INTERNATIONAL JOURNAL OF CLIMATOLOGY, VOL. 18, 165–182 (1998)o

A STUDY OF NATIONAL TREND AND VARIATION IN UK FLOODS

ALICE J. ROBSON*, TANYA K. JONES†, DUNCAN W. REED AND ADRIAN C. BAYLISS

Institute of Hydrology, Crowmarsh Gifford, Wallingford, Oxon, OX10 8BB, UK

Received 27 February 1997
Revised 26 June 1997
Accepted 9 July 1997

ABSTRACT

Trends in UK flood behaviour are examined using extensive peaks-over-threshold and annual maxima data from 890 gauging stations. The analysis focuses on changes that have occurred at the national scale; data were therefore pooled across all available sites. Two annual series are examined, representing flood size and frequency of flood occurrence. Results are presented for the 40-year period 1941–1980, for the 50-year period 1941–1990 (annual maxima only), and for some limited data for 1870–1995. Three main tests for trend (linear regression, normal scores regression and Spearman's correlation) were applied using (i) conventional theoretical significance levels, and (ii) a permutation approach. Test results were generally significant under the conventional approach but were not significant under the permutation approach. This difference occurs because some of the independence assumptions required for the conventional approach are not satisfied. There is a high year-to-year variation inherent in the data and, relative to this variation, any underlying trends are not significant. Regional and seasonal effects are considered but no significant trends emerge.

Although no trends in national flood behaviour are seen, the year-to-year variations in UK floods appear to show some systematic behaviour (fluctuations) when smoothed using locally weighted regression. Further investigation suggests that these fluctuations are driven climatically; similar fluctuations are evident in UK annual rainfall values. The influence of climatic variation on the pattern of flood occurrences also means that, at the national scale, progressive changes in the UK flood regime, e.g. arising from land-use change, are hidden. Over the period 1940 to 1980 smoothed curves suggest increased fluctuation, with emergence of some possible cyclic tendencies. However, in the context of the limited data going back to 1870, this increased fluctuation is not atypical. Overall, no evidence for climate change emerges for the periods for which data are currently available, although climate clearly has a marked effect on floods. The observed climate-linked fluctuations indicate that inferences from short flood records may differ appreciably from those obtained from longer series. This could have potentially important implications for flood-risk studies and trend detection. ©1998 Royal Meteorological Society.

KEY WORDS: floods; peaks-over-threshold, (POT); annual maxima; trend(s); non-stationarity; climate fluctuations; climate change; UK; smoothing techniques; trends; flood data sets, UK; locally weighted regression smoothing (loess).

1. INTRODUCTION

In the aftermath of a major flood event, it is common to question whether floods are generally becoming more frequent or more severe. Today, such worries are enhanced because of widespread awareness of potential climate change effects. If climatic effects are causing observable changes to the flood regime, then clearly it is important that these be recognized. It is common practice in flood management and flood estimation studies to assume that the underlying flood regime is stationary. If this is not the case, then implications for future flood analyses will need to be assessed and allowed for.

Climatic effects, such as changing rainfall patterns, are not the only possible causes of change to flood regime, even when considering change at a national scale (Beven, 1993). Amongst other things, the UK landscape has undergone widespread and progressive urbanization throughout the twentieth century (e.g. the

---

*Correspondence to: A. J. Robson, Institute of Hydrology, Crowmarsh Gifford, Wallingford, Oxon, OX10 8BB, UK.
† deceased, May 1996

CCC 0899-8418/98/020165-18 $17.50

area of urban land increased by 50 per cent between 1930 and 1990; Fuller *et al.*, 1994). There is convincing evidence that urbanization can markedly affect flood regime. Clear trends in flood frequency and/or flood size have been identified for a number of UK sites (Hollis, 1974; Walling, 1979; McCartney, 1996; Robson and Reed, 1996). Land drainage is another potential, although much debated, contributor to increased flood flows (Ward and Robinson, 1989, section 7.10). In evaluating possible changes in UK flood regime, it is relevant to consider possible causes, because each may have a different implication for future flood management.

An analysis of UK flood trends at the national scale has not previously been undertaken. This paper concentrates on changes that have taken place since 1940 and considers whether there have been trends in UK floods at a national level. Attention is focused on the size of floods and the frequency of flood occurrence. The work complements an extensive study of trend and step-change for individual UK stations (Robson and Reed, 1996).

2. THE DATA RESOURCE

The Institute of Hydrology holds two major UK flood data sets. The first of these is the peaks-over-threshold (POT) data base (Bayliss and Jones, 1993), which now contains peak flows for 890 catchments. The stations cover mainland UK and Northern Ireland, with an average record length of 18 complete water years and a total of nearly 88 000 flood peaks. The POT data have, in almost all cases, been extracted manually from continuous chart records and are carefully quality controlled. Further details about the data base and methods of abstraction are provided by Bayliss and Jones (1993). The second floods data base contains annual maximum flows from 1000 stations. Many of these records were derived alongside the POT data, but the data base also includes records for stations that were unsuited to POT analysis. Note that the annual maximum data base used here contains more extensive information on recent floods than the POT data base.

The study of non-stationarity in POT and annual maxima records for individual stations presented in Robson and Reed (1996) identified the presence of step-change in some records. Often these jumps could be related to alterations to gauging structures and/or rating equations. Data for 12 of the stations with probable gauging problems have been omitted from this study.

The analysis described in this paper is based on annual series, because standardization across sites is then relatively straightforward. Two annual series are considered, the number of floods per year and anual maxima.

2.1. Number of floods per year

The number of floods per year is derived from the peaks-over-threshold data base; it is referred to here as POT flood counts. Some standardization between the sites is required so that inconsistencies arising from differences in the abstraction thresholds are minimized. For this, new higher thresholds were selected which correspond to an average of 3 peaks year^{-1} at each station over its period of record. For a small number of stations, the abstraction threshold was too high and the POT records contain fewer than 3 peaks year^{-1}. Note that throughout this paper only complete water years of POT record have been used. In the UK, the water year is generally taken to begin on 1 October. Standardized thresholds have been determined on the basis of the full record available at each site; they have not been recalculated for shorter time spans.

The distribution of the POT records through time is rather uneven (Figure 1). Although there are data going back to the last century, only nine station records pre-date 1930. Most of the data are concentrated between 1960 and 1985. In analysing for national trend, a compromise between length of record and representativeness of the data is required. The sparsity of information in the early years means that the available stations may not be representative of the UK as a whole; data problems or a site-specific effect could also have an unduly strong influence. Set against this is the importance of looking for trend across as long a period as possible. For this paper, a 40-year time span (1941–1980) was selected, representing the best balance between length and reliability. Data from before 1941 were judged to contain too few stations to be representative. Data after 1980 are very unevenly distributed across the country and this could distort conclusions.

2.2. *Annual Maxima*

The second annual series considered is that of the annual maximum instantaneous flows. Again, some standardization between sites is required because of the high between-site variability of annual maxima. Such variability largely reflects the differences in catchment size, rainfall and physiographic characteristics. In order that the analysis is not overwhelmed by the records for the largest catchments, a scaled rank transformation of the annual maxima is used. For each site, the annual maxima are replaced by the ranks of the observations. These ranks are then scaled to lie between 0 and 1. Thus for observation i, at site j

$$A_{ij} = (\text{rank}_j(a_{ij}) - 0{\cdot}5)/n_j$$

where A_{ij} is the transformed annual maximum, a_{ij} is the annual maximum and rank_j gives the ranked position of an annual maximum value amongst the n_j annual maxima observed for site j.

Other choices of standardization method are possible: two others were considered. These were: (i) ranking the data, but not scaling them; (ii) scaling the data by the mean annual maximum at each site. These alternative procedures were found to give substantially the same results as the scaled ranks described above. For simplicity, results are presented for the one standardization only.

The annual maxima records are also distributed unevenly in time (Figure 1), reflecting the evolution of the UK gauging network. The maximum length of record likely to be representative and reliable was the 50-year 1941–1990 period. The period 1981–1990 could be included because the annual maxima records are more up to date that the POT records. A 40-year series (1941–1980) was also derived to allow comparison with POT counts.

