Documentary evidence of past floods in Europe and their utility in flood frequency estimation


aDepartment of Architecture and Civil Engineering, University of Bath, Bath, BA2 7AY, UK
bSchool of Environmental Sciences, The University of Liverpool, Rozby Building, Liverpool, L69 7ZT, UK
cIrstea, Hydrology-Hydraulics Research Unit, 5 rue de la Doua, 69100 Villeurbanne, France
dTechnical University of Madrid, Department of Civil Engineering: Hydraulic and Energy, Madrid, Spain
eDepartment of Civil Engineering, Polytechnique Institute of Castelo Branco, Portugal
fInstitute of Meteorology and Water Management. Podlesna 61; 01-673 Warsaw, Poland
gInstitute of Geography, Masaryk University, Kotlárská 2, 611 37 Brno, Czech Republic
hDepartment of Civil, Chemical, Environmental and Materials Engineering (DICAM), University of Bologna, Bologna, Italy
iDepartment of Irrigation, Drainage and Landscape Engineering, Faculty of Civil Engineering, Czech Technical University in Prague, Thakurova 7, 166 29 Prague 6, Czech Republic
jNorwegian Resources and Energy Directorate, Middelthunsgate 29, 0368 Oslo, Norway
kDepartment of Civil Engineering, Faculty of Engineering, Dokuz Eylul University, Buca, 35160, Izmir, Turkey
lLaboratory of Hydrology, Lithuanian Energy Institute, Breslaujos 3, 44403 Kaunas, Lithuania
mDepartment of Land and Water Resources Management, Faculty of Civil Engineering, Slovak University of Technology Bratislava, Radlinskeho 11, 813 68 Bratislava, Slovak Republic
nGFZ German Research Centre for Geosciences, Telegrafenberg, 14473 Potsdam, Germany
oThe Office of Public Works, Trim, Co. Meath, Ireland
pInstitute of Hydraulic Engineering and Water Resources Management, Vienna University of Technology, Vienna, Austria
qUniversity of Ljubljana, Faculty of Civil and Geodetic Engineering, Jamova 2, SI-1000 Ljubljana, Slovenia
rInstitute of Geophysics, Polish Academy of Sciences. Ksiecia Janusza 64, 01-452 Warsaw, Poland
sDion. Toumazis & Associates, 4 Romanos Str. 1070 Nicosiam Cyprus

Preprint submitted to Journal of Hydrology May 22, 2014
Abstract

This review outlines the use of documentary evidence of historical flood events in contemporary flood frequency estimation in European countries. The study shows that despite widespread consensus in the scientific literature on the utility of documentary evidence, the actual migration from academic to practical application has been limited. A detailed review of flood frequency estimation guidelines from different countries showed that the value of historical data is generally recognised, but practical methods for systematic and routine inclusion of this type of data into risk analysis are in most cases not available. Studies of historical events were identified in most countries, and good examples of national databases attempting to collate the available information were identified. The conclusion is that there is considerable potential for improving the reliability of the current flood risk assessments by harvesting the valuable information on past extreme events contained in the historical data sets.

Keywords: flood frequency estimation, historical events, Europe,

1. Introduction

The reliable estimation of extreme flood events is challenging, but necessary for the design and operation of vital infrastructure such as flood defences, bridges, culverts and dams, and for more general flood risk management and
planning, e.g. emergency planning, flood risk mapping, and for defining flood insurance premiums. In practice, this information is obtained using flood frequency estimation techniques. Through statistical analysis of observed events, a probabilistic behaviour of flood events is inferred which is then extrapolated to provide estimates of the likely magnitude of future extreme events (e.g., the magnitude of the flood expected to be exceeded on average once every 100-year is estimated from a 40-year record). By nature, extreme flood events are rare and seldom observed locally and as a result hydrologists have little chance of gathering an adequate sample of recorded events to make confident predictions. This naturally raises the question of how best to extrapolate to extreme events, when no or only short series of recent events are available. As floods occur in almost all regions of the world, reliable flood estimation is a generic and shared problem. In Europe, the last couple of decades have witnessed a number of high-magnitude low-frequency flood events (Kundzewicz et al., 2013), causing widespread damage and destruction. But flooding in Europe is not a recent phenomenon, and there are multiple accounts of damaging flood events across the continent going back centuries (e.g., Glaser et al., 2004, 2010; Baptista et al., 2011). While the occurrence of extreme floods is a shared problem across Europe (and beyond), the lack of cross-boundary cooperation (national and regional) has lead to individual countries investing in research programmes to develop national procedures for flood frequency estimation. As a result, no standardised European approach or guidelines to flood frequency estimation exist. Where methods do exist they are often relatively simple and their ability to accurately predict the effect of environmental change (e.g., urbanisation,
land-use change, river training and climate change) is unknown (Castellarin et al., 2012; Madsen et al., 2012). Also, the problem of consistent estimates of extreme floods for trans-boundary rivers is rarely considered (Pappenberger et al., 2012). The COST Action ES0901 *European procedures for flood frequency estimation* represents a novel opportunity to develop closer understanding of the methods of flood frequency employed across Europe. The Action is undertaking a pan-European comparison and evaluation of different methods available for flood frequency estimation under the various climatologic and geographic conditions found across Europe, and different levels of data availability. The availability of such procedures is crucial for the formulation of robust flood risk management strategies as required by the Directive of the European Parliament and of the Council on the Assessment and Management of Flood Risks (2007/60/EC).

Currently, flood frequency is most commonly based on systematic instrumental data, collected from established networks of gauging stations operated and maintained by a variety of station authorities/bodies across Europe. These gauging stations are of various forms and complexity depending on the level of data accuracy required. A more detailed discussion of availability, length and types of flood data records as well as procedures for flood frequency estimation procedures used across Europe is provided by Castellarin et al. (2012).

