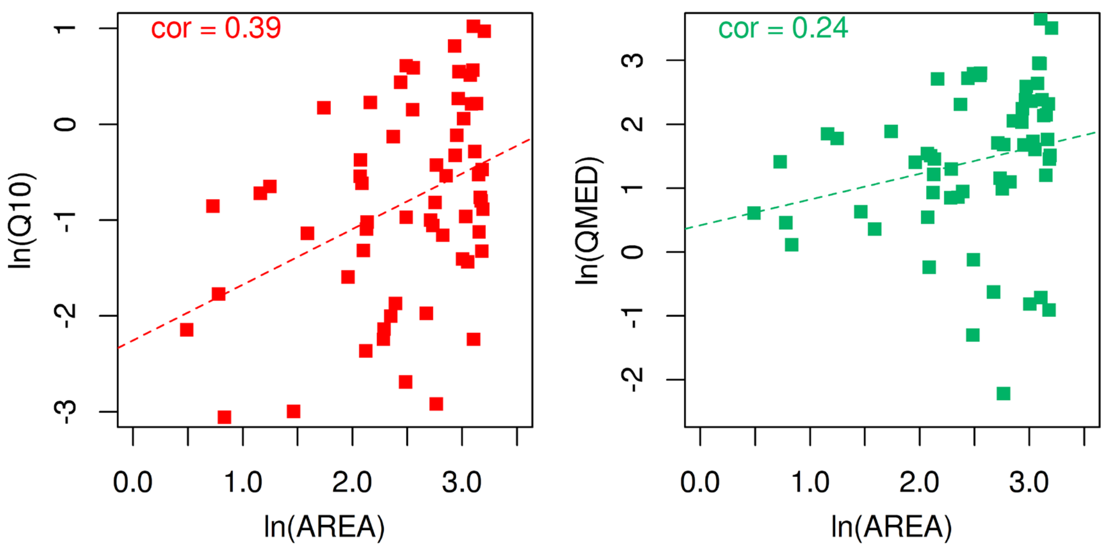

Figure 20 - Relationship between AREA, QMED and Q10

The two scatter graphs in Figure 20 show the relationship between AREA, QMED and Q10:

- Left-hand graph: this plots ln(AREA) (from 0.0 to 3.0 on the x-axis) against ln(Q10) (from -3 to 1 on the y-axis).
- Right-hand graph: this plots ln(AREA (from 0.0 to 3.0 on the x-axis) against ln(QMED) on the y-axis (from -2 to 3).

These show moderate correlations between ln(AREA) and either Q10 (cor = 0.39) or ln(QMED) (cor = 0.24). This demonstrates that both Q10 and ln(AREA) provide somewhat useful and independent information for estimating QMED: each is useful separately, but use of both together is advantageous over either one.

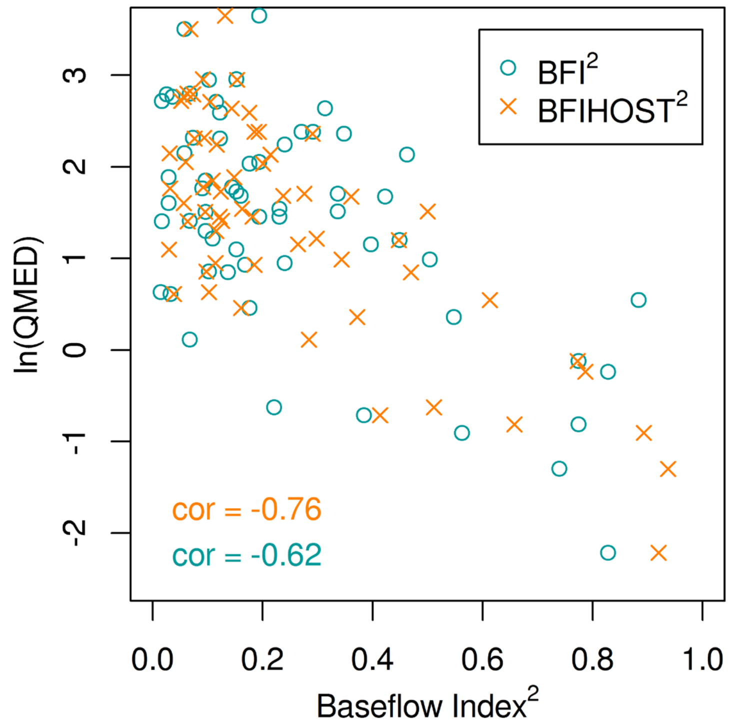

Figure 21 - Relationship between QMED and two different baseflow indices

The scatter graph in Figure 21 shows the relationship between QMED and two different baseflow indices:

- BFI^2 (green circles).
- $BFIHOST^2$ (orange crosses).

The x-axis plots the baseflow index^2 from 0.0 to 1.0. The y-axis plots ln(QMED) from -2 to 3.

After gauged flow statistics, gauged BFI and coarse-mapped SPRHOST and BFIHOST are the next most effective predictors of QMED. Perhaps unexpectedly, a model with √Q10 and $BFIHOST^2$ has a lower negative log-likelihood function than one with √Q10 and BFI^2 – the gauged baseflow index is less effective for estimating QMED than one estimated from soil type. However, Figure 21 does show that there is apparently more correlation between ln(QMED) and $BFIHOST^2$ (cor = −0.76) than there is between ln(QMED) and gauged BFI^2 (cor = −0.62).One possible reason for the stronger correlation is that the interpretation of the HOST classification system, used to generate HOST-based catchment descriptors, is measuring one or more other characteristics as well as baseflow. This is possible, as the standard error of the BFIHOST equation is 0.089 (Boorman and others, 1995). Candidates for the other measured characteristics may include catchment slope, as certain soil classes correlate with either steep slopes or flat lands, and soil permeability, which would relate to the fraction of a large rainfall event (like RMED) typically becoming runoff – which is what SPRHOST is intended to estimate (Robson and Reed 1999). While there is generally a good correlation between SPRHOST and BFIHOST, the requirement for and existence of the descriptor RESHOST demonstrates that there are certain situations where the value of one is a poor guide to the value of the other. As the standard error of the SPRHOST equation is 10 (Boorman and others, 1995), the BFIHOST descriptor may also hold some information on gauged SPR that gauged BFI does not.

Stage 2 models, combining one flow measure with one baseflow or soils measure, effectively combine information on soil permeability and catchment slope with direct information on parts of the flow regime above baseflow, typically caused by the types of rainfall events that cause flood peaks in those catchments, and indirect information on catchment area. It is therefore expected that these models have highly negative log-likelihood and low model structural error (σ_η).