3. STATISTICAL METHODS

3.1. *Tests for trend*

Three main statistical tests were used for identifying trend in the flood records. Multiple testing was used in order to give a clearer indication of whether trend is present. In general, tests were chosen that are robust in the

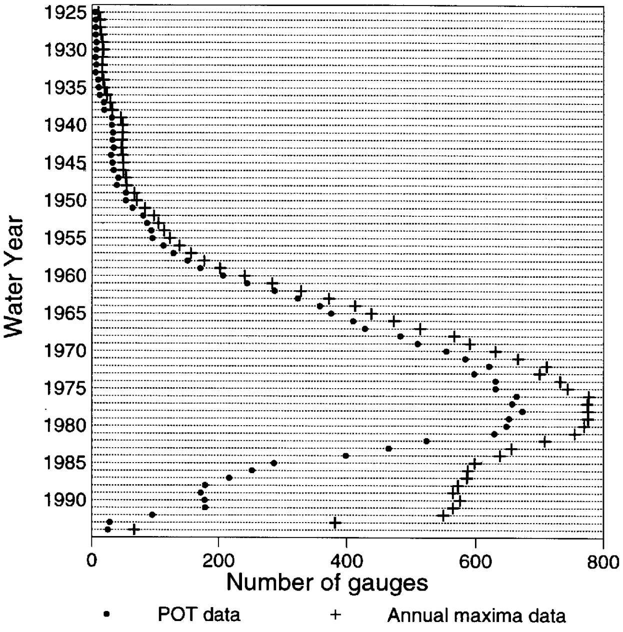

Figure 1. Number of gauging stations operational for water-years since 1925 for peaks-over-threshold (POT) and annual maxima data

sense that they make relatively few distributional assumptions. These tests were applied to the set of all site annual values (for the variable and period under consideration).

- (i) *Linear regression* This test requires the assumptions of normality and independence of observations. The test is included here because it is such a widely used technique.
- (ii) '*Normal scores' linear regression* A robust non-parametric test that is based on linear regression but involves transforming the data so that the assumption of normality is satisfied. Errors are assumed to be independent and distributed identically. The transformation requires ordering the data values and replacing them by the corresponding Normal score statistics, i.e. the *i*th largest observation is replaced by the expected value of the *i*th largest value of a sample of the same size drawn from a Normal distribution (e.g. Sprent, 1989, Chapter 3). Ties were handled by adding a small random component to tied values.
- (iii) *Spearman's rank correlation* A robust non-parametric test that looks for monotone trend by examining the correlation between the ranked data and the year (e.g. Siegel and Castellan, 1988, section 9.3).

The above tests were applied in two modes. Firstly, they were applied directly to the data and tested using conventional theoretical significance levels. These will be referred to as the 'basic' tests. Secondly, they were reapplied using permutation techniques. Under the permutation approach, the same test statistic is used, but test significance levels are determined directly from the data (Good, 1994; Sprent, 1989; Maritz, 1981). Under the null hypothesis of no change, any permutation of the data is equally likely. The null distribution of the test statistic therefore can be determined by randomly permuting the data a large number of times. If the observed test statistic lies in the extremes of the permutation distribution then the null hypothesis is rejected. Because the tests are sample dependent, distributions must be calculated for each flood series. The permutation approach avoids making distributional assumptions and is able to preserve important features of the data such as spatial and time dependencies. For the floods data, this is achieved by permuting the data in year-blocks; all the observations falling in one water-year are permuted together. This avoids the need to rely on assumptions of independent observations, as is implicitly required for the basic tests.

Independence is likely to be violated on two counts. Firstly, some spatial dependency between the sites is to be expected because of catchment nesting and spatial dependencies in extreme rainfall inputs (Reed and Stewart, 1994). Secondly, and perhaps more importantly, there appears to be important structure in the annual values that is rather more complex than is assumed in the basic tests. The most notable aspect of this is the marked year-to-year differences in the data (Figure 2), about which there is individual station variation. This structure indicates that within-year observations are not fully independent, i.e. there are marked spatial dependencies. Permutation tests are able to incorporate these dependencies whereas the basic tests will not.

Some further tests for trend were carried out as an intermediary between linear regression and permutation to help confirm whether dependencies in the data were important. Firstly, an analysis of covariance was used, taking time as the covariate. A random effects component was included in the model to reflect the year-to-year

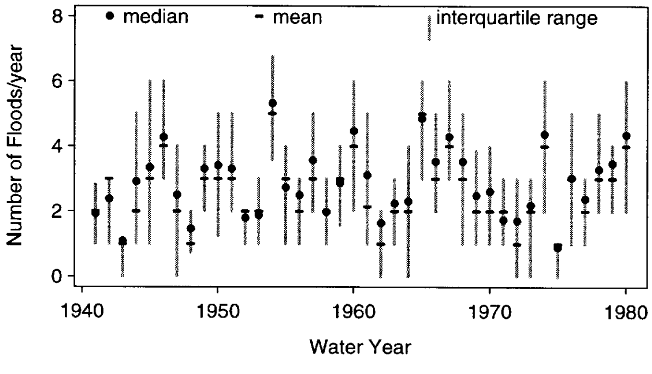

Figure 2. The distribution of number of floods per station for each water-year, 1941 to 1980. Year-to-year differences are large

variations (see Cooper (1969) or Littel *et al*. (1991) for descriptions of random effects models). The basic model fitted to the data was thus

$$Y_{it} = \alpha + \beta t + R_t + \varepsilon_{it}$$

where Y_{it} is the response for site i at time t, α is a national intercept term, β is the fitted national trend, R_t is the yearly random effect assumed to be distributed $N(0, {\sigma_R}^2)$, and ε_{it} is a site error term with distribution $N(0, \sigma^2)$. This model can be thought of as describing a random yearly component, which is common to all observations within a given year, together with a site variation about this component. The null hypothesis is that there is no trend ($\beta = 0$). This was tested using analysis of variance techniques. For this, an F-test was used to test the improvement in fit gained by using a non-zero β, relative to the 'between-year' variance associated with the yearly random effect. For the POT flood counts data, a square-root transformation was used to improve the normality of the residuals. The scaled annual maxima were not transformed. This least-squares approach provides a closer reflection of the underlying error structure than linear regression, although it probably remains an oversimplification (spatial dependence is still not fully allowed for).

A second, simpler, alternative was also considered. This involved a weighted least-square regression analysis of the nationally averaged annual values. Because only national averages are used, assumptions of independence should not hold (dependencies across sites are eliminated). The level of weighting represents the effective number of independent operational sites and can be adjusted to allow for spatial dependencies between sites. Three levels of weighting were used: 1, $\sqrt{n}$ and n (where n is the number of sites operating in a given year). A weighting of 1 represents the extreme of full dependency between stations, a weighting of n represents the opposite extreme of complete independence. A weighting of $\sqrt{n}$ is used as an intermediary.

3.2. *Graphical presentation*

Graphical presentation of the data forms an important component of the analysis. Many of the graphs make use of a smoothing technique known as locally weighted regression smoothing (loess). This is a widely used and relatively robust method (Cleveland, 1979; Cleveland *et al*., 1992, Cleveland and Loader, 1996). The level of smoothing is controlled by a 'span' parameter which dictates the *proportion* of data used within the smoothing window. For the 40-year series a span of 0·2 is used (see section 5). For longer series, the span is adjusted to give a roughly equivalent level of smoothing to the 40-year series. Note that the uneven distribution of the floods data means that the smoothing window covers a much wider time frame in the earlier years than in later years.

When fitting the loess curves to the data, it was possible to make use of a generalized additive model framework (Hastie and Tibshirani, 1990). This approach extends generalized linear regression to the non-linear case. The use of a generalized additive model allows the Poisson-nature of the flood counts series to be incorporated within the model structure. In the terminology of generalized linear models, this was specified using a log link function and a variance proportional to the mean (McCullagh and Nelder, 1983; Chambers and Hastie, 1992). Approximate pointwise confidence intervals can also be derived within the locally weighted smoothing framework. Confidence intervals tend to show decreased width in later years, a result of the increased number of sites.