A well-known consequence of the extrapolation from short series is the high level of uncertainty associated with estimates of design floods with large return periods. For example, estimating the 100-year design flood peak from a 24-year record Stedinger and Griffis (2011) reported a factor of 4-to-1 be-
tween the upper and lower bounds of the 90% confidence interval. Given that
the average record length is typically in the range 20-40 years, hydrologists
have attempted to reduce the uncertainty levels by either: i) bringing addi-
tional gauged data from nearby and comparable catchments into the anal-
ysis (e.g., Hosking and Wallis, 1997), or ii) extending the available records
by bringing flood data from before the beginning of systematic flow record-
ing into the analysis in the form of historical and palaeoflood data (Guo and
Cunnane, 1991), or iii) using rainfall stochastic generators and rainfall-runoff
models to constrain extreme flood assessment by rainfall information (e.g.,
Paquet et al., 2013). The three methods all have merit, but only the second
is the focus of this review.

Realising the importance and utility of long-term datasets, flood hydrol-
ogists have increasingly turned their attention to historical flood information
(Brázdil et al., 1999, 2006; Glaser et al., 2004; Böhm and Wetzel, 2006; Mac-
donald, 2006; McEwen and Werritty, 2007; Glaser et al., 2010; Herget and
Meurs, 2010; Kobold, 2011; Santos et al., 2011; Brázdil et al., 2012), and
how best to incorporate documentary evidence of such historical floods into
flood frequency estimation (e.g., Stedinger and Cohn, 1986; Williams and
Archer, 2002; Benito et al., 2004; Gaume et al., 2010; Macdonald and Black,
2010; Gaál et al., 2010). However, the application of non-instrumental data
into flood risk analysis is not new, as is evident from already existing guid-
ance documents such as the Flood Studies Report (FSR) (NERC, 1975) in
the UK, a French handbook for flood risk assessment with historical data
(Miquel, 1984), the guidelines for flood frequency estimation in Germany
(DVWK, 1999), and the methodological guide to implement the Floods Di-
rective in Spain (MARM, 2011). For the purpose of this study we propose three definitions are adopted for the broad classification of different types of hydrological data.

- **Instrumental:** long records, where records have been kept using available technologies, e.g. gauging stations or stage-boards (c. 1850-present)

- **Documentary:** data derived from sources which are intermittent e.g. documentary descriptions or flood levels marked on bridges (c. AD 1000-present). Documentary evidence most often refers to historical events that occurred decades, centuries or even millennia ago, but it can also relate to more recent events in locations where no instrumental data are available.

- **Palaeoflood:** flood signatures recorded within depositional sequences, often sedimentary (channel cut-offs and lakes), though recent work has also witnessed flood signatures retrieved through dendrochronological approaches (Pleistocene present). As with documentary evidence, geomorphological evidence can also refer to recent flood events.

Regarding the historical and palaeoflood data we can add the following definitions:

- **Perception threshold:** level or discharge above which contemporary society considered the event sufficiently severe to record information about it, e.g. epigraphic markings (Macdonald, 2006) or a written account in news media or a specialist publication.
• Censored data: unmeasured floods known to have occurred above or below the perception threshold, despite not knowing their exact magnitude. Several researchers have shown that just knowing that a flood exceeded a perception threshold can add significant value to the flood frequency analysis (e.g., Stedinger and Cohn, 1986; Cohn and Stedinger, 1987; Payrastre et al., 2011)

An important complication when considering documentary and palaeoflood data is the impact of a changing environment (i.e. changes in climate and land-use, or river engineering works) on the characteristics of the flood series, and how to include this impact in future predictions.

The importance of data for assessing both the hydrology and impact of past events has been recognised as an integral part of flood risk management by the EU Flood Directive. The information collected in the Preliminary Flood Risk Assessment (PFRA) documents developed by the individual EU Member States starts with readily available or easily derivable information, such as records and studies on long term developments. Member States describe flood events that occurred in the past, which had significant adverse impacts, and for which the likelihood of similar future events is still relevant, reporting the frequency or recurrence of these events. The likely impact of climate change on the occurrence and impact of floods shall be taken into account in the review of the PFRA. For this, information beyond the instrumental records is acknowledged as being able to reduce the uncertainty of the assessment.

A key part of the COST Action ES0901 is to improve understanding of the barriers to new approaches to flood estimation. The results and discussions
presented in this paper are mainly based on responses from a questionnaire circulated among COST Action participants on the use of historical floods and documentary evidence in flood frequency estimation. Specifically, this paper will undertake, first, a review of the general challenges for the incorporation of documentary evidence within flood frequency estimation. The focus of this paper is not to address the issues of data sources and information, which have previously be examined in detail by others, such as Brázdil et al. (2006, 2012), but to examine the use and application of historical records and information in flood frequency analysis; specifically. Second, challenges with the application of historical information within a changing environment will be assessed. Then, a review of the use of historical information in flood frequency estimation across Europe is undertaken by examining the detailed questionnaire responses which represent the position and statements of the individual countries. Finally, the paper will conclude by considering the current barriers to further application and potential developments.

2. Challenges for broader application of historical information

As documentary evidence most often predates the installation of gauging stations, and is not directly supported by other instrumental sources (using a limnimetric scale e.g. stageboards), it generally provides indirect information on peak flood discharge, often in the form of a water level marker (Figure 1), or information that a specific location had been flooded, damaged or destroyed, or that the water level had reached a level relative to a structure (e.g. it had reached the top of the doorframe).