Table 9 also shows that points on the flow duration curve, especially higher ones like Q10, are very effective predictors of QMED. Although gauged flow quantiles are clearly very effective in estimating QMED, some gauging is evidently required to obtain them. However, *Q*10 (for example) can be gauged in far less time than QMED to the same level of uncertainty. Furthermore, the accuracy with which Q10 is estimated does not depend on the measurement accuracy with which flows higher than Q10 are gauged. Although this is technically also true of QMED, it can be seen on Figure 20 that the magnitude of Q10 is typically around one-tenth of the magnitude of QMED and therefore considerably easier to gauge to a lower uncertainty. If gauging is impossible, it becomes necessary to use an estimated Q10 value, which can be derived from a pooling procedure based directly on HOST classes and average annual runoff (Holmes and others, 2005), noting that the Q10 estimation procedure has its own uncertainty. It is currently unknown which QMED estimation method would have higher uncertainty: one based entirely on catchment descriptors, or one based on catchment descriptors and estimated flows (in place of gauged flows).

It is unfortunate that the QMED-from-Q10 estimation method developed here is based on so few small catchments, as it has a far smaller model error than the QMED-from-descriptors approach. However, a similar estimation method, called the ‘QMED linking equation’ already exists and is implemented in WINFAP 4 (Wallingford HydroSolutions 2016a). The QMED linking equation is reproduced in Equation 20.

Equation 20 – QMED linking equation

$$QMED = 1.762Q5^{0.866}(1 + GRADQ5)^{-0.775}DPSBAR^{0.265}0.2388^{BFI^2}$$

In Equation 20, GRADQ5 is the gradient of the flow duration curve between Q5 and Q10 under the assumption of a log-normal approximation. The QMED linking equation is based on Q5 and Q10, the daily mean flows exceeded on 5% and 10% of days on average, along with DPSBAR and gauged BFI, and is calibrated to a data set of 336 catchments from 139 to 4,400 km^2. As it incorporates information about the flow regime directly, there is no obvious reason why it would apply less to smaller catchments. Further work should review the performance of the QMED linking equation as applied to the project data set, including the proportion of unexplained variance in QMED in small catchments relative to larger catchments. Given that it was developed using a more diverse set of catchments and its factorial standard error (1.31) is larger than that of many two- and three-parameter models in Table 9, it is almost certain that a more developed QMED linking-type equation specific to small catchments would reduce QMED estimation uncertainty in small catchments further.

Despite the strong performance of both the QMED linking equation and the best equations developed in this section compared to purely catchment-descriptor based QMED estimation methods, it is important to note that a value of QMED estimated by any method incorporating gauged data cannot be used as a substitute ‘gauged’ QMED value for donor transfer, due to the uncertainty that still remains in the estimate: a factorial standard error of 1.3 still corresponds to a 95% confidence interval of 60 to 167% of the calculated value. For a hypothetical donor typical of those available in the NRFA peak flow data set, version 4.1 (record length = 40, β = 0.2), the 95% confidence interval of QMED is just 88 to 113% of the gauged value.

7. Donor transfer in small catchments

7.1 Revisiting donor transfer procedures

The existing FEH statistical method recommends using donor transfer in all cases and recommends geographical proximity between the donor and target catchment centroids as the main factor for donor selection. Donor transfer gives an adjusted QMED value as shown in Equations 21 and 22.

Equation 21 – FEH statistical QMED adjustment by donor transfer: general form

$$QMED_{s,adj} = QMED_{s,cds}\left(\frac{QMED_{g,obs}}{QMED_{g,cds}}\right)^{\alpha}$$

Equation 22 – FEH statistical QMED model error correlation, used for donor transfer

$$\alpha \; = \; 0.4598e^{-0.0200d_{ij}} + (1 - 0.4598)e^{-0.4785d_{ij}}$$

In Equations 21 and 22, subscripts s and g refer to the target site and a nearby gauged site respectively, subscripts adj, cds and obs refer to adjusted, estimated (from catchment descriptors) and observed (gauged) QMED respectively, and d_{sg} in Equation 22 refers to the geographical distance between the target and gauged catchment centroids, in kilometres.

Kjeldsen and others (2008) imply that location of the donor and target on the same river network may be advantageous, but also that gains will be marginal and that geographically close catchments are more likely to be on the same network anyway. The form of Equation 22 implicitly requires that the donor should have low sampling error in ln(QMED), as derived on pages 42 to 43 of Kjeldsen and others (2008), which in practice means a long AMAX record. More recent work (Kjeldsen and others, 2014) has presented a framework for weighted donor transfer from multiple donor catchments and suggests six donors as the typical optimal trade-off between increasing the amount of gauged information contributing to an estimate and the decreasing relevance of each additional piece of information relative to the last.

In practice in small catchments, donor transfer is often ignored or factors other than proximity are prioritised. This overlooks a key feature of the method used to fit the FEH statistical equation: that it accepts that four catchment descriptors are not enough to capture all variations in ln(QMED) and, accepting this, that it attempts to give local consistency of estimates at the cost of imposing long-range spatial patterns in residuals. In other words, the method is designed to cluster positive and negative model residuals, and the optimisation procedure sets the distance over which the clusters form and recede. This is shown most obviously in Figure 2.3 of ‘Making better use of local data in flood frequency estimation’ (Environment Agency and others 2017), while the effect of donor transfer on negating these patterns is shown in Figures 2.4 and 2.5.

Because the overarching pattern in residuals is coerced to relate to geographical location over any other factor, hydrological similarity is, on average, less important than proximity when selecting a donor. Indeed, if a donor is selected according to similarity in catchment descriptors, then it is implied that model residuals are correlated with those descriptors. Selecting a donor because of similarity in AREA, for instance, implies an overall correlation between model residuals and AREA, which has not been observed (Vesuviano and others, 2016). Furthermore, the structure of equation 20 shows that the best donors are those that show similar levels of over or underestimation to the catchment of interest (model error correlation, discussed in Section 3.2 of this report).

However, it is noted that spatial correlation in residuals is a general pattern, not a fixed rule, so donor transfer in practice only acts to reduce the spatial correlations between residuals – it doesn’t necessarily reduce every residual across the whole UK. Furthermore, not every residual can map perfectly onto the overarching pattern of positive and negative values. Small catchments are more heterogeneous as a group (i.e. they are more varied, as uncommon small-scale hydrological features have the potential to occupy more of the catchment’s area and dominate the hydrological response), and potentially more spatially homogeneous when considered individually (i.e. they are often too small to include areas with vastly differing hydrological properties). Therefore, they are not typical of the majority of larger catchments used in developing the existing FEH statistical model and may be less likely to fit the typical pattern of residuals.

Table 10 demonstrates the results of applying automated donor transfer procedures to the 93 small catchments used to calibrate the new small catchments QMED model. QMED is always estimated using the existing FEH statistical equation and selection is based entirely on the proximity of donor and target catchment centroids.

Table 10 - Residuals in ln(QMED) estimates in relation to distance-only donor transfer (existing FEH model, 93 small catchments)

No. donors	No. residuals reduced (vs 0 donors)	No. residuals increased (vs 0 donors)	Mean ln-residual	Mean (ln-residual^2)
0	N/A	N/A	−0.0085	0.3482
1	43	40	−0.0210	0.3473
2	50	43	−0.0245	0.3667
6	51	42	−0.0211	0.3651

Table 10 shows that the smallest mean residual is achieved without donor transfer, and that the spread of residuals is larger for two or six donors than for zero or one. Though this result may be taken to imply that donor transfer is unnecessary or actively unhelpful in small catchments, it could also indicate that it is the automated selection procedure that is unhelpful. This supports the premise that QMED residuals for small catchments may not follow the overarching spatial pattern of the residuals in general.

