An updated national-scale assessment of trends in UK peak river flow data: how robust are observed increases in flooding?

J. Hannaford, N. Mastrantonas, G. Vesuviano and S. Turner

ABSTRACT

A cluster of recent floods in the UK has prompted significant interest in the question of whether floods are becoming more frequent or severe over time. Many trend assessments have addressed this in recent decades, typically concluding that there is evidence for positive trends in flood magnitude at the national scale. However, trend testing is a contentious area, and the resilience of such conclusions must be tested rigorously. Here, we provide a comprehensive assessment of flood magnitude trends using the UK national flood dataset (NRFA Peak Flows). Importantly, we assess trends using this full dataset as well as a subset of near-natural catchments with high-quality flood data. While headline conclusions are useful for advancing national flood-risk policy, for on-the-ground flood-risk estimation it is important to unpack these local changes to determine how climate-driven trends compare with those from the wider dataset that are subject to a wide range of human disturbances and data limitations. We also examine the sensitivity of reported trends to changes in study time window using a 'multitemporal' analysis. We find that the headline claim of increased flooding generally holds up regionally to nationally, although we show a much more complicated picture of spatio-temporal variability. While some reported trends, such as increased flooding in northern and western Britain, appear to be robust, trends in other regions are more mixed spatially and temporally – for example, trends in recent decades are not necessarily representative of longer-term change, and within regions (e.g. in southeast England) increasing and decreasing trends can be found in close proximity. While headline conclusions are useful for advancing national flood-risk policy, for flood-risk estimation it is important to unpack these local changes, and the results and methodological toolkit provided here could provide such supporting information to practitioners.

Key words | flood, non-stationarity, peak flow, trend, variability

J. Hannaford (corresponding author)
G. Vesuviano
S. Turner
Water Resources and Hydro-climatic Risks,
UK Centre for Ecology and Hydrology,
Wallingford,
UK
E-mail: *jaha@ceh.ac.uk*

J. Hannaford
Irish Climate Analysis and Research Units,
Maynooth University,
Maynooth,
Ireland

N. Mastrantonas
Forecast Department, European Centre for Medium-Range Weather Forecast (ECMWF),
Reading,
UK
and
Faculty of Geosciences, Geoengineering and Mining,
Technische Universität Bergakademie Freiberg (TUBAF),
Freiberg,
Germany

HIGHLIGHTS

- Up-to-date national assessment of trends in flood magnitude for the UK.
- We examine long-term trends in a multitemporal context (sensitivity to time period).
- We compare sites at near-natural catchments with the wider network of disturbed sites.
- Results generally confirm the robustness of previously highlighted increases in flooding.
- However, we provide significantly greater spatial and temporal detail.

doi: 10.2166/nh.2021.156

INTRODUCTION

In early 2020, the UK experienced one of the most severe nationally significant flood events of recent decades (Parry *et al.* 2020; Sefton *et al.* in press). These floods came only 3 months after similarly devastating – and record-breaking – flooding in northern and central England (Muchan *et al.* 2019). The summer of 2019 also saw more localised, but dramatic, flooding in similar areas, notably Yorkshire and Lincolnshire.

The term 'unprecedented' has been widely used in connection with these flood events, but one does not have to look far back to find previous 'unprecedented' flooding events. For example, the winter of 2015–2016 saw record-breaking floods in northwest Britain: the highest daily rainfall for the UK and new peak river flow maxima for England were recorded on several rivers (Barker *et al.* 2016). In the winter of 2013–2014, sustained flooding affected large areas of the UK and caused severe impacts in southern England. While this event was less exceptional in terms of flood peaks, the duration and geographical extent were remarkable (Muchan *et al.* 2015).

The first two decades of the 21st century, in general, have been characterised by many major flood events (see also Hannaford 2015). Inevitably, the current floods – like all floods in recent years – have prompted widespread discussion about flood-risk management, and reasonable claims that floods are increasing due to human-induced global warming.

Certainly, future projections suggest an increase in flood severity in the UK (see Watts *et al.* (2015) and Reynard *et al.* (2017) for reviews). While there are wide uncertainty ranges, the majority of studies suggest increases in fluvial flood risk for much of the UK under anthropogenic warming scenarios. River flooding is already one of the most severe and most likely hazards on the UK's National Risk Register (Cabinet Office 2017), and the most recent Climate Change Risk Assessment (Committee on Climate Change 2016) presents compelling evidence that climate change may lead to significantly increased risks from fluvial flooding by mid-century.

Given such future projections, it is reasonable to assume that recent unprecedented flood events are evidence of the impacts of anthropogenic warming becoming manifest in flood regimes. Attribution studies, using large climate model ensembles, indicate that anthropogenic warming signatures can be detected in the 2013–2014 (Schaller *et al.* 2016) and 2015–2016 (Otto *et al.* 2018) floods. However, these anthropogenic effects are weak relative to natural variability (hence the need for very large climate ensembles), and it is a more open question as to whether they are detectable as long-term changes in flooding. To determine this, it is necessary to interrogate long-term records of observations (typically, records of annual maximum (AMAX) streamflow) to detect emerging trends in flooding as a necessary first step towards attribution. Trend detection and attribution are a vital foundation for adaptation, including flood design. Traditionally, flood-risk estimation methods assume stationarity – that is, a statistical process with parameters, for example, mean and variance, which do not shift over time (e.g. Slater *et al.* 2020). It is important to quantify any apparent non-stationarity in flood records to underpin the development of robust approaches to flood design that can incorporate observed changes to flood regimes over time.

This research need has motivated assessments of flood trends from the 1990s onwards, yielding a large body of previous work on non-stationarity in flooding in the UK. The literature up to the early 2010s was reviewed in detail by Hannaford (2015). Generally, evidence was found for increases in high flows in northern and western Britain, a pattern found by, for example, Hannaford & Marsh (2008), who used a Benchmark network of near-natural catchments to enable quantifcation of climate-driven trends. Since this review was published, several other studies have also quantified trends in flood series. Harrigan *et al.* (2018a) examined trends in the Benchmark network but focused only on high flow indicators based on daily river flows (e.g. the Q5 flow, the flow exceeded 5% of the time in any year) rather than AMAX peak flows. The AMAX data has been studied by Prosdocimi *et al.* (2014), Brady *et al.* (2019) and Prosdocimi *et al.* (2019), who also found national-scale evidence of increasing flood trends using techniques that allow spatially coherent trends to be characterised. European wide studies (e.g. Blöschl *et al.* 2019) have, further, demonstrated increasing flood magnitude in the UK is part of a much larger-scale pattern of increasing peak flows across northwest Europe.

More recently, there has been growing interest in non-stationary flood frequency estimation. Following the record-breaking 2015 floods, Spencer *et al.* (2018) applied non-stationary methods to Cumbria, yielding design flood estimates 15–25% higher than comparable stationary analyses. Faulkner *et al.* (2020a) generalised this to the national scale, finding that non-stationary analysis was necessary at around a third of the gauging network, increasing design flood estimates in many cases, with 11–30% increases being widespread. Further developments in non-stationary flood frequency estimation were advanced for England and Wales by Faulkner *et al.* (2020b) who found that a non-stationary model was preferred to a stationary models 22% of the time, and 36% of the time when physical covariates (e.g. catchment rainfall, climate indices such as the North Atlantic Oscillation, NAO) were added. Including non-stationarity typically made little difference to design flows, although occasionally led to substantial differences.

While there is a solid research base highlighting evidence of positive trends, there are several gaps in research. While there has been an increasing focus on identifying spatially coherent trends (following Prosdocimi *et al.* 2014 who cautioned against single-site analyses), trends at individual sites remain important. The local scale is the scale at which flood frequency analysis is performed (even if pooling methods are used) and at which the impacts of non-stationarity are felt. The present study was, therefore, motivated by the need to provide at-site estimates of non-stationarity to help guide flood-risk managers to make decisions about the necessity of using non-stationary methods. Here, we analyse trends for over 700 sites individually but, for brevity, we focus on the regional- to national-scale picture emerging from these at-site trends. Importantly, the analysis does not pool across sites statistically to model spatially coherent trends (cf. Prosdocimi *et al.* 2019). That is, our results are site-based but for presentation summarised on national maps and via regional averages. The individual at-site results are important in themselves and can be explored in detail in the form of a 'trend explorer' website (Griffin 2020).

A key focus of this study is testing the resilience of the reported headline message of positive trends in flooding. Trend detection is a contentious area and the barriers to observation-based trend analyses are widely reported, e.g.: (i) the low signal-to-noise ratios commonly seen in hydrological datasets (Wilby 2006); (ii) the confounding effect of human disturbances on river flow regimes (Whitfield *et al.* 2012); and (iii) sensitivity to chosen study period (Svensson *et al.* 2006; Merz *et al.* 2012). A recent review (Slater *et al.* 2020) outlines these and many other issues associated with trend detection.

Addressing (ii) above, separating climate-driven changes from direct human influences (urbanisation, dam construction, major abstractions) is a major challenge. Some past studies have focused on climate-driven trends, using the UK Benchmark Network (UKBN), but more typically (e.g. Robson *et al.* 1998; Prosdocimi *et al.* 2019), national assessments have lumped all sites together, including rivers with a wide range of anthropogenic impacts. Here, we contrast both approaches. We apply trend analysis to the Benchmark network and then put these results in the context of trends in the wider UK National River Flow Archive (NRFA) dataset containing a diverse range of impacted and non-impacted flood regimes. Studies in the USA have shown substantial differences between 'reference' networks (analogous to the UKBN) and sites influenced by urbanisation and reservoirs (Hodgkins *et al.* 2019). Addressing (iii), previous UK research has mostly focused on trends in fixed periods – either the whole record-of-record or some standard period within to enable fair comparison between sites. However, it is widely known that interdecadal variability can cause apparent trends (Robson *et al.* 1998; Hannaford *et al.* 2013). Prosdocimi *et al.* (2019) and Griffin *et al.* (2019) examined the sensitivity of flood frequency estimates to moving windows, finding strong evidence of sensitivity to the chosen period. Prosdocimi *et al.* (2019) generally found robust evidence of increasing flood trends, despite progressively shortening the reference period. However, this was generally for shorter (post-1976) periods. Here, we quantify sensitivity to study period over long timescales (in some cases, where record lengths allow, back to the 1920s) using multitemporal trend testing methods (Hannaford *et al.* 2013) that have not routinely applied to flood data in the UK. Crucially, in a departure from previous work, this considers *all* possible study periods, by varying start and end years of analysis – thereby providing a much broader context for trends in any fixed period.

