UK Centre for Ecology & Hydrology

Uncertainty in FEH methods

A short guide for FEH users

Adam Griffin, Lisa Stewart, Gianni Vesuviano (UKCEH)
Tracey Haxton (Wallingford HydroSolutions)

Issue Number 1.1
Date 12/08/2021

Title Uncertainty in FEH methods

UKCEH reference 06813

UKCEH contact details FEH Support Team

UK Centre for Ecology & Hydrology, Maclean Building,
Crowmarsh Gifford, Wallingford, Oxfordshire OX10 8BB

t: 01491 838800
e: fehweb@ceh.ac.uk

Authors Adam Griffin, Lisa Stewart, Gianni Vesuviano (UKCEH)
Tracey Haxton (Wallingford HydroSolutions Ltd)

Approved by James Miller

Signed

Date 23/07/2021

Contents

1 What is uncertainty?

In FEH methods, the standard methods of flood frequency estimation give, for each return period *T*, a single value of flow *Q* corresponding to the *T*-year event, Q_T. This is obviously just an estimate; the true value may differ from it. The level of *uncertainty* is how accurate and/or precise we believe the estimate to be.

2 How do we measure uncertainty?

Within the FEH methods, we focus on two key approaches to measuring uncertainty in the median annual maximum flood (*QMED*) and the growth curve X_T ($Q_T/QMED$): factorial standard error and confidence intervals.

2.1 Factorial standard error

Factorial standard error (fse) is used to describe how much measured values *X* differ from estimated values $\hat{X}$. It is defined as the exponential of the standard error (se)

$$fse = e^{se} = e^{\frac{\sigma}{\sqrt{N}}} \qquad (1)$$

where σ is the sample standard deviation of $\hat{X}$. In estimating *QMED*, we measure the sample standard deviation of the error $\log(QMED) - \log(\widehat{QMED})$.

Factorial standard error is used because the error of Q_T estimates is assumed to increase exponentially as flow gets bigger since, for example, we have less information on the rarest *T*-year events i.e. there is greater uncertainty associated with a 100-year event than a 2-year event. Also, the *QMED* catchment descriptor equation was developed assuming that log(*QMED*) has normally-distributed error, so *QMED* is assumed to have increased error as *QMED* gets bigger. Typically, the true standard deviation is not known. Instead, the sample variance, s^2, is often used, or the standard error is estimated directly via other means.

2.2 Sample Variance

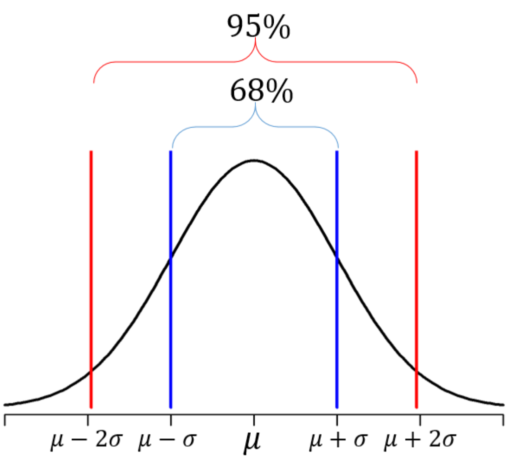
The sample variance, s^2, is a measure of the variability of a time series. For a time series Z_i (such as the AMAX series) the sample variance is given by

$$s^2 = \frac{1}{N-1}\sum_{i=1}^{N}(Z_i - \mu)^2 \qquad (2)$$

where *N* is the number of values and μ is the mean of the values. The sample standard deviation *s* is the positive square root of the sample variance.