Different quantitative methods have attempted to extract the information
contained in historical data using a variety of approaches. The most common approach is to consider a perception threshold for a historical period or sub-period, with the assumption that each flood exceeding this threshold has been recorded (e.g. NERC, 1975). As the consequences are important, this can sometimes be aided by thresholds within the environment of known exceedance. An example is the flooding of the Lincolnshire Plains by the River Trent in Central England when a low lying moraine (Spalford Bank) is overtopped, which is known to occur at flows in excess of 1000 m$^3$s$^{-1}$ (Macdonald, 2013). Having established the threshold, the number of exceedance events during a period can then be retrieved from historical records. A more detailed approach involves the use of hydraulic formulae (e.g. Manning equation) or one or two dimensional hydraulic models (St Venant equations) to convert historical flood levels into historical discharges (Lang et al., 2004a). As shown by Neppel et al. (2010) it is important to ensure that the hydraulic model calibrates with flood marks and rating curves (when available) and reassess the hydrological homogeneity of discharge estimates at several places. Hydraulic studies should provide a discharge estimate, but also a range of possible values within an interval, based on a sensitive analysis or an uncertainty analysis.

Several statistical approaches were developed in the past to improve the flood frequency curve estimation by extracting the information contained in the different types of historical records discussed above. In the USA, Bulletin 17 B (USWRC, 1982) proposed the weighted moments (WM) technique for incorporating historical information in a flood frequency analysis. The WM technique is a straightforward method that is noticeable for ease of im-
plementation. Stedinger and Cohn (1986) developed a maximum likelihood estimator (MLE), which was more flexible, efficient and robust than the WM technique. Moreover, it allowed the introduction of binomial censored data into the likelihood function; however, MLEs present numerical problems in some occasions. To avoid this drawback, while maintaining the efficiency of MLE technique, the expected moments algorithm (EMA) was developed (Cohn et al., 1997). Reis and Stedinger (2005) proposed a Bayesian technique based on Markov Chain Monte Carlo methods (BMCMC) that improves previous techniques by providing the full posterior distributions of flood quantiles. Likewise, the BMCMC technique allows for the introduction of uncertainty into historical peak discharge estimates. The WM technique was adapted to the case of probability weighted moments (PWM), to produce the partial probability weighted moments (PPWM) approach (Wang, 1990). The EMA technique was also adapted to the PWM case, providing the expected probability weighted moment (EPWM) estimator, which improves the estimation of the shape parameter, but has also shown some bias (Jeon et al., 2011).

An example of how the inclusion of historical events can help flood frequency estimation to better represent the probabilistic behaviour of flood events can be seen in Figure 2. It shows the results at the Tortosa gauging station located on the River Ebro in Spain, a comparison between two Generalised Extreme Value (GEV) distributions fitted to i) a sample of 31 annual maximum flood peaks recorded at the gauging station (instrumental) by the method of L-moments, and ii) the same sample of instrumental events, but enhanced with seven historical flood events by the method of PPWM. From
the frequency plot in Figure 2 it is clear that the GEV distribution fitted to
the instrumental record only, would result in severe under-estimation of the
real flood risk at the site of interest. However, the inclusion of the historical
records estimated from a set of flood marks recorded at a house close to
the reach improved the estimation of extreme return period floods, as their
magnitude was unknown from the short instrumental record.

Most of these analytical developments have been undertaken within the
academic field. However, extending these improvements to routine practical
use is not trivial, principally because of the mathematical complexity of most
techniques. For instance, classical MLEs are efficient for sufficiently long
records, but may produce numerical problems in application to case studies
when sample size is small (El Adlouni et al., 2007); a significant drawback for
recommending this technique for practical application. Bayesian techniques
also present critical steps, such as the estimation of prior distributions and
the computation of posterior distributions which are not always straightforward. The elegant statistical models based on censored data sources and
solved using likelihood functions, sometimes combined with Bayesian statistics (Reis and Stedinger, 2005), can provide very good results. Nevertheless,
this review suggests that whilst these models exist, there is limited evidence
that they have migrated from the academic field into operational guidelines.
Potential barriers to the broader application of these approaches may reflect
the complex computational requirements and site specific characteristics that
may be best combined with specific methods, though the survey undertaken
in this study did not contain information on why certain approaches are not
applied. These problems lead to the use of the more simplistic, but robust,
methods in practice, as recommended by operational guidelines, such as the WM technique in the United States and the PPWM in Spain.

In addition to providing formal input into quantitative flood frequency estimation, documentary evidence of past events can be helpful in communicating flood risk to non-specialist stakeholders (McEwen et al., 2013) and for better understanding variations in flood seasonality (Macdonald, 2012).

The transformation of information from descriptive accounts of past events into more easily understood groups of flood magnitude has seen the use of indices, often using a scale dividing the events into a set of qualitative classes (Sturm et al., 2001; Llasat et al., 2005) for flood severity, see Brázdil et al. (2006, 2012); for example class 1 (low to intermediate events: damage and flooding are limited to restricted areas), class 2 (high events: flooded area and debris flow are important, structures such as dikes and roads have been destroyed for several hundred of meters), class 3 (extreme events: damage or destruction of important structures and flooding on the whole plain). Although a useful tool for categorising and visualising flood magnitude, this approach has yet to be useful in the estimation of flood frequency, and is unlikely to present any advances as the approach removes individual event information and groups the events, thereby reducing the potential value of the data.

3. Assessment of environmental change

There is some discussion provided as to means of accounting for the impact of environmental change on flood occurrence, with several countries undertaking comparison to nearby stations, for non-homogeneity and trend
studies. However, in a review of existing guidance in European countries on how to include considerations of environmental change in flood frequency estimation, Madsen et al. (2012) found that generally little or no guidance is provided for how to deal with trend or non-homogeneity when identified, and how this knowledge should be incorporated into flood estimation. This is clearly an area where much more effort is required to translate scientific research into operational guidelines.

