Department for Environment Food & Rural Affairs
Llywodraeth Cymru Welsh Government
Cyfoeth Naturiol Cymru Natural Resources Wales
Environment Agency

delivering benefits through evidence

Development of interim national guidance on non-stationary fluvial flood frequency estimation – practitioner guidance

FRS18087/IG/R2

Flood and Coastal Erosion Risk Management Research and Development Programme

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

Acting to reduce the impacts of a changing climate on people and wildlife is at the heart of everything we do.

We reduce the risks to people, properties and businesses from flooding and coastal erosion.

We protect and improve the quality of water, making sure there is enough for people, businesses, agriculture and the environment. Our work helps to ensure people can enjoy the water environment through angling and navigation.

We look after land quality, promote sustainable land management and help protect and enhance wildlife habitats. And we work closely with businesses to help them comply with environmental regulations.

We can't do this alone. We work with government, local councils, businesses, civil society groups and communities to make our environment a better place for people and wildlife.

Published by:
Environment Agency, Horizon House, Deanery Road, Bristol, BS1 5AH

http://www.gov.uk/government/organisations/environment-agency

ISBN: 978-1-84911-468-4

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

Further copies of this report are available from our publications catalogue:
http://www.gov.uk/government/publications

or our National Customer Contact Centre:
T: 03708 506506

Email: enquiries@environment-agency.gov.uk

Author(s):
Duncan Faulkner and Sarah Warren

Dissemination status:
Publicly available

Keywords:
Stationarity, non-stationarity, flood frequency estimation, climate change, trend

Research contractor:
JBA Consulting
1 Broughton Park, Old Lane North, Broughton, Skipton, North Yorkshire, BD23 3FD
www.jbaconsulting.com

Environment Agency's Project Manager:
Dr Sean Longfield

Theme Manager:
Dr Sue Manson, Dr Sue Manson, Incident Management and Modelling Theme Manager

Project number:
FRS18087

Evidence at the Environment Agency

Scientific research and analysis underpins everything the Environment Agency does. It helps us to understand and manage the environment effectively. Our own experts work with leading scientific organisations, universities and other parts of the Defra group to bring the best knowledge to bear on the environmental problems that we face now and in the future. Our scientific work is published as summaries and reports, freely available to all.

This report is the result of research commissioned and funded by the Joint Flood and Coastal Erosion Risk Management Research and Development Programme. The Joint Programme is jointly overseen by Defra, the Environment Agency, Natural Resources Wales and the Welsh Government on behalf of all Risk Management Authorities in England and Wales:
http://evidence.environment-agency.gov.uk/FCERM/en/Default/FCRM.aspx.

You can find out more about our current science programmes at:
https://www.gov.uk/government/organisations/environment-agency/about/research.

If you have any comments or questions about this report or the Environment Agency's other scientific work, please contact research@environment-agency.gov.uk.

Professor Doug Wilson
Director, Research, Analysis and Evaluation

Executive summary

This interim guidance aims to introduce practitioners to the concepts of non-stationary flood frequency analysis and guide them as to when and how they should do it, and how to interpret the results.

The guidance has been written for analysing river flows as part of the appraisal of flood risk management schemes and is intended for use by experienced hydrologists familiar with the Flood Estimation Handbook methods. It could also be used by project managers involved in appraisal of flood alleviation schemes with some knowledge of flood hydrology, who need to learn more about issues of trend and non-stationarity.

The guidance outlines important concepts including, stationarity, non-stationarity, trend testing, hypothesis testing, statistical distributions used to model UK flood frequency, and parameter estimation. It goes on to explain the use of covariates (an observed variable used in a statistical model to help predict the main variable of interest) in non-stationary flood frequency analysis, and how they can be applied in the England and Wales.

The main focus of this report is to present practitioners with a step-by-step guide on how to carry out non-stationary flood frequency analysis, illustrated with examples in 3 case studies.

Acknowledgements

Particular thanks are extended to the Environment Agency's project team: Richard Macilwaine, Peter Spencer, Sean Longfield and Parveen Mann.

Contents

List of figures

List of tables

1 Introduction

1.1 What's this document about?

This interim guidance aims to introduce practitioners to the concepts of non-stationary flood frequency analysis and guide them as to when and how they should do it, and how to interpret the results. It illustrates the concepts with 3 case studies.
The guidance has been written for analysing river flows as part of the appraisal of flood risk management schemes. It will be updated following feedback on its use and further developments in the underlying science.

1.2 Who does this apply to?

Experienced hydrologists familiar with the Flood Estimation Handbook methods, needing to find out more about trend and non-stationarity and how to allow for them in flood frequency analysis.

Project managers involved in appraisal of flood alleviation schemes with some knowledge of flood hydrology, needing to learn more about issues of trend and non-stationarity.

You should apply this guidance:

- in the appraisal of all projects submitting a short form business case or outline business case to the Environment Agency for assurance and approval after 1st July 2021. Projects submitting before this date could also be assessed against this guidance to check that it would not lead to different decisions provided this would not unduly slow completion or add significantly to the cost.

- to your FCERM strategy if you have not already submitted it to the Environment Agency for assurance and approval. For existing approved plans and strategies we would not normally expect this advice to be applied until the next review, unless specific investment projects within them are planned before this. In these cases, new project appraisals should adopt the new guidance.

You should use this guidance to understand how and when non-stationarity should be included in the appraisal of FCERM projects schemes and strategies. This guidance should be used in conjunction with the full appraisal guidance available here.

2 Purpose of this guidance

This interim guidance aims to introduce practitioners to the concepts of non-stationary flood frequency analysis and guide them as to when and how they should do it, and how to interpret the results. It illustrates the concepts with 3 case studies.

The guidance has been written for experienced hydrologists to use in their analysis of river flows as part of the appraisal of flood risk management schemes. It will be updated following feedback on its use and further developments in the underlying science.

2.1 Audience

Read this guidance if you are:

- a hydrologist familiar with the Flood Estimation Handbook methods, interested to find out more about trend and non-stationarity and how to allow for them in flood frequency analysis
- a project manager with some knowledge of flood hydrology, needing to learn more about the issues
- working on appraisal of capital schemes

Flood frequency estimation is unavoidably a topic that involves statistics, because we are interested in the probability that floods will occur. Some of the concepts may be unfamiliar. This guidance keeps jargon to a minimum and the main terms used are referenced in the glossary. The science report explains the statistical concepts in more depth.

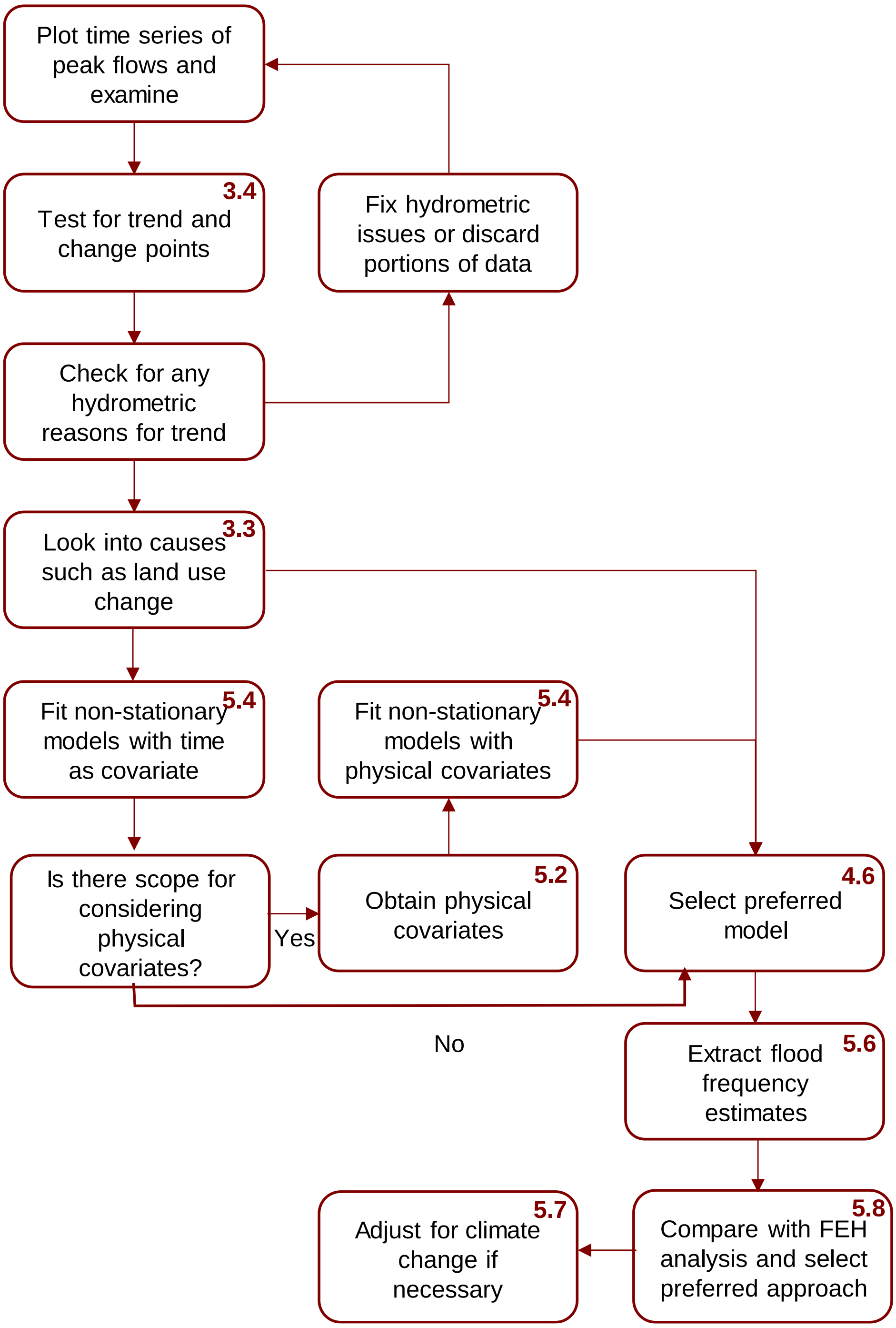
2.2 What to expect

This guidance introduces concepts and techniques that are widespread in academia but until recently have been applied very little in flood risk management. Section 3 introduces some basic concepts, explaining what is meant by non-stationarity, what causes it and when to consider non-stationary analysis. Section 4 goes on to explain more concepts such as distributions, parameters, and covariates. It also contains some important material on how to understand flood probabilities in a non-stationary context and how to select statistical models.

Section 5 provides a step-by step guide to non-stationary analysis. Read it alongside section 4.

Three case studies that illustrate different aspects of non-stationary analysis are given in section 6.

Figure 2-1 Process flow chart giving an overview of the steps to be followed during a flood estimation study where non-stationarity is a potential issue. The numbers in bold refer to sections of this guidance.



2.3 Related documents

There are 2 main documents that support this guidance, listed below. They're the first place that you should look if you have any questions on either the theoretical or the practical aspects of non-stationary flood frequency estimation. This guidance was produced as part of a project (Development of interim national guidance on non-stationary fluvial flood frequency estimation), which also developed methods for non-stationary flood frequency estimation, carried out tests for trend, developed a software package and applied non-stationary methods at a national scale.

2.3.1 Science report

The report, 'Development of interim national guidance on non-stationary fluvial flood frequency estimation, gives the scientific background to the methods that are summarised in this guidance. Refer to the report for more information on where the methods come from, how they were developed, how they relate to wider scientific literature, and for information on options that were considered but are not yet recommended for implementation.

The report also includes an analysis of trend across the gauging station network in England and Wales, and nationwide results from applying non-stationary flood frequency analysis.

2.3.2 User guide for the R package, nonstat

We have developed tools to allow hydrologists to easily carry out trend and change point detection and non-stationary flood frequency analysis using annual maximum flow data.

These tools are implemented using R. R is a freely-available programming language and environment for statistical computing and graphics (see https://www.r-project.org/about.html). R packages are collections of R functions, data, and compiled code in a well-defined format that make data analysis more user-friendly.

The user guide explains how to use the R package, called nonstat, to implement the methods outlined in this guidance.

2.3.3 In summary

If you want to ask 'Why?', refer to the science summary report.

If you want to ask 'How?', refer to the user guide for the nonstat package.

3 Stationarity and its absence

3.1 Definitions

3.1.1 Stationarity

If the processes that produce floods are stationary, we would expect that the probability of a particular flow occurring would not change over time. The relevant time scale is usually a few decades, since this is the typical length of river flow records and the typical design life of flood alleviation schemes. We would expect short-term fluctuations in flood probability, for example due to seasonal variations.

Since we do not have perfect knowledge of flood-generating processes, we adopt a functional definition of stationarity, which describes a data set rather than the underlying physical processes. A time series of flood flows is stationary if its statistical properties do not change over the relevant timescale.

3.1.2 Non-stationarity

If the statistical properties of a data set do change over the relevant timescale, you can regard the data set as non-stationary.

For example, the mean, the median or the variance of annual maximum flows might change during the period of record.

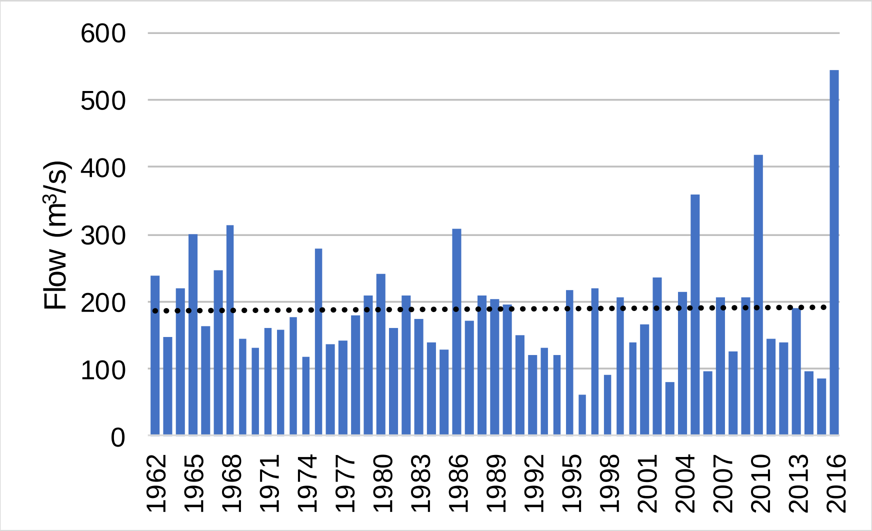
It is normally necessary to use a statistical significance test, or a range of tests, to diagnose non-stationarity.

3.1.3 Trend

Trend is closely related to the above functional definition of non-stationarity. Trend is often thought of as a progressive change in the size of a variable ("floods are getting bigger"). This is one example of non-stationarity. Another might be an increase in the variability of floods, without any change in their average ("big floods are getting bigger and small floods are getting smaller").

