PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A

Reinforcing flood–risk estimation

Duncan W. Reed

Phil. Trans. R. Soc. Lond. A 2002 **360**, 1373-1387
doi: 10.1098/rsta.2002.1005

10.1098/rsta.2002.1005

THE ROYAL SOCIETY

Reinforcing flood-risk estimation

By Duncan W. Reed

DWRconsult, 32 Ilges Lane, Cholsey, Wallingford, Oxfordshire OX10 9PA, UK

Flood-frequency estimation is inherently uncertain. The practitioner applies a combination of gauged data, scientific method and hydrological judgement to derive a flood-frequency curve for a particular site. The resulting estimate can be thought fully satisfactory only if it is broadly consistent with all that is reliably known about the flood-frequency behaviour of the river. The paper takes as its main theme the search for information to strengthen a flood-risk estimate made from peak flows alone. Extra information comes in many forms, including documentary and monumental records of historical floods, and palaeological markers. Meteorological information is also useful, although rainfall rarity is difficult to assess objectively and can be a notoriously unreliable indicator of flood rarity. On highly permeable catchments, groundwater levels present additional data. Other types of information are relevant to judging hydrological similarity when the flood-frequency estimate derives from data pooled across several catchments. After highlighting information sources, the paper explores a second theme: that of consistency in flood-risk estimates. Following publication of the *Flood estimation handbook*, studies of flood risk are now using digital catchment data. Automated calculation methods allow estimates by standard methods to be mapped basin-wide, revealing anomalies at special sites such as river confluences. Such mapping presents collateral information of a new character. Can this be used to achieve flood-risk estimates that are coherent throughout a river basin?

Keywords: flood risk; pooled frequency analysis; consistency; uncertainty; hydrological similarity; digital methods

1. Introduction

Flood-frequency analysis supports flood-risk estimation: the task of evaluating the probability of a flood arriving that is greater than can be safely accommodated. The risk is usually expressed as the probability of such an occurrence within the intended design life of a scheme, or as the typical probability in any year. The reciprocal of this annual exceedance probability defines a 'return period' (in years) of such an occurrence. Though a useful shorthand for the practitioner, others connect return period with regularity, as in the return of Halley's Comet, and fail to recognize its expression of an ever-present risk. It is characteristic that flood-risk estimates are most often required for events of a magnitude or frequency well beyond the range that is supported by formal observations, making for inherently uncertain estimates.

One contribution of 18 to a Discussion Meeting 'Flood risk in a changing climate'.

Flood danger and damage arise principally from high flow velocities and water levels. The hydrologist generally works with a river's flow rate ($\text{m}^3\,\text{s}^{-1}$), which captures neither feature. The flow rate is chosen as the key variable because it is less sensitive to local effects, e.g. where flood behaviour is influenced by engineered structures. The implication of a particular flow for water levels and velocities is left for the hydraulician to assess in a subsequent step.

River flow is the output from a physical system (the catchment) characterized by complex processes, threshold effects and heterogeneity. The most variable conditions impinging on flood formation are usually those in the upper soil layer. In all terrain but the most permeable (e.g. underlain by chalk) or artificially impermeable (e.g. overlain by buildings), prevailing moisture conditions in the upper soil layer largely determine the proportion of incident rainfall available for rapid run-off to a watercourse. In addition to its permeability, porosity and slope, the layer's behaviour is influenced by transient factors, including: prior wetness, the moisture profile within lower layers, and effects arising from land use, vegetation and soil-crusting.

In most climatic zones, the principal flood-generating input to the catchment is rainfall. The flood hydrologist's 'rainfall' has an important distinction from the meteorologist's 'storm': only precipitation within the catchment boundary can be relevant to flood generation. It is easy to deduce from this that an unusually slow-moving storm, or a quasi-stationary weather system, may herald an exceptional flood. However, it is difficult to interpret what the distinction implies for flood risk.

Flood-frequency analysis is generally taken to denote a statistical analysis of floods: their magnitudes (i.e. peak flows) and/or their frequency (i.e. occurrence rates in time). Flood-frequency analysis is a natural choice because flood-risk estimation is an inherently statistical problem. Nevertheless, its use is sometimes heavily criticized (e.g. Klemeš 2000), most especially for its neglect of the physics of catchment processes of flood formation. The implication, attributed more than actual, is that the analyst believes that the available flood data provide a sufficient basis for assessing flood risks. The neglect to represent physical processes, even in concept, would indeed be unacceptable in a study aimed at understanding flood mechanisms. But flood-risk estimation is focused on judging the flood peak that has a given (small) probability of exceedance. Statistical methods take precedence not because they are simple or sufficient, nor because physical processes are irrelevant, but because the methods make efficient use of the most pertinent data: observed flood occurrences.

The fruits of applying statistics can sometimes disappoint. A specialist will often know more about a problem or phenomenon than is reflected in the formal data. Bayesian methods are designed to exploit prior knowledge effectively, where this can be summarized in distributional form. Rosbjerg & Madsen (1996) use a Bayesian approach to constrain the analysis of rainfall extremes at a specific site to be consistent with a regional analysis of rainfall extremes (i.e. pooled across many sites); however, a pooled frequency analysis (see § 3) is as much useful artefact (to control undue variability between estimates at individual sites) as prior knowledge. More typically, the expert brings an appreciation of the kinds of additional information that may reveal more about the flood-frequency behaviour of a river.

