The FEH 2025 statistical method update

Gianni Vesuviano, Adam Griffin

Issue number 1.3
2025-07-22

UK Centre for Ecology & Hydrology

Contents

Executive Summary

This report presents updates to the FEH statistical method resulting from improvements to four catchment descriptors used prominently in the method: *SAAR*, *FARL*, *BFIHOST* and *URBEXT*, and 19 years of additional data collection and period-of-record review at UK river gauging stations.

The structure and application procedure of the FEH statistical method are unchanged; this update concerns the choice of catchment descriptors to use in the method and their corresponding numerical coefficients. Due to the high performance of the modified 2008 *QMED* equation in WINFAP 5, the 2025 equation offers only a marginally improved fit to the data. However, estimates of the current flood frequency in ungauged catchments should be improved, due to better digital representation of those catchments through improved catchment descriptors. Pooled estimates of rarer floods are improved more substantially due to the increased typical record length; uncertainty in *L*-moments continues to reduce as record length increases far beyond the point at which uncertainty in *QMED* becomes small. This method update reiterates that donor stations should be selected by centroid-centroid distance only, since model error, which the donor transfer is intended to correct, is the part of *QMED* that cannot be estimated by the equation and its associated catchment descriptors i.e. *AREA*, $SAAR_{(9120)}$, $BFIHOST_{19(SCALED)}$ and $FARL_{(2015)}$.

Underlying data

The underlying data for the 2025 *QMED* estimation and Pooling methods includes more stations suitable for *QMED* estimation and Pooling, with longer records (typically 19 years extra). For ungauged catchments, where the “current” characteristics of a catchment are most representative, the use of up-to-date catchment descriptors will improve accuracy in most places, although the underlying FEH Local principle of confirming with local expert knowledge is always advised. For gauged catchments, using period-of-record estimates of, for example, *SAAR*, can also improve estimates of flood frequency over the period of record.

QMED

The 2025 *QMED* equation uses the same format and equivalent catchment descriptors as the previous version in WINFAP 5, giving a small improvement to R^2 (0.945 to 0.948) and *fse* (1.4313 to 1.4307). There are only 31 stations where the 2025 estimate of *QMED* is more than 125% of the 2008 *QMED* estimate, and 4 stations where the 2025 estimate of *QMED* is less than 75% of the 2008 *QMED* estimate, with no obvious spatial pattern. In the gauged and ungauged datasets, all of the largest changes were due to differences in *FARL* or *BIFHOST* descriptors, suggesting the change is motivated by changes in the underlying maps of water bodies – updated to use data from LCM2015 (Rowland *et al.* 2015).
This method update formalizes the recommendation that urbanization adjustments

should be applied to all catchments, not just those in which urbanization exceeds a threshold, in order to avoid discontinuities in results on either side of the threshold.

In many applications, *QMED* estimation via catchment descriptors is performed at ungauged catchments. Given there is no “period of record” to refer to in these cases, the most appropriate values of catchment descriptors and urbanization are the most up-to-date values available. Where flow records exist, the best way to model these cases is to take a representative average of the catchment’s history using average urbanization and catchment descriptor values (e.g. *SAAR*). For gauged catchments with long known records, estimates of historical flood frequency should be performed using the full period of record if available, unless there is any reason not to use the full record, such as reservoir construction.

Pooling

The 2025 pooling approach follows the same principles as those developed in 2008. The similarity distance metric (SDM) has been updated to a 5-descriptor form, and the equations for pooled *L*-moments have been updated. There is a single pooling approach for all catchment sizes, including those treated as “small catchments” in the small catchments update to the 2008 method. This means there are potentially some considerable changes to flood frequency curves for the smallest catchments, which is reflected in the ungauged dataset (more than 88% of which are small catchments). Urban adjustments for pooled *L*-moments have also been updated, and should be used for all catchments, rather than just “urban” catchments, to avoid discontinuities at the threshold between “rural” and “urban” catchments.

For the 2025 pooling update, we see a better fitting model illustrated by the lower Pooled Uncertainty Metric (PUM) at all selected return periods (30-1000 years).

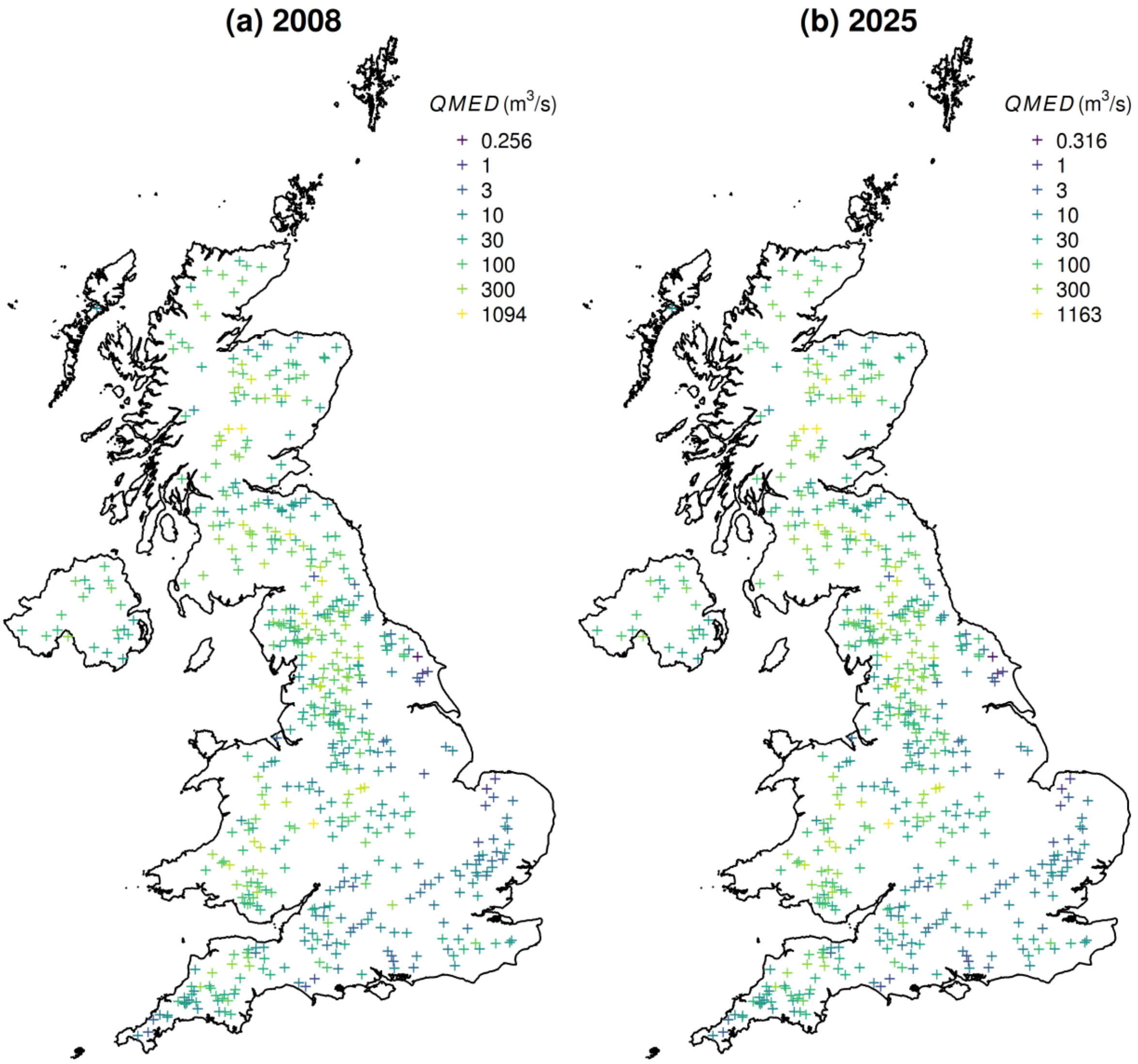

Figure 1: Subset of Figure 16, showing changes in catchment-descriptor *QMED*

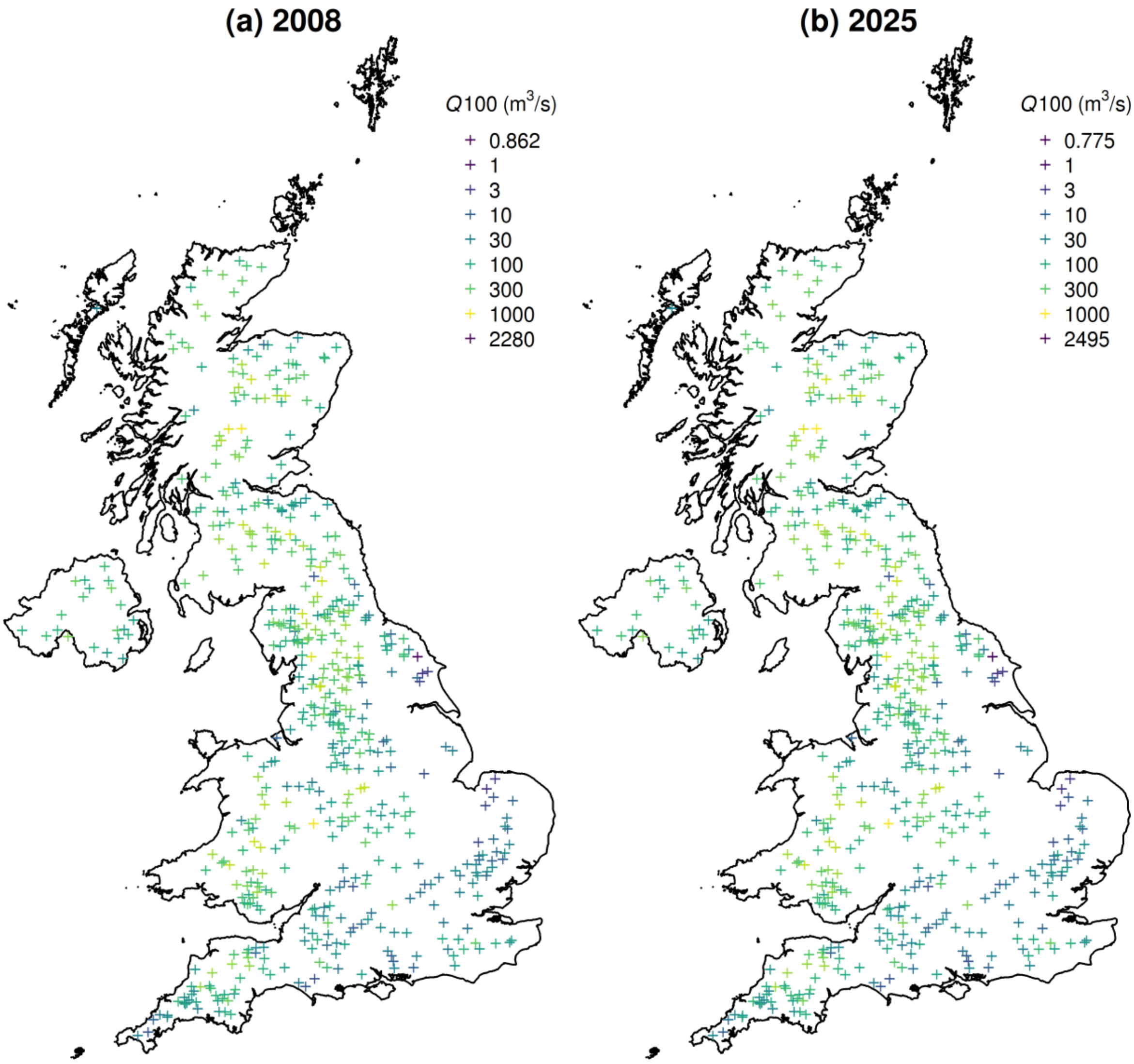

Figure 2: Subset of Figure 35, comparing pooled 100-year flow estimates.

Summary of 2025 method for ungauged catchments

Method aspect	2008	2025	Comments
QMED	$R^2 = 0.945$ $fse = 1.431$	$R^2 = 0.948$ $fse = 1.431$	Based on 602 stations ('08) 626 stations ('25)
Donor transfer	6 donors $R^2 = 0.960$ $fse = 1.374$	8 donors $R^2 = 0.960$ $fse = 1.373$	Based on 626 stations
QMED, urbanized, after donor transfer	$R^2 = 0.946$ $fse = 1.426$	$R^2 = 0.946$ $fse = 1.428$	Based on 884 (626 rural + 258 urban) stations
Pooling SDM & rec length	500 station-years PUM (100 years) = 0.2602	800 station years PUM (100 years) = 0.2479	Based on 378 stations
Pooling, urbanized *L*-moments	PUM (100 years) = 0.2820	PUM (100 years) = 0.2705	Based on 152 stations

Ungauged QMED equation

$$QMED = 6.8247 AREA^{0.8499} 0.1780^{\left(\frac{1000}{SAAR_{9120}}\right)} FARL_{2015}{}^{3.0450} 0.0321^{BFIHOST_{19SCALED}{}^{2}}$$

Donor transfer with 8 donors, closest centroid-centroid distance (previously 6):

$$\boldsymbol{r_{\eta,ij} = 0.4814 e^{-0.0333 d_{ij}} + (1 - 0.4814) e^{-0.4610 d_{ij}}}$$

Ungauged QMED urbanization, applied at all levels of urbanization:

$$QMED_{urban} = QMED_{rural}(1 + 0.3 URBEXT_{2015})^{1.8838} \times$$

$$\left(1 + 0.3 URBEXT_{2015}\left(\frac{PR_{IMP}}{67.0674 - 63.8200 BFIHOST_{19SCALED}} - 1\right)\right)^{3.5200}$$

Pooling-group formation, 800-year length (previously 500):

$$SDM5_{ij} = \sqrt{\begin{aligned}&1.74\left(\frac{\ln AREA_i - \ln AREA_j}{1.3207}\right)^2 + 1.63\left(\frac{\frac{1000}{SAAR_{9120,i}} - \frac{1000}{SAAR_{9120,j}}}{0.3566}\right)^2 + 0.26\left(\frac{FARL_{2015,i}^2 - FARL_{2015,j}^2}{0.0976}\right)^2 \\ &+ 0.55\left(\frac{FPEXT_i - FPEXT_j}{0.0439}\right)^2 + 0.82\left(\frac{\frac{1}{BFIHOST_{19SCALED,i}} - \frac{1}{BFIHOST_{19SCALED,j}}}{0.6610}\right)^2\end{aligned}}$$

(within pooling-group weighting slightly modified)

Urban adjustments to pooled L-moments, applied at all levels of urbanization:

$$L\text{-CV}_{urbanized} = 0.5269^{URBEXT_{2015}} L\text{-CV}_{pooled}$$

$$L\text{-SKEW}_{urbanized} = L\text{-SKEW}_{pooled}$$

UK Centre for Ecology & Hydrology

1. Introduction

The previous FEH statistical method (Kjeldsen *et al.,* 2008) was released 17 years ago, as of 2025. Since then, major advances in the underlying data have been made. However, only small tweaks have been made to the FEH statistical method, including an extension to urban catchments (Kjeldsen, 2010), the ability to use multiple donors for *QMED* adjustment (Kjeldsen *et al.,* 2014) and a small-catchment modification to the pooling similarity distance measure (Vesuviano *et al.,* 2024).

In 2018, the FEH team began the process to develop updates or replacements for the key catchment descriptors used in FEH methods: *BFIHOST*, *SAAR*, *FARL* and *URBEXT*. The first of these, $BFIHOST_{19}$, was released in 2019 (Griffin *et al.,* 2019), while $SAAR_{9120}$, $FARL_{2015}$, $URBEXT_{2015}$ and $BFHOST_{19SCALED}$ will be released with version 14 of the NRFA Peak Flow dataset in autumn 2025, simultaneously with WINFAP 5.3. WINFAP, Wallingford HydroSolutions' commercial implementation of the FEH statistical method, has already used $BFIHOST_{19}$ directly in place of *BFIHOST* in the *QMED* equation for ungauged sites since version 5.0, though both the *QMED* model and the description of sampling error within the optimization function that was used to calibrate the *QMED* model were developed using *BFIHOST*.

Mapping and representation of water bodies has been greatly improved since the development of *FARL* in the late 1990s, with several UKCEH datasets (Land Cover Map 2015: Rowland *et al.,* 2017; Land Cover Map 2019: Morton *et al.,* 2020; and Lakes50k: Hughes *et al.,* 2004) providing options to update this descriptor in UK catchments, and CORINE providing a similar opportunity to augment transboundary catchments with improved representation of water bodies in the Republic of Ireland.

Similarly, the repeated regeneration of Land Cover Map at several time points during the 21st century provides opportunities to estimate urban and suburban coverage within catchments at these same time points, through new *URBEXT* descriptors linked to different instances of Land Cover Map.

The National River Flow Archive undertakes an ongoing annual update and data review programme, which adds new quality-controlled peak and daily mean flow data to station flow records across the UK every year (annual update), and quality-controls the entire flow holdings every few years (data review).

Outside of UKCEH, more recent versions of standard-period average annual rainfall (*SAAR*) have been produced by the Met Office and Met Éireann for the period 1991-2020, corresponding more closely with the AMAX flow recording periods for most stations suitable for this analysis than the 1961-1990 *SAAR* used in the 2008 model. Both estimates (Met Office and Met Éireann) are released under licences that allow commercial use (OGL and CC-BY respectively.

In order to take advantage of these multiple major updates to the data underpinning the FEH methods, a programme of work was developed to update the index flood, donor transfer and growth curve parts of the FEH statistical method, for rural and urban catchments. These updates will be incorporated into WINFAP 5.3, which will be released concurrently with the new $SAAR_{9120}$, $FARL_{2015}$, $URBEXT_{2015}$ and $BFHOST_{19SCALED}$ catchment descriptors.

1.1 Important note on terminology

Throughout this report, references to the “2008” method refer to the FEH statistical method as implemented in WINFAP 5.2. This is substantially based on the method presented by Kjeldsen *et al.* (2008) but includes updates to urbanization procedures (Kjeldsen, 2010), donor transfer (Kjeldsen *et al.*, 2014), pooling in small catchments (Vesuviano *et al.*, 2024) and use of the $BFIHOST_{19}$ catchment descriptor (Griffin et al., 2019) in place of *BFIHOST*.

2. Improved Data

For a full description of the new catchment descriptors, see the dedicated catchment descriptors report (Vesuviano *et al.*, 2025).

2.1 AMAX flow data

Station flow records are updated yearly across the UK, typically with a one-year lag. The latest peak flow data are published by the National River Flow Archive (NRFA) approximately every September, containing annual maximum and peak-over-threshold flows up to the previous September for around 900-1000 stations (minor updates and corrections may also be published between September releases). The stations are also assigned a suitability level, which relates to the accuracy of monitored flows up to *QMED* ("suitable for *QMED*") and above *QMED* ("suitable for pooling"). A station's suitability can change at any update as a result of data review. Stations may also enter or leave the data set at any update.

This recalibration used version 12.1 of the NRFA Peak Flow data set, released in November 2023, while the previous calibration used HiFlows-UK 1.1, released in August 2005.

Kjeldsen *et al.* (2008) manually removed some catchments that were unsuitable for the study but otherwise met the criteria for inclusion. In addition, they manually removed AMAX from some records that spanned a portion before and a portion after reservoir construction, and manually edited *FARL* to pre-reservoir values if only the pre-reservoir AMAX were kept. All changes made to NRFA Peak Flow 12.1 for this study are detailed in this report's Appendix.

In total, 884 catchments were available for this study, although not all were suitable for every stage.

It is noted that the NRFA Peak Flow data set version 13 was released in September 2024. However, this has 909 catchments in common with version 12.1, so any changes to the reported results that might appear from the use of version 13 for calibration will be very minor compared to the changes shown between the current and updated methods in this report.

2.2 *BFIHOST*

BFIHOST was first created in 1995 to estimate soil permeability indirectly, by relating gauged baseflow index at a station to soil type in the upstream catchment through bounded linear regressions (Boorman *et al.,* 1995). Griffin *et al.* (2019) revised the *BFIHOST* estimation procedure, generating a new catchment descriptor, $BFIHOST_{19}$. This was developed to improve upon *BFIHOST* in clay-dominated, peat-dominated and ephemeral catchments, and in catchments with rare soil types (by combining the soil classes with the poorest representation across the UK with the most similar common class) and soil types whose coefficients needed to be constrained in the original model (by using beta regression to eliminate the need for bounding values to the range 0-1). $BFIHOST_{19}$ is available through the FEH Web Service for the entire UK river network, plus small parts of the Republic of Ireland.

Figure 3 compares *BFIHOST* and $BFIHOST_{19}$, showing that 68.4% of $BFIHOST_{19}$ values (mean 0.480, median 0.436) are lower than corresponding *BFIHOST* values (mean 0.498, median 0.464). Figure 3 also shows that $BFIHOST_{19}$ values approach 1 more slowly than do *BFIHOST* values. This is an expected consequence of beta regression, vs linear regression that requires bounding.

$BFIHOST_{19SCALED}$ is a further update to $BFIHOST_{19}$ that is intended to remove the effects of water bodies from $BFIHOST_{19}$, with the intention of producing an ln(*QMED*) model that does not double-count them, improving the “physical consistency” of the model. To generate $BFIHOST_{19SCALED}$ for each catchment, area corresponding to HOST class 30 was removed, and the percentage of the remaining area covered by each of HOST classes 1-29 was calculated. $BFIHOST_{19SCALED}$ was then calculated in exactly the same way as $BFIHOST_{19}$, following Griffin *et al.* (2019).

Figure 4 plots $BFIHOST_{19SCALED}$ against $BFIHOST_{19}$ for all 884 study catchments, clearly showing that the differences between them tend to increase as $FARL_{2015}$ reduces. The relationship is not perfect, as $FARL_{2015}$ is not only a measure of water body coverage, but also water body location and connection to the river network. Furthermore, $FARL_{2015}$ is based on LCM, not HOST, data. Nevertheless, it clearly demonstrates the removal of water body effects from $BFIHOST_{19}$.

$BFIHOST_{19SCALED}$ should always be lower than $BFIHOST_{19}$, but the starred catchment in Figure 4, Wandle at South Wimbledon (NRFA 39003) has a higher value of $BFIHOST_{19SCALED}$ than $BFIHOST_{19}$. This is because the positioning of the gauging station on the digital river network was improved after the generation of the $BFIHOST_{19}$ descriptor, so the current river network snapping method excludes one tributary that was included in 2018 when $BFIHOST_{19}$ was first calculated for the UK river network.

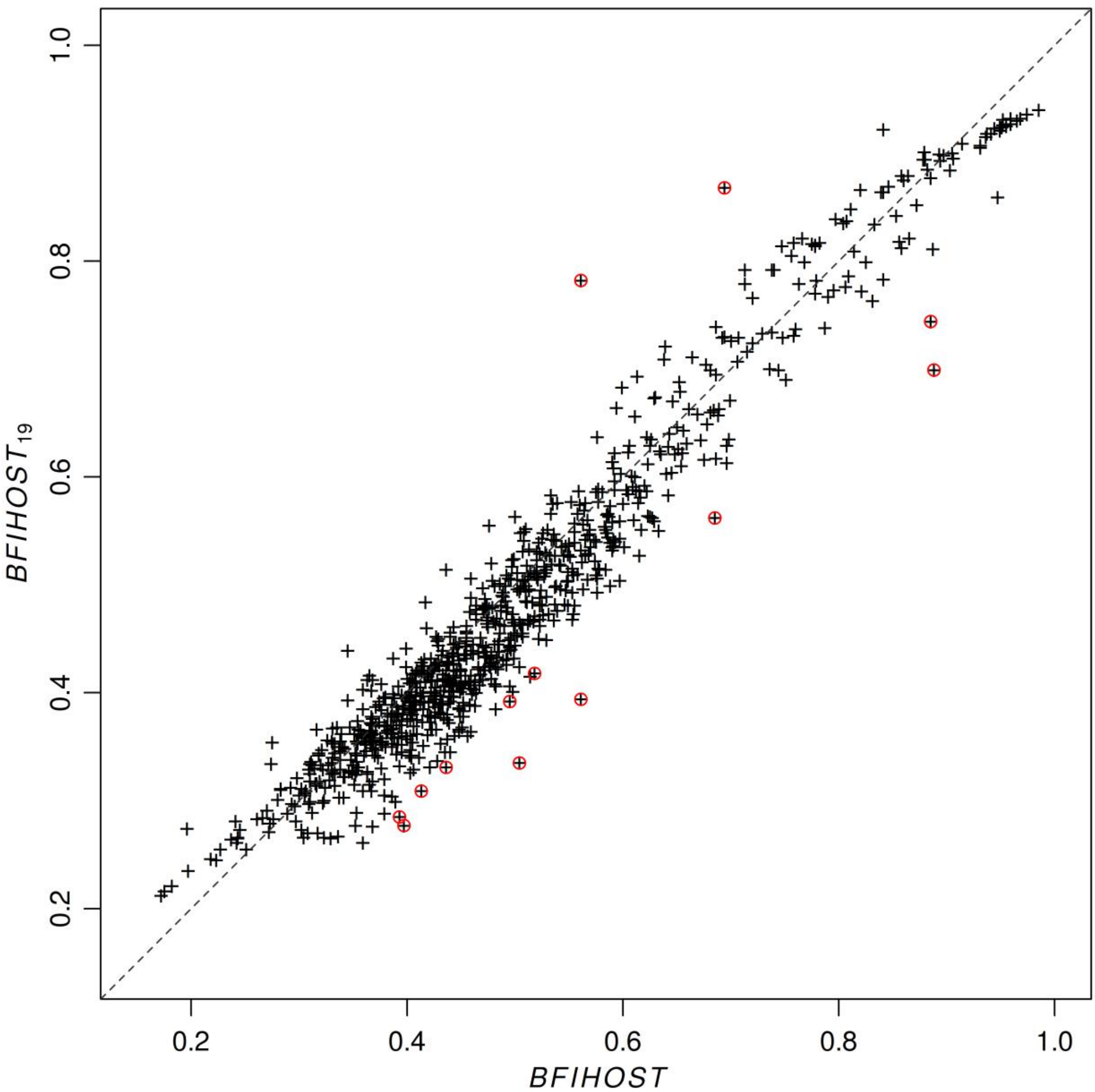

Figure 3 ***BFIHOST*** **vs.** ***BFIHOST***$_{19}$ **for all 884 study catchments. Absolute differences greater than 0.1 circled.**

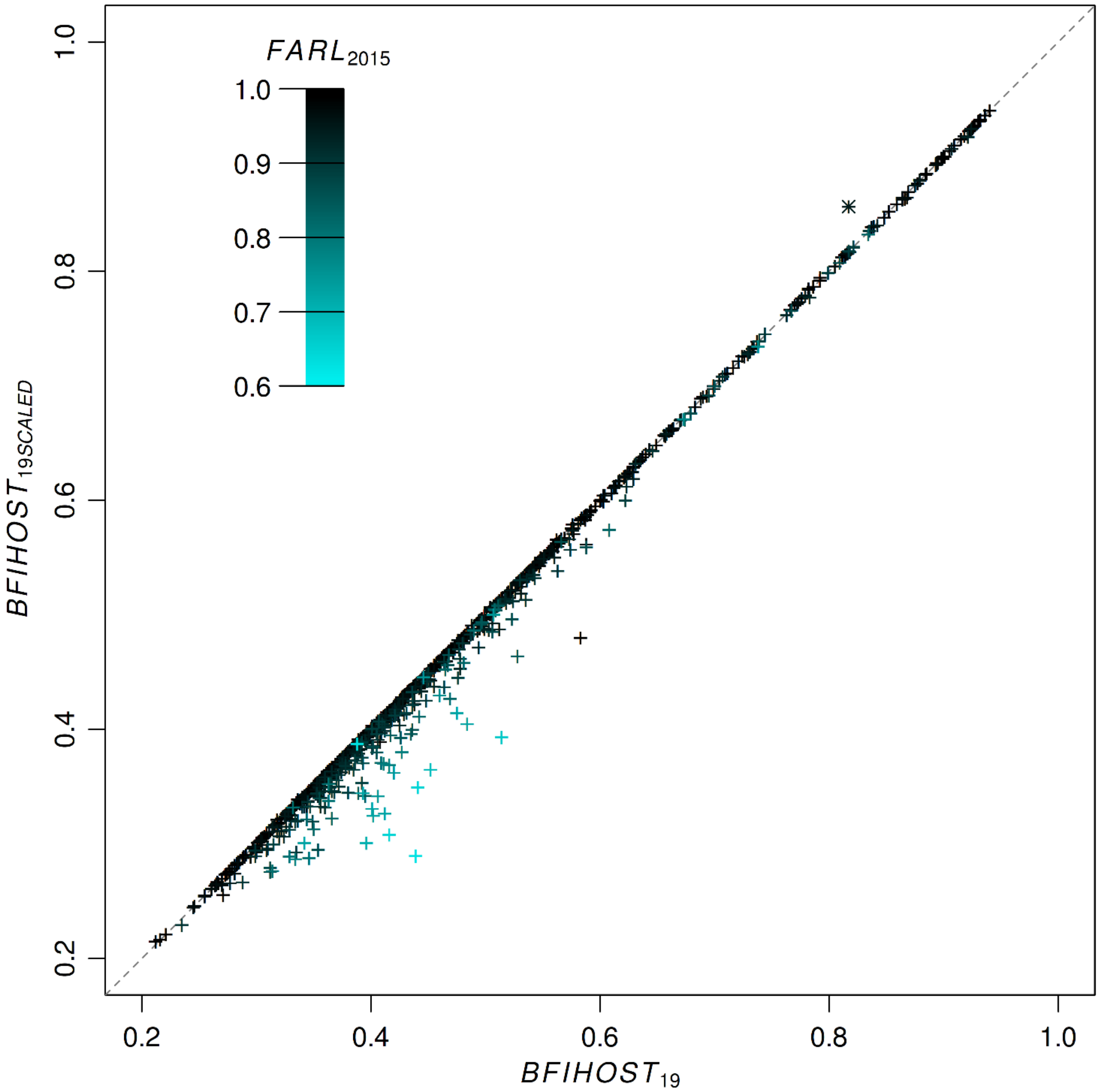

Figure 4 ***BFIHOST*$_{19SCALED}$ vs *BFIHOST*$_{19}$ for 626 essentially rural catchments. Catchment *FARL* indicated by colour.**

2.3 *SAAR*

The *SAAR* descriptor used in the FEH methods maps the mean annual rainfall over the period 1961-1990 (Spackman, 1993) and is an update of the Flood Studies Report (FSR) descriptor/map *AAR*, which maps the same for the period 1941-1970 (this descriptor is now called $SAAR_{4170}$ e.g. on the FEH Web Service). Independently of UKCEH and/or NERC/UKRI, the Met Office produced a map of 1991-2020 mean annual rainfall for the whole UK, including Northern Ireland with its coordinates transformed to the British National Grid (Met Office, 2023). This map is gridded at 1-km resolution but is offset from the *SAAR* grid by 500 metres in both horizontal and vertical axes. It is, however, aligned with both the *BFIHOST* and $BFIHOST_{19}$ grids.

A spatially extended version of these data, provided directly by the Met Office, were used by UKCEH to produce a new catchment descriptor, $SAAR_{9120}$, including at 50-metre resolution along the UK river network. Some modifications to the data were required: first, the data in Northern Ireland were transformed from the British National Grid (EPSG:27700) to the Irish National Grid (EPSG:29903), then 1991-2020 mean annual rainfall from CEH-GEAR (Keller *et al.*, 2015) were used to infill very small parts of the headwaters of the Rivers Erne and Finn.

$SAAR_{9120}$ values are compared to *SAAR* values for all 884 catchments in Figure 5. Per-catchment, $SAAR_{9120}$ takes values from 93.4% to 128.7% of *SAAR*, the mean and median being +8.0% and +7.6% respectively. This is partly but not completely due to methodological differences between the NRFA and Met Office datasets. For all but 14 catchments, $SAAR_{9120}$ is higher than *SAAR*. The greatest proportional increases from *SAAR* to $SAAR_{9120}$ are in the Scottish Lowlands, central and south Wales, Devon and Dorset, and the far north of England. The 10 catchments where $SAAR_{9120}$ is greatest as a proportion of *SAAR* all have relatively high mean altitudes (189-431 m), eight are in or near the Scottish Borders, one is in Bannau Brycheiniog and one is in North Yorkshire.

Figure 6 plots $SAAR_{9120}$ as a proportion of the 1961-90 mean annual rainfall derived from the same HadUK-Grid dataset. This is used in place of *SAAR*, as the *SAAR* grid is offset from $SAAR_{9120}$ by 500 metres, while the 1961-90 mean annual rainfall used to generate Figure 6 is aligned exactly. The 1961-90 mean annual rainfall is extremely similar to *SAAR*, although spatial variations in the *SAAR* descriptor are smoother. (See Appendices for more details)

This study also uses another *SAAR* descriptor for calibration only: $SAAR_{POR}$. This is the average annual rainfall for each catchment over the time period that has valid annual maximum flows. The intention behind this decision is that:

1. models are developed using the average annual rainfall that is most representative of the period over which AMAX were collected,
2. hydrological studies use average annual rainfall values that are more representative of current conditions.

