Reassessing English flood frequency relationships in light of widespread new instrumental stage maxima

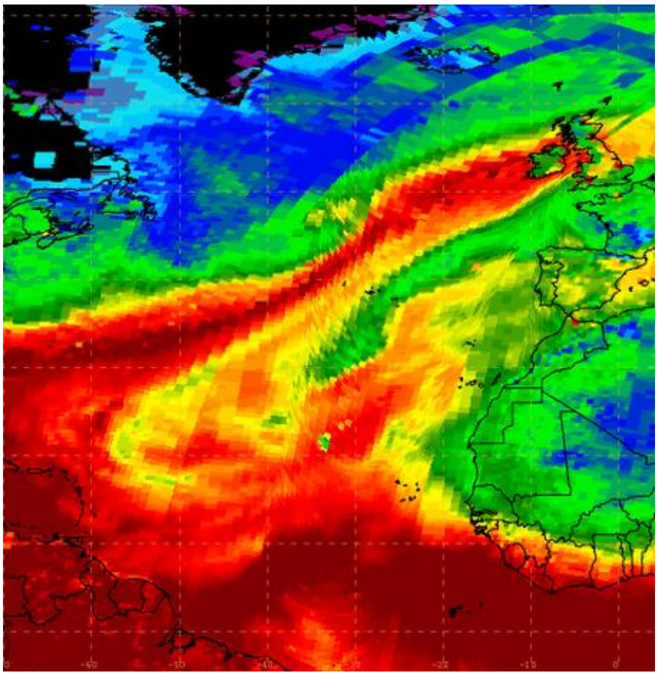
JAMES MATTHEW THORNTON

Reassessing English flood frequency relationships in light of widespread new instrumental stage maxima

James M. Thornton

Flood frequency estimation typically involves applying well-established statistical methods to instrumental peak discharge data. However, even notwithstanding any potential climate change trends, such analyses may have a tendency towards underestimation. The shortness of the available records is the proposed cause, whilst the prominence of flood frequency results in practical applications heightens the concern. Following the passage of the extra-tropical storms Desmond and Eva in December 2015, previous river stage maxima were widely surpassed across northern England. Indeed, preliminary analyses indicated that peak flows on the Rivers Eden, Tyne and Lune were higher than any previously recorded in England and Wales (CEH, 2016). Herein, the effects of including these exceptional observations on flood frequency estimates produced using established methods are investigated. Annual maxima series were extended at 155 stations, and models fitted on a single-site basis with and without the additional data were compared. Predictably, return period flow estimates generally increased with the additional data; the mean 1-in-100-year change (i.e. across all sites) was +7%. Spatial patterns of change correspond closely with the event footprint, whilst associations between change and both record length and catchment area were found to be only weak. The 'enhanced single-site' method was then applied at a subset of stations (without inclusion of the latest data). Interestingly, these estimates were not substantially higher that those produced from the same samples using the single-site method, implying a certain dependence between the pooling group stations. In a final set of analyses, the estimates were found, in many cases, to demonstrate less sensitivity to the choice of statistical distribution than to the sample or method used. Overall, these findings lend some support to the notion that standard methods may underestimate flood frequency. However, the difficulty of disproving probabilistic predictions reduces the confidence with which this assertion can be made. Going forward, reinvigorating efforts to incorporate longer-term hydrological data more routinely into hazard assessments may prove fruitful. It could also be appropriate for flood frequency estimates to be updated more frequently as new instrumental data become available.

Reassessing English flood frequency relationships in light of widespread new instrumental stage maxima



A thesis submitted for the degree of Master of Science by Research

James Matthew Thornton

Department of Geography
Durham University

May 2016

CONTENTS

ABBREVIATIONS

AEP	Annual (flow) exceedance probability
AM	Annual maxima
ASD	Above site datum
CEH	Centre for Ecology & Hydrology
EA	Environment Agency
FEH	Flood Estimation Handbook
FSR	Flood Studies Report
GEV	Generalised Extreme Value (distribution)
GL	Generalised-Logistic (distribution)
NAO	North Atlantic Oscillation
NRFA	Natural River Flow Archive
POT	Peaks over threshold
PWM	Probability Weighted Moments
QMED	Median annual maximum flood (flow)
SEPA	Scottish Environment Protection Agency

FIGURES

TABLES

EQUATIONS

STATEMENT OF COPYRIGHT

ACKNOWLEDGMENTS

This research was designed and conducted independently. Professor John Wainwright and Dr. Nick Cox made comments on draft sections of this thesis, for which I am grateful. I also acknowledge helpful discussions with former colleagues at JBA Risk Management. Permission to use the flood estimation software JFes was kindly granted by JBA Consulting. Finally, I thank my parents for their support.

Chapter 1

Introduction

"Current approaches using flood frequency analysis and flood risk assessment based on 40–50 year long flow records... are not fit for purpose"

Professor Mark Macklin, Aberystwyth University
News Article, University of Cambridge website[1]
December 2015

"What I mean in terms of a complete rethink is that we need to look, first of all, at the historic data we are using and ask ourselves if that is still valid to predict the future"

Mr. David Rooke, Environment Agency
House of Commons Environment, Food and Rural Affairs Committee[2]
January 2016

1.1. The December 2015 floods in northern England

Several flood events of national significance have affected the United Kingdom (UK) in recent years, raising interest in flood risk-related issues. The meteorology, hydrology and impacts of these events have been comprehensively reviewed in the academic literature (see e.g. Met Office, 2014; Dale and Marsh 2002; Marsh and Hannaford, 2007; Chatterton et al., 2010, 2016), and efforts to learn lessons have been made (e.g. Pitt, 2008). Yet despite this growing collective experience, floods continue to be associated with adverse outcomes for people and property, as the events of December 2015 attest.

Throughout much of late Autumn 2015, a pronounced north-south gradient in North Atlantic sea-surface temperatures contributed to a strong south-westerly flow of warm, moisture-laden air towards Northern Europe (Met Office, 2015a). The blocking effect of an enduring high-pressure system over continental Europe to the east displaced this jet from its usual more southerly passage (*Ibid.*). Within this synoptic context, several notable rain-bearing systems were brought to northern

[1] Macklin (2015).
[2] House of Commons (2016a).

Europe during November. As Figure 1.1 illustrates, much of Wales, the north-west of England and Scotland experienced high monthly rainfall totals. In fact, across the UK as a whole, November rainfall totals amounted to 145% of the 1981–2010 average (Met Office, 2015b). Thus, catchments became increasingly saturated – conditions which, in time and given subsequent intense rainfall, would prove favourable for flood generation.

The most intense and long-lasting rainfall episodes during December were broadly associated with a succession of three named systems[3] – Desmond, Eva and Frank – which passed through the British Isles on around the 5th, 23rd–24th and 29th–30th December respectively. The image on the title page (p. 2) shows the 'conveyor belt' of moisture that was brought to the British Isles on the 5th December. Evidence that such 'atmospheric rivers' are linked with winter flooding in the UK has been presented previously (Lavers et al., 2011), and these features do seem to have contributed to the floods presently under discussion. Figure 1.2 shows the exceptional nature of the rainfall during the month, which became the UK's wettest in the record extending back to 1910. In addition to this exceptional monthly total, the national 24-hour rainfall record was broken by the 341.4 mm observed at the Honister Pass, Cumbria, in the 24 hours preceding 1800 GMT on 5th December (Met Office, 2016d). Although this station was only installed relatively recently and is situated in an exposed, elevated and notoriously 'wet' location, this total is nonetheless impressive.

Rainfall patterns in Figure 1.2 (December) were extremely similar to those in Figure 1.1 (November). As already mentioned, wet antecedent conditions meant there was high potential for a large proportion of incident December rainfall to be converted to runoff.

[3] In September 2015, a pilot project entitled Name our Storms was established by Met Éireann and the UK Met Office (Met Office, 2015c). This initiative hopes to increase public awareness of the adverse weather typically associated with extra-tropical cyclone passage. Assigning names to storm systems also provides commentators with useful terms of reference. That said, the diversity and complexity of the rainfall fields associated with extra-tropical cyclones (compared to tropical cyclones, for example) should not be overlooked.

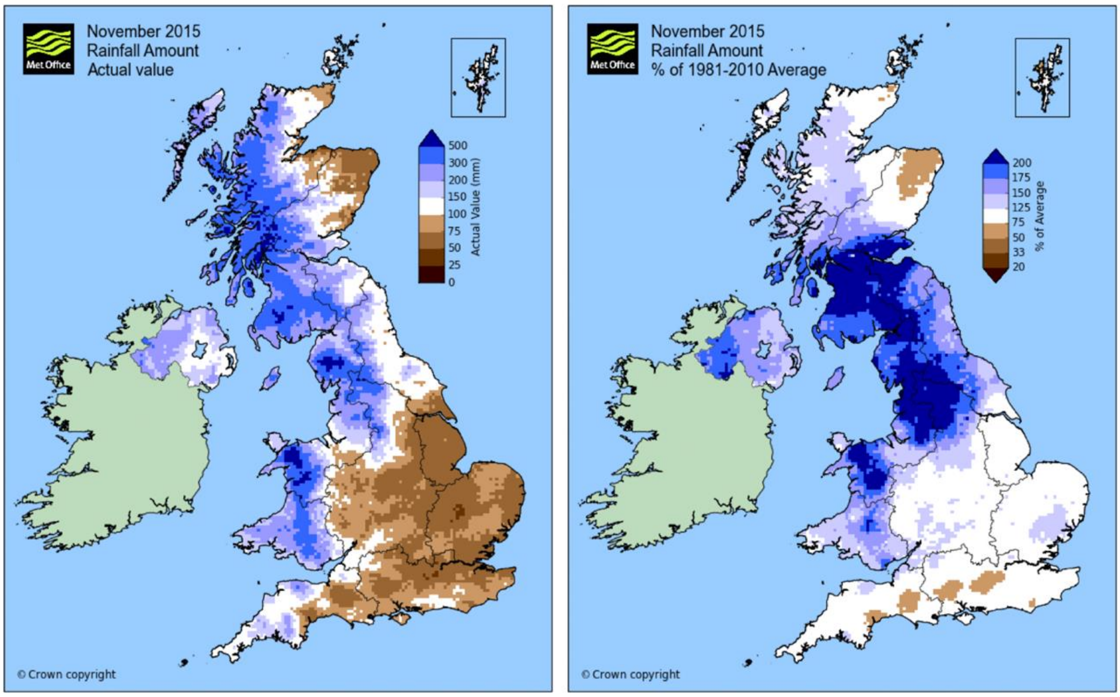


FIGURE 1.1. UK November rainfall totals (left) and as a percentage of the 1981-2010 average (right). Source: Met Office (No date).

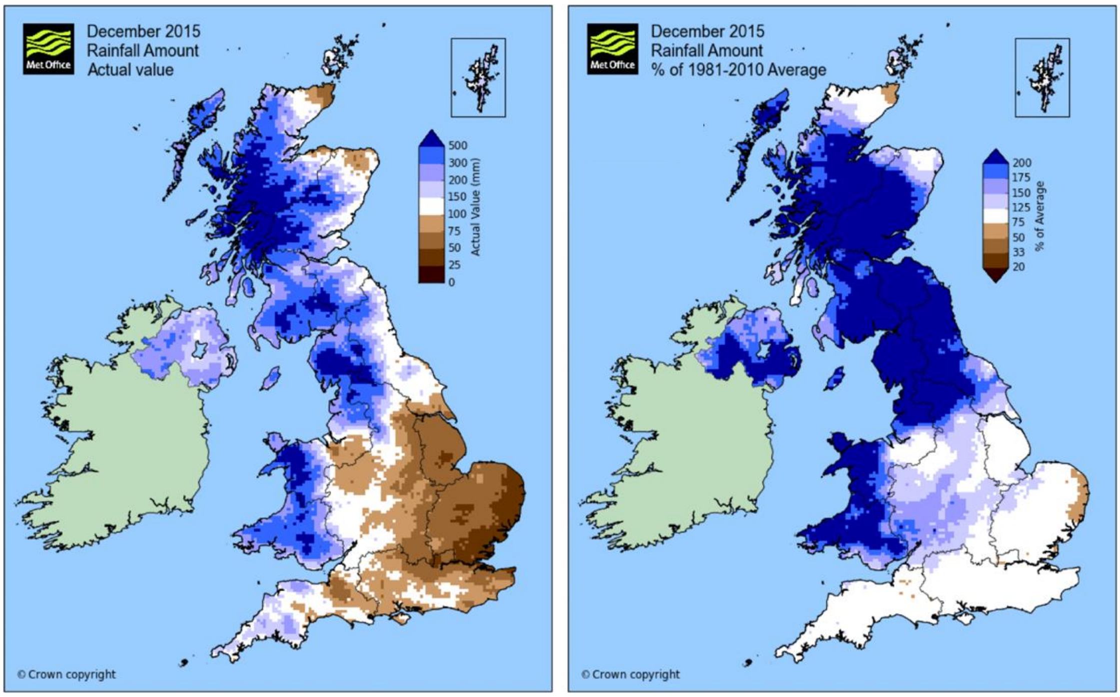


FIGURE 1.2. UK December rainfall totals (left) and as a percentage of the 1981-2010 average (right). Source: Met Office (No date).

On various occasions during December, river levels rose to new maxima at many locations across northern England (CEH, 2016). These locations are highlighted in Figure 1.3.

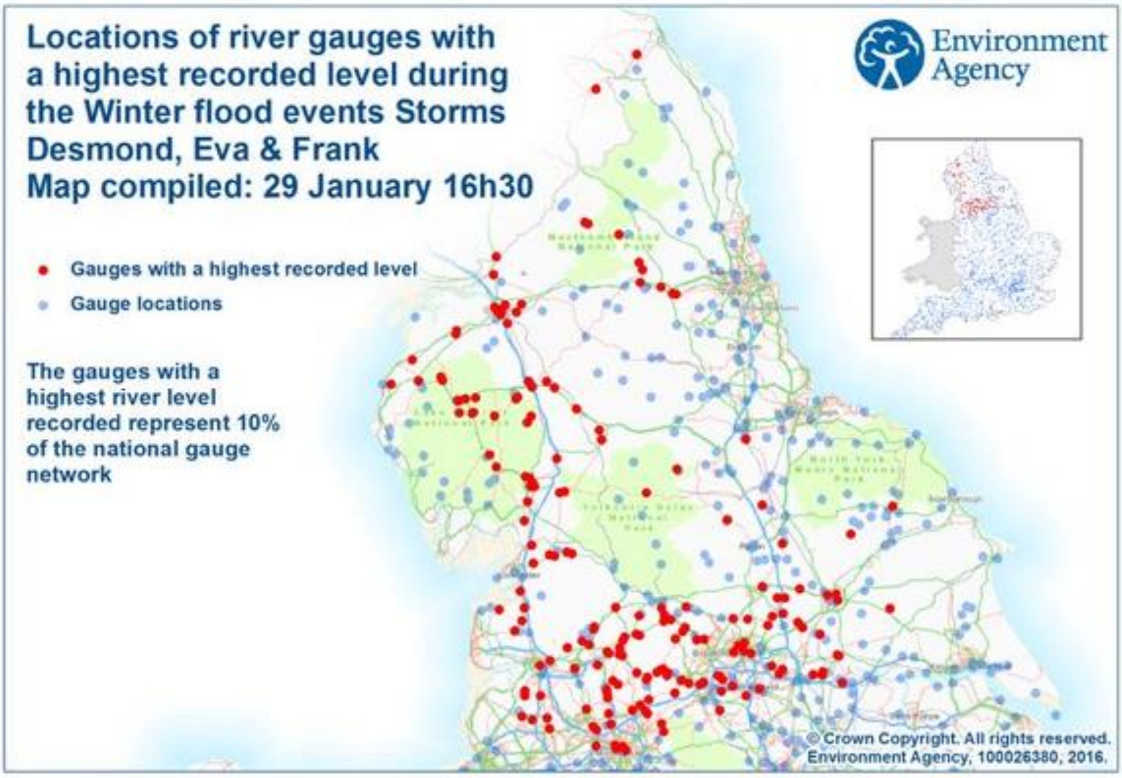


FIGURE 1.3. Locations of river gauges with a highest recorded level during the Winter 2015 flood events, which are said to represent 10% of the entire English network. Source: EA (2016a).

Preliminary analyses by the Centre for Ecology & Hydrology (CEH) further underlined the exceptional nature of event flows on certain major rivers. Figure 1.4, for example, expresses the mean December 2015 flows relative to the long term mean flows for the same month.

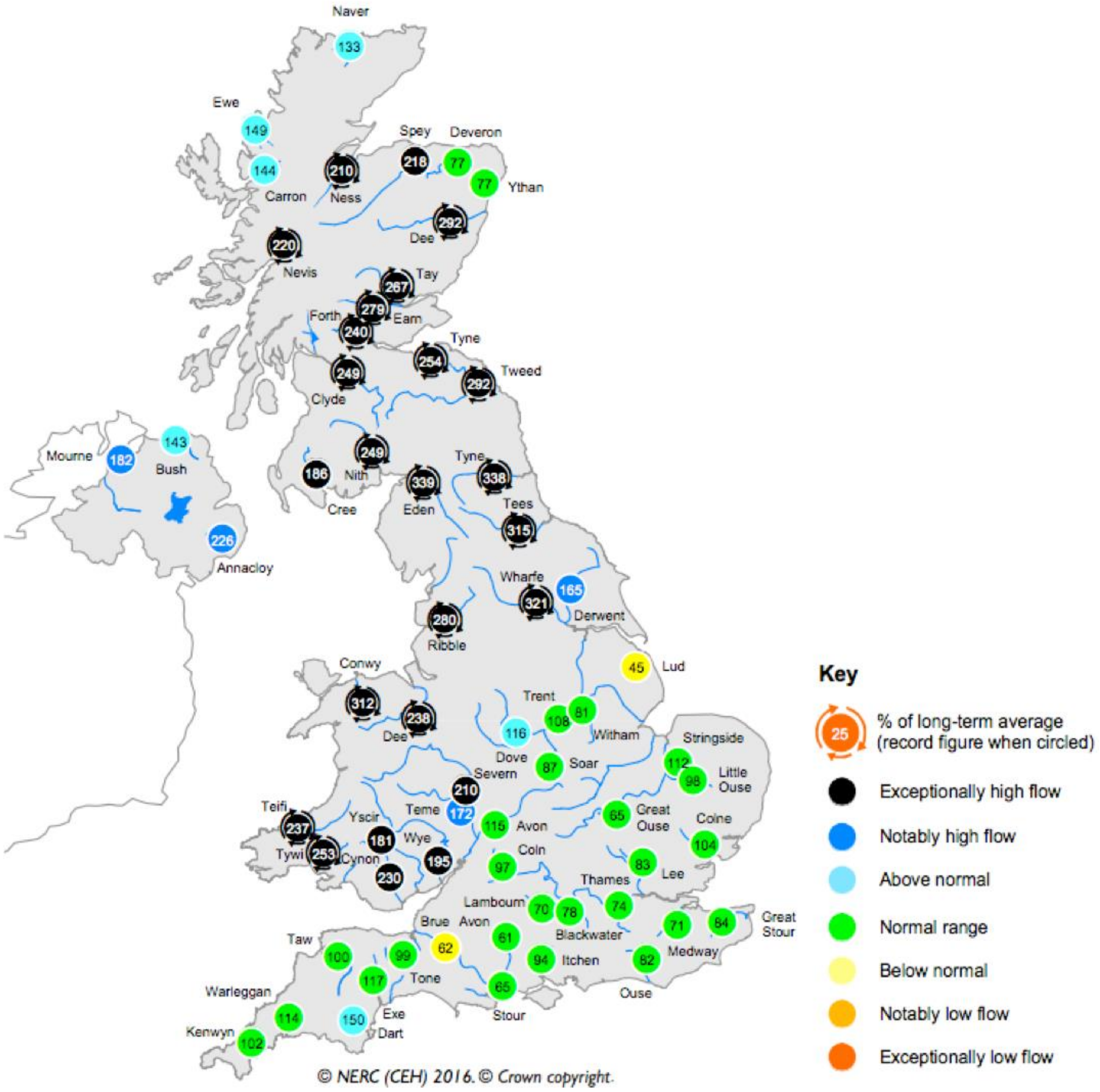


FIGURE 1.4. December 2015 mean flows for selected rivers expressed as a percentage of the long-term average of December mean flows. The period of record on which these percentages are based varies from station to station. Source: CEH (2016).

Whilst monthly mean flows are of general interest, peak flows correlate most closely with flooding. Those on the Rivers Eden, Lune and Tyne were identified as being particularly noteworthy; at approximately 1,700 $m^3 s^{-1}$ in each case, in fact they constitute the highest ever recorded peaks in England and Wales (CEH, 2016) (the largest ever Scottish flow is slightly higher). The return periods of these flows, i.e. the inverse of the annual probabilities of exceedance, were estimated using standard procedures (see Section 2.1.2) to be approximately 1-in-300-years for the Eden, 1-in-150-years for the Lune, and 1-in-100-years for the Tyne.

Many rivers burst their banks, causing widespread floodplain inundation. At least 16,000 homes were flooded in England alone (House of Commons, 2016b). Certain communities experienced repeated disruption, with Glenridding in Cumbria being a notable example. The response of the Environment Agency (EA) to the failure of pumps on the River Foss in York, North Yorkshire, was particularly contentious in some quarters; a decision was made to essentially 'sacrifice' a number of homes that normally benefit from the scheme in order to protect a larger number located elsewhere that were thought likely to flood without intervention.

Early estimates of the expected economic damage costs were in the region of £5–5.8bn (KPMG, 2015), although these now appear slightly exaggerated. The sum of insurance claims relating to domestic and commercial property was expected to reach approximately £1.3bn (ABI, 2016). National infrastructure was also adversely affected, including numerous roads (DfT, 2015) and bridges. The collapse of the 300-year old bridge at Tadcaster, North Yorkshire on 29th December was much publicised. Its repair is expected to take approximately 12 months and cost approximately £3m (BBC, 2016).

In terms of scale and impacts, the December 2015 floods only sit alongside those of Autumn 2000 and Summer 2007 in recent national history. Such notable flood events naturally stimulate a great deal of public and political debate, and those of December 2015 were no exception in this regard. As they have done previously, the affordability of domestic flood insurance (Priestley and Edmonds, 2016) and the presence and performance of flood defences featured as prominent topics once again. Arguably foremost, however, was discussion around the extent to which the frequency or severity of flooding might be changing; especially when these events are considered within the context of the recent sequence, it is unsurprising that there exists a widely held perception that flood risk is increasing (Hannaford, 2015). At this stage, it is important that risk be clearly distinguished from hazard; the standard definition of "risk = probability × consequence" is employed throughout this thesis.

Of all the various possible causal or contributory factors, there is particular interest in whether anthropogenic climate change might have had any impact on the hazard, in terms of either likelihood or severity, associated with recent major floods. (As a point of terminology, it may be noted that likelihood is often referred to as probability or frequency, and severity is often referred to as magnitude or intensity). This interest may be related to the increased prominence of climate change matters in general (e.g. due to increased media coverage), or a growing societal appreciation of the more specific probable impacts of anthropogenic warming on the hydrological system (Bates et al., 2008; Wentz et al., 2007). As Section 2.2 shall reveal, questions relating to the detection, attribution and future prediction of trends in river flood probabilities are not straightforward ones to answer, although good progress is being made, mostly via computer simulation, in certain areas.

1.2. Introducing flood frequency analysis and its attendant challenges

An obvious response to a non-imminent natural threat is to seek to quantify it. In particular, one might seek to estimate the probabilities with which certain hazard intensity levels could be expected to be exceeded, at locations of interest, within a specified period of time. Such exercises are usually known as long-term hazard assessments and are routinely undertaken, in various different guises, with respect all types of natural hazard. By taking the resultant information on hazard frequency and severity, more comprehensive risk assessments can be developed. In turn, the outputs of these risk assessments can aid informed, robust decision making concerning mitigation schemes and other possible interventions. Both hazard and risk assessment results might be described as probabilistic predictions because they relate to the aleatory uncertainty (i.e. random or natural variability) of what could happen in future.

This thesis is primarily concerned with the hazard posed by fluvial flooding. In fluvial hazard (and therefore risk) assessment, estimating the discharge levels associated with certain annual probabilities of exceedance, or vice versa, plays a central role. (The terms discharge and flow are used interchangeably throughout to refer to the volumetric rate of water flow through a cross-sectional area). This activity is commonly referred to as flood frequency analysis, or flood frequency estimation. Hydrologists have a small number of fundamentally different approaches to flood frequency estimation at their disposal (Merwade et al., 2008; Rogger et al., 2012). Those that involve fitting statistical distributions to instrumentally measured peak flow series are particularly well established, and remain widely used in practice (Cunnane, 1987; Potter and Lettenmaier, 1990; Kjeldsen et al., 2008a).

In this regard, distribution fitting is usually necessary in the first place because the available records are seldom long (and therefore complete) enough for high flow quantiles to be reliably estimated from the empirical distributions alone. In the simplest terms, future flows are not necessarily constrained to the range of past observations. Thus, in enabling the extrapolation of flow frequency-magnitude relationships beyond their observed ranges to probability levels holding more practical interest, statistical modelling of this nature has the potential to be extremely useful.

However, approaching flood frequency estimation from this perspective is not without its challenges. For instance, as with statistical approaches in general, considerable reliance is placed on the data as opposed to any underlying physical mechanisms or theory. Consequently, should these data not be sufficiently representative of the underlying processes, the resultant estimates may be somewhat dubious. Placing such reliance on data is especially problematic when dealing with extremes, as one must when assessing flood hazard, because the number of relevant observations naturally tends to be limited – extremes are, by definition, rare. Additionally, reliable high river flow measurements, from which under the statistical approach most flood hazard and risk assessment follows, are notoriously difficult to make in practice (see Section 2.1.3). Hence, such measurements are often subject to considerable uncertainty (Coxon et al., 2015).

Furthermore, most statistical approaches demand that observations can be assumed to be mutually independent and identically distributed (i.i.d.). In other words, they should be independent samples from a single underlying probability distribution. The concept of stationarity – that the underlying stochastic process generating the observations must not be changing over time (Lins, 2012) – is embedded within this assumption. If the process in question were non-stationary, then no single underlying distribution would exist. (It is worth mentioning that few environmental datasets can ever be described as truly stationary in the strictest sense, *Ibid.*). Moreover, it is known that in reality, clustering of floods (i.e. non-Poissonian occurrence) can occur on various timescales due physical feedback mechanisms (Merz et al., 2014), calling into question the assumption of independence. It has been shown that not accounting for this 'persistence' can lead to systematic underestimation of temporally-aggregated flood risk (e.g. over the duration of an annual insurance contract) (Serinaldi and Kilsby, 2016). In summary, given the physical complexity and number of different processes in the hydro-meteorological system, the assumptions required are fairly stringent.

Finally, the model fitting process itself may be complicated by the choice of different statistical distributions and parameter estimation methods that are available; given limited data and no well-established theoretical basis, it can be difficult to identify which might be most appropriate (Kidson and Richards, 2005; Merz and Blöschl, 2008).

The upshot of the complex interaction of such factors, including the possible contravention of the assumptions, is that the overall uncertainty associated with return level or 'design flow' estimates (that is, the flow levels associated with certain annual probabilities of occurrence, or inversely return period) is known to be considerable (Kjeldsen, 2015). It is also reasonably well established, and logical, that the degree of uncertainty normally increases with rarity (return period), as a greater degree of extrapolation is required. Having said that, quantifying and combining the uncertainty contributed by all sources to provide an overall figure, perhaps expressed as confidence intervals around the central estimates, remains problematic. One reason for this is that both aleatory (random) and epistemic (lack of knowledge-based) uncertainties are present (Beven et al., 2011). That said, it is possible for confidence intervals representing certain sources of uncertainty to be produced, and presenting such information where possible should be considered good practice. Madsen et al. (1997a,b) and Collier (2011), for example, estimated confidence intervals as a measure of uncertainty and used the information to identify most appropriate quantile estimation approach depending on the circumstances. Even if confidence intervals can be generated, in the face of multiple sources of uncertainty, including difficult-to-quantify or even non-quantifiable elements (e.g. high flow measurement uncertainties), all participants in the flood estimation process should be encouraged to appreciate and communicate the assumptions, challenges and uncertainties involved as thoroughly as possible.

1.3. Could there be a bias towards underestimation?

Given the challenges associated with flood frequency estimation outlined above, it is perfectly reasonable to expect such estimates to be fairly uncertain. However, one would hope that they do not demonstrate any particular bias, since biases generally pose more problems than unbiased uncertainties (which may even cancel one another out in particularly fortunate situations). However, interestingly, and perhaps from the perspective of flood risk management worryingly, at least two arguments or explanations as to why UK fluvial flood hazard estimates might be biased towards underestimation may be proposed – and to a certain extent already have been by others. These suggestions are now outlined in turn.

1.3.1. “Hazard may be increasing due to anthropogenic climate change”

The first reason, which is now well rehearsed, involves anthropogenic climate change. The standard methods that are routinely used to estimate flood frequency in the UK, which is introduced in Section 2.1.2, require the assumption of stationarity to be made (IH, 1999). Therefore, at a general level, any changes in the frequency or severity of high river flows, irrespective of the cause, would invalidate this assumption (Milly et al., 2008, 2015). More specifically, it has variously been suggested on the basis of thermodynamic arguments (essentially that warmer air can hold more moisture; Wentz et al., 2007) and possible changes in atmospheric circulation patterns and other meteorology (see, for example, Lavers et al. (2013) with respect to ‘atmospheric rivers’) that both total rainfall and extreme rainfall could be expected to increase in a warmer world; an expectation with which climate model projections appear to concur (see Section 2.2.2). Indeed, some increases in intense rainfall have already been identified (Jones et al., 2013). The extension may then be made that if extreme rainfall is expected to increase or is already increasing, then fluvial flood hazard could be expected to broadly follow suit (although the many non-linear catchment processes between rainfall and flooding must be emphasised; Laizé and Hannah, 2010). As such, if there is a material upward trend in peak river flow frequency or severity, it is clear that some form of underestimation would result under the assumption of stationarity, since the present-day probability of exceedance of a given flow level would be higher than that estimated according to the integrated, effectively ‘unordered’ past record (Milly et al., 2008).

In the aftermath of the December 2015 floods, the Deputy Chief Executive of the Environment Agency, David Rooke, suggested that as a result of the transition from a world of “known extremes to unknown extremes”, a “complete rethink” of flood preparation and mitigation measures was required (BBC, 2015; House of Commons, 2016a, p.39). These pronouncements appear to strongly question the performance of the existing models (specifically in the ‘direction’ of underestimation), and invoke climate change as a contributory factor in the events (and hence perhaps a factor in the model ‘failure’ which led to the degree of surprise). The notion of a “complete rethink” in particular suggests that observations from the event may have been somewhat inconsistent with the established view, and perhaps therefore might act as ‘leverage points’ to fundamentally change the prior understanding.

Physics-based atmospheric modelling methods, that enable the contribution of anthropogenic climate change to the likelihood or severity of individual weather events (including flood-causing rainfall), are beginning to emerge (Herring et al., 2014; Pall et al., 2011; Schaller et al., 2016). Suhc information is vital if which post-event debates are to be better informed. Some suggest that positive associations between anthropogenic greenhouse gas emissions and increased flood hazard are already detectable. Indeed, a study conducted in near real time proposed a link between anthropogenic climate change and the meteorology of early December 2015 (van Oldenborgh et al., 2015). This field of research is considered further in Section 2.2.3.

Finally for the time being, it is noteworthy that the longest river flows series do not (yet?) provide evidence for any notable long-term trends in either flood or severity (Hannaford and Marsh, 2008; Hannaford, 2015; Wilby et al., 2008). Recent research on this topic is reviewed in greater depth in Section 2.2.1.

1.3.2. "Even notwithstanding climate change, flood frequency may be underestimated from short records"

The second suggested reason as to why flood hazard might be systematically underestimated does not require any increasing trends to be invoked. It is simply that even if the records are perfectly stationary, because most instrumental UK river flow records are short (typically around 40–50 years long) and extremes by definition rare, the peak flow samples from which flood frequency estimates are conventionally derived are likely to contain disproportionately few extreme flows (Macdonald, 2013). Hence, whilst the statistical methods employed do permit some extrapolation beyond the range of observations as discussed, they may not be capable of raising the flood frequency curves enough. Put alternatively, lower-magnitude observations may be incapable of providing sufficient insight into characteristics of the typically heavy distribution tails.

Assuming an absence of long-term trends, some apparent support for this suggestion is provided by studies which have sought to first extend flood records using non-instrumental (a.k.a non-systematic) sources, and then to reassess flood frequency relationships including the additional data. Besides the benefit of a general reduction in uncertainty (Macdonald et al., 2014), the effect of incorporating these longer records is often that flow frequency-magnitude curves increase (e.g. Black and Fadipe, 2009). In other words, the estimated frequency with which a given flow level might be expected often decreases when longer series are used.

It perhaps follows, then, that the apparently exceptional and 'unexpected' nature of the December 2015 floods are simply a reflection of the possibility that the short flow records upon which existing flood frequency estimates are based may not represent the full range of natural variability possible (CEH, no date). From this line of reasoning and the emerging supporting evidence, the very fitness-for-purpose of standard methods has been called into question (Macklin, 2015).

Finally, it is important to emphasise that these two alternative suggestions are not mutually exclusive. On the contrary, if upward trends are emergent and the more fundamental 'short record' problem is also present, the overall degree of underestimation would be even more pronounced; a combination likely having the most considerable implications for flood risk management policy.

1.4. The difficulty of disproving probabilistic predictions

The quotation of Mr. Rooke that is reproduced at the beginning of this thesis alludes to the fact that whenever events that appear to be exceptional or are 'surprising' occur, there is often a desire to re-evaluate those existing models whose purpose it was to assign a probability to (or, if you will, to 'predict') such events. In particular, one might attempt to ascertain whether or not the probabilities assigned initially to the occurrence of such floods were reasonable. However, taking the case of just a single location for simplicity, given the nature of probabilistic predictions, so long as some non-zero probability was assigned to the flow levels that was attained in the event beforehand, the prior model cannot easily be overturned using the single additional observation. Quite simply, the event that occurred might indeed truly be extremely rare, or in other words have a very low probability of occurrence within the time frame in question. Under such circumstances, that it happened at all within the period of observation could be put down to bad luck. (Such a truly rare event occurring in a relatively short observed period also has the potential to distort the flood frequency relationship in the other direction, i.e. make the extreme intensity level attained seem more likely than it really is). The only scenario in which the probabilistic prediction would be invalidated if is the highest (/lowest probability) predicted intensity level was surpassed by an observation (i.e. zero probability was assigned to its occurrence). This can only occur when flood frequency relationships are bounded, which typically they are not.

Of course, in the case of a 'high extreme', it may still be that too low a probability had previously been assigned to the event or level subsequently observed. In other words, there many have been some underestimation. It can be tempting to suggest that high outliers or low-probability events are either consistent with or even themselves demonstrate prior underestimation, neglecting the subtleties associated with probabilities predictions. This should be avoided since observations of events or their outcomes from across time and space are generally required to build any case that a prior probabilistic prediction could be invalid. For instance, should several events each having very low individual estimated occurrence probabilities occur within a relatively short period of time at a single location, then (assuming temporal independence) the joint probability of the observed sequence happening by chance under the prior model might become so small that is can be formally rejected.

If nothing else, the rather philosophical discussion presented in this section hopefully highlighted the challenge of testing or 'validating' predictions which seek to deal with uncertain future outcomes. In the next section, the present study is introduced.

1.5. The present study

1.5.1. Aim and objectives

Within the context of these hypotheses that flood hazard may be somewhat underestimated, and fully acknowledging that any sort of definitive answer is likely to be elusive using additional data from only a single event, the aim of this thesis is:

To explore the effect of including the latest instrumental river flow observations from northern England in December 2015 on flood frequency estimates produced using established, widely used methods.

These river flows represent new record maxima in many locations.

The study's more specific research objectives are:

1. To explore the effects of including December peak flow observations on flow frequency-magnitude relationships (a.k.a. flood frequency estimates) produced using the single-site statistical method at as many stations as possible;

2. To investigate whether any associations between change in flood frequency estimates observed in Objective 1 and i) flow record length and ii) catchment area might be apparent;

3. To compare flood frequency estimates produced using the 'enhanced single-site' method with those produced on a single-site basis (with and without the additional data) (again from Objective 1) and observed data where possible, and;

4. To assess the sensitivity of flood frequency estimates to choice of statistical distribution relative to variability associated with the specific data or methodology employed (from Objectives 1–3).

1.5.2. Rationale

Estimating flood frequency on a single-site basis is conceptually straightforward (see Section 2.1.2 for details). It was therefore an obvious starting point when seeking to investigate the possible implications of the exceptional December 2015 floods. Specifically, statistical models were fitted within and without the latest annual river flow maxima (AM) and then compared. Using this method, the necessary estimates could be produced relatively quickly at a large number of sites, thus enabling overall patterns of change – including spatial ones – to be identified. **(Objective 1).**

Potential relationships between degree of change and record length and catchment area were investigated to consider whether there might be any underlying factors driving the sensitivity (i.e. change) of the estimates. In particular, it is established that single-site estimates (especially of long return period flows) can be quite sensitive to the particular characteristics of the record. Estimates produced from shorter records, which effectively represent smaller samples, are thus likely to be more sensitive to the additional data than longer records, i.e. a negative association between change and record length was expected. Expectations with respect to any relationship between change and catchment area were less clear. **(Objective 2).**

Due to the sensitivity that single-site estimates can demonstrate to the characteristic of the often short records, it is not recommended by official UK guidance (IH, 1999 and subsequent updates) unless the record length is more than double the 'target' return period, *T* (assuming there is only one, or if not the longest). Where this is not the case, a form of spatially pooled analysis is recommended. This involves essentially extending the record at the target site using data from other, hydrologically similar sites. For flood estimation at gauged sites, in an attempt to achieve a balance between the benefit of the target site data (i.e. the assured relevance of the data to the study location) and pooling (i.e. the increased sample size, and hence lower sampling uncertainty), the 'enhanced single-site' method is endorsed. Again, further information is provided in Section 2.1.2. As part of the work to address this objective, some model-to-data comparisons were carried out. These complement the model-to-model comparisons, of which one must always be a little wary (**Objective 3).**

Finally, there remains much debate regarding what the most appropriate statistical distribution to describe peak flows might be. Therefore, attempts were made to explore how large the variability (or uncertainty) associative with the choice of 'equally valid' statistical distributions might, in typical cases, be relative to the variability (or uncertainty) associated with the data available or method followed **(Objective 4)**.

Following a detailed review of the subject in Section 2.2, for ease of implementation and to provide consistency with current practice, it was decided that this study should proceed under the assumption of stationarity. Further justification of this decision is given in Section 2.2. However, this does not preclude the first possible reason proposed above from being responsible for any underestimation. Rather, making this assumption means that irrespective of whether hazard estimates are revised upwards with the inclusion of the additional event flow data, the revised estimations may still represent underestimations if any sort of increasing trend is emerging.

In comparison with studies that seek to extend records of flooding backwards, extending forwards as here has three benefits. Firstly, uncertainties associated with the magnitude of the recent instrumentally measured flows, whilst by no means negligible (see Section 2.1.3), are likely to be lower than those associated with flows reconstructed from historical or palaeoflood data. Secondly, the climatic and catchment land use conditions under which recent flows have been produced are likely to be more relevant with regard to future flooding than those under which generated historical or palaeofloods. Since in many cases the largest floods can exert a major influence over the fitted distribution and hence the estimates produced, this question of relevance is often of major concern. Thirdly, employing approaches that are widely and routinely used for flood estimation by practicing hydrologists in the UK water industry has the additional benefit that any findings, conclusions or recommendations may have more direct relevance than they might if non-standard methods had been used.

That attempts can now be made to update English flood hazard estimates so rapidly after high-impact flood events is largely thanks to the Environment Agency's (EA) Open Data initiative (Open Data Institute, 2016). Whereas previously river observations only became freely via the National River Flow Archive (NRFA), which is operated by the Centre for Ecology & Hydrology (CEH), after a significant delay, continuous river level observations from gauging stations across England are now publically available in real time. As such, despite still being provisional (having not yet undergone the rigorous quality assurance needed for their re-release on the NRFA), when combined with other data from the NRFA, observations from December 2015 provide an excellent resource for further improving our collective understanding of flood frequency in the UK, forming the foundation of the analyses conducted herein.

As persisting with the assumption of stationarity implies, this research specifically does not set out to revisit the question of whether any firmer conclusions to be made concerning trends in empirical flow series can be made in light of the new data. Previously reported limitations of empirical change detection, to which attention is drawn in Section 2.2.1, likely still apply; in particular, should the latest data be included, the high 'outlier' that would be present at the very end of the series in many locations might lead to spurious trend results (Wilby et al., 2008). As alluded to above, further research, especially the probabilistic attribution of flood-causing weather events to anthropogenic climate change, holds the promise to the address such problems associated with empirical analysis (although are subject to their own uncertainties). Several other potentially interesting topics also lay out of scope, including flood frequency estimation at ungauged locations, the consideration of complex multivariate stochastic models that can be applied to produce plausible future rainfall and river flow data in time and space, other types of flooding such as pluvial flooding, and the notion of flood risk.

In summary, although flood frequency estimation is inherently an extremely challenging endeavour, the levels of investment that depend on it necessitate that it be addressed and continual improvement pursued. By means of discussion of the recent literature throughout, this study seeks to position itself at the forefront of contemporary research and practice in relation to the pertinent topic of flood frequency estimation in the content of natural variability and possible anthropogenic climate change impacts. It is hoped that the findings will in some way complement the ongoing the National Flood Resilience Review, which was established following the flooding of December 2015 and is due to be published in Summer 2016 (Defra, 2015). Overall, the significance and novelty of the proposed work lies in its timely application of the very latest observations to the important and challenging research question of how well flood frequency estimates can be made; no existing studies are known of which have explored changes to flood frequency across many locations in light of any recent exceptional floods.

1.5.3. Thesis structure

The remainder of this thesis is structured as follows:

Chapter 2 consists of a review of relevant literature, and is arranged into three main sections. In Section 2.1, further background to the field of fluvial flood hazard and risk assessment is provided, Specifically, the continued need for long-term assessment is emphasised, and standard statistical methodologies for flood frequency estimation that are widely used in UK practice are introduced in more detail. The measurement of high river flows and associated sources of uncertainty is also discussed. In Section 2.2, recent research tackling potential non-stationarity in UK peak river flows – past, present and future – is evaluated. Both empirically-based and simulation-based studies are considered. The view on how the present work should proceed in this regard, to assume stationarity, was only reached after considering this material. Thereafter, Section 2.3 focuses on some of the historically-informed flood frequency assessments that have been undertaken, and other relevant research. Taken together, the findings of such studies suggest that even irrespective of any potential climate change-induced trends, flood hazard could be underestimated using conventional approaches which rely heavily on relatively short records.