Smoothing techniques such as loess tend to produce rather unreliable fits at the start and end. This is because the window of points used in the smoothing becomes rather small at the end-points. To help improve this, any available data in the 5 years either side of the main period was used as 'lead in' and 'tail off' phases. This approach has been used throughout the paper whenever smoothing is applied.

4. RESULTS

4.1. *Preliminary analysis*

The above trend tests were applied to the annual flood counts and annual maxima series. Data were pooled across all available sites so as to give a national overview. For both flood counts and annual maxima, the magnitudes of the observed trends were generally rather small (Table I, Figure 3). Trends were significant or near-significant for

Table I. Regression gradients and test significance levels for POT flood counts and scaled annual maxima. Statistical tests are described in section 3.1; values $\leqslant 0{\cdot}05$ are significant

	POT flood counts (1941–1980)	Scaled annual maxima (1941–1980)	Scaled annual maxima (1941–1990)
Regression gradient[a]	0·0060	0·0012	0·0010
Basic tests:			
linear regression	0·019	0·000	0·000
Normal scores regression	0·016	0·001	0·000
Spearman's correlation	0·001	0·000	0·000
Permutation tests:[b]			
linear regression	0·76	0·59	0·44
Normal scores regression	0·73	0·66	0·45
Spearman's correlation	0·79	0·55	0·45
Analysis of covariance	0·79	0·66	0·44
Weighted regression:			
weight 1	0·47	0·72	0·55
weight $\sqrt{n}$	0·68	0·68	0·53
weight n	0·77	0·59	0·49

[a] The regression gradients are in units of floods year^{-1} for the POT data. For annual maxima, the gradients are only useful for comparative purposes.

[b] The permutation tests give very different results to the basic (non-permuted) tests because they allow for structural dependencies in the data.

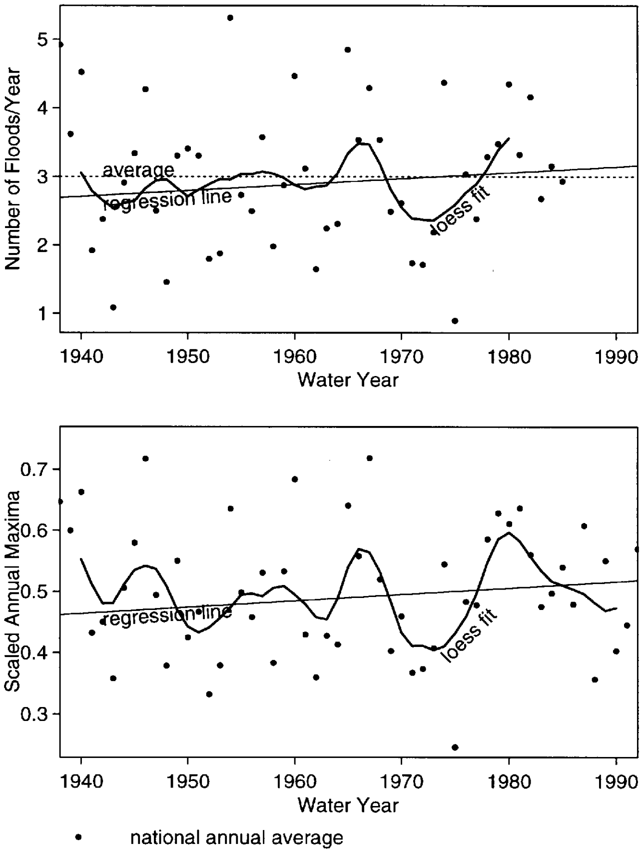

Figure 3. Time series showing variations of number of floods per year and scaled annual maxima. Note that regression line and loess curves are fitted to all site values (not the national averages). The dotted line shows the standardized threshold value (3 floods year^{-1})

the basic tests, but were not at all significant for the permutation and analysis of covariance tests (Table I). The marked difference between the basic and permutation tests, and the agreement between the permutation, analysis of covariance and weighted regression approaches, indicate that the dependencies in the data are particularly important. This is further confirmed by reapplying the permutation tests without preserving the yearly grouping of the data, from which results similar to the basic tests are achieved. It is concluded that the basic tests do not reflect the data structure adequately, even those tests that are 'robust'.

Figure 3 shows the fitted trends relative to yearly variations. Note that in this and the following graphs, regression lines and loess curves have been fitted to the lumped data (annual values for all sites). However, the points shown on the graphs are for nationally averaged annual values (it is not possible to present individual site values clearly). The loess curves strongly suggest that there are fluctuations over 5–10 year periods for both annual maxima and flood counts. In this paper, fluctuation will be used to describe a systematic behaviour in the yearly variations, e.g. where groups of flood-rich or flood-poor years group together to produce a quasi-cyclical effect. Note that in Figure 3, any trend or drift is visually insignificant in comparison with the apparent fluctuations.

4.2. *Seasonal effects*

In this section, floods data are examined briefly to see if any trends emerge when seasonal subsets are considered. Climate change scenarios for the UK suggest that winters will become wetter and summers drier (Hulme, 1996). There is evidence from recent studies of rainfall which appears to fit in with such scenarios (Green *et al*., 1996; March, 1996; Mayes, 1996).

A winter–summer split of the data was made using two 6-month periods, November to April, and May to October. Perhaps the most striking feature is the predominance of winter flooding, with only about one in five floods occurring in the summer period (Figure 4). The plots show roughly parallel behaviour in winter and summer. The results of applying permutation tests are shown in Table II. There are no statistically significant trends in flood counts.

4.3. *Regional effects*

Regional differences in changes to flood behaviour might occur for a variety of reasons. Under climate change, north-western areas are expected to receive increased rainfall relative to the south-east (Hulme, 1996; Arnell, 1996). Regional differences relating to land-use change are also possible, e.g. a higher rate of urbanization has occurred in the south-east. The limited number of sites in the early years means that detailed regional subdivision is not advisable. A simple north-west/south-east divide was investigated.

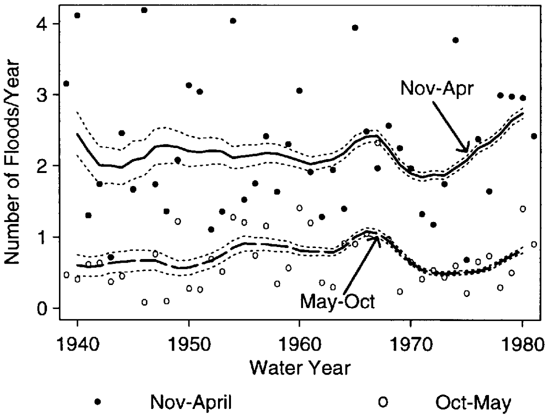

Figure 4. Seasonal flood variations. Loess fits together with approximate 95 per cent confidence intervals are shown. Only a small proportion of floods occur in summer months

Table II. Permutation test results for winter and summer seasons (1941–1980)

	Regression gradient (floods year^{-1})	Linear regression (significance level)	Normal scores regression (significance level)	Spearman's correlation test (significance level)
Winter (November–April)	0·013	0·40	0·43	0·44
Summer (May–October)	−0·007	0·38	0·37	0·45

The north-west/south-east split of the data used the Tees-Exe line as a dividing axis. With this split, around 60 per cent of sites lie to the north-west and 40 per cent to the south-east. The graphs indicate some regional variation for flood counts; flood frequency appears to have increased in the south-east, whereas there has been little net change in the north-west (Figure 5). Regional differences in annual maxima are less obvious, although divergence during the dry spell in the mid-1970s can be seen. There are no significant regional trends for either flood counts or annual maxima (Table III).

4.4. Summary

It is concluded that there is no significant evidence against the null hypothesis of no trend. However, the graphs highlight a notable year-to-year variation in the data. This year-to-year variation is statistically significant when tested using random effects covariance analysis. The smoothed data also suggest that this variation may have become more systematic in later years. In particular, possible cyclic fluctuations appear. In the following sections, the observed fluctuations are investigated further.