Is there a trend in Figure 3-1? This (made up) time series of annual maximum flows shows no increase in the mean over the period of record. A linear trend line fitted by regression is horizontal. Yet the variability of floods is increasing. The 3 largest floods all occur towards the end of the record. So there appears to be a trend in the variance.

Figure 3-1 Made up time series of annual maximum flows



3.2 Assumption of stationarity

The Flood Estimation Handbook (FEH) methods assume stationarity in both rainfall and flood flows. This is known as the ‘identically distributed’ assumption: each annual maximum rainfall or flow at a particular site is assumed to come from an identical statistical distribution.

FEH methods assume that observations of past flood events can represent the behaviour of future events. They are normally applied together with an allowance for the potential impacts of climate change, which are assumed to occur in the future.

If this assumption is not correct, it calls into question the results of flood frequency estimation.

3.3 Causes of non-stationarity in flood flows

3.3.1 Climate, land use or other

Non-stationarity in flood time series may occur due to climatic change or climate variability. Changes within a catchment such as urbanisation, deforestation and changes in agricultural practices may also result in non-stationary flood series. Most of this guidance applies equally to non-stationarity from any of these causes. There has been little research on the causes of non-stationarity in UK flood data. This guidance does not cover the important issue of trend attribution.

Other causes of non-stationarity include construction of reservoirs or flood alleviation schemes, which may include providing flood storage. These are likely to lead to sudden changes in flood flows (step changes), usually reductions. The methods in this guidance don’t apply in these cases.

3.3.2 Climate change: past, present and future

Our current limited knowledge of the causes of past trends in flooding limits our ability to estimate future flows. Until recently, allowances for the impact of climate change on floods have treated climate change as if it were purely a future phenomenon. The procedure has been to make an estimate of the present-day design flow (or rainfall), which is assumed to equally apply to past conditions, and then to adjust it for potential future conditions.

Climate change allowances are based on a baseline period of 1961 to 1990. Although this is now receding into history, the way in which allowances are applied assumes that the entire period spanned of hydrometric records, up to the present day, can be treated as the baseline.

If some of the observed trends in flooding are due to changes in climate, then it is no longer valid to treat the entire observed period as a constant baseline. The implication would be that some of the expected climate change impacts on flooding have already occurred. This may well be the case, for example, given that climate change allowances predict an increase in peak flows by the 2020s compared with 1961 to 1990. If this is correct, it may then be valid to apply a reduced allowance for the impacts of climate change in future decades.

The difficulty is that we are not certain what is causing the trends. Even where trends are seen in rainfall as well as flow, it is difficult to know whether they are due to climate change caused by human activity, which is likely to continue, or natural climate variability, which can change direction.

For this reason, this guidance does not offer a definitive way of estimating future flood frequency under non-stationary conditions. Instead, it suggests 2 approaches to consider.

3.4 Trend tests

3.4.1 Data checking

The longer the flow record, the more confident you can be in the results of a trend test. As a general guide, 40 years should be a minimum length to consider. Even with this record length, it is very important to realise that a trend that looks significant and convincing might disappear or even reverse after another decade or so.

Some apparent trends in flood flow data can be misleading. You should check that any trends are not due to hydrometric issues such as changes in flow measurement structures or in channel geometry that have not been accounted for in a rating equation.

You should check whether any step changes can be explained by interventions such as construction of reservoirs or flood alleviation schemes. If so, it is normally best to ignore the pre-change part of the record.

3.4.2 Visual assessment

Start any assessment of trend by plotting a time series of the AMAX flows. Is there an apparent trend? Does it affect only the highest floods or all floods? Is there any obvious hypothesis for what has caused the trend?

3.4.3 Hypothesis testing

In statistical trend testing the null hypothesis, labelled H0, is that there is no trend.

The tests output a p-value, or probability, and if this is less than a chosen significance level then H0 is rejected. The conventional approach is then to (provisionally) accept a single alternative hypothesis H1, that is, that a statistically significant trend exists. It is common to adopt a 0.05 significance level, which means that if $p<0.05$, there is less than a 5% chance of obtaining a trend at least as extreme as the one in your data set, if the null hypothesis is true. This formal but rather abstract concept is sometimes informally interpreted as meaning there is a 95% probability that trend is present, although this interpretation is rather loose.

It is important to be aware that sometimes your conclusion from a significance test can be wrong. You might wrongly:

- reject the null hypothesis, that is, that there is no trend present, but you conclude that there is a trend. This type of error is more likely if you choose a high significance level
- accept the null hypothesis, that is, although there is a trend you conclude that there is not. This type of error is more likely if you choose a low significance level. If you are concerned about missing trends, consider increasing the significance level, for example to 0.10.

3.4.4 Non-parametric test for trend

The Mann-Kendall test is a very widely used method for testing for consistent trend in a data set. It is non-parametric, which means that it makes no assumption about the statistical distribution of the data. The test is not dependent on the magnitude of the data, but is based on the proportion of increases and decreases between pairs of values. A consequence is that it tests the statistical significance of the trend but does not directly measure the strength of the trend.

You can apply the Mann-Kendall test to a station or a group of stations using the **MK.test** function in the R package, nonstat.

The test outputs a score, labelled Z. Positive values of Z indicate increasing trends, while negative ones refer to decreasing trends. Z scores are standardised, so you can compare them between different stations or different periods of record. If the absolute value of Z exceeds 1.96, the trend is statistically significant at a 0.05 significance level. If it exceeds 1.645, the trend is significant at a 0.10 significance level.

The Mann-Kendall test assumes that each data value is independent of the others. This assumption may not be made if there are flood-rich and flood-poor periods, or on some groundwater-dominated catchments where high baseflow persists for more than one year. Refer to the science report for a way of overcoming this limitation.

The science report presents a nationwide analysis of trend, applying the Mann-Kendall test in a multi-temporal fashion. This selects numerous subsets of each annual maximum flow series, selecting every reasonable combination of start and end years.

The results indicate how sensitive trend test results are to the period of record available for analysis.

3.4.5 Test for change point: Pettitt

Pettitt’s test is designed to detect a sudden change in the average of a time series. It outputs the time of the change as well as its statistical significance level, in the form of a p-value. It is another non-parametric test. The null hypothesis H0 is that there is no difference between the means of the earlier and later portions of the AMAX flow series. Refer to the section above on hypothesis testing for guidance on interpreting p values.

The Pettitt test is included in the R package, nonstat: the **Pettitt.test** function.

3.4.6 Test for change points: PELT

One limitation of Pettitt's test is that it can only detect a single change point. It can also classify some gradual trends as sudden changes. An alternative test is PELT: Pruned Exact Linear Time. See Killick and others (2012), listed in ‘Related documents’. PELT tries to find the optimal segmentation in a time series. It can detect one or more change in the mean, the variance or both.

PELT, as implemented in the nonstat package PELT.test function, provides the following results for each change point found:

1. year that the change occurred
2. direction of change in the mean of the time series (positive or negative)
3. the arithmetic difference between the means before and after the change point
4. the percentage difference in the means
5. the same as 2 to 4 above, for the change in the standard deviation

The test does not output significance levels for the change points.

The PELT test requires an assumption that the data follow a known distributional form. In the nonstat package, the AMAX flows are assumed to follow a log-normal distribution. A minimum segment length is required, which prevents false positive changes at short timescales. This is set to 10 years, which means that the **PELT.test** function will not detect any change points within 10 years of the start or end of the AMAX series. This means that the record length must be at least 20 years (twice the minimum segment length).

3.4.7 What to do if you find a step change

See if you can identify a physical cause for the change. Look into the hydrometric history of the gauging station. Seek information on any sudden changes in the catchment. Refer to ‘Data checking’ above.

If you cannot find a cause, investigate whether the change takes place at a time when there was a known shift between a flood-poor and flood-rich period, either locally or nationally. For instance, some AMAX records show step changes in the late 1990s, with the Easter 1998 or autumn 2000 floods marking the end of a long flood-poor period.

3.4.8 Test for differences between different segments of a time series

Another statistical test that can help in testing for non-stationarity is to predetermine a potential split point in the data set, for example by dividing the time series into 2 equal halves. There are then 2 statistical tests that can be applied:

- **Mann-Whitney U test**, for significance of changes between the distributions of AMAX flows in the earlier and later periods. The test determines whether 2 independent samples were selected from populations having the same distribution. The null hypothesis is that the distributions of the 2 populations are identical
- **Brown-Forsythe test** (Brown and Forsythe, 1974), for significance of changes between the variances of AMAX flows in the earlier and later periods. The null hypothesis is that the samples of AMAX flows are drawn from populations with equal variance

Both tests output a statistical significance level, in the form of a p-value. They are not included in the R package but can be implemented in R using other code.

3.4.9 Parametric tests for trend

The tests described above do not give a full picture of the stationarity or non-stationarity characteristics of a data set. For instance, the Mann-Kendall test may not detect a significant trend when the variance of floods is increasing rather than their mean. This is the case for the data set in Figure 3-1, no trend is detected at a 0.05 significance level.

Parametric tests are a useful complement because they include information on the magnitude of floods rather than only their relative rankings. You can view the fitting of a non-stationary flood frequency distribution as a parametric test for trend: if the data are fitted better by a non-stationary distribution than a stationary distribution, this is an indication that there may be some type of trend present.

In addition, the strength of a trend can be quantified using an approach known as Thiel-Sen. This estimates the slope of the trend line. Refer to the science report for results of the Thiel-Sen approach applied to flood peak series throughout England and Wales. This method is not included in the R package.

3.4.10 Increasing or decreasing trend?

Much of the interest in non-stationarity focuses on increasing trends. But the methods described in this guidance equally apply to decreasing trends.

If you detect a significant decreasing trend, it would be wise to try and investigate its possible causes. One hypothesis would be that floods have reduced as a result of constructing reservoirs, flood storage schemes or natural flood management measures. In other cases, the trend may be due to natural climatic cycles.

3.5 When to consider non-stationary analysis

The guidance is currently intended to be used only in the planning and appraisal of flood risk management schemes. Non-stationary methods may in future be approved for other aspects of the Environment Agency's work.

The methods described in this guidance can be applied only at gauged sites. There has been research into applying non-stationary analysis across a pooling group, but this has not yet developed a method recommended to be generally applied at ungauged sites. Use your judgement to consider when it is appropriate to transfer analysis at a gauged location to an ungauged one.

Visual assessment and the Mann-Kendall test are a useful initial screening step. Neither are guaranteed to detect non-stationarity. So, if you have other reasons to suspect non-stationarity, consider applying non-stationary flood frequency analysis.

If you carry out non-stationary analysis, always do so alongside an equivalent stationary analysis. Use the ways of selecting statistical models explained in this guidance to help decide which type of analysis is preferable.

In summary, consider non-stationary analysis:

- for the planning and appraisal of flood risk management schemes
- when there is a flow gauging station not far from your sites of interest
- when you think relevant data sets may be non-stationary

However, do not abandon stationarity automatically.

4 Concepts of non-stationary flood frequency analysis

This chapter introduces the concepts. Refer to chapter 5 for information on how to put them into practice.

4.1 Distributions

4.1.1 Distributions used in UK flood frequency

Flood frequency analysis involves fitting a statistical distribution to flood peak data. Out of the many ‘extreme value’ distributions available, methods used in the UK (the FEH) most commonly fit the generalised logistic (GLO) or generalised extreme value (GEV) distributions, or simpler versions of these.

The preferred distribution is then adopted as the estimate of the flood frequency curve.

The fitted distribution is sometimes referred to as a statistical model.

4.1.2 Single-site or pooled?

The non-stationary methods described in this guide can be applied only at an individual gauging station.

Although pooled application of non-stationary methods has been explored, it is not yet recommended for use by UK practitioners.

4.2 Parameters and how they might vary

4.2.1 Number of parameters

The GLO and GEV distributions both have 3 parameters: the location, scale and shape. The simpler 2-parameter versions of these distributions are known as the logistic and Gumbel distributions.

4.2.2 Non-stationary parameters

In conventional flood frequency analysis, for a particular gauging station, the parameters are thought of as fixed quantities that we are trying to estimate.

In non-stationary analysis, one or more of the parameters is not fixed. It might be changing over time, or changing in response to changes in some variable other than flow.

The nonstat package enables you to fit 4 types of statistical model:

- stationary

- non-stationary, with the location parameter varying (that is, the average flood magnitude is changing)
- non-stationary, with the scale parameter varying (that is, the variability of flood magnitudes is changing)
- non-stationary, with the location and scale parameters varying

In all cases, the shape parameter is assumed to be constant because there is too much error in estimating it to allow a covariate to be included.

4.3 Covariates: temporal and physical

4.3.1 Time as a covariate

If a parameter of the distribution of AMAX flows is varying, it must be varying in accordance with some quantity other than river flow. This quantity is known as a covariate.

Most simply, the covariate is time, usually expressed in terms of the water year. For instance, the location parameter of the distribution might increase as time goes on.

4.3.2 Physical covariates in addition to time

Physical covariates may help remove some of the year-to-year variability in AMAX flows, allowing time-based trends to be better identified and better fit of the distribution. Examples of potentially relevant physical covariates include:

- annual or seasonal rainfall over the catchment
- temperature
- large-scale indices of atmospheric circulation, such as the North Atlantic Oscillation (NAO) or East Atlantic pattern (EA)
- urban extent for the catchment

If you include physical covariates alongside time in fitting a statistical model, it is preferable that the physical covariates themselves have no time trend. If they do have a trend, it needs to be removed first to reduce correlation between the covariates. The nonstat package can do this detrending.

4.3.3 Physical covariates instead of time

Time itself has no physical influence on flooding. As a covariate, time is merely a substitute for some changing physical quantity that is influencing floods.

A model that includes only physical covariates can provide a more physically meaningful description of non-stationarity.

Within this approach, to model a non-stationary flood series, it is preferable to include at least one physical covariate that exhibits some easily modelled trend over time. An example might be the extent of urbanisation in a catchment, which can be typically expected to show a consistent increase over time. Additionally, urbanisation can be reasonably predicted into the future under a range of scenarios.

A risk associated with this approach is confusing correlation with causation. In other words, there is a temptation to think that because floods are correlated with some covariate, there is a physical process that links the covariate with river flow. In principle, it would be possible to include any covariate with a trend, whether or not it had any physical connection with the processes that cause floods. For example, flood magnitudes might be correlated with mobile phone ownership, since both have tended to increase over the last 20 years.

This could lead to a false sense of confidence about the ability to estimate the future evolution of the flood frequency curve. We might end up with a covariate for which we can confidently predict future values, but which is no more useful than the water year as a way of explaining observed trends in flood magnitudes. Therefore, it would be vital to demonstrate a strong causal relationship, not just a statistical association, for physical covariates if you were going to use them for future predictions. In practice, this is difficult to do without an in-depth investigation, for instance using rainfall-run-off modelling.

The case study on the River Eden in section 6.1 provides an example of using physical covariates.