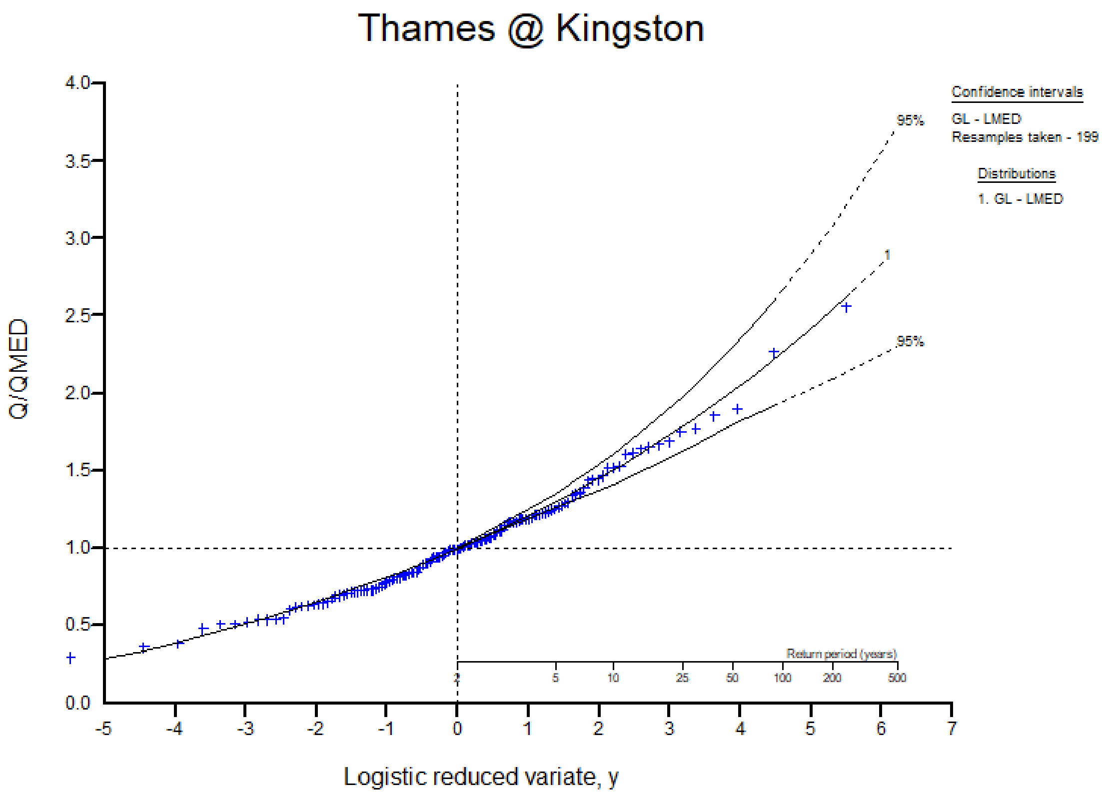
2.3 68-95 rule

In *QMED* uncertainty estimation, we assume that the error of log(*QMED*) is normally distributed. If a value *X* is normally distributed, then if μ is the mean, and σ is the standard deviation, then 68% of the samples of *X* lie in the interval $(\mu - \sigma, \mu + \sigma)$ and 95% of the samples will lie in the interval $(\mu - 2\sigma, \mu + 2\sigma)$.



In practice, if a sample has mean, m, and standard deviation, s, then 68% of samples will lie in the interval $(m - s, m + s)$ and 95% of the samples will lie in the interval $(m - 2s, m + 2s)$.

2.4 Confidence Intervals



Confidence intervals (such as those in the figure above) are used for Q_T to describe how likely it is that the estimate is close to the true value. Typically, we use the 95% confidence interval. This is the interval that we are 95% sure contains the true value of Q_T. The narrower the interval, the more certain we are of the estimate.

It is difficult to know an exact value for these intervals, so there are various ways to approximate them. If we know the fse, then we estimate an approximate 95% confidence interval for *QMED* by

$$\left(\frac{QMED_{EST}}{fse^2}, QMED_{EST} \times fse^2\right) \quad (3)$$

This can alternatively be described in terms of the standard error:

$$(\log(QMED_{EST}) - 2se, \log(QMED_{EST}) + 2se) \quad (4)$$

Alternatively, we can use bootstrapping (Efron & Tibshirani, 1985) or Monte Carlo methods (e.g. Metropolis & Ulam, 1949) to estimate the confidence interval.

