Detecting changes in winter precipitation extremes and fluvial flood risk

This item was submitted to Loughborough University’s Institutional Repository by the/an author.


Additional Information:

- The book chapter, Detecting changes in winter precipitation extremes and fluvial flood risk by Wilby, R.L., Fowler, H.J. and Donovan, B. was published in: Beven, K. and Hall, J. (eds). Applied Uncertainty Analysis For Flood Risk Management © 2014 Imperial College Press. It is available here with the kind permission of Imperial College Press.

Metadata Record: https://dspace.lboro.ac.uk/2134/16270

Version: Accepted for publication

Publisher: © Imperial College Press

Rights: This work is made available according to the conditions of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) licence. Full details of this licence are available at: https://creativecommons.org/licenses/by-nc-nd/4.0/

Please cite the published version.
NOTICE: this is the author’s version of a work that was accepted for publication by Imperial College Press.

Changes resulting from the publishing process, such as peer review, editing, corrections, structural formatting, and other quality control mechanisms may not be reflected in this document. Changes may have been made to this work since it was submitted for publication. A definitive version was subsequently published in:


which has been published in final form at:
http://www.worldscientific.com/worldscibooks/10.1142/p588

This book chapter may be used for non-commercial purposes in accordance with Imperial College Press Terms and Conditions for self-archiving.
CHAPTER 22

DETECTING CHANGES IN WINTER PRECIPITATION EXTREMES AND FLUVIAL FLOOD RISK

Robert L. Wilby
Department of Geography, Loughborough University, UK

Hayley J. Fowler
School of Civil Engineering & Geosciences, Newcastle University, UK

Bill Donovan
Environment Agency, Bristol, UK

22.1 Introduction
There is a widely held perception that flood risk has increased across Europe during the last decade (EEA, 2005). Following extensive flash flooding in England, the Pitt Review (2008) concluded that: “The Summer 2007 floods cannot be attributed directly to climate change, but they do provide a clear indication of the scale and nature of the severe weather events we may experience as a result”. The review further asserted that, “timely decisions will allow organisations the flexibility to choose the most cost-effective measures, rather than being forced to act urgently and reactively. Early action will also avoid lock-in to long-lived assets such as buildings and infrastructure which are not resilient to the changing climate”. This echoes the position held by the Stern Review
Detection and attribution studies

*Detection* is the process of demonstrating that climate has changed in some defined statistical sense, without providing the reason(s) for the change (Hegerl et al., 2007). A change is detected in observations when the likelihood of an observation (such as an extreme temperature) lies outside the bounds of what might be expected to occur by chance. Changes may not be found if the underlying trend is weak compared with the “noise” of natural climate variability (e.g., McCabe and Wolock,
1991; Wolock and Hornberger, 1991). Conversely, there is always a small chance of spurious detection (perhaps due to an outlier event at one end of the observational record). The change might not necessarily occur linearly, indeed, there is evidence of abrupt, step changes in global climate such as the widespread changes in precipitation that occurred in the 1960s (Baines and Folland, 2007; Narisma et al., 2007). We should also keep in mind that climate records from neighbouring sites are often strongly correlated, a feature that is exploited by studies that pool data to maximise the strength of the signal (e.g., Fowler and Kilsby, 2003a; Pujol et al., 2007). However, trends are highly susceptible to false tendency and/or causation (see Legates et al., 2005; Sparks and Tryjanowski, 2005). This can arise because data are not homogeneous, being affected by a host of non-climatic influences such as encroachment of urban areas, changes in observer, instrumentation, monitoring network density, station location or exposure (see Davey and Pielke, 2005; Kalnay and Cai, 2003).

Attribution is the process of establishing the most likely cause(s) of detected changes at a defined level of statistical confidence (Hegerl et al., 2007). The problem for climate change studies is that there is no physical control, unlike epidemiological studies of causation. So attribution ultimately comes from comparing earth system observations against an earth systems model. Thus, attribution of observed climate change to human influence is accepted when there is consistency between a modelled response to combined human plus natural forcing and the observed pattern of climate behaviour. Climate models are used to isolate the unique “fingerprints” of different external forcings such as greenhouse gas concentrations, variations in solar radiation, or sulphate aerosols originating from volcanic eruptions. In one of the classic studies of this kind, Stott et al. (2001) showed that rising global mean temperatures in the second half of the twentieth century could only be explained by combining the human influence on atmospheric composition with natural forcings by solar output and volcanism.

