Review of flood frequency estimation in groundwater-dominated catchments

Final report

April 2022

www.jbaconsulting.com

JBA Project Manager

Sarah Warren
1 Broughton Park
Old Lane
Broughton
Skipton
BD23 3FD

Revision History

Revision Ref/Date	Amendments	Issued to
Draft / 8 February 2022	None	Clare Waller
Final / 1 April 2022	Minor corrections and improvements	Clare Waller

Contract

This report describes work commissioned by the Environment Agency. The client's representative for the contract was Clare Waller. Duncan Faulkner of JBA Consulting carried out this work.

Prepared by .... Duncan Faulkner MSc DIC MA FCIWEM C.WEM CSci

Head of Hydrology

.... Kirstie Murphy BSc (Hons) MSc

Analyst

Reviewed by .... Maxine Zaidman BSc PhD C.WEM CEnv FCIWEM

Technical Director | Hydrology & Hydrometry

Purpose

This document has been prepared as a Draft Report for the Environment Agency. JBA Consulting accepts no responsibility or liability for any use that is made of this document other than by the Client for the purposes for which it was originally commissioned and prepared.

JBA Consulting has no liability regarding the use of this report except to the Environment Agency.

Copyright

JBA consulting

Executive summary

This scoping study aims to identify the main challenges associated with fluvial flood frequency estimation in catchments where groundwater exerts a major influence on river flow, and to propose work needed to address these challenges.

For the purpose of this study a groundwater-dominated catchment is defined as one in which the river flow is dominated by contributions from groundwater, i.e. baseflow.

The proportion of baseflow may be quite different during everyday flows and flood conditions, but it is most straightforward to identify groundwater-dominated catchments using a statistic derived from daily mean flow data, the baseflow index (BFI). A dominance of baseflow implies a close connection between surface water and groundwater. On ungauged catchments BFI can be estimated from soil characteristics. The original FEH catchment descriptor BFIHOST has been supplemented with an improved version, BFIHOST19. We propose a threshold of BFIHOST or BFIHOST19 = 0.66 above which a catchment is classified as groundwater-dominated.

Most UK catchments above this threshold are in the south and east of England. The majority, and those with the highest BFIHOST, are dominated by outcrops of chalk. Others are found in areas of Jurassic limestones and Triassic sandstones. Although they can behave more similarly to lower-permeability catchments, the study also covers karst catchments, not all of which have a high BFIHOST.

Fluvial floods on these types of catchment can arise from a variety of physical processes. These are classified into flash flooding, baseflow-dominated flooding and karst flooding. Combinations of processes can also occur, particularly on catchments with mixed geology or some urban land cover. The report presents some notorious examples of floods that demonstrate all these mechanisms.

The core of the report discusses some challenges of flood frequency estimation associated with groundwater-dominant catchments. While not all unique to such catchments they tend to be acute and can lead to large uncertainty. The six challenges, which are inter-related, are:

- Difficulty defining the contributing catchment area
- Other changes in processes between smaller and larger floods
- Ephemeral streams
- Annual maximum flows that are not floods
- Non-stationarity
- Long-lasting floods

These are illustrated using case studies on six groundwater-dominated rivers.

One consequence of these challenges is that the methods of the Flood Estimation Handbook (FEH) show some evidence of higher uncertainty in groundwater-dominated catchments. Both the Statistical method and the ReFH2 rainfall-runoff method are affected in similar ways, although ReFH2 also faces the difficulty of predicting hydrograph shapes and volumes on rivers where the flow may depend on a complex combination of antecedent conditions.

The report presents a wide-ranging list of suggestions for improvements to FEH methods and alternative methods. It concludes with some recommendations for both current practice and research.

Contents

1 Introduction 1
1.1 Project scope 1
1.2 Previous studies on flood estimation in permeable catchments 1
1.3 Guide to report contents 2
2 What and where are groundwater-dominated catchments? 3
2.1 Definition 3
2.2 Extent 4
2.3 Other characteristics of groundwater-dominated catchments 8
3 Flood-generating processes and example events 10
3.1 Processes 10
3.2 Sources of information on observed floods 13
3.3 Examples of flash floods 14
3.4 Examples of baseflow-dominated floods 16
3.5 Examples of snowmelt floods 19
3.6 Examples of mixed floods or those with uncertain flow processes 19
3.7 Examples of floods on karst catchments 20
4 Challenges of flood frequency estimation on groundwater-dominated catchments 22
4.1 Difficulty defining the contributing catchment area 22
4.2 Other changes in processes between smaller and larger floods 30
4.3 Ephemeral streams 32
4.4 Annual maximum flows that are not floods 38
4.5 Non-stationarity 46
4.6 Long-lasting floods 50
5 Performance of FEH methods on groundwater-dominated catchments 51
5.1 FEH Statistical method 51
5.2 ReFH2 rainfall-runoff method 56
5.3 Concluding comments on performance of FEH methods on groundwater-dominated catchments 59
6 Alternative methods for flood frequency estimation in permeable catchments 60
6.1 Introduction 60
6.2 Alternatives within the FEH family 60
6.3 Other approaches 64
7 Conclusions and recommendations 69
7.1 Conclusions 69
7.2 Recommendations for current practice 69
7.3 Recommendations for research 71

Acronyms

AEP	Annual Exceedance Probability
ALTBAR	Mean catchment altitude (m above sea level)
AMAX	Annual Maximum
BFI	Base Flow Index
BFIHOST	Base Flow Index estimated from soil type
BGS	British Geological Survey
BHS	British Hydrological Society
UKCEH	UK Centre for Ecology and Hydrology
CS	Continuous Simulation
DPSBAR	FEH index of mean drainage path slope
FARL	FEH index of flood attenuation due to reservoirs and lakes
FEH	Flood Estimation Handbook
FPEXT	FEH index describing floodplain extent
GEV	General Extreme Value distribution
GL	General Logistic distribution
HOST	Hydrology of Soil Types
MAFF	Ministry of Agriculture Food and Fisheries (now part of Defra)
NERC	Natural Environment Research Council
NRA	National Rivers Authority
NRFA	National River Flow Archive
PDM	Probability Distributed Model
POT	Peaks Over a Threshold
PROPWET	FEH index of proportion of time that soil is wet
QMED	Median Annual Flood (with AEP 50% or return period 2 years)
R&D	Research and Development
ReFH	Revitalised Flood Hydrograph method
SAAR	Standard Average Annual Rainfall (mm)
SPR	Standard percentage runoff
SPRHOST	Standard percentage runoff estimated from soil type
URBEXT	FEH index of fractional urban extent

1 Introduction

1.1 Project scope

This is the report of a scoping study commissioned by the Environment Agency to help identify potential work needed to improve fluvial flood frequency estimation in catchments where groundwater exerts a major influence on river flow.

The scope of the project includes:

1. Introduction to the problem – definition of permeable catchments, their extent across the UK and what are the problems associated with high flow estimation in them, drawing on specific case studies covering a variety of geological settings across England.
2. Conceptual overview of the flood flow generating mechanisms in permeable catchments:
    a. whether these vary depending on specific geology
    b. impacts of urbanisation
    c. consideration of to what extent permeable catchments can be grouped by their characteristics to reflect typical flood flow generating mechanisms
    d. the role of antecedent conditions.
3. Reference to recent and historic evidence of extreme fluvial floods in permeable catchments, and consideration of whether the historical information could be incorporated into flood estimates for other locations.
4. A discussion of the existing "permeable adjustment" in the Flood Estimation Handbook (FEH) statistical method, its adoption and usage, and its limitations, and international equivalents.
5. A discussion of the performance of the ReFH2 method for permeable catchments.
6. A list and description of potential alternative methods, considering:
    a. methods that could fit within the standard FEH family
    b. non-standard methods.
7. Recommendations on the usage of current available methods and for any further work needed to improve flood estimation in permeable catchments.

The project is a desk-based study collating and reviewing existing readily-available information, to support discussions over current recommendations for flood estimation in permeable catchments, and future research to improve methods.

1.2 Previous studies on flood estimation in permeable catchments

There has long been concern about the ability of standard methods of flood frequency estimation to cope with highly permeable catchments. In the UK context there are several previous investigations that have aimed to make improvements. There is typically a long interval between these studies, perhaps reflective of the long intervals between major floods on such catchments.

In 1977, Harold Potter at the Institute of Hydrology carried out a search for historical floods on chalk catchments in England. The aim was to improve the definition of the flood frequency relationship on such catchments in light of two observations, namely that the year-to-year variability of peak discharges is low on chalk catchments and that exceptionally high peak flows are occasionally reported. The results are summarised in Flood Studies Supplementary Report 4 (FSSR4) (Institute of Hydrology, 1977) and details of the historical events are available in hard copies still held by UKCEH. FSSR4 also included an alternative flood growth curve for chalk streams, with lower growth factors than the standard Flood Studies Report growth curves for the same regions.

In 1996, the Ministry of Agriculture, Fisheries and Food commissioned the Institute of Hydrology to carry out a review of floods and flood frequency estimation in permeable catchments. The review was prompted by the prolonged flooding of Chichester from the River Lavant in January 1994. The aims were to:

- Define the extent of permeable catchments
- Review methods for estimating flood frequency
- Review and update information on flood occurrences
- Recommend further research to improve flood frequency estimation

The present report draws on some material from the project report, Bradford and Faulkner (1997). Some of the findings were incorporated into the Flood Estimation Handbook (FEH) when it was released two years later in 1999.

A follow-on project at CEH investigated flood volumes and durations in permeable catchments (Bradford and Goodsell, 2000). It recommended fitting frequency distributions to annual maximum volumes over a range of durations, derived from daily mean flow data.

1.3 Guide to report contents

We have assumed that readers will have some familiarity with the concepts and terminology of the FEH. There is a list of the many acronyms used on page iv. Readers are also expected to have a basic awareness of geology and hydrogeology, although some technical terms are defined.

Section 2 provides a definition of a groundwater-catchment and illustrates their extent within the UK.

Section 0 discusses the physical processes that lead to fluvial flooding on such catchments and provides some examples of notorious floods.

Section 4 outlines why flood frequency estimation can be difficult or uncertain in groundwater-dominated catchments. It illustrates the challenges in six case studies:

- River Yeo at Cheddar: page 23
- River Ver: page 25
- River Wey at Upwey: page 27
- River Lud at Louth: page 31
- River Lavant at Chichester: page 33
- River Itchen: page 47

Section 5 explores the performance of the FEH methods in groundwater-dominated catchments.

Section 6 provides a long list of possible alternative methods, drawing on ideas from previous research and professional practice, and covering both variants of FEH methods and more radically different alternatives.

Finally, Section 7 draws the findings together with some conclusions and recommendations, both for current practice and future research.

References are listed at the end of the report apart from those that relate to specific flood events or locations, which are provided at the end of each case study in section 4.

2 What and where are groundwater-dominated catchments?

2.1 Definition

Previous work has tended to refer to "permeable" catchments but we do not recommend this terminology because all catchments are permeable to some degree, and understanding of sub-surface flow processes is important throughout hydrology.

For the purpose of this review, a groundwater-dominated catchment is defined as one in which the river flow is dominated by contributions from groundwater, i.e. baseflow.

The proportion of baseflow may be quite different during everyday flows and flood conditions, but it is most straightforward to identify groundwater-dominated catchments using a statistic derived from daily mean flow data, the baseflow index (BFI). It is an estimate of the proportion of the long-term average flow volume in a river that originates from baseflow.

BFI values are available for all gauging stations that are held on the National River Flow Archive (NRFA). The way in which BFI is calculated does not involve any identification of physical processes, instead being based on separating the annual river flow hydrograph into faster and slower components. However, in practice high BFI values are usually found on catchments with extensive outcrops of aquifer, as shown on Figure 2-1. A dominance of baseflow implies a close connection between surface water and groundwater.

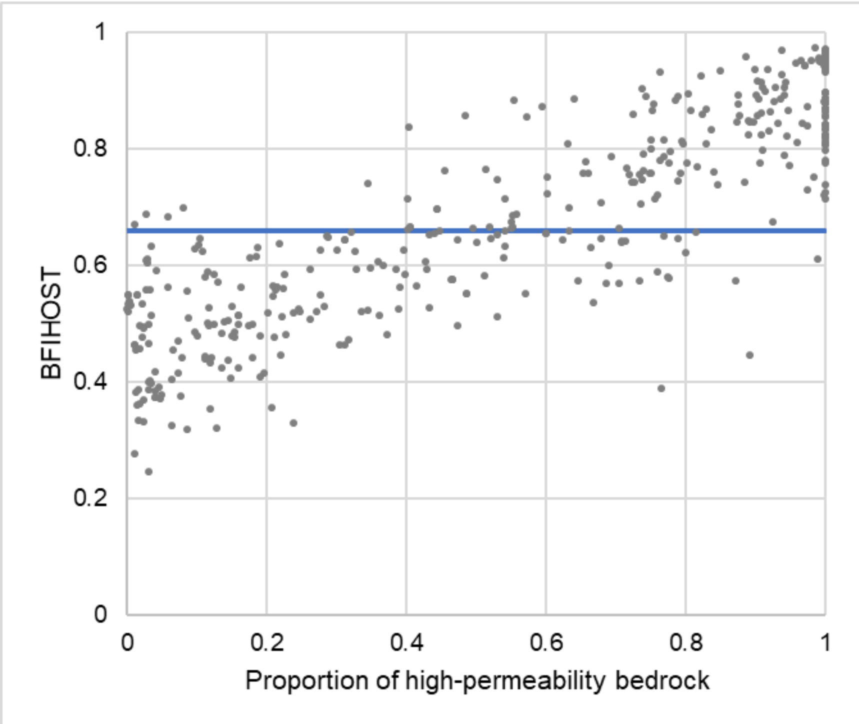


Figure 2-1 Relationship between BFIHOST and proportion of high-permeability bedrock for catchments in the NRFA. Catchments with more than 50% coverage of superficial deposits are excluded. The threshold of BFIHOST=0.66 is marked.

The calculation of BFI assumes stationarity; in reality, the balance of fast and slow flow components may change, for example due to intensification of rainfall or land use change.

On ungauged catchments BFI can be estimated from soil characteristics via the HOST (Hydrology of Soil Types) classification. The original FEH catchment descriptor BFIHOST has recently been supplemented with an improved version, BFIHOST19 (Griffin et al, 2019).

Since the scope of this project includes ungauged catchments, we use BFIHOST rather than BFI to define a threshold.

We propose a threshold of BFIHOST or BFIHOST19 = 0.66 above which a catchment is classified as groundwater-dominated. Within this range, the contribution to runoff from baseflow is at least twice that from more rapid runoff. The threshold is shown on Figure 2-1. Lowering it would start to bring in more catchments that contain little high-permeability bedrock, and increasing it would start to exclude a number of catchments with high permeability geologies.

Further support for choosing this threshold comes from a review of the ability of the FEH Statistical method to estimate QMED, which shows a deterioration above BFIHOST≈0.65 (see Section 5.1).

Because BFIHOST19 values are not yet published for all catchments on the NRFA website, much of the analysis described in this report uses BFIHOST. If BFIHOST19 were to be used in its place, we expect this would result in more catchment being classed as above the threshold of 0.66. HOST classes 8, 10 and 14 all have BFIHOST below 0.66 and BFIHOST19 above 0.66 (Griffin et al, 2019). Taken together, these classes represent about 3.5% of the UK land area. Classes 11 and 16 move in the other direction, falling under the threshold with BFIHOST19, but these classes cover only 0.9% of the land area.

An alternative way to identify highly permeable catchments is to look for catchments that have low rates of runoff. The FEH (Institute of Hydrology, 1999) used a threshold of SPRHOST below 20% to identify so-called permeable catchments. We recommend BFIHOST in preference because (a) BFI can be more reliably estimated from HOST data than SPR and (b) SPRHOST is no longer used in the current versions of the FEH methods.

2.2 Extent

We have mapped the extent of groundwater dominance in two ways: at UK gauging stations that are presented on the National River Flow Archive (NRFA) (Figure 2-2) and at all points on the FEH drainage path network (Figure 2-3).

The NRFA website[1] allows filtering by BFIHOST but not BFIHOST19. 269 out of 1600 (17%) gauged catchments in the UK have BFIHOST>0.66. Figure 2-2 shows the locations of the gauges that exceed the threshold. As expected, most are in the south and east of England. There are two in Wales, four in Scotland and none in Northern Ireland.

One limitation of using gauging stations is that their coverage is uneven between different geologies and landscapes. In highly permeable outcrops there is a lower density of watercourses. Few headwater valleys in the chalk are gauged, not least because they may be dry for long periods. Most gauges on chalk streams are found further downstream where there is some influence from urban runoff, low-permeability superficial deposits or other outcrops, which tends to reduce the BFI. Also, there are few gauges in lowland catchments, which can also be permeable. The proportion of points in the FEH dataset of catchments for England and Wales with BFIHOST>0.66 is 29%, which is higher than the proportion of gauged catchments (22% in England and Wales). This demonstrates that highly permeable catchments are slightly unrepresented in the gauging network.

For both the gauging station sample and all network points, points where BFIHOST>0.66 correspond well with known geologies that give rise to groundwater-dominated river systems. Applying a BFIHOST threshold of 0.66 therefore appears to work well at identifying catchments that are groundwater-dominated and for which it is worth asking whether conventional methods of flood frequency estimation remain applicable.

The maps also show selected groups of moderately to highly permeable solid geological formations that occur in the vicinity of the catchments with BFIHOST over the threshold. All

[1] https://nrfa.ceh.ac.uk/

are made up of sedimentary rocks. The groups shown overlap with most, but not all, of the chosen catchments. Parts of some of the solid formations are covered by superficial deposits, as shown on the map. These may be much less permeable, particularly where they are glacial till.

Most catchments with the highest BFIHOST are dominated by outcrops of chalk, in particular the White Chalk sub-group which has large outcrops in the South Downs, the North Downs, the Chilterns, Norfolk, Lincolnshire and East Yorkshire. Bradford and Faulkner (1997) found that, whether highly permeable catchments are identified using soil characteristics or runoff data, they tend to be dominated by Cretaceous chalk rather than other aquifer formations such as Jurassic limestones or Permo-Triassic sandstones. This may be because rivers flowing over chalk receive more baseflow than those passing over other aquifers. Additionally chalk dominates over large areas, so many catchments can be entirely of chalk. Jurassic limestones and Permo-Triassic sandstones are more localised, and may therefore contribute to the flow regime of part of a catchment but not all of it. The outcrop of the Chalk is larger than that of all the other major UK aquifers added together (Allen et al., 1997).

Chalk aquifers have much lower specific yield[2] than sandstone for instance, and so tend to respond more rapidly to recharge over a single winter season. Another distinctive characteristic of chalk formations is their thickness. The typical depth of the Chalk under outcrops areas in England is 100-400m. Most of the groundwater flow relevant to river floods occurs in the shallow part of the aquifer, as discussed in section 3.1.

Chalk soils fall into HOST class 1 and limestone into HOST class 2, both of which are allocated the maximum possible values for BFIHOST (1.0) and very high values for BFIHOST19 (0.95 and 0.88 respectively) (Griffin et al, 2019).

Another geological formation with large numbers of groundwater-dominated catchments is the Oolitic Limestone, of Jurassic age, which outcrops in the Cotswolds and continues as a ridge running north-east up to the Humber Estuary. Another cluster is seen in Nottinghamshire, on catchments containing a mixture of the Zechstein Group (limestone and dolomite) and a group of Triassic rocks including sandstone and conglomerate. A further cluster occurs on the border of Shropshire and Staffordshire, again with largely Triassic geology.

Figure 2-3 indicates that groundwater-dominated catchments are more widespread in Norfolk and Northern Cambridgeshire than might be expected from the solid geology. These low-lying areas with managed drainage and pumping have extensive superficial deposits of marine alluvium and peat. These soils fall into HOST class 9, which is also found in major river valleys throughout the UK (Boorman et al., 1995). HOST Class 9 is characterised by mineral soils with shallow groundwater and a gleyed layer within 0.4m depth. For this class, BFIHOST is assigned a value of 0.73, and BFIHOST19 0.68. The shallow groundwater in these areas is expected to result in interactions with surface water. It could be that these relatively high values are less well representative of the Fenland soils, because the estimation of BFIHOST coefficients may have been affected by the shortage of gauged catchments in the Fens. The drainage of soils will add another complicating factor. It could be that these catchments combine high storage in dry conditions (and therefore high baseflow) with flashy characteristics in wet conditions. Although there are few, if any, flow gauges measuring runoff from catchments in this area, comparison of pumped volumes against rainfall volumes can indicate higher runoff rates than might be expected from the very high BFIHOST values, because the peat is underlain by less permeable bedrock and there is no gravity drainage.

---

[2] A measure of how much water the aquifer can store. Specific yield is the volume of water released from storage in an unconfined aquifer per unit surface area per unit drop in the water table.

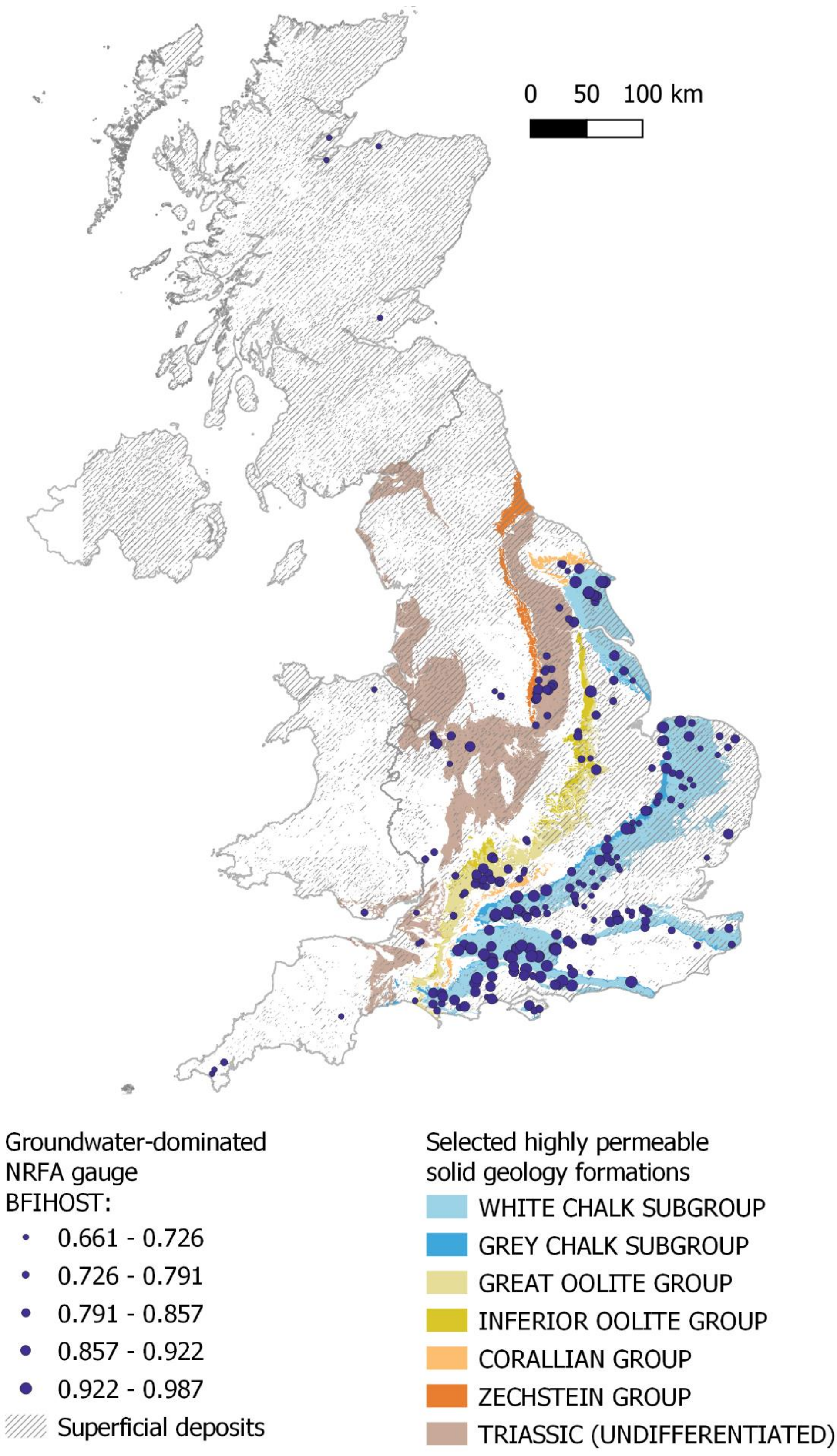


Figure 2-2 Locations of all gauging stations in the NRFA on catchments with BFIHOST>0.66. Contains British Geological Survey materials © UKRI 2022.

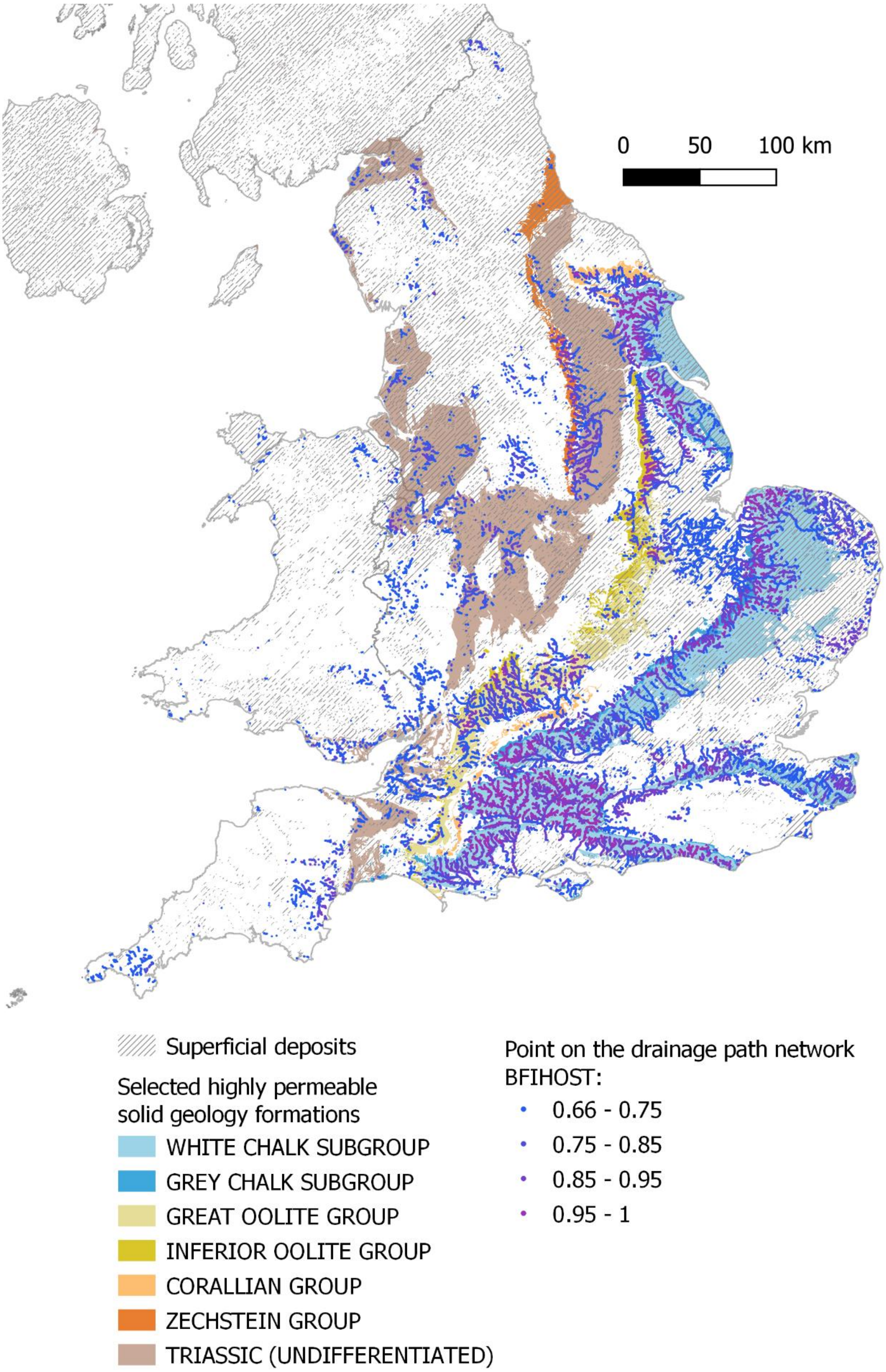


Figure 2-3 Points on the FEH drainage path network in England and Wales with BFIHOST >0.66. Data from the FEH CD-ROM, v1 © NERC (CEH) 1999. © Crown copyright. All rights reserved. Contains British Geological Survey materials © UKRI 2022.