2.4.1 Bootstrapping

Bootstrapping is a way of using the observed data to quantify the uncertainty of the growth curve. It is performed by taking a large number of copies of the time series, concatenating and shuffling the combined list, and splitting it up into the same number of "possible" time series. We compute the growth curve for each of the possible time series, and for each return period we compute the 95% confidence interval as being bounded by the 2.5% and 97.5% quantiles.

Alternatively, the standard error can be based on the sample standard deviation of the bootstrapped samples, using Equation 4 for the 95% confidence intervals. Note that this symmetric method can lead to unexpected results where the lower confidence interval becomes flat/decreasing as return period increases. An advantage of bootstrapping is that it requires no choice of distribution.

2.4.2 Monte Carlo uncertainty estimation

An alternative to bootstrapping is to use Monte Carlo methods to estimate the standard error for both single-site and pooled flood frequency estimates. In this method, a Generalised Logistic (GLO) distribution is fitted using the estimated parameters (probably L-moment or pooled L-moment estimates). Many time series of the same length as that of the target catchment are sampled from this GLO distribution, and are used to either compute the standard error (using the sample standard deviation for Q_T) or to compute confidence intervals using the top and bottom 2.5 percentiles.

Note that unlike other methods, this can lead to different values of standard error for different return periods. It also requires a choice of extreme value distribution (e.g. GLO, GEV, GPa).

2.4.3 Delta method

The delta method is an algebraic method of estimating the standard deviation of Q_T (or *QMED* or the growth curve). This method is typically used in the theoretical development and justification of new methods of estimating *QMED* and Q_T e.g. with the inclusion of historical data. For the GLO distribution, recall that $Q_T = \xi + \frac{\alpha}{\kappa}(1 - (T-1)^{\kappa})$.

We can estimate the standard deviation by computing

$$s^2 \approx \nabla(\mathrm{Q_T})^T \boldsymbol{V} \, \nabla(Q_T) \quad (5)$$

where ***V*** is the covariance matrix of $(\hat{\xi}, \hat{\alpha}, \hat{\kappa})$ and $\nabla(Q_T)^T$ is the vector of derivatives of Q_T

$$\nabla(Q_T) = \left[\frac{\partial Q_T}{\partial \xi}, \frac{\partial Q_T}{\partial \alpha}, \frac{\partial Q_T}{\partial \kappa}\right] \quad (6)$$

which is normally computed using numerical solvers.

3 Uncertainty within the FEH Statistical Methodology

3.1 Uncertainty for *QMED*

3.1.1 GLO-fitted QMED (median)

For a gauged catchment the factorial standard error of *QMED*, based on an observed AMAX series fitted using the GLO distribution, is given by

$$fse = e^{\frac{2\alpha}{\sqrt{N}}} \tag{7}$$

where α is the GLO scale parameter, and *N* is the number of recorded AMAX values. Here it can be seen that as record length increases, fse decreases.

For stations with short records, climatic adjustment should be used (Institute of Hydrology, 1999: Chapter 20).

3.1.2 Catchment descriptor equation

The *QMED* catchment descriptor equation was fitted using 602 stations (Environment Agency, 2008). *QMED* based on the catchment descriptor equation has an fse of 1.431, comparing against the “observed” *QMED* at the stations. This is a fixed value that describes the model as a whole, not uncertainty at any particular station.

3.1.3 Channel dimensions model

The channel dimensions model uses values of channel width and flow to estimate *QMED*. This model has an fse of 1.60 as documented in FEH Local (Environment Agency, 2017).

3.1.4 Flow variability model

The WINFAP package presents a ‘flow variability’ model which uses *Q*5 and *Q*10 (gauged daily mean flows exceeded 5% and 10% of the time) to update the estimate of *QMED* (WHS, 2016a). For gauging stations where the uncertainties in estimating high flows are large, this equation allows users to estimate *QMED* using observed data for in-bank, non-flood flows. This model, as documented in the WINFAP 4 *QMED* linking equation document (WHS, 2016b) has an fse of 1.31.