Differences between $SAAR_{6190}$ and $SAAR_{9120}$ are explained in the catchment descriptor report but are due to: raingauge density and location changes, changes in Met Office programs/coding languages used, changes in Met Office interpolation methods, and most importantly changes due to climate and anthropogenic changes.

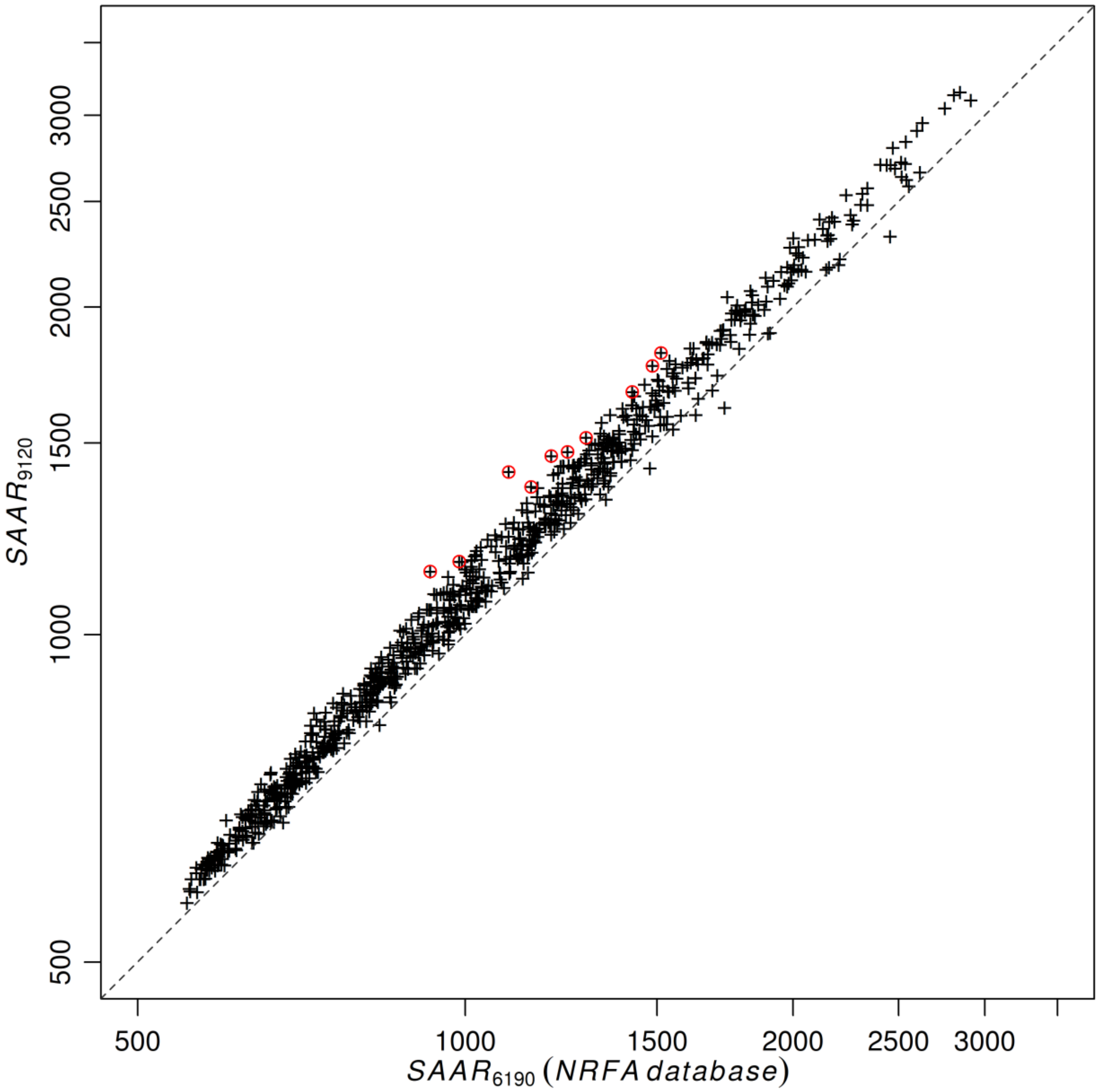

Figure 5 **$SAAR_{9120}$ vs *SAAR* for all 884 catchments. Ten largest percentage increases circled.**

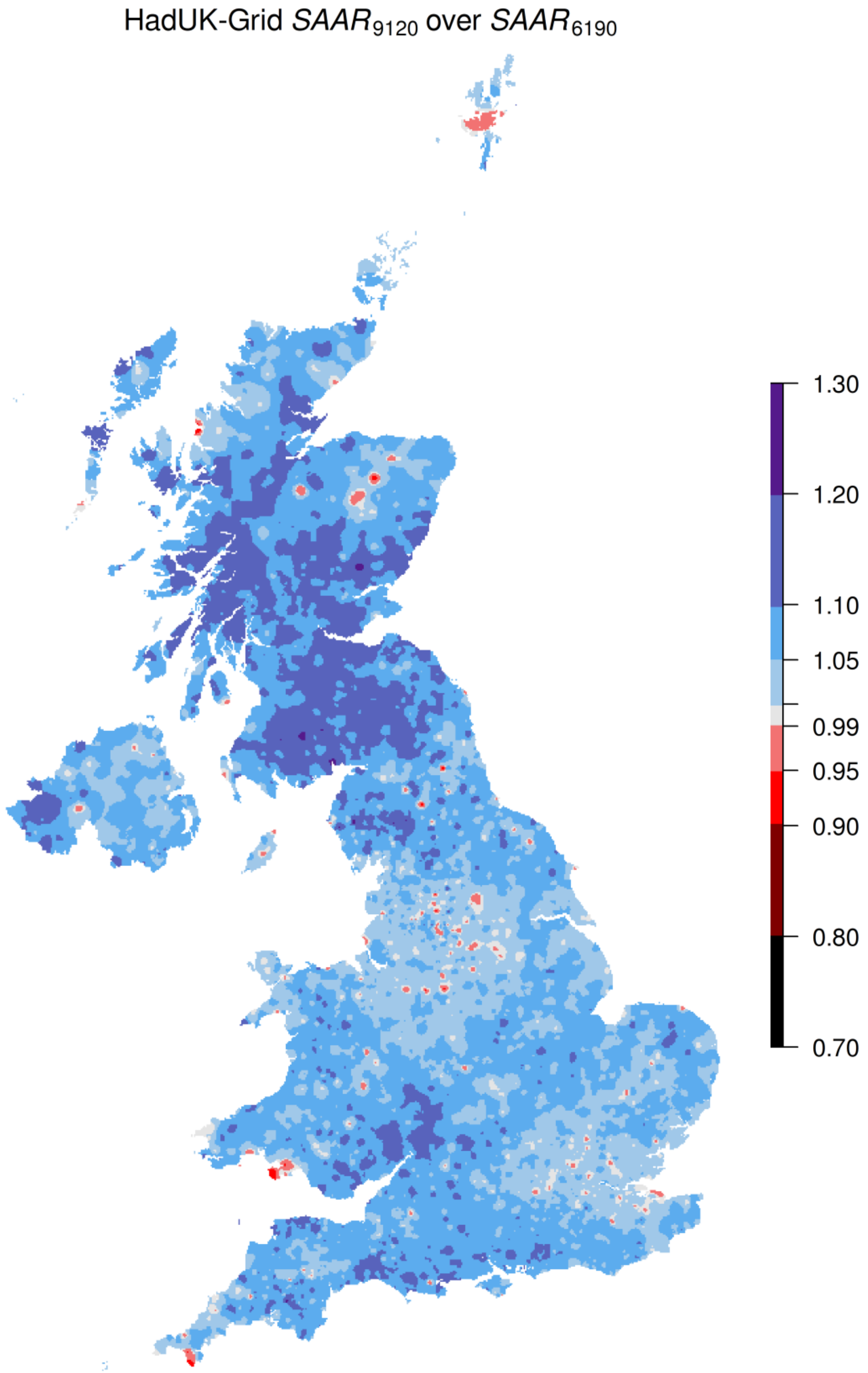

Figure 6 Map of $SAAR_{9120}$ as a proportion of 1961-90 mean annual rainfall, both derived from HadUK-Grid.

2.4 *FARL*

The *FARL* descriptor quantifies flood attenuation in a catchment due to the effects of reservoirs and lakes. It does not quantify the effects of floodplains – *FPEXT* does, as it quantifies only the proportion of a catchment inundated by a 100-year flood.

The original *FARL* descriptor identifies lakes and reservoirs from the IHDTM, meaning that lakes are always represented as collections of 50-metre squares. This results in inaccurate representations of lakes, although the inaccuracy is less for the larger lakes that have the greatest effect on *FARL*. More importantly, the 50-metre resolution of the IHDTM risks incorrectly joining lakes that are separated by less than 50 metres, and assigning lakes as “on-line” if they are near, but not connected to, a river.

Vector representations of lakes solve all three problems associated with the existing raster dataset. A comparison of alternative vector datasets suggested that UKCEH Land Cover Map 2015 (LCM2015: Rowland *et al.*, 2017) vector data be used to represent lakes. The current *FARL* calculation method was use with this to generate $FARL_{2015}$.

Figure 7 compares $FARL_{2015}$ to the existing *FARL* descriptor. There is a close, but not exact, match for most catchments. This is expected as most lake areas should differ between the raster *FARL* and vector $FARL_{2015}$ datasets, and could differ in either direction. Most catchments in which *FARL* and $FARL_{2015}$ most differ are those in which sand/gravel extraction areas are identified as lakes in LCM2015 but not the IHDTM.

It is noted that Moriston at Invermoriston (NRFA station 6003) uses the NRFA Peak Flow version 13 values for *FARL* and $FARL_{2015}$, despite all other catchments using NRFA Peak Flows version 12.1 values. This is because the error in the version 12.1 value was discovered during this study.

For NRFA catchments that use a “pre-reservoir” value of *FARL* – because of pre-reservoir peak flow data – there is an equivalent estimate of $FARL_{2015}$ that models the catchment with the reservoir removed. Both this study and the NRFA Peak Flow dataset use pre-reservoir values for catchments where the AMAX are recorded pre-reservoir. However, post-reservoir values are available from the FEH Web Service, which includes all on-line water bodies.

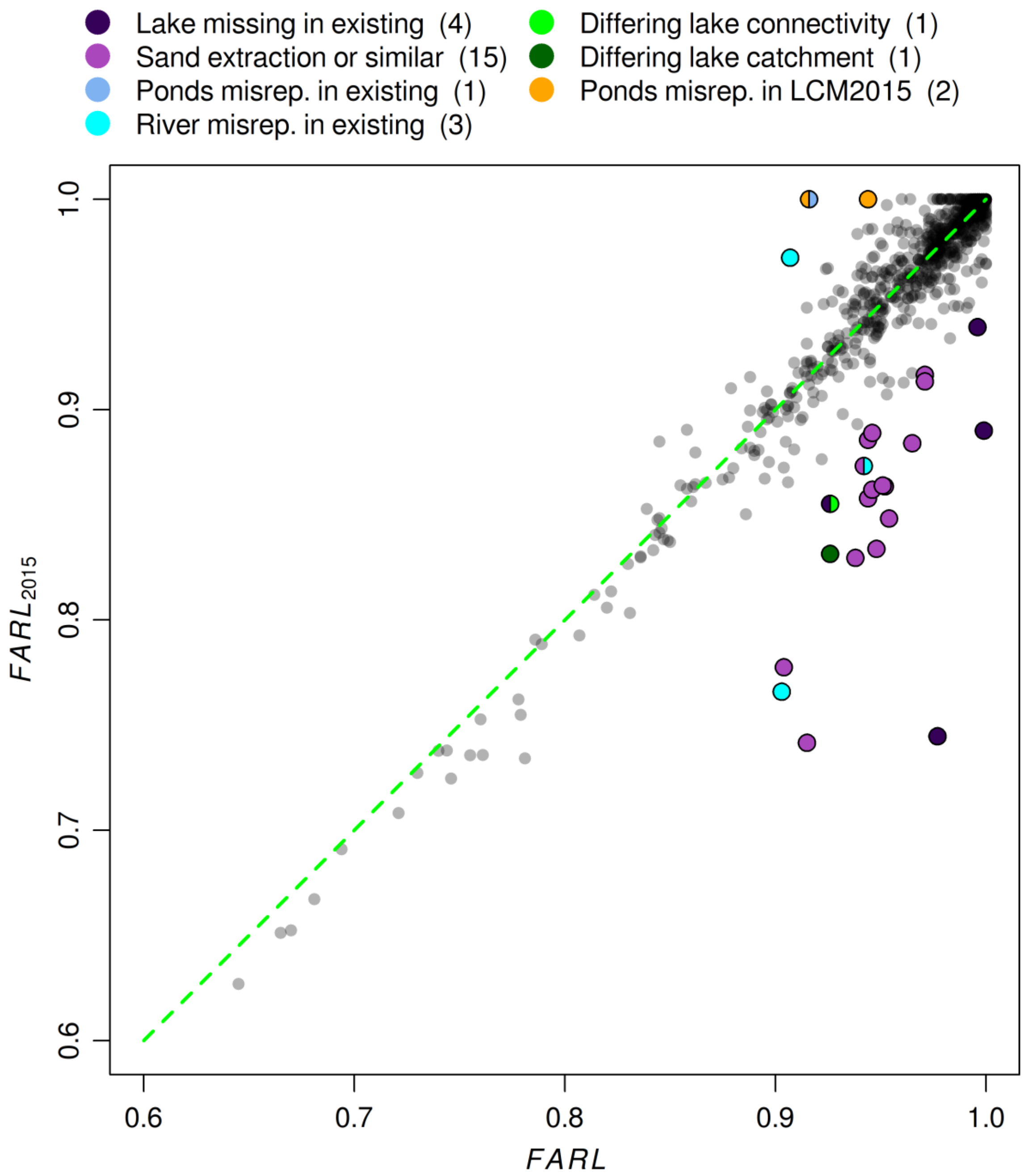

Figure 7 ***$FARL_{2015}$* vs *FARL* for all 884 catchments. Reasons for large differences highlighted.**

2.5 *URBEXT*

$URBEXT_{1990}$ was introduced in the Flood Estimation Handbook to quantify the fraction of a catchment that is urbanized, based on mapping performed in Great Britain in 1988-90 (Fuller *et al.,* 1994) and Northern Ireland in 1989-90 (Brand & Mitchell, 1993). The 2008 FEH statistical method uses $URBEXT_{2000}$ (Bayliss *et al.,* 2006), which is based on UKCEH Land Cover Map 2000 (Fuller *et al.,* 2002). $URBEXT_{2000}$ is a weighted mean of the urban, suburban and inland bare ground fractions within a catchment, whereas $URBEXT_{1990}$ considers urban and suburban fractions only. Land Cover Map (LCM) was updated in 2007, 2015, and every year from 2017 onwards. This allows *URBEXT* values to be generated for several snapshots in time, using LCM as the source for urban and suburban coverage. However, the Land Cover Maps do not share consistent methodologies and hence are not fully comparable.

The FEH statistical method assumes stationarity in flow records, so it is typically reasonable to assume that the level of urbanization at the temporal mid-point of the flow record is the most representative single value for the whole flow record. The urban expansion factor (*UEF*) was developed by Bayliss *et al.* (2006) to estimate *URBEXT* for any year from 1945 to 1999, given only $URBEXT_{2000}$, and is represented by the equation $UEF = 0.7851 + 0.2124\tan^{-1}((Year - 1967.5)/20.32)$. This study extends the applicability of *UEF* to 2015 by calculating Great Britain-wide values for $URBEXT_{1990}$ and $URBEXT_{2015}$ from LCM1990 and LCM2015. The updated *UEF* (Figure 8) is given by:

$$0.7492 + 0.3927\tan^{-1}\left(\frac{Year - 1978.82}{48.7345}\right)$$

In this study, *UEF* is combined with *URBEXT* estimated from LCM2015 ($URBEXT_{2015}$) to generate $URBEXT_M$: a temporary/study-only descriptor that quantifies *URBEXT* at the temporal midpoint of each catchment's AMAX record. Catchments in this study are considered rural or urban according to thresholds in $URBEXT_M$ values. Similarly to $SAAR_{POR}$, $URBEXT_M$ is an approximation of the "typical" conditions during the period over which calibration AMAX data were collected. Hence, $URBEXT_M$ is considered more suitable for model calibration and $URBEXT_{2015}$, representing urbanization in the year 2015, is considered more suitable for hydrological studies that are intended to model more recent catchment conditions. LCM2015 is used in preference to more recent LCMs, as 2015 is the most recent LCM with both automated and manual quality control – more recent LCMs have automated quality control only.

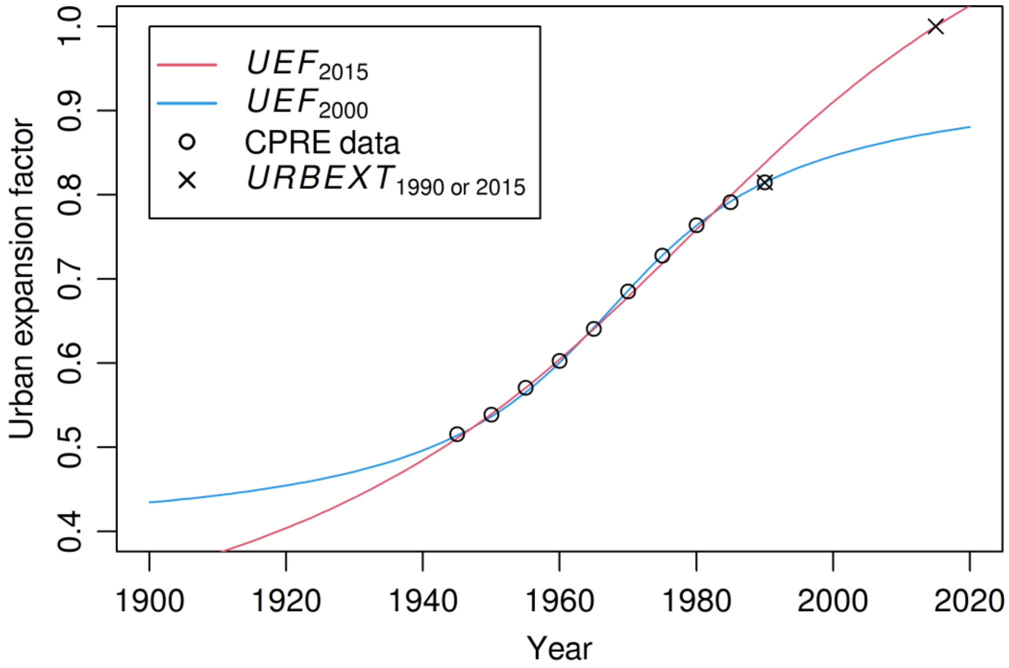

Figure 8 **Urban Expansion Factor (*UEF*). UEF_{2015} is derived for use in this study, UEF_{2000} is from Bayliss *et al.* (2006), rescaled to match rescaled CPRE (1993) data.**

3. *QMED* model

The "2008" *QMED* model was introduced by Kjeldsen *et al.* (2008), replacing the original FEH model introduced nine years earlier (Robson & Reed, 1999). The model implemented in WINFAP 5 modifies the original slightly, replacing *BFIHOST* in the original equation with $BFIHOST_{19}$; for simplicity, this update is referred to as the 2008 method throughout this report. The 2008 *QMED* model is presented below:

$$QMED = 8.3062AREA^{0.8510}0.1536^{\left(\frac{1000}{SAAR}\right)}FARL^{3.4451}0.0460^{BFIHOST_{19}^{\ 2}}$$

This model is intended to estimate the median annual maximum flow in essentially rural catchments only. Hence, for this stage of the study, only catchments marked as suitable for *QMED* estimation and with an $URBEXT_M$ value of 0.03 or less were used in this recalibration. Table 1 compares the dataset of essentially rural catchments used in this recalibration against that used in the original calibration, showing the much greater quantity of AMAX data available now.

Table 1 Summary of "essentially rural catchment" AMAX data sets available to Kjeldsen *et al.* (2008) and this study.

	HiFlows-UK 1.1	NRFA Peak Flow 12.1	Increase (%)
Number of suitable gauges	602	626	+4.0%
Shortest record length	4	6	+50.0%
Longest record length	117	136	+16.2%
Mean record length	32.7	45.3	+38.5%
Number of AMAX events	19679	28339	+44.0%
Final (water) year	2002-03	2021-22	+19 years
Gauge-pairs with shared record length ≥ 40 years	11062	87068	+687%

3.1 *QMED* model method

The development of the *QMED* model requires several intermediate steps:

1. Estimate target *QMED* values from gauged AMAX data,
2. Estimate sampling uncertainty in target *QMED* values, comprising:
    a. Estimate correlation between ln(*QMED*) and inter-catchment centroid distance for pairs of gauges,

   b. Estimate scale parameter of at-site generalized logistic (GLO) distributions as a function of catchment descriptors,
3. Estimate final model and model error simultaneously.

These stages are detailed in Kjeldsen *et al.* (2008), so this report only highlights differences from that method i.e. anything that is not mentioned should be assumed unchanged.

Estimate target *QMED*

For consistency with all of the other stages of method development, AMAX series in this study were log-transformed before estimating target *QMED*, which is still the median of the ordered series.

Estimate correlation between ln(*QMED*) and inter-catchment centroid distance

The *QMED* model accounts for inter-site correlations in *QMED*. To maximize the number of station pairings, gauges in Northern Ireland were projected from the Irish (EPSG:29903) to the British (EPSG:27700) National Grid. Given the extra 19 years of flow data available to this study, far more gauges than before have 40+ years of common record: the total number of gauge-pairs is now 87068, compared to 11062 previously. To better characterize the correlation between pairs of ln(*QMED*) values, the number of bootstrap samples for each gauge-pair was increased from 1000 to 10000.

There are some general relations between ln(*QMED*) and centroid-centroid distance, but any potential trend is very scattered, and the relationship is best visualized as existing between two porous boundaries, rather than around one line (Figure 9). This is no different than what was observed and published by Kjeldsen *et al.* (2008).

The fitted double-exponential relationship between correlation in paired ln(*QMED*) values and catchment centroid-centroid distance, shown as a red line on Figure 9, is:

$$r_{\varepsilon,ij} = 0.3098e^{-0.0046d_{ij}} + (1 - 0.3098)e^{-0.0891d_{ij}}$$

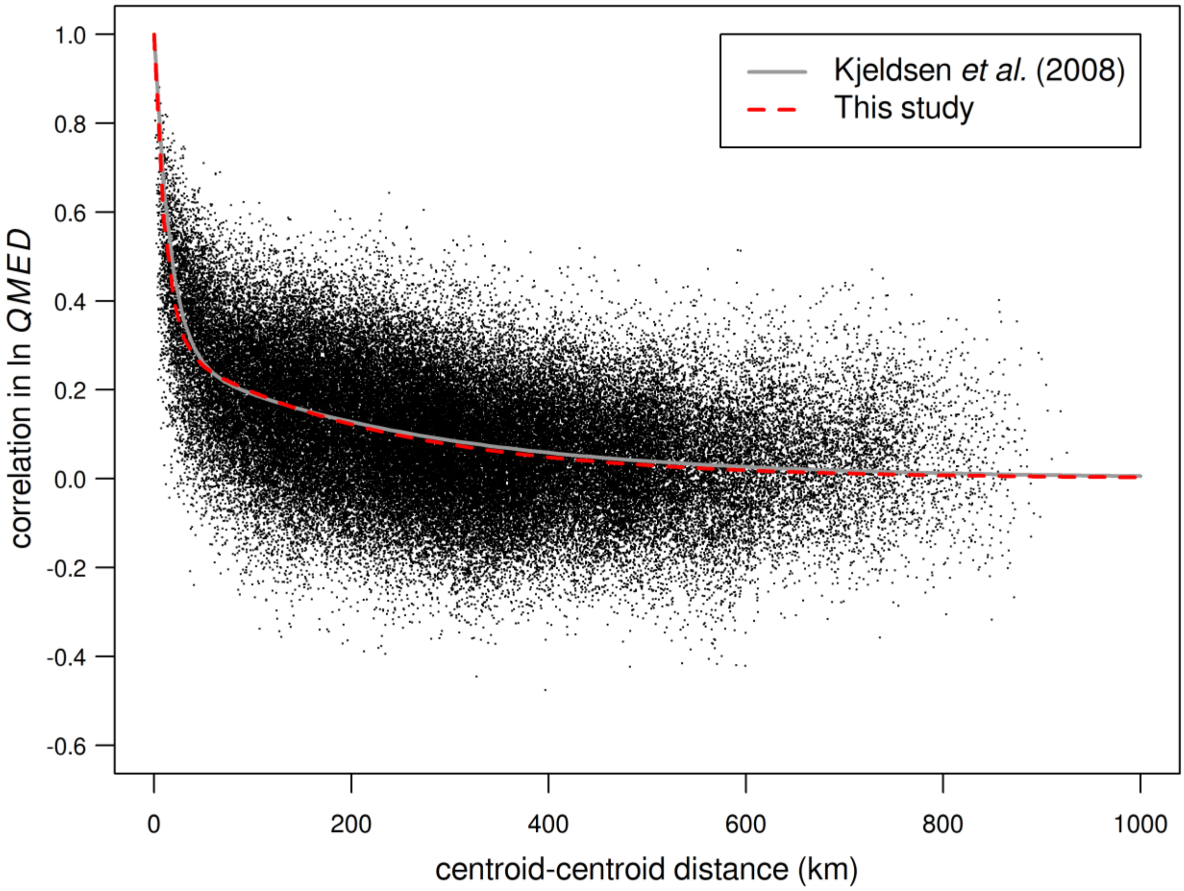

Figure 9 Correlation in paired ln(*QMED*) values by catchment centroid-centroid distance.

The fitted values for ϕ_1, ϕ_2 and ϕ_3 are different from those in Kjeldsen *et al.* (2008), but the overall relationship is broadly similar, visually. A very slightly lower correlation between ln(*QMED*) pairs is implied in this study compared to the 2008 model, except for catchment centroid-centroid distances between 59 and 140 km.

Estimate scale parameter (β) of generalized logistic distribution

Uncertainty in *QMED* is related to record length and β (the scale parameter of a generalized logistic distribution fitted to the AMAX series). This study performed a stepwise linear regression to identify which descriptors should be used to model β. Based on this regression, a four-parameter model for β as a function of ln(*DPLBAR*), 1000/$SAAR_{POR}$, ln($FARL_{2015}$) and $FPEXT^{0.5}$ was chosen (Table 2).

Table 2 Summary statistics for four-parameter model of β, used to model sampling error of variance in ln(*QMED*).

Variable	Coefficient	Standard error	*t*-value	*p*-value
Intercept	0.166873	0.011484	14.532	2.14×10^{-41}
ln(*DPLBAR*)	−0.019588	0.003417	−5.733	1.54×10^{-8}
$1000/SAAR_{POR}$	0.180575	0.008771	20.587	3.47×10^{-72}
$\ln(FARL_{2015})$	−0.199675	0.039335	−5.076	5.09×10^{-7}
$FPEXT^{0.5}$	−0.328186	0.038983	−8.419	2.63×10^{-16}

Variance of errors = 0.003629 on 621 degrees of freedom, R^2 = 0.4422

This model explains over 44% of the variance in β, a significant improvement over the 28% explained by the model in Kjeldsen *et al.* (2008). A better representation of β should result in a better representation of sampling error, hence a better calibration of the ln(*QMED*) model.

Estimate final model and model error simultaneously

This study used the same negative log-likelihood function as Kjeldsen *et al.* (2008):

$$-\ln(L_k) = \frac{1}{2}\ln\left[\det\left(\sigma_\eta^2 G\right)\right] + \frac{1}{2}(y - X\theta)^T\left(\sigma_\eta^2 G\right)^{-1}(y - X\theta)$$

The optimized model fit was found when $-\ln(L_k)$ was minimized.

3.2 Choice of descriptors for ln(*QMED*) model

The 2008 *QMED* model recalibration uses four descriptors, three of which have since been updated: ln(*AREA*), 1000/*SAAR*, ln(*FARL*) and $BFIHOST^2$. Given the very high R^2 values achieved in the original model, it is unlikely that any other four descriptors could explain much more variance in ln(*QMED*) than these, or their updated versions. A stepwise linear regression was performed to verify that there were no better descriptors.

The best four-descriptor model for ln(*QMED*) was found to include ln(*AREA*), $1000/SAAR_{POR}$, $FARL_{2015}$ and $BFIHOST_{19SCALED}^2$: updated versions of the same four descriptors currently in use, using the same transformations in three cases. After further study, it was decided to use a natural logarithm transformation for $FARL_{2015}$. This choice reduced model R^2 by 0.0001, but massively improved the *p*-value of the model's intercept coefficient, from approximately 3×10^{-7} to 10^{-65}, and maintained the model structure developed by Kjeldsen *et al.* (2008).

3.3 Results

Table 3 presents summary statistics for the recalibrated ln(*QMED*) model. The model's *fse* and R^2 are 1.4307 and 0.948 respectively. Kjeldsen *et al.* (2008) reported an *fse* and R^2 of 1.4313 and 0.945 for their ln(*QMED*) model. This indicates that the recalibrated model explains slightly more variance in ln(*QMED*), with a slightly smaller model structural error.

Table 3 Summary statistics for recalibrated ln(*QMED*) model.

Variable	Coefficient	Standard error	*t*-value	*p*-value
Intercept	1.920553	0.100058	19.194	8.07×10^{-65}
ln(*AREA*)	0.849896	0.011080	76.708	5.84×10^{-319}
$1000/SAAR_{POR}$	−1.725942	0.089253	−19.338	1.43×10^{-65}
$\ln(FARL_{2015})$	3.044989	0.243965	12.481	4.66×10^{-32}
$BFIHOST_{19SCALED}{}^2$	−3.438080	0.118887	−28.919	3.96×10^{-117}

$\sigma_\eta^2 = 0.1283$ $df = 621$ $R^2 = 0.9483$ $-2\ln(L) = -393.7$

$$QMED = 6.8247 AREA^{0.8499} 0.1780^{\left(\frac{1000}{SAAR_{9120}}\right)} FARL_{2015}{}^{3.0450} 0.0321^{BFIHOST_{19SCALED}{}^2}$$

$$r_{\eta,ij} = 0.4814 e^{-0.0333 d_{ij}} + (1 - 0.4814) e^{-0.4610 d_{ij}}$$

To emphasize that both the 2025 and 2008 *QMED* equations should never be used without donor transfer, no comparisons between the 2025 and 2008 *QMED* methods are presented until Section 4, on donor transfer.

3.4 Sensitivity study

In order to identify possible "outlier" catchments with unusual influence over the recalibration, a sensitivity study was performed with 100 alternative calibration datasets, each excluding 125 catchments (20% of the total), selected at random.

Figure 10 presents the optimized coefficients and R^2 for the 100 models calibrated to 20%-reduced subsets of the full calibration dataset. For all coefficients, the mean of optimized values is more than 20 times the standard deviation. The stability and small range of the fitted values suggests that no catchment influences the model parameter fitting much more than any other.