In summary, we address the following research questions.

1. What is the evidence for non-stationarity in UK flooding at the national to regional scales and how do patterns of trend vary across the country?
2. How do near-natural Benchmark catchments compare to the wider network including catchments with human disturbances?
3. Are trends sensitive to the study period used, and how much difference do variations in the study period make to the overall conclusions derived in (1)?

METHODOLOGY

The methodology we adopt is based on the standard NRFA trend testing approach outlined by Harrigan *et al.* (2018a). In brief, the methodology is as follows. We apply monotonic trend tests to the UK-wide floods dataset and examine at-site trends and spatial patterns using several fixed study periods. We also subset this dataset according to the membership of the Benchmark network and the membership of standard hydrometric regions. We then examine sensitivity to the study period by using a 'multitemporal' analysis that quantifies trends between all possible start and end years in a record. We apply this to the regional groupings of stations and to a selection of very long (>70 years) hydrometric records to provide context for the recent, fixed study periods. The following sections detail this process.

Station selection criteria

To understand the long-term changes in UK flooding, the primary dataset used is the NRFA Peak Flow Dataset Version 8, released in September 2019 (NRFA 2020). The data are used with the statistical flood estimation methods set out in the Flood Estimation Handbook (FEH), which is the basis for current flood estimation in the UK. From this dataset, we use the AMAX archive, containing the largest observed instantaneous (i.e. based on the original 15-min stage data, rather than a daily maximum) flow in each water year.

The NRFA Peak Flow Dataset Version 8 consists of 935 gauging stations across the UK, of which 878 are classed as 'Suitable for Pooling' or 'Suitable for QMED' estimation according to the indicative suitability criteria applied in the dataset (NRFA 2020), and these were selected as the basis for analysis here. 'Suitable for Pooling' means either the highest AMAX flow or the 8-year event is likely to be within 30% of its true value, while 'suitable for QMED' means that the median AMAX (QMED) is likely to be within 30% of its true value.

The Peak Flow Dataset was then processed based on the following missing data criteria to mitigate the impacts of inevitable gaps in data:

(i) No more than 10% of missing data.
(ii) 27 or more years of data (≥27 AMAX).

This resulted in a dataset of 753 stations, which are mapped in Figure 1.

Two set periods (short and long) were chosen for trend analysis, following Harrigan *et al.* (2018a), optimising spatio-temporal distribution of stations, for better comparison of trends across the UK. A 31-year (short) period was selected for the calendar years 1987–2017 and a 51-year (long) period from the calendar years 1967–2017. However, a few years tolerance was made with the start and end years to increase the sample size (e.g. in reference to the short period, stations were able to begin in 1988 or 1989 if no data were available in 1987 and end in 2016 if no data were available for 2017). Stations were accepted for each period if at least 27 valid AMAX were available and 10% or less of AMAX values were missing during that period.

The resulting dataset gives good spatial coverage across the UK, although it is important to note the sparser coverage in Scotland (especially in the west) which simply reflects the currently available Peak Flows network in these areas.

The dataset was stratified in three ways to allow analyses to take place: (i) single sites, (ii) hydroclimatic regions and (iii) across hydrometric sub-networks, specifically the UKBN 2.0 (hereafter, UKBN2) (Harrigan *et al.* 2018a), also highlighted in Figure 1.

The regional classification was undertaken based on the nine hydroclimatic zones of Harrigan *et al.* (2018b). Results were summarised for all the stations in each region (Figure 1). Furthermore, for the multitemporal analysis described below, a regional median flow record was created as follows. Each AMAX in each record was standardised by subtracting the mean of the 1987–2017 AMAX values for that record and

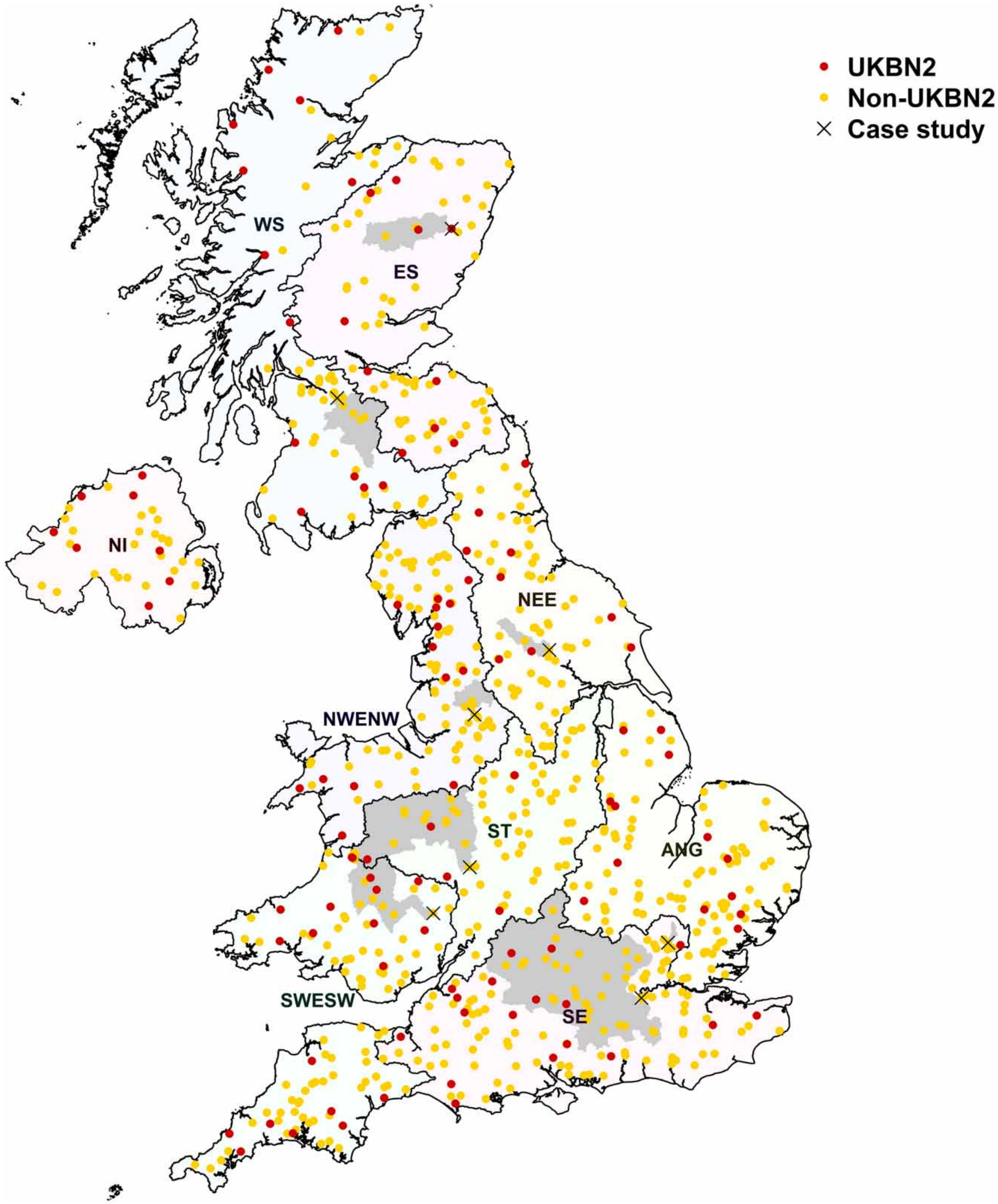

Figure 1 | Location map showing gauging stations of study catchments, both Benchmark Catchments (UKBN2) and the wider NRFA Peak Flows network. Location of the long hydrometric records also shown with shaded catchment areas. The hydrological regions (after Harrigan *et al.* 2018b) also shown.

dividing the result by the standard deviation of the 1987–2017 AMAX values for that record. Then, within each region, all records with at least 27 years of valid AMAX data and suitability for either pooling or QMED estimation were accepted. Records with more than 10% missing AMAX values were permitted to maximise the number of AMAX available to contribute to each year. For each year, the regional median AMAX was the median value of all standardised AMAX accepted for that year.

UKBN2 stations are classed as near-natural and with generally good quality data and, thus, are appropriate for identifying climate-driven hydrological trends. However, given the difficulties of finding stations of good quality across the full flow range, the network is stratified into several categories depending on hydrometric performance and artificial influences on the flood and low-flow regimes (Harrigan *et al.* 2018a). Thus, stations that received a score of 2 (suitable) or 1 (caution) for high flows were included. The latter are more likely to be subject to some degree of disturbance, but typically the degree of influence is not well known. There are 16 'caution' sites (compared to 98 'suitable'), and they were included to ensure good geographical coverage.

Trend analysis

Monotonic trends were assessed using the Mann–Kendall (MK) test (Mann 1945; Kendall 1975), a non-parametric rank-based approach that is widely supported for use in streamflow analysis (e.g. Hannaford & Marsh 2008; Murphy *et al.* 2013). The magnitude of trends was estimated using the robust Theil–Sen approach (Theil 1950; Sen 1968), with trend magnitude expressed as a percentage change compared to the long-term mean (the Theil–Sen average, TSA; Harrigan *et al.* 2018a).

The MK Z statistic (MKZ) follows the standard normal distribution with a mean of zero and a variance of one. A positive (or negative) value of MKZ indicates an increasing (or decreasing) trend. The probability of Type 1 errors set at the 5% significance level allowed the evaluation of statistical significance. A two-tailed MK test was chosen; hence, the null hypothesis of 'no trend present' (increasing or decreasing) is rejected when MKZ is outside ± 1.96 using traditional statistical testing.