The paper takes as its main theme the search for information to strengthen a flood-risk estimate made from peak flows alone. Section 2 reviews site-specific information that can reinforce flood-risk estimation at a particular site. Once ideas of transferring and pooling flood data have been introduced, § 4 considers information to support the identification of hydrologically similar catchments. The discussion widens in § 5 to introduce the theme of consistency in derived estimates. Section 6 anticipates the challenges presented by digital methods of flood-risk estimation.

2. Site-specific information to support flood-risk estimates

It is characteristic that flood-risk estimates are required at return periods beyond those which the length of formal record of flow data can support. This section reviews various kinds of site-specific information to support flood-risk estimates at a particular site. Rationales and mechanisms for transferring information from hydrologically similar catchments are discussed in §§ 3 and 4.

(*a*) *Historical flood data*

Where long-established communities lie close to the river, evidence of past floods can be found in contemporary publications, and sometimes in defined flood-marks or flood-stones (Archer 1992). Bayliss & Reed (2001) discuss ways of organizing a study of historical floods, and review methods for incorporating historical data into a flood-frequency analysis (e.g. Salas *et al*. 1994).

(*b*) *Palaeological information*

Palaeoflood hydrology is the identification and interpretation of past floods from geomorphological evidence. Riparian soils and deposits can be examined to explore whether they are of fluvial origin. In areas undisturbed by the plough, it may be possible to identify a layer or layers associated with specific floods and to estimate their dates (Baker 2000). The subject gains from a multidisciplinary approach to researching botanical and geophysical markers of past floods (Zawada 1997). Slack-water deposits, found where tributaries join the main river, are particularly prized (Kochel & Baker 1988).

To be of greatest use in flood-risk estimation, the palaeoflood investigation will identify or estimate

(i) distinctive markers of the levels reached in past floods;

(ii) the water level above which the palaeoflood series is thought to be reliably complete;

(iii) the start date from which an earlier large flood would have left a mark, had one occurred;

(iv) the flood flow corresponding to the threshold level; and

(v) the age of each flood (e.g. by radiocarbon dating of organic material entrained in deposits).

The first three steps are essential, and the fourth desirable. Although a defining characteristic of such investigations, the fifth step is necessary only insofar as it strengthens the credibility of the palaeoflood series. Because of the uncertainties involved, attempts to estimate the peak flows associated with particular palaeofloods can be unproductive. Thus, when combining palaeoflood data (or historical data) with a frequency analysis of conventionally gauged data, the priority is to establish the long-term frequency of exceedance of a well-defined threshold. In geomorphologically active areas, it is important to check that the palaeoflood data relate to the current catchment configuration.

Other geomorphological approaches are possible. Present-day channel dimensions often reflect typical flood size (Wharton & Tomlinson 1999), while bedrock features or boulders can suggest the largest historical flood consistent with the landscape within which the channel sits (Clarke 1996). In the absence of more detailed information, general classifications of riparian soils may be useful in identifying areas historically at risk of flooding.

(*c*) *Rainfall data*

In many climatic regions, floods stem principally from heavy or prolonged rainfall. Rainfall data can be helpful in interpreting particular floods (particularly extreme ones), in calibrating and validating rainfall-run-off models, and in characterizing a catchment's climate. However, rainfall data can be remarkably uninformative about flood frequency. There are a number of reasons for this.

Firstly, rainfall data generally relate to individual sites, which may not represent conditions across the catchment as a whole. Secondly, whereas the frequency of a flood (i.e. an objective measure of its rarity) can be determined by reference to one feature, usually the peak flow at the site of interest, the rarity of flood-producing rainfall is more elusive. As will be seen in § 5, two features, usually the rainfall depth and duration, need to be specified before the frequency of a heavy rainfall event can be assessed. A given catchment is more likely to flood in response to rainfall that is extreme over a period resonant with its natural (temporal) characteristics than in response to a much longer or shorter event. Thus, it is unclear over which duration to evaluate the rainfall rarity: that yielding the greatest rarity for the particular event, that typically relevant to flood generation on the particular catchment, or an intermediate duration. Finally, heavy rainfall, even of a duration relevant to flood generation on the catchment, may not result in a flood. Climatic effects, usually with a strong seasonal signature, are generally responsible for this. Particularly in southern and eastern Britain, large deficits of soil moisture develop in most summers. These deficits reduce or eliminate the flood response to heavy rainfall.

An example highlights the mismatch between rainfall extreme and run-off extreme on some catchments. The 100 largest floods from a 25-year flow record on Harpers Brook, near Corby in eastern England, are found to be concentrated in winter/spring, with a mean flood date of 18 February and a (circular) standard deviation of 67 days. The 100 largest one-day rainfalls in the same period are abstracted from the raingauge most central to the 74.3 km^2 catchment. The rainfall extremes are found to be concentrated in summer/autumn, with a mean event date of 15 August and a standard deviation of 96 days. The heavy rainfalls and floods occupy opposite seasons!