The effectiveness of the donor transfer procedure in small catchments is tested here using the existing FEH statistical method, as the larger number of catchments included in calibration of that method was enough to allow residuals to be coerced to a consistent spatial pattern across the UK. Observation of the blue circles on Figure 1 suggests that the spatial correlation fitted to either the retuned FEH equation or the new small catchment equation residuals could be compromised by the sparseness of catchments. In any case, the residuals for either equation are fitted in such a way that spatial correlation between them reduces below 0.01 at a centroid-centroid distance of less than 2 km (as shown by the fitted φ values in equations 15 and 17) – effectively, no spatial correlation was fitted between residuals for either the retuned FEH equation or the new small catchment equation. Consequently, the donor transfer relationship for the existing FEH statistical method, reproduced in Equations 21 and 22, is not suitable for either the retuned equation or new small catchment equation (model 4ABAB).

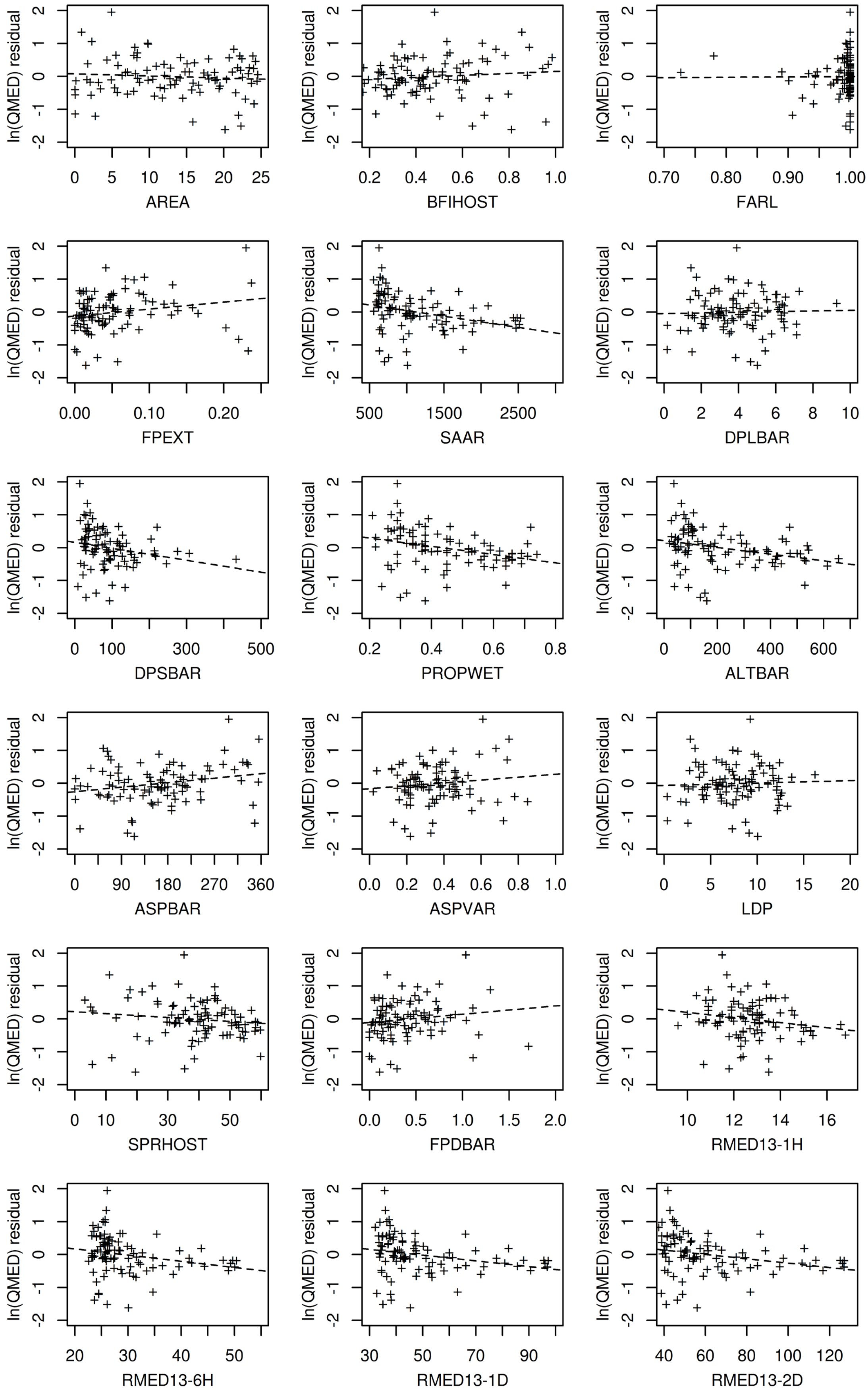

Figure 22 - Correlations between residuals in ln(QMED) and catchment descriptors for the existing FEH QMED equation

Given that no spatial correlation between residuals is evident when only this sparse data set of small catchments is considered, potential correlations were investigated between residuals and catchment properties. The eighteen scatter graphs in Figure 22 show several apparent correlations between residuals in ln(QMED) estimated by the existing FEH equation and eighteen catchment descriptors. There appear to be negative associations with wetter catchments (as measured by SAAR, PROPWET or any RMED13 descriptor), and steeper and higher catchments (although these may also be wetter catchments). Indeed, this is to be expected somewhat, as the coefficients fitted to the (1000/SAAR) term are very different between the existing FEH statistical equation and the retuned FEH equation. Therefore, the relationship between the raw SAAR value and the fitted term is noticeably different (Figure 9).

In addition, there appear to be mildly positive associations with flood plain descriptors FPEXT and FPDBAR. Therefore, selecting suitable donors could beneficially consider catchment wetness and rainfall, and potentially flood plain characteristics. It is noted that there is no serious correlation between residuals in ln(QMED) and AREA. This reinforces the findings of SC090031/R2 (‘Performance of FEH peak flow estimation methods in small catchments) and Vesuviano and others (2016) and is further confirmed by the similarity in the coefficients given to AREA in the FEH statistical equation and the retuned FEH equation (Figure 9).