Different types of non-stationarity can be considered within historical records, as the frequency distribution could change during the period for which historical and palaeoflood data are recorded: i) the changes related to non-homogeneity problems (historical data availability, transformation of indirect information to discharge estimate); ii) climatic variability over long time scales could limit the utility of historical data under a stationarity framework to some hundreds of years in the past (Hosking and Wallis, 1997). This topic remains an open field of research, with present interest amplified by the perspective of climate change for the 20th and 21st centuries; iii) channel changes (natural and anthropogenic) over long timeframes (e.g., Brázdil et al., 2011a). As a means of minimising the potential impact of these climatic non-homogeneities, historical records used for flood frequency analysis are not extended back beyond around 400 years in Spain. This practice limits the influence of past climatic changes; as a greater frequency of extreme flood events are found in the period 1540-1640 (Benito et al., 2003). Similar timeframes are recommended in a number of academic papers (e.g. Parent and Bernier, 2003; Macdonald, 2013), but this often focuses on concerns relating to data quality and quantity prior to this (as discussed above) rather
than climatic variability, with several studies commenting on the longer time-frame providing greater climatic variability, and therefore a more uncertain climate range (e.g. Macdonald et al., 2006). These issues become even more important when attempting to merge gauged flow data with palaeoflood data stretching back millennia, though it could be argued that climatic variability over millennial timescales incorporates sufficient variability that climate phases become less significant. While some researcher have embraced the use of palaeoflood data (Baker et al., 2002), others remain more sceptical of their practical utility, especially when regional flood frequency methods are available (e.g. Hosking and Wallis, 1986). Notably, Neppel et al. (2010) identified large error associated with historical flood magnitude estimation could lead to a reduction in the precision of design flood estimates when compared to estimates using gauged data only, supporting the view that palaeoflood data should be handled carefully when included into a flood frequency analysis.

Lang et al. (2004b) proposed a statistical test based on the Poisson process for the detection of changes in peak-over-threshold series. It has been applied to several historical series in France and Spain (Barriendos et al., 1999) and in central Europe (Glaser et al., 2004). The power of the test is limited when the number of historical floods is low. On the contrary, including low to intermediate historical floods increases the risk of non-homogeneity, as such floods can be strongly influenced by anthropogenic changes. It is therefore recommended to check the validity of the rating curves used for historical floods.

The development of slackwater deposits as a tool in the reconstruction of palaeoflood series has expanded extensively over the last couple of decades
Werritty et al. (2006); Jones et al. (2010); Huang et al. (2012); Dezileau et al. (2014), with a number of review papers (e.g. Benito and Thorndycraft, 2006) and books (Gregory and Benito, 2003) addressing the topic in detail.

Lakes can act as efficient repositories for sediments eroded from within the catchment and that are transported through the fluvial system (Mackereth, 1966). The sediments reaching a lake are dependent on a number of variables which may vary through time and space; see Schillereff et al. (2014) for a full review. The sediments that reach the lake may be laid down providing a sedimentary record of high-magnitude flows which appear as distinct lamina-tions of coarse material. An increasing number of studies have examined lake sediment sequences with the intention of determining flood histories (Noren et al., 2002; Gilli et al., 2013; Wilhelm et al., 2013). The sediments preserved within the lake can contribute valuable information on flood frequency and potential magnitude of single events over timeframes reaching several millennia (Noren et al., 2002). For example, Swierczynski et al. (2013) derived a 7,000-year flood chronology for the lake Mondsee in Upper Austria. Even the seasons of the palaeofloods could be precisely determined by the micro-stratigraphic position of a detrital layer within the annual succession of lake deposition. This flood chronology shows a striking variability in the flood occurrence from decadal to millennial time scales. There is a period of more than 200 years (21 B.C. – 216 A.D.) without any flood documented, whereas the average frequency is 0.04 floods/year yielding 9 floods for such a time interval.
4. Questionnaire on use of historical data in flood frequency estimation

As part of the COST Action ES0901 *European procedures for flood frequency estimation* a review was undertaken examining if, and how different European countries incorporate historical information into flood frequency analysis. Responses were collected from 15 European countries, representing the different participant countries of the COST Action; all participant countries were invited to contribute through the completion of a questionnaire, which was initially distributed to COST participants, who completed or passed onto colleagues better placed to do so. The questionnaire applied the definitions detailed above so as to distinguish between historical and instrumental data series. A summary version of the questionnaire responses is provided in Table 1.

**TABLE 1**

The following three sub-sections summarise the information collected from the questionnaires. In particular: i) the length of existing historical data series, ii) the accessibility to historical flood data, and iii) summaries of specific guidelines developed in European countries.

4.1. Data availability

Each country was asked to provide details of the sites and locations where the most complete historical series are available. This information is used to provide an indication of the types and use of historical records as a series of national summaries, but cannot be considered as an exhaustive inventory.

For each reported case-study the ratio between the length of the instru-
mental record and the total time from the end of the instrumental record until the first recorded historical flood event was calculated. The average of the ratios calculated from the case studies within each country are reported (Table 2) together with the number of case-studies and the oldest recorded flood event. Note that the oldest flood refers to the oldest flood event associated with an estimate of peak flow; in some countries, older events were recorded but could not be assigned an estimate of the discharge.

TABLE 2

The average ratios are all below 0.50 suggesting that additional information of extreme floods can be found as far back in time as twice the period covered by the instrumental record. The countries listed in Table 2 are representative of North, South, East and West Europe, indicating that historically augmented flood estimation could be useful across the continent. While no quantitative assessment of the benefit of the extended data series were conducted as part of this review, several previous studies have highlighted the utility of such series. For example, Macdonald et al. (2013) found that extending a 40-year instrumental record with documentary evidence of flooding dating back to 1772 resulted in an almost 50% reduction on the uncertainty of the estimated design flood with a return period of 100 years. Similar conclusions have been reached by other researchers such as Payrastre et al. (2011). Thus, the data series listed in Table 2 represent an important resource for providing more reliable estimates of flood risk across Europe.