The MK test requires data to be independent (i.e. free from serial correlation or temporal autocorrelation), as positive serial correlation increases the likelihood of Type 1 errors or incorrect rejection of a true null hypothesis (Kulkarni & von Storch 1995). All indicators were checked for positive lag-1 serial correlation at the 5% level using the autocorrelation function on detrended series. The linear trend used to detrend the original time series was estimated using the robust Theil–Sen estimator, which is also used for characterising trend magnitude.

Block bootstrapping was used to overcome the presence of serial correlation and involves the application of the MKZ statistic to block resampled series that preserve any short-term autocorrelation structure. Following guidance from Önöz & Bayazit (2012) regarding the optimal block length given the sample size and magnitude of temporal autocorrelation coefficient, a block length of 4 years was chosen and applied only when a series had statistically significant serial correlation – this occurred for 7,055 of the 231,245 time series analysed across all stations and periods used in this study. In these cases, a robust estimate of the significance of the MKZ statistic was generated from a distribution of 10,000 resamples, where the null hypothesis of no trend is rejected when MKZ calculated from original data are higher than the 9,750th largest (statistically significant increasing trend) MKZ value or lower than the 250th smallest (statistically significant decreasing trend) MKZ value from the resampled distribution under a two-tailed test at the 5% level (Murphy *et al.* 2013). While this provides a mechanism for addressing short-term serial correlation, it should be noted that in common with previous similar studies using these methods (Harrigan *et al.* 2018a), we do not address the issue of long-term persistence (e.g. Cohn & Lins 2005), which can significantly impact the interpretation of trends. This is unlikely to be a major issue in our study given the low number of positive serial correlation tests.

Decadal-scale variability and multitemporal analysis

As is widely noted in the literature (e.g. MacDonald & Sangster 2017), decadal-scale variability results in flood-rich and flood-poor periods that can affect the robustness of trends if fixed short temporal periods are selected. To illustrate decadal climate variability (DCV) in AMAX records, the standardised time series were smoothed using locally weighted regression (LOESS). Smoothing was achieved using a LOESS filter with a 15-year span (following Harrigan *et al.* 2018a). These time-series LOESS plots were grouped by hydroclimatic region (Figure 1), with plots starting from a 1961 cut-off as there is high variation in start dates and network density tails off significantly before this. For each region, individual LOESS plots are shown, as well as the regional median series.

To examine sensitivity to the study window, we adopt the approach of Hannaford *et al.* (2013) who argue that trends in any fixed period need to be put into a longer-term context, given the confounding role of decadal-scale hydrological variability that hampers the interpretation of linear trends. These authors advocate a multitemporal approach whereby trends are evaluated for all possible study periods, that is, varying the start and end year of the analysis and looking at the sensitivity of the results to such changes.

In the multitemporal approach, trends are calculated for all AMAX series for all possible start and end years (with a minimum period length of 27 years) for a total of 231,245 periods. These individual station-by-station multitemporal analyses are not reported in this paper given the sheer amount of data, but can be explored in Griffin (2020). For brevity in this paper, we show multitemporal analyses for each regional median series. Matrix plots are produced, showing start years along the x-axis and end years along the y-axis, where each cell corresponds to a single trend result, coloured according to the MKZ statistic. As with the LOESS plots, given the wide range of start dates, the multitemporal analyses were started in 1961.

Long hydrometric records

The multitemporal approach is even better suited to longer series to understand how representative the post-1960- and 1970 periods typically used in trend analysis are of much longer-term variability. As noted above, multitemporal analyses by a station for the full period of record are available for all sites used in this study, using these same graphics (see Griffin 2020).

To examine changes over a much longer period, nine of the longest available NRFA Peak Flow records were selected, with approximately one per region selected to give good spatial coverage (bearing in mind the low number of available sites with pre-1960 start dates in the dataset). These were selected by comparing the longest records in each region and appraising them for long-term consistency and quality, while still maintaining as long a record as possible. The selection is presented in Table 1. In some cases, records extend back many decades (generally to the 1920s), but the longest available records from western Scotland began in 1955 – in this case, little is added to the regional-scale multitemporal analysis, but it is included for completeness. For all plots used in this paper, for presentation purposes only the post-1920 period is shown, to avoid plots being dominated by whitespace, even though the full Thames record extends to 1882 and the Wye to 1908.

For the Dee and the Clyde, no AMAX data were available in the NRFA Peak Flows series from 2005 onwards, as they have yet to be updated. For these sites, AMAX were extracted from a separate source, the NRFA Highest Instantaneous Flows (HIFs). This dataset is based on the same source 15-min data from the UK measuring agencies, but has not undergone the rigorous QC as the NRFA Peak Flows data. In many cases, however, the recent data are identical, and a comparison of the long-term HIF and AMAX series was made to ensure that they were sensibly identical (within a few percent), before transplanting post-2005 HIFs to ensure that these series are as up-to-date as possible.

It should be noted that all long hydrometric stations inevitably have quality and homogeneity issues. The Dee is considered a Benchmark catchment, but most sites feature a range of human disturbances as well as some data homogeneity issues, as noted in the comment section of the table. The long records can therefore been seen as indicative of long-term flood variability, but they should be treated with caution.

RESULTS

Results per station

Of the full set of 753 stations, 587 met the criteria for the 'short' (1987–2017) period, and 334 met the criteria for the 'long' (1967–2017) period. MKZ scores ranged from −2.6 to +4.0 over the short period and from −3.1 to +4.5 over the long period. Overall, MKZ scores were slightly more likely to be positive over the long period than the short period: for the long period, 65.0% of MKZ scores were positive.

For the short period, 60.4% of MKZ scores were positive. The difference is even greater when considering the percentage of MKZ scores positive at 10% (5%) significance

Table 1 | Details of selected long hydrometric records, derived from the NRFA website (www.nrfa.ac.uk)

NRFA ID	River	Name	Area	Start year	End year	Catchment Area (km^2)	Mean elevation (m)	Base Flow Index[a] (BFI)	Standard-period annual average rainfall (SAAR)[b] (mm)	Comment on record homogeneity
12001	Dee	Woodend	ES	1929	2017	1,370	513.0	0.53	1,108	–
27001	Nidd	Hunsingore Weir	NEE	1933	2017	484.3	195.4	0.48	962	Trends calculated on digital record post-1966
38002	Ash	Mardock	ANG	1939	2017	78.7	93.7	0.53	619	Old station (pre-1979) subject to bypassing
39001	Thames	Kingston	SE	1882	2017	9,948	108.9	0.63	706	AMAX derived from naturalised series. Complex station history (see NRFA) but series used in long-term flood trend studies (Marsh & Harvey 2012)
54001	Severn	Bewdley	ST	1921	2017	4,325	175.5	0.53	912	–
55002	Wye	Belmont	SWESW	1908	2017	1,895.9	298.0	0.46	1,230	Flows prior to 1932 are considered unreliable but not rejected from Peak Flow Dataset
69025	Irwell	Manchester Racecourse	NWENW	1941	2017	557	213.6	0.62	1,260	1970 Flood Attenuation Storage works increased channel conveyance and flood storage basin built 2000
84005	Clyde	Blairston	WS	1955	2017	1,704.2	279.0	0.44	1,139	–

[a]The BFI is a measure of the proportion of the river runoff that derives from stored sources; the more permeable the rock, superficial deposits and soils in a catchment, the higher the baseflow and the more sustained the river's flow during periods of dry weather (https://nrfa.ceh.ac.uk/derived-flow-statistics).

[b]The average annual rainfall over the catchment for 1961–1990. This statistic is derived from the SAAR map for 1961–1990, a 1-km grid based on data from the Met Office, rather than a catchment rainfall series (https://nrfa.ceh.ac.uk/rainfall-statistics).

level: 20.4% (15.3%) for the long period, but 10.2% (6.5%) for the short period. This suggests a general UK-wide increase in flows over the last 50 years, but with modest numbers of significant trends, with more significance seen when looking over a longer period.

Spatially, positive MKZ scores are clustered in the north and south-west of England, north Wales and Northern Ireland for the short periods, and the north-west of a diagonal line running from south-west England (Dorset) to north-east England (roughly the North York Moors) for the long periods (Figure 2). Negative MKZ scores are clustered around south Wales, central England, London and west Scotland for the short periods, and London and eastern England for the long periods. These patterns mean that some areas, like south Wales and the tip of Cornwall in south-west England, show opposite trends for the long and short periods (in both cases, positive and negative, respectively). It should be noted that there is considerable spatial 'marbling' of positive and negative MKZ scores for all periods, where strongly positive values may occur in an area where the general trend is negative, and vice versa.

The relative slope of the Theil–Sen function ranged from −72 to +117% for the short period and −60 to +99% for the long period. TSA always followed the same sign as the MKZ score for each station, except in cases where one or the other was zero, and tended to be greater where the MKZ score was greater, with some exceptions where shallower relative slopes could be considered more significant than steeper relative slopes. The similarity between the sign of MKZ score and the sign of TSA meant that clusters of negative and positive TSA followed the same spatial patterns as positive and negative MKZ scores.