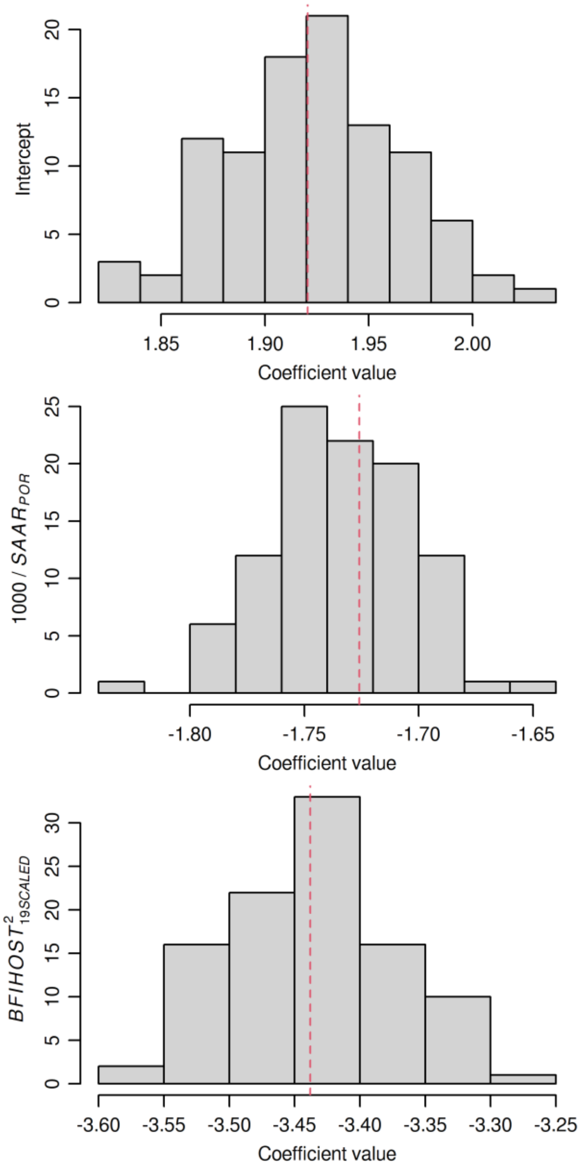

Figure 10 **Optimized coefficients and R^2 for ln(*QMED*) model when 125 catchments are randomly excluded, 100 times. Vertical red lines show values when no catchments are excluded.**

4. Donor transfer

Donor transfer is a part of the FEH statistical methodology that is intended to improve the modelled estimate of *QMED* by transferring information about the model error from suitable donors. Because it is only the information on the residuals that is used, not information on hydrological processes, there is no requirement for a donor catchment to be hydrologically similar in any way to the catchment of interest. Vesuviano *et al.* (2024) show no relationship between model error and catchment descriptors, while Kjeldsen & Jones (2007) and Kjeldsen *et al.* (2008) demonstrated, on two separate datasets, with two separate ln(*QMED*) models, that selecting donors based on similarity in catchment descriptors (ln(*AREA*), ln(*SAAR*) and *BFIHOST*) actually worsens *QMED* estimates.

Some important notes on donor transfer are presented below:

- The ln(*QMED*) model before donor transfer explains 94.8% of the variance in ln(*QMED*), so the most variance that any donor transfer procedure can explain is 5.2%.
- The structure of the donor transfer equation means that it is necessary to identify a catchment with a similar model error, i.e. “true model estimate of *QMED*” + “model error” − “similar model error” = “true estimate of *QMED* plus smaller model error”.
- Model error is not related to any of the catchment descriptors in the ln(*QMED*) model (Figure 11). If it were, we would not expect the regression lines in Figure 11 to be horizontal.
- Consequently, it is impossible to estimate the sign or magnitude of model error for any particular catchment from the properties of that catchment.
- However, model error *is* correlated with catchment centroid-centroid distance (shown as spatial clustering on Figure 12, and on pages 7-10 of Environment Agency, 2017), so donor transfer based on centroid-centroid distance will on average reduce model error, with the caveat from the first bullet point that the model already explains 94.8% of the variance in ln(*QMED*), so it is impossible for any donor transfer procedure to explain more than an additional 5.2% of the variance in ln(*QMED*).
- Appropriate use of donor transfer requires that the model errors at donor stations are “as expected”, so if the catchment descriptors do not accurately describe the catchment, the model error will not be as expected. This is distinct from rejecting donor catchments because of differences between the donor and target catchments, but does mean that it may be acceptable to reject a donor catchment if its catchment descriptors do not describe the catchment accurately.

The relationship between model error correlation and centroid-centroid distance is shown in Figure 13. It is very similar to Kjeldsen *et al.* (2008) for distances of a few kilometres, then slightly weaker for greater distances.

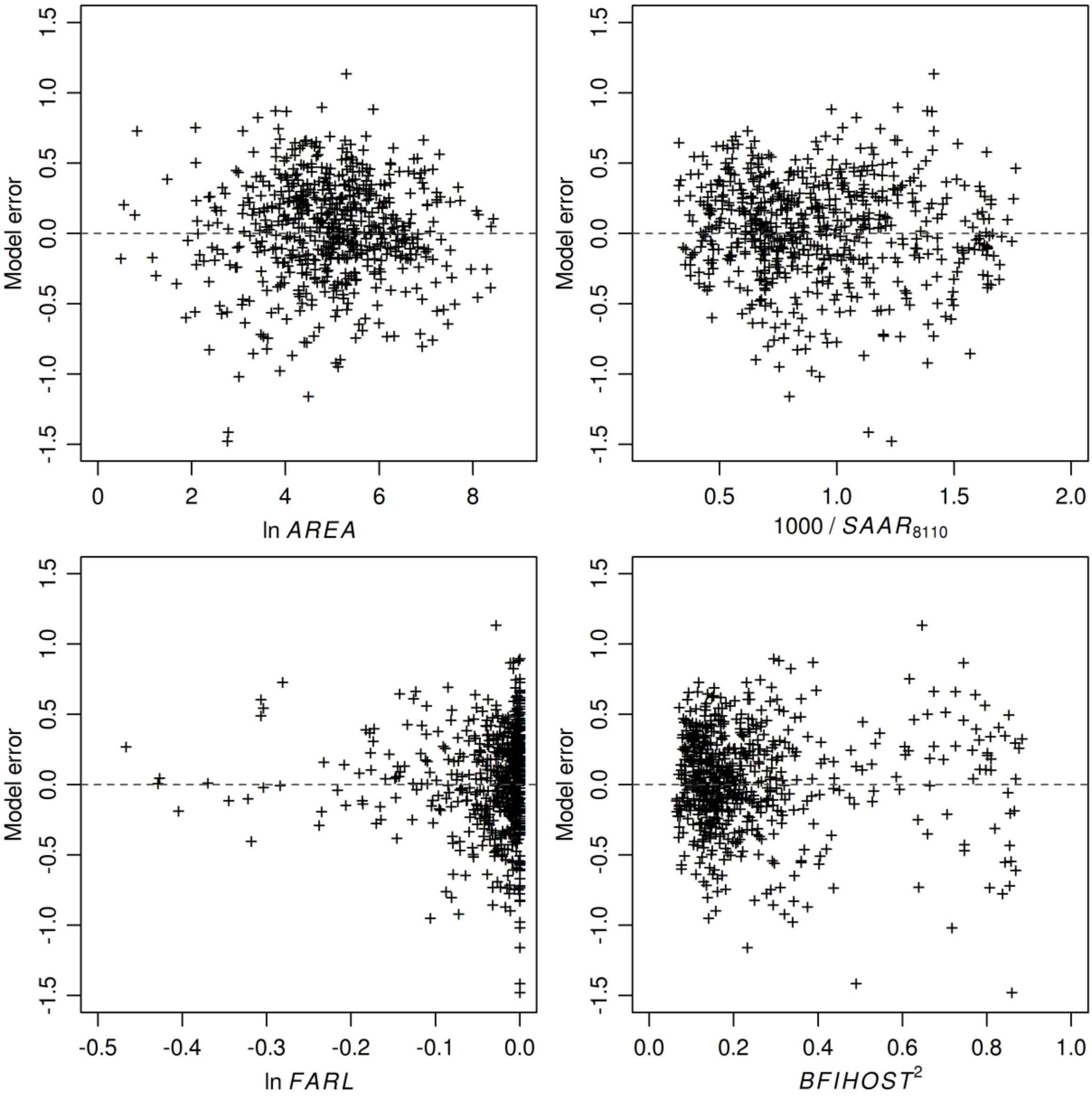

Figure 11 Recalibrated ln(*QMED*) model error vs. catchment descriptors in model.

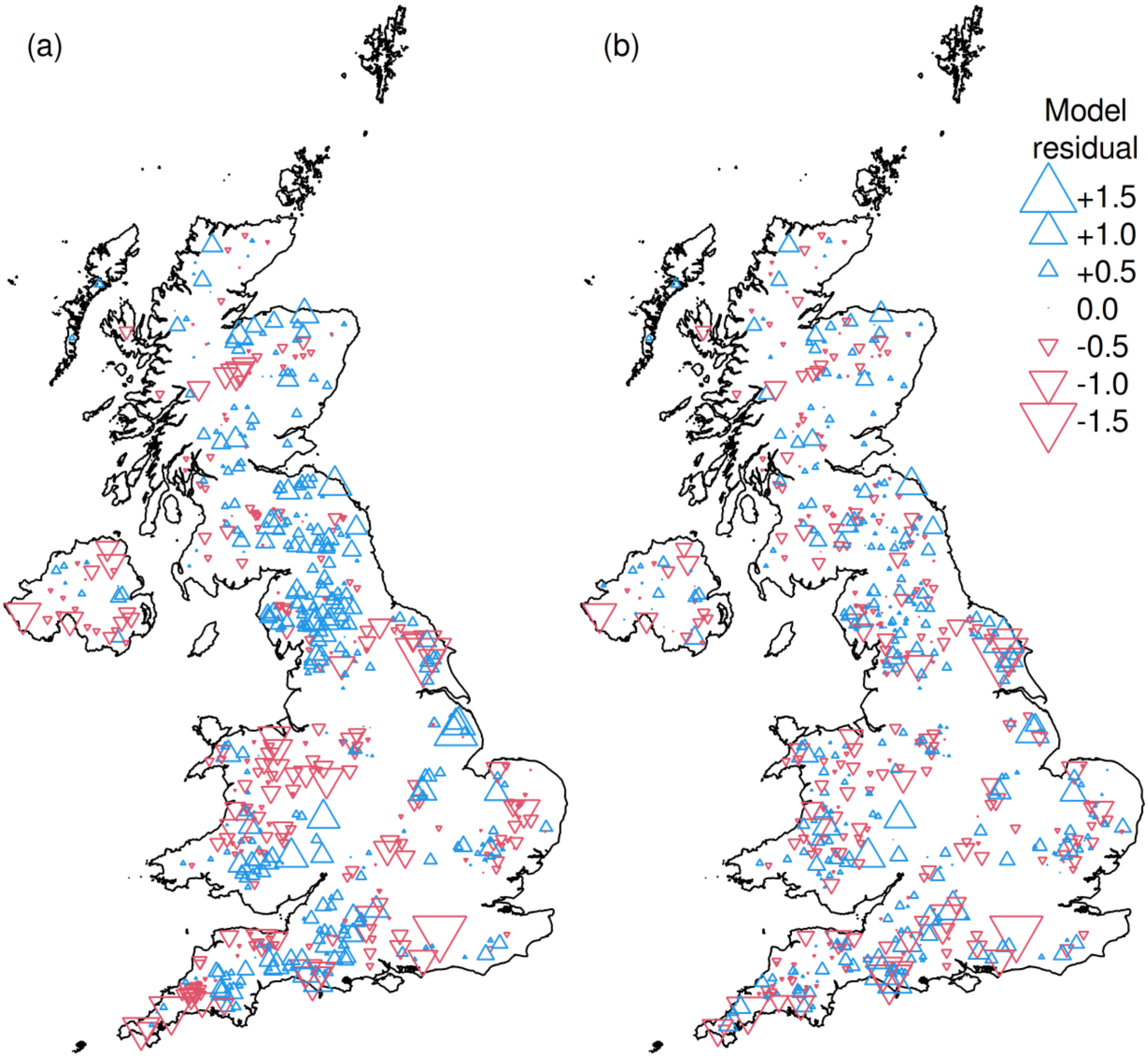

Figure 12 **Spatial variation of recalibrated 2025 ln(*QMED*) model error across UK, before (a) and after (b) donor transfer, on 626 calibration catchments.**

Figure 12 (a) clearly shows spatial clusters of same-sign model errors that do not follow any obvious catchment descriptor (e.g. both north Wales and Cumbria are wet, but one has a cluster of positive model errors and the other a cluster of negative model errors). Using distance to select donors typically results in the selection of donors with the same sign of error as the catchment of interest, leading to a reduction in absolute model error and spatial clustering after donor transfer (Figure 12(b)).

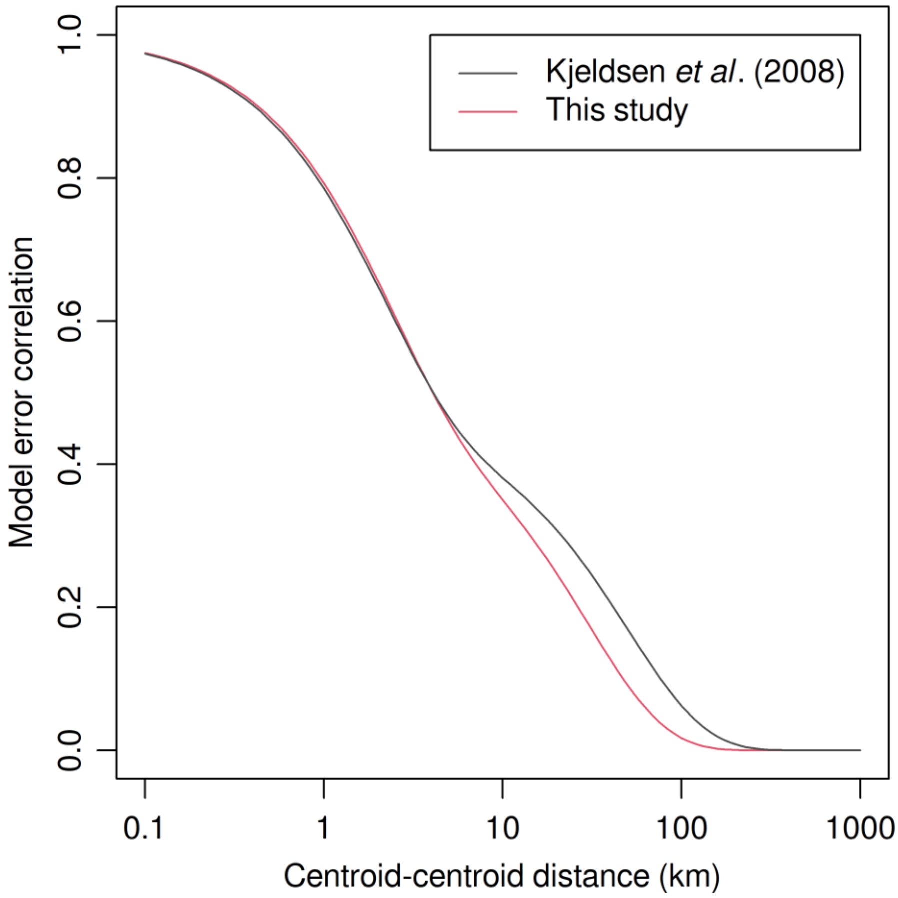

Figure 13 Model error correlation vs. catchment centroid-centroid distance, found by Kjeldsen *et al.* (2008) and in this study.

In practice, use of multiple donors is strongly recommended, since it is possible for the closest single donor to have the opposite sign of error to the catchment of interest, but this is much less likely to be the case for the majority of a group of nearby donor catchments.

In order to investigate the optimal number of stations to use in donor transfer, multiple donor transfer is conducted according to the procedure in Kjeldsen *et al.* (2014), summarized below for one catchment with *n* donors:

- Calculate centroid-centroid distance between the catchment of interest and all gauged catchments (excluding the catchment of interest if it is gauged),
- Sort the gauged catchments by centroid-centroid distance and keep the closest *n* catchments,
- Calculate the model error variance vector $\boldsymbol{b}$ between the gauged catchment and *n* donors,
- Calculate the model error covariance matrix $\boldsymbol{\Omega}$ between each pair of donors,
- Calculate the optimal weight for each donor: $\boldsymbol{\alpha} = \boldsymbol{\Omega}^{-1}\boldsymbol{b}$,

- Calculate the weighted adjustment: $\sum_{i=1}^{n} \boldsymbol{\alpha}\left(\ln QMED_{i,gauged} - \ln QMED_{i,cd}\right)$,
- Add the weighted adjustment to the catchment-descriptor value of ln(*QMED*) for the catchment of interest to obtain the donor-transfer value of ln(*QMED*).

Model error correlation between catchments *i* and *j*, $r_{\eta,ij}$, is calculated using the equation below, derived from calibration of the ln(*QMED*) model:

$$r_{\eta,ij} = 0.4814e^{-0.0333d_{ij}} + (1 - 0.4814)e^{-0.4610d_{ij}}$$

where d_{ij} is the distance between two catchment centroids – either the gauged catchment and a donor, or a donor and another donor. In the case that there is only one donor, **Ω** takes the value of 1. Catchments in Northern Ireland were allowed to take donors from Great Britain, and vice versa, as centroid co-ordinates of gauges in Northern Ireland were transformed from the Irish National Grid (EPSG:29903) to the British National Grid (EPSG:27700).

Figure 14 plots the variation of *fse* and R^2 for the ln(*QMED*) model as *n*, the number of donors, was increased from zero to twenty. This shows a clear improvement in model performance as the number of donors was increased to eight, followed by a plateau, then slight decrease as the number of donors was increased further. The maximum R^2 achieved is 0.9598, hence only 4% of variation in ln(*QMED*) is still unexplained after donor transfer, down from 5.2% before donor transfer.

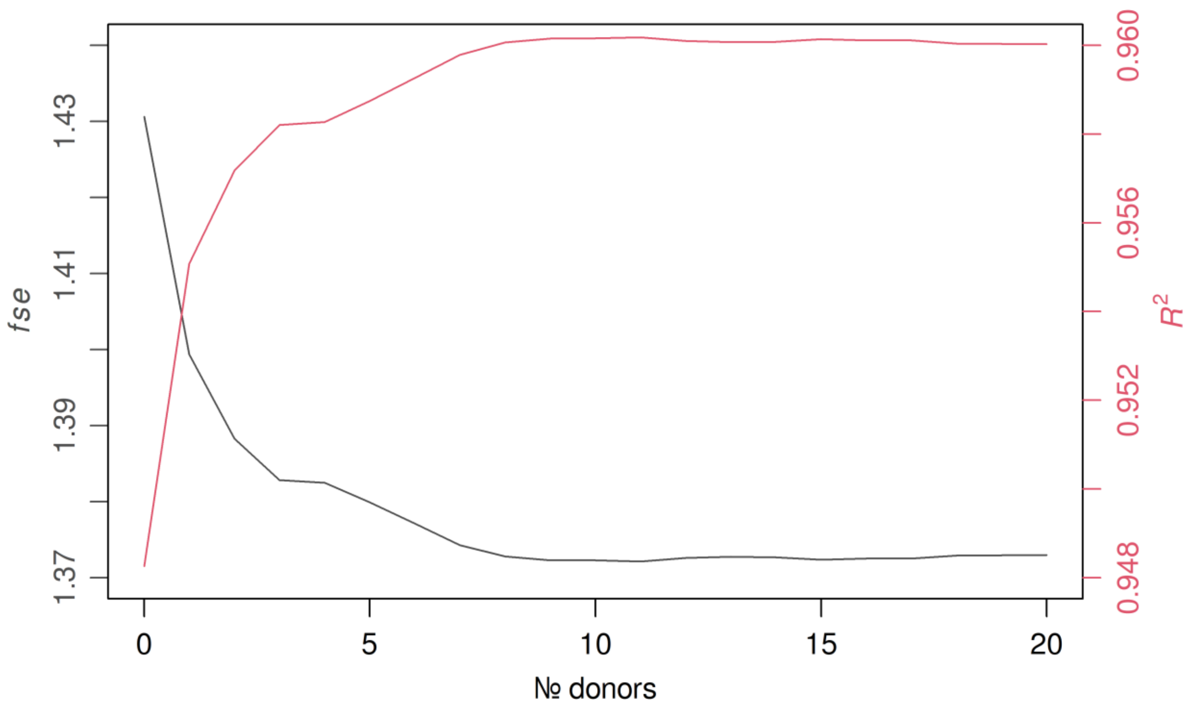

Figure 14 Variation in sample *fse* and R^2 of ln(*QMED*) model as number of donors is increased from zero to twenty.

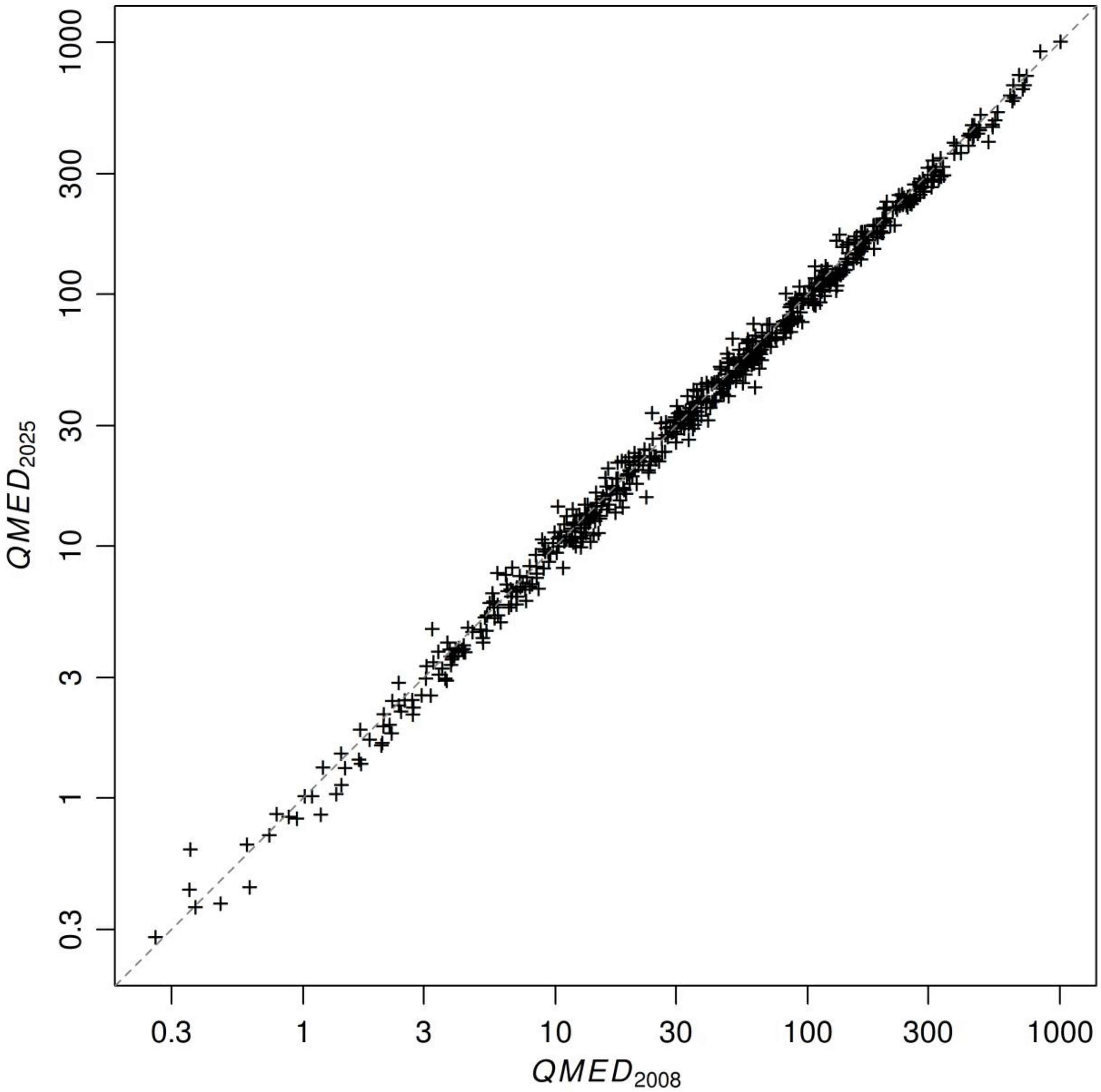

Figure 15 ***QMED*** **after donor transfer, estimated by the 2008 method and the method developed in this study (2025), for 626 essentially rural catchments.**

Figure 15 plots *QMED* after donor transfer, as estimated by the 2008 method (with Kjeldsen *et al.*'s 2014 modification) and the method developed here. For the vast majority of catchments, the change in *QMED* is much less than the uncertainty in either method. Figure 16 plots 2008 *QMED*, recalibrated (2025) *QMED,* and the percentage change from 2008 *QMED* to 2025 *QMED* spatially. While there seems to be a spatial pattern in (the usually small) changes to modelled *QMED*, the fact that the recalibrated model's errors are random with respect to gauged ln(*QMED*) (see Figure 12) suggests that the 2008 equation is systematically in error with respect to the considerably newer gauged estimates of ln(*QMED*) that were used to calibrate the 2025 model.

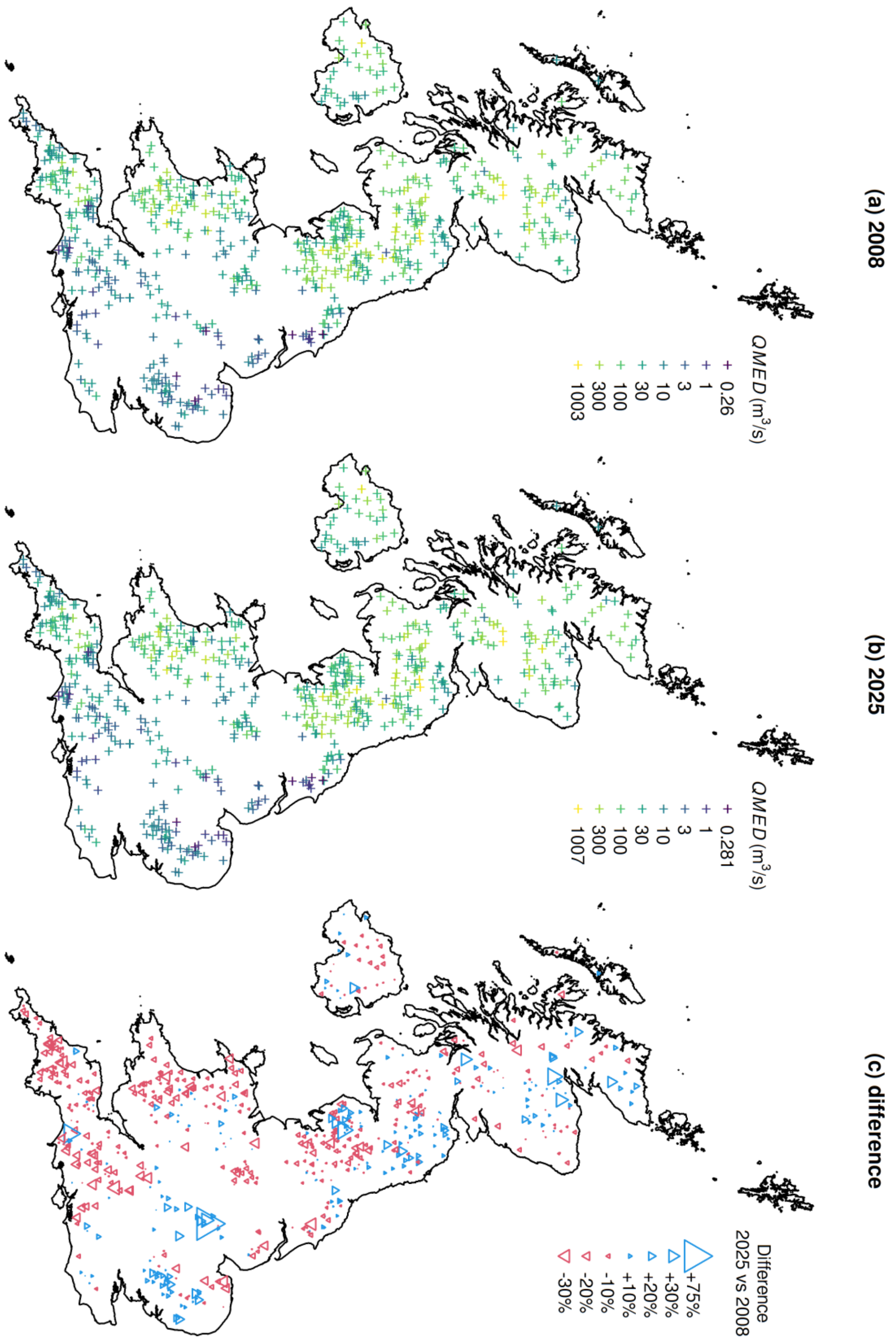

Figure 16 2008 *QMED* (a), recalibrated 2025 *QMED* (b), and change from 2008 to 2025 *QMED*, after donor transfer (626 essentially rural catchments).

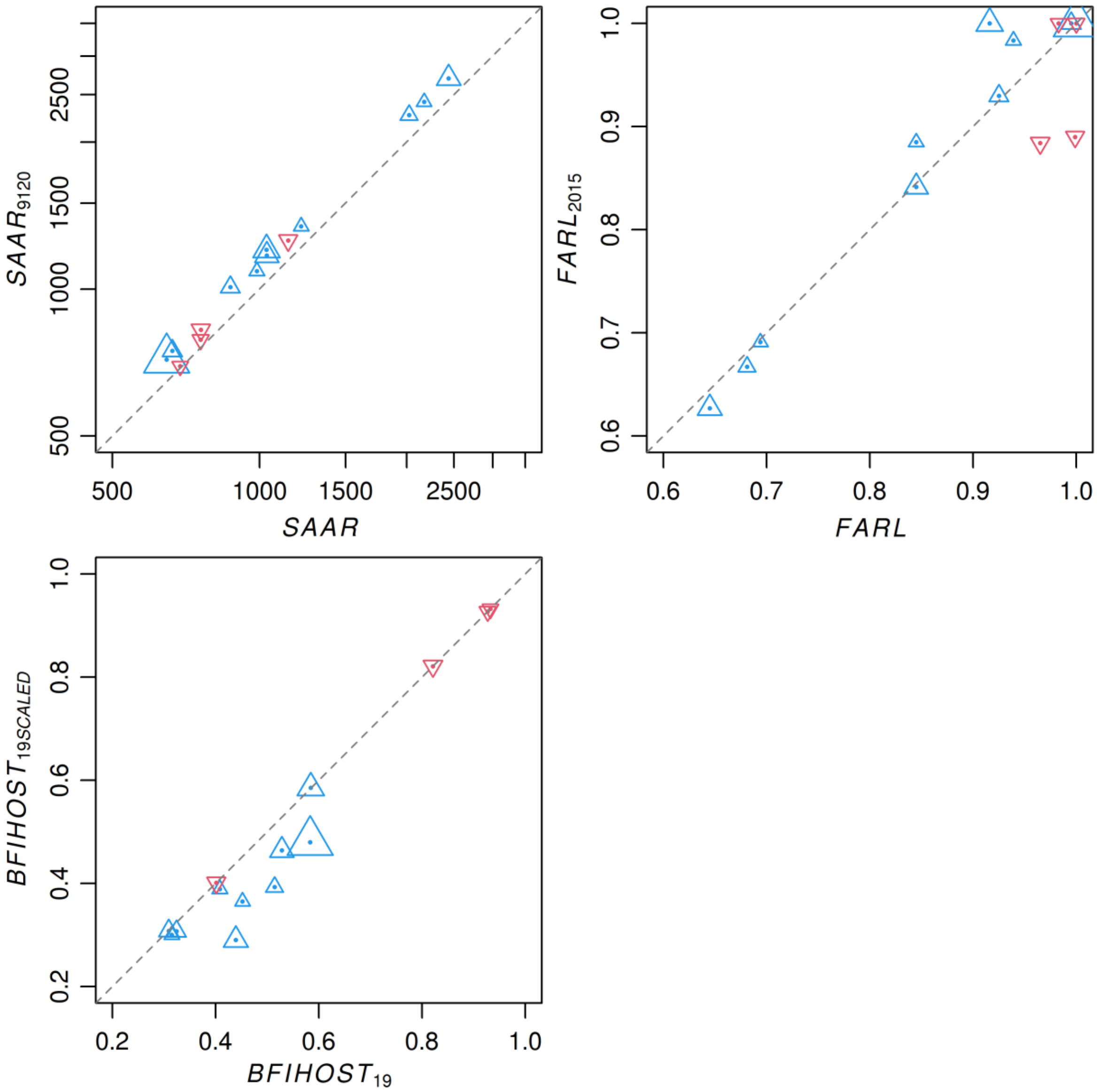

Figure 17 **Catchment descriptors for catchments where 2025 *QMED* after donor transfer is more than 125% (blue triangles) or less than 75% (red triangles) of 2008 *QMED* estimate after donor transfer), for 626 essentially rural catchments.**

There are 10 stations where the 2025 estimate of *QMED* is more than 125% of the 2008 *QMED* estimate, and 4 stations where the 2025 estimate of QMED is less than 75% of the 2008 *QMED* estimate. Figure 17 compares relevant catchment descriptors for these 14 catchments, showing that decreases in *QMED* can be due to decreases in *FARL* whereas increases in *QMED* can often be attributed to decreases in *BFIHOST*. There are two catchments with little change in *SAAR*, *FARL* or *BFIHOST*, but a 25% reduction in *QMED*. In both cases, the reduction derives from donor transfer and the 2025 estimate is closer to the gauged estimate.