4.4 Understanding probabilities and return periods in a non-stationary context

4.4.1 Return period, AEP and encounter probability

Return period (T) is commonly used by practitioners, despite official recommendations to express flood rarity in different terms.

The reciprocal of the return period (on the annual maximum scale), 1/T, is the annual exceedance probability, AEP.

Both return period and AEP are used in this guidance. Outputs from the R package are expressed in terms of return period.

In practice, it can be more useful to express the exceedance probability associated with a period of time longer than a year. For example, you may be interested in the probability of a flood defence scheme being overtopped during its design life, or the probability of a housing development flooding during a 25-year mortgage term. This quantity is known as the encounter probability. It is the chance of encountering a particular flood flow (or higher) during a particular period.

You can calculate the encounter probability P using:

$$P = 1 - \left(1 - {}^{1}/_{T}\right)^{N} \quad \text{Equation 1}$$

where T is the return period of the flood and N is the length of the period, both expressed in years.

Refer to Sayers (2016) for more information.

4.4.2 Non-stationary return periods and the integrated flow estimate

Return period is an awkward concept under non-stationary conditions, when the average interval between floods of a particular size might be changing over time. When modelling non-stationarity using time as a covariate, you can think of the return period as the reciprocal of the AEP at a particular point in time.

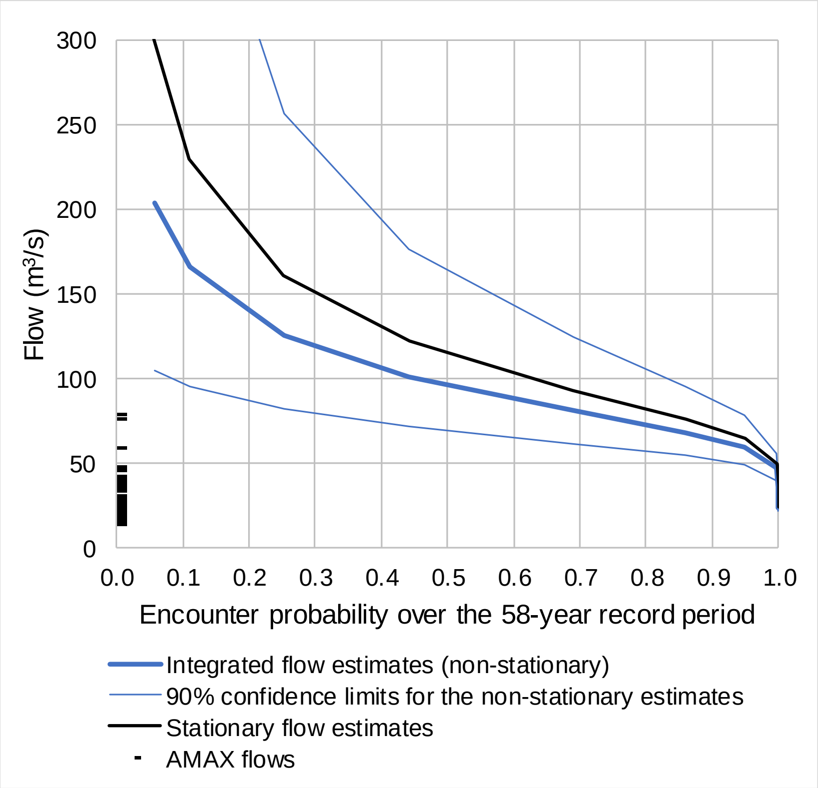
In non-stationary conditions, the encounter probability becomes a more valuable concept. In this guidance, the integrated flow estimate is the flood flow associated with a particular encounter probability (see section 5.6). In non-stationary conditions you need to specify not just the length of the period (N) but also its timing. For instance, you could say "An integrated flow estimate of X m^3/s has an encounter probability P over the 20-year period starting in 1995". See

Figure ***4-1*** for an illustration.

The same terminology can also be used for future periods, although it is more challenging to have confidence in the results.

When including physical covariates, there are extra complications because the AEP of a particular flow varies not only with time but also with the physical covariates. However, the concept of the integrated flow estimate remains valid; refer to section 5.6.

Figure 4-1 Illustrating encounter probability and integrated flow estimates



The example given in

Figure ***4-1*** presents results for a station with a record length of 58 years. Integrated flow estimates have been calculated from a non-stationary model and output by the R package, for a range of return periods. They represent flows expected over the period of record. The return periods are converted to encounter probabilities using Equation 1, applied in Excel.

Figure ***4-1*** shows how encounter probability and flow are related to each other. The plot shows both stationary and non-stationary estimates, along with the confidence limits for the latter (another output from the package). The observed annual maximum flows are plotted on the y axis for reference.

A flow of 100 m^3/s has an encounter probability of 0.44 during the 58-year period, according to the non-stationary results. If you wanted to design a structure that had only a 10% probability of being overtopped during the period, you could look up the flow with an encounter probability of 0.1, which is just under 200 m^3/s according to the non-stationary model.

For this example, the flows from the stationary analysis are higher than from the non-stationary analysis. This will not always be the case.

This type of plot differs from the conventional flood frequency curve which plots flow against return period, or AEP, and typically also shows the annual maximum flows, with their probabilities estimated using a plotting position. There are 2 issues with showing non-stationary estimates on a conventional flood frequency curve:

- The integrated flow estimates are defined relative to a period of time which is typically longer than one year. It is possible to relate them to an instantaneous AEP or return period, as noted above, which enables a more conventional plot, but it is open to misinterpretation. This issue is illustrated in the case studies.
- The annual maximum flows cannot be readily plotted on the same probability scale because the probability of a given flow is typically changing from year to year.

The **res.phys.cov** function in the nonstat package is able to output encounter probability plots similar to that shown above.

4.5 Ways of fitting distributions

4.5.1 L-moments

FEH methods use L-moments to fit extreme value distributions. These are not readily adapted to work in non-stationary conditions.

4.5.2 Maximum likelihood

Non-stationary distributions can be fitted using maximum likelihood estimation (MLE). One drawback of MLE, and a reason why it was not used in the FEH, is that it requires a numerical optimisation that does not always converge to a solution. You may occasionally find that MLE does not give an answer for a particular data set.

4.6 Ways of selecting between rival models

The MLE method can estimate the parameters of a particular statistical model, but in non-stationary flood frequency analysis there can be a large variety of candidate models to choose from. Even if there is only one covariate, there is a need to choose between models in which only the location, only the scale, both or neither vary. If there are several potential covariates the number of candidate models can grow rapidly. Which family of extreme value distributions to fit also needs to be considered.

There are many helpful ways of judging model quality, listed below. The first 3 approaches are statistical measures. They often, but not always, agree with each other in which models they select.

4.6.1 Likelihood ratio testing

A likelihood ratio test assesses which of a pair of statistical models (one more complex than the other) is the better fit.

It is a preferred approach when comparing a small number of candidate models, when the likelihood ratio can be calculated for each nested pair of models. It is impractical when comparing hundreds, which can be the case when several covariates are being considered. This has been implemented in the nonstat package to use when only time is a covariate (the **fit.time** function).

The package indicates which model is preferred according to the likelihood ratio test. No numerical output from the test is provided by the code.

4.6.2 AIC (Akaike information criterion)

The AIC establishes a trade-off between the goodness of fit and the simplicity of the model, measured by the number of parameters. It can be readily compared across a large number of candidate models, the lowest AIC indicating the preferred model. The AIC values themselves have no absolute meaning, but the model with lowest AIC of a group of models fit to the same data will be that which achieves the best trade-off. AIC is outputted by the nonstat package for each fitted model.

4.6.3 BIC (Bayesian information criterion)

This is similar to AIC, but BIC gives more weight than AIC to model simplicity. In calculating AIC, the penalty for the number of parameters k is $2k$; for the BIC, the penalty is $ln(n)k$ where n is the sample size, so if n > 8 then the penalty for a more complicated model is greater and so there is a preference for simpler models. BIC is also outputted by the nonstat package.

If in doubt, you are recommended to prefer BIC over AIC because there are advantages in fitting simpler models.

4.6.4 Hydrological reasoning

It is important that the statistical model makes physical sense. For instance, if rainfall is included as a covariate, it should act to increase one or more of the distribution parameters, so that higher rainfall is associated with higher peak flows.

Another consideration is that the covariates should not be too correlated with each other. This should not be a concern if you include only one physical covariate at a time, along with water year, in the models that you fit. If the physical covariate is correlated with time, this effect will be removed automatically by the package if the default setting to detrend the data based on the time covariate has been used. If you try to fit models with several physical covariates, first check for cross-correlations. Calculate a correlation coefficient for each pair of covariate time series in Excel or R, for instance.

4.6.5 Consistency of model form across locations

A consistent choice of covariates and type of relationship between covariates and parameters is expected for nearby and similar catchments, and particularly for gauges on the same river. For example, if all neighbouring gauges have a significant trend in the location parameter then it may be a good idea to include a trend in the location parameter, even if measures such as likelihood ratios or BIC indicate that models with a fixed location parameter are preferable.

4.6.6 Visual inspection of model fit plotted on probability-probability (P-P) and quantile-quantile[1] (Q-Q) plots

Refer to the science report for a more formal explanation of what P-P and Q-Q plots show. Together they are known as diagnostic plots.

The P-P plot compares the following 2 quantities, calculated for each annual maximum flow in the data set:

a) the non-exceedance probability of the flow, estimated from the statistical model

b) an empirical estimate of the probability made using a plotting position formula, which equally spaces the points along the x axis between 0 and 1

The Q-Q plot compares the following, again calculated for each annual maximum flow:

a) the flow estimated using the statistical model from the empirical probability at step (b) above

b) the measured flow

A well-fitting model will have diagnostic plots where the points lie close to a diagonal line. These indicate that the modelled probabilities and quantiles (flows) match their empirical equivalents closely. The P-P and Q-Q plot essentially contain the same information, but expressed on a different scale, so what looks like a reasonable fit on one scale may look poor on the other. In general, a good fit throughout the whole distribution may be preferable to having a very good fit for some points but a poor fit for others, particularly if the poor fit is in the higher values that are of most interest when considering extreme flows. The Q-Q plots can be useful to reveal where any poor fit occurs in the distribution (it may also highlight if there is a single outlying point which needs investigating for a potential data quality issue).

You can find example P-P and Q-Q plots in the second case study, section 6.2.

These visual inspections are feasible when a small number of candidate models is being compared. For example, you may find them useful to help choose between the GEV and GLO distributions. They are useful for seeing when models have not fitted well, so it is worth checking the plots for your selected model.

[1] A quantile is a statistical term for an estimate of a variable made for a particular probability

4.6.7 Confidence intervals

Compare the confidence intervals of the various fitted models. All other things being equal, give preference to models with narrower confidence intervals, since their results are more certain. Refer to the next section for more information.

4.6.8 Visual inspection of the flow estimates in comparison with the recorded flood peak data

This is the final check: do the model outputs look sensible? The nonstat package produces plots of the flow estimates superimposed on a time series of the annual maximum flow series to help you judge the sensibility of the outputs.

It is often interesting to examine the exceedance probability of the largest flood(s) that have been observed. This judgment can be conceptually more difficult in a non-stationary setting, where a flood that occurred in a particular year might have a different exceedance probability if it occurred earlier in the record, or in a year in which physical covariates such as annual rainfall or NAO were different.

4.7 Uncertainty

4.7.1 Confidence intervals

Uncertainty can be quantified using confidence intervals. The 95% confidence interval is the range within which we are 95% confident that the true answer lies. If we want a higher level of confidence, such as 99%, then we need to use a wider range.

4.7.2 Why consider uncertainty?

It is helpful to be able to quantify the uncertainty in flood frequency estimates so that we can measure whether it can be reduced, for example, by using a different method or by incorporating more data.

It is important to think about the implications that the uncertainty has on the output or outcome of a study. If the range of uncertainty has a significant impact on a decision that needs to be taken, that should act as a prompt to seek ways of reducing the uncertainty.

Non-stationary models of flood frequency are more complex than stationary models, with more parameters to fit, and so their results tend to have wider confidence intervals. In brief, there can be more scope for a non-stationary model to give an inaccurate answer even if statistical measures judge it to be the best fit to the data.

The case study on the River Kennal in section 6.2 has an example of calculating confidence limits and using them to help select a preferred model.

5 How to carry out non-stationary analysis: a step-by-step guide

5.1 General comment

Remember to refer to the nonstat package user guide for instructions on how to use the package. It covers issues such as the required formats for input data and how to understand the inputs and outputs of each function. To help with cross-referencing, this practitioner guidance mentions some function names in the package, which are given in **bold**.

5.2 Data needed

5.2.1 Annual maximum flows

For each gauging station you need an annual maximum series in .am format, as used by the National River Flow Archive.

5.2.2 Physical covariates

You can attempt to fit a non-stationary model using any variables as covariates. Refer to section 4.3 for some recommended covariates. You need to provide a covariate value to pair up with each annual maximum flow.

Useful sources of data for covariates include:

- the Centre for Ecology and Hydrology Gridded Estimates of Areal Rainfall (CEH-GEAR) dataset[2], which provides a 1 km grid of daily rainfalls from 1890 to the present day. Catchment-average time series of rainfall for all NRFA gauges are available from the NRFA website. From these, you can calculate seasonal or annual totals corresponding to the water year in which each annual maximum flow is recorded
- the National Oceanic and Atmospheric Administration (NOAA) website[3] for monthly values of indices of atmospheric circulation dating back to 1950. You can calculate seasonal or annual averages from the monthly figures. Use water years for compatibility with the way annual maximum flows are extracted
- historical maps or digital data sets from which you can estimate how the urban coverage has evolved during the period of record

[2] https://eip.ceh.ac.uk/rainfall

[3] https://www.cpc.ncep.noaa.gov/data/teledoc/telecontents.shtml

5.2.3 Transforming covariate data

It is good statistical practice to transform covariate data sets so that they are all on the same scale before they are included in a non-stationary analysis. To do this, first calculate the mean and standard deviation of the observed covariate values. Then subtract the mean from each observation and divide the result by the standard deviation. The **fit.time** function in the nonstat package carries out this transformation automatically for the time covariate. Users need to carry it out when supplying any physical covariate data.

5.3 Choosing a distribution

5.3.1 GLO or GEV?

The nonstat package can fit both the GEV and GLO distributions. To help manage the complexity of the non-stationary analysis, it can be helpful to select between these distributions on the basis of a stationary analysis, and then carry out the non-stationary analysis using only one distribution.

Try fitting the GEV and GLO distributions and compare the P-P and Q-Q plots (see section 4.6). Check for any errors or warnings, which tend to be more common for the GLO distribution. If all other things are equal you may want to choose the GLO since on average this has been found to give a better fit to flood peak data in the UK.

Refer to the River Kennal case study (section 6.2) for an example of selecting the distribution.