Thereafter, the research methods that were followed in order to address the aim and objectives are described in **Chapter 3**, whilst the results are presented and discussed in **Chapter 4**. Finally, conclusions are drawn and recommendations are made in **Chapter 5**.

Chapter 2

Review

2.1. Established practice in flood hazard and risk assessment

The aim of this section is to demonstrate the continued importance of fluvial hazard and risk assessment. Some typical applications are presented first, before the standard statistical flood frequency estimation techniques that often underpin them are described. Finally, the process of obtaining measurements of high river flows and sources of uncertainty therein are briefly discussed.

2.1.1. The continued need for long-term flood assessment

In contrast to short-term flood forecasting, where the aim is to predict areas that may be inundated and perhaps also timing and severity several hours or days in advance of a flood (Werner et al., 2009), long-term assessments express hazard (or risk) in the absence of an imminent threat. This activity is crucial because, even if perfect short-term forecasting were possible (of course it is not, and may never be; Weerts et al., 2011), flood mitigation and other hydraulic structures such as levees, retaining walls, and bridges have long intended lifespans at the time of their construction. They are therefore very likely to be subjected to extremely high flow levels at some stage, and so must be designed to withstand them. Thus, establishing the design criteria of such structures (for example, the peak flows associated with the desired annual probability of exceedance) constitutes the most direct application of flood frequency analysis (Coles and Tawn, 1994; Kidson and Richards, 2005; Prosdocimi et al., 2014).

Flood frequency results also often provide inflow boundary conditions to hydraulic models (Bates and De Roo, 2000; Horritt and Bates, 2002). These design flows can be efficiently routed over high-resolution digital terrain datasets (often now using 2D codes) in order to produce fluvial flood hazard maps. Indeed, it has been possible for some years now to produce such maps for large areas relatively quickly (Bradbrook et al., 2005). The maps usually show the spatial distribution of maximum water depth associated with flows of certain annual probabilities of exceedance (or, inversely, return periods) (Merz et al., 2007). Figure 2.1 provides an illustration of the general concept, although this particular example is somewhat more sophisticated than many since it contains additional uncertainty information, and may therefore be described as probabilistic as opposed to deterministic.

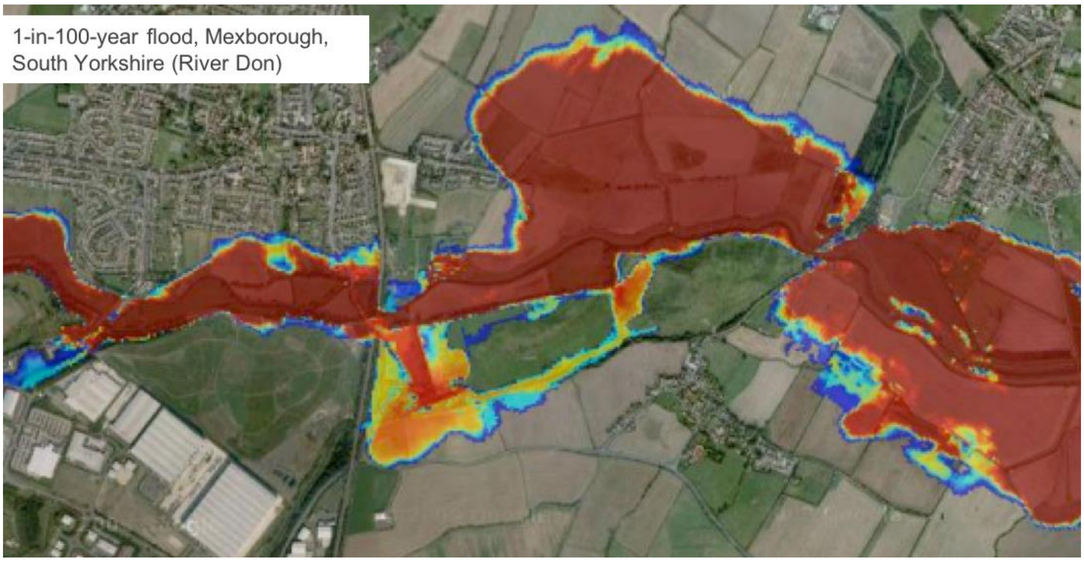


FIGURE 2.1. Example flood hazard map for Mexborough, South Yorkshire. To date, most such maps have been purely deterministic. In contrast, this map shows the likelihood of inundation, ranging from high (red) to low (blue), associated with the 1% Annual Exceedance Probability (AEP) (i.e. 1-in-100-year return period) flow. In this way, the effects of uncertainty, stemming from various sources, can be represented. Source: Beven et al. (2015).

In the UK, planning zones are identified on the basis such maps, those produced by the EA are appropriated for this purpose. Accordingly, such maps also are heavily relied upon when conducting so-called Strategic Flood Risk Assessments (Porter and Demeritt, 2012; see also JBA Consulting, 2014 for an example of such a document).

If risk is defined as the mathematical expectation of loss (Rougier, 2013), then it is clear that its calculation must incorporate not only various plausible future hazard possibilities, underpinned by some sort of flood frequency analysis, but also information on consequences. In this sense, the information typically contained within flood hazard maps – specifically the probability with which a water depth might be exceeded in floodplain locations – supports risk analyses at individual sites (Büchele et al., 2006; Penning-Rowsell et al., 2005). When aggregate risk to a portfolio of spatially distributed assets is of concern, due to spatial correlation in flooding patterns, it may be necessary to map inundation associated with for many plausible, spatially coherent scenario events. Generating such catalogues still typically requires at-site relationships between flows and their probabilities to be established initially, however (see e.g. Thornton et al., 2014). When aggregate risk over time is also of concern, events must also be assigned frequencies that include the effects of any temporal dependence or clustering.

In England, flood risk management is primarily the responsibility of the EA, whose remit includes the delivery of a range of flood-related infrastructure and services. Having the ability to estimate risk, including the reductions that might be achieved under various hypothesised interventions, provides a strong foundation for objective and robust decision making (Hall et al., 2003; EA, 2014). For instance, the likely long-term economic merits (or otherwise) of proposed flood defences can be demonstrated in this fashion (Harvey et al., 2012).

Another outlet for flood risk modelling is the estimation of fair, risk-reflective prices for the flood element insurance policies at individual properties, along with associated decision making by insurance companies around capital allocation, reinsurance purchases etc. A wide variety of different approaches, with differing degrees of complexity, may be pursued to this end. Essentially, be it at an individual property or global level, risk modelling helps the (re)insurance industry to assume what may be termed the 'residual' risk, which is that remaining once the mitigating effects of flood defences and other measures have been accounted for (Dawson et al., 2011).

In summary, flood risk managers in various spheres rely heavily upon flood hazard and risk assessment results. It follows that these results, and hence subsequent decisions, depend greatly on the reliability of the flood frequency estimates which underpin them. Some of the generic challenges in this regard were introduced in Section 1.2. In the next section, the specific methods by which these estimates are generated by practitioners in the UK water industry are introduced.

2.1.2. Standardised statistical methods for flood frequency estimation in the UK

The UK has a relatively long history of standardisation of (mostly statistical) flood estimation methods for practical applications. This began in earnest with the publication of the Flood Studies Report (NERC, 1975), which has now been superseded by the Flood Estimation Handbook (FEH) (IH, 1999) and its subsequent updates (Kjeldsen et al., 2008b).

The FEH statistical method remains dominant to this day. It is based upon the 'index flood' approach, with the median of the annual maximum (AM) flow distribution, *QMED*, typically being adopted as the index flood. Importantly, this quantity can be estimated reasonably well from relatively short AM series – as few as 13 data points are normally sufficient. In order to estimate the probabilities of more extreme flow levels, such as the 1-in-100-year flow, dimensionless growth factors relevant to the site in question must then be determined. When compiled together, these factors constitute a growth curve. Finally, in order to estimate design flows, the index flood can simply be multiplied by the growth factors corresponding to the required return periods.

The first maxim of flood estimation according to the FEH is that "gauged is best". In other words, whenever flood estimations are required for a gauged site, measurements from that particular site should assume a central role. As already stated, estimating *QMED* in such circumstances is normally straightforward. With regards to growth curve estimation, when the AM series length records exceeds twice the target return period, T (assuming there is a single target probability; when there is not, the longest target return period should be taken), then only data from the site in question are required. Several alternative statistical distributions might be selected, although having been identified as most suitable for the majority of UK sites, the Generalised Logistic (GL) distribution, coupled with the L-moments fitting method of Hosking and Wallis (1997), is recommended as the default (IH, 1999). The GL distribution and method of L-moments are elucidated in Section 3.4.1.

Where the record at the target site is not so long, and especially where the number of years is less than T (as is usually the case given the length of most flow series), it is recommended that additional data be incorporated into the growth curve estimation procedure. To achieve this, several other gauged catchments that are hydrologically similar to the target site, defined as such according to a set of 'catchment descriptors', must first be identified. These catchments need not be geographically close to the target site, as Figure 2.2, which shows a pooling group that might be identified for the River Dee at Polhollick (12003), demonstrates. Then, a pooled growth curve can be developed from the combined dataset with the contribution made by each catchment weighted according to its degree of similarly with the target station.

FIGURE 2.2. The pooling group for estimation floods for the River Dee at Polhollick (12003; location indicated with a cross) consists of data from 20 gauged catchments (dots) considered hydrologically similar to the study site. The FSR regions (NERC, 1975) are shown in colour. Source: Blöschl (2013).

In both single-site and pooled cases, it is possible for the fit of the model to the observations can be inspected graphically (e.g. Figure 2.3). Data plotting positions are usually determined using the Gringorten formula (Gringorten, 1963; Equation 1). The exceedance probability, *P(X)*, is estimated as:

$$P(X) = \frac{r - 0.44}{n + 0.12} \qquad [1]$$

where n is the total number of years in the series and r is the rank position (arranged in descending order of magnitude).

The formula overcomes the problem that r/n is a poor estimator when n is small.

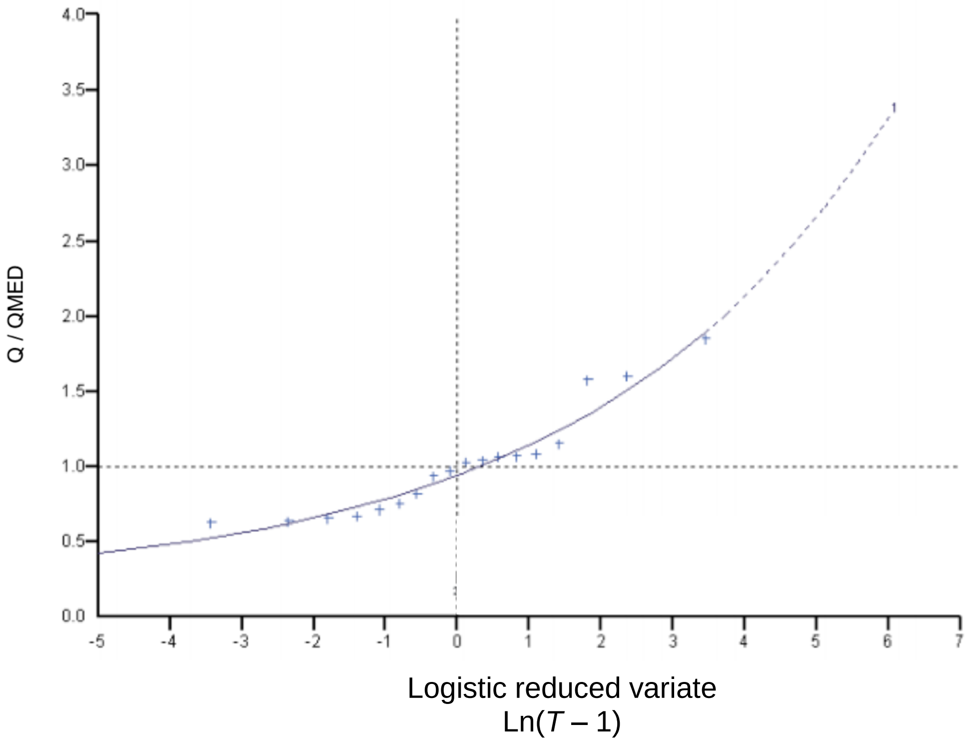


FIGURE 2.3. Example flood frequency curve for the River Meig, Highland (4005). The positions of the observations are determined using the Gringorten formula. Source: WINFAP-FEH 3 User Manual (2009).

In reality, flow gauges are rarely positioned conveniently near sites of interest. Therefore, flood estimation at ungauged sites is frequently required, presenting a notable additional challenge (Kjeldsen and Jones, 2009). To address this, empirical formulae have been derived which enable the estimation of *QMED* at ungauged sites, given parameters describing the physical characteristics of the catchment draining to that point. The latest version of this equation, which was developed by Kjeldsen et al. (2008b), is:

$$QMED = 8.3062\, AREA^{0.8510} 0.1536^{\left(\frac{1000}{\mathrm{SAAR}}\right)}\, FARL^{3.4451} 0.0460^{\mathrm{BFIHOST}^2} \quad [2]$$

where *AREA* is catchment drainage area (in km^2), *SAAR* is standard annual average rainfall 1961-1990, *FARL* is an index of flood attenuation due to reservoirs and lakes, and *BFIHOST* is a baseflow index derived from HOST soil type data (Boorman et al., 1995).

Once *QMED* has been estimated, a pooled growth curve can then be developed for the ungauged site in the same fashion as already described, using gauged sites as proxies. This process of pooling or 'regionalisation', which amounts to a substitution of space for time, is presently the primary means by which standard methods attempt to deal with both the problem of short records and the need for flood estimation at ungauged sites. However, the focus of the present thesis is on the theoretically more straightforward challenge of flood estimation at gauged sites, to which discussion now returns.

The implicit assumption in pooling observations from hydrologically similar catchments is that the resultant group of catchments are homogeneous in terms of their flood generating mechanisms (IH, 1999; Archer et al., 2007). Therefore, even where considered similar enough to be included by the method, introducing data that is not native to the target site might compromise the overall relevance of the data to that site somewhat. In this regard, a careful balance between the amount of data and its relevance is required.

A number of cases exist in the literature where pooled estimates have been rejected for different reasons (but all essentially related to lack of relevance) in favour of either single-site estimates or those produced based on alternative (e.g. historical) data. For instance, Archer et al. (2007) rejected pooled results for the River Tyne at Bywell, suggesting that reservoir effects in the pooling group may be rather different to those at the study site. Black and Fadipe (2009), meanwhile, suggested that due to the unusual (summer) seasonality of flooding in the Spey catchment in northeast Scotland confounds pooling attempts. Similarly, in assessing pooled flood frequency results for the River Ouse at York (compared to some historical data), Macdonald and Black (2010) suggested that the lack of sufficiently large analogue catchments presents a problem for pooling here.

Concerns about the assumption of homogeneity are particularly evident in urbanised catchments, which naturally tend to form a focus for flood risk management activity due to the concentrations of people and property they contain. Because of the diversity, even uniqueness of flood generation processes in such catchments, identifying sufficiently analogous catchments is often extremely difficult (Reed, 2002). Consequently, the FEH does not recommend pooling of any sort (i.e. including the enhanced single-site variant, which is introduced below) in urbanised catchments. This is clearly an unfortunate situation for the hydrologist or flood risk manager, who must resort to making somewhat crude 'urban adjustments' to as-rural estimates (Kjeldsen et al., 2010) or must applying entirely different methods. More generally, it seems that the established concept of 'uniqueness of place' (Beven, 2000) may always limit the utility of pooling somewhat. Put another way, if the standard similarity criteria were tightened so to that relevance or homogeneity was more or less guaranteed, then the pooling 'groups' may end up with few or no additional catchments in them. Certainly, pooling should never by thoughtlessly undertaken since it by no means represents a perfect solution to the short record problem.

In an attempt to strike an improved balance, the enhanced single-site method has been developed (Kjeldsen et al., 2008b). This approach, which acknowledges the particular value of data from the target site by assigning them a much greater weighting compared to other records in the pooling group, is now recommended for use at gauged sites.

A number of additional considerations must also be made under the statistical approach. For instance, there is some debate as to what specific 'cut' of a river flow time-series is best used as a basis for flood frequency estimation. The choice normally comes down to either using block annual maxima data (AM) data or, alternatively, all peaks in exceedance of a specified threshold (Peaks Over Threshold, POT). To give a flavour of the particular considerations, individual AM data points are on the one hand are more likely to conform to the assumption of independence. On the other, taking only the AM may cause flows which are not annual maxima but are nevertheless potentially important to be discounted. Taking a POT approach can ensure that more potentially relevant peaks are included. However, in this case, the choice of threshold (to which results can be sensitive; Beguería, 2005) is arbitrary. The present study employs the AM approach.

As mentioned previously, there is also a degree of choice regarding the statistical distribution and parameter estimation method to be employed. It has been suggested, for instance, that distributions founded in Extreme Value Theory (EVT), such as the Generalised Extreme Value (GEV) distribution, have a stronger theoretical basis than certain alternatives – annual series of block maxima are believed to converge to a GEV distribution as the sample size increases (Faranda et al., 2011). Despite this, as already mentioned, GL distributions often appear to fit the UK peak data better (IH, 1999; Kjeldsen and Jones, 2004). In terms of fitting methods, the main alternative to L-moments, which is the approach recommend by the FEH and hence employed in this study, is Maximum Likelihood Estimation (e.g. Martins and Stedinger, 2000).

Alternatives to statistical approaches to estimating present-day flood frequency are available. In many, the degree of process representation may be much greater than simply distribution fitting, although calibration using empirical data is still typically required. For example, return period peak flow estimates may be produced by following the FSR/FEH rainfall-runoff method, which was reworked a decade ago to give the Revitalised Flood Hydrograph (ReFH) model (Kjeldsen et al., 2005). This method requires the specification of hyetographs for storms of design return period in terms of total volume, duration and the profile shape (again using statistical modelling). Then, given some assumed antecedent conditions, a conceptual rainfall-runoff model is applied. The T-year flood assumed to be simply a product of the T-year rainfall event.

To overcome some of the limitations associated with such an approach, more advanced methods that involve continuous simulation of river flows, usually driven by some stochastically generated rainfall data either at individual sites or across multiple sites, (e.g. Calver et al., 2009; Grimaldi et al., 2012; Wheater et al., 2006) are increasingly favoured. Although both the ReFH approach and the continuous simulation approaches aim to represent effect of antecedent catchment conditions (e.g. soil moisture volumes) on flood frequency and severity, continuous approaches have the benefit of being less sensitive to uncertain initial conditions, since they are continually updated and so can be 'spun up' (Svensson et al., 2013). A possible detraction is that the rainfall generator must function well over the full range of intensities, including lower ones (*Ibid.)*

Both approaches can potentially benefit from fact that longer rainfall records are often available than river flow records (Jones et al., 2004). Being less than straightforward to conduct, continuous simulation in particular remains very much in the research domain. Although standard tools (previously a spreadsheet, although this has now been superseded) are available to ease the implementation of the ReFH method, due to the additional components in the modelling chain (and therefore extra sources of uncertainty), it is only recommended in circumstances which require a design volume (i.e. a full hydrograph).

2.1.3. Uncertainty in the measurement of high river flows

Hydrometric measurements are the responsibility of different monitoring authorities in each nation of the UK; the EA are responsible in England, the Scottish Environmental Protection Agency (SEPA) in Scotland, and so on. Once their quality is assured, the observations are compiled into the NRFA.

As already discussed, instrumental observations of river flows play a fundamental role in flood frequency estimation. It is therefore appropriate to consider some of the sources of uncertainty associated with such measurements. As one might expect, the measurement of very high flows, which are naturally most influential in flood estimation, is particularly challenging. Some specific sources of uncertainty are briefly highlighted.

According to Freer et al. (2013), the UK's hydrometric monitoring network was established primarily for low-flow (i.e. drought) monitoring. Perhaps because many gauging structures are designed for this purpose, and more generally because most natural channels do not provide sufficient 'control' in times of flood, out-of-bank flows of flood can bypass gauging stations. This often results in the underestimation of flood levels (Wilby et al., 2008). Incidentally, this is another possible explanation for the underestimation of flood frequency when compared to reality, although not when comparing one modelled estimate to another.

A second major source of uncertainty stems from the fact at the majority of monitoring stations, discharge is not continuously measured directly. Rather, only water stage (i.e. level) is recorded. This necessitates the application of pre-calibrated rating equations, which based on available and relevant contemporaneous stage-discharge measurements at each site, to associate stage and discharge. These relationships are often constructed from measurements made at low to moderate levels, and so in order to estimate flood discharges, extrapolation beyond their calibrated ranges is frequently necessary.

Furthermore, many stations may exhibit 'multi-stage behaviour', whereby the equation governing the stage–discharge relationship varies according to the stage range (Reitan and Petersen-Øverleir, 2009). In the example shown in Figure 2.4, for instance, a small change point seems to approximately coincide with the level of bankfull capacity, which is not atypical.

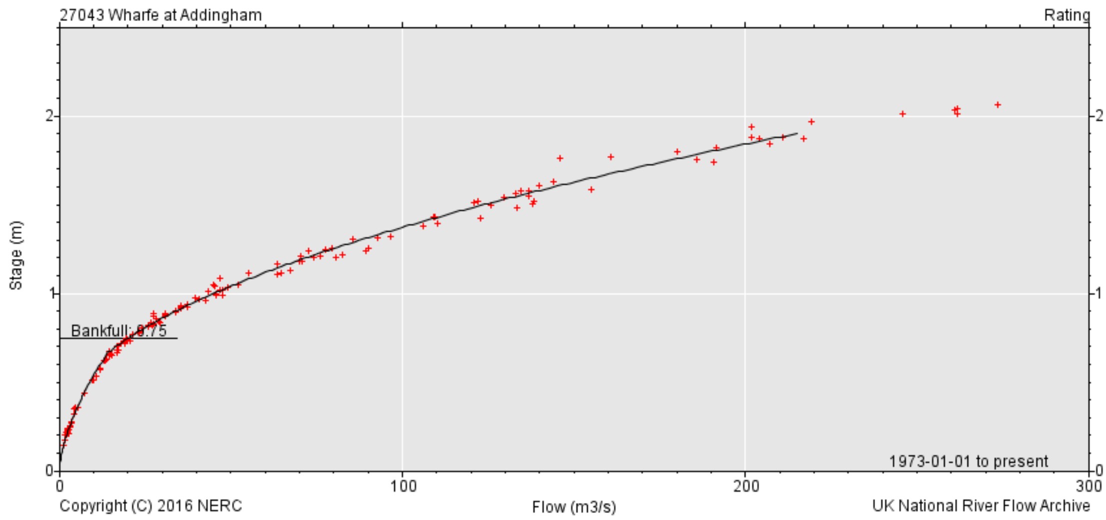


FIGURE 2.4. Rating curves relating water level with peak discharge for the River Wharfe at Addingham, West Yorkshire (27043). Here, a small but noticeable shift in the stage-discharge relationship occurs at approximately bankfull stage. Calibration points are show as small red crosses. This rating appears relatively well constrained; others demonstrate much more 'scatter' and complexity. Source: CEH.

Changes in channel geomorphology, instrumentation location and type over time, meanwhile, can impinge on flow record consistency (i.e. stationarity) (Hannaford, 2015). Unsurprisingly, the construction of reservoirs and dams can have has an even more pronounced effect, usually rendering any 'pre-construction' observations irrelevant (Archer et al., 2007). Alternatively, attempts may be made to 'correct' for past changes.

Finally, instrumentation failure always remains a possibility, with such failures surely more likely in times of flood. Consider Figure 2.5, for example, which shows the data returned from the station on the River Derwent at Malton, North Yorkshire (27858) during late 2015, obtained as part of the present study. Unlike most others, ultrasonic instrumentation is installed at this station and so rating curves are not required. (Discharge can be measured directly with such instrumentation; Herschy (2009; Ch. 12) provides a description of the method). The constant discharge measurements of 200 $m^3 s^{-1}$ are clearly erroneous (and were removed from the study).

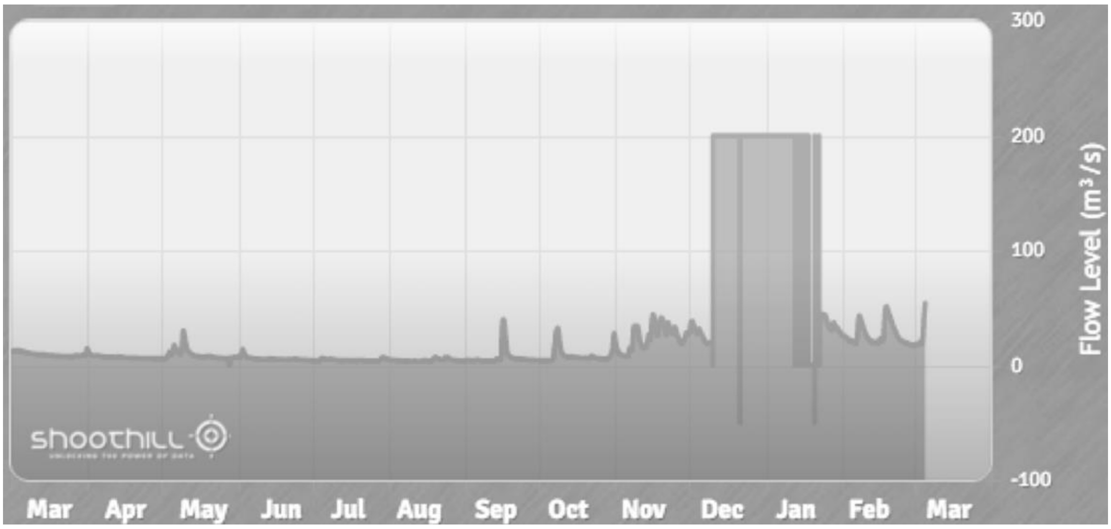


FIGURE 2.5. Time-series plot of discharge series for the River Derwent at Malton, North Yorkshire (27858) from March 2015 onwards. It is apparent that the data returned during late December and January, that is for much of the period when flows were known to be highest, are erroneous.

Coxon et al. (2015) recently quantified the overall uncertainty associated with UK instrumental discharge measurements at 500 stations. Although a wide range of overall uncertainty in discharge measurements (10–397%) was reported, at 80% of locations 'high-flow' estimates (defined to be *QMED*; so not particularly high in the context of this study) were found to be associated with uncertainties of less than 40%. The methodology employed was incapable of extrapolating beyond the limits of the stage-discharge measurements, however, and larger uncertainties would surely be expected in the region considered in this study. Another important finding of the work was that uncertainty was dominated by local sources (*Ibid.*), which complicates generalisation attempts.

The NRFA goes to some lengths to provide carefully document information on the characteristics of each catchment and station, and to identify data of questionable quality. Despite this, the practical difficulties associated with thoroughly assessing measurement uncertainty and record consistency can be considerable. This especially the case when the number of stations is high and the scientific focus is placed elsewhere, as it is here. Overall, it is apparent that the measurements upon which much subsequent hydrological science is based are inherently uncertain. River flow data, and most especially of all flood observations, should never simply be treated as 'correct', as they are in some studies (Clarke, 1999).

2.1.4. Summary

This section has re-emphasised the continued need for long-term flood hazard and risk assessment, which is distinguished from short-term forecasting. Standard statistical approaches to flood frequency estimation have described, along with some of their limitations and alternatives. Finally, with an emphasis on the attendant sources of uncertainty, the process by which river flow measurements are typically made has been explained. This material provides solid background from which the present study may develop.

2.2. Potential non-stationarity in UK peak flows

This section reviews recent research that has explored the challenging topic of whether flood hazard in the UK might be changing, that is to say, is non-stationary. It is clear that the mere occurrence of record breaking floods such as those of December 2015 does not constitute evidence for increasing hazard; records will continue to be broken whether trends are real or perceived (Matalas, 1997). Hence, some form of more advanced analysis is required.

In general, non-stationarity may take the form of abrupt shifts, cycles or monotonic trends. Of these (provided they happened in the past and were captured in the records), abrupt shifts in a time-series are usually the easiest changes to detect. This can be done either visually, or by searching for jumps in the mean calculated across rolling windows, for instance. Therefore, in actuality, most attention is paid the identification of cycles and trends. In the case of fluvial flood risk management, any possible changes in the probabilities of high river flow occurrence, as opposed to other components of the hydrograph (such as changes in seasonal or low flow distributions), are of most concern. Readers who are interested in such other aspects are referred to Hannaford and Marsh (2006) and Hannaford and Buys (2012), for example.

The topic is particularly pertinent for the reason given in Chapter 1; namely that whilst the assumption of stationarity is routinely made when conducting standard statistical flood frequency analyses, there is a growing perception that UK flood hazard or risk might be increasing, and that this could be as a result of anthropogenic climate change. If there is indeed an increasing trend in the frequency of peak flows (i.e. hazard), then flood risk in both the immediate and longer-term future might be being underestimated. It was felt that no study in the general field of flood hazard and risk assessment could overlook the topic of stationarity. Having said that, following the review presented below, it was decided that non-stationarity would not be explicitly considered in the subsequent research. Justification for this decision is presented in due course.

Conceptually, the numerous factors that could be responsible for changes in peak flows (and other elements of the hydrograph besides) are easily identifiable. They include natural weather and climate variability (Scaife et al., 2008), changes in rainfall patterns and other meteorological phenomena associated with anthropogenic climate change (e.g. temperature and snow accumulation) (Huntingford et al., 2014; Watts et al., 2015; Wilby et al., 2008), urbanisation and other land use changes (Beven et al., 2008; Wheater and Evans, 2009), and direct modification of the hydrological system (such as channel engineering, reservoir construction and abstraction; see e.g. Marsh and Harvey, 2012).

As has already been mentioned, in the aftermath of particular events, there is often intense debate about the extent to which past greenhouse gas emissions might have contributed (BBC, 2014; Hannaford, 2015; Watts et al., 2015), although care must be taken not to conflate meteorology (specifically precipitation) and flooding. Robust detection of past changes and subsequent attribution to underlying causal mechanisms, coupled with reliable projections of future changes, could lead to both more informed debate and improved policy responses. However, as this section reveals, meeting these aspirations is, at present, highly demanding.

2.2.1. Empirical analyses

Empirical analyses can be conducted to explore past changes. Jones et al. (2013) conduced regional frequency analyses on extreme rainfall over the period (1961–2009). Some evidence was found for increases in frequency, with variability in the changes depending on region, seasonality and event duration. In certain locations, decreases in return period estimates for a given volume were sharp. The authors suggested that “these result may have significant implications for flood defence design and planning” (p. 1178). However, catchment characteristics and non-linear physical processes mean that peak river flow responses need not necessarily follow changes in rainfall (Laizé and Hannah, 2010). For example, with respect to the Thames catchment, Hannaford (2015) states that despite increases in air temperature, tendency for winter rainfall and increased annual runoff all being observed, there is no trend in fluvial flood magnitude.

Accordingly, many studies have attempted to find evidence for past changes in more direct indicators of flooding (normally peak flows – either AM or POT – but sometimes also duration above a certain threshold, and so on). Various statistical approaches are available to interrogate instrumental flow records; Robson (2002) and, more recently, Chiverton et al. (2015) provide useful summaries. Least-squares linear regression with focus on the gradient, and the Mann-Kendall test (Kendall, 1975), are two often-favoured methods of investigate trends.

The study of Hannaford and Marsh (2008), which used such approaches, makes an important contribution. It reported that in near-natural 'benchmark' catchments[16] with only moderate record lengths (<55 years; most of which started in the 1960s), some upward trends in flood frequency and severity were apparent. Catchments in northern and western areas, which are under maritime influence, demonstrated this most clearly. In contrast, records from lowland catchments were found to demonstrate no obvious trends (*Ibid.*).

However, in such short records, the number of extreme observations are limited. Moreover, the extremes tend to exhibit a high degree of inter-annual variability. Consequently, establishing whether such trends are genuine is often extremely difficult on an empirical basis (Hannaford and Marsh, 2008; Pattison and Lane, 2011). Dixon et al. (2006) even showed that trends that can be considered statistically significant can be undermined if the sample start or end dates are arbitrarily changed. The results of such trend analyses are also known to be highly sensitive to precisely where high outliers are located temporally (Wilby et al., 2008). Prosdocimi et al. (2014), who conducted a more recent national scale examination of both peak flows and precipitation, summarise the situation nicely. They state that "one striking feature of the estimated trends is that the high variability found in the data leads to very inconclusive test results" (p.1125). An additional consideration that must be made when conducing such tests across a network of stations, with a view to arriving at an 'overall' statement, is their spatial correlation. Tests of 'field significance', which account for correlation between sites, must be applied to address this (Wilks, 2006). Guerreiro et al. (2013), for example, assessed the significance of changes in rainfall on the Iberian Peninsula in this fashion.

---

[16] Most research on hydrological change to date has been undertaken in 'near-natural' benchmark catchments (Bradford and Marsh, 2003). Whilst this approach is useful for distinguishing between natural and anthropogenic changes, any findings might have only general relevance to flood risk. More work on urbanising and urbanised catchments in future would therefore seem sensible

Attribution of any detected change on an empirical basis is even more problematic. Whilst inherent hydrological complexities mean that changes in several drivers may induce similar hydrograph responses, only one realisation exists with which to work (i.e. what actually happened), and so deconvolution to establish cause and effect is rarely possible. A notable example where empirical attribution has been possible is the recent study by Prosdocimi et al. (2015), who did manage to attribute observed changes in high-flow frequency and magnitude with urbanisation in a catchment in North West England by using a nearby non-urbanised catchment as the control.

In view of the difficulties that confront analyses of fairly short records, longer instrumental flow records have the potential to provide useful additional insight. Fortunately, some such records are available in the UK, although they are few in number. Their analysis has yielded two particularly important findings. The first is that there is very little, if any, evidence for significant long-term trends in UK peak river flows (Hannaford and Marsh, 2008; Hannaford, 2015). The AM record from the River Thames at Teddington (Figure 2.6), which is one of the longest continuous flow records anywhere in the world, is a case in point; even when the series is 'naturalised' to remove effects of historical abstractions, it is evident that the very largest peaks occurred in the earlier portion of the record.

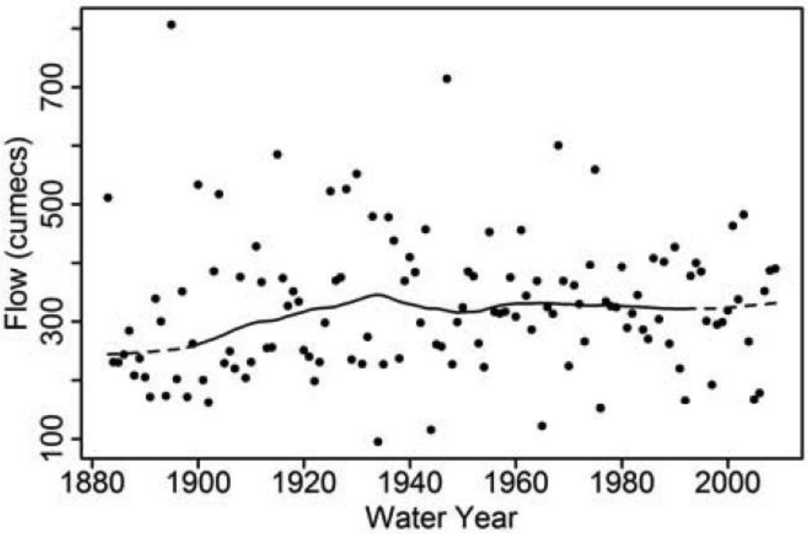


FIGURE 2.6. Naturalised annual maxima (AM) daily mean river flows for the River Thames at Teddington. Naturalised refers to normalisation to account for abstraction and other anthropogenic processes. The smoother line is a locally weighted regression smoothing curve (LOESS) (Cleveland, 1979). Source: Marsh and Harvey (2012).

The second important feature revealed by longer records is that flood peaks often tend to demonstrate some clustering into so-called 'flood-rich' and 'flood-poor' periods on multi-decadal timescales (Black, 1995; Hannaford and Marsh, 2008). These fluctuations are believed to be the signature of some form of climatic variability, with changes in westerly airflow associated with variability in the North Atlantic Oscillation (NAO) – which is the dominant mode of climatic variability in the UK – a possible contender. Indeed, an association between positive NAO phase and high flows has already been exposed (Biggs and Atkinson, 2011). That said, Jones and Quinn (2013) found little change in flood generating synoptic systems since the 1930s that could easily explain these intervals, which demonstrates the degree of complexity involved.

Whilst on the topic of flood-rich periods, it may be noted that one possible explanation for the upward trends identified in some of the shorter records is that the opening of many gauging stations in the 1960s could have coincided with the start of one; since this period, NAO has generally been positive (Hannaford and Marsh, 2008). Nevertheless, variability on shorter frequencies is strong, and the recurrence of any flood-rich periods remains poorly understood. Should it become possible at some point to forecast with reasonable skill whether the future (on an appropriate timescale) might be relatively flood-rich or poor, then one could theoretically adjust future flood probabilities relative to those given by the integrated records such that they are more appropriate to the specific timescale in question.

Lastly, Hannaford (2015) reviewed whether there might be any specific evidence for climate change impacts specifically in empirical UK river flow series. It was stated that although some of the regime changes reported above are consistent with future projections (see the next section; Section 2.2.2.) are present, others appear to be in contradiction. From this, it was ultimately concluded that any observed changes cannot generally be attributed to climate change.

2.2.2. Coupled climate-hydrological model simulations

In contrast to empirical approaches, the application of climate models allows synthetic experiments to be conducted (e.g. changes in land use on flood frequency) and possible future changes in flooding to be explored (Prudhomme et al., 2010; Cloke et al., 2013). Focussing on the latter, a long chain of coupled models is normally required. Firstly, due to the coarseness of Global Climate Model (GCM) grids, Regional Climate Models (RCMs) (which are produced by means of dynamical downscaling, and hence are nested within GCMs) are better suited to hydrological applications. Given an assumed future emissions scenario (or scenarios), future rainfall data may be generated. It is normally necessary to further downscale these precipitation data so that they are 'meaningful', i.e. there is some spatial variability, at the catchment scale. The opportunity to correct biases in the models' output rainfall, as diagnosed by their inability to replicate observed statistics over the historical control period, is often also taken at the downscaling stage (Kay et al., 2009; Lafton et al., 2013).

It is then possible to force previously calibrated rainfall-runoff models by these simulated future rainfall data. In this way, continuous time-series of future river flows can be generated, theoretically in a large number of catchments. Thereafter, changes in flood frequency expected by some future (simulated) time period of interest can be calculated and compared to statistics representing the present. For instance, if 'the 2080s' are the period of interest, then flow quantiles from the period 2069–2099 might be estimated. The 'present' estimates might be produced from 1980–2010 observations (Prudhomme and Reynard, 2009). Of course, the usual challenges associated with flood frequency estimation come into play, and so the significance of any changes should ideally be viewed in light of the estimation uncertainty. An example of a national scale UK project that follows this general outline is called Future Flows. Transient, 1-km resolution climate projections were used to generate daily flow time-series in 281 catchments for the period 1951–2098, for 11 ensemble members (the role of ensembles is explained shortly) (Prudhomme et al., 2012; Haxton et al., 2012).

In terms of results, many such studies indicate that peak flow return periods are expected to 'shorten' in future (Fowler and Kilsby, 2007; Bell et al., 2012; Kay and Jones, 2012; Quinn and Horswell, 2014; Smith et al., 2014). Wilby et al. (2008) describe the difference between this prediction and the lack of trends in most observed series so far as an "emerging mismatch" (p. 2511).

Numerous limitations and sources of uncertainty are associated with this approach. They include irreducible uncertainty with respect to the evolution of future greenhouse gas emissions, uncertainty in GCM structure, the severe inability of RCMs to replicate observed rainfall over the historical period (Cloke et al., 2013), structural and parametric hydrological model uncertainty (e.g. New et al., 2007), and uncertainty in flood frequency analysis (Cameron, 2006; Smith et al., 2014). The failure of the models to replicate observed rainfall, for example, necessitates the use of bias correction factors that are conditioned on a past control period. Despite knowledge of the highly non-linear nature of the climate system, it is assumed that the adjustment factors will hold for the future (Beven, 2011). Moreover, multiple valid corrections can produce highly contrasting hydrological outputs (Fowler et al., 2007).

These uncertainties are often compounded as they propagate through the modelling chain, as the conceptual diagram of Wilby and Dessai (2010), reproduced in Figure 2.7 illustrates.

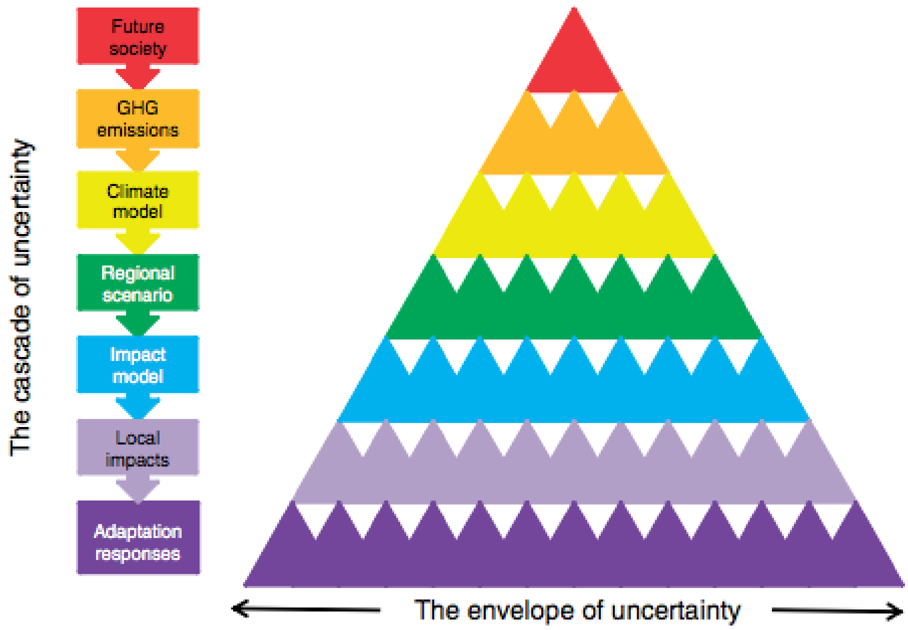


FIGURE 2.7. The cascade and envelope of uncertainty in future flood impact studies. The increasing number of triangles as one proceeds down the cascade reflects the increasing number of permutations at each level. Source: Wilby and Dessai (2010).