4.2. Central depository of historical data

No centralised database exists as a depository for flood information at a European scale. But a variety of laudable national/regional/local and
individual databases exists. However, there is no common agreed format, and the databases often include either/or both qualitative and quantitative information with limited quality control on the information uploaded. The purpose of existing data varies, which often reflects the structure and types of information collected, the result is that some disciplines may feel insufficient or ‘the wrong’ type of data may be present, reflecting the varied uses of historical information, from those examining social impacts of past floods to those interested in using the information in flood frequency estimates, as such some disciplines may consider important information to be absent. These databases tend to be funded through a variety of different mechanisms, with few receiving continuous central support; as such they are funded initially, but then become reliant on individuals or professional societies for continuation, good examples being the British Hydrological Society Chronology for British Hydrology Events (BHS CBHE), as described by Black and Law (2004), or the French national Historical Database BDHI currently in development in the framework of the EU Flood Directive (Lang et al., 2012). Whilst a valuable resource the full potential of these databases cannot be realised in pan-European flood frequency estimation at present, due to the absence of a standardised method for construction and minimum data requirements. The National Disaster Archive compiled by the Disaster & Emergency Management Presidency (AFAD) in Turkey, for example, provides tabular and spatial information (date, location) about the entire spectrum of historic disaster events (e.g., floods, droughts, earthquakes, landslides, forest fires, nuclear accidents, etc.) associated with figures of deaths, injuries, affected populations, etc. However, this is not immediately utilizable
in flood frequency analyses due to the lack of data describing the physical characteristics of the events, such as flood levels and discharges.

Recent efforts by a group of researchers from the Slovak Academy of Sciences started with mapping of all historical flood marks and collecting historical reports of floods in Slovakia. Their results are continuously published, e.g. recent studies by Pekárová et al. (2011, 2013) give the overview of the history of floods and extreme events in Slovakia and in the upper Danube River Basin at Bratislava.

These databases provide pockets of knowledge, but large areas of Europe remain ungauged. The use of geospatial databases for the visualisation of information and capability to embed images within such databases presents an important development, permitting flood levels and additional information beyond a basic descriptive account to be housed within each flood account, empowering the researcher to more rapidly and easily access required information. One of the principal constraints to the wider application of historical information in flood frequency analysis has been the time requirements for collecting the necessary data; well developed and constructed geospatial databases present a valuable step towards removing these constraints.

4.3. Practical guidelines for inclusion of historical data

A number of countries were identified as possessing practical guidelines for inclusion of historical flood information into flood frequency estimation, including: Austria, France, Germany, Ireland, Italy, Slovakia, Spain and the United Kingdom.

Austria
In Austria historical information, where available, was included in the development of national maps of flood discharge (Merz et al., 2008). The historical information was included in flood frequency estimation procedure based on the use of likelihood functions of censored information and Bayesian modelling techniques as described by Merz and Blöschl (2008) and Viglione et al. (2013).

**France**

Miquel (1984) presented a methodological guide for the inclusion of historical data in flood frequency analysis. It was based on a Bayesian approach to peak-over-threshold (POT) values with an a posteriori estimate of the flood distribution, by combining with the Bayes theorem and a priori distribution based on instrumental data and historical POT values. Parent and Bernier (2003) presented an application of this model, using a MCMC algorithm for computation. Naulet et al. (2005) used a maximum likelihood approach on annual maximum values, with different sub-periods (each one being related to a threshold of perception according to documentary sources availability) and different types of data (censored, censored with uncertainties, binomial censored). Lang et al. (2010) and Neppel et al. (2010) applied an error model on discharge estimate, accounting for random errors (sampling uncertainties) and systematic errors (water level and rating curve errors). They showed that ignoring the rating curve errors may lead to an unduly optimistic reduction in the final uncertainty in estimation of flood discharge distribution. Gaume et al. (2010) and Payrastre et al. (2011) presented a Bayesian framework allowing the use of regional information of historical floods at un-
gauged sites. They also provided results on the usefulness of historical data in flood frequency analysis regarding the type of data (censored, censored with uncertainties, binomial censored).

**Germany**

The German Association for Water, Wastewater and Waste (DWA) and its predecessor DVWK have published guidelines which give recommendations for the use of historical sources and data: DWA (2008): Guidelines on how to exploit and interpret historical sources for determining extreme flood discharges. DVWK (1999): Guidelines for integrating large historical flood magnitudes in flood frequency analysis are based on the methods presented in Bulletin 17B (USWRC, 1982). This publication was superseded by the more recent guidelines on flood estimation which devotes a separate chapter to the integration of large historical flood magnitudes in flood frequency analysis (DWA, 2012). Three alternative approaches are offered to consider historical data in the parameter estimation of the frequency distribution. One of them is based on the definition of a set of likelihood functions representing the actual nature of the available flood information, i.e.: i) discharge of historical information known, ii) discharge is known to fall within an interval (upper and lower bound specified), or iii) event is known to have exceed a perception threshold, but the actual discharge value is unknown.

**Ireland**

In Ireland, the generally accepted approach to incorporating historical flood data follows that put forward by Bayliss and Reed (2001) in a similar man-
ner to that described for the UK. With the imminent release of the Flood Studies Update (FSU) methodologies in 2014, growth curve analysis will use L-moment methods to derive growth curves, with the EV1 and LN2 distributions being the preferred distributions for use at gauged locations. It is envisaged that methods of incorporating historical information will move towards the use of L-moment based methods in the future. The central source of information on historical floods will remain the Irish flood hazard mapping website, floodmaps.ie.