When using donor adjustment, centroid-centroid proximity should be the primary metric used, but local knowledge of record quality and length (and mitigating factors such as influential reservoirs) may strongly suggest using other stations. This is valuable secondary information to be made use of when reviewing donor groups.

4.1 Ungauged catchments

The study dataset of 1563 ungauged catchments contains 1267 catchments with both $URBEXT_{2000}$ and $URBEXT_{2015}$ less than 0.03. Of these, there are 116 stations where the 2025 estimate of *QMED* is more than 125% of the 2008 *QMED* estimate, and 15 stations where the 2025 estimate of QMED is less than 75% of the 2008 *QMED* estimate.

Figure 18 and Figure 19 are equivalent to Figure 16 and Figure 17 for these 1267 catchments. As these catchments are ungauged, it is not possible to show the effect of model error clustering before and after donor transfer. Figure 19 shows that increased ungauged *QMED* estimates can result from an increase in rainfall ($SAAR_{9120}$ greater than *SAAR*), decrease in flow attenuation ($FARL_{2015}$ greater than *FARL*) or decrease in baseflow index ($BFIHOST_{19SCALED}$ less than $BFIHOST_{19}$), but that reduced ungauged *QMED* estimates only consistently result from increased flow attenuation ($FARL_{2015}$ less than *FARL*). Urbanization has little effect in this dataset, as all catchments are essentially rural.

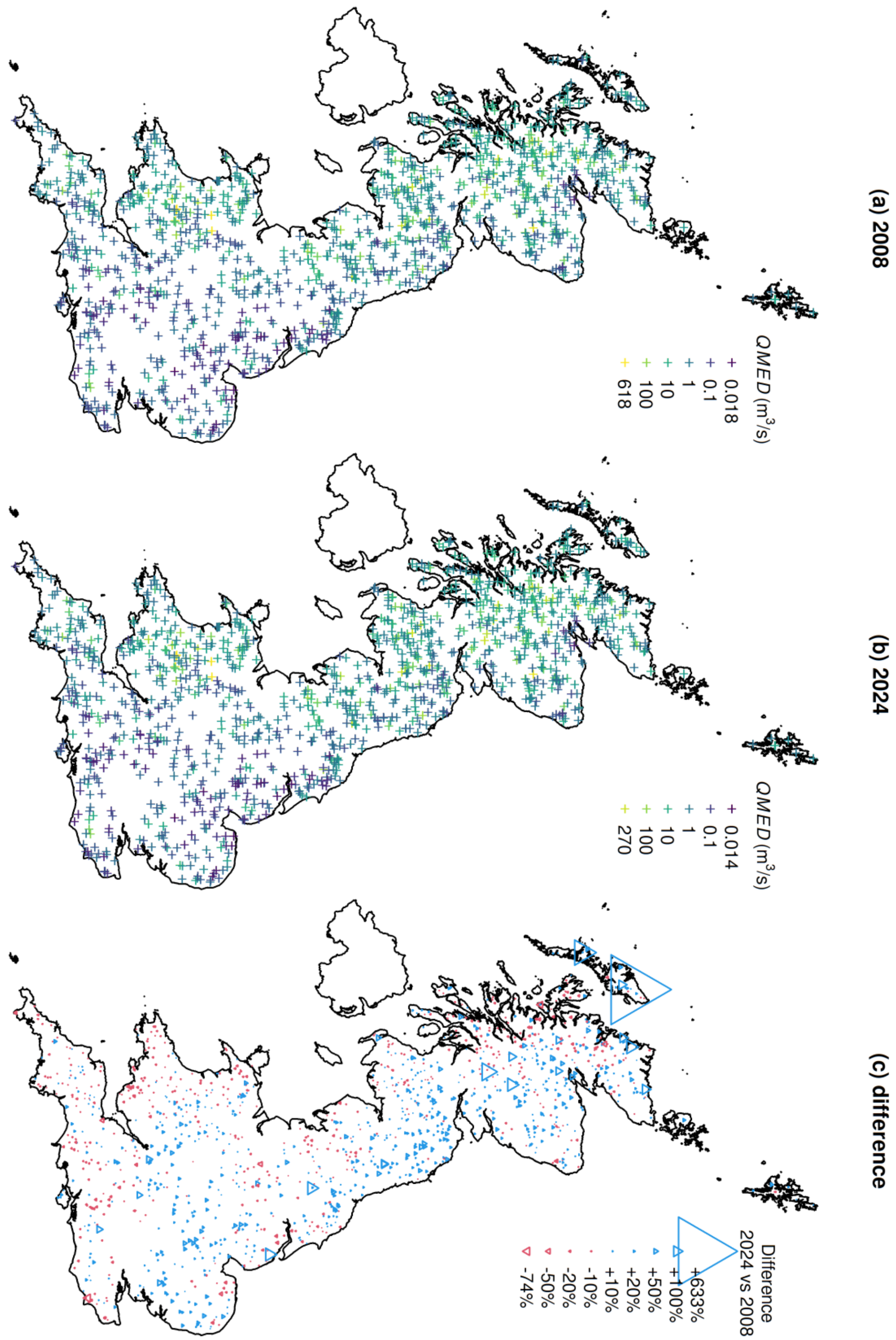

Figure 18 **2008 *QMED* (a), recalibrated 2025 *QMED* (b), and change from 2008 to 2025 *QMED*, after donor transfer (1267 ungauged essentially rural catchments).**

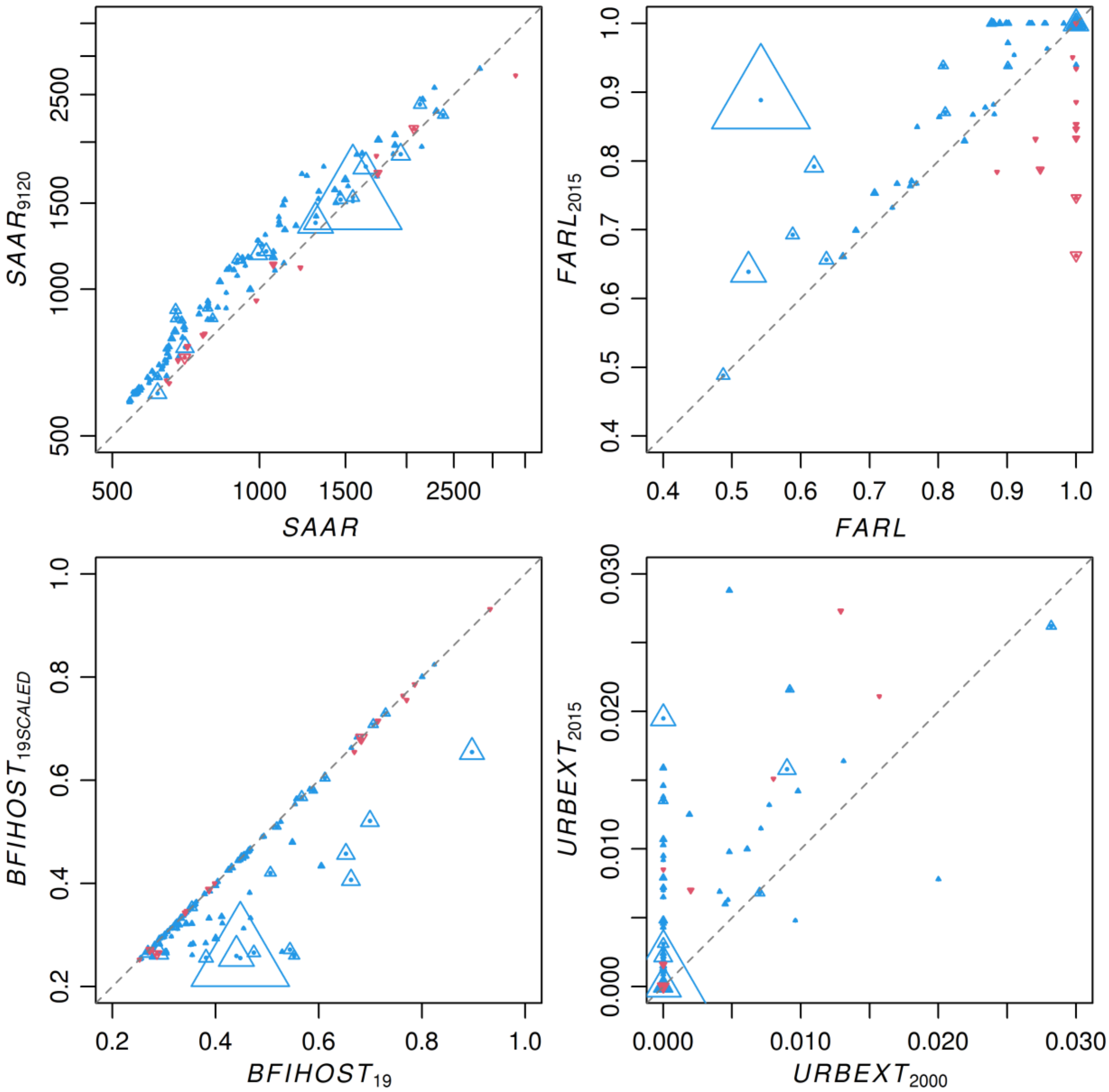

Figure 19 **Catchment descriptors for catchments where *QMED* after donor transfer is more than 125% (blue triangles) or less than 75% (red triangles) of 2008 *QMED* estimate after donor transfer), for 626 essentially rural catchments.**

Figure 19 shows differences due to changes in catchment descriptors in ungauged catchments. As for the gauged locations, the biggest differences are due to differences in *FARL* vs $FARL_{2015}$, and $BFIHOST_{19}$ vs $BFIHOST_{19SCALED}$, both of which may result from the inclusion of new reservoirs or better descriptions of existing ones. These catchments with large changes are also very small catchments (< 25 km^2).

5. Urbanization (*QMED*)

The 2008 *QMED* equation is only intended to estimate the median annual maximum flow in essentially rural catchments ($URBEXT_{2000}$ < 0.03). An urban adjustment factor (*UAF*) is used to adjust these estimates upwards in urbanized catchments, in proportion to $URBEXT_{2000}$ and $BFIHOST_{19}$ value. The *UAF* was calibrated on 200 catchments, separately from the ln(*QMED*) model, a couple of years later (Kjeldsen, 2010). Table 1 compares the subset of rural catchments used by Kjeldsen (2010) to calibrate the 2008 method against what is available for this study.

Table 4 Summary of "urbanized catchment" AMAX data sets available to Kjeldsen *et al.* (2008) and this study*.

	HiFlows-UK 1.1	NRFA Peak Flow 12.1	Increase (%)
Number of suitable gauges	206	258	+25.2%
Shortest record length	3	6	+100%
Longest record length	120	139	+15.8%
Mean record length	35.9	49.0	+36.3%
Number of AMAX events	7401	12636	+70.7%
Final (water) year	2002-03	2021-22	+19 years

*Including six (HiFlows-UK 1.1) and four (NRFA Peak Flow 12.1) catchments that were excluded from most later stages of each study.

This study used LCM2015 (Rowland *et al.*, 2017) to derive $URBEXT_{2015}$, to describe catchment urbanization. This also allowed the consideration of "urban" and "suburban" land uses separately. Figure 20 compares $LCMURB_M$ (LCM2015 "urban" land use fraction, scaled to the mid-point of the AMAX record using *UEF*, the urban expansion factor) and $LCMSUB_M$ (the same for "suburban" land use) for the 884 catchments (626 essentially rural plus 258 urbanized catchments) in the full study dataset. As expected, $LCMURB_M$ is less than $LCMSUB_M$ in the vast majority of catchments; two of the four exceptions are on the River Tame and encompass greater Birmingham, while the other two are essentially rural (urban and suburban each less than 1%).

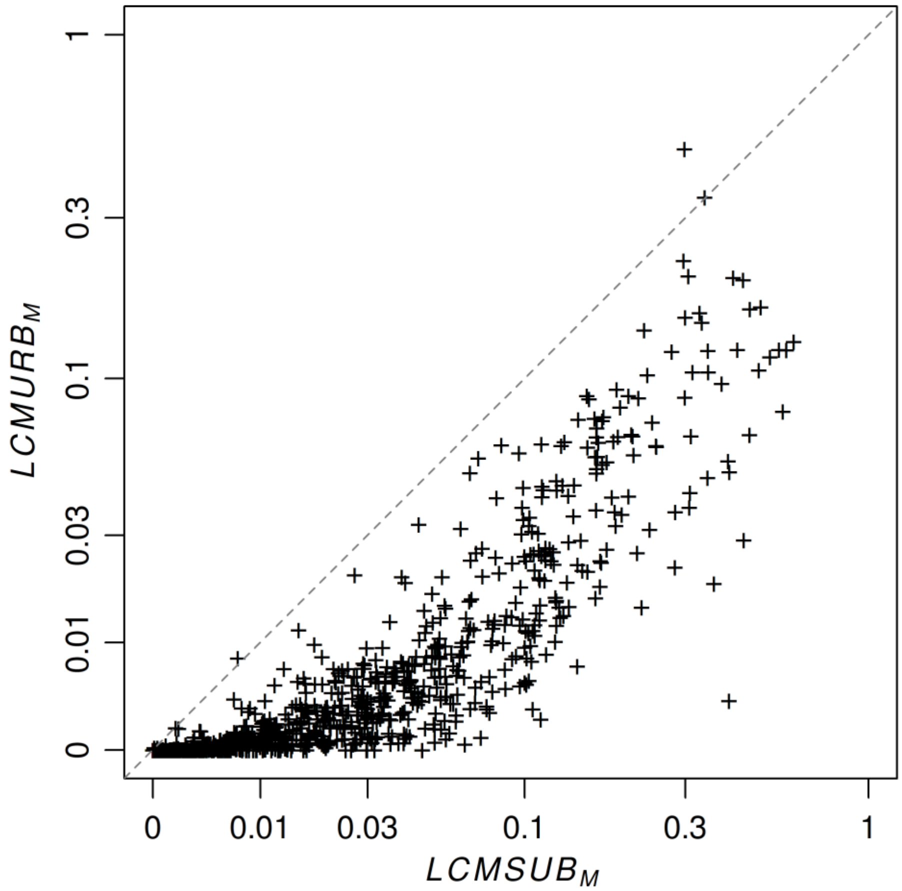

Figure 20 Comparison of *LCMURB*$_M$ and *LCMSUB*$_M$ for 884 catchments.

5.1 Recalibration of urbanization adjustment factor (*UAF*)

The *UAF* (Robson & Reed, 1999; Kjeldsen, 2010) extends the applicability of the ln(*QMED*) model to urbanized catchments. Its general form is of a multiplying factor applied to the (rural) *QMED* estimate, itself composed of two terms. This is presented below in its WINFAP 5.2 form (Wallingford HydroSolutions, 2022):

$$QMED_{urban} = QMED_{rural}(1 + IF.URBAN)^{1.25}\left(1 + IF.URBAN\left(\frac{PR_{IMP}}{69.366 - 65.686BFIHOST_{19}} - 1\right)\right)^{1.33}$$

In the above equation, *IF.URBAN* is typically equivalent to 0.4701$URBEXT_{2000}$ and PR_{IMP} is typically 70. A recalibration of all numeric (not variable) constants in this equation was performed as part of this study.

The first part of the recalibration was to determine the relationship between *SPRHOST* and $BFIHOST_{19SCALED}$ via linear regression, as this forms the part of the equation in the fraction under "PR_{IMP}". This regression used 906 catchments (884 catchments suitable for *QMED* estimation plus a further 22 in the NRFA Peak Flow dataset version 12.1 marked as unsuitable for *QMED* estimation, but with no reason to question the *SPRHOST* and $BFIHOST_{19SCALED}$ values.

The second part of the recalibration was to fit the two exponents in the equation, through linear regression of ln($QMED_{urban}$) – ln($QMED_{rural}$) against the two bracketed terms. For this part of the study, the 258 catchments marked as suitable for *QMED* estimation and with $URBEXT_M \geq 0.03$ were used, excluding four circled in Figure 21, where $QMED_{rural}$ was at least three times bigger than $QMED_{urban}$. The quantity of 254 suitable catchments compares favourably to the 115 used by Robson & Reed (1999) and 200 used by Kjeldsen (2010). $QMED_{urban}$ was taken as the gauged value of *QMED* at each gauge, while $QMED_{rural}$ was estimated from the recalibrated ln(*QMED*) model.

Here, *IF.URBAN* was taken as equal to $0.3URBEXT_M$, not $0.4701URBEXT_M$. This is based on more recent research by Miller *et al.* (2017), which suggests that *URBAN*, measured from a 50k map, is approximately equal to *URBEXT*, computed digitally from Land Cover Map, and not 1.567 times larger, as assumed previously.

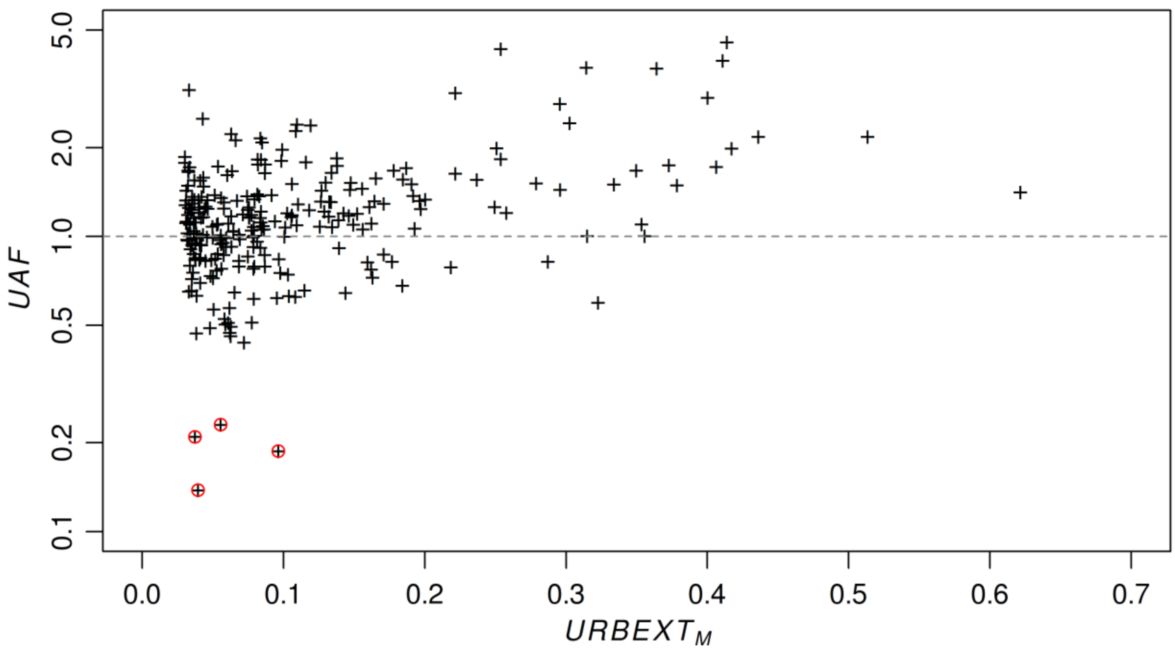

Figure 21 Urban adjustment factor (*UAF*) vs. $URBEXT_M$ for 254 urbanized catchments. Red circled catchments excluded from *UAF* model recalibration.

The final fitted linear regression between $QMED_{urban}$ / $QMED_{rural}$ and the two urbanization terms is given in Table 5. “First term” and “second term” relate to each of the top-level bracketed terms and their position in the equation. This regression has an R^2 of 0.3601, effectively equal to the 0.36 explained by Kjeldsen *et al.*’s (2008) equation. Wallingford HydroSolutions has not published an R^2 for the equation used in WINFAP 5.2, but it would not be comparable as that equation was calibrated only to catchments with $URBEXT_{2000} \geq 0.15$.

Table 5 Summary statistics for recalibrated *UAF* model (254 catchments, $URBEXT_M \geq 0.03$).

Variable	Coefficient	Standard error	*t*-value	*p*-value
Intercept	1	0	∞	0
First term	1.883802	0.868086	2.170	0.0309
Second term	3.519960	0.649906	5.416	1.42×10^{-7}

df = 252 $R^2 = 0.3601$

$$QMED_{urban} = QMED_{rural}(1 + 0.3URBEXT_M)^{1.8838} \times \left(1 + 0.3URBEXT_M\left(\frac{PR_{IMP}}{67.0674 - 63.8200BFIHOST_{19SCALED}} - 1\right)\right)^{3.5200}$$

It is noted that the above equation can give unrealistically high *UAF*s for very urbanized, very permeable catchments, much like the original *UAF*. In practice, *UAF* values should be capped at 10 – WINFAP 5.3 implements this.

Comparison with 2008 method and recalibration

Combining the 2025 model for rural ln(*QMED*) with the updated *UAF* gives the following six-parameter equation:

$$QMED = 6.8247AREA^{0.8499}0.1780^{\left(\frac{1000}{SAAR_{9120}}\right)}FARL_{2015}{}^{3.0450}0.0321^{BFIHOST_{19SCALED}{}^2} \times (1 + 0.3URBEXT_{2015})^{1.8838}\left(1 + 0.3URBEXT_{2015}\left(\frac{70}{67.0674 - 63.8200BFIHOST_{19SCALED}} - 1\right)\right)^{3.5200}$$

Applying the above, with donor transfer, to the 884 study catchments gives an *fse* of 1.4256 and R^2 of 0.9464, while applying the 2008 *QMED* and 2010 urbanization procedure gives marginally higher *fse* and lower R^2 values of 1.4278 and 0.9459 respectively. It is noted that donor transfer can be applied before or after urbanization of *QMED*; the final value is identical either way.

Figure 22 plots *QMED* after donor transfer, as estimated by the 2008 method and the 2025 method developed here, for 258 urbanized catchments. Similarly to the essentially rural catchments, there is little change in *QMED*, and percentage changes generally decrease as *QMED* increases. The spatial map (Figure 23) shows that the 2025 method tends to estimate slightly higher *QMED* values than the 2008 method across the whole of the urbanized UK, except in parts of the midlands and southeast England. There are 31 stations where the 2025 estimate of

QMED is more than 125% of the 2008 *QMED* estimate, and 4 stations where the 2025 estimate of QMED is less than 75% of the 2008 *QMED* estimate, with no obvious spatial pattern. Figure 24 compares relevant catchment descriptors for these 35 catchments, showing that decreases in *QMED* again seem to be mainly due to decreases in *FARL,* however, increases in *QMED* appear to be attributable mainly to increases in *SAAR* and/or *URBEXT*. Large differences in *QMED* are not related to changes in *BFIHOST*.

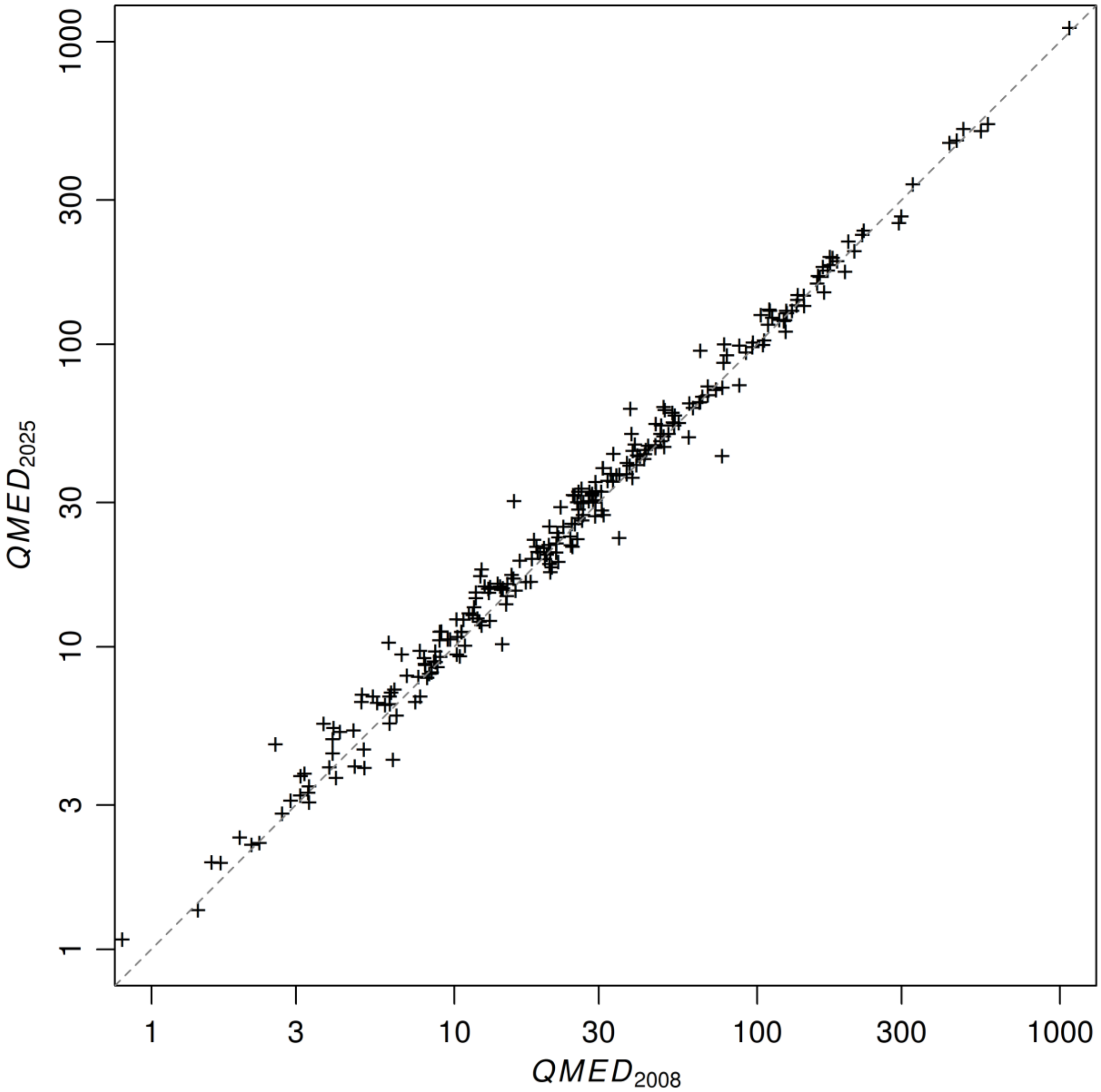

Figure 22 **Urbanized *QMED* after donor transfer, estimated by the 2008 method and the method developed in this study (2025), for 258 urbanized catchments.**

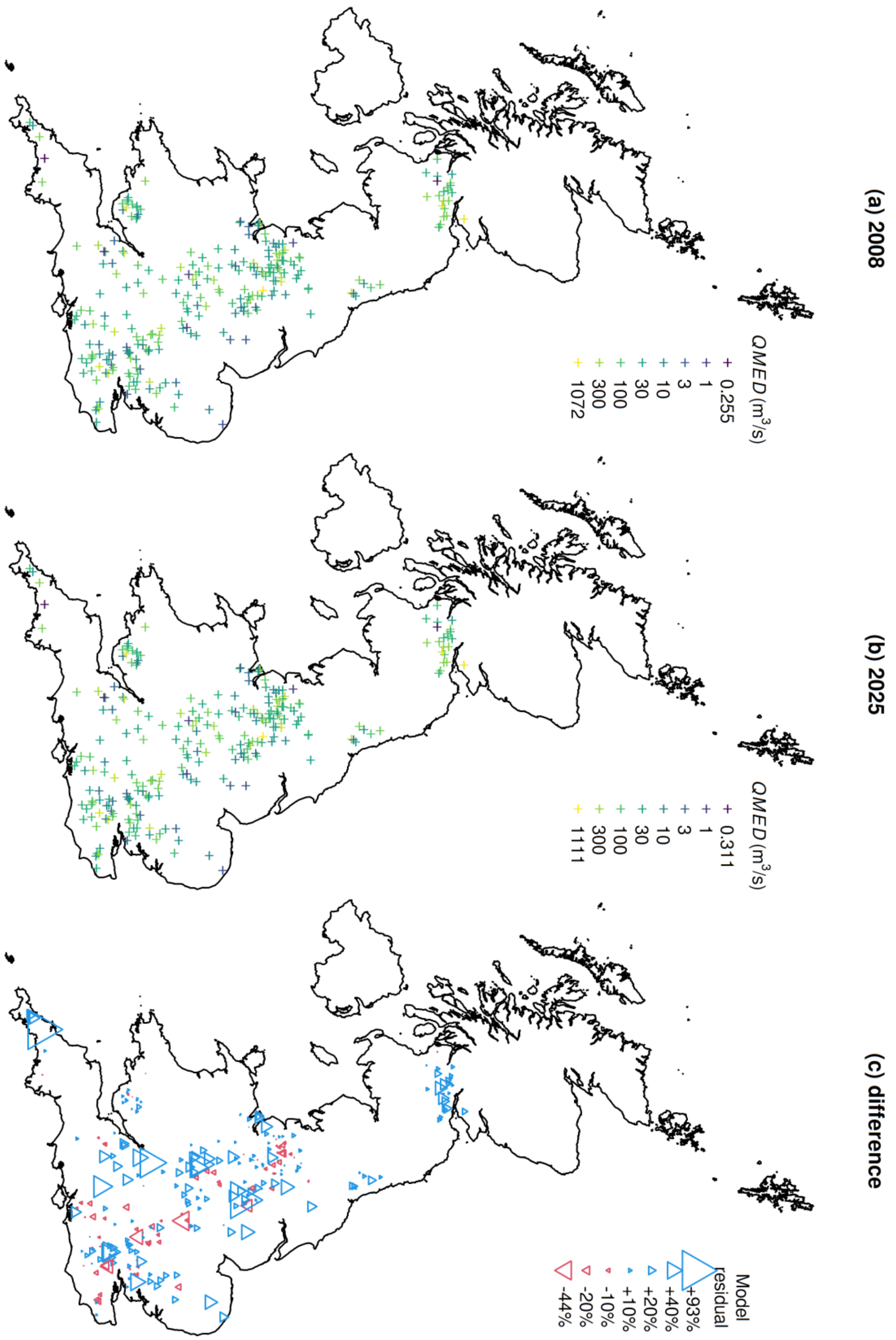

Figure 23 **2008 urbanized *QMED* (a), recalibrated 2025 urbanized *QMED* (b), and change from 2008 to 2025 urbanized *QMED* (c), after donor transfer (258 catchments).**

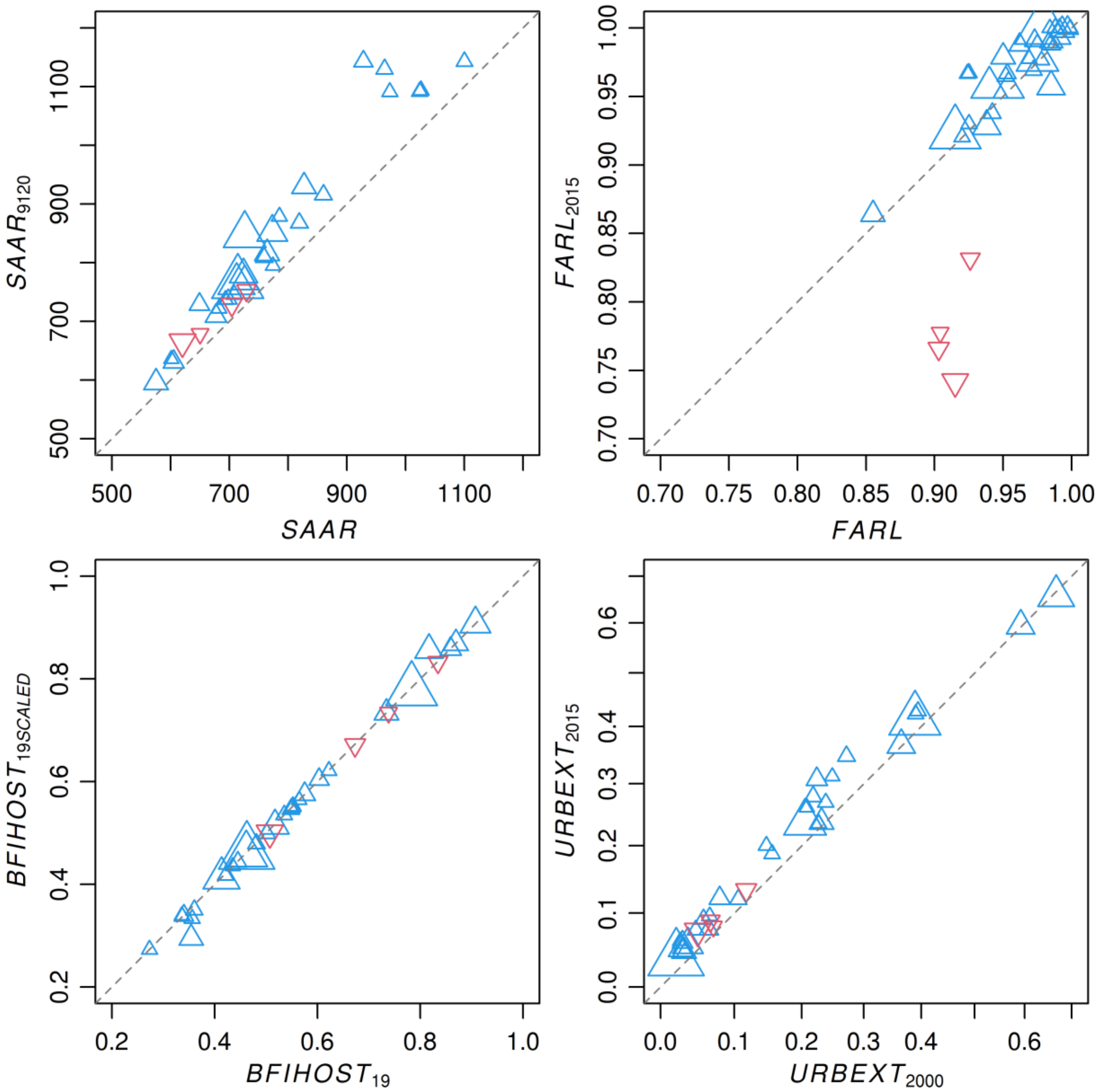

Figure 24 **Catchment descriptors for catchments where urbanized *QMED* after donor transfer is more than 125% (blue triangles) or less than 75% (red triangles) of 2008 urbanized *QMED* estimate after donor transfer, for 258 urbanized catchments.**

Table 6 compares *bias* and *fse* in modelled ln(*QMED*) for different subsets of catchment size, showing that the 2025 model outperforms the 2008 model in *fse* terms in various categories of small catchments, by a relatively larger margin than it does across larger catchments, and much more so as catchment size decreases. While the 2025 model does not outperform the 2008 model in terms of *bias*, values are small in all cases, and the 2025 model's lower *fse* indicates that its results have a smaller spread of errors relative to the low typical error. The combination of 2025

QMED and 2025 urbanization models therefore offers advantages for *QMED* estimation in smaller and more urbanized catchments.