Kay et al. (2009) evaluated the significance of the contribution made by no less than six different sources of uncertain to overall uncertainty, finding those related to GCM structure to be particularly notable.

In order to represent some of the uncertainty and variability, ensemble simulations are often conducted. In addition to the divergent results (both between and within catchments) that different ensemble members are prone to produce, the computational burden associated with this approach is high. When national scale rather than only catchment scale analysis is required, and (or) when more than one plausible emissions scenario must be considered, and (or) several future time slices are of interest, both the degree of variability in the results and the computational challenge may grow exponentially.

As an alternative to full simulation, approaches based on 'response surfaces' have been recently pioneered (Prudhomme et al. 2013). These surfaces, which are produced via sensitivity analysis, reduce the computational burden involved. Figure 2.8 shows expected changes (and their ranges) in 1-in-20-year peak flows by major river-basin region for the 2080s generated in such a fashion.

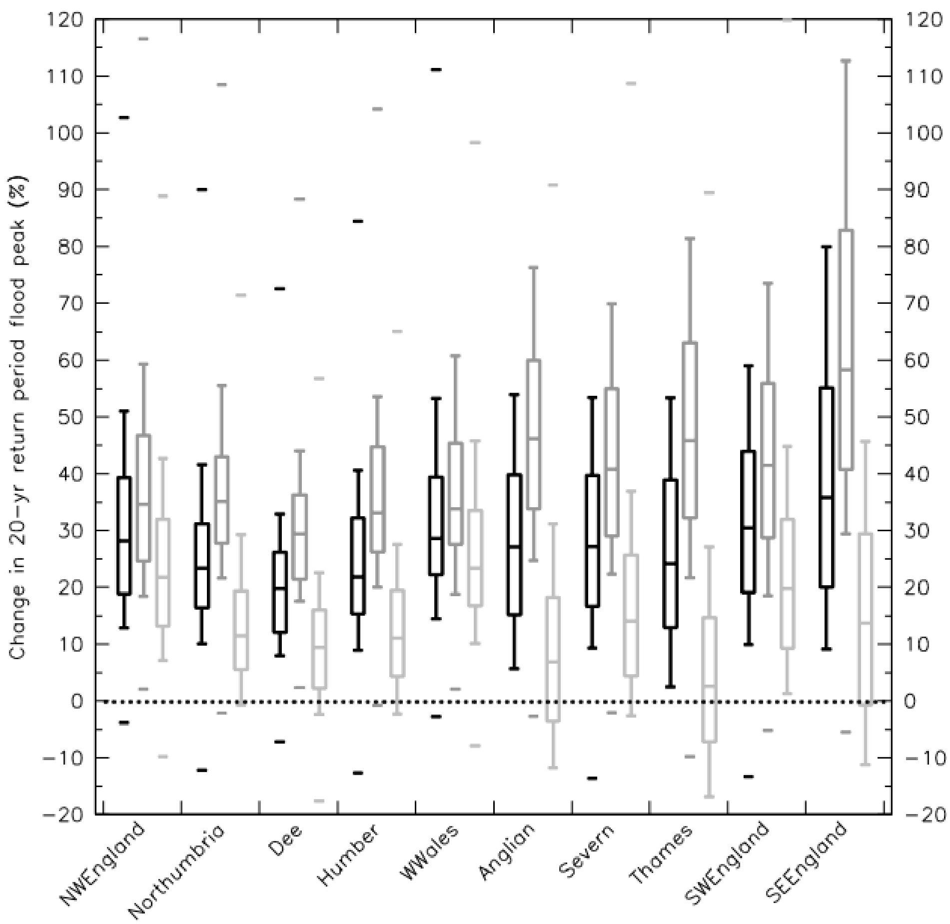


FIGURE 2.8. Central estimates of regional climate change flood impacts ranges for the 1-in-20-year annual probability of exceedance (black) for river-basin regions in England and Wales under the medium UKCP09 emissions scenario (Murphy et al., 2009).The upper and lower additional grey box-plot diagrams for each region show the results with the +2 and -2 standard deviation response surfaces added to the central estimate for each projection. This reflects the uncertainty from using composite response surfaces to represent what is actually a range of possible catchment responses classified as the same response type. Whiskers show the 10th – 90th percentile range, boxes the interquartile range (IQR) and median. Source: Kay et al. (2014).

As the few examples that have been cited demonstrate, simulation-based assessments of future flood hazard tend to focus on fairly distant future periods. This is presumably because climate change impacts on high flows do not emerge clearly from natural variability in the shorter term. It does, however, mean that this approach is capable of providing little assistance to problems requiring a relatively short-term view of future flood hazard.

Even when the problem at hand does relate to the longer-term future, due to the complexity of the modelling chain, the computational resources required, and the large uncertainty/variability typically associated with the results, coupled climate-hydrological modelling does not represent a viable means by which practicing hydrologists might generate future flood frequency estimates.

Rather, the established statistical method based on past flood event data prevails. Having said this, the results of model-based simulations of the type described above do play some part in ensuring that hydraulic design and critical floodplain infrastructure is resilient to possible, albeit highly uncertain changes in flood hazard that might be realised over their intended lifespans (EA, 2011). Specifically, official guidance specifies relatively crude uplift factors that should be applied to 'baseline' design flow estimated, under the assumption of stationarity, in the traditional manner (Wilby and Keenan, 2012). Until recently, a uniform 20% uplift was advocated for any period between 2025 and 2115, and for any location in the UK (Defra, 2006); an approach which was considered precautionary (Reynard et al., 2009).

Following further modelling work (Prudhomme and Reynard, 2009), the guidance has been updated. In particular, it is now advised that these allowances be varied both spatially and according to the nature of the application at hand (for example whether or not the design relates to 'essential infrastructure') (EA, 2016b). Especially where they are fairly precautionary, it could perhaps be suggested that these guidelines provide some 'leeway' for any potential contravention of the initial assumption of stationarity; essentially, when such allowances are applied, any potential non-stationarity is 'dealt with' in a phase following flood estimation. Having said that, it is established that over-design can, in its own way, be just as costly and under-design.

2.2.3. Probabilistic event attribution

Recently, an extremely novel method of exploring the links between climate change and contemporary flood events has begun to emerge. These so-called probabilistic event attribution studies (e.g. Kay et al. 2011; Pall et al., 2011; Schaller et al., 2016) also rely on the coupling of atmospheric (weather) and hydrological models. They attempt to quantify the contribution made by past greenhouse gas emissions to the weather that led to recent flood events, and thereby attribute a proportion of the 'risk' directly to this cause (note that in these studies, risk is not always used in the strictest sense defined earlier, but rather is used to refer to probability or likelihood). As a consequence of the large degree of natural variability (independent of any anthropogenic influence), accessing enough computational capacity to run the very large number of high-resolution simulations necessary for meaningful conclusions as to any anthropogenic contribution to be drawn from such studies has been an important consideration in the past. Massey et al. (2015) describe an innovative solution.

The most recent study of this nature is that of Schaller et al. (2016). In this work, a hydrological model was forced by many simulated realisations of the weather of Winter 2013/2014 under both 'counterfactual' pre-industrial conditions and present day atmospheric concentrations. Anthropogenic emissions were found to have increased the 30-day peak flow on the lower Thames associated with the 1-in-100-year monthly precipitation by a best estimate of 21%. Despite the variability, this result was considered to represent a relatively strong climate change 'fingerprint' on the balance of probability (JBA Trust, 2016). Interestingly, a much less pronounced impact on peak flows – which correlate much more closely with flooding – was detected (best estimate: +4%).

Nevertheless, as the previous statement alludes to, wide uncertainty bands are associated with these results; the 30-day best estimate statistic reported above, for instance, was associated with an uncertainty range of -12% to +133% (Schaller et al., 2016). Boundary condition uncertainties and inherent variability contribute to this range, however (i.e. the uncertainty is not only that associated with anthropogenic impacts).

By their nature, these experiments support only quite nuanced conclusions regarding the contribution that anthropogenic climate change has made to past individual floods. These conclusions can be rather difficult for non-experts to interpret. Furthermore, because the methodology has so far only been applied to a few events, it is not yet possible to make more generalised statements about the extent to which flood hazard might have increased due to past greenhouse gas emissions (i.e. across all flood types, or flood-causing meteorological setups). Consequently, at this early stage in the development of such research, practising hydrologists and flood risk managers might, with some justification, be unclear as to how best respond to such findings.

2.2.4. Summary

The large and growing literature on the subject of possible non-stationarity in UK fluvial flood hazard attests to the considerable importance of this topic. From the review that has been presented here, it is apparent that the insight empirical assessments are capable of providing is often limited by the records being short and 'noisy'. That said, the longest instrumental records provide little indication for any notable long-term trends in either the frequency or severity of high flows. Simulation-based approaches can be applied to explore possible future changes, although their results tend to be highly uncertain, and moreover typically only relate to relatively distant future periods. For these and more practical reasons, the approach is not yet seen as viable means by which routine flood frequency estimates, including for future hydraulic design, might be made. Such modelling work has, however, informed climate change allowances that are often combined with flow estimates produced conventionally. Probabilistic attribution studies show much promise, and in cases suggest that anthropogenic emissions are already contributing to enhanced flooding, although this is by no means a clear picture across all events that have been assessed in this way. An important additional point is that introducing non-stationarity requires more complex models, and hence yet more uncertainty (Montanari and Koutsoyiannis, 2014; Serandi and Kilsby, 2015). Finally, maintaining the assumption of stationarity enables existing industry practice to be followed as closely as possible. For these reasons, the issue of non-stationarity is placed largely to one side for the remainder of this thesis. In the next section (Section 2.3), attention turns to an arguably even more fundamental reason as to why flood frequency might be routinely or systematically underestimated.

2.3. Is flood frequency underestimated per se?

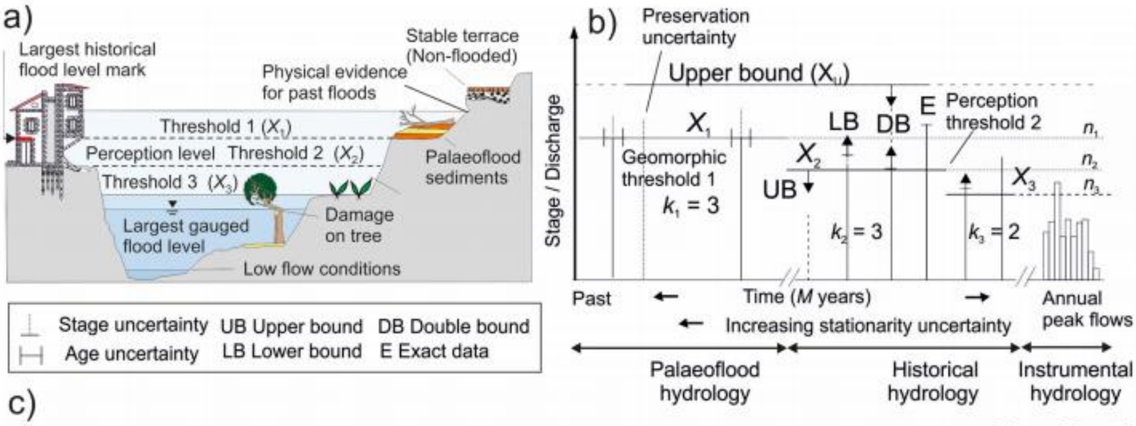
The possibility that flood frequency relationships may be underestimated 'per se', i.e. even notwithstanding any possible increasing trends associated with anthropogenic climate change, was introduced in Chapter 1. One could even contend that this possibility has been somewhat overlooked at the expense of the highly politicised 'hot topic' of climate change. This section, and to a certain extent the subsequent research, seeks to redress this.

Much of the work that is cited here in support of the hypothesis that flood frequency might be underestimated on the basis of short records involved extending records back in time. If this possible explanation is to be proposed, therefore, then the assumption of long term stationarity over the studies periods must be established at the outset. More precisely, it must be assumed that flood hazard was not significantly 'higher' or 'greater' in the historical periods to which these studies relate.

2.3.1. Support for the short-record underestimation hypothesis: historical and palaeohydrology

Historical flood hydrology may be defined as the study of floods that occurred before the advent of instrumental (a.k.a. systematic) measurement, but which were recorded by humans in other ways (e.g. documentary records or epigraphic marks) in such a way that useful information may be recovered (Bayliss and Reed, 2001; Brázdil et al., 2006, Stedinger and Cohn, 1986). Palaeoflood hydrology, meanwhile, involves reconstructing floods that were neither measured systematically nor recorded by humans in other ways, but which left traces in the natural landscape. Baker (1987) presented a relatively early review of some of the approaches that might be taken in palaeoflood studies, whilst Baker et al. (2002) sought to deconstruct some of the perceived limitations associated with such an approach. A more recent summary of the state of the discipline, albeit with a strong focus on the United States, is provided by the same author (Baker, 2008).

Both disciplines theoretically enable longer-term flood chronologies, which have potential to bring considerable benefits to the task of flood frequency estimation, to be reconstructed. Of course, the temporal scales supported by systematic (i.e. instrumental), historical and palaeoflood methods can crossover to some extent. For instance, depending on the arrangement of monitoring networks, very recent floods in ungauged locations may fall best under the 'historical' definition. With reference to floods in Continental Europe, Benito et al. (2004) evaluated all three data types and their potential for integration. Figure 2.9 provides an excellent comparison of the approaches, whilst Figure 2.10, in a somewhat simpler fashion, shows a classic illustration of the additional insight that reconstructed flow data can provide.



c)

Flood data source	Resolution	Flood stage/discharge	Record length
Instrumental hydrology: Annual maximum/partial duration series from gauged records	Daily	Flood water level known / eventually problems with large flow water levels	decades
Historical hydrology: Documentary archives, recorded marks, observed water level, newspapers	Annual-daily	Flow level known or not; censored data with stage above perception level	centuries
Palaeoflood hydrology: Slackwater flood deposits, silt lines, boulder berms, erosion lines, scars and damage on trees	Annual-decades or unknown	Minimum water stage known; censored data with stage above geomorphic level	millennia

FIGURE 2.9. Sources of quantitative flood information. (a) Sketch of a cross section showing various flood level indicators from palaeofloods (sediments and damage on trees), and documentary-based floods (i.e. those able to cause damage or socio-economic disruption). For historical hydrology, only floods exceeding a flood level related to a perception threshold (X_i) over a period of n_i years ($n_1 > n_2 > n_3$) are recorded. Palaeofloods from stratigraphic records are related to geomorphic thresholds. (b) Organization of historical and palaeoflood data, using the described thresholds (X_i), and multiple types of observations to support flood frequency analysis. K_i corresponds to the number of flood peaks during the last n_i years that exceeded the X_i threshold but not the X_{i-1} threshold. Upper bound level (X_u) may be used to limit the maximum discharge. Data types: E: flood peak is known. LB: flood was bigger than X_i which is known; UB: the upper flood level of known magnitude (X_u) was not exceeded over a certain time period. DB: flow level was within the interval given by X_u and X_i. (c) Data source characteristics, timing, stage information, and typical temporal framework of systematic (instrumental) and non-systematic data (palaeoflood and documentary evidence). Source: Benito et al. (2015).

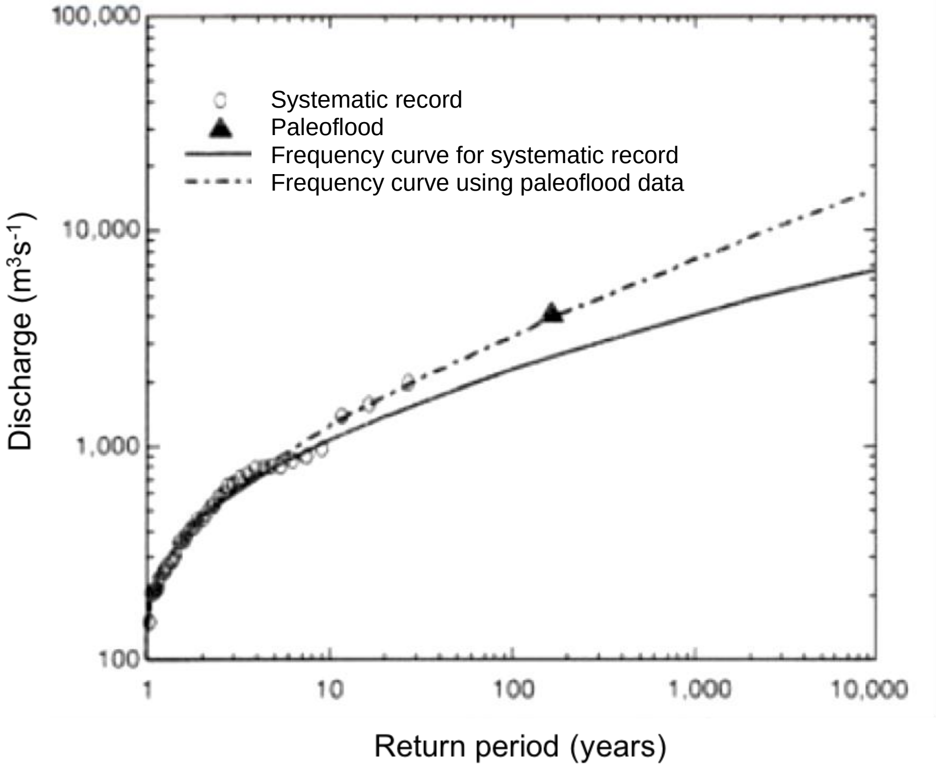


FIGURE 2.10. Illustrative flood frequency curves for the systematic gauge record alone (solid line) compared to a curve that incorporates paleoflood data (dashed) for the San Juan River, Utah, United States. Source: Orchard (2001).

As the date of some of these citations reveals, the realisation that extended records can add value to flood hazard assessments exercises is not an entirely new one. Indeed, in a UK context, the inclusion of historical data was advocated in the FSR (NERC, 1975). In a similar spirit, the sixth 'maxim' of flood estimation, as set out in the FEH, is that that "there is always more information" (EA, 2012). However, despite this and the beginnings of a movement towards a standardised approach for the inclusion of historical data in flood estimation at the turn of the century (Bayliss and Reed, 2001), there has been little concerted uptake of these methods by practising hydrologists. Possible reasons include the time-consuming nature of searching for the necessary data and a lack of dedicated software programs to ease their integration.

That said, a small but steadily increasing number of UK-focussed studies have sought to investigate the impacts of incorporating alternative forms of hydrological data pertaining to longer time periods. It must of course be noted that where historical flows can be reconstructed, they are likely to be somewhat uncertain. Even still, the benefit of the longer record might represent a 'net gain' (Macdonald et al., 2014).

Including such data in the model fitting process can be technically challenging. Archer (2010) describes an adaptation of the maximum likelihoods method of Stedinger and Cohn (1986) which facilitates the inclusion of historical data. Specifically, the distribution that best fits a combination of gauged and historical information can be identified. Importantly, not only historical floods whose flows could be estimated but also those that could not be fully quantified but which were believed to have exceeded a certain threshold can be dealt with. Of course (where flows can be quantified), as with systematic measurements, frequency-magnitude plotting positions of historical observations can be determined using the Gringorten formula (Gringorten, 1963), allowing the resultant model fits to be inspected visually. Further information on the graphical procedures which enable gauged and historical data to be plotted on the same scale, and other considerations related to the inclusion of historical data, are provided by Bayliss and Reed (2001). Next, some key findings arising from such work are summarised below.

Many years ago now, Archer (1987) extended the gauged record for the River Wear at Durham, which began in 1958, by estimating historic flood discharges from 1771 onwards. The extended catalogue suggested that using the gauged record alone was likely to lead to the severe underestimation of flood frequency at this location. Bayliss and Reed (2001) presented a similar case study of the River Avon at Evesham, Worcestershire (54002). Here, textual sources describing historical floods were evaluated for authenticity and completeness, and a ranked flood series was produced. Approximate flow magnitudes were then assigned using stage-discharge relationships developed from the systematic records. Once again, the inclusion of historical data indicated that conventional methods seem to underestimate flood frequency.

Slightly more recently, Black and Fadipe (2009) used Manning's equation to reconstruct discharges for the notable floods of 1968 and 1829 from available water level records at four sites in the Spey Catchment, Scotland. Compared to estimates produced by standard pooled analysis of the instrumental records, the 1-in-100-year flow was found to increase by more than 50% at three of the four sites upon inclusion of the historical data. Remarkably, Archer (2010) reported that including historical data increased flood estimates at all 12 of the locations he considered in the northeast of England, although the gauged record covering a relatively flood-poor period was proposed as a possible explanation for this. Finally, Foulds and Macklin (2016) took a somewhat different approach, producing multi-centennial length flood records from lichen-dated torrential sedimentary deposits from upland areas across the UK. Their subsequent analysis indicated that twenty-first century floods are relatively unremarkable in terms of both frequency and magnitude when placed within this longer context. The strong conclusion that "reliance solely on mid-late twentieth century flood series is underestimating current risk in the UK uplands" (p. 268) was drawn.

For balance, it must be noted that the estimates do not necessarily always increase. For example, Archer et al. (2007) found that single-site analysis results for the River Tyne at Bywell, Northumbria (23001) broadly aligned with a historically derived frequency-magnitude curve. (In this case, the pooled analysis results were not considered appropriate, and were rejected, with gravel extraction and reservoir construction (Kielder Water) in the past highlighted as factors which complicate flood estimation in this catchment. More generally, it is increasingly recognised that the impacts of such local factors on flood estimation can be large more generally, and efforts are being made to better include such information; Dixon et al., in press). Examples of studies which have reported reductions in flood frequency estimates with longer-term data even exist. Macdonald and Black (2010), for instance, reassessed flood frequency for the River Ouse at York. Introducing data going as far back as the year 1200, they suggested that historically informed estimates were both "lower and considered to be more credible" than those produced by conventional means (p. 1152). Finally, Macdonald et al. (2014) considered the Sussex Ouse at Lewes, East Sussex. Here, historical data was again found to slightly decrease the frequency-magnitude curve.

Such results emphasise the more general benefit that inclusion of historical data often brings – a reduction in estimation uncertainty. The uncertainty around the 1-in-100-year flow estimate was reduced by 40% relative to the systematic uncertainty in the case presented by Macdonald et al. (2014). Essentially, an important suggested benefit of employing longer records is their ability to 'iron out' some of the variability in conventionally made estimates that is simply a function of whether there happened to be relatively many, or relatively few, large peaks in the instrumented period. Irrespective of whether flood hazard estimates increase or decrease, it is clear that in many cases, and despite their own (sometimes perceived) uncertainties, making use of longer records can reduce uncertainties in the estimation of high flow frequencies.

Yet, even with the best of intentions, incorporating such additional information across many locations (perhaps in some form of national scale assessment) would be extremely difficult from a practical perspective. In additional to the effort that sourcing and compiling historical data or palaeo-records demands, another concern which accompanies the inclusion of data on past floods, especially very old ones, is that the climatic regime and catchments conditions under which historical and palaeoflood were produced might not be relevant to the future. In other words, that some long-term stationarity may exist (e.g. the emergence of Europe from the Little Ice Age). This concern is represented as 'increasing stationarity uncertainty' in Figure 2.10.

2.3.2. Another means by which longer flow records might be generated: continuous simulation

The general problem associated with short records in flood frequency estimation – that whatever characteristics the limited sample happens to capture may exert a strong influence on the results – has been well appreciated for some time, and has already been discussed at some length in this thesis. In particular, a form of regionalisation (pooling) and the augmentation of systematic records using alternative, longer-term data to effectively extend the number of station-years worth of data available at locations at which estimates are required have been summarised.

A third broad approach, which has not yet been considered, involves generating synthetically extended continuous flow time-series using hydrological (a.k.a rainfall-runoff) models (Lamb and Kay, 2004; Calver et al., 2009). Rainfall records tend to be much longer than river flow records in the UK (Jones et al., 2004), perhaps because rainfall is by far the easier quantity to measure. As such, by forcing rainfall-runoff models with these longer series directly, flow series can be extended back in time. Examples include the studies of Jones and Lister (1998) and Jones et al. (2007).

Alternatively, synthetic (stochastic) rainfall data may be generated from existing records and then used to force runoff models. This approach may enable some of the variability that is as yet unobserved but is nevertheless possible variability even under present conditions to be encapsulated (Burton et al., 2008; Wilks and Wilby, 1999). Where simulations are long enough, flood frequency statistics can be directly extracted from the empirical distributions. Further still, climate forgings may be incorporated into the stochastic generators (Cameron, 2000; Kilsby et al., 2007). Some drawbacks appear to be limiting the uptake of this method for practical flood estimation also. Certainly discharge predictions produced by both conceptual and physically-based, spatially distributed hydrological models for given ('known') rainfall inputs are always somewhat uncertain. Model structure and parameter uncertainties (Beven, 1993, 2006; Wagener et al., 2003; Gupta et al., 2005), as well as the challenge of their transfer to ungauged locations (Blöschl, 2013; Wagener and Wheater, 2006), all contribute to overall predictive uncertainty.

From the discussion hitherto, it is clear that no single ideal or preferable method for flood frequency estimation exists. Rather, each has its own particular benefits and weaknesses, which is an important conclusion in itself. Presently, each of the 'alternative' methods are less straightforward to implement than the established statistical one. This is a crucial consideration given the time and budgetary pressures practising hydrologists often work under. In response, comparing the results given as many different methods as possible, as Macdonald (2013) suggest, represents an eminently sensible approach. In this context and spirit, regularly revisiting results given by the established statistical methods as new systematic data becomes available – the approach taken by this study – retains the potential to yield valuable insight.

2.3.3. Updating estimates with new instrumental observations of exceptional floods: an analogy to the present study?

The study by Miller et al. (2013) may be analogous to the present one. Flood frequency in Cumbria was reassessed following the severe 2009 floods using the extended instrumental series, with particular consideration given to the effects of lakes on river levels. As Figure 2.11 shows, a considerable shift in even the enhanced single-site analysis (i.e. pooled) analysis results, which should be less sensitive to the addition of a single observation than single-site analyses, was seen.

A criticism that might be levelled at 'reactive' studies of this nature, including the present one, is that by reassessing flood frequency so rapidly (only) following floods that are known to have been extraordinary or record breakers, a bias is introduced in the sense that estimates are almost certain to rise (Beran, 2002; Luo, 1987). Having said that, although the inclusion of record breakers may bias estimates upwards, their exclusion may bias them downwards (Archer, 2010). Thus, not including the latest data – effectively pretending that the events did not happen – simply enables the status quo to prevail.

Indeed, especially where the records are relatively short and so the estimates unstable, the growth curves may increase sharply immediately flooding a notable flood, as shown, but then decrease steadily if reassessed subsequently as time passes without any further exceptional floods. If there was some feeling that flood estimates are generally stable and robust (i.e. approximate well the underlying distribution) in the first place, then excluding additional data could be acceptable. Unfortunately, as many of the examples cited show, changes to the data and methods continue to induce large increases. As such, and whilst acknowledging its potential drawbacks, the experimental design proposed – which hinges on the inclusion of the additional data – was deemed appropriate.

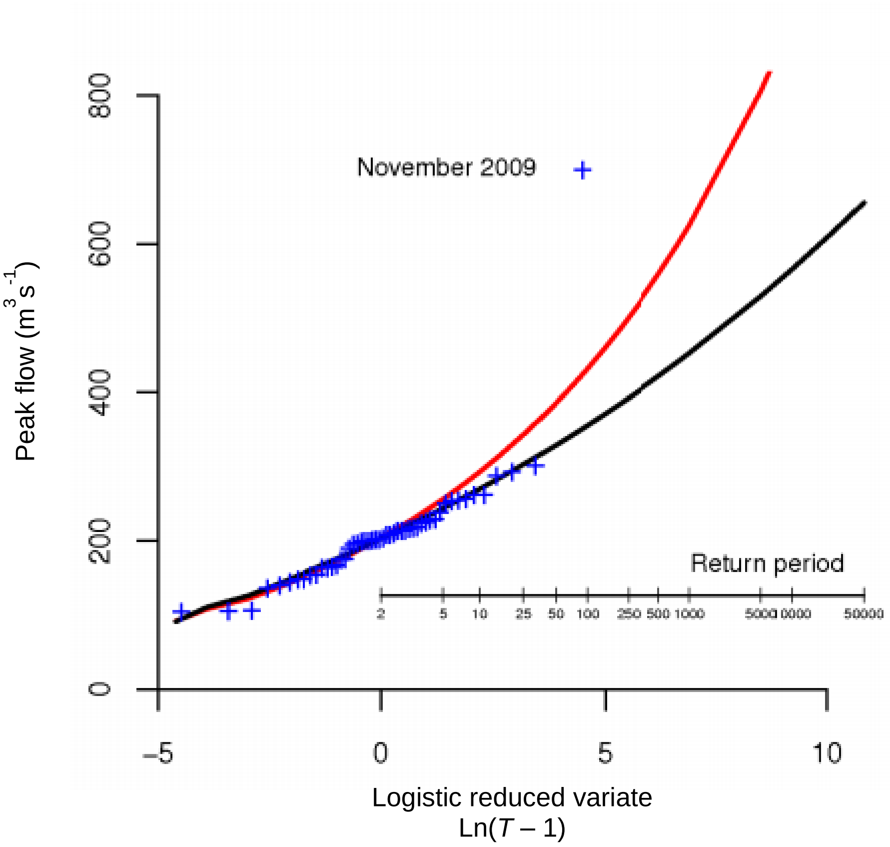


FIGURE 2.11. Flood frequency curves (enhanced single-site method) for the River Derwent at Camerton (75002) prior to the November 2009 event (black) and including the event (red). Source: Miller et al. (2013). The authors note that the plotting positions of such large floods can be extremely uncertain.

A final, important related point of note, which is emphasised by both Matalas (1997) and Beran (2002), is that the occurrence of record breakers is not itself evidence of non-stationarity.

Chapter 3

Research methods

3.1. Introduction and overall workflow

This chapter describes the methods that have been followed in order to address the objectives. Figure 3.1 provides a simple illustration of the overall workflow that was followed.

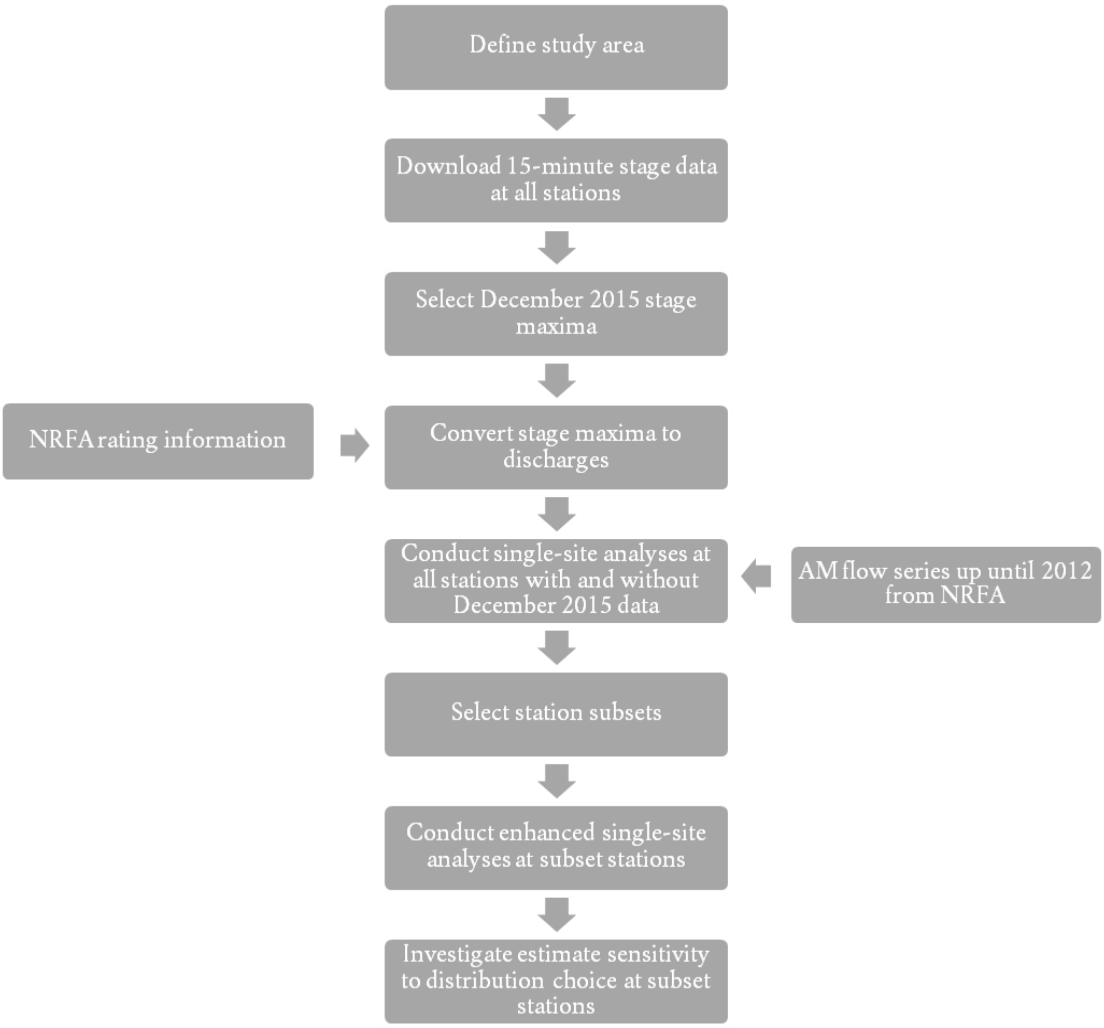


FIGURE 3.1. The workflow that was followed in order to address the objectives of the study.

3.2. Study area

Following initial assessment of the patterns of extreme UK rainfall and river discharges in late 2015, a study area intended to encompass all parts of northern England that may have been affected was defined. This area is shown in Figure 3.2.

Topographically, the area is divided along a broadly north-south axis by the Pennine hills. It includes the uplands of the Lake District (which rise to the highest point in England, Scafell Pike at 987 m), the Yorkshire Dales and North York Moors. Slope angles are steep in places. Vegetation cover across the study is rather varied, reflecting the diverse quality of the agricultural land (Natural England, 2016). The region is predominantly underlain by Carboniferous limestones, mudstones and sandstones, although some older metamorphic rocks are present in the Lake District (BGS, 2016).

Several major urban conurbations including Manchester, Liverpool, Leeds, Sheffield and Newcastle lie within the area, as do numerous smaller towns and villages. In steeper regions such as the Calder Valley in West Yorkshire, settlements are clustered in the valley bottoms, where their industrial growth was aided by abundant water supplies. Even in flatter areas, many settlements have developed alongside rivers.

FIGURE 3.2. The region considered in the present study, which was affected most severely by the flooding in December 2015. Other parts of the UK, notably northern Wales and southern Scotland were also affected by flooding over the winter, but flooding in these areas is not considered here. All original maps in this thesis were produced open-source software, QGIS; http://www.qgis.org/en/site/

3.3. Hydrometric data

3.3.1. Obtaining and pre-processing December 2015 river observations

The length, spatial distribution, and archiving standards of environmental datasets in the UK are generally excellent, and hydrological data is of no exception. Although the present study focuses on river flows, readers may be interested that much past rainfall data can be obtained from the Met Office[21]. A gridded precipitation dataset called CEH-GEAR (Keller et al., 2015) has also recently been developed with hydrological applications in mind (with the benefit that 'gaps' in time and space are interpolated).

River measurements are the responsibility of different monitoring authorities in each nation of the UK. Therefore, limiting the scope of this study to England meant the necessary 'event' flow data (which are distinguished from past series and other information, e.g. rating equations) could be obtained from a single source – the EA.

15-minute instantaneous measurements at all returning gauging stations within the study area were downloaded from the EA's Real Time flood-monitoring API[22]. In this way, 15-minute time-series of river observations spanning December were produced, initially for 166 stations. The database management system PostgreSQL[23] facilitated efficient data processing and storage. While direct measurements of river discharge were available at a small number of sites at which ultrasonic flow gauges had been installed, at the vast majority of monitoring locations, only stage was recorded. Stage measurements were made in metres relative to a local site datum.

Once the data acquisition phase was complete, the stage time-series were plotted and visually inspected. In this way, some clearly erroneous data such as that caused by the instrumental failure at Malton (27858), highlighted in Section 2.1.3, was identified. This station was removed from the study. Following this, at each site, the maximum stage observation (or, where ultrasound instrumentation was installed, the peak discharge) during the entire month were identified. The peak discharge data measured directly via ultrasound were then put to one side whilst the further processing of the water level observations that was necessary was conducted.

[21] https://data.gov.uk/metoffice-data-archive
[22] https://environment.data.gov.uk/flood-monitoring/doc/reference
[23] http://www.postgresql.org/

3.3.2. Matching EA and NRFA station reference numbers

NRFA collates, checks, and eventually publishes river flow observations from all UK measuring authorities. Daily mean discharges[24], AM series, and rating equations[25] corresponding to a large number of stations are all available from this source. Because flow data are only added to the NRFA database some months or even years after having been observed, the data used in this study were obtained directly from the EA, as previously described.

Perhaps a little strangely, the EA and CEH do not refer to monitoring stations using a common reference system. (All monitoring station references quoted here are NRFA references). For a given station to be included, the December event peak flow, the appropriate rating equation and the previous AM series all had to be available. This necessitated the matching of EA stations which provided event data with the corresponding station on the NRFA database. Initially this proved problematic, although a lookup table was eventually obtained. 11 of the 166 stations providing peak December observations could not be matched with a NRFA station reference, contained obviously erroneous data, or else did not have rating or previous AM data available. Thus, data from 155 stations were taken forwards.

3.3.3. Estimating December 2015 peak flows

For stations at which only peak water levels were available, appropriate site and stage-specific rating curves were applied to estimate peak discharge. At many stations, rating specific to different stages are provided. In such circumstances, the most appropriate given the observed stage was applied. The necessary rating information was obtained from each station's NRFA webpage[28]. A worked example is provided below, whilst the full set of equations used are given in Table A1 (see Appendix A).

The example given is for the Sheepmount station on the River Eden in Carlisle, Cumbria (76007). The peak December 2015 level recorded at this station, of 7.806 m above the site datum (ASD), was measured on the 6th December following the intense precipitation associated with Storm Desmond.

[24] http://nrfa.ceh.ac.uk/news-and-media/news/nrfa-releases-expanded-data-download-facility

[25] Various hydrological data including AM series and peak flow rating equations from individual gauging stations can be readily obtained using the interactive map and search functionality of the NRFA website, http://nrfa.ceh.ac.uk/data/search

[28] see e.g. http://nrfa.ceh.ac.uk/data/station/peakflow/76007

A time-series plot from the GaugeMap[29] website (Figure 3.3), which is a very useful resource, corroborates the validity of the reading.

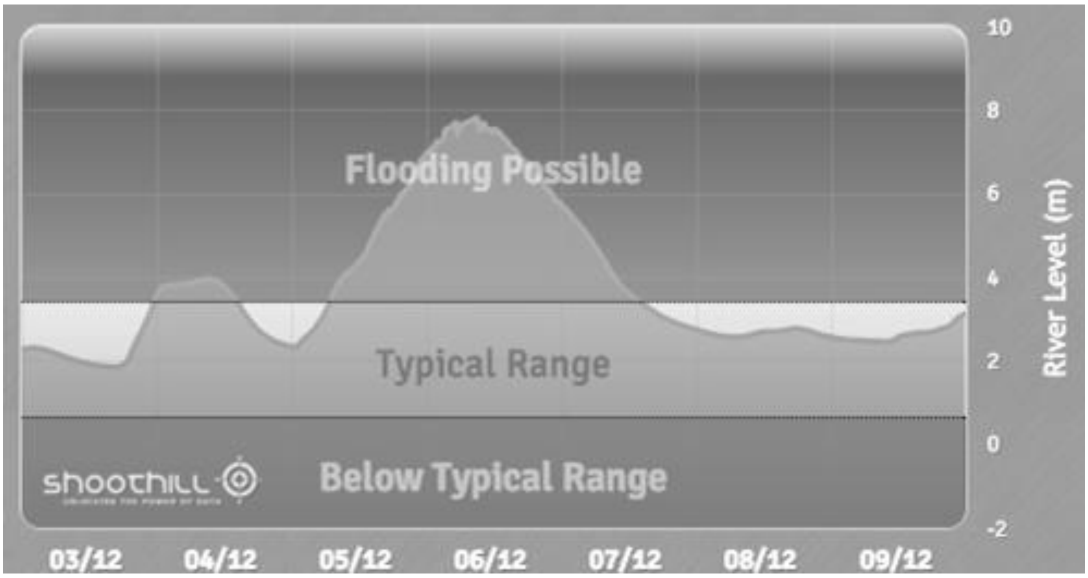


FIGURE 3.3. Time-series plot of river level for the River Eden at Sheepmount, Carlisle, Cumbria (76007) in early December 2015. The GaugeMap website continuously plots the same real-time water level data that were used in this study, and maintains an easily searchable history.

Figure 3.4 illustrates the rating curve at this site. In this case (and due to the severity of the event in many others also), it was necessary to apply the rating relationship considerably beyond its calibration range to estimate the flow level associated with the maximum stage. Wherever this was necessary, a note was made in Table A1; see Appendix A).

[29] www.gaugemap.co.uk

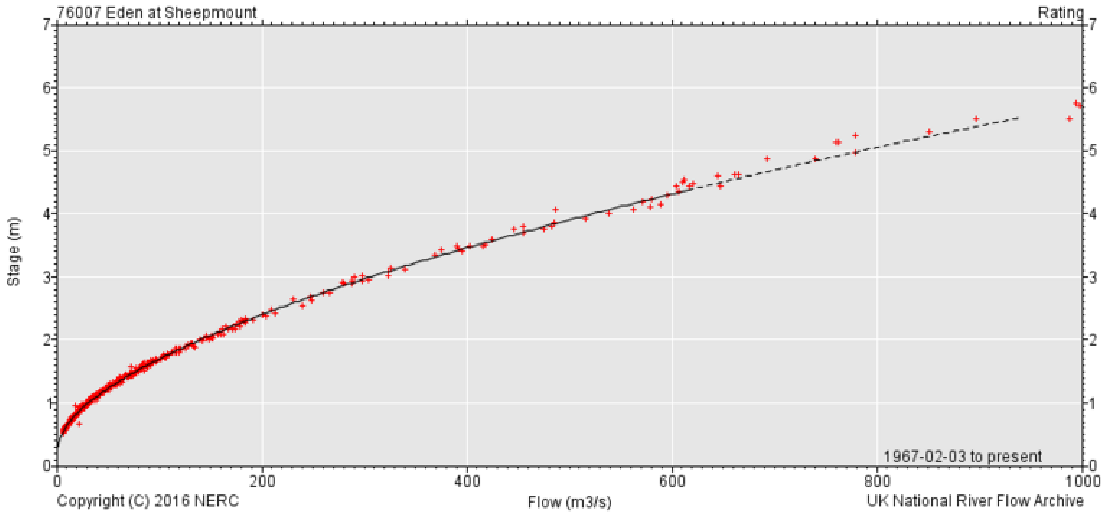


FIGURE 3.4. Rating curve relating water level with peak discharge for the River Eden at Sheepmount, Carlisle, Cumbria (76007). Calibration points are show as small red crosses. Source: CEH.