(*d*) *Catchment models*

The strong influence of catchment wetness/dryness on flood behaviour is one motivation to develop catchment models to represent the relationship between rainfall and river flow explicitly. Such models are useful in many contexts: in short-term flow forecasting, in investigating land-use effects, and in interpreting particular flooding episodes. But does a rainfall-run-off model offer new information about flood frequency? The answer probably depends on how 'good' the model is, and on the manner in which its principal parameters have been calibrated.

Beven (2001) demonstrates the immensity of the task faced in catchment modelling, and the sheer size of the literature devoted to it. Relatively few of these publications focus on the use of catchment models for flood-frequency estimation. Without entering into a detailed classification, two main approaches can be distinguished, between which other possibilities lie. For brevity, only the most distinct approaches, the design event method and continuous simulation modelling, are discussed here.

(i) *Design event method*

The design event approach uses a rainfall-run-off model to transform a 'design rainfall' into the required 'design flood', e.g. a hydrograph with the desired 100-year flood peak. One of the earliest and most developed design event methods is the *Flood studies report* rainfall-run-off approach (NERC 1975; Houghton-Carr 1999).

Conventionally, the rainfall-run-off model is calibrated against rainfall and run-off observations in heavy rainfall episodes. Difficult though such flood-event modelling is, the greater challenge lies in devising a scheme for assigning suitable 'design inputs', typically the depth, duration and temporal profile of the rainfall event, and the pre-event catchment wetness, from which to synthesize the required design flood. If the flood-frequency estimates are to be credible, the assignment scheme has itself to be calibrated. This is a major undertaking. Elements of the scheme (e.g. the rule for assigning pre-event wetness) are adjusted until, on average (when applied to many catchments), the flood-frequency curve synthesized by the design event method agrees with that obtained by a single-site or pooled analysis of peak flows. Monte Carlo methods can offer a way of organizing the search.

Strengths of the design event method are that it generates a hydrograph rather than just a peak flow estimate, and that an estimate of the 10 000-year flood (say) is based on an extrapolated estimate of rainfall frequency rather than on a gross extrapolation of peak flow data. The critical weakness of the design event method is the absence of a strong rationale for choosing a particular scheme for assigning the design inputs. While there may be some gain in calibrating the scheme across hydrologically similar catchments, the choice of the *structure* of the assignment scheme is likely to remain both arbitrary and influential.

(ii) *Continuous simulation modelling*

Conceptual models developed to represent catchment processes and land-use effects in a physically realistic manner have many virtues. Their use can implement, and extend, understanding of hydrological processes across a range of time-scales and space-scales. A particular strength of continuous simulation modelling is the ability to represent moisture conditions in the upper soil layer explicitly and continuously.

Flood-frequency estimation based on continuous simulation modelling (Calver *et al*. 1999) uses a wide range of information, extensive time-series of rainfall and flow data and detailed physiographic data defining soil properties and stream features, in addition to flood peak data. The approach is especially appropriate where a catchment is expected to undergo land-use change. In conjunction with stochastic models designed to generate long series of spatially consistent rainfall data (see Northrop *et al*. 1999), such models can also help to project the sensitivity of flood risk to changed climatic conditions.

However, the continuous simulation approach can only be considered to yield best estimates if, when applied to relevant historical conditions, the flood-risk estimates are broadly consistent with what is known about the flood-frequency behaviour of the river. If, in order to attain this consistency, parameters of the rainfall-run-off model parameter are recalibrated, it is questionable whether the catchment model is informing the flood-risk estimate or the observed flood-frequency behaviour is informing the catchment modelling.

(*e*) *Groundwater levels*

Because they flood only infrequently, highly permeable catchments present a particular problem for flood-frequency estimation. Groundwater-level data help to identify the conditions under which heavy rainfall (or snowmelt), should it occur, will trigger flooding. An empirical approach is possible, imitating techniques used for analysing rainfall-induced landslides (Glade *et al*. 2000), which are similarly episodic. However, an approach based on continuous simulation modelling is likely to be more convincing if a scheme can be devised to calibrate the model jointly against observed river flows and groundwater levels. This is because a model-based approach is likely to be more effective (than a purely empirical approach) when ground conditions are known to be strongly influential (e.g. well-drained soils), where flooding episodes are infrequent but sustained (e.g. groundwater-fed streams), and when there is detailed information about a relevant non-flood variable (e.g. groundwater levels). This rationale for preferring a continuous simulation modelling approach on highly permeable catchments may apply in other cases, where particular features (e.g. land use) are known to have an effect on flood frequency beyond that represented in the available flood peak data.

3. Information transfers and pooled frequency analysis

Pooled frequency analysis uses data from sites believed to be similar to the subject site. There are three motivations to pool flood data.

(i) To avoid over-reliance on the relatively small sample of data available at an individual site.

(ii) In the belief that physically similar catchments in a similar climatic setting share a common flood regime.

(iii) To meet the practical requirement to estimate flood frequency at an ungauged site.

(*a*) *Index flood approach*

The index flood approach recognizes that catchments can be physically similar without being identical. The index flood is taken to be a property of the particular site, while the flood growth curve is evaluated from the pooled flood data. In other words, the ratio of rare to commonplace flood magnitudes, the 'flood growth factor', is assumed to be similar for hydrologically similar catchments. Stedinger (2000) summarizes the index flood approach thus:

> [It] 'substitutes space for time' by using regional information to compensate for having relatively short records at each site.