2.3 Other characteristics of groundwater-dominated catchments

In seeking to understand the characteristics of flood flows on catchments with high BFIHOST, it is important to be aware that BFIHOST is correlated with some other physical properties of catchments in the UK. For instance, very large catchments are unlikely to have high baseflow because they will contain a mixture of geology. The fact that the chalk outcrops in the south and east of England means that groundwater-dominated catchments will tend to be lower-lying and to have lower rainfall than some others.

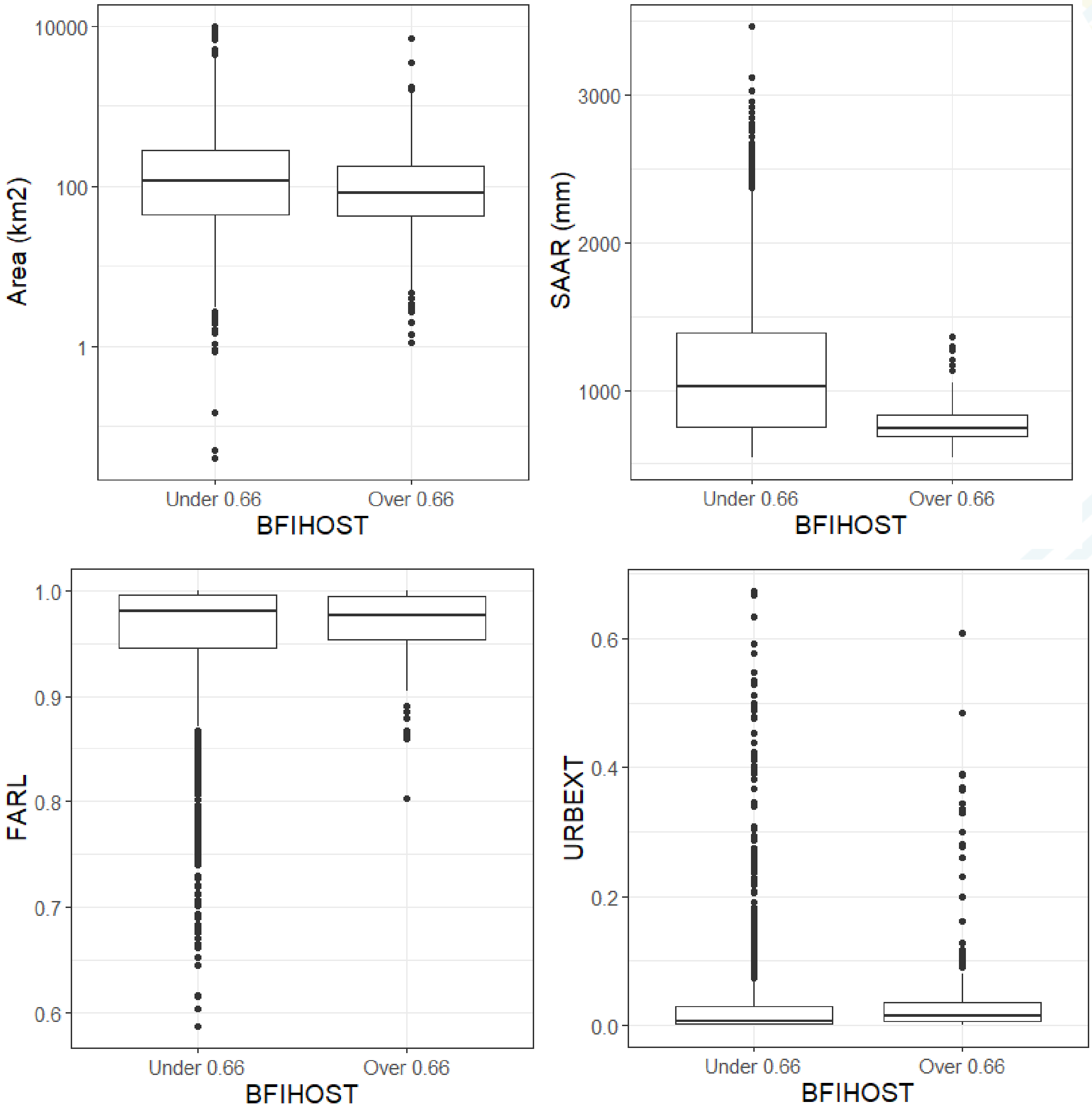


Figure 2-4 Box and whisker plots illustrating distributions of FEH catchment descriptors across the NRFA dataset, distinguishing between catchments under and over the threshold of BFIHOST=0.66

Figure 2-4 quantifies some of these associations. Compared with gauged catchments below the BFIHOST threshold, those above tend to be slightly smaller on average, to have much lower annual average rainfall, SAAR (and a much smaller spread of SAAR values), little influence of lakes and reservoirs (FARL) and a similar degree of urbanisation. This last point is worth noting: although most groundwater-dominated catchments are largely rural, some are heavily urbanised. Even a modest amount of urban development could have a large impact on flood runoff processes in areas of highly permeable soils. The location of the urban area in the catchment can be an important factor to consider.

3 Flood-generating processes and example events

3.1 Processes

3.1.1 Overview

This study concerns fluvial flooding on groundwater-dominated catchments rather than direct flooding from groundwater. The two phenomena are not always easy to distinguish. Fluvial flooding results from significantly enhanced rates of baseflow discharge into the river network. Groundwater flooding is emergence of groundwater in locations distinct from the river network, but which may drain into the river network by overland flow routes. The crux is the definition of the river network, and whether this should include ephemeral watercourses in valleys that are dry for most of the year, or even for longer periods.

The discussion in this section is based on Bradford and Faulkner (1997), augmented and updated, with additional references cited. A key reference is the Aquifer Properties Manual (Allen et al., 1997).

Floods on groundwater-dominated catchments can be classified as follows.

	Type	Typical season	Characteristics
1	Flash floods	Summer	Rainfall intensity exceeds infiltration capacity leading to overland flow. Limited groundwater component. Can occur on steep scarp slopes or in urban areas leading to high flow velocity, debris transport and danger to life.
2	Baseflow-dominated floods	Late winter or spring	Prolonged floods which tend to occur following high recharge during autumn and winter. Can lead to disruption of transport networks for weeks or months.
(a)	High water table		Mainly in headwater valleys with ephemeral streams (winterbournes).
(b)	High water table with groundwater surge		Accelerated baseflow due to rapid movement of groundwater e.g. through shallow fissured zones in chalk aquifers.
(c)	As above plus quick runoff		Above in combination with intense rainfall and/or rapid snowmelt, potentially exacerbated by frozen soils.
3	Karst flooding	Any	More similar to floods on less permeable catchments but catchment boundaries may be very different from topographic boundaries and may vary with event size.

Sections 3.1.2 to 3.1.6 expand on the classification outlined above. Section 3.3 provides examples of floods for each of the flow processes in the classification.

3.1.2 Flash floods

Flash floods (Type 1) are controlled by processes at or near to the ground surface rather than by the underlying geology. Infiltration capacity is normally very high in soils that overlie highly permeable substrates, due to presence of networks of fractures. It can be reduced when the soil surface is compacted, extremely dry or frozen. In such conditions, the intensity of extreme rainfall is more likely to exceed the infiltration capacity. Refer to section 1.2.1 of Bradford and Faulkner (1997) for a discussion of water movement through the unsaturated zone of Chalk soils.

In urban areas the infiltration capacity is much lower and so rapid runoff can easily be generated, potentially exacerbating flash flooding. Urban development in areas with highly permeable soils can lead to a major increase in peak flows unless it is accompanied by suitable mitigation measures.

Case study 4 on page 31 discusses flood frequency estimation on a chalk catchment with evidence of occasional flash floods.

3.1.3 Baseflow-dominated floods

Baseflow-dominated floods (Type 2) are controlled by subsurface conditions. They are most likely to occur in the period January to March, when groundwater levels are at their seasonal maximum.

River flows in groundwater-dominated catchments are largely governed by the position of the water table in relation to the ground surface and shallow zones of higher transmissivity[3] and specific yield. These zones have developed in limestone and chalk aquifers within the range of water table fluctuation (including in periglacial periods when conditions were wetter) due to enlargement of fissures by dissolution. Such aquifers can be described as dual-porosity systems, although in practice both the flow and the storage of water are dominated by the fissure system rather than by the rock matrix. When the water table is unusually high, the transmissivity can increase dramatically, leading to a sudden increase in baseflow. Refer to case study 5 on the River Lavant, page 33. The process has been likened to an overflow pipe within a cistern (Allen et al., 1997).

This dual porosity feature provides a convincing hypothesis to explain an important characteristic of fluvial flood frequency in some groundwater-dominated catchments, which is an unusually high variability in annual maximum flows. Chapter 4 discusses the consequences of this feature for methods of flood frequency estimation.

These upper zones of high transmissivity are particularly well developed in chalk aquifers along valley axes, particularly where there are major structural features in the formation such as synclines, anticlines and faults, and near the contact with overlying less permeable formations. Structural complexity is greater in the chalk of the southernmost counties of England, i.e. the Hampshire Basin and South Downs because it has been affected by Alpine tectonics (Allen et al., 1997). This may explain why there is some (limited) evidence of greater variability of annual maximum flows on some chalk streams in southern counties compared with those on groundwater-dominated catchments further north such as the North Downs, Chilterns or Yorkshire Wolds. Figure 3-1 indicates that most catchments that show a high coefficient of variation (L-CV) of the annual maximum series are in southern counties. This relationship is not seen for the L-skewness.

---

[3] A measure of how much water an aquifer can transmit: the product of hydraulic conductivity and saturated aquifer thickness.

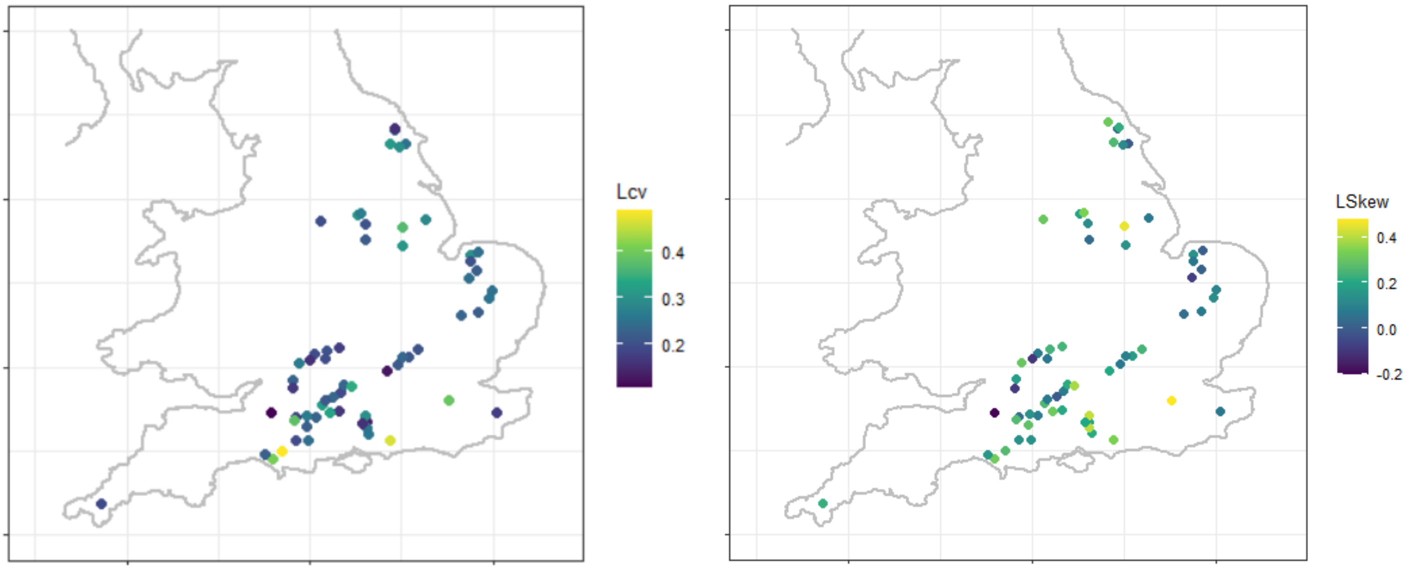


Figure 3-1: Maps of (a) L-CV (coefficient of variation) and (b) L-skewness of annual maximum floods for all UK gauging stations classed as suitable for pooling with BFIHOST>0.66 and at least 20 annual maximum flows. Data source: NRFA peak flow dataset, v10.

Once the water table reaches the surface, which is most likely to happen along valley bottoms, any subsequent rainfall over that part of the catchment will generate a quick runoff response. This will be superimposed on the baseflow component of river flow, although it may be difficult to distinguish it from the accelerated baseflow discussed above or from runoff from less permeable parts of the catchment. Quick runoff can be observed even in rural catchments consisting entirely of highly permeable formations with no drift cover, as shown in case study 3 on page 27.

Quick runoff in conjunction with seasonally high water tables may also occur in areas where soils are frozen, where it may be augmented by snowmelt.

Bradford and Goodsell (2000) proposed that major winter floods on chalk catchments may have three components:

1. Groundwater derived from the main, deeper part of the aquifer.
2. Groundwater from the shallow, high-transmissivity part of the aquifer.
3. Direct runoff, which will depend on the area of the valley with the water table at or close to ground level.

3.1.4 Mixed flood-generating processes

Catchments of mixed geology or with significant urbanisation can show a mixture of flood mechanisms, with a rapid response superimposed on the baseflow component of the hydrograph. There are several examples in the Chiltern Hills, where large urban areas such as High Wycombe, Luton and Hemel Hempstead have been developed on chalk geology. Hydrographs on rivers such as the Wye, the Gade and the Upper Lee show this type of mixed response.

A mixed flood response can also be seen even on some rural catchments containing only highly permeable geology, as seen in the case study on the River Wey (page 27).

3.1.5 Karst

A rather different type of groundwater flow process that can strongly influence river flows is karst behaviour (type 3). Karst catchments in the UK tend to occur mainly in Carboniferous limestone although there are also important karst features in the chalk, Jurassic limestone and in other soluble rocks (Allen et al., 1997; Maurice et al., 2006). Karst catchments do not necessarily show a dominance of baseflow and so may not be identified using a threshold based on BFI or BFIHOST. This is partly because flow through karst systems tends to be much more rapid than through other aquifers.

Some karst catchments are identified using the proposed BFIHOST threshold, such as the Axe at Wookey (station 52001) which drains a partly karst area of the Mendip Hills and has a BFIHOST of 0.69. Many areas of karst in the UK are closely intermingled with outcrops of low-permeability formations and drift cover, such as in the Yorkshire Dales, and gauged catchments in these areas tend to have low average values of BFI and BFIHOST. Differences between the topographic and hydrological catchment boundaries can make BFIHOST values misrepresentative. Despite this, karst catchments have been included in this report for completeness.

Case study 1 on page 23 is an example of a karst catchment at Cheddar.

3.1.6 Additional research on flood mechanisms on groundwater-dominated catchments

Several years after the work by Bradford and Faulkner (1997), the NERC LOCAR research programme (LOwland CAtchment Research) got under way. It included an intensive field monitoring campaign in two neighbouring chalk catchments: the Rivers Pang and Lambourn in Berkshire. One of the many published outputs was a paper by Griffiths et al. (2006) on streamflow generation in the Pang and Lambourn catchments. It found that:

- Interaction between surface water and groundwater is predominantly controlled by the position of the surface water body relative to the groundwater flow system, which varies seasonally.
- The position of the groundwater flow system is dependent on the hydrogeological characteristics of soil and substrate materials, and the influence of surrounding topography.
- Groundwater flow systems generally consist of multiple systems of different orders of magnitude, often occurring within a nested hierarchical order.
- Local topographic effects and variability in hydrological characteristics of riverbed sediments, riparian soils and underlying geology can affect interactions with groundwater.

Other publications on the same catchments examined the spatial distribution of groundwater flooding during the winter 2000-2001 (Finch et al., 2004) and controls on the discharge of chalk streams (Bradford, 2002). Their findings were broadly similar to those of Bradford and Faulkner (1997), emphasising the importance of a shallow layer of solution-enlarged fractures in the chalk, with high permeabilities, along the main streams and dry valleys or near the contact with overlying, less permeable formations. Both point and diffuse discharges occur from this layer.

3.2 Sources of information on observed floods

There are papers, publications and databases that provide comprehensive lists of historic and contemporary flood events, some of which deal specifically with flood events in groundwater-dominated catchments. These include:

- Flood Studies Supplementary Report 4 (Institute of Hydrology, 1977)
- Bradford and Faulkner (1997)

- Chronologies of flash floods compiled by David Archer[4] (Archer et al, 2019)
- The British Hydrological Society's Chronology of British Hydrological Events[5]

In addition, specific flood events are widely described and reported on by the Environment Agency and other regulators, local authorities and local action groups.

Some examples of different types of fluvial floods in groundwater-dominated catchments (selected from the above sources) are cited in the sections that follow. Included with the examples are flood events inferred by examination of the NRFA peak flows dataset[6], identifying events with the highest peak flow as a ratio to QMED on groundwater-dominated catchments.

3.3 Examples of flash floods

3.3.1 Langtoft (Yorkshire Wolds), 1892

Date: 3 July 1892.

Geology: Chalk (unconfined).

Cause: Torrential rain (possible waterspout?). Intense rainstorm on warm, humid day produced flash flood in two streams in Briggate valley, which join just above village. Considerable debris scoured from hills.

Locations affected: Langtoft. 65 houses damaged, loss of crops, damage to farm buildings.

Comments: Similar event occurred on 10/4/1657.

References:

- Climatological Observers Link No 311, March 1996

3.3.2 Louth, 1920

River: Lud, flows east on margins of Lincolnshire Wolds. Catchment area 52 km² to the gauge at Louth.

Date: 29 May 1920.

Geology: Chalk (some boulder clay in valleys).

Cause: Cyclonic rainfall associate with warm, moist SW air-stream meeting cool, moist SE air-stream led to thunderstorms for about 3 hours. Average total rainfall was estimated as 98mm, but reports of 116mm at Elkington Hall and 153mm at Hallington.

Locations affected: Louth, 24 drowned, 46 houses destroyed and 173 others made uninhabitable. 1250 people made homeless. £100k damage. Major soil erosion caused by 18-30 m wide torrents in otherwise normally dry valleys.

Comments: One of the most disastrous floods in 20th century in England. Lud rose 5 m in 15 minutes, reaching a peak flow of 152m³/s in about 45 minutes at Louth. Flood wave estimated to be 60m wide. See also case study on page 31.

References:

- Newnham,E.V. 1921. Report on the thunderstorm which caused disastrous floods at Louth on May 29th, 1920. Professional Notes No 17, Met. Off., London.
- Crosthwaite, P.M. 1921. The Louth Flood of 1920. Trans. Inst. Water Eng., 26.

---

4 https://www.jbatrust.org/how-we-help/publications-resources/rivers-and-coasts/uk-chronology-of-flash-floods-1/

5 https://www.cbhe.hydrology.org.uk/

6 https://nrfa.ceh.ac.uk/peak-flow-data

- Anon. 1921. The Disaster at Louth. Meteorological Magazine, 55.
- British rainfall 1920. HMSO.
- Rodda, J.C., Downing.R.A., & Law, F.M. 1976. Systematic Hydrology. Newsnes – Butterworth, London.
- Latter, P.R. 1932. The Louth flood. The Lincolnshire Magazine, 1, No 2.
- Robinson, D.N.1995. The Louth Flood. Louth Naturalists', Antiquarian and Literary Society.
- Clark, C. and Arellano, A.L.V. (2004) The Louth storm and flood after 80 years. Weather, 59: 71-76. https://doi.org/10.1256/wea.25.03

3.3.3 Martinstown, 1955

River: Wey, Dorset. The event pre-dates river flow measurements.

Date: 18 July 1955.

Geology: mainly Chalk, but also Oxford Clay.

Cause: Intense, local and prolonged storm. Dry summer; air temperature on 17 July about 29°C. Cold air from a shallow depression moving from NW France initiated thunderstorms that became almost stationary over the south Dorset hills between Bridport, Dorchester and Weymouth. The initial storms continued for 4 hours with intensities of up to 50 mm/h. This rainfall was 279 mm (unofficial measurement 355mm at Hardy Monument).

Locations affected: Weymouth, Upwey, Broadwey and Coryates. Shallow flooding along Winterbourne valley. Osmington Mill stream rose by 3.6m. Severe erosion of coastal cliffs occurred and properties were flooded.

Comments: Largest recorded British daily rainfall until the record was broken in 2009. This is a partly permeable catchment. The heaviest rainfall occurred on the Oxford Clay area rather than the Chalk, but groundwater levels in Chalk rose 12m at Upwey over the following week. See also case study on page 27.

References:

- Bleasdale, A. 1974. Rainfall and flooding in Dorset, July 18 1955. British Rainfall, 1968. HMSO (1974).

3.3.4 Hawley (Kent) 1968

River: Darent. Catchment area 101km^2 to the gauge at Otford and 187 km^2 to Hawley.

Date: 16 September 1968.

Geology: Chalk.

Cause: Extreme rainfall over 14-15 September 1968 caused severe flooding across southeast England. Over 170mm of rain fell on parts of the Darent catchment in 48 hours, much of it concentrated into short intense bursts. Annual exceedance probabilities (AEPs) for these rainfall totals are lower than 0.05% for the 48-hour duration at some locations.

Locations affected: Flows are estimated to have peaked at 23m^3/s on the Darent at Otford (source: NRFA files). At Hawley the peak flow was originally estimated in the range 30-50m^3/s (NRFA files). It has recently been re-evaluated at about 26m^3/s by simulation from the rainfall data using a tailored PDM rainfall-runoff model that matched the reported peak flow at Otford.

As a ratio of QMED these peak flows are about 6 at Otford and 9 at Hawley. The AEP of the peak at Hawley has been estimated as about 0.3%.

Comments: Most of the Darent catchment is underlain by an outcrop of chalk. When groundwater levels are low, river flow is lost to the chalk through seepage. This tends to happen in the autumn, with losses ceasing over winter as the groundwater level rises.

Despite this seasonality, the September 1968 flood was the highest on record by a long way. Part of the headwaters is made up of less permeable strata (sandstones and mudstones) with some urbanisation around Sevenoaks. In contrast to the chalk, runoff from these areas can be rapid, happening over hours rather than weeks.

References:

- Salter, P. M. and Richards, C.J. (September 1974). A memorable rainfall event over Southern England (Part 1). The Meteorological Magazine. No. 1226, Vol 103.
- JBA Consulting (2019). Darent and Cray hydrological assessment, final report. Report to Environment Agency.

3.3.5 Bagnor (Berkshire) 2007

River: Winterbourne Stream, West Berkshire. Tributary to the River Lambourn. Catchment area 45 km^2 to the gauge at Bagnor.

Date: 20 July 2007.

Geology: Chalk catchment in the Berkshire Downs with extensive drift cover of clay with flints and rural land cover (BFIHOST = 0.76)

Cause: Little information found about flow processes during this flood. The preceding period was unusually wet and so groundwater levels may have been higher than usual for the time of year.

Comments: Peak flow of 2.9m^3/s was over 7 times QMED. Time to peak appears to be less than a day, followed by a sharp fall in flow and then a gradual baseflow recession. The flood occurred at the same time as extensive flooding in many other catchments.

3.4 Examples of baseflow-dominated floods

3.4.1 Lewes, 1974

River: Winterbourne, tributary of Ouse.

Date: 21-26 November 1974.

Geology: Chalk.

Causes: High groundwater levels and exceptional rainfall. September 1974 had 2.7 times the average monthly rainfall. This was followed in October by rainfall 1.25 times the average. Although the early part of November was dry there was an exceptional rainfall from 11-22 November when 182mm rainfall occurred, or 155% of the monthly average over a 12-day period. A rapid rise in water levels occurred at Newmarket IB borehole on 15 October and the Winterbourne started to flow on 19 October. Water levels rose again on 14 November and on 23rd rose by 6m in 24 hours to become artesian until 27 November. By 20 November river flows were increasing rapidly and the peak flow was reached on 23 November.

Locations affected: Lewes (Pinwell Road area).

Comments: This flood was smaller than that in November 1960, which resulted in many properties in the valley being flooded for nearly two weeks. Another flood in December 1960 was similar to the flood in November 1974. Flood prevention measures were implemented after the 1960 flood, which, together with reduced tide levels in the Ouse, reduced the impact of the 1974 flood. Reports exist of floods on the Winterbourne in 1915, 1925, 1935, 1960 (twice).

References:

- H.R.Potter, 1979 (based on a report by Southern Water Authority, November 1975).

3.4.2 Sleaford, 1977

River: Slea. Catchment area 48.4km^2 to the gauge at Leasingham Mill.

Date: February 1977.

Geology: Lincolnshire Limestone.

Cause: High and sustained groundwater discharge. The area received 133% of the average annual rainfall between 1 October 1976 and 16 February 1977, with 50mm occurring from 18-25 February, mainly as short intense periods of rainfall. Several flow peaks occurred between 18-25 February, and the highest occurred on 1 March at Sleaford.

Locations affected: Sleaford. Fields and gardens flooded locally. Sandbag defences and dug diversion channels prevented more serious impacts.

Comments: Preceded by drought of 1976. Analysis of the Cranwell rainfall record indicated an AEP of 3%. Flow hydrograph shows short runoff events superimposed on massive peak of groundwater discharge. The rainfall over 9-11 February of 50mm produced runoff of only 1.5% of rainfall volume, indicating the relative insignificance of direct runoff and the importance of high and sustained groundwater flow, despite the low groundwater levels in the preceding year.

References:

- Taylor,H.R. 1977. The succession of high flows that occurred from 19 February to 3 March, 1977. Anglian Water Authority, Lincs River Division., Report on Flooding No 4.

3.4.3 Canterbury, 1988

River: Nail Bourne, Kent.

Date: February 1988.

Geology: Chalk.

Cause: High groundwater discharge following a sustained period of several months of above-average rainfall which culminated in January 1988 with over 150mm rainfall. Groundwater levels in Chalk were very high.

Locations affected: Bridge Village south-east of Canterbury. Ten properties flooded for several days, until flow diverted, but area affected for 3 weeks.

Comments: Flows remained very high for several months following the event. Previous floods reported in 1940s and 1960s, 1977 and 1993.

References:

- W.S.Atkins, 1996. Nail Bourne Flood Alleviation Scheme. Environment Agency.

3.4.4 Chichester, 1994

River: Lavant. Catchment area 86km^2 to the gauge at Graylingwell.

Date: January 1994

Geology: mainly White Chalk.

Cause: High groundwater levels followed by a sequence of heavy rainfall events. A succession of winter storms over a six-week period produced high groundwater levels (about 24.5m above January average at Chilgrove) and a more rapid catchment response. Flooding was aggravated by the restricted capacity of the culverted channels through Chichester. Over the period 29/9/93 to 10/1/94 the total rainfall at Chilgrove was nearly 600mm and the net rise in groundwater level 40m. Gravel extraction and pit infilling just east of Chichester may have contributed to the extent of flooding.

Locations affected: There was local flooding in the Westhampnett area and some city basements due to high groundwater levels in the underlying alluvium prior to significant river flows in early January. Thereafter, flooding was primarily caused by flows exceeding the capacity of the channel and culverts in the urban area of Chichester and upstream villages. Over-banking caused flooding in the Westhampnett area and this flow, added to flows in the adjacent Pagham Rife, caused flooding at points of constriction downstream.

A flood of similar peak flow occurred in December 2000.

References:

- Holmes,C.G 1994. The West Sussex floods of December 1993 and January 1994.
- Marsh,T. 1994. The dry valleys strike back. BHS Newsletter No 41.
- Posford Duvivier. 1994. River Lavant flood investigation. NRA.
- Midgley,P. & Taylor,S.M. 1995. Chichester, 1994: the impact of man on a groundwater flood. BHS 5th Nat. Hyd. Symp., Edinburgh.

The case study on page 33 explores flood frequency estimation on the Lavant.

