Estimating flood peaks and hydrographs for small catchments:

R2 – performance of FEH peak flow estimation methods in small catchments

FCERM Research & Development Programme

Research Report

Date: March 2024

Version: SC090031/R2

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

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):
Elizabeth (Lisa) Stewart, Gianni Vesuviano, Andy Young, Tracey Haxton

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

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

Every year in the UK, many flood studies are carried out in small catchments, typically classed as draining areas of less than 25 km^2. The aim of this project is to examine how appropriate existing flood estimation methods are when applied to small catchments, and to develop new and improved techniques that can be used in small fluvial catchments and plots. This report describes an analysis of the performance of three recent methods of flood estimation on an 'extended' data set of peak flows for 217 small UK catchments and makes recommendations for the next stage of the project.

Three methods for estimating QMED, the median annual flood, in small catchments are described: the revitalised flood hydrograph (ReFH1 and ReFH2) method, the improved FEH statistical method and MacDonald and Fraser's (2013) catchment descriptor equation. It is noted that since the latter method does not estimate the flood growth curve, only the ReFH and FEH statistical methods can be used to estimate flood peaks of longer return period.

Details of the 'extended' data set developed during Task 1 of the project are provided and the range of catchment types is discussed. QMED values are available for a total of 217 small catchments, while peak flows for estimating higher return period floods are available for 192 catchments. The median values of BFIHOST (indexing permeability) and SAAR (average annual rainfall) within the small catchment data set are very similar to those of the NRFA peak flow data. However, the study data set is somewhat less urbanised than the NRFA peak flow data set, which is not highly urbanised in general.

Comparisons of the performance of the three key methods when used to estimate QMED are presented, based on separate analyses for rural and urban catchments within a 'high-quality' subset of the data as well as the full 'extended' data set. For rural catchments, the best performing method overall was found to be ReFH2, particularly in wetter and less permeable catchments, although the FEH statistical method came a close second overall and performed well in permeable catchments. The recently upgraded ReFH2 method was found to be a substantial improvement on its predecessor, ReFH1. Estimated error in QMED was not found to be related to catchment area in the data set and some of the largest positive errors were found for catchments of around 20 km^2. The results suggest that further work on the specifics of donor transfer, for example, the influence of donor catchment area, may be worthwhile.

In urban catchments, ReFH2 and the FEH statistical method were found to outperform the ReFH1, and MacDonald and Fraser methods as expected, although the analysis was limited by the relatively small number of urban catchments in the 'high-quality' data set. ReFH2 can be recommended to use in urban and/or permeable catchments. The results of the analysis suggest that the FEH statistical method may overestimate QMED in small urban catchments in high-SAAR areas.

The FEH methods of QT estimation were compared for small rural catchments only, since the enhanced single-site method, which gives a higher weight to at-site data in the pooling-group and was used as the benchmark, is not currently recommended for urban catchments. The results suggest that there is little evidence to prefer one FEH method over another, although ReFH2 was again found to be a great improvement on ReFH1, maintaining a consistent level of performance for return periods up to 1,000 years. The analysis found that the FEH statistical method generates similar flood growth curves for every small catchment, reflecting the importance of catchment area in selecting pooling-groups. This is an important finding which suggests that alternative pooling procedures may offer benefits for flood estimation in small catchments.

Finally, it is recommended that future work should consider the design storm inputs to ReFH2 in small urban catchments, particularly whether winter or summer storms and antecedent conditions are most appropriate to heavily urbanised catchments.

The completion of this task marks a breakpoint in the contract to allow for discussion of the recommendations for the remainder of the project. A separate report presents a summary of the findings of this evaluation and sets out a recommended plan for developing improvements for a range of small catchment applications.

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

1.1 Scope

Phase 2 of project SC090031 'Estimating flood peaks and hydrographs for small catchments' aims to develop improved techniques for flood estimation in small UK catchments, including peak flows and hydrographs. The project is led by the Centre for Ecology & Hydrology (CEH), and involves practical input and guidance from JBA Consulting, Wallingford HydroSolutions (WHS) and David MacDonald (independent consultant). The overall objective of the project is to allow hydrologists and engineers to produce flood estimates for small catchments, down to plot scale, with less uncertainty than is currently possible.

Phase 1 of project SC090031 was a scoping study, the results of which are reported by Faulkner and others (2012). The phase 1 analysis concluded that the then-existing FEH methods (the FEH statistical and ReFH1 methods) were more appropriate for flood estimation in small catchments than other widely used techniques such as ADAS 345 (ADAS, 1982) or IH 124 (Marshall & Bayliss, 1994). However, the flood response of highly permeable and/or urbanised catchments was identified as being more difficult to characterise. The main recommendations of phase 1 led to the definition of the following objectives for phase 2:

- The development of an expanded data set of small catchments peak flow data (Task 1).
- The development of improved methods to model flood flows in small ungauged catchments and plots (Task 2).
- Further review and recommendations (Task 3).

Task 1 of the current project was concerned with developing an extended data set of peak flows for catchments of up to 40 km^2 in area. This report is the outcome of Tasks 2.1.1 and 2.1.3 (Evaluation of existing methods/recommendations) and describes an analysis of the performance of the most recent versions of existing FEH methods on the 'extended' data set of small catchment peak flows. The method of estimating the median annual flood (QMED) proposed by MacDonald and Fraser (2013) is also evaluated, as it was designed to be used on small catchments and is the most up-to-date alternative to the FEH. The 'extended' data set used in this analysis includes only small catchments with identifiable drainage paths and does not include the three plot-scale areas analysed earlier in the project (Haxton and others, 2014; Siddaway and Faulkner, 2015).

Due to the 'index flood' structure of both the FEH statistical method and MacDonald and Fraser's method, performance of the methods is evaluated initially for the index flood, QMED. This event has an annual exceedance probability (AEP) of 50%, with a corresponding return period of two years. The analysis of the FEH methods is then extended to flood peaks with longer return periods, paying special attention to Q30 and

Q100 (the so-called 30-year and 100-year floods, with 3.33% and 1% AEP respectively), as these are typical in development planning and drainage design.

Gauges within the ‘extended’ small catchment data set identified as having the highest quality data are considered separately from the data set as a whole, to evaluate the performance of the methods on catchments where there is more confidence in the benchmarking data.

1.2 Report structure

Section 2 of this report outlines the current methods that have been tested during this part of the study. The analysis has been carried out on the new flood peak data set for small catchments developed during Task 1 of the project, and this is described in detail in Section 3. Section 4 presents the evaluation of the current small catchment methods for estimating QMED, the median annual flood, distinguishing between different types of catchment. The methods are further evaluated in terms of their performance for higher return period estimation, in Section 5. Finally, the conclusions from this part of the study are presented in Section 6, together with some suggestions for further methodological development.

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

2. Methods

2.1 Overview

Flood estimation methods in the UK typically follow one of two approaches. The statistical index flood methodology links peak flow rates to measurable catchment and/or climate properties, while the rainfall-runoff approach models the flow response to a set of design inputs, including rainfall depth, duration, temporal distribution and antecedent catchment conditions. The choice of approach reflects the specifics of each individual application, for example, the size of the catchment and whether a design flood hydrograph or peak flow rate is required. Some guidance on selecting the most appropriate method is given in FEH Volume 1 (Reed, 1999) and the Environment Agency's Flood Estimation Guidelines (Environment Agency, 2022).

The analysis described in this report has compared the application of several variants of the current FEH statistical and rainfall-runoff methods to the 'extended' data set of small catchment flood peaks collated in Task 1 of this project. In addition, the QMED estimation equation proposed by MacDonald and Fraser (2013) has been included in the comparison as this is a recent method which has been specifically calibrated to small UK catchment data. Older methods for estimating flood peaks in small catchments such as IH124 (Marshall and Bayliss, 1994) and ADAS 345 (ADAS, 1982) were evaluated during the first phase of this project and found to be inferior to the FEH methods (Faulkner and others, 2012). For this reason, they have not been considered further.

2.2 Revitalised flood hydrograph method (ReFH1 and ReFH2)

The revitalised flood hydrograph (ReFH) method (Kjeldsen and others, 2005; Kjeldsen, 2007) was developed at CEH some years after the FEH was released as a replacement for the FSR/FEH rainfall-runoff method. The method uses an event-based rainfall-runoff model, the revitalised flood hydrograph (ReFH) model, to convert design storm events of appropriate duration and rarity into corresponding design flood hydrographs of equivalent rarity. The ReFH model consists of three components: a loss model, a routing model and a baseflow model. The ReFH model has four parameters which control hydrological losses (C_{max}, maximum soil capacity), routing (T_p, time-to-peak) and baseflow (BR, baseflow recharge and BL, baseflow lag) and two initial conditions (C_{ini}, initial soil moisture and BF_0, initial baseflow). The original version of ReFH, now known as ReFH1, was calibrated using the FEH99 rainfall depth-duration-frequency (DDF) model (Faulkner, 1999), and parameter estimation equations to allow the method to be applied to ungauged sites were provided by Kjeldsen (2007). Recent collaborative research by Wallingford HydroSolutions (WHS) and the Centre for Ecology & Hydrology (CEH) has led to the development of an upgraded version known as ReFH2 (Wallingford HydroSolutions, 2015). Design hydrographs from ReFH2 can be generated from either the FEH99 rainfall DDF model or from the new FEH13 rainfall model (Stewart and others, 2013). New sets of catchment descriptor equations for estimating model parameters at ungauged sites have been derived, to use with FEH99 in England, Wales, Scotland and Northern Ireland, and to use with FEH13 across the whole UK. These are used together with initial conditions which are specified according to the country of application and the rainfall model being used. To allow ReFH2 to be applied on the plot scale, alternative parameter estimation equations, which do not require drainage network descriptors, have also been developed (Wallingford HydroSolutions, 2015).

The development of ReFH2 has specifically addressed some of the issues identified with ReFH1 (Faulkner and Barber, 2009), particularly the performance of the method in urban catchments and permeable catchments. In addition, the definition of the ReFH1 adjustment coefficient, α, which effectively adjusts the estimate of C_{ini} so that the resultant peak flows approximate those derived from the FEH statistical method, has been reviewed. The α factor is no longer required when ReFH2 is used with the FEH13 rainfall model.

A framework for including the impact of urbanisation explicitly within the ReFH rainfall-runoff model proposed by Kjeldsen and others (2013) has been incorporated into ReFH2. The percentage runoff in ReFH2 is defined as a weighted sum of the contributions from the rural and urban parts of the catchment, with percentage runoff from the rural fraction estimated as in ReFH1. Runoff from the urban part of the catchment is further divided into that resulting from rain falling on pervious and impervious surfaces, with impervious surfaces assumed to have a constant runoff coefficient with a default value of 70% following Packman (1980). The total net urban runoff is routed by convolution with a separate unit hydrograph for which a shorter time-to-peak can be specified.

2.3 Improved FEH statistical method

The improved FEH statistical method (Kjeldsen and others, 2008) is an update to that published in Volume 3 of the FEH (Robson & Reed, 1999) and is implemented in the WINFAP-FEH software package. The method is widely used to estimate catchment flood peaks for a wide range of areas and return periods and a brief outline of the main aspects follows.

Index flood

The flood frequency estimation procedure consists of two stages. First, the index flood, defined as QMED, the median annual maximum flood, is estimated either from annual maximum peak flow data (AMAX) or from catchment descriptors. The most up-to-date version of the catchment descriptor equation as published in Kjeldsen and others (2008) is shown in Equation 1:

Equation 1 – FEH catchment descriptor equation for QMED

$$QMED_{CDS} = 8.3062AREA^{0.8510}0.1536^{(1000/SAAR)}FARL^{3.4451}0.0460^{BFIHOST^{\wedge}2}$$

AREA is catchment area (km^2), SAAR is standard average annual rainfall (mm), FARL is the FEH index of attenuation due to lakes and reservoirs and BFIHOST is baseflow index derived from the HOST soil classification.

Donor transfer

The estimate of QMED derived from Equation 1 can be adjusted using a QMED value estimated from flood peaks at a nearby gauging station using the method of donor transfer. Current guidance recommends that the donor catchment should be the closest to the site of interest measured centroid-to-centroid. In practice, however, more flexible approaches to donor transfer are often followed (Environment Agency, 2015).

Kjeldsen and others (2014) propose a modification to the donor transfer method to allow multiple gauged sites to contribute to the adjustment procedure, selected as the *n* closest gauged sites to the site of interest. Each donor site's contribution is weighted according to its distance from the site of interest and its distance from each other donor site. Typically, using six donor sites was found to be an optimal trade-off between the volume of information provided to the estimation procedure and the relevance (geographical distance) of that information. However, multiple donor transfer is impractical for manual calculation and therefore has not been widely adopted.

Urban adjustment

Flood frequency is known to be affected by the level of urbanisation in a catchment. Therefore, QMED in ungauged urban catchments should also be adjusted for urbanisation as shown in Equation 2:

Equation 2 – FEH QMED adjustment for ungauged urban catchments

$$QMED = UAF \times QMED_{RURAL}$$

UAF is an urban adjustment factor calculated from catchment descriptors.

Growth curve

Separately from the index flood, a growth curve (z_T) must be determined if any flood peak estimates other than QMED are desired. The growth curve is a dimensionless relationship that expresses QT over QMED, where QT is the *T*-year flood peak (see Equation 3):

Equation 3 – FEH statistical growth curve equation

$$QT = Z_T \times QMED$$

This is typically estimated by using L-moments (Hosking & Wallis, 1997) to parameterise an appropriate statistical distribution. For estimating growth curves at ungauged sites, a set of gauged catchments that are hydrologically similar to the subject site are chosen. This 'pooling-group' of catchments is selected on the basis of the AREA, SAAR, FARL and FPEXT catchment descriptors. The annual maximum series at each catchment is standardised by its median value (QMED) and the L-moments of each individual series are pooled to form a single set of L-moments typical of the catchment of interest. If the catchment is urban, these L-moments will be adjusted to reflect the fact that flood growth curves are generally less steep in urbanised catchments (that is, urbanisation increases QMED but tends not to affect very rare events). Based on the properties of the pooled L-moments, an appropriate statistical distribution is chosen to compute the growth curve: in the UK, this has been shown to typically correspond to the generalised logistic (GLO) distribution (Kjeldsen & Prosdocimi, 2015). The weighted averages of the higher-order L-moments L-CV and L-SKEW are used to compute the dimensionless growth curve, and QT flood peak estimates are found by multiplying the estimated QMED with the value of the growth curve corresponding to a probability of exceedance ($1 - 1/T$).

Enhanced single-site analysis

The FEH's basic recommendation is to use pooled analysis to construct the flood growth curve unless there is a flood peak record of at least twice the length of the required return period in years at the site of interest. If some flood peak data are available at the site of interest, then these may be enough to estimate common floods, such as QMED. However, the data are unlikely to be suitable to estimate rarer floods, such as *Q*100. Kjeldsen and others (2008) propose using a new method, enhanced single-site analysis, which provides a framework in which to combine gauged data at the site of interest with gauged data from other sites, in order to create a larger data set that is suitable for estimating rarer floods. In enhanced single-site analysis, the index flood is derived from the gauged series at the site of interest. A pooling-group is then formed, to give at least 500 annual maxima when combined with the at-site data. The L-moments of the series in the pooling-group are pooled, with enhanced weighting given to the L-moments calculated from data at the site of interest. The pooled L-moments are used to parameterise the growth curve for the site of interest.

2.4 MacDonald and Fraser's equation

MacDonald and Fraser (2013) published an improved equation for the median annual flood in small catchments, developed with the specific aim of reducing the error in QMED estimates relative to the improved FEH statistical method for catchments between 0.5 and 25 km^2 in area. This is shown in Equation 4:

Equation 4 - MacDonald and Fraser's equation for the median annual flood in small catchments

$$QMED = 6.120AREA^{0.758}0.288^{(1000/SAAR)}0.042^{BFIHOST^2}$$

The authors do not publish any corresponding procedure for generating growth curves, and there is no guarantee that the improved FEH statistical method is equally appropriate when this QMED regression equation is used. Therefore, the improved FEH statistical procedure remains the only index-flood method that has been developed and tested for estimating longer return period flood peaks in the UK.

The development of MacDonald and Fraser's small catchment QMED equation used a set of 135 catchments with known descriptors and gauged AMAX series for calibration. A total of 104 of these catchments were essentially rural ($URBEXT_{2000}$ < 0.03), while 31 were at least slightly urbanised. Gauged QMED was regressed against individual catchment descriptors to determine those that were most correlated. Then, multiple linear regression was performed with various combinations of those descriptors, to find which set explained the most variation in QMED. The performance of the final model (Equation 4) was evaluated against the performance of the existing FEH equation (Equation 1) using the same 135 catchments as were used for calibration. The reported factorial standard error (fse) of MacDonald and Fraser's equation (Equation 4) is a slight improvement over the improved FEH statistical QMED equation (1.64 before urban adjustment factor, 1.55 after), though they do note that their equation was evaluated against its own calibration data set and so was likely to have a slight inbuilt advantage.

Notably, the MacDonald and Fraser equation does not account for differing levels of urbanisation. While the improved FEH statistical QMED estimate would be multiplied by an urban adjustment factor in non-rural catchments, MacDonald and Fraser recommend that their equation is used in the form of Equation 3 for both urban and rural catchments as the addition of the $URBEXT_{2000}$ descriptor to the multiple linear regression had no effect on the proportion of variance explained or the factorial standard error. Furthermore, no mechanism for donor transfer is proposed. It is unlikely that the donor transfer mechanism developed for the improved FEH statistical method would behave optimally if applied in its existing form, as the residuals of MacDonald and Fraser's equation are, by definition, different from the residuals of the FEH equation and therefore may not follow the same spatial patterns.

3. Data

In order to evaluate the results produced by flood estimation methods, it is necessary to determine benchmark ‘least uncertain’ target values, which are assumed to be as close as possible to the ‘true’ values. At any site, the closest and least-uncertain estimate for a specified flood peak normally derives entirely or in part from actual observed flood peak data at that site, with records as short as five years being shown to reduce uncertainty considerably over a purely ungauged estimate (Kjeldsen, 2015). The most commonly available observed flow data are series of annual maxima (AMAX) and series of peak-over-threshold (POT) data. As their names imply, AMAX series feature one entry (the largest value) per water year (starting on 1 October and ending on the following 30 September), while POT series include every independent flood peak that exceeds a threshold magnitude. Although POT series normally provide many times more information on large events than AMAX series, the current methods used for flood frequency estimation in the UK do not formalise the integration of POT data into single-site or enhanced single-site analyses. Furthermore, extracting POT series is formally more complex than extracting AMAX series, since rules need to be defined to ensure that only truly independent flood peaks are extracted, and an appropriate threshold must be chosen. See Bayliss & Jones (1993) for a discussion on using POT data for flood frequency analysis in the UK.

The flood peak estimation models discussed in Section 2 require a range of catchment descriptors as input data. Therefore, at each of the gauged sites identified for benchmarking, catchment-average values of these descriptors are required in order to generate ungauged estimates of QMED. These values are readily available: the FEH Web Service, and its predecessor the FEH CD-ROM, map catchment-average values of all required descriptors (and more) on a 50-metre grid for every UK catchment larger than 0.5 km^2. Where catchments are smaller than this, catchment-average values can feasibly be manually computed from the underlying grids.

In this study, ungauged estimates of the median annual flood (QMED) are benchmarked against at-site observed QMED, taken as the median of the at-site AMAX series. Rarer flood peaks (QT) are benchmarked against enhanced single-site analyses that pool the full at-site AMAX series, at an enhanced weight, with full AMAX series from several hydrologically similar catchments, at a reduced weight, for a target pooled record length of 500 years (Kjeldsen and others, 2008).

3.1 Data sources

A total of 217 gauged small catchments across the UK, with areas up to 40.9 km^2, were identified in Task 1 of this project as having suitably robust annual maxima records to permit at least estimation of QMED and, in some cases, pooling of the entire record to estimate peak flows at higher return periods (QT estimates). While the project is still focused mainly on developing methods for catchments up to 25 km^2, a ‘transition zone’ of catchments slightly larger than 25 km^2 is also considered, to ensure seamless integration between any proposed small catchment methods and the existing generic methods.

The set of benchmark catchments comprises:

- 151 catchments identified by JBA, in Task 1 of the project, through a review of numerous sources
    - 119 of the 151 suitable catchments are present in the current NRFA peak flow data set (v3.3.4, NRFA, 2014). Annual maxima (AMAX) and peaks-over-threshold (POT) series were provided for each of these. For stations that remained open beyond the last update to the data set, the freely available AMAX and POT series were extended to either the end of calendar year 2014 (POT) or water year 2013 (AMAX), or the station closing date, if sooner. Catchment descriptor (CD3) files, taken directly from the NRFA peak flow data set v3.3.4 were provided, unedited, for all catchments. These files were used to inform the QMED and pooling suitability of each catchment individually
    - the remaining 32 catchments are not present in the current NRFA peak flow data set. AMAX series, to water year 2013 or the station closing date, were provided for all of these. POT series, to the end of calendar year 2014 or the station closing date, were provided for 27 stations only. The FEH CD-ROM version 3 was used to export CD3 files for each of the 32 catchments manually. All these catchments were considered suitable for estimating QMED only
- 135 catchments identified by MacDonald & Fraser (2013) for their review of QMED estimation in catchments under 25 km^2. Every catchment is associated with catchment-average descriptor values exported from the FEH CD-ROM version 2, a QMED value and the number of annual maxima from which that QMED value is calculated. In addition:
    - 66 of these 135 catchments were also identified in Task 1 of the current project. In all 66 cases, flood peak records, catchment descriptor values and QMED/pooling suitability, as found in Task 1, are used in the study data set. MacDonald and Fraser’s records for these catchments are not used, to avoid duplication of existing data
    - the remaining 69 catchments are unique to MacDonald and Fraser’s (2013) study. Of these, 64 are of sufficient quality for inclusion in this study. AMAX series have been identified for 39 of the 64 accepted additional catchments, allowing them to be used in QT analysis
- a further two catchments identified by MacDonald & Fraser (2013) but rejected by them solely for being over 25 km^2. These are both present in the NRFA peak flow data set, where they are marked ‘suitable for pooling’. In addition, one of these two catchments is heavily urbanised, and this class of catchment is relatively rare. AMAX, POT and CD3 files for both catchments were added to the study data set from the NRFA peak flow data set.

The 217 suitable catchments are mapped in Figure 1, showing the available record length and distinguishing between those catchments that have full AMAX series, for enhanced single-site analysis and those that have single QMED values, for benchmarking QMED only. A summary of the records available are shown in Tables 1, 2 and 3: ‘high-quality’ QMED data in Table 1, ‘extended’ QMED in Table 2, and QT estimation in Table 3 (‘high-quality’ and ‘extended’ QT sets). ‘High-quality’ in this case means that the station is considered suitable for QMED estimation (or pooling) by the NRFA. The minimum requirement for this is that the gauged value of QMED (or the third largest AMAX value in the case of pooling) has been assessed as within 30% of the true value. Stations suitable for pooling often benefit from a much smaller than 30% uncertainty in QMED. Further information about these data sets is given in Sections 4 and 5 and a full list of the gauging stations used in the analysis is provided in Appendix A. Appendix B gives further information about quality control and standardisation.

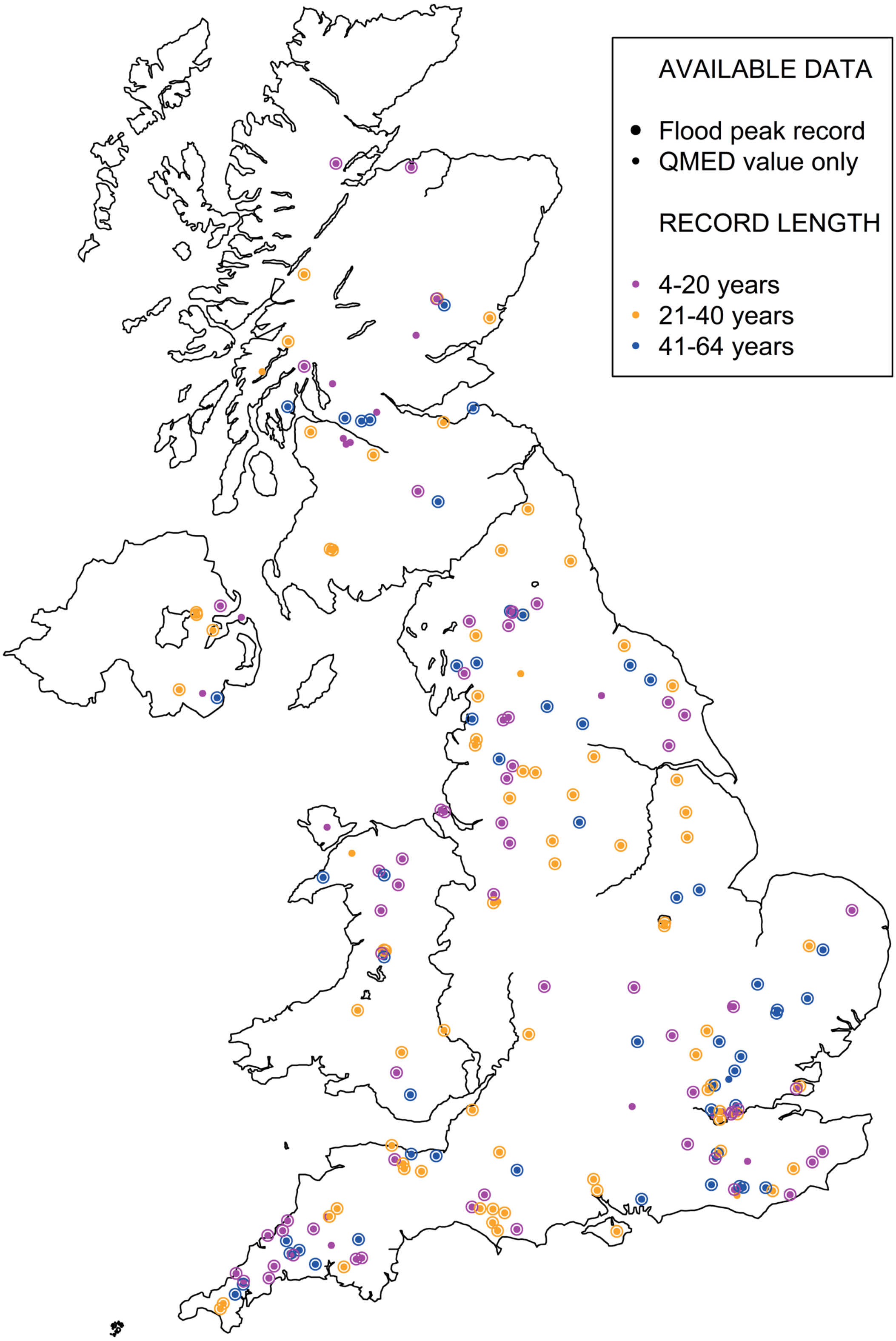

Figure 1 - Location map of sites identified for phase 2 study

The map in Figure 1 shows the 217 suitable catchments. Available record length is shown by purple dots (4-20 years), yellow dots (21-40 years) and blue dots (41-64 years). QMED value data is available where locations are marked with a coloured dot only, whilst flood peak record data is available for those marked with a dot with a surrounding circle.

Table 1 – 'High-quality' QMED data available to the analysis

Data type ('high-quality' QMED)	Number of stations
Updated NRFA peak flow data	119
Additional stations identified in Task 1.2	32
Added to NRFA peak flow data shortly after Task 1.2	2
Total number of stations	153

Table summary: A summary of the 'high-quality' QMED data records available.

Table 2 – 'Extended' QMED data available to the analysis

Data type ('extended' QMED)	Number of stations
All data listed in Table 1 (above)	153
MacDonald & Fraser stations	64
Total number of stations	217

Table summary: A summary of the 'extended' QMED data records available.

Table 3 – 'High-quality' QT and 'extended QT' data available to the analysis

Data type	Data type	Number of stations
'High-quality' QT	Gauged AMAX data suitable for pooling	58
'Extended' QT	Gauged AMAX data	192

Table summary: A summary of the 'high-quality' QT and 'extended QT' data records available.

Summary statistics of the available QMED values and AMAX series are presented in Table 4. This shows that the total number of stations and annual maxima available for the current phase of the study is approximately three times the amount available for phase 1. For direct comparison with phase 1, the summary statistics for phase 2 are shown for both the full data set of suitable catchments and the subset of catchments under 25 km^2 in area. Even comparing the phase 1 and phase 2 data sets on a like-for-like basis, there are more than twice as many catchments and almost twice as many annual maxima available in phase 2.

Table 4 - Summary of AMAX data set, with comparison to phase 1 study

Statistic	Phase 1 (< 25 km^2)	Phase 2 (<25 km^2)	Phase 2 (<40.9 km^2)
Number of gauges	72	158	217
Number of above gauges which are suitable for QT estimation	N/A	(*133*)	(*192*)
Shortest record length (years)	9	4	4
Longest record length (years)	58	64	64
Mean record length (years)	30.9	24.8	27.4
Total number of AMAX	2225	3925	5946
Number of above AMAX which are suitable for QT estimation	N/A	(*3444*)	(*5465*)

As no flood peak records (except for a pre-calculated QMED value) could be identified for 25 of the gauges identified solely by MacDonald and Fraser, enhanced single-site benchmarking of QT estimation was not possible at 25 stations. For this reason, Table 4 also specifies how many catchments have information available for estimating rarer flood peaks and how much information is available. This is bracketed and in italics.

3.2 Data description

Figures 2 to 6 show the distribution of record length, BFIHOST, SAAR and $URBEXT_{2000}$ for the 217 sites contained in the study data set. Table 5 summarises the median and range of these catchment descriptors in order to put them into context with the NRFA peak flow data set overall.

The median values of BFIHOST and SAAR are very similar between the study data set and NRFA peak flow data set. Furthermore, the range of all three descriptor values is slightly larger in the study data set. However, the study data set is somewhat less urbanised than the NRFA peak flow data set, which is itself not highly urbanised in general. Figure 1 shows a sparse representation of catchments in the heavily urbanised West Midlands in the study data set. Unfortunately, many small, urbanised catchments in that area of the UK had to be rejected because of insufficient data quality.

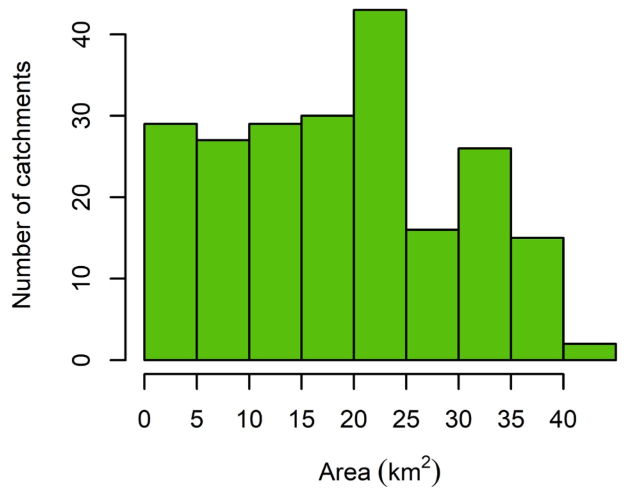

Figure 1 - Histogram of catchment area for the study data set

The histogram in Figure 2 shows the distribution of catchments on the y-axis against their area in km^2, derived from the IHDTM on the x-axis.

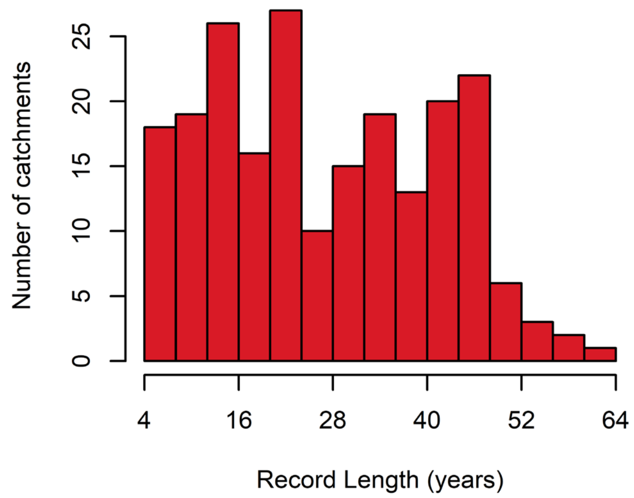

Figure 3 - Histogram of record length for the study data set

The histogram in Figure 3 shows the distribution of catchments on the y-axis by record length on the x-axis.

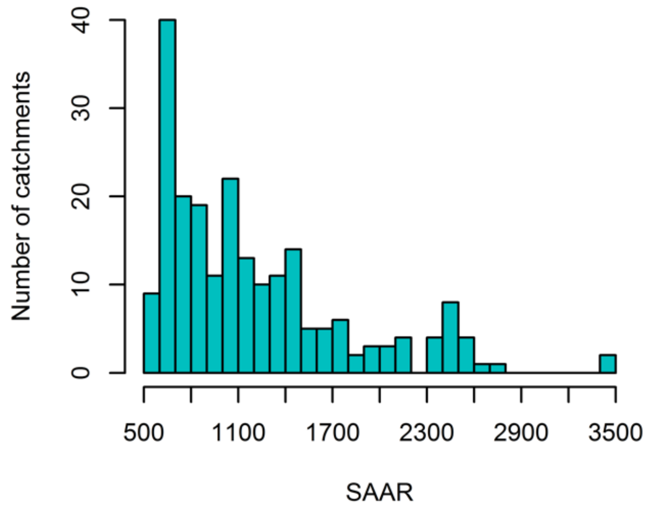

Figure 4 - Histogram of SAAR for the study data set

The histogram in Figure 4 shows the distribution of catchments on the y-axis against average annual rainfall in the standard period (1961-1990) in millimetres (SAAR) on the x-axis.

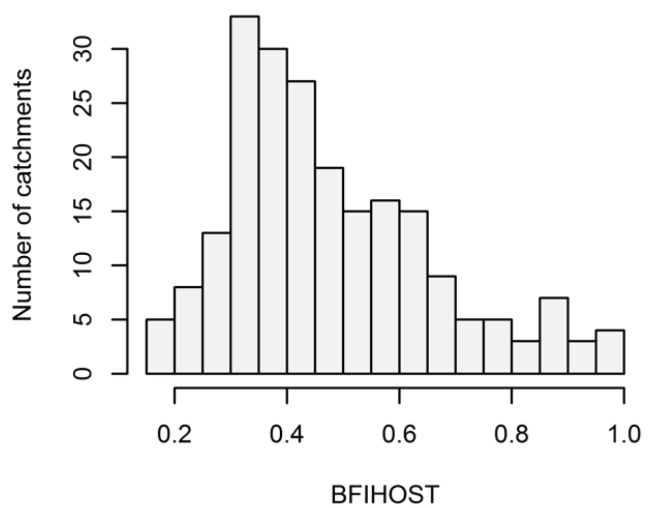

Figure 5 - Histogram of BFIHOST for the study data set

The histogram in Figure 5 shows the distribution of catchments on the y-axis against the base flow index (BFI) on the x-axis.

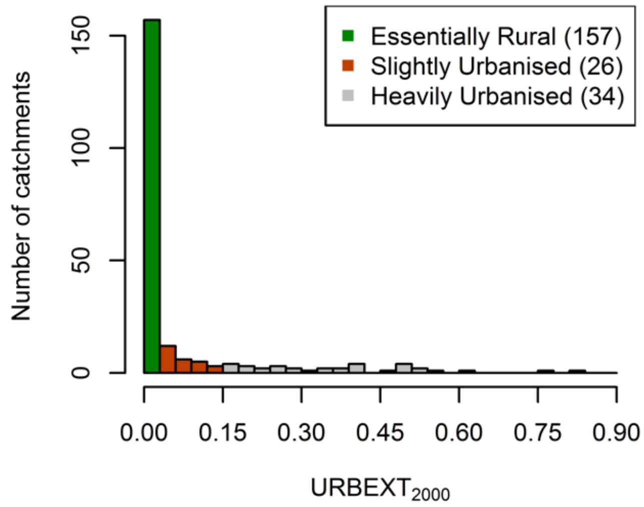

Figure 6 - Histogram of $URBEXT_{2000}$ for the study data set

The histogram in Figure 6 shows the distribution of catchments on the y-axis against their index of urban and suburban landcover in the year 2000 ($URBEXT_{2000}$) on the x-axis.