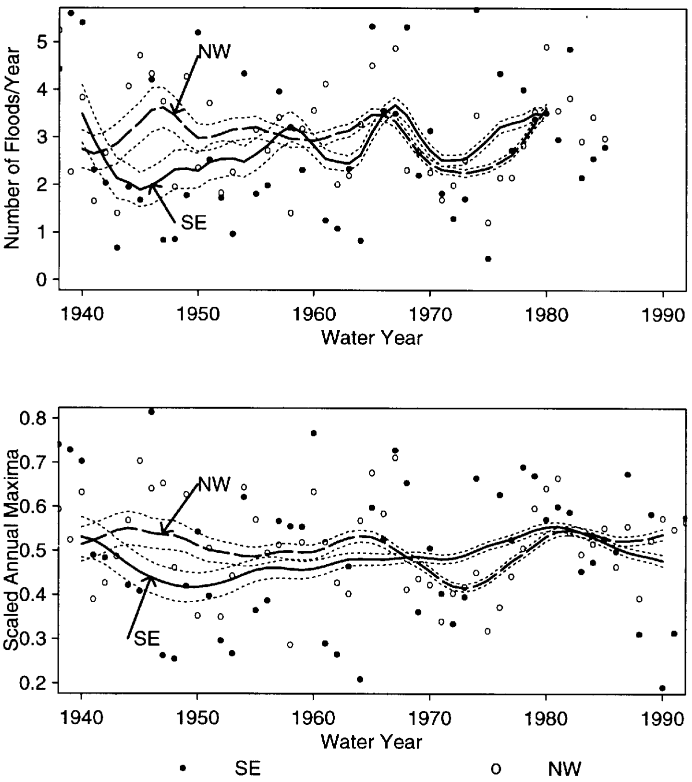

Figure 5. Regional variations for a north-west/south-east divide of the POT flood counts and annual maxima data. Annual means are shown together with loess curves and 95 per cent confidence intervals. The upper graph suggests a possible increase in flood occurrence for south-eastern areas

Table III. Permutation test results for regional data sets showing regression gradients and test significance levels. North-west/south-east data sets are based on a regional divide of the data using the Tees–Exe line as an axis

	Regression gradient	Linear regression	Normal scores regression	Spearman's correlation
POT flood counts (1941–1980):				
north-west	−0·008	0·66	0·59	0·83
south-east	0·028	0·31	0·29	0·42
Annual maxima:				
north-west (1941–1980)	−0·001	0·51	0·54	0·78
north-west (1941–1990)	0·001	0·53	0·49	0·41
south-east (1941–1980)	0·005	0·12	0·19	0·21
south-east (1941–1990)	0·001	0·46	0·52	0·62

5. DATA DISTRIBUTION AND SMOOTHING EFFECTS

The very uneven time-spread of the data (Figure 1) means that care must be taken in applying and interpreting the fitted smoothed curves. For example, Figure 3 shows what appears to be a progressively increasing fluctuation, suggesting that the flood regime is becoming more variable on an annual or longer time-scale. However, the picture could be confounded because the increase in data density over time means that there are differences in the degree of smoothing applied at different times.

The aim of this section is to consider whether the increased fluctuation over time is real. There are two main points that need to be considered:

(i) How much of an effect does the uneven distribution of the data through time have on the perceived fluctuation?
(ii) How dependent are the results on the choice of smoothing technique and the degree of smoothing?

The above issues are addressed in the following paragraphs. Results are presented in full for the annual flood count series. The findings for annual maxima were essentially similar and only selected results are shown.

To help investigate the question (i), a *thinned* POT data set was produced. the thinned dataset contains only sites where POT records began no later than 1941 and ended no earlier than 1980. This results in a relatively uniform data density across the period (the distribution is not perfectly even because data for some sites are missing in some years). There were 25 sites with sufficiently long records to be included in the thinned data set. Together these provided an average of 23 data points per year with relatively even geographical spread. Similar data sets were produced for annual maxima for 40- and 50-year periods, with averages of 28 and 18 stations per year.

The difference in the smoothed fits for the full and thinned data sets proved to be surprisingly small (Figure 6). The main features are closely replicated by the thinned data with, if anything, the thinned data sets showing a greater increase through time in terms of the amplitude of fluctuation. Thus the uneven distribution of data does not seem to be implicated in the increasing fluctuation. No trend is seen in the thinned 40-year data flood counts or annual maxima series. However, for the 50-year thinned annual maxima, a trend can be seen relative to the full 50-year series, despite the general similarity. Permutation tests were applied to the thinned data sets (Table IV). As expected, there were no significant results for the 40-year series, however, the 50-year thinned annual maxima series shows a significant positive trend.

A further evaluation of the effect of data density variation was carried out by using only the nationally averaged annual values when fitting the loess curve (Figure 7). The difference between this and the fit to the set of all site annual values was very small. Estimates of the confidence intervals for the two series illustrate the influence of multiple observations on confidence limits. Because assumptions of independence are not fully met for the set of all site annual values, the associated confidence intervals will be an underestimate. The confidence intervals associated with nationally averaged values are probably more realistic.

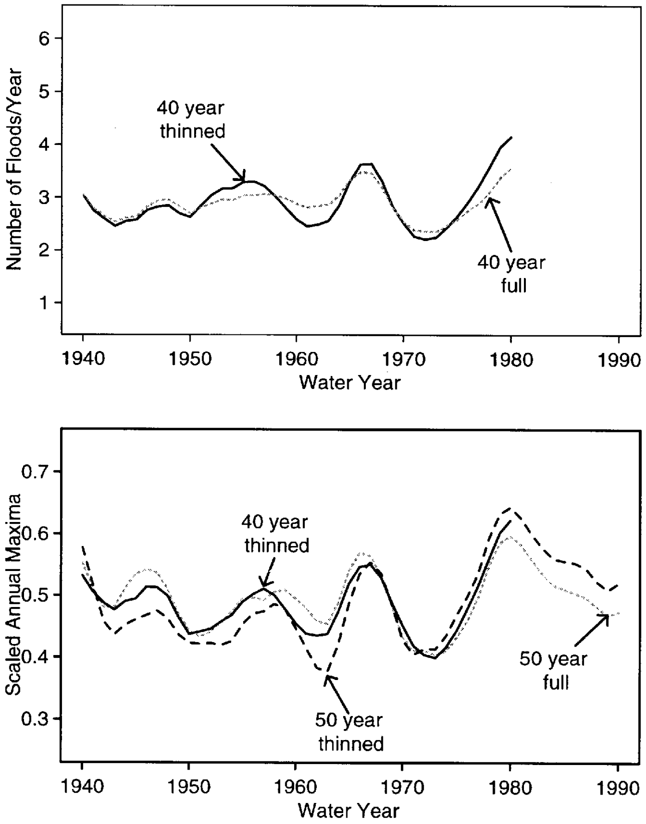

Figure 6. Comparison of loess fits to the full and thinned data sets. The uneven data distribution of the full data set has minimal effect on the smoothed curves

Table IV. Permutation test results for thinned data series. There is a significant positive trend in the 50-year annual maxima

	Regression gradient	Linear regression (significance level)	Normal scores regression (significance level)	Spearman's correlation test (significance level)
POT flood counts (1941–1980)	0·020	0·30	0·27	0·28
Annual maxima:				
1941–1980	0·001	0·53	0·69	0·52
1941–1990	0·003	0·04	0·07	0·04

In order to investigate question (ii), i.e. the effect of smoothing method, a number of alternatives to locally weighted smoothing were considered. These were *Kernel smoothing* (Hardle, 1990), *Supersmoothing* (Friedman, 1984; Statistical Sciences, 1995, section 7.11), and *Cubic smoothing splines* (Hastie and Tibshirani, 1990, section 2.10). Kernel smoothing computes a weighted average at each point in time using a 'kernel' function to provide the weights. A fixed time width is used to determine the smoothing window (instead of a fixed proportion of the data as occurs in loess). Supersmoothing uses a variable span approach, which is advantageous because it can handle changes in the error variance or curvature across the range. It is not such a robust technique as loess. The spline fit is produced by piecewise fitting of polynomials; as with loess, it was possible to use a generalized additive model framework (see section 3.2).