**Italy**

The gauging network for systematic river-stage monitoring in Italy was largely installed in the twentieth century, therefore Italian streamflow records are usually much shorter than 100 years (Calenda et al., 2009). In this context, historical and non-systematic information on flood events is a valuable resource. Historical evidence of flooding in Italy has been recorded (e.g., Aldrete, 2007), and national databases of historical disasters (mainly landslides and floods) have been established (Guzzetti et al., 1996, 2004). Nevertheless, these databases contain predominantly descriptive information such as: triggering mechanisms, economic losses and casualties, but little information related to peak discharge. Consequently, although basin authorities routinely use information on historical floods for geographically delineating the most vulnerable areas and acknowledge the value of this information for improving flood frequency estimation (see e.g., AdB-Po, 1999), no evidence of practical use of historical floods in flood frequency estimation was identified in Italy at a national level, though examples were found at regional and local scales.
For example an application to the Piedmont region reported by Claps and Laio (2008) and Laio et al. (2011), and local application by Calenda et al. (2009) on the River Tiber.

**Czech and Slovak Republics**

There are several methods for inclusion of historical flood data in flood frequency estimation in the Czech and Slovak Republics, which were published in reports e.g. Dub and Nemec (1969), Kašpárek (1984) and Novický et al. (1992). These methods are based on corrections of systematic errors by estimation of statistical parameters (coefficient of variability, skewness) of applied distribution functions. The German guidelines for using historical floods, published in DVWK (1999), was applied by Szolgay et al. (2008). Recent studies in Slovakia used a Bayesian framework to include both local and regional information about historical floods at ungauged sites, and to provide results on the usefulness of different types of historical data in flood frequency analysis (Gaál et al., 2010, 2013).

Flood frequency analysis in the Czech Republic is based on combination of floods derived from documentary evidence and systematic hydrologic measurements, which permits the creation of 500-year series: examples include the Vltava (Prague), Ohře (Louny) and Elbe (Děčín) series in Bohemia (Brázdil et al., 2005). In Moravia (eastern Czech Republic), similar compiled series are available for the River Morava, starting as early as 1691 (Brázdil et al., 2011b). More recently, knowledge of historical floods coupled with flood plain information in Prague was used for the estimation of hydraulic parameters, permitting the calculation of peak discharges of past disastrous
floods during the pre-instrumental period (Elleder et al., 2013).

**Spain**

In Spain, the use of historical records is generally recommended when possible, by fitting a GEV distribution by the PPWM method. In addition, historical records were used in some Mediterranean basins (3) to improve: i) the results of the regional flood frequency analysis, and ii) estimates of high return period quantiles along the Mediterranean East coast of Spain (Jiménez-Álvarez et al., 2012).

The 92nd Region is located in the northeast of Spain, including the rivers of the left bank of the River Ebro with heads in the central Pyrenees (Figure 3). In this region the regional coefficient of skewness (L-CS) estimated from instrumental records was improved by the use of historical information. It was seen that two high flood events that occurred in the 20th century affected most of this region (1907 and 1982). However, they were not recorded, as the former occurred before the existence of a gauging station network in Spain, while the latter exceeded the maximum capacity of the gauging stations. Values of at-site L-CS were improved by the use of a GEV distribution fitted with historical information by the PPWM method. The regional L-CS value was updated by a weighted mean of at-site L-CS with weighting factors dependent on the uncertainty of at-site estimations.

The 72nd and 82nd regions are located in the eastern part of Spain, including the lower parts of the Júcar and Segura catchments that are affected by rare and heavy rainfall events coming from the Mediterranean Sea (Figure 3). These events are caused by cut-off lows occurring in spring and autumn,
when cold air in the upper part of the troposphere moves from northern
latitudes to the south over the warm Mediterranean Sea, generating heavy
convective rainfall events and, consequently, intense flood events. However,
there is a lack of information recorded about these flood events; either they
occurred in the past before a gauging station was installed, or they were not
recorded, as they exceeded gauging station capacity. This lack of informa-
tion can result in potentially severe underestimation of higher return period
quantiles. Estimates with only instrumental records can lead to magnitudes
around 5 to 10 times smaller for the 500-year return period. As floods come
from two types of rainfall events, a Two-Component Extreme Value (TCEV)
distribution (Rossi et al., 1984) fitted by MLE is recommended. In these
regions, the use of historical information in flood frequency is crucial to
achieve reliable estimation of higher return period quantiles. In Spain, the
use of historical information to improve flood frequency analyses is recom-
mended (MARM, 2011). A large catalogue of historical floods is supplied by
the Spanish civil defence organization.

United Kingdom

The use of historical record has been called for since the mid-1970s, ini-
tially through the early work of the Flood Studies Report (NERC, 1975)
and Potter (1978). More recently, Bayliss and Reed (2001) provided the first
approach designed specifically for practitioners on how to augment instru-
mental datasets with documental evidence of historical records. However,
the uptake of this approach has been piecemeal and slow, in part as practi-
tioners still require a user-friendly tool for incorporating historical data into
flood frequency analysis. Current methods widely employed for incorporating historical flood information into flood assessments often consist of a conventional flood frequency plot, with the historical levels/discharges marked on, but importantly not included within the statistical analysis. The use of an informal graphical plotting approach was advocated by Reed and Robson (1999) to permit greater confidence among practitioners in the application of historical data. By contrast, Macdonald et al. (2006) and Macdonald and Black (2010) have advocated the use of L-Moments, as they permit greater flexibility and retained an approach practitioners were already familiar with in dealing with pooled data, compared to more mathematically involved Maximum-Likelihood approaches (Macdonald et al., 2013). Each of the approaches considered a preference for a Generalised Logistic distribution model to represent the flood growth curve. An interesting use of historical information was reported by Williams and Archer (2002) who used historical flood data to assess the return period of a recent large event.