Correlation between catchment area and record length is similar in magnitude to that between SAAR and BFIHOST or $URBEXT_{2000}$, but positive, indicating that the smallest catchments in the study data set are unable to provide data in the same quantities as the relatively larger catchments. FPEXT is moderately correlated with SAAR ($r^2 = -0.47$), the physical basis of which may lie in the trend for the highest-SAAR catchments to be more likely to be upland (that is, near the sources of rivers) and steep. FPEXT is also moderately correlated with $URBEXT_{2000}$ ($r^2 = 0.34$), reflecting the heavy urbanisation of the lower Thames basin and, unfortunately, highlighting the link between heavy development and floodplains.

A comprehensive list of catchment descriptor values and other properties is presented in Appendix A.

Table 5 - Catchment descriptor summaries for the study data set and NRFA peak flow data v3.3.4

Statistic	NRFA peak flow	Study data set
Count	962	217
BFIHOST (minimum)	0.172	0.172
BFIHOST (maximum)	0.974	0.985
BFIHOST (median)	0.468	0.437
SAAR (minimum)	555	552
SAAR (maximum)	2913	3473
SAAR (median)	971	1043
$URBEXT_{2000}$ (minimum)	0.0000	0.0000
$URBEXT_{2000}$ (maximum)	0.5917	0.8110
$URBEXT_{2000}$ (median)	0.0089	0.0050

Table summary: the median and range of the catchment descriptors.

A correlation matrix between several properties of the study data set is presented in Figure 7. This shows mainly weak or very weak correlations between each pair of properties in the data set, with only nine (of 28) greater than 0.3. Four of these are between PROPWET and one other descriptor: BFIHOST ($r^2 = -0.4046$); SAAR ($r^2 = 0.8359$); $URBEXT_{2000}$ ($r^2 = -0.3356$) and FPEXT ($r^2 = -0.4229$). The moderate correlation between PROPWET and various other catchment descriptors may be interpreted as it being either superfluous or a reasonable alternative/guide to various other catchment characteristics simultaneously.

Two other noteworthy correlations are SAAR-BFIHOST ($r^2 = -0.37$) and SAAR-$URBEXT_{2000}$ ($r^2 = -0.37$). Both are negative, generally relating higher SAAR values to lower BFIHOST and $URBEXT_{2000}$ values and explained by the lack of heavy urbanisation in mountainous regions. Correlation between BFIHOST and $URBEXT_{2000}$ is almost non-existent ($r^2 = 0.001$).

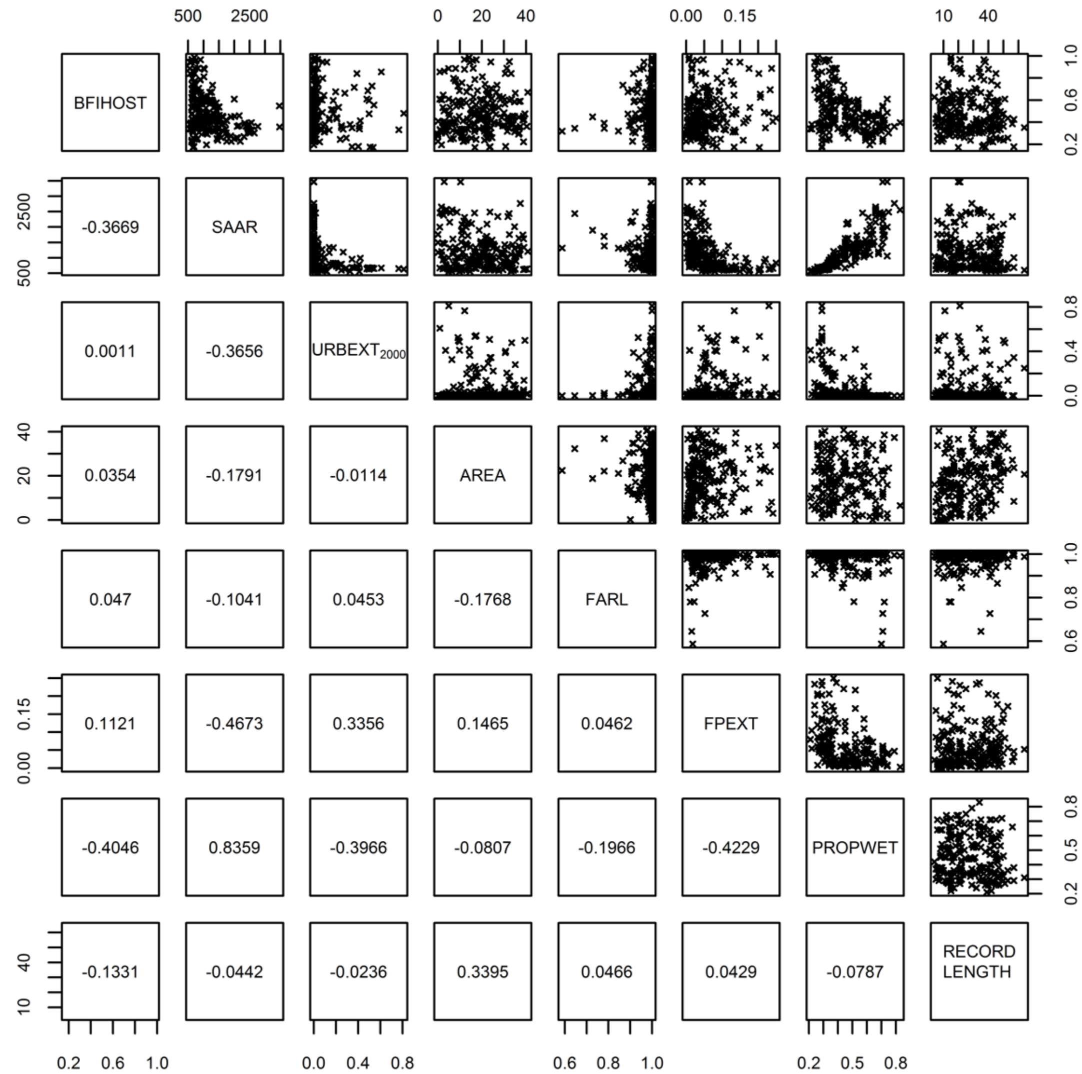

Figure 7 - Correlation matrix between properties of the study data set

4. Evaluating QMED estimation methods

4.1 Introduction

This section evaluates the performance of the existing FEH methods (FEH statistical and ReFH methods) and the MacDonald and Fraser (MF) method for estimating QMED, the median annual flood. The data set of small catchment QMED values and flood peak AMAX values has been divided into a ‘high-quality’ subset and an ‘extended’ data set as detailed in Table 6. Section 4.2 summarises the available data and Section 4.3 outlines the methods that have been evaluated. The statistics used to evaluate the different methods are presented in Section 4.4 and an overview of the performance of the methods for all catchments is given in Section 4.5. The analyses applied to the ‘high-quality’ and ‘extended’ data sets are described in Sections 4.6 and 4.7 respectively, and the main conclusions drawn from the analysis are outlined in Section 4.8.

4.2 Data summary

A data set of 217 small, gauged catchments is summarised in Section 3 and described fully in Appendix A. For all 217 catchments, data are sufficient to allow ungauged QMED estimation via any of the index flood methods being considered, and gauged QMED estimation to act as a benchmark. However, the rainfall-runoff methods (ReFH1 and ReFH2) are unsuitable for using on the 10 catchments with values of FARL < 0.9 unless reservoir routing is applied. This was not considered practical within an automated analysis and therefore the following analysis directly compares the rainfall-runoff and index flood methods on the 207 catchments without significant attenuation due to lakes and reservoirs.

Within these 207 catchments, 114 are marked in the NRFA peak flow data set v3.3.4 as ‘suitable for QMED estimation’, the criterion for which is discussed in Section 3.1. In addition, a further 32 were rigorously checked in Task 1 of this project and found to meet the same criterion. These 146 QMED-suitable catchments form a ‘high-quality’ data set, which is smaller and less comprehensive than the 207 catchment ‘extended’ data set but suffers from lower uncertainty around the value that the methods should target.

Table 6 shows the total number of catchments in each data set and, as urbanisation is an important factor in characterising flood response, the numbers of catchments in three sub-classes are also shown. In the text, catchments with $URBEXT_{2000}$ < 0.03 are considered to be ‘essentially rural’, catchments with $URBEXT_{2000} \geq 0.15$ are ‘heavily urbanised’ and catchments with intermediate $URBEXT_{2000}$ values are ‘moderately urbanised’. ‘Heavily urbanised’ and ‘moderately urbanised’ catchments together are termed ‘urbanised’, with no prefix.

Table 6 - Number of catchments per data set

Data set	Essentially rural (URBEXT$_{2000}$ < 0.03)	Moderately urbanised (0.03 ≤ URBEXT$_{2000}$ < 0.15)	Heavily urbanised (URBEXT$_{2000}$ ≥ 0.15)	Total
'High-quality'	103	19	24	146
'Extended'	150	24	33	207

4.3 Method summary

An overview of the FEH statistical and ReFH methods of flood peak estimation, as well as the QMED equation of MacDonald and Fraser (2013), is given in Section 2. Here several variants of the current methods are tested for their accuracy in modelling QMED values in small catchments. The methods are:

- improved FEH statistical method (Kjeldsen and others, 2008), with the updated urbanisation adjustment procedure (Kjeldsen, 2010) and no donor transfer
- improved FEH statistical method (Kjeldsen and others, 2008), with the updated urbanisation adjustment procedure and donor transfer from one site
- improved FEH statistical method, with the updated urbanisation adjustment procedure and donor transfer from six sites (Kjeldsen and others, 2014)

In all cases, donors were selected from the 848 QMED-suitable members of the NRFA peak flows data set v3.3.4, purely on the distance between the centroid of the catchment of interest and the centroid of the donor catchment. Although hydrological judgement is often used in selecting donors, the large number of catchments considered here necessitated an automatic selection procedure. A small centroid-to-centroid distance implies proximity and therefore may imply similar soil types, the same hydrometric area and that the same events caused the annual maxima. It is for this reason that Kjeldsen and others (2014) found six to be the typical optimum number of donors.

- MacDonald and Fraser's 'Improved method for estimating the median annual flood for small ungauged catchments in the United Kingdom' (MacDonald & Fraser, 2013). Donor transfer and urban adjustment are not implemented
- the revitalised flood hydrograph (ReFH1) method (Kjeldsen, 2007)
- the revitalised flood hydrograph (ReFH2) method (Kjeldsen and others, 2013), which includes updated modelling of the urban component of the hydrograph as described by Kjeldsen and others (2013), assuming winter seasonality for all catchments and all other parameterisation as default. This analysis pre-dates the recent releases of ReFH2.1 and 2.2, and so uses design event inputs from the

FEH99 rainfall depth-duration-frequency (DDF) model. Appendix C provides a preliminary comparison of the performance of ReFH2 with both the FEH99 rainfall model and the FEH13 rainfall model.

4.4 Quantifying methods of QMED estimation

The performance of the methods is evaluated quantitatively through three statistics: relative error, ln-error and root-mean-square error (RMSE).

Relative error measures bias in the estimate relative to the size of observed QMED (Equation 5). This would be equivalent to bias if the observed QMED were the true QMED value.

Equation 5 – Relative error equation

$$\sum_{i=1}^{n}\left[\left(QMED_{MOD,i} - QMED_{OBS,i}\right)/QMED_{OBS,i}\right]/n$$

A measure of bias on the natural logarithmic scale is given by ln-error (Equation 6):

Equation 6 – Ln-error equation

$$\sum_{i=1}^{n}\left[ln\left(QMED_{MOD,i}\right) - ln\left(QMED_{OBS,i}\right)\right]/n$$

Unlike relative error, performance in ln-error provides a measure of error that is evenly balanced between over- and under-estimation. For example, a modelled QMED estimate that is half the gauged value will have the same absolute value of ln-error as a modelled estimate that is twice the gauged QMED.

RMSE (Equation 7) measures the spread of errors:

Equation 7 – RMSE equation

$$\left(\sum_{i=1}^{n}\left[ln\left(QMED_{MOD,i}\right) - ln\left(QMED_{OBS,i}\right)\right]^{2}/n\right)^{0.5}$$

The exponential of RMSE is the factorial standard error, a common measure of uncertainty for hydrological variables. In all cases, the MOD and OBS subscripts refer to modelled and observed (that is, median of AMAX) values, and *n* refers to the number of catchments being considered.

4.5 Performance overview

Figure 8 plots the QMED or 2-year flood peak value found by the six methods being considered, against the observed median of annual maxima. This plot gives a very broad overview of what is discussed in more detail throughout the following sections but generally shows good agreement between all methods and observed QMED, with more scatter at lower observed QMED values and a possible tendency towards underestimation at higher QMED values.

All 217 catchments, including those with FARL < 0.9, are plotted in Figure 8 to indicate how the tested methods perform in catchments with significant lake and/or reservoir attenuation. As previously stated, the rainfall-runoff methods are unsuitable to use on these catchments unless reservoir routing is incorporated. Performance metrics, reported in tables throughout Sections 4.6 to 5.5, exclude catchments with FARL < 0.9 so that all index-flood and rainfall-runoff methods can be compared over identical data sets. Grey crosses are used to show the performance of catchments with FARL < 0.9 in figures throughout Sections 4.6 to 5.5. However, they are not considered in calculating the correlation coefficients shown on these figures.

Four catchments are identified on Figure 8 as potential outliers (25808 Burnt Hill at Moor House, 205034 Woodburn at Control, 27032 Hebden Beck at Hebden and 39082 Graveney at Longley Road), as all methods, whether rainfall-runoff or statistical, FEH or non-FEH, produce estimates of QMED that are poor, but that agree with each other. All these catchments are in the ‘high-quality’ data set, suggesting that the QMED values that all methods struggle to model are not massively in error. As no method was able to produce a close estimate of QMED for catchments 25808 or 205034, this is likely to be due to the severe mismatch between the nominal and DTM-derived catchment areas. DTM area has been accepted throughout this study. Catchment 39082 was not used in the calibration of any method tested here and is more extremely heavily urbanised than any catchment used to calibrate any FEH method ($URBEXT_{2000}$ = 0.811). Catchment 27032 is deemed suitable for pooling, and QMED is gauged, not extrapolated from a rating curve. However, the presence of karst may affect flood flows, and dye tracing has shown that water entering Mossdale Caverns (in the topographic catchment) is transferred out (Faulkner, pers. comm.)

The presence of these four outlier catchments in the data sets is likely to have some effect on all statistics reported in the following sections. The outlier catchments are discussed in more detail in later sections as their effects become more apparent.

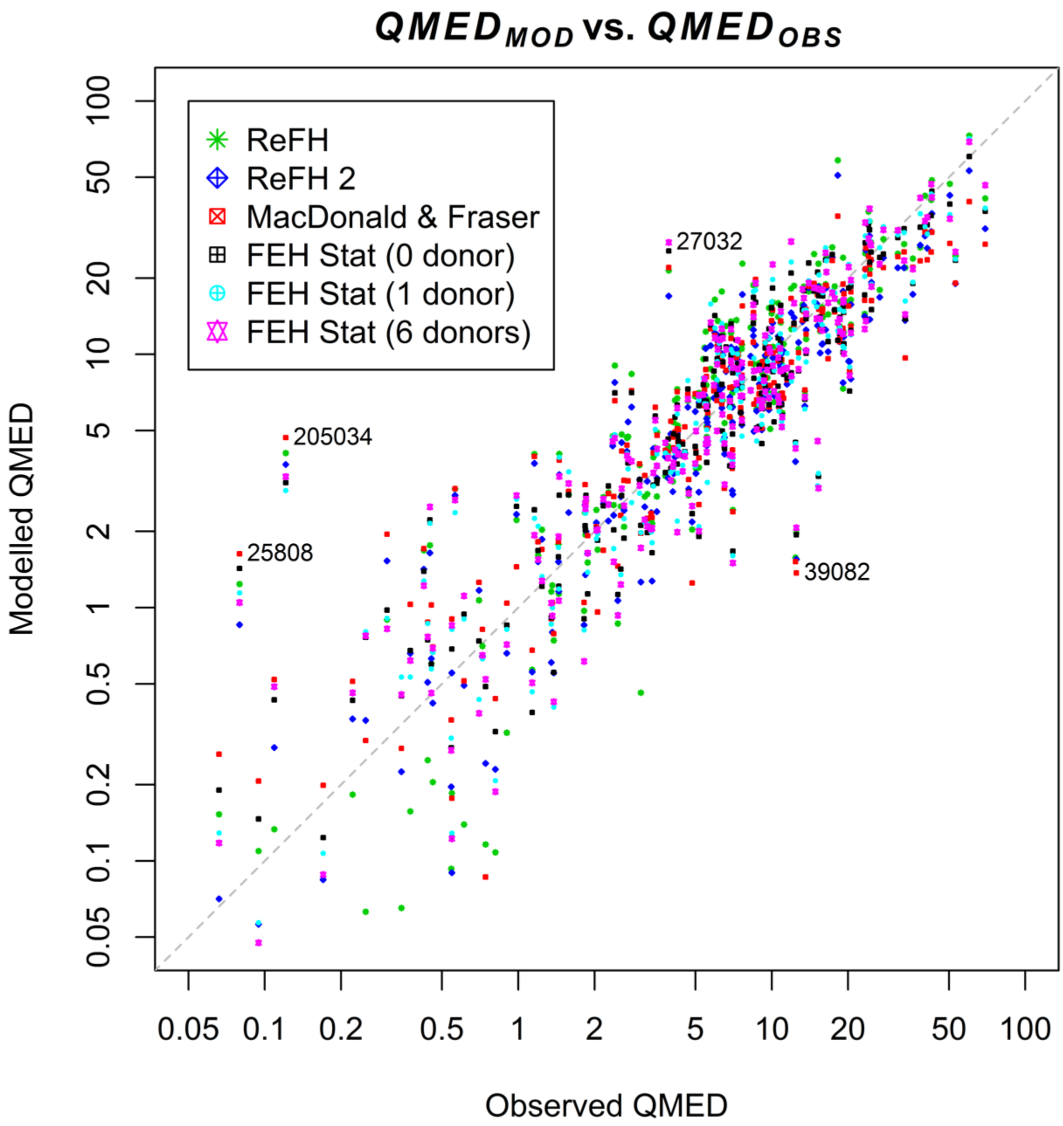

Figure 8 - Modelled vs. observed *QMED* at 217 sites, including those with FARL < 0.9

The scatter graph in Figure 8 plots the QMED or 2-year flood peak value found by the six methods being considered (0.05 to 100, on the y-axis), against the observed median of annual maxima (0.05 to 100, on the x-axis):

- ReFH: green stars
- ReFH2: dark blue diamonds
- MacDonald & Fraser: red squares
- FEH Stat (0 donor): black squares
- FEH Stat (1 donor): light blue circles
- FEH Stat (6 donors): pink stars

4.6 Analysis of high-quality data set

In rural catchments

A total of 103 catchments in the 'high-quality' data set are considered essentially rural, while 43 are moderately to heavily urbanised (Table 6). Subsets of catchments with high- and low-BFIHOST and SAAR values are also considered individually within the two groups of urbanisation level, in order to identify any potential trends or behaviours specific to wet, dry, permeable or less permeable catchments. A value of SAAR of 800 mm was chosen to distinguish between wet and dry catchments. A value of BFIHOST equal to 0.65 was chosen to distinguish between permeable and impermeable catchments as used in phase 1 of this study.

Figure 9 and Tables 7, 8, 9 and 10 quantify the relative error, ln-error, RMSE and FSE of QMED, as estimated by the six methods being considered, in 'high-quality' rural catchments as a whole, and when subset by BFIHOST and SAAR. These show that the best performing method in terms of relative error, ln-error, RMSE and FSE overall is ReFH2. ReFH2 is also the best performing method in wetter (SAAR ≥ 800) and less permeable (BFIHOST < 0.65) catchments in terms of relative error, ln-error and RMSE. In addition, it is the most consistent performer in terms of ln-error across the whole group of QMED-suitable catchments and all four subsets of the group, showing no obvious dependency on either BFIHOST or SAAR. The worst method in terms of relative error and ln-error is MacDonald and Fraser's, while the worst in terms of RMSE is ReFH1. This shows that ReFH1's unremarkable level of mean error on either relative or ln-scale masks a large spread of individual errors.

Table 7 - Relative error in QMED estimated by six current and recent flood estimation methods in rural catchments

Subset	All (103 sites)	SAAR ≥ 800 (82 sites)	SAAR < 800 (21 sites)	BFIHOST ≥ 0.65 (14 sites)	BFIHOST < 0.65 (89 sites)
ReFH1	0.7172	0.8914	0.0372	0.2984	0.7831
ReFH2	0.5138	0.5712	0.2896	0.6761	0.4883
MacD & Fraser	0.8931	0.9504	0.6696	1.1818	0.8477
Stat. (0 donors)	0.6555	0.7961	0.1065	0.7766	0.6364
Stat. (1 donor)	0.5727	0.6832	0.1411	0.6257	0.5644
Stat. (6 donors)	0.5712	0.6851	0.1261	0.6377	0.5607

The best performing method in each subset was identified as:

- ReFH2 for all sites
- ReFH2 for SAAR ≥ 800
- ReFH1 for SAAR < 800
- ReFH1 for BFIHOST ≥ 0.65
- ReFH2 for BFIHOST < 0.65

Table 8 – Ln-error in QMED estimated by six current and recent flood estimation methods in rural catchments.

Subset	All (103 sites)	SAAR ≥ 800 (82 sites)	SAAR < 800 (21 sites)	BFIHOST ≥ 0.65 (14 sites)	BFIHOST < 0.65 (89 sites)
ReFH1	0.1115	0.1948	-0.2137	-0.3121	0.1782
ReFH2	0.0201	0.0242	0.0042	0.1367	0.0018
MacD & Fraser	0.1707	0.1508	0.2484	0.4348	0.1291
Stat. (0 donors)	0.1180	0.1870	-0.1513	0.2560	0.0963
Stat. (1 donor)	0.0823	0.1338	-0.1189	0.1276	0.0751
Stat. (6 donors)	0.0627	0.1112	-0.1269	0.0830	0.0595

The best performing method in each subset was identified as:

- ReFH2 for all sites
- ReFH2 for SAAR ≥ 800
- ReFH2 for SAAR < 800
- Stat. (6 donors) for BFIHOST ≥ 0.65
- ReFH2 for BFIHOST < 0.65

Table 9 – RMSE in QMED estimated by six current and recent flood estimation methods in rural catchments

Subset	All (103 sites)	SAAR ≥ 800 (82 sites)	SAAR < 800 (21 sites)	BFIHOST ≥ 0.65 (14 sites)	BFIHOST < 0.65 (89 sites)
ReFH1	0.7521	0.7477	0.7693	1.1671	0.6636
ReFH2	0.6754	0.6739	0.6810	0.9132	0.6298
MacD & Fraser	0.7317	0.7365	0.7130	0.9393	0.6935
Stat. (0 donors)	0.6904	0.6943	0.6750	0.8494	0.6620
Stat. (1 donor)	0.6907	0.6944	0.6761	0.8574	0.6607
Stat. (6 donors)	0.6907	0.6939	0.6782	0.9135	0.6487

The best performing method in each subset was identified as:

- ReFH2 for all sites
- ReFH2 for SAAR ≥ 800
- Stat. (0 donors) for SAAR < 800
- Stat. (0 donors) for BFIHOST ≥ 0.65
- ReFH2 for BFIHOST < 0.65

Table 10 – FSE in QMED estimated by six current and recent flood estimation methods in rural catchments

Subset	All (103 sites)	SAAR ≥ 800 (82 sites)	SAAR < 800 (21 sites)	BFIHOST ≥ 0.65 (14 sites)	BFIHOST < 0.65 (89 sites)
ReFH1	2.121	2.112	2.158	3.213	1.942
ReFH2	1.965	1.962	1.976	2.492	1.877
MacD & Fraser	2.079	2.089	2.040	2.558	2.001
Stat. (0 donors)	1.995	2.002	1.964	2.338	1.939
Stat. (1 donor)	1.995	2.003	1.966	2.357	1.936
Stat. (6 donors)	1.995	2.002	1.970	2.493	1.913

The best performing method in each subset was identified as:

- ReFH2 for all sites

- ReFH2 for SAAR ≥ 800
- Stat. (0 donors) for SAAR < 800
- Stat. (0 donors) for BFIHOST ≥ 0.65
- ReFH2 for BFIHOST < 0.65

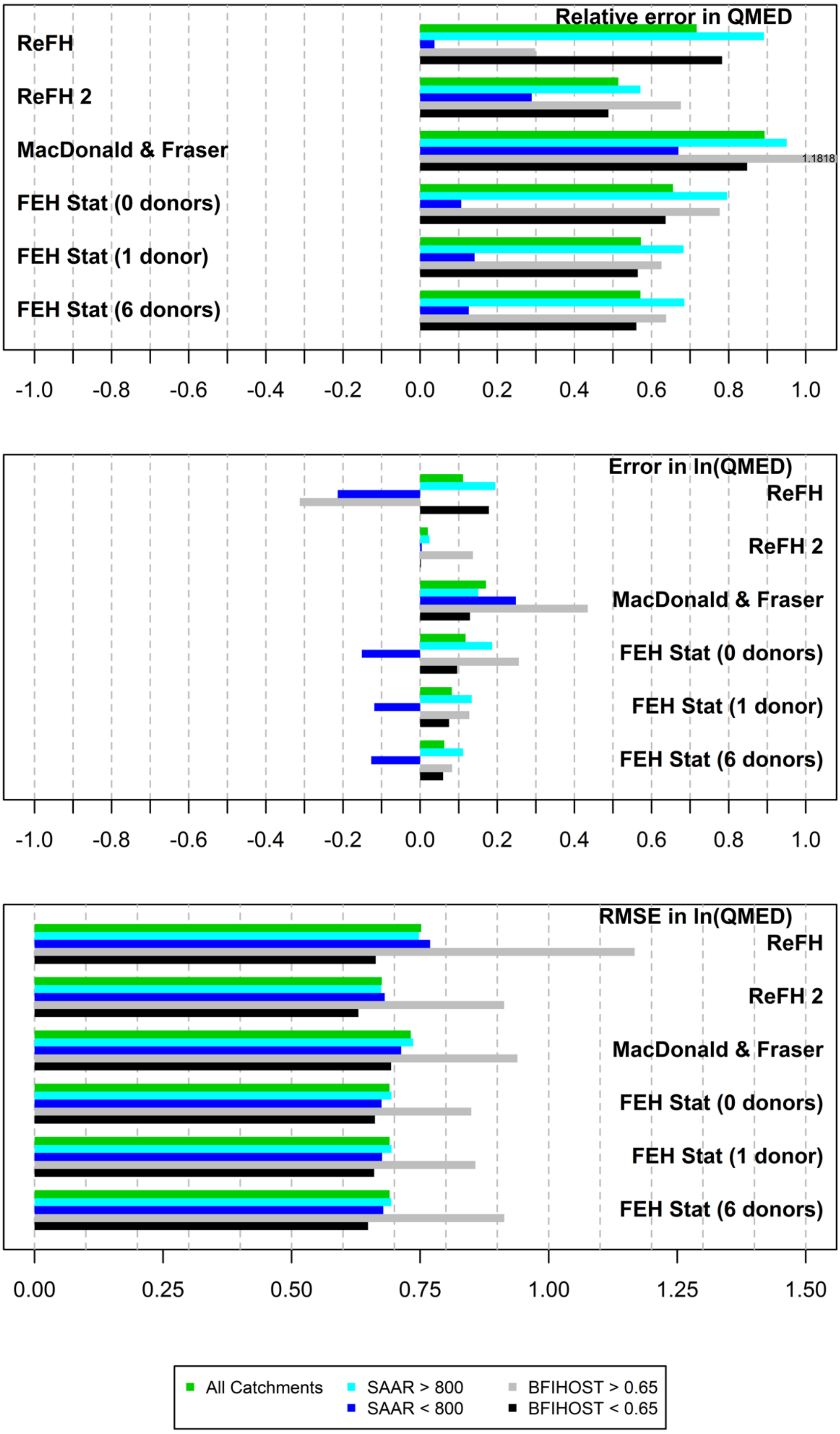

Figure 9 - Relative error, ln-error and RMSE in QMED estimated by six current and recent flood estimation methods (rural catchments)

The three graphs in Figure 9 quantify the relative error (top graph), ln-error (middle graph) and RMSE (bottom graph) of QMED, as estimated by the six methods in ‘high-quality’ rural catchments as a whole, and when subset by BFIHOST and SAAR:

- ReFH
- ReFH2
- MacDonald & Fraser
- FEH Stat (0 donors)
- FEH Stat (1 donor)
- FEH Stat (6 donors)

Results are shown for:

- All catchments (green bars)
- SAAR >800 (light blue bars)
- SAAR < 800 (dark blue bars)
- BFIHOST >0.65 (grey bars)
- BFIHOST <0.65 (black bars)

Across all methods, the smallest relative error is given for the subset of catchments with BFIHOST < 0.65. The largest relative errors are most commonly associated with the subsets of catchments with SAAR ≥ 800 or BFIHOST ≥ 0.65, while the largest RMSEs are most common for the subset with BFIHOST ≥ 0.65. RMSE follows a similar pattern for all methods, and all methods are relatively consistent, with each having consistent RMSE values across the full data set and three subsets (drier catchments, wetter catchments and less permeable catchments).

Across all models and catchment subsets, relative error in QMED is generally larger (more positive) than error in ln(QMED), though only five of 60 error bars in Figure 9 show negative values. This suggests that over-estimation is more severe when $QMED_{AMAX}$ is small – an error of, for example, 1 m^3/s is many times larger than the smallest $QMED_{AMAX}$ in this data set and positive relative errors are always numerically larger than positive ln-errors. Figure 10 plots error in ln(QMED) against QMED from AMAX and supports this assertion, with the addition that ReFH1 can also severely underestimate QMED in catchments where it is low. The grey regression line on Figure 10 considers all 103 rural catchments, while the green line excludes the three circled catchments that are identified as outliers on Figure 8 and are rural. Excluding these outliers reduces the strength of correlations, but does not eliminate them, or even reduce them below 0.05 in most cases.

Despite the reasonable assumption that QMED and therefore error in ln(QMED), may be related to catchment area, there is in fact little evidence of trend between error in ln(QMED) and AREA (Figure 11). Some of the largest positive errors in estimated QMED are for catchments around 20 km^2.

The strength of the correlation when MacDonald and Fraser’s method is used is particularly noteworthy: for gauged values of QMED above approximately 20 m^3/s, MacDonald and Fraser’s method **always** underestimates. This is not true of ReFH or the

FEH statistical method variants, though all are far more likely to underestimate for catchments with higher gauged QMED values. It should be noted that increasing the number of donors in the FEH statistical method, from zero to one, reduces the correlation between ln-error and gauged QMED. Correlation and mean ln-error are reduced further by increasing the number of donors from one to 6. However, the differences between the performance of the FEH statistical method with one and six donors are marginal, for the data set as a whole and for each subset of catchments.

For each method individually, RMSE is always highest for the subset of catchments with BFIHOST ≥ 0.65. Figure 10 highlights catchments with BFIHOST ≥ 0.65. It is likely that the high RMSE for this subset may be a function of sampling error – the subset contains only 14 catchments.

It is important to note that the overall factorial standard error of the FEH statistical method, with any number of donors, is significantly higher than the value reported for the FEH QMED equation (1.431, Kjeldsen and others, 2008), meaning that the spread of errors in QMED is larger for small catchments, even if the average error is still near-zero. This strongly suggests that there is scope to explore either recalibrating or changing the form of the FEH QMED equation for small catchment-specific applications.

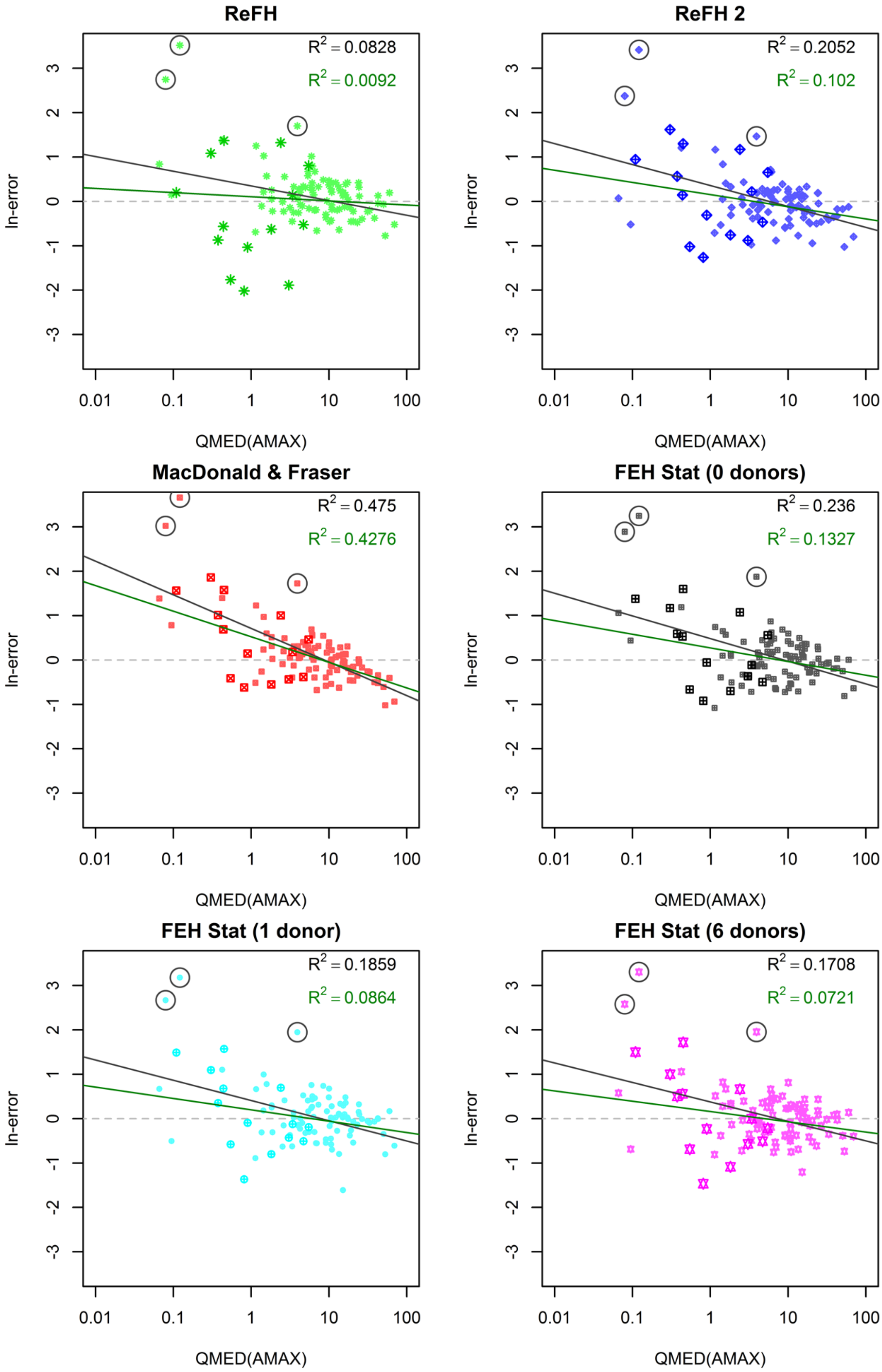

Figure 10 - ln-error in QMED estimate against QMED value from annual maxima.