3.1.5 Donor method (one donor)

$$fse = e^{\left(\sqrt{s^2(1-\alpha_d^2)}\right)} \tag{8}$$

Where $\alpha_d = 0.4598e^{-0.02\times d} + (1-0.4598)e^{-0.4785\times d}$ and *d* is the distance between the target catchment and the donor in km, *s* is the standard error of the *QMED* catchment descriptor equation. In this donor method (and for multiple donors), we assume there to be no uncertainty in the measurement of gauged *QMED* at the donor site. Note that this value of α is linked to the *QMED* donor adjustment formula

$$QMED = QMED_{CD}\left(\frac{QMED_{donor,obs}}{QMED_{donor,CD}}\right)^{\alpha_d} \tag{9}$$

Based on this approach to donor adjustment, the fse gets smaller the closer the donor is to the target catchment.

3.1.6 Donor method (multiple donors)

When multiple donors are used to improve the estimate of *QMED* in a similar fashion to above, the product of several adjustments is used giving:

$$QMED = QMED_{CD} \prod_{j=1}^{D} \left(\frac{QMED_{donor\ j,obs}}{QMED_{donor\ j,CD}}\right)^{\alpha_{d,j}} \quad (9b)$$

The description of fse is more complex, but works in a similar way to the single donor case:

$$fse = e^{\left(\sqrt{s^2 - \boldsymbol{b}^T \boldsymbol{\Omega}^{-1} \boldsymbol{b}}\right)} \quad (10)$$

where ***b*** is the subject-donor covariance vector, and ***Ω*** is the between-donor covariance matrix (Kjeldsen *et al.*, 2014). This becomes the same as Equation 8 if only one donor is used.

3.1.7 Combined ungauged and short-gauged estimate

If only a short gauged record is available to estimate *QMED*, then a lower uncertainty can be achieved by taking a weighted average of two *QMED* estimates: the median of the gauged AMAX series and a statistical estimate – ideally the multiple donor method (Kjeldsen, 2015).

$$\ln QMED = (1 - \omega) \ln QMED_g + \omega \ln QMED_s \quad (11)$$

$$\omega = \frac{s_g^2}{s_g^2 + s_s^2} \quad (12)$$

Where *s* refers to standard error, and subscripts *g* and *s* refer to gauged and statistical respectively. The factorial standard error of this estimate is

$$fse = e^{(s_g s_s)/\sqrt{s_g^2 + s_s^2}} \quad (13)$$

3.2 Uncertainty for Growth Curve

3.2.1 Basic Single-site Analysis

If there is enough data, i.e. more than 14 years of AMAX, we can use direct computations of standard error, bootstrapping or Monte Carlo simulation to determine uncertainty for Q_T estimates directly, or for *QMED* and the growth curve separately.

3.2.2 Donor method (one and multiple donors)

The fse is calculated based on the covariance between the target site and the donor(s), and between the donors if there is more than one. It combines the error of the pooled approach with the donor method for *QMED* estimation, and so produces a generalised estimate of uncertainty of the flood frequency curve. It can be calculated based on bootstrapping, but typically uncertainty increases as the average distance to the donors increases, but also as the variability between the donors increases.

There has been work into trying to describe fse across England and Wales under the donor method. Kjeldsen (2015) estimated fse at fixed return periods with and without the use of one donor (replicated in Table 1).

Table 1: List of fse for different return periods (from Kjeldsen, 2015).

Return Period	fse (0 donor)	fse (1 donor)
2	1.47	1.42
5	1.48	1.43
30	1.52	1.47
100	1.54	1.50

An average fse model incorporating return period was published in Environment Agency (2017) for ungauged pooled analysis using 0, 1, 2 and 6 donors, and is suitable for return periods from 2 to 2000 years. The 6-donor equation is replicated here as Equation 14.

$$fse \approx 1.406 + 0.0011y + 0.0040y^2 \quad (14)$$

where $y = -\log\left(-log\left(1 - \frac{1}{T}\right)\right)$. We use this approximation because calculating an exact fse for 6 donors is a complex calculation involving inverting a 6x6 matrix. Slight differences in fse between Table 1 and Equation 13 are due to the exact selection of catchments used in each case, as well as the equation being fitted to minimise error across many return periods.