3.4.5 Widespread in south-east England, 2014

River: Widespread in south-east England, including Newbury (Berkshire) and Croydon (London)

Date: December 2013- February 2014

Geology: High groundwater levels were recorded in Chalk aquifers throughout southern England. Chalk wells in Dorset and West Sussex recorded rises of over 25m during December. In January, in the southern Chalk, new monthly maxima were recorded at six index boreholes. Between January and February, groundwater emerged at a total of six boreholes, in West Sussex, Wiltshire, Dorset, Berkshire, Surrey and Kent. Levels fell during February, however some boreholes still reached monthly maxima in February. In the Chilterns, water levels peaked from mid-March onwards and levels remained high in the North Downs. Areas of Buckinghamshire, the Berkshire/Oxfordshire border and Hampshire were still impacted by groundwater flooding in April, and flood alerts remained until early June.

Some rivers draining these aquifers recorded their highest flows in periods of records dating back to the 1950s-1960s. The Thames at Kingston, with substantial permeable areas in its catchment, recorded its highest flow since 1974. The Kennet at Theale recorded its highest flow since 1961, and the River Itchen the highest flow since 1958.

Cause: An exceptional number of storm events crossed the UK during the 2013/2014 winter period. The clustering and persistence of the storms was linked to an unusually strong North Atlantic jet stream and six Atlantic depressions tracked further south than usual. This resulted in one of the most exceptional periods for winter rainfall in 248 years in England and Wales. The two-month (December and January) total rainfall for the southeast and central southern England region was 372mm; the wettest of any two-month period since 1910.

The persistent rainfall caused most UK catchments to be saturated by mid-December 2013, and the following seven to eight weeks saw groundwater, fluvial, tidal and pluvial flooding. Flood events were complex, with high water levels in rivers preventing drainage. Groundwater emergence combined with fluvial flooding and drainage failures affected large areas.

Locations affected: Southern and southeast England. The inundation lasted for many weeks and caused extensive disruption to transport and other infrastructure. In areas of Buckinghamshire, Berkshire/Oxfordshire border and Hampshire, impacts of groundwater flooding lasted until mid-July. Due to saturated ground water was unable to drain away and consequentially sewers continued to surcharge and some roads remained submerged in mid-July.

Widespread groundwater-fed flooding also occurred in south-east England during the winter 2000-01, in the aftermath of severe fluvial flooding in Autumn 2000.

References:

- McKenzie, A., 2015. Groundwater flooding research and mapping in the UK. European Geologist European Geologist, p.49.
- Slingo,J., Blecher,S., Scaife,A., McCarthy,M., Saulter,A., McBeath,K., Jenkins,A., Huntingford,C., Marsh,T., Hannaford,J., Parry,S., 2014. The recent storms and floods in the UK. Exeter, UK Met Office, 27 pp
- Morris, S.E., Cobby, D., Zaidman, M. and Fisher, K., 2018. Modelling and mapping groundwater flooding at the ground surface in Chalk catchments. Journal of Flood Risk Management, 11, pp.S251-S268.
- Muchan, K., Lewis, M., Hannaford, J. and Parry, S., 2015. The winter storms of 2013/2014 in the UK: hydrological responses and impacts. Weather, 70(2), pp.55-61.
- Thorne, C., 2014. Geographies of UK flooding in 2013/4. The Geographical Journal, 180(4), pp.297-309.

3.5 Examples of snowmelt floods

3.5.1 Till, 1841

River: Till, Wiltshire (Salisbury Plain). Tributary of River Wylye. Catchment area 72 km^2.

Date: 16 January 1841.

Geology: Chalk.

Cause: Snow melt flood with frontal rain superimposed on high baseflow following a wet autumn in 1840. Severe frost and snow together with thunderstorms for the first 4 days of January and the Till rose to bank level. Cold weather then returned with heavy snow on 9 January, together with glazed ice and heavy rain on 10-11 January. A rapid thaw began on the 16th (or possibly earlier) accompanied by heavy rainfall from an eastward moving warm front.

Some authors (e.g. Cross, 1967) state that frozen ground was a contributory factor to the severity flood. This is disputed by Clark (2004) who examined meteorological data and concluded that the data used by Cross (1967) from Salisbury was unrepresentative and took no account of thaws before 16 January.

Locations affected: Villages of Tilshead, Orcheston, Maddington, Shrewton, and Winterbourne Stoke. 3 drowned, 200 homeless, 72 houses destroyed (weak, local construction materials). £10K damage. Further damage downstream at Salisbury as Avon flooded.

Comments: River was already bank-full by early January, partly due to high groundwater levels from preceding autumn. Flood peak, estimated to be 2.3m above bank-full condition, was reached within 6 hours but returned to bank-full within the next 18 hours.

References:

- Cross, D.A.E. 1967. The great Till flood of 1841. Weather, 22.

3.6 Examples of mixed floods or those with uncertain flow processes

3.6.1 Heighington, 2007

Date: June 2007

River: Heighington Beck. Catchment area 21km^2 to the gauge at Heighington, rural.

Geology: Oolitic limestone catchment in Lincolnshire with minimal superficial deposits (BFIHOST = 0.95).

Cause: Prolonged intense rainfall, following an unusually wet early summer.

Locations affected: No information found. The event was selected on the basis of its remarkably high peak flow rather than from information on impacts.

Comments: Peak flow of $5.3m^3/s$ was over 8 times QMED, and nearly three times as large as the second highest flood on record. Flood took several days to reach its peak. It occurred at the same time as extensive flooding in many other catchments. Little information found about flow processes during this flood.

3.7 Examples of floods on karst catchments

3.7.1 Cheddar, 1968

River: Yeo

Date: 10-11 July 1968

Geology: Carboniferous limestone with karst, most drainage via caves.

Cause: The period leading up to 10 July saw a succession of showers, leading to a reduction of the soil moisture deficit. A northerly-moving depression was orographically enhanced as the air lifted over the Mendip ridge. Around 85% of the rain fell in the four hours between 8pm and midnight. The intensity exceeded the infiltration capacity of the ground in areas of the Mendips, leading to localised overland flow. The capacity of swallets (where streams disappear underground) was overwhelmed, so water flowed down the system of usually dry valleys, and into Cheddar Gorge. At the bottom of the gorge, water initially flowed into the caves but around midnight the flow direction reversed, with water from the caves joining the flow down the gorge. This indicates a rapid response within the karst system. The flood was prolonged: the next morning, flood water was still deep enough to submerge cars.

Locations affected: Cheddar Gorge and village. Road blocked by boulders, dozens of homes and shops flooded. One person drowned.

Comments: This was an exceptionally severe flood. It is described in insightful and comprehensive detail in a report by the Wessex Cave Club. The authors, as well as being cavers, were specialists in meteorology and geomorphology, and one of them (Malcolm Newson) went on to become Professor of Physical Geography at Newcastle University.

The catchment rainfall depth has been estimated as 97mm in 4 hours. According to the FEH13 rainfall depth-duration-frequency model, the corresponding AEP is 0.1%.

A flow of $10.5m^3/s$ was measured emerging from the spring at Cheddar. The discharge down the gorge was estimated as $10m^3/s$. The 1970 report is not quite clear but it seems likely that these two discharges are separate from each other and can be added to give a total of about $20m^3/s$. This is much lower than a subsequent estimate by Black and Veatch (2011). From photographs and anecdotal evidence the flow along the gorge was estimated as approximately 30-$50m^3/s$, in addition to the discharge through the underground river. Taking the latter as $10m^3/s$ as in the 1970 report would give a total discharge of 40-$60m^3/s$, very much higher than the figure in the 1970 report.

References:

- JBA Consulting (2018). Cheddar Flood Modelling Project Hydrology inception report and method statement. Report to Environment Agency.
- Hanwell, J.D. and Newson, M.D. (1970). The Great Storms and Floods of July 1968 on Mendip. Wessex Cave Club.
- Black & Veatch (2011). River Yeo and River Axe System Critical Asset Survey. Report to Environment Agency.

4 Challenges of flood frequency estimation on groundwater-dominated catchments

This chapter identifies some ways in which flood frequency estimation is more difficult or uncertain on groundwater-dominated catchments. Ultimately all the challenges can be thought of as stemming from the same root: the balance between surface and underground flow processes, and how it changes between events. The following sections discuss five manifestations of this phenomenon:

- Difficulty defining the contributing catchment area
- Changes in processes between smaller and larger floods
- Ephemeral streams
- Annual maximum flows that are not floods – a consequence of some of the previous issues
- Long-lasting floods

Each of the challenges is illustrated using one or more case studies. These are deliberately focused and mostly brief, the intention being to provide a real-world illustration of the challenge. Some of the case studies go on to explain how the challenge has been tackled within a real-world project; these should not be taken as comprehensive or authoritative accounts.

The topic of annual maximum flows that are not floods is treated in more detail, as required in the project scope.

Chapter 5 goes on to discuss how well the FEH methods cope with these challenges.

4.1 Difficulty defining the contributing catchment area

On some groundwater-dominated catchments it can be difficult defining which portions of the catchment contribute to river flow during flood conditions. The contributing area may be different from the topographic catchment due to groundwater flow across topographic catchment boundaries (Case studies 1 and 2), and it may vary with rainfall intensity or with antecedent groundwater conditions (Case study 1). This can cause great difficulty when applying FEH methods, or most equivalents, which treat catchment area as a constant.

In some events, even with catchment-wide rainfall, river flow may arise from runoff over only a small portion of the catchment (Case study 3), for example built-up areas, zones of lower permeability or valley bottoms where the groundwater level is high (see 3.1.3). A more complex situation is when these rapid runoff responses are superimposed on a more gradual baseflow hydrograph, leading to questions over joint probability (see 6.3.2).

To improve understanding of contributing areas it can be helpful to:

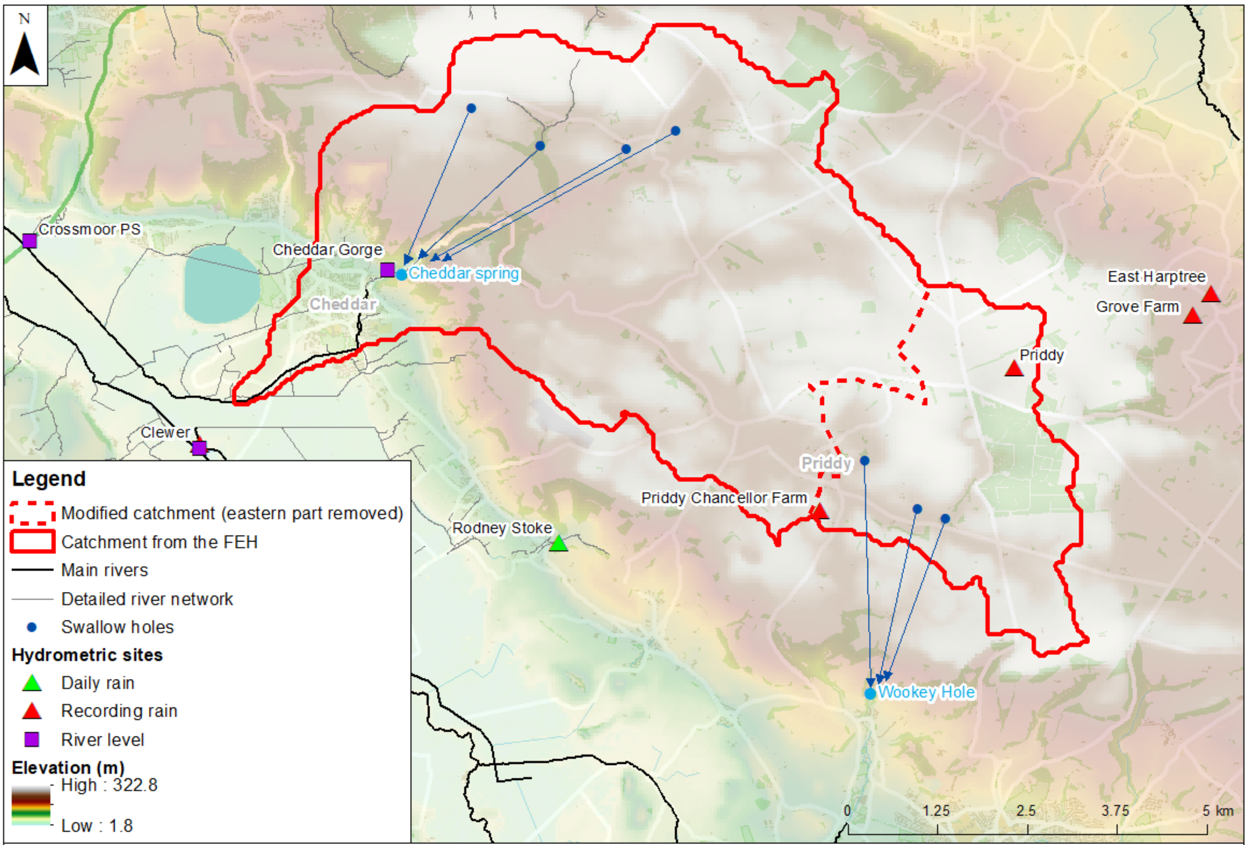
- When applying FEH methods, rely even more than usual on local flow records, which reflect the true contributing catchment area. This would have value even if the quality of high flow measurement, or length of record is less than ideal.
- Compare flow records for neighbouring catchments, for example to determine whether flow magnitudes appear to be broadly consistent (or not) with catchment area
- Develop an understanding of the geology and hydrogeology. This can be helped by reviewing geological maps, hydrogeological maps, geological memoirs, groundwater modelling studies, field visits, literature from caving clubs, project reports and academic publications.

Case study 1: River Yeo at Cheddar, Somerset

Highlights:

- Karst with variable contributing catchment area.
- Threshold changes in flood response

Geology: Carboniferous limestone. BFIHOST: 0.76



In the upper catchment, rainwater sinks into the limestone at swallow holes (“swallets”). It then passes through a deep cave system until the underlying lower-permeability sandstone reaches the surface, leading to a large spring at the base of Cheddar Gorge. One impact of this is that the catchment limits are defined by the groundwater flow direction, rather than the surface topography. The locations of the main swallets are shown in the above mapping. The group of three swallets to the south, in the Priddy area, although they lie within the topographic surface water catchment of Cheddar, have all been proved (for example by dye tracing) to flow south out of the Cheddar catchment, discharging into the River Axe from Wookey Hole. This means that under all but exceptional conditions this eastern portion of the surface water catchment does not contribute to flow at Cheddar.

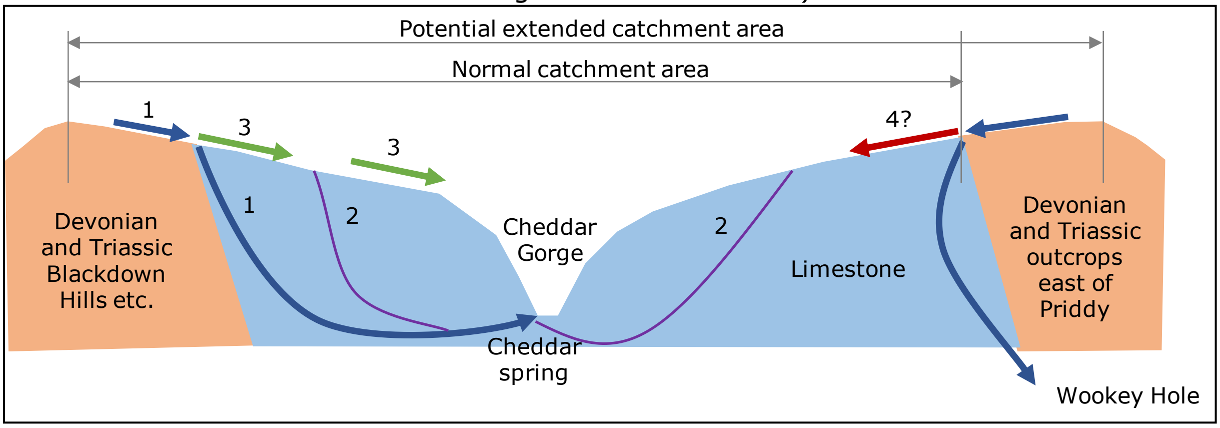
A conceptual model envisages four processes that can contribute to flood flows on the Yeo through Cheddar:

1. Runoff from the lower-permeability outcrops on the crest of the Mendip Hills, which enters the cave systems at swallow holes and flows underground rapidly to emerge at the Cheddar spring.

Case study 1: River Yeo at Cheddar, Somerset

2. Percolation from rain over the limestone outcrop that forms the majority of the catchment area, flowing more slowly via small fractures that eventually feed into the cave system.
3. In exceptional conditions, surface runoff either from overflowing of swallets (as reported in 1968 – see 3.7.1) or from rainfall over the limestone outcrop.
4. In even more exceptional conditions, the process described at no. 3 may extend into the area east of Priddy, which does not normally contribute flow to Cheddar despite lying in its surface water catchment.

Simplified conceptual sketch of processes 1 to 4 that could bring flood water to Cheddar (not to scale and not intended to be a straight-line cross section).



References

- JBA Consulting (2019). Flood estimation report: Cheddar Flood Modelling Project. Report for Environment Agency.
- BGS (2018) Mendip caves and karst. https://www2.bgs.ac.uk/mendips/caveskarst/Caves_1.htm (Accessed 2018)

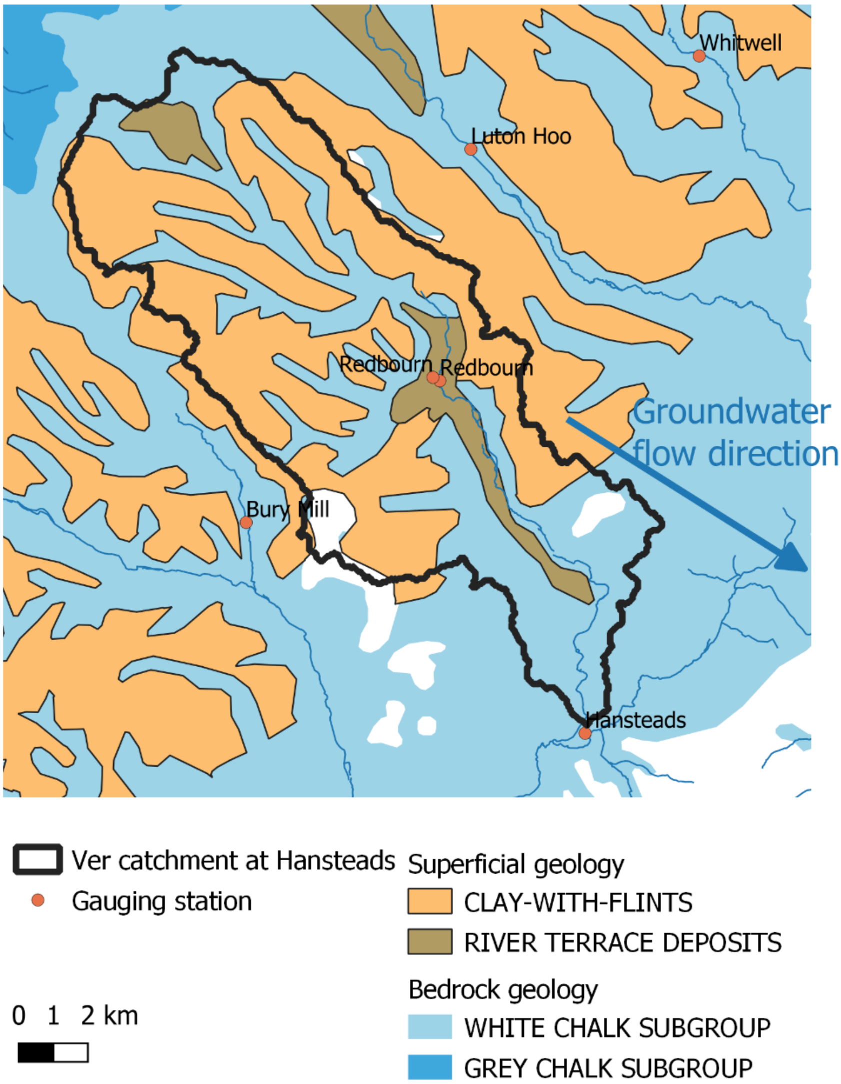
Case study 2: River Ver, Hertfordshire

Highlights:

- Difference between groundwater and topographic catchments
- Extreme error in QMED estimation

Geology: White Chalk sub-group (Seaford and Lewes nodular formations).

BFIHOST: 0.65-0.70



Away from the river valleys, the Chalk in the Ver catchment is overlain by extensive superficial deposits, mainly consisting of clay with flints. River flows are reduced by groundwater abstractions.

The topographic catchment is said to significantly exceed the groundwater catchment. Groundwater contours indicate that the direction of flow in the Chalk aquifer diverges from that of the River Ver from Redbourn downstream, with groundwater flowing south-east, ultimately out of the topographic catchment. This may be influenced by abstraction. Much of the course of the river appears to be perched above the water table, and thus the river may lose water to the Chalk aquifer where hydraulic connectivity permits. In the area surrounding the lower part of the catchment there is evidence of karst features, with tracer tests showing rapid transfer of water between catchments.

Case study 2: River Ver, Hertfordshire

There are gauging stations on the River Ver at Redbourn and, further downstream, at Hansteads, also known as Colney Street. The median annual maximum flow QMED at Redbourn is $0.37m^3/s$ compared with $5.80m^3/s$ estimated from catchment descriptors. At Conley Street QMED is $1.26m^3/s$ compared with $9.69m^3/s$ from catchment descriptors. The catchment descriptor estimates include an urban adjustment, although this is minor at Redbourn). The FEH regression equation for QMED, plus the urban adjustment, overestimates by a factor of over 15 at Redbourn and nearly 8 at Colney Street. Another gauge, also at Redbourn but on the Redbourn tributary, also shows extreme overestimation.

These are remarkably large errors, perhaps larger than anything seen elsewhere in the UK.

Although the Redbourn gauge is not included in the NRFA peak flows dataset, Colney Street is recommended as suitable for pooling. The fact that the FEH regression overestimates at all three gauging stations in the catchment indicates that it is likely to be a genuine feature rather than resulting from any gross errors in rating curves. Potential explanations for this phenomenon include some or all of:

1. A groundwater catchment that is much smaller than the topographic catchment area.

2. A greater degree of connection between surface water and groundwater than indicated by the HOST soil classes. The extensive clay deposits reduce the BFIHOST for the catchment to well below what would be expected for a chalk catchment. The river valley consists of alluvium over chalk and therefore the river can be expected to be in close hydraulic connection with the groundwater, and probably losing water to the aquifer.

3. Artificial reduction in peak flows owing to abstraction lowering the groundwater table.

In summary, there are several potential convincing physical explanations for the extreme over-prediction of QMED. A related concern would be whether the Ver catchment might ever experience conditions that would lead to a flood response more like that seen in impermeable catchments. One possible scenario would be heavy rainfall and/or snowmelt on frozen ground; another might be a downpour of great intensity than the infiltration capacity of the soils, perhaps after hot dry weather which had led to the formation of a surface crust. Such conditions could lead to a flood flow very much larger than any in the observed record.

References

- JBA (2018) Flood estimation report: River Ver. Report to Environment Agency.
- IGS (1984) Hydrogeological Map of the Area Between Cambridge and Maidenhead.
- Shand, P, Tyler-Whittle, R, Besien, T, Lawrence, A R, and Lewis, O H. (2003) Baseline Report Series: 6. The Chalk of the Colne and Lee river catchments. BGS/EA.
- Allen et al. (1997) The physical properties of major aquifers in England and Wales. BGS/EA.

Case study 3: River Wey at Upwey, Dorset

Highlights:

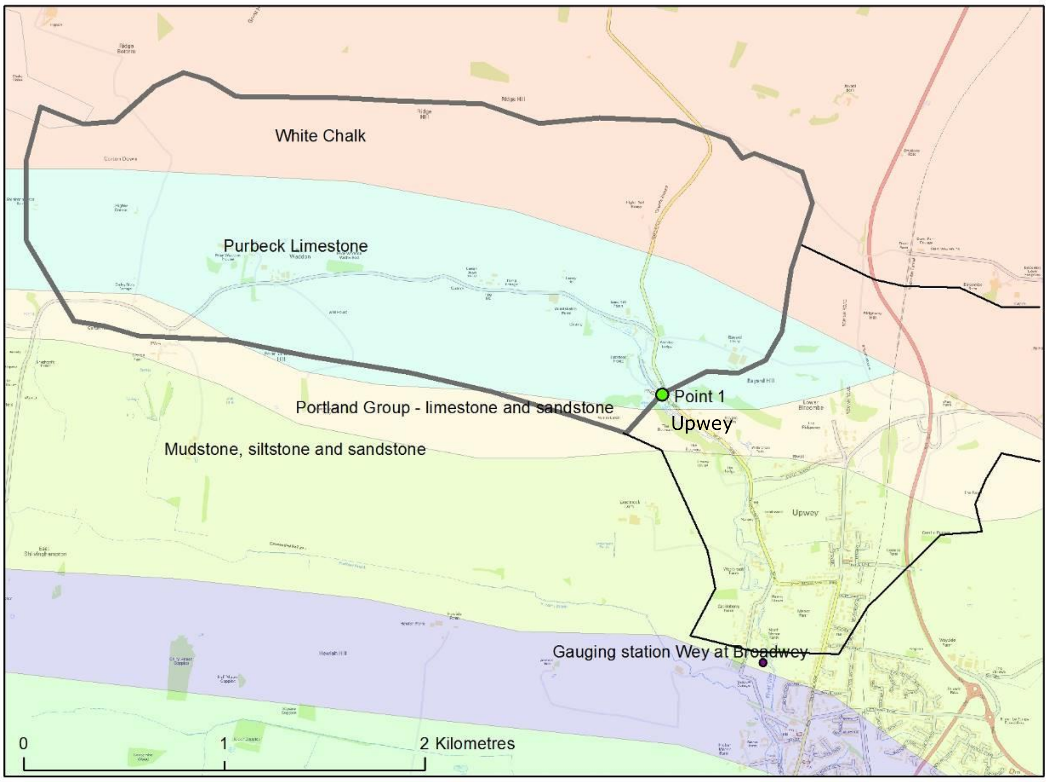
- Runoff arising from part of catchment
- Mixed flood responses

Geology: White Chalk sub-group (Seaford and Newhaven formations) and Purbeck Limestone group. BFIHOST: 0.83

The River Wey at Upwey is a small steep catchment draining the southern flanks of a Chalk ridge, with Purbeck Limestone cropping out in a valley south of a major fault. Further down the catchment the surface geology becomes less permeable, with a narrow band of Portland Group (limestone and mudstone) and then an outcrop of mudstone, siltstone and sandstone.

The source of the River Wey is a spring called the Wishing Well, located at the boundary between the Purbeck Limestone and the Portland Beds, just south of Point 1 on the map. The spring is thought to derive most of its water from the Chalk north of the fault.

Shortly downstream of the spring, the Wey is joined by a tributary stream. This stream flows over the Purbeck Limestone. Its topographic catchment boundary is shown in grey on the map. It seems possible that the catchment area is rather different to that feeding the Wishing Well spring, but no further information on this is readily available. The uncertainty over the hydrological catchment is mitigated by the presence of a gauging station at Broadwey, marked on the map.