The six scatter graphs in Figure 10 plot ln error in QMED (y-axis from -3 to 3) against QMED from AMAX (x-axis from 0.01 to 100) for the six methods:

- ReFH: top left graph
- ReFH2: top right graph
- MacDonald & Fraser: middle left graph
- FEH Stat (0 donors): middle right graph
- FEH Stat (1 donor): bottom left graph
- FEH Stat (6 donors): bottom right graph

Solid lines show least-squares fit. Smaller, paler coloured points show sites with BFIHOST < 0.65. Solid green lines show least-squares fit against 100 rural catchments (excluding the three circled).

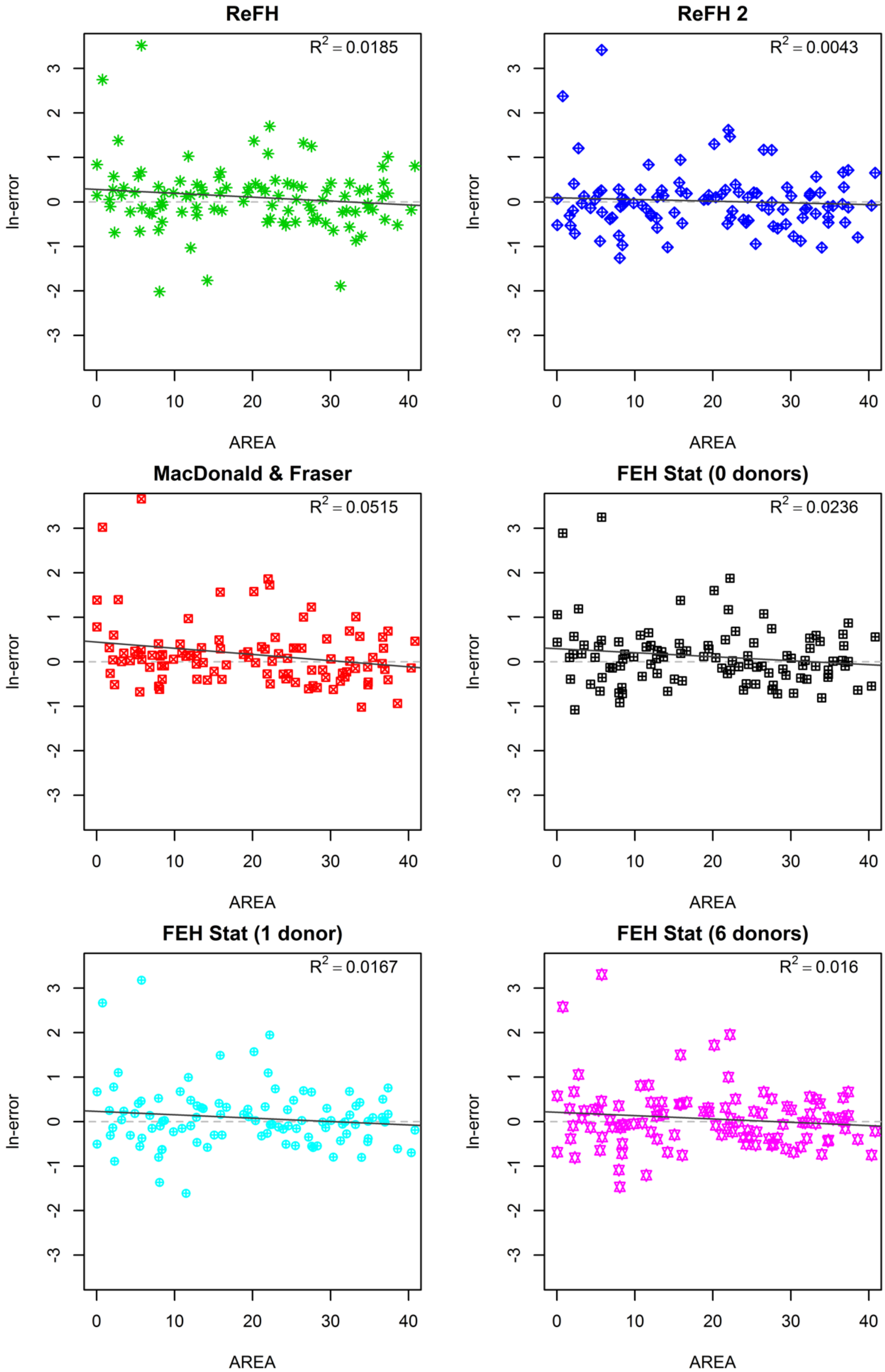

Figure 11 - Ln-error in QMED estimate against AREA

The six scatter graphs in Figure 11 plot ln-error in QMED (y-axis from -3 to 3) against AREA (x-axis from 0 to 40) for the six methods:

- ReFH: top left graph
- ReFH2: top right graph
- MacDonald & Fraser: middle left graph
- FEH Stat (0 donors): middle right graph
- FEH Stat (1 donor): bottom left graph
- FEH Stat (6 donors): bottom right graph

Solid lines show least-squares fit, while R^2 values on each plot show the proportion of variance explained by the least-squares fit.

In general, ReFH2 is shown to improve on ReFH1, with smaller errors and RMSE for each subset of catchments individually. ReFH2 is notable in that its performance on all 103 rural catchments, and on each subset within the 103, is of a more consistent quality than ReFH1. This provides evidence that ReFH2 is more suitable to use in permeable catchments, although (in common with all other tested methods) the average error and RMSE associated with using ReFH2 in permeable catchments is significantly larger than that associated with any other catchment subset or the 'high-quality' data set as a whole.

The FEH statistical method and ReFH2 are similar in terms of performance, though ReFH2 is slightly better in terms of RMSE and ln-error. In all circumstances, the FEH statistical method has a larger ln-error with zero donors than with one or six donors, and in most circumstances, a larger relative error also. This result implies that donor transfer may still serve a useful purpose for small catchments. It should be noted that using six donors does not always result in a smaller error than using one donor.

ReFH2 performs more strongly than the FEH statistical method when considering the propensity to over or underestimate QMED within either the low-SAAR or high-SAAR catchment subsets. This could indicate an unexplained dependency in the statistical method's QMED equation or donor transfer procedure that is somewhat cancelled out when all catchment types are grouped together. In particular, all variants of the FEH statistical method show negative ln-errors for catchments with SAAR < 800 and positive ln-errors for catchments with SAAR ≥ 800. However, when catchments are subset by BFIHOST, SAAR is more mixed and the differences in ln-errors between subsets are reduced, especially when one or six donors are used.

MacDonald and Fraser's equation is generally the most likely to overestimate either QMED or ln(QMED) within the 103 'high-quality' rural catchments considered here. It is also most likely to overestimate QMED within any of the four catchment subsets. Its ln-error is generally larger than that of the other methods, with some exceptions for high-SAAR catchments. Though it shows less of a dependency on SAAR than the FEH statistical method in small catchments, a consistent bias away from the zero ln-error line means that this does not translate into a low RMSE relative to other methods.

Using donors in the FEH statistical method is not shown to noticeably affect uncertainty (RMSE) in estimating QMED, except to increase it in permeable catchments. This is potentially an artefact of the small number of 'high-quality' permeable catchments available for this study. Using donors is, however, shown to reduce the correlation between error in ln(QMED) and the AMAX estimate of QMED for rural catchments overall (Figure 10). As the study data set follows a similar distribution to the NRFA peak flow data set in terms of BFIHOST and SAAR, any patterns in RMSE relating to these two catchment descriptors are unexpected. It is also surprising that, if residuals in SAAR and BFIHOST are geographically clustered, donor transfer from an increased number of nearby sites (with presumably similar SAAR and BFIHOST values) doesn't significantly help to counteract the regression residual at the site of interest. This unusual effect could be related to the limited number of low-SAAR (21) and high-BFIHOST (14) catchments, six of which overlap both groups. Although catchment area is the most significant term in the FEH QMED equation, and all the catchments in this study are outliers in terms of area, error in estimated ln(QMED) has already been shown to be unrelated to area (Figure 11). Therefore, the suggestion that regression residuals are less likely to follow a spatial pattern for small catchments than for large ones is unlikely to be true. Here, it is however worth reiterating that the differences in performance between the FEH statistical method with zero, one and six donors are marginal, especially between one and six donors. This implies that multiple donor transfer is neither particularly useful nor harmful for estimating QMED in small ungauged catchments.

Figure 12 plots the ln-error in estimated QMED against the ratio of areas between each of the 103 'high-quality' rural study catchments and its closest donor in the NRFA peak flow data set v3.3.4. When six donors are used, the geometric mean of donor areas is considered. For one donor, a limited relationship between donor/study catchment area ratio and ln-error can be seen, which appears to be heavily influenced by two or three catchments with a donor/target area between 10 and 100, and an ln-error around 2 to 3. This relationship is almost eliminated when six donors are used. Therefore, any potential gain that could be made by only permitting donors with similar areas to the study catchment could be offset by the increased distance between the study catchment and donor(s). It should be considered that rejecting a donor based on size difference implies some dependency between regression residuals and catchment area; this is not likely to be significant, given that AREA is included in the FEH QMED regression equation.

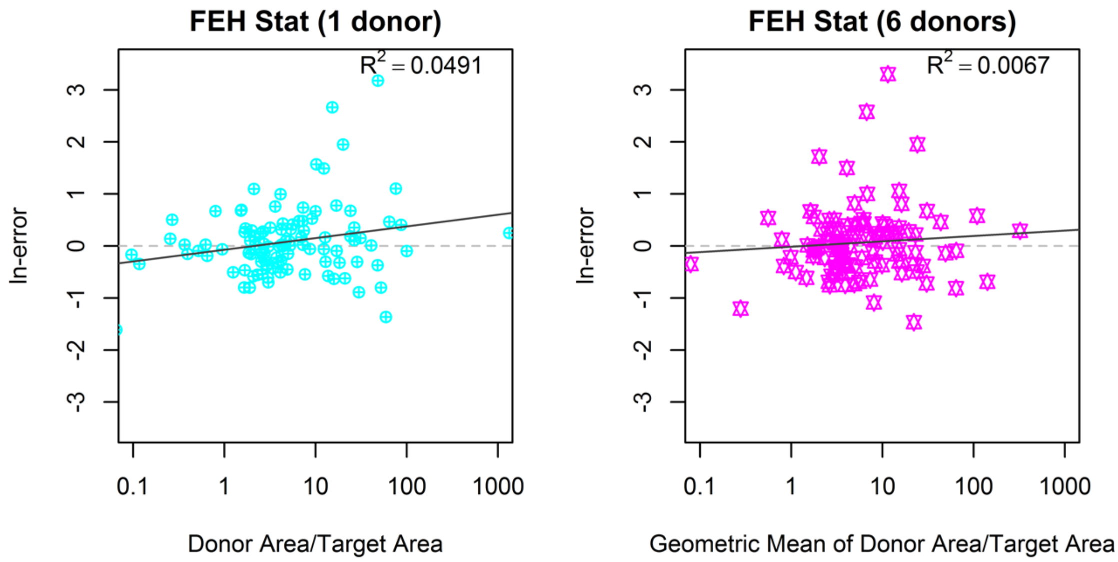

Figure 12 - Ln-error in QMED estimate against ratio of areas between donor catchment(s) and study catchment

Figure 12 shows two scatter graphs:

- left-hand graph: This plots ln-error on the x-axis (from -3 to 3) by donor area/target area on the y-axis (from 0.1 to 1000) for FEH stat (1 donor)
- right-hand graph: This plots ln-error on the x-axis (from -3 to 3) by geometric mean of donor area/target area on the x-axis (from 0.1 to 1000) for FEH stat (6 donors)

Solid lines show least-squares fit, while R^2 values on each plot show the proportion of variance explained by the least-squares fit.

Figure 13 plots the change in ln-error resulting from donor transfer against the centroid-to-centroid distance from each of the 103 'high-quality' rural study catchments to its donor; the geometric mean of distances is considered when six donors are used. A negative change in ln-error indicates a reduction in the difference between gauged and estimated QMED. Figure 13 shows practically no relation between centroid-to-centroid distance and improvements in QMED estimation while, if regression residuals were strongly spatially clustered, larger improvements would be expected for catchments with closer donors. However, the change in ln-error due to donor transfer is potentially not the best metric to use to assess this, as the donor transfer procedure includes weighting to reduce the influence of donors with increasing distance, resulting in the change in ln-error tending towards zero.

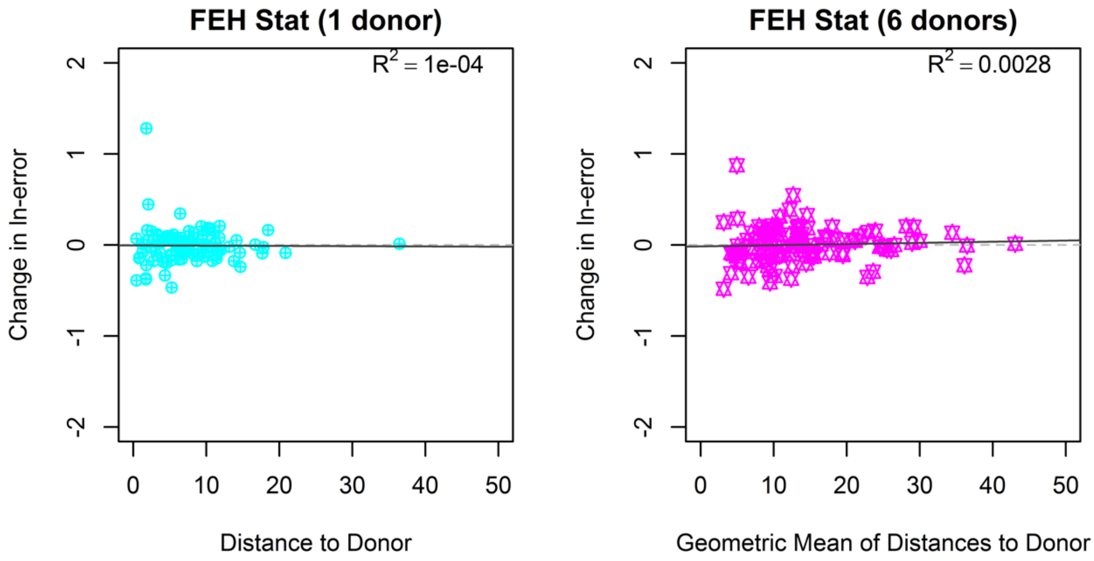

Figure 13 - Ln-error in QMED estimate against distance from donor catchment(s) to study catchment

Figure 13 shows to scatter graphs:

- left-hand graph: This plots change in ln-error on the x-axis (from -2 to 2) by distance to donor area on the y-axis (from 0 to 50) for FEH stat (1 donor)
- right-hand graph: This plots change in ln-error on the x-axis (from -2 to 2) by geometric mean of distances to donor on the x-axis (from 0 to 50) for FEH stat (6 donors)

Solid lines show least-squares fit, while R^2 values on each plot show the proportion of variance explained by the least-squares fit.

Urban catchments

Figure 14 and Tables 11, 12 and 13 quantify the relative error, ln-error and RMSE of QMED, as estimated by the six methods being considered, in urbanised catchments, as a whole and when subset by BFIHOST and SAAR. Due to the relatively small number of urbanised catchments overall (43) and the extremely small number of urbanised catchments with high values of SAAR (12) and BFIHOST (10), the following observations may be subject to considerable sampling uncertainty.

Table 11 – Relative error in QMED estimated by six current and recent flood estimation methods in urban catchments

Subset	All (43 sites)	SAAR ≥ 800 (12 sites)	SAAR < 800 (31 sites)	BFIHOST ≥ 0.65 (10 sites)	BFIHOST < 0.65 (33 sites)
ReFH1	0.1385	0.2246	0.1052	-0.3043	0.2727
ReFH2	0.0974	0.0485	0.1163	-0.0318	0.1366
MacD & Fraser	0.1247	0.1741	0.1056	-0.1622	0.2117
Stat. (0 donors)	0.1393	0.2304	0.1041	0.1265	0.1432
Stat. (1 donor)	0.1210	0.1842	0.0965	0.1675	0.1069
Stat. (6 donors)	0.1522	0.1923	0.1367	0.1999	0.1378

The best performing method in each subset was identified as:

- ReFH2 for all sites
- ReFH2 for SAAR ≥ 800
- Stat. (1 donor) for SAAR < 800
- ReFH2 for BFIHOST ≥ 0.65
- Stat. (1 donor) for BFIHOST < 0.65

Table 12 – Ln-error in QMED estimated by six current and recent flood estimation methods in urban catchments

Subset	All (43 sites)	SAAR ≥ 800 (12 sites)	SAAR < 800 (31 sites)	BFIHOST ≥ 0.65 (10 sites)	BFIHOST < 0.65 (33 sites)
ReFH1	-0.1259	0.1135	-0.2186	-0.8218	0.0850
ReFH2	-0.0858	-0.0189	-0.1117	-0.2892	-0.0242
MacD & Fraser	-0.1166	0.0602	-0.1851	-0.5503	0.0148
Stat. (0 donors)	-0.0418	0.1284	-0.1077	-0.0845	-0.0288
Stat. (1 donor)	-0.0593	0.0671	-0.1082	-0.0705	-0.0559
Stat. (6 donors)	-0.0328	0.1014	-0.0847	-0.0560	-0.0257

The best performing method in each subset was identified as:

- Stat. (6 donors) for all sites

- ReFH2 for SAAR ≥ 800
- Stat. (6 donors) for SAAR < 800
- Stat. (6 donors) for BFIHOST ≥ 0.65
- MacD & Fraser for BFIHOST < 0.65

Table 13 – RMSE in QMED estimated by six current and recent flood estimation methods in urban catchments

Subset	All (43 sites)	SAAR ≥ 800 (12 sites)	SAAR < 800 (31 sites)	BFIHOST ≥ 0.65 (10 sites)	BFIHOST < 0.65 (33 sites)
ReFH1	0.7840	0.4402	0.8818	1.2533	0.5699
ReFH2	0.6125	0.3628	0.6852	0.7638	0.5587
MacD & Fraser	0.7557	0.4403	0.8469	1.1059	0.6112
Stat. (0 donors)	0.5893	0.4099	0.6454	0.6544	0.5681
Stat. (1 donor)	0.5905	0.4351	0.6406	0.6903	0.5567
Stat. (6 donors)	0.5902	0.3802	0.6536	0.7114	0.5481

The best performing method in each subset was identified as:

- Stat. (0 donors) for all sites.
- ReFH2 for SAAR ≥ 800
- Stat. (1 donor) for SAAR < 800
- Stat. (0 donors) for BFIHOST ≥ 0.65
- Stat. (6 donors) for BFIHOST < 0.65

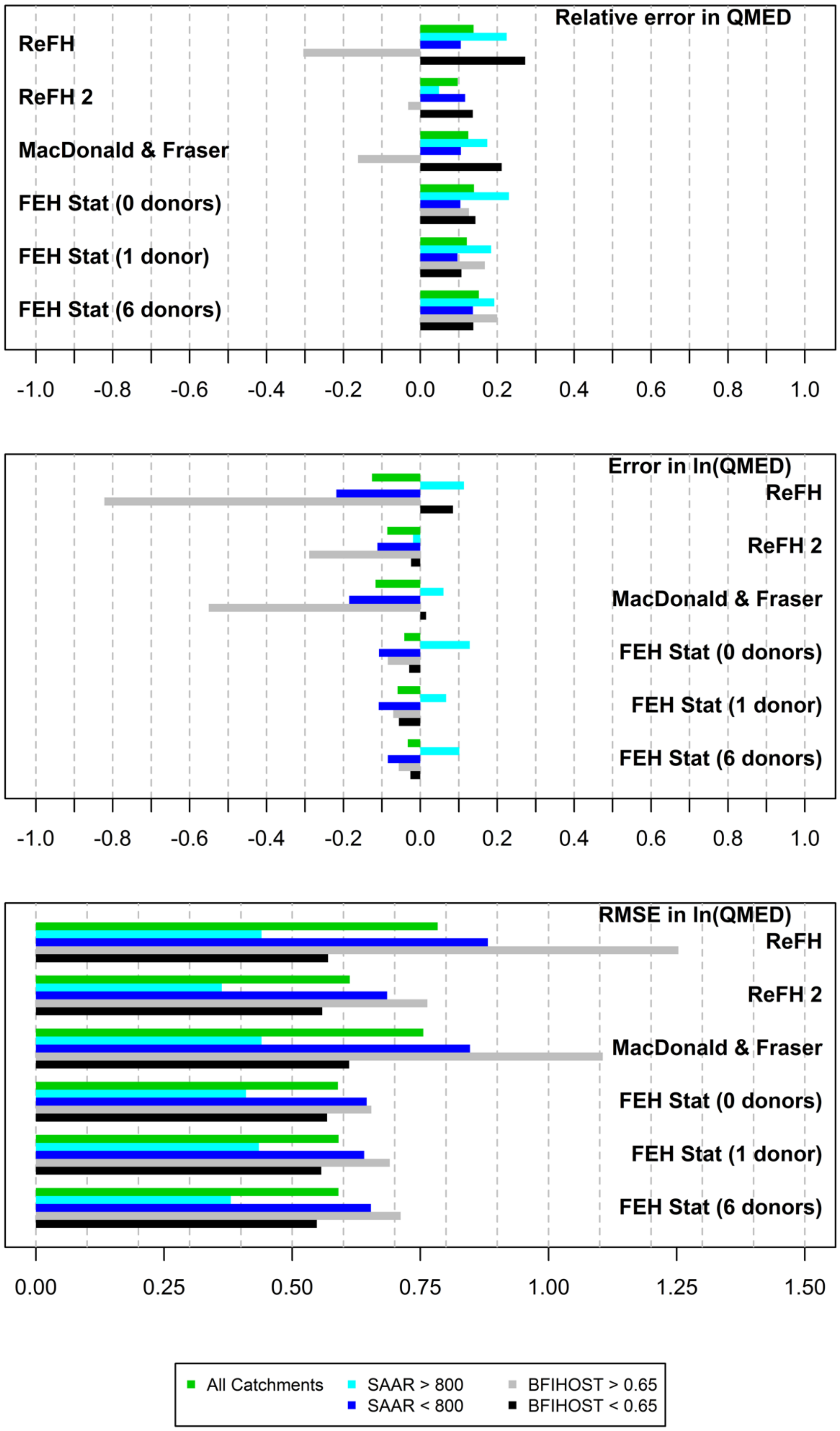

Figure 14 - Relative error, ln-error and RMSE in QMED estimated by six current and recent flood estimation methods (urban catchments)

The three graphs in Figure 14 quantify the relative error (top graph), ln-error (middle graph) and RMSE (bottom graph) of QMED, as estimated by the six methods in urbanised catchments, as a whole and when subset by BFIHOST and SAAR:

- ReFH
- ReFH2
- MacDonald & Fraser
- FEH Stat (0 donors)
- FEH Stat (1 donor)
- FEH Stat (6 donors)

Results are shown for:

- All catchments (green bars)
- SAAR >800 (light blue bars)
- SAAR < 800 (dark blue bars)
- BFIHOST >0.65 (grey bars)
- BFIHOST <0.65 (black bars)

Across the full group of 43 urbanised comparison catchments, the best method in terms of relative error is ReFH2, the best method in terms of ln-error is the FEH statistical method with six donors and the best methods in terms of RMSE is the FEH statistical method with zero donors, although all variants of the FEH statistical method give similar overall results for all 3-performance metrics. The worst performing method in terms of ln-error and RMSE is ReFH1, while the worst performing method in terms of relative error is the FEH statistical method with six donors, although all six methods perform very similarly in terms of relative error over the 43-catchment data set. Performance in ln-error is considered a more robust measure than performance in relative error as it is evenly balanced between over and underestimation: for example, a modelled QMED that is half of the gauged value will have the same magnitude of ln-error as a gauged QMED that is half of the modelled value.

In contrast to rural catchments, relative errors are consistently around the +0.1-0.2 range across all methods and catchment subsets, while ln-errors are negative across all methods and catchment subsets (excluding high-SAAR catchments and, using ReFH1 or MacDonald and Fraser’s equation, less permeable catchments).

In terms of RMSE and ln-error, MacDonald and Fraser’s method performs far worse than the FEH statistical method in high-BFIHOST and low-SAAR catchments, but only slightly worse in terms of RMSE (and sometimes better in terms of ln-error) in high-SAAR and low-BFIHOST catchments. This is unexpected, as a greater proportion of MacDonald and Fraser’s calibration catchments came from the lower-SAAR, generally higher-BFIHOST parts of the UK, compared with both the FEH statistical calibration group and the current study data set. This may reflect the small sample size of both the low-SAAR and high-BFIHOST groups but may also reflect the fact that more of south east England is impermeable than is often assumed.

Significant overestimations in both QMED and ln(QMED) are less likely in the group of 43 ‘high-quality’ urban catchments than the group of 103 ‘high-quality’ essentially rural catchments. Overall, RMSE is reduced significantly for all methods apart from ReFH1 and MacDonald and Fraser’s equation, opening a performance gap between current FEH methods (ReFH2/FEH statistical) and others (ReFH1/MacDonald and Fraser’s equation). The fact that MacDonald and Fraser’s equation, in particular, is less likely to overestimate in urban catchments is contradictory to MacDonald and Fraser’s own observations that no relationship between QMED and $URBEXT_{2000}$ could be found for their data set. It is less surprising that the performance of ReFH1 changes in urbanised catchments, as its calibration data set featured only seven heavily urbanised and 17 moderately urbanised catchments. It is perhaps more surprising that the performances of ReFH2 and the FEH statistical method change, as both have explicit adjustment procedures for urbanisation. These methods have the lowest error in rural catchments and the error metrics in the smaller urban catchments are generally lower than their rural equivalents. This would suggest that the urban procedures within the methods might overcompensate for urbanisation, that is, that the overall smaller errors in urbanised catchments result from the errors in the as-rural estimates being added to opposite-sign errors resulting from the urbanisation adjustments.

Although ReFH2’s performance in terms of relative and ln-error is inconsistent between rural and urban catchments, it is an improvement on ReFH1 in both cases. It is worth noting that the urban routing part of ReFH2’s modelling framework is new and therefore less tested than the rural routing, baseflow and loss parts. There is therefore high potential for this to be studied and reconfigured for small catchments. It is reiterated that only winter storm profiles, and ‘as-rural’ initial wetness and baseflow values were used in ReFH2 in this analysis, including for urban catchments. Considering cases where it may be more appropriate to use summer and urbanised parameters more closely could provide opportunities for further improvement.

The subset of low-SAAR catchments contains 31 of the 43 urbanised study catchments, so the error metrics and overall observations are similar as for the full group of 43 urbanised catchments. All methods appear to perform particularly well in small, urbanised, high-SAAR catchments, with RMSE values approaching that reported during development of the FEH statistical method on catchments up to almost 10,000 km^2. This is, however, for a subset of just 12 catchments. It is unfortunate that there are so few heavily urbanised, high-SAAR small catchments available to study as it is difficult to attribute any observations to the value of SAAR and to explore the apparently high performance of the methods further. A further complication stems from the fact that all high-SAAR catchments are low-BFIHOST catchments, and all high-BFIHOST catchments are low-SAAR catchments. Therefore, not only is it practically impossible to definitively assign any observations to the effects of high SAAR or high BFIHOST, but it is practically impossible to separate the effects of the two catchment descriptors.

Figure 15 plots ln-error in modelled QMED against observed QMED for all six methods, showing that there may again be some trend between the errors in modelled QMED values and the estimated QMED values themselves.

MacDonald and Fraser's method shows the weakest trend, despite not accounting for urbanisation at all. The lack of trend, however, does not mean a lack of error: underestimation of QMED at a few sites is more common and greater than overestimation of QMED at a larger group of sites. For all methods, the downward trends are somewhat mitigated by excluding two catchments with $QMED_{AMAX}$ of approximately 12.5 m^3/s (circled), on which QMED is greatly underestimated. However, even after performing a regression without these two catchments (dark green lines), correlations remain above 0.05 for all variants of the FEH statistical method. Both excluded catchments have at least one unusual characteristic: one (28115) is very permeable, with a BFIHOST of 0.841, while the other (39082 – identified as an outlier in Figure 8) is more heavily urbanised than any catchment used to calibrate the methods, and so heavily urbanised that the methods for modelling it in ReFH2 are not yet fully developed ($URBEXT_{2000}$ = 0.811).

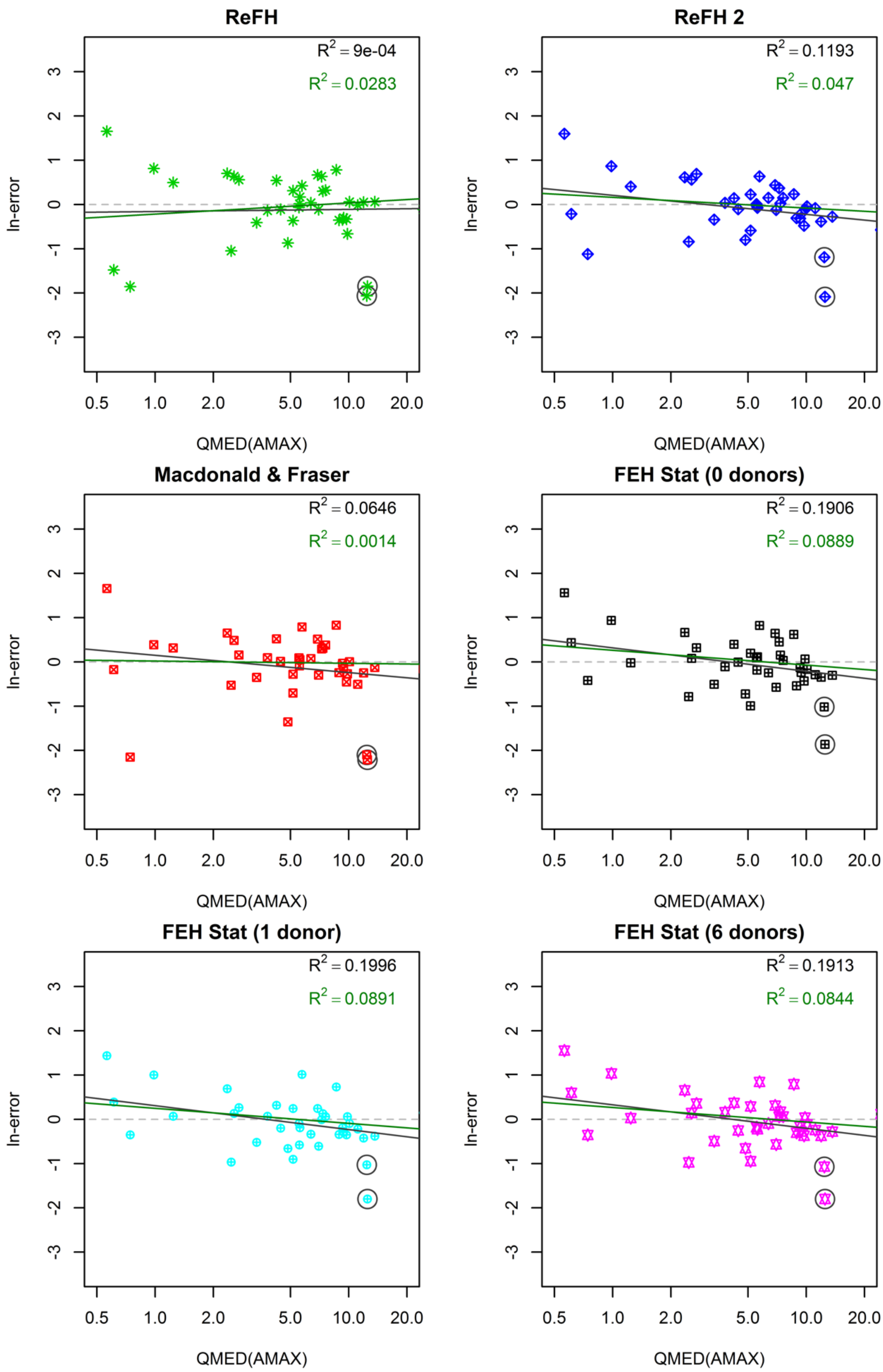

Figure 15 - Ln-error in QMED estimate against QMED value from annual maxima

The six scatter graphs in Figure 15 plot ln-error in modelled QMED (y-axis from -3 to 3) against QMED(AMAX) (x-axis from 0.5 to 20) for the six methods:

- ReFH: top left graph
- ReFH2: top right graph
- MacDonald & Fraser: middle left graph
- FEH Stat (0 donors): middle right graph
- FEH Stat (1 donor): bottom left graph
- FEH Stat (6 donors): bottom right graph

Solid black lines show least-squares against 43 urbanised catchments. Solid green lines show least-squares against 41 urbanised catchments, excluding two circled. R^2 values on each plot show the proportion of variance explained by the least-squares fit.

A regression of ln-error in QMED against $URBEXT_{2000}$ in urbanised catchments (Figure 16) shows notable negative correlation (that is, a tendency to underestimate increasingly for greater $URBEXT_{2000}$ values) for all six methods. For ReFH2 and ReFH1, the negative correlations can be eliminated by excluding catchments with $URBEXT_{2000}$ values above 0.6 (2 circled) and catchments with $URBEXT_{1990}$ values above 0.5 respectively. Further study of individual points on Figure 16 suggests that excluding catchments with $URBEXT_{2000}$ values above 0.6 would eliminate or reduce correlation between ln-error in QMED and $URBEXT_{2000}$ when using the FEH statistical method. In essentially rural catchments, ln-error in QMED is less dependent on $URBEXT_{2000}$ as expected (Figure 17). There is evidence to show that the methods may perform differently in moderately urbanised catchments when compared to heavily urbanised catchments. However, there are too few catchments in each of the 'moderately' and 'heavily' urbanised groups to allow a robust comparison (see Table 6).