3.2.3 Pooled analysis

As above, the uncertainty associated to a pooling-group estimate of growth curve can be obtained either via bootstrapping or Monte Carlo simulation.

For pooling-group methods, one may come across a “Pooled Uncertainty Measure” (PUM) which was used as a metric to determine the performance of the pooling-group approach. Dependent on return period T, it is given by

$$PUM_T = \sqrt{\frac{\sum_{i=1}^{N} w_i\left(\log x_{T,i} - \log x_{T,i}^{(p)}\right)^2}{\sum_{i=1}^{N} w_i}} \quad (15)$$

where $x_{T,i}$ is the at-site growth curve, $x_{T,i}^{(p)}$ is the pooled growth curve, and w_I is a series of per-station weights based on record length. PUM is one value for the pooling-group method, not a measure of uncertainty for a given pooling-group. It is also not a standard measure of uncertainty outside of this application.

3.3 Combined Uncertainty

If *QMED* and the growth curve X_T are computed separately, then the uncertainty of Q_T can be related to the uncertainty of *QMED* and X_T, but it is not simply the sum or product of the two uncertainties:

$$Var\left(\widehat{Q_T}\right) = Q_T^2 Var\left(\widehat{QMED}\right) + QMED^2 Var\left(\widehat{X_T}\right) + 2\, QMED\, X_T Cov\left(\widehat{QMED}, \widehat{X_T}\right) \quad (16)$$

Where $\hat{X}$ is the estimate of the true value *X*. *Var*(*QMED*) can be computed using fse, but the covariance term is highly complex to compute, involving joint probabilities of both *QMED* and X_T. In the theoretical case where *QMED* is completely independent of the growth curve, the “Cov” term is zero.

4 How is uncertainty implemented in WINFAP?

4.1 Single-site Analysis

The growth curve 95% confidence intervals are based on the standard error computed using bootstrapped samples using Equation 4 to give the curves. Currently, sampling error, which could be included in estimates of at-site uncertainty, is not included in WINFAP due to insufficient information of measurement precision and accuracy at the gauging station.

4.2 Enhanced Single-site Analysis

Enhanced Single-site analysis (ESS) uses a combination of gauged flow measurements along with pooling-group estimates. The associated uncertainty is complex, as it has to combine the measurement uncertainty of the gauged records with the uncertainty of the modelled part. This is quite complex (see Combined Uncertainty), but in principle could be estimated using Monte Carlo methods. In general, we can assume that pooled uncertainty is higher than single-site uncertainty, and both are expected to be higher than enhanced single-site uncertainty.

4.3 Pooled analysis

Uncertainty is not currently shown for pooled analysis in WINFAP 4.

5 How is uncertainty implemented in ReFH2?

The ReFH2 rural design event model was assessed relative to the enhanced single-site statistical model for return periods from 2 to 1000 years in 285-420 catchments, depending on return period (Wallingford HydroSolutions, 2019). The fse of the ReFH2 model using FEH13 rainfall inputs was similar to but slightly higher than that of the pooled statistical method, also assessed relative to the enhanced single-site estimates. However, the direct comparison of uncertainty between ReFH2 and the pooled statistical method is not completely valid, because the baseline enhanced single-site method is not fully independent of the pooled statistical method. Further, ReFH2 estimates are less biased than pooled statistical estimates, relative to the baseline on average. Uncertainty in a peak flow estimate is not shown within the software.

6 Uncertainty in FEH rainfall models

6.1 Kriging variance

In the FEH99 and FEH13 DDF models, *RMED* (the 2-year return period rainfall event) is used as an index rainfall by which other events at a rain gauge are standardised. *RMED* is estimated as the median of gauged annual maxima at each rain gauge station and then extended to the rest of the UK via kriging (georegression), which smoothly “fills in the gaps”. This introduces some uncertainty at ungauged sites between gauging stations. In FEH Vol. 2, the standard deviation of *RMED* (2-year return period rainfall event) is approximated by the square root of the kriging variance.