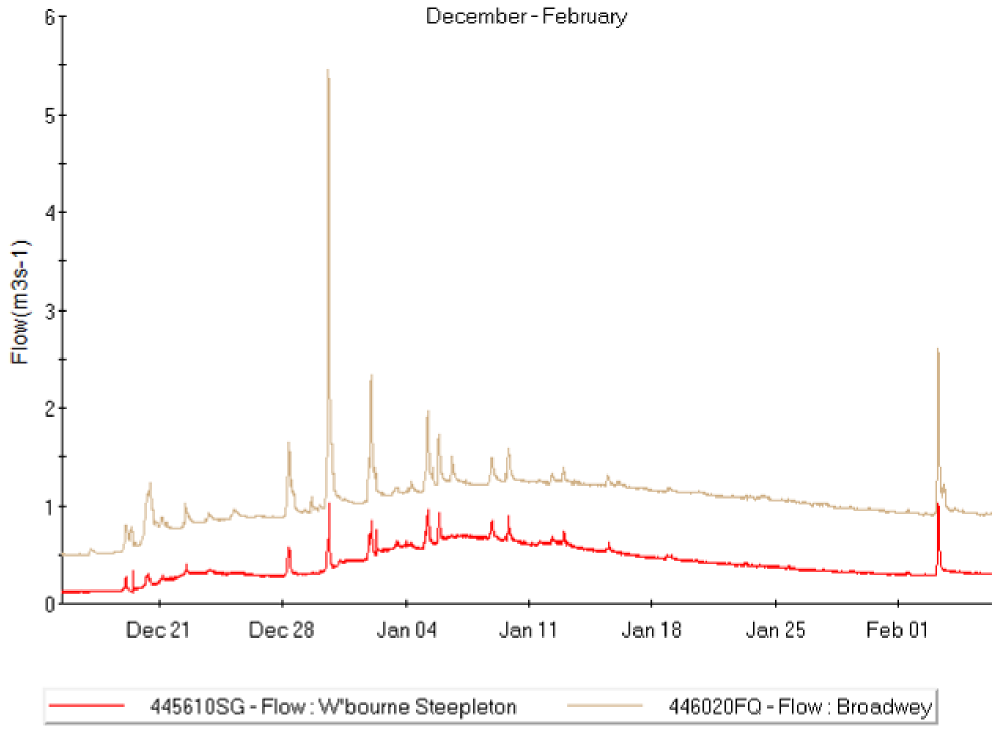
In assessing whether and how to transfer flow statistics (such as QMED) from the Broadwey gauge to the ungauged catchment at Upwey, it is important to consider the change in geology between the two locations.

Case study 3: River Wey at Upwey, Dorset

The peak flows at Broadwey generally consist of a very brief (few hours duration) hydrograph superimposed on a much more gradual (weeks to months) hydrograph, typical of a groundwater flow response. For the highest peaks, the magnitude of the brief "spikes" is 4-5 times that of the underlying groundwater flow hydrograph at the time. An understanding of the nature of the spikes is vital in order to assess flow magnitudes.

The explanation that initially appears most likely is that the brief spikes are due to rapid runoff from a lower-permeability area of the catchment which responds to short-duration intense rainfall. This area may consist mostly of the band of mudstone, siltstone and sandstone that crops out immediately upstream of the gauging station, along with the small built-up area of Upwey village.

An alternative hypothesis is that the spikes could represent runoff from valley bottoms and riparian areas in the chalk and limestone part of the catchment, which start to produce rapid runoff when the groundwater level is high. There is some evidence to support this latter suggestion since the highest peak flows all occur at times of the year when the groundwater-dominated hydrograph is also high, mostly in December or January. There do not appear to be any major peaks at times when the groundwater outflow is low, as might be expected if the spikes are due to runoff from areas of the catchment that are permanently low-permeability. Further evidence in support comes from the comparison with the flow record on the nearby South Winterbourne at Winterbourne Steepleton, a catchment consisting entirely of Chalk formations. The hydrograph at Winterbourne Steepleton shows a similar pattern of brief spikes superimposed on a slowly-responding groundwater hydrograph, although the spikes are less pronounced.



Flow data for two nearby rivers: the Wey at Broadwey and South Winterbourne at Winterbourne Steepleton, December 1993 to February 1994

Case study 3: River Wey at Upwey, Dorset

If it is correct that floods at Broadwey occur mainly due to seasonally enhanced rapid runoff from valley bottoms in the chalk and limestone areas, then it can be assumed that similar flow processes will prevail in the catchment draining to Upwey. This helps to justify transferring the QMED adjustment factor and flood growth curve from Broadwey to Upwey.

This conclusion is strengthened by the observation that there are also similar spikes in the record from a water level gauge in Upwey (not shown).

References

- Arkell, W.J.; Wright, C.W.; Osborne White, H.J (1947) The geology of the country around Weymouth, Swanage, Corfe and Lulworth. Geological Survey of Great Britain, DF341-343A

4.2 Other changes in processes between smaller and larger floods

As we have just discussed, the contributing catchment area may vary between events. Other types of change in runoff processes may also occur between events. For example, areas which usually do not produce rapid runoff may start to do so when the soil surface is compacted, extremely dry or frozen, or when the rainfall intensity is higher than the infiltration capacity. Case study 4 presents an example of this, on the River Lud.

One consequence of the way in which the balance of runoff processes varies between events is that annual maximum floods on groundwater-dominated catchments can show an unusual degree of variability. Unusually high variability is seen on the River Lud and also on the Lavant (case study 5). Figure 4-1 shows that the L-coefficient of variation (L-CV) of annual maximum floods tends to be higher on such catchments than on less permeable catchments, across the full distribution of its values. A t-test shows that there is a significant difference between the mean L-CV on the two groups of catchments.[7]

This contrasts with the finding in FSSR4 (Institute of Hydrology, 1977), made on the basis of a much shorter period of record, that highly permeable catchments tend to have lower variability in peak flows.

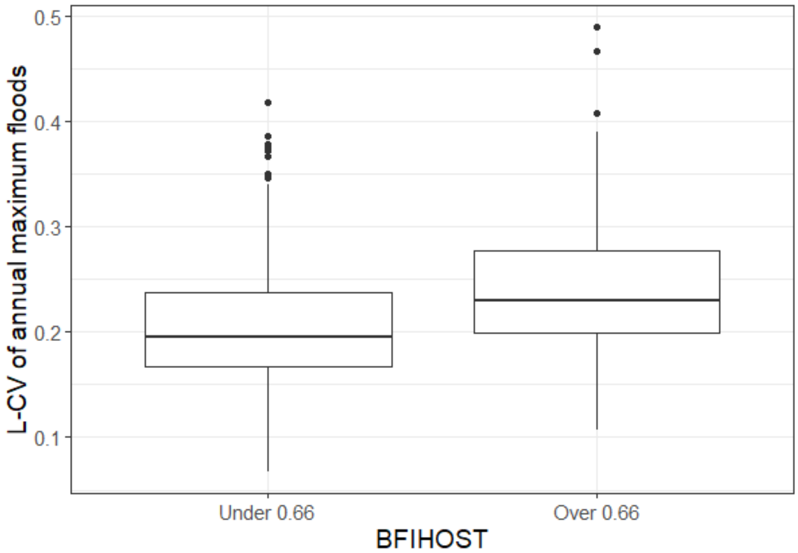


Figure 4-1: Box and whisker plot of L-CV for all gauges classed suitable for pooling and with at least 20 annual maximum flows, partitioned into groundwater-dominated and other. Data source: NRFA peak flow dataset, v10.

The high L-CV seen on some groundwater-dominated catchments can be due to the occasional occurrence of "monster floods" which are far higher than others in the record. An internal research project by JBA in 2013 (unpublished) found that such floods are more prevalent in catchments with BFIHOST>0.80 than elsewhere.

A concern is that these may be able to occur on other groundwater-dominated catchments, even if they have not done so during the gauged record. This highlights the importance of considering longer-term flood history. While some historical reviews readily reveal examples of monster floods (for example on the River Lud), others (for example on the River Itchen[8]) have found none for at least the last 200 years. It is not known whether this difference is down to luck or structural differences between the catchments.

[7] No significant difference was found in the means of the L-skewness between the two groups of catchments.

[8] JBA Consulting (2018). Flood estimation report: River Itchen. Report to Environment Agency.

Case study 4: River Lud, Lincolnshire

Highlights:

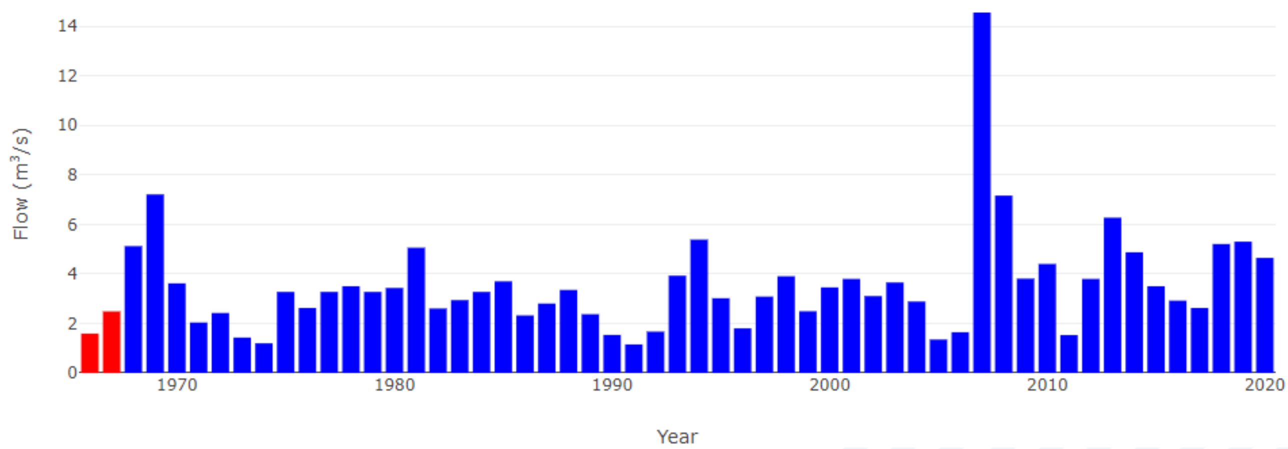
- Change in runoff processes for flash floods
- Value of historical information

Geology: White Chalk and Grey Chalk sub-groups. BFIHOST: 0.82

Lincolnshire has a small outcrop of chalk along a north-south oriented ridge. The Welton Chalk formation (part of the White Chalk) outcrops on the higher ground, with the Ferriby Chalk formation (part of the Grey Chalk) appearing in the valleys, which also have some drift cover consisting of till deposited in an ice age.

There is a gauging station at Louth. The median of the annual maximum flows (QMED) is 3.3m^3/s. The highest flow recorded at the gauge is 14.6m^3/s in June 2007. This is over four times QMED, unusually high for a British catchment. The measurement is uncertain because the gauge is not classed as suitable for pooling: the rating does not consider bypassing or drowned flow and cannot be validated beyond QMED.

Refer to section 3.3 for a description of the flood of May 1920. The estimated peak flow in 1920 was about 150m^3/s, about 45 times QMED. (An even higher estimate of 160-170m^3/s is given by Clark and Arellano, 2004). This is a phenomenally high ratio and suggests a complete transformation in the process of runoff production.



Another flood in January 1857 had an estimated peak flow of 40-60m^3/s.

It seems likely that nearly all annual maximum flows at Louth are due to subsurface flow. Occasionally the rainfall intensity exceeds the infiltration capacity of the chalk escarpments causing rapid surface runoff and leading to much higher peak flows.

One consequence of this switch in behaviour is that it becomes difficult to justify fitting a single statistical distribution to the peak flows when they arise from a mixture of processes.

Another is the importance of learning from historical information when estimating flood frequency on groundwater-dominated catchments. A study that ignored the 1920 flood might obtain very different results from one that included it.

If someone had carried out a flood frequency analysis at Louth in 2006, prior to the 2007 flood, ignoring the longer-term history, they might have obtained a fairly shallow flood frequency curve. Some groundwater-dominated catchments are characterised by many similar annual maximum flows. Growth curves fitted to such data must be treated with caution: there is always the possibility of a much larger flood for example if the groundwater level exceeds a critical elevation, or if there is a severe convective storm. Similar comments apply to pooled flood growth curves.

Case study 4: River Lud, Lincolnshire
References • Bradford and Faulkner (1997) and references therein • Clark, C. and Arellano, A.L.V. (2004) The Louth storm and flood after 80 years. Weather, 59: 71-76. https://doi.org/10.1256/wea.25.03

4.3 Ephemeral streams

Ephemeral streams are a special case of the change in processes between events. In chalk and limestone areas it is common for valleys to be dry in the summer and to have streams flowing in wetter periods. Some streams remain dry for longer periods: the Assendon Stream near Henley has had only five flow episodes since the 1960s (JBA Trust, 2016).

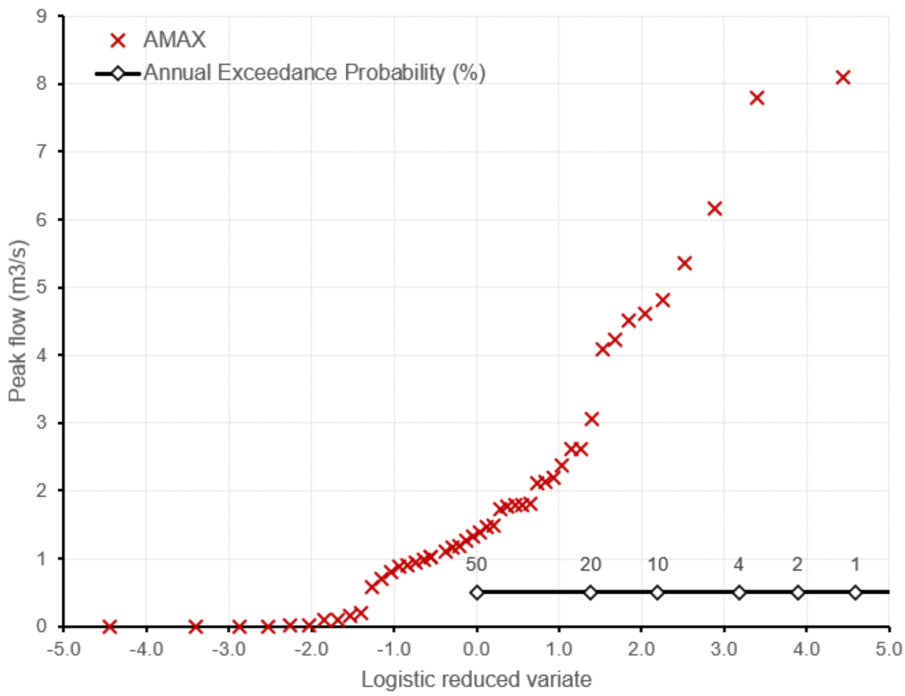
The River Lavant is an example of a gauged ephemeral stream and forms the subject of case study 5.

Ephemeral streams can record annual maximum flows of zero in some years, or very low annual maximum flows that are due to baseflow alone. The following section discusses how to handle these low outliers.

Case study 5: River Lavant, Sussex
Highlights: • Ephemeral stream with some zero annual maximum flows • Value of groundwater level records • Prolonged, high-volume floods
Geology: White Chalk. BFIHOST: 0.94
The course of the upper River Lavant follows the east-west trending Singleton anticline, which is thought to act as a barrier to groundwater movement. There is a gauging station at Graylingwell. The hydrograph is baseflow-dominated, with a gradual response strongly linked to groundwater level. Superimposed on this are smaller more rapid fluctuations in flow, which occur in response to recent rainfall, generally taking just a few hours to peak. These may indicate rapid runoff from rainfall in riparian areas where the groundwater table is close to the surface. The river dries up almost every summer and autumn, and occasionally remains almost dry all year. Five of the 50 annual maximum flows are zero[9] and another four are only marginally above zero. When plotted on a frequency plot, as shown over the page, these annual maximum flows from dry years appear to follow a different flood frequency distribution from the more substantial flows. The influence of these low flows can be suppressed using methods discussed in Section 0. Flood frequency estimation on the Lavant benefits from additional sources of information to provide a longer-term context to the flow record: a very long record of groundwater level at Chilgrove borehole and a chronology of historical floods at Chichester. Both extend back to before the mid-19th century, although the groundwater level data is available only at a monthly resolution before 1960 and so may not identify the peak levels accurately. There is a close correlation between annual maximum flows and groundwater levels, as shown over the page. It is possible to incorporate both these additional data sources to improve estimation of the flood frequency curve. Because standard UK methods of flood frequency method do not permit the simultaneous removal of low floods and addition of historical floods, a flood frequency curve for the Lavant was fitted using the Bulletin 17C method from the USA, which is described in Section 4.4.5.

9 They are missing from the NRFA dataset, but from comparison with flow recorded on nearby catchments, and with annual maximum groundwater levels at Chilgrove, it appears very likely that the river was dry during these years rather than the gauge being inoperable.

Case study 5: River Lavant, Sussex

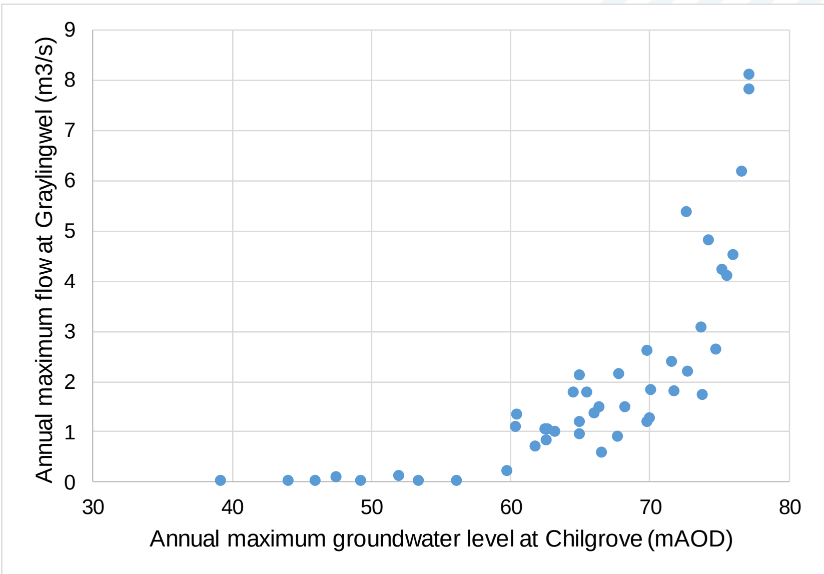


Annual maximum flows for the Lavant at Graylingwell, plotted against return period estimated from the Gringorten plotting position

The plot below shows that in years when the water table measured at Chilgrove borehole in the Lavant catchment does not exceed 60m, there is virtually no flow in the Lavant. As the annual maximum groundwater table rises above 60m there is a steadily increasing maximum flow on the Lavant, and this effect accelerates for groundwater levels above 70m. Hydrogeological investigations have identified a zone of high transmissivity in the chalk at Chilgrove above 70m, the ground surface elevation being about 77m.

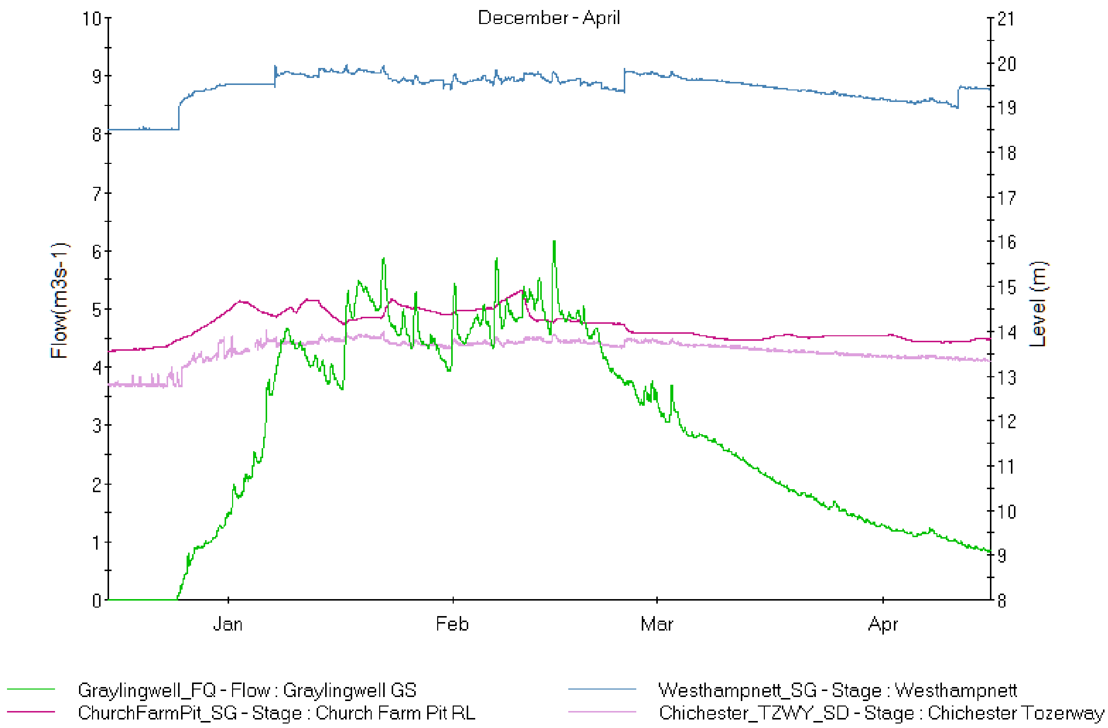
Relationship between annual maximum flow on the River Lavant and annual maximum groundwater level at Chilgrove.

Source: JBA Consulting (2017)



Other challenges of flood frequency estimation on the Lavant include the length of the flood hydrograph. As illustrated over the page, the flow at Graylingwell can remain high for months. This is typical of baseflow-dominated rivers.

Case study 5: River Lavant, Sussex



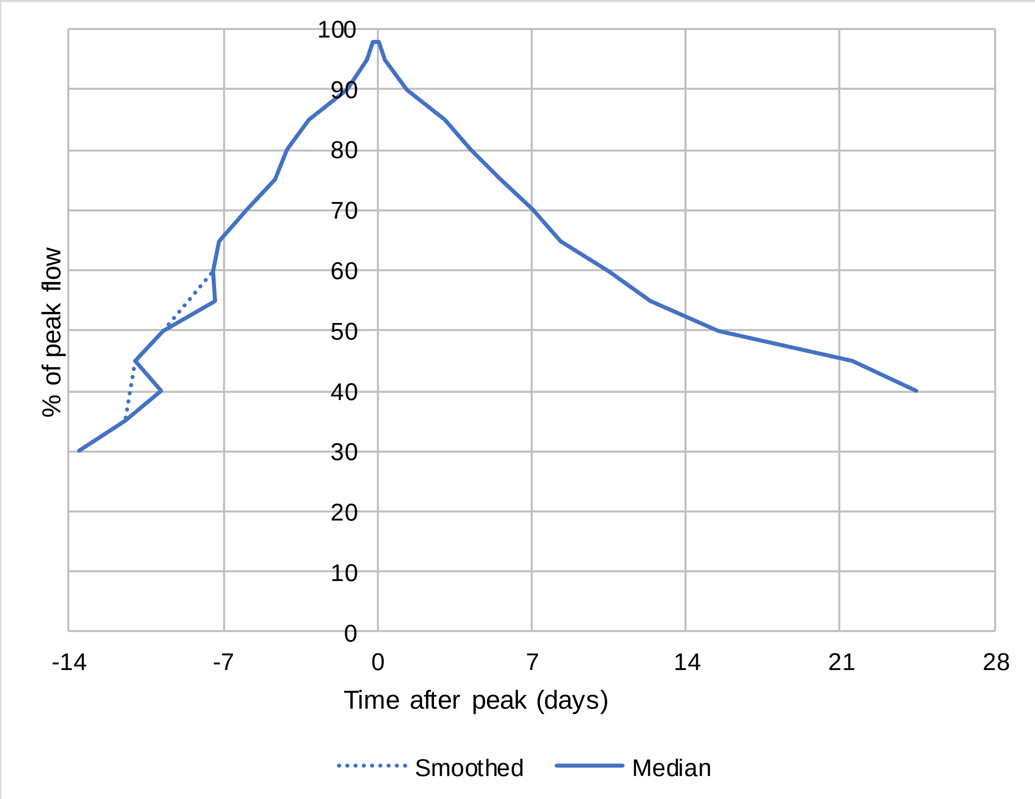
Flow and water level time series for the River Lavant during the flood of winter 2013-14.

Source: JBA Consulting (2017)

A characteristic hydrograph shape for the Lavant was developed empirically by analysis of the shapes of observed flood hydrographs. This used the method developed by Archer et al. (2000) which averages the width of dimensionless hydrographs. The largest 12 annual floods recorded at Graylingwell since 1986 were included in the analysis. The decision to analyse only the larger events was based on an observation that the hydrograph width appears to vary with peak flow, with larger events showing signs of narrower hydrographs. Some of the smallest floods consist of long periods of fairly steady baseflow, hence have long durations.

The plot below shows the resulting hydrograph, derived by taking the median of the observed widths. The hydrograph was smoothed slightly to remove some time-reversals which are a feature caused by the number of events available for averaging decreasing at lower percentiles. The hydrograph is prolonged, with the flow remaining above 75% of the peak for ten days.

Case study 5: River Lavant, Sussex



Characteristic dimensionless hydrograph shape for the River Lavant at Graylingwell

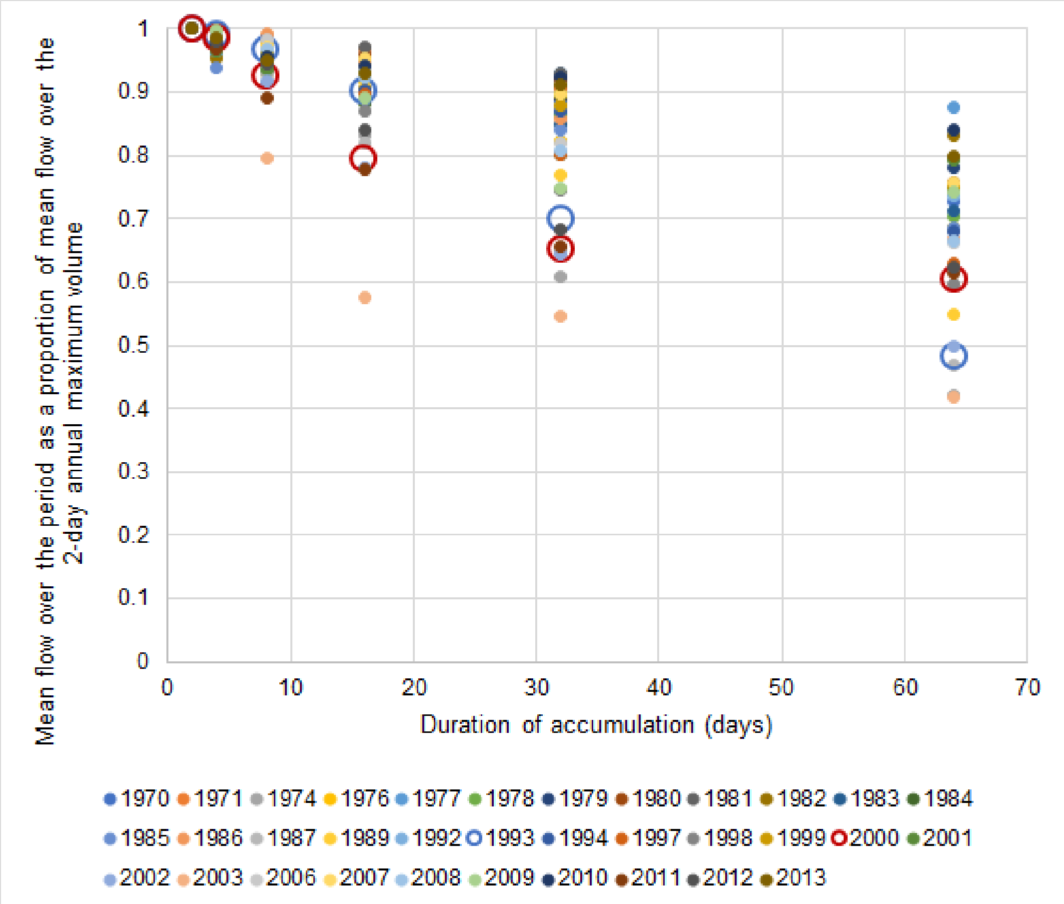
Source: JBA Consulting (2017)

A series of design flood hydrographs was developed by multiplying the estimated peak flow for each AEP by the dimensionless hydrograph shape. To provide an independent check on the resulting flood volumes, a flood volume frequency analysis was carried out. A set of GEV distributions was fitted to annual maximum flow volumes for a range of durations between 2 and 64 days.

The graph over the page shows how the volume of the annual maximum flow changes as the duration over which it is accumulated increases. Volumes are normalised by the volume over the 2-day duration. Even for very long accumulations, over 2 months, in some events the normalised volume drops little compared with that for the highest-volume 2-day period during the same event. This implies that the flow is nearly constant over the whole period.

The data points for the floods in water years 1993-94 and 2000-01 are highlighted. These events see a relatively rapid drop in mean flow as the duration increases, which is consistent with the finding that larger floods tend to have narrower normalised hydrographs. This is expected given that groundwater levels can rise and fall quite rapidly once the water table rises into the near-surface zone of the Chalk, and artesian conditions do not tend to persist for months.