Table 6 WINFAP 5 and 2025 *QMED* and *UAF* model *bias* and *fse* (factorial standard error) for different subsets of small catchment.

Catchment subset	№ of catchments	WINFAP 5 *bias*	2025 *QMED* & *UAF* model *bias*	WINFAP 5 *fse*	2025 *QMED* & *UAF* model *fse*
≤ 25 km^2	74	−0.0315	−0.0535	1.650	1.604
≤ 40 km^2	128	−0.0106	−0.0343	1.613	1.565
≤ 60 km^2	209	−0.0127	−0.0264	1.574	1.552
≤ 80 km^2	289	−0.0087	−0.0153	1.523	1.506

6. Pooling

The 2008 pooling similarity distance measure (SDM) was introduced by Kjeldsen *et al.* (2008). This is presented below:

$$SDM_{ij} = \sqrt{3.2\left(\frac{\ln AREA_i - \ln AREA_j}{1.28}\right)^2 + 0.5\left(\frac{\ln SAAR_i - \ln SAAR_j}{0.37}\right)^2 + 0.1\left(\frac{FARL_i - FARL_j}{0.05}\right)^2 + 0.2\left(\frac{FPEXT_i - FPEXT_j}{0.04}\right)^2}$$

This has remained until now the default method for forming pooling-groups, except (since 2021) for studies conducted on catchments under 40 km^2. This study redeveloped the SDM with the latest and most suitable AMAX and catchment descriptors. Only catchments marked as suitable for pooling and with an $URBEXT_M$ value of 0.03 or less were used in this redevelopment. The requirement for pooling-suitable and initially essentially rural catchments meant that only 386 catchments were available for use here, somewhat less than the 602 presumably used by Kjeldsen *et al.* (2008). However, the average quality of the gauged data available now is far higher, and 257 (42.7%) of the 602 catchments used by Kjeldsen *et al.* (2008) are now explicitly marked as not suitable for pooling. Due to the typically longer record lengths now available, current estimates are subject to less sampling uncertainty than those made in 2008. However, each record is either comparable to or shorter than the return periods typically of interest.

The pooling method creates flexible regions, called pooling-groups, on a catchment-by-catchment basis. Catchments are selected for a pooling-group according to their similarity to the catchment of interest in key catchment descriptors, measured by the similarity distance measure (SDM) presented above, and added to the pooling-group until the total length of catchment AMAX records exceeds 500 years. Weighted mean values of the gauged second and third *L*-moment ratios (Hosking & Wallis, 1997), *L*-CV and *L*-SKEW, are calculated and used to parameterize a pooled generalized logistic (GLO) distribution, which is then taken as appropriate for the site of interest.

The weights applied to each catchment descriptor in the SDM and the target number of years in the pooling-group were set by Kjeldsen *et al.* (2008) in order to minimize the pooled uncertainty measure (PUM). This is the weighted mean difference between at-site growth factor and pooled growth factor at the M sites used to develop the method, at a specified return period of T years:

$$PUM_T = \left(\frac{\sum_{i=1}^{M} w_i\left(\ln x_{T_i} - \ln x_{T_i}^P\right)^2}{\sum_{i=1}^{M} w_i}\right)^{½}$$

$$w_i = \frac{n_i}{1 + n_i/16}$$

Where x_T is the at-site growth factor, x_T^P is the pooled growth factor and n_i is the number of AMAX at site *i*. Kjeldsen *et al.* (2008) evaluated PUMs for periods of 20, 50 and 100 years.

The weight assigned to each catchment contained within the pooling-group is based on its similarity to the target catchment and on the sampling error of its *L*-moment ratios (which is based on record length).

Additional filtering of stations

The NRFA rates 386 stations in the study dataset as "suitable for pooling". However, these ratings are not static, and a station can gain a new suitability rating due to changes either at the station or in its upstream catchment. In addition, organizations that work closely with certain AMAX records may develop their own detailed understanding of that data. For this part of the study, Wallingford HydroSolutions (WHS) and Cyfoeth Naturiol Cymru (CNC) both provided lists of stations that they routinely, but not always, excluded from pooling-group formation. WHS's list contained 19 catchments, all of which were very permeable (minimum, mean and maximum $BFIHOST_{19}$ = 0.799, 0.895 and 0.933 respectively). In these cases, the very high $BFIHOST_{19}$ was an indicator of one or more of three potential issues: a significant difference between the topographic and contributing catchment, an ephemeral catchment, or a large fraction of groundwater-driven AMAX. CNC provided a list of seven stations, one overlapping with WHS's list. The main reasons for exclusion were bypassing and non-modular high flows. Of the 25 catchments highlighted by WHS and CNC, 20 were rated as suitable for pooling.

On further review, eight pooling-suitable stations were excluded: all of those that had problems other than high $BFIHOST_{19}$ values. Stations that were excluded purely for high $BFIHOST_{19}$ values were kept in the case that $BFIHOST_{19}$ might be selected for a new SDM.

6.1 Redeveloped pooling method

Redevelopment of the pooling method requires several intermediate steps:

1. Determine appropriate catchment descriptors that influence at-site *L*-moments,
2. Estimate target growth factors for several return periods from gauged at-site AMAX data,
3. Estimate pooled growth factors, using similarity distance measure (SDM) to select pooling-group, and optimize weights within SDM to minimize pooled uncertainty measure (PUM) across several important return periods,
4. Perform a variogram analysis to refine the weights assigned to each catchment within pooling-groups.

Step 3 was performed 20 times independently, using target pooling-group lengths of 100, 200, 300… to 2000 years. Pooling-group members were initially weighted using the system developed by Kjeldsen *et al.* (2008), as the calculation of weights in Step 4 cannot be performed without first specifying an SDM. However, the authors stated that the specific weights assigned to each catchment in a pooling-group had very little effect on the PUM calculated across all catchments, so Step 3 was repeated after refining the within-pooling-group catchment weighting, but Steps 3 and 4 were not performed in a repeated cycle.

Selection of catchment descriptors

Catchment descriptors, and multiple transformations thereof, were tested for their correlations with *L*-moment ratios of the at-site AMAX series The train function from the caret R package (Kuhn *et al.,* 2023) was used to perform stepwise regression, with a goal of maximizing R^2. Each catchment in this regression was weighted by *w*, giving more prominence to longer records with less uncertain *L*-moments.

For the second *L*-moment ratio, a five-term regression containing ln(*DPLBAR*), 1000/$SAAR_{POR}$, $FARL_{2015}^2$, *FPEXT* and $BFIHOST_{19SCALED}$ gave an R^2 of 0.4733. The position of $BFIHOST_{19SCALED}$ as the fifth-most important descriptor can support the exclusion of *BFIHOST* from the four-term regression created in 2008. However, our results also highlight the validity of a five-term regression including $BFIHOST_{19SCALED}$, which removes the possible need to exclude catchments with high $BFIHOST_{19SCALED}$ from the calibration. In the SDM, ln(*DPLBAR*) was replaced with ln(*AREA*), which reduced R^2 to 0.4676; this was done as it is much easier to estimate ln(*AREA*) than ln(*DPLBAR*) accurately “by hand” in the case that a user-defined catchment is not well represented on the FEH Web Service.

For the third *L*-moment ratio, R^2 was maximized overall, to a value of 0.0384, by a one-term regression containing 1/*DPSBAR*. Due to the essential randomness of *L*-SKEW, at least relative to catchment descriptors, it was proposed to base the 2025 SDM only on those catchment descriptors that explain some variance in t_2. This is presented below:

$$SDM5_{ij} = \sqrt{x_1\left(\frac{\ln AREA_i - \ln AREA_j}{1.3207}\right)^2 + x_2\left(\frac{\frac{1000}{SAAR_{9120,i}} - \frac{1000}{SAAR_{9120,j}}}{0.3566}\right)^2 + x_3\left(\frac{FARL_{2015,i}^2 - FARL_{2015,j}^2}{0.0976}\right)^2 + x_4\left(\frac{FPEXT_i - FPEXT_j}{0.0439}\right)^2 + x_5\left(\frac{\frac{1}{BFIHOST_{19SCALED,i}} - \frac{1}{BFIHOST_{19SCALED,j}}}{0.6610}\right)^2}$$

where *x* terms in both SDMs represent weights to be optimized to minimize PUM.

Target growth factors

At-site growth factors were calculated at 30-, 100-, 200- and 1000-year return periods for all 378 suitable study catchments, assuming generalized logistic distributions, parameterized by the *L*-moments of gauged AMAX series, in all cases. The x_1-x_5 weights in SDM_5 were set to minimize the sum of the 30-, 100-, 200- and 1000-year PUM.

Optimize SDM weights and pooling-group length

As stated, the optimization to minimize PUM by adjusting the weights in SDM_5 was performed at 20 different pooling-group lengths. Optimal SDM weights (x_1-x_5) of 1.74, 1.63, 0.26, 0.55 and 0.82 were found with an optimal pooling-group length of 800 years. PUMs achieved by SDM_5 are compared against PUMs from the existing SDM in Table 7. Reductions in PUM, indicating an improvement in performance, were found at a range of return periods from 30 to 1000 years.

Table 7 PUMs obtained for 2008 SDM and SDM_5 (best new SDM).

SDM	PUM			
	30 years	100 years	200 years	1000 years
Current (2008)	0.1795	0.2602	0.3101	0.4350
SDM_5	0.1703	0.2480	0.2963	0.4175

Variogram analysis

Given optimal values for x_1-x_5, it was possible to perform the variogram analysis in Stage 4. The justification and procedure are described in detail in pages 57-62 of Kjeldsen *et al.* (2008). Key numerical differences relate to the calculated sampling variances of *L*-CV and *L*-SKEW, and the distance-based correlations between *L*-CV and *L*-SKEW at pairs of stations. Figure 25 presents the fitted relationships between sample length and variance in *L*-CV and *L*-SKEW.

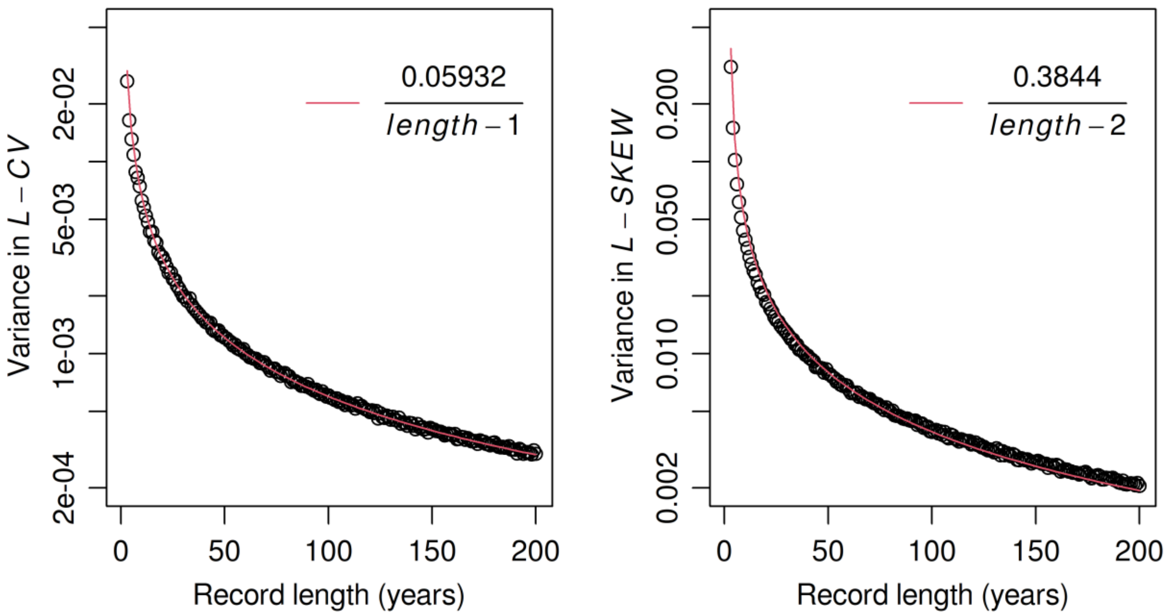

Figure 25 Variance in *L*-CV and *L*-SKEW for a typical UK AMAX record (*L*-CV = 0.2075, *L*-SKEW = 0.1834) as a function of record length.

Distance-based correlations between *L*-CV and *L*-SKEW were estimated using the same procedure as distance-based correlations in *QMED* i.e. all pairs of stations with 40 or more years of overlapping record were identified (35,115 pairs when considering only pooling-suitable stations), 10,000 bootstraps were generated for each pair, then correlations in *L*-CV and *L*-SKEW were related to catchment centroid-centroid distance. The fitted double-exponential relationships between correlation in paired *L*-CV or *L*-SKEW values, and catchment centroid-centroid distance, were:

$$\mathrm{cor}\left(t_{2,i},t_{2,j}\right) = 0.2807e^{-0.0076d_{ij}}+(1\ -\ 0.2807)e^{-0.0641d_{ij}}$$

$$\mathrm{cor}\left(t_{3,i},t_{3,j}\right) = 0.1404e^{-0.0120d_{ij}}+(1\ -\ 0.1404)e^{-0.0811d_{ij}}$$

Generating pooling-groups for each of the 906 NRFA Peak Flow 12.1 catchments (suitable for pooling, *QMED* and neither) revealed that the maximum SDM_5 of any station in any pooling-group was 4.06. Hence, the variogram was constructed from 100 bins of width 0.04 similarity, as in Kjeldsen *et al.* (2008). The fitted relationships are below:

L-CV: $\gamma = 0.02100\left(1\ -\ e^{-\frac{SDM}{9.6966}}\right)$

L-SKEW: $\gamma = 0.02311\left(1\ -\ e^{-\frac{SDM}{0.3693}}\right)$

Note that the above form of γ for *L*-CV has a value of zero at SDM = 0. This means that the weighting for an at-site record is not additionally enhanced beyond the already high value that is achieved by having an SDM of zero.

Substituting these new values for variance, correlation and γ into the pooling-group formation procedure reduced PUM values very marginally (Table 8), indicating a slight improvement in the method.

Table 8 PUMs obtained for SDM5 before and after within pooling-group reweighting.

SDM	PUM			
	30 years	100 years	200 years	1000 years
SDM_5 (original PG weights)	0.1703	0.2480	0.2963	0.4175
SDM_5 (reweighted PG)	0.1702	0.2479	0.2962	0.4173

The final SDM for the redeveloped pooling method is:

$$SDM5_{ij} = \sqrt{\begin{aligned}&1.74\left(\frac{\ln AREA_i - \ln AREA_j}{1.3207}\right)^2 + 1.63\left(\frac{\frac{1000}{SAAR_{9120,i}} - \frac{1000}{SAAR_{9120,j}}}{0.3566}\right)^2 + 0.26\left(\frac{FARL_{2015,i}^2 - FARL_{2015,j}^2}{0.0976}\right)^2 \\ &+ 0.55\left(\frac{FPEXT_i - FPEXT_j}{0.0439}\right)^2 + 0.82\left(\frac{\frac{1}{BFIHOST_{19SCALED,i}} - \frac{1}{BFIHOST_{19SCALED,j}}}{0.6610}\right)^2\end{aligned}}$$

6.2 Small catchments method

The small-catchment SDM (Vesuviano *et al.,* 2024), implemented in WINFAP 5 since 2021, contains two terms weighted equally:

$$SDM_{SC} = \sqrt{\left(\frac{\ln AREA_i - \ln AREA_j}{1.264}\right)^2 + \left(\frac{\ln SAAR_i - \ln SAAR_j}{0.349}\right)^2}$$

One consequence of using this SDM for small catchments is an increased diversity in pooling-groups, as the weighting applied to ln(*AREA*) in this SDM does not overwhelm the pooling-group selection procedure, as it does in the 2008 SDM.

Given the reduced weighting applied to *AREA* in SDM_5, the continuing need for an SDM specific to small catchments was tested here with the latest catchment descriptors and NRFA data set. All 378 essentially rural, pooling-suitable catchments were available for selection, but the PUM was only calibrated on the 44 of these that had *AREA* < 40 km^2. A calibration dataset of 44 catchments is somewhat smaller than the 57 used to calibrate the Small Catchments SDM (Vesuviano *et al.,* 2024). However, that dataset included a number of "de-urbanized" urban catchments.

Considering that equal weights may not necessarily offer the lowest PUM on this dataset, an optimization procedure to identify the PUM-minimizing weights and pooling-group length was undertaken, following the same procedure as previously.

Table 9 compares the PUMs achieved by the optimized small catchments SDM on this dataset of 44 small pooling-suitable catchments against those achieved by SDM_5 on the same dataset, showing that there is no longer a need for a separate small-catchment SDM.

Table 9 PUMs achieved on dataset of 44 small, pooling-suitable small catchments by two SDMs (re-optimized small-catchment and *SDM5*).

SDM	PUM			
	30 years	100 years	200 years	1000 years
Re-optimized small-catchment	0.2238	0.3057	0.3550	0.4760
SDM5	0.2070	0.2889	0.3390	0.4635

While this appears to contradict findings from a study performed just seven years ago (Vesuviano *et al.,* 2024), the apparent need for a small-catchment SDM may have arisen from the very high weight applied to *AREA* relative to other descriptors in the 2008 SDM, which is severely reduced in *SDM5*. It is also plausible that the "apparent" correlations between catchment descriptors and *L*-moment ratios may have changed simply due to sampling variability, in which case the lower sampling variabilities associated with the now-longer AMAX records should take precedence.

Whichever factor is dominant, the assumption that pooling-group methods in general allow easy incorporation of new AMAX data is challenged.

This recommendation comes soon after the publication of different findings (Stewart et al., 2024) in an EA project. Although publication of the small catchments study (Stewart et al., 2024) was completed in 2024, the analysis and method development for that study was undertaken in 2017, and the method implemented into WINFAP 5.0 in September 2021.

6.3 Simpler alternatives to pooling

Several simpler alternatives to pooling were tested: national-average *L*-moments (*L*-CV = 0.2075, *L*-SKEW = 0.1834), regional-average *L*-moments (using the same regions as the Flood Studies Report), and catchment-descriptor regressions for *L*-moment ratios (those used to select descriptors for the pooling SDM).

Table 10 compares PUMs achieved by the simpler alternatives to pooling against those achieved by pooling with the 2008 SDM and with *SDM5*. Unsurprisingly, national *L*-moments give the worst performance, followed by regional *L*-moments. The 2008 pooling method is the third-worst performer, potentially due to the changes in *L*-moments, and therefore changes in the relationships between *L*-moments and catchment descriptors, following 19 additional years of data collection.

It is noteworthy that all methods shown in Table 10 achieve significantly lower 100-year PUMs than those reported by Kjeldsen *et al.* (2008), who used a 602-catchment calibration dataset, including many records that would now be considered unsuitable for pooling due to inappropriately high uncertainties in the highest AMAX that have the highest influence over *L*-moments. This strongly demonstrates the importance of data quality review.

Table 10 Comparison of PUMs achieved using simpler alternatives to pooling against those achieved by pooling with different SDMs.

Method	PUM			
	30 years	100 years	200 years	1000 years
National *L*-moments	0.2017	0.2828	0.3323	0.4554
Regional *L*-moments	0.1888	0.2681	0.3166	0.4372
Pooling (2008)	0.1795	0.2602	0.3101	0.4350
Regression *L*-moments (5-par)	0.1738	0.2537	0.3034	0.4279
SDM_5 (all rural pooling)	0.1702	0.2479	0.2962	0.4173

6.4 Overall patterns

Figure 26 plots the 30-, 100-, 200- and 1000-year pooled growth factors using the 2025 method minus the same growth factors found using the 2008 method for catchments above 40 km^2 and small-catchment (SC) method for catchments up to 40 km^2 – note that these are *not* percentage differences. Much like differences in *QMED*, most changes are small and evenly distributed. However, there is a cluster of catchments in northeast England where changing from the 2008 or SC method to the 2025 method consistently reduces the *T*-year growth factors, and potentially another cluster in southwest England where changing from the 2008 or SC method to the 2025 method consistently increases the *T*-year growth factors, though by a smaller amount.

Figure 27 is equivalent to Figure 26 for 1278 ungauged rural catchments. As over 88% of these are under 40 km^2, the small-catchment pooling method is used in the majority of cases. Many small-catchment growth factors are large, and larger than 2025 growth factors (see Figure 33 for more information). However, the general spatial pattern of positive and negative differences between growth factors follows the gauged dataset, there are just more ungauged catchments in the areas of the country in which small-catchment growth factors exceed 2025 growth factors.

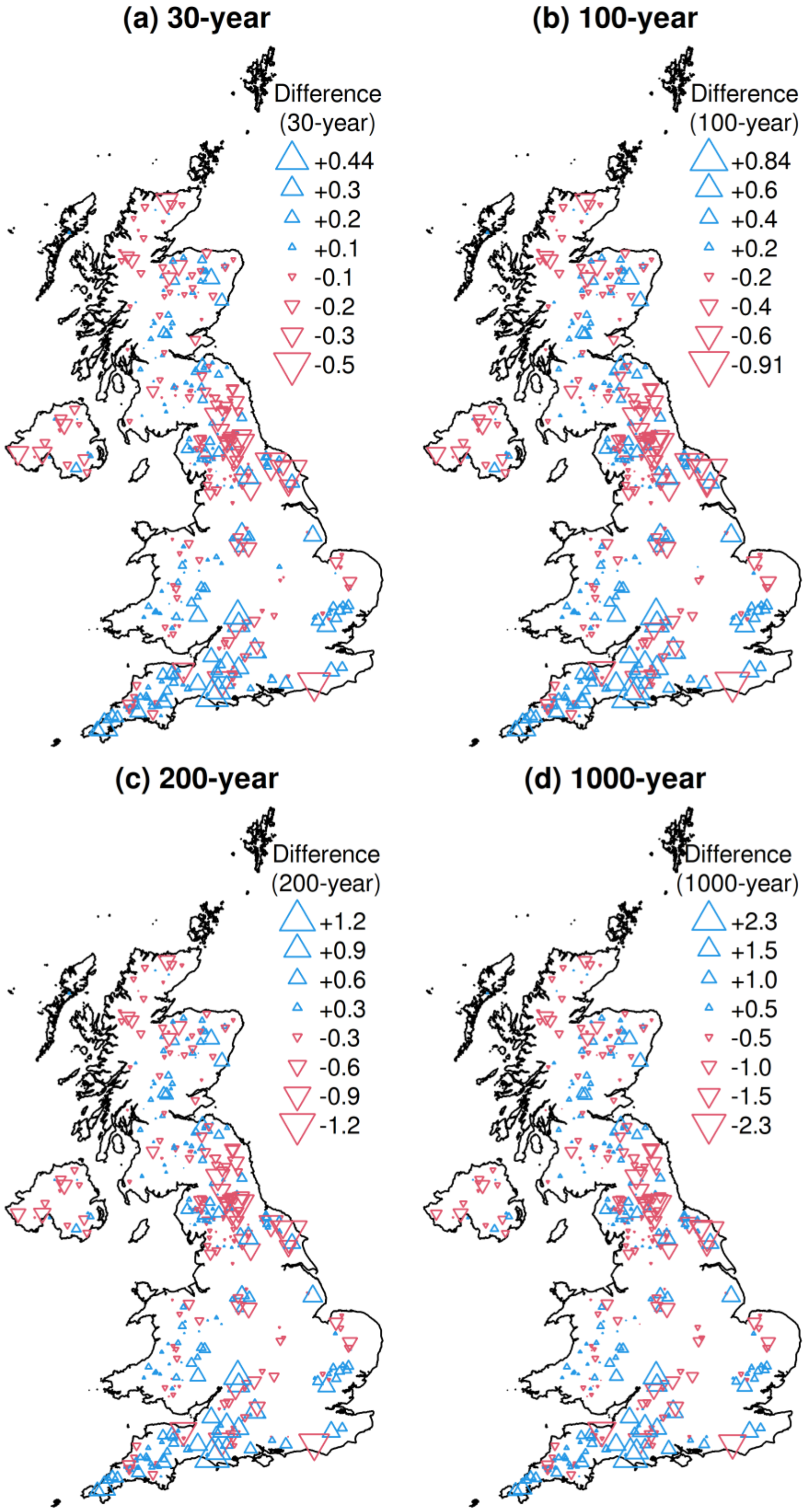

Figure 26 **Difference between 2025 and 2008 methods for 30- (a), 100- (b), 200- (c) and 1000- (d) year growth factors for 378 essentially rural catchments. Positive values show 2025 results are larger.**

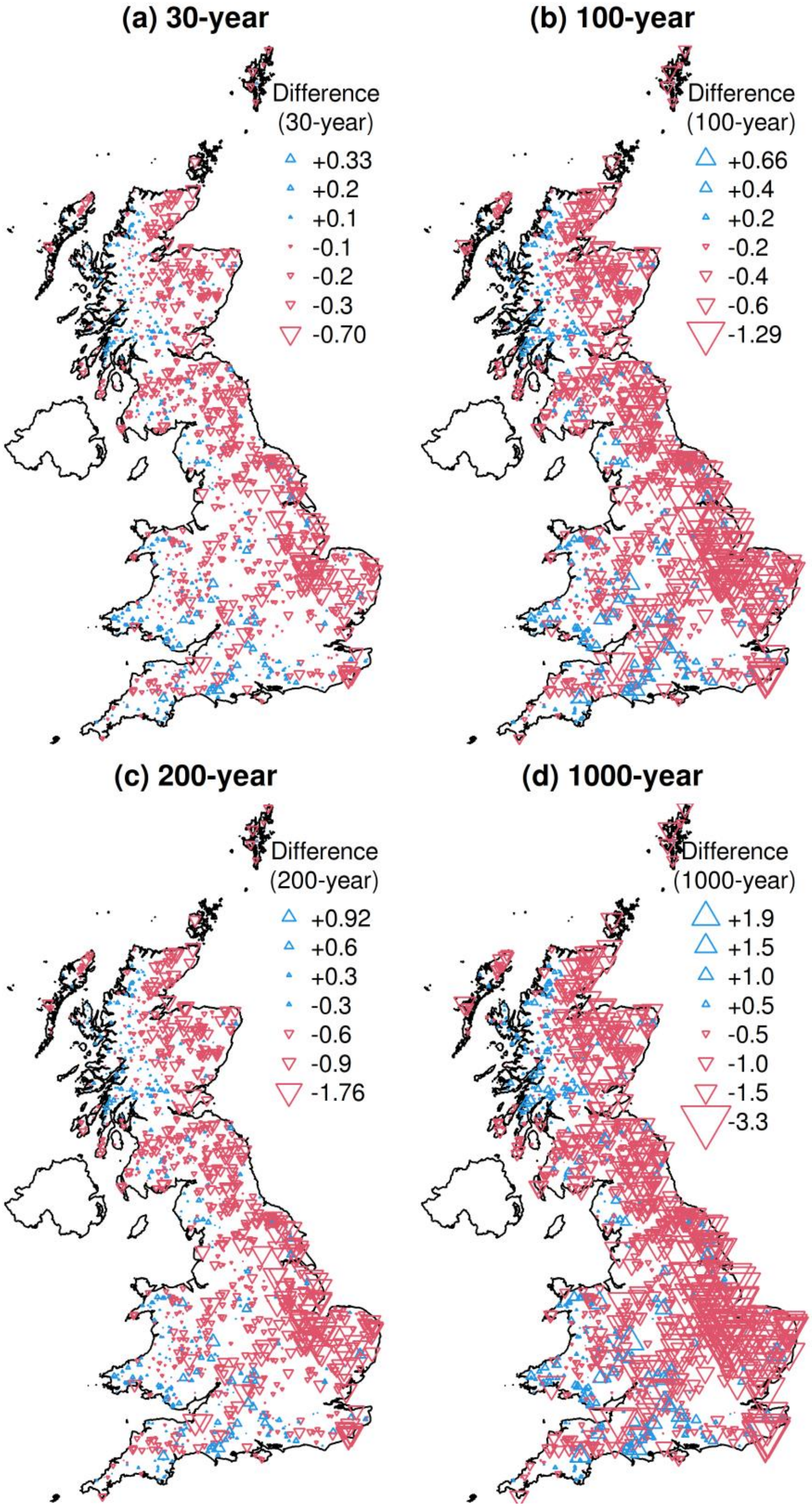

Figure 27 **Difference between 2025 and 2008 methods for 30- (a), 100- (b), 200- (c) and 1000- (d) year growth factor for 1278 ungauged essentially rural catchments. Positive values show 2025 results are larger.**

7. Urbanization (pooling)

By default, the pooling method incorporates only essentially rural catchments into pooling-groups, specifying an $URBEXT_{2000}$ value of 0.03 as the threshold between rural and urban catchments, but research using ReFH2 suggests that the effects of urbanization may not be apparent until a much higher $URBEXT_{2000}$ threshold value (Stewart *et al.*, 2024). In recognition of this, WINFAP 5 does allow the use of more urbanized catchments in pooling-groups (though not by default).

Conceptually, all surfaces are likely to saturate and generate surface runoff in a similar way to impermeable surfaces when exposed to very large rainfall events. Hence, given the positive relationship between urbanization and *QMED*, there should be an equal and opposite relationship between urbanization and growth factors, so that the *T*-year flood peak (*QMED* × *T*-year growth factor) is entirely unrelated to catchment permeability or impermeability, for some high value of *T*. Figure 28 plots 30-, 100-, 200- and 1000-year growth factors against $URBEXT_M$ for the 152 pooling-suitable catchments in the study dataset that have $URBEXT_M \geq 0.03$.

Figure 28 shows that the relationship between a catchment's urbanization level and its flood frequency curve is not strong. However, it is negative on average, and the gradient increases with return period. This does suggest that catchments do behave more impermeably under rarer floods.

While Figure 28 shows specific return periods, Figure 29 shows the relationship between *L*-moments and $URBEXT_M$ for the same 152 catchments.

There seems to be a stronger relationship between $URBEXT_M$ and *L*-CV than between $URBEXT_M$ and *L*-SKEW. However, both are weak: The correlation between $URBEXT_M$ and ln(*L*-CV) is −0.109, while the correlation between $\ln(URBEXT_M + 1)$ and ln(*L*-SKEW + 1) is just 0.044.

For each of *L*-CV and (1 + *L*-SKEW), Figure 30 plots the at-site value divided by the ungauged pooling-group estimate, against $URBEXT_M$, for the same 152 catchments. Similarly to Figure 29, this shows a gradient against $URBEXT_M$ for *L*-CV but not *L*-SKEW. However, the choice of *y*-axis in these two cases imply the forms of equations that can "urbanize" or "de-urbanize" the *L*-moments, following Kjeldsen (2010).