In addition to the fixed short and long periods, analyses were also performed for an arbitrary 'full' period-of-record. While this means the at-site results are less comparable in space, it does give a view of trends over the whole available period (as would be used by many if not most practitioners). For the full period, MKZ ranged from −2.9 to +5.9, and TSA

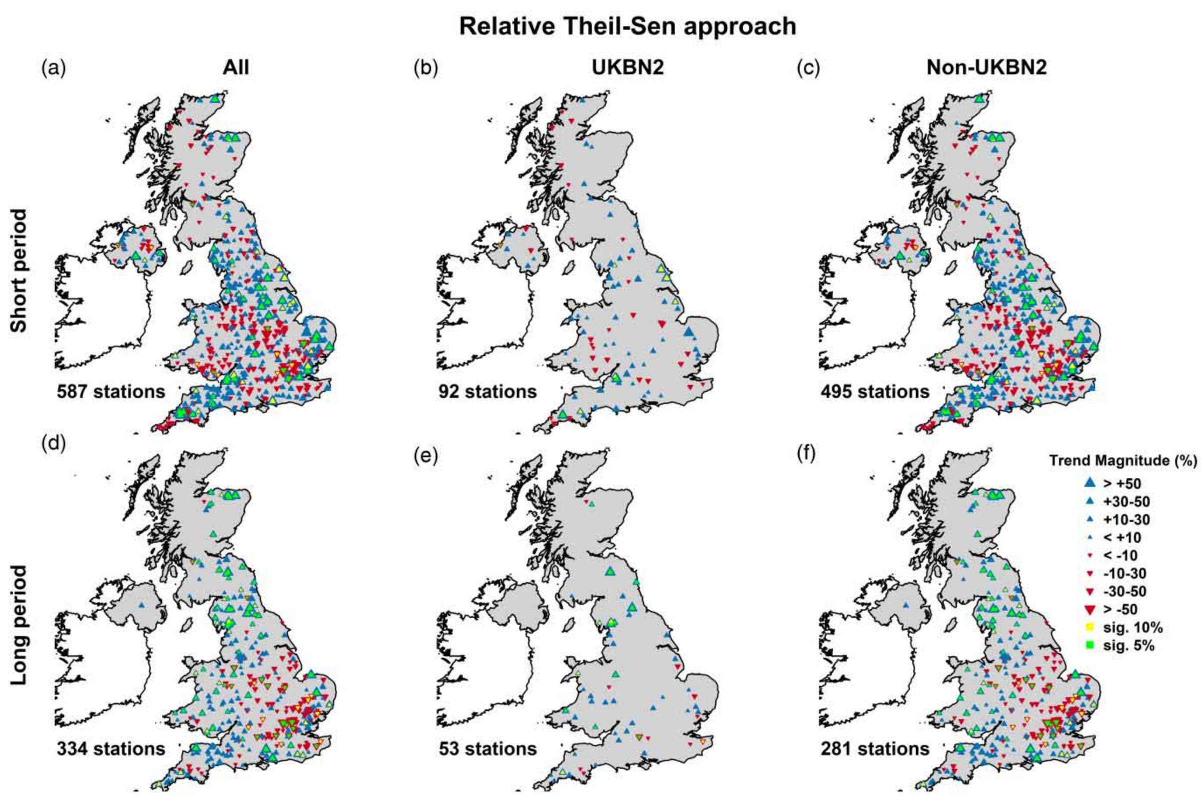

Figure 2 | Maps showing analysis results for the two fixed study periods (short and long) for three subsets: the whole dataset, those in the UKBN2 and those not in UKBN2. Symbols show the trend magnitude and significance according to the legend.

ranged from −85% to +107%. In total, 66.9% of trends were positive (15.9 and 22.2% at the 5 and 10% significance level), and 32.4% were negative (3.1 and 5.2% at the 5 and 10% significance level). Overall, spatial patterns in MKZ and TSA are similar for the full period (not shown) and long period, although positive TSA are generally higher for the full period. The use of a full period also greatly increases coverage in Northern Ireland. The broad similarity occurs because most 'full' periods start within ±10 years of 1967 and end in 2017.

Benchmark catchments vs. the wider network

The nature of UKBN2 can lead to apparent compromises when comparing to fuller datasets (i.e. the wider non-UKBN2) given the 'sparser' spatial coverage of UKBN2, which is inevitably a small subset of the UK network given the exacting requirements of UKBN2 status. Despite these difficulties, comparisons between UKBN2 and the full set can be drawn in the sense of spatial trend patterns (Figure 2). Similarities exist, for example, the positive trends apparent in areas of northern and western Britain that are present in both datasets to some extent. Conversely, important differences exist between the UKBN2 and full set. The significant decreasing trends prominent in the national dataset in south-east England are not present in the UKBN2 results. There are relatively few gauged catchments in the UK, which have near-natural flow regimes, and of these, even fewer are gauged by stations with the ability to measure the full range of flows accurately. This is particularly so with catchments in central, southern and eastern England, where dense populations inevitably lead to widespread artificial impacts. These negative south-eastern trends are not apparent in the UKBN2. This may suggest that the trend is not apparent in the (more) natural catchments in these areas, and large-scale influences on the non-UKBN2 catchments may be causing this trend. Alternatively, of course, it may be due to the lack of coverage of UKBN2 in this area. The implications of this are revisited in the discussion.

Results per region

A summary of trend analysis results broken down per region is presented in Tables 2 and 3. Considering regional records, 79.5% of MKZ scores are positive, 19.2% are negative and 1.3% are zero. In total, 24.0% (16.9%) are positive at 10% (5%) significance, and 0.02% (0%) are negative at 10% (5%) significance. The 'most positive' region is East Scotland (98.6% positive MKZ scores and no significant negative values), and the 'least positive' is Anglia (47.8% positive and 50.2% negative MKZ scores). Southeast England is the only region where no MKZ scores are significantly positive, and it is the only region where any MKZ scores are significant negative (one is at the 10% level). Anglia and Severn Trent are the two next least-extreme regions, with 1.0 and 0.6% significant positive (at 10%) MKZ scores and none significant at 5%.

Sensitivity to time window

For each full record, the maximum possible number of sub-records was produced, where each sub-record consisted of

Table 2 | Summary of trend results per regions – short period

Region	No. of records	Positive (%)	Sign. pos. at 10% (%)	Sign. pos. at 5% (%)	Negative (%)	Sign. neg. at 10% (%)	Sign. neg. at 5% (%)	Zero (%)
WS	22	36.36	4.55	4.55	63.64	4.55	4.55	0
ES	24	66.67	16.67	8.33	33.33	0	0	0
NEE	63	74.60	14.29	9.52	25.40	0	0	0
ST	64	46.88	6.25	6.25	50.00	1.56	1.56	3.12
ANG	99	62.63	10.10	5.05	34.34	2.02	1.01	3.03
SE	120	56.67	9.17	5.00	42.50	3.33	1.67	0.83
SWESW	87	49.43	11.49	8.05	47.13	3.45	1.15	3.45
NWENW	79	79.75	10.13	6.33	17.72	0	0	2.53
NI	30	60.00	10.00	6.67	40.00	6.67	3.33	0

Table 3 | Summary of trend results per regions – long period

Region	No. of records	Positive (%)	Sign. pos. at 10% (%)	Sign. pos. at 5% (%)	Negative (%)	Sign. neg. at 10% (%)	Sign. neg. at 5% (%)	Zero (%)
WS	9	77.78	22.22	22.22	22.22	11.11	11.11	0
ES	17	100.00	76.47	70.59	0	0	0	0
NEE	33	81.82	27.27	21.21	18.18	3.03	3.03	0
ST	29	48.28	13.79	10.34	51.72	13.79	10.34	0
ANG	75	41.33	5.33	2.67	57.33	6.67	1.33	1.33
SE	76	48.68	7.89	5.26	50.00	15.79	10.53	1.32
SWESW	52	86.54	23.08	17.31	13.46	0	0	0
NWENW	42	90.46	42.86	28.57	9.52	2.38	2.38	0
NI	1	100.00	0	0	0	0	0	0

all data from any one year to any later year present in the record. MKZ scores were calculated for every unique sub-record that contained at least 27 valid AMAX and no more than 10% missing AMAX values (sub-records differing only in the presence or absence of missing values at either the start or end were not considered unique, since they were identical after missing data were removed). In all, analyses were performed for 231,245 sub-records of 753 full records. Within these sub-records, there were considerably more positive MKZ scores (65.1%) than negative (33.6%) or zero (1.3%) scores. Furthermore, 18.3% (12.9%) of all MKZ scores were significantly positive at the 10% (5%) level compared with just 4.0% (2.2%) that were significantly negative at the same level. The percentage of positive MKZ scores at a station correlated somewhat with the MKZ score for the long periods but only weakly with the MKZ score for the short periods, further indicating that trend analyses on longer time series can be more informative of the expected behaviour of the station.

Plots of decadal-scale variability for each region show how the presence and timing of flood-rich and flood-poor periods can affect the robustness of the trends if fixed short temporal periods are selected (Figure 3). It should be noted that each of the nine hydroclimatic regions encompasses a different number of sites (ranging from 22 in WS to 119 in SE), and the agreement across sites in each region also differs. In some regions, there is close agreement across the sites (e.g. SE and ANG), whereas others are much more heterogeneous (e.g. WS and SWESW). Nevertheless, the decadal-scale variability appears to capture the main features regionally, highlighting flood-rich and flood-poor periods. There are generally increasing trends in northern regions, despite flood-poor periods present in the early 1970s and, in some regions, for example, eastern Scotland, a tailing off in the last decade. There is less variation across the southern and eastern regions, despite variation at some sites in the earlier periods, and here the trends appear to be more neutral. The decadal-scale variability indicates that flood-rich and flood-poor periods can affect the robustness of the trends depending on which temporal periods are selected. Taking the NEE region, for example, using the 'long' period suggests the positive trends are much more significant compared to the shorter, or equally the full time-series as the flood-rich and flood-poor periods across the period-of-record skew the apparent trends. Caution should be taken when describing these trends with special notice given to the period chosen.