5.4 Fitting candidate models

5.4.1 Time as a covariate

With time as a covariate, for a particular distribution there are 4 candidate models that can be fitted:

- stationary
- location varying with time
- scale varying with time
- location and scale varying with time

All 4 are fitted at the same time by **the fit.time** function in the nonstat package.

5.4.2 Physical covariates: all combinations (Option 1)

With more than one potential covariate, the number of candidate models increases rapidly. One approach is to examine all possible combinations of covariates, for each of the location and scale parameters. There can be thousands of combinations, if there are many covariates. The **fit.phys.cov** function in the package can fit these combinations for up to 7 covariates. One drawback of this approach, apart from the time taken for the calculations, is the possibility that the models which are judged statistically to have the best fit are excessively complex, with too many covariates to be physically interpretable.

5.4.3 Physical covariates: one plus time (Option 2)

A simpler approach is to allow just one covariate per model parameter, or 2 covariates if one of them is time. This limits the number of combinations and avoids selecting excessively complex models. The combinations included in Option 2 include those that use only time as covariates, so it is possible that the preferred model will end up being one that has no physical covariates, or indeed a stationary model.

5.4.4 Physical covariates: one on its own (Option 3)

Simpler again is to allow only one covariate per model parameter. Option 3 fits models that use each covariate on its own. Within this option there cannot be a different covariate for the location and scale parameters.

5.5 Choosing a preferred model

Consider all of the approaches described in section 4.6, if feasible and applicable, when selecting models. Some will not be feasible if you are comparing a large number of candidate models.

Remember always to consider the stationary model (which is always included in the model combinations fitted by the nonstat package) alongside non-stationary options.

5.6 Extracting results from non-stationary analysis

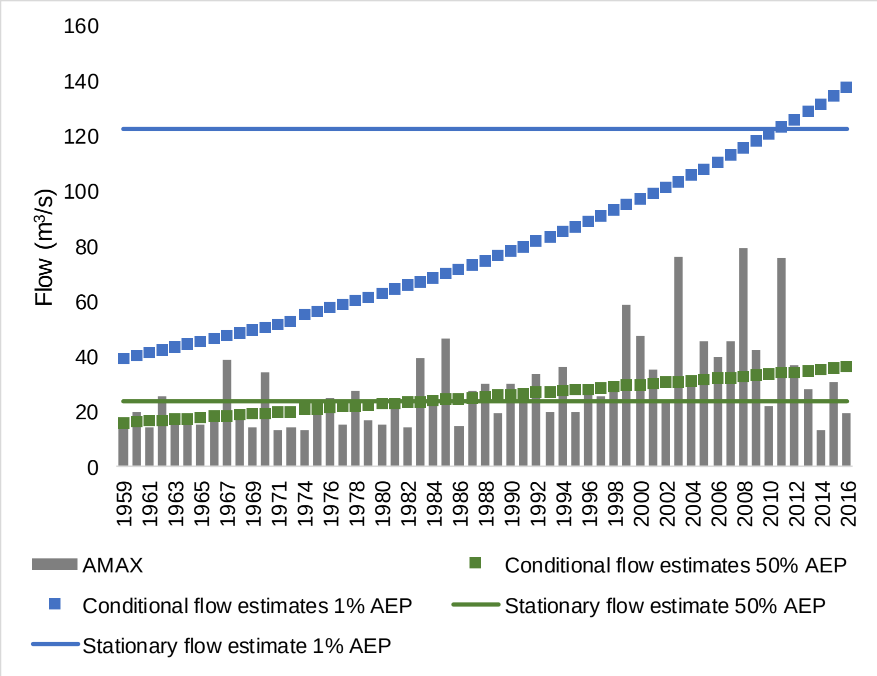
5.6.1 Time as a covariate

When you model non-stationarity using time as a covariate, the design flow estimates will be changing steadily over time (Figure 5-1). The **fit.time** function outputs a table of results with a set of flow estimates for each year of the observed record. These results are known as conditional flow estimates. For instance, the conditional estimate for 2015 is the expected flow under the (clearly unrealistic) conditions that the year is always 2015.

Figure 5-1 illustrates conditional flow estimates based on only time as a covariate. The symbols show the conditional flow estimate for each year of record at an example station, for 2 values of AEP. There is a time trend in the location parameter, which manifests itself as an increasing trend in the conditional estimates. There is also a trend in the scale parameter, manifested as an increase in the ratio between the 1% and 50% flow estimates.

The stationary flow estimates are included for comparison.

Figure 5-1 Illustrating conditional flow estimates based on only time as a covariate



If you are estimating the return period of an observed flood:

The return period or probability of a particular event is not a straightforward concept under non-stationary conditions. The same event may have a different probability of occurring nowadays than it did a few years ago.

Look at the conditional flow estimates for the water year when the flood occurred. Compare the estimates with the observed peak flow of the flood to estimate its probability of occurrence in the year when it occurred. There is an example of this in the case study on the River Kennal.

You can also look at conditional flow estimates for other years to estimate the probability of the same flood occurring nowadays, or at other times.

If you are estimating flood frequency in present-day conditions:

If the flow record goes up to the present day (give or take a water year), you can adopt the conditional flow estimates for the last year of record as the present-day results.

If you are estimating flood frequency for scheme design:

Refer to the guidance on adjusting for climate change in section 5.7. There are 2 options available.

5.6.2 Physical quantities as covariates

Design flow estimates from a non-stationary model that uses physical covariates will change not only over time, if water year is included as a covariate, but also with the value of the other covariates. Figure 5-2 shows an example where both time and NAO are covariates.

For instance, if the covariate is annual rainfall, then the 1% AEP design flow given 1,200 mm of rainfall is the expected flow under the (clearly hypothetical) conditions that the annual rainfall is always 1,200 mm.

Because we know the values of all covariates for past years, we can calculate a conditional flow estimate for each year. For instance, the conditional estimate for 2015 is the expected flow under the (again hypothetical) conditions that the year is always 2015 and the annual rainfall is always that observed in 2015.

The conditional flow estimate may be useful when examining the probability of past floods, but it is less informative when thinking about design.

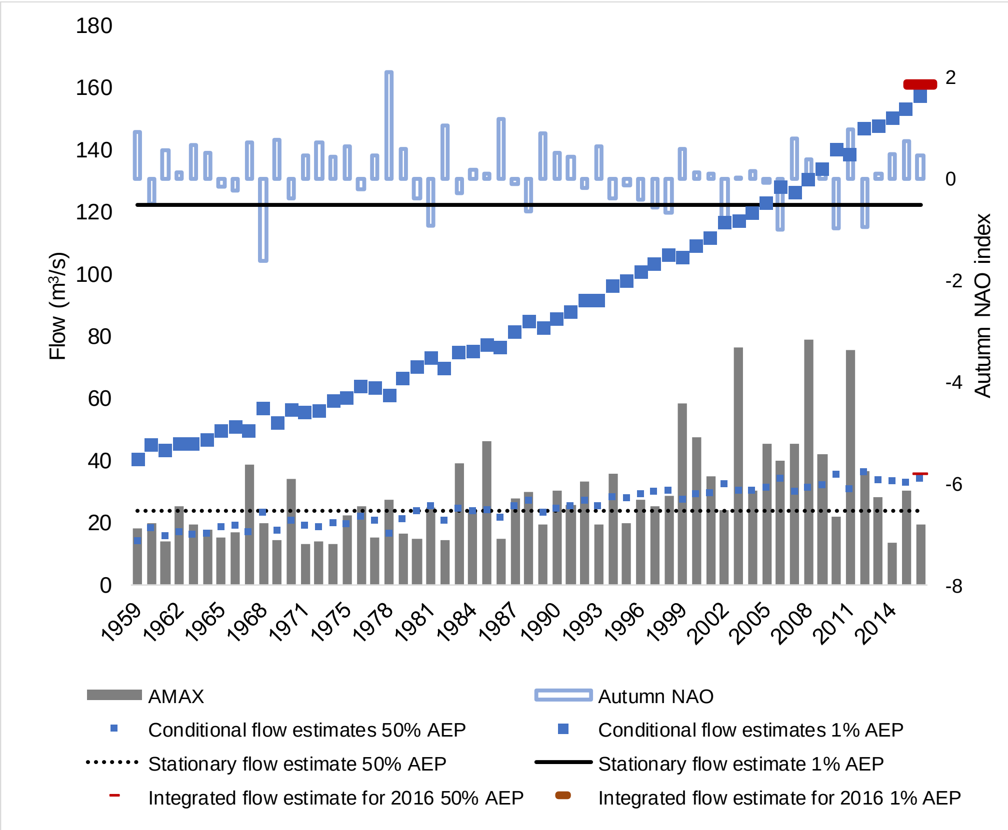
Figure 5-2 illustrates conditional flow estimates based on both time and physical covariates. The symbols show the conditional flow estimate for each year of record at an example station, for 2 values of AEP. The interannual variation in the conditional estimates is due to the influence of a physical covariate, the autumn NAO. More negative values of autumn NAO are associated with larger annual maximum flows.

As in Figure 5-1, time still has a strong influence as a covariate at this site, resulting in a large increase in the conditional estimates over the period of record.

The stationary flow estimates are included for comparison. An explanation for the red symbols is given later.

When examining results from the R package, you can investigate the effect that each individual covariate has on the flood frequency distribution by examining the parameters output by the package. For each covariate and each distribution parameter, the **fit.phys.cov** function outputs a gradient term which expresses how the covariate affects the parameter value. These gradient terms are named μ_1, μ_2,… for the location parameter and ϕ_1, ϕ_2, … for the scale parameter. If the gradient is positive, the parameter increases as the covariate increases.

Figure 5-2 Illustrating conditional flow estimates based on both time and physical covariates



5.6.3 Integrated flow estimate

The integrated flow estimate[4] removes the dependence on a particular value of the covariates. It is calculated by averaging the probabilities corresponding to the conditional flow estimates, over a sample or a statistical distribution of covariate values.

Refer to the science report for a formal mathematical definition. For an informal explanation, refer to the following process, which assumes that you are wanting to obtain the integrated flow estimate relevant to the observed period. These calculations are carried out automatically within the **res.phys.cov** function in the nonstat package.

1. Choose a trial value for the flow estimate as a starting point.
2. For each year in your annual maximum flow series, calculate the probability of the flow exceeding the value from step 1 given the value of the observed

[4] The integrated flow estimate is more formally known as the marginal return level.

covariate(s) in that year. This probability comes from the fitted non-stationary model.

3. Average all the probabilities you calculated at step 2.
4. Repeat from step 1 until your averaged probability is close enough to the exceedance probability you require.

The integrated flow estimate is defined relative to a period of time, usually longer than a year. It is usually calculated by averaging over all the covariate values observed within the period of river flow measurements, as above. It is the flood flow associated with a particular encounter probability (see section 4.4).

When the water year is not included as a covariate, the integrated flow estimate does not vary over time and so can be understood as a temporally stationary estimate.

In theory, the integrated estimate can also be calculated by averaging over a different distribution of covariate values, for instance one that is intended to represent future conditions. This is not currently recommended.

Because the integrated flow estimate is defined for a period of time rather than an instantaneous point in time, it can be misinterpreted if it is plotted on a time series graph like that shown in Figure 5-2. It is more appropriate to plot it as illustrated in Figure 4-1, which uses the same data set as Figure 5-2.

The case study on the River Eden in section 6.1 provides another example to illustrate the integrated flow estimate.

5.6.4 Single-year integrated flow estimate

If the covariates include both water year and physical variables, it is possible to calculate an integrated flow estimate by averaging the probabilities corresponding to the observed physical covariate values, but setting the water year covariate to a single value, such as the final year of record. This gives the single-year integrated flow estimate. The **res.phys.cov** function can calculate this for the last year of record.

If the flow records runs up to, or nearly up to, the present day, this estimate is representative of the present-day expected flow for a particular exceedance probability, without being conditional on any particular value of a covariate such as annual rainfall. You could probably assume it is representative of the short-term future too.

The single-year integrated flow estimate can be more easily compared with alternative estimates such as those from a model that uses only water year as a covariate. The red symbols on Figure 5-2 show the single-year integrated flow estimates for the last year of record, 2016. They are similar but not identical to the conditional flow estimates for 2016, because the latter are conditional on the value of the autumn NAO matching that observed in 2016.

5.6.5 Extracting results

Follow the advice listed for the situations listed below. This assumes that time is included as a covariate alongside physical quantities.

If you are estimating the return period of an observed flood:

If you want to know the annual probability of a flood of that magnitude around the year it occurred, calculate a single-year integrated flow estimate for the year when the flood occurred. The probability for any particular year will depend on the physical covariates for that year, but that dependence is removed by calculating the integrated flow estimate.

If you want to know the probability of a flood of that magnitude occurring at some unspecified time during the whole of the observed record, use the integrated flow estimates that correspond to the period of record.

If you are estimating flood frequency in present-day conditions:

Use the single-year integrated flow estimate, calculated for the most recent year of record.

If you are estimating flood frequency for scheme design:

Refer to the guidance on adjusting for climate change in section 5.7. You can apply option 1 using the single-year integrated flow estimate, calculated for the most recent year of record. Option 2 is more difficult to apply when there are physical covariates in addition to water year. It could be applied by calculating a single-year integrated flow estimate, setting the year to the mid-point of the 1961 to 1990 baseline period. The R package cannot currently do this.

5.7 Adjusting for future climate change

5.7.1 Background

Refer to the background information about causes of non-stationarity in section 3.3.

Current guidance on accounting for climate change in flood risk management investment decisions is given in: ‘Adapting to Climate Change: Advice for Flood and Coastal Erosion Risk Management Authorities’.

Adjusting non-stationary flood frequency estimates for climate change is not straightforward. Here is some guidance on what not to do.

5.7.2 What not to do

Do not project non-stationary analysis into the future, for example, setting the time covariate to a future water year.

Do not assume that physical covariates can help with estimating future flood frequency. We could only be confident of this if we had a non-stationary model that included all the important physical factors and relationships that influence the effect of climate change on flood flows.

For example, annual rainfall is often a useful covariate to explain the variation of annual maximum flows. Because we can obtain estimates of the impact of climate change on annual rainfall, it might be tempting to think that we can use these in a non-stationary model to estimate future flood frequency. This would only be valid if the only way that climate change affects river floods is via changes in total annual rainfall. This is not the case as changes in storm intensity and evapotranspiration are also likely to be influential.

Avoid assuming that observed trends can be attributed to climate change in the absence of any other obvious causative factors (such as urbanisation). Without in-depth research we just don’t know enough to deduce this, but we can be tempted to think we do because we like to find patterns and stories to explain observations.

5.7.3 Some ways forward

There are 2 suggested options for adjusting non-stationary flood estimates for climate change:

Option 1: Applying climate change allowances in full to a present-day estimate of flood frequency. The thinking behind this is that because we cannot be confident of the extent to which climate change is causing trends, we will assume that it is not until we have evidence to the contrary. This approach reduces the risk of underestimating the impact of future climate change.