The two equations generated are:

$$L\text{-CV}_{urbanized} = 0.5269^{URBEXT_M} L\text{-CV}_{pooled} \qquad R^2 = 0.1287$$

$$(L\text{-SKEW}_{urbanized} + 1) = 1.0174^{URBEXT_M} (L\text{-SKEW}_{pooled} + 1) \qquad R^2 = 0.0005$$

Use of the *L*-SKEW equation is not recommended as it has near-zero predictive power. Use of the *L*-CV equation *is* recommended as, while it is also weak, Figure 29 and Figure 30 do show some relationship between *L*-CV and urbanization (unlike *L*-SKEW and urbanization).

Figure 31 plots the 30-, 100-, 200- and 1000-year pooled growth factors using the 2025 method minus the same growth factors found using the 2008 method for catchments above 40 km^2 and the small-catchment (SC) method for catchments up to 40 km^2 – note that these are *not* percentage differences. While most changes are small, there appear to be more negative than positive changes, with the gap widening as return period increases. This is likely due to the lack of adjustment to *L*-SKEW in the 2025 method, which makes the flood frequency distributions "straighter" at extreme return periods. However, no relationship was found between urbanization and *L*-SKEW, so an adjustment cannot be justified.

Figure 32 is equivalent to Figure 31 for 285 ungauged urbanized catchments. As 81% of these are under 40 km^2, the 2025 pooling method is compared against the small-catchment pooling method in the majority of cases. Many small-catchment growth factors are large (see Figure 33 for more information).

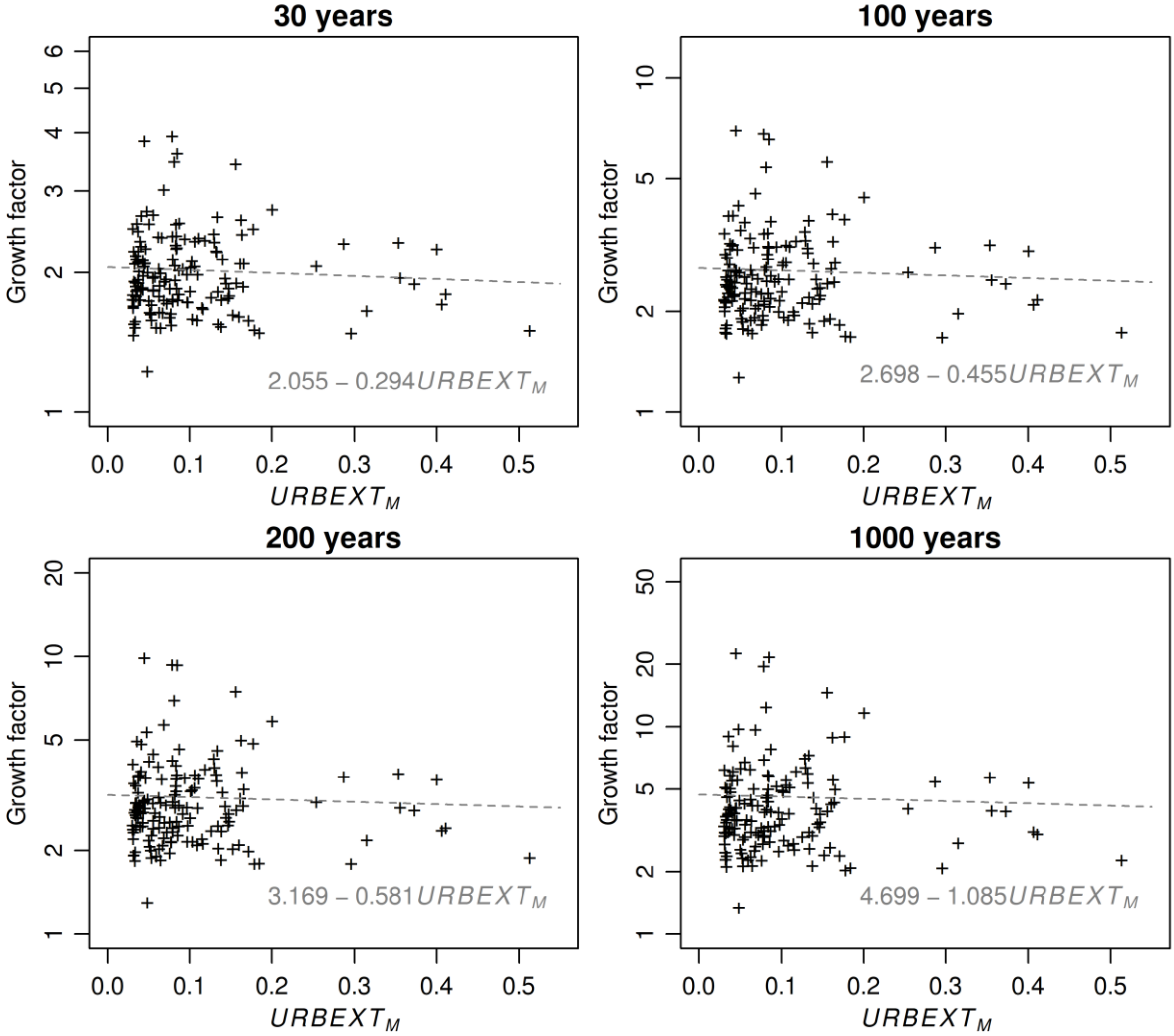

Figure 28 30-, 100-, 200- and 1000-year growth factors for 152 pooling-suitable catchments with $URBEXT_M \geq 0.03$.

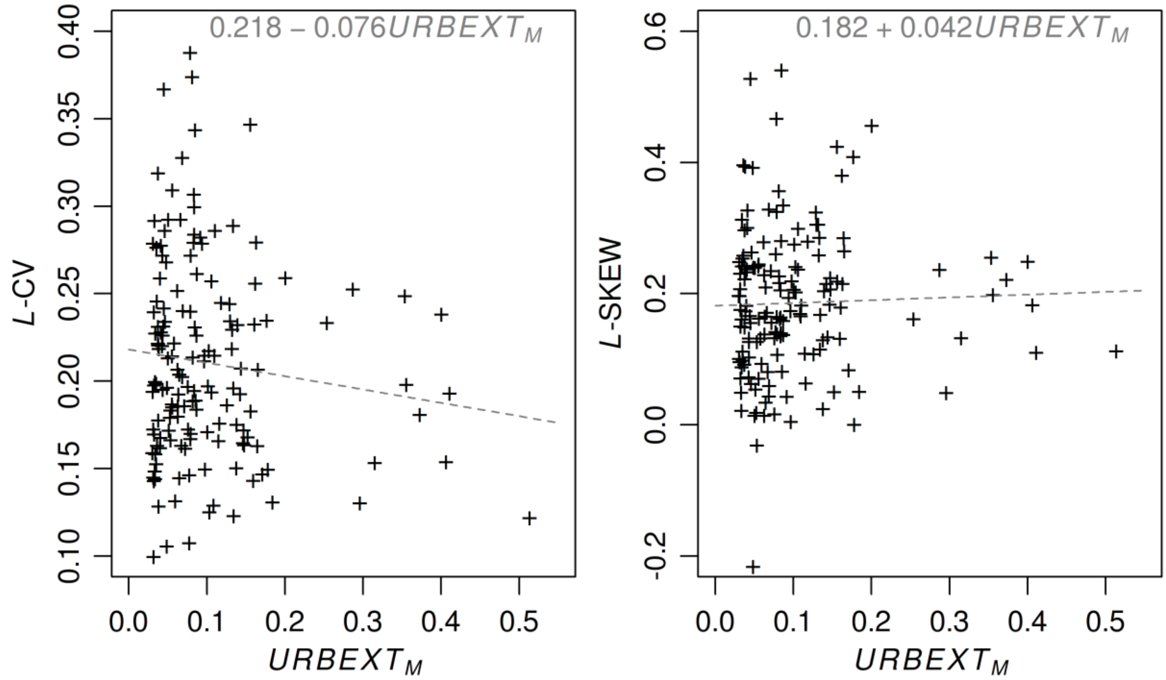

Figure 29 ***L*-CV and *L*-SKEW against *URBEXT*$_M$ for 152 pooling-suitable catchments with *URBEXT*$_M$ ≥ 0.03.**

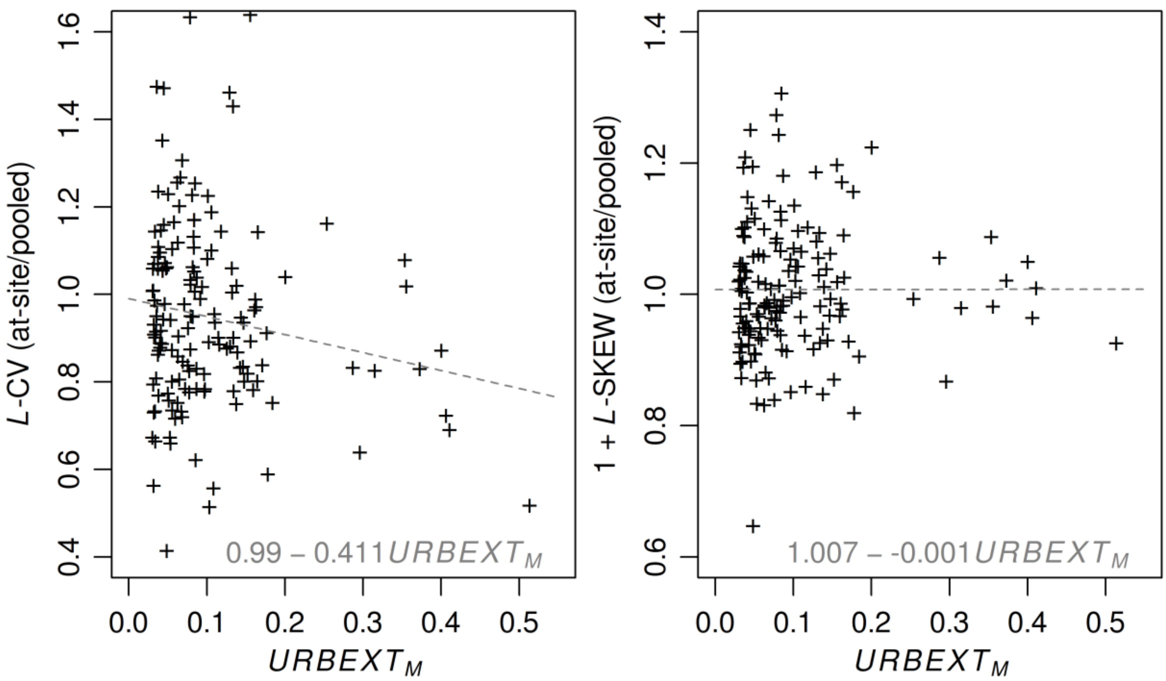

Figure 30 **Ratio of at-site-to-pooled *L*-CV and *L*-SKEW against *URBEXT*$_M$ for 152 pooling-suitable catchments with *URBEXT*$_M$ ≥ 0.03.**

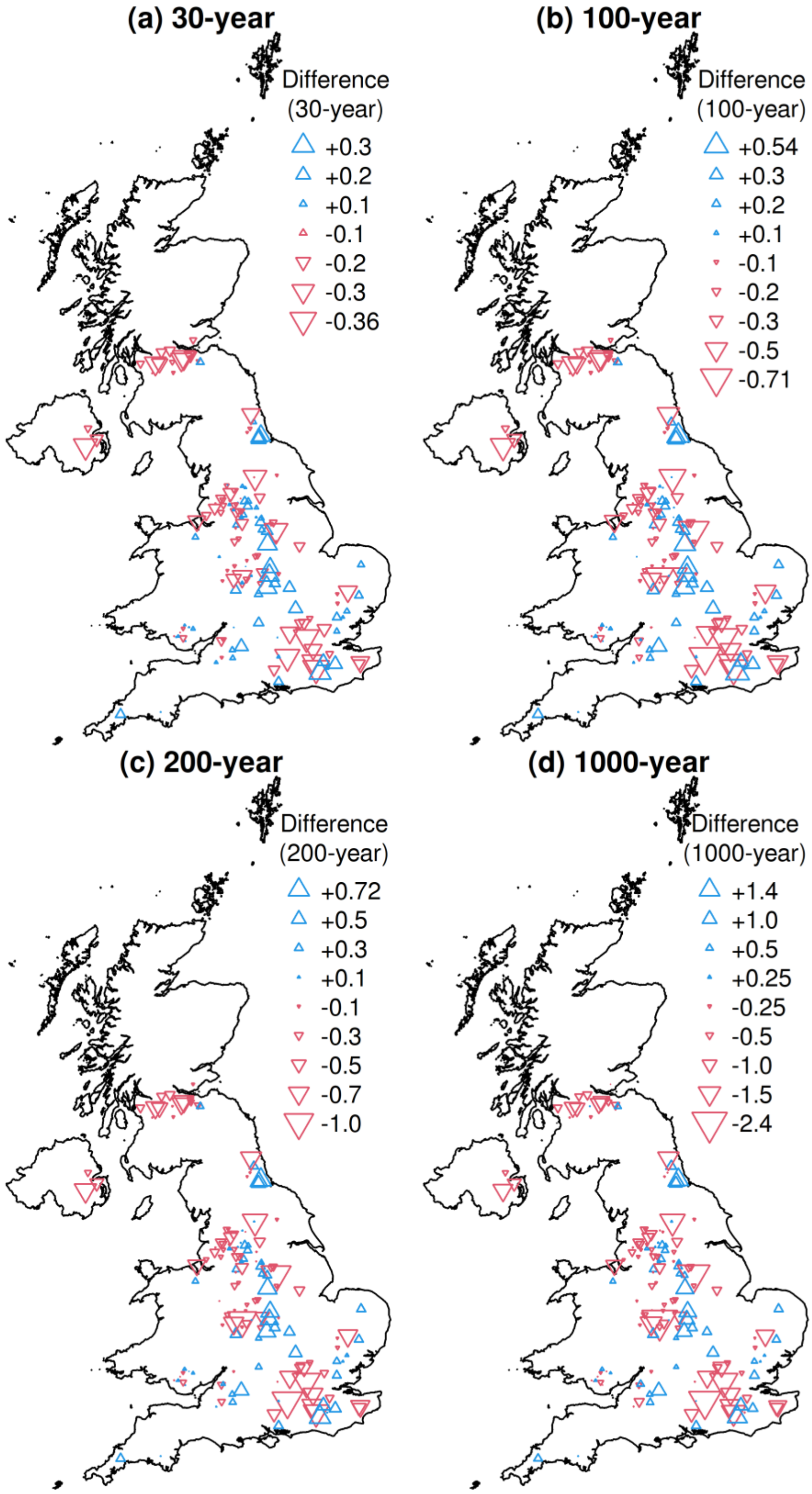

Figure 31 **Difference between 2025 and 2008 methods for 30- (a), 100- (b), 200- (c) and 1000- (d) year growth factor for 152 urbanized catchments. Positive values show 2025 gives larger values.**

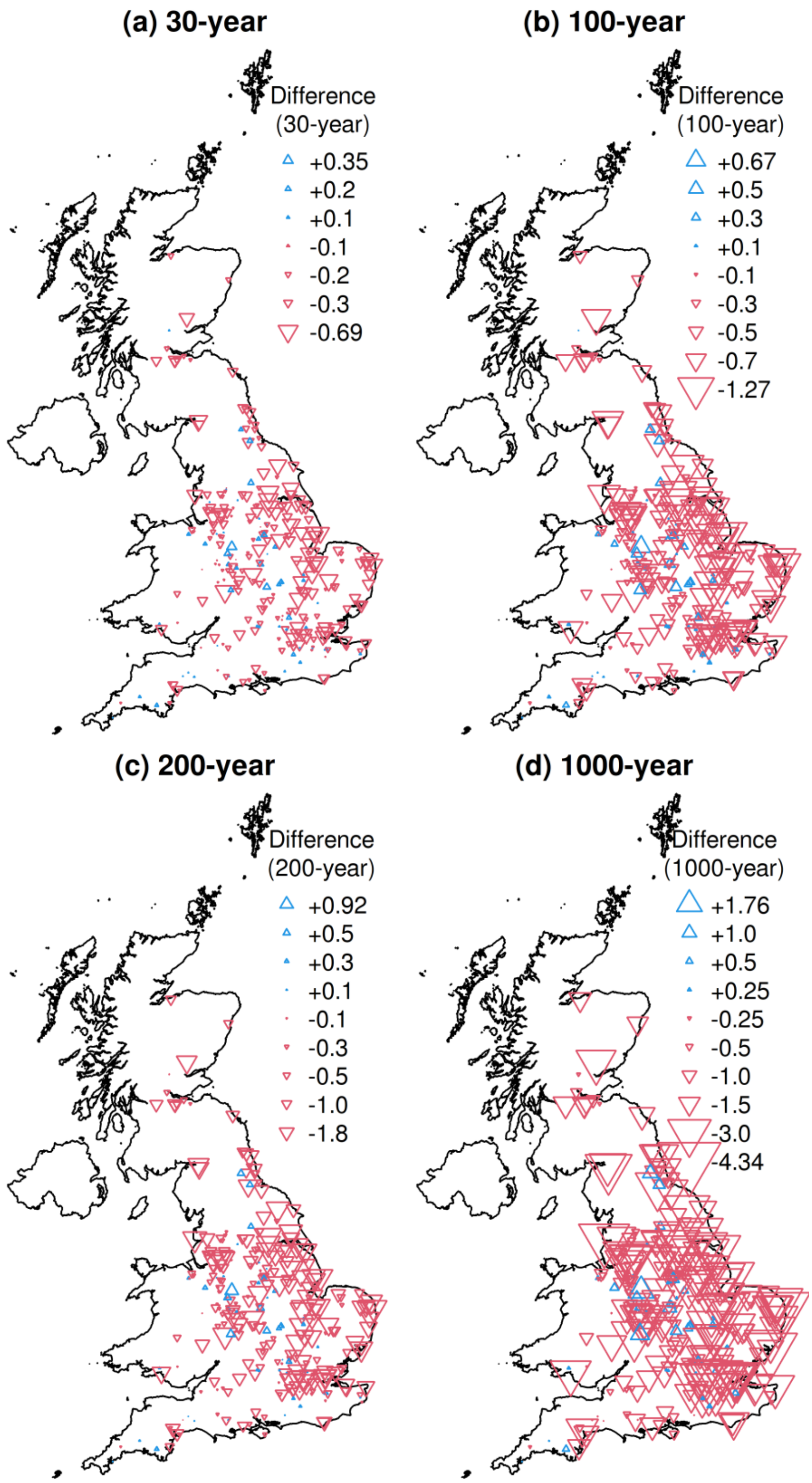

Figure 32 **Difference between 2025 and 2008 methods for 30- (a), 100- (b), 200- (c) and 1000- (d) year growth factor for 285 ungauged urbanized catchments. Positive values show 2025 gives larger values.**

Figure 33 plots the 2025 1000-year growth factor, calculated using the methods developed in Sections 6 and 7 against the 1000-year growth factor using the methods of Kjeldsen *et al.* (2008) for catchments over 40 km^2 and Vesuviano *et al.* (2024) for catchments under 40 km^2. This shows close correspondence between 2008 or small-catchment and 2025 growth factors for values up to about 5. Above this, there are many catchments where the 2025 growth factor can take any value between about 2 and 7, while the 2008 method's growth factors fall into clear bands between about 5 and 7. These clear bands suggest the repeated selection of similar pooling-groups using the small-catchment method. It is important to note that over 87% of the ungauged catchments are small, so over 87% of the points shown correspond to small-catchment (i.e. Vesuviano *et al.*, 2024) growth factors. Small-catchment pooling-groups are selected using only two catchment descriptors, while 2025-method pooling-groups are selected using five catchment descriptors. Table 9 demonstrates that the 2025 method is more suitable than the small-catchment method on the small catchments in the NRFA Peak Flow v12.1 gauged dataset. Unfortunately, the ungauged dataset does not contain "true" values against which to compare the 2025-, 2008- and SC-method growth factors.

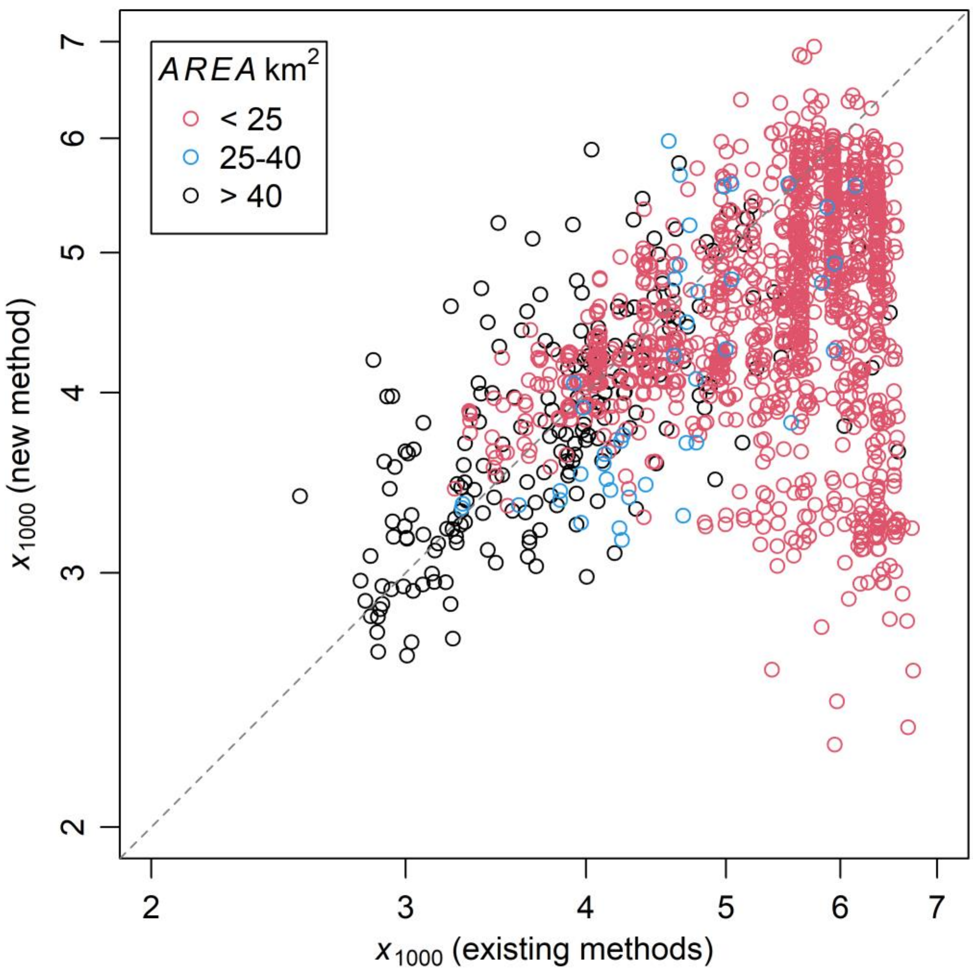

Figure 33 1000-year growth factors: 2025 vs 2008 method for 1563 ungauged catchments.

8. Summary

The previous FEH statistical method was released in 2008. Since then, sporadic updates have been made to extend the methodology, but very little work has been done to integrate major advances in the underlying data. Since 2018, four of the catchment descriptors used in the FEH statistical method have been updated. These were released publicly between 2019 and 2025. Collection and quality control of annual maximum (AMAX) flows from the mid-2000s to the early 2020s has greatly increased the quantity and quality of model calibration data available.

This study assessed the catchment descriptors used to estimate *QMED*, the median annual flood, finding that updated versions of the descriptors currently used to model it, in the same transformations, were the most appropriate choices. Recalibration of the *QMED* model slightly improved the fraction of variance explained and greatly improved the representation of sampling error and inter-site correlation, while negligibly affecting the model structural error. A sensitivity study indicated that no individual catchment had an unduly high influence on the model parameter fitting.

Donor transfer was re-assessed and, as a result, the recommended number of donors was increased from six to eight. It was confirmed that donor transfer is a mandatory part of the *QMED* estimation procedure, that catchment descriptors should not be used at all to select or exclude potential donors, unless they clearly misrepresented the potential donor catchment, and that selection should be based on catchment centroid-to-centroid distance only.

The urban adjustment factor was also recalibrated. Changes were made to the numerical constants but not the structure of the model. However, *URBAN* was redefined to be equal to *URBEXT*, based on research published between the release of the 2008 method and updates, and now. This change will have been partly compensated for by changes to the numerical constants.

A re-assessment of the pooling method recommended the inclusion of $BFIHOST_{19SCALED}$ in the similarity distance measure (SDM), as a fifth descriptor in addition to updated versions of the existing ones. Optimal weights within this new SDM were very different from those in the 2008 SDM, with much less emphasis given to *AREA*. The recommended pooling group length was increased to 800 years. The reduced importance of *AREA* in forming pooling-groups removes the need for a small catchment-specific SDM, which is no longer recommended.

The effects of urbanization on flood frequency relationships were difficult to extract from random variability. A recalibrated model is recommended for adjusting *L*-CV to account for catchment urbanization. However, it is no longer recommended to adjust the pooled *L*-SKEW, as it has almost zero observed relationship with catchment urbanization.

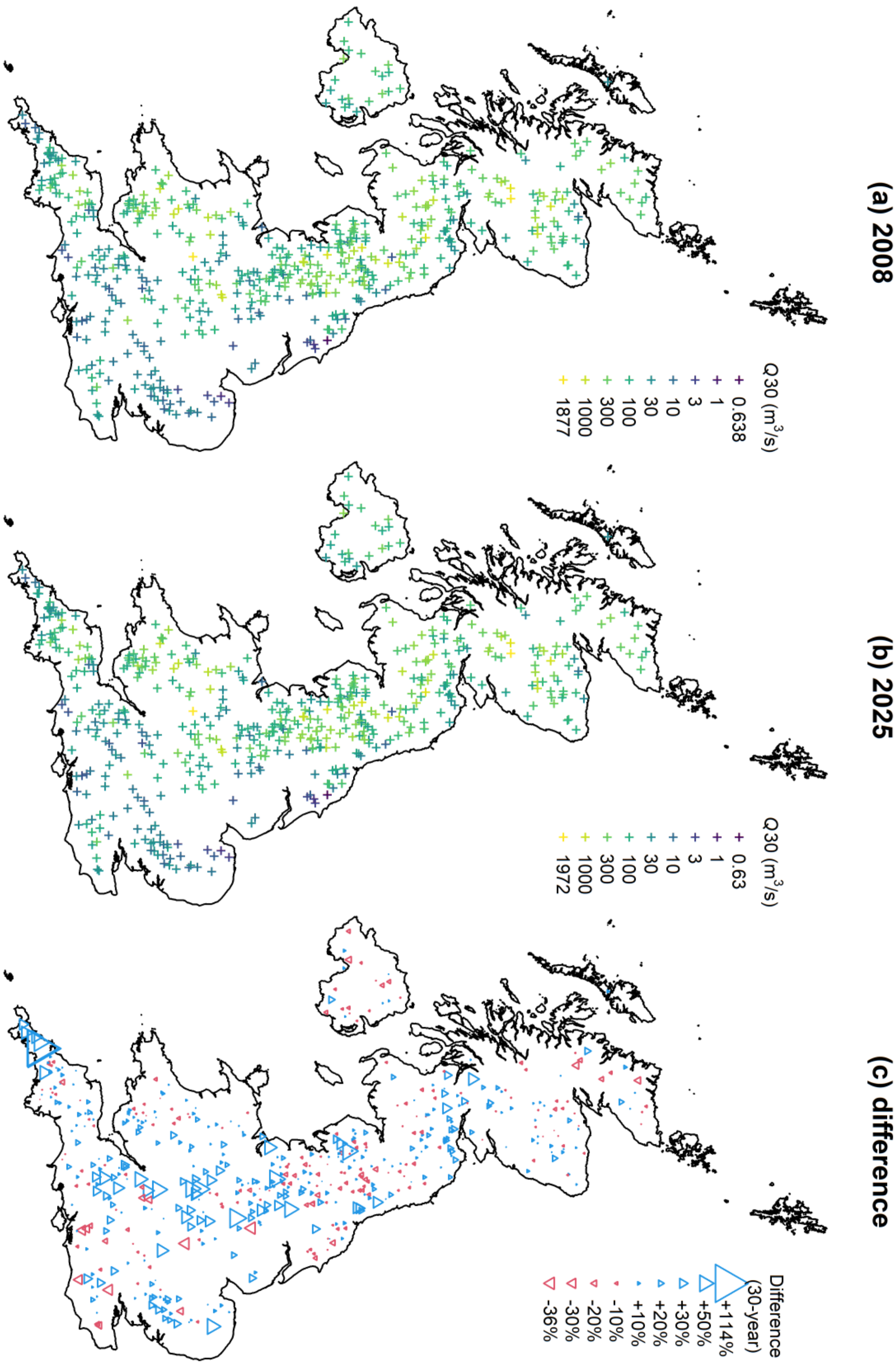

Figure 34 **2025 values divided by 2008 values of 30-year peak flow for 530 pooling-suitable catchments.**

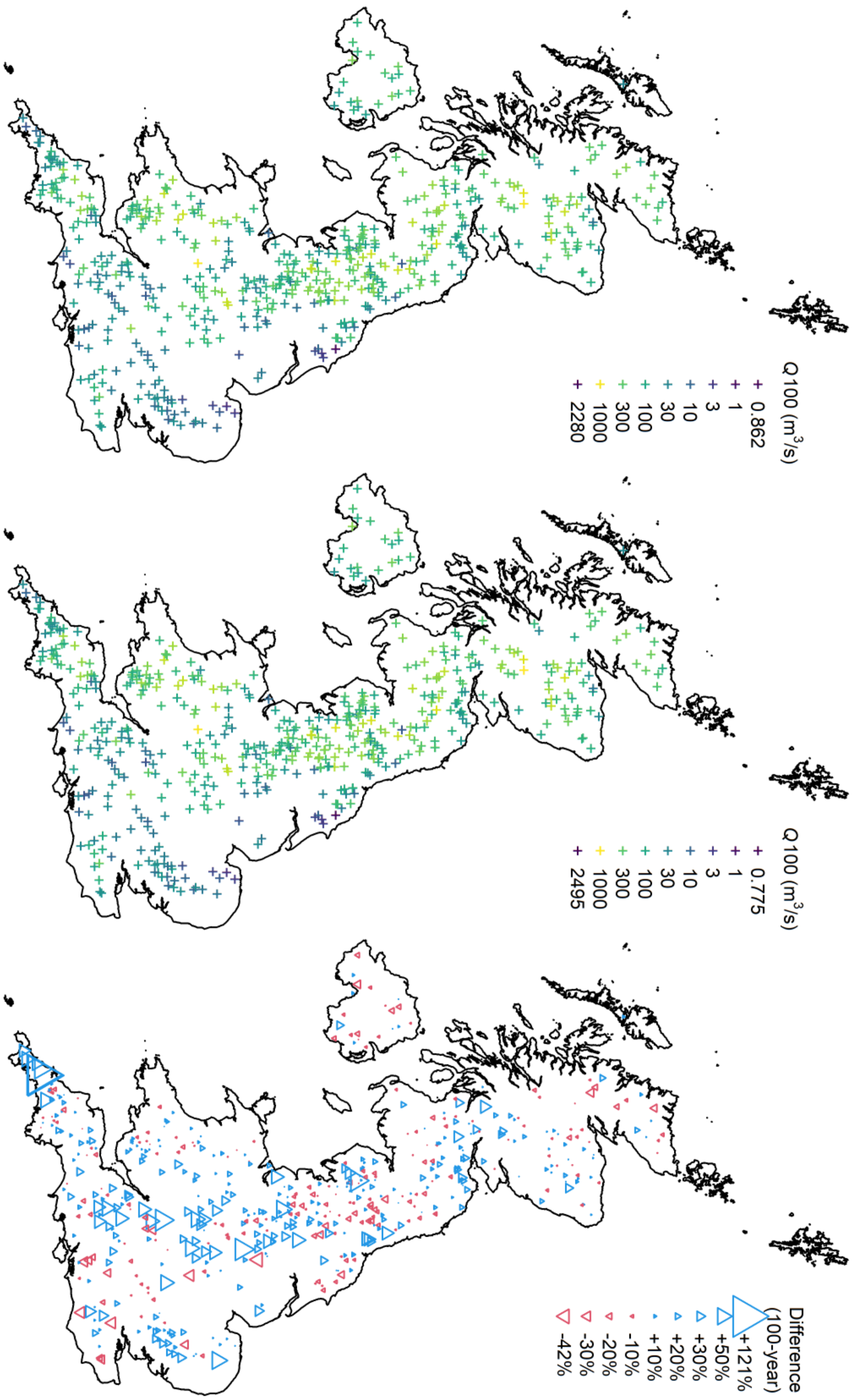

Figure 35 **2025 values divided by 2008 values of 100-year peak flow for 530 pooling-suitable catchments.**