The impact of such DCV on detected trends is demonstrated through the multitemporal plots in Figure 4. Significant trends are temporally clustered. For example, in West Scotland, almost all periods starting in or after 1975 have a negative trend, whereas almost all periods starting in or before 1972 have a positive trend. In Northern Ireland, negative trends are associated with a small range of start years (late 1970s) and end years (mid-2000s). In south-east England and Severn Trent, negative trends are associated with distinct periods of start and end years. In Anglia, negative trends are associated with a larger range of start years, but positive trends are associated with the most recent periods. East Scotland, north-east England and north-west

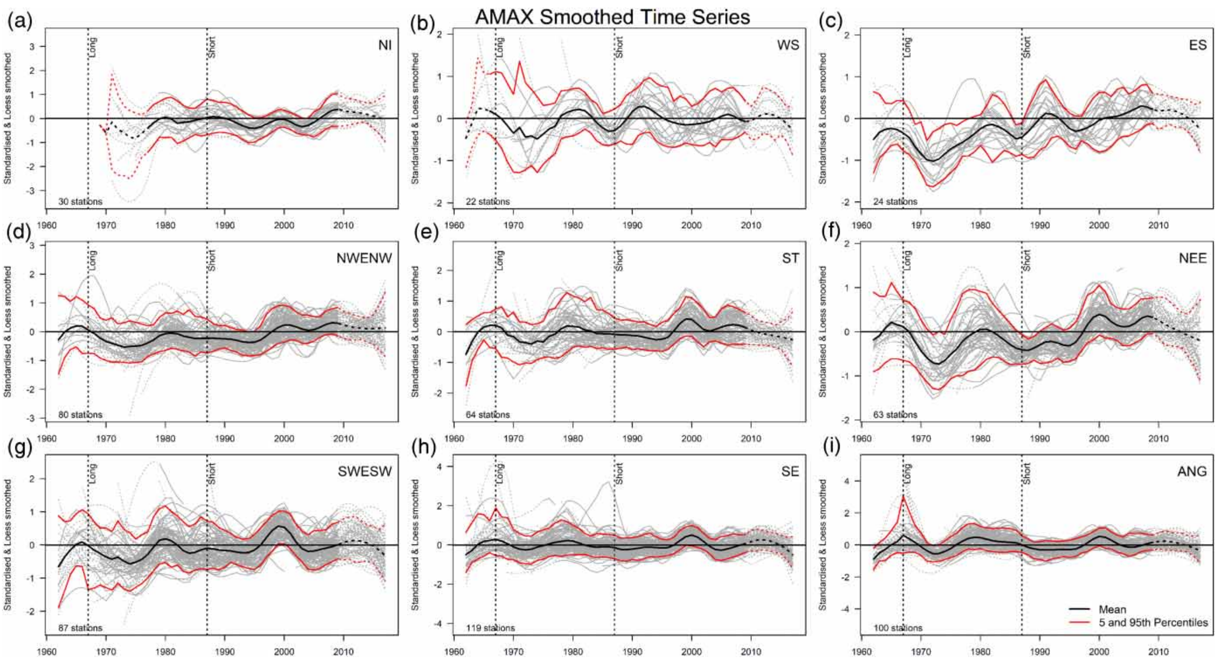

Figure 3 | Time-series DCV plots for each region, showing LOESS curves with a 15-year span. Each grey line corresponds to a standardised LOESS plot for each individual station, with the black line showing the regional median time series. Red lines denote 5th and 95th percentiles based on the individual sites. Please refer to the online version of this paper to see this figure in colour: http://dx.doi.org/10.2166/nh.2021.156.

England/North Wales show similar temporal patterns, with more significant positive trends for periods beginning around 1970 or in the early 1980s, and less significant positive trends for periods beginning between or after those two periods. South-west England/South Wales shows similar but less significant positive characteristics, with trends turning negative for many periods with a start year of 1979. Over the short period, MKZ scores range from −0.37 to +1.36. All are positive except Severn Trent and west Scotland (both negative). Over the long period, MKZ scores range from −0.44 to +3.48. All are positive, except Anglia (negative) and the south-east (zero).

Selected long records

The DCV in the long hydrometric records is shown in Figure 5. These plots confirm that the DCV that influences trends in the regional series post-1960 is also prevalent in the earlier decades. As noted, the 1970s are a flood-poor period across most series, whereas the late 1990s/early 2000s are generally flood-rich, which has an influence on the short- and long-period trends. Earlier variations are clear across records. The late 1950s and early 1960s were generally flood-poor, while the late 1940s and early 1950s were more flood-rich. These are generalisations; however, there is significant variability between rivers. The Wye shows a pronounced long-trend suggestive of heterogeneity in the flow record, although nonetheless, the oscillations superimposed on this trend are broadly in line with other sites.

The multidecadal plots for these long records are shown in Figure 6. The increasing (and often significant) trends in the more northern and western catchments are apparent, but only in the Wye are trends ending in recent (post-1990) years apparent across (most) start years, reflecting the upward trend in this record. The Nidd shows that significant trends ending in recent years extend back to the 1940s. Trends with earlier start and end dates were generally (non-significant) negative. The Dee shows a broadly similar pattern, except with significant negative trends earlier in the record. Moreover, trends ending in recent years are positive and significant if they started before 1960, but trends that start after that date show more evidence of decreases – this reflects the downturn in AMAX seen for

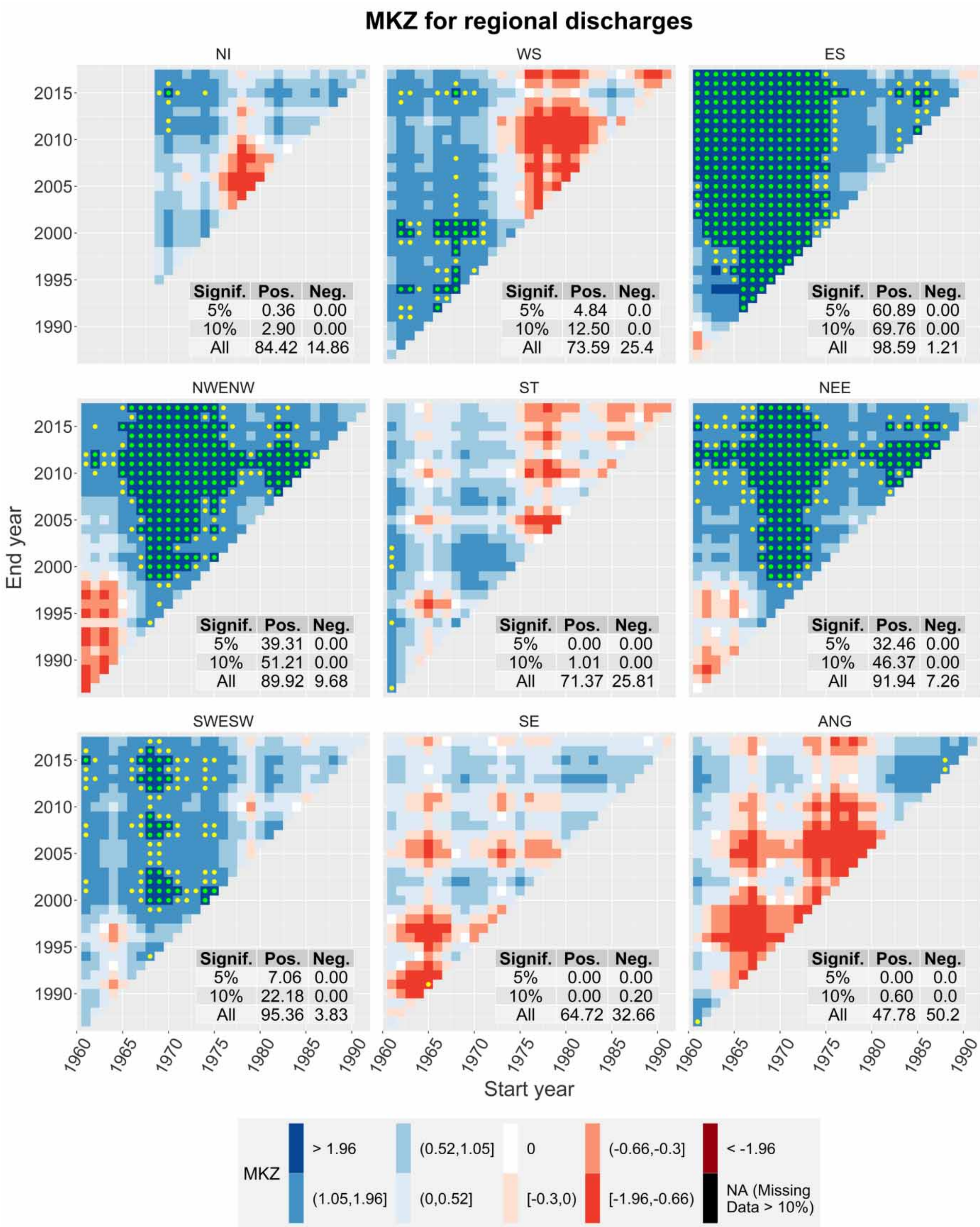

Figure 4 | Multitemporal trend plots for each region (applied to regional median series). Each cell in each matrix shows the result of a trend analysis applied to that combination of start (*x*-axis) and end (*y*-axis) years. The trend result is coloured according to the legend, with dots denoting significant trends at 5% (green) and 10% (yellow). Summary statistics are shown for the proportion of results from all cases (start and end years) in terms of positive/negative direction and either 5 or 10% significance. Please refer to the online version of this paper to see this figure in colour: http://dx.doi.org/10.2166/nh.2021.156.