The index flood is typically defined as the mean (QBAR) or the median (QMED) of annual maxima. The latter is adopted here. In the index flood approach, the T-year flood, Q_T, is estimated as the product of the index flood and the T-year growth factor, x_T,

$$Q_T = \mathrm{QMED}\, x_T.$$

The index flood approach is a two-stage procedure, requiring estimation of the index flood, QMED, and estimation of the flood growth factor, x_T.

(*b*) *Index flood estimation*

At an ungauged site, QMED is initially estimated from a generalized model relating the index flood to numerical descriptors of catchment or channel properties. Such models are, however, inherently uncertain. The model recommended in the *Flood estimation handbook* (Institute of Hydrology 1999) has a factorial standard error of 1.55 (Robson 1999*b*), and is typically outperformed by a QMED estimate made from just two years of gauged flood data (Robson 1999*a*). Where 15 years of gauged data are available, QMED is typically well estimated, with a factorial standard error of 1.1. Thus, in UK conditions, there is considerable potential to improve estimates by local transfer of information from gauged (donor) to ungauged (subject) site. The transfer procedure recommended in Institute of Hydrology (1999) takes the relative error in the generalized estimate at the donor site to be indicative of that expected at the subject site.

Hydrological intuition suggests that data transfers will always be valuable where a donor lies close to the subject site, and their catchments are known to be strongly similar. Where these ideal conditions are not approached, the practitioner has to apply considerable judgement in the selection of donor catchments. Relevant criteria are the quality of QMED estimate available at the donor catchment, the degree of hydrological similarity (between donor and subject catchments), and the margin by which the donor stands in merit above the next-best donor. The performance of different transfer schemes (including the null option of leaving the generalized estimate unchanged) can be tested by cross-validation, i.e. treating each gauged site in turn as if it were ungauged. For reasons discussed in § 6, interest in the performance of these adjustment procedures is set to grow.

(*c*) *Flood growth estimation*

L-moment methods provide a powerful approach to pooled frequency analysis (Hosking & Wallis 1997). L-moment estimators are based on linear (L for linear) combinations of the ordered (i.e. ranked) data values. In extreme value applications, they outperform estimation methods based on conventional moments, and offer comparable performance to maximum likelihood estimation. In pooled analyses, dividing by the index flood standardizes the flood peak data at each site. The workhorses of the approach are the L-moment ratios, which are dimensionless quantities summarizing the variation (strictly, the coefficient of L-variation), skewness and kurtosis of the distribution of extreme values.

The L-moment tool kit provides measures for identifying potentially discordant sites, testing group heterogeneity, and assessing the goodness-of-fit of candidate distributions for flood growth. The discordancy measure is designed principally as a screening technique, to draw attention to possibly erroneous data.

It is tempting to use the L-moment ratios in the judgement of hydrological similarity (see § 4). However, a catchment may appear discordant to others in the 'pooling group' (Reed *et al.* 1999) because its record includes an exceptionally large event, a so-called 'outlier', rather than because the catchment is intrinsically different. It is important to scrutinize a catchment with unusual L-moment ratios for possible data error. However, where the data are valid, any decision to exclude a catchment from the pooling group must be based on hydrological properties, not on statistical properties of the sample. As discussed earlier, a motivation to pool flood data from hydrologically similar catchments is explicitly to avoid over-reliance on the limited sample of data available at an individual site. In a well-constructed pooling group, a particular datum value is unlikely to dominate the analysis unduly. Should the outlier be in the flood series at the subject site rather than in the broader pool, it will be advisable (if at all possible) to develop historical or palaeoflood data (see § 2) to put the exceptional event into a longer-term context.

Particularly on highly permeable catchments, an annual maximum may sometimes be uncharacteristically small, representing a seasonal peak in river flow rather than a flood. Such unusually small 'floods' are an artefact of using the annual maximum series to summarize flood behaviour. Robson & Faulkner (1999) present a technique for adjusting the L-moment ratios to remove the distorting effect of small annual maxima. Such difficulties are avoided in LH-moment methods (Wang 1997), which give greater weight to the larger annual maxima. While lacking the simplicity and clarity of L-moment methods, LH-moments present an interesting extension, given that the practical objective in flood-risk estimation is to obtain a best estimate of the upper tail of the distribution of flood peaks.

Judging hydrological similarity between catchments is crucial to selecting donor catchments (for index flood estimation) and pooling groups (for flood growth estimation). Thus, data that help to establish or refute hydrological similarity provide an important kind of auxiliary information.

4. Information to support the selection of hydrologically similar catchments

Geographical position has long been used as a surrogate for hydrological similarity when defining regions for pooled frequency analysis. Neighbouring catchments usually share a similar climate, and often have similar soils. Moreover, a boundary drawn on a map makes it clear within which pooling group an ungauged catchment lies.