5. Discussion

Despite general agreement in the scientific literature on the utility of historical flood information in flood frequency estimation, the survey undertaken has shown that there is only a limited transfer of methods from academia into practical guidance. A few good examples of guidelines and depositories for historical flood data were identified, but no single unified approach or database is evident. Depositories were identified both as part of larger government hydrometric databases, but also existing independently from official government databases, and operated mainly by volunteers and
populated by citizen science efforts (e.g. UK BHS CBHE). The lack of practical guidelines and fragmented access to historical information are practical barriers towards more operational use of these data sources to support current risk mapping efforts and decision-making problems. In addition, it is also clear that the inclusion of historical information is not always straightforward, requiring a greater degree of scrutiny before application than typically required for instrumental data. In particular, it should be recognised that historical information is fundamentally different from quality controlled streamflow measurements obtained from gauging stations. For example, the degree of certainty associated with discharge estimates from historical information requires special consideration. Research has shown that simply ignoring uncertainties on discharge estimates will favour the use of historical information, as sampling uncertainty is reduced by increasing the length of the flood period. Nevertheless, it is important to correctly describe the uncertainties on peak discharge for the instrumental, historical and palaeoflood data, including errors on water level $H$, on the rating curve $Q(H)$, on the threshold of perception and on the starting date of the historical period. The latter should not be systematically the date of the oldest flood in the historical data set (Strupczewski et al., 2013), but should include a period prior to this. The Bayesian framework appears to be a suitable statistical tool, enabling inclusion of several kinds of data (e.g. single values, intervals, number of exceedances) and able to include errors/uncertainties on discharge estimates (i.e., systematic error on water levels and on the rating curve transformation) into flood frequency analysis.

While this review has found that there is largely consensus in the sci-
cientific literature as to the usefulness of historical data in flood frequency estimation, the methods have overwhelmingly focussed on extending at-site estimates. Few studies have reported on the use of historical information in a regional context. A notable exception is the procedure for certain geographical regions of Spain, where the occurrence of very extreme events in the past has resulted in a set of regional flood frequency curves adjusted upwards to represent the worst case, even if no actual events has been observed at a particular site. This is potentially a very interesting methodological development, recognising the limitations of fitting current statistical models to datasets that are known not to include potentially very extreme events, similar to events that have occurred in other locations within the region.

By contrast, Hosking and Wallis (1997) argue that historical information is of limited use in regional flood frequency estimation; their reservations are based on i) concerns about the accuracy and completeness of the historical information (historical data are most often found in old and large human settlements and not at a representative sample across all possible catchments), ii) representativeness of catchment within a region where historical data are available, and iii) using data so far in the past that the underlying frequency distribution might have changed too much (non-stationarity). A regional model combining both regional and historical data was presented by Jin and Stedinger (1989) combining the index flood method with a GEV distribution where the model parameters are estimated using a combination of probability weighted moments and a maximum likelihood procedure. Gaume et al. (2010) also presented a maximum-likelihood approach to combining regional and historical data within the framework of the index flood method. Sur-
prisingly, no or only little further development of these procedures appears to have been reported in the literature, but this is an area where further research is still required to develop a new generation of risk tools to effectively allow regional models to use historical information, and to define procedures to enable the transfer of historical data between catchments.

The potential of historical information in public awareness of flood risk is considerable, historical events are tangible, with epigraphic markings providing an example of how communities have preserved evidence from past events to educate future generations of flood risk, which may not be witnessed within any single lifetime. Increasingly recognition of the non-quantitative information contained within historical flood accounts is being recognised, providing detailed descriptions of the social and cultural responses to extreme events, responses that inherently shape current flood risk management approaches through learned knowledge within communities. This informal knowledge is increasingly being sought and embedded within local flood risk management plans, as recognition of the value of local lay knowledge has developed (McEwen et al., 2013).

The development of national approaches in individual countries has resulted in no-single approach being applied at a European level, constraining the potential for cross border information transfer, and at worst leading to misunderstanding and poor communication to the public (e.g. flood maps with different flood extents at the boundary). Future research must address several key themes:

- construction of a single database framework within which data can be stored and managed, with both extraction, uploading (preferably
through approaches advocated by citizen science) and geospatial presentation capabilities;

- move towards organisation data sharing across boundaries, with greater free access to data for benchmark sites;

- development of a computationally simple user interface toolbox, within which hydrological series comprising of different data types, lengths and completeness can be assessed together;

- development of a set of practices for the treatment of data uncertainty associated with historical records; and,

- a forum for the sharing and review of best practice at a European level.

Inevitably an assessment of the data has to be made by the individual undertaking the analysis and the purpose for which the data is compiled, but the above proposals would facilitate a more rapid and structured approach to the compilation and analysis of the data, overcoming a number of the obstacles currently cited as prohibiting expansion in the application of historical data.

6. Conclusions

There is increasing recognition that historical records of flooding provide a valuable means by which extreme rare events can be better understood, facilitating more enlightened flood frequency analysis where interest is focused on extreme events (events with a return period in excess of 100 years).

As evidenced within this research (Table 1 and 2), a number of examples of historical flood analysis are present within most European countries, with
a number of countries if not actively incorporating historical flood records into flood frequency analysis considering how they can be used, in compliance with the EU Floods Directive (2007/60/EC). Whilst no single approach is uniformly applied to historical flood frequency analysis across Europe, a number of national and regional approaches exists. As historical evidence is often found in connection with large rivers, the use of this information could be a key driver in both academic and practical investigations of transboundary flood management.