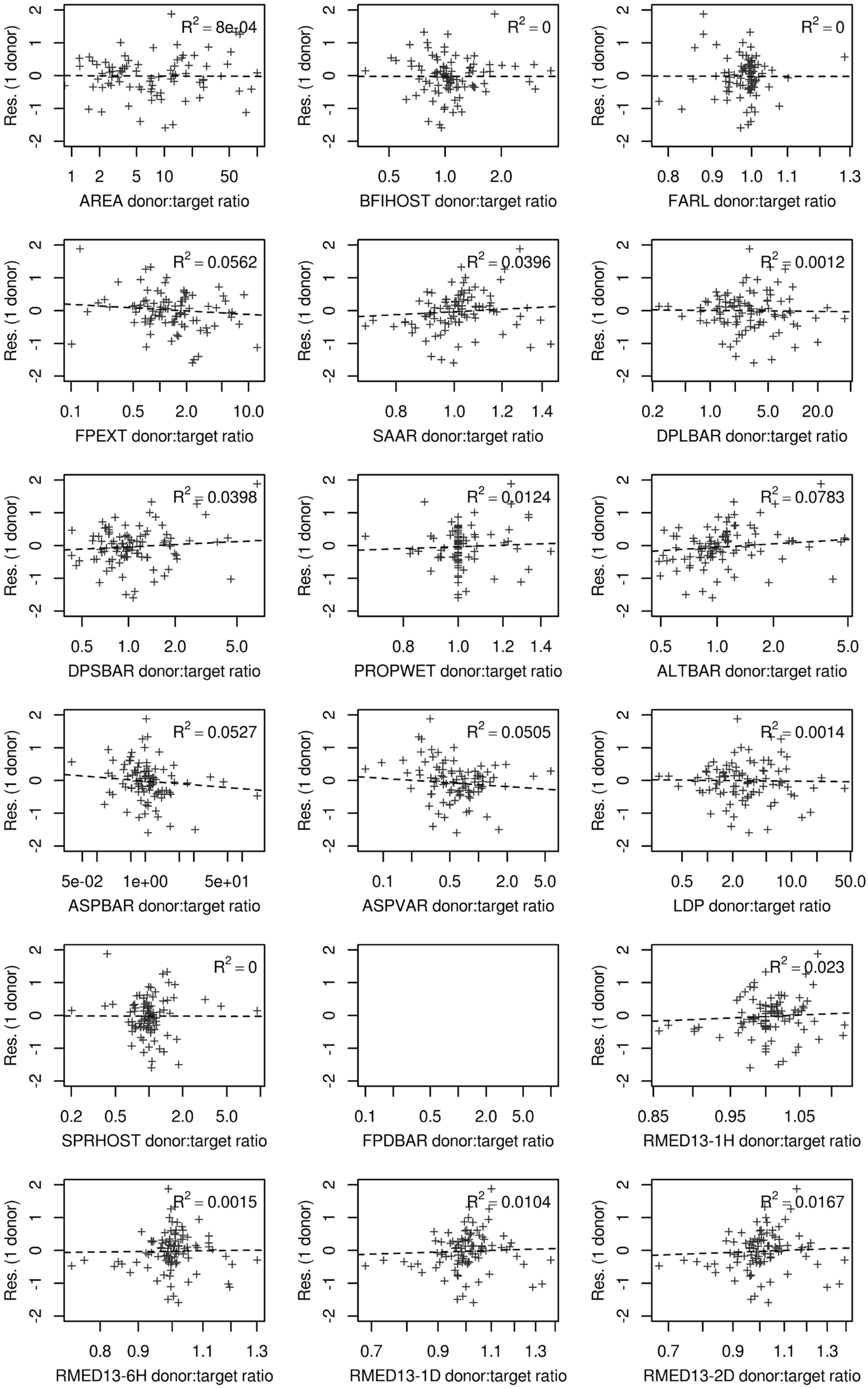

Figure 23 - Correlations between residuals in ln(QMED) and ratios of catchment descriptors (donor:target) for the existing FEH statistical equation (note: NRFA Peak Flows dataset v4.1 did not include FPDBAR)

The eighteen scatter graphs in Figure 23 plot model residuals against the ratio of catchment descriptor values between a donor and the catchment of interest, for seventeen different descriptors (note that ratios of FPDBAR values are not plotted as these are not made generally available with the NRFA's peak flow data set). Figure 23 shows that underestimation becomes more prevalent if SAAR at the donor is more than 1.2× SAAR at the target. Similarly, overestimation of QMED tends to become more prevalent if SAAR at the donor is less than 0.8× SAAR at the target. This is consistent with Figure 22, which shows increasing overestimation with increasing SAAR. A similar consistency between FPEXT value at the target, FPEXT value at the donor, and over- or underestimation is also visible.

Given that hydrologists will tend to use judgement in selecting donors, two modified versions of the automated donor selection procedure are tested on the 93-catchment data set of catchments up to 25 km^2. The first places bounds on the maximum permitted dissimilarity between donor and target in terms SAAR. Specifically, the donor catchment SAAR must be within ±20% of the target catchment SAAR. The second places the same bounds on SAAR but also specifies that the donor catchment AREA must not be more than five times or less than one-fifth of the target catchment AREA. For target catchments under 5 km^2, this rule is relaxed to allow donors up to 25 km^2. The existing FEH (2008) equation is used to estimate QMED, and all other aspects of the automated donor selection procedure are unchanged – potential donors are first filtered by similarity, then selected in order of centroid-centroid distance.The results of the first modified donor transfer procedure are shown in Table 11 and the results of the second in Table 12.

Table 11 - Residuals in ln(QMED) estimates in relation to distance and SAAR-based donor transfer (existing FEH model, 93 small catchments)

No. donors	No. residuals reduced (vs 0 donors)	No. residuals increased (vs 0 donors)	Mean ln-residual	Mean (ln-residual^2)
0	N/A	N/A	−0.0085	0.3482
1	40	53	0.0212	0.3536
2	50	43	−0.0268	0.3703
6	50	42	0.0224	0.3791

Table 32 - Residuals in ln(QMED) estimates in relation to distance, AREA and SAAR-based donor transfer (existing FEH model, 93 small catchments)

No. donors	No. residuals reduced (vs 0 donors)	No. residuals increased (vs 0 donors)	Mean ln-residual	Mean (ln-residual^2)
0	N/A	N/A	−0.0085	0.3482
1	46	46	0.0014	0.3711
2	54	37	−0.0056	0.3813
6	52	39	0.0215	0.3690

Perhaps, surprisingly, a donor transfer procedure in which SAAR is filtered (Table 11) offers lower performance than one that is not filtered at all (Table 10) in terms of the number of residuals reduced, mean residual and spread of residuals, despite the observed trend relating SAAR and model error. However, this finding is considered less surprising once the amount of variation in model error at any particular subset of similar SAAR values is considered (Figure 22). This finding is consistent with page 68 of the Environment Agency's Flood Estimation Guidelines (LIT 11832: Environment Agency, 2022), which state that "It is not essential for donor and subject sites to be similar in catchment area, BFIHOST and other catchment descriptors included in the QMED regression equation".

Despite the observed lack of correlation between model error and AREA, a donor transfer procedure in which potential donors are first filtered by SAAR and AREA appears at first to show promising results. However, there are both advantages and disadvantages to filtering potential donors. The first advantage is that the number of residuals reduced is greater when donors are filtered, meaning that there is a slightly better chance of reducing the model residual at an arbitrary site. The second advantage is that the mean residual is closer to zero when either one or two donors are used. However, the spread of residuals is greater when donors are first filtered by SAAR and AREA, meaning that the lower mean residual is caused by a more even balance between total over and underestimation across all sites, not necessarily by existing estimates. In fact, viewed in the context that filtered donor transfer reduces the model error at slightly more sites, the only way that the spread of errors could be increased while the majority of errors are made smaller is by making the model errors much larger at a minority of sites. The main criticism of filtered donor transfer is that, unlike distance-only donor transfer, there is no number of donors that gives a lower spread of residuals than is achieved with zero donors.