With digital catchment data now available, other pooling schemes can be considered. Similarity in soils and similarity in climate are clearly important criteria, with similarity in topography (including catchment size), land use and stream topology also being relevant. In developing objective criteria for pooling-group construction, Jakob *et al.* (1999) considered different combinations of six variables, preselected from a wider set of catchment descriptors and associated variables. Their final recommendation was to pool catchments according to proximity (i.e. their distance from the subject catchment) in a three-dimensional space representing catchment size, wetness and soils. Burn & Goel (2000) explore a scheme for pooling flood data according to proximity in both attribute (i.e. catchment descriptor) and geographic space.

Unnatural features are difficult to build into an objective pooling scheme. Partly urbanized catchments differ from one another in many more ways than in their natural state: in the type of development, its position relative to drainage paths, its position relative to soil types, and the flood control measures adopted. Because of such factors, flood estimation on mixed-land-use catchments is particularly problematic.

Flood magnitude data are central to flood-frequency estimation, yet are invariably subject to considerable measurement error. In contrast, the dates of floods are usually known without error. Here, ‘seasonality’ is taken to summarize the seasonal distribution (of event dates) in terms of its location (e.g. centred on the mean day of flooding) and dispersion (e.g. the standard deviation about this mean day). Fisher (1993) and Mardia & Jupp (2000) present formal methods for analysing seasonality, including two-sample tests of homogeneity. When making intercatchment comparisons, it is helpful to define the flood seasonality from a peaks-over-threshold series, with the threshold set to yield a particular mean number of floods per year. A significant difference in the seasonal distributions of floods can be interpreted as evidence of hydrological dissimilarity. However, the lack of significant difference should be interpreted more cautiously: as evidence only of possible similarity. As discussed in § 2, the flood regime of a catchment is informed both by the seasonality of flooding and the extent to which this concurs/contrasts with the seasonality of maximum rainfalls. Castellarin *et al.* (2001) use both factors to distinguish ‘regions of influence’ (Burn 1990) for flood-frequency analysis.

5. Discussion

(*a*) *Why is flood-risk estimation so uncertain?*

Modelling the magnitude–frequency relationship of extreme outputs from a physical system is a fundamentally different task from modelling the system itself (i.e. identifying the relationship that output variables take with input variables). Whereas the system model can be calibrated (and validated) by reference to measured inputs and outputs, the magnitude–frequency relationship of the extreme outputs can only be inferred. The relationship is unmeasurable and, for practical purposes, unverifiable. Feedback is not even-handed: there is much greater scope to invalidate an underestimate than an overestimate. When an extreme flood occurs, its magnitude may be measurable but its frequency can only be inferred. Non-stationary effects, i.e. abrupt change, systematic long-term trend or long-term variation, add to uncertainty in flood-risk estimation. Such effects are difficult to diagnose, and may reflect changes in measurement practice, progressive land-use change, long-term variation in climate, or global climate change.

The inherent difficulty of flood-risk estimation can be expressed in another way. It can be argued that estimating the 100-year flood is fundamentally the same problem as estimating the 2-year flood. As indicated in § 3, a good estimate of the 2-year flood (i.e. QMED) can generally be obtained if 15 years of flood data are to hand. Under this reasoning, estimating the 100-year flood might be judged no more difficult than finding a (quasi-stationary) 750-year flood record close to the subject site!

(*b*) *Roles of uncertainty estimation*

Given the inherent uncertainty in flood-risk estimation, one would expect uncertainty estimation to fulfil an important role. Resampling, or permutation sampling, is a practical way of assessing the uncertainty arising from having gauged data only for particular years. In pooled analyses, resampling in year-blocks provides a way of respecting inter-site dependence: a feature of flood and rainfall extremes in all but the sparsest networks. Such assessments are especially useful where record lengths are short or appear to show trends (Robson *et al*. 1998).

Another role that uncertainty estimation can fulfil is to judge the relative merit of competing methods. Where methods are general procedures applicable at any site, there is scope to assess performance by cross-validation and to select the technique yielding (for example) the lowest root-mean-square error. In simpler problems than flood-risk estimation, a map of kriging variance can highlight weak-spots in a gauge network, indicating where new investment in monitoring is likely to reduce uncertainty most effectively. These are examples of situations in which the assessment of uncertainty supports effective decision making.

These examples contrast with the case where an uncertainty analysis shows an estimate to be highly uncertain but provides no information as to how the estimate can be improved or which alternate estimation method might be preferable. In essence, the uncertainty assessment is then a statement of ignorance. Such assessments are appropriate and useful where the decision is between allowing and avoiding a new hazard. However, most flood-risk estimates support decisions of a different character. These are investment decisions, where the choice is typically between relocation, tolerating the current level of flood risk, making improvements, or building a new flood defence. In such cases, a simple statement of ignorance may be as effective as a sophisticated one.

Faced with inherent uncertainty in flood-risk assessments, and limited resources for flood defence improvements, an alternative strategy is to strive for flood-risk estimates that are consistent from site to site. In this way, although the standard of flood defence provided may not be the one intended, the relative protection afforded to different communities in the same river basin is as intended.