At this example site, only a single rating curve is given, which simplifies matters. Incidentally, the stage-discharge relationship here appears much better constrained than many. The information given indicates that the relationship is applicable providing the stage, *h*, exceeds 0.99 metres ASD. Discharge, *Q*, may be estimated as follows:

$$Q = 56.612 \times (h - 0.2980)^{1.699} \quad [3]$$

Given the observed peak December stage, $h = 7.806$

$Q = \underline{1{,}739.5 \text{ m}^3 \text{ s}^{-1}}$

This process was completed at all 'stage only' stations. The resultant flow estimates were then compiled, along with the few directly measured discharge maxima, into Table A1 (see Appendix A).

3.3.4. Extending the annual maxima (AM) series

Next, the existing annual maximum river flow (AM) series for the 155 stations were downloaded from the NRFA website[30]. Naturally, these series had variable start dates. Rejected records, i.e. those which for whatever reason were believed to be of questionable quality, were clearly indicated as such in the incoming data.

Presumably because of the amount time it takes to thoroughly check the quality of the observations and update the database, in January 2016 (when the research was begun), the most recent AM data available from this source were from the 2011/2012 hydrometric year. (That the latest flows were not already on the database necessitated the stage-discharge conversion work described above). Fortunately, the absence of observations between 2011/2012 and 2015/2016 does not affect the analyses from a methodological perspective, although important floods during that in that period being missed remains a possibility.

The AM series length distribution at the 155 study stations (not including any 2015 data) is illustrated by Figure 3.5. The mean record length was 39.7 years. Figure 3.6 shows record length by location, in which there appears to be little obvious spatial pattern (apart from perhaps that some of the longest records are towards the south of the regions, where populations are highest).

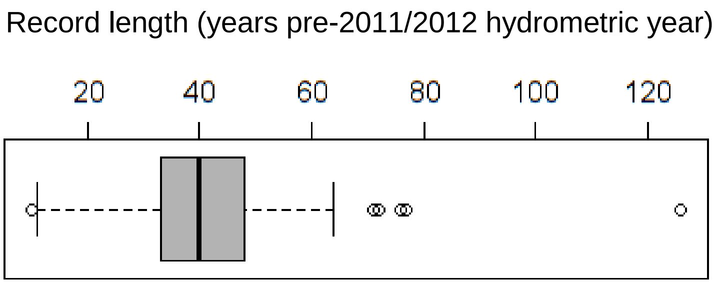


FIGURE 3.5. The distribution of AM series length (up to and including the 2011/2012 hydrometric year) at all 155 stations included in the study. The thicker line shows the median value, and the grey box the interquartile range (IQR). Whiskers encompass all points which lie beyond the IQR but within 1.5IQR of Q1 and Q3. Remaining outliers are plotted as circles.

[30] These data were obtained packaged into a single folder, called WINFAP-FEH v3.3.4 Data; see http://nrfa.ceh.ac.uk/winfap-feh-files

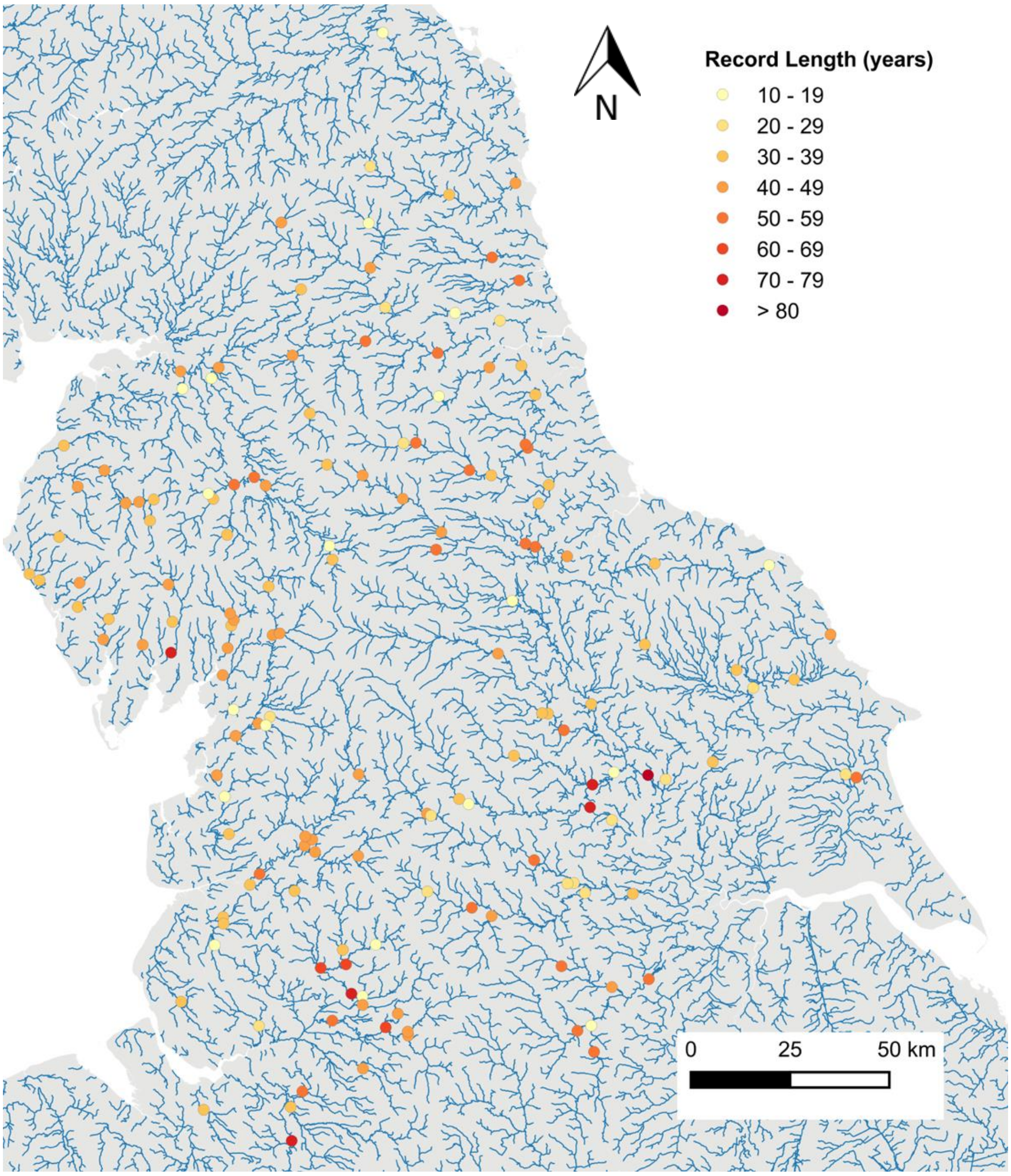


FIGURE 3.6. Spatial distribution of AM series length (up to and including the 2011/2012 hydrometric year) at all 155 stations included within the study. Watercourses are shown in blue.

To facilitate the intended analyses, it was necessary to assume that the peak December 2015 flows represent the annual maxima for the present hydrometric year (i.e. 1st October 2015–30th September 2016) at all stations. At the time of writing, this period is still ongoing. Therefore, there is a possibility that in time these data could be revealed to not be the true 2015/2016 AM flows, although the likelihood of them being widely surpassed is thought to be extremely slim. Having made this assumption, two AM series could be generated at each station – one with the additional 2015/2016 AM data, the other without.

3.4. Flood frequency analyses

3.4.1. Single-site analyses at all stations with and without the latest observations

At each of the 155 stations, statistical flood frequency analyses were conducted – first using the AM series without the additional 2015/2016 data points and then with them (Objective 1). In accordance with recommended UK practice, the Generalised Logistic (GL) distribution was fitted to the data series, with a variant of the method of L-moments method used for parameter estimation (Hosking and Wallis, 1997; IH, 1999).

L-moments are linear combinations of probability weighted moments (PWM). The theory of PMW is summarised by Greenwood et al. (1979), who defined them as:

$$\varphi_r = E\{ X [F(X)]^r\} \quad [4]$$

where φ_r is the rth-order PWM and $F(x)$ is the cumulative distribution function (cdf) of the random variable X. (φ is used here in place of the more usual β to prevent any confusion with Equation 8.

Unbiased estimators (b_i) of the PWM are given by Hosking and Wallis (1997) as:

$$b_r = n^{-1} \sum_{i=r+1}^{n} \frac{(j-1)(j-2)\dots(j-r)}{(n-1)(n-2)\dots(n-r)} x_{j:n} \quad [5]$$

where n is the sample size and $x_{j:n}$ represents an ordered sample $x_{1:n} \leq x_{2:n} \leq \dots x_{n:n}$ from the distribution of X.

The sample L-moments (λ_r) are then linear combinations of sample PWMs, calculated as:

$$\lambda_1 = b_0$$

$$\lambda_2 = 2b_1 - b_0$$

$$\lambda_3 = 6b_2 - 6b_1 + b_0$$

$$\lambda_4 = 20b_3 - 30b_2 + 12b_1 - b_0 \quad [6]$$

Sample L-moment ratios, τ_r, are based on the sample L-moments. Since λ_1 relates to the mean (or 'L-location') and λ_2 relates to dispersion ('L-scale'), λ_2/λ_1 gives a quantity analogous to the coefficient of variation, known as L-CV (τ). Higher order L-moment ratios are given by the following formula:

$$\tau_r = \lambda_r / \lambda_2, r = 3, 4\ldots \tag{7}$$

τ_3 is known as L-SKEW and τ_4 L-KURTOSIS, although only L-CV and L-SKEW were used in this instance. Like ordinary product moments, L-moments can be used to summarise probability distributions (and samples thereof).

For a GL distribution, the relationship between return period, T (expressed in years), and corresponding peak flows, Q, is given is defined using the inverse of the cumulative distribution function (or quantile function) as follows (reproduced from Miller et al., 2013):

$$Q_T = \xi + \frac{\alpha}{\kappa}\left(1 - (T - 1)^{-\kappa}\right) = \xi\left[1 + \frac{\beta}{\kappa}\left(1 - (T - 1)^{-\kappa}\right)\right] = \xi z_T \tag{8}$$

where ξ, α, $\beta = \alpha/\xi$ and κ are GL model parameters, and z_T is the growth factor is the growth factor at at return period, T defined by the term within the square brackets.

The location parameter ξ is defined as QMED. The α and β parameters, which control the growth curve, are estimated using L-CV and L-SKEW. These estimators are given in full by, for instance, Kjeldsen and Jones (2006), and so are not needlessly repeated here. In the single-site case, L-CV and L-SKEW are obtained directly from the AM series. In the pooled case (which is relevant to Section 3.4.2), they are obtained via a weighted average of the collection of L-moment ratios (according to catchment similarity, plus any additional weight that may be assigned to the target site in the case of the enhanced single-site method).

In this way, flows estimates associated with the following annual exceedance probability percentages (% AEP; with the equivalent return periods in parentheses) were obtained:

- **50%** (1-in-2-years)
- **20%** (1-in-5-years)
- **10%** (1-in-10-years)
- **4%** (1-in-25-years)
- **2%** (1-in-50-years)
- **1%** (1-in-100-years)
- **0.2%** (1-in-500-years)

The model fitting process was carried out using JFes, an application developed by JBA Consulting[31] which largely replicates other more established (but expensive to licence) flood estimation programs such as WINFAP-FEH. Unfortunately, the software did not support the calculation of confidence intervals around the estimates; ideally, the change in estimate given the inclusion of the latest data would have been assessed in light of the confidence intervals around the 'before' and 'after' models. Given more time, minimum confidence intervals could have been generated by applying Bootstrap or Jack-Knife resampling (Hall et al., 2004) in software such as R. That said, as has already been discussed, understanding and comprehensively quantifying uncertainty around flood estimates very much remains as ongoing research challenge (Kjeldsen, 2015). The following figures, which again relate to the Sheepmount station, seek to illustrate the key inputs and outputs of the process.

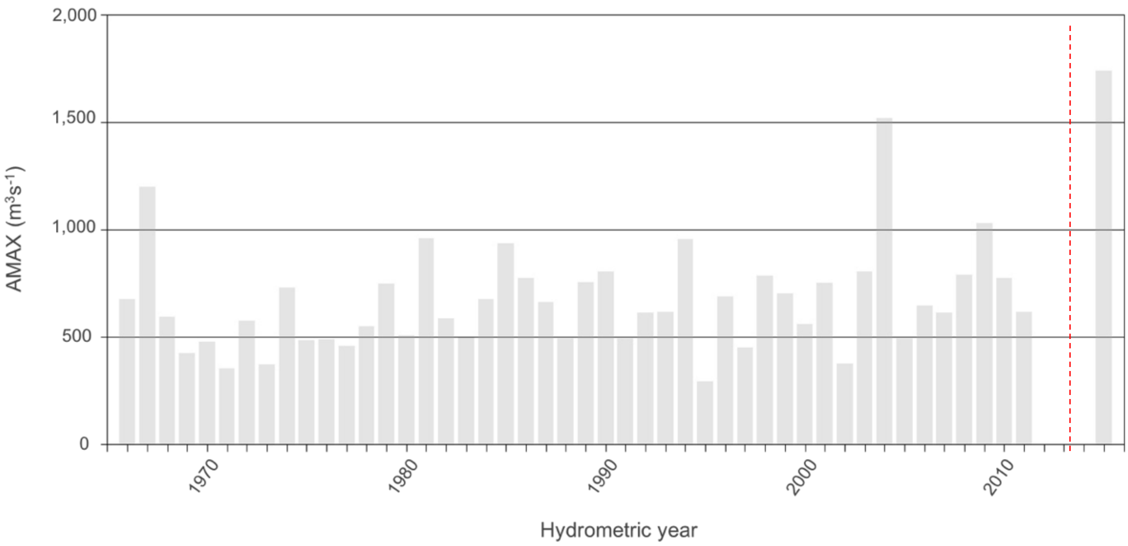


FIGURE 3.7. Annual maxima (AM) river flows for the River Eden at Sheepmount, Carlisle, Cumbria (76007). The dashed red line donates the break point in between the samples with and without the latest data.

[31] http://www.jbaconsulting.com/project/flood-estimation-software-jfes

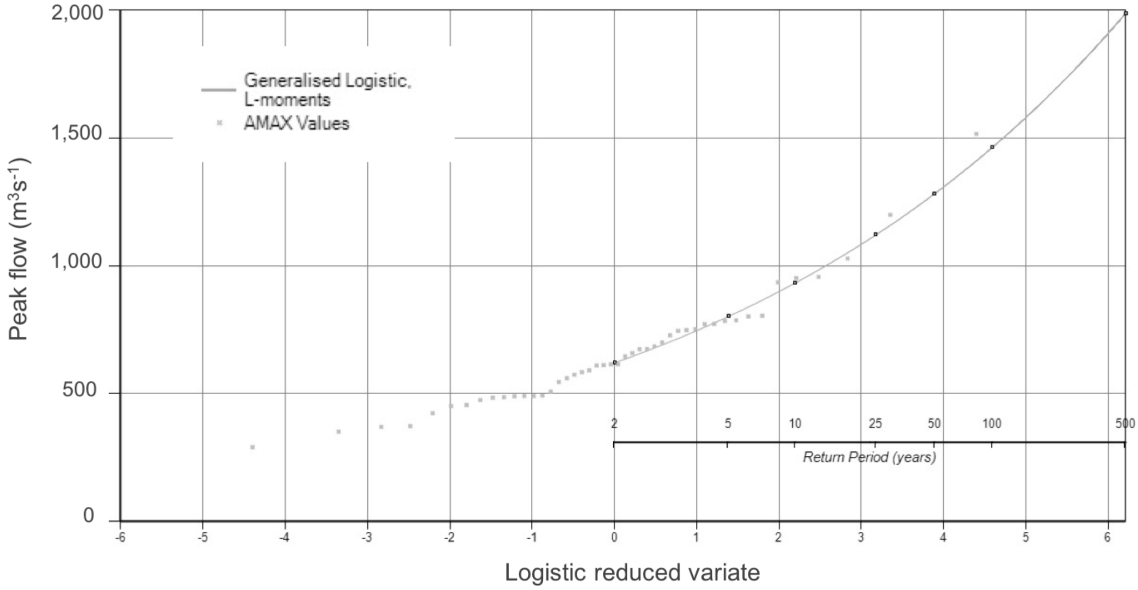


FIGURE 3.8. Fitted Generalised Logistic (GL) distribution to annual maxima (AM) river flows without the additional 2015/2016 data point (i.e. to the left of the dashed line in Figure 6.6) for the River Eden at Sheepmount, Carlisle, Cumbria (76007). The L-moments method fitting method was used. Data points are plotted using the Gringorten formula (Gringorten, 1963).

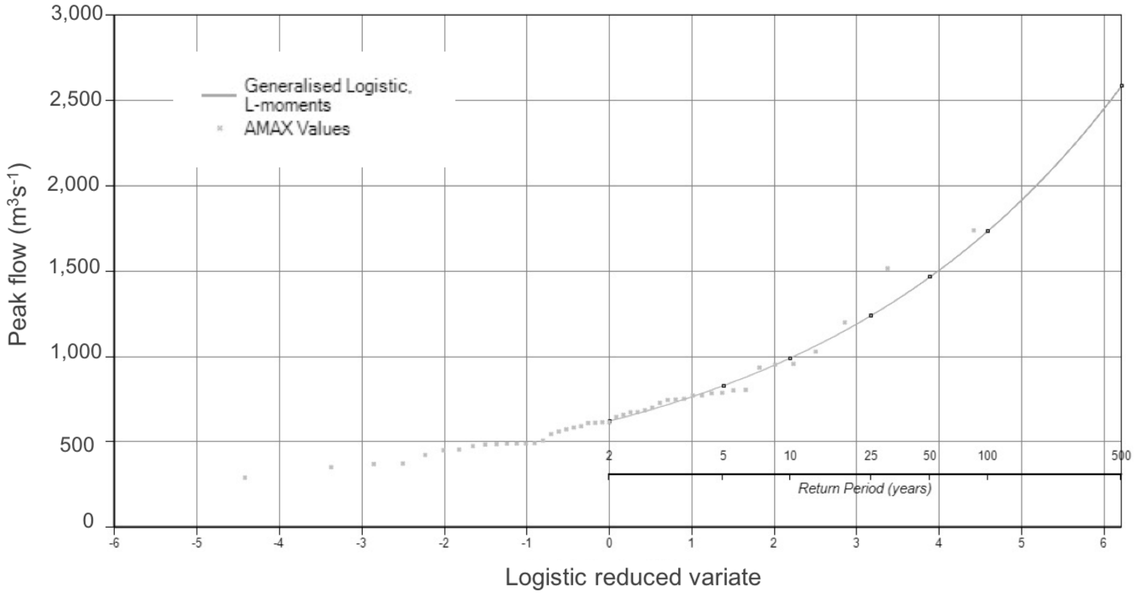


FIGURE 3.9. Fitted Generalised Logistic (GL) distribution to annual maxima (AM) river flows with the additional 2015/2016 data point (i.e. the entire data series shown in Figure 3.6) for the River Eden at Sheepmount, Carlisle, Cumbria (76007). The L-moments method fitting method was used. Data points are plotted using the Gringorten formula (Gringorten, 1963).

Having derived these return period flow estimates, the percentage change in 1-in-100-year flow, $\delta Q_{100(\%)}$, was calculated at each station. This probability level, which is extreme but still perceptible, is often of particular interest to practitioners practical applications, for example as a standard for flood defence design. Moreover, it appears have become something of a standard reference probability level when changes in flood frequency estimates due to the use of different data and methods are reported in the literature. Calculating this measure of change would therefore eventually facilitate straightforward comparison with other reported changes (e.g. those of Black and Fadipe, 2009). $\delta Q_{100(\%)}$ was given by:

$$\delta Q_{100(\%)} = \left[\left(\frac{Q_{100b}}{Q_{100a}}\right) - 1\right] \times 100 \tag{9}$$

where Q_{100a} is the 1-in-100-year flow estimate without the additional event data and Q_{100b} is 1-in-100-year flow estimate with the additional data.

Using data from the above example, this gives:

$$\delta Q_{100(\%)} = \left[\left(\frac{1736.68}{1464.16}\right) - 1\right] \times 100$$

$$= \underline{18.6\%}$$

At this point, the single-site analyses had been completed.

It was expected that estimates at stations with shorter records would be more sensitive to the addition of the new AM peaks than those produced from longer series. To investigate whether there might indeed be any relationship between the calculated changes and record length, a simple scatter plot was produced, the least-squares regression lines found and corresponding significance (p-)values determined, all using R software[32]. A similar process was followed to explore whether there might be any association between change and catchment area, although the expectation in this case was less clear. In this case, the common log (i.e. $\log_{10}$) of catchment area was taken to improve the normality of its distribution. This phase of work addressed Objective 2.

[32] https://www.r-project.org/

Whilst these single-site estimates are considered useful in their own right (and indeed formed the major part of this study), it is acknowledged that this approach is not typically recommended for flood estimation in practice unless the length of peak flow series available is long relative to the target return period (>2*T* being a 'rule of thumb'). As previously discussed, where records are not this long, the results given by single-site analysis may be highly sensitive to the random presence or absence of large floods in the sample (relative to the true distribution). In other words, estimates of the underlying annual flow probability distribution on a single-site basis may not be entirely reliable. For this reason, and remaining mindful of the corresponding assumptions and potential drawbacks associated with spatially pooled analyses and other relevant matters (e.g. choice of statistical distribution, which is addressed in due course), a reduced number of stations at which further analyses might be conducted were identified. It was hoped that these detailed analyses, which address Objectives 3 and 4, would yield some additional insight into the wider challenges associated with flood frequency estimation.

3.4.2. Enhanced single-site analyses at selected stations

This section describes the work undertaken to address Objective 3. The enhanced single-site method employed was introduced briefly in Section 2.1.2. Here it is merely reiterated that this approach is the recommended means by which estimates of longer return period flows estimated should be obtained at gauged locations, and that the approach may be thought of as an attempt to get the 'best of both'; the relevance afforded by single-site analysis and the increased sample size afforded by pooling. Although a pooled growth curve is derived using this approach, an enhanced weighting is given to the data from the target site. The 'revised method', which is described by Kjeldsen et al. (2008a,b) was followed. Advice to continue adding sites to the pooling group until the combined dataset exceeded 500 years was adhered to.

Within the scope of this project, it was not possible to conduct enhanced single-site analyses at all 155 stations. Rather, two subset groups of stations were identified. It was hoped that these stations would give representative indications of the broad range of behaviour (across all stations) observed during the single-site phase. The following steps were carried out to select two groups of stations, 'Group A' and 'Group B':

1. The difference in 1-in-100-year growth factor (i.e. the dimensionless value one must multiply *QMED* by in order to estimate the 1-in-100-year flow) produced when the additional data was added was calculated from the single-site results. (Using the growth factor rather than the change in flow itself removes any influence of changes in *QMED*, which might be more pronounced, but are still rather small, at stations with shorter records)
2. The growth factors were ordered from largest to smallest
3. 10 stations at which a reasonably large positive change in the 1-in-100-year growth factor had been observed were selected in a quasi-random fashion into what was entitled 'Group A'. To ensure some diversity, multiple stations in the same catchments were not chosen. Additionally, only those with relatively long record lengths (>~40 years) were taken, which made it less likely that the pronounced change was simply due to volatility in the single-site estimates associated with a short record.
4. A further 10 stations having a much more neutral response to the additional data in the single-site phase (i.e. the difference in growth factor was approximately zero) were then selected into 'Group B'. Perhaps not coincidentally, Group B stations generally had longer record lengths.

The selected stations are listed in Table 1, whilst their locations are shown in Figure 3.10.

Station	Ref	Group	Easting	Northing
Caton	72004	A	352860	465290
Low Briery	75009	A	328558	524216
Pooley Bridge	76015	A	347236	524959
Temple Sowerby	76005	A	360452	528312
Beetham Weir	73008	A	349620	480590
Elland	27029	A	412400	422000
Sedgwick	73005	A	350883	487419
Portinscale	75005	A	325195	523885
Armley	27028	A	428100	434000
Southwaite Bridge	75004	A	313090	528090
Sunderland Bridge	24001	B	426500	537800
Skelton	27009	B	456800	455400
Sea Cut at Scarbrough	27033	B	502800	490800
Portwood	69027	B	390700	391870
Barnard Castle	25008	B	404700	516600
Burn Hall	24005	B	425900	538700
Woodhouse Mill Regulator	27025	B	443200	385700
Ashton Weir	69007	B	377240	393560
South Park	25004	B	428400	512900
Doncaster	27021	B	457000	404000

TABLE 1. Stations that were selected into the two subset groups for further analysis.

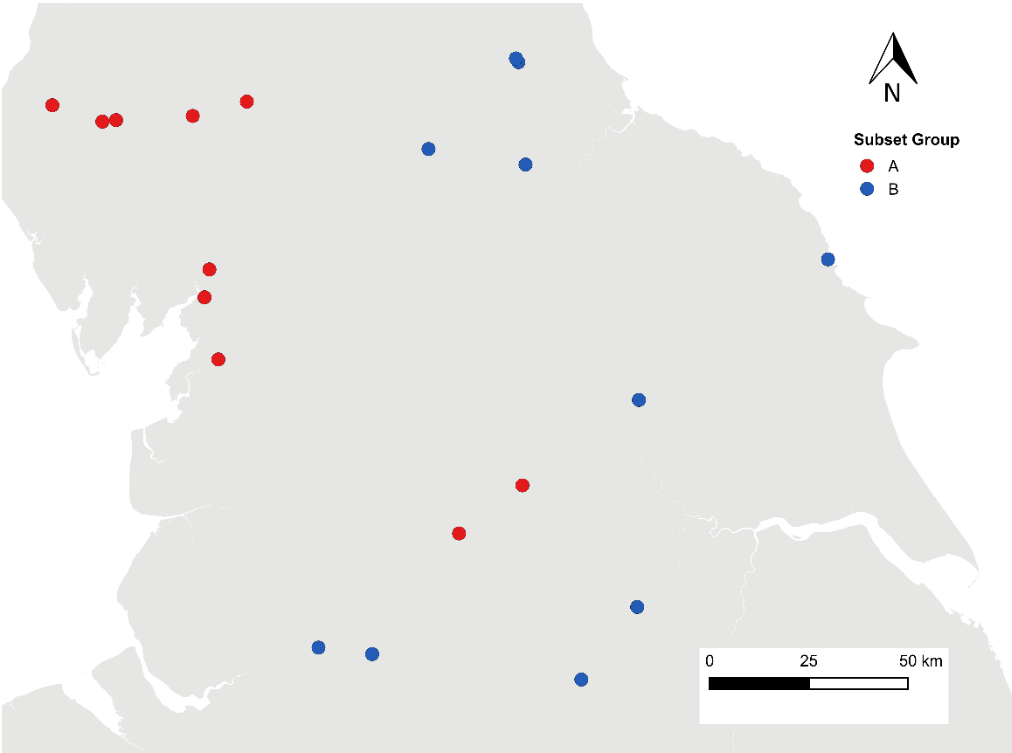


FIGURE 3.10. Locations of the stations that were selected into the two subset groups.

The enhanced single-site analyses were only conducted without the additional 2015/2016 data included – including the latest data in these analyses was unfortunately not possible in the time available. Although conducting these analyses with the latest peaks included might have been extremely interesting, an important feature of the enhanced single-site method is it should be able to overcome the short, unrepresentative record problem, and so conducting the analyses without the additional data provides a useful test in this regard.

In producing these estimates, the default weightings assigned by JFes to all stations in the pooling group, including the higher weighting given to the target site, were maintained. The results of these enhanced single-site analyses were then compared with the two corresponding single-site estimates. A little model-to-data comparison was also possible.

Finally, it must be highlighted that some catchments in the subsets were defined as urban. Although pooling of any sort is not recommended in urban catchments, flood risk estimates are still required in such catchments and so these stations were kept in for the purposes of this study.

3.4.3. Sensitivity to the choice of statistical distribution

Finally, consideration is given to one of the other sources of uncertainty in design flood estimates – the choice of statistical distribution (Objective 4). Here, the intention was to compare, albeit on a relatively visual and qualitative basis, how variability flood frequency estimates associated with the data and method used (record length, whether pooling was undertaken etc.) might compare to that associated with the choice of statistical distribution assumed to 'best' approximate the underlying flow probability distribution. This is a valid line of enquiry because, although the GL distribution used thus far comes recommended in official guidance, different distributions may be preferred at different stations, as Kjeldsen et al. (2008b) demonstrate in a far more thorough assessment than was possible in this work.

Extending this point slightly, it may be that based purely on limited and (as was discussed in Section X.X) uncertain data, it may not even be possible to formally identify a preferred model. (In this context, preferred model means a preferred distribution, but in other cases it can mean both model structure and parameter values). Competing models may seem to fit such data similarly well. In hydrological modelling, where admittedly there are often many more (and sometimes spatially distributed) parameters, this concept has become known as equifinality (Beven, 2006).

As such, exploring alternative distributions fitted to the same data is a logical course of action. The GEV distribution, which may have a stronger theoretical basis in Extreme Value Theory, is a particularly attractive alternative to the GL distribution. Its cumulative distribution function is given by:

$$F(x;\,\mu,\sigma,\xi) = \begin{cases} \exp\left\{-\left[1 + \xi(x-\mu)/\sigma\right]^{-1/\xi}\right\}, \\ 1 + (x-\mu)/\sigma > 0,\ \xi \neq 0, \\ \exp\{-\exp[-(x-\mu)/\sigma]\},\ \xi = 0. \end{cases} \quad [10]$$

where μ, σ and ξ are the location, scale and shape parameters respectively.

If $\gamma > 0$, the distribution is said to be heavy-tailed. If $\gamma < 0$, the distribution has a bounded upper tail. The special case of $\gamma = 0$ gives a Gumbel distribution, which has a thin, unbounded tail (Katz et al., 2002). The third possible choice that was considered was the Gumbel distribution.

For this final stage of analysis, the number of stations considered was further reduced, with five stations from Groups A and a further five from Group B being selected at random. The L-moments fitting method was used in conjunction with these distributions also, and once again JFes was used.

Chapter 4

Results and discussion

4.1. Single-site analyses at all stations

4.1.1. Distribution of change in Q_{100}

The return period peak flow estimates produced at all stations within the study region on a single-site basis, with and without the 2015/2016 AM data, are presented in Table B1 (see Appendix B). Figure 4.1 shows the distribution of change in the 1-in-100-year flow level (δQ_{100}) across all stations brought about by the introduction of the additional data. Figure 4.2, meanwhile, shows the spatial pattern in this measure of change.

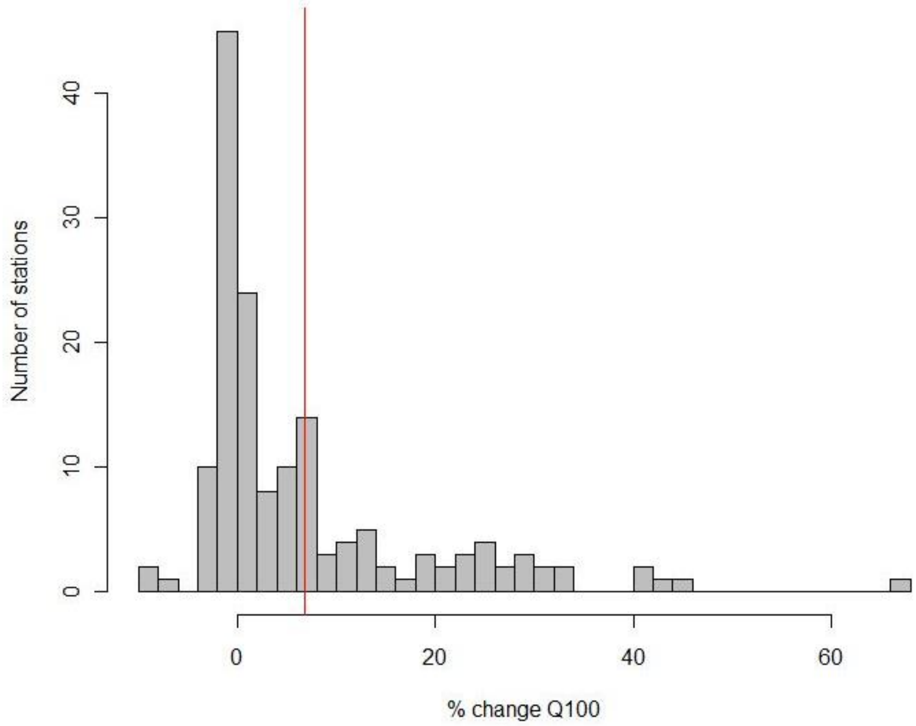


FIGURE 4.1. Distribution of percentage change in the 1-in-100-year return period flow (% change Q_{100}) across at all stations within the study region. The red line shows the mean change (+6.8%).

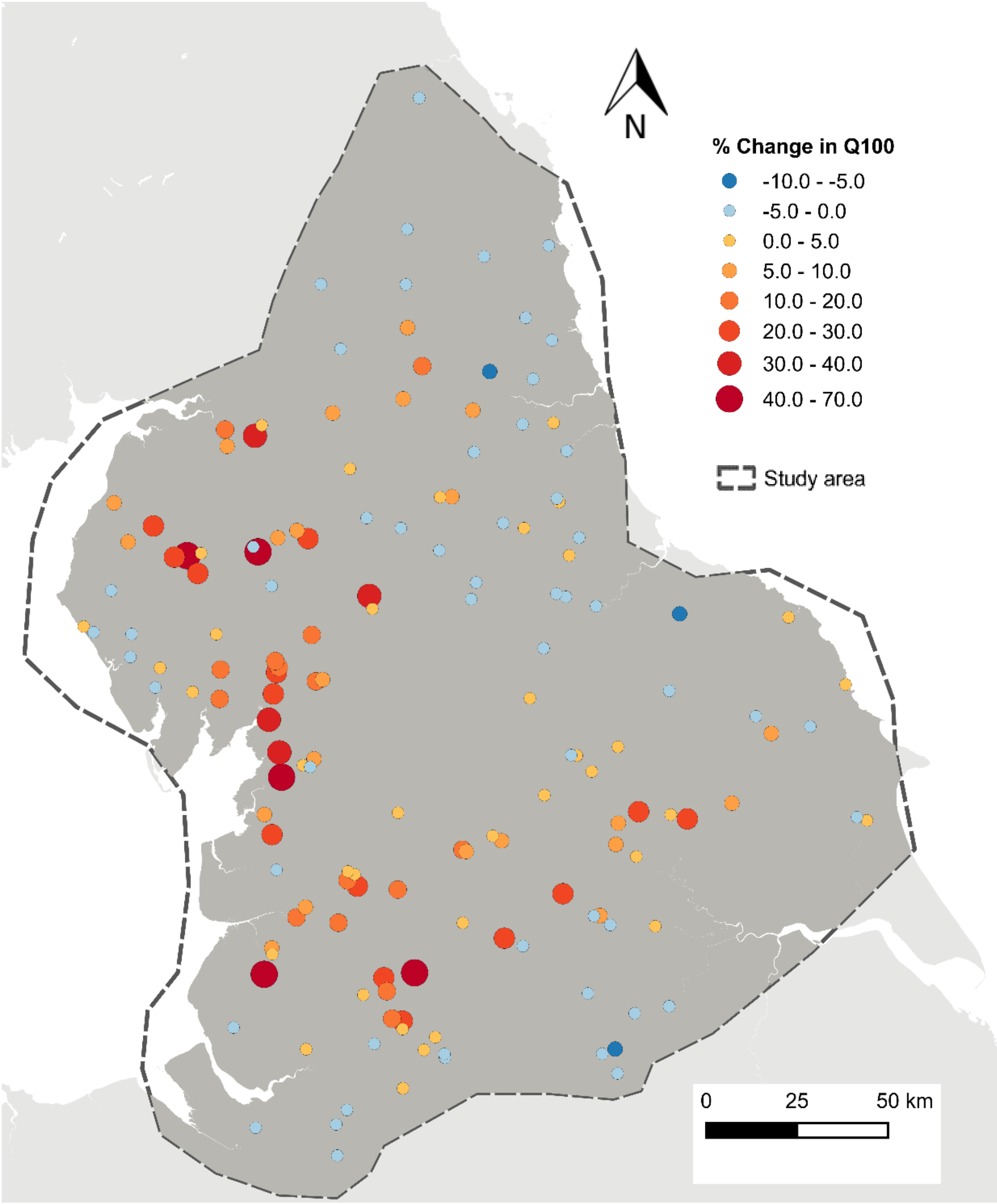


FIGURE 4.2. Spatial pattern of percentage change in the 1-in-100-year return period flow (% change Q_{100}) across at all stations within the study region.

When considering these results, it is worth firstly mentioning that a small proportion of any change in Q_{100} is likely to be associated a slight shift in *QMED* in response to the additional data. Shorter records will be more sensitive to this. However, as Figures 3.5 and 3.6 showed, record lengths are reasonable at all study stations. Although the factors are not specifically not separated here, by far the biggest contributor to δQ_{100} is change in the 1-in-100-year growth factor, z_{100}.

At many stations (58 of the 155), a slight reduction in Q_{100} was observed. Indeed, the modal class in Figure 4.1 is in the negative region. The distribution is positively skewed however, and the mean change across all stations of +6.8% represents a general moderate increase. Dispersion is fairly high, with the 5th and 95th percentiles being −2.4% and +30.6% respectively.

The absolute minimum and maximum change figures are not reported individually because to do so might give undue prominence to potential outliers. Indeed, during the course of the analyses, it was eventually realised that the peak stage observation of the station with the largest purported increase – 16.245 m at the Lune at Caton (72004) – was probably erroneous. A further detailed inspection of the time-series revealed that this gauge initially failed for a short period on the 7th December as stage exceeded 7.5 m, which is a level higher than any previously recorded. It then failed more 'catastrophically' on the 12th December, returning the anomalous data. Regrettably, this problem was not detected during the initial data checking phase. More fortunately, given the large number of study sites, the mean change quoted above (across all sites) is not unduly affected by this anomaly; discounting this station gives a mean change of +6.4%, cf. 6.8%). This situation demonstrates that although the time it takes for thorough quality checking to be undertaken before data are 'officially' published on the NRFA can be frustrating, it is essential when dealing with the difficult to measure quantity of river stage.

Two possible explanations might be proposed with respect to the stations at which little change or slight reductions in Q_{100} were observed. Firstly, these stations may simply have been 'missed' somewhat by the event, which is to say were not affected by particularly high flows. If this was the case, then the results would seem to suggest that even the addition of a single year's data without a significant flood can bring about noticeable (although small) reduction in design flood estimates under the single-method. In other words, a fairly high degree of sensitivity seems to be apparent.

A second possible explanation, however, is that some stations had already experienced an influential flood previously. Under such circumstances, although the December 2015 peaks might still be notable positive outliers in the records more generally, much of the 'potential' for a positive increase might have already been realised. Returning to Figure 2.11 (Miller et al., 2013), which shows change in the estimated flood frequency relationship at a station in Cumbria (the River Derwent at Camerton; 75002) following the 2009 floods, provides an excellent illustration of this point. Should this station have been affected by a high flow in December 2015, any change would have been far less marked than if the 2009 flood had not occurred. (Unfortunately, this particular station did not provide data to the present study, and so this example is merely hypothetical). Of course, the extent to which concept affects the present results depends very much on the spatial patterns of extremes in previous events. Spatial patterns are discussed a little further below, but it may briefly be remarked now that given the recent apparent upturn in serve (albeit often fairly localised) flooding in Cumbria in recent years (e.g. 2005, 2009, 2012, 2015), it is perhaps surprising that this region still appears as something of a 'hot spot' in Figure 4.2.

The majority of sites registered positive increases, which was a largely expected result. This general direction of change is broadly consistent with those reported in previous studies that have reassessed flood frequency based on records extended backwards in time, as opposed to forwards as here (see Section 2.3.2). Having said that, whilst being of the same order, the magnitudes of the changes are perhaps not quite so large as some of those reported by historically-based studies. For example, at one of the sites studied by Black and Fadipe (2009), inclusion of historical data raised the Q_{100} estimate by 116%. None of the changes observed here were so large. Nevertheless, this absence of such large increases seems logical because in taking an historical approach, one essentially has flexibility to go as far back in time as is necessary (or possible) until any high 'outlier' floods are found or more generally it is felt that a reasonably representative catalogue of peak flows has been attained. This study simply used the singular additional data points at each station 'as they were'.

The possible implication of this, then, (assuming long-term stationarity) is that even with the exceptional December 2015 peaks now included – it should be remembered that they will be included in future pooling groups also – the full range of possible natural variability may still not be contained within the instrumental measurement series. This suggestion aligns very well with the views and evidence presented by Foulds and Macklin (2016), who posit that the North Atlantic climate system may have the capacity to produce even wetter and more prolonged flood-rich periods than those hitherto observed in the twenty-first century.

It is clear from Figure 4.2 that some spatial pattern exists in δQ_{100}. Given the appreciation of the characteristics of the event that was developed during preliminary research, the pattern is mostly as expected; stations with reasonably large positive changes are distributed across the west and south of the study area, and more specifically across Cumbria, Lancashire and parts of Yorkshire.

Stations at which reductions were observed are confined to the extreme south, west and northeast of the area. Naturally, there is close correspondence between Figure 1.3, which shows the locations at which previous record maxima were exceeded, and the patterns of change in Figure 4.2. (In making this comparison, it is worth noting that because not all EA stations have an accessible AM series and rating information, many more stations are shown in Figure 1.3 than could be included in this study).

The spatial heterogeneity in change could simply be the product of the stochastic characteristics of the events in question. The specific timings of tributary inputs, patterns of floodplain storage and such factors vary in space within individual events (and furthermore between events), affecting flood peaks (e.g. Porter, 2011). Likewise, channel engineering, defence construction and suchlike have the potential to affect these natural patterns, in the latter case perhaps increasing downstream peaks (although there seems to be few published studies that have sought to quantify this effect). In summary, it is unsurprising that in real events, flows severities (expressed either as return levels or return periods) can vary quite considerably from place to place. (For this reason, so one should not really describe any event in terms of a single return period, unless the value corresponds to an average across affected sites or a particular region. Furthermore, the quantity (e.g. monthly rainfall, instantaneous peak discharge, etc.) should also be explicitly stated).