Figure 24 plots absolute model residuals after SAAR and AREA-filtered donor transfer against absolute model residuals after distance-only donor transfer. As implied by the statistics shown in Table 12, filtered donor transfer offers marginal improvements over distance-only donor transfer mainly when the model residual is already small. When the model residual is larger, that is, when the existing FEH statistical method performs less well, filtered donor transfer is generally less helpful than distance-only donor transfer in reducing model residuals. Another disadvantage of filtering potential donors by similarity in catchment descriptors is that the pool of suitable donors is reduced. In this study, applying rigid bounds on the permissible AREA and SAAR of potential donors only allowed donor transfer for 92 catchments and multiple donors for 91 catchments. Both catchments where multiple donor transfer was made impossible were in Northern Ireland, which has only 39 potential donors in total.

For catchments larger than 25 km^2, limiting the range of donors could be detrimental as the existing FEH statistical method fits spatial correlations to model residuals, so excluding some donors from consideration means that the ones selected will generally be further away from the target. Tables 13 and 14 repeat the studies shown in Tables 10 and 11, but consider all 847 QMED-suitable catchments in the NRFA database. Urban target catchments are considered deurbanised, but donors must be rural. Figure 25 plots absolute model residuals after SAAR and AREA-filtered donor transfer against absolute model residuals after distance-only donor transfer.

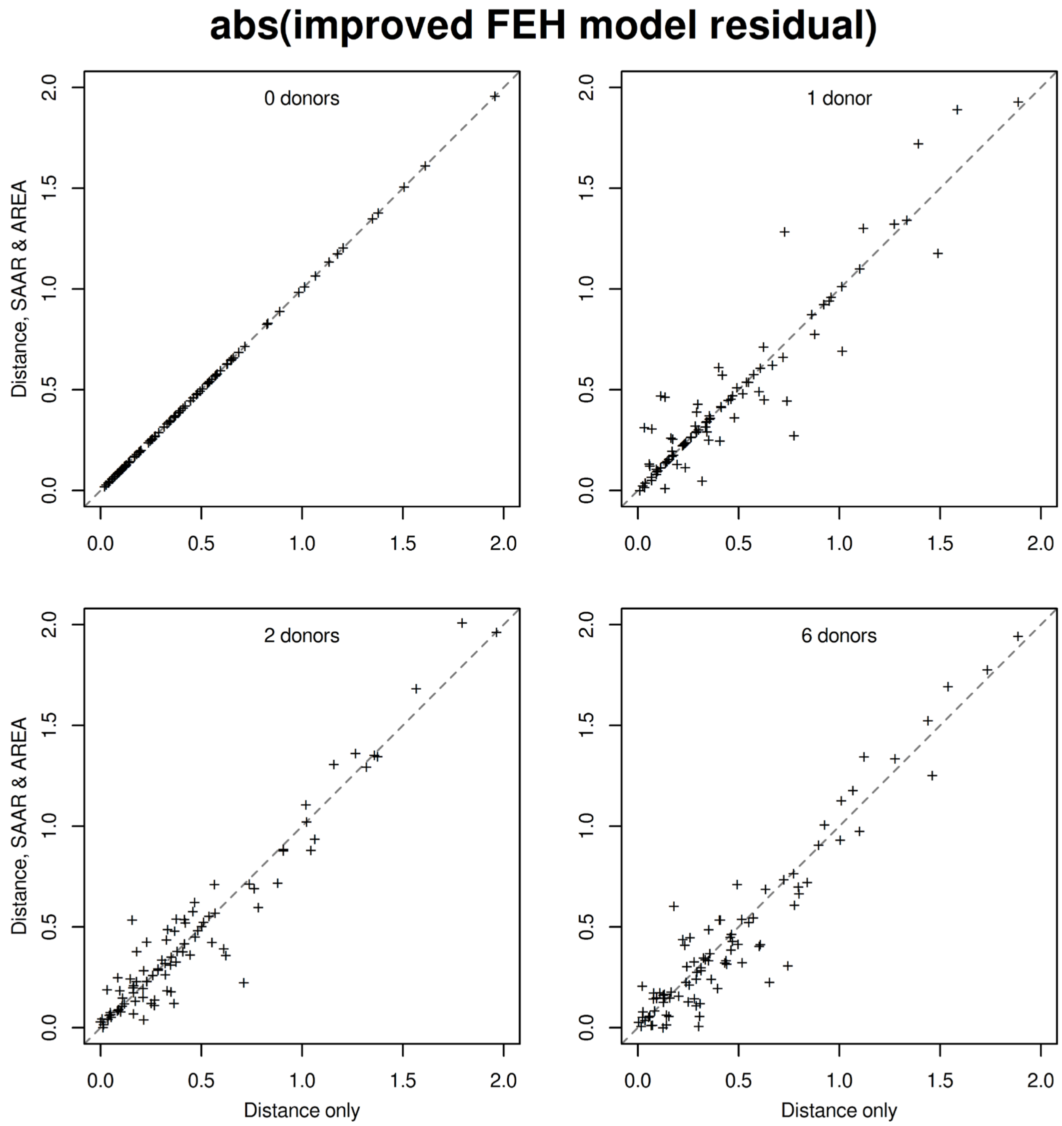

Figure 24 - Comparison of absolute model errors achieved by distance-only and filtered donor transfer (existing FEH statistical model, small catchments)

The x-axes and y-axes both plot the magnitude of the existing FEH statistical QMED model residual from 0.0 to 2.0. The x-axis plot absolute residual when donors are selected by distance only, while the y-axis plots absolute residual when donors are filtered by SAAR and AREA, then selected by distance. The four graphs compare different numbers of donors:

- Top-left graph: 0 donors.
- Top-right graph: 1 donor.
- Bottom-left graph: 2 donors.
- Bottom-right graph: 6 donors.

Tables 13 and 14 show almost the opposite to Tables 10 and 12. Mean residual is lower when donors are not filtered by AREA and SAAR. However, the numbers of residuals reduced and the spread of residuals may point towards the benefits of a modified donor transfer method for larger catchments, although Figure 25 shows this effect to be very inconsistent, with both distance-only and filtered donor transfer outperforming each other at different sites, and much smaller model residuals (therefore much smaller improvements due to donor transfer) at the vast majority of sites. It is crucial to note that calculation of α is still based on Equation 22, so if a more similar but more distant catchment is used, rather than a less similar but closer one, α is still reduced.