In all of the above smoothing approaches, the degree of smoothing is controlled by a single (method-dependent) parameter. Choosing a sensible value for this parameter is rather subjective, representing a

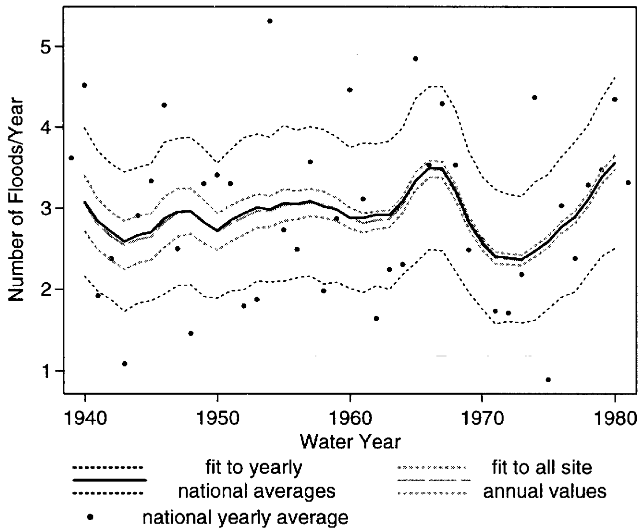

Figure 7. Comparison of a loess fit to national annual averages and to all site annual values. Fits are very similar but the confidence limits for the yearly national averages are much wider

compromise between bias and variance: as the level of smoothing increases the bias increases, but the variance decreases. An example of the effect of changing the loess span parameter is shown in Figure 8. It can be seen that larger span values result in a greater degree of smoothing. In this case, a value of 0·2 was selected as being reasonable; this is the span parameter used for the other 40-year series shown in this paper. Note that at all levels of smoothing, there appears to be increased fluctuation towards the end of the period.

The smoothing techniques described above were applied to the full and thinned flood series. Figure 9 compares the three alternative techniques with loess for the full POT counts data. For this, the value of the smoothing parameter was chosen, by eye, to give a degree of smoothing roughly equivalent to the loess curve (for which a span of 0·2 was used). The results for the loess, kernel and spline methods are seen to be very similar. The supersmoothed fit shows the same broad features but differs in detail, particularly over the first decade. Visually, the supersmoothed fit is too irregular to be appealing. Similar findings occurred for other levels of smoothing and for applications to the thinned data sets.

Overall, the smoothing techniques seem relatively robust; the choice of smoothing technique and the uneven structure of the data do not substantially affect the main features of the fitted curves. In this sense, the fluctuations seen from the graphs are ‘real’. The smoothed data show what appears to be an increase in variability and a tendency towards cyclic behaviour. However, to test for this statistically would be non-trivial and longer records would almost certainly be required. The difficulties of identifying cyclic (as opposed to chaotic) behaviour are well known, even when records are comparatively long (Burroughs, 1992). Furthermore, smoothing techniques may tend to produce apparent cyclic behaviour even if there is no such behaviour in the underlying data. A more productive approach is to consider whether the increased fluctuation can be explained in terms of causative factors, and in particular, climatic variations.

6. THE LINK BETWEEN CLIMATE AND FLOOD FLUCTUATION

Almost all of the floods time series plots in this paper show a marked fluctuation when smoothed. Climatic variation is the most obvious source of the year-to-year variation in flood series; other factors that affect floods, such as urbanization, tend to change in a slow and progressive way rather than to fluctuate. Here, the investigation is kept at a fairly qualitative level and seeks to establish whether commonality can be detected. In particular, the link between flood and climatic variation is explored using rainfall data. A more detailed analysis of the

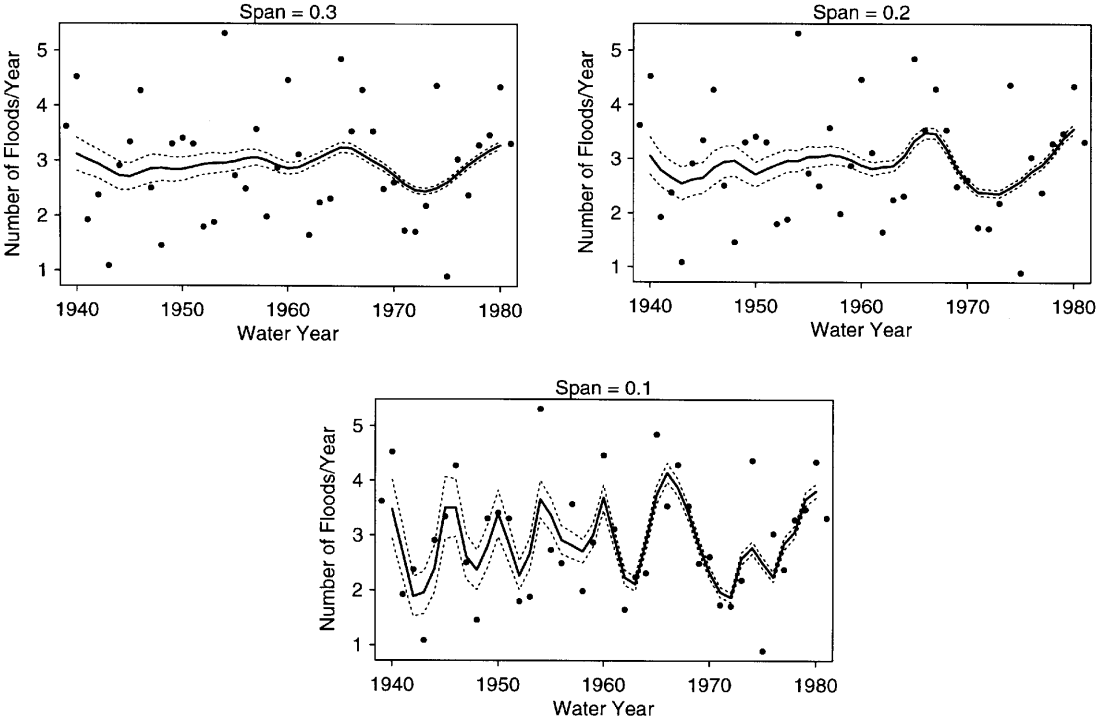

Figure 8. Effect of the span smoothing parameter on loess fits to the data. In all cases there appears to be more marked fluctuation in the second half of the record

interaction between climate and flood regime is beyond the scope of this paper, because detailed hydrological modelling of catchments at a national scale would be required.

The most striking feature of the flood series is the difference between the flood-rich periods of the 1960s and 1980s and the comparatively dry 1970s (Figure 3). These differences can be explained in terms of changes in general UK wetness (UK Review Group on Acid Rain, 1990; Kozuchowski, 1993; Mayes, 1996). The UK experiences variations in the predominance of different weather types (Lamb, 1972; UK Review Group on Acid Rain, 1990) and these affect rainfall patterns and thus floods. Typically, the mix of types appears to vary over the time-scale of decades and affects many aspects of climate (Jones and Kelly, 1982; Karl and Riebsame, 1989; Latif and Barnett, 1994; Rasmusson and Arkin, 1993). Related long-term cycles in flood and other environmental data series have also been noted (Barrett, 1990; Sirios, 1993; Robson and Neal, 1996).

Annual rainfall statistics for Great Britain are used to further explore the effects of climatic variation on flooding. The rainfall statistics are based on long monthly series for England and Wales, and for Scotland (Woodley, 1996). Even though this rainfall series is a rather crude measure of flood-producing potential, it provides a rough guide as to likely conditions. When plots of the total annual rainfall and floods data are examined, there is clearly some harmony between the two time series (Figure 10). The rainfall data show roughly the same kind of underlying fluctuation as the floods data.

Permutation tests for trend were applied to the rainfall data to check for any trends. No trend is found in the 40-year data series although there is a suggestion of trend in the 50-year rainfall data (Table V). The nature of this suggested trend becomes clearer when winter–summer splits of the rainfall data are examined. An increasing trend that is strongly significant emerges for the 50-year winter rainfall series. Unfortunately, there is no 50-year winter flood series with which the winter rainfall data can be compared (seasonal splits are not possible for annual maxima data). Conceivably, this winter rainfall trend could provide an explanation for the trend seen in the thinned 50-year annual maxima series.