Having said that, the patterns of change seem to make good hydrological sense; there is general consistency along watercourses (and within catchments), certainly in terms of the direction of change. There are a few locations where reductions were produced in close geographical proximity to positive increases, but these tend to be on tributaries of larger rivers. Perhaps the most plausible explanation for this observation is that these smaller catchments were spared the most extreme event rainfall, although once again possible influence of previous floods on the starting estimates (i.e. Q_{100a}) remains possible.

4.1.2. Scatter plots

To try and investigate whether there might be any underlying controls on the observed changes besides simply the characteristics of the events in question, Figures 4.3 and 4.4, which show scatter plots of percentage change in δQ_{100} vs. record length and catchment area respectively, may be considered.

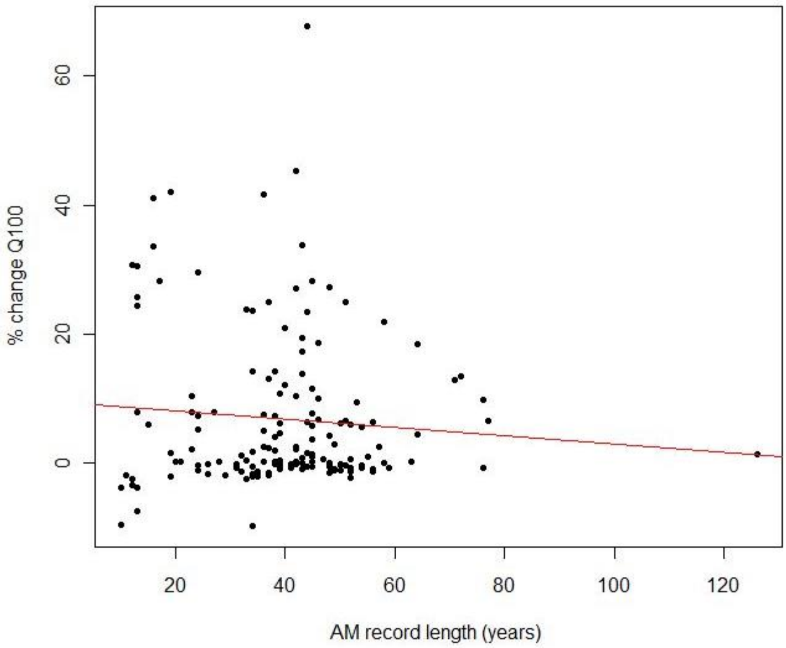


FIGURE 4.3. Scatter plot of percentage change 1-in-100-year flow against the length of the annual maxima (AM) series (years) used in the model fitting (up to and including the 2011/2012 hydrometric year) at all stations. The red line shows the least squares regression line. $R^2 = 0.0068$. ($p = 0.308$).

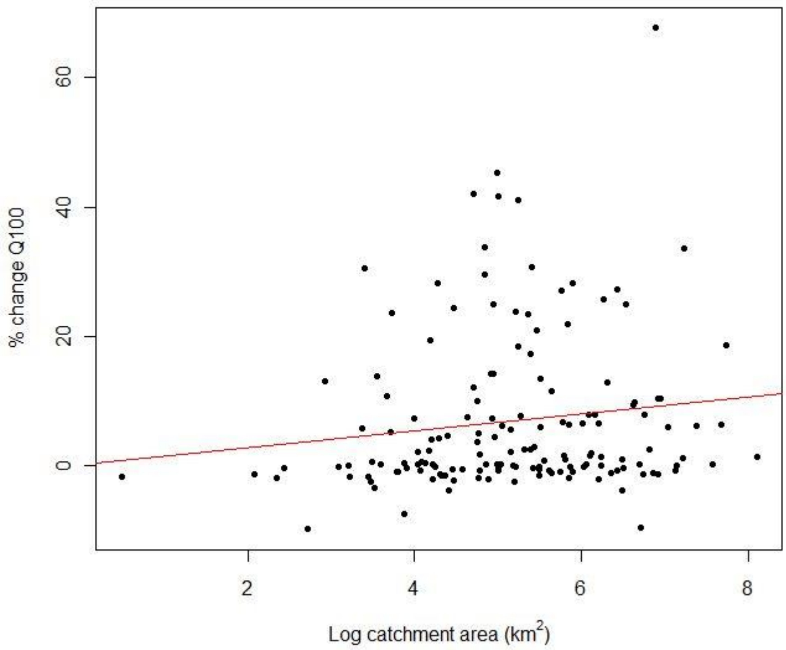


FIGURE 4.4. Scatter plot of percentage change 1-in-100-year flow against the $\log_{10}$ catchment area (km^2) at all stations. $R^2 = 0.0185$ ($p = 0.091$).

The most obvious relationship present is that in Figure 4.3, which shows that (to some extent) as record length increases, δQ_{100} decreases. In other words, there is a weakly negative relationship.

This direction of trend was as expected at the outset, and is likely to reflect the fact that when records are long, each new AM data point is effectively averaged out across all the others and so has lesser influence than in shorter records. It may also reflect the possibility that larger floods are more likely to have been recorded at stations with longer records in the first place, meaning the starting estimates, Q_{100a} may have likely to be higher (if there should indeed be some sort of 'short record underestimation problem'). However, the high p-value indicates that there is no statistical significance to this relationship at any reasonable confidence level. The correlation could have been somewhat disrupted by the previously mentioned possibilities of some stations simply being missed by the events, or the degree of change being limited by previous major floods.

Interestingly, there are some stations with quite long records at which reasonably large positive shifts have been observed, countering the general trend. More specifically, at five stations with what might be considered to be relatively long records of >50 years, an increase of more than 10% in Q_{100} was produced. These results confirm that return period estimations corresponding to double the record length can be highly sensitive using the single-site approach.

Figure 4.4 suggests that there may be some weak positive association between catchment area and degree of change, although the p-value once again does not support any statement of statistical significance. This general trend may be a reflection of the widespread nature of the intense rainfall in November and December 2015. Perhaps in this event more so than predecessors, very high flows were might have been recorded at the outlets of the some of the larger catchments thanks to large proportions of the total catchment areas making 'strong' contributions . In contrast, perhaps in previous events only some sub-catchments were affected by the most intense rainfall, and so the overall hydrograph response at large catchment outlets was dampened by the lesser contributions from elsewhere. Of course, the probability of rainfall being extreme (howsoever defined) over a wide area is much lower than an extreme of the same level occurring only in one locality, and so this suggestion is loosely supported by statements regarding the exceptional nature of the December 2015 event.

Overall, that the relationships in Figures 4.3 and 4.4 are both rather weak suggests that these hypothesised explanatory variables are not significant controls on the sensitivity of flood estimates (although they demonstrably are on the estimates themselves), at least within the explanatory variable ranges considered in this study.

4.1.3. Summary

The fairly high sensitivity that single-site estimates displayed to only small changes in the sample (in terms of number of observations, if not magnitude) confirms the findings of previous studies. However, given the drawbacks associated with other flood frequency estimation methods, the possibly of single-site analyses providing the preferred estimate on some occasions should never be entirely discounted. In particular, the individual characteristics of the site may only be captured well using the at-site data itself. Therefore, assuming these new single-site estimates might be preferable to the alternatives at some of the study sites at least, the magnitude of the increases in 'baseline' flood estimates (i.e. the degree of previous underestimation) may be large enough to impinge the appropriateness of design standards even once any climate change allowances (see Section 2.2.2) have been made. However, these differences are on the whole lower than those reported by historical reassessments and also those produced by the major methodological overhaul and data update described by Kjeldsen (2008b).

The general observation that many stations saw reasonably large positive increases whilst others saw a relatively modest reductions seem to suggest that flood frequency estimates may demonstrate a sort hysteresis whereby they can increase very 'quickly' , for example with the addition of only one data point, but might then 'fall' somewhat more slowly as additional years without major flood are added to the series. Such an effect was also described by Archer et al. (2007). One must presume that these estimates would eventually converge to the underlying probability distribution. Should most estimates eventually settle higher than their initial/present starting values, this would substantiate the underestimation hypothesis. The question that follows is how long might it take for stability to be achieved; without the luxury of a long time to observe (and the true distribution being unknown), this may never become clear. It is certainly likely that the timescales will be too long to be practical.

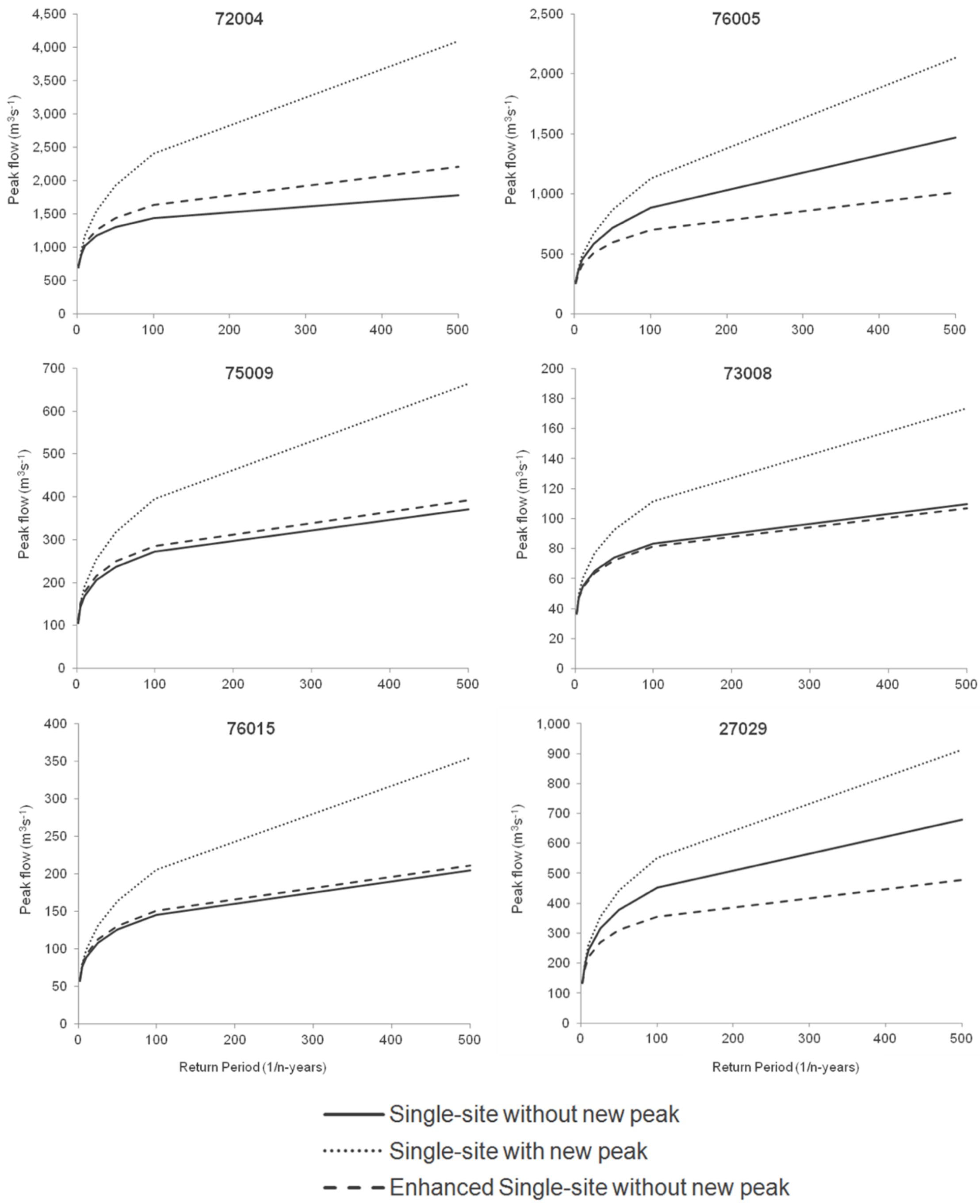
Finally, these updated results have been made under the assumption of stationarity. Although long instrumental peak flow datasets provide little indication of any increasing trend in flood frequency or severity (Hannaford, 2015), and so the assumption made is defensible, recent attribution studies have found a detectible anthropogenically-driven increase in hazard associated with anthropogenic climate change with respect to certain events (e.g. Schaller et al., 2016). Likewise, climate model-based simulations fairly consistently project future increases in flood frequency. Hence, if climate-induced trends are indeed beginning to emerge, then even the updated estimates presented herein may in fact be underestimates. This is because under stationarity, no provision is made for the possibility that the expected frequency of a flood of a given magnitude 'next year' is any different (higher) than that estimated from the available record as a whole. Furthermore, the impact of urbanisation on changing flood frequency (associated with reduced catchment permeability, increased water transit times, etc.) is now empirically detectable sometimes (Prosdocimi et al., 2015). Hence, future urbanisation has the potential to make a significant contribution to increasing flood hazard, despite the existence of legislation attempting to limit this..

In the next section, the results of the enhanced single-sites that were conducted at selected stations are considered.

4.2. Enhanced single-site analyses at selected stations

4.2.1. Group A and Group B enhanced single-site vs. single-site estimates

Figure 7.5 shows the results of the enhanced single-site analyses conduced at the 10 Group A stations, whilst Figure 7.6 shows the results for the 10 Group B stations. In both cases, the previously obtained single-site analysis results are also plotted for comparison. These results are not presented in tabulated form in this thesis.


72004
76005
75009
73008
76015
27029
Peak flow (m^{3}s^{-1})
Return Period (1/n-years)
Single-site without new peak
Single-site with new peak
Enhanced Single-site without new peak

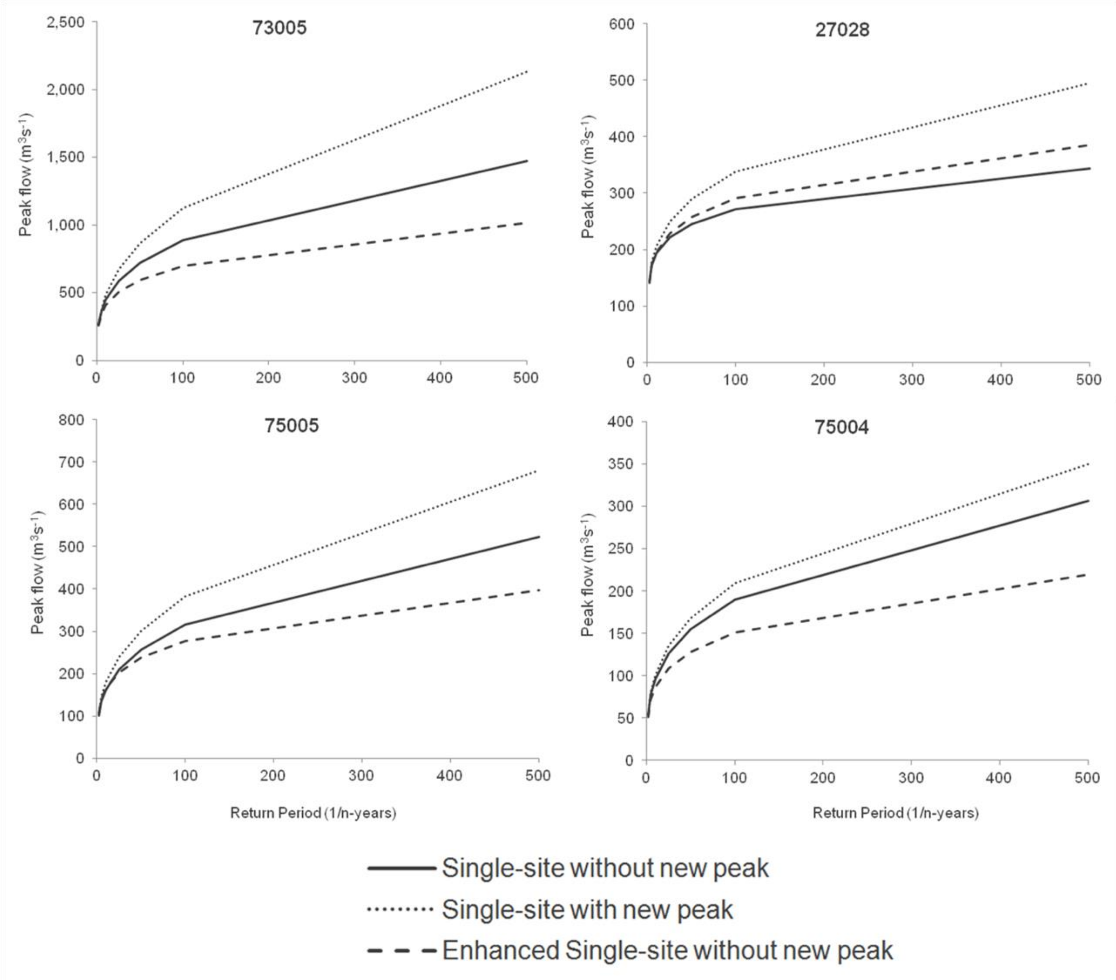


FIGURE 4.5. Comparisons of flow frequency-magnitude relationships (two single site analyses and one enhanced single-site analysis in each case) at Group A stations (i.e. a subset of those at which a significant change in growth curve was observed with the addition of new 'peak' data from December 2015 in the single-site case). The NRFA station reference number is labelled.

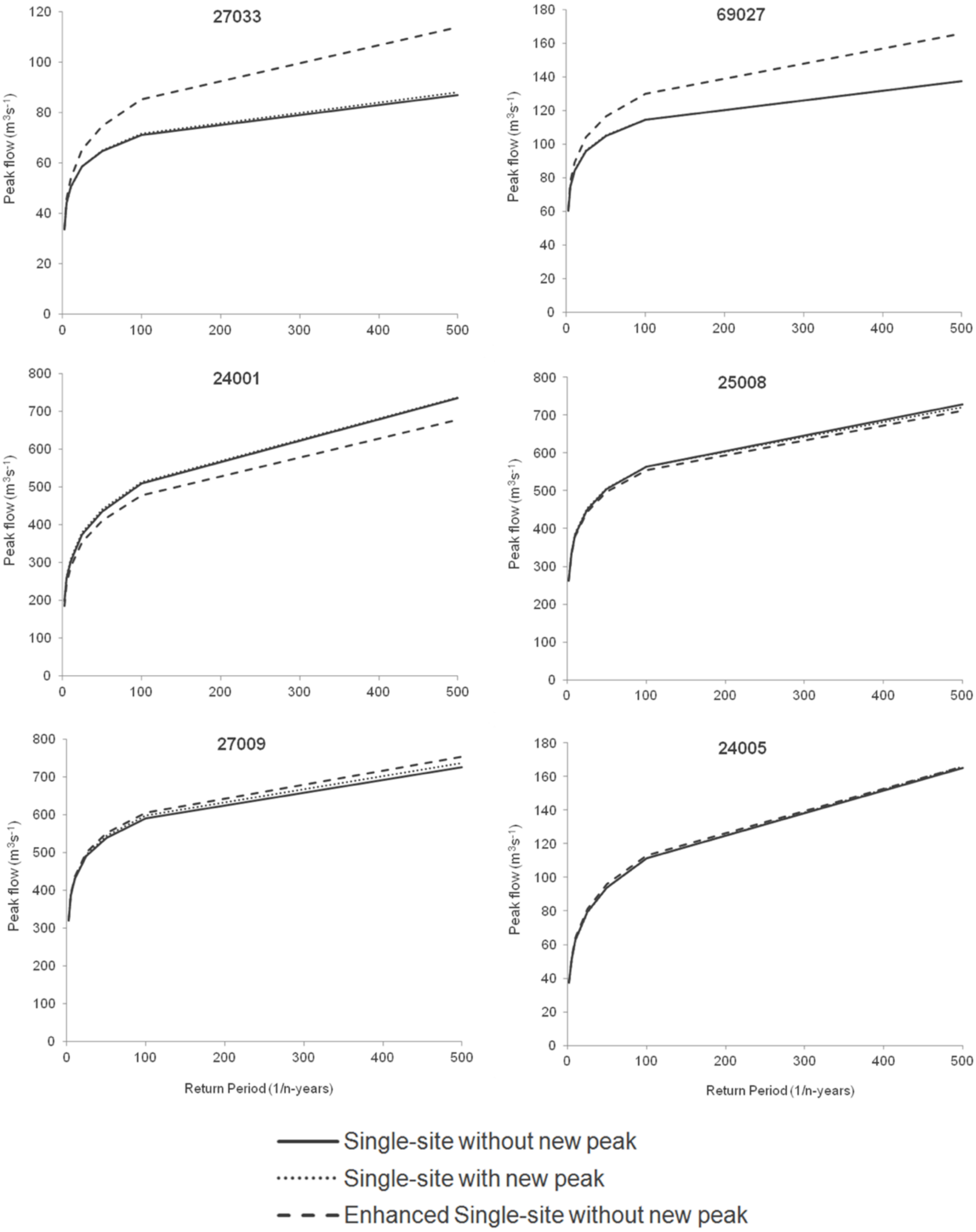
The uppermost-left plot in Figure 4.5 (the Group A stations), Station 72004, provides a 'classic' example of a result that might be expected if conducing analyses on short, unrepresentative records using the single-site basis is associated with something of an underestimation problem. In this case, even with the higher weight assigned to the data from the target station, pooling raises the single-site flood frequency estimate to a level more similar to that produced via single-site analysis with the new flood peak included. Such a result could be explained in terms of the pooling method ensuring that a more representative empirical distribution for estimation, which may often happen to be higher. Indeed, it was expected at the outset that in many cases the enhanced single site curves might be more alike those produced using the single-site method but including the latest peaks (assuming for the moment that these estimates are reasonable and have not been forced 'too high', which is a major uncertainty in itself). This expectation – that enhanced single-site curves without the latest data should be higher than single-site curves also without the latest data – is also conditional on the pooling group catchments being sufficiently independent from the target site. If they are not, and all records in the pooling group under represent large floods somewhat, then the result of pooling may simply be to extended 'moderate' records (perhaps similarly to those that would be obtained from resampling from the observed series, but with a potential reduction in relevance); it is well established that inter-site dependence within a pooling group can limit the effective number of sites, hence the amount of meaningful data in the pooed set, and ultimately the accuracy of estimations (Collier, 2011; Hosking and Wallis, 1998).

However, in all other Group A sites, the enhanced curve is either lower or else barely noticeably higher than the original single-site estimate. This limited difference may be due to the site record is long. Under these circumstances, both estimates should describe the underlying distribution reasonably well, and hence match one another closely. Another hypothetical explanation for a close match is some catchment characteristic makes the single-site distribution more 'extrapolatable' in some way. In this case, where lower magnitude observations are informative of high magnitude observations, again both estimates might approximate well the underlying distribution (assuming the pooling group is acceptable), and thus could be interpreted as confirming one another. At all locations within the group, the single-site curve with the new flood peak is the highest, although this is not particularly surprising given the basis upon which the Group A stations were identified.

The main conclusion that can be drawn from these results is that enhanced single site-analyses do not consistently raise estimated hazard relative to single-site estimates produced using the same dataset at stations now known to be capable of being affected by large, influential floods. On the contrary, in six of the 10 cases it is lower. In these cases, the pooled results seem to be in contradiction to any notion of an underestimation hypothesis, suggesting that the single-site records in these locations even without the December 2015 flood were disproportionally flood rich. However, the magnitude of the shifts generated when the latest data in included do call the lower pooled results into question somewhat.

It may therefore be tentatively proposed that perhaps the pooling groups are not independent enough (in terms of their flood records) to be able to augment the single-site data with a greater proportion of flood peaks in a given time interval. This may be a function of the large spatial 'footprints' of the frontal meteorological systems that typically cause the highest flows (as opposed to more locally intensive, convective storms), i.e. many stations may be affected to similar relative levels (in terms of catchment area etc.), and so pooling may be too similar to resampling. To address this, a minimum geographical separation distance could be imposed on stations in pooling groups, although this would likely be at the expense of the similarity of their catchment descriptors relating to soil types, for instance. It should also be mentioned that Hosking and Wallis (1988) investigated this phenomena, finding any bias to be unchanged by the presence of inter-site dependence. They report that whilst accuracy is reduced in such cases, concerns around homogeneity are still often the dominant source of uncertainty in pooled flood estimates.

When conducting such 'model-to-model' comparisons, one must also remain mindful that there is a danger that neither may actually be representing the true distribution well. This difficulty comes with the territory when the models cannot be (in)validated with reasonable confidence (as far as a model ever can be); in flood estimation, the true distribution is always somewhat unknown (and without a very long time to observe it) unknowable. Plotting the at-site observations (the 'realisations' from the distribution) against the fitted model is the best that can be achieved, see Section 4.2.2.


27033
69027
24001
25008
27009
24005
Peak flow (m3s-1)
Return Period (1/n-years)
Single-site without new peak
Single-site with new peak
Enhanced Single-site without new peak

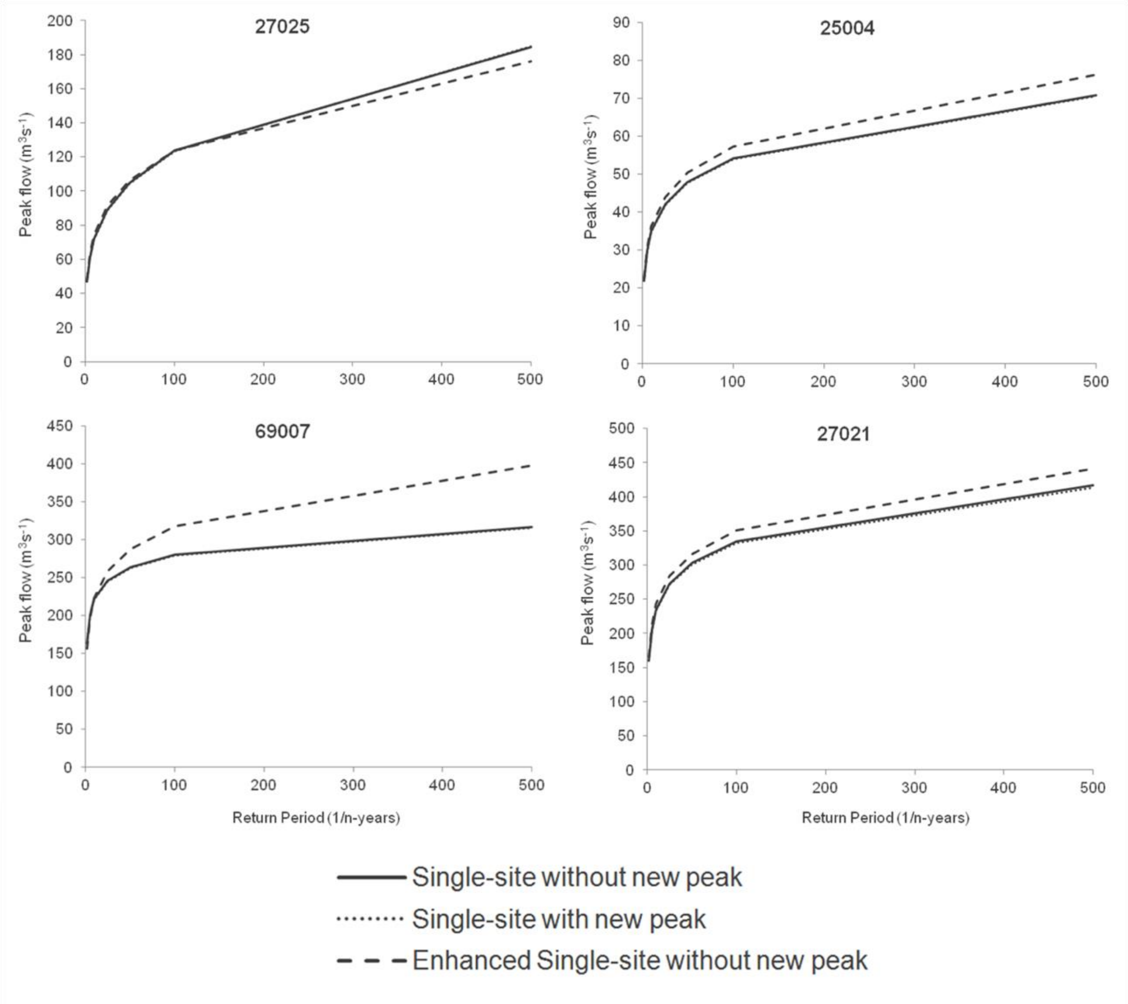


FIGURE 4.6. Comparisons of flow frequency-magnitude relationships (two single site analyses and one enhanced single-site analysis in each case) at Group B stations (i.e. a subset of those at little change in growth curve was observed with the addition of new 'peak' data from December 2015 in the single-site model fitting). The NRFA station reference number is labelled.

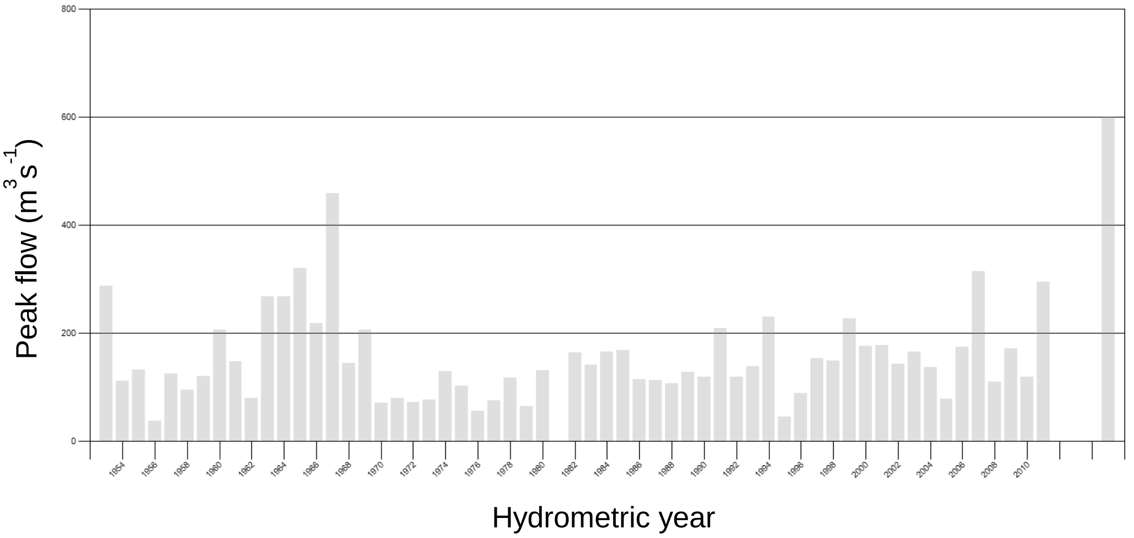
The most conspicuous feature of the Group B station curves, where flood frequency magnitude relationships were hardly affected by the inclusion of the 2015/2016 peak flows (or more precisely, the 1-in-100-year growth factor was hardly affected), is the degree of similarity between each of the three curves. The enhanced single-site results are slightly higher than the other estimates in some cases. However, in light of the findings at the Group A stations, one must wonder whether these estimates would appear high enough when presented with an extreme flow at those locations. Of course, there may be some physical characteristics of these catchments that limits their flood potential somewhat (e.g. karstic geology), and so all three estimates (since they generally confirm one another) may be entirely reasonable. Exploring such hypotheses, which are in many ways closely related to the concept of the maxim physically plausible flood, fell beyond the scope of this investigation, but could form attractive avenues for future research. The extent to which these results might be generalised is discussed a little further in general terms in due course.

Finally for the present section, it may be noted that another possible and rather obvious explanation for general the degree of similarity between the enhanced single-site curve and the single-site curve fitted without the 2015/2016 peaks in both Groups A and B exists. It is simply that the enhanced weighting that JFes automatically applied to the target site data under the enhanced single-site method might have been high. Although it would be at the expense of data relevance, if prior knowledge suggests that some larger difference between the curves may be more appropriate (e.g. the pooled result 'should be higher', as proposed by the underestimation hypothesis), then there may be some benefit in lowering this weighting, or at least exploring the sensitivity of the difference to it.

4.2.2. Flood frequency model-data comparison (in brief)

As has hopefully already been made clear, testing models against the data they are supposed to represent as far as possible is a crucial activity. In this case, given some questions around the utility of the pooled estimates, which it must be remembered come recommended by official guidance, how well the enhanced method seems to fit the data at certain locations was considered. In contrast to the single-site case, when a poor fit may simply be put down to the limited record, should then enhanced single-site curve fit appear poor, there may be an argument for rejecting it in favour of the single-site curve (which will naturally fit the at site data better, but perhaps 'too well'). In this endeavour, it is important to remember that data plotting positions using the Gringorten formula, for example, can be themselves highly uncertain; an observation which Miller et al. (2013) makes, but which is often overlooked.

In the present study, it was possible to find cases where the enhanced single-site estimates (conducted without the additional peaks, as all enhanced estimates were) do not seem to fit the data well, perhaps due to data of limited relevance being introduced. To aid this discussion, results from two example stations are plotted below (Figure 4.7 and 4.8). Although both these stations (which incidentally have reasonably long records were) affected by high flows in December 2015, it is clear that the model fit is poor more generally, i.e. it is not only the latest peaks which do not fit well.



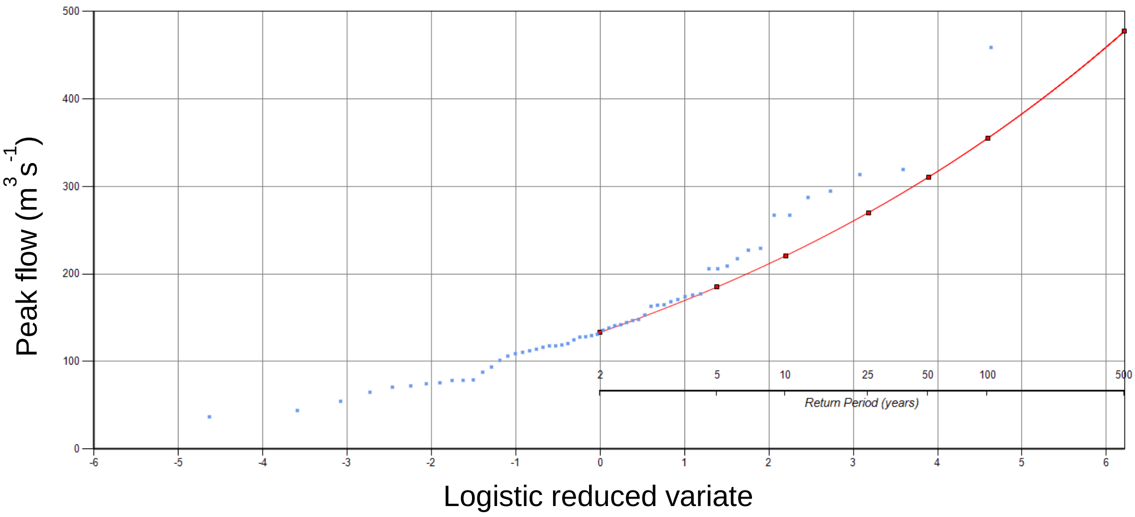


FIGURE 4.7. AM series (top) and enhanced single-site model without additional data (bottom) for the River Calder at Elland, West Yorkshire (27029). The red curve shows the fitted model and the blue points are the AM observations (plotted using the Gringorten formula).

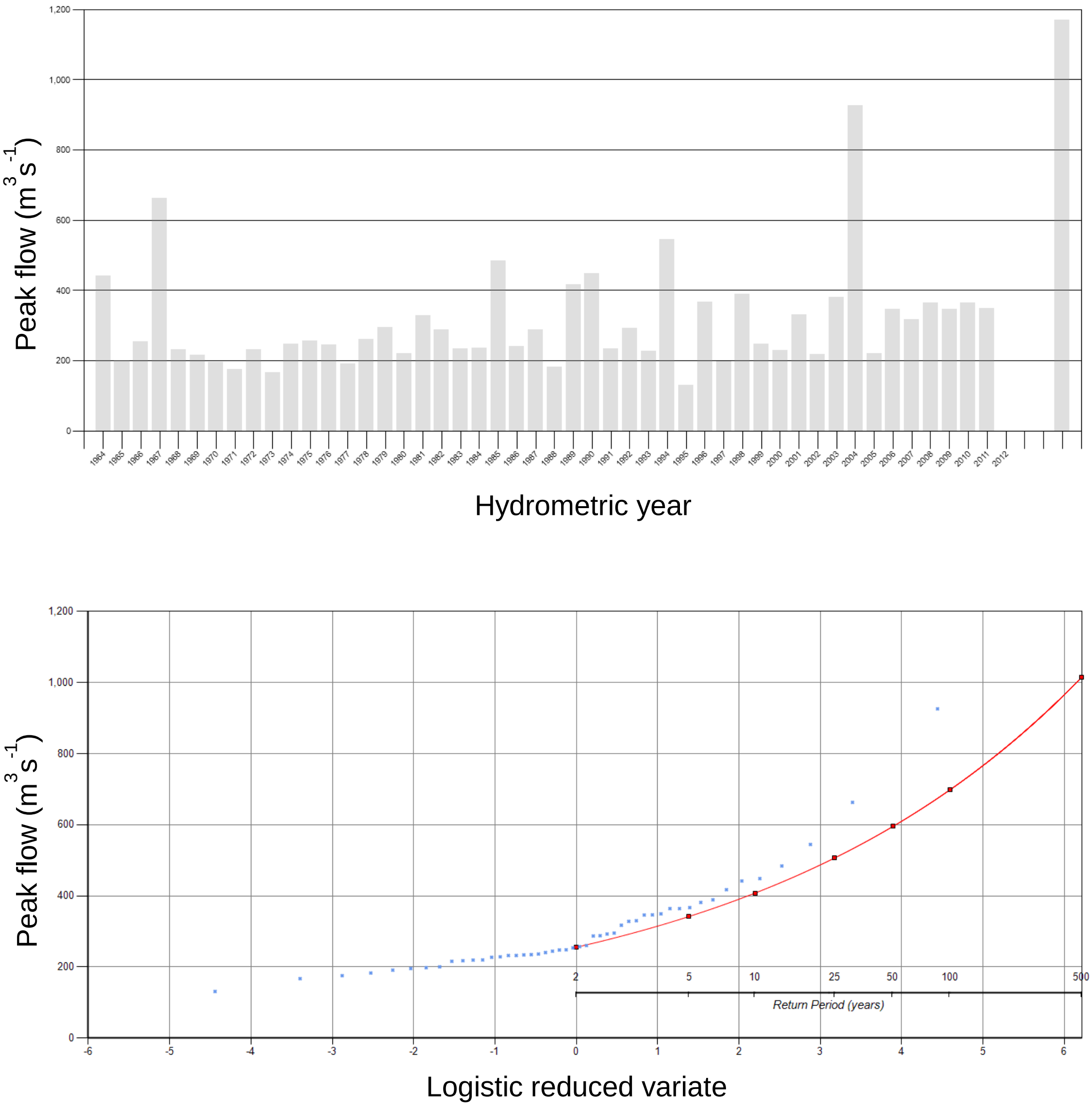


FIGURE 4.8. AM series (top) and enhanced single-site model without additional data (bottom) for the River Eden at Temple Sowerby, Cumbria (76005). The red curve shows the fitted model, and the blue points are the AM observations (plotted using the Gringorten formula).

In these examples, the ‘direction’ of apparent mismatch is potentially worrying; the model, shown by the red line, sits underneath the data. Indeed, since enhanced single site analyses were including the latest peaks was not possible, these peaks are not actually plotted in the lower pane in either Figure 4.7 or 4.8. It is apparent that if they were, the mismatch would become even greater (and the axes would need adjusting).

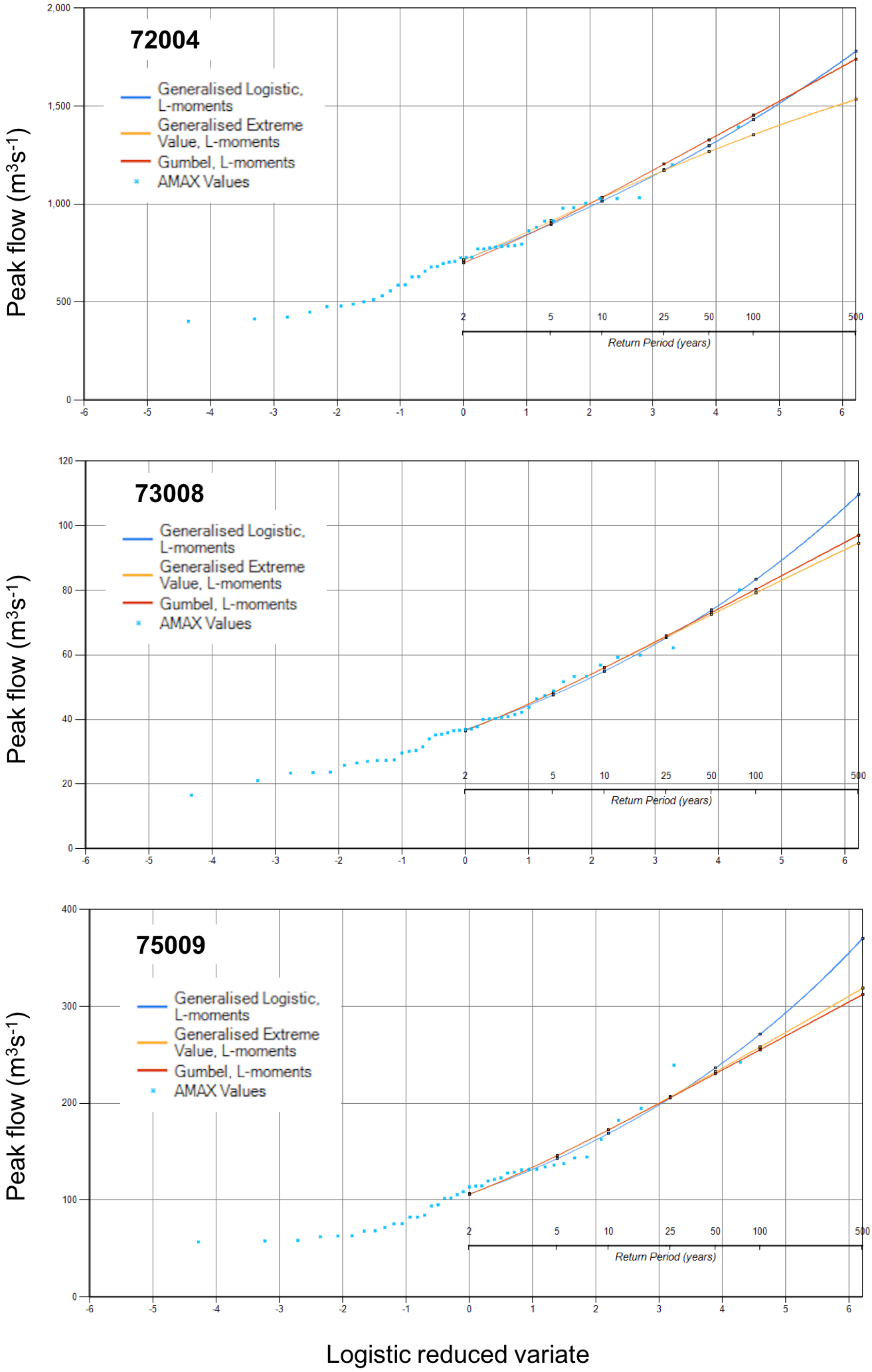
Contrastingly, in cases where the pooled model-data comparison is poor but the pooled curve sits above the data, then this might be considered reassuring for reasons already deliberated at length. These alternative responses to perhaps similar mismatches in magnitude demonstrate that much scope (or even need) exists to introduce prior knowledge and judgement, perhaps via Bayesian methods, to problems where uncertainty is deep-rooted and pervasive such as flood estimation.