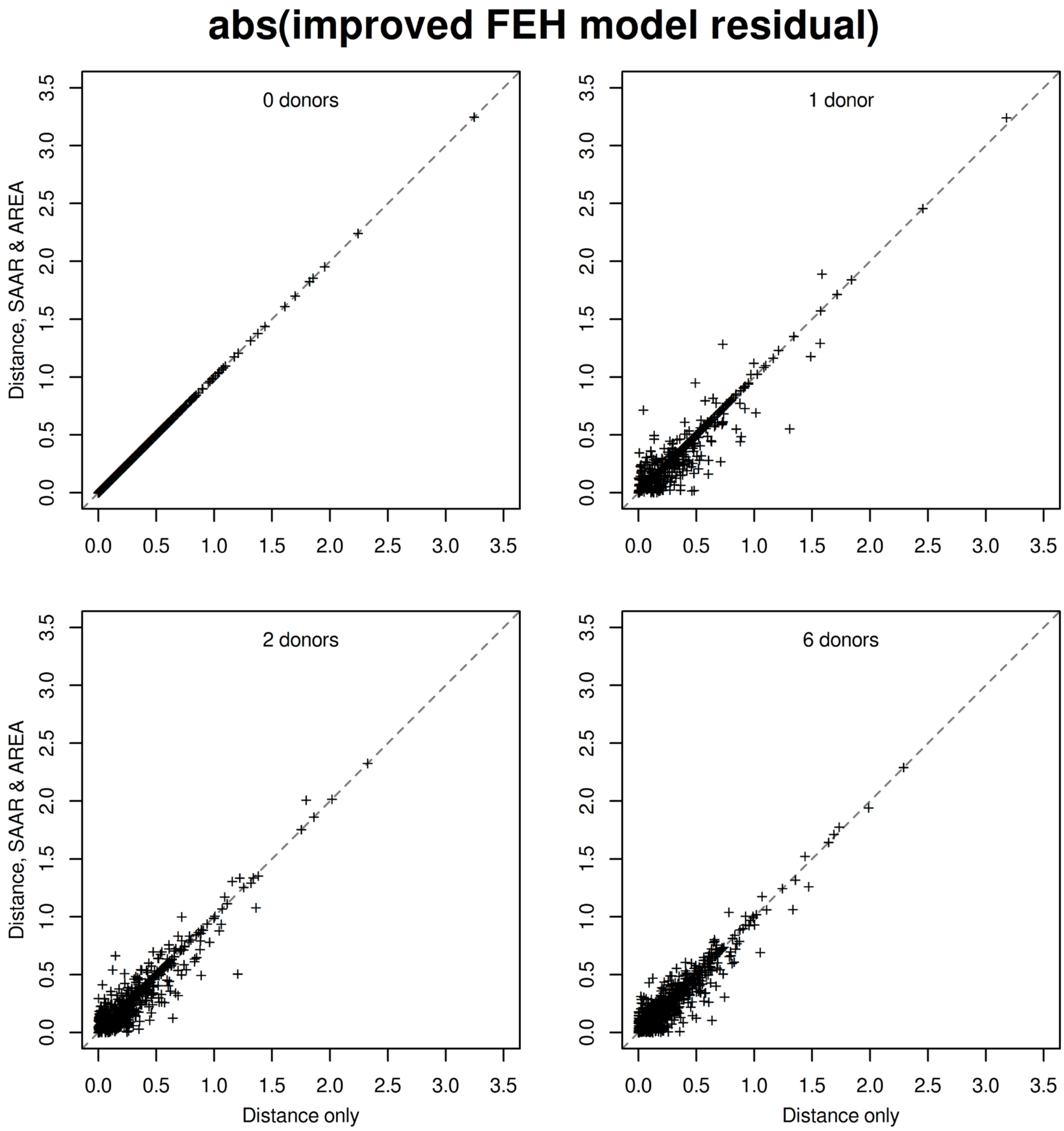

Figure 25 - Comparison of absolute model errors achieved by distance-only and filtered donor transfer (existing FEH statistical model, full NRFA database version 4.1, suitable for pooling)

Figure 25 is analogous to Figure 24, but considers the full NRFA Peak Flows v4.1 data set.

The x-axes and y-axes both plot the magnitude of the existing FEH statistical QMED model residual from 0.0 to 3.5. The x-axis plot absolute residual when donors are selected by distance only, while the y-axis plots absolute residual when donors are filtered by

SAAR and AREA, then selected by distance. The four graphs compare different numbers of donors:

- Top-left graph: 0 donors.
- Top-right graph: 1 donor.
- Bottom-left graph: 2 donors.
- Bottom-right graph: 6 donors.

Table 43 - Residuals in ln(QMED) estimates in relation to distance-only donor transfer (QMED-suitable catchments, full NRFA database version 4.1)

No. donors	No. residuals reduced (vs 0 donors)	No. residuals increased (vs 0 donors)	Mean ln-residual	Mean (ln-residual^2)
0	N/A	N/A	0.0096	0.1685
1	483	364	−0.0084	0.1532
2	487	360	−0.0187	0.1535
6	497	350	−0.0228	0.1472

Table 54 - Residuals in ln(QMED) estimates in relation to distance, SAAR and AREA-based donor transfer (QMED-suitable catchments, full NRFA database version 4.1)

No. donors	No. residuals reduced (vs 0 donors)	No. residuals increased (vs 0 donors)	Mean ln-residual	Mean (ln-residual^2)
0	N/A	N/A	0.0096	0.1685
1	510	336	−0.0036	0.1463
2	502	341	−0.0117	0.1347
6	494	346	−0.0146	0.1264

The eighteen scatter graphs in Figure 26 show the model residual of the retuned (< 25 km^2) FEH statistical QMED equation against the values of eighteen different catchment descriptors. For the retuned FEH equation, there are no strong relationships between model residual and any catchment descriptor. While this does reinforce how this equation applies particularly well to small catchments (or at least the 93 catchments used in its calibration), it does not help identify what catchment properties may be useful for identifying suitable donors. Tables 15 and 16 test distance-only donor transfer against that filtered by AREA and SAAR for the retuned FEH equation. These show that, while any type of donor transfer has an overwhelmingly negative effect on the estimates given by retuned FEH equation, the mean and spread of the retuned FEH equation residuals with zero donors is already smaller than that of the existing FEH equation with any donor selection scheme and any number of donors.