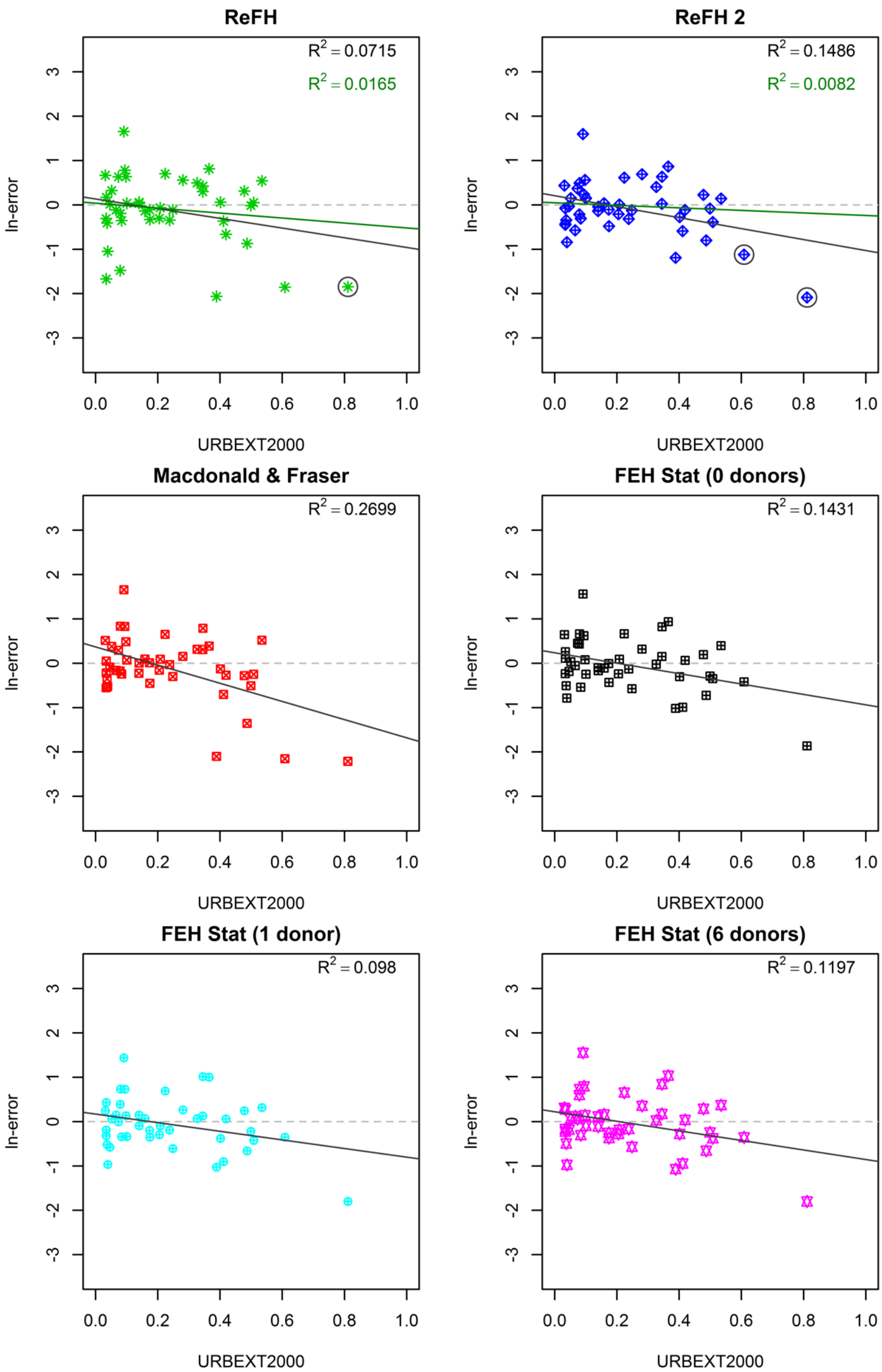

Figure 16 - Ln-error in QMED estimate against $URBEXT_{2000}$ (≥ 0.03)

The six scatter graphs in Figure 16 plot ln-error in modelled QMED (y-axis from -3 to 3) against $URBEXT_{2000}$ (≥ 0.03) on the x-axis (from 0.0 to 1.0) for the six methods:

- ReFH: top left graph
- ReFH2: top right graph
- MacDonald & Fraser: middle left graph
- FEH Stat (0 donors): middle right graph
- FEH Stat (1 donor): bottom left graph
- FEH Stat (6 donors): bottom right graph

Solid black lines show least-squares fit against all 43 urbanised catchments. Solid green lines show least-squares fit against 41 (or 42) catchments where ReFH2 (or ReFH) is applicable. R^2 values on each plot show the proportion of variance explained by the least-squares fit.

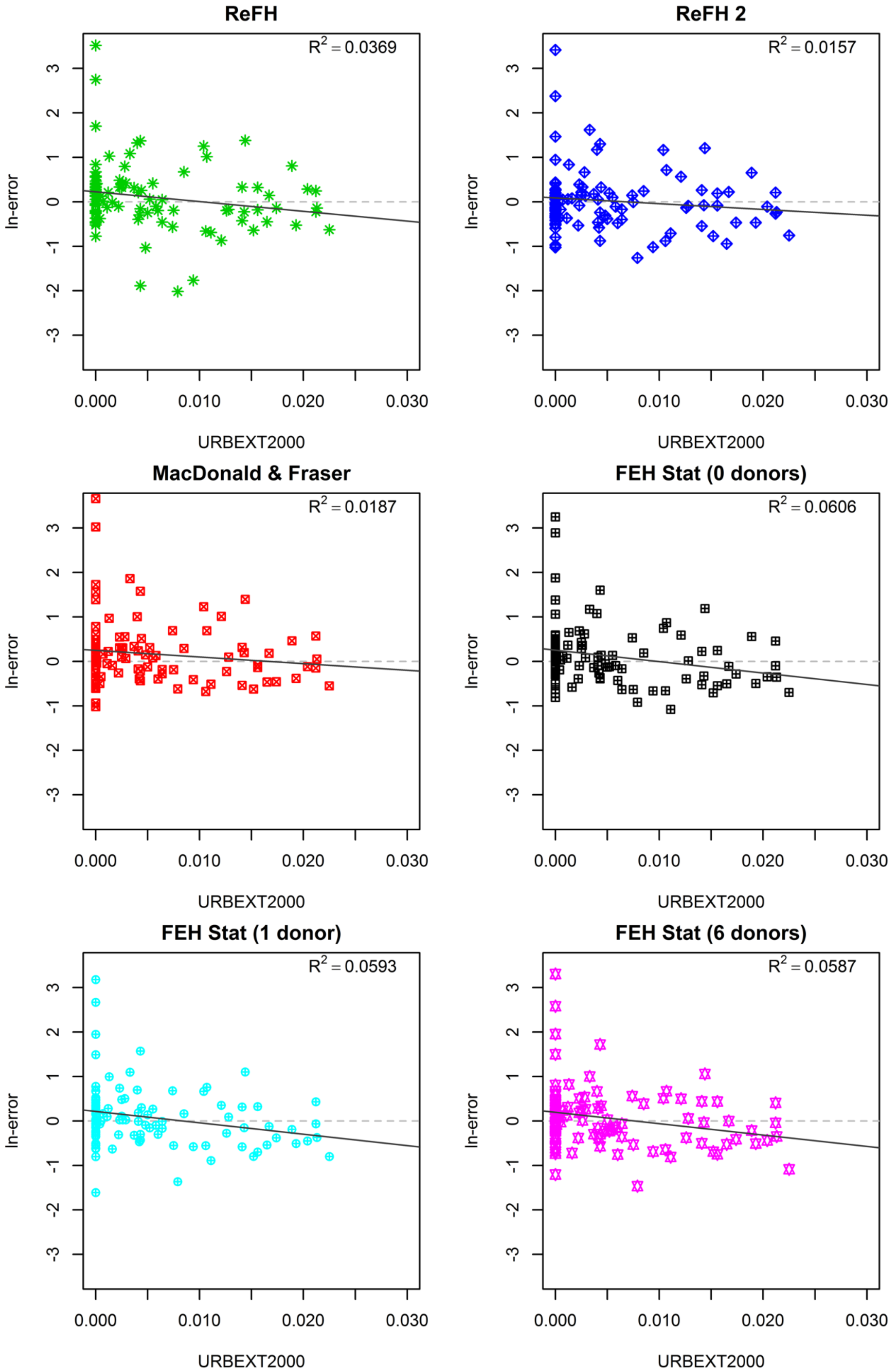

Figure 17 - Ln-error in QMED estimate against $URBEXT_{2000}$ (< 0.03)

The six scatter graphs in Figure 17 plot ln-error in modelled QMED (y-axis from -3 to 3) against $URBEXT_{2000}$ (< 0.03) on the x-axis (from 0.000 to 1.030) for the six methods:

- ReFH: top left graph
- ReFH2: top right graph
- MacDonald & Fraser: middle left graph
- FEH Stat (0 donors): middle right graph
- FEH Stat (1 donor): bottom left graph
- FEH Stat (6 donors): bottom right graph

Solid lines show least-squares fit, while R^2 values on each plot show the proportion of variance explained by the least-squares fit.

4.7 Analysis of ‘extended’ data set

Rural catchments

Figure 18 and Tables 14, 15 and 16 quantify the relative error, ln-error and RMSE of QMED, as estimated by the six methods being considered, in rural catchments as a whole, and when subset by BFIHOST and SAAR. Overall, and for each catchment subset, the relative and ln-error bars in Figure 18 are shifted leftwards in comparison to Figure 9, that is, overestimation of QMED is reduced. Relative error is reduced for every catchment subset and modelling method. Ln-error is also reduced near-universally, with only ReFH2 decreasing (slightly) in performance and only the low-SAAR subset being less-well modelled. RMSE is reduced in 27 of the 30 cells in Tables 14, 15 and 16 relative to Table 7. It is surprising that there should be such a systematic effect of adding new catchments. It may suggest that less well-maintained or sited gauges, or gauges with highly extrapolated rating curves, have a tendency to over-report high flow rates. If this is true, then the apparent reduction in error could be spurious. Including more of these lower quality gauges does seem to improve the performance metrics of all methods almost universally across the full group and subsets of rural catchments.

Table 14 – Relative error in QMED estimated by six current and recent flood estimation methods in rural catchments

Subset	All (150 sites)	SAAR ≥ 800 (122 sites)	SAAR < 800 (28 sites)	BFIHOST ≥ 0.65 (21 sites)	BFIHOST < 0.65 (129 sites)
ReFH1	0.5117	0.6327	-0.0157	0.0811	0.5818
ReFH2	0.3193	0.3499	0.1861	0.3970	0.3067
MacD & Fraser	0.6231	0.6376	0.5597	0.8843	0.5806
Stat. (0 donors)	0.4450	0.5416	0.0239	0.5084	0.4346
Stat. (1 donor)	0.3946	0.4807	0.0192	0.4224	0.3901
Stat. (6 donors)	0.3973	0.4915	-0.0134	0.3967	0.3974

The best performing method in each subset was identified as:

- ReFH2 for all sites
- ReFH2 for SAAR ≥ 800
- Stat. (6 donors) for SAAR < 800
- Stat. (6 donors) for BFIHOST ≥ 0.65
- ReFH2 for BFIHOST < 0.65

Table 15 – Ln-error in QMED estimated by six current and recent flood estimation methods in rural catchments

Subset	All (150 sites)	SAAR ≥ 800 (122 sites)	SAAR < 800 (28 sites)	BFIHOST ≥ 0.65 (21 sites)	BFIHOST < 0.65 (129 sites)
ReFH1	0.0583	0.1314	-0.2602	-0.4292	0.1377
ReFH2	-0.0419	-0.0363	-0.0662	-0.0066	-0.0477
MacD & Fraser	0.1023	0.0802	0.1985	0.3217	0.0666
Stat. (0 donors)	0.0499	0.1093	-0.2089	0.1019	0.0415
Stat. (1 donor)	0.0285	0.0877	-0.2298	0.0090	0.0316
Stat. (6 donors)	0.0156	0.0797	-0.2638	-0.0522	0.0266

The best performing method in each subset was identified as:

- Stat. (6 donors) for all sites

- ReFH2 for SAAR ≥ 800
- ReFH2 for SAAR < 800
- ReFH2 for BFIHOST ≥ 0.65
- Stat. (6 donors) for BFIHOST < 0.65

Table 16 – RMSE in QMED estimated by six current and recent flood estimation methods in rural catchments

Subset	All (150 sites)	SAAR ≥ 800 (122 sites)	SAAR < 800 (28 sites)	BFIHOST ≥ 0.65 (21 sites)	BFIHOST < 0.65 (129 sites)
ReFH1	0.6862	0.6633	0.7783	1.0972	0.5930
ReFH2	0.6043	0.5898	0.6637	0.8171	0.5621
MacD & Fraser	0.6496	0.6403	0.6887	0.8483	0.6111
Stat. (0 donors)	0.6232	0.6111	0.6734	0.8012	0.5892
Stat. (1 donor)	0.6287	0.6097	0.7052	0.8411	0.5868
Stat. (6 donors)	0.6333	0.6117	0.7202	0.8813	0.5831

- The best performing method in each subset was identified as:
- ReFH2 for all sites
- ReFH2 for SAAR ≥ 800
- ReFH2 for SAAR < 800
- Stat. (0 donors) for BFIHOST ≥ 0.65
- ReFH2 for BFIHOST < 0.65

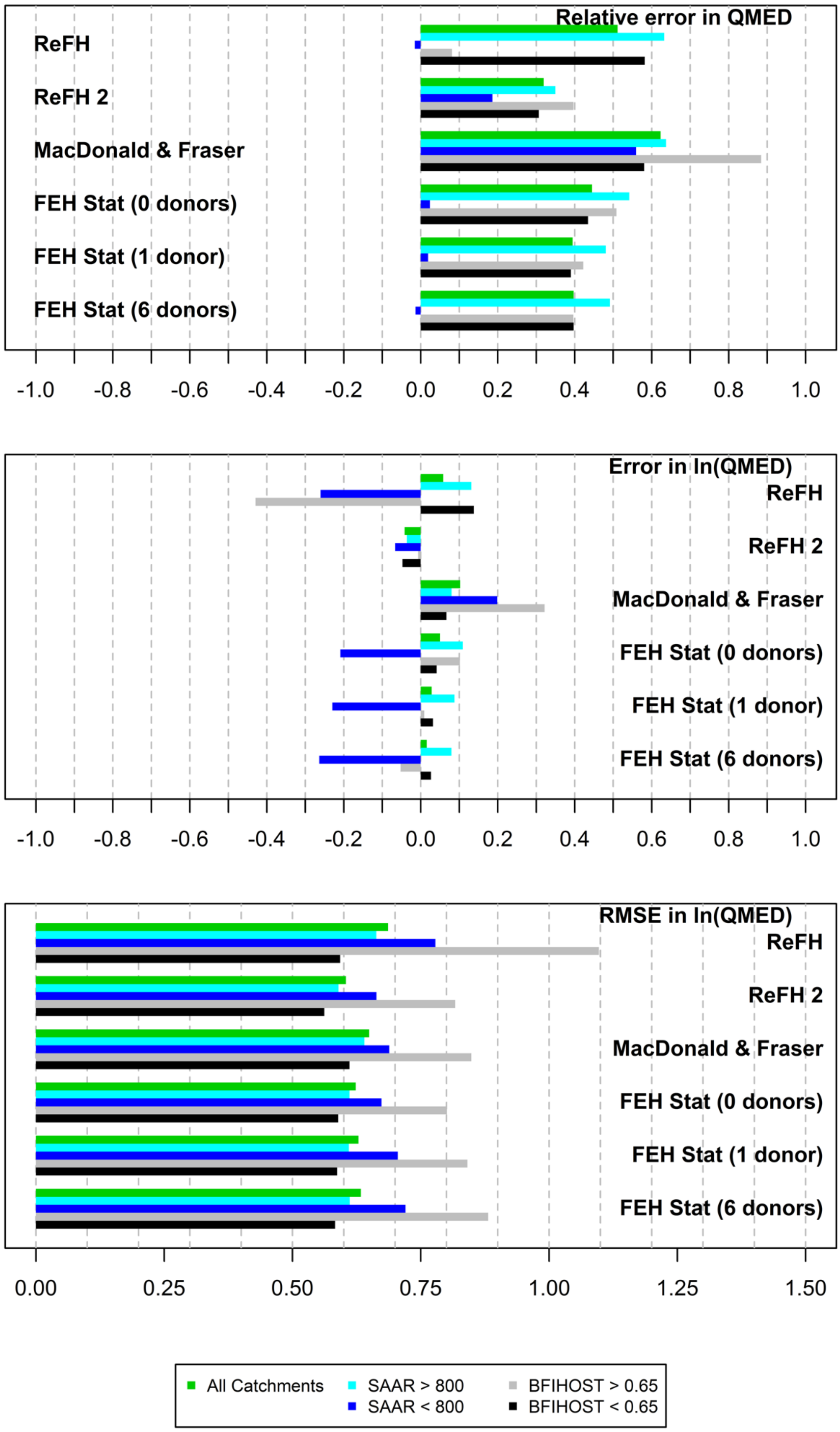

Figure 18 - Relative error, ln-error and RMSE in QMED estimated by six current and recent flood estimation methods (rural catchments)

The three graphs in Figure 18 quantify the relative error (top graph), ln-error (middle graph) and RMSE (bottom graph) of QMED, as estimated by the six methods in rural catchments, as a whole and when subset by BFIHOST and SAAR:

- ReFH
- ReFH2
- MacDonald & Fraser
- FEH Stat (0 donors)
- FEH Stat (1 donor)
- FEH Stat (6 donors)

Results are shown for:

- All catchments (green bars)
- SAAR >800 (light blue bars)
- SAAR < 800 (dark blue bars)
- BFIHOST >0.65 (grey bars)
- BFIHOST <0.65 (black bars)

About half of the catchments used to calibrate MacDonald and Fraser's equation are known to be outside the NRFA peak flow data set v3.3.4 and therefore to have uncertain data quality.

Urban catchments

Figure 19 and Tables 17, 18 and 19 quantify the relative error, ln-error and RMSE of QMED, as estimated by the six methods being considered, in urbanised catchments as a whole, and when subset by BFIHOST and SAAR. In terms of relative error, the best performing method is ReFH2, while the best performing in terms of ln-error and RMSE is the FEH statistical method, with six and one donor(s) respectively. The most consistent method in terms of relative error across all catchment subsets is ReFH2. QMED does not seem to be systematically overestimated at fewer 'high-quality' gauges. Not unexpectedly, ReFH1 severely underestimates QMED in high-BFIHOST catchments. However, so does MacDonald and Fraser's equation, which incorporates BFIHOST, though not any measure of urban extent. As it possesses such a large spread of relative and ln-errors, ReFH1 cannot be recommended for use in urbanised catchments.

In contrast to rural catchments, evaluating a larger group of urban catchments does not affect relative error, ln-error or RMSE systematically. Additionally, changes in these metrics are relatively small. In general, relative error is improved slightly for ReFH1, ReFH2 and MacDonald and Fraser's equation overall and for low-SAAR and low-BFIHOST subsets. Ln-error is improved for the FEH statistical method overall and in the same two subsets of catchments. RMSE is generally reduced for the FEH statistical method, except in the high-BFIHOST catchments, of which there are now 13. RMSE is also reduced across all methods for the low-BFIHOST subset, and across all methods except ReFH1 for the low-SAAR subset.

Table 17 – Relative error in QMED estimated by six current and recent flood estimation methods in urban catchments

Subset	All (57 sites)	SAAR ≥ 800 (16 sites)	SAAR < 800 (41 sites)	BFIHOST ≥ 0.65 (13 sites)	BFIHOST < 0.65 (44 sites)
ReFH1	0.1061	0.2298	0.0578	-0.3790	0.2494
ReFH2	0.0804	0.0516	0.0916	-0.0575	0.1211
MacD & Fraser	0.1107	0.2317	0.0635	-0.2028	0.2034
Stat. (0 donors)	0.1673	0.2949	0.1175	0.1877	0.1613
Stat. (1 donor)	0.1306	0.1818	0.1106	0.2361	0.0994
Stat. (6 donors)	0.1641	0.2148	0.1443	0.2544	0.1374

The best performing method in each subset was identified as:

- ReFH2 for all sites
- ReFH2 for SAAR ≥ 800
- MacD & Fraser for SAAR < 800
- ReFH2 for BFIHOST ≥ 0.65
- Stat. (1 donor) for BFIHOST < 0.65

Table 18 – Ln-error in QMED estimated by six current and recent flood estimation methods in urban catchments

Subset	All (57 sites)	SAAR ≥ 800 (16 sites)	SAAR < 800 (41 sites)	BFIHOST ≥ 0.65 (13 sites)	BFIHOST < 0.65 (44 sites)
ReFH1	-0.1706	0.0666	-0.2632	-1.0026	0.0752
ReFH2	-0.0975	-0.0595	-0.1123	-0.3359	-0.0270
MacD & Fraser	-0.1110	0.0921	-0.1903	-0.5550	0.0202
Stat. (0 donors)	-0.0179	0.1683	-0.0905	-0.1020	0.0070
Stat. (1 donor)	-0.0434	0.0798	-0.0915	-0.0799	-0.0327
Stat. (6 donors)	-0.0148	0.1280	-0.0706	-0.0723	0.0021

The best performing method in each subset was identified as:

- Stat. (6 donors) for all sites

- ReFH2 for SAAR ≥ 800
- Stat. (6 donors) for SAAR < 800
- Stat. (6 donors) for BFIHOST ≥ 0.65
- Stat. (6 donors) for BFIHOST < 0.65

Table 19 – RMSE in QMED estimated by six current and recent flood estimation methods in urban catchments

Subset	All (57 sites)	SAAR ≥ 800 (16 sites)	SAAR < 800 (41 sites)	BFIHOST ≥ 0.65 (13 sites)	BFIHOST < 0.65 (44 sites)
ReFH1	0.8529	0.5325	0.9491	1.4718	0.5499
ReFH2	0.6218	0.4765	0.6700	0.8426	0.5396
MacD & Fraser	0.7167	0.4797	0.7901	1.0421	0.5869
Stat. (0 donors)	0.5956	0.4490	0.6438	0.7724	0.5322
Stat. (1 donor)	0.5812	0.4069	0.6364	0.7953	0.5007
Stat. (6 donors)	0.5849	0.3737	0.6490	0.8125	0.4982

The best performing method in each subset was identified as:

- Stat. (1 donor) for all sites
- Stat. (6 donors) for SAAR ≥ 800
- Stat. (1 donor) for SAAR < 800
- Stat. (0 donors) for BFIHOST ≥ 0.65
- Stat. (6 donors) for BFIHOST < 0.65

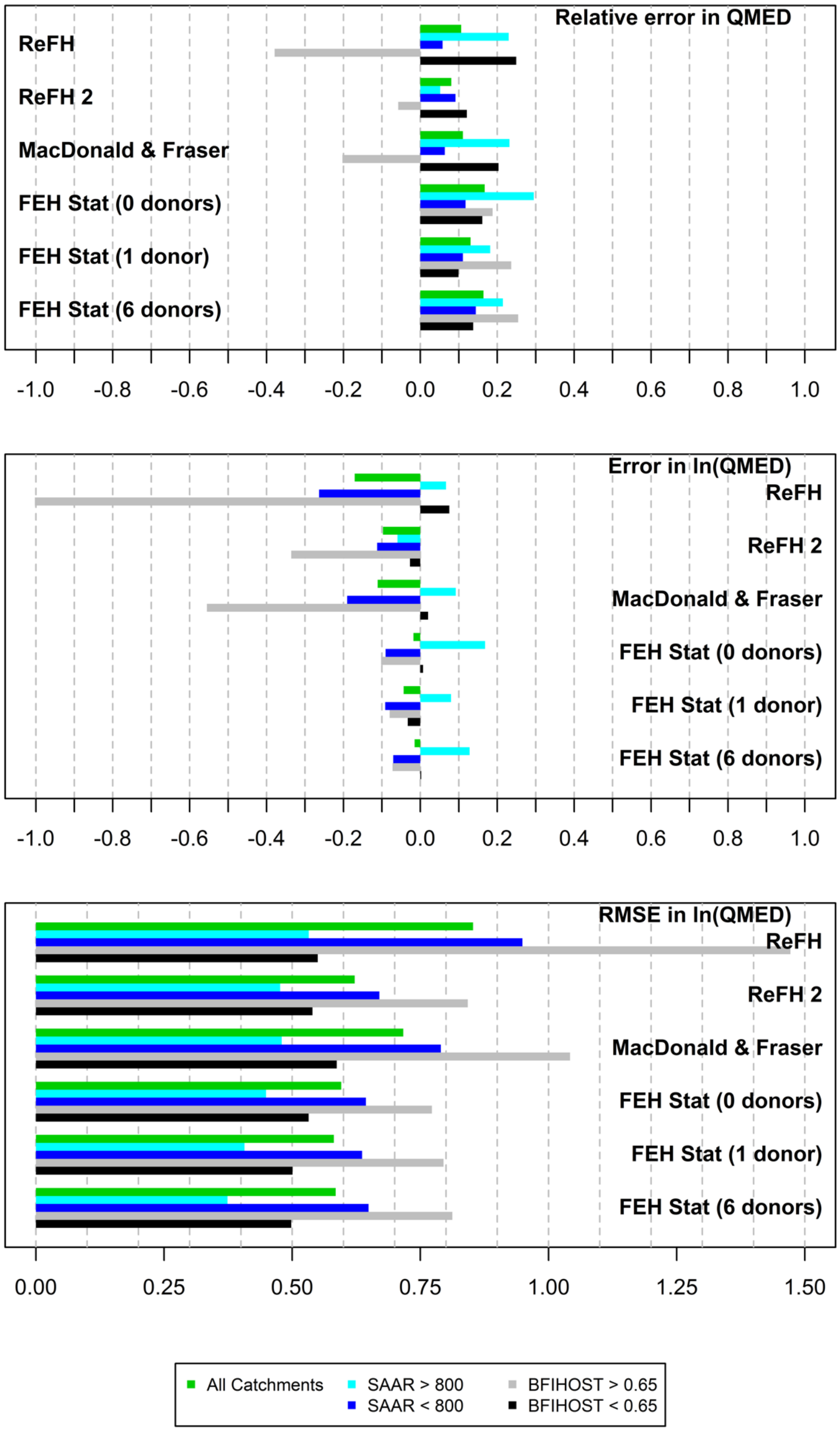

Figure 19 - Relative error, ln-error and RMSE in QMED estimated by six current and recent flood estimation methods (urban catchments)

The three graphs in Figure 19 quantify the relative error (top graph), ln-error (middle graph) and RMSE (bottom graph) of QMED, as estimated by the six methods in urbanised catchments, as a whole and when subset by BFIHOST and SAAR:

- ReFH
- ReFH2
- MacDonald & Fraser
- FEH Stat (0 donors)
- FEH Stat (1 donor)
- FEH Stat (6 donors)

Results are shown for:

- All catchments (green bars)
- SAAR >800 (light blue bars)
- SAAR < 800 (dark blue bars)
- BFIHOST >0.65 (grey bars)
- BFIHOST <0.65 (black bars)

4.8 Summary of results

A total of 146 catchments from the ‘extended’ data set of 217 small, gauged catchments (207 without the influence of lakes and reservoirs) described in Section 3 was classified as ‘high-quality’ and analysed separately. Six methods of QMED estimation were applied to the data sets and the results were compared, focusing particularly on three statistics: relative error, ln-error and root-mean-square error (RMSE). The results for essentially rural and urbanised catchments were analysed separately. The six methods were:

- ReFH1
- ReFH2 (using FEH99 rainfall inputs, assuming winter seasonality with parameters set to default values)
- MacDonald and Fraser’s improved method for estimating QMED (MacDonald and Fraser, 2013).
- FEH statistical method with no donor transfer
- FEH statistical method with data transfer from one donor site
- FEH statistical method with data transfer from six donor sites

The main conclusions arising from comparing methods of estimating QMED are summarised below.

- For essentially rural catchments, the best performing method overall was found to be ReFH2, although the performance of the FEH statistical method was a close second. ReFH2 performed particularly well in wetter and less permeable catchments.
- ReFH2 was found to be a substantial improvement on ReFH1, even in permeable catchments, although the FEH statistical method performed better in these cases. However, only 14 rural catchments were classified as permeable, making it difficult to draw firm conclusions involving permeable rural catchments.
- Error in design flood estimates was not found to be related to catchment area in the data set. Some of the largest positive errors in estimated QMED were found for catchments of around 20 km^2.
- The results suggest that donor transfer continues to serve a useful purpose in the FEH statistical method for small catchments. Using six donors may not always reduce error compared with using a single donor in small catchments.
- Only 43 of the catchments in the ‘high-quality’ data set are urbanised (that is, not essentially rural), making it difficult to draw definitive conclusions about the performance of the six methods in urban catchments. However, ReFH2 and the FEH statistical method perform better than the ReFH1 and MacDonald and Fraser methods as expected.
- The FEH statistical method is recommended for estimating QMED in urbanised catchments ($URBEXT_{2000} \geq 0.03$). However, ReFH2 performs almost equally well and is recommended in urban catchments where SAAR ≥ 800 mm.
- The FEH statistical method may overestimate QMED in small urban catchments with high average annual rainfall, and so should be compared with ReFH2 results in these situations.
- The analysis of urban catchments in the ‘extended’ data set did not find any substantial differences in the error statistics when compared to the ‘high-quality’ urban set.

5. Evaluating QT estimation methods

5.1 Data summary

The study data set of 217 small, gauged catchments has been described in detail in Section 3. Gauged AMAX records are available for 192 catchments, while only pre-computed QMED values are available for the other 25. Therefore, it is possible to generate at-site L-moments for 192 catchments. Enhanced single-site analyses, which can provide benchmark estimates of flood peaks rarer than QMED, require at-site L-moments, but only apply to catchments with $URBEXT_{2000} < 0.03$. They are therefore possible in 140 catchments. Direct comparisons between index flood and rainfall-runoff analyses mean that catchments must have FARL ≥ 0.9. There are 133 catchments meeting all these criteria.

43 of these 133 catchments are flagged in the NRFA peak flow data set v3.3.4 as suitable for pooling. Therefore, there is confidence in high-level gaugings at these sites (judged likely to be within 30% of the true value) and consequently in enhanced single-site QT estimates that pool data from records at these sites. None of the additional sites checked in Task 1 of this project were found to match this level of confidence.

Table 20 illustrates the division of catchments in each data set by urbanisation level. As in Section 4, catchments with $URBEXT_{2000} < 0.03$ are 'rural', catchments with $URBEXT_{2000} \geq 0.15$ are 'heavily urbanised' and catchments with intermediate $URBEXT_{2000}$ values are 'moderately' urbanised. 'Heavily urbanised' and 'moderately urbanised' catchments together are 'urbanised', with no prefix.

Table 20 - Number of catchments per data set

Data set	Essentially rural ($URBEXT_{2000} < 0.03$)	Moderately urbanised ($0.03 \leq URBEXT_{2000} < 0.15$)	Heavily urbanised ($URBEXT_{2000} \geq 0.15$)	Total
'high-quality'	43	7	5	55
'Extended'	133	21	28	182

5.2 Method summary

The same methods that were described for estimating QMED in Section 4.3 will be evaluated in this section for estimating QT, apart from MacDonald and Fraser's method, which has no procedure for estimating flood peaks other than QMED. In the following sections, the FEH statistical method is used to refer to pooled analysis that treats the subject site as if it were ungauged. Enhanced single-site analysis (assuming a generalised logistic distribution) is used as the benchmark against which the performance of the methods is evaluated. Comparisons of the performance of ReFH2 used with the FEH99 and FEH13 rainfall models for estimating QT are presented in Appendix C.

5.3 Quantifying methods of estimating QT

The equations given in Section 4.4 for quantifying QMED estimation apply equally to estimating QT, with ‘observed’ values now corresponding to those obtained through enhanced single-site analysis. However, it is not appropriate to refer to the values of QT estimated by the tested methods as being in error, as the enhanced single-site estimate is itself an estimate of QT (albeit the best estimate according to current practice), which pools data from outside the site of interest. In the QT studies presented in this section, the terms ‘relative difference’, ‘ln-difference’ and ‘root mean square difference’ (RMSD) are used, to make it clear that the tested methods are benchmarked against an estimate of QT derived not entirely from at-site gauged data.

It is important to note that the benchmark values of QT and the method estimates of QT are not strictly independent. The enhanced single-site and FEH statistical (pooled) estimates are based on almost identical pooling-groups, the differences being the greatly reduced weighting given to non-site data in the enhanced single-site method, the exclusion of at-site data in the FEH statistical method, and the addition of extra sites to the FEH statistical pooling-group if necessary, to provide 500 years of pooled data. ReFH2, with FEH99 rainfall, features an optional adjustment coefficient, α, normally used for catchments in England, Wales and Northern Ireland (but not Scotland). This is applied to the ratio of C_{ini}/C_{MAX}, and was calibrated to match the magnitude of the T-year flood peak to the enhanced single-site estimate of the same flood peak at 546 calibration sites up to 1,000 km^2 (Wallingford HydroSolutions, 2015). Despite these dependencies, using enhanced single-site estimates as benchmarking values is considered less problematic than using at-site AMAX series values: while these values would be independent of the test estimates, the sampling uncertainty associated with even the longest records would increase the risk of mis specifying the benchmark QT values.

ReFH2 used with FEH13 rainfall does not apply the adjustment coefficient, as that combination was found to match the enhanced single-site estimate approximately as well as the FEH statistical method (without at-site data) for QMED and rarer flood peaks on a set of 328 catchments (493 at QMED only) distributed throughout Great Britain. Therefore, ReFH2 with FEH13 rainfall is independent of the FEH statistical method and the enhanced single-site method.

5.4 Analysis of ‘high-quality’ data set

According to current practice, the best benchmark estimates of rare flood peaks derive from a combination of at-site data (where these exist) and pooled data. ‘High-quality’ benchmark estimates are those that incorporate at-site data that is of ‘high-quality’ all the way up to the largest values. In the NRFA peak flow data set, the highest quality catchments are marked as suitable for pooling. The NRFA peak flow data set v3.3.4

contains 43 rural small catchments that are marked suitable for pooling, which form the data set considered in this section.

Figure 20 plots modelled versus benchmark QT for a selection of return periods from five to 200 years for the 43 ‘high-quality’ rural catchments with gauged annual maxima records and FARL ≥ 0.9. Benchmark QT values are determined through FEH enhanced single-site analysis (Kjeldsen and others, 2008), combining 4 to 64 years of at-site annual maxima data, at an enhanced weight, with annual maxima from hydrologically similar catchments, at a reduced weight, for a target pooling-group length of 500 years.

Figure 20 shows a good correspondence between all the tested methods and between each method and the enhanced single-site estimate. As observed throughout Section 4, correspondence is lower for catchments with observed QT values below 3 m^3/s. As far fewer Q200 than *Q*5 flood peaks are under 3 m^3/s, the spread in the differences, and therefore RMSD (root-mean-square difference), reduces with increasing return period.

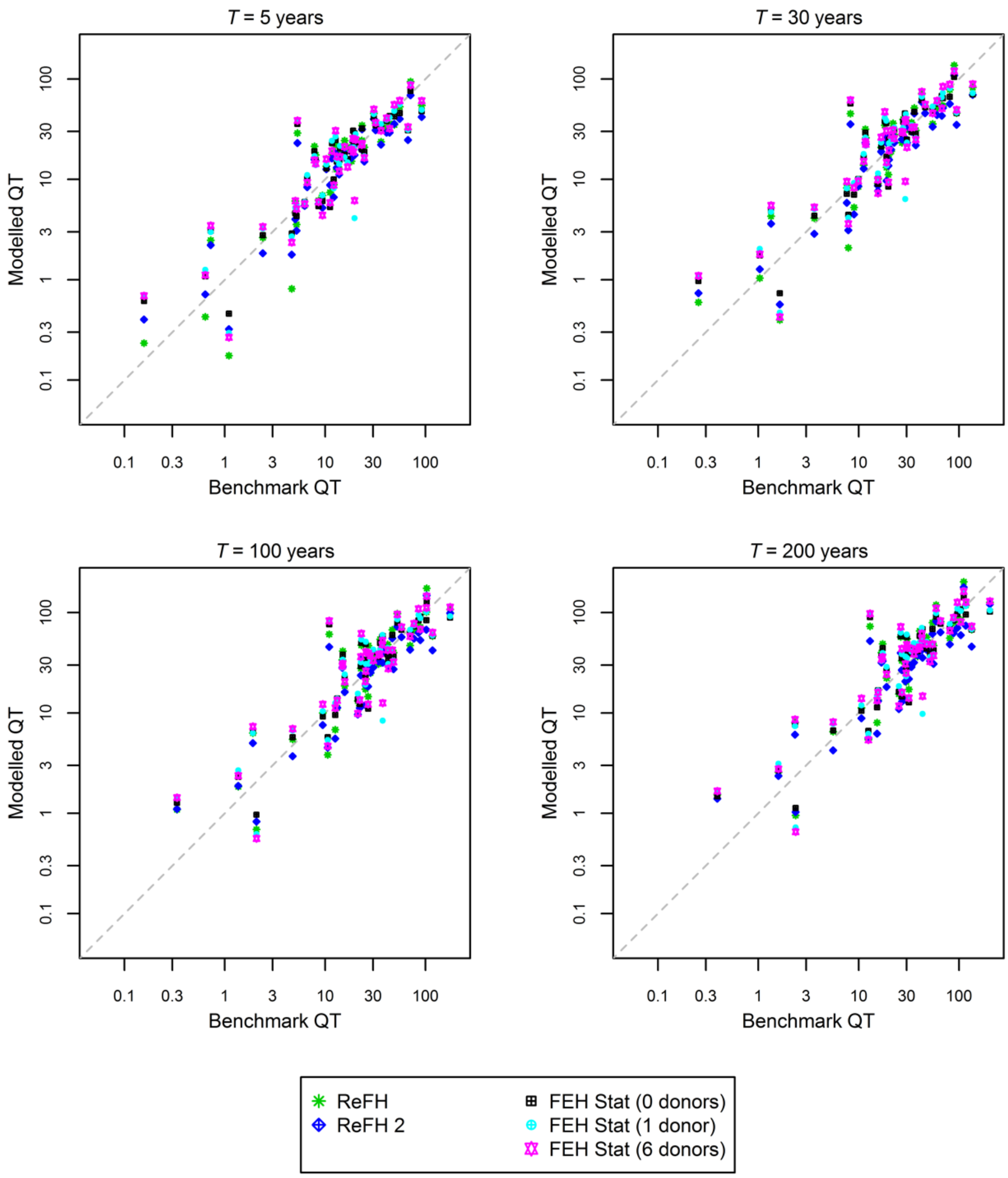

Figure 20 - Modelled versus benchmark QT at 43 sites with $URBEXT_{2000}$ < 0.03 and FARL ≥ 0.9 for 5-, 30-, 100- and 200-year return periods

The four scatter graphs in Figure 20 show the modelled QT on the y-axes (from 0.1 to 100) against benchmark QT on the x-axes (from 0.1 to 100) for a selection of return periods:

- T = 5 years: top left graph
- T = 30 years: top right graph
- T = 100 years: bottom left graph
- T = 200 years: bottom right graph

The graphs plot results for five methods:

- ReFH: green asterisks

- ReFH2: blue diamonds
- FEH stat (0 donors): black squares
- FEH stat (1 donor): light blue dots
- FEH stat (6 donors): pink stars

For the FEH statistical method variants, relative difference, ln-difference and RMSD do not vary with return period, showing that, on average, growth curves are neither steeper nor flatter than the benchmark enhanced single-site values.