4.2.3. Summary

Overall, it appears that enhanced single-site results may not be capable of consistently raising flood frequency-magnitude curves to levels that appear to be more consistent with the latest observed flows. In this sense, the results presented in this section may be broadly consistent with the underestimation hypothesis. However, the difficulty of testing these types of predictions and reliance on model-model comparisons (since it is often expected that a better model may in fact appear to fit the data poorly) severely restrict the confidence strongly with which any firm conclusions can be drawn. More generally, the previous discussion has highlighted that pooled methodology may be associated with certain drawbacks, especially where the strong influence of local site conditions or lack of independence in the group records challenge the validity of the underlying assumptions.

4.3. Sensitivity to the choice of statistical distribution

Lastly, sensitivity of single-site estimations to the choice of statistical distribution is briefly considered. Results, which are not tabulated in this thesis, are only shown for five stations in each of Group A and Group B (Figures 4.9 and 4.10 respectively).


72004
Generalised Logistic, L-moments
Generalised Extreme Value, L-moments
Gumbel, L-moments
AMAX Values
Return Period (years)
Peak flow (m3s-1)
73008
Generalised Logistic, L-moments
Generalised Extreme Value, L-moments
Gumbel, L-moments
AMAX Values
Return Period (years)
Peak flow (m3s-1)
75009
Generalised Logistic, L-moments
Generalised Extreme Value, L-moments
Gumbel, L-moments
AMAX Values
Return Period (years)
Peak flow (m3s-1)
Logistic reduced variate

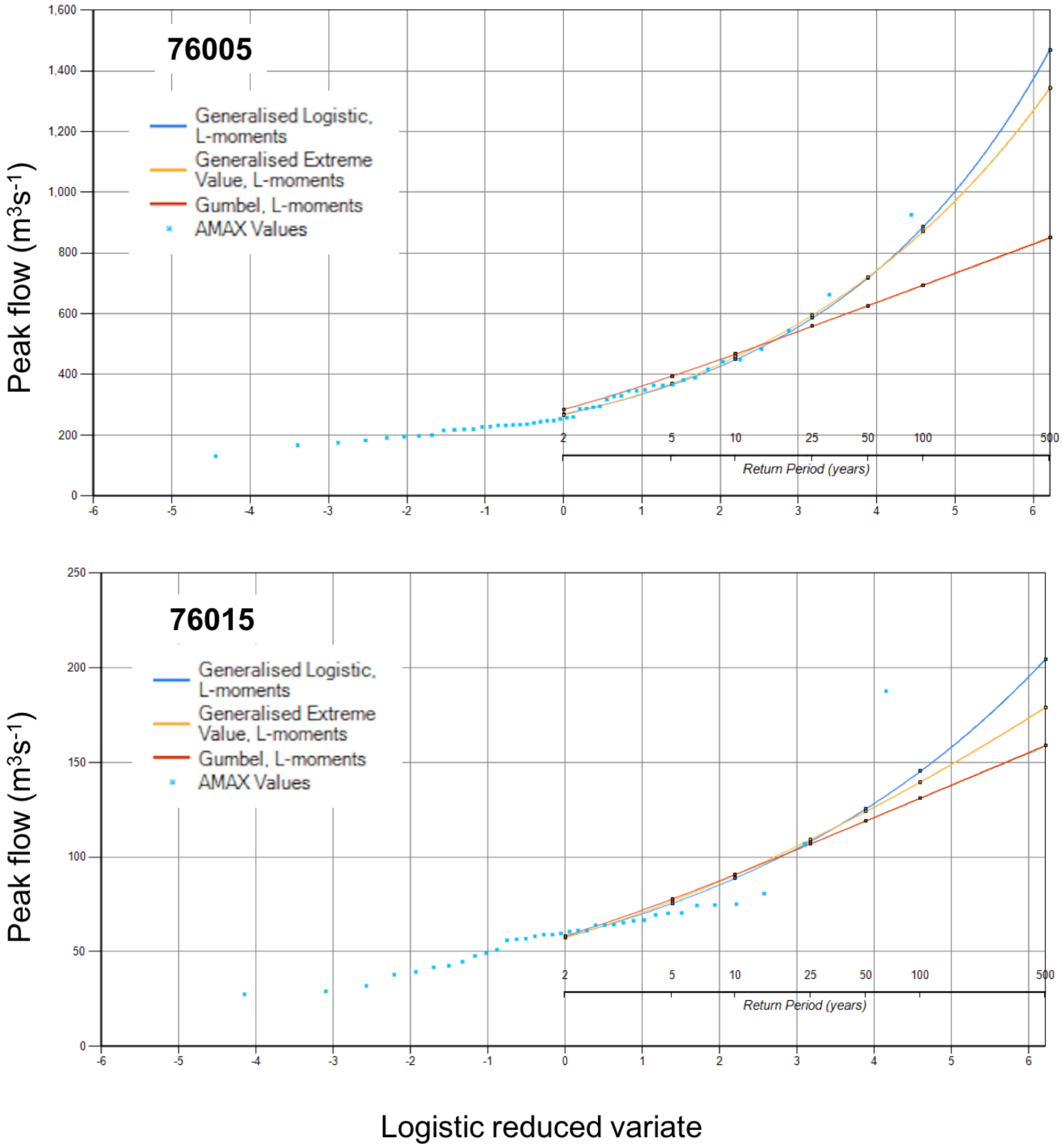
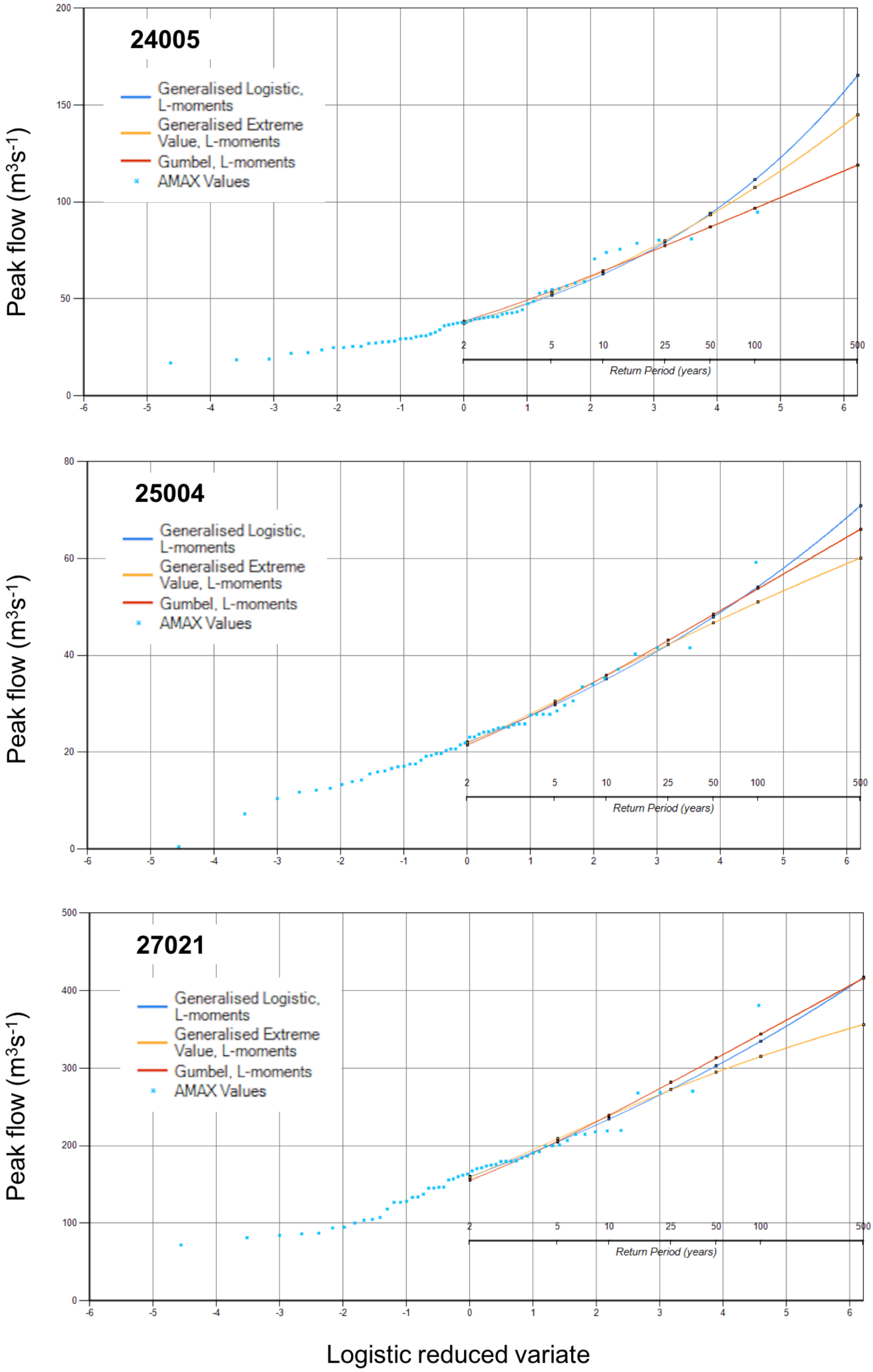


FIGURE 4.9. Illustration of the impact of choice of distribution of flow frequency magnitude relationships at Group A stations (single-site only). Generalised Logistic (GL), Generalised Extreme Value (GEV) and Gumbel (GEV reduces to the special case of a Gumbel when the shape parameter, ξ = 0) distributions were fitted in case. The 'without additional peak' series were used for simplicity. The L-moments fitting method was employed in all cases. The NRFA station reference number is labelled.


24005
Generalised Logistic, L-moments
Generalised Extreme Value, L-moments
Gumbel, L-moments
AMAX Values
Peak flow (m3s-1)
Return Period (years)
2 5 10 25 50 100 500
25004
Generalised Logistic, L-moments
Generalised Extreme Value, L-moments
Gumbel, L-moments
AMAX Values
Peak flow (m3s-1)
Return Period (years)
2 5 10 25 50 100 500
27021
Generalised Logistic, L-moments
Generalised Extreme Value, L-moments
Gumbel, L-moments
AMAX Values
Peak flow (m3s-1)
Return Period (years)
2 5 10 25 50 100 500
Logistic reduced variate

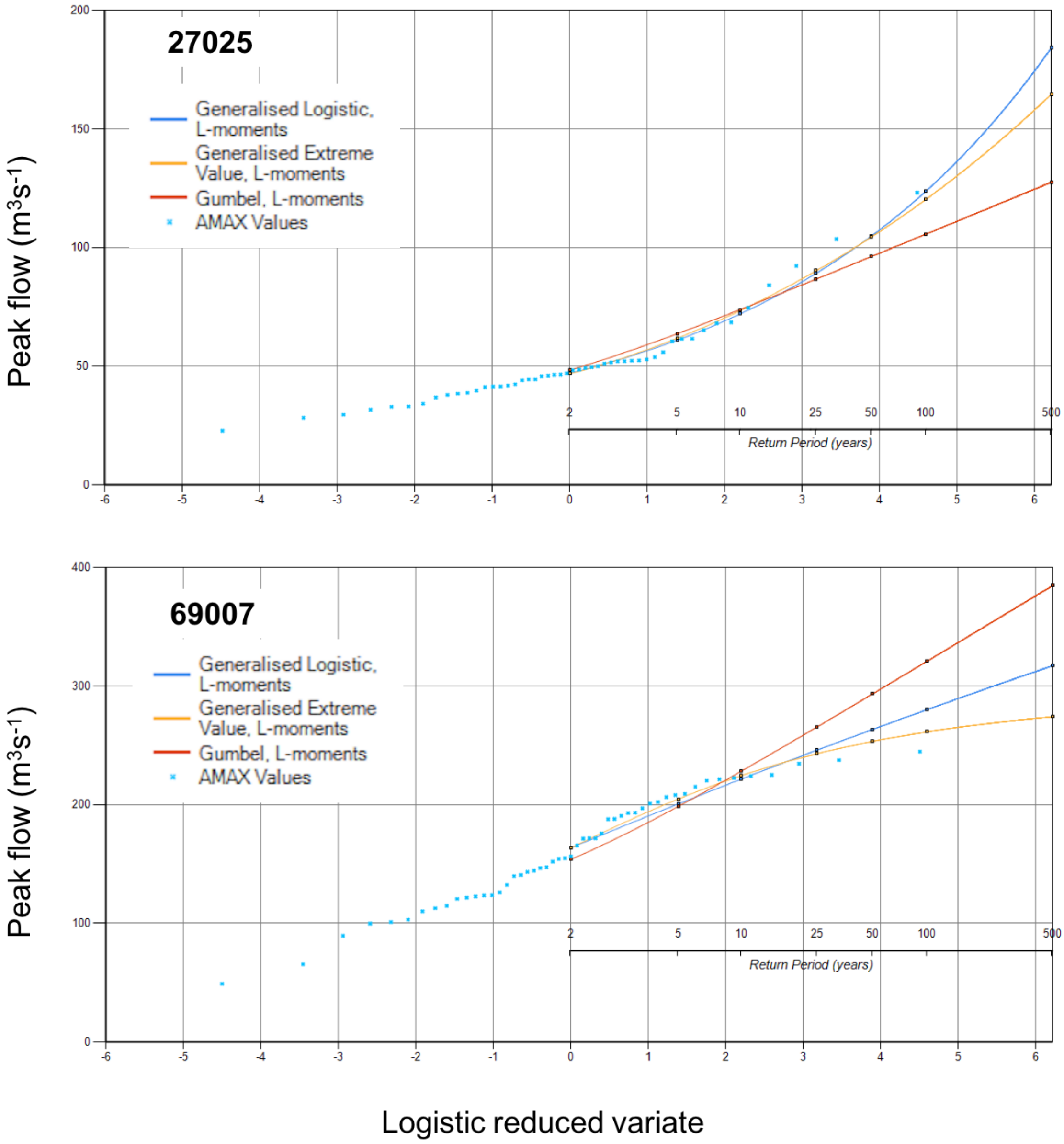


FIGURE 4.10. Illustration of the impact of choice of distribution of flow frequency magnitude relationships at Group B stations (single-site only). Generalised Logistic (GL), Generalised Extreme Value (GEV) and Gumbel (GEV reduces to the special case of a Gumbel when the shape parameter, $\xi = 0$) distributions were fitted in case. The 'without additional data' series were used for simplicity. The L-moments fitting method was employed in all cases. The NRFA station reference number is labelled.

These graphs show that broadly speaking, the results demonstrate only limited sensitivity the choice of statistical distribution. Moreover, little systematic difference is apparent between Group A and Group B stations. Certainly variability in estimates at the 1-in-100-year level associated with the different choice of distribution is lower than that associated with that produced in many cases given additional data (Section 4.1) and whether or not pooling was undertaken or not (Section 4.2). These results accord with those of Kjeldsen et al. (2008) who showed that across 600 UK stations, there was little to choose between the GL and GEV distributions, with the GL being marginally preferred overall. Given all the other difficulties encountered in many flood estimation, this at least provides a little reassurance.

4.5. More general discussion

4.5.1. Regionalisation and comparisons with ReFH results

A key question posed by the main results (these being firstly the observation that the flood-estimates increased overall upon the addition of single new AM observations, and secondly the tentative suggestion that pooled analyses conducted without the additional data may not be capable of sufficiently elevating the estimates) is how spatially transferrable they might be not only to those locations within the study area that did not happen to be severely affected in the events of December 2015, but also to other parts of the UK (particularly upland and western regions).

More specifically, there is an outstanding question regarding whether any factors might prevent similarly extreme floods (relative to the prior records) to those of December 2015 from occurring at other locations in future. It could simply be that purely by chance, major floods have happened to 'miss' certain gauges over the short instrumental period. In other words, the lack of major 'positive outlier' floods in some locations over the instrumental period may be entirely unrelated to locational or physical catchment characteristic. Under these circumstances, such locations clearly maintain the potential to be affected by notable, high impact floods in future. Should it be possible in future to distinguish such locations from those at which flood potential is limited by some physical control, (this would, of course, not be on a binary basis), then it may be possible to somehow raise the frequency magnitude curves on the basis of data from locations where distribution shifting floods have already been observed. In this way, the 'expectedness' or appreciation of the possibility of such floods at these historically fortunate locations might be increased.

Practically, this could involve conducting some form of alternative regionalisation of results to 'less affected' stations by spatially interpolating/extrapolating the annual flow probability distributions, perhaps via their parameters. In this process, one could assign additional weight based on either record length (upon the assumption that these will be more representative) or some more explicit measure of how many large floods have affected that stations per period of time, or even how much the distribution has been raised recently. Such an approach could also facilitate flood estimation in ungauged locations. Alternatively, established regression equations used for flood estimation in ungauged locations could be reassessed in light of the latest data. In a similar way, Merz and Blöschl (2008) call for more hydrologically-informed reasoning in flood frequency estimation.

Making comparisons with results given by the ReFH method may also be enlightening at this juncture. Interestingly, the ReFH method gives growth curves than are generally steeper than those produced by standard pooling of peak flow observations (Kjeldsen et al., 2005). To deal with this mismatch (essentially the possibility of the different methods producing conflicting results at the same location), an explicit recommendation is made in the relevant official guidance to effectively calibrate any ReFH results downwards such that the difference between them and standard pooled results is minimised. To this end, the necessary factors are provided (*Ibid.*). An illustration of the typical situation is given in Figure 4.11.

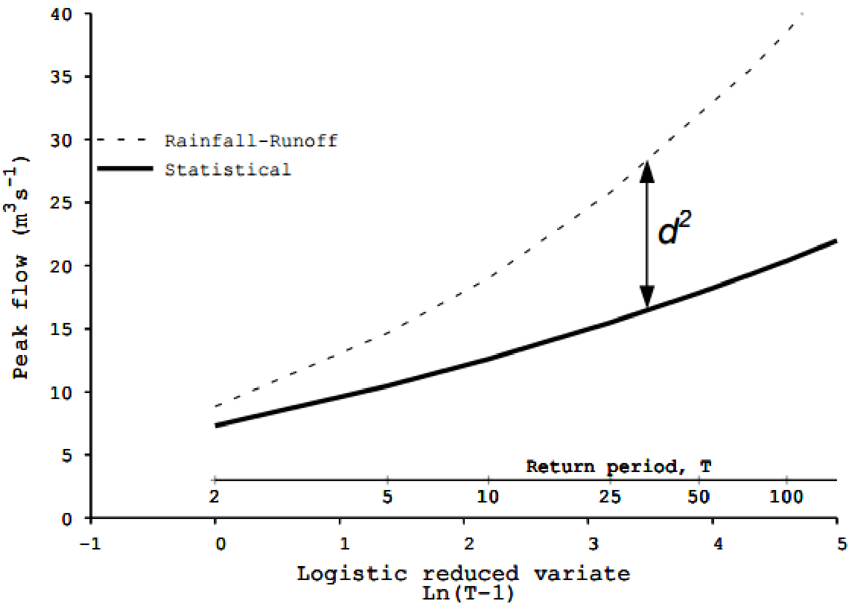


FIGURE 4.11. Illustration of typical flood frequency curves produced by the ReFH method compared to those produced via statistical (pooled) analysis of peak river flow observations. In response to the general mismatch, adjustment factors are provided which can be applied to estimates produced via the ReFH method to improve agreement with the statistical method. In the diagram, the adjustment factor, d^2, is seen to vary with return period. Source: Kjeldsen et al. (2005).

In light of this situation, it would be interesting to explore further whether a proportion of the 'overestimation' is associated with the fact that rainfall records are much longer than river flow records, and hence are more likely to be representative. If this effect makes some contribution, then it may be that the ReFH methods represent less of an overestimation, but rather the pooled analyses something of an underestimation. Perhaps the adjustment should even be the other way around.

4.5.2. Stability of predictions over time

A key theme that has run through this thesis is the difficulty of objectively evaluating alternative 'predictions' that relate to aleatory outcomes drawn from an unknown distribution.

According to one school of thought, it is perfectly obvious and reasonable for the addition of new, relevant data to the flood estimation problem to bring about changes in the estimates when the information available to constrain the underlying probability distribution is somewhat limited at the outset, and information gleaned from elsewhere or from physical theory cannot be easily transferred; from this perspective, the changes are likely to be thought of as improvements.

On the contrary, it may be suggested that the results of a hazard assessment that is 'good' and robust in the first instance (which must surely be the aim), then it should not change significantly with the addition of new data from any single event[36]; if it does, then assuming the new model is an improvement, then the underlying distribution must have been poorly estimated beforehand. Any changes become problematic if the previous model was used as a basis for decision making and its uncertainty and hence instability was not fully acknowledged and accounted for.

Since it is generally impossible to invalidate probabilistic predictions, the stability of flood frequency-magnitude relationships over time might therefore be proposed as a potentially useful indicator of whether a particular model can be considered 'good'.

[36] For analogous discussion in relation to seismic hazard assessment, where the mismatch between the recurrence timescales of the natural phenomena and the lengths of instrumental records is often even more problematic than it is in flood hazard assessment, readers are referred to Stein et al. (2012) and Stein et al. (unpublished).

4.5.3. Simulating spatially coherent plausible future flood events

In this thesis, focus has been placed on the assessment of flood hazard at individual sites. However, if one is managing portfolio of spatially distributed assets or is interested in flooding probabilities over a wide area more generally, one must consider both spatial and temporal correlations in hazard. For this purpose, advanced multivariate extreme value statistical methods are normally applied (Heffernan and Tawn, 2004; Keef et al., 2009, 2013; Wyncoll and Gouldby, 2015). Based on modelled historical extremal dependence structures, these approaches can be applied to simulate many unobserved by nonetheless plausible future flood events or time-series that, when combined into a catalogue, may be considered more-or-less representative of the 'full range' of future flood hazard intensities in time and space.

Neal et al. (2013), for example, simulate the inundation associated with many scenarios of joint fluvial flood flows on tributaries at Carlisle, Cumbria that were generated in such a way. On a larger, typically national or even continental scale, much stochastic modelling underpins catastrophe models, which are chains of hybrid statistical-physical used for calculating a range of flood risk metrics within the (re)insurance industry. Speight (2013) provides useful background information on this topic before applying it for the quantification of fluvial and coastal flood risk to an insured portfolio of static caravans.

4.5.4. Some other factors affecting flood hazard and risk

Flood hazard, and more precisely quantifying the flood frequency element of it, has formed the focus of this thesis. However, this task should not be viewed in isolation. Other factors (which are naturally also uncertain) must come together to ultimately produce a more useful statement of hazard in the form, for instance, of an expected probability of exceedance of a given water depth or velocity at a given location within a particular period.

One important source of uncertainty in producing such a dataset is related to the fact that in hydraulic modelling for national scale flood hazard assessments at least, river channels are not explicitly represented. Therefore, their capacities must typically be assumed, with this volume subtracted from the inflows. Representation of floodplain friction, which can change on a seasonal basis as well as during an individual event (as vegetation is flattened), represents another uncertain parameter in hydraulic simulation, and moreover one to which flood depths and extents can be sensitive. A final point to highlight in this extremely brief summary, is that the actions of flood defences have perhaps the most significant impact on flood hazard of all. As Figure 4.12 reveals, these actions can be extremely dynamic in time; £2.3bn of capital investment in inland and coastal flood defences has recently been announced by the Government, making the trend over recent years that is shown likely to continue (EA, 2015).

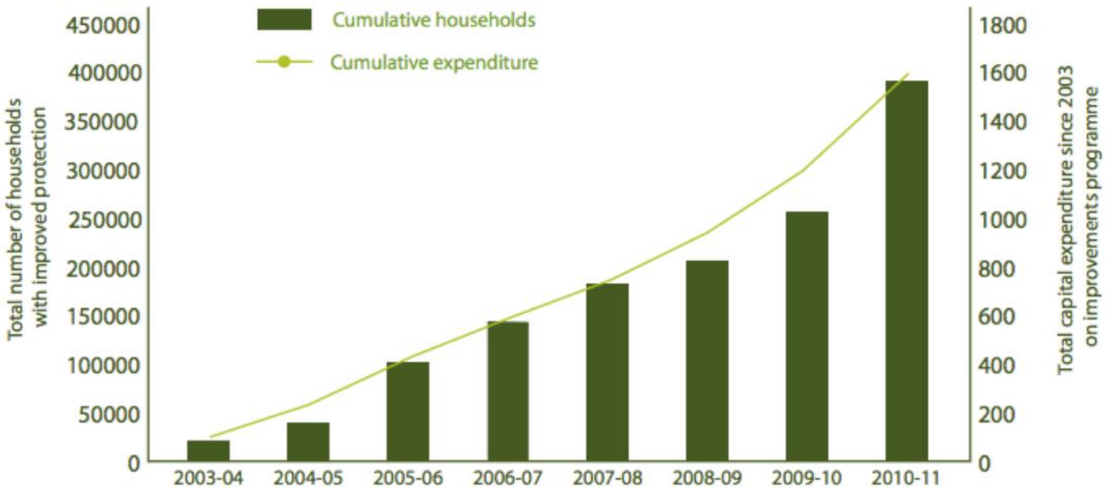


FIGURE 4.12. Cumulative number of households benefiting from reduced likelihood of flooding since 2003-2004 in England. Source: EA (2009).

In terms of exposure, surprisingly, a considerable number of properties continue to be built in areas of known floodplain in the UK; recent estimates suggest perhaps as many as 20,000 each year (Committee on Climate Change, 2015). This rapidly increasing exposure may be partly responsible for some of the recent (real or perceived) increase in flood risk that was discussed in the introduction. On a regional level, and in the shorter term, it is fair to suggest that changes in hazard and risk will be dominated by the balance between construction of new properties in 'at risk' areas and flood defences designed to protect them.

In summary, attention is drawn to such factors and considerations so that the 'materiality' of any uncertainty or variability in flood frequency estimates, or potential changes in hazard overtime, can be evaluated within the wider context. Whilst estimated flood frequency relationships are clearly a major factor in flood hazard assessments, and are known to be a major source of overall uncertainly (compared to the hydraulic modelling stage, for instance), they are by no means the only one. Other factors might be similarly influential and uncertain, and accordingly the amount of effort invested in each should be proportionate. Ultimately, this is another reason (besides the difficulties of testing them) why caution should be urged with respect to suggestions that flood frequency relationships might in cases have been, and still be, underestimated.

4.5.5. A final note: other important types of flooding in the UK

This research has also focussed solely fluvial flooding. Nevertheless, appreciation of the significance of pluvial flooding, which can be especially problematic in an urban context, is increasingly (Blanc et al., 2012). Indeed, in certain recent floods, a considerable proportion of the total monetary damage has been apportioned to non-fluvial sources (Pitt, 2008, Chatterton et al., 2010; 2016). More generally, the close physical linkage between rainfall, runoff and eventual river flows means that pluvial flooding and fluvial flooding are often highly correlated in both time and space.

With respect to coastal flooding, although interactions between the astronomical tidal cycle and meteorological surge (the effects of low pressure and wind-driven waves) can be complex (Horsburgh and Wilson, 2007), the underlying driver of large-scale, low-pressure rain-bearing storm systems is common to all three primary flood types (convective storms excluded). Increases in global mean sea-level are one of the most certain responses of the earth-system to anthropogenic climate change, with increases already eminently observable (Church and White, 2006). This may increase the expectation of coastal flooding. Having said that, since the devastating coastal floods of 1953 (Waverley, 1954), the coastline of the UK has been rather well defended against coastal flooding, and accordingly significant coastal inundation has been limited in recent years (Penning-Rowsell, 2015).

Chapter 5

Conclusions and recommendations

5.1. Conclusions

In December 2015, new river flow maxima were observed widely across northern England. This thesis has presented the first broad scale assessment of the implications of these events on flood frequency estimates produced using established statistical methods. By applying rating equations to peak river stage measurements, annual maxima series were first extended at 155 stations. Then, flood frequency relationships fitted to data series with and without the additional data were compared. Finally, a series of further analyses were undertaken at selected stations.

The research was framed in particular by the suggestion that even notwithstanding any possible climate change related trends in flood hazard, flood frequency estimates based on short instrumental records may have a tendency towards underestimation. As part of the broad review of recent research into the general field of UK fluvial flood hazard that has been presented (Chapter 2), the topical issue of non-stationarity was considered in detail, although it did not ultimately form a primary research focus.

As was expected, inclusion of the new flow data led to increases in flood frequency estimates (produced using the single-site method) at many stations. The fairly high sensitivity that these estimates demonstrated to the addition of the data more generally, including at locations with relatively long record lengths, confirms the established view that the single-site method may not be particularly well suited the estimation of high flow quantiles in the UK given typical record lengths. On the other hand, in contrast to pooled analyses, conducting single-site analysis does at least provide assurance that the data employed for flood estimation are relevant to the location in question.

Having said this, the increases in flood frequency estimates that were produced, as quantified by the change in 1-in-100-year flow, were generally not as large as some of those that have been reported by studies which extended records of flooding historically before re-evaluating flood frequency relationships on this basis. Assuming long-term historical stationarity, this situation may indicate that even with the undoubtedly exceptional December 2015 flow peaks now included, the instrumental records may not represent the full range of possible natural variability with respect to flooding. Consequently, a certain degree of underestimation may persist if the updated instrumental series are used alone

Relationships between change in the 1-in-100-year flow produced on a single-site basis and both record length and catchment area were found to be weak. It seems, therefore, that these factors do not strongly control the sensitivity of single-site flood estimates to additional data, at least within the record length and catchment area ranges studied.

Pooled analyses, meanwhile, represent the primary means by which UK water industry practice seeks to overcome the short record problem. Here, a variant of the pooling approach known as enhanced single-site method was employed to explore the extent to which the longer, theoretically more representative records afforded by pooling might elevate flood frequency estimates produced on a single-site basis (assuming some might suffer from a degree of underestimation). Based on these comparisons, it could be contended that the enhanced method may not be capable of consistently raising hazard estimates to levels that are apparently more consistent with the latest observations (the need for such hedging language is explained shortly).

Additionally, some model-data comparisons were made which, at least in the specific cases presented, appear to support this suggestion. In the final phase of analysis, the impact of the choice of statistical distribution of flood frequency estimates, keeping all else equal, was considered qualitatively. Sensitivity of the estimates to choice of statistical distribution was found to be low, especially when considered relative to the differences produced by using different data and/or methods. This provides some reassurance given the plethora of other uncertainties and challenges associated with the task of flood estimation.

Throughout the thesis, the benefits and limitations associated with the various different flood frequency estimation approaches have been presented and, as far as possible, evaluated. In this regard, it is apparent that no single, optimal approach to flood frequency estimation exists. Hence, comparing estimates provided by several different methods perhaps represents an ideal approach, albeit one that might not be straightforward to carry out in practice.

Overall, it may be tentatively suggested that the results of this study provide some support for one (or both) of the underestimation hypotheses presented in Chapter 1, with potential implications for flood risk management policy.

Having said that, a great deal of caution must be exercised when interpreting these results for two primary reasons – hence the need for hedging. Firstly, it can be extremely difficult (if not impossible) to distinguish between two alternative probabilistic predictions in general, and by using data from only an individual event more particularly; even if an extremely low initial probability was assigned, and so an observed event appears 'too surprising', rare or exceptional (and so underestimation is suspected), provided some probability was assigned to it at the outset, it is not easy to reject the prior model solely on this basis. Secondly, although uncertainty in flood frequency estimates are known to contribute significantly to the overall uncertainty associated with hazard and risk assessments, there are numerous other, perhaps similarly uncertain and influential factors involved in such assessments. As such, uncertainties and potential biases in flood frequency relationships should be evaluated in their light whenever possible. Of course, uncertainty in the December 2015 flow measurements (especially due to flow bypassing the gauging stations and rating curve extrapolation) should also not be overlooked. On this point, not being able to estimate some form of confidence intervals around the estimates in the time available represents a major limitation of the study.

Another significant limitation of the study is that it was not possible to conduct enhanced single-site analyses with the latest December 2015 peaks included. The findings of Miller et al. (2013), however, suggest that a high degree of sensitivity in results could have been expected even with this method. It might also have been interesting to consider the sensitivity of the estimates to (likely) measurement uncertainty around the latest high flows (although their exceptional nature means that quantifying this uncertainty precisely would not be possible).

Finally, it should be noted finally that the revised flood frequency estimates presented herein were produced under the assumption of stationarity. Accordingly, even if any potential underestimation associated with past natural variability not being captured in short records is placed to one side, they may still represent underestimations if increases in flood frequency and/or severity are emergent. Although such trends are not yet detectible in empirical records, both recent event attribution studies and model-based future projections indicate that flood hazard has already increased perceptibly as a result of anthropogenic emissions, and that this signal is anticipated to emerge more clearly from the noise over coming years and decades.

5.2. Recommendations

Several recommendations may be proposed as a result of the literature review and original research phases of this work. They are as follows:

- It may be appropriate to reinvigorate methods which would enable historical (documentary and epigraphic) and palaeoflood data to be included more routinely in flood frequency estimation for practical purposes

- Statistical flood frequency models should be updated more regularly using the latest instrumental observations (but not only following major or record-breaking floods)

- The instability over time and the potential biases of central flood frequency estimates produced in the traditional fashion, and uncertainty around these estimates, should be acknowledged as fully as possible when used for design purposes. The uncertainties considered should include not only the theoretically quantifiable (e.g. confidence intervals related to sample uncertainty) but also the less quantifiable (e.g. uncertainty stemming from errors in the measurement of high flows)

- Advice to pool observations from hydrologically similar sites (either in 'full' pooling or in enhanced single-site analysis) should be followed with care, with an appreciation of the inherent assumptions.; due to hydrological uniqueness, pooling does not always provide a favourable solution to the problem of short, unrepresentative records

- In light of the extreme December 2015 observations, and perhaps also historical 'great flood' data, it may also be apposite to reassess the more complex spatio-temporal flood hazard models that are used operationally, presently, such models also rely entirely on the information contained within short instrumental series

REFERENCES

ABI (2016). https://www.abi.org.uk/News/News-releases/2016/01/New-figures-reveal-scale-of-insurance-response-after-recent-floods. Accessed 3rd February 2016.

Acreman, M. C. (1989). Extreme historical UK floods and maximum flood estimation. *Water and Environment Journal*, *3*(4), 404-412.

Archer, D. R., Leesch, F., & Harwood, K. (2007). Assessment of severity of the extreme River Tyne flood in January 2005 using gauged and historical information. *Hydrological Sciences Journal*, *52*(5), 992-1003.

Archer, D.R. (1987). Improvement in flood estimates using historical flood information on the River Wear at Durham. Proceedings of the BHS First National Hydrology Symposium, Hull, September, 1987. British Hydrological Society.

Archer, D. R. (2010). Applying historical information to Flood risk assessment in north-east England. *BHS* Third International Symposium, Managing Consequences of a Changing Global Environment, Newcastle, July 2010.

Baker, V. R. (1987). Paleoflood hydrology and extraordinary flood events. *Journal of Hydrology*, *96*(1), 79-99.

Baker, V. R. (2008). Paleoflood hydrology: Origin, progress, prospects. *Geomorphology*, *101*(1), 1-13.

Bardossy, A., Coeur, D., & Bobée, B. (2004). Use of systematic, palaeoflood and historical data for the improvement of flood risk estimation. Review of scientific methods. *Natural Hazards*, *31*(3), pp.623-643.

Bates, B.C., Z.W. Kundzewicz, S. Wu and J.P. Palutikof, Eds., 2008: *Climate Change and Water. Technical Paper of the Intergovernmental Panel on Climate Change*, IPCC Secretariat, Geneva.

Bates, P. D., & De Roo, A. P. J. (2000). A simple raster-based model for flood inundation simulation. *Journal of Hydrology*, 236(1), 54-77.

Bayliss, A. C., & Reed, D. W. (2001). The use of historical data in flood frequency estimation. http://nora.nerc.ac.uk/8060/1/BaylissRepN008060CR.pdf. Accessed 9th February 2016.

BBC (2014). http://www.bbc.co.uk/news/uk-politics-25656426. Accessed 24th February 2016.

BBC (2015). http://www.bbc.co.uk/news/uk-35188146. Accessed 15th January 2016.

BBC (2016). http://www.bbc.co.uk/news/uk-england-york-north-yorkshire-35226240. Accessed 15th January 2016.

Beguería, S. (2005). Uncertainties in partial duration series modelling of extremes related to the choice of the threshold value. *Journal of Hydrology*,*303*(1), 215-230.

Bell, V. A., Kay, A. L., Cole, S. J., Jones, R. G., Moore, R. J., & Reynard, N. S. (2012). How might climate change affect river flows across the Thames Basin? An area-wide analysis using the UKCP09 Regional Climate Model ensemble. *Journal of Hydrology*, *442*, 89-104.

Benito, G., Lang, M., Barriendos, M., Llasat, M. C., Francés, F., Ouarda, T., ... & Bobée, B. (2004). Use of systematic, palaeoflood and historical data for the improvement of flood risk estimation. *Review of scientific methods. Natural Hazards, 31*(3), 623-643.

Benito, G., Brázdil, R., Herget, J., & Machado, M.J. (2015). Quantitative historical hydrology in Europe. *Hydrology and Earth System Sciences, 19*, 3517–3539.

Beran, M. (2002). On breaking weather records: Part 1: The past. *Weather, 57*(8), 303-309.

Beven, K. (1993). Prophecy, reality and uncertainty in distributed hydrological modelling. *Advances in Water Resources*, *16*(1), 41-51.

Beven, K. (2006). A manifesto for the equifinality thesis. *Journal of Hydrology*,*320*(1), 18-36.

Beven, K. (2011). I believe in climate change but how precautionary do we need to be in planning for the future? *Hydrological Processes*, *25*(9), 1517-1520.

Beven, K. (2000). Uniqueness of place and process representations in hydrological modelling. *Hydrology and Earth System Sciences Discussions*,*4*(2), 203-213.

Beven, K., Lamb, R., Leedal, D., & Hunter, N. (2015). Communicating uncertainty in flood inundation mapping: a case study. *International Journal of River Basin Management*, *13*(3), 285-295.

Beven, K., Smith, P. J., & Wood, A. (2011). On the colour and spin of epistemic error (and what we might do about it). *Hydrology and Earth System Sciences, 15*(10), 3123-3133

Beven, K., Young, P., Romanowicz, R., O'Connell, E., Ewen, J., O'Donnell, G., Homan, I., Posthumus, H., Morris, J., Hollis, J. & Rose, S. (2008). Analysis of historical data sets to look for impacts of land use and management change on flood generation. Defra R&D Final Report FD2120. Defra, London.

BGS (2016). http://mapapps.bgs.ac.uk/geologyofbritain/home.html. Accessed 6th February 2016.

Biggs, E. M., & Atkinson, P. M. (2011). A characterisation of climate variability and trends in hydrological extremes in the Severn Uplands. *International Journal of Climatology*, *31*(11), 1634-1652.

Black, A. R. (1995). Major flooding and increased flood frequency in Scotland since 1988. *Physics and Chemistry of the Earth, 20*(5), 463-468.

Black, A. R., & Fadipe, D. (2009). Use of historic water level records for re-assessing flood frequency: Case study of the Spey catchment. *Water and Environment Journal*, *23*(1), 23-31.

Blanc, J., Hall, J. W., Roche, N., Dawson, R. J., Cesses, Y., Burton, A., & Kilsby, C. G. (2012). Enhanced efficiency of pluvial flood risk estimation in urban areas using spatial–temporal rainfall simulations. *Journal of Flood Risk Management*, *5*(2), 143-152.

Blöschl, G. (Ed.). (2013). Runoff prediction in ungauged basins: synthesis across processes, places and scales. Cambridge University Press.

Boorman, D. B., Hollis, J. M., & Lilly, A. (1995). Hydrology of soil types: a hydrologically-based classification of the soils of United Kingdom. Institute of Hydrology.

Bradbrook, K., Waller, S., & Morris, D. (2005). National floodplain mapping: datasets and methods–160,000 km in 12 months. *Natural hazards*, *36*(1-2), 103-123.

Bradford, R. B., & Marsh, T. J. (2003, June). Defining a network of benchmark catchments for the UK. In *Proceedings of the Institution of Civil Engineers - Water and Maritime Engineering* (Vol. 156, No. 2, pp. 109-116). Thomas Telford Ltd.

Brázdil, R., Kundzewicz, Z. W., & Benito, G. (2006). Historical hydrology for studying flood risk in Europe. *Hydrological Sciences Journal*, *51*(5), 739-764.

Büchele, B., Kreibich, H., Kron, A., Thieken, A., Ihringer, J., Oberle, P., Merz, B. and Nestmann, F. (2006). Flood-risk mapping: contributions towards an enhanced assessment of extreme events and associated risks. *Natural Hazards and Earth System Sciences*, *6*(4), pp.485-503.

Burton, A., Kilsby, C. G., Fowler, H. J., Cowpertwait, P. S. P., & O'Connell, P. E. (2008). RainSim: A spatial–temporal stochastic rainfall modelling system. *Environmental Modelling & Software*, *23*(12), 1356-1369.

Calver, A., Stewart, E., & Goodsell, G. (2009). Comparative analysis of statistical and catchment modelling approaches to river flood frequency estimation. *Journal of Flood Risk Management*, *2*(1), 24-31.

Cameron, D. (2006). An application of the UKCIP02 climate change scenarios to flood estimation by continuous simulation for a gauged catchment in the northeast of Scotland, UK (with uncertainty). *Journal of Hydrology*, *328*(1), 212-226.

Cameron, D., Beven, K., & Naden, P. (2000). Flood frequency estimation by continuous simulation under climate change (with uncertainty). *Hydrology and Earth System Sciences Discussions*, *4*(3), 393-405.

CEH (2016). http://www.ceh.ac.uk/news-and-media/blogs/briefing-note-severity-december-2015-floods-preliminary-analysis. Accessed 23rd January 2016.