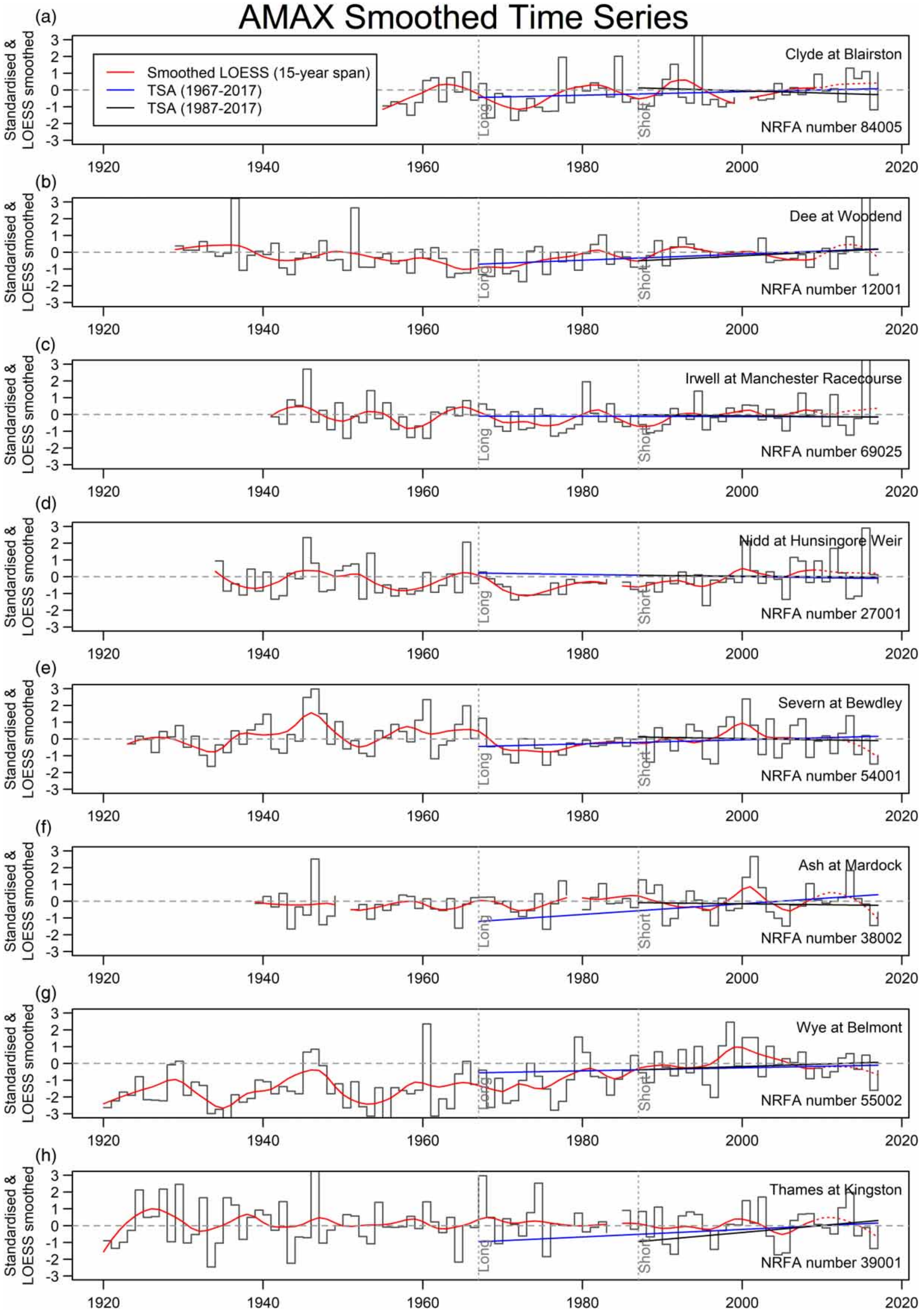

Figure 5 | DCV in the long hydrometric records. The black bars are the AMAX time series, the red line the LOESS series with a 15-year span. Blue and black lines show fixed trends over the long and short study periods. Please refer to the online version of this paper to see this figure in colour: http://dx.doi.org/10.2166/nh.2021.156.

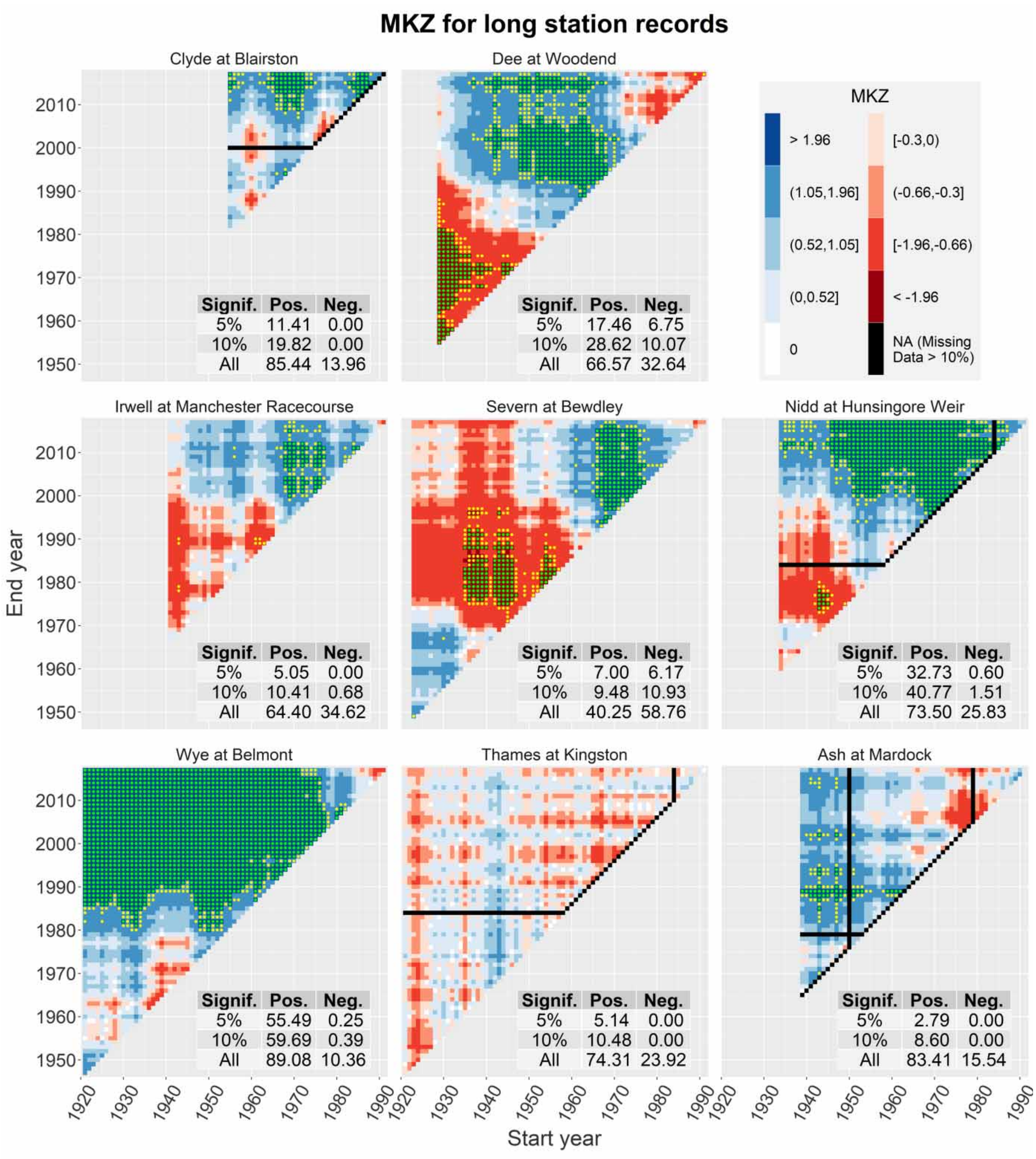

Figure 6 | Multitemporal trend plots for each long record, with the same description as Figure 4 but for longer periods – extending to the start of each record, or 1920 in the case of the Thames and the Wye. Summary statistics are shown for the proportion of results from all cases (start and end years) in terms of positive/negative direction and either 5 or 10% significance. Black lines represent the impact of periods of missing data.

the Dee in Figure 5. The Irwell shows similar variations between negative trends in earlier periods and positive trends later, but with fewer significant results. The results from the 'lowland' catchments in south-east England, the Ash and the Thames, are somewhat different from the other catchments. The Thames shows very mixed patterns of generally weak and non-significant trends through the series, with variations between positive and negative trends. The Ash is similarly mixed, although most trends were positive and occasionally significant (if starting earlier in the period and only for some end years).

DISCUSSION

The results of the present study accord with previous research, reaffirming generally positive trends as being the main outcomes for large areas of the UK, especially northern and western regions. In this sense, this study builds on the headline conclusions of the review of Hannaford (2015) and agrees with subsequent work on more updated flood records, as cited in the introduction to this paper. We also find that – in general terms – these positive trends are mostly resilient to changes in the study approach. In particular, a broadly similar message emerges from the full series of more 'noisy' anthropogenically influenced stations compared to those in the UKBN. Furthermore, changes to the time period of analysis – especially changes to the start date of analysis, up to the most recently available data – do not especially change the finding that there is more compelling evidence for an increase in flood magnitude in the UK than for a decrease or no change. These headline findings add to a growing evidence base that suggests that traditional flood frequency analysis approaches, which assume stationarity, may be called into question (e.g. Faulkner *et al.* 2020a, 2020b).

However, beneath this headline message, this study has examined trend responses for over 700 individual catchments and has examined sensitivity to time window by computing trends for over 200,000 possible start and end dates. Unsurprisingly, it reveals a much more complex picture of spatial and temporal variability in flood magnitude.

First, the national and regional picture is more nuanced than the 'increasing in north and west' headline. For the long fixed period, which represents the best trade-off between study period length and spatial coverage, there is generally an increasing trend across northern and western Britain, which accords well with previous studies (noting again the relative sparsity of gauges in western Scotland in this study). The number of significant trends is, however, relatively modest. There are also increasing trends in southern England and parts of eastern England that are often significant. This agrees with the findings of Faulkner *et al.* (2020a) who observed increases in similar areas, and Brady *et al.* (2019) and Prosdocimi *et al.* (2019) whose regional trend models also point to increases in parts of eastern England. Here, however, with the added detail of a site-by-site assessment, the regional picture in lowland England is shown to be very mixed, and there is evidence of decreasing flood trends (which are often significant) in much of central-eastern England.

Interestingly, the results from the Benchmark network correspond well with the full dataset of trends subjected to anthropogenic influences and hydrometric data constraints, suggesting at face value that the overall patterns in the wider dataset are mostly climate-driven. Of course, this is at the gross regional-to-national scale and, at any individual site, anthropogenic influences, like the effect of impoundments, or data artefacts, such as changes to ratings, may be generating spurious trends in any one station within the wider dataset. Nevertheless, for such big picture assessments, the full dataset leads to similar conclusions. Turning this argument on its head, however, the Benchmark network (110 sites suitable for high-flow analysis) can also be seen to be of limited representativeness of the full range of spatial variability in trends. There is little evidence of any negative trends in eastern England in the Benchmark dataset. At face value, this may suggest that such trends are human rather than climate-driven, but the spatial coherence of these negative trends in the full Peak Flows dataset implies that it is more likely that they are a climate-driven trend that is simply not captured in the UKBN2 dataset due to the lack of catchments in this 'hotspot'. Caution is therefore needed in extrapolating from such 'climate-sensitive' reference hydrometric networks (RHNs), and this underscores the benefit of looking at both RHNs and the wider network, despite the confounding influences in the latter. Studies from other countries with RHNs, notably in North America, have made similar arguments (Burn & Whitfield 2016; Hodgkins *et al.* 2019) although typically using the mismatch to underscore the importance of RHNs. This argument is probably more valid in North America where there is a high number of genuinely 'pristine' catchments to incorporate into RHNs, but is more challenging in the UK where human influences are so prevalent. We conclude that the best approach is to assess trends in both, and divergences between the two warrant further investigation, focusing on the attribution of observed changes – we return to this later in this commentary.