Both detection and attribution require accurate characterisation of internal climate variability, ideally over several decades or longer. Internal climate variability is normally defined by long control simulations of coupled ocean-atmosphere global climate models without
external forcing because instrumental records are too short. Alternatively, climatologies of the last millenium may be reconstructed from proxy evidence such as tree rings or lake sediments. Whenever an observation lies outside the range of internal variability, climate change may be detected; whenever a pattern of climate anomalies is only explained by inclusion of human influences in model simulations, attribution is established. However, uncertainties in the representation of physical processes in climate models, in reconstructing climate from proxy data, and in the effects of the forcing agents mean that attribution results are most robust when based on several models (Zhang et al., 2006).

Evidence of human influences on the climate system has been accumulating steadily over the last two decades (Table 22.1). Early studies focused on large-scale, long-term changes in variables such as seasonal or annual global mean temperatures (e.g., Stott and Tett, 1998; Tett et al., 1999). However, since the IPCC Third Assessment Report (IPCC, 2001) there has been a proliferation of studies covering a much wider range of variables, and some are now even attributing climate changes at sub-continental scales (see Barnett et al., 2005). For example, Karoly and Stott (2006) assert that the observed warming in annual mean central England temperatures since 1950 is very unlikely due to natural factors but is consistent with the expected response to anthropogenic forcing.

With the possible exception of rainfall reductions over south-west Australia (Timbal et al., 2005), attribution of rainfall trends to human influence is not yet possible below the scale of the global land area, broken into latitudinal zones (Zhang et al., 2007). However, changes in moderately extreme precipitation events are, in theory, more robustly detectable than changes in mean precipitation (Hegerl et al., 2004; 2006) because as precipitation increases (under higher temperatures and greater water holding capacity of the atmosphere) a greater proportion falls in heavy and very heavy rainfall events (Katz, 1999; Trenberth et al., 2003; Pall et al., 2007). Disproportionate increases in heavy rainfall have been widely reported for the observed climate record (Groisman et al., 2005) but rates of change and/or regional patterns of observed and simulated rainfall extremes show little similarity in early studies (e.g. Kiktev et al.,
This is partly due to the inability of climate models to adequately resolve extreme precipitation at sub-grid box scales, the scale mismatch between point observations and gridded climate model output, and the difficulty of defining statistically robust “extreme” indices (Hegerl et al., 2006). These points will be revisited below.

Table 22.1 Examples of climate change detection and attribution studies.