Option 2: Applying climate change allowances to a 1961 to 1990 baseline estimate. This could give more accurate future flood estimates on catchments where observed trends are mainly due to climate change. It is explained here for the case where time is the only covariate in the non-stationary analysis. The steps are:

a) Estimate the flow from the non-stationary model, setting the water year covariate to 1975, the mid-point of the 1961 to 1990 baseline period. This gives the conditional flow estimate for the year 1975.

b) Compare the result with the stationary estimate, representative of the whole period of record.

c) If the 2 estimates are significantly different, allowing for their confidence intervals, go to step (d). If not, it may be better to adopt option 1 above.

d) Ask if the difference between the 2 estimates could be accounted for by anything other than climate change. In the unlikely event that you are confident enough that the answer is no, go to step (e), otherwise revert to option 1. Bear in mind the cautionary advice above in ‘What not to do.’

e) Adopt the non-stationary estimate for the 1961 to 1990 period as a baseline and then adjust it using the currently recommended change factors for the desired future epoch. For a present day estimate, use the 2020s epoch change factor.

Refer to the case study on the River Kennal for an example of these approaches.

Guidance on climate change allowance for flood peaks is in the process of being updated, using the UKCP18 probabilistic projections for river basin regions. As for previous work, the baseline period for the projections was 1961 to 1990, with a longer baseline for the hydrological modelling, 1961 to 2001. The intention is for the outputs of the project to be made available via a web tool.

Substantially more research would be needed to derive climate change allowances suitable for different baseline periods, such as one more representative of present-day conditions.

5.8 Comparing with FEH estimates

5.8.1 Single-site versus pooled

Aside from the fundamental difference in the assumption of stationarity, one of the main differences between a non-stationary flood frequency analysis, carried out according to this guidance, and an FEH analysis, is that FEH analysis tends to use pooling groups.

There are 2 main advantages to carrying out pooled analysis rather than single-site analysis:

- it reduces the estimation error, by averaging the analysis over several stations
- it permits flood frequency estimation at ungauged sites

Although pooled application of non-stationary methods is feasible, it is not yet recommended for use by UK practitioners. The science report explains how non-stationary pooled analysis was explored and gives recommendations for further work.

5.8.2 Other differences

Another reason why FEH results may differ is that FEH methods use L-moment ratios to fit flood frequency curves, whereas the methods described in this guidance use maximum likelihood estimation. This can sometimes lead to significant differences.

To compare non-stationary flood estimates against an equivalent FEH at-site analysis, carry out a single-site analysis of the same AMAX series using FEH methods (WINFAP). This should **always** be done to provide a comparison between the results from a ‘benchmark’ stationary FEH, and non-stationary methods described in the report.

6 Case studies

Each of the 3 case studies illustrates a different aspect of non-stationary flood frequency analysis. To avoid repetition, issues such as choice of distribution, understanding confidence limits or handling physical covariates are not repeated; instead each is dealt with in just one case study. Look at the first 2 case studies, on the Rivers Eden and Kennal, to cover most aspects. The third, on the Little Ouse, deals with how to handle decreasing trends.

All calculations described in the case studies were carried out using the R package, apart from some simple additional analysis of the results using Excel. The graphs shown are either those output by the R package or produced in Excel from the results files from the package. Please note that results from the final release of the package may differ slightly from those in the case studies.

6.1 Eden at Temple Sowerby (station 76005)

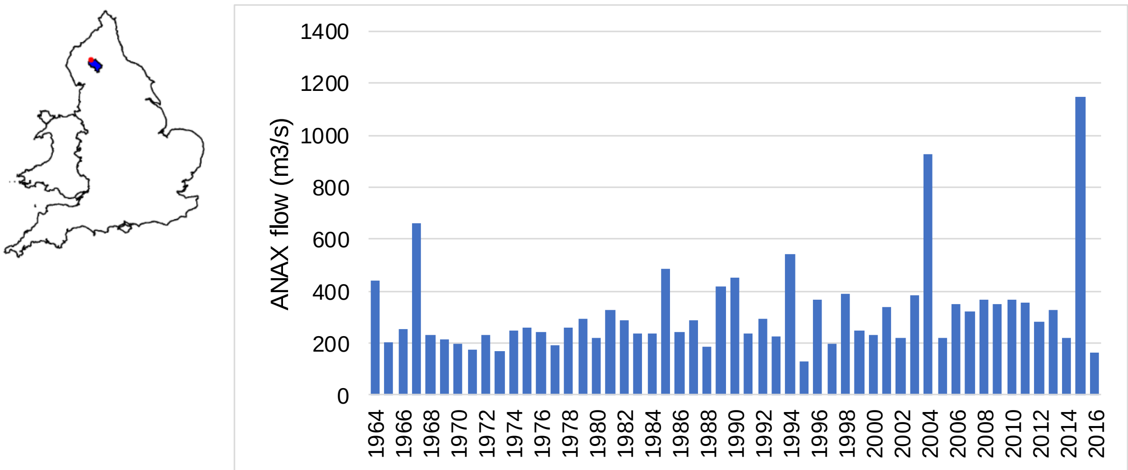
The main purpose of this case study is to illustrate how covariates can be chosen. Both time and physical quantities are considered, separately and together, as candidate covariates.

6.1.1 Catchment and data

The Eden at Temple Sowerby is a 616 km^2 catchment draining a rural upland area in the Northern Pennines. There has been no appreciable urban development or other large-scale land use change during the period of record.

There are 53 annual maximum flows in the data set used for this project, from 1964 to 2016.

Figure 6-1 Catchment location and flood peak time series, Eden at Temple Sowerby (station 76005)



6.1.2 Trends

Non-parametric test

The Mann-Kendall test shows signs of an increasing trend over the period of record. For the record as a whole, there is a trend of fairly high significance, with p=0.06, indicating that there is only a 6% chance of obtaining a trend at least as extreme as the one seen in the data set, under a null hypothesis of no trend.

As the p value is just greater than 0.05, there is no statistically significant trend at the 0.05 significance level.

There is a significant trend at the 0.05 significance level for most subsets of the record that begin in the 1960s to 1970s and end in the 2010s.

Change point tests

Neither the Pettitt nor the PELT test detect any significant step changes.

6.1.3 Non-stationary analysis: time as a covariate

For the purpose of this case study, only the GEV distribution was considered.

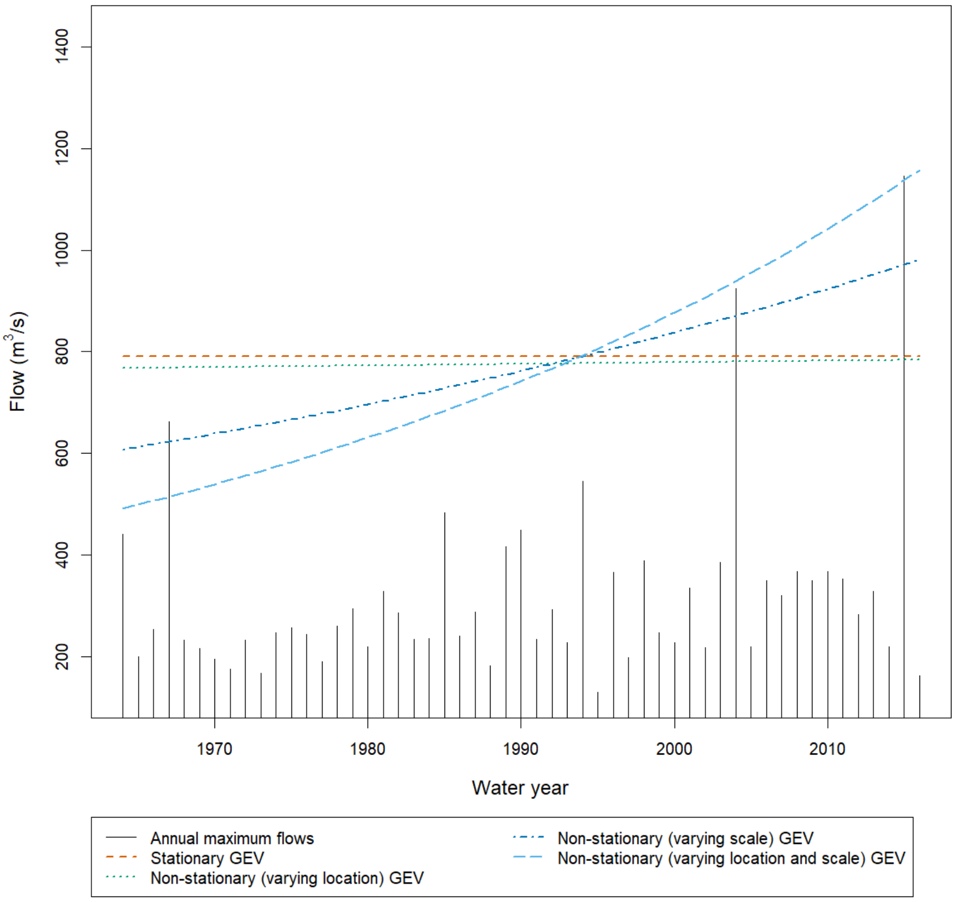
Four candidate models were fitted: stationary, location varying with time, scale varying with time and both location and scale varying with time.

The highest quality model, according to both likelihood ratios and BIC, is the stationary model.

The varying location model produces similar results to the stationary model, but the varying scale models show a strong trend, so that the estimated flow for low AEPs increase greatly between the start and the end of the record, as shown in the plot below.

In the case of the varying location and scale model, this increase amounts to more than a doubling of the estimate over the period of record. According to this model, the severe flood of December 2015 had an AEP of 2%, less severe than that of the March 1968 flood even though the peak flow in 2015 was 75% higher than in 1968. These sort of results are difficult to believe, lending further weight to the preference for the stationary model over the non-stationary options with only time as a covariate.

Figure 6-2 Plot of time-varying model results for the 2% AEP, Eden at Temple Sowerby (station 76005)



6.1.4 Non-stationary analysis: physical covariates

Non-stationary models were fitted using the following covariates, both in addition to time and on their own:

- catchment-average rainfall (annual, autumn and winter)
- North Atlantic Oscillation (NAO) (winter, summer and autumn)
- East Atlantic pattern (EA) (winter)

Each covariate was allowed to influence the location, the scale and both parameters in separate candidate models. This gave 88 candidate models in total. The covariates were detrended by the nonstat package.

The model with the lowest BIC used the following covariates:

- location parameter: time and annual rainfall
- scale parameter: none (constant)

It is also interesting to examine other models with low BIC and/or AIC statistics, since there is little difference in the value of either of these statistics for the best-ranking models. The model with the second lowest BIC, and lowest AIC, had these covariates:

- location parameter: time and annual rainfall
- scale parameter: time

Annual rainfall appears to be a physically reasonable covariate to include on a rural catchment with low evaporation like the Eden. Water years with high rainfall will lead to higher soil moisture and more run-off.

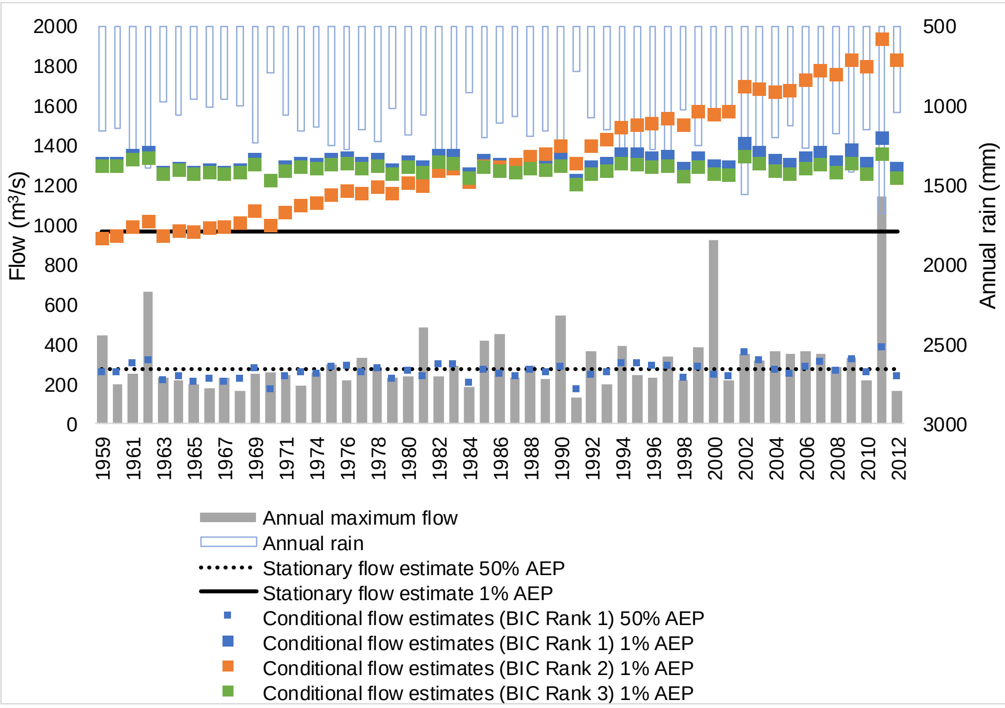
BIC and AIC values for some of the fitted models are listed below. The stationary model performs much less well than the highest-quality models according to these statistical measures.

Table 6-1 Statistical quality of example models, Eden at Temple Sowerby (station 76005)

Covariates for location	Covariates for scale	AIC	BIC	BIC rank (lowest = 1)
Time and annual rainfall	None	629.1	639.0	1
Time and annual rainfall	Time	628.6	640.5	2
Annual rainfall	None	634.1	642.0	3
None (stationary model)		652.3	658.3	21

The graph below compares the stationary flow estimates with the conditional estimates from the top 3 ranking models given above. Only one set of conditional estimates is shown for the 50% AEP; the others are very similar.

Figure 6-3 Time series plot of conditional flow estimates from the physical covariate models, Eden at Temple Sowerby (station 76005)



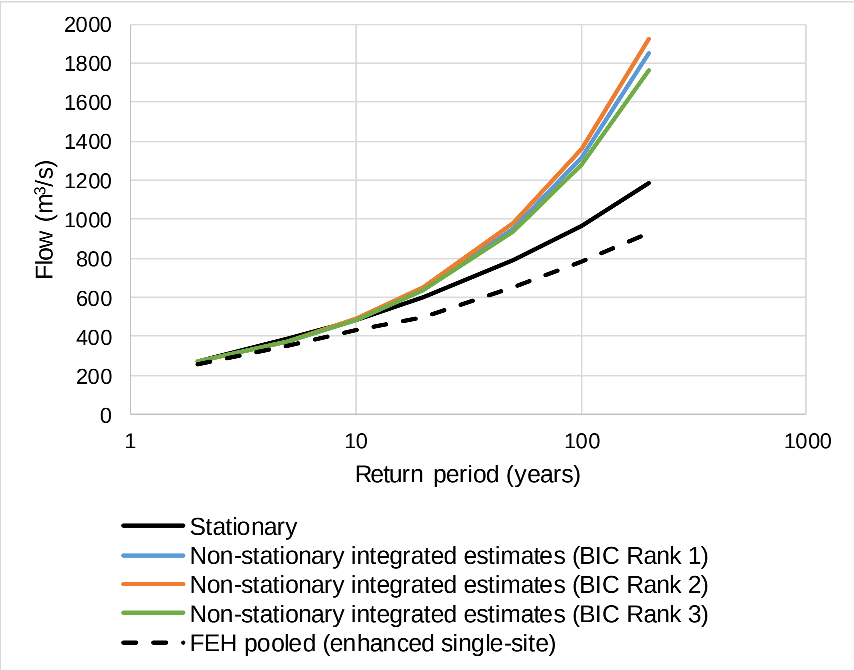
The above plot illustrates some important points.