(*c*) *Achieving consistency in rainfall-frequency estimates*

The idea of consistency is illustrated first for extremes of rainfall. Hydrologists consider a three-way relationship between rainfall depth, duration and frequency (i.e. rarity). Specification of any two variables should be enough to determine the third. The challenge for the analyst is to develop a rainfall depth–duration–frequency relationship that best respects and integrates the available data. Extensive records are available from daily-read raingauges. However, the analysis of rainfall extremes for shorter durations is typically supported by far fewer gauges and rather shorter records. In these circumstances, it is reasonable to analyse the depth–frequency relationship separately for each duration. It is known that the D-hour maximum rainfall depth of T-year return period cannot exceed the $(D + 1)$-hour depth of that return period. Yet such anomalies are possible unless the D-hour and $(D+1)$-hour rainfall-frequency analyses are explicitly reconciled. Thus, consistency requires that rainfall depth–frequency curves for different durations should never intersect.

It is less easy to define consistency requirements for rainfall estimates at different sites. Ideally, inter-site differences should reflect physical differences in rainfall characteristics rather than sampling effects or differences in measurement practice. A typical inconsistency arises where the depth–duration–frequency analysis at one site is distorted by the chance occurrence of an exceptional rainfall within the period of record. It is inconsistent that rainfall estimates at this site should in consequence differ from those at climatologically similar sites. This problem is overcome by the use of spatially focused methods (see Stewart *et al.* 1999), which ensure that rainfall growth changes smoothly and systematically from site to site. The other ingredient in the index variable approach (see § 3) is the estimation of the index variable itself. Prudhomme & Reed (1999) map the median annual maximum one-day rainfall by modified residual kriging (Martinez-Cob 1996), promoting systematic changes (from site to site) that reflect topographic and marine influences on rainfall extremes.

(*d*) *Achieving consistency in flood-frequency estimates*

In addition to climatological factors, flood frequency is affected by many catchment factors, including catchment size, topography and soil properties. Consistency requirements for flood estimates are therefore more difficult to specify than for rainfall estimates. In general, flood flows are expected to increase down-river, as additional tributaries join. However, flood attenuation effects in lakes and floodplains, or sometimes within the river channel, provide an exception to this rule, as do instances where a river transits permeable strata.

River confluences are special places. A classic blunder is to assume that the 100-year flood downstream of the confluence is the sum of the influent 100-year floods. In reality it is less, because of incomplete dependence between the tributary floods. (Because particular storms have particular ‘footprints’ on catchments, there is incomplete dependence even for confluent catchments that are identical: the so-called ‘areal reduction’ effect.) Koltun & Sherwood (1998) study pairwise dependence in river flooding. Typical requirements for consistency are that the design flood downstream of a confluence should be less than the sum of those on the confluent tributaries but no smaller than the larger of the two.

Pooled frequency analysis based on hydrologically similar catchments (e.g. Jakob & Reed 2000) promotes intercatchment consistency in flood growth curves, although this is strongly dependent on sound judgements of hydrological similarity. Consistency is especially difficult to achieve in estimates of the index flood (Morris 2001).

6. Challenges presented by digital methods

The theme of this paper is that ancillary information should be used to improve flood-frequency estimates. The *Flood estimation handbook* (Institute of Hydrology 1999) introduced fully digital methods to UK flood hydrology. This means that flood estimates by a specific method can be calculated relatively rapidly for many sites. Concurrent advances in information handling make it possible to construct and display maps of flood estimates for all but the smallest watercourses (Morris 2001). Both the eye and the algorithm are good at identifying the unusual.

Inter-nodal differences in flood estimates arise in two main ways. First, they can reflect effects arising from catchment properties represented within the estimation method. Most obviously, drainage area changes abruptly at confluences. More subtle changes occur where permeable and impermeable catchments conflow, or where a largely rural catchment is joined by a heavily urbanized tributary. An inter-nodal difference may also reflect a feature of the drainage network itself, such as the attenuating effect of flow through a lake. Less helpfully, an abrupt change in the flood estimate can reflect an incidental feature of the flood-estimation procedure: for example, a change in the donor catchment used to adjust estimates of the index flood (see § 3). What is a valuable refinement for flood design at a particular site becomes an anomaly when viewed more widely (Morris 2001).

The 'digital revolution' presents a number of challenges. One is to develop effective ways of reconciling estimates at different sites within a river network, so that their interrelationship is physically meaningful. Sauquet *et al.* (2000) apply geostatistical methods to meet this challenge, albeit for the easier problem of mapping mean flows. Another consequence is that analysts will become more aware of dependence effects at confluences.

Will these developments lead to sophisticated interpolations of workmanlike flood estimates, to steadily refined recommendations to practitioners in how they pool or transfer information, or to more profound changes? The author believes the lattermost. Flood estimation will never be quite the same again once the first digital maps of relatively sophisticated flood-estimation procedures are published. Two observations are offered to support this prevision of challenging times.

First, an exceedingly simple method, based on three catchment descriptors and flood data from some 30 UK gauging stations, generated a national flood-risk map with a high political and societal impact. *The Institute of Hydrology report 130* method (Morris & Flavin 1996) provides part of the hybrid *Indicative floodplain map*, widely disseminated by the Environment Agency to promote public awareness of flood risk. The maps have been made possible principally by technological advances in surveying and digital terrain modelling, and in geographical information systems. Visualization of the flood-risk estimates was key to their release. Now that the *Indicative floodplain map* has been enshrined in statutory guidance (DTLR 2001), there is a need to enhance the content and consistency of the maps.