Figure 21 shows how relative difference, ln-difference and RMSD in this group of catchments varies from two to 1,000 years. Differences are measured relative to the enhanced single-site estimate in all cases.

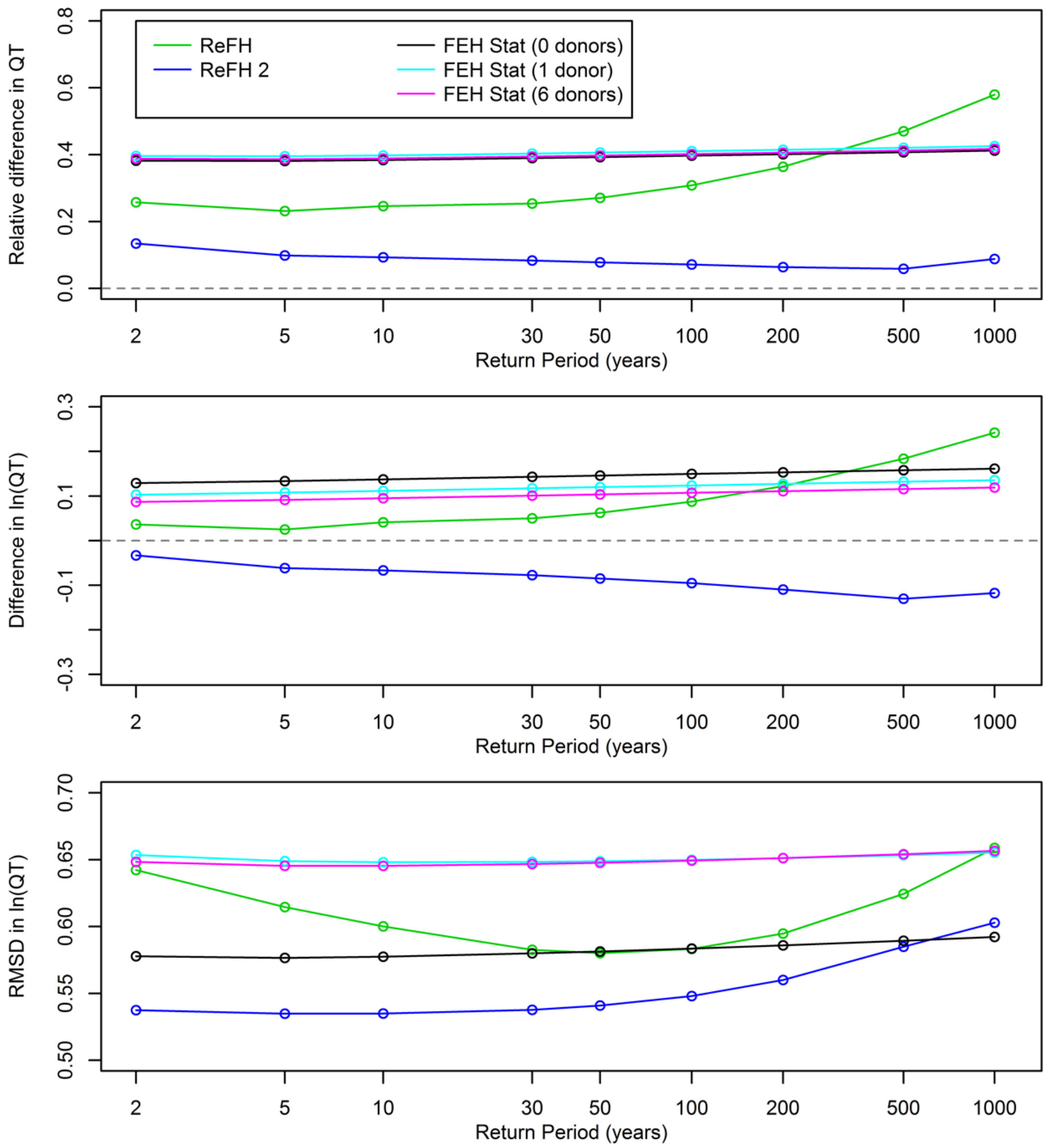

Figure 21 - Relative difference, ln-difference and RMSD in QT estimated by five current and recent methods (rural catchments)

The three line graphs in Figure 21 show how relative difference (top graph), ln-difference (middle graph) and RMSD (bottom graph) in this group of catchments varies within 1000 years:

- top graph: The x-axis plots the return period (up to 1000 years). The y-axis plots relative different in QT (from 0.0 to 0.8)
- middle graph: The x-axis plots the return period (up to 1000 years). The y-axis plots relative difference in ln(QT) from -0.3 to 0.3
- bottom graph: The x-axis plots the return period (up to 1000 years). The y-axis plots RMSD in ln(QT) from 0.50 to 0.70

The graphs plot results for the five methods:

- ReFH: green line
- ReFH2: blue line
- FEH stat (0 donors): black line
- FEH stat (1 donor): light blue line
- FEH stat (6 donors): pink line

Unusually, despite the number of donors not affecting relative or ln-difference, adding either one or six donors does increase RMSD dramatically. This is dominated by the effect of one site (25003), which receives donor adjustment mainly from one (25808) at which estimated QMED is 18 times larger than gauged (as a result of DTM area being 15 times greater than nominal area).

Although using multiple donors should be expected to mitigate the problem of transfer from single problematic sites, the combined weight of donors 2 to 6 is less than the weight of the main donor in this case. Since there is no strong justification for donor transfer in small catchments (Figure 13), and donor transfer increases the spread of differences without noticeably reducing the 'typical' difference, it is difficult to recommend donor transfer in small catchments in its current form. If it is to be beneficial, then it is perhaps wise to limit the donor pool to catchments with gauged QMED above a certain lower limit, for example 3 m^3/s.

This is shown in more detail in Figure 22, which plots Q30 and Q100 growth factors side-by-side and shows little difference between them, aside from the vertical axis scale. For ReFH1 and ReFH2, some variation with return period can be observed: ReFH1 overestimates by progressively larger amounts as return period increases, while ReFH2 underestimates (on ln-scale) by progressively larger amounts. ReFH2 shows a smaller variation with return period than ReFH1 and reports consistently lower RMSE over all return periods from two to 1,000 years – both improvements.

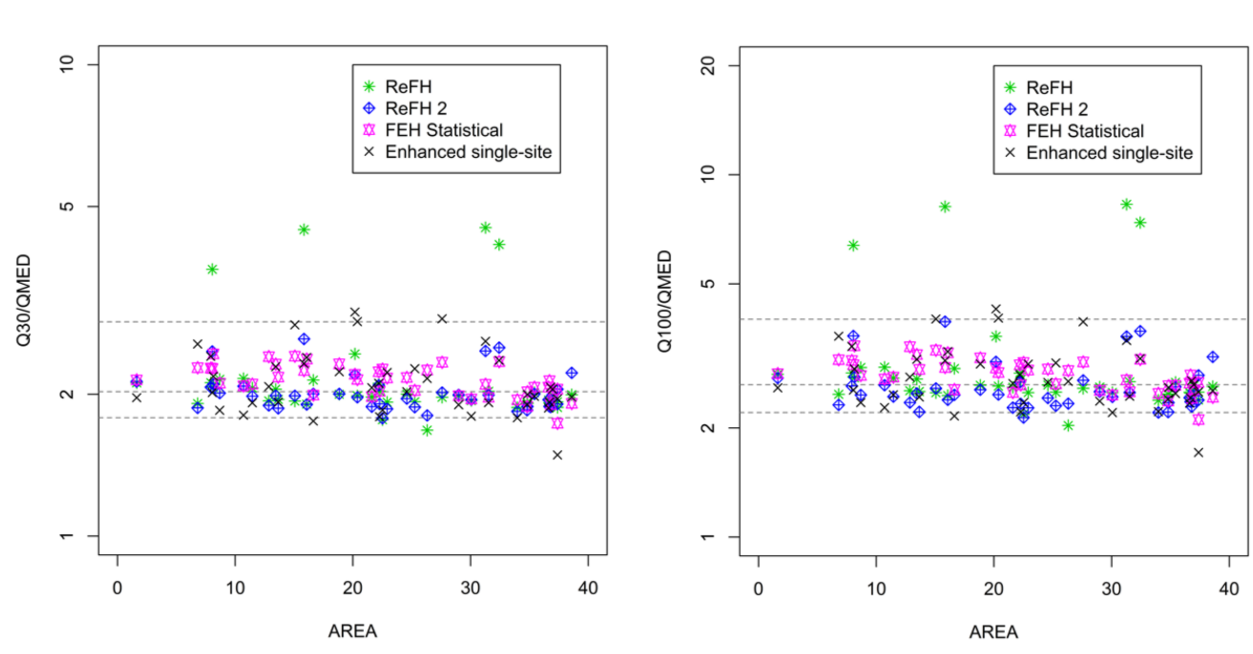

Figure 22 - Q30/QMED and Q100/QMED growth factors for rural catchments with FARL ≥ 0.9 and at-site AMAX records suitable for pooling

Figure 22 plots Q30 and Q100 growth factors side-by-side:

- left-hand graph: The x-axis plots AREA from 0 to 40. The y-axis plots Q30/QMED from 1 to 10
- right-hand graph: The x-axis plots AREA from 0 to 40. The y-axis plots Q100/QMED from 1 to 10

Results are plotted for:

- ReFH: green asterisks
- ReFH2: blue diamonds
- FEH statistical: pink stars
- Enhanced single-site: black crosses

As observed previously, relative difference is more positive than ln-difference, suggesting a dependency between QMED and difference. ReFH2 estimates lower QT peak values than the variants of the FEH statistical method, and ReFH1 does not perform at all well outside its calibration limit of 150 years.

Performance at T = 30 years

Tables 21, 22 and 23 and Figure 23 report relative difference, ln-difference and RMSD for all tested methods at a return period of 30 years, for rural catchments, including subsets of drier (SAAR < 800), wetter (SAAR ≥ 800), permeable (BFIHOST ≥ 0.65) and less-permeable (BFIHOST < 0.65) catchments. They are therefore analogous to Table 7 and Figure 9 but for Q30 rather than QMED. The wetter subset and less permeable subset each contain 38 members, corresponding to 88% of the total group. In addition, the wetter subset and less permeable subset have 36 members in common. Therefore, it is difficult to separate the effect of BFIHOST and the effect of SAAR, and any observations specific to drier or more permeable catchments may be subject to a large degree of sampling uncertainty. A further complication is that the Q30 and Q100 benchmark values considered in this section are subject to more uncertainty than the QMED benchmark values considered in Section 4.

Tables 21, 22 and 23 and Figure 23 show that the best performing method in terms of relative difference and RMSD is ReFH2, while the best in terms of ln-difference is ReFH1. ReFH2 also has the lowest RMSD within each subset of catchments, the lowest ln-difference within wetter and more permeable catchments, and the lowest relative difference in both the wetter and less permeable catchment subsets.

Although it is not always the strongest performer in terms of ln-difference, ReFH2 is the most consistent method across all subsets of catchments. In terms of RMSD, Figure 23 shows that all methods could improve the way that they model permeable catchments. Despite the small sample size, this is believed to be a legitimate effect, as it was observed at QMED with larger samples in previous sections.

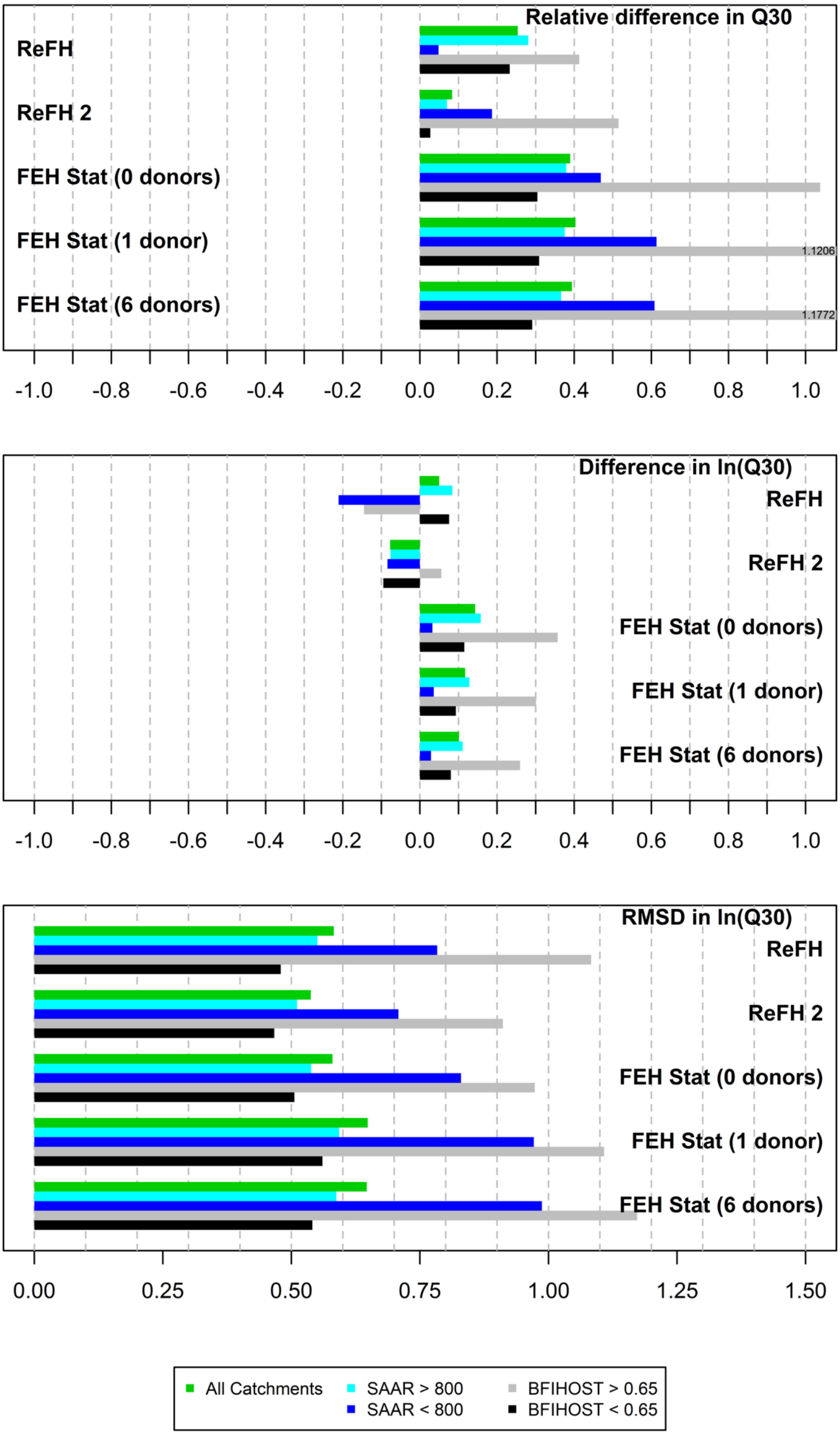

Figure 23 - Relative difference, ln-difference and RMSD in Q30 estimated by five current and recent flood estimation methods (rural catchments)

The three graphs in Figure 23 quantify the relative difference in Q30 (top graph), ln-difference in Q30 (middle graph) and RMSE in Q30 (bottom graph), as estimated by five methods in rural catchments:

- ReFH.
- ReFH2.
- FEH Stat (0 donors).
- FEH Stat (1 donor).
- FEH Stat (6 donors).

Results are shown for:

- All catchments (green bars).
- SAAR >800 (light blue bars).
- SAAR < 800 (dark blue bars).
- BFIHOST >0.65 (grey bars).
- BFIHOST <0.65 (black bars).

Table 21 – Relative difference in Q30 estimated by five current and recent flood estimation methods in rural catchments.

Subset	All (43 sites)	SAAR ≥ 800 (38 sites)	SAAR < 800 (5 sites)	BFIHOST ≥ 0.65 (5 sites)	BFIHOST < 0.65 (38 sites)
ReFH	0.2539	0.2809	0.0482	0.4130	0.2329
ReFH2	0.0836	0.0700	0.1870	0.5149	0.0269
Stat. (0 donors)	0.3897	0.3793	0.4690	1.0375	0.3045
Stat. (1 donor)	0.4033	0.3757	0.6133	1.1206	0.3089
Stat. (6 donors)	0.3943	0.3661	0.6082	1.1772	0.2913

The best performing method in each subset was identified as:

- ReFH2 for all sites.
- ReFH2 for SAAR ≥ 800.
- ReFH for SAAR < 800.
- ReFH for BFIHOST ≥ 0.65.
- ReFH2 for BFIHOST < 0.65.

Table 22 – In-difference in Q30 estimated by five current and recent flood estimation methods in rural catchments.

Subset	All (43 sites)	SAAR ≥ 800 (38 sites)	SAAR < 800 (5 sites)	BFIHOST ≥ 0.65 (5 sites)	BFIHOST < 0.65 (38 sites)
ReFH	0.0500	0.0837	-0.2118	-0.1451	0.0757
ReFH2	-0.0774	-0.0765	-0.0844	0.0555	-0.0949
Stat. (0 donors)	0.1432	0.1578	0.0324	0.3569	0.1151
Stat. (1 donor)	0.1173	0.1280	0.0360	0.3008	0.0932
Stat. (6 donors)	0.1010	0.1105	0.0285	0.2597	0.0801

The best performing method in each subset was identified as:

- ReFH for all sites.
- ReFH2 for SAAR ≥ 800.
- Stat. (6 donors) for SAAR < 800.
- ReFH2 for BFIHOST ≥ 0.65.
- ReFH for BFIHOST < 0.65.

Table 23 – RMSD in Q30 estimated by five current and recent flood estimation methods in rural catchments.

Subset	All (43 sites)	SAAR ≥ 800 (38 sites)	SAAR < 800 (5 sites)	BFIHOST ≥ 0.65 (5 sites)	BFIHOST < 0.65 (38 sites)
ReFH	0.5826	0.5507	0.7836	1.0829	0.4793
ReFH2	0.5377	0.5111	0.7081	0.9111	0.4668
Stat. (0 donors)	0.5800	0.5385	0.8299	0.9733	0.5059
Stat. (1 donor)	0.6484	0.5929	0.9715	1.1081	0.5605
Stat. (6 donors)	0.6468	0.5875	0.9871	1.1725	0.5408

The best performing method in each subset was identified as:

- ReFH2 for all sites.
- ReFH2 for SAAR ≥ 800.
- ReFH2 for SAAR < 800.
- ReFH2 for BFIHOST ≥ 0.65.

- ReFH2 for BFIHOST < 0.65.

Using donors in the FEH statistical method always causes ln-difference to reduce, except when moving from zero to one donor in drier catchments. This is despite the extremely poor choice of donor for catchment 25003. The increase in RMSD resulting from donor transfer cannot be attributed to the poor choice of donor for catchment 25003. Indeed, donor transfer increases both relative error and RMSD in the drier and permeable catchment subsets, neither of which contain catchment 25003. This continues to show the limited utility of donors in small catchments.

For all methods, RMSD decreases notably from QMED to Q30. This may result from the significant increase in quality (and decrease in size) of the data set as catchments that are not suitable for pooling are excluded from the QT analysis. This is not shown consistently in Figure 21 as the RMSD of QMED there represents that of the 43 ‘high-quality’, pooling-suitable catchments rather than the 103 ‘high-quality’ QMED-suitable catchments considered in Section 4.6.

Performance at T = 100 years

Tables 24, 25 and 26 and Figure 24 report relative difference, ln-difference and RMSD for all tested methods at a return period of 100 years. These appear similar to Tables 21, 22 and 23 and Figure 23, the main differences being an increase in relative difference for low-SAAR catchments in ReFH1 and ReFH2, a large increase in relative difference in permeable catchments in ReFH1, corresponding shifts in ln-difference for these methods and catchment subsets, and a slight compression of RMSD values across the four catchment subsets for all methods. The FEH statistical method variants improve their modelling of permeable catchments relative to Q30, although relative and ln-differences are still very high.

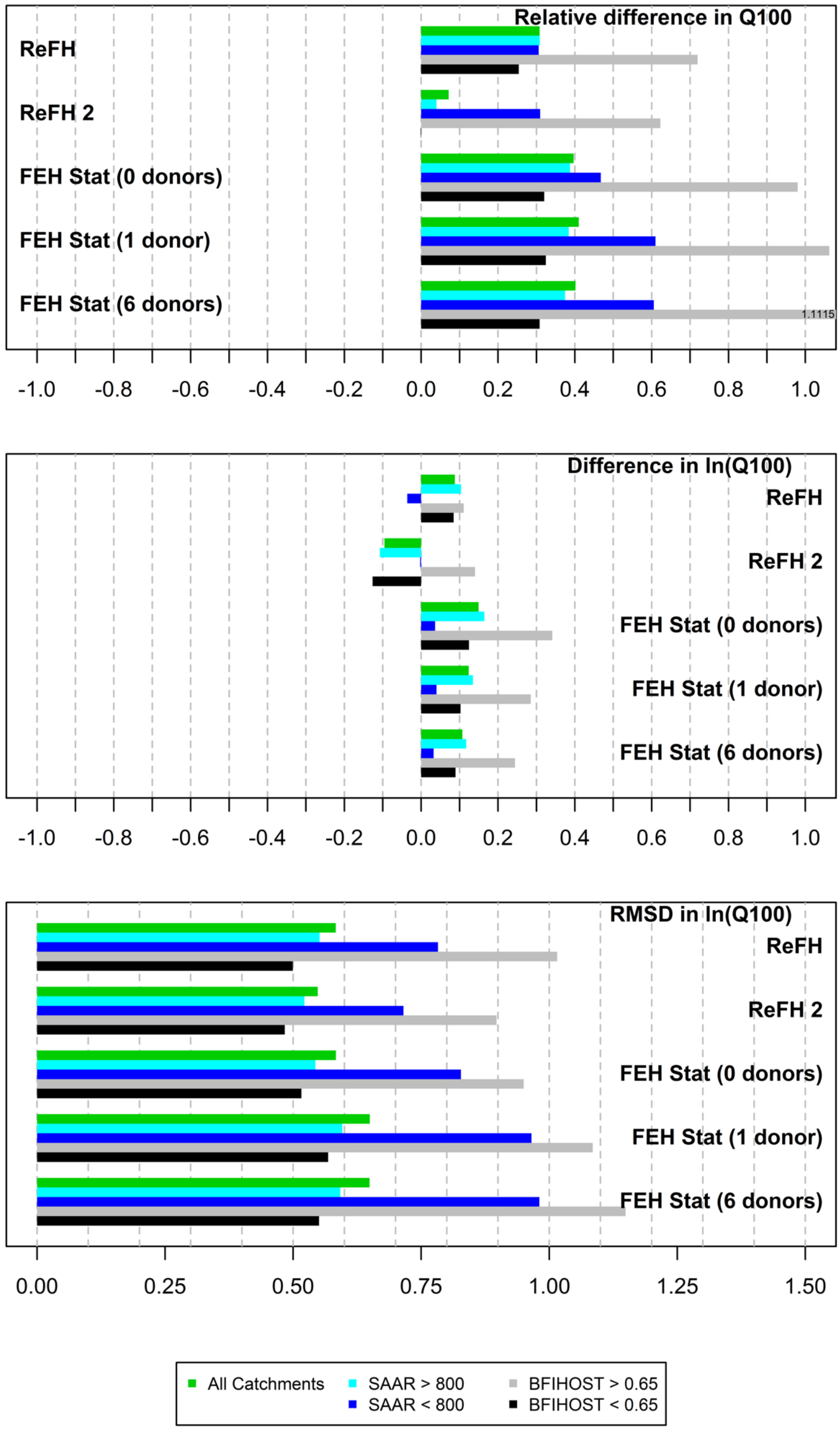

Figure 24 - Relative difference, ln-difference and RMSD in Q100 estimated by five current and recent flood estimation methods (rural catchments)

The three graphs in Figure 24 quantify the relative difference in Q100 (top graph), ln-difference in Q100 (middle graph) and RMSE in Q100 (bottom graph), as estimated by five methods in rural catchments:

- ReFH.
- ReFH2.
- FEH Stat (0 donors).
- FEH Stat (1 donor).
- FEH Stat (6 donors).

Results are shown for:

- All catchments (green bars).
- SAAR >800 (light blue bars).
- SAAR < 800 (dark blue bars).
- BFIHOST >0.65 (grey bars).
- BFIHOST <0.65 (black bars).

Table 24 – Relative difference in Q100 estimated by five current and recent flood estimation methods in rural catchments

Subset	All (43 sites)	SAAR ≥ 800 (38 sites)	SAAR < 800 (5 sites)	BFIHOST ≥ 0.65 (5 sites)	BFIHOST < 0.65 (38 sites)
ReFH	0.3082	0.3086	0.3058	0.7193	0.2541
ReFH2	0.0715	0.0401	0.3098	0.6228	-0.0011
Stat. (0 donors)	0.3970	0.3877	0.4676	0.9798	0.3203
Stat. (1 donor)	0.4104	0.3841	0.6100	1.0625	0.3246
Stat. (6 donors)	0.4016	0.3747	0.6056	1.1115	0.3082

The best performing method in each subset was identified as:

- ReFH2 for all sites.
- ReFH2 for SAAR ≥ 800.
- ReFH for SAAR < 800.
- ReFH2 for BFIHOST ≥ 0.65.
- ReFH for BFIHOST < 0.65.

Table 25 – ln-difference in Q100 estimated by five current and recent flood estimation methods in rural catchments

Subset	All (43 sites)	SAAR ≥ 800 (38 sites)	SAAR < 800 (5 sites)	BFIHOST ≥ 0.65 (5 sites)	BFIHOST < 0.65 (38 sites)
ReFH	0.0875	0.1038	-0.0362	0.1108	0.0844
ReFH2	-0.0954	-0.1076	-0.0026	0.1402	-0.1264
Stat. (0 donors)	0.1496	0.1645	0.0362	0.3415	0.1243
Stat. (1 donor)	0.1237	0.1347	0.0398	0.2853	0.1024
Stat. (6 donors)	0.1074	0.1173	0.0323	0.2443	0.0894

The best performing method in each subset was identified as:

- ReFH for all sites.
- ReFH for SAAR ≥ 800.
- ReFH2 for SAAR < 800.
- ReFH for BFIHOST ≥ 0.65.
- ReFH for BFIHOST < 0.65.

Table 26 – RMSD in Q100 estimated by five current and recent flood estimation methods in rural catchments

Subset	All (43 sites)	SAAR ≥ 800 (38 sites)	SAAR < 800 (5 sites)	BFIHOST ≥ 0.65 (5 sites)	BFIHOST < 0.65 (38 sites)
ReFH	0.5833	0.5517	0.7829	1.0155	0.4993
ReFH2	0.5480	0.5220	0.7155	0.8972	0.4837
Stat. (0 donors)	0.5835	0.5433	0.8275	0.9502	0.5163
Stat. (1 donor)	0.6499	0.5961	0.9653	1.0850	0.5683
Stat. (6 donors)	0.6493	0.5920	0.9808	1.1491	0.5508

The best performing method in each subset was identified as:

- ReFH2 for all sites.
- ReFH2 for SAAR ≥ 800.
- ReFH2 for SAAR < 800.
- ReFH2 for BFIHOST ≥ 0.65.

- ReFH2 for BFIHOST < 0.65.

5.5 Analysis of ‘extended’ data set

Figure 25 plots modelled against benchmark QT for return periods from five to 200 years for the ‘extended’ data set of 133 rural catchments. Figure 25 shows a generally good correlation between benchmark and modelled QT for all methods, although correlation tends to be better when the observed events are larger. ReFH1 appears to underestimate the 5-year flood peak more than the other four methods, particularly so when this is small ($\lesssim$ 3 m^3/s). Outliers highlighted in Figure 8 remain outliers at longer return periods. Two further outliers are identified, both of which are characterised by enhanced single-site QT values below 1 m^3/s for T up to 200 years. One (25809) is exceptionally small, the other (26802) may not align with its groundwater catchment according to the NRFA (2015).

Figure 26 quantifies the relative difference, ln-difference and RMSD of QT, as estimated by the five methods being considered, in the 133 rural catchments in the ‘extended’ data set. This shows that that both the relative and ln-differences in FEH statistical estimates of QT are broadly independent of return period, up to at least 1,000 years.

ReFH2 tends to show reducing levels of relative difference, but increasingly negative levels of ln-difference, as T increases from two to 1,000 years. The decrease in relative difference is shown to be due to one extreme outlier (205034), whose relative difference decreases from approximately 29 to 16 as T increases from two to 1,000 years. This catchment is an extreme outlier in all methods and, as such, later figures in this section (Figure 27 and Figure 28) present superimposed results with the outlier excluded. It has already been stated that the DTM area of this catchment is approximately 40 times its measured area and this will certainly affect the performance of the methods. The catchment is also very heavily forested.

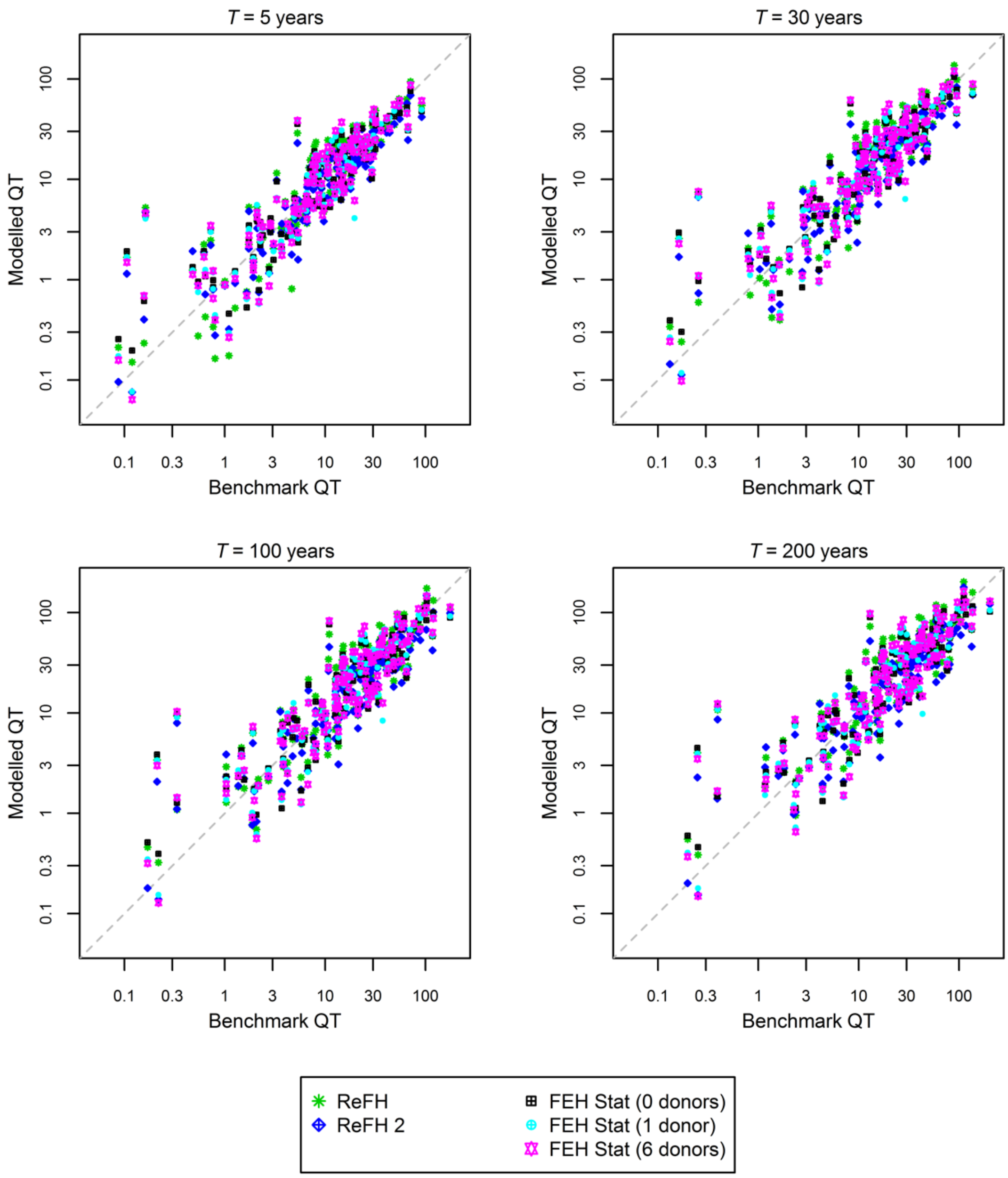

Figure 25 - Modelled vs. benchmark *QT* at 133 sites with URBEXT2000 < 0.03 and FARL ≥ 0.9 for 5-, 30-, 100- and 200-year return periods

The four scatter graphs in Figure 25 show the modelled QT on the y-axes (from 0.1 to 100) against benchmark QT on the x-axes (from 0.1 to 100) for a selection of return periods:

- T = 5 years: top left graph.
- T = 30 years: top right graph.
- T = 100 years: bottom left graph.
- T = 200 years: bottom right graph.

The graphs plot results for five methods:

- ReFH: green asterisks.
- ReFH2: blue diamonds.
- FEH stat (0 donors): black squares.
- FEH stat (1 donor): light blue dots.
- FEH stat (6 donors): pink stars.

ReFH1 shows little relation between relative or ln-difference and T until approximately 150 years – the limit of the adjustment coefficient that was introduced specifically to force correspondence between rainfall and flood return periods and, by extension, the limit of the method's applicability range. After this point, the method grows too severely with return period.