6.2 Growth curves

When considering a gauged series of annual maximum rainfall, confidence intervals for rainfall growth curves (for a fixed duration and location) could be computed via bootstrapping, as described in Sections 2.4.1 and 4.1. When considering estimates from a DDF model, there is also modelling uncertainty from the DDF model to be taken into account. The current implementation of the FEH13 rainfall DDF model iin the FEH Web Service does not include uncertainty estimation.

7 Notes on sources of error

7.1 Sampling Error

In addition to having uncertainty about how well the model fits the observed values, we can also consider the uncertainty in how well the model fits the true flow. This comes from the fact that we don’t have an infinite number of observations to perfectly describe the true flow. This interacts in a complicated way with the total uncertainty of *QMED* and Q_T.

7.2 Measurement Error

In addition to sampling error, there is measurement error, which also is a part of the possible uncertainty relating to flood frequency estimation. This includes shortcomings in precision and accuracy of flow measurement, which can also lead to bias.

References

Efron B. and Tibshirani R., 1985. The bootstrap method for assessing statistical accuracy. *Behaviormetrika*, 12, 1-35.

Environment Agency, 2008. *Improving the FEH statistical procedures for flood frequency estimation*. Science Report SC050050. Bristol: Environment Agency.

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

Institute of Hydrology, 1999. *Flood Estimation Handbook* (5 volumes) Wallingford: Institute of Hydrology.

Kjeldsen, T. R., Jones, D. A., and Morris, D. G., 2014. Using multiple donor sites for enhanced flood estimation in ungauged catchments, *Water Resources Research*, 50, 6646– 6657.

Kjeldsen T.R., 2015. How reliable are design flood estimates in the UK? *Journal of Flood Risk Management*, 8(3), 237-246.

Metropolis N. and Ulam, S., 1949. The Monte Carlo method. *Journal of the American Statistical Association*, 44(247), 335-341.

Wallingford HydroSolutions, 2016a. *WINFAP 4*. Wallingford HydroSolutions Ltd. (https://www.hydrosolutions.co.uk/software/winfap-4/)

Wallingford HydroSolutions, 2016b. *WINFAP 4 QMED linking equation*. Wallingford HydroSolutions Ltd. (https://software.hydrosolutions.co.uk/winfap4/QMED-Linking-Equation-Technical-Note.pdf)

Wallingford HydroSolutions, 2019. *ReFH2 Science Report: Evaluation of the Rural Design Event Model*. Wallingford HydroSolutions Ltd. https://www.hydrosolutions.co.uk/app/uploads/2019/10/ReFH2-Science-Report-Evaluation-of-the-Rural-Design-Event-Model.pdf

BANGOR
UK Centre for Ecology & Hydrology
Environment Centre Wales
Deiniol Road
Bangor
Gwynedd
LL57 2UW
United Kingdom
T: +44 (0)1248 374500
F: +44 (0)1248 362133

EDINBURGH
UK Centre for Ecology & Hydrology
Bush Estate
Penicuik
Midlothian
EH26 0QB
United Kingdom
T: +44 (0)131 4454343
F: +44 (0)131 4453943

LANCASTER
UK Centre for Ecology & Hydrology
Lancaster Environment Centre
Library Avenue
Bailrigg
Lancaster
LA1 4AP
United Kingdom
T: +44 (0)1524 595800
F: +44 (0)1524 61536

WALLINGFORD (Headquarters)
UK Centre for Ecology & Hydrology
Maclean Building
Benson Lane
Crowmarsh Gifford
Wallingford
Oxfordshire
OX10 8BB
United Kingdom
T: +44 (0)1491 838800
F: +44 (0)1491 692424

enquiries@ceh.ac.uk

www.ceh.ac.uk