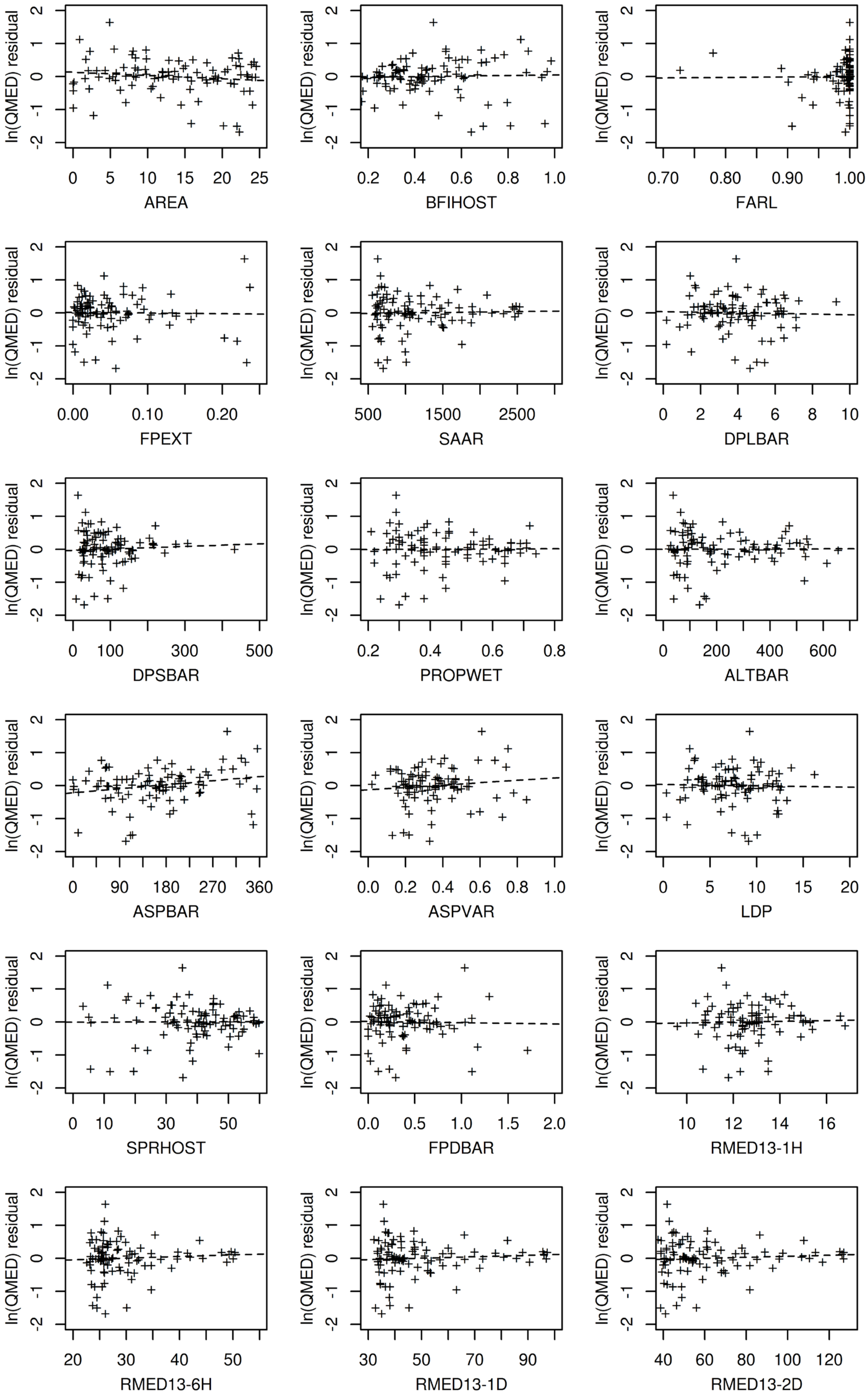

Figure 26 - Correlations between residuals in ln(QMED) and catchment descriptors for the retuned FEH statistical QMED equation

Table 6 - Residuals in ln(QMED) estimates in relation to distance-only donor transfer (retuned FEH equation, 93 small catchments)

No. donors	No. residuals reduced (vs 0 donors)	No. residuals increased (vs 0 donors)	Mean ln-residual	Mean (ln-residual^2)
0	N/A	N/A	0.0025	0.2963
1	41	52	−0.0099	0.3042
2	42	51	−0.0134	0.3252
6	44	49	−0.0100	0.3250

Table 76 - Residuals in ln(QMED) estimates in relation to distance, AREA and SAAR-based donor transfer (retuned FEH equation, 93 small catchments)

No. donors	No. residuals reduced (vs 0 donors)	No. residuals increased (vs 0 donors)	Mean ln-residual	Mean (ln-residual^2)
0	N/A	N/A	0.0025	0.2963
1	41	51	0.0097	0.3347
2	42	49	0.0031	0.3495
6	38	53	0.0302	0.3428

7.2 Recommendations for donor transfer

The standard procedure for donor transfer in the existing ‘improved’ FEH QMED equation has been revisited specifically for small catchments and tested with the existing FEH model and new small catchments model (model 4ABAB). As there is little spatial correlation between residuals in ln(QMED) for this data set of small catchments, the main motivation for always selecting the closest donor is weakened. By comparing residuals to catchment descriptors, a procedure whereby potential donors were ‘filtered out’ based on how dissimilar they were to the target catchment, in terms of SAAR and AREA, was tested.

For the existing FEH model, this was found to result in slightly lower mean ln-residuals and in reduced magnitudes of ln-residual at slightly more sites, with the use of two donors being better than one in terms of number of residuals reduced. However, further investigation found that the ‘filtered’ donor selection method offered the most improvement over distance-only donor transfer at sites where the model residual was already small before donor transfer. At sites where the model residual was larger, that is, where the existing FEH statistical method performed less well, filtered donor transfer was generally less helpful than distance only donor transfer in reducing model residuals.

For model 4ABAB, neither distance-only nor ‘filtered’ donor transfer, with any number of donors, was found to reduce either the mean or spread of model residuals or reduce the model residual at more than half of the test sites. This means that, according to any assessment metric, donor transfer is actively unhelpful for model 4ABAB. However, even without donor transfer, the mean residual for model 4ABAB is near-zero and the spread of residuals is considerably lower than can be achieved with any number of donors using the existing FEH model.

In situations where both the target catchment and potential donors are larger (above 25 km^2), it is recommended that a close donor catchment, based on the distance between catchment centroids, is selected. This is because the existing FEH statistical method was designed to constrain any variance not explained by the catchment descriptors into a consistent spatial pattern. Spatially consistent variation in residuals is clear when viewed across the UK as a whole (Environment Agency, 2017; Figure 2.3). However, evidence from Table 14 suggests that some judgement based on hydrological similarity is justified, that is, that the most suitable donor will be nearby **and** hydrologically similar. However, as Figure 19 shows strong correlation between error in estimated ln(QMED) and less-commonly-used catchment descriptors (for example, FPEXT, DPSBAR), it is likely that further research into trade-offs between proximity and hydrological similarity will reveal the importance of catchment descriptors other than those most commonly used for assessing hydrological similarity.

8. Conclusions and recommendations

8.1 Conclusions

- The best performing catchment descriptor-based models for ln(QMED) in small catchments all contain ln(AREA), one HOST descriptor and one rainfall descriptor. These are the same for both the existing FEH model and MacDonald and Fraser's small catchment QMED model, suggesting that the same high-level descriptive statistics still capture whatever the different rare event flood-generating processes may be between small and larger catchments.
- The selection of rainfall descriptor is complicated by the strong correlations between all rainfall descriptors. The ranking of rainfall descriptors in terms of their effect on model log-likelihood can almost certainly be influenced by the choice of calibration catchments. Therefore, while a larger calibration data set might result in a QMED model with a different rainfall descriptor, the value of QMED predicted by that model is unlikely to be very different.
- Incorporation of gauged flow statistics, specifically Q10 and BFI in this study, greatly reduces model structure error in the predicted QMED equation.

8.2 Recommendations

QMED estimation

- Additional research carried out on a further screened data set and reported in Section 3 of the 'Small catchments overview report' (R0), leads to the recommendation to continue using the existing (2008) FEH statistical QMED equation to estimate QMED in small ungauged catchments. To avoid any doubt, this equation is repeated below (Equation 23).