For all current methods (ReFH2 and FEH statistical variants), RMSD is shown to increase with return period, while for ReFH1, it is shown to be remain approximately constant until $T \approx 100$ years, before increasing rapidly. All current methods report lower RMSD than ReFH1 over the full two to 1,000 year range.

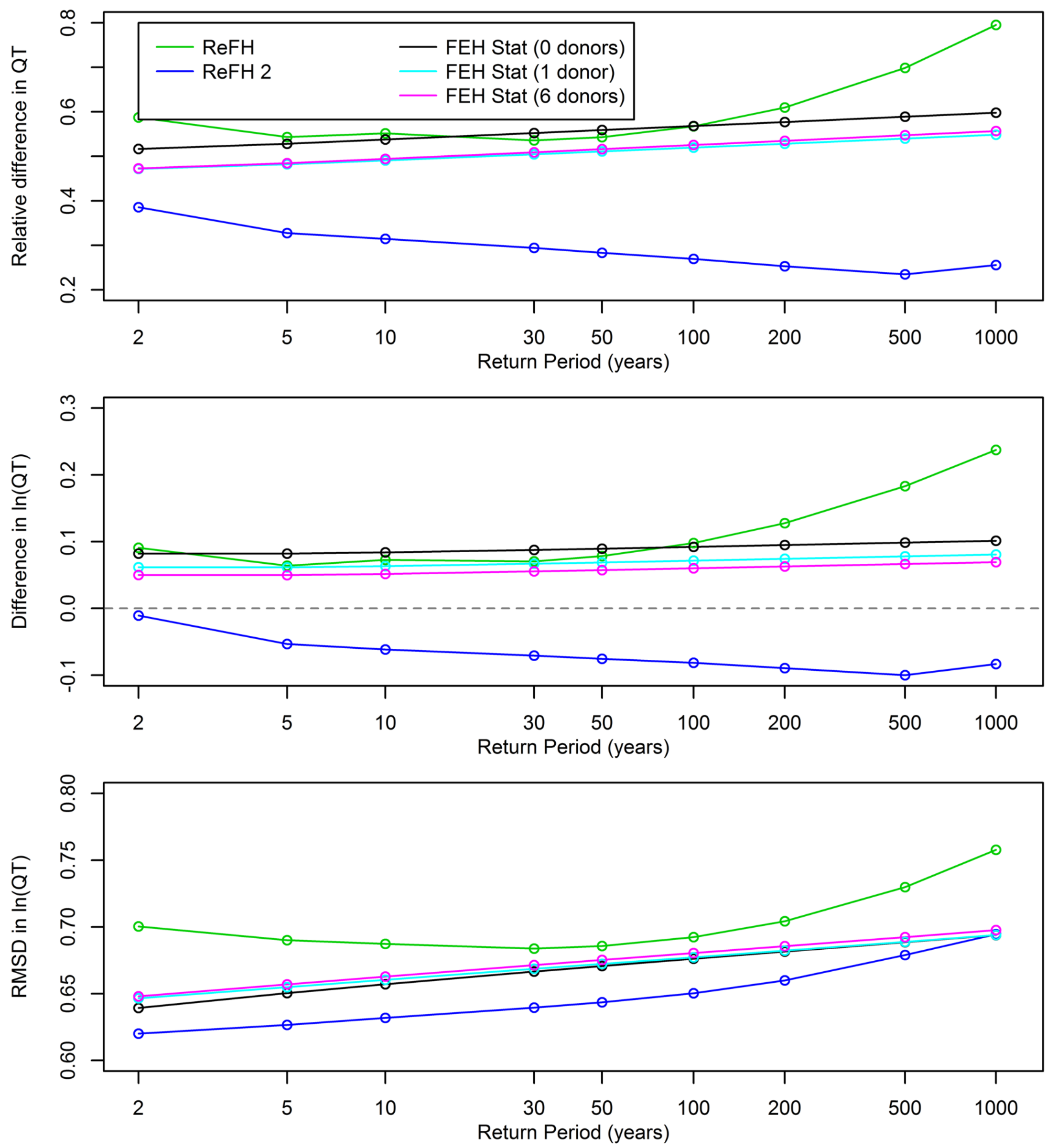

Figure 26 - Relative difference, ln-difference and RMSD in QT estimated by five current and recent methods (rural catchments)

The three line graphs in Figure 26 show how relative difference (top graph), ln-difference (middle graph) and RMSD (bottom graph) in this group of catchments varies within 1000 years:

- Top graph: The x-axis plots the return period (up to 1000 years). The y-axis plots relative different in QT (from 0.0 to 0.8).
- Middle graph: The x-axis plots the return period (up to 1000 years). The y-axis plots relative difference in ln(QT) from -0.3 to 0.3.
- Bottom graph: The x-axis plots the return period (up to 1000 years). The y-axis plots RMSD in ln(QT) from 0.60 to 0.80.

The graphs plot results for the five methods:

- ReFH: green line.
- ReFH2: blue line.
- FEH stat (0 donors): black line.
- FEH stat (1 donor): light blue line.
- FEH stat (6 donors): pink line.

For most of the range of tested return periods, the lowest RMSD is reported by ReFH2. There is a crossover point with the FEH statistical method at a return period of around 1,000 years, after which the performance has not been quantified.

Interestingly, even though donor transfer is ineffective at improving QMED estimates, and even worsens the performance of the FEH statistical method in the 'high-quality' data set, it appears to be slightly beneficial here, reducing the average relative and ln-differences without affecting RMSD.

Table 27 – Relative difference in Q30 estimated by five current and recent flood estimation methods in rural catchments.

Subset	All (133 sites)	SAAR ≥ 800 (108 sites)	SAAR < 800 (25 sites)	BFIHOST ≥ 0.65 (18 sites)	BFIHOST < 0.65 (115 sites)
ReFH	0.5358	0.6735	-0.0591	0.2363	0.5826
ReFH2	0.2942	0.3480	0.0616	0.3501	0.2854
Stat. (0 donors)	0.5523	0.6886	-0.0365	0.4680	0.5655
Stat. (1 donor)	0.5045	0.6260	-0.0203	0.3820	0.5237
Stat. (6 donors)	0.5090	0.6375	-0.0461	0.3529	0.5335

The best performing method in each subset was identified as:

- ReFH2 for all sites.
- ReFH2 for SAAR ≥ 800.
- Stat. (1 donor) for SAAR < 800.
- ReFH for BFIHOST ≥ 0.65.
- ReFH2 for BFIHOST < 0.65.

Table 28 – Ln-difference in Q30 estimated by five current and recent flood estimation methods in rural catchments.

Subset	All (133 sites)	SAAR ≥ 800 (108 sites)	SAAR < 800 (25 sites)	BFIHOST ≥ 0.65 (18 sites)	BFIHOST < 0.65 (115 sites)
ReFH	0.0703	0.1449	-0.2517	-0.1127	0.0990
ReFH2	-0.0708	-0.0463	-0.1766	0.0049	-0.0827
Stat. (0 donors)	0.0875	0.1693	-0.2658	0.1192	0.0826
Stat. (1 donor)	0.0669	0.1428	-0.2607	0.0216	0.0740
Stat. (6 donors)	0.0554	0.1338	-0.2833	-0.0374	0.0699

The best performing method in each subset was identified as:

- Stat. (6 donors) for all sites.
- ReFH2 for SAAR ≥ 800.
- ReFH2 for SAAR < 800.
- ReFH2 for BFIHOST ≥ 0.65.
- Stat. (1 donor) for BFIHOST < 0.65.

Table 29 – RMSD in Q30 estimated by five current and recent flood estimation methods in rural catchments.

Subset	All (133 sites)	SAAR ≥ 800 (108 sites)	SAAR < 800 (25 sites)	BFIHOST ≥ 0.65 (18 sites)	BFIHOST < 0.65 (115 sites)
ReFH	0.6837	0.6922	0.6460	0.8287	0.6582
ReFH2	0.6395	0.6283	0.6856	0.7922	0.6122
Stat. (0 donors)	0.6666	0.6627	0.6832	0.7531	0.6520
Stat. (1 donor)	0.6686	0.6609	0.7009	0.8010	0.6455
Stat. (6 donors)	0.6714	0.6645	0.7006	0.8409	0.6408

The best performing method in each subset was identified as:

- ReFH2 for all sites.
- ReFH2 for SAAR ≥ 800.
- ReFH for SAAR < 800.
- Stat. (0 donors) for BFIHOST ≥ 0.65.

- ReFH2 for BFIHOST < 0.65.

Performance at T = 30 years

Tables 27, 28 and 29 and Figure 27 report relative difference, ln-difference and RMSD for all tested methods at a return period of 30 years, for rural catchments, including subsets of drier, wetter, permeable and less permeable catchments. The presence or absence of the outlier catchment, 205034, has a quantitative effect on relative difference, ln-difference and RMSD, both overall and in the high-SAAR and low-BFIHOST subsets, but does not change the results qualitatively. It is noted that, as the numbers of drier catchments and permeable catchments are increased to 25 and 18 respectively, with an overlap of 8, it is now possible to make tentative assessments of the methods in relation to BFIHOST and SAAR values.

At Q30, the best method in terms of relative difference and RMSD overall, in wetter catchments and in less permeable catchments is ReFH2. This is also the best method in terms of ln-difference in the drier, wetter and permeable catchment subsets. ReFH2's strong performance in high-SAAR, low-BFIHOST catchments may reflect the proportion of its calibration data set located in Scotland, where lower BFIHOST and higher SAAR values are common. On the other hand, it should be noted that ReFH2 is also frequently a strong performer in lower-SAAR and higher-BFIHOST catchments.

Neither relative nor ln-differences in ReFH2 estimates of Q30 are shown to depend very strongly on either SAAR or BFIHOST, compared with other methods. This means that the ReFH2 method performs most consistently regarding these catchment properties. The FEH statistical method appears to show more dependency on SAAR at Q30 than it does at QMED. This reinforces the idea that the effect of SAAR in the FEH statistical method may be too powerful when the method is applied to small catchments.

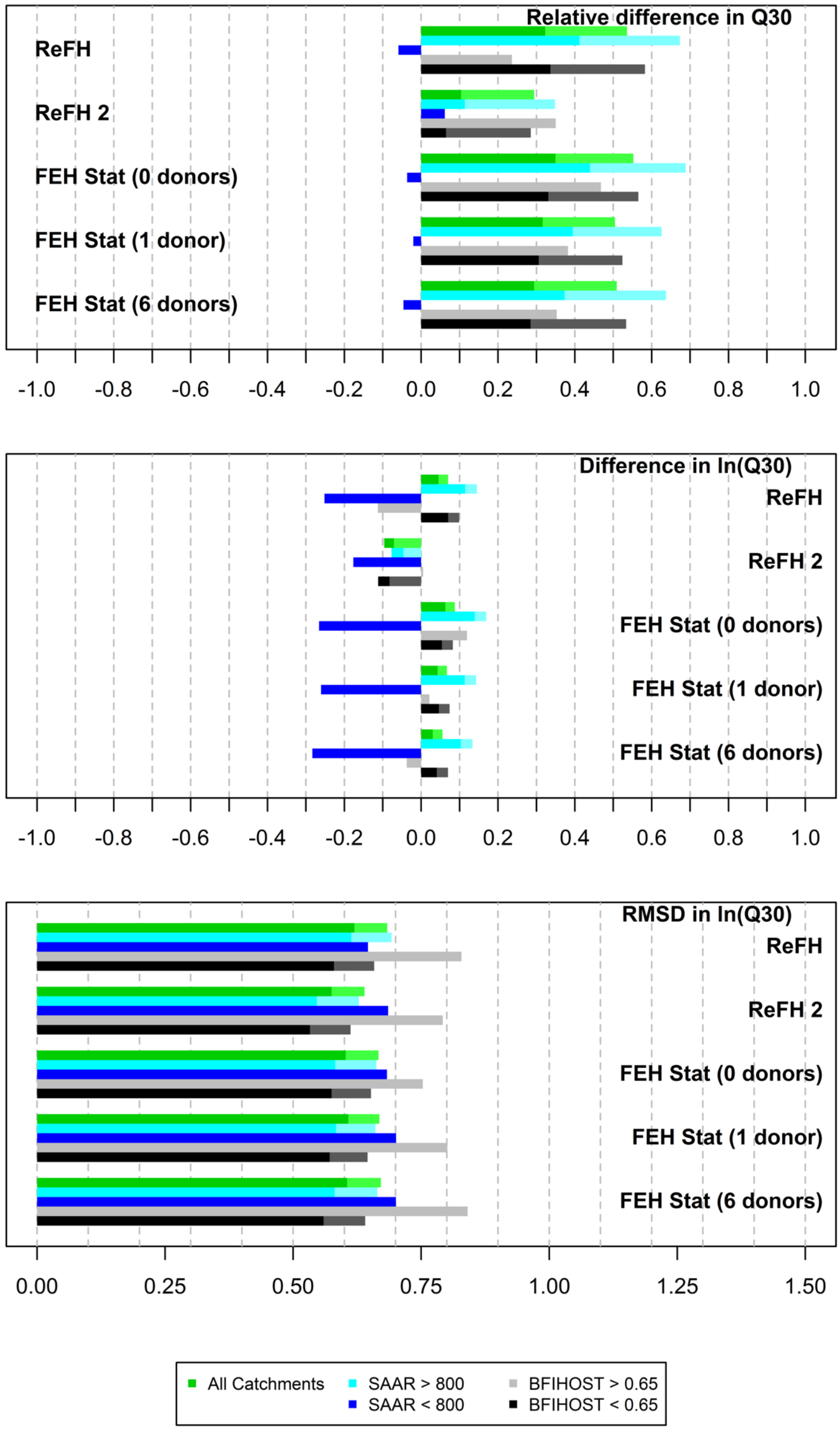

Figure 27 - Relative difference, ln-difference and RMSD in *Q*30 estimated by five current and recent flood estimation methods (rural catchments)

The three graphs in Figure 27 quantify the relative difference in Q30 (top graph), ln-difference in Q30 (middle graph) and RMSD in Q30 (bottom graph), as estimated by five methods in rural catchments:

- ReFH.
- ReFH2.
- FEH Stat (0 donors).
- FEH Stat (1 donor).
- FEH Stat (6 donors).

Results are shown for:

- All catchments (green bars).
- SAAR >800 (light blue bars).
- SAAR < 800 (dark blue bars).
- BFIHOST >0.65 (grey bars).
- BFIHOST <0.65 (black bars).

Darker sections on bars show results excluding 205034.

Similarly to its performance at QMED, the reasonable overall performance of ReFH1 (particularly in terms of relative difference) masks its weaker performance in more permeable catchments. It is not surprising that ReFH1 does not perform well in the permeable catchment subset (BFIHOST ≥ 0.65), as only three permeable catchments were used in its calibration.

In contrast to the 'high-quality' data set, using donors has a negligible effect on RMSD in the 'extended' data set of rural catchments. In the group of permeable catchments, however, using donors increases RMSD. Within each catchment subset, increasing the number of donors from zero to one improves the relative and ln-difference in Q30 marginally. Taken together, this further shows the limited utility of donor transfer in small catchments.

While ReFH2 has the lowest RMSD overall, no variant of the FEH statistical method is far behind. In wetter and less permeable catchments, RMSD increases from QMED to Q30, while in more permeable catchments, RMSD decreases from QMED to Q30. The fact that RMSD does not remain consistent from QMED to Q30 suggests that the description of flood growth factors can be improved in all methods.

Performance at T = 100 years

Tables 30, 31 and 32 and Figure 28 report relative difference, ln-difference and RMSD for all tested methods at a return period of 100 years.

At Q100, the best performing method within each subset of rural catchments, except permeable, according to each performance statistic, is the same at Q100 as at Q30. This means that ReFH2 performs best in wetter and less permeable catchments but does not perform poorly in drier or more permeable catchments. The best method in terms of both relative difference and RMSD is ReFH2, while the best method in terms of ln-difference is FEH statistical with six donors. However, the strong overall performance of the FEH statistical method in terms of ln-difference masks the dependencies on BFIHOST and SAAR that can be seen in Tables 30, 31 and 32 and Figure 28. This suggests that ReFH2 may be more suitable for estimating the 100-year flood peak in small catchments.

As at Q30, donor transfer increases the RMSD of the FEH statistical method in the high-BFIHOST subsets but has a negligible effect overall. It is worth noting that estimating Q100 for the subset of high-BFIHOST catchments is not highly problematic for any modelling method in terms of RMSD. This is very different from what was observed at QMED.

Figure 29 plots Q100 over QMED against AREA for all 133 catchments considered in this section. This shows modelled Q100/QMED growth factors ranging from around 2.0 to 9.6 and enhanced single-site growth factors from around 1.6 to 4.5. The three horizontal dashed lines show the 5th, 50th and 95th percentiles of the enhanced single-site growth factors. It is immediately clear that the range of growth factors output by the FEH statistical method is both very small and largely centred around the 50th percentile of enhanced single-site growth factors. A possible trend, where FEH statistical growth factors are smaller in larger catchments, is also apparent. It is not surprising that the range of FEH statistical growth factors is so low, as AREA is the main factor by which catchments are selected for pooling. As a result, all small catchments have similar pooling-groups, which change gradually with increasing catchment area. The plausibility of all these catchments having similar growth curves is questionable, therefore there is a need to investigate modifications to the FEH statistical method for small catchments, to avoid similar sites always being chosen for pooling regardless of the subtleties of each individual small catchment.

ReFH2 gives similar growth factors to the FEH statistical method, but over a slightly wider range, and ReFH1 gives similar growth factors to ReFH2, but with larger values in the top few percentiles. As both are rainfall-runoff models, a plot of rainfall growth factors versus runoff growth factors is given in Figure 30. It is noted that ReFH1 and ReFH2 rainfall growth factors differ in all catchments where $URBEXT_{1990}$ exceeds zero. This is because design storm duration (and therefore depth) is controlled by SAAR and time-to-peak, which is URBEXT-dependent in ReFH1 but not ReFH2.

Figure 30 shows generally close correspondence between rainfall growth factors and flood growth factors. It also shows that the outlying 100-year flood growth factors given by the original ReFH method do not result from similarly outlying rainfall growth factors. Each outlying site is initialised in ReFH with a very large available soil moisture capacity (C_{max} – C_{ini}) and therefore percentage runoff is very low regardless of return period. As a result, even an increase of only a few percentage points may equate to a doubling or tripling of the percentage runoff value which, coupled with an approximate tripling of the rainfall depth, leads to these outlier flood growth factors of 6.4 to 9.6. Significantly, BFIHOST exceeds 0.84 in all the outlying catchments. This is well above the limit of recommended applicability for the original ReFH method, but not for ReFH2. Figure 30 demonstrates that the unrealistic growth factors that may be generated if ReFH is used to estimate a series of QT values in permeable catchments do not occur in ReFH2.

Table 30 – Relative difference in Q100 estimated by five current and recent flood estimation methods in rural catchments.

Subset	All (133 sites)	SAAR ≥ 800 (108 sites)	SAAR < 800 (25 sites)	BFIHOST ≥ 0.65 (18 sites)	BFIHOST < 0.65 (115 sites)
ReFH	0.5672	0.6928	0.0245	0.3921	0.5946
ReFH2	0.2694	0.3042	0.1189	0.3969	0.2494
Stat. (0 donors)	0.5680	0.7123	-0.0550	0.4384	0.5883
Stat. (1 donor)	0.5196	0.6488	-0.0386	0.3549	0.5454
Stat. (6 donors)	0.5254	0.6619	-0.0640	0.3249	0.5568

The best performing method in each subset was identified as:

- ReFH2 for all sites.
- ReFH2 for SAAR ≥ 800.
- ReFH for SAAR < 800.
- Stat. (6 donors) for BFIHOST ≥ 0.65.
- ReFH2 for BFIHOST < 0.65.

Table 31 – ln-difference in Q100 estimated by five current and recent flood estimation methods in rural catchments.

Subset	All (133 sites)	SAAR ≥ 800 (108 sites)	SAAR < 800 (25 sites)	BFIHOST ≥ 0.65 (18 sites)	BFIHOST < 0.65 (115 sites)
ReFH	0.0979	0.1650	-0.1937	0.0388	0.1072
ReFH2	-0.0814	-0.0673	-0.1423	0.0417	-0.1006
Stat. (0 donors)	0.0921	0.1795	-0.2852	0.1044	0.0902
Stat. (1 donor)	0.0716	0.1530	-0.2802	0.0068	0.0817
Stat. (6 donors)	0.0600	0.1440	-0.3028	-0.0522	0.0776

The best performing method in each subset was identified as:

- Stat. (6 donors) for all sites.

- ReFH2 for SAAR ≥ 800.
- ReFH2 for SAAR < 800.
- Stat. (1 donor) for BFIHOST ≥ 0.65.
- Stat. (6 donors) for BFIHOST < 0.65.

Table 32 – RMSD in Q100 estimated by five current and recent flood estimation methods in rural catchments.

Subset	All (133 sites)	SAAR ≥ 800 (108 sites)	SAAR < 800 (25 sites)	BFIHOST ≥ 0.65 (18 sites)	BFIHOST < 0.65 (115 sites)
ReFH	0.6924	0.7019	0.6498	0.7807	0.6775
ReFH2	0.6503	0.6374	0.7036	0.7903	0.6256
Stat. (0 donors)	0.6762	0.6722	0.6933	0.7476	0.6644
Stat. (1 donor)	0.6773	0.6695	0.7101	0.7979	0.6564
Stat. (6 donors)	0.6804	0.6736	0.7093	0.8386	0.6522

The best performing method in each subset was identified as:

- ReFH2 for all sites.
- ReFH2 for SAAR ≥ 800.
- ReFH for SAAR < 800.
- Stat. (0 donors) for BFIHOST ≥ 0.65.
- ReFH2 for BFIHOST < 0.65.

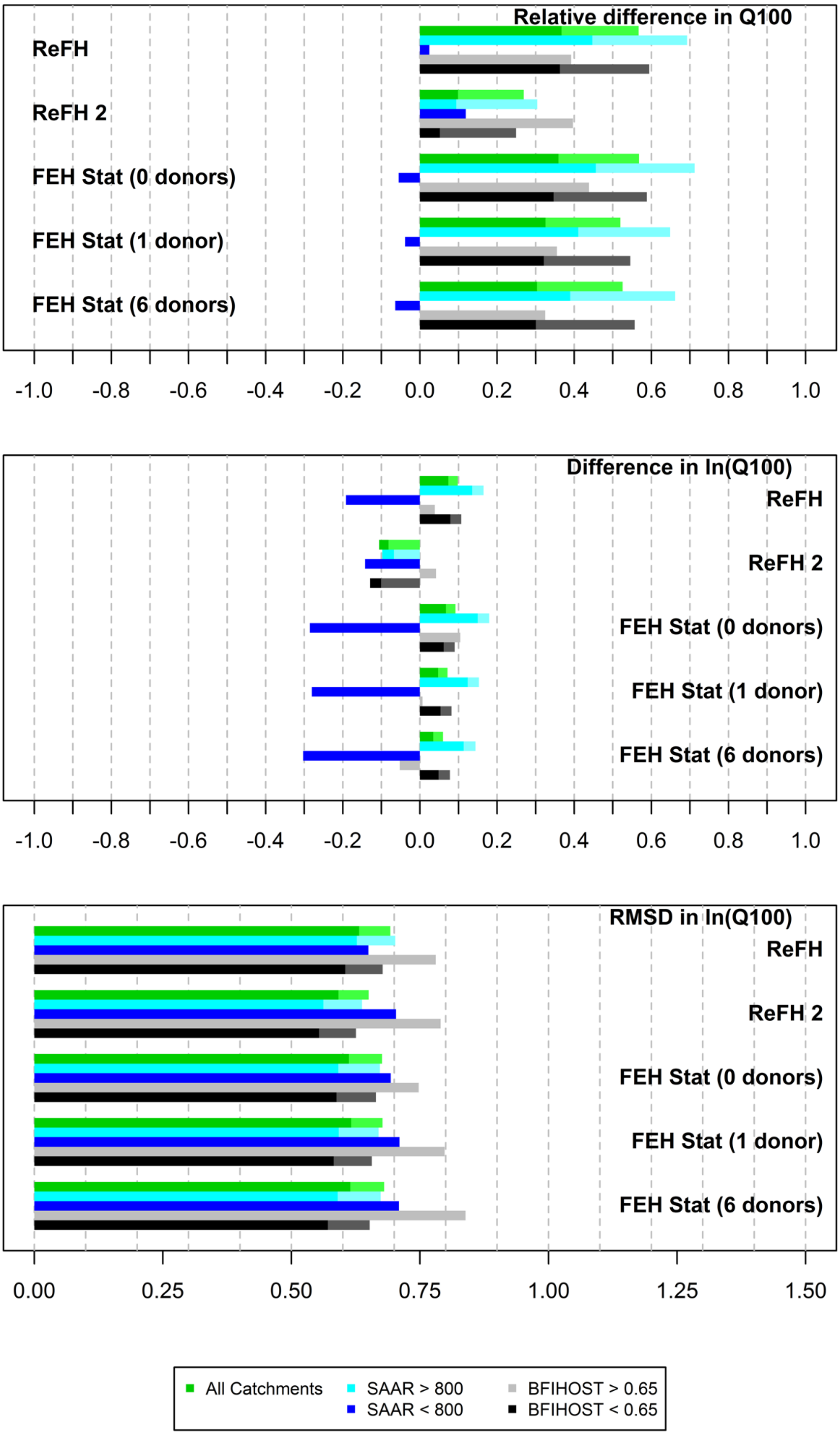

Figure 28 - Relative difference, ln-difference and RMSD in Q100 estimated by five current and recent flood estimation methods (rural catchments)

The three graphs in Figure 28 quantify the relative difference in Q100 (top graph), ln-difference in Q100 (middle graph) and RMSD in Q100 (bottom graph), as estimated by five methods in rural catchments:

- ReFH.
- ReFH2.
- FEH Stat (0 donors).
- FEH Stat (1 donor).
- FEH Stat (6 donors).

Results are shown for:

- All catchments (green bars).
- SAAR >800 (light blue bars).
- SAAR < 800 (dark blue bars).
- BFIHOST >0.65 (grey bars).
- BFIHOST <0.65 (black bars).

Darker bars show results excluding 205034.

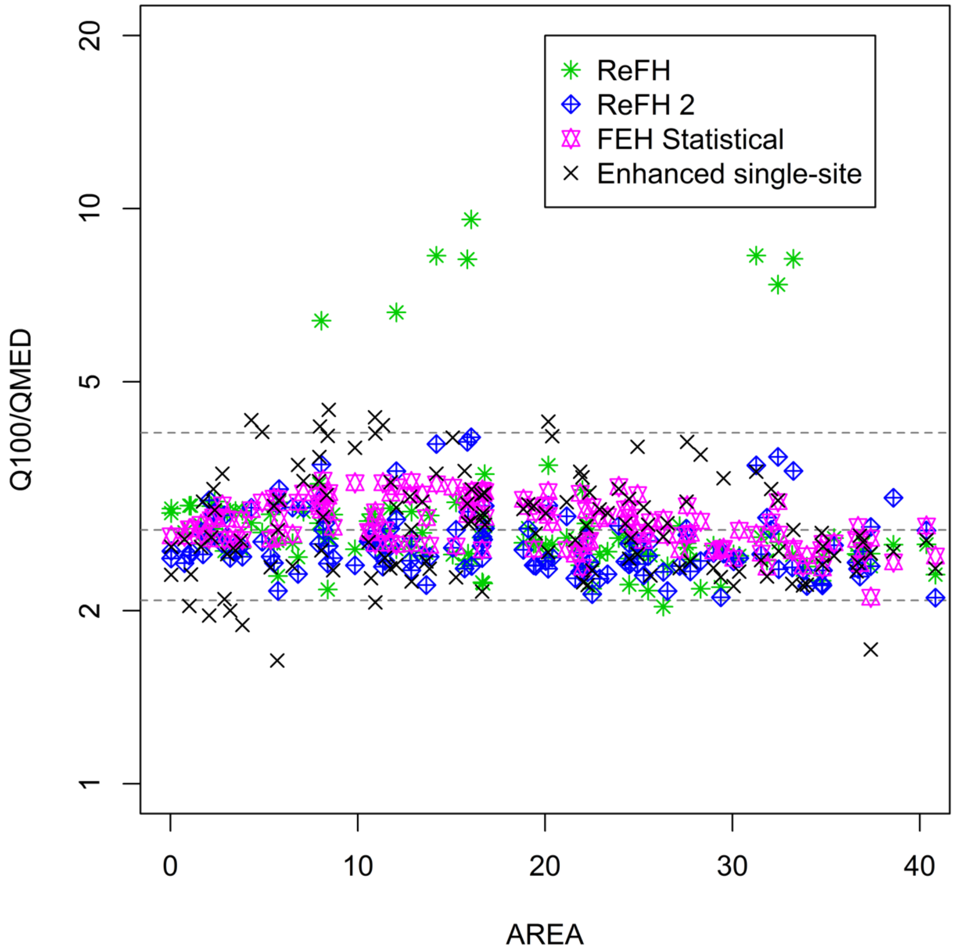

Figure 29 - Q100/QMED growth factor for 133 rural sites with FARL ≥ 0.9 and at-site AMAX records

Figure 29 is a scatter graph plotting Q100 over QMED (y-axis – from 1 to 20) against AREA (x-axis – from 0 to 40) for all 133 catchments for four methods:

- ReFH: green asterisks.
- ReFH2: blue diamonds.
- FEH statistical: pink stars.
- Enhanced single-site: black crosses.

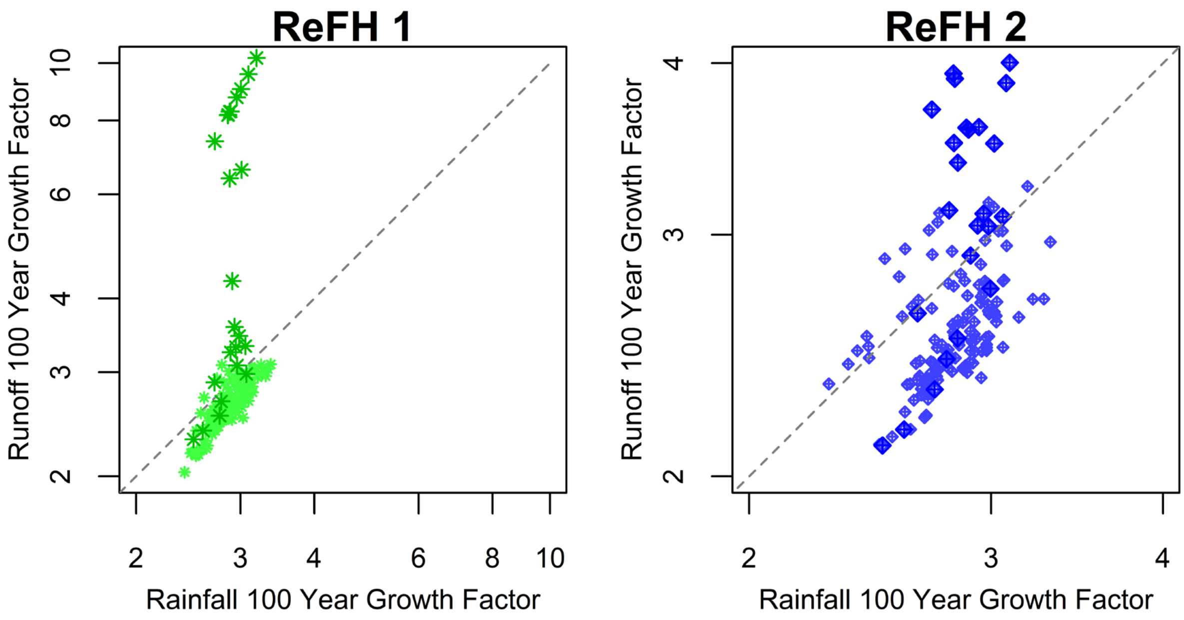

Figure 30 - Comparison of 100-year rainfall and flood growth factors for ReFH and ReFH2

Figure 30 plots 100-year rainfall and flood growth factors for ReFH and ReFH2 side-by-side:

- Left-hand graph (ReFH1): The y-axis plots runoff 100 year growth factor from 2 to 10. The x-axis plots rainfall 100 year growth factor from 2 to 10.
- Right-hand graph (ReFH2): The y-axis plots runoff 100 year growth factor from 2 to 4. The x-axis plots rainfall 100 year growth factor from 2 to 4.

Smaller, paler points show sites with BFIHOST < 0.65.

5.6 Summary of results

The set of 217 small, gauged catchments described in Section 3 was broken down into a ‘high-quality’ data set (43 catchments) and an ‘extended’ data set of lesser quality (133 catchments). Five methods of estimating QT were applied to the data sets and the results were compared, focusing particularly on three statistics: relative difference, ln-difference and root-mean-square difference (RMSD). The results for essentially rural and urbanised catchments were analysed separately. The five methods were:

- ReFH1
- ReFH2 (using FEH99 rainfall inputs, assuming winter seasonality with parameters set to default values)
- FEH statistical method with no donor transfer
- FEH statistical method with data transfer from one donor site
- FEH statistical method with data transfer from six donor sites

Only essentially rural catchments were considered, as the procedure for benchmark QT estimates (enhanced single-site method) is not valid for urbanised catchments. The main conclusions arising from the comparison of QT estimation methods are summarised below.

- The best performing method in terms of RMSD was found to be ReFH2 for return periods up to 500 years.
- In terms of ln-difference, ReFH2 underestimates on average while the FEH statistical method overestimates on average.
- ReFH2 is shown to improve greatly upon ReFH1. ReFH1 is shown to be unsuitable for estimating flood peaks rarer than about 150 years, to give very steep flood growth curves in certain circumstances and to be unsuitable in permeable catchments under any circumstance.
- For the FEH statistical method, ln-difference does not vary noticeably with return period. This means that, overall, QMED, Q30 and Q100 (for example) are all biased by approximately the same amount. Further, if one is unbiased, then all are.
- All methods overestimated Q30 and Q100 considerably in permeable catchments in the ‘high-quality’ data set (sample size 5) However, this was not true for the ‘extended’ data set (sample size 18).
- Most values of the error statistics considered were lower for the ‘extended’ data set than for the ‘high-quality’ data set. A possible explanation for this may be that the ‘extended’ data set includes more catchments that are similar to those used in the calibration of the FEH methods.
- Donor transfer is shown to reduce bias in estimated QT, except possibly in drier catchments. However, donor transfer does not decrease RMSD.
- The FEH statistical method generated similar flood growth curves for every small catchment, because area is the main criterion by which pooling-group members are selected. Alternative growth curve estimation methods for small catchments, such as estimating distribution parameters directly from catchment descriptors, should be considered.
- There is no definitive method to generate a benchmark flood peak estimate (except QMED) in catchments that are not essentially rural. It is therefore very difficult to assess the quality of a QT estimate in a non-rural catchment.

6. Conclusions and recommendations

A number of variants of current FEH methods together with the QMED equation proposed by MacDonald and Fraser (2013) have been applied to the data set of small, gauged catchments constructed during Task 1 of the current project. Details of the catchments divided into subsets described as ‘high-quality’ and ‘extended’ are provided in Section 3 and Appendix A of this report. The ‘extended’ data set of 217 catchments range in area from 0.04 to 40.9 km^2. Gauged QMED values are available at all 217 catchments, while gauged annual maximum series totalling 5,465 station-years are available at 192 catchments. Within this ‘extended’ data set, 153 catchments are defined as having ‘high-quality’ QMED estimates and 58 are defined as having ‘high-quality’ annual maximum records, using flags in the NRFA peak flow data set version 3.3.4 as guidance. In Section 4, these methods are applied to estimate QMED first on 146 catchments (those members of the data set with ‘high-quality’ QMED and FARL ≥ 0.9) and then on all 207 catchments with FAR*L* ≥ 0.9. In both cases, gauged QMED is used as a benchmark. In Section 5, the methods are applied to estimate longer return period floods from five to 1,000 years, first on 43 rural catchments with gauged annual maxima records considered suitable for pooling and FARL ≥ 0.9, then on 133 rural catchments with gauged annual maximum records and FARL ≥ 0.9. Enhanced single-site analyses are used to provide benchmark values.