The analyses presented so far have concentrated on data from 1941 onwards. However, another 70 years of data exist prior to this (Figure 11). Such data can offer further insight into the links between climate and floods,

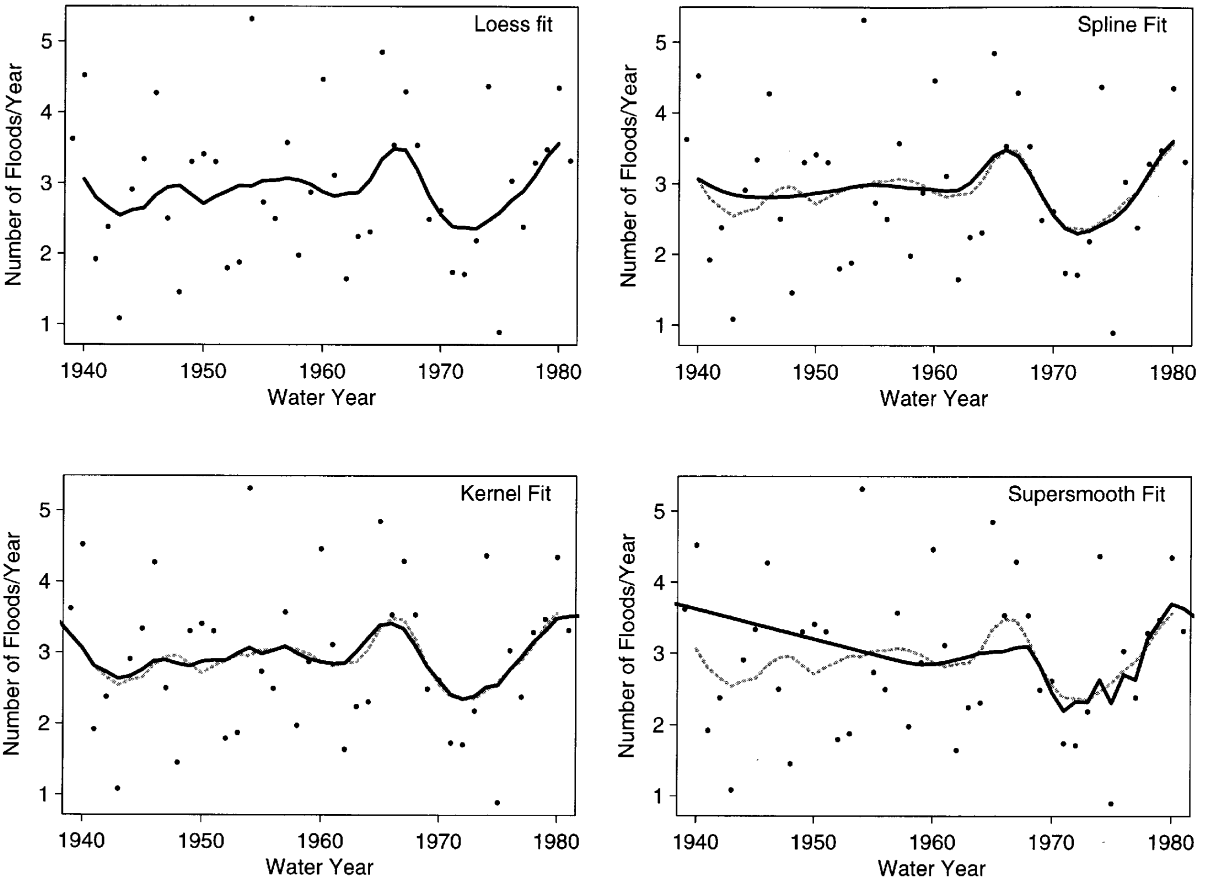

Figure 9. Effects of using alternative smoothing techniques. For each, the level of smoothing is chosen to be equivalent to the loess span parameter of 0·2. Dashed lines show the loess fit for comparison

even though many reservations must apply. Data are inevitably less reliable the further back one goes. The early floods series are based on only a very few sites which are neither geographically nor hydrologically representative. Inference is therefore only really valid on a site by site basis. Note that results for single sites, presented in Robson and Reed (1996), suggest that there are positive trends in flood frequency for around half of the sites operating back to 1900 and earlier. The long rainfall series is not perfect either (Woodley, 1996), but is likely to be more consistent through time than the floods data, given that significantly more rain-gauges than flood stations are available in the early years. A further problem is posed by the changing regional distribution of the flood gauging network. This changed distribution is not reflected in the rainfall averages used here. Given all these caveats, the resemblance between rainfall and POT flood counts for the period 1870 to 1990 seems remarkable (Figure 12). Both the rainfall and the PT flood counts show a major peak around 1880, and lesser peaks around 1915, 1930, 1940 and 1980 (peaks in the 1960s are less closely matched). The troughs are also in reasonable agreement. The correlation between the rainfall and floods data is 0·54 (highly significant; Figure 13). Both rainfall and flood series suggest that there may have been gradual increases since 1900 (Figure 12). Application of permutation tests to the 1870–1990 and the 1900–1990 series do not identify any significant change for rainfall, but a borderline increase for POT counts is seen (Table VI).

The long flood series data help to put the perceived recent increased fluctuation into perspective. It would seem that the fluctuations seen since 1960 may not be very different to those occurring between 1910 and 1940. (The 1880 peak in the flood series is larger still, but is based solely on data for the Thames at Kingston.) The period 1940–1960 was probably unusually quiet in terms of flood fluctuation. The implications of these observations are

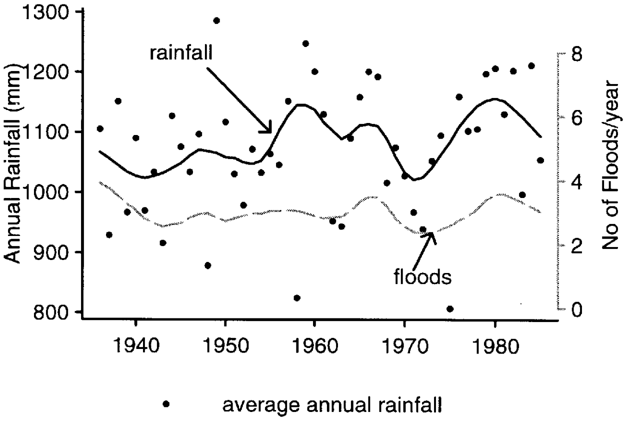

Figure 10. Comparison of flood and rainfall data. Similarities between the two series are apparent

Table V. Permutation test results for Great Britain rainfall

	Regression gradient	Linear regression (significance level)	Normal scores regression (signficance level)	Spearman's correlation test (significance level)
Annual rainfall 1941–1980	1·41	0·35	0·38	0·23
Annual rainfall 1941–1990	1·77	0·10	0·11	0·06
Winter rainfall (November–April):				
1941–1980	2·10	0·06	0·11	0·07
1941–1990	2·26	0·002	0·005	0·006
Summer rainfall (May–October):				
1941–1980	−0·49	0·67	0·76	0·61
1941–1990	−0·50	0·51	0·57	0·45

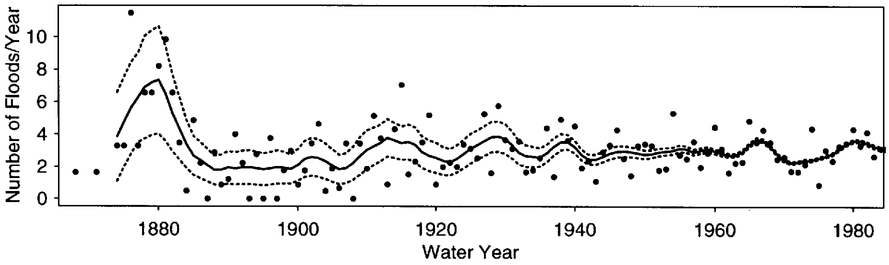

Figure 11. Loess fits to the entire data flood series showing 95 per cent confidence intervals. Prior to 1930 only a very few stations existed and these were not regionally or areally representative; data quality problems are also likely for some of the early records. The fluctuations since 1960 do not seem unusual relative to the earlier parts of the records

clear. If the changes seen in the early data are believed, then the increased fluctuation since 1960 is neither new nor uncommon.