Case study 5: River Lavant, Sussex



Relationship between flood volume (normalised) and duration. Very small and zero annual maximum volumes are excluded from this plot.

Source: JBA Consulting (2017)

Over a duration of 32 days (the longest for which the characteristic hydrograph was defined), the volumes of the design flood hydrographs were within 15% of the volumes estimated by the volume frequency analysis, for a range of AEPs between 1% and 50%. An exact match between the two sets of volumes was not expected because the frequency analysis of peak flows incorporated historical information whereas the analysis of volumes did not.

In conclusion, the flood volume frequency analysis indicated that the approach of combining a peak flow with a normalised characteristic flood hydrograph gave a design flood hydrograph with a realistic volume over a long duration.

References:

JBA Consulting (2017). Flood estimation report: Lavant at Chichester. Report to Environment Agency.

4.4 Annual maximum flows that are not floods

4.4.1 Introduction to the problem

As seen on the River Lavant in case study 5, some groundwater-dominated catchments can record annual maximum flows (AMAX) of zero, or other low values that are clearly not associated with floods. The project scope requires a discussion of methods for accounting for this feature when fitting a flood frequency curve.

4.4.2 On what sort of catchments do these occur?

We have analysed the NRFA peak flow dataset to investigate the prevalence of low AMAX. Figure 4-2 shows the distribution of AMAX values across the whole dataset, each expressed as a ratio of the median AMAX for the gauging station, QMED.

4.3% of AMAX across the UK are below half of QMED, which is the threshold that the FEH uses to define a low AMAX. On catchments with BFIHOST<0.66 the proportion is 3.7% and on those with higher BFIHOST it is 8.1%. So there is a higher proportion of low AMAX on groundwater-dominated catchments.

0.2% of all AMAX are below 0.1 QMED.

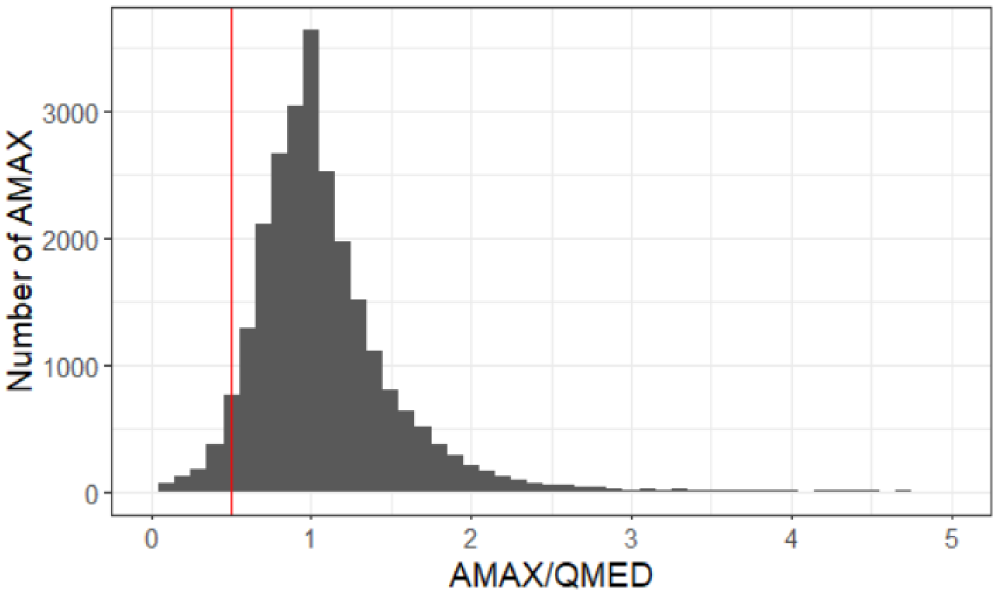


Figure 4-2: Histogram of annual maximum flows , expressed as a ratio to QMED. Data source: NRFA peak flow dataset, v10, gauges suitable for pooling. The red line shows the threshold of 0.5QMED.

There is little sign of a relationship between the ratio of AMAX/QMED and BFIHOST (Figure 4-3).

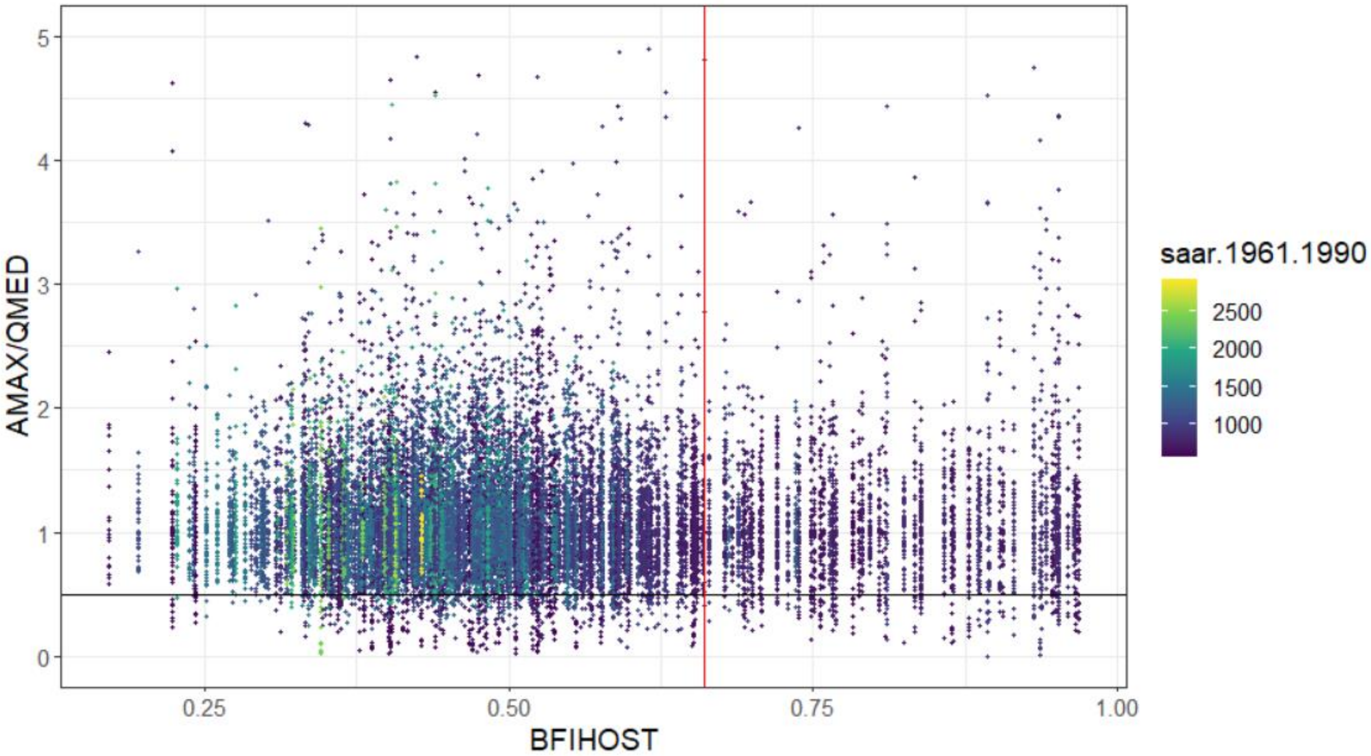


Figure 4-3: Annual maximum flows as a ratio to QMED, plotted against BFIHOST. Colour indicates SAAR (mm). Peak flows above 5QMED are not plotted. Data source: NRFA peak flow dataset, v10, gauges suitable for pooling.

There is a clearer relationship with SAAR (Figure 4-4). Very low AMAX seem to occur mainly on low-SAAR catchments. This makes intuitive sense: such catchments are more likely to build up large soil moisture deficits in some years, leading to little runoff. Using SAAR as a threshold more reliably identifies catchments with AMAX<0.5 QMED: for example, where SAAR<800mm, 8.6% of AMAX are below 0.5QMED, compared with 2.6% for higher SAAR. The outliers on the SAAR plot with SAAR nearly 2500mm and very low AMAX are all on Haweswater Beck, in years during which there was no spilling from Haweswater Reservoir, so they do not represent a natural process.

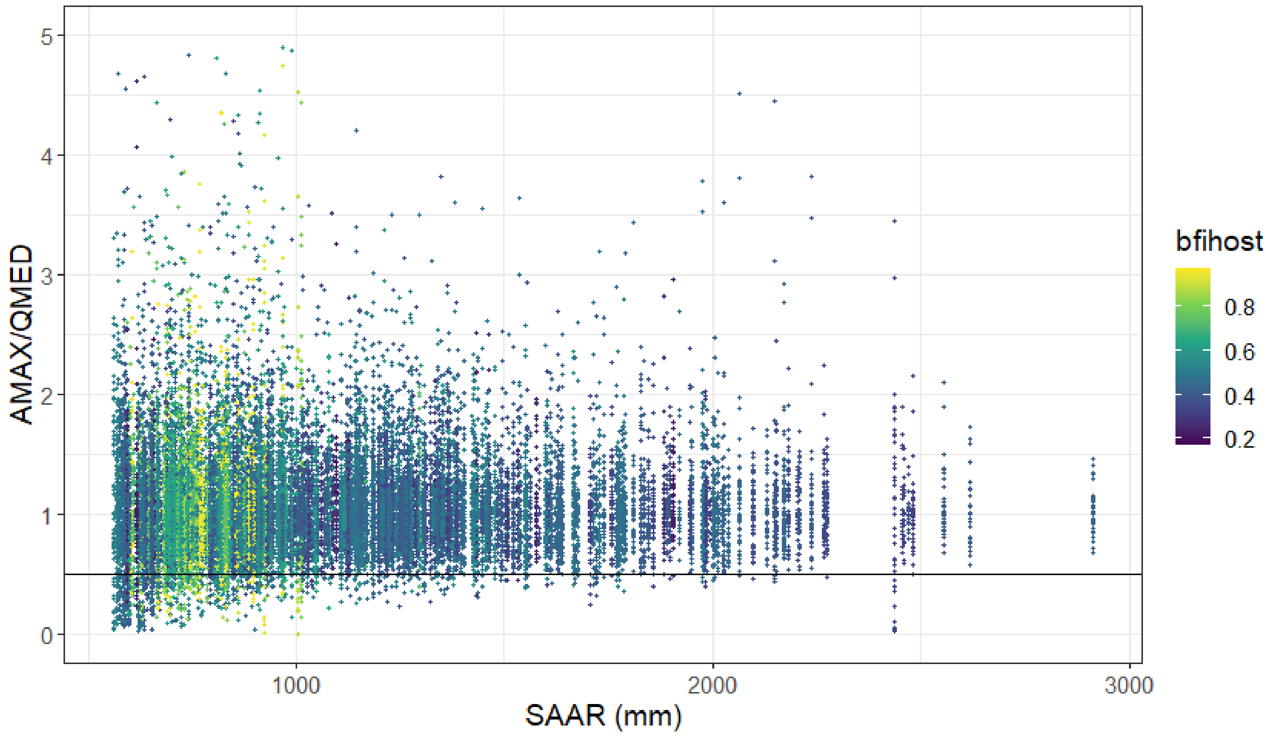


Figure 4-4: Annual maximum flows, as a ratio to QMED, plotted against SAAR. Colour indicates BFIHOST. Peak flows above 5QMED are not plotted. Data source: NRFA peak flow dataset, v10, gauges suitable for pooling.

Unusually low AMAX seem to be a phenomenon that tend to occur on low-rainfall catchments, or perhaps specifically on catchments that are more prone to drought conditions. Most groundwater-dominated catchments have low rainfall, so there is also some association with BFIHOST. But there are plenty of surface water-dominated catchments with low rainfall, where these AMAX are also found.

The FEH suggested a threshold of SPRHOST<20% to define catchments where adjustment is worthwhile, but we recommend that any adjustment method should be applied irrespective of catchment type, wherever low AMAX are found.

4.4.3 The FEH "permeable adjustment"

Including non-flood AMAX in a flood frequency analysis is not logical. It could introduce bias in the flood frequency curve. Occasionally it can result in a fitted growth curve which is bounded above (i.e. the growth factor reaches an upper limit).

Volume 3 of the FEH presents a method for removing the influence of AMAX below 0.5QMED. It refers to this as a "permeable adjustment". Faulkner and Robson (1999) provide some background to the method and discuss alternative approaches.

The threshold of 0.5QMED is arbitrary. Bradford and Goodsell (2000) suggest that a better threshold might be identified locally by comparison of groundwater levels and river flows to identify the groundwater level at which flows start to show a marked increase due to water entering the shallow, high-transmissivity zone of the aquifer. On the Rivers Lavant and Ems the corresponding flow is higher than 0.5QMED; in fact slightly higher than QMED. A drawback of using a higher threshold is that the sample size is reduced, leading to a larger standard error in fitting the flood growth curve. The adjustment method can be applied using any reasonable threshold; it is not specific to 0.5QMED.

The adjustment can be applied either in a single-site analysis or as part of a pooled analysis. It involves three steps:

1. Fitting a Generalised Logistic (GL) growth curve to the AMAX that fall above the threshold. These data are referred to as flood years. The growth curve gives the conditional distribution of flood sizes, given that a year has a flood.
2. Adjusting the growth curve to convert it from one applicable only in flood years to one applicable in all years. This involves stretching the growth curve along the return period axis and a slight re-scaling to ensure that the growth curve retains a value of 1.0 for a return period of 2 years (equivalent to an AEP of 50%).
3. Calculating the equivalent L-moments for the parameters of the adjusted growth curve. This step is only needed if the adjustment forms part of a pooled analysis.

Step 2 involves numerical solution of an equation and so cannot be applied analytically.

The adjustment works for situations where the number of flood years is at least half of the total number of years.

The adjustment was developed only for the GL distribution. With some straightforward manipulation of the mathematics, as shown below, it can also be applied to the 2-parameter Logistic distribution, which does not need a numerical solution. This is worthwhile in cases when the shape parameter (k′) of the GL distribution for flood years is close to zero. Without this, the adjusted scale parameter can take a physically impossible negative value.

Applying the FEH adjustment to the Logistic distribution

The notation is the same as that used in the FEH. Equation numbers are those used in the FEH, Volume 3.

β' is the estimated scale parameter of the Logistic distribution for flood years. We know this from fitting the distribution using L-moment ratios.

β^* is the estimated scale parameter of the Logistic distribution for all years. This is what we are trying to find.

ω is the proportion of years in which a flood occurs

t_2 is the sample L-CV of the annual maximum flows in the flood years

C is a constant defined in Eqn 19.10.

x_T is a growth factor for return period T (′ for flood years and * for all years)

$\beta' = t_2$ (Eqn 15.29)

$C^{-1} = x_{2\omega}' = 1 + \beta' \ln(2\omega - 1)$ (Eqn 15.28)

$1 + \beta^* \ln(T-1) = C(1 + \beta' \ln(\omega T - 1))$ (Eqn 19.10)

Rearrange to find β^*:

$$\beta^* = \frac{C[1 + \beta' \ln(\omega T - 1)] - 1}{\ln(T - 1)}$$

This equation can be solved for any return period T greater than 2 years, giving almost identical results. 50 years is a suggested return period to use.

It would, in theory, also be possible to develop a version of the adjustment for other distributions. This would mainly be a matter of mathematical manipulation which should not require a great deal of time to carry out.

The FEH adjustment is not applied in WINFAP. Several consultants have developed in-house code to implement it, and it is also available in a spreadsheet developed by Wallingford

HydroSolutions[10] and in the UKFE package available in the R language (Hammond, 2020). We have not attempted to audit these different software implementations, but from an initial look it appears their results may not all agree, even at individual sites. Further complications can arise when attempting to incorporate adjusted L-moments into a pooled or enhanced single-site flood frequency curve, especially when de-urbanisation of the L-moments is needed.

4.4.4 Effect of the FEH adjustment

The FEH presents examples of the difference that the adjustment can make for single-site growth curves. In three out of four cases, the adjusted growth curves have more positive skewness, leading to increased flood estimates for low AEPs. Since that analysis, record lengths at most gauges have increased by over 25 years and this is likely to have reduced the sensitivity of fitted growth curves to non-flood AMAX. In addition, it is more typical to estimate growth curves by pooled analysis. Although the original FEH method accounted for BFIHOST in selecting pooling groups, the current method (Kjeldsen et al., 2008) does not and so a pooling group for a groundwater-dominated catchment will not necessarily include other catchments with similar geology. This will tend to dilute the influence of low AMAX on pooled growth curves. With these reasons in mind, it is worth asking how much difference the FEH adjustment makes using current datasets.

Figure 4-5 shows the effect of the FEH permeable adjustment using current datasets. We have applied it for estimating the single-site growth curve at the six gauging stations on groundwater-dominated catchments that contain the largest numbers of AMAX below 0.5QMED. The proportion of non-flood years varies from 15% to 34% across these gauges. The adjustment was applied using the UKFE package (Hammond, 2020).

In five out of the six cases the adjustment, as implemented in UKFE, has little effect. In the sixth, at gauge 44013 (Piddle at Little Puddle), the effect is to reduce the skewness, leading to lower estimates at low AEPs. It seems that the typical effect of the adjustment has reversed since the FEH was written. This may be due to the occurrence of high AMAX on many groundwater-dominated catchments in the intervening decades, for example in 2000 and 2014. Typical record lengths for these gauges are 40-60 years, although 44013 is shorter, at 28 years.

If the adjustment has only a modest effect on most single-site calculations, it can be expected to have negligible effect on a typical pooled analysis, because it will be diluted by the inclusion of other gauges with no non-flood years.

[10] https://www.hydrosolutions.co.uk/software/free-downloads/whs-permeable-adjustment-worksheet/

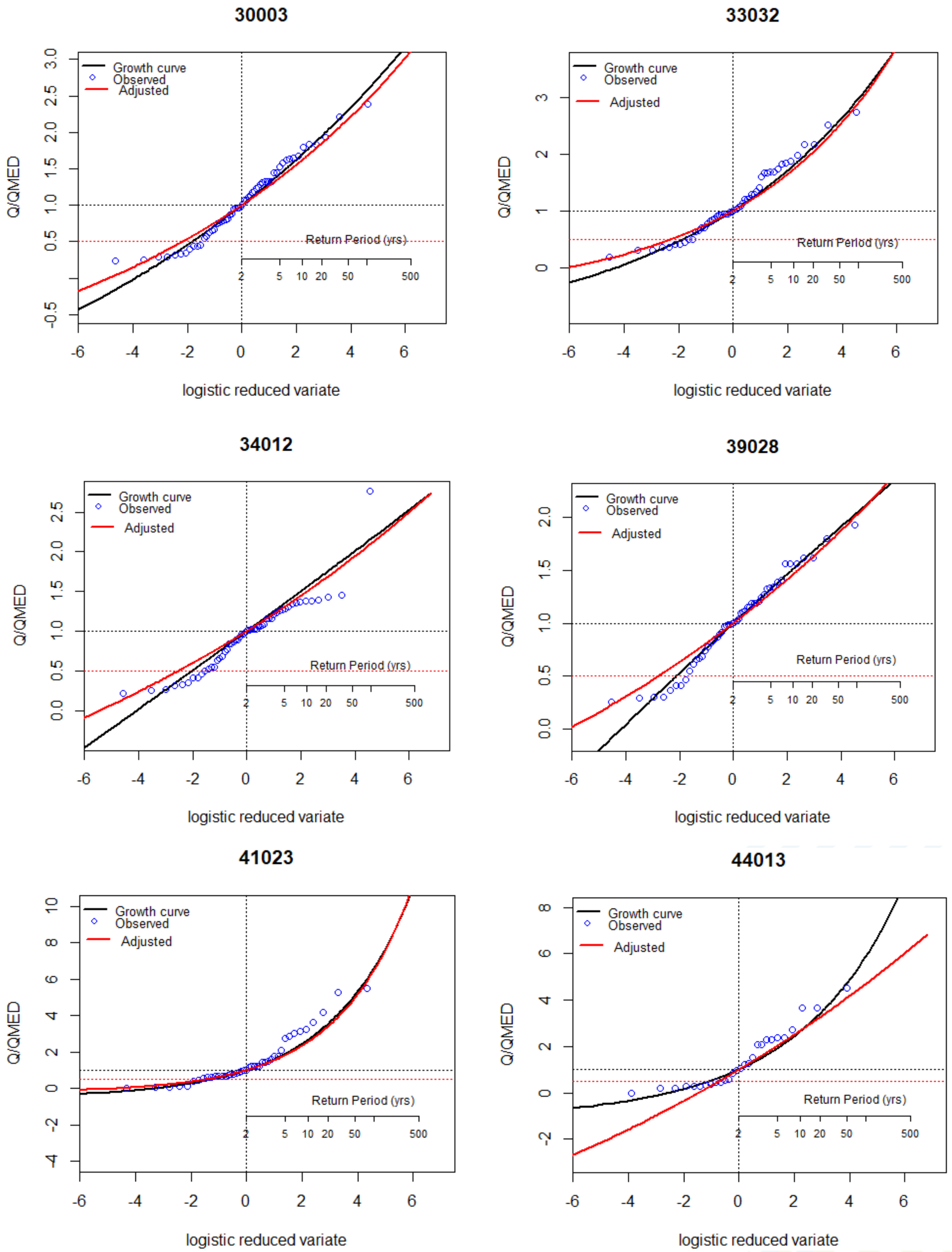


Figure 4-5: Effect of the FEH "permeable adjustment" at six gauging stations. The horizontal dotted red line shows the 0.5QMED threshold below which AMAX are removed.

We have also applied the adjustment over the entire NRFA peak flow dataset. The results are summarised in Figure 4-6, which shows the ratio of the 1% AEP flow estimates before and after the adjustment. Ratios above 1 indicate that the adjustment increases the estimates. The largest reduction in the estimated flow was -20% at gauge 44013 (which is included in Figure 4-5). The largest increase was +12% at gauge 44003. Smaller changes were seen for higher probability floods. As should be expected from the analysis in section 4.4.2, the adjustment can have a noticeable effect on some catchments with low BFIHOST.

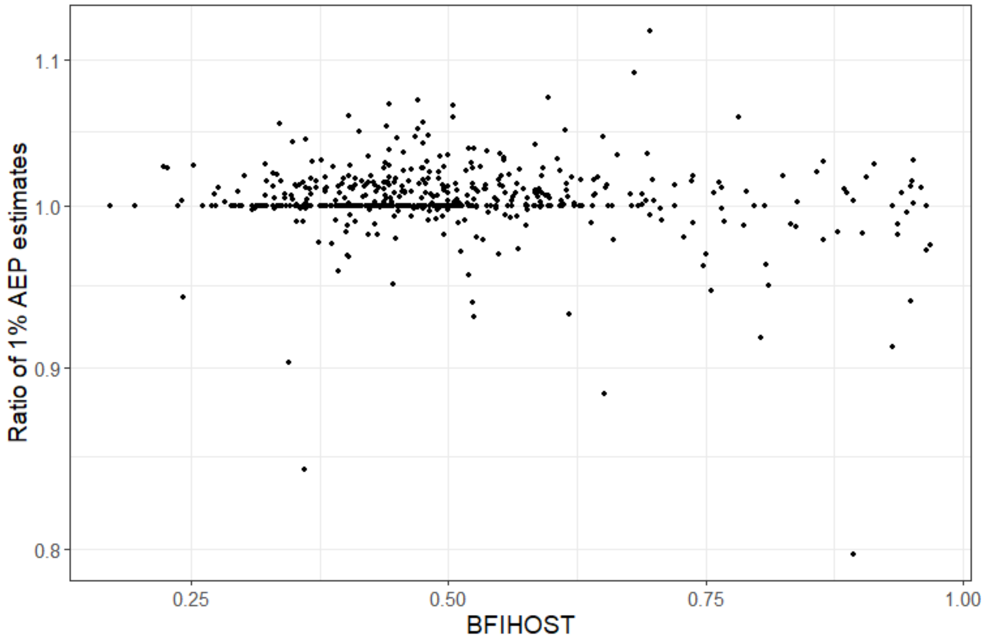


Figure 4-6: Effect of the FEH "permeable adjustment" on the estimated 1% AEP flow at all gauges, plotted against BFHOST. Data source: NRFA peak flow dataset, v10, gauges suitable for pooling.

Because the adjustment has little effect on the vast majority of current AMAX series, even when fitting a single-site growth curve, it may not be worth investing much in attempting to improve it or extend its applicability beyond the GL distribution.

4.4.5 Alternative adjustment methods

Alternative approaches to removing the influence of low AMAX include (Faulkner and Robson (1999)):

- Using 2 or even 5-year maxima rather than annual maxima. This would remove the influence of long periods of low flows. However, it would dramatically cut the effective sample size.
- Carrying out frequency analysis based on peak-over-threshold (POT) data, which would only identify floods. Unfortunately it is often not possible to derive POT data on groundwater-dominated catchments because independent peaks are difficult to identify on the subdued hydrographs. An alternative which can work well is to identify peaks from rolling average flow series, which can show more distinct independent peaks.
- Fitting the growth curve to AMAX flows using partial probability-weighted moments. Bradford and Faulkner (1997) found that this was unsatisfactory in a UK context.
- Fitting the growth curve to AMAX flows using linear higher-order moments, or LH-moments.

Internationally, methods have been developed for the same purpose, for application on arid catchments. In the USA, Bulletin 17C is the current guidance document for flood frequency estimation, issued by the US Geological Survey (England et al., 2019).

The Bulletin 17C method first removes zero flows and "potentially influential low floods" (PILFs) from the AMAX series. The same method is also recommended in Australian Rainfall and Runoff (ARR) (Ball et al., 2019). The recommended way of detecting PILFs is to use the multiple Grubbs-Beck test which uses a significance test to evaluate a null hypothesis that all AMAX in a sample are drawn from the same population of identically distributed data. The alternative hypothesis is that the k-th smallest observation in the dataset is unusually small compared to that population. If that observation is declared a PILF then all smaller observations are declared a PILF. This provides a more objective way of detecting PILFs than the FEH approach of setting an arbitrary threshold of 0.5QMED, although both US and Australian guidance also allow for a user-defined threshold.

Once PILFs are removed, the Bulletin 17C method uses the expected moments algorithm (EMA) to fit a flood frequency distribution. This has the additional capability to incorporate non-systematic data such as historical floods. Essentially it can remove occasional floods below a low threshold and add occasional floods above a high threshold. ARR uses a different, Bayesian, method for curve fitting.

The Bulletin 17C method is implemented in the HEC-SSP software and the ARR method in TUFLOW-FLIKE.

In Alberta, Canada, a simple conditional probability method is recommended for removing zero flows from flood frequency analysis (Alberta Transportation, 2001). It involves first excluding the zero flow years from the series and fitting a conditional probability distribution, G(x), to the remainder - that is, conditional on the stream not being dry for the year. The probability of zero flows, p_0, is then estimated as the proportion of zero-flow years in the original record. Finally the unconditional probability distribution F(x) is converted from the conditional probability distribution G(x) using the formula:

$$F(x) = p_0 + (1-p_0)G(x)$$

In practice the conversion can be applied to the return period rather than the probability, i.e. the desired return period T of the design flood can be converted to a return period T´ measured in terms of years of non-zero flow, using

$$T' = (1-p_0)T$$

The resulting value of T´ is then substituted into the conditional probability distribution. This enables calculation of the design flood for the return period T.

This conditional probability method is like the FEH "permeable adjustment", although much simpler because it does not attempt to find parameters for an adjusted flood frequency curve. It could not be incorporated into a pooled analysis like the FEH adjustment can.

4.4.6 Suggestions for improved practice

We have the following suggestions for improvements to the procedure for removing the influence of non-flood years:

- Do not refer to it as a "permeable adjustment". For one thing, it is equally relevant in some catchments with low permeability. For another, the name might give some analysts the impression that if they apply the adjustment they will have dealt with all issues associated with flood estimation on groundwater-dominated catchments. This is a long way from the truth.
- Apply the same method to all catchments rather than restricting it to groundwater-dominated catchments.
- Test whether the adjustment improves the fit of flood frequency curves, for example using tail fit scores.

- Consider a more objective way of identifying non-flood years, such as the multiple Grubbs-Beck test. We do not recommend this as a high priority since the adjustment has such a small effect on the vast majority of catchments.
- If, after the above, the adjustment is found to be worthwhile, incorporate it (with any alterations made) in all software that implements the FEH Statistical method, including WINFAP.
- Prioritise other improvements to flood estimation in groundwater-dominated catchments over this one.