The analyses described in Sections 4 and 5 result in several conclusions and recommendations. Some can be implemented immediately, while others require further investigation.

Firstly, the analysis has shown that the current FEH methods, that is, the FEH statistical method and ReFH2, outperform ReFH1 and the MacDonald and Fraser QMED equation in small catchments. For this reason, the two latter methods should no longer be used. In general, the performance of the FEH statistical method and ReFH2 has been found to be similar.

Secondly, ReFH2 and the FEH statistical methods give similar flood peak estimates in many situations. General agreement between the methods has a beneficial effect in reducing uncertainty. In circumstances where the two methods give dissimilar flood peak estimates, the following recommendations based on the observed performance of each method in different catchment subsets, apply on which result should normally be prioritised.

It is recommended that ReFH2 should be prioritised in estimating QMED in small rural catchments ($URBEXT_{2000}$ < 0.03). This is because it gives the most consistent errors relative to high and low SAAR and BFIHOST values, and the lowest root-mean-square error over both the full data set of 146 rural small catchments and the higher-quality data set of 103. It is a consistently strong performer in subsets where it is not the absolute best method. The ReFH2 design method has been shown to improve greatly on its predecessor, ReFH1.

The FEH statistical method should be prioritised for estimation of QMED in drier urban catchments ($URBEXT_{2000} \geq 0.03$, SAAR < 800). However, ReFH2 performs almost equally well and is more strongly recommended for wetter urban catchments (SAAR ≥ 800). The FEH statistical method tends to overestimate QMED by larger and larger amounts as SAAR increases beyond 1,200.

Error in design flood estimates was not found to be related to catchment area in the data set. Some of the largest positive errors in estimated QMED were found for catchments of around 20 km^2 in area.

The results suggest that donor transfer continues to serve a useful purpose in the FEH statistical method for small catchments, although using six donors does not always reduce error compared to using a single donor. Further work should consider the relative importance of proximity compared to other hydrological properties (as indicated by catchment descriptors) in selecting donors.

The analysis of essentially rural catchments within the ‘extended’ data set produced slightly surprising results in that most values of the error statistics considered were lower than for the ‘high-quality’ data set. This suggests that there are benefits to be gained from using the ‘extended’ data set in the next stages of the project.

The growth factors generated by ungauged FEH statistical pooling-groups are, on average, shown to closely match those given by the enhanced single-site method, for both the full and ‘high-quality’ data sets of rural catchments. However, as area is the main factor used in selecting pooling-groups for ungauged catchments, many of the same catchments appear in many pooling-groups and the overall range of 100-year FEH statistical growth factors is small. Further investigation should consider if any alternative criteria are more important than area for selecting pooling-groups in ungauged small catchments.

There are several situations noted where the tested flood estimation methods tend to overestimate QMED and rarer flood peaks. This probably reflects differences in the flood-generating mechanisms in the permeable and impermeable parts of small catchments. There is a general overestimation of QMED and rarer flood peaks in high-BFIHOST catchments, across all methods. This may be because the methods possess an unexplained dependency on baseflow specific to small catchments, or it may be because BFIHOST does not accurately measure what it is intended to in certain small, high-BFIHOST catchments. It should be noted that the digitised HOST data from which BFIHOST is derived exist at 1-km resolution and express the proportional coverage of each HOST class per 1 km^2 cell. Therefore, the actual HOST class coverage of some small catchments may not match with the average HOST class coverage recorded for the cells and part-cells lying within those catchment boundaries. It may therefore be beneficial to incorporate alternative measures of baseflow into the tested flood estimation methods or to extend the baseflow estimation by testing and (where successful) incorporating additional descriptors that now exist, such as elevation, urban/suburban cover and different types of vegetation cover. The FEH Local project has investigated the possibility of generating a high-resolution replacement for BFIHOST but has concluded that it is unlikely to be possible to do this for the entire UK with currently available soil data.

Furthermore, all methods tend to overestimate both QMED and rarer flood peaks in catchments where the observed values are small (< 3 m^3/s – see Figures 4.1 and 5.1). This could reflect the effect of small errors in IHDTM catchment boundary definition leading to errors in estimating catchment area for very small catchments. Another possible explanation lies in the possibility that surface and groundwater catchments may not coincide for small drainage areas. In the longer term, a higher resolution DTM may become available for hydrological applications as currently being investigated as part of the FEH Local project. This would be expected to define very small catchments much more accurately than the current 50 m DTM.

There may be a possible overdependence on SAAR in the FEH statistical method (for example, Figure 9), resulting in overestimating QMED and QT in high-SAAR catchments. This could reflect the influence of some land-use class more commonly found in specifically small, high-SAAR catchments that is not currently accounted for. Although research in the current project could not identify a correlation between forest cover and QMED, other research (Wan Jaafar and Han, 2012) has identified a potential relationship between QMED and agricultural land-use, which may be worth investigating. This issue lies outside the scope of the current project.

Contrary to some perceptions, the ratio of areas between the donor and catchment of interest appears to be of little importance on a broad scale (Figure 12). There usually appears to be a greater change in performance in moving from zero to one donors than from one to six donors. Donors appear to be most beneficial in low-BFIHOST catchments (which form most of this data set). Further work beyond this project could investigate individual catchments and their circumstances in detail so that situations in which using one or more donors is particularly beneficial can be identified.

In this study, only winter storms and initial conditions were used with ReFH2. This decision was based on a brief investigation of error in QMED as the $URBEXT_{2000}$ value signifying the boundary between winter and summer conditions was tested at various values. However, Figure 16 shows a clear trend towards underestimation by ReFH2 (and ReFH1) as $URBEXT_{2000}$ becomes large (although excluding catchments with $URBEXT_{2000} \geq 0.6$ almost eliminates this trend). This suggests that there is scope for exploring whether using summer design storms should be recommended for heavily urbanised catchments. The validity of the assumption of an impervious fraction of 0.3 within ReFH2 for urban catchments should also be considered. For all methods other than ReFH2, Figure 16 shows a significant trend between $URBEXT_{2000}$ and ln-error. Therefore, further calibration of the urban adjustments in the FEH statistical method could be justified by this. Figure 16 underlines that MacDonald and Fraser’s method should not be used in heavily urbanised catchments.

According to the mechanism for dividing total rainfall into urban and rural fractions in ReFH2, all rainfall is routed to the urban sub-model once $URBEXT_{2000}$ exceeds 0.64 (although urban area can be manually overwritten in the event modelling screen). ReFH2 therefore does not by default differentiate between otherwise identical catchments with differing levels of extreme urbanisation. Formalising the method for $URBEXT_{2000}$ values above 0.64 represents one opportunity for improving ReFH2 and its guidance.

The analysis of urban catchments was hampered by the fact that there is no standard procedure for performing an enhanced single-site analysis if the catchment of interest is even slightly urbanised. This means that it was impossible to generate the same standard of benchmark QT estimates for catchments with $URBEXT_{2000} \geq 0.03$ as for catchments with $URBEXT_{2000} < 0.03$. This underlines the fact that generalised methods of flood frequency estimation are not recommended for using in urban catchments.

References

AGRICULTURAL DEVELOPMENT AND ADVISORY SERVICES (ADAS), 1982. 'The Design of Field Drainage Pipe Systems' (ADAS Reference Book 345). London: MAFF Agricultural Development and Advisory Services Land and Water Service.

AHMAD, M.I., SINCLAIR, C.D. AND WERRITTY, A., 1988. 'Log-logistic flood frequency analysis' Journal of Hydrology, 98 (3-4), 205 to 224.

BAYLISS, A.C. AND JONES, R.C., 1993. 'Peaks-over-threshold flood database: Summary statistics and seasonality (IH report 121)' Wallingford: Institute of Hydrology.

BAYLISS, A., 1999. 'Catchment Descriptors (Flood Estimation Handbook volume 5)' Wallingford: Institute of Hydrology.

BOORMAN, D.B., HOLLIS, J.M. AND LILLY, A., 1993. 'Hydrology of soil types: a hydrologically-based classification of the soils of the United Kingdom (IH report 126)' Wallingford: Institute of Hydrology.

DALRYMPLE, T., 1960. 'Flood-Frequency Analyses, Manual of Hydrology: Part 3. US Geological Survey Water-Supply Paper 1543-A' Washington, DC: United States Government Printing Office.

ENVIRONMENT AGENCY, 2022. 'Flood estimation guidelines' LIT 11832. Bristol: Environment Agency'.

FAULKNER, D., 1999. 'Rainfall frequency estimation (Flood Estimation Handbook volume 2)' Wallingford: Institute of Hydrology.

FAULKNER, D. AND BARBER, S., 2009. 'Performance of the Revitalised Flood Hydrograph method' Journal of Flood Risk Management, 2 (4), 254 to 261.

FAULKNER, D., KJELDSEN, T., PACKMAN, J. AND STEWART, L., 2012. 'Estimating flood peaks and hydrographs for small catchments: Phase 1' Bristol: Environment Agency.

HAXTON, T., VESUVIANO, G., STEWART, L. AND YOUNG, A., 2014. 'Evaluation of ReFH2 for plot-scale estimation of QMED' Project note (SC090031). Wallingford: Wallingford HydroSolutions.

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

HOUGHTON-CARR, H., 1999. 'Restatement and application of the Flood Studies Report rainfall-runoff method (Flood Estimation Handbook volume 4)' Wallingford: Institute of Hydrology.

KJELDSEN, T.R., 2007. 'Flood Estimation Handbook Supplementary Report No. 1. The revitalised FSR/FEH rainfall-runoff method' Wallingford: Centre for Ecology & Hydrology

KJELDSEN, T.R., 2010. ‘Modelling the impact of urbanization on flood frequency relationships in the UK’ Hydrology Research, 41 (5), 391 to 405.

KJELDSEN, T.R., 2015. ‘How reliable are design flood estimates in the UK?’ Journal of Flood Risk Management, 8 (3), 237 to 246.

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

KJELDSEN, T.R. AND PROSDOCIMI, I., 2015. ‘A bivariate extension of the Hosking and Wallis goodness-of-fit measure for regional distributions’ Water Resources Research 51 (2), 896 to 907.

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., MILLER, J.D. AND PACKMAN, J.C., 2013. ‘Modelling design flood hydrographs in catchments with mixed urban and rural land cover’ Hydrology Research, 44 (6), 1040 to 1057.

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.

MARSHALL, D.C.W. AND BAYLISS, A.C., 1994. ‘Flood estimation for small catchments (IH report 124)’ Wallingford: Institute of Hydrology.

NERC, 1975. ‘Flood Studies Report (5 volumes)’ London: Natural Environment Research Council.

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 17 August 2020].

NRFA, 2015. ‘NRFA Station Peak Flow Data’ 26802 - Gypsey Race at Kirby Grindalythe.

PACKMAN, J.C., 1980. ‘The effects of urbanisation on flood magnitude and frequency (IH report 63)’ Wallingford: Institute of Hydrology.

REED, D., 1999. ‘Overview (Flood Estimation Handbook volume 1)’ Wallingford: Institute of Hydrology.

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

SIDDAWAY, A. AND Faulkner, D., 2015. ‘Investigations of plot-scale runoff data Project note (SC090031)’ Skipton: JBA Consulting.

STEWART, E.J., JONES, D.A., SVENSSON, C., MORRIS, D.G., DEMPSEY, P., DENT, J.E., COLLIER, C.G. AND ANDERSON, C.W., 2013. ‘Reservoir Safety – Long Return Period Rainfall’ Bristol: Environment Agency.

WALLINGFORD HYDROSOLUTIONS, 2015. ‘The Revitalised Flood Hydrograph Model ReFH 2: Technical Guidance’ Wallingford: Wallingford HydroSolutions.

WAN JAAFAR, W.Z. AND HAN, D., 2012. ‘Catchment characteristics for index flood regionalisation’ Proceedings of the ICE – Water Management, 165 (3), 179 to 189.

List of abbreviations

ADAS Agricultural Development and Advisory Services –

part of the former Ministry of Agriculture Fisheries and Food. ADAS 345 refers to ADAS Reference Book 345 (ADAS, 1982).

AEP Annual exceedance probability.

The chance that a flood of a given magnitude or larger will occur in any water year, normally expressed as a percentage.

AMAX Annual maximum/annual maxima.

A flood peak record that features only the largest event within each water year, with no lower limit on the magnitude of any event. Water years run from 1 October to the following 30 September.

AREA AREA (in capitals) refers to the FEH digital catchment descriptor.

This measures the area of a catchment according to the Integrated Hydrological Digital Terrain Model (IHDTM). All other catchment descriptors are based on the shape and boundary of the catchment as represented in the IHDTM.

BFIHOST An FEH catchment descriptor, which estimates a catchment's baseflow index from the HOST classifications of the soils present.

CD3 A file format for storing FEH digital catchment descriptors, used by WINFAP-FEH. Can be exported from the FEH Web Service and FEH CD-ROM version 3.

DTM Digital terrain model. The DTM in this report refers to the Integrated Hydrologic Digital Terrain Model (IHDTM). The IHDTM enforces drainage direction and flow paths according to rivers mapped by Ordnance Survey and Land & Property Services.

FARL An FEH catchment descriptor measuring flood attenuation by reservoirs and lakes.

Further information is given by Bayliss (1999).

FEH Flood Estimation Handbook.

FEH13 Flood Estimation Handbook 2013.

In this context, referring to the latest and current recommended model for design rainfall depths, published in the joint Defra/Environment Agency report 'Reservoir Safety – Long Return Period Rainfall' (Stewart and others, 2013).

FEH99 Flood Estimation Handbook 1999.

In this context, referring to the 6-parameter model for design rainfall depths, published in volume 2 of the FEH (Faulkner, 1999).

FPEXT An FEH catchment descriptor measuring the fraction of a catchment that is estimated to be inundated by the 100-year flood. Further information is given by Kjeldsen and others (2008).

FSR Flood Studies Report.

GLO Generalised logistic distribution.

The default distribution for AMAX series in the UK. A re-parameterised version of the log-logistic distribution (Ahmad and others, 1988).

HOST Hydrology of soil types.

A hydrologically based classification and mapping of soils in the UK, in which every soil type is assigned to one of 29 classes. See Institute of Hydrology report 126 (Boorman and others, 1995) for further information.

IH Institute of Hydrology.

MF MacDonald and Fraser, referring to the QMED estimation method detailed ‘An improved method for estimating the median annual flood for small ungauged catchments in the UK’ (MacDonald and Fraser, 2013).

NRFA National River Flow Archive.

POT Peaks-over-threshold.

A flood peak record that features every independent flow maximum above a specified lower limit.

PROPWET

An FEH digital catchment descriptor.

Measures the approximate fraction of time that soil moisture deficit within a catchment is less than 6 mm. Derivation of this descriptor is complex and described by Bayliss (1999).

QMED Median annual flood, equivalent to the 2-year flood and the flood with AEP 50%. This is the index flood for the FEH statistical method.

*Q*T (*Q*5, *Q*30, *Q*100, *Q*200)

T-year flood, equivalent to the flood with AEP (100/T) %.

ReFH1 The original revitalised flood hydrograph method.

A rainfall-runoff flood estimation method, now superseded by ReFH2. Described in Section 2.2 of this report and in detail by Kjeldsen (2007).

ReFH2 The second version of the revitalised flood hydrograph method.

The current rainfall-runoff flood estimation method in the UK. Described in Section 2.2 and in detail by Kjeldsen and others (2013) and Wallingford HydroSolutions (2015).

RMSD Root-mean-square difference.

An error statistic that is functionally identical to root-mean-square error (Equation 7). The term 'difference' is substituted for 'error' in comparing the QT estimation methods, as the benchmark values of QT are themselves estimates, therefore the tested methods cannot be said definitively to be in error if they simply do not match the benchmark values.

RMSE Root-mean-square error.

SAAR Standard-period annual average rainfall.

An FEH catchment descriptor quantifying the mean annual total rainfall that fell in a catchment over the period 1961 to 1990.

UAF Urban adjustment factor.

A quantity, calculated from $URBEXT_{2000}$ and BFIHOST, which is applied to an ungauged FEH statistical estimate of QMED. Its purpose is to increase the QMED estimate with urbanisation.

URBEXT ($URBEXT_{1990}$, $URBEXT_{2000}$)

Urban extent.

An FEH catchment descriptor that measures the fraction of urban development within a catchment. $URBEXT_{1990}$ is based upon urban areas identified in the Land Cover Map 1990 while $URBEXT_{2000}$ is based upon the Land Cover Map 2000. Both methods differentiate between urban and suburban land use, assuming that only half of a suburban area is developed. $URBEXT_{2000}$ also includes inland bare ground, assuming that it is between suburban and urban areas in terms of permeability.

WHS Wallingford HydroSolutions.

Developer, publisher and retailer of ReFH2 software, WINFAP-FEH, FEH Web Service and, previously, ReFH1 software.

WINFAP Software to perform the FEH statistical method.

Appendix A

Table 33 - all 217 stations in the study data set (Section 3)

Station No.	Station	Watercourse	Record (years)	QMED (m^3/s)	Easting	Northing	AREA (km^2)	BFI-HOST	SAAR (mm)	$URBEXT_{2000}$	FARL	FPEXT	DPLBAR (km)	DPSBAR (m/km)	PROP-WET	Pooling	QMED	Task 1	MacD & Fraser
4009	Peffery	Strathpeffer STW	19	5.43	249250	858650	15.89	0.487	1060	0.0085	0.973	0.0378	3.66	152.4	0.74	NO	YES	YES	NO
4063	Pyl Brook	West Barnes Lane (Kings College)	11	6.50	523200	167950	17.34	0.556	652	0.5423	0.995	0.0873	5.20	28.9	0.29	NO	YES	NO	YES
4186	Graveney	Stretham (Abercarne Road)	11	7.00	529250	170300	12.14	0.332	648	0.7663	1.000	0.1336	4.53	31.4	0.29	NO	YES	NO	YES
7009	Mosset Burn	Wardend Bridge	16	9.14	303950	855850	28.30	0.606	803	0.0000	0.998	0.0601	8.03	61.3	0.42	NO	YES	YES	NO
13017	Colliston Burn	Colliston	21	3.37	360900	746650	8.40	0.546	750	0.0000	0.996	0.0678	2.72	35.9	0.36	NO	YES	YES	YES
15002	Newton Burn	Newton	24	6.91	323000	760550	16.65	0.461	1200	0.0000	1.000	0.0153	7.67	198.5	0.68	NO	YES	NO	YES
15004	Inzion	Loch of Lintrathen	44	6.41	327950	755850	24.49	0.529	1081	0.0000	0.997	0.0292	6.49	187.4	0.53	NO	YES	NO	YES
15027	Garry Burn	Loakmill	16	7.54	307450	733950	22.57	0.574	947	0.0116	0.999	0.0591	5.92	110.9	0.46	NO	YES	NO	YES
15809	Muckle Burn	Eastmill	20	7.63	322300	760450	16.69	0.481	1132	0.0000	0.960	0.0196	5.75	160.8	0.68	NO	YES	NO	YES
17012	Red Burn	Castle Cary	20	20.10	278800	678050	21.74	0.329	1172	0.1392	0.995	0.1097	5.12	55.4	0.58	NO	YES	NO	YES
18020	Loch Ard Burn	Duchray	12	1.35	246800	698700	0.86	0.609	2000	0.0000	1.000	0.0058	1.12	155.8	0.74	NO	YES	NO	YES
19010	Braid Burn	Liberton	38	3.81	327250	670750	15.39	0.514	770	0.1586	0.947	0.0326	6.20	113.8	0.49	NO	YES	YES	YES
20002	West Peffer Burn	Luffness	49	3.53	348850	681150	26.31	0.471	616	0.0023	0.996	0.1279	5.62	30.4	0.33	YES	YES	YES	NO
21001	Fruid Water	Fruid	15	19.10	308850	620700	22.17	0.392	1699	0.0000	0.780	0.0113	5.36	221.2	0.72	NO	YES	YES	YES
21017	Ettrick Water	Brockhoperig	49	69.47	323400	613150	38.59	0.421	1740	0.0000	1.000	0.0120	5.88	241.5	0.72	YES	YES	YES	NO

Station No.	Station	Watercourse	Record (years)	QMED (m^3/s)	Easting	Northing	AREA (km^2)	BFI-HOST	SAAR (mm)	$URBEXT_{2000}$	FARL	FPEXT	DPLBAR (km)	DPSBAR (m/km)	PROP-WET	Pooling	QMED	Task 1	MacD & Fraser
22003	Usway Burn	Shillmoor	28	19.22	388650	607750	21.87	0.302	1056	0.0000	1.000	0.0061	9.28	205.0	0.45	NO	YES	YES	YES
23018	Ouse Burn	Woolsington	31	2.56	419550	570000	10.48	0.312	669	0.0975	0.978	0.1296	3.05	30.4	0.45	YES	YES	YES	YES
24006	Rookhope Burn	Eastgate	20	24.62	395250	539050	36.62	0.293	1126	0.0001	0.994	0.0177	7.24	119.4	0.59	YES	YES	YES	NO
25003	Trout Beck	Moor House	41	15.16	375900	533550	11.46	0.227	1904	0.0000	1.000	0.0412	3.41	91.9	0.64	YES	YES	YES	YES
25011	Langdon Beck	Langdon	28	15.88	385200	530850	12.79	0.237	1463	0.0011	1.000	0.0125	4.24	123.4	0.59	NO	YES	YES	YES
25012	Harwood Beck	Harwood	45	33.27	384950	530900	24.58	0.261	1577	0.0000	1.000	0.0212	5.57	121.0	0.59	YES	YES	YES	YES
25019	Leven	Easby	36	5.54	458500	508650	15.07	0.525	830	0.0043	1.000	0.0194	5.53	128.0	0.38	YES	YES	YES	YES
25808	Burnt Hill	Moor House	8	0.08	374600	523150	0.75	0.294	1499	0.0000	1.000	0.0000	0.88	116.9	0.64	NO	YES	YES	NO
25809	Bog Hill	Moor House	9	0.07	377250	532650	0.05	0.228	1757	0.0000	1.000	0.0000	0.17	100.4	0.64	NO	YES	YES	NO
25810	Sike Hill	Moor House	6	0.09	377200	533200	0.04	0.275	1757	0.0000	1.000	0.0000	0.18	79.9	0.64	NO	YES	YES	NO
26555	Ings Beck	South Newbald	15	0.54	490950	436100	14.20	0.985	704	0.0094	1.000	0.0083	3.99	73.7	0.26	NO	YES	YES	NO
26802	Gypsey Race	Kirby Grindalythe	15	0.11	490400	467450	15.85	0.959	757	0.0000	1.000	0.0305	3.85	57.2	0.32	YES	YES	YES	YES
26803	Water Forlornes	Driffield	15	0.44	502250	458250	32.43	0.949	721	0.0074	1.000	0.0159	7.85	65.4	0.31	YES	YES	YES	NO
27010	Hodge Beck	Bransdale Weir	41	9.42	462750	494400	18.84	0.341	987	0.0007	1.000	0.0094	5.25	149.8	0.40	YES	YES	YES	YES
27032	Hebden Beck	Hebden	48	3.92	402550	464350	22.20	0.252	1433	0.0000	0.997	0.0207	6.16	99.1	0.62	YES	YES	YES	NO
27038	Costa Beck	Gatehouses	42	1.38	477600	483800	7.98	0.774	722	0.0220	0.990	0.1253	3.12	36.0	0.40	NO	YES	NO	YES
27047	Snaizeholme Beck	Low Houses	34	13.22	383300	488250	10.93	0.304	1733	0.0000	0.977	0.0208	3.83	204.2	0.62	NO	YES	NO	YES
27051	Crimple	Burn Bridge	42	4.54	428350	451900	8.15	0.309	855	0.0058	1.000	0.0133	2.54	62.9	0.34	YES	YES	YES	YES
27073	Brompton Beck	Snainton Ings	34	0.81	493550	479450	8.06	0.887	721	0.0079	1.000	0.2373	4.53	47.7	0.39	YES	YES	YES	YES

Station No.	Station	Watercourse	Record (years)	QMED (m^3/s)	Easting	Northing	AREA (km^2)	BFI-HOST	SAAR (mm)	$URBEXT_{2000}$	FARL	FPEXT	DPLBAR (km)	DPSBAR (m/km)	PROP-WET	Pooling	QMED	Task 1	MacD & Fraser
27081	Oulton Beck	Farrer Lane	28	2.36	436450	428100	25.10	0.535	677	0.2235	0.997	0.0486	6.86	40.4	0.32	YES	YES	YES	YES
27082	Cundall Beck	Bat Bridge	17	7.02	441950	472350	22.96	0.654	635	0.0175	0.999	0.2062	5.36	14.8	0.34	NO	YES	NO	YES
27852	Little Don	Langsett Reservoir	22	19.27	421300	400400	21.13	0.320	1316	0.0013	0.846	0.0074	4.54	122.3	0.43	NO	YES	NO	YES
28033	Dove	Hollinsclough	35	4.67	406350	366850	7.93	0.403	1346	0.0000	1.000	0.0075	3.18	166.9	0.52	YES	YES	YES	YES
28041	Hamps	Waterhouses	29	26.66	408200	350250	36.97	0.301	1085	0.0041	1.000	0.0326	8.45	86.3	0.44	YES	YES	YES	NO
28070	Burbage Brook	Burbage	57	4.29	425950	380350	8.45	0.426	1006	0.0000	1.000	0.0310	2.63	85.3	0.38	NO	YES	YES	YES
28115	Maun	Mansfield the Dykes	22	12.40	455950	363600	30.56	0.841	714	0.3886	0.915	0.0539	7.11	43.4	0.36	NO	YES	YES	NO
29009	Ancholme	Toft Newton	40	1.83	503250	387650	29.52	0.625	616	0.0044	0.997	0.2063	5.60	12.2	0.26	NO	YES	YES	NO
29013	Moor Beck	Clapgate Farm	29	0.22	496650	411100	7.10	0.797	637	0.0798	0.984	0.0863	3.19	25.0	0.26	NO	YES	YES	NO
30013	Heighington Beck	Heighington	38	0.61	504150	369600	24.03	0.945	605	0.0790	0.963	0.1200	6.35	19.0	0.26	YES	YES	YES	YES
30014	Pointon Lode	Pointon	42	2.61	512850	331250	10.94	0.338	591	0.0143	1.000	0.1046	6.03	29.0	0.22	NO	YES	YES	YES
31023	West Glen	Easton Wood	42	1.88	496550	325800	4.32	0.320	641	0.0000	1.000	0.0516	1.87	32.8	0.27	NO	YES	YES	YES
31025	Gwash South Arm	Manton	36	10.21	487550	305150	23.93	0.306	663	0.0064	0.995	0.0266	7.27	61.1	0.30	NO	YES	YES	YES
31026	Egleton Brook	Egleton	36	1.14	487850	307350	2.30	0.533	645	0.0111	1.000	0.0936	1.75	41.0	0.30	NO	YES	YES	YES
32029	Flore	Experimental Catchment	5	2.54	465550	260450	8.34	0.430	624	0.0016	1.000	0.0861	2.50	38.9	0.30	NO	YES	YES	YES
33030	Clipstone Brook	Clipstone	7	10.94	493250	225550	40.35	0.362	640	0.0156	0.975	0.0824	6.85	34.0	0.31	NO	YES	YES	NO
33045	Wittle	Quidenham	46	1.16	602650	287750	27.55	0.534	608	0.0104	0.974	0.1771	4.77	15.1	0.31	NO	YES	YES	NO
33048	Larling Brook	Stonebridge	32	0.30	592750	290650	21.99	0.694	635	0.0033	0.907	0.2331	5.44	8.8	0.24	NO	YES	YES	NO
33052	Swaffham Lode	Swaffham Bulbeck	45	0.38	555300	262850	33.25	0.841	567	0.0121	0.998	0.2017	6.82	25.6	0.26	NO	YES	YES	NO

Station No.	Station	Watercourse	Record (years)	QMED (m^3/s)	Easting	Northing	AREA (km^2)	BFI-HOST	SAAR (mm)	$URBEXT_{2000}$	FARL	FPEXT	DPLBAR (km)	DPSBAR (m/km)	PROP-WET	Pooling	QMED	Task 1	MacD & Fraser
33064	Whaddon Brook	Whaddon	15	0.25	535900	246650	14.56	0.943	558	0.1171	0.997	0.1492	4.98	24.1	0.24	NO	YES	NO	YES
33065	Hiz	Hitchin	21	0.35	518500	228950	12.00	0.968	621	0.0342	1.000	0.0475	3.71	58.0	0.30	NO	YES	YES	NO
33813	Mel	Meldreth	17	0.55	537800	246550	4.86	0.886	552	0.0346	0.996	0.1743	3.39	21.6	0.24	NO	YES	NO	YES
34051	Spixworth Beck	Spixworth	13	2.47	623750	316450	21.37	0.744	622	0.0390	0.996	0.1315	5.25	14.8	0.29	NO	YES	YES	NO
36009	Brett	Cockfield	44	4.00	591400	252450	25.62	0.395	598	0.0052	1.000	0.1129	5.57	18.5	0.28	NO	YES	YES	NO
36010	Bumpstead Brook	Broad Green	47	6.83	568950	241800	27.58	0.387	588	0.0075	0.999	0.0447	4.61	34.1	0.27	YES	YES	YES	NO
36011	Stour Brook	Sturmer	47	6.38	569650	244050	34.28	0.382	592	0.1003	0.999	0.0603	6.65	33.5	0.26	YES	YES	YES	NO
37033	Eastwood Brook	Eastwood	39	5.16	585850	188850	9.85	0.340	555	0.4113	0.995	0.0797	3.73	29.6	0.21	NO	YES	YES	YES
37052	Prittlewell Brook	Prittlewell	15	1.24	583300	187000	6.36	0.380	559	0.3264	1.000	0.0543	2.97	28.4	0.22	NO	YES	YES	NO
38007	Canons Brook	Elizabeth Way	64	7.00	543150	210350	20.74	0.352	601	0.2483	0.988	0.0520	4.28	29.3	0.31	NO	YES	YES	YES
38012	Stevenage Brook	Bragbury Park	41	2.71	527450	221150	35.11	0.663	634	0.2806	0.968	0.0639	5.41	31.7	0.30	NO	YES	YES	NO
38014	Salmon Brook	Edmonton	53	5.52	534350	193750	22.85	0.258	665	0.2926	0.978	0.0562	6.85	48.3	0.29	NO	YES	NO	YES
38020	Cobbins Brook	Sewardstone Road	43	7.60	538650	199950	38.87	0.223	616	0.0517	0.997	0.0582	7.80	44.4	0.29	YES	YES	YES	NO
39005	Beverley Brook	Wimbledon Common	50	11.14	521650	171750	39.49	0.476	630	0.4992	0.994	0.1379	7.12	26.6	0.29	YES	YES	NO	NO
39017	Ray	Grendon Underwood	50	4.99	468050	221100	21.15	0.238	622	0.0037	0.982	0.1584	4.55	28.0	0.32	NO	YES	YES	YES
39036	Law Brook	Albury	17	0.55	504550	146750	16.05	0.888	819	0.0084	0.960	0.0173	4.87	85.7	0.36	NO	YES	NO	YES
39049	Silk Stream	Colindeep Lane	36	13.65	521700	189550	30.76	0.182	685	0.4014	0.972	0.0848	5.28	40.1	0.29	NO	YES	YES	NO
39054	Mole	Gatwick Airport	53	10.10	526000	139800	32.33	0.437	816	0.1399	0.943	0.1719	6.60	33.5	0.36	NO	YES	YES	NO

Station No.	Station	Watercourse	Record (years)	QMED (m^3/s)	Easting	Northing	AREA (km^2)	BFI-HOST	SAAR (mm)	$URBEXT_{2000}$	FARL	FPEXT	DPLBAR (km)	DPSBAR (m/km)	PROP-WET	Pooling	QMED	Task 1	MacD & Fraser
39055	Yeading Brook West	North Hillingdon	20	4.24	508500	184500	16.82	0.172	657	0.5347	0.999	0.2032	6.47	16.1	0.29	YES	YES	YES	YES
39082	Graveney	Longley Road	21	12.50	527850	170550	4.93	0.480	630	0.8110	1.000	0.2299	3.91	13.8	0.29	NO	YES	YES	NO
39086	Gatwick Stream	Gatwick Link	39	9.75	528500	141750	32.63	0.597	830	0.1745	0.946	0.1036	8.99	47.9	0.36	NO	YES	YES	NO
39092	Dollis Brook	Hendon Lane Bridge	57	7.36	524050	189450	23.72	0.178	689	0.3444	0.991	0.0466	6.31	49.4	0.29	NO	YES	YES	YES
39095	Quaggy	Manor House Gardens	51	5.16	539450	174800	33.50	0.610	644	0.4780	0.997	0.0833	7.40	36.7	0.27	YES	YES	YES	NO
39096	Wealdstone Brook	Wembley	38	11.95	519250	186250	23.45	0.175	664	0.5080	0.997	0.1333	4.85	25.7	0.29	NO	YES	YES	YES
39116	Sulham Brook	Sulham	16	0.70	464200	174050	3.03	0.408	657	0.0017	1.000	0.2409	1.51	46.0	0.29	NO	YES	NO	YES
39126	Red	Redbourne	22	0.56	510450	211750	22.31	0.643	702	0.0909	0.993	0.0576	4.65	30.1	0.30	NO	YES	YES	NO
39134	Ravensbourne East	Bromley South	21	4.85	540550	168650	9.80	0.685	680	0.4869	0.993	0.0678	3.47	24.9	0.27	NO	YES	YES	YES
39135	Quaggy River	Chinbrook Meadows	13	0.98	541000	171950	14.50	0.715	674	0.3650	0.998	0.0526	5.78	43.2	0.27	NO	YES	YES	YES
39813	Mol	Ifield Weir	10	3.23	524450	136350	13.08	0.675	827	0.1972	0.890	0.0911	2.94	42.4	0.36	NO	YES	NO	YES
39830	Beck	Rectory Road	7	2.06	536800	169900	9.15	0.728	673	0.5249	0.937	0.0599	3.64	25.9	0.29	NO	YES	NO	YES
39831	Chaffinch Brook	Beckenham	7	2.17	535950	168500	9.29	0.597	673	0.4980	1.000	0.0760	3.23	32.7	0.29	NO	YES	NO	YES
39832	Wandle	Carshalton	34	0.74	527900	164750	0.90	0.855	668	0.6086	1.000	0.0418	1.44	33.8	0.29	NO	YES	YES	NO
40021	Hexden Channel	Hopemill Bridge Sandhurst	32	12.75	581300	129000	32.06	0.406	778	0.0126	1.000	0.0347	6.99	76.0	0.35	NO	YES	YES	NO
40555	Old Mill Stream Tributary East Stour	Aylesford Stream	10	3.35	602350	141250	17.96	0.686	753	0.0369	0.991	0.0922	6.00	38.3	0.34	NO	YES	YES	NO
40556	Cradlebridge Sewer	Cradlebridge	10	3.16	594950	133700	5.80	0.248	712	0.0213	1.000	0.0953	2.23	36.7	0.34	NO	YES	YES	NO