Firstly, **models that have a similar quality according to statistical measures like AIC or BIC may have quite a different model form**. Compare the 1% AEP conditional estimates for models ranking 1 and 2. The difference between the models is that the rank 2 model includes time as a covariate for the scale parameter. This acts to steepen the flood frequency curve as time goes on, so by the end of the record the ratio of the 1% AEP to the 50% AEP is much higher than at the start. Take this as a reminder not to rely only on statistical measures to judge model suitability.

Secondly, **non-stationary models do not necessarily have a consistent time trend in their results**. The rank 3 model does not include time as a covariate. For the rank 1 model, although time is a covariate for the location parameter, its effect is minor compared with the interannual variation which is controlled by the annual rainfall. You can see this from the position of blue symbols on the plot: they jump around from year to year without much discernable long-term trend. These results may not all look like what you expect from a non-stationary analysis.

Thirdly, **conditional flow estimates are not necessarily helpful ingredients for a design process.** They are defined relative to hypothetical circumstances. The integrated flow estimates are more useful because they are averaged over the distribution of the covariates. See the plots below.

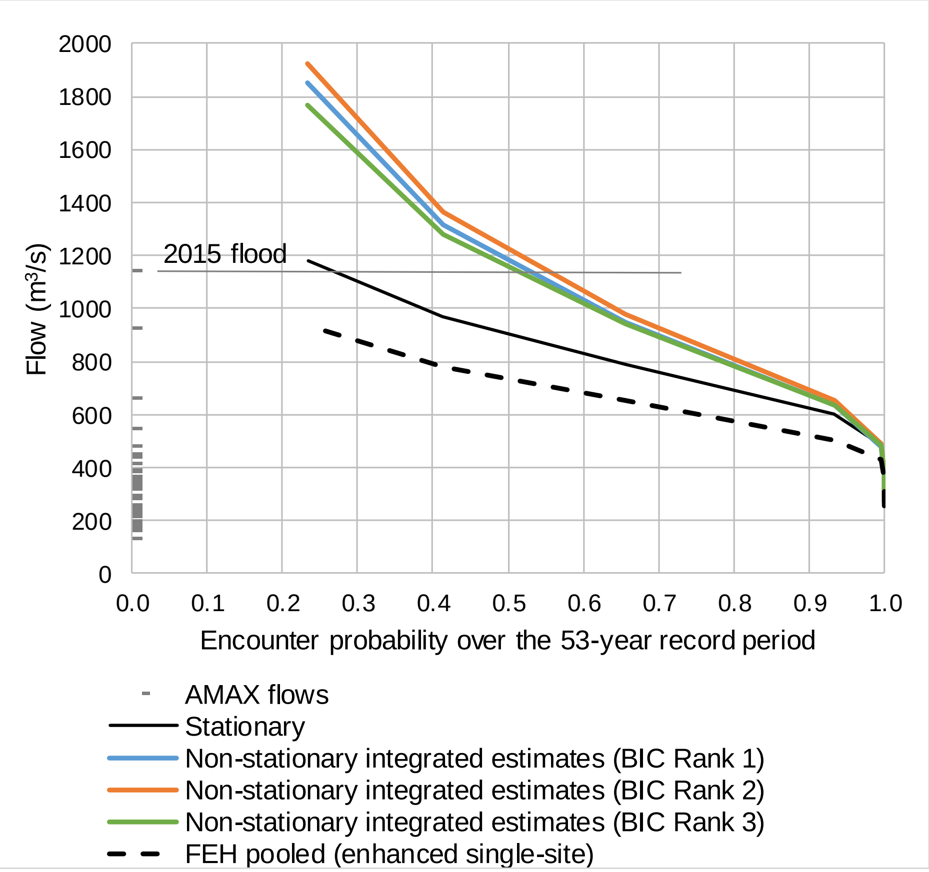
Figure 6-4 Frequency plot of integrated flow estimates from the physical covariate models, Eden at Temple Sowerby (station 76005)



Understanding the frequency plot

Figure 6-4 is included to provide a link with conventional flood frequency plots which will be familiar to many practitioners. However, it needs to be interpreted with care. The concept of return period is not easily applied under non-stationary conditions (see section 4.4). You should not interpret it in the usual way as the average time between floods of a given size, because this ‘expected waiting time’ is not a constant value under non-stationary conditions. The integrated flow estimates are defined relative to their encounter probability over the entire period of record and so it is more appropriate to plot them against this probability, as shown below.

Figure 6-5 Encounter probability plot, Eden at Temple Sowerby (station 76005)



Understanding the encounter probability plot (Figure 6-5)

As on the frequency plot (Figure 6-4), the 3 non-stationary models generate similar results. Although their conditional flow estimates appear quite different, depending whether the time trend is included in the scale parameter, the integrated estimates are similar. All 3 non-stationary models show higher flow estimates than the stationary model, with the difference growing for low probabilities.

The plot does not show results for encounter probabilities below 23%; this corresponds to an annual probability of 0.5% (from equation 1).

The plot includes results from a conventional FEH statistical analysis, with an enhanced single-site growth curve. This gives lower estimates than any of the single-site analyses. The FEH results were calculated using equation 1 applied to design flood estimates from WINFAP.

Another way of exploring the results is to examine the encounter probability associated with the highest flow on record, 1,146 m^3/s during Storm Desmond on 5 December 2015. According to the stationary model shown in the above plot, the probability of this flow occurring during the period of record was 0.25. The non-stationary models estimate a much higher encounter probability, of about 0.50 to 0.55.

6.1.5 Preferred model

In terms of integrated flow estimates, there is little to choose between the 3 non-stationary models. The model with BIC rank 3 is the simplest, with only annual rainfall as the covariate for the location parameter, and a constant scale parameter. This is suggested as preferable. The lack of significant trend over time may be surprising, although it is consistent with the findings of the Mann-Kendall test.

The integrated flow estimates from the preferred model can be interpreted as temporally stationary estimates, because the one covariate included in the model (detrended annual rainfall) is stationary. Because the preferred model does not include water year as a covariate, there is no need to calculate a single-year integrated flow estimate. The integrated flow estimates can be more straightforwardly interpreted in terms of an annual exceedance probability, or return period, as on Figure 6-4.

This choice of model is consistent with the findings from the national-scale analysis at some surrounding catchments. For example, for the Eamont at Udford, a tributary that joins the Eden shortly downstream of Temple Sowerby, the lowest-BIC model has an identical form, with only annual rainfall as the covariate for the location parameter, and a constant scale parameter. The same is found in several catchments in west Cumbria.

Further work could examine diagnostic plots and confidence limits, which are illustrated in the other case studies.

6.1.6 Results and final comments

The table below compares the design flows from the recommended model with the stationary estimates. Results are provided for a range of probabilities, expressed in 3 ways: return period, AEP and encounter probability over the period of record.

Table 6-2 Design flows from the recommended model with the stationary estimates, Eden at Temple Sowerby (station 76005)

Return period (years)	2	5	10	20	50	100	200
AEP (%)	50%	20%	10%	5%	2%	1%	0.5%
Encounter probability over 53-year length	100%	100%	100%	93%	66%	41%	23%
Flow (m^3/s) from preferred non-stationary model (BIC rank 3)	268	374	484	637	940	1,282	1,767
For comparison: flow (m^3/s) from stationary model	272	387	486	602	791	969	1,183

Some of the findings from this case study may seem counter-intuitive. The non-parametric trend test shows a p-value of 0.06, indicating no trend at a 0.05 significance level. When time is considered as a covariate, the preferred model is a stationary one, with no covariates. Yet when physical covariates are included, there is a clear preference for models that include those covariates. The chosen model yields much higher flow estimates for low-probability floods than the stationary model, and yet this model has no consistent trend over time. It can be regarded as another version of a stationary model, one with a covariate that does not have a time trend.

The Eden has experienced 2 major floods this century, in 2005 and 2015, with flows of 925 and 1,146 m^3/s respectively. The results above indicate that neither flood is as exceptional as might be thought based on a stationary analysis with no covariates. The AEPs of the 2 floods are estimated as 5 to 10% and 1 to 2% respectively, compared with 1 to 2% and 0.5 to 1% respectively from the model with no covariates.

For low-probability floods, design flows from the preferred model are strikingly higher than those from a conventional FEH pooled flood frequency analysis.

6.2 Kennal at Ponsanooth (station 48007)

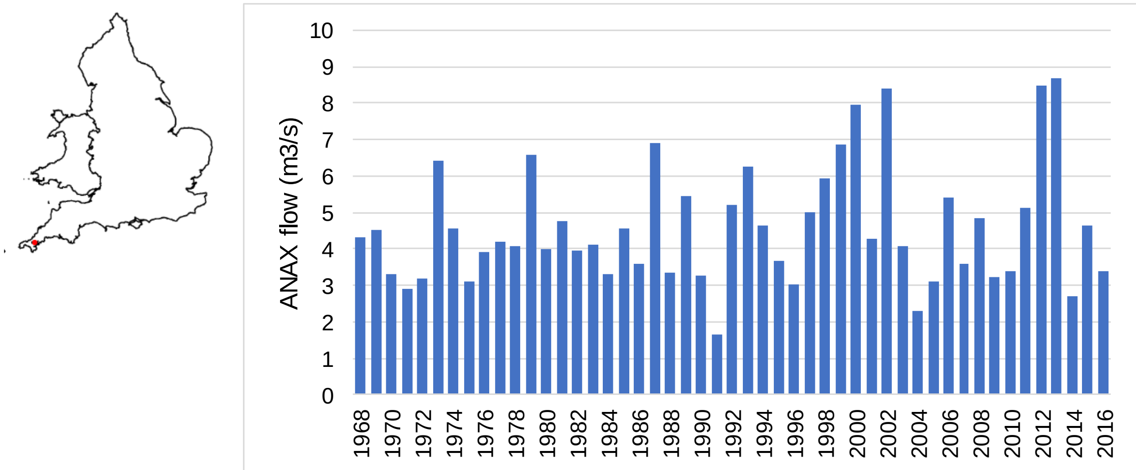
The main purpose of this case study is to illustrate the use of diagnostic plots and confidence intervals to help choose between different distributions and types of model fit. The only covariate considered for this case study is time. The case study also illustrates how to incorporate an allowance for climate change.

6.2.1 Catchment and data

The River Kennal at Ponsanooth is a 26 km^2 catchment draining a rural area of Cornwall. The flow regime is affected by Stithians Reservoir, built before the start of the gauged record, which produces significant attenuation (FARL at the gauge is 0.87). 63% of the land use in the catchment is grassland.

There are 49 annual maximum flows in the data set used for this project, from 1968 to 2016.

Figure 6-6 Catchment location and flood peak time series, Kennal at Ponsanooth (station 48007)



6.2.2 Trends

Non-parametric test

The Mann-Kendall test shows a moderate increasing trend, with a Z score of 0.90. This has a low significance (p=0.36), which indicates that there is a 36% chance of obtaining a trend at least as extreme as the one in the data set, under a null hypothesis of no trend.

The fact that the highest 4 floods have all occurred in the second half of the record gives an impression of an increase.

Change point tests

Neither the Pettitt nor the PELT test detect any significant step changes.

Other tests

Another statistical test detects a significant difference between the variance of annual maximum flows in the years up to 1990 and the variance for 1991 onwards. This is the Brown-Forsythe test (Brown and Forsythe, 1974). The null hypothesis is that the samples of AMAX flows are drawn from populations with equal variance. At a 0.05 significance level, the null hypothesis was rejected for the Ponsanooth data set.

6.2.3 Non-stationary analysis: time as a covariate

Both the GEV and GLO distributions were considered. For each distribution, 4 candidate models were fitted: stationary, location varying with time, scale varying with time and both location and scale varying with time.

Compare the 2 pairs of diagnostic plots below (Figure 6-7 and Figure 6-8), both for the stationary fit. On these plots, good-fitting models have points that plot close to the diagonal line. Refer to section 4.6 for guidance on interpreting them.