Equation 23 - QMED equation

$$QMED = AREA^{0.8510} 0.1536^{\frac{1000}{SAAR}} FARL^{3.4451} 0.0460^{BFIHOST^2}$$

- If gauged flow statistics are available, the QMED linking equation, included in WINFAP 4, is recommended over all purely catchment-descriptor based methods for estimating QMED. This equation is repeated below for convenience (Equation 24).

Equation 24 – QMED linking equation

$$QMED = 1.762 Q5^{0.866} (1 + GRADQ5)^{-0.775} DPSBAR^{0.265} 0.2388^{BFI^2}$$

Donor transfer

- Donor transfer is not recommended for use with the retuned FEH equation, as both the distance-only and 'filtered' methods of donor selection increased the mean and spread of residuals in ln(QMED) estimates and increased the magnitude of residuals at more than half of the 93 test small catchments.
- Donor transfer is recommended if the existing FEH statistical method is used to estimate QMED in small catchments. Although there are some advantages to filtering potential donors by AREA and SAAR, they are outweighed by disadvantages, particularly if the estimate of QMED given before donor transfer is poor (which will not be known if the catchment of interest is ungauged).
- Donor transfer is unnecessary when the QMED linking equation is used, as it incorporates gauged at-site flow data from the site of interest directly.
- For larger catchments (that is, both target and potential donor > 25 km^2), donor selection should normally be based on centroid-centroid distance between the catchments. However, there may be some potential benefits in considering hydrological similarity when selecting donors. These are more difficult to assess, as model residuals are smaller, both before and after donor transfer.

Further work

- Further work is required to identify which catchment descriptors are most strongly linked to residuals in estimated ln(QMED) for larger catchments, and over what ranges, as it is similarity between these descriptors that will be most important in determining suitable donors to counteract modelling error in the QMED equation.
- Given the importance of gauged flow statistics, further work should evaluate the performance of the QMED linking equation in small catchments specifically, and whether a similar, small catchment-specific equation incorporating gauged flow statistics can further reduce uncertainty in small catchments.

References

BOORMAN, D.B., HOLLIS, J.M. AND LILLY, A., 1995. 'Hydrology of soil types: a hydrologically-based classification of the soils of the United Kingdom (IH report 126)' Wallingford: Institute of Hydrology.

ENVIRONMENT AGENCY., 2017. 'Making better use of local data in flood frequency estimation' Bristol: Environment Agency.

ENVIRONMENT AGENCY., 2022. 'LIT 11832: Flood Estimation Guidelines' Bristol: Environment Agency.

FAULKNER, D.S., FRANCIS, O. AND LAMB, R., 2012. 'Greenfield run off and flood estimation on small catchments' Journal of Flood Risk Management, 5 (1), 81 to 90.

GUSTARD, A., BULLOCK, A. AND DIXON, J.M., 1992. 'Low flow estimation in the United Kingdom' Wallingford: Institute of Hydrology.

HOLMES, M.G.R., YOUNG, A.R., GOODWIN, T.H. AND GREW, R., 2005. 'A catchment-based water resource decision-support tool for the United Kingdom' Environmental Modelling & Software, 20 (2), 197 to 202.

HOSKING, J.R.M. AND WALLIS, J.R., 1997. 'Regional Frequency Analysis: An Approach Based on *L*-Moments' Cambridge: Cambridge University Press.

KJELDSEN, T.R., STEWART, E.J., PACKMAN, J.C., FOLWELL, S.S. AND BAYLISS, A.C., 2005. 'Revitalisation of the FSR/FEH rainfall-runoff method' Bristol: Environment Agency.

KJELDSEN, T.R., JONES, D.A. AND BAYLISS, A.C., 2008. 'Improving the FEH statistical procedures for flood frequency estimation' Bristol: Environment Agency.

KJELDSEN, T.R. AND JONES, D.A., 2009. 'An exploratory analysis of error components in hydrological regression modeling' Water Resources Research, 45 (2), W02407.

KJELDSEN, T.R., JONES, D.A. AND MORRIS, D.G., 2014. 'Using multiple donor sites for enhanced flood estimation in ungauged catchments' Water Resources Research, 50 (8), 6,646 to 6,657.

MACDONALD, D.E. AND FRASER, R.J., 2013. 'An improved method for estimating the median annual flood for small ungauged catchments in the United Kingdom' Journal of Flood Risk Management, 7 (3), 251 to 264.

MOSTELLER, F., 1946. 'On Some Useful 'Inefficient' Statistics' Annals of Mathematical Statistics, 17 (4), 377 to 408.

NATIONAL RIVER FLOW ARCHIVE (NRFA)., 2014. 'NRFA peak flow data' [online] Wallingford: National River Flow Archive. Available from:

http://www.ceh.ac.uk/data/nrfa/data/peakflow_data.html [Accessed 6 March 2015] [Accessed 19 August 2020].

ROBSON, A. AND REED, D., 1999. ‘Statistical procedures for flood frequency estimation (Flood Estimation Handbook volume 3)’ Wallingford: Institute of Hydrology.

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.

WALLINGFORD HYDROSOLUTIONS., 2016a. ‘WINFAP 4 QMED Linking equation’ [online] Wallingford: Wallingford HydroSolutions. Available from: http://software.hydrosolutions.co.uk/winfap4/QMED-Linking-Equation-Technical-Note.pdf [Accessed 19 April 2017] [Accessed 19 August 2020].

WALLINGFORD HYDROSOLUTIONS., 2016b. ‘WINFAP 4 Urban Adjustment Procedures’ [online] Wallingford: Wallingford HydroSolutions. Available from: http://software.hydrosolutions.co.uk/winfap4/Urban-Adjustment-Procedure-Technical-Note.pdf [Accessed 19 April 2017] [Accessed 19 August 2020].

List of abbreviations

AMAX	Peak instantaneous annual maximum flow
BFI	Baseflow index
FEH	Flood Estimation Handbook
GLO	Generalised logistic distribution
HOST	Hydrology of soil types
IHDTM	Integrated Hydrological Digital Terrain Model
ML	Maximum likelihood
MORECS	Met Office Rainfall and Evapotranspiration Calculation System
NRFA	National River Flow Archive
POT	Peaks-over-threshold
PRUAF	Percentage runoff urban adjustment factor
QMED	Median annual flood, normally estimated as median of AMAX series
RMED	Median annual rainfall, normally estimated from a rainfall model
SPR	Standard percentage runoff

Would you like to find out more about us or your environment?

Then call us on

03708 506 506 (Monday to Friday, 8am to 6pm)

Email: enquiries@environment-agency.gov.uk

Or visit our website

www.gov.uk/environment-agency

incident hotline

0800 807060 **(24 hours)**

floodline

0345 988 1188 **(24 hours)**

Find out about call charges (https://www.gov.uk/call-charges)

Environment first

Are you viewing this onscreen? Please consider the environment and only print if absolutely necessary. If you are reading a paper copy, please don’t forget to reuse and recycle.