Station No.	Station	Watercourse	Record (years)	QMED (m^3/s)	Easting	Northing	AREA (km^2)	BFI-HOST	SAAR (mm)	$URBEXT_{2000}$	FARL	FPEXT	DPLBAR (km)	DPSBAR (m/km)	PROP-WET	Pooling	QMED	Task 1	MacD & Fraser
40809	Pippingford Brook	Paygate	17	9.20	547950	134300	24.16	0.413	859	0.0061	0.913	0.0182	7.21	92.8	0.36	NO	YES	NO	YES
41001	Nunningham Stream	Tilley Bridge	33	11.00	566150	112850	16.79	0.378	804	0.0131	1.000	0.0664	4.86	67.6	0.34	NO	YES	NO	YES
41016	Cuckmere	Cowbeech	45	8.72	561150	115050	19.10	0.471	855	0.0273	0.966	0.0434	5.00	78.5	0.34	NO	YES	NO	YES
41020	Bevern Stream	Clappers Bridge	45	13.66	542250	116150	35.42	0.355	886	0.0128	0.993	0.0757	8.70	46.7	0.34	YES	YES	YES	NO
41021	Clayhill Stream	Old Ship	45	4.07	544850	115300	7.10	0.252	805	0.0000	1.000	0.0509	2.89	27.2	0.34	NO	YES	YES	YES
41028	Chess Stream	Chess Bridge	48	6.84	521750	117300	24.92	0.497	849	0.0135	0.983	0.0971	6.69	47.0	0.34	NO	YES	NO	YES
41037	Winterbourne Stream	Lewes	37	0.46	540300	109550	17.41	0.966	904	0.0098	1.000	0.0074	4.03	124.4	0.34	NO	YES	NO	YES
41801	Hollington Stream	Hollington	6	2.03	578800	110050	3.47	0.366	781	0.4094	1.000	0.0094	2.25	83.4	0.34	NO	YES	NO	YES
41806	North End Stream	Allington	15	0.72	538450	113800	2.37	0.646	929	0.0000	1.000	0.0211	1.31	130.3	0.34	NO	YES	NO	YES
42017	Hermitage Stream	Havant	48	9.30	471050	106750	17.35	0.245	785	0.2380	0.991	0.0748	4.09	32.7	0.34	NO	YES	YES	YES
42019	Tanners Brook	Millbrook	31	3.47	438800	113250	14.15	0.368	793	0.2498	0.978	0.0408	5.00	56.3	0.33	NO	YES	NO	YES
42020	Tadburn Lake Stream	Romsey	28	3.01	436250	121250	19.61	0.607	782	0.0537	0.983	0.0630	5.98	48.4	0.34	NO	YES	NO	YES
43019	Shreen Water	Colesbrook	41	13.51	380750	127850	30.36	0.565	884	0.0152	0.993	0.0630	5.76	51.7	0.35	NO	YES	YES	NO
43555	Brit	Beaminster RI	6	5.57	348000	101100	9.69	0.453	992	0.0459	0.985	0.0177	2.16	114.8	0.38	NO	YES	YES	NO
44006	Sydling Water	Sydling St Nicholas	40	0.90	363250	99650	12.06	0.879	1030	0.0048	0.944	0.0162	3.33	128.9	0.38	NO	YES	YES	YES
44008	South Winterbourne	Winterbourne Steepleton	35	0.45	362900	89750	20.17	0.811	1012	0.0043	1.000	0.0149	5.01	93.7	0.38	YES	YES	YES	NO
44009	Wey	Broadwey	37	1.82	366600	83950	7.95	0.783	894	0.0225	1.000	0.0153	2.95	117.7	0.38	NO	YES	YES	YES
44013	Piddle	Little Puddle	21	3.04	371650	96800	31.27	0.889	1004	0.0043	1.000	0.0152	5.70	115.3	0.38	YES	YES	YES	NO

Station No.	Station	Watercourse	Record (years)	QMED (m^3/s)	Easting	Northing	AREA (km^2)	BFI-HOST	SAAR (mm)	$URBEXT_{2000}$	FARL	FPEXT	DPLBAR (km)	DPSBAR (m/km)	PROP-WET	Pooling	QMED	Task 1	MacD & Fraser
44801	Hooke	Hooke	22	1.45	353850	99950	11.76	0.597	1030	0.0013	0.923	0.0183	3.51	81.0	0.38	NO	YES	YES	YES
44807	Win	Winfrith	9	1.20	380550	84900	16.78	0.786	894	0.0047	1.000	0.0150	3.82	108.5	0.36	NO	YES	NO	YES
45006	Quarme	Enterwell	9	9.76	291950	135550	20.03	0.514	1419	0.0008	1.000	0.0095	5.45	153.9	0.54	NO	YES	NO	YES
45816	Haddeo	Upton	21	3.52	298850	128900	6.81	0.590	1210	0.0050	1.000	0.0114	3.07	81.0	0.37	YES	YES	YES	YES
45817	Unnamed Stream	Upton	21	1.36	298850	128950	1.74	0.603	1207	0.0022	1.000	0.0172	1.40	67.8	0.54	NO	YES	YES	YES
45818	Withiel Florey Stream	Bessom Bridge	22	4.26	298050	132600	9.85	0.578	1270	0.0000	1.000	0.0056	2.60	103.5	0.42	NO	YES	YES	YES
46005	East Dart	Bellever	50	38.51	265700	77550	22.27	0.363	2095	0.0004	1.000	0.0420	6.46	95.0	0.46	YES	YES	YES	YES
46801	Erme	Erme	9	23.37	264000	63250	15.25	0.257	2110	0.0000	0.992	0.0389	4.00	74.9	0.47	NO	YES	NO	YES
46806	Avon	Avon Intake	17	24.73	268050	64100	13.84	0.370	2152	0.0000	0.903	0.0381	4.80	90.9	0.47	NO	YES	NO	YES
47009	Tiddy	Tideford	45	5.96	234400	59550	37.37	0.591	1276	0.0107	1.000	0.0237	8.19	121.2	0.48	YES	YES	YES	NO
47016	Lumburn	Lumburn Bridge	13	7.02	245900	73250	20.58	0.597	1285	0.0044	1.000	0.0145	4.94	83.8	0.48	NO	YES	NO	YES
47021	Kensey	Launceston Newport	12	18.58	233000	85180	34.83	0.584	1298	0.0174	0.998	0.0218	8.33	101.3	0.50	YES	YES	YES	NO
47022	Tory Brook	Newnham Park	21	7.33	255100	57650	13.45	0.431	1403	0.0141	0.942	0.0233	4.90	106.0	0.48	YES	YES	YES	NO
47025	Wolf	Germansweek	21	11.49	244550	94250	11.34	0.411	1188	0.0007	1.000	0.0066	3.13	78.1	0.50	NO	YES	NO	YES
47804	Hennard Stream	Moors Mill	14	7.62	242450	93850	7.17	0.398	1150	0.0000	1.000	0.0059	2.30	67.3	0.50	NO	YES	NO	YES
48001	Fowey	Trekeivesteps	45	17.32	222650	69750	36.80	0.445	1636	0.0026	0.938	0.0435	7.03	92.4	0.47	YES	YES	YES	NO
48004	Warleggan	Trengoffe	45	9.98	215900	67350	25.26	0.499	1445	0.0025	0.978	0.0350	6.10	93.8	0.45	YES	YES	YES	NO
48005	Kenwyn	Truro	46	5.60	182050	45000	19.11	0.601	1100	0.0342	0.988	0.0096	4.91	90.4	0.42	NO	YES	YES	YES
48006	Cober	Helston	30	5.49	165450	27250	40.83	0.671	1206	0.0189	0.979	0.0337	7.75	74.9	0.44	NO	YES	YES	NO
48007	Kennal	Ponsanooth	46	4.24	176150	37650	26.83	0.736	1294	0.0103	0.866	0.0258	6.28	65.4	0.44	YES	YES	YES	NO

Station No.	Station	Watercourse	Record (years)	QMED (m^3/s)	Easting	Northing	AREA (km^2)	BFI-HOST	SAAR (mm)	$URBEXT_{2000}$	FARL	FPEXT	DPLBAR (km)	DPSBAR (m/km)	PROP-WET	Pooling	QMED	Task 1	MacD & Fraser
48008	St Austell	Molingey	11	7.25	200700	49500	30.25	0.601	1272	0.0734	0.961	0.0343	5.19	111.0	0.45	NO	YES	YES	NO
48009	St Neot	Craigshill Wood	12	8.47	218400	66150	22.91	0.463	1512	0.0023	0.982	0.0224	7.12	78.2	0.45	YES	YES	YES	YES
48555	Allen (Truro)	Idless	17	6.08	182300	47300	24.48	0.616	1081	0.0055	1.000	0.0198	4.12	86.5	0.45	NO	YES	YES	NO
48556	Par	Luxulyan	17	6.91	204350	58200	25.75	0.541	1385	0.0312	0.972	0.1169	5.01	54.7	0.45	NO	YES	YES	NO
48801	Cober	Trenear	27	2.40	167550	31050	26.53	0.672	1265	0.0040	0.976	0.0351	4.38	65.0	0.44	NO	YES	YES	NO
49003	De Lank	De Lank	48	13.99	213350	76550	21.61	0.379	1628	0.0000	0.998	0.0636	4.78	76.0	0.45	YES	YES	YES	YES
49005	Bollingey Stream	Bolingey Cocks Bridge	4	6.52	176850	52900	16.08	0.627	1044	0.0060	0.991	0.0229	3.72	81.4	0.42	YES	YES	YES	NO
49006	Camel	Camelford	8	9.12	210650	83850	12.86	0.576	1418	0.0042	1.000	0.0122	3.89	57.5	0.46	YES	YES	YES	NO
49555	Valency	Boscastle Anderton Ford	8	6.99	214150	91150	5.53	0.534	1327	0.0106	1.000	0.0063	1.70	76.5	0.46	NO	YES	YES	NO
49556	Coastal Stream	Port Isaac	8	0.42	199800	80400	2.77	0.503	999	0.0144	1.000	0.0027	1.50	135.3	0.45	NO	YES	YES	NO
50009	Northlew	Norley Bridge	24	18.51	250150	99950	20.16	0.446	1195	0.0014	1.000	0.0231	4.44	77.3	0.50	NO	YES	NO	YES
51002	Horner Water	West Luccombe	33	10.60	289800	145850	20.38	0.539	1485	0.0001	0.978	0.0028	6.30	213.9	0.54	YES	YES	YES	YES
51003	Washford	Beggearn Huish	47	6.12	304000	139450	36.70	0.588	1151	0.0028	0.982	0.0048	6.84	194.4	0.38	YES	YES	YES	NO
52015	Land Yeo	Wraxall Bridge	35	3.41	348300	171550	23.33	0.669	906	0.0167	0.933	0.0579	6.12	73.3	0.35	NO	YES	YES	YES
52016	Currypool Stream	Currypool Farm	44	2.68	322050	138200	15.70	0.586	934	0.0000	1.000	0.0375	4.75	133.8	0.35	NO	YES	YES	YES
52020	Gallica Stream	Gallica Bridge	8	20.28	357050	109950	16.61	0.389	950	0.0014	0.971	0.0181	4.08	86.4	0.38	NO	YES	NO	YES
52025	Hillfarrance Brook	Milverton	22	10.67	311350	127000	27.75	0.633	1009	0.0141	0.996	0.0230	5.21	135.5	0.35	NO	YES	YES	NO
52026	Alham	Higher Alham	29	1.44	367950	140900	4.90	0.610	1006	0.0041	1.000	0.0071	1.53	90.0	0.37	NO	YES	YES	YES
54022	Severn	Plynlimon Flume	37	15.03	285250	287200	8.69	0.323	2483	0.0000	1.000	0.0098	2.90	180.2	0.66	YES	YES	YES	YES

Station No.	Station	Watercourse	Record (years)	QMED (m^3/s)	Easting	Northing	AREA (km^2)	BFI-HOST	SAAR (mm)	$URBEXT_{2000}$	FARL	FPEXT	DPLBAR (km)	DPSBAR (m/km)	PROP-WET	Pooling	QMED	Task 1	MacD & Fraser
54026	Chelt	Slate Mill	23	9.42	389150	226450	31.31	0.443	726	0.2049	0.975	0.1034	11.01	79.5	0.33	NO	YES	YES	NO
54060	Potford Brook	Sandyford Bridge	26	2.38	363600	322000	22.37	0.645	677	0.0013	0.998	0.1328	5.37	24.7	0.34	NO	YES	NO	YES
54062	Stoke Brook	Stoke	13	0.45	363750	328000	10.92	0.757	698	0.0224	0.939	0.0985	5.20	26.3	0.34	NO	YES	NO	YES
54087	Allford Brook	Childs Ercall	21	0.17	366650	322750	2.92	0.863	663	0.0000	1.000	0.1083	1.59	15.2	0.34	NO	YES	NO	YES
54090	Tanllwyth	Tanllwyth Flume	28	2.27	284250	287650	1.10	0.328	2462	0.0000	1.000	0.0080	1.20	155.0	0.66	NO	YES	NO	YES
54091	Severn	Hafren Flume	34	5.92	284350	287650	3.48	0.303	2514	0.0000	1.000	0.0122	2.21	159.2	0.66	NO	YES	YES	YES
54092	Hore	Hore Flume	34	6.35	284550	287250	3.19	0.330	2531	0.0000	1.000	0.0015	2.22	214.5	0.66	NO	YES	YES	YES
54097	Hore	Upper Hore Flume	14	3.90	283150	286900	1.61	0.303	2649	0.0000	1.000	0.0031	1.39	224.3	0.66	NO	YES	NO	YES
54555	Bow Brook	Feckenham	12	8.92	400500	261150	22.44	0.365	690	0.0837	0.980	0.0491	6.06	52.1	0.28	NO	YES	YES	NO
55010	Wye	Pant Mawr	56	50.40	284300	282550	27.18	0.386	2341	0.0000	1.000	0.0183	4.87	211.8	0.66	NO	YES	YES	NO
55015	Honddu	Tafalog	29	16.68	327700	229350	24.93	0.573	1315	0.0000	1.000	0.0066	5.48	257.1	0.54	NO	YES	NO	YES
55033	Wye	Gwy Flume	33	8.93	282450	285350	3.84	0.330	2575	0.0000	1.000	0.0156	2.08	200.4	0.66	NO	YES	NO	YES
55035	Iago	Iago Flume	15	1.85	282500	285400	1.01	0.335	2461	0.0000	1.000	0.0025	1.10	186.6	0.66	NO	YES	NO	YES
57010	Ely	Lanelay	41	41.06	303350	182650	38.90	0.455	1620	0.0340	1.000	0.0444	6.57	117.9	0.47	NO	YES	YES	NO
57017	Rhondda Fawr	Tynewydd	13	24.30	293250	198650	16.64	0.317	2458	0.0156	0.999	0.0117	3.26	217.4	0.57	YES	YES	YES	NO
58010	Hepste	Esgair Carnau	24	11.92	296950	213350	10.94	0.261	2079	0.0000	1.000	0.0397	3.75	78.4	0.62	NO	YES	NO	YES
60012	Twrch	Ddol Las	37	13.39	265050	243950	19.50	0.419	1531	0.0025	1.000	0.0324	6.43	161.0	0.65	NO	YES	YES	YES
64011	Cerist	Llawr Cae	18	5.37	281950	316300	5.35	0.459	2159	0.0000	1.000	0.0037	2.00	433.3	0.66	NO	YES	YES	NO
65005	Erch	Pencaenewydd	42	10.85	240000	340450	19.39	0.439	1477	0.0012	0.991	0.0711	7.11	96.0	0.56	NO	YES	YES	YES
65008	Nant Peris	Tan-Yr-Alt	21	33.60	260850	357950	10.32	0.548	3465	0.0000	0.996	0.0444	3.19	487.9	0.71	NO	YES	NO	YES

Station No.	Station	Watercourse	Record (years)	QMED (m^3/s)	Easting	Northing	AREA (km^2)	BFI-HOST	SAAR (mm)	$URBEXT_{2000}$	FARL	FPEXT	DPLBAR (km)	DPSBAR (m/km)	PROP-WET	Pooling	QMED	Task 1	MacD & Fraser
66801	Upper Conway	Blaen y Coed	6	14.68	280450	345100	11.73	0.228	2196	0.0000	0.911	0.0541	4.31	83.8	0.71	NO	YES	NO	YES
67003	Llyn Brenig	Llyn Brenig Outflow	10	15.28	297450	353850	22.45	0.319	1317	0.0000	0.587	0.0182	3.99	72.2	0.70	NO	YES	NO	YES
67010	Gelyn	Cynefail	41	16.42	284350	341950	12.87	0.251	2000	0.0000	0.969	0.0322	3.74	127.8	0.71	NO	YES	YES	YES
67013	Hirnant	Plas Rhiwedog	12	24.08	294600	334950	32.47	0.415	1756	0.0000	1.000	0.0182	6.21	222.9	0.71	NO	YES	YES	NO
68010	Fender	Ford Lane	9	5.54	328050	388000	18.00	0.432	774	0.2078	0.999	0.1024	4.02	29.0	0.38	NO	YES	YES	YES
68011	Arley Brook	Gore Farm	6	6.09	369650	379900	33.76	0.437	831	0.0212	0.998	0.2498	7.32	12.4	0.37	NO	YES	YES	NO
68014	Sandersons Brook	Sandbach	5	1.45	375350	365250	3.77	0.394	742	0.0229	0.986	0.1597	2.75	13.3	0.37	NO	YES	NO	YES
68021	Arrowe Brook	Acton Lane	8	4.45	325500	389300	17.87	0.513	750	0.1735	0.996	0.1385	4.47	24.8	0.38	YES	YES	YES	NO
69019	Worsley Brook	Eccles	25	5.74	375300	397950	24.09	0.349	956	0.3450	0.941	0.2202	5.29	22.1	0.44	NO	YES	YES	YES
69034	Musbury Brook	Helmshore	8	5.02	377450	421250	3.03	0.345	1454	0.0000	1.000	0.0083	1.79	159.4	0.51	NO	YES	NO	YES
69042	Ding Brook	Naden Reservoir	21	1.58	385000	417450	2.18	0.401	1488	0.0000	1.000	0.0138	1.29	139.9	0.57	NO	YES	YES	YES
69046	Bradshaw Brook	Bradshaw Tennis Club	14	23.50	373300	412200	36.87	0.345	1384	0.0234	0.781	0.0223	7.45	101.2	0.51	NO	YES	YES	NO
69047	Roch	Littleborough	21	8.25	394000	416500	14.80	0.475	1353	0.0354	0.890	0.0384	3.71	138.5	0.57	YES	YES	YES	NO
71003	Croasdale Beck	Croasdale Flume	18	10.07	370650	454650	10.71	0.276	1882	0.0000	1.000	0.0160	3.61	156.8	0.60	YES	YES	YES	YES
71005	Bottoms Beck	Bottoms Beck Flume	14	15.52	374500	456550	10.58	0.280	1511	0.0000	0.999	0.0459	3.51	90.8	0.60	NO	YES	NO	YES
71013	Darwen	Ewood	41	27.57	367700	426250	39.08	0.424	1340	0.1393	0.938	0.0356	6.52	95.1	0.51	YES	YES	YES	NO
72007	Brock	U/S A6	34	31.41	351150	440550	31.53	0.319	1361	0.0000	1.000	0.0535	9.69	108.8	0.60	YES	YES	YES	NO
72014	Conder	Galgate	47	17.70	348150	455350	28.99	0.443	1183	0.0064	0.975	0.0822	7.39	93.4	0.60	YES	YES	YES	NO
72817	New Mill Brook	Hollowforth Hall	30	20.51	350350	436350	33.19	0.387	1110	0.0212	0.992	0.0930	8.50	35.8	0.52	NO	YES	YES	NO

Station No.	Station	Watercourse	Record (years)	QMED (m^3/s)	Easting	Northing	AREA (km^2)	BFI-HOST	SAAR (mm)	$URBEXT_{2000}$	FARL	FPEXT	DPLBAR (km)	DPSBAR (m/km)	PROP-WET	Pooling	QMED	Task 1	MacD & Fraser
73006	Cunsey Beck	Eel House Bridge	41	7.66	336950	494050	18.77	0.448	1897	0.0022	0.727	0.0522	5.08	119.2	0.71	NO	YES	YES	NO
73009	Sprint	Sprint Mill	45	42.55	351450	496100	34.80	0.453	2011	0.0000	0.997	0.0612	10.22	224.3	0.71	YES	YES	YES	NO
73015	Keer	High Keer Weir	21	12.24	352250	471900	30.06	0.486	1158	0.0029	0.976	0.0746	6.20	83.0	0.60	YES	YES	NO	NO
73803	Winster	Lobby Bridge	12	8.46	342350	488550	22.03	0.538	1508	0.0000	0.991	0.0608	5.01	119.2	0.71	NO	YES	NO	YES
76001	Haweswater Beck	Burnbanks	35	18.22	350850	515950	32.34	0.345	2438	0.0000	0.645	0.0154	6.21	293.6	0.71	YES	YES	YES	NO
76011	Coal Burn	Coalburn	37	1.84	369350	577750	1.63	0.196	1096	0.0000	1.000	0.0736	1.30	47.2	0.62	YES	YES	YES	YES
76811	Dacre Beck	Dacre Bridge	14	53.01	346050	526300	33.97	0.457	1428	0.0000	0.999	0.0724	7.61	101.4	0.64	YES	YES	YES	NO
80003	White Laggan Burn	Loch Dee	33	6.60	246800	578050	5.70	0.385	2469	0.0000	0.996	0.0168	2.10	246.3	0.69	NO	YES	YES	NO
80004	Green Burn	Loch Dee	22	4.15	248050	579050	2.62	0.365	2383	0.0000	0.998	0.0172	1.42	189.9	0.69	NO	YES	NO	YES
80005	Dargall Lane	Loch Dee	28	4.10	245050	578700	2.07	0.355	2435	0.0000	1.000	0.0230	1.64	307.5	0.69	NO	YES	YES	YES
84002	Calder	Muirshiel	21	16.31	230900	663750	12.06	0.271	2316	0.0000	0.988	0.0398	4.46	95.2	0.61	NO	YES	YES	NO
84016	Luggie Water	Condorrat	46	24.35	273950	672550	35.32	0.327	1089	0.0658	0.995	0.0526	5.76	56.0	0.58	YES	YES	YES	NO
84023	Bothlin Burn	Auchengeich	42	8.63	267800	671600	34.85	0.313	1029	0.0940	0.912	0.1365	9.24	38.4	0.58	YES	YES	YES	NO
84026	Allander Water	Milngavie	41	36.01	255800	673750	30.36	0.369	1424	0.0404	0.896	0.0501	7.90	101.3	0.61	NO	YES	YES	NO
84029	Cander Water	Candermill	37	20.52	276400	647000	25.50	0.399	1033	0.0165	0.985	0.0419	5.86	52.9	0.58	NO	YES	YES	YES
84034	Auldhouse Burn	Spiers Bridge	15	6.20	254600	659050	17.17	0.478	1329	0.1761	0.924	0.0771	6.45	56.7	0.61	NO	YES	NO	YES
84035	Kittoch Water	Waterside	15	19.10	259600	656200	16.80	0.337	1184	0.2647	0.978	0.0525	4.66	56.8	0.58	NO	YES	NO	YES
84036	Earn Water	Letham	15	15.80	256650	654850	20.89	0.431	1481	0.0005	0.908	0.0528	7.84	75.9	0.61	NO	YES	NO	YES
86001	Little Eachaig	Dalinlongart	42	42.83	214250	682050	31.84	0.393	2340	0.0005	1.000	0.0270	5.10	277.6	0.71	NO	YES	YES	NO
87801	Allt Uaine	Intake	20	8.50	226250	711300	2.89	0.358	3473	0.0000	1.000	0.0087	1.31	368.6	0.74	NO	YES	NO	YES

Station No.	Station	Watercourse	Record (years)	QMED (m^3/s)	Easting	Northing	AREA (km^2)	BFI-HOST	SAAR (mm)	$URBEXT_{2000}$	FARL	FPEXT	DPLBAR (km)	DPSBAR (m/km)	PROP-WET	Pooling	QMED	Task 1	MacD & Fraser
89004	Strae	Glen Strae	29	60.11	214650	729400	37.38	0.362	2766	0.0000	0.995	0.0468	7.20	324.4	0.79	YES	YES	YES	NO
89007	Abhainn a' Bhealaich	Braevallich	25	40.22	195700	707600	23.60	0.303	2488	0.0000	0.923	0.0426	5.25	128.6	0.75	NO	YES	NO	YES
91802	Allt Leachdach	Intake	34	6.35	226150	778100	6.52	0.397	2555	0.0000	0.992	0.0031	2.68	407.5	0.83	NO	YES	NO	YES
101005	Eastern Yar	Budbridge	31	4.70	453050	83450	24.28	0.707	841	0.0193	0.996	0.0458	6.53	87.1	0.33	NO	YES	YES	YES
102001	Cefni	Bodffordd	16	9.60	242900	376850	21.01	0.448	1061	0.0007	0.964	0.1029	6.17	29.0	0.45	NO	YES	NO	YES
203038	Rocky	Rocky Mountain	18	10.80	324300	326550	6.80	0.327	1610	0.0000	1.000	0.0162	2.31	213.1	0.53	NO	YES	NO	YES
203046	Rathmore Burn	Rathmore Bridge	32	10.82	319750	385350	22.51	0.430	1043	0.0000	1.000	0.0726	6.46	57.8	0.52	YES	YES	YES	YES
203049	Clady	Clady Bridge	32	23.24	319950	383750	29.38	0.367	1079	0.0000	1.000	0.0599	8.00	58.1	0.52	NO	YES	YES	NO
205015	Cotton	Grandmere	13	2.80	352450	381750	22.53	0.489	863	0.0765	0.998	0.2169	5.76	24.3	0.52	NO	YES	NO	YES
205034	Woodburn	Control	11	0.12	337300	390050	5.74	0.410	1177	0.0000	0.901	0.0514	2.17	53.9	0.52	NO	YES	YES	NO
205101	Blackstaff	Eason's	22	9.86	331750	372400	14.15	0.414	990	0.4192	0.997	0.1657	2.95	54.2	0.52	NO	YES	YES	YES
206004	Bessbrook	Carnbane	30	10.59	307400	329250	34.76	0.584	1055	0.0204	0.917	0.0441	6.68	97.3	0.53	NO	YES	YES	NO
206006	Annalong	Recorder	48	15.33	334800	323350	13.66	0.336	1720	0.0000	0.980	0.0236	3.20	275.8	0.53	YES	YES	YES	YES

Notes:

- 9 catchments are listed with Station No. xx555 or xx556, where xx indicates the hydrometric area of the catchment. This numbering system is used only in the ‘Small Catchments’ series of reports for catchments that are not part of the NRFA numbering system.
- Catchments 4063 and 4186 are located in the Thames Basin (hydrometric area 39) rather than the Conon group (hydrometric area 4). These station numbers were supplied by David MacDonald and are assumed to be local.

Appendix B

Data standardisation

The data supplied for this study came from several sources. Catchment descriptors for 119 of the sites identified in Task 1, plus the two added sites, came from the NRFA peak flow data set v3.3.4. Flood peak series for these sites derive mainly from the NRFA peak flow data set, but also include added values for stations that were still open as of December 2014 (or were closed after September 2012). Catchment descriptors for the remaining 32 sites identified in Task 1 were exported from the FEH CD-ROM version 3, while flood peak data were requested from and supplied by the relevant measuring authority for each site. Catchment descriptor and QMED data for the 130 sites identified by MacDonald & Fraser (2013) and accepted in this study, 64 of which overlap with those identified in Task 1, were supplied in a single Excel spreadsheet. This states that the catchment descriptor values were exported from the FEH CD-ROM version 2. The QMED values were calculated from paper and digitised flood peak records, then input manually.

To avoid any potential inconsistencies between these various data sources, the easting and northing values for all 64 accepted sites identified only by MacDonald and Fraser were checked against the FEH CD-ROM version 3 and modified wherever a discrepancy was noted. The DTM-aligned site co-ordinates were then input to a proprietary CEH software program written to query, in bulk, the digital terrain model and catchment descriptor grids underpinning the FEH CD-ROM. The catchment descriptor values output by this program replaced those given in the spreadsheet.

Catchment descriptor (CD3) files produced by the NRFA contain two measures of area: nominal area and DTM area. While nominal area has the potential to better represent the catchment's true area, DTM area was always selected to use in these analyses, for three reasons. First, catchment-average values for all other descriptors in each CD3 file (for example, SAAR, BFIHOST) are derived from the DTM, using the DTM boundary and therefore DTM area. The DTM area is therefore more coherent with the other catchment descriptors than the nominal area. Second, the replacement descriptors for MacDonald and Fraser's 64 accepted unique catchments were automatically derived from the DTM, again using DTM area. Third, CD3 files exported from the FEH CD-ROM, such as those that practitioners may use for ungauged site analyses, are able to report DTM area only.

Catchment descriptor values for the two added catchments were taken directly from CD3 files included in the NRFA peak flow data set (NRFA, 2014). DTM area was again taken rather than nominal area.

Ten of the 151 catchment descriptor files accepted in Task 1 had missing $URBEXT_{2000}$ values. Four of these 10 catchment descriptor files also had missing FPEXT, centroid easting and centroid northing values. In one case, this was due to mis specified gauging station co-ordinates in the CD3 file. These co-ordinates were corrected and $URBEXT_{2000}$, FPEXT, centroid easting and centroid northing were infilled from the FEH CD-ROM version 3. In the other three cases, the catchments were very small, did not align with the DTM river network, and so relevant catchment descriptor values could not be exported. FPEXT for all three very small catchments was set to zero, following the reasoning that all are very small, steep upland catchments. Catchment centroid co-ordinates were estimated, taking catchment area and the topography of the land surrounding the gauging station into consideration. As these three catchments are very small, the estimate centroids are unlikely to be in error by more than 100 to 200 metres.

Rare cases of missing catchment descriptor values is a known feature of the NRFA peak flow data set v3.3.4 and results from the true locations of stations not aligning with the DTM river network. For these few stations, true catchment descriptor values cannot be derived automatically, either by a user or the NRFA. Missing values were manually infilled using data from the FEH CD-ROM version 3, by aligning these stations with the DTM river network.

Appendix C

Comparison of FEH99 and FEH13 rainfall models in ReFH2

C.1 Method summary

As discussed in Section 2, ReFH2 is the most up-to-date version of the FEH's rainfall-runoff method for hydrological design. When used in design mode, the method requires inputs to characterise the design rainfall (depth, duration, return period and temporal profile) and antecedent catchment conditions. The second release of ReFH2 (v. 2.1) allows the user to apply design storm inputs from the new FEH13 depth-duration-frequency (DDF) model (Stewart and others, 2013) as well as from the original FEH99 model (Faulkner, 1999) but was made available after the bulk of the analytical work described in this report had already been carried out. The FEH13 DDF model is considered to provide improved design rainfall estimates across the full range of return periods for which ReFH2 can be applied, and therefore using it would be expected to result in more accurate estimates of peak flow at ungauged sites. This section presents preliminary comparisons between applying ReFH2 with FEH99 and FEH13 on the small catchment data set.

C.2 Performance overview

Figure 31 plots modelled against observed QMED for ReFH2 with both FEH99 and FEH13 as input rainfall models. All catchments are considered, except those with FARL < 0.9, as the modelling method is unsuitable for use in catchments with significant attenuation.

In general, the choice of FEH99 or FEH13 rainfall appears to have minimal effect on the modelled value of QMED. Using FEH13 appears to correct some underestimations that can occur with FEH99 for observed values of QMED under 1 m^3/s. Additionally, the 'cloud' of modelled QMED values appears to sit slightly higher for the FEH13 rainfall model than the FEH99 model. No obvious difference in the spread of errors is apparent in Figure 31. For the four 'outlier' sites identified in previous sections, using the FEH13 model both improves and worsens QMED estimates at two sites each. A fifth outlier, 33813, is identified. However, identification as an outlier in this section does not necessarily suggest problems with the gauged data, as only results from two similar methods (rather than six methods falling into two broad classes) are shown here.

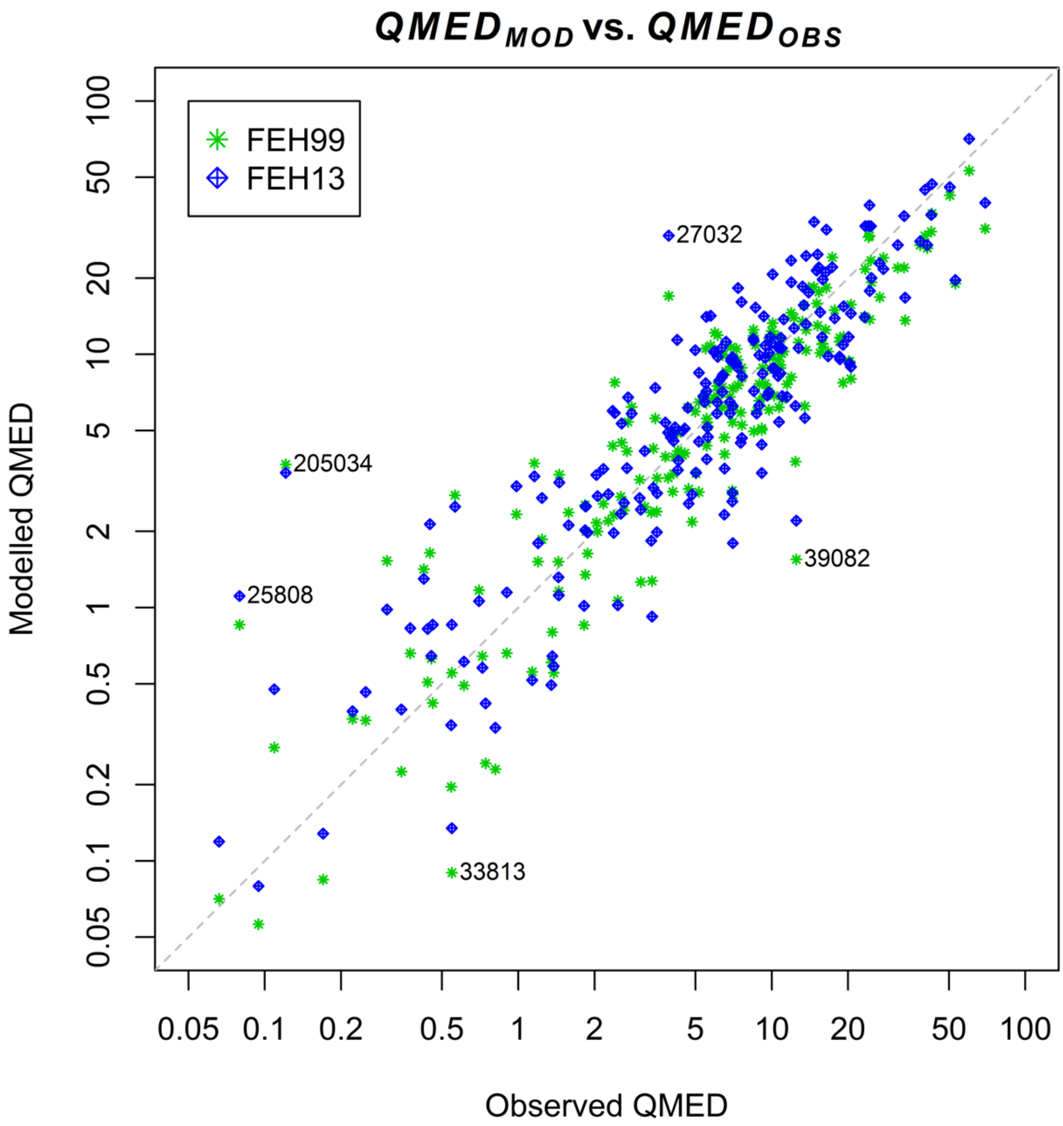

Figure 31 - Modelled vs. observed QMED at 207 sites, excluding those with FARL < 0.9

The scatter graph in Figure 31 plots the modelled QMED (y-axis – from 0.05 to 100) by the observed QMED (x-axis – 0.05 to 100):

- FEH99: green stars.
- FEH13: blue diamonds.

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.