The second observation concerns how the spatial element of new datasets heightens the emphasis on digital mapping, which in turn leads to changes in analysis methods. An example is taken from the adjacent field of hydraulic modelling. Advances in the remote sensing of water levels are prompting changes in the way that wave heights are estimated in the open sea (Hwang *et al.* 1998). It seems inevitable that the analysis of flows in rivers and floodplains will similarly adapt to using direct measurements of water levels (Moore *et al.* 2000) throughout a river system, once the technology is developed to provide data of appropriate quality and resolution. Such advances may not contribute strongly to the understanding of physical processes. However, they may nevertheless rearrange the applied science of river modelling.

7. Conclusion

Flood-frequency estimation is inherently uncertain. The practitioner applies a combination of gauged data, scientific method and hydrological judgement to derive a flood-frequency curve for a particular site. The resulting estimate can be thought fully satisfactory only if it is broadly consistent with all that is reliably known about the flood-frequency behaviour of the river. Extra information comes in many forms, including documentary and monumental records of historical floods, and palaeological markers. Approaches that pool flood data from hydrologically similar catchments are supported by other kinds of information.

Catchment modelling has an important role to play, especially when seeking to represent the impact of land-use change on flood frequency. Uncertainty assessments fulfil various roles, depending on the application to which the flood-risk estimate is to be put. In some decisions, it is necessary to apply due caution in the presence of uncertainty. In other cases, an important goal is to obtain flood-risk estimates that, while uncertain, are consistent from site to site. It is suggested that this challenge can now be met by exploiting digital catchment data and automated calculation methods: to identify anomalies and to develop new estimation and adjustment procedures.

At critical sites, best estimates of flood risk will continue to require individual judgement. Key steps will be to determine the ancillary information that is most relevant, and to decide how this is to be integrated with flood-risk estimates from semi-standard or automated procedures.

David Morris, Max Beran and an anonymous reviewer are thanked for helpful suggestions.

References

Archer, D. R. 1992 *Land of singing waters—rivers and great floods of Northumbria.* Stocksfield, UK: Spredden Press.

Baker, V. R. 2000 Paleoflood hydrology and the estimation of extreme floods. In *Inland flood hazards* (ed. E. E. Wohl), pp. 359–377. Cambridge University Press.

Bayliss, A. C. & Reed, D. W. 2001 The use of historical data in flood frequency estimation. Report to MAFF, Centre for Ecology and Hydrology, Wallingford. Available at http://www.nwl.ac.uk/feh/historical_floods_report.pdf.

Beven, K. 2001 *Rainfall-runoff modelling: the primer.* Wiley.

Burn, D. H. 1990 Evaluation of regional flood frequency analysis with a region of influence approach. *Water Resources Res.* **26**, 2257–2265.

Burn, D. H. & Goel, N. K. 2000 The formation of groups for regional flood frequency analysis. *Hydrol. Sci. J.* **45**, 97–112.

Calver, A., Lamb, R. & Morris, S. E. 1999 River flood frequency estimation using continuous runoff modelling. *Proc. Inst. Civil Engrs Water Maritime Energy* **136**, 225–234.

Castellarin, A., Burn, D. H. & Brath, A. 2001 Assessing the effectiveness of hydrological similarity measures for regional frequency analysis. *J. Hydrol.* **241**, 270–285.

Clarke, A. O. 1996 Estimating probable maximum floods in the upper Santa Ana Basin, southern California, from boulder size. *Environ. Engng Geosci.* **2**, 165–182.

DTLR 2001 Development and flood risk. Planning policy guidance note 25, Department for Transport, Local Government and the Regions, London.

Fisher, N. I. 1993 *Statistical analysis of circular data.* Cambridge University Press.

Glade, T., Crozier, M. J. & Smith, P. 2000 Applying probability determination to refine landslide-triggering rainfall thresholds using an empirical antecedent daily rainfall model. *Pure Appl. Geophys.* **157**, 1059–1079.

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

Houghton-Carr, H. 1999 Restatement and application of the Flood Studies Report rainfall-runoff method. In *Flood estimation handbook*, vol. 4. Wallingford, UK: Institute of Hydrology.

Hwang, P. A., Teague, W. J., Jacobs, G. A. & Wang, D. W. 1998 A statistical comparison of wind speed, wave height and wave period derived from satellite altimeters and ocean buoys in the Gulf of Mexico region. *J. Geophys. Res.* **103**, 10 451–10 468.

Institute of Hydrology 1999 *Flood estimation handbook*. Wallingford, UK: Institute of Hydrology.

Jakob, D. & Reed, D. W. 2000 Improving the accuracy of estimates: a new method for statistical flood frequency estimation in the UK. In *Proc. Int. Symp. River Flood Defence*, Hydraulic engineering report no. 9/2000, Kassel, Germany.

Jakob, D., Reed, D. W. & Robson, A. J. 1999 Selecting a pooling-group. In *Statistical procedures for flood frequency estimation* (ed. A. Robson & D. Reed). *Flood estimation handbook*, vol. 3, pp. 153–180. Wallingford, UK: Institute of Hydrology.

Klemeš, V. 2000 Tall tales about tails of hydrological distributions. *J. Hydrol. Engng* **5**, 227–239. (See also discussion in *J. Hydrol. Engng* **7**, 192–193.)