Climatic variations clearly have a marked effect on flooding. The increased level of fluctuation seen in the smoothed floods data suggests that the UK climate has been more variable since 1960 than it was between 1940 and 1960. However, in the context of the longer data series, this increased variation does not appear to be unusual.

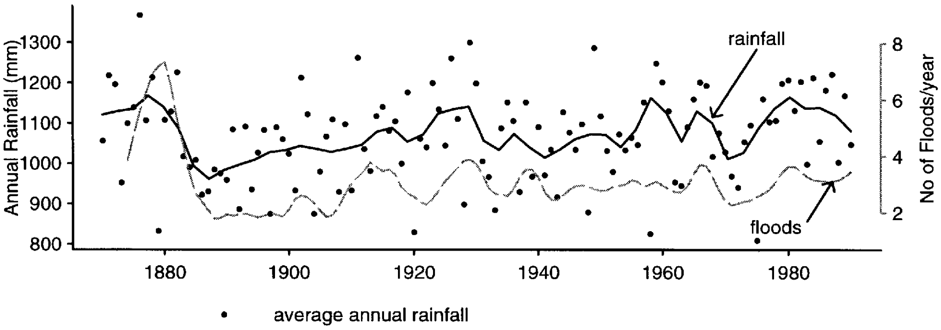

Figure 12. Comparison of rainfall and flood series since 1870. The frequency of flooding is linked to rainfall fluctuations

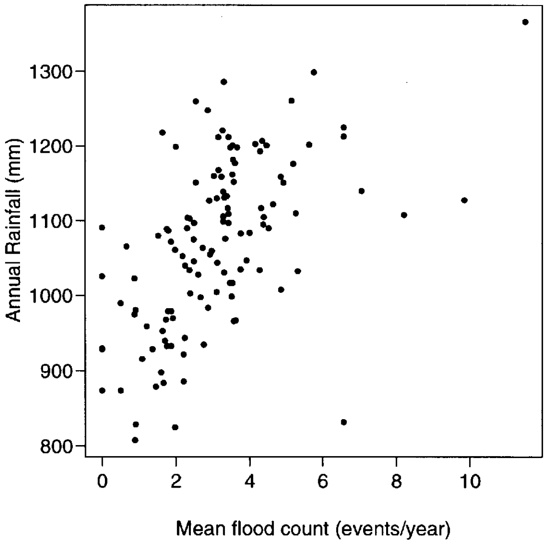

Figure 13. Correlation between annual rainfall and mean number of floods per year. The correlation coefficient is 0·54 (highly significant)

Table VI. Permutation test results for long time series data. There are very few floods data for 1870–1940, so the results should be interpreted cautiously

	Regression gradient	Linear regression (significance level)	Normal scores regression (significance level)	Spearman's correlation test (significance level)
POT flood counts:				
1870–1995	0·009	0·07	0·03	0·21
1900–1995	0·010	0·12	0·05	0·17
Annual maxima:				
1870–1995	0·0008	0·09	0·11	0·36
1900–1995	0·0010	0·10	0·10	0·33
Annual rainfall:				
1869–1995	0·38	0·18	0·24	0·10
1900–1995	0·62	0·13	0·17	0·07

7. IMPLICATIONS AND CONCLUSIONS

7.1. The difficulty of identifying trend

No significant trends emerged in the main analysis of POT flood counts for 1941–1980 and annual maxima series for 1941–1990. It is apparent that any general trend is very difficult to detect because of the significant year-to-year variations in the flood records. Any underlying trend would either have to be very striking (and widespread), or data sets would have to be much longer, for statistical detection to be possible.

The confounding effect of climatic variation means that trends associated with land-use change can neither be easily identified nor readily dismissed. For example, there is some suggestion of increased flooding frequency in the south-east in comparison with the north-west (seen visually but not statistically proven). Such regional differences might reflect land-use change: greater urbanization has taken place in the south-east. However, they could also have arisen from climatic effects.

7.2. The importance of year-to-year fluctuations

Even though trend has not been detected, the observed year-to-year fluctuations in the data could have important consequences for flood design and trend analyses. The yearly variations in flood occurrences and annual maxima were found to be statistically significant. They arise because of the spatial dependencies between sites and because of yearly differences in climate. This additional 'structure' in the data means that care needs to be taken when testing for trend. Even 'robust' statistical tests such as Spearman's rank correlation do not reflect this structure sufficiently. For such tests, results were found to be misleadingly significant due to the incorrect assumption of independence between observations. Permutation tests were found to be preferable because they avoid the need to assume independence.

The yearly variations additionally appear to show some dependencies across periods of 5–10 years, giving rise to quasi-cyclical tendencies. Note that this has not been formally proven (longer series would be required) and could relate to the use of smoothing techniques. Such fluctuations mean that inferences made from short data sets will depend on the climatic conditions during the period, and there is potential for errors to be introduced depending on whether the period considered is unusually flood-rich or flood-poor. There may also be some effects where a period shows unusually small fluctuations (e.g. for the apparently quiescent 1940–1960 period). Great care needs to be taken whenever short records are used for trend detection or for flood design. Even for longer series, flood estimation and flood risk studies may need to take fuller account of climatic conditions during the observation period.

7.3. The issue of climate change

Although the evidence presented here provides no conclusive proof that climatic change has affected UK flood behaviour, this does not mean that climatic effects can be disregarded. On the basis of this analysis, the case remains unproven.

A number of studies suggest that there have been recent changes in UK rainfall and hydrological regimes (Grew and Werrity, 1995; Arnell, 1996; Green *et al*., 1996; Marsh, 1996; Mayes, 1996). Some of these studies point to the very wet conditions seen during winter months in north and west Scotland in the 1990s, particularly since 1993. Long Scottish rainfall series, dating back to before 1870, show that recent winter rainfall totals fall outside the previous observed range (Green *et al*., 1996). For north and west Scotland, some more recent POT records are available, extending up to 1993. Although trends occur at some specific sites (Robson and Reed, 1996), no significant trends are found for the period to 1993 when data are pooled across this region.

Climate-change scenarios suggest that changes in the winter–summer distribution of rainfall will vary regionally, and there is some evidence that such changes are occurring (Marsh, 1996; Mayes, 1996). The limitations of the flood series in terms of examining small regions over long periods, mean that changes of this nature are not identified easily.

Finally, although no evidence of climate change has been found, the results may assist in efforts to assess potential climate effects under various scenarios. Firstly, the fluctuations seen in the flood series could provide a means of characterizing flood behaviour for relatively wet and dry periods. Secondly, the correlation observed between annual rainfall and flood occurrences could provide a baseline estimate for the degree to which flood frequency may change.

7.4. *Importance of long reliable records*

The analyses presented here have concentrated on the periods 1941–1980 (flood counts) and 1941–1990 (annual maxima). *The sensitivity of these analyses to climatic conditions during the period of record, provides a strong reason for continuing data collection*. Inclusion of more recent years would provide a sensible means of establishing whether the lack of trend continues to the present day. Examination of the much longer, but rather sketchy, data sets that go back to 1870, provides an added perspective from which the fluctuations seen since 1960 no longer appear atypical. Long data records are clearly invaluable when it comes to resolving these issues, even given their imperfections.

ACKNOWLEDGEMENTS

This study forms part of the Flood Estimation Handbook research programme, funded by the Ministry of Agriculture, Fisheries and Food, the Environment Agency, the Northern Ireland Department of Agriculture, and (on behalf of four Scottish agencies) the Scottish Office Environment Department. The authors thank David A. Jones for his advice. Comments from Terry Marsh and Alan Gustard were also most useful. The research reported in this paper was begun by Tanya Jones who tragically died in May 1996.

REFERENCES

Arnell, N. 1996. *Global Warming, River flows and Water Resources*, Wiley, Chichester, pp. 1–224.