The multitemporal approach also reveals significant complexity and demonstrates that these headline maps of spatial changes for the full period, or static time slices, like the short and long periods, could look different if study periods were shifted by a few years: the results from a standard period analysis make up only one pixel in each of the multitemporal plots, which show large variations in trend magnitude, direction and significance. The increasing trends in northern and western Britain are mostly resilient to changes in the study period, at least with starting dates from the early 1970s onwards, in terms of the direction of trends. However, the analysis shows that the strength and significance of trends varies considerably. Importantly, the multitemporal approach shows that the decreasing trends in eastern England discussed above are largely a function of the chosen study period. In most of the English lowlands of central, southern and eastern England, there is little compelling evidence of long-term trends, but positive and negative (and mostly non-significant) trends throughout the series. Combined with the spatial heterogeneity (positive and negative trends on the maps in relatively close proximity) in these regions, this suggests that the few significant trends are unlikely to be either spatially or temporally representative.

The longer records also show significant variation over many decades. Again, multitemporal analyses reveal that for the more northern/western rivers, recent increases in flood magnitude can be traced back to start dates before the 1960s. However, the strength of trends varies significantly over time, and decreasing trends were witnessed in early periods. This does little to 'disprove' any climate change impact per se (as anthropogenic warming impacts would be expected to be most prevalent in recent decades) but further points to the challenge of identifying robust long-term trends against a backdrop of pronounced multidecadal variability.

The prevalence of positive trends, especially in northern and western Britain, suggests that the conclusions of past studies, of increasing trends up to the early 2000s, remain valid through to the late 2010s. The question of what is driving these coherent changes is an important one. Positive trends in rainfall and river flow have previously been identified (see, e.g. Hannaford (2015) and Spencer *et al.* (2018) and references therein) and attributed to variability in large-scale atmospheric–oceanic drivers, most notably the NAO which is the leading mode of variability in the region and Atlantic Multidecadal Oscillation (AMO), a much lower frequency mode of sea-surface temperature variability which has also been shown to influence decadal patterns of trends in flooding in Western Europe (Hodgkins *et al.* 2017). The variations in DCV shown here, in the regional plots, appear consistent with AMO variability – a flood-rich period in the 1990s/early 2000s, with a flood-poor period in the late 1960s/early 1970s. Longer records like the Dee at Woodend also show earlier flood-rich periods in the 1930s, which is also consistent with AMO variability (Hodgkins *et al.* 2017). Other studies have demonstrated flood-rich and flood-poor periods influenced by these and other atmospheric–oceanic drivers over several centuries (MacDonald & Sangster 2017).

For practitioners, this presents a question: are observed increases in flooding a short-term variation or longer-term climate change? Put simply, will trends increase at similar rates in future under anthropogenic warming, or will trend magnitudes change (or even reverse) as a result of the AMO or other modes of DCV? Current trends are driven by a combination of both anthropogenic warming and DCV, but while the thermodynamic component behind increased flooding is likely to continue in a warming world, future changes could be modulated by circulation-driven changes in the AMO or other modes of DCV.

To conclude, we echo previous calls (Merz *et al.* 2012) for improved attribution of flood changes. Event-based attribution (Schaller *et al.* 2016; Otto *et al.* 2018) adds to our confidence that flood trends have some anthropogenic component, but there is a need to generalise this to long-term trends if such attribution is to support flood-risk estimation. There are two levels to this. First, separating climate-driven trends from human interventions. While we have shown little difference here between the UKBN and impacted sites at the large-scale, the picture is more complicated in individual regions and catchments. More comprehensive trend detection and attribution studies are needed to examine the relative contribution of climate drivers relative to catchment-scale changes (e.g. urbanisation and land cover/land management changes) that have been uncovered in some localised cases but rarely at larger scales (e.g. Rust *et al.* 2014; Prosdocimi *et al.* 2015). Secondly, for 'climate driven' trends, there remains a need to separate anthropogenic forcing from DCV, which is challenging as the two are intrinsically linked, with anthropogenic warming influencing natural patterns of large-scale circulation, and vice versa (e.g. Deser *et al.* 2017). As this remains a long-term challenge, one pragmatic approach from a flood-risk estimation perspective is to build large-scale climate predictors in non-stationary flood frequency analysis (e.g. Steriou *et al.* 2019; Faulkner *et al.* 2020b). Such approaches have met with some success, although single descriptor dipole patterns like the NAO are a blunt instrument, and a range of studies are demonstrating an influence on UK rainfall and river flows of sea-surface temperature and pressure variations on an Atlantic (Lavers *et al.* 2013; Barnes *et al.* In review) or even global (Svensson & Hannaford 2019) scale. More work is needed to understand the atmosphere-ocean mechanisms that drive flood variability in the UK, on a range of timescales, to support such enhanced flood-risk estimation approaches.

CONCLUDING REMARKS

This paper has provided an up-to-date assessment of flood trends at the national scale. Our results are comparable with previous studies, but we demonstrate the resilience of these findings to important methodological considerations. However, we also show significant granularity in the regional and national picture and sensitivity to chosen study periods. To this end, we add a considerable value for flood practitioners who must balance local-scale information with this wider national picture. Given the variation in trend responses, we recommend that trend analysis should be undertaken in catchments of interest as a part of flood frequency estimation studies. We provide the outputs of this study in an accessible format and in an interactive tool that allows closer appraisal (Griffin 2020). However, significant obstacles to application remain, not least around the perennial question of attribution of observed changes.

ACKNOWLEDGEMENTS

We thank the UK National River Flow Archive and the UK measuring agencies for their stewardship of the national Peak Flows dataset and supporting data and metadata. We thank the following people for valuable review and feedback: Sean Longfield, Peter Spencer (Environment Agency), Duncan Faulkner (JBA Consulting) and Adam Griffin (UKCEH). We thank Neil MacDonald (U. Liverpool) and two anonymous referees for their reviews, which led to significant improvements to this paper.

AUTHOR CONTRIBUTIONS

J.H. designed and led the study. N.M., G.V. and S.T. conducted the statistical analyses and produced the figures. J.H. led the preparation of the manuscript, with contributions from all authors.

FUNDING

This work was primarily supported by the UK Natural Environment Research Council (award no. NE/R016429/1) as part of the UK-SCAPE programme delivering National Capability (Task 1.6 'Water', Analysis and Exploitation). Further support was provided by the Environment Agency Flood and Coastal Erosion Risk Management Research and Development Programme, within the project: 'Developing interim guidance for flood frequency estimations (no. FRS18087/IG/R1).

DATA AVAILABILITY STATEMENT

All the data used in this study are freely available from the UK National River Flow Archive (NRFA). https://nrfa.ceh.ac.uk/.

REFERENCES

Barker, L., Hannaford, J., Muchan, K., Turner, S. & Parry, S. 2016 The winter 2015/2016 floods in the UK: a hydrological appraisal. *Weather* **71** (12), 324–333.

Barnes, A., Svensson, C. & Kjeldsen, T. In review *North Atlantic Air Pressure and Temperature Conditions Associated with Heavy Rainfall in Great Britain*.

Blöschl, G., Hall, J., Viglione, A., Perdigão, R. A. P., Parajka, J., Merz, B. & Živković, N. 2019 Changing climate both increases and decreases European river foods. *Nature* **573** (7772), 108–111. https://doi.org/10.1038/s41586-019-1495-6.

Brady, A., Faraway, J. & Prosdocimi, I. 2019 Attribution of long-term changes in peak river flows in Great Britain. *Hydrological Sciences Journal* **64** (10), 1159–1170. https://doi.org/10.1080/02626667.2019.1628964.

Burn, D. H. & Whitfield, P. H. 2016 Changes in floods and flood regimes in Canada. *Canadian Water Resources Journal* **41**, 139–150.

Cabinet Office 2017 *National Risk Register of Civil Emergencies – 2017 Edition*. Cabinet Office, London.

Cohn, T. A. & Lins, H. 2005 Nature's style: naturally trendy. *Geophysical Research Letter* **23**, L23402.

Committee on Climate Change 2016 *UK Climate Change Risk Assessment 2017 Evidence Report*. Available from: www.theccc.org.uk/UK-climate-change-risk-assessment-2017/ (accessed 7 April 2021).

Deser, C., Hurrell, J. W. & Phillips, A. S. 2017 The role of the North Atlantic Oscillation in European climate projections. *Climate Dynamics* **49**, 3141–3157.

Faulkner, D., Warren, S., Spencer, P. & Sharkey, P. 2020a Can we still predict the future from the past? Implementing non-stationary flood frequency analysis in the UK. *Journal of Flood Risk Management* **13**, e12582. https://doi.org/10.1111/jfr3.12582.

Faulkner, D., Griffin, A., Hannaford, J., Sharkey, P., Warren, S. & Shelton, K. 2020b *Development of Interim National Guidance on Non-stationary Fluvial Flood Frequency Estimation*. Science Report FRS18087/IG/R1. Environment Agency, Bristol.

Griffin, A. 2020 Trend explorer. Available from: https://shiny-apps.ceh.ac.uk/river-flow-trends-explorer/.

Griffin, A., Vesuviano, G. & Stewart, E. 2019 Have trends changed over time? A study of UK peak flow data and sensitivity to observation period. *Natural Hazards and Earth System Sciences* **19**, 2157–2167. doi:10.5194/nhess-19-2157-2019.

Hannaford, J. 2015 Climate-driven changes in UK river flows: a review of the evidence. *Progress in Physical Geography* **39** (1), 29–48. https://doi.org/10.1177/0309133314536755.