Figure 6-7 Diagnostic plots for GEV stationary fit, Kennal at Ponsanooth (station 48007)

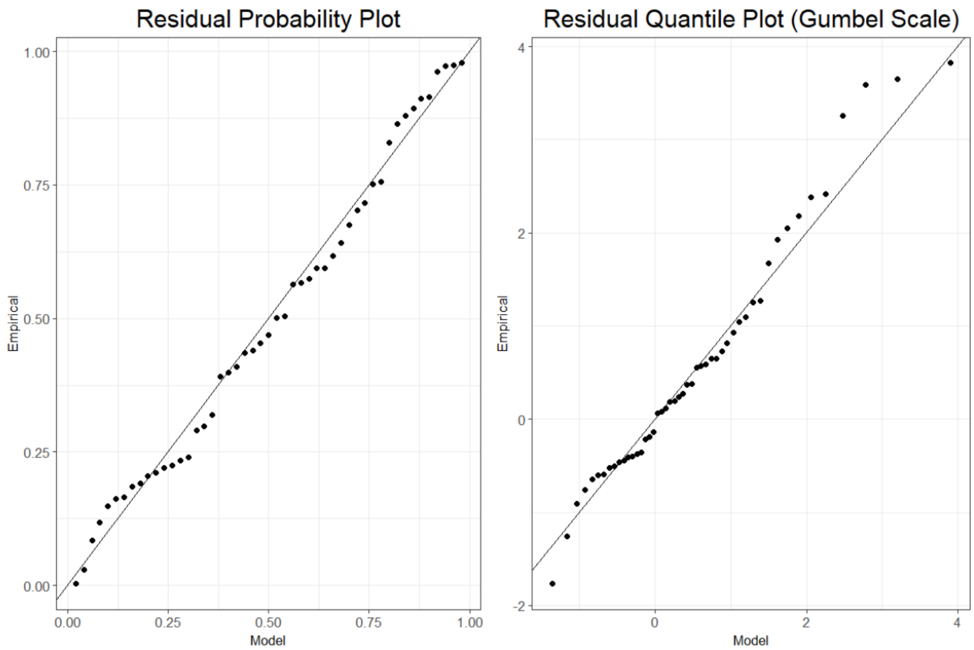
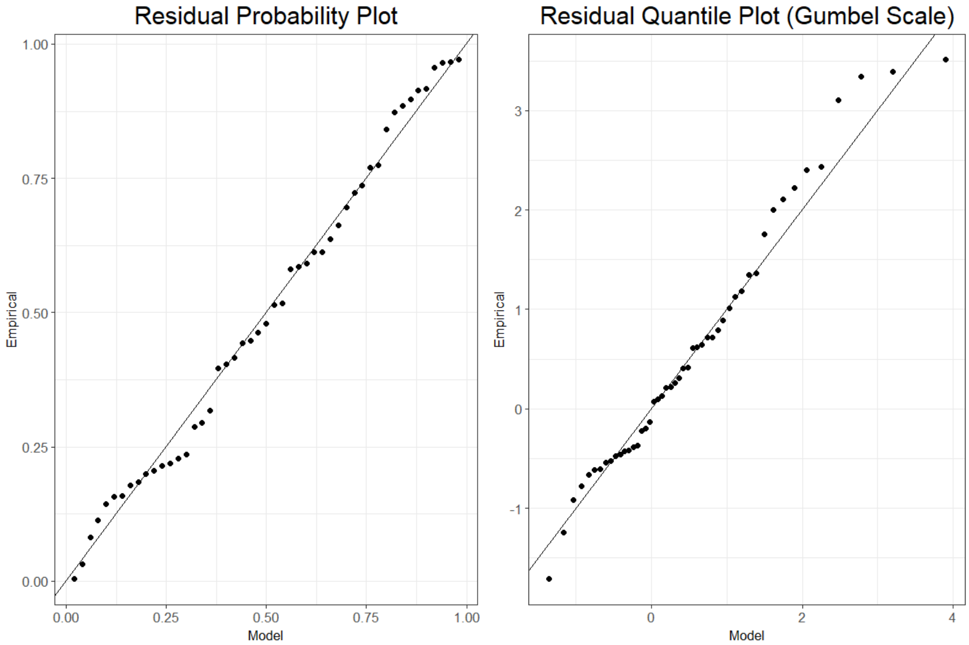


Figure 6-8 Diagnostic plots for GLO stationary fit, Kennal at Ponsanooth (station 48007)



Selecting the preferred distribution

There is little difference between the 2 distributions in terms of goodness of fit. The probability-probability plots (left) and the quantile-quantile plots (right) both compare modelled quantities, on the x axis, with their equivalents directly from data, on the y axis.

The Q-Q plots for the GEV and GLO are very similar. Both the stationary GEV and GLO models appear to underestimate most of the higher quantiles (this can be seen from the higher points lying to the left of the diagonal line).

On the P-P plot, some of the points in the middle appear to plot slightly closer to the diagonal line when they are modelled using the GLO distribution. For this reason, the GLO was selected.

The 4 Q-Q plots below (Figure 6-9) compare the stationary GLO model with 3 non-stationary versions of the GLO.

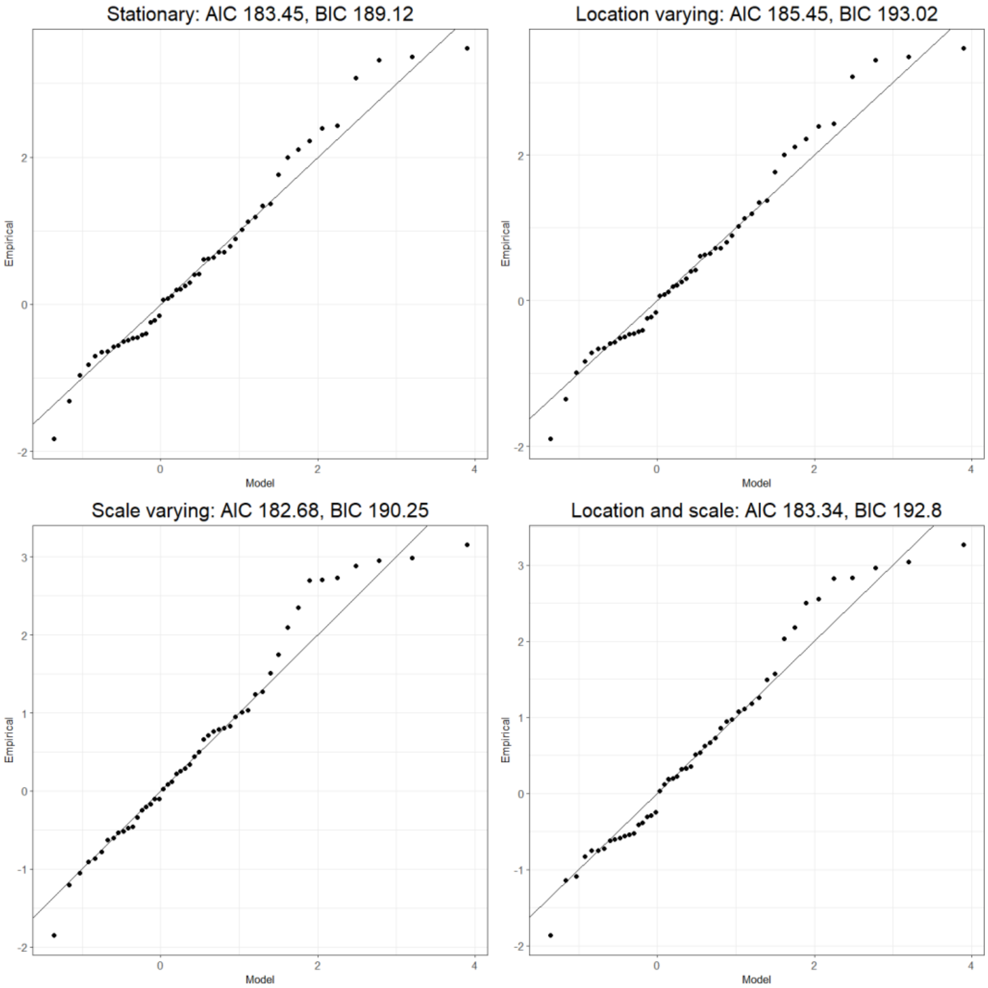
6.2.4 Discussion of model fits

AIC and BIC values for each candidate model are included on the Q-Q plots below. There is little difference in the goodness of fit of the models. The model with the lowest BIC is the stationary version. This model is also preferred when likelihood ratios are compared (the information is given in the output from the package). In contrast, the AIC is lowest for the model in which the scale parameter varies with time. A drawback of the scale-varying model is that its fit does not appear to be as good as some other models, as judged from the Q-Q plot. The general underestimation of higher quantiles, noted earlier for the stationary model, appears more marked for the scale-varying model.

The points plotted from the stationary and location-varying models appear to fit closer to the diagonal line.

The next step in selecting a preferred model is to examine the results and their confidence intervals.

Figure 6-9 Comparing Q-Q diagnostic plots for non-stationary model types, Kennal at Ponsanooth (station 48007)



6.2.5 Results and impacts of climate change

Discussion of results

The non-stationary conditional flow estimates all show an increasing trend over time. The gradient of the trend line increases with the return period, due to the fact that the (increasing) scale parameter has more influence on longer return periods.

The 90% confidence intervals are similar in width for the 5-year flood, but for the 100-year, at the end of the record the non-stationary confidence interval is much wider. The interval is (8.4,16.2) m^3/s for the stationary estimate and (9.0,22.6) m^3/s for the non-stationary estimate. The stationary estimate falls within the 90% confidence limits of the non-stationary result, and vice versa. This can be taken as indicating no significant difference between the stationary and non-stationary results, in present-day conditions. However, the 2 flow estimates may well give rise to substantial differences in predicted flood damages and costs of any flood protection measures.

One way to help choose between the alternatives is to examine the local and regional pattern of non-stationarity. For all neighbouring catchments around the Kennal, a stationary model is preferred over any time-varying model. It seems that the apparent non-stationarity at Ponsanooth is something of a local anomaly. Further investigation could look at comparison of record coverage at nearby gauges, and local trends in rainfall and land use.

For the purpose of the rest of this case study, to illustrate ways of adjusting for possible climate change impacts, we will adopt the non-stationary estimates. For the present-day, the 100-year estimate is 13.5 m^3/s.

Figure 6-10 Plot of GLO model results and confidence intervals for 5-year return period, comparing stationary and non-stationary (varying scale), Kennal at Ponsanooth (station 48007)

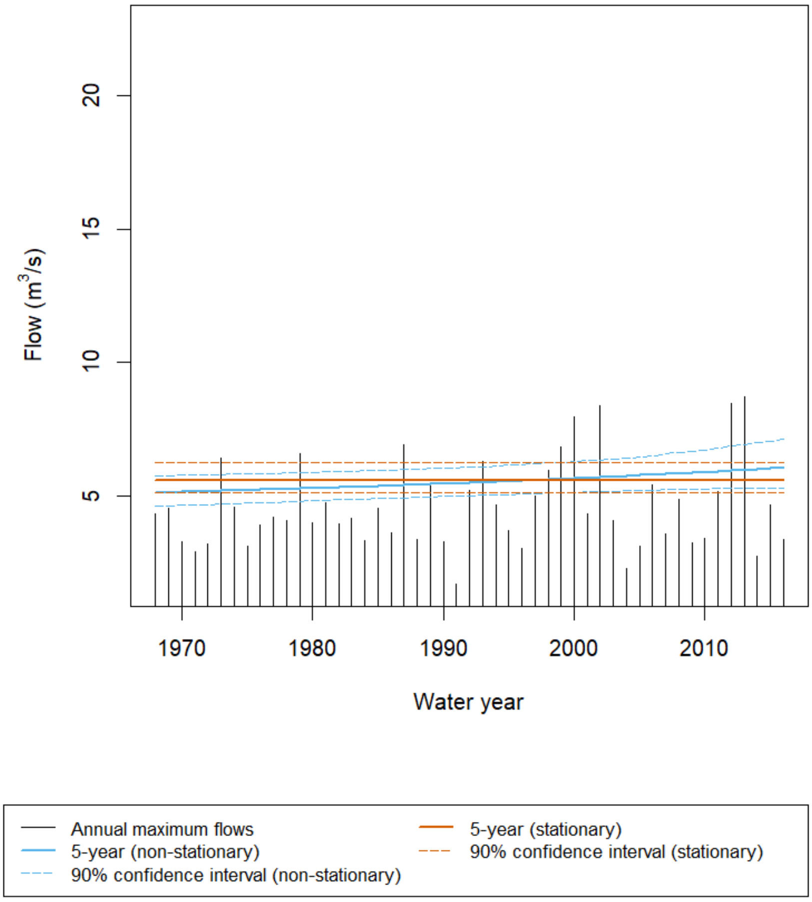
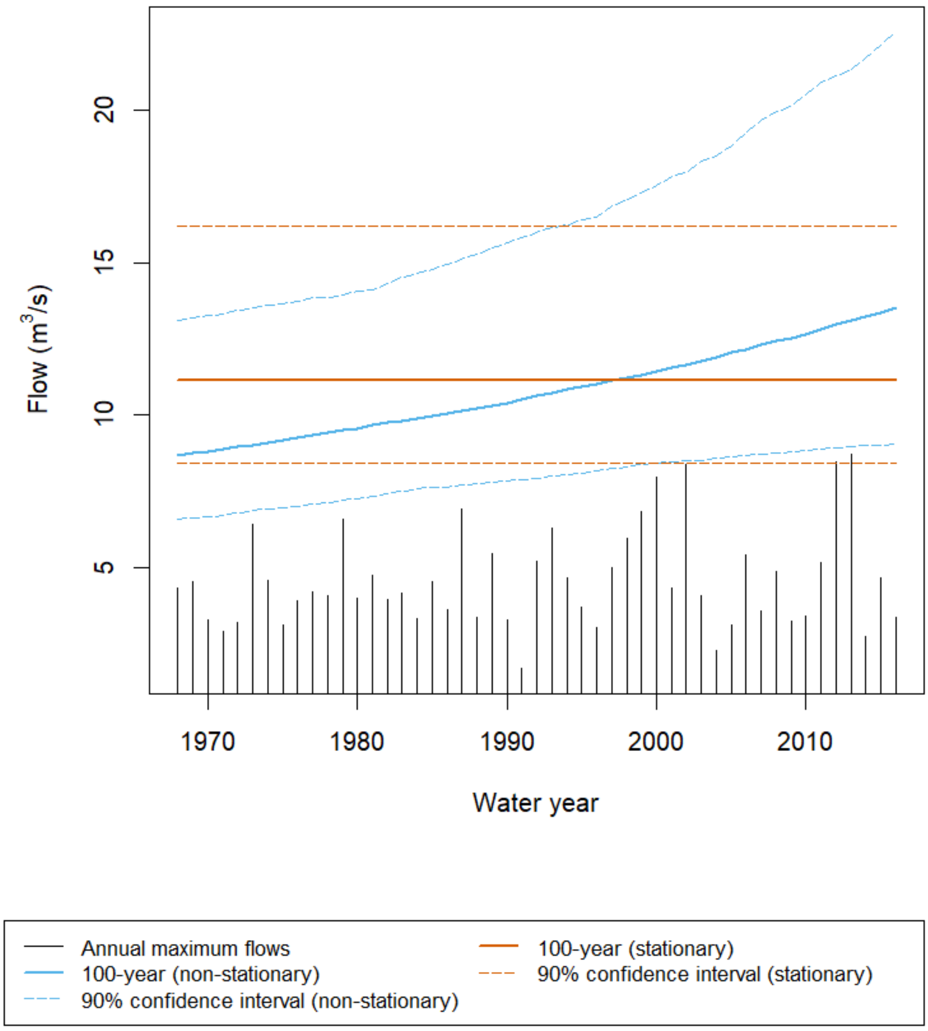


Figure 6-11 Plot of GLO model results and confidence intervals for 100-year return period, comparing stationary and non-stationary (varying scale), Kennal at Ponsanooth (station 48007)[5]



Return periods of observed floods

The largest measured flood in the 49 years of record occurred on Christmas Eve 2013, with a peak flow of 8.7 m^3/s. This was only marginally higher than a flood the previous year, which peaked at 8.5 m^3/s.

For the year 2013, the non-stationary model gives a conditional flow estimate for a return period of 20 years of 8.8 m^3/s. So, the model results imply that both the 2012 and 2013 floods had a return period of about 20 years. The stationary model gives an estimated return period of a little over 30 years.

Impacts of climate change

In most practical applications of flood frequency analysis for the planning and assessment of flood schemes, it will be necessary to consider future conditions,

[5] The annotations, all flows in m^3/s, are to help explain the calculation of climate change allowances.

therefore allowing for the impacts of climate change. Refer to the guidance in section 5.7, which introduces 2 approaches for a sensitivity analysis of climate change impacts: adjusting a present-day flood estimate and adjusting a 1961 to 1990 baseline estimate.