Arnell, N. W., Brown R. P. C. and Reynard, N. S. 1990. *Impact of Climatic Variability and Change on River Flow Regimes in the UK*, Institute of Hydrology Report No. 107, Wallingford.

Barrett, J. H. 1990. *An extreme value analysis of the flow of Burbage Brook*. MSc Dissertation, Department of Probability and Statistics, University of Sheffield.

Bayliss, A. C. and Jones, R. C. 1993. *Peaks-over-threshold Flood Database: Summary Statistics and Seasonality*, Institute of Hydrology Report No. 121, Wallingford, 61 pp.

Beven, K. 1993. 'Riverine flooding in a warmer Britain', *Geog. J.*, **159**(2), 157–161.

Burroughs, W. J. 1992. *Weather Cycles, Real or Imaginery*? Cambridge University Press, Cambridge, 201 pp.

Chambers, J. M. and Hastie, T. J. (eds). 1992. *Statistical Models*, S. Wadsworth and Brooks/Cole Advanced Books and Software, California, 608 pp.

Cleveland, W. S. 1979. 'Robust locally weighted regression and smoothing scatterplots', *J. Am. Stat. Soc*., **74**(386), 829–836.

Cleveland, W. S. and Loader, C. 1996. 'Smoothing by local regression: principles and methods', In Hardle, W. and Shimek, M. G. (eds), *Statistical Theory and Computational Aspects of Smoothing*, Physica Verlag, Heidelberg.

Cleveland, W. S., Gross, E. and Shyu, W. M. 1992. 'Local regression models', in Chambers, J. M. and Hastie, T. J. (eds), *Statistical Models*, S. Wadsworth and Brooks/Cole Advanced Books and Software, California, pp. 309–376.

Cooper, B. E. 1969. *Statistics for Experimentalists*, Pergamon Press, Oxford, 336 pp.

Friedman, J. H. 1984. *A Variable Span Smoother*, Technical Report 5, Laboratory for Computational Sciences, Department of Statistics, Stanford University, CA.

Fuller, R. M., Sheail, J. and Barr, C. J. 1994. 'The Land of Britain, 1930–1990: a comparative study of field mapping and remote sensing techniques', *Geogr. J*., **160**(2), 173–184.

Green, S., Sanderson, F. J. and Marsh, T. J. 1996. 'Evidence for recent instability in rainfall and runoff patterns in the Celtic regions of western Europe', *Proceedings, Celtic Hydrology Symposium*, Rennes, July 1996, pp 73–83.

Grew, H. and Werrity, A. 1995. 'Changes in flood frequency and magnitude in Scotland', *BHS Fifth National Hydrology Symposium*, Edinburgh, pp. 3.1–3.9.

Good, P. 1994. *Permutation Tests, Springer Series in Statistics*, Springer-Verlag, Berlin, 229 pp.

Hardle, W. 1990. *Smoothing Techniques: with Implementation in S*, Springer-Verlag, Berlin, 261 pp.

Hastie, T. J. and Tibshirani, R. J. 1990. *Generalized Additive Models*, Chapman and Hall, London 335 pp.

Hollis, G. E. 1974. 'The effect of urbanization on floods in the Canon's Brook, Harlow, Essex', in Gregory, K. and Walling, D. E. (eds), *Fluvial Processes in Instrumented Watersheds*, Institute British Geographers, Special Publication No. 6, pp 123–140.

Hulme, M. 1996. *The 1996 CCIRG Scenario of Changing Climate and Sea-level for the United Kingdom*, Climatic Impacts LINK Technical Note 7, Climatic Research Unit, University of East Anglia, Norwich, 29 pp.

Jones, P. D. and Kelly, P. M. 1982. 'Principal component analysis of the Lamb catalogue of daily weather types. Part I: annual frequencies', *J. Climatol*., **2**, 147–159.

Karl, T. R. and Riebsame, W. E. 1989. 'Impact of decadal fluctuations in mean precipitation and temperature on runoff: a sensitivity study over the United States', *Climatic Change*, **15**(3), 423–447.

Kozuchowski, K. M. 1993. 'Variations of hemispheric zonal index since 1899 and its relationship with air temperature', *Int. J. Climatol.*, **13**, 853–864.

Lamb, H. H. 1972. 'British Isles weather types and a register of the daily sequence of circulation patterns', *Geophys. Mem.*, **16**, 1861–1971.

Latif, M. and Barnett, T. P. 1994. 'Causes of decadal climate variability over the North Pacific and North America', *Science*, **266**, 634–637.

Littel, R. C., Freund, R. J. and Spector, P. C. 1991. *SAS System for Linear Models*, 3rd edn, SAS Institute Inc., Cary, NC, 329 pp.

McCartney, L. 1996. 'An investigation of the upward trend in flood frequency at the White Cart Water at Hawkhead'. BSc final-year project, Department of Geography, University of Kingston, London.

McCullagh, P. and Nelder, J. A. 1983. *Generalized Linear Models*, Chapman and Hall, London, 261 pp.

Maritz, J. S. 1981. *Distribution-free Statistical Methods*. Chapman and Hall, London 264 pp.

Marsh, T. J. 1996. 'The 1995 UK drought—a sign of climatic instability?' *Proc. Inst. Civ. Eng.* Water, Marit. En., **118**, 189–195.

Mayes, J. 1996. 'Spatial and temporal fluctuations of monthly rainfall in the British Isles and variations in the mid-latitude westerly circulation', *Int. J. Climatol.*, **16**, 585–596.

Rasumsson, E. M. and Arkin, P. A. 1993. 'Global view of large-scale precipitation variability', *J. Climate*, **6**(8), 1495–1522.

Reed, D. W. and Stewart, E. J. 1994. 'Inter-site and inter-duration dependence on rainfall extremes', in Barnett, V. and Turkman, K. F. (eds), *Statistics for the Environment. 2: Water Related Issues*, Wiley, Chichester, pp. 125–143.

Reynard, N. S., Reed, D. W., Marsh, R. J. and Naden, P. S. 1996. 'Trends and changes in rainfall patterns, runoff and flow frequencies', in Samuels, P. (ed). *Implications of Climate Change for Flood and Coastal Defence Research*, Advisory Report to MAFF, Hydraulic Research, Wallingford.

Robson, A. J. and Neal, C. 1996. 'Water quality trends at an upland site in Wales, UK, 1983–1993', *Hydrol. Proc.*, **10**, 183–203.

Robson, A. J. and Reed, D. W. 1996. *Non-stationarity in UK Flood Series*, Flood Estimation Handbook Note 25, Report to MAFF, Project FD 0409, 37 pp. + appendices.

Siegel, S. and Castellan, N. J. 1988. *Non-parametric Statistics for the Behavioral Sciences*, McGraw-Hill, New York, 399 pp.

Sirios, A. 1993. 'Temporal variation of sulfate and nitrate concentration in precipitation in Eastern North America, 1979–1990', *Atmos. Environ.*, **27A**, 945–963.

Sprent, P. 1989. *Applied Non-parametric Statistical Methods*, Chapman and Hall, London, 259 pp.

Statistical Sciences 1995. *S-PLUS Guide to Statistical and Mathematical Analysis*, Version 3.3, Statsci, a division of MathSoft, Inc. Seattle.

UK Review Group on Acid Rain 1990. *Acid deposition in the UK, 1986–1990*, Third Report, Department of Environment, HMSO, London, 124 pp.

Venables, W. M. and Ripley, B. D. 1996. *Modern Applied Statistics with Splus*, Springer-Verlag, Berlin, 462 pp.

Walling, D. E. 1979. 'The hydrological impact of building activity: a study near Exeter', in Hollis, G. E. (ed.), *Man's Impact on the Hydrological Cycle in the United Kingdom*, Geo Abstracts Ltd, Norwich, pp. 135–151.

Ward, R. C. and Robinson, M. 1989. *Principles of Hydrology*, 3rd edn, McGraw-Hill, New York, 365 pp.

Woodley, 1996. 'A review of two national rainfall series', *Int. J. Climatol.*, **16**, 677–687.