Hannaford, J. & Marsh, T. J. 2008 High-flow and flood trends in a network of undisturbed catchments in the UK. *International Journal of Climatology* **28** (10), 1325–1338. https://doi.org/10.1002/joc.1643.

Hannaford, J., Buys, G., Stahl, K. & Tallaksen, L. M. 2013 The influence of decadal scale variability on trends in long European streamflow records. *Hydrology and Earth System Sciences* **17**, 2717–2733. https://doi.org/10.5194/hess 17 2717 2013.

Harrigan, S., Hannaford, J., Muchan, K. & Marsh, T. J. 2018a Designation and trend analysis of UKBN2. *Hydrology Research* **49** (2), 552–567. https://doi.org/10.2166/nh.2017.058.

Harrigan, S., Prudhomme, C., Parry, S., Smith, K. & Tanguy, M. 2018b Benchmarking ensemble streamflow prediction skill in the UK. *Hydrology and Earth System Sciences* **22** (3). https://doi.org/10.5194/hess-22-2023-2018.

Hodgkins, G., Whitfield, P., Burn, D., Hannaford, J., Renard, B., Stahl, K., Fleig, A., Madsen, H., Mediero, L., Korhonen, J., Murphy, C., Crochet, P. & Wilson, D. 2017 Climate-driven variability in the occurrence of major floods in North America and Europe. *Journal of Hydrology* **552**, 704–717.

Hodgkins, G. A., Dudley, R. W., Archfield, S. A. & Renard, B. 2019 Effects of climate, regulation, and urbanization on historical flood trends in the United States. *Journal of Hydrology* **573**, 697–709.

Kendall, M. G. 1975 *Rank Correlation Methods*, 4th edn. Charles Griffin, London, UK.

Kulkarni, A. & von Storch, H. 1995 Monte Carlo experiments on the effect of serial correlation on the Mann-Kendall test of trend. *Meteorologische Zeitschrift* **4** (2), 82–85.

Lavers, D., Prudhomme, C. & Hannah, D. 2013 European precipitation connections with large-scale mean sea-level pressure (MSLP) fields. *Hydrological Sciences Journal* **58** (3), 310–327.

MacDonald, N. & Sangster, H. 2017 High magnitude flooding across Britain since AD1750. *Hydrology and Earth System Sciences* **21**, 1631–1650. https://doi.org/10.5194/hess-21-1631-2017.

Mann, H. B. 1945 Nonparametric tests against trend. *Econometrica* **13** (3), 245–259.

Marsh, T. J. & Harvey, C. 2012 The Thames flood series: a lack of trend in flood magnitude and a decline in maximum levels. *Hydrology Research* **43** (3), 203–214. https://doi.org/10.2166/nh.2012.054.

Merz, B., Vorogushyn, S., Uhlemann, S., Delgado, J. & Hundecha, Y. 2012 HESS opinions 'more efforts and scientific rigour are needed to attribute trends in flood time series'. *Hydrology and Earth System Sciences* **16**, 1379–1387. https://doi.org/10.5194/hess-16-1379-2012.

Muchan, K., Lewis, M., Hannaford, J. & Parry, S. 2015 The winter storms of 2013/2014 in the UK: hydrological responses and impacts. *Weather* **70** (2), 55–61.

Muchan, K., Turner, S., Hannaford, J., Downes, E., Griffin, A., Barker, L. & Matthews, B. 2019 *Briefing Note. Severity of the November 2019 Floods – Preliminary Analysis*. UK Centre for Ecology & Hydrology, Wallingford, UK.

Murphy, C., Harrigan, S., Hall, J. & Wilby, R. L. 2013 Climate-driven trends in mean and high flows from a network of reference stations in Ireland. *Hydrological Sciences Journal* **58** (4), 755–772.

National River Flow Archive 2020 *Peak Flows Dataset V8*. Available from: https://nrfa.ceh.ac.uk/peak-flow-dataset (accessed 25 March 2020).

Önöz, B. & Bayazit, M. 2012 Block bootstrap for Mann–Kendall trend test of serially dependent data. *Hydrological Processes* **26** (23), 3552–3560.

Otto, F., van der Wiel, K., van Oldenborgh, G., Philip, S., Kew, S., Uhe, P. & Cullen, H. 2018 Climate change increases the probability of heavy rains in Northern England/Southern Scotland like those of storm Desmond – a real-time event attribution revisited. *Environmental Research Letters* **13**, 024006.

Parry, S., Barker, L., Sefton, C., Hannaford, J., Turner, S., Muchan, K., Matthews, B. & Pennington, C. 2020 *Briefing Note: Severity of the February 2020 Floods – Preliminary Analysis*. UK Centre for Ecology & Hydrology, Wallingford, UK.

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

Prosdocimi, I., Kjeldsen, T. R. & Miller, J. D. 2015 Detection and attribution of urbanization effect on flood extremes using nonstationary flood-frequency models. *Water Resources Research* **51**, 4244–4262. doi:10.1002/2015WR017065.

Prosdocimi, I., Dupont, E., Augustin, N., Kjeldsen, T. R., Simpson, D. & Smith, T. 2019 Areal models for spatially coherent trend detection: the case of British peak river flows. *Geophysical Research Letters* **46**, 13054–13061. https://doi.org/10.1029/2019GL085142.

Reynard, N. S., Kay, A. L., Anderson, M., Donovan, B. & Duckworth, C. 2017 The evolution of climate change guidance for fluvial flood risk management in England. *Progress in Physical Geography* **41** (2), 222–237.

Robson, A. J., Jones, T. K., Reed, D. W. & Bayliss, A. C. 1998 A study of national trend and variation in UK floods. *International Journal of Climatology* **18**, 165–182. doi:10.1002/(SICI)1097-0088(199802)18:2 < 165::AID-JOC230 > 3.0.CO;2-#.

Rust, W., Corstanje, R., Holman, I. P. & Milne, A. E. 2014 Detecting land use and land management influences on catchment hydrology by modelling and wavelets. *Journal of Hydrology* **517**, 378–389. http://dx.doi.org/10.1016/j.jhydrol.2014.05.052.

Schaller, N., Kay, A. L., Lamb, R., Massey, N. R., van Oldenborgh, G. J., Otto, F. E. L., Sparrow, S. N., Vautard, R., Yiou, P., Ashpole, I., Bowery, A., Crooks, S. M., Haustein, K., Huntingford, C., Ingram, W. J., Jones, R. G., Legg, T., Miller, J., Skeggs, J., Wallom, D., Weisheimer, A., Wilson, S., Stott, P. A. & Allen, M. R. 2016 Human influence on climate in the 2014 southern England winter floods and their impacts. *Nature Climate Change* **6** (6), 627–634. https://doi.org/10.1038/nclimate2927.

Sefton, C., Muchan, K. & Parry. In press The 2019–2020 UK floods. For: Weather. doi:10.1002/wea.3993.

Sen, P. K. 1968 Estimates of the regression coefficient based on Kendall's tau. *Journal of the American Statistical Association* **63** (324), 1379–1389.

Slater, L. J., Anderson, B., Buechel, M., Dadson, S., Han, S., Harrigan, S., Kelder, T., Kowal, K., Lees, T., Matthews, T., Murphy, C. & Wilby, R. L. 2020 *Nonstationary Weather and Water Extremes: A Review of Methods for Their Detection, Attribution, Management*. Hydrology and Earth System Sciences Discussions. Available from: https://hess.copernicus.org/preprints/hess-2020-576/ (accessed 7 April 2021).

Spencer, P., Faulkner, D., Perkins, I., Lindsay, D., Dixon, G., Parkes, M., Lowe, A., Asadullah, A., Hearn, K., Gaffney, L., Parkes, A. & James, R. 2018 The floods of December 2015 in Northern England: description of the events and possible implications for flood hydrology in the UK. *Hydrology Research* **49** (2), 568–596. https://doi.org/10.2166/nh.2017.092.

Steirou, E., Gerlitz, L., Apel, H., Sun, X. & Merz, B. 2019 Climate influences on flood probabilities across Europe. *Hydrological Earth System Sciences* **23**, 1305–1322. https://doi.org/10.5194/hess-23-1305-2019.

Svensson, C. & Hannaford, J. 2019 Oceanic conditions associated with Euro-Atlantic high pressure and UK drought. *Environmental Research Communications* **1** (10). https://doi.org/10.1088/2515-7620/ab42f7.

Svensson, C., Hannaford, J., Kundzewicz, W. & Marsh, T. 2006 Trends in river floods: why is there no clear signal in observations? In: *Frontiers in Flood Research (Proceedings of Kovacs Colloquium, Paris, June 2006)*. IAHS Publ. 305, 12 pages.

Theil, H. 1950 A rank-invariant method of linear and polynomial regression analysis. *Proceedings of the Royal Netherlands Academy of Sciences* **53**, 386–392.

Watts, G., Battarbee, R. W., Bloomfield, J. P., Crossman, J., Daccache, A., Durance, I., Elliott, J. A., Garner, G., Hannaford, J., Hannah, D. M., Hess, T., Jackson, C. R., Kay, A. L., Kernan, M., Knox, J., Mackay, J., Monteith, D. T., Ormerod, S. J., Rance, J., Stuart, M. E., Wade, A. J., Wade, S. D., Weatherhead, K., Whitehead, P. G. & Wilby, R. L. 2015 Climate change and water in the UK – past changes and future prospects. *Progress in Physical Geography: Earth and Environment* **39** (1), 6–28.

Whitfield, P., Burn, D., Hannaford, J., Higgins, H., Hodgkins, G. A., Marsh, T. J., Looser, U. & Hydrologic Reference Networks I. 2012 The status of national reference hydrologic networks for detecting trends and future directions. *Hydrological Sciences Journal* **57** (8), 1562–1579.

Wilby, R. L. 2006 When and where might climate change be detectable in UK river flows? *Geophysical Research Letters* **33**, L19407. doi:10.1029/2006GL027552.

First received 2 November 2020; accepted in revised form 12 March 2021. Available online 26 March 2021