CEH (No Date). http://nrfa.ceh.ac.uk/long-records-overview. Accessed 2nd March 2016.

Chatterton, J., Clarke, C., Daly, E., Dawks, S., Elding, C., Fenn, T., Hick, E., Miller, J., Morris, J., Ogunyoye, F. & Salado, R. (2016) The costs and impacts of the winter 2013 to 2014 floods, Environment Agency. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/501784/The_costs_and_impacts_of_the_winter_2013_to_2014_floods_-_report.pdf. Accessed 19th February 2016.

Chatterton, J., Viavattene, C., Morris, J., Penning-Rowsell, E. C., & Tapsell, S. M. (2010). The costs of the summer 2007 floods in England. Environment Agency. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/291190/scho1109brja-e-e.pdf. Accessed 19th February 2016.

Chiverton, A., Hannaford, J., Holman, I. P., Corstanje, R., Prudhomme, C., Hess, T. M., & Bloomfield, J. P. (2015). Using variograms to detect and attribute hydrological change. *Hydrology and Earth System Sciences*, *19*(5), 2395-2408.

Church, J. A., & White, N. J. (2006). A 20th century acceleration in global sea-level rise. *Geophysical Research Letters*, *33*(1).

Clarke, R. T. (1999). Uncertainty in the estimation of mean annual flood due to rating-curve indefinition. *Journal of Hydrology*, *222*(1), 185-190.

Cleveland, W. S. (1979). Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association*, *74*(368), 829-836.

Cloke, H. L., Wetterhall, F., He, Y., Freer, J. E., & Pappenberger, F. (2013). Modelling climate impact on floods with ensemble climate projections. *Quarterly Journal of the Royal Meteorological Society*, *139*(671), 282-297.

Coles, S. G., & Tawn, J. A. (1994). Statistical methods for multivariate extremes: an application to structural design. *Applied Statistics*, 1-48.

Collier, A. J. (2011). Extreme value analysis of non-stationary processes: a study of extreme rainfall under changing climate. https://theses.ncl.ac.uk/dspace/handle/10443/1219.

Committee on Climate Change (2015). Progress in preparing for climate change, 2015 report to Parliament. https://www.theccc.org.uk/wp-content/uploads/2015/06/6.736_CCC_ASC_Adaptation-Progress-Report_2015_FINAL_WEB_250615_RFS.pdf

Coxon, G., Freer, J., Westerberg, I. K., Wagener, T., Woods, R., & Smith, P. J. (2015). A novel framework for discharge uncertainty quantification applied to 500 UK gauging stations. *Water Resources Research*, *51*(7), 5531-5546.

Cunnane, C. (1987). Review of statistical models for flood frequency estimation. In Hydrologic frequency modeling (pp. 49-95). Springer Netherlands.

Dawson, R. J., Ball, T., Werritty, J., Werritty, A., Hall, J. W., & Roche, N. (2011). Assessing the effectiveness of non-structural flood management measures in the Thames Estuary under conditions of socio-economic and environmental change. *Global Environmental Change*, 21(2), 628-646.

Defra (2015). https://www.gov.uk/government/news/a-country-more-flood-resilient

Defra (2006). FCDPAG3 Economic Appraisal Supplementary Note to Operating Authorities – Climate Change Impacts, October 2006. http://www.sdcg.org.uk/Climate-change-update.pdf

DfT (2015). https://www.gov.uk/government/news/40-million-to-help-rebuild-flood-hit-roads-for-people-in-cumbria-and-lancashire. Accessed 2nd January 2016.

Dixon, H., Faulkner, D., Fry, M., Kral, F., Macklin, M., Morris, D., Prosdocimi, I., Reed, D., Rogers, P., Sefton, C., Stewart , L., & Vesuviano, V. (in press). Making better use of local data in flood frequency estimation, Report SC130009/R

Dixon, H., Lawler, D.M., Shamseldin, A.Y., Webster, P., Demuth, S., Gustard, A., Planos, E., Seatena, F. and Servat, E. (2006). The effect of record length on the analysis of river flow trends in Wales and central England. Climate variability and change: hydrological impacts, pp.490-495.

EA (2009). https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/292928/geho0609bqds-e-e.pdf. Accessed 16th April 2016.

EA (2011). Adapting to Climate Change: Advice for Flood and Coastal Erosion Risk Management Authorities. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/297379/geho0711btzu-e-e.pdf. Accessed 12th March 2016.

EA (2014). Flood and coastal risk management: long-term investment scenarios. https://www.gov.uk/government/publications/flood-and-coastal-risk-management-in-england-long-term-investment. Accessed 25th January 2016.

EA (2015). Central Government funding for flood and coastal erosion risk management in England. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/480527/Funding_for_Flood_and_Coastal_Erosion_in_England_Dec_2015.pdf. Accessed 13th February 2016.

EA (2016a). https://twitter.com/johncurtinEA/status/693122676810977280/photo/1. Accessed 11th February 2016

EA (2016b). https://www.gov.uk/guidance/flood-risk-assessments-climate-change-allowances#table-1. Accessed 28th March 2016.

Faranda, D., Lucarini, V., Turchetti, G., & Vaienti, S. (2011). Numerical convergence of the block-maxima approach to the Generalized Extreme Value distribution. *Journal of Statistical Physics*, *145*(5), 1156-1180.

Foulds, S. A., & Macklin, M. G. (2016). A hydrogeomorphic assessment of twenty-first century floods in the UK. *Earth Surface Processes and Landforms,* 41, 256–270.

Fowler, H. J., & Kilsby, C. G. (2007). Using regional climate model data to simulate historical and future river flows in northwest England. *Climatic Change*, *80*(3-4), 337-367.

Fowler, H. J., Ekström, M., Blenkinsop, S., & Smith, A. P. (2007). Estimating change in extreme European precipitation using a multimodel ensemble. *Journal of Geophysical Research: Atmospheres*, *112* (D18).

Freer, J., Beven, K. J., Neal, J., Schumann, G., Hall, J., & Bates, P. (2013). Flood risk and uncertainty. In Risk and uncertainty assessment for natural hazards, edited by: Rougier, JC, Sparks, RSJ, and Hill, LJ. 190-233.

Greenwood, J. A., Landwehr, J. M., Matalas, N. C., & Wallis, J. R. (1979). Probability weighted moments: definition and relation to parameters of several distributions expressable in inverse form. *Water Resources Research*, *15*(5), 1049-1054.

Grimaldi, S., Petroselli, A., & Serinaldi, F. (2012). A continuous simulation model for design-hydrograph estimation in small and ungauged watersheds. *Hydrological Sciences Journal*, *57*(6), 1035-1051

Gringorten, I. I. (1963). A plotting rule for extreme probability paper. *Journal of Geophysical Research*, *68*(3), 813-814.

Guerreiro, S. B., Kilsby, C. G., & Serinaldi, F. (2014). Analysis of time variation of rainfall in transnational basins in Iberia: abrupt changes or trends?. *International Journal of Climatology*, *34*(1), 114-133.

Gupta, H. V., Beven, K. J., & Wagener, T. (2005). Model calibration and uncertainty estimation. *Encyclopedia of Hydrological Sciences*.

Hall, J. W., Meadowcroft, I. C., Sayers, P. B., & Bramley, M. E. (2003). Integrated flood risk management in England and Wales. *Natural Hazards Review*, 4(3), 126-135.

Hall, M. J., Van Den Boogaard, H. F. P., Fernando, R. C., & Mynett, A. E. (2004). The construction of confidence intervals for frequency analysis using resampling techniques. *Hydrology and Earth System Sciences Discussions, 8*(2), 235-246.

Hannaford, J. (2015). Climate-driven changes in UK river flows A review of the evidence. *Progress in Physical Geography*, 39(1), 29-48.

Hannaford, J., & Buys, G. (2012). Trends in seasonal river flow regimes in the UK. *Journal of Hydrology*, 475, 158-174.

Hannaford, J., & Marsh, T. (2006). An assessment of trends in UK runoff and low flows using a network of undisturbed catchments. *International Journal of Climatology*, 26(9), 1237-1253.

Hannaford, J., & Marsh, T. J. (2008). High-flow and flood trends in a network of undisturbed catchments in the UK. *International Journal of Climatology*, 28(10), 1325-1338.

Harvey, H., Hall, J., & Peppé, R. (2012). Computational decision analysis for flood risk management in an uncertain future. *Journal of Hydroinformatics*, *14*(3), 537-561.

Haxton, T., Crooks, S. Jackson, C.R., Barkwith, A.K.A.P., Kelvin, J., Williamson, J., Mackay, J.D., Wang, L., Davies, H., Young, A. & Prudhomme, C. (2012). Future flows hydrology data. NERC Environmental Information Data Centre. http://doi.org/10.5285/f3723162-4fed-4d9d-92c6-dd17412fa37b

Heffernan, J. E., & Tawn, J. A. (2004). A conditional approach for multivariate extreme values (with discussion). *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *66(3)*, 497-546.

Herring, S. C., Hoerling, M. P., Peterson, T. C., & Stott, P. A. (2014). Explaining extreme events of 2013 from a climate perspective. *Bulletin of the American Meteorological Society*, *95(9)*, S1-S104

Herschy, R. W. (2009). Streamflow measurement. CRC Press.

Horritt, M. S., & Bates, P. D. (2002). Evaluation of 1D and 2D numerical models for predicting river flood inundation. *Journal of Hydrology*, *268(1)*, 87-99.

Horsburgh, K. J., & Wilson, C. (2007). Tide-surge interaction and its role in the distribution of surge residuals in the North Sea. *Journal of Geophysical Research: Oceans*, *112* (C8).

Hosking, J. R. M., & Wallis, J. R. (1988). The effect of intersite dependence on regional flood frequency analysis. *Water Resources Research, 24*(4), 588-600.

Hosking, J.R.M. & Wallis, J.R. (1997). Regional frequency analysis: an approach based on L-moments, Cambridge University Press.

House of Commons (2016a). Environment, Food and Rural Affairs Committee. http://data.parliament.uk/writtenevidence/committeeevidence.svc/evidencedocument/environment-food-and-rural-affairs-committee/winter-floods-201516/oral/26721.pdf. Accessed 24th January 2016.

House of Commons (2016b). Briefing Paper. Winter floods 2015-2016. http://researchbriefings.parliament.uk/ResearchBriefing/Summary/CBP-7424#fullreport. Accessed 22nd January 2'16.

Huntingford, C., Marsh, T., Scaife, A. A., Kendon, E. J., Hannaford, J., Kay, A. L., ... & Lowe, J. A. (2014). Potential influences on the United Kingdom's floods of winter 2013/14. *Nature Climate Change, 4(9)*, 769-777.

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

JBA Consulting (2014). http://www.cotswold.gov.uk/media/1346820/Strategic-Flood-Risk-Assessment-SFRA-Level-2-June-2014.pdf

JBA Trust (2016). http://www.jbatrust.org/news/human-influence-climate-2014-southern-england-winter-floods-impacts/. Accessed 16th February 2016.

Jones, M. R., Fowler, H. J., Kilsby, C. G., & Blenkinsop, S. (2013). An assessment of changes in seasonal and annual extreme rainfall in the UK between 1961 and 2009. *International Journal of Climatology*, *33(5)*, 1178-1194.

Jones, P.D., Lister, D.H. & Kostopoulou, E. (2004). Reconstructed river flow series from 1860s to present. Science Report SC040052/SR. https://crudata.uea.ac.uk/cru/data/riverflow/flowreconstruction/ReconstructedFlowsTR.pdf. Accessed 27th April 2016.

Jones, P. D., & Lister, D. H. (1998). Riverflow reconstructions for 15 catchments over England and Wales and an assessment of hydrologic drought since 1865. *International Journal of Climatology*, *18*(9), 999-1013.

Jones, P. D., Lister, D. H., Wilby, R. L., & Kostopoulou, E. (2006). Extended riverflow reconstructions for England and Wales, 1865–2002. *International Journal of Climatology*, *26*(2), 219-231.

Katz, R. W., Parlange, M. B., & Naveau, P. (2002). Statistics of extremes in hydrology. *Advances in Water Resources*, *25*(8), 1287-1304.

Kay, A. L., & Jones, D. A. (2012). Transient changes in flood frequency and timing in Britain under potential projections of climate change. *International Journal of Climatology*, *32*(4), 489-502.

Kay, A. L., Crooks, S. M., & Reynard, N. S. (2014). Kay, A.L., Crooks, S.M., Davies, H.N., Prudhomme, C. & Reynard, N.S. (2014). Probabilistic impacts of climate change on flood frequency using response surfaces I: England and Wales. *Regional Environmental Change*, 14 (3). 1215-1227.

Kay, A. L., Crooks, S. M., Pall, P., & Stone, D. A. (2011). Attribution of Autumn/Winter 2000 flood risk in England to anthropogenic climate change: A catchment-based study. *Journal of Hydrology*, *406*(1), 97-112.

Kay, A. L., Davies, H. N., Bell, V. A., & Jones, R. G. (2009). Comparison of uncertainty sources for climate change impacts: flood frequency in England. *Climatic Change*, *92*(1-2), 41-63.

Keef, C., Svensson, C., & Tawn, J. A. (2009). Spatial dependence in extreme river flows and precipitation for Great Britain. *Journal of Hydrology*, *378*(3), 240-252.

Keef, C., Tawn, J. A., & Lamb, R. (2013). Estimating the probability of widespread flood events. *Environmetrics*, *24*(1), 13-21.

Keller, V.D.J., Tanguy, M., Prosdocimi, I., Terry, J.A., Hitt, O., Cole, S.J., Fry, M., Morris, D.G. and Dixon, H. (2015). CEH-GEAR: 1 km resolution daily and monthly areal rainfall estimates for the UK for hydrological and other applications. *Earth System Science Data*, *7*(1), pp.143-155.

Kendall, M. G. (1975). Rank Correlation Methods, Oxford University Press, New York.

Kidson, R., & Richards, K. S. (2005). Flood frequency analysis: assumptions and alternatives. *Progress in Physical Geography*, *29*(3), 392-410.

Kjeldsen, T. R. (2010). Modelling the impact of urbanization on flood frequency relationships in the UK. *Hydrology Research*, *41*(5), 391-405.

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

Kjeldsen, T. R., & Jones, D. A. (2004). Sampling variance of flood quantiles from the generalised logistic distribution estimated using the method of L-moments. *Hydrology and Earth System Sciences Discussions*, *8*(2), 183-190.

Kjeldsen, T. R., & Jones, D. A. (2006). Prediction uncertainty in a median-based index flood method using L moments. *Water Resources Research*,*42*(7).

Kjeldsen, T. R., & Jones, D. A. (2009). A formal statistical model for pooled analysis of extreme floods. *Hydrology Research*, *40*(5), 465-480.

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

Kjeldsen, T., Jones, D., Bayliss, A., Spencer, P., Surendran, S., Stefan, L., Webster, P., & McDonald, D. (2008a). Improving the FEH statistical method. In: Flood & Coastal Management Conference 2008, University of Manchester, 1-3 July 2008. Environment Agency/Defra.

KPMG (2015). https://home.kpmg.com/uk/en/home/media/press-releases/2015/12/flooding-economic-impact-will-breach-5bn.html. 2nd January 2016.

Lafon, T., Dadson, S., Buys, G., & Prudhomme, C. (2013). Bias correction of daily precipitation simulated by a regional climate model: a comparison of methods. *International Journal of Climatology*, *33*(6), 1367-1381.

Laizé, C. L., & Hannah, D. M. (2010). Modification of climate–river flow associations by basin properties. *Journal of Hydrology*, *389*(1), 186-204.

Lamb, R., & Kay, A. L. (2004). Confidence intervals for a spatially generalized, continuous simulation flood frequency model for Great Britain. *Water Resources Research*, *40*(7).

Lavers, D. A., Allan, R. P., Villarini, G., Lloyd-Hughes, B., Brayshaw, D. J., & Wade, A. J. (2013). Future changes in atmospheric rivers and their implications for winter flooding in Britain. *Environmental Research Letters*, *8*(3).

Lavers, D. A., Allan, R. P., Wood, E. F., Villarini, G., Brayshaw, D. J., & Wade, A. J. (2011). Winter floods in Britain are connected to atmospheric rivers. *Geophysical Research Letters*, *38*(23).

Lins, H.F. (2012). A Note on stationarity and nonstationarity, WMO. http://www.wmo.int/pages/prog/hwrp/chy/chy14/documents/ms/Stationarity_and_Nonstationarity.pdf. Accessed 4th February 2016.

Luo, C. Z. (1987). Investigation and regionalization of historical floods in China. *Journal of Hydrology*, *96*(1), 41-51.

Macdonald, N. (2013). Reassessing flood frequency for the River Trent through the inclusion of historical flood information since AD 1320. *Hydrology Research*, *44*(2), 215-233.

Macdonald, N., & Black, A. R. (2010). Reassessment of flood frequency using historical information for the River Ouse at York, UK (1200–2000). *Hydrological Sciences Journal–Journal des Sciences Hydrologiques*, *55*(7), 1152-1162.

Macdonald, N., Kjeldsen, T. R., Prosdocimi, I., & Sangster, H. (2014). Reassessing flood frequency for the Sussex Ouse, Lewes: the inclusion of historical flood information since AD 1650. *Natural Hazards and Earth System Sciences*, *14*(10), 2817-2828.

Macklin., M. G. (2015). http://www.cam.ac.uk/research/news/unprecedented-storms-and-floods-are-more-common-than-we-think

Madsen, H., Rasmussen, P.F. & Rosbjerg, D. (1997a). Comparison of annual maximum series and partial duration series methods for modeling extreme hydrologic events 1. At-site modelling. *Water Resources Research*, *33*(4), 747-757.

Madsen, H., Pearson, C.P. & Rosbjerg, D. (1997b). Comparison of annual maximum series and partial duration series methods for modeling extreme hydrologic events 2. Regional modeling. *Water Resources Research*, *33*(4), 759-769.

Marsh, T. J., & Dale, M. (2002). The UK floods of 2000–2001: a hydrometeorological appraisal. *Water and Environment Journal*, *16*(3), 180-188.

Marsh, T. J., & Hannaford, J. (2007). The summer 2007 floods in England and Wales–a hydrological appraisal. NERC/Centre for Ecology & Hydrology. http://nora.nerc.ac.uk/2814/1/N002814BK.pdf. Accessed 5th March 2016.

Marsh, T., & Harvey, C. L. (2012). The Thames flood series: a lack of trend in flood magnitude and a decline in maximum levels. *Hydrology Research*,*43*(3), 203-214.

Martins, E. S., & Stedinger, J. R. (2000). Generalized maximum-likelihood generalized extreme-value quantile estimators for hydrologic data. *Water Resources Research*, *36*(3), 737-744.

Massey, N., Jones, R., Otto, F.E.L., Aina, T., Wilson, S., Murphy, J.M., Hassell, D., Yamazaki, Y.H. & Allen, M.R. (2015). weather@ home—development and validation of a very large ensemble modelling system for probabilistic event attribution. *Quarterly Journal of the Royal Meteorological Society*, *141*(690), 1528-1545.

Matalas, N. C. (1997). Stochastic Hydrology in the Context of Climate Change. *Climatic Change*, *37*(1), 89-101.

Merwade, V., Olivera, F., Arabi, M., & Edleman, S. (2008). Uncertainty in flood inundation mapping: current issues and future directions. *Journal of Hydrologic Engineering*, *13*(7), 608-620.

Merz, B., Thieken, A. H., & Gocht, M. (2007). Flood risk mapping at the local scale: concepts and challenges. In Flood risk management in Europe (pp. 231-251). Springer Netherlands.

Merz, R., & Blöschl, G. (2008). Flood frequency hydrology: 1. Temporal, spatial, and causal expansion of information. *Water Resources Research*,*44*(8).

Merz, B., Aerts, J. C. J. H., Arnbjerg-Nielsen, K., Baldi, M., Becker, A., Bichet, A., ... & Delgado, J. M. (2014). Floods and climate: emerging perspectives for flood risk assessment and management. *Natural Hazards and Earth System Sciences*, *14*(7), 1921-1942.

Met Office (2014). The recent storms and floods in the UK, February 2014. http://www.metoffice.gov.uk/media/pdf/n/i/Recent_Storms_Briefing_Final_07023.pdf. Accessed 17th February 2016.

Met Office (2015a). http://blog.metoffice.gov.uk/2015/12/31/whats-been-happening-to-our-weather/. Accessed 3rd January 2016.

Met Office (2015b). http://www.metoffice.gov.uk/climate/uk/summaries/2015/november. Accessed 3rd January 2016

Met Office (2015c). http://www.metoffice.gov.uk/mobile/news/article/news/releases/archive/2015/name-our-storms. Accessed 2nd January 2016.

Met Office (2015d). http://blog.metoffice.gov.uk/2015/12/07/did-climate-change-have-an-impact-on-storm-desmond/. Accessed 3rd January 2016

Met Office (2016). http://blog.metoffice.gov.uk/2015/12/31/whats-been-happening-to-our-weather/. Accessed 3rd January 2016

Met Office (no date). http://www.metoffice.gov.uk/climate/uk/summaries/anomacts. Accessed 6rd January 2016

Miller, J. D., Kjeldsen, T. R., Hannaford, J., & Morris, D. G. (2013). A hydrological assessment of the November 2009 floods in Cumbria, UK. *Hydrology Research*, *44*(1), 180-197.

Montanari, A., & Koutsoyiannis, D. (2014). Modeling and mitigating natural hazards: Stationarity is immortal! *Water Resources Research, 50*(12), 9748-9756.

Murphy, J. M., Sexton, D., Jenkins, G., Boorman, P., Booth, B., Brown, K., ... & Kendon, E. (2009). UKCP09 Climate change projections. Met Office Hadley Centre, Exeter.

Natural England (2016). http://publications.naturalengland.org.uk/category/5954148537204736. Accessed 19th March 2016.

Neal, J., Keef, C., Bates, P., Beven, K., & Leedal, D. (2013). Probabilistic flood risk mapping including spatial dependence. *Hydrological Processes*, *27*(9), 1349-1363.

NERC (1975). Flood Studies Report, Natural Environmental Research Council. Department of the Environment, London.

New, M., Lopez, A., Dessai, S., & Wilby, R. (2007). Challenges in using probabilistic climate change information for impact assessments: an example from the water sector. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, *365*(1857), 2117-2131.

Newton, D. W. (1983). Realistic assessment of maximum flood potentials. *Journal of Hydraulic Engineering*, *109*(6), 905-918.

Niehoff, D., Fritsch, U., & Bronstert, A. (2002). Land-use impacts on storm-runoff generation: scenarios of land-use change and simulation of hydrological response in a meso-scale catchment in SW-Germany. *Journal of Hydrology*, *267*(1), 80-93.

Open Data Institute (2016). http://theodi.org/ea-going-open-how-and-why. Accessed 20th January 2016.

Orchard, K. L. (2001). Paleoflood Hydrology of the San Juan River, Southeastern Utah, USA. M.S. Thesis, The University of Arizona.

Pall, P., Aina, T., Stone, D.A., Stott, P.A., Nozawa, T., Hilberts, A.G., Lohmann, D. and Allen, M.R. (2011). Anthropogenic greenhouse gas contribution to flood risk in England and Wales in autumn 2000. *Nature*, *470*(7334), 382-385.

Pattison, I., & Lane, S. N. (2012). The relationship between Lamb weather types and long-term changes in flood frequency, River Eden, UK. *International Journal of Climatology*, 32(13), 1971-1989.

Penning-Rowsell, E. C. (2015). A realistic assessment of fluvial and coastal flood risk in England and Wales. *Transactions of the Institute of British Geographers*, *40*(1), 44-61.

Penning-Rowsell, E., Johnson, C., Tunstall, S., Tapsell, S., Morris, J., Chatterton, J., & Green, C. (2005). The benefits of flood and coastal risk management: a handbook of assessment techniques. Middlesex University Press.

Pitt, M. (2008). The Pitt Review: Learning lessons from the 2007 floods. Cabinet Office, London, UK. http://webarchive.nationalarchives.gov.uk/20100807034701/http://archive.cabinetoffice.gov.uk/pittreview/_/media/assets/www.cabinetoffice.gov.uk/flooding_review/pitt_review_full%20pdf.pdf. Accessed 19th February 2016.

Porter, R.A.I-H. (2011). An evaluation of a hydraulic model used to predict the impact of flow attenuation on downstream flood flows. Durham Theses, Durham University. http://etheses.dur.ac.uk/813/1/I-Hsien_Porter_Thesis_-_Final.pdf?DDD14+. Accessed 15th April 2016.

Porter, J., & Demeritt, D. (2012). Flood-risk management, mapping, and planning: the institutional politics of decision support in England. *Environment and Planning A*, *44*(10), 2359-2378.
Potter, K. W., & Lettenmaier, D. P. (1990). A comparison of regional flood frequency estimation methods using a resampling method. *Water Resources Research*, *26*(3), 415-424.

Priestley and Edmonds (2016). Household flood insurance Briefing Paper, January 2016. researchbriefings.files.parliament.uk/documents/SN06613/SN06613.pdf. Accessed 22nd January 2016.

Prosdocimi, I., Kjeldsen, T. R., & Miller, J. D. (2015). Detection and attribution of urbanization effect on flood extremes using nonstationary flood-frequency models. *Water Resources Research*, *51*(6), 4244-4262.

Prosdocimi, I., Kjeldsen, T. R., & Svensson, C. (2014). Non-stationarity in annual and seasonal series of peak flow and precipitation in the UK. *Natural Hazards and Earth System Sciences*, *14*, 1125-1144.

Prudhomme & Reynard, 2009. Regionalised Impacts of Climate Change on Flood Flows: Rationale for Definition of Climate Change Scenarios and Sensitivity Framework R&D Milestone Report FD2020/MR2

Prudhomme, C., Crooks, S., Kay, A. L., & Reynard, N. (2013). Climate change and river flooding: part 1 classifying the sensitivity of British catchments. *Climatic Change*, *119*(3-4), 933-948.

Prudhomme, C., Dadson, S., Morris, D., Williamson, J., Goodsell, G., Crooks, S., Boelee, L., Davies, H., Buys, G., Lafon, T. and Watts, G. (2012). Future Flows Climate: an ensemble of 1-km climate change projections for hydrological application in Great Britain. *Earth System Science Data*, *4*(1), pp.143-148.

Prudhomme, C., Wilby, R. L., Crooks, S., Kay, A. L., & Reynard, N. S. (2010). Scenario-neutral approach to climate change impact studies: application to flood risk. *Journal of Hydrology*, 390(3), 198-209.

Quinn, N. & Horswell, M. (2014) Trends in hydrological intensification in the UK: Implications for flood risk management. In: 6th International Conference on Flood Management, Sao Paulo, Brazil, 16-18 September 2014. http://eprints.uwe.ac.uk/24488. Accessed 13th February 2016.

Reed, D. W. (2002). Reinforcing flood–risk estimation. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, *360*(1796), 1373-1387.

Reitan, T., & Petersen-Øverleir, A. (2009). Bayesian methods for estimating multi-segment discharge rating curves. *Stochastic Environmental Research and Risk Assessment*, *23*(5), 627-642.

Reynard, N.S., Crooks, S., Kay, A.L. & Prudhomme, C. (2009). Regionalised impacts of climate change on flood flows. R&D Technical Report FD2020/TR. http://randd.defra.gov.uk/Document.aspx?Document=FD2020_8852_FRP.pdf. Accessed 2nd March 2016.

Robson, A. J. (2002). Evidence for trends in UK flooding. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, *360*(1796), 1327-1343.

Rogger, M., Kohl, B., Pirkl, H., Viglione, A., Komma, J., Kirnbauer, R., Merz, R. and Blöschl, G. (2012). Runoff models and flood frequency statistics for design flood estimation in Austria–Do they tell a consistent story? *Journal of Hydrology*, *456*, 30-43.

Rougier, J. C. (2013). Quantifying hazard losses. In Risk and Uncertainty Assessment for Natural Hazards, edited by: Rougier, JC, Sparks, RSJ, and Hill, LJ, 19-39.

Scaife, A. A., Folland, C. K., Alexander, L. V., Moberg, A., & Knight, J. R. (2008). European climate extremes and the North Atlantic Oscillation. *Journal of Climate*, *21*(1), 72-83.

Schaller, N., Kay, A.L., Lamb, R., Massey, N.R., Van Oldenborgh, G.J., Otto, F.E., Sparrow, S.N., Vautard, R., Yiou, P., Ashpole, I. and Bowery, A., 2016. Human influence on climate in the 2014 southern England winter floods and their impacts. *Nature Climate Change*.

Serinaldi, F., & Kilsby, C. G. (2015). Stationarity is undead: Uncertainty dominates the distribution of extremes. *Advances in Water Resources, 77*, 17-36.

Serinaldi, F., & Kilsby, C. G. (2016). Understanding Persistence to Avoid Underestimation of Collective Flood Risk. *Water*, *8*(4), 152.

Smith, A., Freer, J., Bates, P., & Sampson, C. (2014). Comparing ensemble projections of flooding against flood estimation by continuous simulation. *Journal of Hydrology*, *511*, 205-219.

Stedinger, J. R., & Cohn, T. A. (1986). Flood frequency analysis with historical and paleoflood information. *Water Resources Research*, *22*(5), 785-793.

Stedinger, J. R., Vogel, R. M., & Foufoula-Georgiou E. (1993) Frequency analysis of 663 extreme events. In: Handbook of Hydrology (Ed. D. R. Maidment), Chapter 664 18, McGraw-Hill, New York.

Stein, S., Geller, R. J., & Liu, M. (2012). Why earthquake hazard maps often fail and what to do about it. *Tectonophysics*, *562,* 1-25.

Stein, S., Liu, M., Spencer, B. D., & Brooks, E. M. (unpublished) Promise and Paradox: why improved knowledge of plate tectonics hasn't yielded correspondingly better earthquake hazard maps. http://www.earth.northwestern.edu/people/seth/Texts/paradox.pdf. Accessed 20th December 2016.

Svensson, C., Brookshaw, A., Scaife, A.A., Bell, V.A., Mackay, J.D., Jackson, C.R., Hannaford, J., Davies, H.N., Arribas, A., & Stanley, S. (2015) Long-range forecasts of UK winter hydrology. *Environmental Research Letters*, *10*(6).

Svensson, C., Kjeldsen, T. R., & Jones, D. A. (2013). Flood frequency estimation using a joint probability approach within a Monte Carlo framework. *Hydrological Sciences Journal*, *58*(1), 8-27.

The Law Society (2016). http://www.lawsociety.org.uk/support-services/advice/practice-notes/flood-risk/. Accessed 14th February 2016.

Thornton, J.M., Thomson, T., Liu, Y., Chaney, S., Dunning, P., Hutchings, S., Taylor, P., & Pickering, C. (2014). National scale high-resolution quantification of fluvial flood risk in Great Britain. In EGU General Assembly Conference Abstracts (Vol. 16, p. 11117).

van Oldenborgh, G. J., Otto, F. E., Haustein, K., & Cullen, H. (2015). Climate change increases the probability of heavy rains like those of storm Desmond in the UK–an event attribution study in near-real time. *Hydrology and Earth System Sciences Discussions*, *12*, 13197-13216.

Wagener, T., & Wheater, H. S. (2006). Parameter estimation and regionalization for continuous rainfall-runoff models including uncertainty. *Journal of Hydrology*, *320*(1), 132-154.

Wagener, T., McIntyre, N., Lees, M. J., Wheater, H. S., & Gupta, H. V. (2003). Towards reduced uncertainty in conceptual rainfall-runoff modelling: Dynamic identifiability analysis. *Hydrological Processes*, *17*(2), 455-476.

Watts, G., Battarbee, R.W., Bloomfield, J.P., Crossman, J., Daccache, A., Durance, I., Elliott, J.A., Garner, G., Hannaford, J., Hannah, D.M., & Hess, T. (2015). Climate change and water in the UK–past changes and future prospects. *Progress in Physical Geography*, *39*(1), pp.6-28.

Waverley, J. A. (1954) Report of the Departmental Committee on Coastal Flooding (London: HMSO).

Weerts, A. H., Winsemius, H. C., & Verkade, J. S. (2011). Estimation of predictive hydrological uncertainty using quantile regression: examples from the National Flood Forecasting System (England and Wales). *Hydrology and Earth System Sciences*, 15,(1).

Wentz, F. J., Ricciardulli, L., Hilburn, K., & Mears, C. (2007). How much more rain will global warming bring? *Science*, *317*(5835), 233-235.

Werner, M., Cranston, M., Harrison, T., Whitfield, D., & Schellekens, J. (2009). Recent developments in operational flood forecasting in England, Wales and Scotland. *Meteorological Applications*, *16*(1), 13-22.

Wheater H.S., Isham, V.S., Chandler, R.E., Onof, C.J. & Stewart E.J. (2006). Improved methods for national spatial-temporal rainfall and evaporation modelling for BSM. Environment Agency. http://evidence.environment-agency.gov.uk/FCERM/Libraries/FCERM_Project_Documents/FD2105_6108_TRP_pdf.sflb.ashx. Accessed 28th March 2016.

Wheater, H., & Evans, E. (2009). Land use, water management and future flood risk. *Land Use Policy*, *26*, S251-S264.

Wilby, R. L., & Dessai, S. (2010). Robust adaptation to climate change. *Weather*, *65*(7), 180-185.

Wilby, R. L., & Keenan, R. (2012). Adapting to flood risk under climate change. *Progress in Physical Geography.* 1-31

Wilby, R. L., & Quinn, N. W. (2013). Reconstructing multi-decadal variations in fluvial flood risk using atmospheric circulation patterns. *Journal of Hydrology*,*487*, 109-121.

Wilby, R. L., Beven, K. J., & Reynard, N. S. (2008). Climate change and fluvial flood risk in the UK: more of the same? *Hydrological Processes*,*22*(14), 2511-2523.

Wilks, D. S., & Wilby, R. L. (1999). The weather generation game: a review of stochastic weather models. *Progress in Physical Geography*, *23*(3), 329-357.

Wilks, D. S. (2006). On "Field Significance" and the False Discovery Rate. *Journal of Applied Meteorology & Climatology*, *45*(9).

WINFAP-FEH 3 User Manual (2009). http://www.e-secure.biz/documents/KLGJU9XGH9/WINFAP-FEH3-UserGuide.pdf. Accessed 16th February 2016.

APPENDICES

Appendix A: December 2015 peak flow estimates, northern England

Station	Ref	Easting	Northing	Maximum December 2015 stage (m)	Site and stage-specific rating equation	Equation applied	Peak Q ($m^3 s^{-1}$)	Comments
Heaton Mill	21806	389977	642381	2.155	2c	Q = 3.024 * (h + 1.551) ^ 2.791	117.06	
Morwick	22001	423400	604500	3.773	Db	Q = 33.106* (h + 0.06400) ^ 1.428	225.86	
Shilmoor	22003	386800	608700	0.781	Ba	Q = 23.862 * (h - 0.01700) ^ 1.921	14.23	
Hartford Bridge	22006	424336	580000	1.312	2b	Q = 33.977 * (h - 0.02200) ^ 2.127	58.40	
Mitford	22007	417500	585800	2.358	4d	Q = 19.476 * (h + 0.000) ^ 2.027	110.83	
Rothbury	22009	406700	601600	2.708	4c	Q = 31.492 * (h - 0.2480) ^ 1.934	179.58	
Stamfordham	22801	408200	571800	2.264	4c	Q = 1.550 * (h - 0.1681) ^ 3.642	22.95	
Bywell	23001	403800	561700	6.959	2c	Q = 72.522 * (h - 0.06100) ^ 1.603	1603.04	
Eddys Bridge	23002	404100	550800	1.382	1b	Q = 11.739 * (h - 0.2500) ^ 2.771	16.55	
Reaverhill	23003	390600	573200	5.072	4a	Q = 32.323 * (h - 0.04500) ^ 1.918	715.52	Rating applied beyond max stage
Haydon Bridge	23004	385600	564700	4.646	9c	Q = 62.303 * (h + 0.000) ^ 1.749	914.60	
Featherstone	23006	367200	561100	2.462	1c	Q = 101.400 * (h - 0.1500) ^ 1.720	428.64	
Rowlands Gill	23007	416814	558095	1.220	1c	Q = 44.000 * (h + 0.000) ^ 1.658	61.18	
Rede Bridge	23008	386800	583200	3.045	3c	Q = 34.400 * (h + 0.000) ^ 1.823	261.90	
Alston	23009	371600	546500	2.472	4b	Q = 3.421 * (h + 0.8524) ^ 3.504	230.27	
Kielder Burn	23011	364400	594600	1.823	1c	Q = 22.530 * (h + 0.000) ^ 1.842	68.10	
Team Valley	23017	424900	558500	0.867	3b	Q = 6.304 * (h + 0.2510) ^ 2.798	8.61	

Woolsington	23018	419500	569900	1.106	6c	Q = 1.390 * (h + 0.2000) ^ 2.654	2.82	
Otterburn	23033	386400	594400	3.454	Bv2c	Q = 142.410 * (h - 2.301) ^ 1.644	179.96	
Sunderland Bridge	24001	426500	537800	3.263	3c	Q = 82.484 * (h - 0.2230) ^ 1.283	343.48	
Stanhope	24003	398300	539100	3.432	2c	Q = 98.400 * (h - 0.9000) ^ 0.8900	224.95	
Bedburn	24004	411800	532200	1.416	1d	Q = 18.350 * (h - 0.1000) ^ 1.848	30.48	
Burn Hall	24005	425900	538700	1.235	Dc	Q = 20.230 * (h + 0.000) ^ 2.090	31.45	
Eastgate	24006	395200	539000	1.276	1d	Q = 17.450 * (h - 0.1630) ^ 1.817	21.20	Rating only goes up to 1980
Witton Park	24008	417300	530900	4.088	2c	Q = 20.355 * (h + 0.000) ^ 1.943	313.93	
Chester-Le-Street	24009	428400	551200	3.434	1c	Q = 89.405 * (h + 0.000) ^ 1.029	318.20	
Broken Scar	25001	425900	513700	3.187	e2	Q = 381.662 * (h - 1.539) ^ 1.001	629.29	
Moorhouse	25003	375900	533600	0.859	2d	Q = 24.293 * (h - 0.1950) ^ 1.770	11.77	
South Park	25004	428400	512900	0.985	1d	Q = 32.806 * (h - 0.2080) ^ 1.785	20.91	
Rutherford Bridge	25006	403400	512200	1.801	Dd	Q = 37.822 * (h - 0.1750) ^ 1.784	90.03	
Barnard Castle	25008	404700	516600	2.542	2d	Q = 95.581 * (h - 0.3020) ^ 1.746	390.76	
Low Moor	25009	436400	510500	6.231	2f	Q = 9.302 * (h - 1.207) ^ 2.533	555.05	
Harwood	25012	384900	530900	1.552	1a	Q = 17.200 * (h + 0.000) ^ 2.514	51.93	
Middleton	25018	395000	525000	2.855	1a	Q = 24.200 * (h + 0.000) ^ 2.424	307.75	
Easby	25019	458500	508700	0.497	1c	Q = 15.000 * (h + 0.000) ^ 2.347	2.91	
Preston-Le-Skerne	25020	429200	523800	1.266	1b	Q = 8.818 * (h + 0.000) ^ 1.464	12.45	
Bradbury	25021	431800	528500	1.307	2b	Q = 4.716 * (h + 0.000) ^ 1.323	6.72	
Foston Mill	26003	509300	454800	0.354	2b	Q = 5.872 * (h + 0.03300) ^ 1.728	1.14	
Wansford Snakeholm Lock	26010	506600	455600	2.063	1a	Q = 5.601 * (h - 1.565) ^ 1.671	1.75	
Hunsingore	27001	442800	453000	2.852	11d	Q = 7.743 * (h + 0.000) ^ 3.265	297.44	Rating applied beyond max stage

Flint Mill	27002	442200	447300	3.743	2b	Q = 74.004 * (h - 0.2620) ^ 1.654	582.42	Rating applied beyond max stage
Beal Weir Bridge	27003	453000	425500	4.130	1b	Q = 7.753 * (h + 2.353) ^ 1.949	296.22	
Hadfields	27006	439000	391000	1.504	3c	Q = 41.337 * (h + 0.5770) ^ 1.107	93.04	
Westwick Lock	27007	435600	466700	3.328	4b	Q = 198.200 * (h - 0.7370) ^ 1.000	513.54	
Skelton	27009	456800	455400	6.634	11d	Q = 0.0004400 * (h + 9.261) ^ 5.047	508.41	
Doncaster	27021	457000	404000	3.934	4b	Q = 8.527 * (h + 1.386) ^ 1.767	163.49	
Rotherham	27022	442500	392300	1.682	1a	Q = 83.920 * (h + 0.1130) ^ 1.545	207.20	
Barnsley	27023	435000	407250	1.088	2c	Q = 18.616 * (h - 0.003000) ^ 2.545	22.91	
Woodhouse Mill Regulator	27025	443200	385700	1.969	4b	Q = 19.112 * (h - 0.2930) ^ 1.403	39.44	
Ilkley	27027	411600	448100	3.334	1	Q = 57.284 * (h + 0.03900) ^ 1.628	414.61	Rating only goes up to 1975
Armley	27028	428100	434000	5.214	6c	Q = 14.694 * (h + 0.3360) ^ 1.932	402.82	Rating applied beyond max stage
Elland	27029	412400	422000	2.757	2c	Q = 50.291 * (h + 0.000) ^ 2.445	600.29	
Adwick	27030	447700	402000	1.226	2b	Q = 31.462 * (h - 0.09500) ^ 2.065	40.57	
Colne Bridge	27031	417400	419900	1.787	3c	Q = 91.924 * (h - 0.4310) ^ 0.8940	120.69	
Sea Cut at Scarborough	27033	502800	490800	0.844	2c	Q = 33.168 * (h + 0.000) ^ 2.082	23.30	
Kilgram	27034	419000	486000	5.256	3c	Q = 59.497 * (h - 0.5520) ^ 1.140	347.62	Rating applied beyond max stage
Kildwick	27035	401100	445700	4.219	2b	Q = 6.144 * (h + 0.9810) ^ 1.989	163.15	Rating applied beyond max stage
Buttercrambe	27041	473100	458700	2.345	2g	Q = 30.329 * (h + 0.000) ^ 1.643	123.03	Rating applied beyond max stage
Addingham	27043	409200	449400	2.463	2e	Q = 77.926 * (h - 0.2390) ^ 2.001	385.74	Rating applied beyond

								max stage
Birstwith	27053	423000	460300	3.445	3c	Q = 35.350 * (h - 0.2740) ^ 1.206	142.18	Rating applied beyond max stage
Broadway Foot	27055	456000	488300	1.631	2e	Q = 24.000 * (h + 0.04000) ^ 1.300	46.78	
Pickering, Ings Bridge	27056	479100	481900	1.145	2d	Q = 17.561 * (h + 0.02300) ^ 1.726	22.96	Rating applied beyond max stage
Ripon	27059	430100	471000	1.299	2e	Q = 20.446 * (h + 0.05600) ^ 1.726	34.54	Rating applied beyond max stage
Skip Bridge	27062	448200	456100	N/A	Ultrasonic station - flow measured directly	N/A	247.05	
Crakehill	27071	442500	473300	5.455	3f	Q = 109.650 * (h - 3.500) ^ 1.000	214.37	
Snainton Ings	27073	493627	479462	0.193	2c	Q = 15.570 * (h + 0.01100) ^ 1.711	1.03	
Methley	27079	440900	425700	3.684	1a	Q = 59.989 * (h - 0.9600) ^ 1.372	237.23	Rating applied beyond max stage
Oulton Lemonroyd	27080	438100	428200	2.908	1c	Q = 45.931 * (h + 0.05700) ^ 1.615	265.72	
Farrer Lane	27081	436500	428100	0.627	2d	Q = 10.536 * (h - 0.09100) ^ 1.672	3.71	
Huntington	27083	461239	454337	3.985	1a	Q = 5.889 * (h - 0.6540) ^ 1.181	24.39	Rating applied beyond max stage
Cross Hills	27084	402100	445200	1.763	2e	Q = 26.405 * (h - 0.2510) ^ 1.747	54.37	
Alma Weir	27086	431600	470900	1.455	2e	Q = 46.181 * (h - 0.2230) ^ 1.707	65.94	
Low Marishes	27087	483300	477400	4.782	1b	Q = 2.221 * (h - 0.6490) ^ 1.694	24.58	
Mytholmroyd	27088	401200	426100	5.646	2b	Q = 15.083 * (h - 0.4830) ^ 1.432	158.26	
Tadcaster	27089	447700	444100	4.508	1c	Q = 7.469 * (h + 0.4830) ^ 2.808	330.97	
Catterick Bridge	27090	422600	499300	2.964	2c	Q = 36.568 * (h - 0.1000) ^ 2.123	341.39	
Briggswath	27092	487316	508248	2.382	1b	Q = 46.152 * (h - 0.4510) ^ 1.270	106.45	
Ashbrook	68001	367010	363310	1.701	14b	Q = 4.108 * (h + 0.5040) ^ 1.978	19.63	