4.5 Non-stationarity

Some groundwater-dominated catchments show evidence of non-stationarity, generally with flood flows increasing over time (Faulkner et al., 2019). A possible explanation is the sensitivity of flood magnitude to recharge, which is strongly related to rainfall over the winter season. Faulkner et al. (2020) found that in some chalk areas of southern England, winter rainfall is a useful covariate for modelling non-stationarity in flooding, whereas elsewhere annual rainfall tends to be preferred. Climate change scenarios indicate a projected increase in winter rainfall, although a statistically significant trend is not yet always evident in observed data.

As illustrated in case study 6, trend testing and non-stationary flood frequency analysis can be useful tools on such catchments.

Case study 6: River Itchen and tributaries, Hampshire

Highlights:

- Non-stationary flood frequency analysis

Geology: White Chalk outcrop of South Downs, largely drift-free. BFIHOST (Itchen at Highbrook and Allbridge): 0.95

The Itchen and some of its headwater tributaries are permanent watercourses although there are some ephemeral reaches in the upper catchment, for example on the River Alre. The groundwater catchment area exceeds the topographical catchment by about 130 km^2.

There are gauging stations on the Itchen and on its three tributaries, the Candover Stream (42009), Cheriton Stream (42008) and River Alre (42007). The gauging station on the Itchen at Allbrook and Highbridge (42010) measures flow from the Chalk portion of the catchment. There are no time series of groundwater levels in the catchment that pre-date the river flow records.

Annual maximum flows at all gauging stations, shown in the time series over the page, show signs of an upward trend since the 1960s or early 1970s. This appears to affect the variability of floods as well as the mean, with large floods occurring in the more recent period of the record, the largest being in winter 2000-2001 and winter 2013-2014. At most gauges the fourth largest floods all fall in the second half of the record.

A Mann-Kendall test for trend reveals an upward trend at all four gauges, although it is only significant at the 5% level at one gauge, 42008. At 42007 and 42009 it is significant at the 10% level. This non-parametric test does not account for the magnitude of the trend or for non-stationarity in variability of floods.

Winter rainfall over the Itchen catchment also shows an upward trend over the period of the river flow record, although it is not statistically significant at the 5% level according to the Mann-Kendal test[11]. There is a close correlation (coefficient ρ=0.72) between winter rainfall and annual maximum flow on the Itchen (gauge 42010).

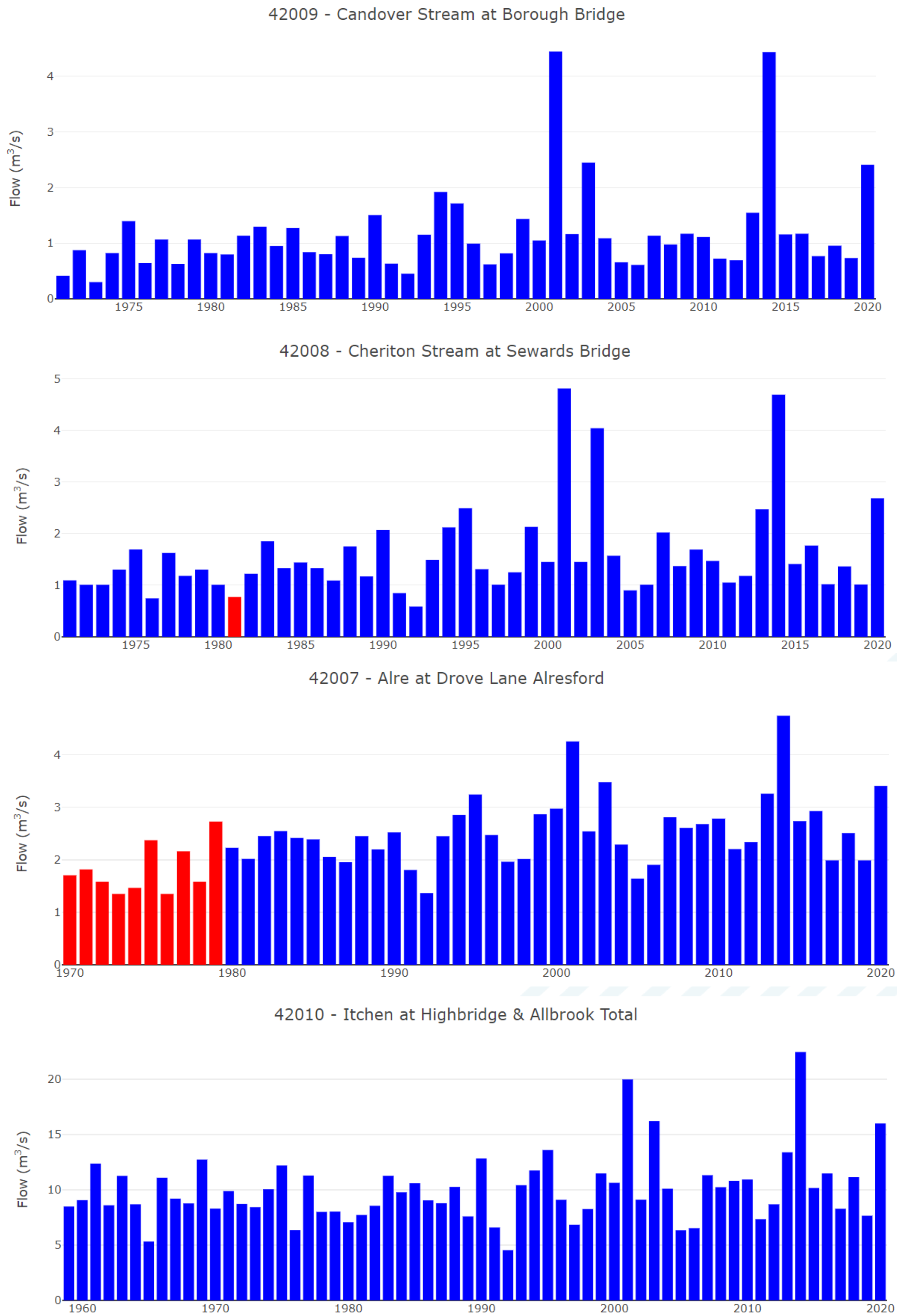
One way to account for the trend in flows is to fit a non-stationary flood frequency distribution. When non-stationarity is modelled using a GL distribution whose parameters can vary with time, at three of the gauges, a stationary distribution provides the best fit. The only gauge where the non-stationary model using time as a covariate fits better than the stationary model is 42009 (Candover Stream), where the non-stationary model has a trend in the location parameter of the GL distribution.

An alternative approach is to model the non-stationarity using physical variables as well as covariates, either instead of or in addition to time. Using the range of covariates and fitting options described in Faulkner et al. (2020)[12], the best fitting models include winter rainfall as a covariate, either for the location parameter, the scale parameter or both. At one gauge (42010), time is also included as a covariate, indicating that winter rainfall does not adequately explain all the trend in annual maximum flows.

[11] This analysis was carried out using CEH-GEAR rainfall data, averaged over the catchment of gauge 42010 for water years 1958-2016.

[12] The results for physical covariates are taken from Faulkner et al. (2020) and so use data up to water year 2016/17 only. The trend tests and the non-stationary analysis using time as a covariate includes some more recent data, up to water year 2019/20.

Case study 6: River Itchen and tributaries, Hampshire



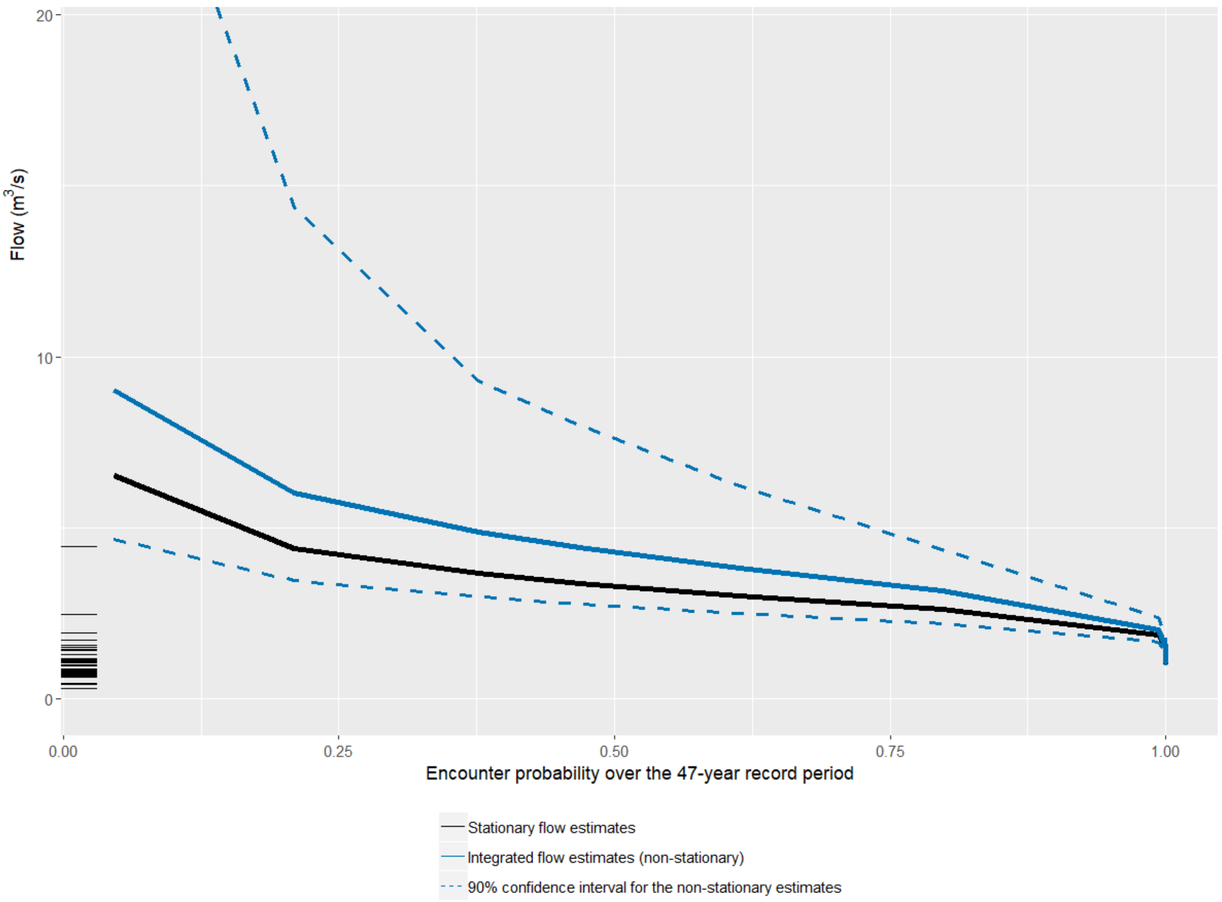
Annual maximum flows on the Itchen (bottom) and its three tributaries. Source: NRFA peak flow dataset, v10. Note differing periods of record.

Case study 6: River Itchen and tributaries, Hampshire

Example results for the non-stationary analysis using physical covariates are shown below for gauge 42009, on the Candover Stream. The graph plots flow against the encounter probability, i.e. the probabiity of exceeding that flow over the record length at the gauge. Flows for a given probability are higher for the non-stationary model.

The confidence interval for this model is very wide because it includes three more parameters than the stationary model. This illustrates one of the drawbacks of non-stationary analysis: there can be a very large estimation error in its results.

Another limitation is that even models that use physical covariates can have limited explanatory power for the future. The degree to which they can be relied on in future conditions depends on the cause of observed trends and on whether the statistical model includes all physical processes that influence the non-stationarity. On a rural groundwater-dominated catchment like the Itchen we can be confident that trends in urbanisation or short-duration rainfall intensity are unlikely to be influential. Winter rainfall is likely to be the main way in which climate change influences flood flows, although potential evaporation (linked to temperature) may also be worth including in a non-stationary model.



References

JBA Consulting (2018). Flood estimation report: River Itchen. Report to Environment Agency.

The trend tests and non-stationarity analysis were carried out using the *nonstat* package in R (Warren and Longfield, 2020).

4.6 Long-lasting floods

An important feature of floods on some groundwater-dominated catchments is their longevity. Baseflow can stay high for months on some rivers, leading to prolonged high river flows.

On such catchments, flood mitigation measures that involve storage are unlikely to be feasible. Additional economic damage may occur from the long term disruption of transport routes and essential infrastructure.

The concept of a design flood hydrograph becomes more difficult to apply when the flood lasts for months. Rainfall-runoff models such as ReFH2 that simulate a flood arising from a single storm event are not appropriate in such cases. Where there is gauged data, a characteristic dimensionless hydrograph shape can be derived empirically. There are two common approaches to this:

1. Lining up normalised flood hydrographs (flow expressed as a ratio to the peak) so that their peaks coincide and then averaging the normalised flow values at each time step.
2. Calculating the widths of observed hydrographs at different percentiles of the peak flow and using the median width at each percentile to construct a characteristic hydrograph shape. This method was developed by Archer et al. (2000) who demonstrated it on the River Frome, a baseflow-dominated catchment on the Cotswold limestone.

Neither approach is necessarily adequate on rivers where short-term peaks are superimposed on a long-term baseflow response. Formetta et al. (2022) give an example of applying approach (2) on a selection of small catchments, including the Redbourn which is groundwater-dominated with occasional rapid runoff from an urban area. Some flood hygrographs are due to sustained high groundwater levels and others to intense runoff from impermeable areas. These contrasting mechanisms are difficult to characterise as a single representative hydrograph shape.

The Lavant case study on page 33 includes an example of a long-lasting flood and describes how method (2) was used to develop a characteristic hydrograph shape for application in conjunction with design estimates of peak flows.

Another approach to building up a suitable hydrograph shape is to carry out frequency analysis of flood volumes. Bradford and Goodsell (2000) investigated flood volumes on permeable catchments and recommended carrying out volume frequency analysis by fitting a GL distribution to a series of annual maximum flood volumes over a given duration. This involves extracting discharge volumes over a period of d consecutive days from daily mean flow data, where d is the duration of interest. The annual maximum volume series is standardised by its median (if required) and the distribution is fitted by L-moments. A volume frequency analysis can form a useful complement to the hydrograph width analysis mentioned above, as illustrated in the Lavant case study.

For some types of flood study it may be appropriate to ignore flood volumes and timing, instead running a hydraulic model in steady-state. This is a simpler approach that is more readily justified for very slowly changing flow rates.

5 Performance of FEH methods on groundwater-dominated catchments

5.1 FEH Statistical method

5.1.1 Introduction

An assessment of the performance of the FEH Statistical method on groundwater-dominated catchments would be worthwhile but the budget available for this project did not permit this. Despite this, some observations are offered in this section.

5.1.2 QMED

There are anecdotal reports of large errors in the FEH regression equation for QMED in some groundwater-dominated catchments. The River Ver case study on page 25 includes some examples.

A systematic comparison of QMED from the current FEH regression against QMED estimated from annual maximum flows (Figure 5-1) bears this out. This is inspired by a figure in Hammond (2021).

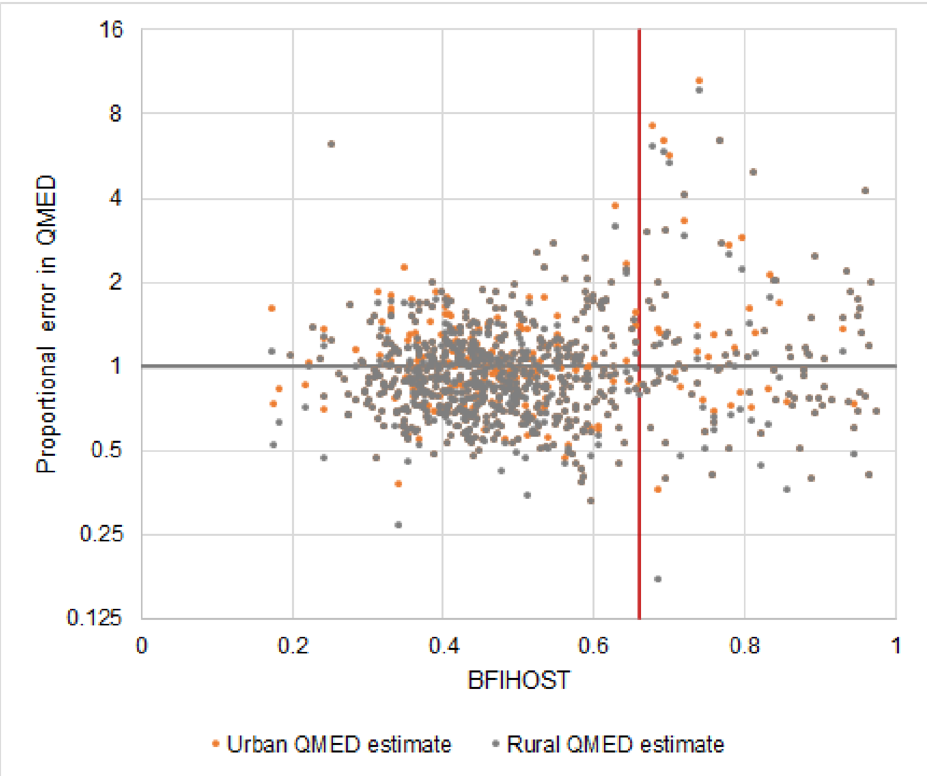


Figure 5-1: Proportional estimation errors in the FEH regression equation for QMED, plotted against BFIHOST. Data source: NRFA peak flow dataset, v10, gauges suitable for QMED with at least 10 AMAX flows. Plot shows both rural and urban catchments, with the effect of urban adjustment shown. Calculations carried out in the UKFE package (Hammond, 2020).

There is a visible increase in the scatter for catchments with BFIHOST>0.66. Nine catchments show overestimation of QMED by more than a factor of 4, and all but one have BFIHOST>0.66. The exception is the Hebden Beck, gauge 27032 which, although it not a high-BFIHOST catchment, is strongly influenced by karst, with the headwaters of the catchment being diverted via a cave system into the neighbouring River Wharfe. So this can also be classed as a groundwater issue.

Even larger errors in QMED estimates have been reported at some other gauges on groundwater-dominated catchments, not classed as suitable for QMED, such as the Ver at Redbourn, described in the case study on page 25, where the reported proportional error is over 15.

After urban adjustment, no gauges show underestimation of QMED by a factor exceeding 4.

The (geometric) mean bias in the estimated QMED is -5% for BFIHOST<0.66 and +17% for BFIHOST>0.66. There is clearly a tendency towards overestimation on groundwater-dominated catchments.

The root mean square error in log QMED is 0.36 for BFIHOST<0.66 and 0.66 for BFIHOST>0.66. The error is nearly twice as high for groundwater-dominated catchments as for others.

The table below lists the gauges at which the FEH regression and urban adjustment overestimate QMED by a factor exceeding 4. At all the gauges but 26016 the BFIHOST is below 0.85, some being below 0.70.

Table 5-1: Gauges that show the largest overestimation of QMED by the FEH Statistical method

Station ref.	No. of AMAX	BFIHOST	URBEXT	$QMED_{AM}$	$QMED_{CD}$ (rural)	$QMED_{CD}$ (urban)	Proport-ional error (urban)	Notes
26016	23	0.96	0.00	0.10	0.43	0.43	4.28	Gypsey Race at Kirby Grindalythe
27032	54	0.25	0.00	4.08	25.6	25.6	6.28	Hebden Beck at Hebden. Karst, as explained above.
38011	58	0.74	0.04	0.42	4.18	4.50	9.83	Mimram at Fulling Mill.
39014	64	0.68	0.11	1.22	7.46	8.90*	7.29	Ver at Hansteads. Some urbanisation.
39027	52	0.72	0.01	2.20	9.08	9.08	4.12	Pang at Pangbourne
39033	58	0.77	0.00	0.40	2.57	2.57	6.40	Winterbourne Stream at Bagnor.
39088	46	0.69	0.06	1.14	6.68	7.37	6.49	Chess at Rickmansworth. Some urbanisation.
39089	47	0.70	0.03	0.61	3.26	3.46	5.66	Gade at Bury Mill.
44008	41	0.81	0.00	0.45	2.22	2.22	4.96	South Winterbourne at Winterbourne Steepleton.
* This is slightly lower than the value given in the case study for the River Ver, perhaps due to the fact that the urban extent was updated in that case study. $QMED_{AM}$ at Hansteads is also slightly different in the case study, which was based on an earlier version of the peak flows dataset.								

Figure 5-2 is a map of the proportional errors in QMED, giving prominence to the results on groundwater-dominated catchments. All eight of the high-BFIHOST gauges in Table 5-1 are on Chalk catchments and all but 44008 are further north than the chalk streams draining to the south coast of England. Four are on neighbouring catchments in Hertfordshire, the Mimram and three tributaries of the Colne (39014, 39088 and 39089). Reasons for overestimation in the Ver catchment are discussed in the case study on page 25, and this may also affect the Mimram.

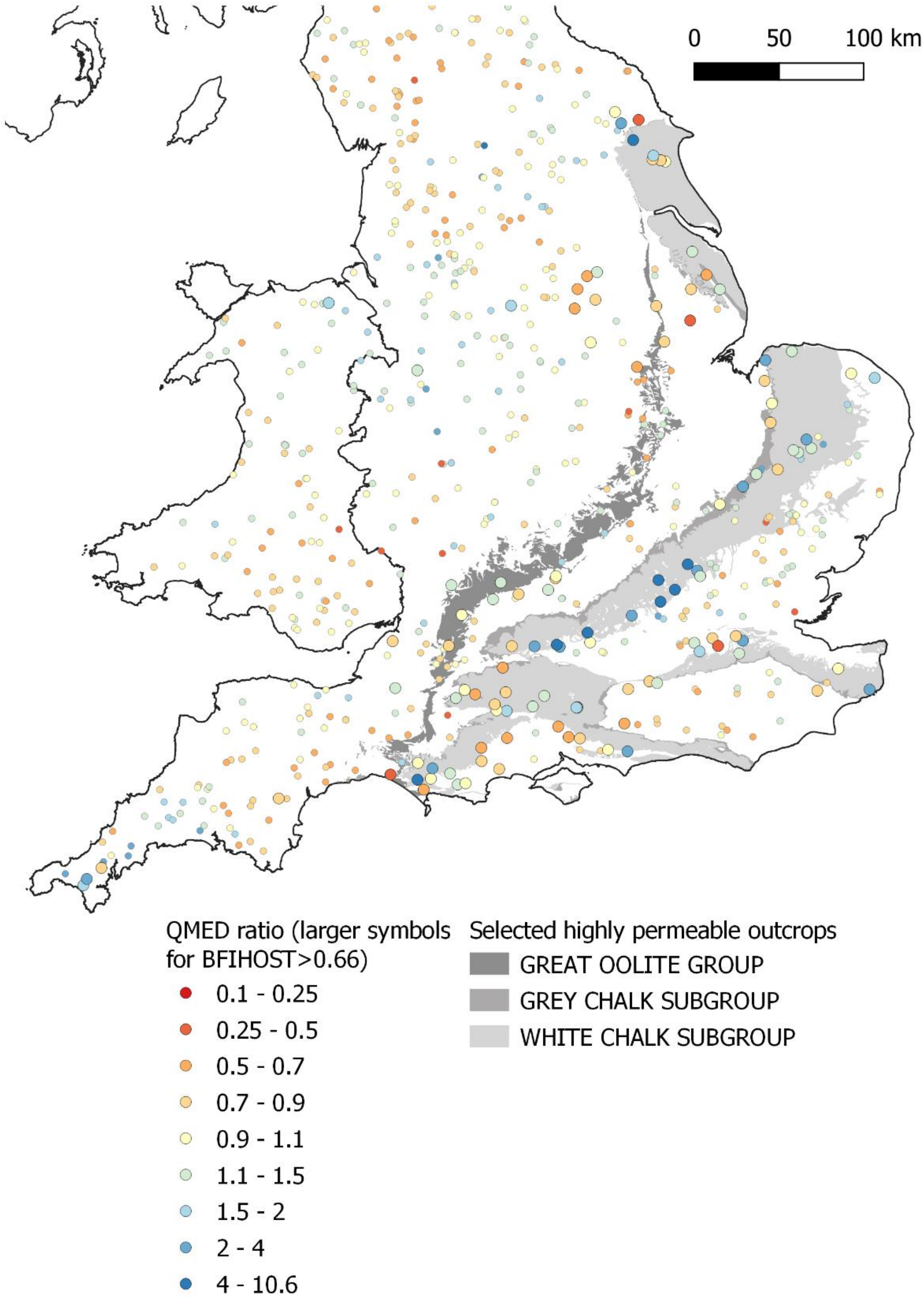


Figure 5-2: Map of proportional error in QMED, highlighing gauges on groundwater-dominated catchments. Results are plotted at gauging station locations, not catchment centroids. Contains British Geological Survey materials © UKRI 2022.

Potential reasons for overestimation of QMED on these catchments include:

- The topographic contributing area overestimating the hydrological catchment.
- The BFIHOST value being unrepresentative of the hydrological response of the catchment.

For example, at 39014 the baseflow index BFI derived from the flow record is 0.88, much higher than BFIHOST of 0.68. the situation is unlikely to change when BFIHOST19 is more widely adopted as an estimator of catchment BFI: it does not appear that BFIHOST19 values are much higher on some of the catchments.

The HOST characteristics are evaluated as spatial averages over the catchment and take no account of the spatial configuration of the soil associations, such as a tendency for highly permeable soils to be concentrated in river valleys, allowing runoff from neighbouring areas to easily percolate into the underlying aquifer. There is more to surface water-groundwater interactions than can be measured by BFIHOST.

In the regression for QMED, the estimated QMED is proportional to $0.046^{BFIHOST^2}$. An increase in BFIHOST from 0.68 to 0.88 would reduce QMED by 60%. This would substantially reduce the overestimation, although not eliminate it.

- Depression of the groundwater table due to abstraction, which is a feature of most if not all of the catchments listed.
- The parameterisation of the QMED equation. The form of the QMED equation may make it more sensitive to errors in BFI than other parameters.

In practice it should be possible to use data transfer to mitigate some of the overestimation of QMED at locations near to these gauges. The spatial clustering of errors in the upper Colne catchment and the Mimram helps with this, although in some other chalk areas such as Hampshire, large overestimation of QMED occurs not far from gauges that show significant underestimation (Figure 5-2).

Another potential route to improve the estimation of QMED at some locations is to use statistics from flow duration curves. Wallingford HydroSolutions (2016) provides an equation to estimate QMED from a combination of flow duration statistics, one catchment descriptor (DPSBAR) and BFI estimated from daily flow data. There is no dependence on BFIHOST. This can be applied at flow gauges that are not able to measure flood flows. When it is applied at the gauges listed in Table 5-1 it gives estimates of QMED that are much closer to those obtained from annual maximum flows.

Further work could look more systematically at relationships between the performance of the FEH QMED regression and catchment properties, in particular geological characteristics.

5.1.3 Pooling

The Flood Estimation Guidelines provide a history of the way in which soils and geology have been used to influence growth curve estimation. It is reproduced here for convenience.

Extract from Flood Estimation Guidelines (Environment Agency, 2022)

It is common for analysts to remove stations from a group due to large differences in BFIHOST. This may be because in the original version of the FEH method, BFIHOST was one of the three descriptors from which the pooling group was selected.

Subsequently, the SC050050 research found that BFIHOST had very little explanatory power for flood growth curves, with ten other catchment descriptors found to be more useful at explaining variation in the L-moment L-CV.

Earlier research, including the FEH and Flood Studies Supplementary Report 4 (1977), consistently reported differences in flood growth curves on permeable and nearby impermeable catchments. They report that there is generally less year-to-year variation on the permeable catchments and hence flatter growth curves.

Despite the findings of SC050050, there is a common perception that highly permeable catchments are likely to have different flood growth curves, perhaps showing greater skewness due to the occurrence of occasional floods that are many times higher than QMED, for example, during conditions in which the catchment acts more like an impermeable catchment … There do not seem to be many of these exceptional floods evident in the gauged period of record and so they may not be reflected in the sample L-moments.

Some more recent research also supports the idea that BFIHOST is worth considering when refining pooling groups (Formetta and others, 2018). The research tested an alternative approach to constructing pooling groups, using flood seasonality in conjunction with just one catchment descriptor, BFIHOST. This procedure was found to provide a more accurate estimate of the growth curve than the current FEH method. Although the new method is not currently recommended for implementation, the findings might act as an encouragement to refine groups with the aim of making them more representative of the geology and soils of the subject catchment. However, the effect on the pooled growth curve is likely to be small.