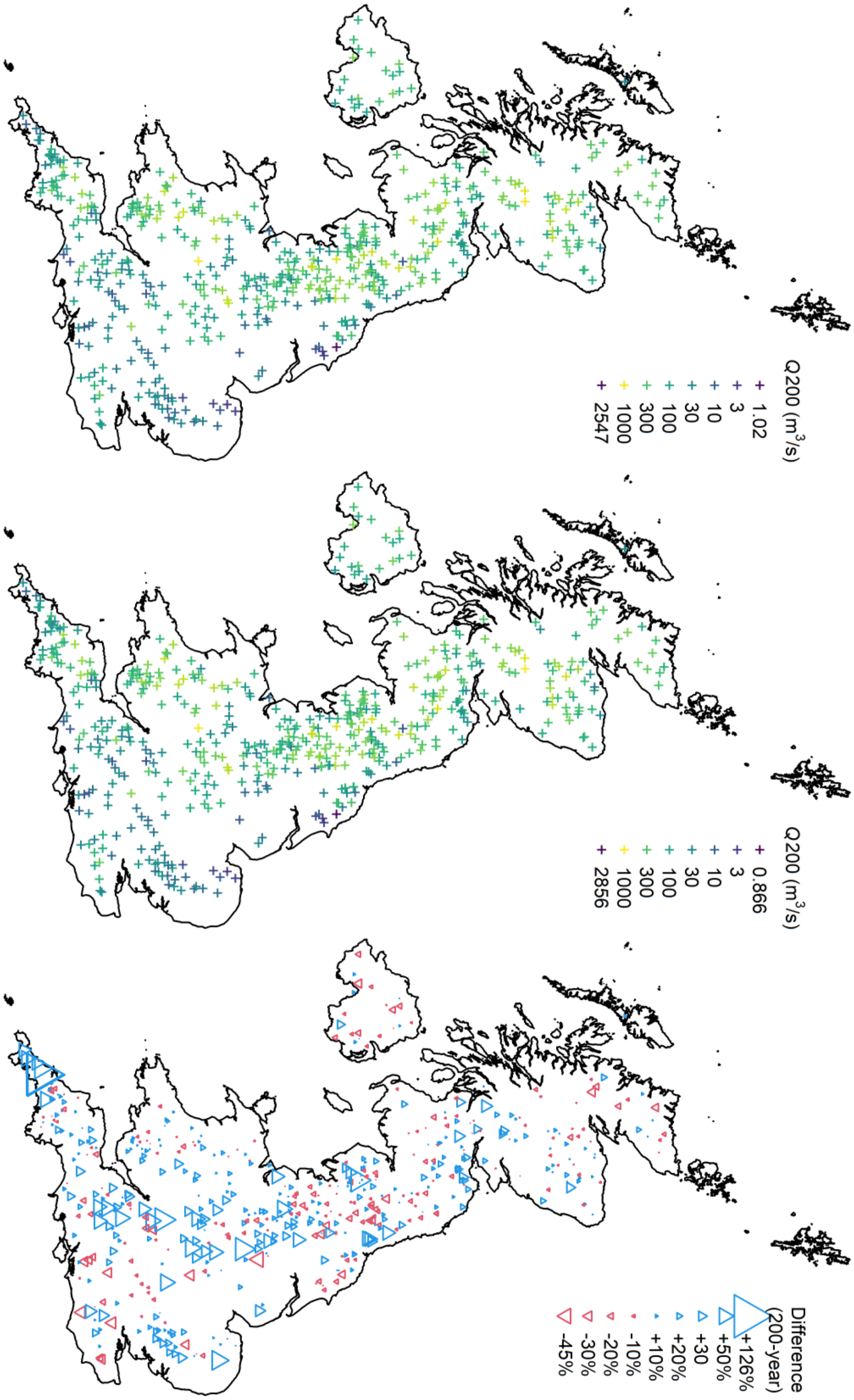

Figure 36 **2025 values divided by 2008 values of 200-year peak flow for 530 pooling-suitable catchments.**

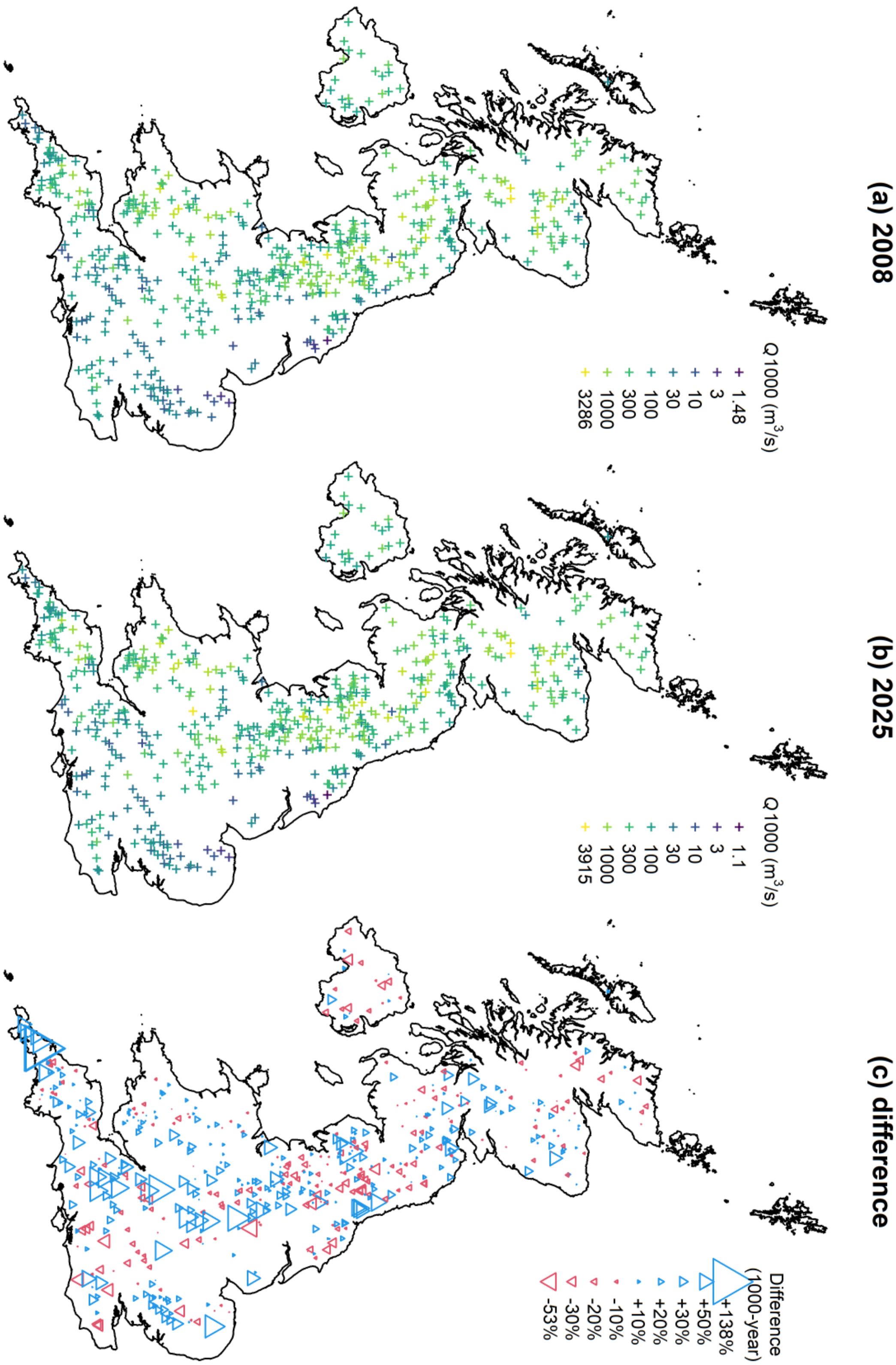

Figure 37 2025 values divided by 2008 values of 1000-year peak flow for 530 pooling-suitable catchments.

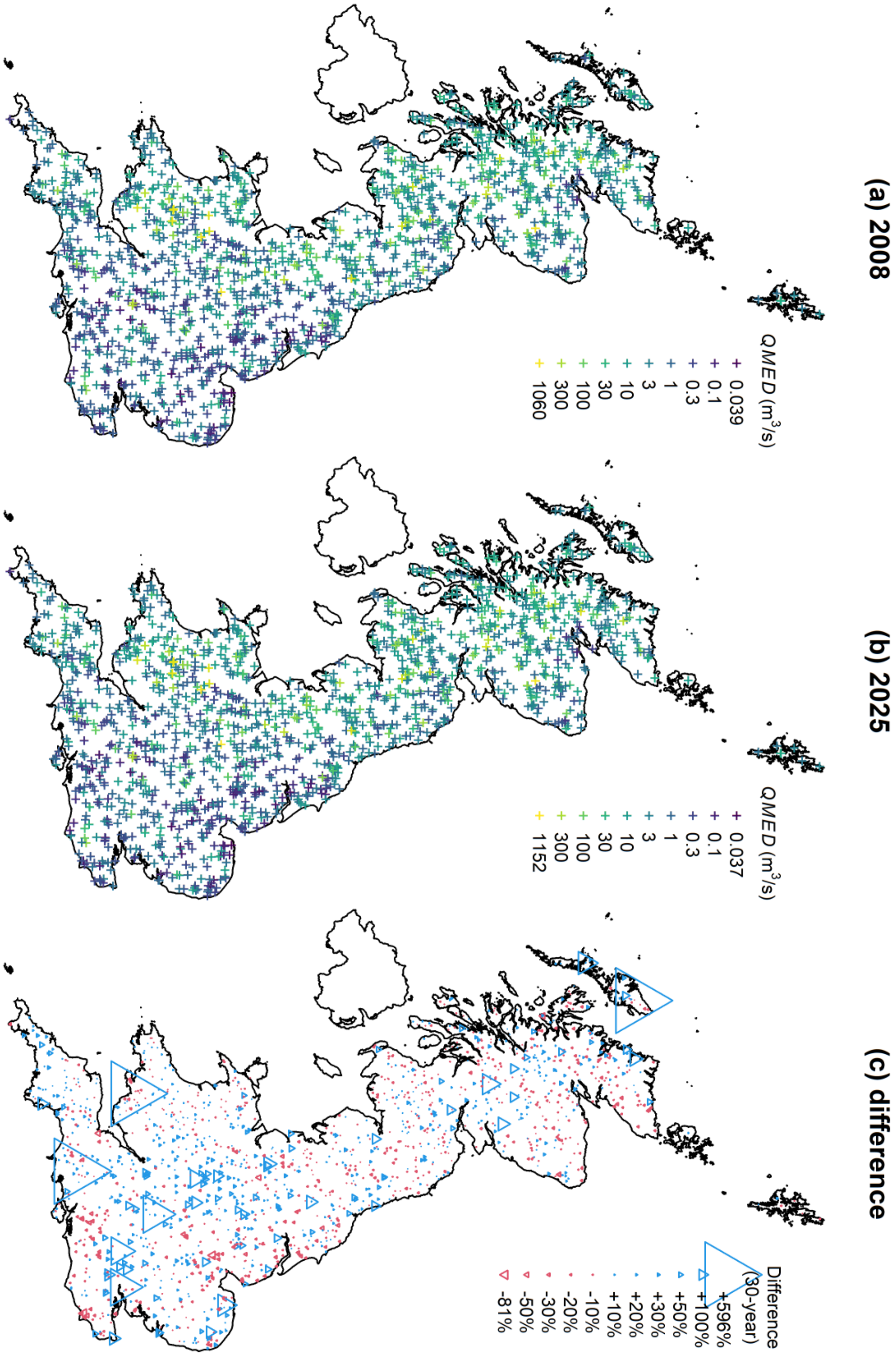

Figure 38 **2025 values divided by 2008 values of 30-year peak flow for 1563 ungauged catchments.**

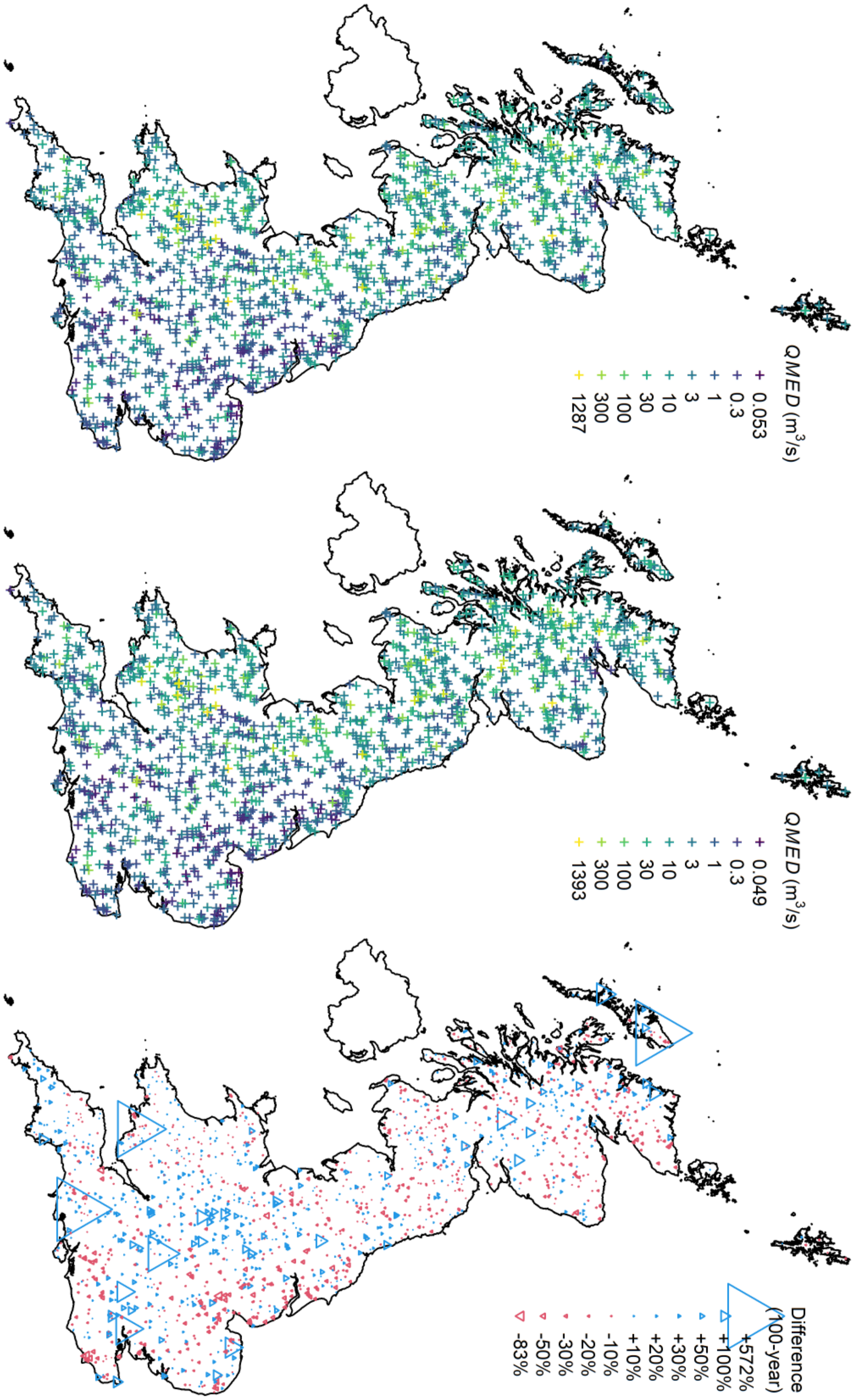

Figure 39 2025 values divided by 2008 values of 100-year peak flow for 1563 ungauged catchments.

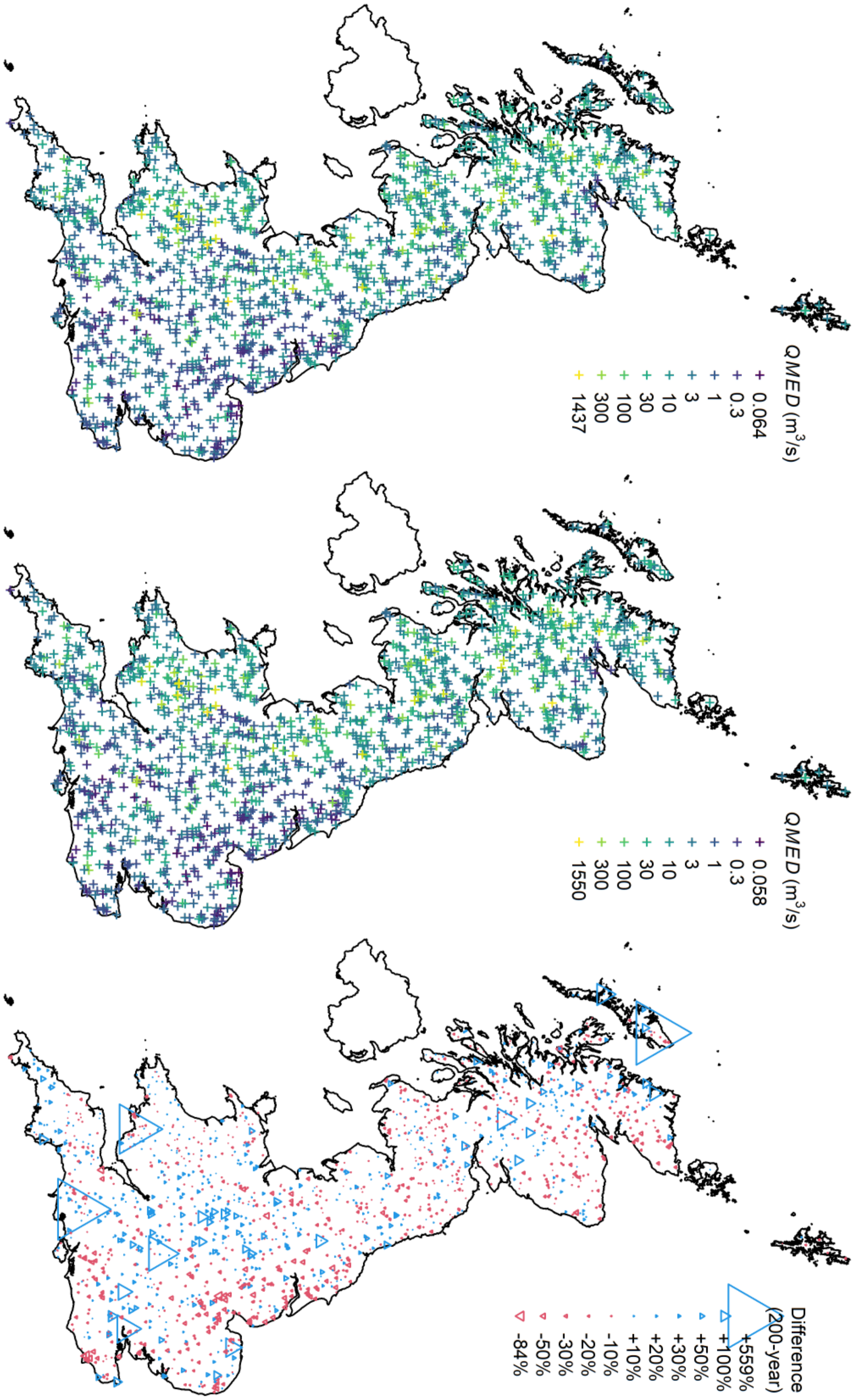

Figure 40 **2025 values divided by 2008 values of 200-year peak flow for 1563 ungauged catchments.**

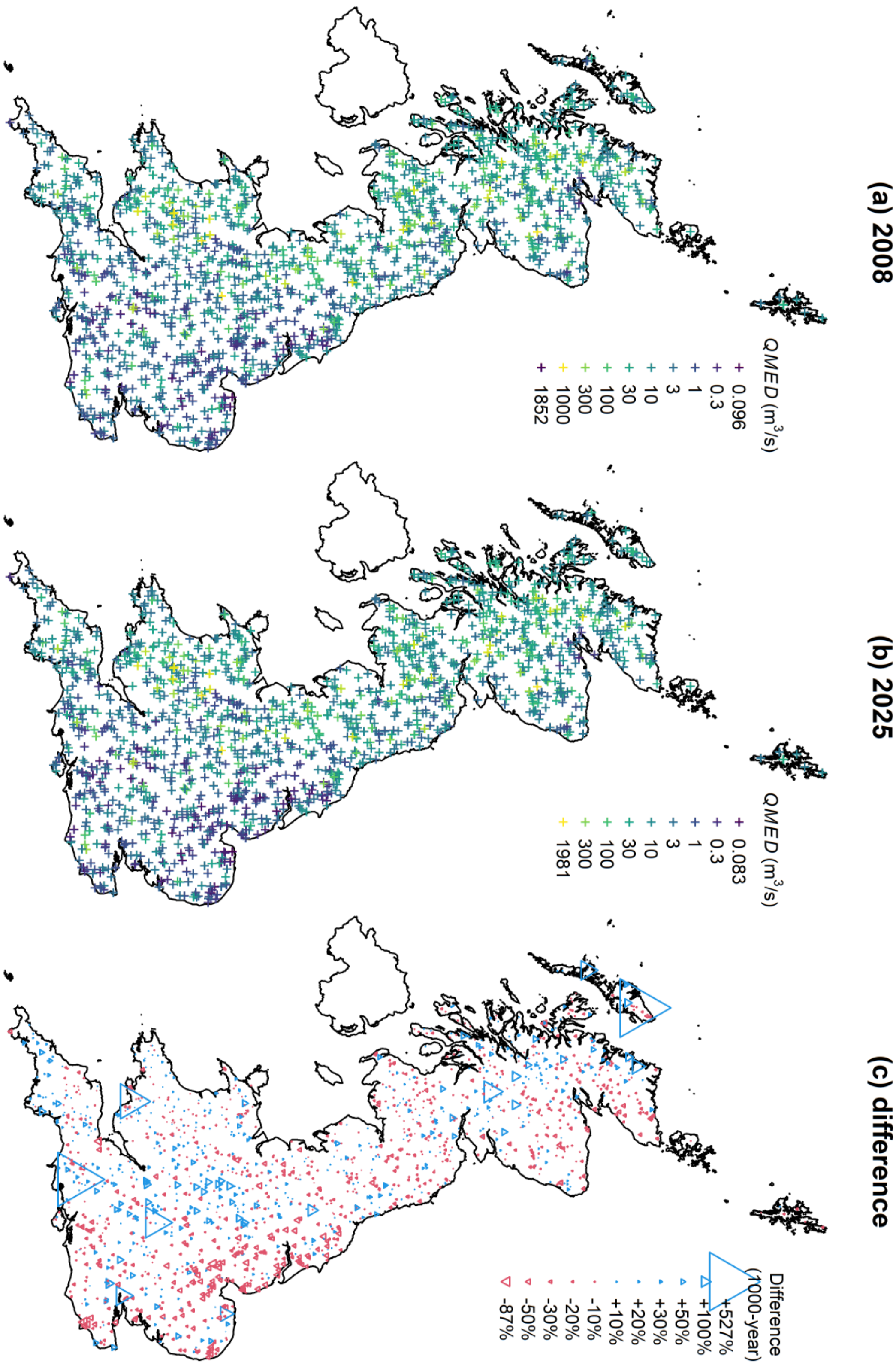

Figure 41 2025 values divided by 2008 values of 1000-year peak flow for 1563 ungauged catchments.

9. Recommendations

- For ungauged catchments, unless local knowledge suggests otherwise, we recommend using the new data (flow data and catchment descriptors) and the 2025 methods for *QMED* and Pooling.
- For ungauged catchments, *QMED* estimation with catchment descriptors should always include donor adjustment, and pooling-groups should be used to calculate growth curves.
- When using donor adjustment, centroid-centroid proximity should be the primary metric used, but local knowledge of record quality and length, cases where catchment descriptors may misrepresent the donor catchment, and mitigating factors such as influential reservoirs, may strongly suggest using other donor stations. This is valuable secondary information to be made use of when reviewing donor groups.
- For gauged catchments, gauged POT or AMAX data can be used to estimate *QMED* if they are of sufficient quality. Donor transfer should not be used in this case. The enhanced single-site method (pooling with at-site AMAX incorporated) should be used to estimate growth curves.
- For pre-reservoir analysis, use pre-reservoir *FARL* and *SAAR* (if pre-1990), but make use of the 2025 *QMED* and Pooling methods. If the NRFA has already rejected post-reservoir data, then *FARL* and $FARL_{2015}$ represent the pre-reservoir value.
- For gauged catchments, regarding future developments, use full period-of-record information if available, or $SAAR_{9120}$ if not. Use 2025 methods in all cases. Use *SAAR* for the period of relevance (9120 for future and current, 6190 for pre-1990 historical), using the HadUK estimates (available from Met Office website).
- As discussed in the FEH Local project, if you have a better knowledge of *AREA*, *FARL*, *BFIHOST* or *FPEXT*, then use that, but compare with the FEH Web Service values.
- Gauged *BFI* and *BFIHOST* (including $BFIHOST_{19}$ and $BFIHOST_{19SCALED}$) don't describe precisely the same thing (gauged *BFI* is observed, *BFIHOST* versions are modelled). Do not use known baseflow index (gauged *BFI*) as a substitute for $BFIHOST_{19SCALED}$ in the FEH statistical method.
- When applying the FEH statistical method, continue to follow the same diagnostic steps currently recommended (interrogate stations with high discordancy, check pooling groups for homogeneity, etc.).
- For small catchments, we recommend using the same *QMED* and Pooling procedures as for larger catchments, and not to use the alternative small catchments methods.

References

Bayliss AC, Black KB, Fava-Verde A, Kjeldsen TR. 2006. *$URBEXT_{2000}$ – A new FEH catchment descriptor. Calculation, dissemination and application*. London: Department for Environment, Food and Rural Affairs.

Boorman DB, Hollis JM, Lilly A. 1995. *Hydrology of soil types: a hydrologically-based classification of the soils of the United Kingdom (Institute of Hydrology Report 126)*. Wallingford: Institute of Hydrology.

Brand MJD, Mitchell G. 1993. CORINE land cover (Ireland) – the northern perspective. Proceedings of the Association for Geographic Information Conference, Birmingham.

CPRE. 1993. *The regional lost land – land use change in England's regions and counties 1945-1990*. Report to CPRE by Environment Information Services.

Environment Agency. 2017. *Making better use of local data in flood frequency estimation (SC130009/R)*. Bristol: Environment Agency.

Fuller RM, Groom GB, Jones AR. 1994. The Land Cover Map of Great Britain: an automated classification of Landsat Thematic Mapper data. *Photogrammetric Engineering and Remote Sensing* 60, 553-562.

Fuller RM, Smith GM, Sanderson JM, Hill RA, Thomson AG. 2002. The UK Land Cover Map 2000: construction of a parcel-based vector map from satellite images. *Cartographic Journal* 39, 15-25.

Griffin A, Young A, Stewart L. 2019. Revising the *BFIHOST* catchment descriptor to improve UK flood frequency estimates. *Hydrology Research* 50.6, 1508-1519.

Hosking JRM, Wallis JR. 1997. *Regional Frequency Analysis: An Approach Based on* L-*Moments*. Cambridge: Cambridge University Press.

Hughes M, Hornby DD, Bennion H, Kernan M, Hilton J, Phillips G, Thomas R. 2004. The development of a GIS-based inventory of standing waters in Great Britain together with a risk-based prioritisation protocol. *Water, Air and Soil Pollution: Focus* 4, 73-84.

Keller VDJ, Tanguy M, Prosdocimi I, Terry JA, Hitt O, Cole SJ, Fry M, Morris DG, Dixon H. 2015. CEH-GEAR: 1 km resolution daily and monthly areal rainfall estimates for the UK for hydrological and other applications. *Earth System Science Data* 7, 143-155.

Kjeldsen TR. 2010. Modelling the impact of urbanization on flood frequency relationships in the UK. *Hydrology Research* 41.5, 391-405.

Kjeldsen TR, Jones D. 2007, Estimation of an index flood using data transfer in the UK. *Hydrological Sciences Journal* 52.1, 86-98.

Kjeldsen TR, Jones DA, Bayliss AC. 2008. *Improving the FEH statistical procedures for flood frequency estimation*. Bristol: Environment Agency.

Kjeldsen TR, Jones DA, Morris DG. 2014. Using multiple donor sites for enhanced flood estimation in ungauged catchments. *Water Resources Research* 50, 6646-6657.

Kuhn M, Wing J, Weston S, Williams A, Keefer C, Engelhardt A, Cooper T, Mayer Z, Kenkel B, R Core Team, Benesty M, Lescarbeau R, Ziem A, Scrucca L, Tang Y, Candan C, Hunt T. 2023. *caret: Classification and Regression Training. R package version 6.0-94* [online]. https://cran.r-project.org/package=caret

Met Office. 2011. *National Meteorological Library and Archive Fact sheet 2 – Thunderstorms*. Exeter: Met Office.

Met Office, Hollis D, McCarthy M, Kendon M, Legg T. 2023. *HadUK-Grid Gridded Climate Observations on a 1km grid over the UK, v1.2.0.ceda (1836-2022)*. NERC EDS Centre for Environmental Data Analysis [data set].

Morton RD, Marston CG, O'Neil AW, Rowland CS. 2020. *Land Cover Map 2019 (25m rasterised land parcels, GB)*. NERC Environmental Information Data Centre [data set].

Robson A, Reed D. 1999. *Statistical procedures for flood frequency estimation (Flood Estimation Handbook volume 3)*. Wallingford: Institute of Hydrology.

Rowland CS, Morton RD, Carrasco L, McShane G, O'Neil AW, Wood CM. 2017. *Land Cover Map 2015 (25m raster, GB)*. NERC Environmental Information Data Centre [data set].

Spackman E. 1993. Calculation and Mapping of Rainfall Averages for 1961-90. *Proceedings of British Hydrological Society Symposium*. Manchester: University of Salford, 15 December 1993.

Stewart E, Faulkner D, Formetta G, Griffin A, Haxton T, Prosdocimi I, Vesuviano G, Young A. 2024. *Estimating flood peaks and hydrographs for small catchments (Phase 2). Report SC090031/R0*. Bristol: Environment Agency.

Vesuviano G, Griffin A, Stewart L, Young A. 2024. *Estimating flood peaks and hydrographs for small catchments: R4 – Estimating the median annual flood (*QMED*) in small catchments*. Bristol: Environment Agency

Vesuviano, G., Wallbank, J., Griffin, A., Miller, J. 2025. New FEH Catchment Descriptors 2025: Updates to FARL, SAAR, URBEXT and BFIHOST. UK Centre for Ecology & Hydrology, Wallingford, UK. Available from: https://www.ceh.ac.uk/services/feh-2025-statistical-updates

Wallingford HydroSolutions. 2022. *Urbanisation procedures – WINFAP Technical Guide* [online]. https://winfapdocs.hydrosolutions.co.uk/Pooled-&-QMED-Analysis/Urbanisation-Procedures/

10. Appendix: manual edits to station/AMAX data

Kjeldsen *et al.* (2008) made several manual edits to station and AMAX data. These were all implemented in this project, except in cases where:

- The station was not in the NRFA Peak Flow dataset v12.1,
- The station was in the NRFA Peak Flow dataset v12.1 but not considered suitable for peak flow estimation,
- The change was implemented by the NRFA sometime between 2008 and 2024 (e.g. due to a new rating, period-of-record review, or manually edited time period-relevant *FARL*), so a further manual edit for this study was no longer required.

Removal of stations

A number of stations were removed from the 2008 study and from this study, for the same reason. The removals and reasons were:

- Station 27032: Karstic limestone catchment where observed flood peak data do not relate to the catchment descriptor values.
- Station 27033: Flood flows diverted into catchment. Observed flood flows do not relate to the catchment descriptor values.
- Station 39027: Effect of groundwater abstraction on QMED is unknown but thought to be significant.
- Station 39033: Effect of groundwater abstraction on QMED is unknown but thought to be significant.
- Station 42007: Highly permeable catchment where AMAX reflect seasonal change in groundwater, and effective catchment area unknown.
- Station 80003: Poor measurement of flood flows so flood peak data rejected for use in this project. (note: AMAX in NRFA Peak Flow 12.1 are unchanged from HiFlows-UK 1.1)
- Station 205034: True drainage area unknown but much smaller than DTM area.

A number of stations that were removed from the 2008 study were not removed from this study, for the following reasons:

- Station 7001: Erroneously elevated flows pre-1978 were corrected in 2017.
- Station 21006: Incorrectly abstracted chart data were completely replaced in 2017 and 2022.
- Station 55010: Full period of record was reviewed in 2020.
- Station 69041: Full period of record was reviewed in 2019.
- Station 72017 (formerly 72807): Full period of record was reviewed in 2020.