Rudheath	68003	366760	371800	2.685	8b	Q = 16.952 * (h + 0.001000) ^ 0.9960	45.35	
Lostock Gralam	68007	369700	375720	2.446	10a	Q = 6.784 * (h - 0.1317) ^ 1.344	20.95	
Bridge Trafford	68020	344830	371110	0.685	6b	Q = 10.090 * (h - 0.1600) ^ 1.334	4.27	
Ashton Weir	69007	377240	393560	2.778	1b	Q = 85.045 * (h - 0.6760) ^ 0.8957	165.44	
Wilmslow	69012	384970	381490	0.931	1b	Q = 14.426 * (h - 0.04000) ^ 1.860	11.64	
Compstall	69015	396240	390780	1.142	8b	Q = 76.106 * (h - 0.4456) ^ 0.9325	54.31	
Marple Bridge	69017	396370	389790	1.353	GBc	Q = 104.836 * (h - 1.075) ^ 0.4600	58.18	
London Road	69020	384900	397520	1.382	6a	Q = 20.511 * (h + 0.002000) ^ 1.659	35.17	
Blackford Bridge	69023	380690	407740	3.362	3c	Q = 15.913 * (h + 1.144) ^ 1.671	196.90	
Farnworth	69024	374340	406820	1.413	5a	Q = 76.932 * (h - 0.02268) ^ 1.570	129.06	Rating applied beyond max stage
Manchester Racecourse	69025	382070	400350	5.668	N/Ac	Q = 32.533 * (h - 0.2830) ^ 1.823	700.29	Rating applied beyond max stage
Portwood	69027	390700	391870	1.611	14b	Q = 41.871 * (h + 0.002000) ^ 1.466	84.39	
Causey Bridge	69030	358760	392240	2.973	14e	Q = 2.739 * (h + 1.021) ^ 1.868	36.39	
Kirkby	69032	339160	398330	1.817	2c	Q = 9.189 * (h + 0.02200) ^ 1.498	22.89	
Broomstairs Bridge	69041	393750	395320	1.739	7a	Q = 31.896 * (h - 0.05200) ^ 1.712	78.08	
Collyhurst Weir	69043	384850	399690	1.146	9b	Q = 28.923 * (h + 0.3790) ^ 1.003	44.16	
Bury Ground	69044	379900	411400	2.178	2b	Q = 128.057 * (h -0.5237) ^ 1.580	283.67	
Rochdale ETW	69803	388200	412700	2.222	2a	Q = 26.198 * (h - 0.01800) ^ 1.601	92.84	
Wanes Blades	70002	347620	412570	N/A	Ultrasonic station - flow measured directly	N/A	84.79	
Croston	70004	349840	417980	2.994	5d	Q = 41.625 * (h + 0.000) ^ 0.3890	63.77	
Littlewood Bridge	70005	349740	419650	3.795	15a	Q = 5.796 * (h - 0.06000) ^ 1.512	42.50	Rating applied beyond max stage
Samlesbury	71001	358920	430490	6.558	7c	Q = 52.088 * (h - 0.6640) ^ 1.692	1047.78	

Whalley Weir	71004	372900	436030	4.642	4e	Q = 1.102 * (h + 0.000) ^ 4.026	532.52	
Henthorn	71006	372180	439170	3.311	5d	Q = 30.158 * (h + 0.1430) ^ 1.956	340.69	
Hodder Place	71008	370410	439980	2.321	1d	Q = 63.159 * (h - 0.2210) ^ 2.107	301.54	Rating applied beyond max stage
New Jumbles Rock	71009	370250	437590	5.044	2c	Q = 0.06240 * (h + 5.431) ^ 4.177	1138.59	Rating applied beyond max stage
Barden Lane	71010	383740	435050	3.362	10e	Q = 179.340 * (h - 2.240) ^ 0.8000	196.64	
Arnford	71011	383880	455580	2.255	4c	Q = 10.256 * (h + 0.5100) ^ 2.634	149.42	
Ewood	71013	367730	426250	2.114	2e	Q = 27.613 * (h + 0.000) ^ 1.044	60.33	
Blue Bridge	71014	356470	427780	3.305	6c	Q = 20.100 * (h + 0.000) ^ 2.000	219.55	
Wray	72003	360490	467950	2.938	4g	Q = 34.412 * (h - 0.5790) ^ 2.142	216.32	
Caton	72004	352860	465290	16.245	3f	Q = 958.340 * (h - 5.130) ^ 0.5400	3518.12	Rating applied beyond max stage
Killington	72005	362200	490660	4.027	8e	Q = 17.408 * (h + 1.281) ^ 2.147	626.87	Rating applied beyond max stage
A6 Bridge	72007	351210	440550	1.371	6b	Q = 27.410 * (h - 0.1980) ^ 2.178	38.80	Rating applied beyond max stage
Wennington	72009	361540	470080	3.071	4d	Q = 7.620 * (h + 0.8210) ^ 2.350	185.72	
Brigflats	72011	363990	491090	3.909	10b	Q = 9.684 * (h + 0.04000) ^ 2.963	566.82	Rating applied beyond max stage
Galgate	72014	348160	455370	2.532	3b	Q = 6.612 * (h + 0.06100) ^ 1.692	33.15	Rating applied beyond max stage
Lunes Bridge	72015	361210	502900	5.120	3c	Q = 95.238 * (h - 1.323) ^ 1.000	361.62	Rating applied beyond max stage
Scorton	72016	350120	449990	1.473	10e	Q = 44.795 * (h + 0.000) ^ 3.366	164.97	
Hornby	72807	358570	468390	2.452	3d	Q = 105.800 * (h + 0.1720) ^ 1.184	331.54	
Low Nibthwaite	73002	329450	488210	1.410	5b	Q = 21.330 * (h - 0.05200) ^ 1.742	36.35	

Sedgwick	73005	350883	487419	4.276	3d	Q = 49.088 * (h + 0.5895) ^ 1.500	526.83	
Eel House Bridge	73006	336969	494044	1.806	12a	Q = 9.215 * (h - 0.4130) ^ 2.126	18.64	Rating applied beyond max stage
Beetham Weir	73008	349620	480590	2.114	1c	Q = 2.073 * (h + 1.000) ^ 3.638	129.21	
Sprint Mill, Kendal	73009	351477	496104	1.920	5b	Q = 38.561 * (h - 0.1400) ^ 1.839	111.34	Rating applied beyond max stage
Newby Bridge	73010	336600	486264	2.480	2c	Q = 37.578 * (h - 0.2750) ^ 2.260	224.41	Rating applied beyond max stage
Mint Bridge	73011	352411	494470	2.688	2b	Q = 21.274 * (h - 0.1100) ^ 2.193	169.74	Rating applied beyond max stage
Victoria Bridge	73012	351789	493070	3.874	6c	Q = 73.820 * (h - 1.185) ^ 1.709	400.26	
Jeffy Knotts	73014	335965	503406	3.871	15b	Q = 1.454 * (h + 0.2900) ^ 3.314	163.90	Rating applied beyond max stage
High River Keer	73015	352320	471890	2.587	2d	Q = 11.187 * (h - 0.5986) ^ 1.148	24.63	Rating applied beyond max stage
Duddon Hall	74001	319560	489570	1.799	5b	Q = 37.044 * (h - 0.08300) ^ 2.134	117.27	Rating applied beyond max stage
Galesyke	74002	313537	503818	1.319	15c	Q = 26.278 * (h - 0.5270) ^ 0.7930	21.84	
Bleach Green Weir	74003	308390	515400	1.641	1b	Q = 18.925 * (h - 0.09000) ^ 1.959	44.71	Rating applied beyond max stage
Braystones	74005	300909	506051	1.910	28N/A (first in the list)	Q = 0.4347 * (h + 2.427) ^ 3.397	63.49	
Calder Hall	74006	303490	504490	1.343	6b	Q = 29.409 * (h - 0.03500) ^ 2.267	54.05	Rating applied beyond max stage
Cropple Howe	74007	313100	497770	2.203	14c	Q = 56.844 * (h - 0.5350) ^ 1.281	109.48	Rating applied beyond max stage
Ulpha	74008	320910	494720	1.576	5c	Q = 47.542 * (h - 0.4150) ^ 1.545	59.87	

Thirlmere	75001	331300	519500	3.262	1d	Q = 18.000 * (h + 0.1500) ^ 1.170	75.66	Rating applied beyond max stage
Ouse Bridge	75003	319823	532151	3.891	12d	Q = 75.350 * (h - 0.9770) ^ 1.550	395.40	
Southwaite Bridge	75004	313090	528090	2.700	20c	Q = 2.539 * (h + 1.790) ^ 2.748	157.41	
Portinscale	75005	325195	523885	3.730	14v1top	Q = 11.542 * (h + 0.000) ^ 2.446	288.86	
Threlkeld	75007	332254	524801	2.566	18b	Q = 22.364 * (h - 0.3840) ^ 1.620	79.16	Rating applied beyond max stage
Low Briery	75009	328558	524216	3.643	5c	Q = 16.930 * (h + 0.5360) ^ 2.326	471.27	Rating applied beyond max stage
Bull Gill	75017	309600	538400	2.510	2c	Q = 23.266 * (h - 0.2090) ^ 1.079	57.18	Rating applied beyond max stage
Burnbanks	76001	350750	515920	0.931	4b	Q = 39.228 * (h - 0.09294) ^ 1.787	28.61	
Udford	76003	357570	530450	2.872	84b	Q = 31.887 * (h + 0.2634) ^ 2.228	406.78	Rating applied beyond max stage
Eamont Bridge, Beehive, River Lowther	76004	352508	528562	2.934	18c	Q = 1.289 * (h + 0.5860) ^ 4.250	271.06	
Temple Sowerby	76005	360452	528312	4.545	26c	Q = 1.222 * (h - 0.1741) ^ 4.654	1170.28	Rating applied beyond max stage
Sheepmount	76007	339000	557100	7.806	11b	Q = 56.612 * (h - 0.2980) ^ 1.699	1739.50	Rating applied beyond max stage
Greenholme	76008	348618	558072	3.472	10f	Q = 14.519 * (h - 0.2990) ^ 2.388	228.80	
Coalburn Beck	76011	369370	577780	0.930	3b	Q = 8.152 * (h - 0.4630) ^ 1.636	2.35	
Kirkby Stephen	76014	377299	509694	2.731	11b	Q = 38.063 * (h - 0.3070) ^ 1.469	139.76	Rating applied beyond max stage
Pooley Bridge	76015	347236	524959	2.429	3c	Q = 108.580 * (h - 0.6980) ^ 1.291	220.49	
Great Corby	76017	346810	555360	5.829	3e	Q = 1027.080 * (h - 3.932) ^ 0.5771	1486.20	

Great Musgrave Bridge	76806	376500	513120	3.201	3d	Q = 0.3523 * (h + 1.603) ^ 4.439	373.72	Rating applied beyond max stage
Cummersdale	76809	339480	552727	3.238	3f	Q = 4.524 * (h + 0.000) ^ 3.509	279.31	
Dacre Bridge	76811	346007	526287	1.743	3b	Q = 25.014 * (h - 0.1537) ^ 1.823	58.21	

TABLE A1. Maximum observed stage levels, ratings curves applied and estimated peak December 2015 flows at the 155 stations across northern England that were included in the present study. For a station to be included, previous AM data had to be available in addition to December 2015 peaks.

Appendix B: Flood frequency estimates (single-site) across northern England with and without December 2015 data

Station	Area (km²)	With December 2015 peaks?	QMED (empirical) (m^3s^{-1})	Return period (1-in-n-years) flow ($m^3 s^{-1}$)						
				2	5	10	25	50	100	500
21806	655.53	F	151.60	129.85	207.96	271.97	374.43	471.34	590.40	986.83
		T	142.24	127.15	200.11	260.63	358.48	451.91	567.60	958.10
22001	578.21	F	147.64	146.51	212.91	266.94	352.91	433.77	532.65	859.20
		T	149.91	148.42	215.09	268.66	353.00	431.55	526.80	836.94
22003	21.87	F	19.22	18.83	28.65	36.62	49.28	61.17	75.68	123.53
		T	19.18	18.49	28.08	35.97	48.64	60.66	75.46	124.98
22006	273.62	F	52.55	56.21	87.33	112.21	151.25	187.48	231.28	373.17
		T	52.68	56.18	86.95	111.56	150.16	185.97	229.27	369.52
22007	282.03	F	100.86	99.99	152.25	194.52	261.41	324.02	400.26	650.30
		T	101.30	100.19	151.85	193.52	259.34	320.83	395.57	640.02
22009	345.99	F	133.00	131.51	189.24	234.25	303.31	366.11	440.76	675.60
		T	135.10	133.27	190.67	234.75	301.56	361.61	432.29	650.97
22801	48.11	F	14.06	14.87	25.97	34.89	48.95	62.05	77.95	129.75
		T	15.32	15.78	26.62	34.93	47.52	58.83	72.14	113.25
23001	2172.36	F	876.43	880.10	1063.29	1185.02	1348.50	1479.81	1620.34	1990.87
		T	880.58	884.41	1080.85	1215.77	1401.99	1555.43	1723.24	2181.89
23002	118.07	F	48.41	42.98	53.45	59.80	67.69	73.59	79.54	93.72
		T	41.22	41.47	52.91	59.42	67.13	72.62	77.94	89.83
23003	1012.97	F	404.00	404.00	511.43	584.52	684.61	766.49	855.47	1096.17
		T	411.17	435.10	555.85	638.20	751.21	843.83	944.65	1218.10
23004	749.90	F	453.51	458.47	550.68	616.65	710.85	790.98	880.97	1138.26
		T	458.64	460.22	561.24	637.04	749.69	849.14	964.39	1311.81

23006	322.97	F	242.68	245.83	293.04	326.71	374.65	415.33	460.92	590.78
		T	243.67	247.39	298.22	335.54	390.02	437.31	491.34	650.35
23007	243.84	F	41.98	44.69	68.29	88.61	122.55	155.92	198.30	347.70
		T	42.34	45.13	68.64	88.69	121.87	154.24	195.06	337.28
23008	345.20	F	138.04	138.08	173.62	198.56	233.60	262.95	295.51	386.60
		T	138.63	139.45	177.41	204.63	243.58	276.78	314.13	421.19
23009	118.62	F	139.19	139.35	180.80	212.55	260.55	303.62	354.21	510.28
		T	142.00	141.61	184.49	217.05	265.91	309.46	360.32	515.70
23011	58.81	F	65.78	65.87	81.46	91.17	103.55	113.02	122.74	146.67
		T	66.05	65.95	81.30	90.84	102.98	112.24	121.74	145.07
23017	62.35	F	13.22	12.82	17.95	21.99	28.25	33.98	40.84	62.65
		T	13.16	12.62	17.70	21.74	28.06	33.91	40.95	63.65
23018	10.48	F	2.56	2.62	3.91	5.02	6.86	8.67	10.96	18.98
		T	2.67	2.62	3.89	4.97	6.77	8.53	10.75	18.53
23033	180.63	F	153.15	156.50	181.62	194.51	208.53	217.74	226.09	242.84
		T	159.41	159.89	183.16	194.53	206.41	213.93	220.54	233.14
24001	661.04	F	186.15	199.04	258.66	304.39	373.61	435.77	508.87	734.68
		T	189.27	201.46	262.58	309.04	378.84	441.10	513.87	736.45
24003	173.41	F	124.54	123.17	151.56	171.33	198.92	221.90	247.25	317.57
		T	125.27	123.96	154.06	175.52	206.07	231.98	261.00	343.66
24004	74.13	F	25.21	25.09	37.10	47.13	63.43	79.08	98.55	164.76
		T	25.30	25.22	37.14	47.03	63.03	78.32	97.25	161.22
24005	178.95	F	37.85	37.66	51.90	62.81	79.32	94.14	111.55	165.31
		T	37.82	37.45	51.55	62.42	78.94	93.85	111.43	166.07
24006	36.62	F	24.62	24.21	29.79	33.51	38.50	42.51	46.81	58.15
		T	24.52	23.93	29.42	33.14	38.24	42.40	46.91	59.11
24008	455.10	F	212.95	209.41	255.77	285.74	325.09	356.02	388.54	471.76

		T	222.54	211.87	259.65	290.51	330.98	362.78	396.19	481.58
24009	1005.00	F	248.07	254.10	309.27	346.43	396.92	437.90	482.15	600.59
		T	254.42	256.80	311.61	347.78	396.08	434.68	475.80	583.41
25001	847.70	F	387.40	393.40	503.98	577.10	674.88	753.10	836.55	1055.35
		T	388.89	388.89	498.72	570.89	666.92	743.38	824.62	1036.23
25003	11.46	F	15.16	16.03	20.57	24.20	29.92	35.24	41.68	62.61
		T	15.14	15.89	20.42	24.05	29.77	35.08	41.52	62.47
25004	224.58	F	22.53	22.03	29.83	35.07	42.19	47.97	54.21	70.89
		T	21.93	21.99	29.69	34.90	41.97	47.73	53.95	70.62
25006	86.81	F	76.76	74.26	96.08	111.65	133.85	152.69	173.83	234.12
		T	76.76	74.19	96.29	111.57	132.77	150.33	169.62	222.70
25008	510.17	F	262.48	262.04	333.34	382.12	449.27	504.45	564.65	728.55
		T	263.64	265.20	337.17	385.82	452.08	506.00	564.36	721.08
25009	1267.10	F	410.18	398.47	493.04	545.95	607.66	651.05	692.60	783.78
		T	413.60	403.49	497.95	550.17	610.52	652.57	692.53	779.17
25012	24.58	F	33.27	33.85	44.18	52.12	64.13	74.94	87.65	126.98
		T	33.39	34.29	44.79	52.74	64.65	75.25	87.61	125.27
25018	242.36	F	214.93	210.48	270.92	311.37	366.02	410.15	457.61	583.70
		T	221.61	213.32	274.14	314.31	367.99	410.91	456.65	576.48
25019	15.07	F	5.54	5.61	9.07	12.17	17.52	22.95	30.01	56.05
		T	4.99	4.99	8.11	10.91	15.74	20.65	27.04	50.64
25020	152.71	F	15.21	14.91	18.19	20.18	22.66	24.51	26.37	30.82
		T	15.20	14.81	18.08	20.09	22.60	24.50	26.42	31.07
25021	75.43	F	5.81	6.09	8.45	10.39	13.50	16.46	20.11	32.32
		T	5.91	6.11	8.43	10.33	13.37	16.26	19.80	31.59
26003	59.40	F	1.74	1.80	2.39	2.73	3.14	3.44	3.72	4.37
		T	1.72	1.78	2.37	2.72	3.14	3.45	3.74	4.43

26010	49.47	F	2.00	1.96	2.55	3.00	3.67	4.26	4.95	7.02
		T	1.95	1.94	2.52	2.96	3.63	4.23	4.92	7.06
27001	493.89	F	120.19	120.19	159.74	188.69	230.83	267.33	308.94	430.98
		T	120.50	120.90	162.65	193.96	240.47	281.52	329.07	472.31
27002	759.03	F	236.61	242.83	299.86	340.94	399.92	450.36	507.25	671.13
		T	237.21	243.22	305.14	351.94	421.92	484.06	556.42	776.36
27003	1936.50	F	276.23	280.11	290.29	293.87	296.77	298.19	299.21	300.62
		T	276.23	280.72	290.90	294.49	297.39	298.82	299.85	301.27
27006	364.99	F	86.74	84.19	116.01	140.21	176.55	208.95	246.84	362.69
		T	87.11	84.40	115.89	139.74	175.46	207.22	244.26	357.09
27007	912.58	F	284.57	278.10	362.87	426.50	521.09	604.62	701.45	993.46
		T	284.67	281.06	368.62	434.38	532.20	618.63	718.86	1021.34
27009	3300.80	F	322.00	320.15	386.75	431.06	490.64	538.53	589.82	725.25
		T	322.00	321.13	388.87	434.14	495.21	544.48	597.39	737.76
27021	1252.88	F	165.55	159.91	205.01	234.25	272.73	303.05	334.98	416.93
		T	163.53	160.02	204.58	233.42	271.31	301.14	332.51	412.89
27022	824.54	F	121.19	134.63	201.36	254.01	335.66	410.62	500.45	787.00
		T	144.73	145.34	211.28	258.86	327.30	385.90	452.11	643.19
27023	119.53	F	28.79	26.38	40.32	53.03	75.31	98.22	128.45	242.28
		T	28.79	26.26	40.03	52.61	74.70	97.47	127.54	241.09
27025	351.10	F	47.75	47.14	61.07	72.09	89.21	104.95	123.83	184.23
		T	47.07	46.88	60.67	71.64	88.75	104.55	123.57	184.77
27027	443.00	F	267.21	258.38	309.68	350.51	414.26	473.19	544.19	772.84
		T	271.02	267.03	326.27	372.60	443.90	508.92	586.37	830.98
27028	687.05	F	140.78	144.01	173.03	193.28	221.58	245.18	271.25	343.71
		T	142.69	143.38	178.41	206.30	249.85	290.12	338.64	494.98
27029	340.75	F	133.46	135.87	197.81	244.92	315.74	378.93	452.84	679.03

		T	135.76	135.99	205.99	263.56	356.01	443.69	551.67	912.71
27030	310.96	F	44.21	41.85	63.85	82.96	115.12	146.97	187.65	332.55
		T	44.05	41.75	63.42	82.28	114.04	145.52	185.78	329.32
27031	244.77	F	94.34	94.45	119.63	134.73	153.38	167.22	181.07	213.79
		T	96.14	95.23	120.20	135.04	153.22	166.60	179.93	211.07
27033	33.00	F	33.93	34.01	44.15	50.50	58.60	64.80	71.19	86.94
		T	33.31	33.65	43.82	50.25	58.54	64.95	71.58	88.16
27034	510.90	F	243.41	245.92	293.43	323.70	362.97	393.49	425.28	505.36
		T	248.18	248.04	296.61	327.49	367.49	398.53	430.82	511.99
27035	283.47	F	67.60	68.89	86.15	100.42	123.45	145.39	172.51	263.79
		T	67.95	69.47	88.49	104.81	131.98	158.62	192.37	310.83
27041	1594.22	F	71.52	69.93	84.02	95.58	114.10	131.63	153.18	225.05
		T	71.58	70.58	85.72	98.31	118.71	138.22	162.42	244.35
27043	429.98	F	262.27	266.45	330.94	371.22	422.64	461.99	502.46	602.25
		T	265.87	270.20	335.24	375.30	425.86	464.15	503.17	597.98
27053	219.28	F	219.28	95.76	117.88	129.88	143.51	152.86	161.64	180.27
		T	98.05	96.92	119.73	132.17	146.41	156.23	165.49	185.30
27055	131.45	F	40.73	36.56	58.87	83.70	135.48	197.85	292.04	742.20
		T	40.02	36.86	58.95	83.37	133.99	194.63	285.78	717.76
27056	67.62	F	14.95	14.68	24.32	31.62	42.54	52.24	63.54	97.93
		T	14.98	15.00	24.59	31.73	42.27	51.52	62.18	94.03
27059	78.28	F	21.88	22.13	30.91	38.50	51.25	63.83	79.86	136.69
		T	22.01	22.54	31.45	39.00	51.41	63.44	78.55	130.73
27062	523.47	F	87.39	89.11	124.49	154.76	205.01	254.16	316.30	533.61
		T	91.32	94.61	139.18	178.39	245.08	311.76	397.66	707.98
27071	1354.42	F	161.70	163.42	188.17	204.80	227.32	245.56	265.21	317.63
		T	162.02	164.97	190.39	207.33	230.11	248.44	268.08	320.00

27073	8.06	F	0.81	0.84	1.06	1.19	1.34	1.45	1.55	1.79
		T	0.82	0.85	1.07	1.19	1.33	1.43	1.53	1.75
27079	950.07	F	232.35	234.92	294.85	334.09	386.16	427.50	471.32	585.01
		T	233.70	234.86	293.26	331.53	382.35	422.73	465.55	576.78
27080	861.58	F	154.80	156.46	190.01	213.17	245.27	271.83	300.96	381.01
		T	156.02	158.70	195.51	221.59	258.54	289.73	324.54	422.99
27081	25.10	F	2.32	2.40	3.45	4.31	5.67	6.94	8.49	13.59
		T	2.32	2.46	3.52	4.36	5.68	6.88	8.34	13.00
27083	126.35	F	11.26	11.66	13.85	15.11	16.61	17.70	18.76	21.16
		T	11.30	11.65	14.58	16.61	19.42	21.74	24.30	31.33
27084	41.01	F	30.85	31.84	40.48	46.16	53.70	59.70	66.07	82.64
		T	31.79	32.57	41.81	47.91	56.06	62.58	69.52	87.72
27086	117.35	F	27.50	27.36	39.88	51.87	73.80	97.24	129.19	256.61
		T	27.56	28.25	41.84	54.67	77.87	102.39	135.51	265.49
27087	475.92	F	14.70	14.12	17.45	20.04	23.99	27.57	31.82	45.09
		T	14.77	14.38	18.04	20.92	25.37	29.44	34.31	49.76
27088	172.96	F	89.60	90.42	123.52	149.45	189.42	225.92	269.46	407.22
		T	91.50	92.88	127.46	154.19	194.92	231.73	275.22	410.79
27089	815.36	F	215.44	226.17	284.67	319.50	362.25	393.78	425.22	498.88
		T	226.72	232.43	291.37	325.62	366.80	396.62	425.86	492.64
27090	497.61	F	323.74	319.41	390.17	433.87	489.14	531.09	573.91	678.25
		T	325.59	321.36	389.34	430.80	482.68	521.67	561.13	656.00
27092	325.25	F	152.23	154.91	188.70	206.79	227.15	240.99	253.87	280.87
		T	151.44	151.43	186.21	205.44	227.65	243.12	257.82	289.67
68001	621.52	F	48.63	48.10	63.50	74.57	90.47	104.04	119.35	163.37
		T	47.72	47.83	63.43	74.51	90.24	103.55	118.43	160.66
68003	412.43	F	53.98	53.64	64.13	71.93	83.41	93.47	105.06	139.59

		T	53.71	53.28	63.70	71.47	82.96	93.07	104.73	139.66
68007	148.28	F	20.33	20.57	23.73	25.39	27.23	28.46	29.58	31.90
		T	20.60	20.59	23.71	25.34	27.15	28.36	29.46	31.73
68020	148.70	F	15.79	16.04	18.83	20.26	21.81	22.83	23.76	25.62
		T	15.66	15.88	18.80	20.24	21.75	22.71	23.56	25.18
69007	667.27	F	156.51	163.81	200.79	221.66	246.17	263.51	280.20	317.17
		T	160.97	163.87	200.47	221.10	245.34	262.48	278.99	315.51
69012	68.38	F	16.50	16.24	22.29	27.11	34.66	41.65	50.08	77.31
		T	16.01	16.07	22.07	26.89	34.48	41.55	50.13	78.07
69015	149.44	F	51.83	54.24	70.81	81.40	95.16	105.88	117.07	145.34
		T	52.12	54.22	70.56	81.01	94.60	105.20	116.26	144.22
69017	184.23	F	62.16	63.51	79.64	89.63	102.31	111.95	121.82	145.93
		T	61.98	63.27	79.23	89.18	101.87	111.56	121.52	146.04
69020	57.19	F	23.31	22.44	29.84	35.49	44.03	51.68	60.66	88.31
		T	23.34	22.74	30.25	35.93	44.42	51.96	60.74	87.42
69023	190.45	F	71.92	71.31	87.01	97.16	110.49	120.96	131.97	160.14
		T	72.38	71.03	89.48	102.75	121.79	138.05	156.39	209.13
69024	141.77	F	65.30	65.67	84.43	97.63	116.21	131.81	149.13	197.72
		T	65.35	66.14	85.76	99.78	119.77	136.75	155.81	210.21
69025	550.62	F	272.00	265.65	333.97	379.24	439.90	488.51	540.44	676.94
		T	273.50	265.55	341.78	395.86	472.58	537.41	609.84	815.05
69027	146.04	F	61.58	60.26	74.97	84.20	96.01	105.08	114.44	137.60
		T	62.21	60.70	75.48	84.71	96.47	105.46	114.70	137.46
69030	147.74	F	27.83	28.07	33.32	36.16	39.38	41.58	43.63	47.99
		T	28.60	28.46	33.70	36.49	39.62	41.74	43.70	47.80
69032	96.57	F	17.60	17.89	21.89	24.76	28.90	32.43	36.41	47.86
		T	17.75	18.06	22.07	24.91	28.95	32.35	36.16	46.94

69041	115.97	F	47.70	48.71	60.16	66.96	75.28	81.40	87.50	101.72
		T	48.78	49.13	61.10	68.32	77.27	83.95	90.65	106.59
69043	72.30	F	28.56	28.56	33.70	36.70	40.34	42.98	45.59	51.59
		T	29.40	30.75	37.16	41.59	47.74	52.82	58.41	73.78
69044	141.12	F	111.80	111.80	136.90	151.92	170.44	184.16	197.88	230.22
		T	113.57	107.99	139.30	161.33	192.36	218.40	247.34	328.55
69803	110.53	F	43.52	43.84	51.35	55.77	61.14	65.06	68.94	77.91
		T	44.08	43.67	54.59	62.77	74.90	85.60	97.98	135.22
70002	190.24	F	45.60	46.81	52.49	55.58	59.11	61.54	63.82	68.69
		T	46.40	46.56	55.65	62.31	72.03	80.45	90.08	118.37
70004	81.42	F	33.51	33.74	44.24	52.13	63.89	74.29	86.36	122.85
		T	33.66	34.22	45.26	53.62	66.14	77.28	90.26	129.81
70005	54.50	F	23.34	23.36	28.69	32.41	37.61	41.95	46.75	60.08
		T	23.41	23.59	29.36	33.51	39.44	44.50	50.19	66.52
71001	1133.93	F	607.20	605.57	727.61	805.95	908.22	988.19	1071.87	1284.43
		T	610.46	608.94	739.76	826.45	942.61	1035.71	1135.18	1396.63
71004	317.29	F	174.45	178.71	225.11	258.20	305.31	345.29	390.09	517.70
		T	175.93	178.38	234.04	278.85	349.51	415.42	495.44	756.58
71006	446.28	F	224.15	228.74	280.27	316.03	365.81	407.15	452.66	578.46
		T	227.76	230.93	283.78	320.37	371.20	413.33	459.65	587.32
71008	258.14	F	225.57	216.67	268.83	306.33	360.09	405.98	457.70	606.32
		T	226.62	218.38	271.25	309.13	363.29	409.40	461.24	609.65
71009	1048.04	F	533.96	538.77	674.08	772.74	915.86	1039.40	1179.89	1590.03
		T	538.06	543.22	691.92	804.91	974.63	1125.93	1302.77	1843.39
71010	110.61	F	81.55	81.02	107.82	126.86	153.90	176.77	202.35	274.93
		T	82.93	81.92	111.49	133.46	165.85	194.22	226.89	324.28
71011	203.22	F	121.13	120.52	130.75	136.29	142.58	146.89	150.93	159.49

		T	121.32	120.90	131.81	137.89	144.95	149.89	154.61	164.92
71013	39.08	F	28.12	27.88	34.78	40.31	49.02	57.13	66.96	98.92
		T	28.25	28.17	35.76	42.07	52.27	62.02	74.10	114.93
71014	136.21	F	87.33	87.78	115.78	136.33	166.32	192.34	222.06	309.49
		T	87.86	88.72	119.99	144.16	180.98	214.24	253.55	376.01
72003	82.77	F	134.57	134.57	179.52	228.90	330.68	452.00	633.55	1487.27
		T	134.71	131.74	177.66	226.73	325.45	440.56	609.61	1378.39
72004	985.37	F	727.30	716.93	897.54	1016.14	1173.88	1299.40	1432.72	1779.72
		T	727.73	698.85	956.14	1179.61	1555.45	1927.51	2402.65	4093.78
72005	219.21	F	262.43	252.45	322.87	367.42	424.88	469.31	515.37	630.53
		T	264.22	253.04	334.55	390.79	468.72	533.16	603.85	798.05
72007	31.53	F	31.41	31.93	41.42	48.75	59.93	70.03	81.97	119.18
		T	31.96	32.20	41.62	48.80	59.60	69.26	80.57	115.26
72009	139.37	F	108.89	112.53	137.73	154.10	175.68	192.71	210.66	256.86
		T	109.26	114.30	141.78	160.10	184.77	204.64	225.95	282.34
72011	194.15	F	287.18	293.93	368.32	417.85	484.47	538.06	595.48	747.19
		T	287.58	296.44	376.79	432.08	508.53	571.62	640.71	829.95
72014	28.99	F	17.70	17.69	22.68	25.79	29.75	32.79	35.90	43.57
		T	17.89	17.87	23.17	26.55	30.95	34.37	37.94	46.96
72015	140.83	F	201.71	196.00	232.36	254.19	281.14	301.15	321.19	368.56
		T	202.10	196.74	239.49	267.59	305.00	334.79	366.45	448.96
72016	88.00	F	90.80	89.98	106.67	119.04	137.25	153.18	171.50	226.01
		T	92.74	91.53	112.47	129.50	156.60	182.08	213.23	316.10
72807	230.70	F	216.09	213.31	264.40	297.69	341.70	376.52	413.31	508.31
		T	217.66	214.80	267.54	302.28	348.59	385.53	424.84	527.50
73002	72.90	F	19.56	20.15	25.88	29.86	35.41	40.02	45.10	59.15
		T	19.65	20.34	26.33	30.54	36.48	41.47	47.01	62.56

73005	212.19	F	155.85	160.55	212.84	252.21	310.88	362.81	423.14	605.66
		T	160.75	160.91	221.48	271.42	351.80	428.20	522.45	838.51
73006	18.77	F	7.87	7.73	10.16	11.93	14.51	16.73	19.26	26.64
		T	7.90	7.81	10.50	12.56	15.68	18.47	21.76	31.87
73008	127.45	F	36.94	36.89	47.56	54.98	65.32	73.92	83.40	109.64
		T	36.97	36.66	49.82	60.41	77.12	92.70	111.61	173.31
73009	34.80	F	43.07	44.12	57.91	67.83	82.08	94.26	108.01	147.58
		T	44.13	44.47	59.81	71.47	88.98	104.59	122.83	178.61
73010	247.81	F	72.43	71.32	92.36	109.30	136.04	160.99	191.30	290.28
		T	72.56	71.47	94.45	113.97	146.18	177.51	216.92	353.62
73011	65.59	F	54.52	56.93	75.10	88.98	109.91	128.65	150.63	218.19
		T	54.68	57.25	77.93	94.91	122.14	147.93	179.67	285.61
73012	183.23	F	146.43	148.53	188.76	219.43	265.60	306.86	355.17	503.34
		T	148.29	149.42	196.81	236.40	300.84	362.71	439.72	701.88
73014	56.59	F	86.15	87.50	116.38	140.48	179.67	217.27	264.04	423.03
		T	86.57	88.87	118.94	143.83	184.03	222.36	269.77	429.49
73015	30.06	F	12.24	11.60	14.11	15.58	17.36	18.66	19.93	22.87
		T	12.29	11.63	14.96	17.26	20.46	23.11	26.03	34.07
74001	86.01	F	119.58	119.31	148.74	171.39	205.80	236.80	273.34	386.72
		T	119.14	119.15	148.17	170.56	204.61	235.34	271.61	384.38
74002	43.99	F	21.18	20.88	25.69	28.75	32.73	35.82	39.04	47.14
		T	21.44	20.92	25.65	28.65	32.54	35.55	38.68	46.52
74003	44.58	F	34.07	33.44	43.88	51.60	62.92	72.80	84.14	117.76
		T	34.24	33.79	44.21	51.81	62.85	72.40	83.26	115.01
74005	129.49	F	70.71	71.22	83.32	91.04	101.09	108.92	117.08	137.72
		T	70.14	70.86	82.88	90.65	100.85	108.87	117.31	138.90
74006	43.93	F	55.47	56.06	80.26	100.54	133.64	165.49	205.22	340.93

		T	55.09	55.87	79.60	99.57	132.24	163.77	203.20	338.42
74007	70.11	F	102.96	103.16	110.90	114.62	118.47	120.88	122.97	126.90
		T	103.46	103.43	111.02	114.63	118.34	120.64	122.63	126.34
74008	48.05	F	68.02	69.10	80.85	88.57	98.83	107.00	115.67	138.25
		T	67.94	68.68	80.39	88.18	98.65	107.08	116.11	139.95
75001	41.88	F	20.92	18.81	30.58	38.50	49.25	57.97	67.39	92.56
		T	21.65	19.11	32.69	42.59	56.93	69.30	83.35	124.33
75003	363.01	F	97.96	100.01	135.25	163.28	207.04	247.46	296.15	452.86
		T	100.64	100.09	141.26	177.47	239.04	300.60	379.90	666.30
75004	116.17	F	51.27	53.96	77.20	96.19	126.54	155.20	190.35	307.08
		T	52.15	54.75	79.82	100.84	135.14	168.17	209.37	350.16
75005	237.26	F	105.84	100.83	134.28	162.93	210.55	257.18	316.18	522.84
		T	106.14	106.14	144.87	179.72	240.17	301.71	382.21	680.80
75007	64.57	F	61.43	62.32	68.83	72.68	77.35	80.78	84.18	92.05
		T	62.39	62.64	69.49	73.60	78.66	82.40	86.15	95.00
75009	146.97	F	113.76	106.53	142.98	168.87	205.63	236.73	271.51	370.18
		T	114.17	105.30	151.47	190.49	254.61	316.73	394.64	663.34
75017	102.40	F	36.70	36.79	42.44	45.56	49.18	51.70	54.09	59.29
		T	36.74	36.96	43.32	47.05	51.59	54.90	58.17	65.74
76001	32.34	F	18.22	16.75	28.96	37.45	49.30	59.16	70.03	100.15
		T	18.85	17.29	29.37	37.60	48.90	58.15	68.22	95.53
76003	407.17	F	195.71	196.91	255.10	295.33	351.14	397.37	448.13	587.87
		T	198.32	198.74	260.65	304.47	366.48	418.79	477.13	641.95
76004	156.20	F	110.23	116.58	161.61	193.33	238.04	275.62	317.40	434.82
		T	110.39	118.20	165.83	199.94	248.72	290.25	336.94	470.62
76005	618.21	F	255.55	269.50	366.51	449.10	585.68	718.79	886.54	1470.09
		T	257.26	270.13	382.09	486.36	672.52	867.21	1127.74	2133.42

76007	2276.03	F	615.49	620.83	802.50	933.43	1121.56	1282.50	1464.16	1987.73
		T	615.62	622.61	830.12	992.07	1240.84	1467.30	1736.68	2585.30
76008	333.43	F	139.13	151.44	183.98	207.06	239.79	267.44	298.32	385.78
		T	142.24	153.18	186.56	209.97	242.86	270.40	300.94	386.36
76011	1.63	F	1.84	1.85	2.36	2.78	3.46	4.13	4.95	7.78
		T	1.85	1.87	2.37	2.79	3.45	4.09	4.87	7.48
76014	66.84	F	84.46	87.87	107.99	119.42	132.91	142.50	151.77	172.43
		T	85.29	88.82	110.01	122.26	136.94	147.52	157.87	181.39
76015	149.24	F	60.17	57.82	75.61	88.85	108.39	125.53	145.30	204.37
		T	60.66	57.51	79.74	99.10	131.79	164.23	205.77	354.23
76017	1371.70	F	575.48	548.45	744.89	892.98	1113.93	1309.70	1537.30	2226.99
		T	580.43	564.97	811.28	1015.60	1346.11	1661.71	2052.63	3372.57
76806	223.10	F	200.34	201.52	245.33	272.96	308.50	335.91	364.27	434.93
		T	220.52	205.69	263.46	305.30	365.65	417.47	476.14	646.12
76809	248.51	F	158.80	156.05	205.91	242.58	296.15	342.70	395.94	552.84
		T	159.25	162.56	217.07	256.71	314.10	363.55	419.69	583.16
76811	33.97	F	54.71	54.65	66.03	73.04	81.88	88.58	95.40	111.99
		T	56.72	55.16	65.82	72.25	80.22	86.15	92.11	106.28

TABLE B1. Return period peak flow estimates produced using the Generalised Logistic (GL) distribution (L-moments method) at 155 stations across northern England, on a single-site basis, with and without the peaks of December 2015. T and F indicate true and false.