Kochel, R. C. & Baker, V. 1988 Paleoflood analysis using slackwater deposits. In *Flood geomorphology* (ed. V. R. Baker, R. C. Kochel & P. C. Patton), pp. 357–376. Wiley.

Koltun, G. F. & Sherwood, J. M. 1998 Factors related to the joint probability of flooding in paired streams. Water-resources investigation report no. 9. Columbus, Ohio: USGS. Available at http://oh.water.usgs.gov/reports/pdf.rpts/wrir.98–4238.pdf.

Mardia, K. V. & Jupp, P. E. 2000 *Directional statistics*. Wiley.

Martinez-Cob, A. 1996 Multivariate geostatistical analysis of evapotranspiration and precipitation in mountainous terrain. *J. Hydrol.* **174**, 19–35.

Moore, T., Zhang, K., Close, G. & Moore, R. 2000 Real-time river level monitoring using GPS heighting. *GPS Solut.* **4**, 63–67.

Morris, D. G. 2001 Appraisal of the FEH flood frequency estimation procedures. In *Proc. Flood Estimation Handbook Forum, CEH Wallingford, 13 November 2001.* (Available at http://www.nwl.ac.uk/feh/FEH_FORUM_presentation_ohp_v1a.doc.)

Morris, D. G. & Flavin, R. W. 1996 Flood risk map for England and Wales. Report no. 130, Institute of Hydrology, Wallingford, UK.

NERC 1975 Synthesis of the design flood hydrograph. In *Hydrological studies* (ed. J. V. Sutcliffe) In *Flood studies report*, vol. 1, pp. 374–485. Natural Environment Research Council, London.

Northrop, P. J., Chandler, R. E., Isham, V. S., Onof, C. & Wheater, H. S. 1999 Spatial-temporal stochastic rainfall modelling for hydrological design. In *Hydrological extremes: understanding, predicting, mitigating* (ed. L. Gottschalk, J.-C. Olivry, D. Reed & D. Rosbjerg), pp. 225–235. Wallingford, UK: IAHS Press.

Prudhomme, C. & Reed, D. W. 1999 Mapping extreme rainfall in a mountainous region using geostatistical techniques: a case study in Scotland. *Int. J. Climatol.* **19**, 1337–1356.

Reed, D. W., Jakob, D., Robson, A. J., Faulkner, D. S. & Stewart, E. J. 1999 Regional frequency analysis: a new vocabulary. In *Hydrological extremes: understanding, predicting, mitigating* (ed. L. Gottschalk, J.-C. Olivry, D. Reed, & D. Rosbjerg), pp. 237–243. Wallingford, UK: IAHS Press.

Robson, A. J. 1999*a* Estimating QMED from flood data. In *Statistical procedures for flood frequency estimation* (ed. A. Robson & D. Reed). *Flood estimation handbook*, vol. 3, pp. 77–99. Wallingford, UK: Institute of Hydrology.

Robson, A. J. 1999*b* Estimating QMED from catchment descriptors. In *Statistical procedures for flood frequency estimation* (ed. A. Robson & D. Reed). *Flood estimation handbook*, vol. 3, pp. 100–128. Wallingford, UK: Institute of Hydrology.

Robson, A. J. & Faulkner, D. S. 1999 Adjusting for permeable catchments. In *Statistical procedures for flood frequency estimation* (ed. A. Robson & D. Reed). *Flood estimation handbook*, vol. 3, pp. 204–211. Wallingford, UK: Institute of Hydrology.

Robson, A. J., Jones, T. K., Reed, D. W. & Bayliss, A. C. 1998 A study of national trend and variation in UK floods. *Int. J. Climatol.* **18**, 165–182.

Rosbjerg, D. & Madsen, H. 1996 The role of regional information in estimation of extreme point rainfalls. *Atmos. Res.* **42**, 113–122.

Salas, J. D., Wold, E. E. & Jarrett, R. D. 1994 Determination of flood characteristics using systematic, historical and paleoflood data. In *Coping with floods* (ed. G. Rossi, N. Harmoncioglu & V. Yevjevich), pp. 111–134. Dordrecht: Kluwer.

Sauquet, E., Krasovskaia, I. & Leblois, E. 2000 Mapping mean monthly runoff pattern using EOF analysis. *HESS* **4**, 79–93.

Stedinger, J. R. 2000 Flood frequency analysis and statistical estimation of flood risk. In *Inland flood hazards* (ed. E. E. Wohl), pp. 334–358. Cambridge University Press.

Stewart, E. J., Reed, D. W., Faulkner, D. S. & Reynard, N. S. 1999 The FORGEX method of rainfall growth estimation. I. Review of requirement. *HESS* **3**, 187–195.

Wang, Q. J. 1997 LH-moments for statistical analysis of extreme events. *Water Resources Res.* **33**, 2841–2848.

Wharton, G. & Tomlinson, J. J. 1999 Flood discharge estimation from river channel dimensions: results of applications in Java, Burundi, Ghana and Tanzania. *Hydrol. Sci. J.* **44**, 1–17.

Zawada, P. K. 1997 Palæoflood hydrology: method and application in flood-prone southern Africa. *S. Afr. J. Sci.* **93**, 111–132.