Let us say that an estimate is needed for the 2050s epoch, using the central allowances. For the south-west river basin, the change factor is 20% relative to a 1961 to 1990 baseline.

Option 1 would be to increase the present-day estimate by 20%, giving a 100-year estimate of 16.2 m^3/s for the 2050s.

Alternatively, taking option 2 (see section 5.7):

a) From the non-stationary model, at the mid-point of the 1961 to 1990 baseline period the 100-year flow is 9.2 m^3/s. The data does not extend as far back as 1961 so this estimate is only approximately representative of the baseline period.

b) This can be compared with a stationary estimate, representative of the whole period of record, of 11.1 m^3/s.

c) The 2 estimates are not significantly different, as 9.2 m^3/s falls within the 90% confidence interval for the stationary estimate. However, there is a practical difference between the estimates, which could lead to different findings from a flood study.

d) One possible explanation for the difference is the impact of climate change, the effect of which could be expected to be more pronounced when looking at the whole period of record (1968 to 2016) than for the early portion of the record in the mid-1970s. Other hypotheses could explain the difference. Urbanisation seems unlikely, since the catchment remains largely rural. Natural climatic variation is a reasonable hypothesis though, and its effects are difficult to separate from progressive climate change. To continue illustrating this approach, we will assume that climate change is the reason for the difference.

e) We then adopt the non-stationary estimate of 9.2 m^3/s as a baseline and adjust it using the change factor for the 2050s, 20%. The resulting 100-year estimate for the 2050s is 11.2 m^3/s.

It now becomes apparent that this option 2 approach can be more risky. It has given an estimate for the 2050s (11.2 m^3/s) that is lower than the present-day non-stationary estimate (13.5 m^3/s). In fact, it is almost the same as the result of the stationary model without any adjustment for climate change (11.1 m^3/s).

Even the upper (90th percentile) change factor for the 2050s, which is +40% in south-west England, is less than the +47% increase between the non-stationary model results in 1975 and 2016.

In this case, it is unsafe to assume that the published climate change allowances can be effectively reduced to allow for the change that has already occurred, because the change in the non-stationary model during the period of record is larger than any of the published allowances. Therefore, the outcome would be to infer that there will be no more increase in flood flows due to future climate change. If non-stationary analysis is to be preferred, the option 1 approach to climate change adjustment is more appropriate.

Final comments

For illustrative purposes this case study has presented results from the non-stationary model. On balance, there are several reasons for preferring the stationary model at Ponsanooth:

- The Mann-Kendall test outputs a p value of 36%, indicating that there is little statistical significance to the trend.
- When comparing flood frequency models, the recommended statistical measures (BIC and likelihood ratio) prefer the stationary model.
- The diagnostic plots indicate a better fit for the stationary model.
- The non-stationary model has much a wider confidence interval when estimating present-day flows.
- The non-stationary model implies a large degree of trend over the gauged period, larger than even the most precautionary change factors in current guidance on impacts of climate change.
- A non-stationary model is preferred by statistical measures at all surrounding catchments.

6.3 Little Ouse at Abbey Heath (station 33034)

The main purpose of this case study is to illustrate how to cope with a situation where there is a distinct trend towards decreasing flood magnitude.

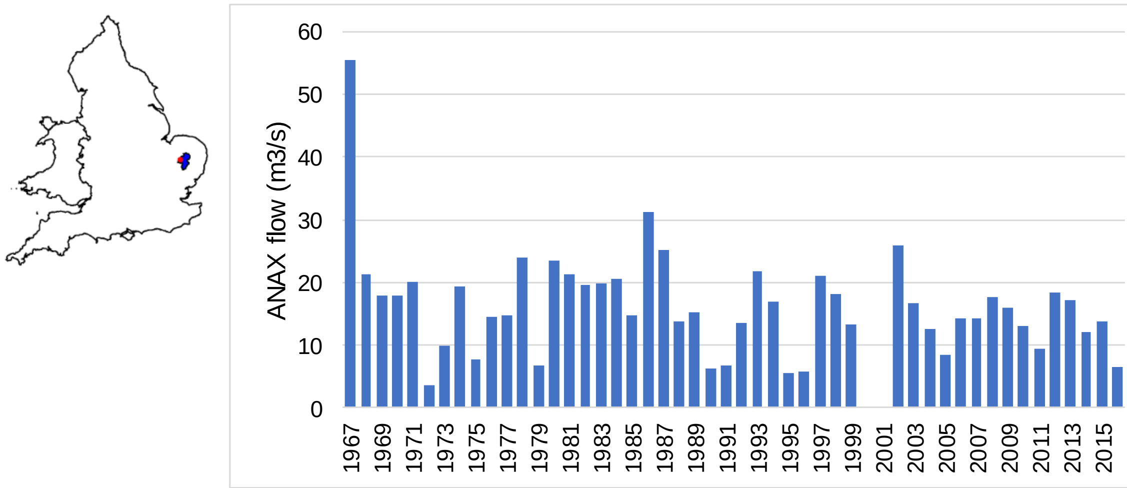
6.3.1 Catchment and data

The Little Ouse at Abbey Heath is a 688 km^2 catchment draining a largely rural catchment in Norfolk with chalk geology, largely overlain with superficial deposits. 63% of the land use in the catchment is arable.

There are 50 annual maximum flows in the data set used for this project, from 1967 to 2016.

The flood of 1968 is a striking outlier and its magnitude is uncertain in light of non-modular flow and bypassing of the gauge.

Figure 6-12 Catchment location and flood peak time series, Little Ouse at Abbey Heath (station 33034)



6.3.2 Trends

Non-parametric test

The Mann-Kendall test shows a decreasing trend. Over the whole period of record to date, the Z score is -1.964, and the corresponding p-value is 0.049. In other words, there is only a 4.9% chance of obtaining a trend at least as extreme as the one in the data set, under a null hypothesis of no trend. This can be interpreted as a statistically significant trend at a 0.05 significance level.

The exceptional flood in September 1968, at the start of the record, will not have contributed much to this finding because of the non-parametric nature of the test (the result would have been identical if the 1968 flood was only marginally bigger than the second highest flood). The more influential feature of the data set regarding trend is the reduction in the frequency of floods exceeding about 20 m3/s over the period of record.

The reason for this reduction is not obvious. Most, but not all, nearby catchments also show a decreasing trend, so the cause may be climatic. However, an upstream gauge on the Thet at Melford Bridge with a longer record back to 1960 shows no trend.

One possibility is that the decrease on the Little Ouse is exacerbated by groundwater abstraction. Another potential explanation is the creation of flood storage in the 1970s due to extraction of sand and gravel, creating the Nunnery Lakes which is now a nature reserve.

Change point tests

Neither the Pettitt nor the PELT test detect any significant step changes.

6.3.3 Non-stationary analysis: time as a covariate

Both the GEV and GLO distributions were considered. Diagnostic plots (not shown here) indicate that the GLO distribution fits marginally better.

The highest quality GEV model, according to both likelihood ratios and BIC, is that which has the scale decreasing with time.

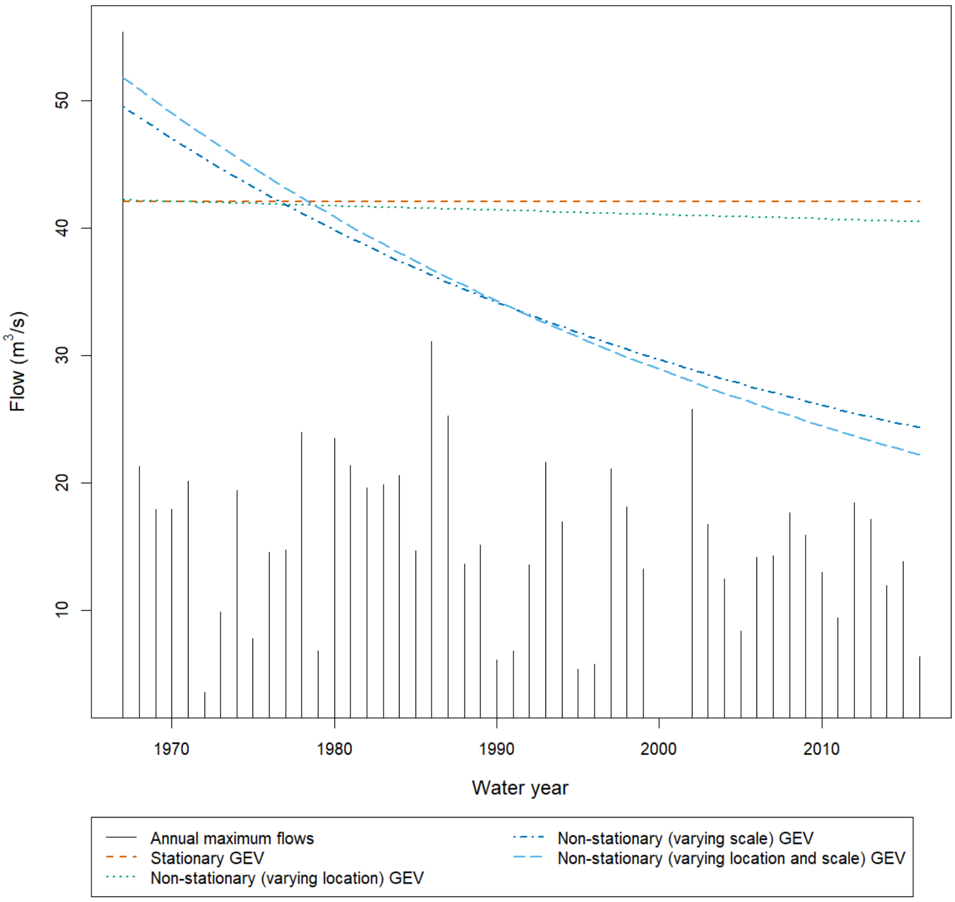
The highest quality GLO model, according to the same measures, is the stationary model.

Diagnostic plots (not shown here) show little difference between the goodness of fit of the various ways of including time as a covariate. All models struggle to fit the extreme flood of 1968.

In a case like this, physical principles need to be an important factor in selecting the preferred model. The fact that some statistical measures point towards the scale parameter decreasing over time (at least for the GEV distribution) does not necessarily mean that the effect is genuine. There are no known trends in catchment land use or climate that might lead to a reduction in the variability of flood magnitude on the Little Ouse. The Nunnery Lakes is a possible explanation, and a hydraulic model could quantify its impact.

A relevant question to ask is whether it is possible for a flood like that of September 1968 to occur nowadays. Few, if any, hydrologists would be willing to rule this out, or to say that such a flood is much less likely to occur today, which is what is implied by the non-stationary models with decreasing scale parameter. The plot below shows how the varying scale model leads to a halving of the estimated 100-year flood over the period of record.

Figure 6-13 Example model results: GEV distribution with time as covariate, AEP 1%, Little Ouse at Abbey Heath (station 33034)



6.3.4 Sensitivity test

Another useful line of enquiry is to test the sensitivity of the analysis to the outlier in 1968. The flood of September 1968 is notorious across some catchments in south-east England, often appearing as an extreme outlier in the annual maximum series. At some gauges, the peak flow is estimated with great uncertainty from extrapolated rating equations.

If the 1968 flood is removed from the analysis, the results change, so that for both the GLO and GEV distributions the highest-quality model is stationary.

This sensitivity to a single flood, with uncertain magnitude, helps to justify the decision to prefer a stationary analysis for the Little Ouse. The effect of the 1968 flood can be limited by carrying out enhanced single-site analysis using FEH methods.

References

Brown MB, Forsythe AB. 1974 'Robust tests for the equality of variances' Journal of the American Statistical Association, 69(346), 364–367. doi: 10.1080/01621459.1974.10482955

Killick R, Fearnhead P, Eckley IA. 2012 'Optimal detection of changepoints with a linear computational cost' Journal of the American Statistical Association 107(500) 1,590 to 1,598

Sayers P. 2016 'Communicating the chance of a flood: The use and abuse of probability, frequency and return period' Available at: https://www.eci.ox.ac.uk/research/water/forum/Communicating the chance of a flood - the use and abuse of retrun period.pdf

Glossary and abbreviations

AEP	Annual exceedance probability.
AIC	Akaike information criterion. A measure of the quality of a statistical model, which establishes a trade-off between the goodness of fit and the simplicity of the model.
AMAX	Annual maximum (for example, the highest river flow in a water year).
BIC	Bayesian information criterion. A measure of the quality of a statistical model, which establishes a trade-off between the goodness of fit and simplicity. Compared with the AIC, the BIC gives more preference to simpler models.
Conditional estimate	Expected value of a variable (such as flow) conditional on covariates being at some specified value. For instance, it might be conditional on the year being 2020, or the total winter rainfall being 300 mm.
Covariate	An observed variable used in a statistical model to help predict the main variable of interest.
EA	East Atlantic pattern: an index of atmospheric variability, like a southwards shifted version of the NAO.
Encounter probability	The probability of an event happening (such as a flow rate being exceeded) at least once over a specified period of time. The AEP is the encounter probability of a flood within a 1-year period.
FEH	Flood Estimation Handbook.
GEV	Generalised extreme value: a statistical distribution fitted to extremes such as floods.
GLO	Generalised logistic: another statistical distribution.
Integrated estimate	Expected value of a variable (such as flow) for a particular probability, without any conditionality on covariate values (contrast with conditional estimate, above). The integrated estimate is calculated by averaging the probabilities corresponding to the conditional flow estimates, over a sample or a statistical distribution of covariate values. More formally in statistics it is known as the marginal return level. Refer to section **Error! Reference source not found.** for more explanation.
Mann-Kendall test	A non-parametric method for testing for the presence of consistent trend in a time series.
MLE	Maximum likelihood estimation. A way of fitting a statistical model by maximising something known as the 'likelihood function'.
Model	In this guidance, all the models mentioned are statistical models, that is, mathematical descriptions of a data set.
NAO	North Atlantic Oscillation: an index of the north-south difference in air pressure between the north and central Atlantic Ocean, associated with changes in the direction and strength of the jet stream.
Non-parametric	A type of statistical method which makes no assumption about the statistical distribution of the data.
Non-stationary	A time series is non-stationary if its statistical properties change over the relevant time scale, for example over the period of record.

PELT	Pruned Exact Linear Time: a test for sudden step changes in a time series.
Pettitt test	A statistical test for detecting a sudden change in the average of a time series.
Stationary	A time series is stationary if its statistical properties do not change over the relevant timescale.
Single-year integrated estimate	An integrated estimate (of flow, for instance) obtained by averaging over the sample of observed physical covariate values, but setting the water year covariate to a single value. Refer to section **Error! Reference source not found.**.
WINFAP	Software that implements the FEH statistical method for flood frequency estimation.

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

Then call us on
03708 506 506 (Monday to Friday, 8am to 6pm)

email
enquiries@environment-agency.gov.uk

or visit our website
www.gov.uk/environment-agency

incident hotline **0800 807060** (24 hours)
floodline **0345 988 1188 / 0845 988 1188** (24 hours)

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