Modification of *FARL* and *FARL*$_{2015}$

There are seven large reservoirs with downstream catchments that recorded a mix of AMAX data from before and after reservoir operation. If either before-operation or after-operation data are marked as rejected by the NRFA, then only the non-rejected period is used. If AMAX from both before and after reservoir operation are acceptable to use, then the longest period is used and the *FARL* and/or *FARL*$_{2015}$ values are modified to account for the reservoir existing or not. Only one value of *FARL* required a change from the value stored by the NRFA:

- Station 6003: *FARL* increased from 0.813 to 0.985 as only pre-reservoir record used.

Nine *FARL*$_{2015}$ values were recalculated ignoring the upstream reservoir, as nine of the accepted AMAX records used in this study were pre-reservoir:

- Station 6003: *FARL*$_{2015}$ increased from 0.8100 to 0.9493 as only pre-reservoir record used.
- Station 21030: *FARL*$_{2015}$ increased from 0.8144 to 1.0000 as only pre-reservoir record used.
- Station 23002: *FARL*$_{2015}$ increased from 0.8391 to 0.9988 as only pre-reservoir record used.
- Station 23003: *FARL*$_{2015}$ increased from 0.9161 to 0.9762 as only pre-reservoir record used.
- Station 23005: *FARL*$_{2015}$ increased from 0.7854 to 0.9857 as only pre-reservoir record used.
- Station 23015: *FARL*$_{2015}$ increased from 0.9142 to 0.9720 as only pre-reservoir record used.
- Station 28002: *FARL*$_{2015}$ increased from 0.8294 to 0.9915 as only pre-reservoir record used.
- Station 48009: *FARL*$_{2015}$ increased from 0.6292 to 0.9622 as only pre-reservoir record used.
- Station 48011: *FARL*$_{2015}$ increased from 0.9132 to 0.9681 as only pre-reservoir record used.

1st October AMAX

Thirty-one stations had one or more water years with no AMAX followed by a year with two AMAX, the first occurring on 1st October. The reason for this is that the water year starts at 09:00 UTC on 1st October, but the NRFA Peak Flow files did not include the time of the peak (so events occurring between 00:00 and 09:00 on 1st October were assigned to the wrong water year). This occurred at 31 stations:

6008, 11002, 12007, 15013, 21024, 28009, 28015, 28022, 28024, 28056, 31005, 31010, 31025, 32003, 32004, 32029, 38001, 39018, 39036, 42009, 45005, 46006, 47011, 54102, 54114, 67008, 69024, 72014, 84011, 84014 and 205008.

In each case, the AMAX on 1st October was moved to the previous water year.

The NRFA Peak Flow dataset, starting from version 13.0.3, includes timestamps, which fixes this issue when used with WINFAP version 5.2 and above.

Other modifications

Other modifications to single AMAX recorded by Kjeldsen *et al.* (2008) were not repeated here, as all AMAX records had been fully reviewed at least once between 2008 and 2024. No other modifications specific to this project were made.

Commonly Excluded Stations

There are a number of stations regularly excluded from FEH statistical analyses by Cyfoeth Naturiol Cymru (CNC) and Wallingford HydroSolutions (WHS). These are listed below with brief reasoning. It should be noted that most of them were also excluded from the pooling calibration dataset used to derive the new pooling methods.

Station	Name	Rejected by	In calibration set?
26013	Driffield Trout Stream at Driffield	WHS	YES
26014	Water Forlornes at Driffield	WHS	YES
26016	Gypsey Race at Kirby Grindalythe	WHS	NO
27073	Brompton Beck at Snainton Ings	WHS	YES
33032	Heacham at Heacham	WHS	NO
34012	Burn at Burnham Overy	WHS	YES
39019	Lambourn at Shaw	WHS	YES
39020	Coln at Bibury	WHS	YES
39035	Churn at Cerney Wick	WHS	YES
41023	Lavant at Graylingwell	WHS	YES
42006	Meon at Mislingford	WHS	YES
42007	Alre at Drove Lane Alresford	WHS	NO
42008	Cheriton Stream at Sewards Bridge	WHS	YES

42009	Candover Stream at Borough Bridge	WHS	NO
42012	Anton at Fullerton	WHS	NO
43004	Bourne at Laverstock	WHS	NO
43008	Wylye at South Newton	WHS	YES
43029	Wylye at Brixton Deverill	WHS	NO
44013	Piddle at Little Puddle	WHS	YES
30013	Heighington Beck at Heighington	CNC	NO
43004	Bourne at Laverstock	CNC	NO
78005	Kinnel Water at Bridgemuir	CNC	NO
79002	Nith at Friars Carse	CNC	NO
201007	Burn Dennet at Burndennet	CNC	NO
203033	Upper Bann at Bannfield	CNC	NO
203043	Oonawater at Shanmov	CNC	NO

WHS Reasoning

- Topographic and contributing catchment issue (where it is explicitly cited in the NRFA description): where these are different, it means that the descriptors used to ascertain 'similarity' and model *QMED* (and hence to estimate the ratio of gauged to modelled *QMED* for donor transfer) may not be correct. In addition, this implies that flows in the river are likely to be more impacted by groundwater mechanisms. Groundwater may dominate or significantly influence AMAX and POT.
- Ephemeral catchments: This is an indication that some rivers within the catchment do not flow for a majority of the time, hence that flow mechanisms in the catchment are not 'typical'.

- Daily flows and responsiveness to rainfall: In most catchments, the 'peak flow' and the distribution of AMAX are a combination of the types of rainfall events and the response of the catchment to these. Catchments are rejected where AMAX flows are clearly dominated by groundwater contributions.

CNC reasonings

- Very high baseflow index values.
- Evidence of non-modular flow for the highest AMAX values, along with a lack of accurate rating for AMAX events that clearly exceeded bankfull.
- Unknown sluice operation leading to unknown impacts on peak flows.

11. Appendix: Description of ungauged dataset

The ungauged dataset is a selection of 1563 points on the IHDTM river network within Great Britain. They were chosen to be representative of the GB river network, covering wide ranges and combinations of *AREA*, *BFIHOST*, *FARL*, *SAAR*, *URBEXT* and *FPEXT*, but a Monte Carlo approach was used to select them – another run of the code with a different randomisation seed would produce a qualitatively similar but subtly different dataset.

Due to the size of the dataset, and a lack of local knowledge in all areas, it is difficult to give precise reasons for large differences between catchment descriptor versions at any ungauged locations, but large-scale patterns in version differences are discussed. Differences between old and new *QMED* and Pooling method results usually come back to differences in the catchment descriptors. Figure 42 shows the distribution several catchment descriptors for the IHDTM river network in Great Britain.

This dataset is not expected to be published or used operationally, but if there is sufficient interest, such a public dataset could be investigated. This should be linked up with work in the Natural Capital and Ecosystem Assessment Programme on water quantity network representativeness.

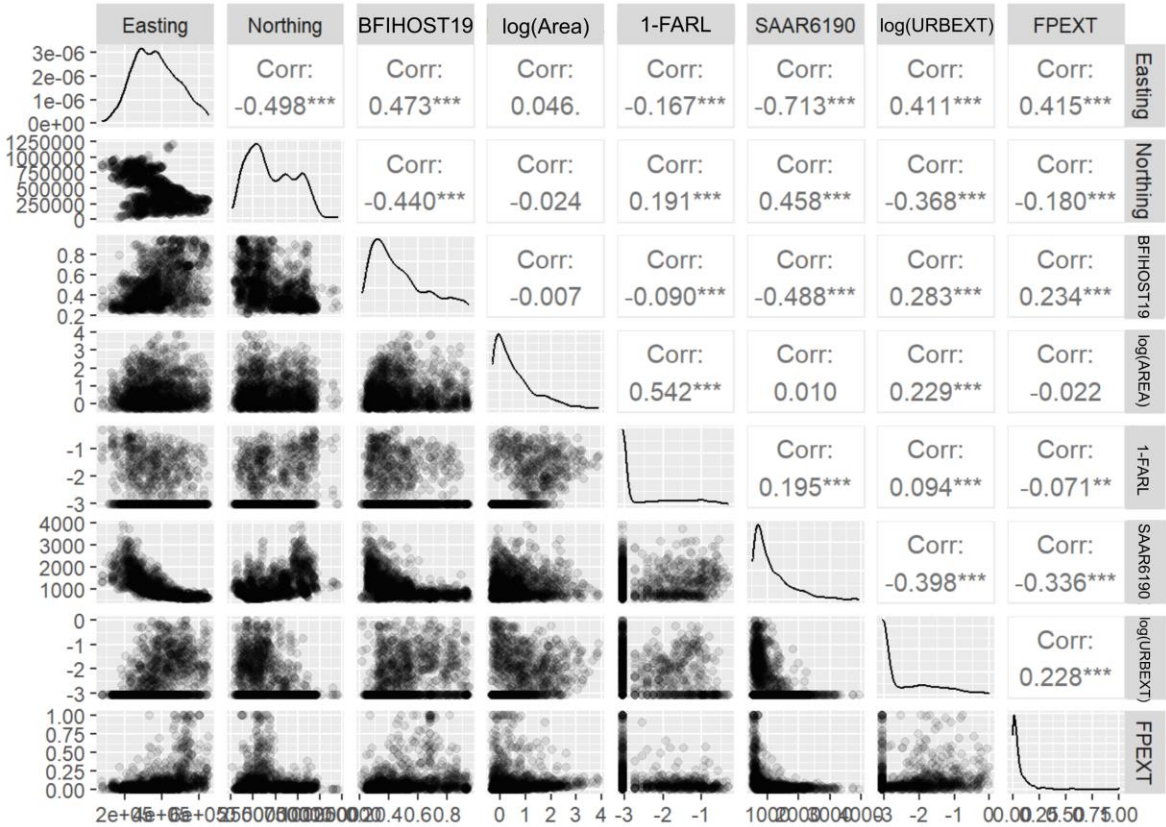

Figure 42: Distribution of catchment descriptors across the IHDTM river network in Great Britain.

12. Appendix: *SAAR* grid offsets

The $SAAR_{6190}$ and HadUK ($SAAR_{9120}$) grids are offset by 500 metres in each direction from each other in a way that we have insufficient information to convert from one to the other. Figure 43 shows an example of this. If we consider $SAAR_{6190}$ to be on the black grid and HadUK to be on the blue, the different grids give subtly different summaries of the underlying “truth”.

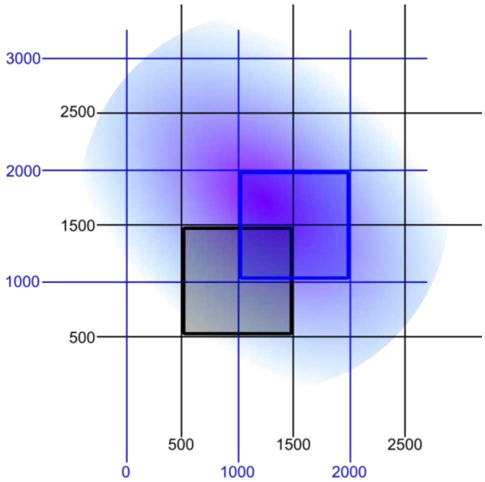

Figure 43 Illustration of grid offsets. *SAAR*6190 (black) and HadUK (blue) are offset by 500 m in each direction, and so give subtly different estimates of the underlying truth (shading).

13. Appendix: Worked Examples

13.1 Ungauged catchment

The catchment used in this example is Dover Beck at Lowdham (NRFA number 28060). While it is a gauged NRFA Peak Flow catchment, its AMAX record is not suitable for *QMED* estimation or pooling, so an ungauged analysis is more appropriate. The relevant catchment descriptors (Peak Flow Dataset V13.1) are:

$AREA$: 62.6725 (km^2) | $BFIHOST_{19SCALED}$: 0.690 | $FARL_{2015}$: 0.9972

$SAAR_{9120}$: 703 (mm) | $FPEXT$: 0.063 | $URBEXT_{2015}$: 0.0557

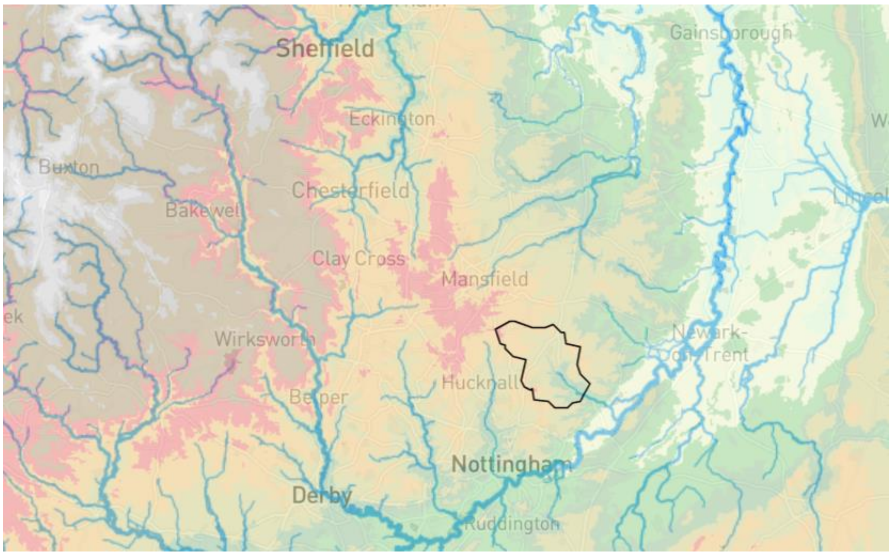

Figure 44: Map of catchment and surrounding area for Dover Beck at Lowdham. Colour indicates elevation.[1]

QMED estimation (rural)

$QMED_{rural}$ is estimated using the FEH statistical 2025 $QMED_{rural}$ equation:

$$QMED_{rural} = 6.8247AREA^{0.8499}0.1780^{\left(\frac{1000}{SAAR_{9120}}\right)}FARL_{2015}{}^{3.0450}0.0321^{BFIHOST_{19SCALED}{}^{2}}$$

[1] Taken from nrfa.ceh.ac.uk © UKCEH 2025. For Great Britain: Contains Ordnance Survey data © Crown copyright and database right 2025.

Inserting the relevant catchment descriptor values gives a $QMED_{rural}$ estimate of

$$QMED_{rural} = 3.805 \text{ m}^3/\text{s}$$

The factorial standard error of the FEH statistical 2025 $QMED_{rural}$ equation is 1.431, so the 95% confidence interval of the $QMED_{rural}$ estimate is

$$1.885 \text{ m}^3/\text{s} < QMED_{rural} < 7.682 \text{ m}^3/\text{s}$$

Donor transfer

Donor transfer uses eight donor catchments, whose AMAX records must be suitable for *QMED* estimation. Catchments are selected based purely on the distance between their centroids and the catchment of interest's centroid. Dover Beck at Lowdham's centroid is located at:

CEasting: 460975 **CNorthing: 351551**

The eight "suitable for *QMED*" catchments with the closest centroids (Peak Flow V13.1) are given in Table 11.

Table 11 Key data for donor catchments for Dover Beck.

Catchment	CEasting (centroid easting)	CNorthing (centroid northing)	Distance (km)	Weight	*QMED* (gauged)	*QMED* (CD, rural)
28055	428773	349408	32.3	0.1379	14.400	13.872
28024	476502	316706	38.1	0.1256	34.366	47.600
28058	422845	349881	38.2	0.0447	10.600	11.281
30005	490684	325071	39.8	0.0554	6.901	4.940
31023	495230	325298	43.2	0.0203	1.777	1.261
28070	426300	382090	46.2	0.0160	4.302	4.641
28008	412867	354823	48.2	0.0076	86.869	74.172
28011	418164	376340	49.5	0.0039	113.921	124.141

Applying donor transfer gives a $QMED_{donor}$ estimate of

$$\boldsymbol{QMED_{donor} = 3.837\ \mathrm{m^3/s}}$$

The factorial standard error of the estimate after donor transfer is 1.373, so the 95% confidence interval of the $QMED_{donor}$ estimate $\frac{QMED}{fse^{1.96}} < QMED < QMED \times fse^{1.96}$ is

$$\boldsymbol{2.061\ \mathrm{m^3/s} < QMED_{donor} < 7.142\ \mathrm{m^3/s}}$$

QMED estimation (urban adjustment)

It is recommended to apply the urban adjustment to $QMED_{donor}$ for all catchments, including essentially rural ones, to avoid sudden step changes in *QMED* estimates if a catchment's urban level increases from below to above the threshold for "essentially rural". The urban adjustment for the FEH statistical 2025 method is

$$QMED_{urban} = QMED_{donor}(1 + 0.3URBEXT_{2015})^{1.8838} \times \left(1 + 0.3URBEXT_{2015}\left(\frac{PR_{IMP}}{67.0674 - 63.82BFIHOST_{19SCALED}} - 1\right)\right)^{3.52}$$

Inserting the relevant catchment descriptor values, and using the recommended value of $PR_{IMP} = 70$, gives a $QMED_{urban}$ estimate of

$$\boldsymbol{QMED_{urban} = 4.713\ \mathrm{m^3/s}}$$

The factorial standard error of the urban-adjusted *QMED* is 1.379, so the 95% confidence interval of the $QMED_{rural}$ estimate is

$$\boldsymbol{2.510\ \mathrm{m^3/s} < QMED_{rural} < 8.848\ \mathrm{m^3/s}}$$

Pooling

The similarity distance measure used in the FEH statistical 2025 method for catchments of any size (i.e. below and above 25 km^2) is:

$$SDM5_{ij} = \sqrt{\begin{aligned}&1.74\left(\frac{\ln AREA_i - \ln AREA_j}{1.3207}\right)^2 + 1.63\left(\frac{\frac{1000}{SAAR_{9120,i}} - \frac{1000}{SAAR_{9120,j}}}{0.3566}\right)^2 + 0.26\left(\frac{FARL^2_{2015,i} - FARL^2_{2015,j}}{0.0976}\right)^2 \\ &+ 0.55\left(\frac{FPEXT_i - FPEXT_j}{0.0439}\right)^2 + 0.82\left(\frac{\frac{1}{BFIHOST_{19SCALED,i}} - \frac{1}{BFIHOST_{19SCALED,j}}}{0.6610}\right)^2\end{aligned}}$$

The recommended pooling-group length is 800 years. Using the above SDM, the pooling-group (based on Peak Flow V13.1) for Dover Beck at Lowdham is given in Table 12.

In practice, proposed pooling groups should be checked for good data quality, known impacts on flood frequency (e.g. reservoirs), very different levels of urbanization, *L*-moment homogeneity, etc. Here we assume the pooling group is made of suitable stations for this example.

Table 12: Key data for pooling group stations for Dover Beck.

Catchment	№ AMAX	SDM_5	*L*-CV	*L*-CV weight	*L*-SKEW	*L*-SKEW weight
30004	60	0.663	0.226	0.086	0.029	0.071
26013	12	0.682	0.276	0.030	0.237	0.032
38002	81	0.731	0.300	0.091	0.072	0.075
34012	56	0.839	0.257	0.073	0.033	0.066
26003	61	0.989	0.249	0.068	−0.004	0.066
33032	54	1.041	0.297	0.063	0.134	0.063
54036	50	1.066	0.331	0.060	0.319	0.062
33054	46	1.079	0.231	0.058	0.189	0.060
7010	17	1.103	0.185	0.034	0.285	0.039
36003	62	1.110	0.312	0.063	0.084	0.065
53023	46	1.133	0.236	0.057	0.155	0.060
39028	54	1.133	0.242	0.060	−0.002	0.063
36007	57	1.134	0.378	0.061	0.108	0.064

Catchment	№ AMAX	*SDM*5	*L*-CV	*L*-CV weight	*L*-SKEW	*L*-SKEW weight
38004	63	1.187	0.329	0.061	0.154	0.065
7011	10	1.220	0.494	0.023	0.554	0.026
36004	55	1.223	0.304	0.057	0.176	0.062
39042	50	1.231	0.200	0.055	0.073	0.061

In Table 12, the *L-CV* values of the pooled sites are first "de-urbanized" by applying the urban adjustment for *L-CV* in reverse:

$$L\text{-CV}_{deurb} = \frac{L\text{-}CV_{gauged}}{0.5269^{URBEXT_{2015}}}$$

This provides an estimate of what *L*-CV would be if $URBEXT_{2015}$ were zero. Using

$$L\text{-CV}_{\text{pooled}} = \sum_{i=j}^{M} L\text{-CV}_j w_{j,L\text{-CV}} \qquad \text{where } w_j = \frac{(c_j + b_j)^{-1}}{\sum (c_k + b_k)^{-1}}$$

$$\text{using } c_{j,L\text{-CV}} = \frac{0.05932}{n_j - 1} \qquad c_{j,L\text{-SKEW}} = \frac{0.3844}{n_j - 2}$$

$$b_{j,L\text{-CV}} = 0.020995\left(1 - \exp\left(-\frac{SDM_j}{9.6966}\right)\right) \qquad b_{j,L\text{-SKEW}} = 0.023184\left(1 - \exp\left(-\frac{SDM_j}{0.3772}\right)\right)$$

The pooled *L*-moment ratios are:

$L\text{-CV}_{\text{pooled}} = 0.279$ $\quad$ $L\text{-SKEW}_{\text{pooled}} = 0.129$

Pooling (urban adjustment)

It is always recommended to apply the urban adjustment to pooled *L*-moments to avoid sudden step changes in Q_T estimates if a catchment's urban level increases from below to above the threshold for "essentially rural". The urban adjustments for the FEH statistical 2025 method are

$$L\text{-CV}_{urban} = 0.5269^{URBEXT_{2015}} L\text{-CV}_{pooled}$$

$$L\text{-SKEW}_{urban} = L\text{-SKEW}_{pooled}$$

Inserting the relevant value for $URBEXT_{2015}$ into the above equations gives urban *L*-moment ratios of

$$\boldsymbol{L\text{-CV}_{urban} = 0.269 \qquad L\text{-SKEW}_{urban} = 0.129}$$

Q_T estimation

The FEH statistical 2025 method uses the generalized logistic (GLO) distribution to estimate events with *T*-year return periods in ungauged catchments (including gauged catchments whose AMAX are unsuitable for statistical analysis). The GLO distribution function is

$$\frac{Q_T}{QMED} = 1 + \frac{\beta}{\kappa}\left(1 - (T - 1)^{-\kappa}\right)$$

where

$$\kappa = -L\text{-SKEW}_{urban}$$

$$\beta = \frac{\kappa L\text{-CV}_{urban}\sin(\pi\kappa)}{\pi\kappa(L\text{-CV}_{urban} + \kappa) - L\text{-CV}_{urban}\sin(\pi\kappa)}$$

Hence, β, κ and the GLO distribution for Dover Beck at Lowdham are:

$$\boldsymbol{\kappa = -0.129}$$

$$\boldsymbol{\beta = 0.278}$$

$$Q_T = 4.713\left(1 - 2.152\left(1 - (T - 1)^{0.129}\right)\right)$$

and Q_{30}, Q_{100}, Q_{200} and Q_{1000} are

$$\boldsymbol{Q_{30} = 10.237\ m^3/s \qquad Q_{100} = 12.928\ m^3/s}$$

$$\boldsymbol{Q_{200} = 14.660\ m^3/s \qquad Q_{1000} = 19.312\ m^3/s}$$

These are equal to the 1-in-30, 1-in-100, 1-in-200 and 1-in-1000 AEP events respectively (3.33%, 1%, 0.5% and 0.1% AEP).

13.2 Catchment suitable for *QMED* estimation

The catchment used in this example is Cynon at Abercynon (NRFA number 57004). Its AMAX record is suitable for *QMED* estimation but not pooling.

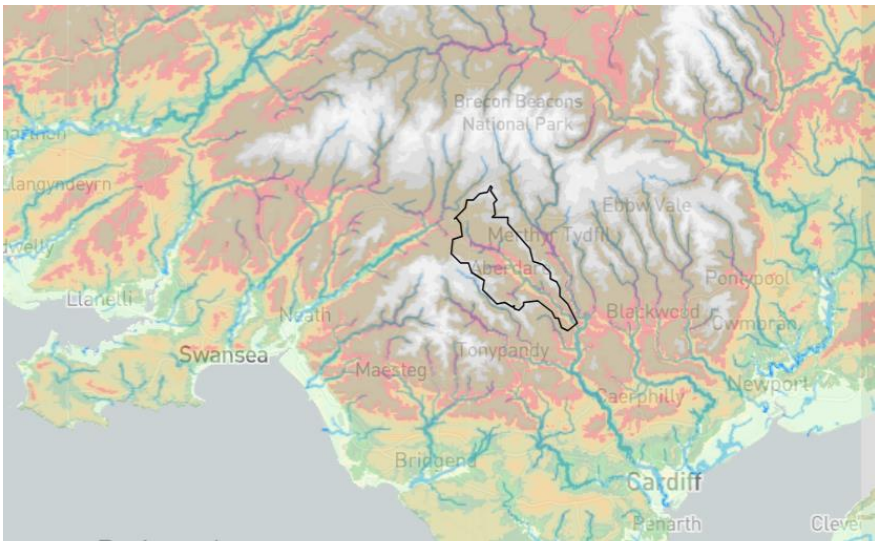

Figure 45 Map of catchment and surrounding area for Cynon at Abercynon. Colour indicates elevation[2]

QMED estimation

This catchment is flagged by the NRFA as "suitable for *QMED*", so *QMED* can be estimated as the median value of the gauged ln(AMAX) record, ignoring excluded years. At this station, there is one excluded year (1960-61) and 61 valid years. Hence, the median is the 31st-largest value.

$$\boldsymbol{QMED = 76.130\ \mathrm{m^3/s}}$$

The estimated QMED value of 76.130 m^3/s should not have donor transfer, urban adjustments, or any other modifications applied.

Uncertainty in a gauged QMED estimate depends on the length and *L*-moment ratios of the AMAX series from which it was derived. Assuming that the AMAX series follows a GLO distribution, the standard error of a gauged *QMED* estimate is

$$se = \frac{2\beta}{\sqrt{n}}$$

where β is the GLO distribution parameter and *n* is the AMAX record length. For Cynon at Abercynon:

$$se = 0.0544$$

[2] Taken from nrfa.ceh.ac.uk © UKCEH 2025. For Great Britain: Contains Ordnance Survey data © Crown copyright and database right 2025.

and the 95% confidence interval around QMED ranges from

$$\frac{QMED}{e^{se^{1.96}}} < QMED < QMED \times e^{se^{1.96}}$$

giving a 95% confidence interval of

$$\mathbf{68.433\ m^3/s} < \boldsymbol{QMED} < \mathbf{84.693\ m^3/s}$$

Pooling, urban adjustment to *L*-moment ratios and *QT* estimation

These stages should be approached as if the site were ungauged, as the at-site AMAX record is unsuitable for pooling, and so the methods should be applied as in Section 13.1.

13.3 Catchment suitable for pooling

The catchment used in this example is Avon Water at Fairholm (NRFA number 84014). Its AMAX record is suitable for *QMED* estimation and pooling.

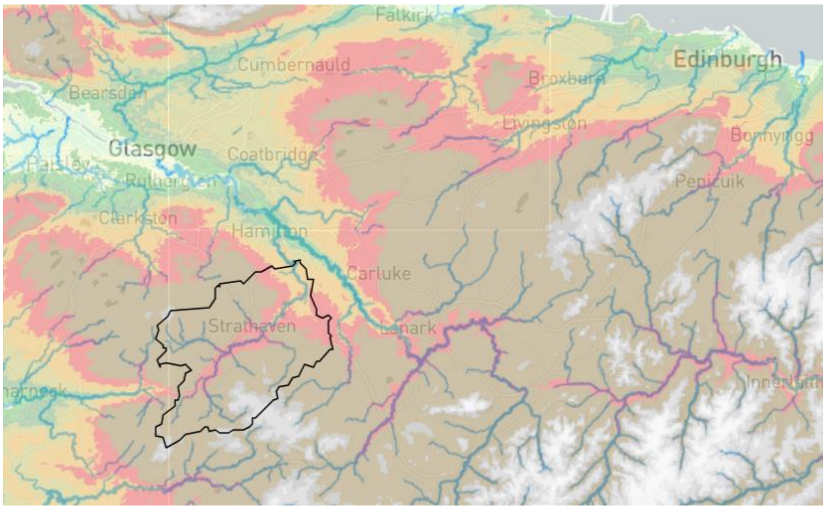

Figure 46 Map of catchment and surrounding area for Avon Water at Fairholm. Colour indicates elevation[3]

QMED estimation

If a catchment is suitable for pooling, *QMED* should be estimated as the log-median of all valid AMAX, as in the "Catchment suitable for *QMED* estimation example". In this case, there is one excluded AMAX and 58 valid AMAX. *QMED* is calculated as

[3] Taken from nrfa.ceh.ac.uk © UKCEH 2025. For Great Britain: Contains Ordnance Survey data © Crown copyright and database right 2025.

the exponentiated mid-point between the ln-transformed 29th and 30th-largest AMAX (the middle of the sorted AMAX series).

AMAX29 = 162.681 m^3/s AMAX30 = 162.547 m^3/s

ln(AMAX29) = 5.09178 ln(AMAX30) = 5.09098

ln(*QMED*) = 5.09138 ***QMED* = 162.614 m^3/s**

In this case, the ln-transformed median and the untransformed median of AMAX values are equal to three decimal places.

The 95% confidence interval is

$$148.770 \text{ m}^3/\text{s} < QMED < 177.934 \text{ m}^3/\text{s}$$

Pooling

As the at-site AMAX series is suitable for pooling, it is included in the pooling-group with an SDM of zero. The formulation of the structural error variance in the FEH statistical 2025 method means that weights for the at-site *L*-moment ratios are calculated as for any other site (there is no “enhancement”). Nevertheless, the SDM_5 of zero ensures that the at-site data are weighted highly relative to pooled data, even without the enhancement.

Table 13: Key data on pooling group stations for Avon Water.

Catchment	№ AMAX	SDM_5	*L*-CV	*L*-CV weight	*L*-SKEW	*L*-SKEW weight
84014	58	0	0.176	0.125	0.136	0.187
8013	30	0.345	0.162	0.047	0.290	0.047
76021	22	0.386	0.193	0.036	0.172	0.038
202001	47	0.445	0.087	0.058	−0.045	0.052
21012	59	0.471	0.144	0.064	0.136	0.055
12008	37	0.494	0.228	0.048	0.168	0.046
76002	25	0.499	0.195	0.037	0.199	0.038

Catchment	№ AMAX	*SDM*5	*L*-CV	*L*-CV weight	*L*-SKEW	*L*-SKEW weight
54038	49	0.508	0.160	0.056	0.188	0.051
78005	43	0.530	0.083	0.051	−0.002	0.048
27035	53	0.535	0.187	0.057	0.409	0.051
236005	40	0.562	0.091	0.048	0.163	0.046
7001	62	0.569	0.159	0.060	0.240	0.053
12007	32	0.577	0.150	0.042	0.237	0.042
8009	70	0.581	0.130	0.063	0.001	0.054
76008	55	0.605	0.141	0.055	0.095	0.050
201005	50	0.640	0.170	0.051	0.395	0.048
72005	53	0.644	0.194	0.052	0.122	0.049
66006	48	0.688	0.231	0.048	0.286	0.046

All stages of this process, including de-urbanizing the pooling-group *L*-CV values, re-urbanizing the pooled *L*-CV values, and fitting the GLO distribution, are equivalent to the “ungauged” and “suitable for *QMED*” cases as described in Section 13.1.

Following those steps we calculate Q_{30}, Q_{100}, Q_{200} and Q_{1000} to be

$$Q_{30} = 279.195 \text{ m}^3/\text{s}$$

$$Q_{100} = 342.456 \text{ m}^3/\text{s}$$

$$Q_{200} = 385.115 \text{ m}^3/\text{s}$$

$$Q_{1000} = 506.618 \text{ m}^3/\text{s}$$

Contact

enquiries@ceh.ac.uk
@UK_CEH
ceh.ac.uk

Bangor
UK Centre for Ecology & Hydrology
Environment Centre Wales
Deiniol Road
Bangor
Gwynedd
LL57 2UW
+44 (0)1248 374500

Edinburgh
UK Centre for Ecology & Hydrology
Bush Estate
Penicuik
Midlothian
EH26 0QB
+44 (0)131 4454343

Lancaster
UK Centre for Ecology & Hydrology
Lancaster Environment Centre
Library Avenue
Bailrigg
Lancaster
LA1 4AP
+44 (0)1524 595800

Wallingford (Headquarters)
UK Centre for Ecology & Hydrology
Maclean Building
Benson Lane
Crowmarsh Gifford
Wallingford
Oxfordshire
OX10 8BB
+44 (0)1491 838800

UK Centre for
Ecology & Hydrology