Acknowledgement

The authors would like to thank COST Action ES0901 European procedures for flood frequency estimation for financial support. Two anonymous reviewers are acknowledged for their valuable comments on an earlier version of the paper.

7. References


Strupczewski, W., Kochanek, K., Bogdanowicz, E., 2013. Flood frequency analysis supported by the largest historical flood. Natural Hazards and Earth System Sciences Discussions 1 (6), 6133–6153.

Swierczynski, T., Lauterbach, S., Dulski, P., Delgado, J., Merz, B., Brauer, A., 2013. Mid-to late Holocene flood frequency changes in the northeastern Alps as recorded in varved sediments of Lake Mondsee (Upper Austria). Quaternary Science Reviews 80, 78–90.


activity and climate change over the last 1400 years recorded by lake sediments in the north-west European Alps. Journal of Quaternary Science 28 (2), 189–199.

<table>
<thead>
<tr>
<th>Country</th>
<th>Routine use of historical flood data</th>
<th>Existing recognised approach</th>
<th>Preferred information method</th>
<th>Central depository of catchment change?</th>
<th>Central website of historical flood data?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Austria</td>
<td>Yes</td>
<td>Yes</td>
<td>Bayesian methods</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Czech Republic</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Finland</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td><a href="http://www.reperescrue-saine.fr/carte-php">http://www.reperescrue-saine.fr/carte-php</a></td>
</tr>
<tr>
<td>France</td>
<td>Not routinely</td>
<td>No, but Bayesian methods</td>
<td>When available</td>
<td>Yes</td>
<td><a href="http://undine.bafg.de">http://undine.bafg.de</a></td>
</tr>
<tr>
<td>Germany</td>
<td>No, but some practical use guidelines available</td>
<td>No but Bayesian Likelihood Not routinely Yes</td>
<td><a href="http://mldikeh.de">http://mldikeh.de</a></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ireland</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td><a href="http://webmap.irpi.cnr.it/">http://webmap.irpi.cnr.it/</a></td>
</tr>
<tr>
<td>Italy</td>
<td>Not routinely</td>
<td>No</td>
<td>Maximum</td>
<td>Yes</td>
<td><a href="http://webmap.irpi.cnr.it/">http://webmap.irpi.cnr.it/</a></td>
</tr>
<tr>
<td>Norway</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td><a href="http://webmap.irpi.cnr.it/">http://webmap.irpi.cnr.it/</a></td>
</tr>
<tr>
<td>Lithuania</td>
<td>No</td>
<td>No</td>
<td>-</td>
<td>No</td>
<td><a href="http://webmap.irpi.cnr.it/">http://webmap.irpi.cnr.it/</a></td>
</tr>
<tr>
<td>Poland</td>
<td>Not routinely</td>
<td>No but Bayesian</td>
<td>Not routinely</td>
<td>Yes</td>
<td><a href="http://webmap.irpi.cnr.it/">http://webmap.irpi.cnr.it/</a></td>
</tr>
<tr>
<td>Slovakia</td>
<td>No, but some practical use guidelines available</td>
<td>Yes</td>
<td>MCMC No</td>
<td>No</td>
<td><a href="http://webmap.irpi.cnr.it/">http://webmap.irpi.cnr.it/</a></td>
</tr>
<tr>
<td>Spain</td>
<td>Yes</td>
<td>Yes</td>
<td>PPWM method</td>
<td>No</td>
<td><a href="http://tuaa.afad.gov.tr">http://tuaa.afad.gov.tr</a></td>
</tr>
<tr>
<td>Turkey</td>
<td>No</td>
<td>No but guidelines</td>
<td>-</td>
<td>No</td>
<td><a href="http://www.trp.dundee.ac.uk/cbhe/welcome.html">http://www.trp.dundee.ac.uk/cbhe/welcome.html</a></td>
</tr>
<tr>
<td>United Kingdom</td>
<td>Not routinely</td>
<td>No but guidelines</td>
<td>Graphical method</td>
<td>Not routinely</td>
<td><a href="http://www.trp.dundee.ac.uk/cbhe/welcome.html">http://www.trp.dundee.ac.uk/cbhe/welcome.html</a></td>
</tr>
</tbody>
</table>

Table 1: Use of historical data in flood frequency estimation.
Table 2: Summary of historical flood records. *Ratio* in column four refers to the average ratio between length of instrumental record and the total length of the historical plus instrumental records.

<table>
<thead>
<tr>
<th>Country</th>
<th>No. studies</th>
<th>Year of oldest recorded</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Czech Republic</td>
<td>8</td>
<td>1118</td>
<td>0.22</td>
</tr>
<tr>
<td>France</td>
<td>13</td>
<td>1601</td>
<td>0.23</td>
</tr>
<tr>
<td>Germany</td>
<td>1</td>
<td>1374</td>
<td>0.31</td>
</tr>
<tr>
<td>Lithuania</td>
<td>2</td>
<td>1427</td>
<td>0.33</td>
</tr>
<tr>
<td>Norway</td>
<td>12</td>
<td>1345</td>
<td>0.47</td>
</tr>
<tr>
<td>Slovakia</td>
<td>5</td>
<td>1012</td>
<td>0.24</td>
</tr>
<tr>
<td>Spain</td>
<td>11</td>
<td>1779</td>
<td>0.38</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>14</td>
<td>1210</td>
<td>0.19</td>
</tr>
</tbody>
</table>
Figure 1: Flood marks on the Loire river at Puy-en-Velay (France).
Figure 2: Improvement of the frequency curve estimation by the use of instrumental record (IR) and historical data (HD) available at the Tortosa gauging station in Spain.
Figure 3: Location of regions in Spain where historical information was used for improving the estimation of the frequency curve.