SC050050 refers to Kjeldsen et al. (2008), which sets out the current version of the FEH statistical method. It would be interesting to see whether the findings about explanatory variables for L-moments would be unchanged if the research was repeated with the contemporary peak flows dataset. The analysis summarised in Figure 4-1 indicates that there is a significant difference between the distributions of L-CV on high and low-BFIHOST catchments. On the face of it, this appears to contradict the findings of Kjeldsen et al. (2008). A possible explanation may be that another catchment descriptor, such as SAAR, is associated with BFIHOST and so is acting as a surrogate for it in the FEH pooling method. In section 6.2.2 we mention some research that has explored this further.

Hammond (2021) quantified the uncertainty of flood frequency curves estimated using the various FEH methods: ungauged pooling, single-site and pooling at a gauge (also known as enhanced single-site). In the latter two cases, QMED was estimated directly from annual maximum flows. Figure 5-3 shows the results for a return period of 100 years (AEP=1%), distinguishing between groundwater-dominated and other catchments. Note the different vertical scales.

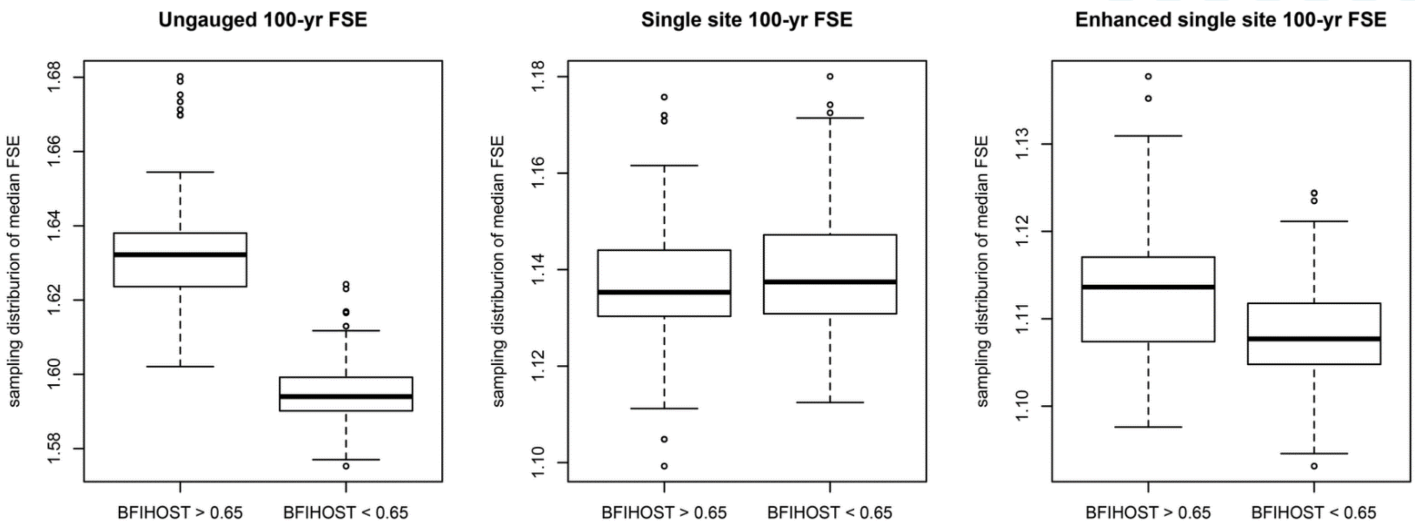


Figure 5-3: National distribution of factorial standard error (FSE) for the 100-year flood (AEP 1%) estimated using three methods, with catchments separated into two groups using a threshold of BFIHOST=0.65. Data source: NRFA gauges suitable for pooling. From Hammond (2021), used with permission.

For ungauged pooling the FSE is distinctly higher for BFIHOST>0.65. The highest FSE across the set of catchments with BFIHOST<0.65 is about equal to the lower quartile of FSE values for BFIHOST>0.65. This will be partly due to the increased FSE in QMED when estimated from catchment descriptors on groundwater-dominated catchments, but the pooling procedure is also expected to contribute to the increased error.

5.2 ReFH2 rainfall-runoff method

5.2.1 Suitability of a design event method

It is common to think of design event methods such as ReFH2 as less appropriate on groundwater-dominated catchments. This is the line taken in the Flood Estimation Guidelines, drawing on the findings of Bradford and Faulkner (1997) and also Webster (1999) who found that the relationship between the return periods of storms and floods became increasingly scattered for more permeable catchments, and concluded that permeable catchments are not really suitable for design flood analysis using an event-based method. More recently, Ledingham et al. (2019) have concluded that inadequate treatment of the seasonality of rainfall and soil moisture seriously reduces the reliability of event-based flood estimation in Britain, particularly so on dry lowland catchments where the seasonality of maximum daily rainfall and maximum flow differ greatly.

Floods on groundwater-dominated catchments tend to be influenced by combinations of processes, such as runoff from intense rainfall combined with high baseflow due to large recharge over the winter season. These are not represented adequately by the design event approach in which the river response depends mainly on rainfall during a single design storm.

Rainfall-runoff models rely directly on an accurate determination of the catchment area. As we have seen, the effective contributing area is not always well known on groundwater-dominated catchments and may vary with flood size.

Despite these theoretical deficiencies, the ReFH2 method appears to give reasonable peak flows for groundwater-dominated catchments, as discussed in the following section. Its performance is much improved over ReFH1. There has been less attention paid to evaluating aspects of ReFH2 apart from peak flows, and these are discussed in 5.2.3.

Unless otherwise stated, all outputs from ReFH2 are produced using v2.3 of the method, which is implemented in v3.1 of the ReFH2 software. The development of the design package for the ReFH2.3 method (i.e. the calibration of initial soil moisture) included 73 catchments with BFHOST>0.65.

5.2.2 Performance: peak flows

Wallingford HydroSolutions (2019a) provides an evaluation of the performance of ReFH2, distinguishing between catchments with BFIHOST above 0.65 and those below. The results in this report are from version 2.2 of the ReFH method. The current version of the method is 2.3, which includes a modification to close the water balance.

Wallingford HydroSolutions (2019a) compares peak flows from ReFH2.2 with flood frequency curves estimated using enhanced single-site analysis. The difference between the two is referred to as the error, and the ratio as the bias. The results show that a larger factorial standard error (FSE) on catchments with BFIHOST above 0.65: for an AEP of 50% the FSE is 1.52 on groundwater-dominated catchments compared with 1.43 for other catchments in England. For the 1% estimates the respective FSE values are 1.56 and 1.47.

For modest AEPs there is little or no bias in the ReFH2.2 results on either type of catchment, although for groundwater-dominated catchments there is a tendency for a positive bias at low AEPs (<1%).

Wallingford HydroSolutions (2019b) provides a less comprehensive evaluation of the performance of ReFH2.3. The general pattern of the performance is like that of v2.2. For estimating QMED, ReFH2.3 (like 2.2) appears to show a tendency towards larger estimation error as BFIHOST19 increases, indicated by the increased scatter on Figure 5-4.

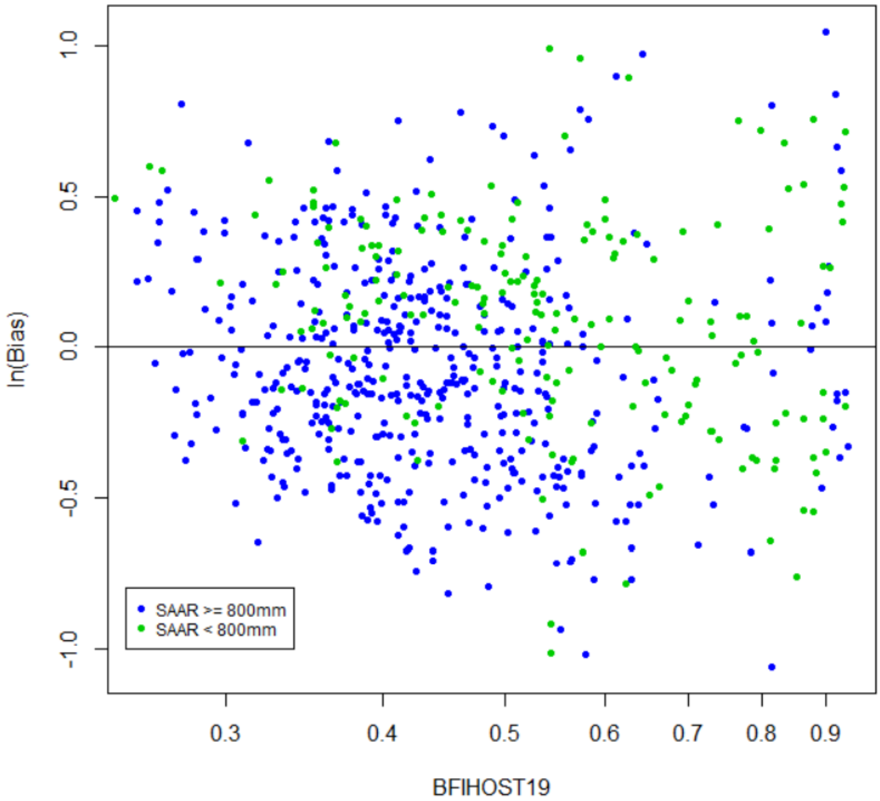


Figure 5-4: Residuals in QMED estimated from ReFH2.3 method using the water balance design package. This is Figure 8 from Wallingford HydroSolutions (2019b), reproduced with permission.

It is important to realise that the dataset plotted in Figure 5-4 excludes catchments where QMED is over- or underestimated by a factor exceeding 3. These were excluded from the process of calibrating the initial soil moisture Cini when ReFH2 was developed (Wallingford HydroSolutions, 2019c). 15 such catchments were excluded, classed as water balance violations. This explains why the highest and lowest values of ln(Bias) visible on the figure are about ±1, indicating over or under-estimation by a factor of e (2.78). It is not a fair comparison to consider these results as equivalent to the evaluation of the Statistical method in Figure 5-1.

As an illustrative example, we have run ReFH2 for the River Ver at Hansteads (gauge 39014). This is the gauge at which the Statistical method showed the largest overestimation of QMED (Table 5-1). Using the default model parameters, ReFH2 estimates a peak 2-year flow of 7.7m^3/s. This is comparable with the estimate from the Statistical method and is 6.3 times higher than the median of the annual maximum flows, 1.22m^3/s.

It is apparent that, despite the results plotted in Wallingford HydroSolutions (2019b), ReFH2.3 can show the same extreme overestimation of QMED as the Statistical method on a handful of groundwater-dominated catchments. This is to be expected since both methods are relying on the same set of catchment descriptors, none of which take into account features such as the disparity between topographic and hydrological catchments.

5.2.3 Performance: runoff volumes and hydrographs

Another way to judge the performance of ReFH2 is to examine the hydrographs and volumes of runoff that it simulates. As discussed in 4.6, it is not realistic to expect ReFH2 to produce realistic hydrographs on some baseflow-dominated catchments where high flows can persist for weeks or months, or where a rapid runoff response is superimposed on baseflow (although it may be possible to approximate this response by treating part of the catchment as urban, with a much shorter time to peak).

We have run ReFH2 on a notional groundwater-dominated catchment with area 130km^2, BFIHOST19 = 0.95 and SAAR = 700mm. There is some urbanisation in the catchment.

The design rainfall lasts for 22 hours, the direct runoff for 66 hours and the baseflow output continues for a total of 506 hours, or 21 days. It is worth noting that although the graphical output from ReFH2 terminates the hydrograph when the direct runoff ends (Figure 5-5), the tabular output can include a much longer period of runoff, until the baseflow returns to something approaching its initial value, BF_0.

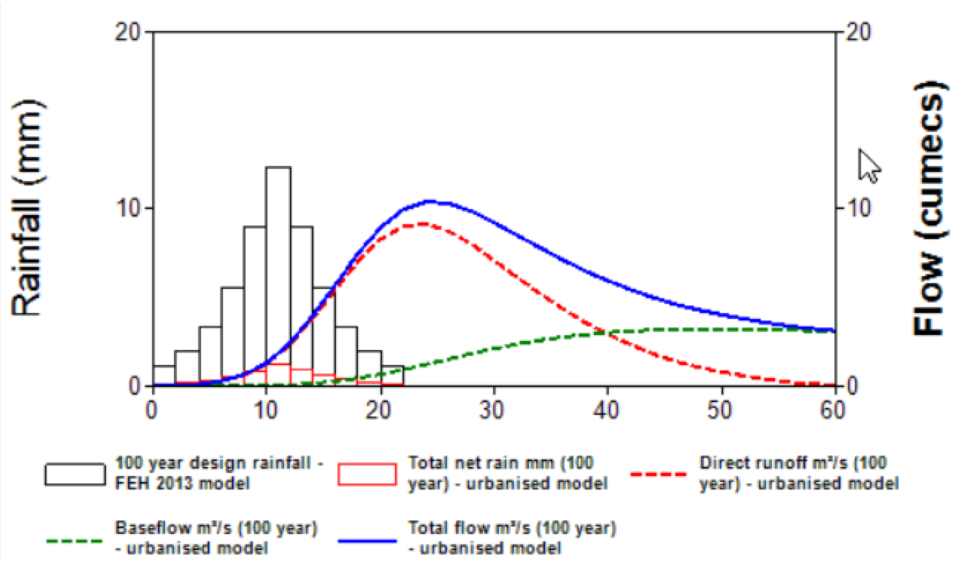


Figure 5-5: ReFH2 output for a catchment with very high BFIHOST19 (AEP 1%)

For this catchment the value of BF_0 is zero: in other words, there is no baseflow at the start of the flood event. In the winter season, the value of BF_0 is proportional to (Kjeldsen, 2007):

$$63.8\ (C_{ini} - 120.8) + 5.54SAAR$$

The initial soil moisture C_{ini} is 49mm and so the above quantity is negative, leading to a negative value for BF_0, which is reset to zero. It does not seem satisfactory for ReFH2.3 to simulate floods on a groundwater-dominated catchment with no baseflow at the start of the event (Figure 5-5). The simulated hydrograph will inevitably look nothing like most actual floods.

On catchments with high BFIHOST19 ReFH2 uses a very low value for C_{ini} (expressed as a ratio of the maximum soil moisture capacity, C_{max}). This arises from the calibration of the design event to match observed values of QMED which is explained in Wallingford HydroSolutions (2019c). In effect, the model is representing such catchments as having very little water stored in the unsaturated zone, even in winter conditions. While this could conceivably be realistic, it is not true of the amount of water stored in the saturated zone, and yet the initial baseflow is set on the basis of C_{ini}. There may be potential to remove this link and improve the estimation of BF_0.

Looking at the volumes of runoff from the simulation:

- Direct runoff as a percentage of rainfall increases from 9% at AEP>50% to 11% at AEP=0.1%
- Total flow as a percentage of rainfall increases from 24% at AEP>50% to 37% at AEP=0.1%.

Since the model does not represent evapotranspiration, the remaining water, which comprises by far the majority of the rainfall, is presumably going into storage and is not released, even 20 days after the rainfall ceases.

The percentage runoff, measured using direct runoff only, is clearly very low. It seems surprising that it barely increases with reducing flood probability.

5.3 Concluding comments on performance of FEH methods on groundwater-dominated catchments

Within the scope of this project it has not been possible to carry out a systematic comparison of the performance of FEH methods on groundwater-dominated catchments. There is some relevant published information in Wallingford HydroSolutions (2019a). It shows that, for catchments with BFIHOST>0.65, the FEH Statistical and ReFH2.2 methods, both applied as if catchments were ungauged, perform similarly for AEPs 50% and 3%. They show similar bias and FSE.

For smaller AEPs ReFH2.2 performs less well than the Statistical method, although this needs to be interpreted with caution because the comparison is with an enhanced single-site growth curve which is (a) inevitably similar to a pooled growth curve and (b) itself an estimate with increasing uncertainty as the exceedance probability reduces.

Both methods show an occasional tendency for extreme overestimation of peak flows on some groundwater-dominated catchments.

Our investigation has raised some concerns about the ability of the ReFH2 method to simulate realistic flood hydrographs and volumes in groundwater-dominated catchments. These may not be surprising given the comments in 5.2.1. There are two situations in which it is common to use the ReFH2 method:

- To provide a hydrograph shape, scaled to a peak flow estimated from the Statistical method.
- To provide a complete flood hydrograph, in settings where flood risk is influenced by the hydrograph volume rather than only its peak. This is a feature of many lowland watercourses, some of which have headwaters that rise in chalk or limestone areas, such as the Yorkshire and Lincolnshire Wolds.

Both of these applications rely on characteristics of the hydrograph other than its peak. For this reason it would seem wise to give more prominence to evaluating the performance of ReFH2 using aspects other than peak flows.

6 Alternative methods for flood frequency estimation in permeable catchments

6.1 Introduction

This chapter provides a wide-ranging list of suggestions for improvements and alternatives. They are not all necessarily recommended for pursuing. Some fairly modest exploratory analysis may help in identifying promising avenues for investigation.

The suggestions in this section are lettered sequentially to help with cross-referencing from section 7.3 which attempts to set out some prioritised recommendations.

6.2 Alternatives within the FEH family

6.2.1 Developments that could improve all methods

Catchment descriptor improvements

A. Consider using hydrological contributing area in place of topographic catchment.

   For some catchments the contributing area will vary with rainfall intensity or season, but it may be possible to at least estimate an expected contributing area at a 50% AEP, for estimating QMED. This information could be obtained from hydrogeological maps, groundwater models, cave surveys and local knowledge. It may be increasingly difficult to work out how the contributing area changes for more extreme floods.

B. Reconsider BFIHOST19

   Compare it with BFI on catchments that have high BFI (Figure 6-1) and seek to reduce the scatter. It may not be possible to improve on the work that has already been done.

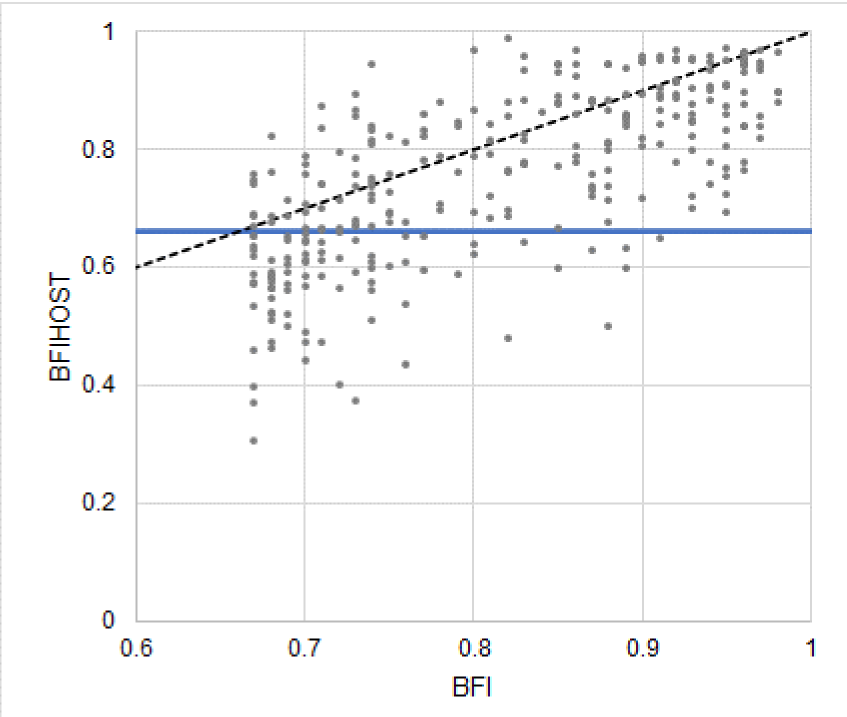


Figure 6-1: BFIHOST compared with BFI for catchments with BFI>0.66. Data source: NRFA.

C. Reconsider BFI as a measure of the contribution to river flow from storage.

Littlewood (2008) found that the estimation of BFI on an example catchment was highly sensitive to the choice of time step, changing by 40 percentage points. An alternative index, SFI (slow flow index) was less sensitive and showed more sign of convergence to a stable value as the time step decreased. SFI can be calculated when the IHACRES rainfall-runoff model is fitted (requiring flow, rainfall and temperature data) and so is not as straightforward to evaluate as BFI, which requires only flow data. Littlewood (2008) concluded that at least some of the uncertainty in statistical relationships between model parameters and catchment attributes is caused by this dependence on time step.

D. Consider alternative descriptors that allow for the spatial configuration of soils, geology, land use or topography.

   These might help to characterise the response of catchments with low-permeability drift cover on high ground but a close surface water - groundwater connection in river valleys (e.g. the River Ver, as shown in the map on page 25) or those with permeable headwaters and highly responsive lower catchments (e.g. the River Wandle in London).

Increasing the sample size

E. Use low-cost methods such as temporary flow gauges, developing ratings for level gauges or non-contact flow measurement technologies, to acquire flow series for a more gauges on groundwater-dominated catchments. This could usefully focus on areas where FEH methods for ungauged catchments show the largest errors. The extra data will be useful both for local studies and for improving generalised methods on ungauged catchments.

F. Consider including river flow data from chalk areas in Northern France to build up a larger sample of groundwater-dominated catchments.

6.2.2 Developments of the Statistical method

The main shortcomings of the Statistical method on groundwater-dominated catchments are:

- Bias and large standard error in QMED estimation, with occasional extreme over-estimation
- Limited ability to account for changes in processes (including expansion of contributing area and transition to overland runoff) with flood magnitude

To some extent these can already be accounted for by careful application of the current methods, following advice already available in the Flood Estimation Guidelines about understanding the geology and hydrogeology and paying particular attention to historical information. Not all practitioners or reviewers will pay adequate attention to this advice, and on ungauged catchments the options are more limited.

It would be worthwhile trialling some changes aimed at improving the performance of the Statistical method. These could include:

G. Improve the FEH QMED regression equation from Kjeldsen (2008).

   i. Seek to further understand the reasons for the extreme overestimation of QMED seen on some catchments and the differences in estimation errors between nearby catchments.

   ii. Consider some of the alternative catchment descriptors suggested above.

   iii. Seek to create an alternative QMED regression to be applied only on high-BFIHOST catchments.

   There have been previous attempts to improve QMED estimation on permeable catchments which show some promise. Two are mentioned below.

- A JBA internal R&D project in 2013 (unpublished) investigated improvements to flood estimation in permeable catchments. After testing a range of approaches, it fitted a revised QMED equation to a set of largely rural catchments with BFIHOST>0.80. The preferred model used three catchment descriptors: AREA, PROPWET and BFIHOST. The FSE was 1.60, an improvement on the FSE of 1.72 for the FEH equation when applied over the same set of catchments.
- Hammond (2017) fitted a regression model for QMED, using similar methods to those in Kjeldsen (2008), to 80 essentially rural catchments with BFIHOST>0.65. The model excludes FARL although adds URBEXT2000 as an explanatory variable, despite the maximum URBEXT in the dataset being 0.03. Over the dataset used in its calibration, the revised model shows a statistically significant lower FSE than the current FEH regression equation for QMED on rural catchments (1.60 compared with 1.70). It appears to avoid the occasional extreme overestimation of QMED seen in the FEH equation. This finding is similar to that from the earlier JBA work.

Both of these findings indicate that there is some potential for improving the estimation of QMED by focusing separately on groundwater-dominated catchments.

H. Develop specific guidance on data transfer on groundwater-dominated catchments

  i. This could consider topics such as where to ignore the default recommendations and how to choose between competing donors with very different adjustments. A systematic comparison of donor adjustment strategies focused only on groundwater-dominated catchments may help to reveal a preferable strategy for donor selection and weighting.

  ii. If an alternative QMED regression is developed, this should be accompanied by a revised spatial error model and hence revised guidance on weighting of donor adjustments.

I. Reconsider pooling group composition

  i. Reconsider allowing for BFIHOST19 in selecting pooling groups.

  Internal research by JBA in 2013 found no benefit from including BFIHOST in selecting pooling groups, consistent with Kjeldsen (2008). It concluded that any information that BFIHOST provides about the variability of the L-moments is also provided by the descriptors already used in pooling, due to the significant correlations between BFHOST and both FARL and FPEXT. This finding could be revisited but perhaps more worthwhile would be to focus specifically on groundwater-dominated catchments.

  ii. Explore the development of a specific pooling procedure for groundwater-dominated catchments, like that now available for small catchments.

  This was investigated by JBA in 2013, who found that quite different variables were included when the procedure used by Kjeldsen (2008) was repeated only for permeable catchments. In place of AREA and SAAR, for instance, DPSBAR, ALTBAR and/or PROPWET were significant. This may have been partly an artefact of the small sample size.

  One variant could be introducing a feature that encourages growth curves for all highly permeable catchments to make some allowance for the possibility of "monster" floods, even where there is no local memory or evidence of such an event. One way of achieving this was explored by JBA in 2013, using a single pooling group for all permeable catchments, with weights of the gauges varying according to the FEH similarity distance measure. The procedure was found to give a slight improvement in the pooled uncertainty measure in comparison with the FEH procedure.

A related possibility would be to introduce a way of accounting for structural differences in hydrogeology, which may affect the potential for catchments to generate occasional "monster" floods. For instance, the different characteristics of northern and southern provinces of the Chalk appear to affect the L-CV (Figure 3-1). The idea of separate regional analysis of data from chalk catchments in southern and eastern England was tested by Bradford and Faulkner (1997), but there was no attempt to quantify any reduction in estimation error.

J. Develop specific guidance on growth curve estimation on groundwater-dominated catchments

   i. Develop recommendations on where to give more priority to single-site growth curves and how to transfer information from single-site (or enhanced single-site) curves to nearby ungauged points.

   ii. Further explore the value of historical flood data and related sources of evidence such as long-term records of groundwater level.

   Much work has already been done on developing historical flood chronologies in groundwater-dominated catchments and incorporating the information into growth curve estimation, so a useful task could be to collate this information, identify any gaps that be filled and investigate any patterns between catchments. The "FEH Local" Technical Guidance 12_17 (Environment Agency, 2017) provides advice on how to incorporate groundwater level data in flood frequency estimation, including the importance of checking for the effects of changes in abstraction.

We have not included any recommendations for improvements to the procedure for removing the influence of non-flood years because we suggest this is a less important priority. Some suggestions are made in Section 4.4.6.

6.2.3 Developments of the design event method

The ReFH2 method shows similar shortcomings to those of the Statistical method for estimating peak flows on groundwater-dominated catchments, although comprehensive information about its performance is less readily available, for example due to the exclusion of catchments on which ReFH2 was found to give the poorest estimates of QMED (see Section 5.2.2). We recommend that the Statistical method continues to be regarded as the first choice for estimation of peak flow frequency on purely groundwater-dominated catchments, so the suggestions below deal with other aspects.

K. Evaluate the range of catchment geologies for which it is reasonable to expect ReFH2 to generate realistic hydrograph shapes and volumes.

L. Improve the estimation of the initial baseflow, BF_0

   It can currently take a value of zero on a baseflow-dominated catchment (see Section 5.2.3). This could include revisiting the decision to set BF_0 on the basis of the initial soil moisture C_{ini}.

M. Re-examine the volumetric performance on groundwater-dominated catchments.

   Is it realistic that on some such catchments the modelled percentage of direct runoff increases very little with reducing probability of exceedance, and that the majority of the rainfall goes into long-term storage, not to be released as modelled baseflow even after several weeks?

N. Develop guidance on application of design event methods such as ReFH2 to catchments with mixed geology or land use, or those in which rapid runoff from riparian areas is superimposed on a much slower baseflow response.

It would be particularly valuable to focus on applications where volumes and durations are important considerations, for example the design of flood storage or pumping arrangements. Example catchments for tests might include the River Darent in Kent (mixed geology and land use), the Bentley Ings Drain in South Yorkshire (limestone headwaters draining to low-lying clay with some urbanisation), the River Wey in Dorset (case study on page 27) or catchments which run off the Lincolnshire Wolds into the River Ancholme or the Humber Estuary (chalk headwaters feeding water into low-level systems managed by Internal Drainage Boards).

6.3 Other approaches

There are other approaches to flood frequency estimation that fall outside the FEH family of methods. This section mentions several that have been applied or considered for groundwater-dominated catchments.

6.3.1 Continuous simulation

Continuous simulation (CS) has been applied quite widely on groundwater-dominated catchments in the UK. Two advantages relevant to such catchments are that there is no requirement to separate flood and base flows, and no requirement to link rainfall and flood AEPs (Calver et al, 2000).

In the 1990s MAFF commissioned research into CS modelling of floods from combined surface and subsurface sources (Calver et al., 2000). It tested various rainfall-runoff models and a model of groundwater levels and concluded on a positive note about the prospects of reproducing flood frequency characteristics, even in ungauged catchments and for low AEPs. The research fed into a national-scale assessment of flood estimation by continuous simulation (Calver et al., 2005), which has seen little if any subsequent implementation.

Examples of flood studies carried out for the Environment Agency applying continuous simulation on catchments with a major groundwater influence include:

- River Frome, an Oolitic Limestone catchment (JBA Consulting in 1998, updated in 2016). The rainfall-runoff model HYSIM was used to extend the river flow series back to 1923 using observed rainfall as input.
- River Kennet, a chalk catchment in Berkshire for which Thames Region of the Environment Agency commissioned research into CS. The model used stochastically generated daily rainfall as an input and the project paid attention to frozen ground and snowmelt effect. It concluded that CS showed promise but more work was needed to derive convincing stochastic rainfall inputs (Halcrow in 2006).
- Bentley Ings, a small catchment with limestone headwaters, an urban mid-catchment and a lower pumped catchment. CS was used in the design of a replacement pumping station here (JBA Consulting in 2010, featured in a paper by Lamb et al., 2016).
- Seven chalk catchments in Hampshire (JBA Consulting in 2015). The project concluded that CS cannot be recommended outright for the case study catchments, at least not without modifying the (point) stochastic rainfall model to simulate realistic variability of depths over several seasons. Where there is a long observed series containing some extreme events, the project recommended that single-site analysis seems the most robust method of estimating a flood frequency curve, and that this should be combined with an assessment of historical flooding, with particular attention paid to any mention of extreme events, such as those due to rainfall on frozen soil.

- Salisbury, which lies at the confluence of four chalk rivers (JBA Consulting in 2018). Like the Hampshire project, this also applied a lumped stochastic rainfall model. The model output needed to be scaled up to match observed rainfall over accumulation periods of 2 months and 4 months, the latter being the approximate critical duration for winter recharge leading to flood flows. CS was found to work well on this project, giving good agreement with observed annual maximum flows (e.g. Figure 6-2). Limitations included the lack of ability to estimate the probability of frozen ground and snowmelt.

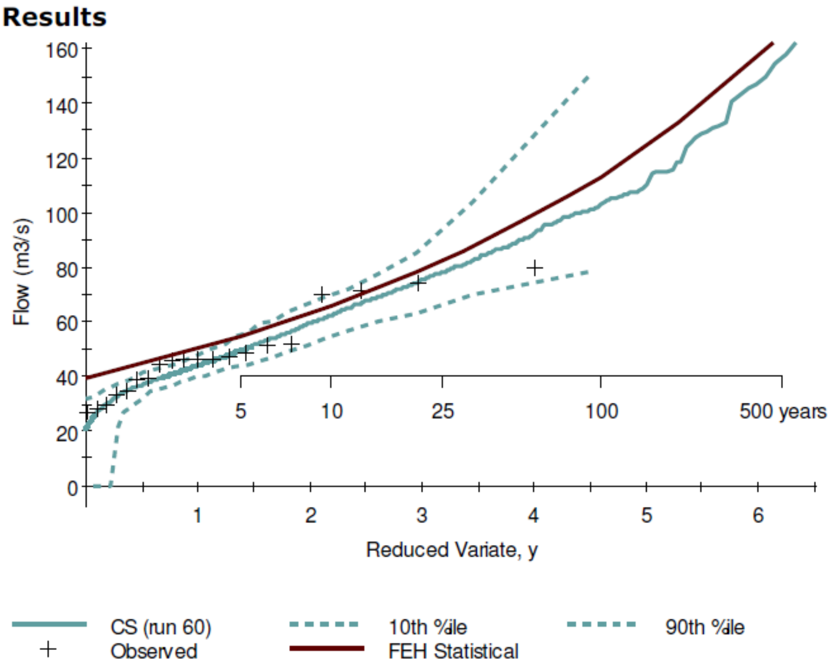


Figure 6-2: Flood frequency curves from CS and FEH for the Avon at Salisbury, a chalk catchment

- River Darent, a catchment with complex heterogeneous geology and land cover (JBA Consulting in 2018).
- Thames catchment flood storage, applying a spatio-temporal stochastic rainfall to produce spatially consistent simulations of river flow for numerous sub-catchments of the Thames, testing the effect of proposed distributed flood storage measures (JBA Consulting and Jacobs, ongoing). As in the Salisbury study, the rainfall outputs needed adjusting before they gave a realistic match to observed rainfall over the long accumulation periods that are associated with flooding on the Thames.

Research is under way to address some of the limitations of CS identified in the projects listed above. Newcastle University and JBA are collaborating to improve the performance of a spatio-temporal rainfall model over long accumulation periods, and to develop a user-friendly tool to implement it.

On catchments with a flood response entirely dominated by groundwater, the benefits of CS are arguable. The modelled flood growth curve tends to match the rainfall growth curve over the critical duration of the catchment. The potential benefits are clearer in situations where a flood hazard can arise from a variety of combinations that cannot be confidently represented with a single design event. This can be the case on catchments with mixed geology, or highly permeable catchments that contain urban areas from which rapid runoff occurs. CS is also a powerful technique for assessing options for flood storage because it provides a much more rigorous test of options than a single design flood event.

We have two suggestions for improvements in addition to the R&D already underway on rainfall modelling:

O. Develop some case studies or other ways to illustrate the value of continuous simulation on catchments with a mixed flood response.

P. Investigate how to incorporate frozen ground and/or snowmelt into continuous simulation modelling.

6.3.2 Groundwater models

Numerical or conceptual models have been developed for many major aquifer units in the UK. They have typically been run for periods of 40 years or more, at a daily time step.

It may be possible to make more use of their outputs to aid with fluvial flood estimation, for example on ephemeral streams or other watercourses without continuous flow measurements. Possibilities would include using a groundwater model to estimate percentiles on the flow duration curve, from which QMED could be estimated, or to improve the way in which flood frequency information is transferred from gauged to ungauged sites, by understanding similarities and differences in the flow regimes.

6.3.3 Joint probability analysis

Another approach to dealing with mixed flood responses is to attempt to fit a marginal statistical distribution to variables that characterise the different responses, before combining the distributions to represent the statistics of the overall river flow. For example, separate distributions might be fitted to extreme values of:

- Baseflow and rapid runoff
- Groundwater level and short-term rainfall intensity
- Hydrologically effective rainfall and short-term rainfall intensity
- River flow volume and peak river flow

Most of the above pairs of variables are likely to show some dependence and it is necessary to allow for this when combining the probability distributions. A complicating factor is that the degree of dependence may vary with the event magnitude. This can be handled by statistical methods such as the Heffernan and Tawn model, which is widely used in flood risk management (e.g. Keef et al., 2011) or by copulas.

An example of a joint probability investigation on a groundwater-dominated catchment is a flood study for the River Misbourne carried out by JBA in 2015 for the Environment Agency. The premise of the method was that long duration river response (controlled by baseflow contributions during periods of elevated groundwater levels) and short duration flood responses are separately and explicitly considered. Statistical analysis was applied to annual maximum series derived from the separated records. A hydraulic model of the river and floodplain system was then run with:

- Long-duration design hydrographs only
- Short-duration responses only
- Both responses concurrently

After some experimentation, separate extents for the long and short-duration design events were overlapped to produce a single outline that reflects both mechanisms. This assumed complete dependence between the occurrence of both responses, i.e. a long-duration response of AEP *x* combines with a short-duration response of AEP *x* to give a flood impact also of AEP *x*. This may not necessarily be a valid assumption, but was adopted as a conservative compromise in that analysis. The multivariate event modeller software (MEM) provides a tool for investigating joint probability problems and can fit the Heffernan and Tawn model. This leads to a suggestion:

Q. Explore the application of multivariate methods to flood frequency analysis on a catchment with a mixed flood response.

6.3.4 Sewer modelling of karst

Karst catchments can show a dual flood response mechanism where the behaviour of the catchment can switch to overland flow in rare occasions such as at Cheddar in 1968 (3.7.1). It may be challenging to represent this switch in a model, unless the karst system is represented hydraulically.

A precedent for this is a modelling study of the karst around Gort in County Galway, Ireland (Jennings O'Donovan & Partners and Southern Water Global, 1998; Lees, 1998). The karst was modelled in detail using three components: a stochastic rainfall model, a rainfall-runoff and recharge model and a karst groundwater model to predict the water level in turloughs. These were thought to comprise the first ever regional numerical model of a karst system. Karst conduits were represented using urban drainage modelling software, HYDROWORKS. Pipe diameters and gradients were modified during model calibration.

This approach would be feasible on some other karst catchments where there is enough calibration data available. Its cost-effectiveness would need to be justified. A simpler approach would be to represent the karst system using a notional maximum capacity, for example with the functionality in ReFH2 that allows export of water from a catchment via sewers.

6.3.5 Using evidence from geomorphology

Most methods for flood estimation on ungauged catchments rely on a knowledge of the catchment area. This is not always easy to define on groundwater-dominated catchments, particularly where watercourses originate from springs.

An alternative approach would be to use geomorphological evidence such as the channel size. Unfortunately this does not seem promising for groundwater-dominated catchments. The FEH Local research (Dixon et al., 2017) developed an improved regression for estimating QMED on the basis of bankfull channel width. Some large prediction errors were found on highly permeable catchments. It also developed an alternative regression that combines channel width with catchment descriptors. The report extract below gives a commentary.

> "There may be a tendency for chalk or limestone streams to have different geometric characteristics, perhaps being shallower and wider, or to have a channel-forming flood flow higher than QMED. According to Harvey (1969), the recurrence interval of the bankfull discharge on a baseflow-dominated stream, the Wallop Brook in Hampshire, was considerably higher than 2 years, being between 5 and 10 years at most sites. One explanation put forward was that, on a baseflow stream, the annual flood may not be competent to cause scour of the banks."
>
> Source: Dixon et al., 2017.

Another complicating factor may be that significant flow could be conveyed below the channel bed.

Despite these drawbacks, there are cases on groundwater-dominated catchments where evidence from channel size can be useful in either increasing confidence in a flood estimate made by another method or in identifying the need for further work. An example of the former is given in a case study for Slaplake Brook, Devon, in Technical Guidance 12_17 (Environment Agency, 2017). When QMED was estimated from the equation that incorporates both channel width and catchment descriptors, the result was identical to an estimate made only from catchment descriptors.

7 Conclusions and recommendations

7.1 Conclusions

We propose a threshold of BFIHOST or BFIHOST19 > 0.66 for initially identifying groundwater-dominated catchments. Within this range, the contribution to runoff from baseflow is at least twice that from more rapid runoff. It is important to recognise that virtually all catchments show some groundwater influence on flooding. The degree of influence can depend critically on the spatial configuration of the geology, which is not represented by catchment-average measures such as the FEH descriptors. For instance, even a small outcrop of limestone in a river valley can include swallowholes or cave entrances that divert some or all of the flow of the river away from the topographic catchment, drastically affecting downstream peak flows.

Many of the findings from this project reinforce what is already widely known, or suspected, about flood frequency estimation in groundwater-dominated catchments. It can be difficult and uncertain. On the other side of the coin, it can provide an opportunity for hydrological skills to shine, which can make for highly fulfilling work.

It is interesting to compare the findings with those reported from a project with similar aims that was commissioned 25 years ago (Bradford and Faulkner, 1997), along with earlier work (Institute of Hydrology, 1977). The extra quarter century or more of hydrological records has confirmed some ideas but cast some doubt on others. For instance, the adjustment of flood growth curves to remove the influence of non-flood years now has little effect on the results at the vast majority of sites. Flood statistics on groundwater-dominated catchments, once thought to show below-average variability, now show a statistically significantly higher L-CV than those on other catchments. This is mainly a result of the extensive large floods that have occurred in the years since 2000.

There are many examples of good practice in flood studies that seek to develop a good understanding of the geology and the hydrogeological processes and to exploit local data sources to their fullest. There are some types of catchments or studies in which it is more difficult to be confident in how to overcome the challenges described in this report. These include ungauged locations, catchments with mixed geology or land use and studies in which flood volumes need to be quantified accurately, such as where chalk streams flow into lowland pumped catchments.

There is potential to make progress with reducing some of the uncertainties and improving guidance, as outlined below.

7.2 Recommendations for current practice

The Flood Estimation Guidelines already give some recommendations for flood estimation on groundwater-dominated catchments. There is nothing from the findings of this project that indicates a need to fundamentally modify those recommendations. The first and most important step should remain as developing an understanding of the geology, hydrogeology and flow processes in flood conditions.

The suggestions in this section are numbered sequentially to help with cross-referencing.

The findings in section 5.1.2 indicate the need to take great care when estimating QMED at ungauged locations with high BFIHOST19. While some further investigation would be needed to develop definitive recommendations, we suggest:

1. Avoiding choosing donors on catchments with much lower BFIHOST19
2. Looking at QMED adjustment factors on several surrounding catchments with similar geology
3. Using the following types of information to help guide the choice of donor(s):
    a. Groundwater flow directions, and differences between topographic and hydrological catchments. This information is available from sources including

the NRFA and BGS geological[13] and hydrogeological maps[14] and reports on groundwater modelling studies.

   b. Whether streams are perched above the water table or not.
   c. Spatial configuration of geology, such as whether the valley bottoms are drift-free.
   d. Differences between BFI and BFIHOST19. This information should be available even at flow gauges without suitable annual maximum flow data.
   e. Consistency of QMED adjustment factors on surrounding catchments with similar geology.

4. Seeking additional local sources of flow data, such as gauges that provide flow duration curve statistics, from which QMED can be estimated. These could be treated as additional candidate donor stations.
5. Using judgement to decide how and whether to apply the distance moderation factor for data transfer.

This study has provided some more evidence of the tendency for groundwater-dominated catchments to have greater variability in flood magnitudes (and therefore, most likely, steeper growth curves), particularly in the structurally more complex chalk aquifers in southern counties of England. Without further research it does not seem appropriate to make any firm recommendations for a change in practice for estimation of flood growth curves. Some suggestions for practitioners to consider are:

6. Giving more priority than usual to local flow and river level data, for example relaxing the data quality criteria for a gauge to be classed as suitable for fitting flood growth curves.
7. Valuing historical information, pre-dating the gauged record, even more highly than on less permeable catchments (this is already stated in the Guidelines).
8. Comparing single-site growth curves (augmented with historical data) with pooled curves and consider which is more consistent with the evidence contained in this report and elsewhere.
9. Considering modifying pooling groups to give priority to other groundwater-dominated catchments, particularly if they are in the same aquifer unit. This may only be beneficial if it adds some longer records to the pooling group, otherwise it may lead to information being replicated because flood peak series from within the same aquifer are expected to be highly correlated.

Other suggestions focus on the unique challenges that hydrograph shapes and volumes present on groundwater dominated catchments:

10. Understand the degree to which the project outcomes will be sensitive to flood volume. In a flood mapping study on a baseflow-dominated catchment it might be suitable to ignore flood volumes and run a hydraulic model in steady state. On the other hand, for examining flood storage options or for a catchment in which permeable headwaters flow into a low-lying river, volumes could be critical.
11. Investigate the relative roles of baseflow and rapid runoff and how they combine. Consider separate frequency analysis of these two types of response, using a multivariate statistical model to allow for the dependence between their marginal (individual) distributions.
12. Consider whether an empirical analysis of hydrograph shapes might give a more realistic design flood hydrograph than the ReFH2 method, which can produce

[13] https://webapps.bgs.ac.uk/data/maps/maps.cfc?method=listResults&mapName=&series=E50k&scale=&pageSize=100

[14] https://www2.bgs.ac.uk/groundwater/datainfo/hydromaps/hydro_maps_scanviewer.html

hydrographs with an initial baseflow of zero on some catchments where in reality baseflow is dominant during floods.

Finally, on catchments where both highly permeable and less permeable (or urban) areas contribute significantly to flood runoff, or when testing options for flood storage on any catchments, we suggest:

13. Consider continuous simulation as an alternative to FEH methods.

7.3 Recommendations for research

We have identified many possible avenues for further investigation in Section 6. Some of these are more promising, and more likely to be cost-effective, than others. We suggest that they could be grouped into four work packages. We have included some indicative cost estimates as requested, but it is worth considering that useful progress could be made on some of these tasks via routes that do not cost the Environment Agency anything, such as MSc dissertation projects.

Any new methods or changes in methods will need to be robust to the challenges, change and uncertainties that climate change may bring. On groundwater-dominated catchments the impact of climate change may be different to elsewhere, with more dependence on winter rainfall and soil moisture deficits.

The labelling of the tasks within each work package below cross-references to the headings in Section 6, which provides more detail about each task.

1. Explorations and tests on a small sample of catchments (~£30K)

   A suitable starting point might be to look at a handful of gauged catchments where the FEH ungauged catchments methods give particularly poor estimates of flood frequency. Investigations on these catchments might lead to useful insights that could be missed by analysis across larger datasets in which they could be dismissed as aberrations or outliers. This could help to set some useful foundations for a wider analysis.

   The investigations could include the following:

   A. Consider using hydrological contributing area in place of topographic catchment.

   C. Reconsider BFI as a measure of the contribution to river flow from storage.

   D. Consider alternative descriptors that allow for the spatial configuration of soils, geology, land use or topography

   G.i. Seek to further understand the reasons for the extreme overestimation of QMED seen on some catchments and the differences in estimation errors between nearby catchments.

   Q. Explore the application of multivariate methods to flood frequency analysis on a catchment with a mixed flood response.

   It may be worth seeking data from flow gauges outside the NRFA peak flows dataset, using findings from reviews of flow derivation methods. Even uncertain flow measurements might provide useful information.

2. Analysis across all groundwater-dominated catchments (~£35K)

   While it would be possible to start this straightaway, it might be wise to leave it until after the explorations listed above. The scope would need to be refined at that point, but potential tasks include:

   F. Consider including river flow data from chalk areas in Northern France to build up a larger sample of groundwater-dominated catchments.

   G.iii. Seek to create an alternative QMED regression to be applied only on high-BFIHOST catchments.

H. Develop specific guidance on data transfer on groundwater-dominated catchments.

I.ii. Explore the development of a specific pooling procedure for groundwater-dominated catchments, like that now available for small catchments.

J.i Develop recommendations on where to give more priority to single-site growth curves and how to transfer information from single-site (or enhanced single-site) curves to nearby ungauged points.

3. Tests and refinement of commercial methods

   Currently in the UK the industry standard version of the design event method is developed commercially rather than with public funding. If this funding model is to continue, it would be up to Wallingford HydroSolutions to consider whether and how to implement the three suggestions that this report makes that are specific to the ReFH2 method. These are:

   K. Evaluate the range of catchment geologies for which it is reasonable to expect ReFH2 to generate realistic hydrograph shapes and volumes.

   L. Improve the estimation of the initial baseflow, BF_0.

   M. Re-examine the volumetric performance on groundwater-dominated catchments.

4. Collation of existing information and development of guidance (~£20K)

   These could be done at any stage because they do not depend on the development of new or improved methods.

   J.ii. Further explore the value of historical flood data and related sources of evidence such as long-term records of groundwater level.

   N. Develop guidance on application of design event methods such as ReFH2 to catchments with mixed geology or land use, or those in which rapid runoff from riparian areas is superimposed on a much slower baseflow response.

   O. Develop some case studies or other ways to illustrate the value of continuous simulation on catchments with a mixed flood response.

References

Alberta Transportation (2001). Guidelines on flood frequency analysis. Alberta Transportation, Transportation and Civil Engineering Division, Civil Projects Branch. Alberta, Canada.

Allen, D.J., L.J. Brewerton, L.M. Coleby, B.R. Gibbs, M.A. Lewis, A.M. MacDonald, S.J. Wagstaff and A.T. Williams (1997). The physical properties of major aquifers in England and Wales. British Geological Survey.

Archer, D., Foster, M., Faulkner, D. and Mawdsley, J. (2000). The synthesis of design flood hydrographs. In: Flooding Risks and Reactions. Proceedings of the Water Environment 2000 Conference, 5 October 2000. Institution of Civil Engineers, London, pp. 45-57.

Archer, D., O'Donnell, G., Lamb, R., Warren, S., Fowler, H.J. (2019). Historical flash floods in England: New regional chronologies and database. J. Flood Risk Management 12 (Suppl. 1): e12526. https://doi.org/10.1111/jfr3.12526.

Ball, J., Babister, M., Nathan, R., Weeks, W., Weinmann, E., Retallick M. and Testoni, I. (Editors) (2019). Australian Rainfall and Runoff: A Guide to Flood Estimation, © Commonwealth of Australia (Geoscience Australia).

Boorman, D.B., Hollis, J.M. and Lilly, A. (1995) Hydrology of Soil Types. A Hydrologically-Based Classification of the Soils of the United Kingdom. Institute of Hydrology Report 126.

Bradford (2002) Volume-duration growth curves for flood estimation in permeable catchments. Hydrology and Earth System Sci., 6(5), 935-947.

Bradford, R.B. and Faulkner, D.S. (1997). Review of Floods and Flood Frequency Estimation in Permeable Catchments. MAFF R&D Project FD0423, Institute of Hydrology, Wallingford.

Bradford, R.B. and Goodsell, G. (2000). Flood Volumes and Durations in Permeable Catchments. Report FD1605 for MAFF by CEH Wallingford.

Calver, A., Crewett, Davies, Lamb and Crooks (2000). Modelling floods from combined surface and subsurface sources. Institute of Hydrology Report to MAFF.

Calver, A., Crooks, S, Jones, D, Kay, A, Kjeldsen, T, Reynard, N. (2005). National river catchment flood frequency method using continuous simulation. Technical Report FD2106/TR. Department for Environment, Food and Rural Affairs.

Dixon, H., Faulkner, D., Fry, M., Kral, F., Lamb, R., Macklin, M., Vesuviano, G. (2017). Making better use of local data in flood frequency estimation: Report – SC130009/R. Environment Agency.

England, J.F., Jr., Cohn, T.A., Faber, B.A., Stedinger, J.R., Thomas, W.O., Jr., Veilleux, A.G., Kiang, J.E., and Mason, R.R., Jr., (2019). Guidelines for determining flood flow frequency—Bulletin 17C. U.S. Geological Survey Techniques and Methods, book 4, chap. B5, 148 p., https://doi.org/10.3133/tm4B5.

Environment Agency (2017). Technical Guidance 12_17: Using local data to reduce uncertainty in flood frequency estimation.

Environment Agency (2022). Flood estimation guidelines. Technical guidance 197 08, issue 8.

Faulkner, D.S. and Robson, A.J. (1999). Estimating floods in permeable drainage basins. In: Hydrological Extremes: Understanding, Predicting, Mitigating. IAHS Publ. No. 255, pp. 245-250.

Faulkner, D., Griffin, A., Hannaford, J., Sharkey, P., Warren, S., Shelton, K., Vesuviano, G., Mastrantonas, N. and Stewart, L. (2020). Development of interim national guidance on non-stationary fluvial flood frequency estimation – science report. FRS18087/IG/R1. Environment Agency.

Faulkner, D, Warren, S, Spencer, P and Sharkey, P. (2019). Can we still predict the future from the past? Implementing non-stationary flood frequency analysis in the UK. Journal of Flood Risk Management https://doi.org/10.1111/jfr3.12582.

Finch, J.W., Bradford, R.B and Hudson, J.A. (2004). The spatial distribution of groundwater flooding in a chalk catchment in southern England. Hydrol. Process. 18, 959–971.

Formetta, G., Bell, V. and Stewart, E. (2018). Use of Flood Seasonality in Pooling-Group Formation and Quantile Estimation: An Application in Great Britain. Water Resour. Res. https://doi.org/10.1002/2017WR021623.

Formetta, G., Griffin, A., Haxton, T., Stewart, E. and Young, A. (2022). Estimating design hydrographs in small catchments. Report – SC090031/R6. Environment Agency (in press).

Griffiths J., A. Binley, N. Crook, J. Nutter, A. Young, S. Fletcher (2006). Streamflow generation in the Pang and Lambourn catchments, Berkshire, UK. J. Hydrology, 330, 71-83.

Hammond, A. (2020). UKFE: UK Flood Estimation. R Package Version 0.1.3. The Comprehensive R Archive Network, Vienna.

Hammond, A. (2021). Sampling uncertainty of UK design flood estimation. Hydrology Research 52 (6): 1357–1371. doi: https://doi.org/10.2166/nh.2021.059.

Institute of Hydrology (1977). Flood Studies Supplementary Report 4. IH, Wallingford.

Institute of Hydrology (1999). Flood Estimation Handbook. IH, Wallingford.

JBA Trust (2016). Emergence of an ephemeral chalk stream at Assendon, Oxfordshire, in 2014. Available from https://www.jbatrust.org/how-we-help/publications-resources/.

Jennings O'Donovan & Partners and Southern Water Global (1998). An Investigation of the Flooding Problems in the Gort-Ardrahan Area of South Galway. Report to the Office of Public Works.

Keef, C., Rob Lamb, Jonathan Tawn, Paul Dunning, Crispian Batstone and Mark Lawless (2011). The risk of widespread flooding – Capturing spatial patterns in flood risk from rivers and coasts. Science Report SC060088/R1, Environment Agency.

Kjeldsen, T.R., Jones, D. A. and Bayliss, A.C. (2008). Improving the FEH statistical procedures for flood frequency estimation. Science Report SC050050, Environment Agency.

Ledingham, J., David Archer, Elizabeth Lewis, Hayley Fowler and Chris Kilsby (2019). Contrasting seasonality of storm rainfall and flood runoff in the UK and some implications for rainfall-runoff methods of flood estimation. Hydrology Research 50 (5): 1309–1323. doi: https://doi.org/10.2166/nh.2019.040.

Lees, M.J., Price, N., Wheater, H.S. and Peach, D (1998). A rainfall-runoff simulation model for South Galway, Ireland. Hydrology in a Changing Environment. Volume III. British Hydrological Society.

Littlewood, I. G. (2008). Characterisation of river flow regimes for environmental and engineering hydrology: unit hydrographs for rainfall-streamflow modelling. Folia Geographica: Series Geographica-Physica 39: 5-36.

Maurice, L.D., Atkinson, T.C., Barker, J.A., Bloomfield, J., Farrant, A., Williams, A. (2006). Karstic behaviour of groundwater in the English Chalk. Journal of Hydrology, 330 (1-2). 63-70. https://doi.org/10.1016/j.jhydrol.2006.04.012.

Wallingford HydroSolutions (2016). WINFAP 4 QMED Linking equation.

Wallingford HydroSolutions (2019a). ReFH2 Science Report: Evaluation of the Rural Design Event Model

Wallingford HydroSolutions (2019b). ReFH2 Science Report: Closing a Water Balance.

Wallingford HydroSolutions (2019c). ReFH2 Science Report: Model Parameters and Initial Conditions for Ungauged Catchments.

Warren, S. and Longfield, S. (2020). Development of interim national guidance on non-stationary fluvial flood frequency estimation – R package nonstat user guide. FRS18087/IG/R3. Environment Agency.

Webster, P. (1999). Factors affecting the relationship between the frequency of a flood and its causative rainfall, in Hydrological Extremes: Understanding, Predicting, Mitigating (Proc. IUGG99 Symposium HS1, Birmingham, July 1999). IAHS Publ. No. 255.

Offices at

Coleshill
Doncaster
Dublin
Edinburgh
Exeter
Haywards Heath
Isle of Man
Limerick
Newcastle upon Tyne
Newport
Peterborough
Saltaire
Skipton
Tadcaster
Thirsk
Wallingford
Warrington

Registered Office
1 Broughton Park
Old Lane North
Broughton
SKIPTON
North Yorkshire
BD23 3FD
United Kingdom

+44(0)1756 799919
info@jbaconsulting.com
www.jbaconsulting.com
Follow us:

Jeremy Benn Associates Limited

Registered in England 3246693

JBA Group Ltd is certified to:
ISO 9001:2015
ISO 14001:2015
ISO 27001:2013
ISO 45001:2018