<table>
<thead>
<tr>
<th>Realm</th>
<th>Spatial/temporal domain</th>
<th>Sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temperature</td>
<td>Mean temperature trends for North America</td>
<td>Karoly et al. (2003)</td>
</tr>
<tr>
<td></td>
<td>Mean temperature trends in Australia</td>
<td>Karoly and Braganza (2005)</td>
</tr>
<tr>
<td></td>
<td>Annual mean Central England temperatures</td>
<td>Karoly and Stott (2006)</td>
</tr>
<tr>
<td></td>
<td>Four indices of temperature extremes over US</td>
<td>Meethi et al. (2007)</td>
</tr>
<tr>
<td></td>
<td>Decadal mean temperatures over six continents</td>
<td>Stott (2003)</td>
</tr>
<tr>
<td></td>
<td>European summer temperature anomaly of 2003</td>
<td>Stott et al. (2004)</td>
</tr>
<tr>
<td>Precipitation</td>
<td>Global mean temperature changes</td>
<td>Stott et al. (2001)</td>
</tr>
<tr>
<td></td>
<td>Volcanic influence on global mean precipitation over land areas</td>
<td>Gillett et al. (2004b)</td>
</tr>
<tr>
<td></td>
<td>Twentieth century annual precipitation trends</td>
<td>Lambert et al. (2004)</td>
</tr>
<tr>
<td></td>
<td>Decreasing totals over SW Australia</td>
<td>Zhang et al. (2007)</td>
</tr>
<tr>
<td></td>
<td>Trends in winter rainfall over SW Australia</td>
<td>Timbal et al. (2005)</td>
</tr>
<tr>
<td>Troposphere</td>
<td>Sea level pressure</td>
<td>Gillett et al. (2003)</td>
</tr>
<tr>
<td></td>
<td>Tropopause height changes</td>
<td>Santer et al. (2003)</td>
</tr>
<tr>
<td></td>
<td>Atmospheric moisture content</td>
<td>Santer et al. (2007)</td>
</tr>
<tr>
<td></td>
<td>Near surface humidity</td>
<td>Willett et al. (2007)</td>
</tr>
<tr>
<td>Oceans</td>
<td>Heat content of oceans</td>
<td>Barnett et al. (2005)</td>
</tr>
<tr>
<td></td>
<td>Arctic sea ice extent</td>
<td>Gregory et al. (2002)</td>
</tr>
<tr>
<td></td>
<td>Wave conditions in the North Sea</td>
<td>Pfizenmayer and von Storch (2001)</td>
</tr>
<tr>
<td>Hydrosphere</td>
<td>Trend in global drought</td>
<td>Burke et al. (2006)</td>
</tr>
<tr>
<td></td>
<td>Global hydrological cycle since 1550</td>
<td>Tett et al. (2007)</td>
</tr>
<tr>
<td></td>
<td>Trend in arctic river discharges</td>
<td>Wu et al. (2005)</td>
</tr>
<tr>
<td>Biosphere</td>
<td>Length of the growing season</td>
<td>Christidis et al. (2007)</td>
</tr>
<tr>
<td></td>
<td>Warming over different vegetation types</td>
<td>Dang et al. (2007)</td>
</tr>
<tr>
<td></td>
<td>Canadian forest fires</td>
<td>Gillett et al. (2004)</td>
</tr>
</tbody>
</table>

Attributing individual extreme weather events to human activities presents additional challenges. It is often said that this kind of event is consistent with expected climate changes or might become more commonplace in the future. Formal attribution can only be accomplished in a probabilistic sense. In much the same way that smoking increases the risk of lung cancer by X%, the risk of an extreme meteorological event is said to be Y% more likely as a consequence of past greenhouse gas emissions than without. For example, the exceptional European heatwave in the summer of 2003 resulted in an area-average temperature anomaly of 2.3°C compared with the 1961–1990 average. By simulating
probability distributions of summer temperatures with and without anthropogenic emissions it was shown that the risk of a 2003 heatwave has increased by a factor of at least two due to human influence (Stott et al., 2004). Attributing changes in flood risk to carbon dioxide emissions is not yet feasible using the same approach because of the very large uncertainty in modelled precipitation changes at the river catchment scale (Prudhomme et al., 2003). For example, using results from the EU PRUDENCE project, Fowler and Ekström (2009) showed the extent to which the present generation of regional climate models (RCMs) is unable to reproduce properties of observed extreme summer rainfalls.

### 22.3 Changes in UK flood risk indices

Long-term changes in UK fluvial flood risk have been reviewed elsewhere (see Wilby et al., 2008). Several studies report increasing winter precipitation, larger multi-day rainfall totals, and higher contributions from intense daily events since the 1960s (Table 22.2). Others suggest strong regional gradients with larger winter increases in the north and west of the UK, and at higher elevation sites. These patterns translate into increased Winter mean runoff, especially in the upland areas of western England and Wales. The most widely accepted explanation is that the changes were forced by a strongly positive phase of the North Atlantic Oscillation (NAO) since the 1960s. This displaced storm tracks northwards and strengthened westerly moisture advection from the Atlantic over north-west Europe (Haylock and Goodess, 2004). Whether or not recent variability in the NAO is itself a manifestation of anthropogenic forcing remains an open question (Hegerl et al., 2007). However, when longer UK rainfall-runoff records are analysed, many of the trends found in shorter series cease to be significant (CEH and UKMO, 2001; Robson et al., 1998; Robson, 2002; Wilby et al., 2008). As mentioned above, trend detection for extreme events is far from straightforward because the outcome can depend on the chosen metric, period of record, power of the statistical test, and confounding factors:

- **Flood risk metrics.** A variety of indices have been used to detect changes in flood risk. For example, the sample studies in Table 22.2 employ series of monthly, seasonal and annual rainfall/river flow,
annual maxima of daily rainfall intensities/river flows or N-day rainfall totals/maximum flows, proportional contribution of heavy events to total rainfall, and annual counts of peaks over threshold flows. Analyses may be performed using point or area average data, individual records, data from networks of stations, via pooling of (rainfall) maxima by region or elevation, and using gridded extreme indices (e.g. Alexander et al., 2006). In some cases, more exotic indicators may be applied such as frequencies of flood generating atmospheric circulation patterns (e.g. Bárdossy and Filiz, 2005), changes in the timing of extreme events (e.g. Fowler and Kilsby, 2003b); or standardised precipitation indices (e.g. Seiler et al., 2002).

In any event, the chosen metric should be meaningful in terms of flood generation mechanisms and interpretable from a policy perspective (e.g. sub-daily data are needed for flash flooding). Fortuitously, model experiments show that changes in indices of extreme precipitation are stronger than corresponding changes in mean precipitation (Hegerl et al., 2004). Widely employed indicators of extreme events (e.g. number of days with precipitation greater than 10mm, or the fraction of total rainfall due to events exceeding the 95th percentile amount) are acknowledged to be “not as extreme” as they could be (Frich et al., 2002; Tebaldi et al., 2006). Here, the trade-off is between having extremes that are severe enough to have repercussions for society, yet are not so rare that there are insufficient events to enable detection of change.

- **Period of record.** When longer rainfall and river flow records are analysed, many trends found in shorter series cease to be significant. This can be due to the influence of outliers (at the start or the end of the record), or simply the large multi-decadal variability of the UK rainfall regime. It is widely recognised that extreme statistics derived from conventional 30 year climatological periods are subject to large uncertainty (Kendon et al., 2008; Lane, 2008). Following a detailed review of the 2000/01 UK flooding, it was recommended that records of at least 50 years duration should be used for investigating possible climate change signals in rainfall and river flows (CEH and UKMO, 2001; Kundzewicz and Robson, 2004). Unfortunately, continuous river flow records of this length, with accompanying
meta-data, are rare. Out of more than 1000 flow records held by the Global Runoff Data Centre, less than 200 worldwide have continuous daily flow series longer than 40 years that extend into the late 1990s (Kundzewicz et al., 2005).

Table 22.2 Observed changes in UK precipitation and runoff (from Wilby et al., 2008).

<table>
<thead>
<tr>
<th>Region</th>
<th>Period (Metrics)</th>
<th>Key Findings</th>
<th>Sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>9 homogeneous rainfall regions</td>
<td>1766–2000 (monthly rain)</td>
<td>Significantly wetter winters in west Scotland; 1990s unusually wet</td>
<td>Alexander and Jones (2001)</td>
</tr>
<tr>
<td>102 stations across Scotland</td>
<td>1914–2004, 1961–2004 (daily, seasonal rain and snow cover)</td>
<td>Increased winter rainfall since 1960s across all regions but up to 70% in north; increase in spring rainfall in west since 1910s; large reductions in autumn/spring snow cover; increase in heavy winter rain in north and west</td>
<td>Barnett et al. (2006)</td>
</tr>
<tr>
<td>28 rainfall stations and 15 gauging stations</td>
<td>1881– onwards</td>
<td>Increase in annual rainfall maxima since 1960s; no long-term trend in rainfall maxima, flood peaks, overall flood volume or duration</td>
<td>CEH &amp; UKMO (2001)</td>
</tr>
<tr>
<td>204 stations across UK</td>
<td>1961–2000 (daily rain)</td>
<td>Significant increase in annual maxima, 5- and 10-day rainfall events over Scotland and northern England, especially in the 1990s</td>
<td>Fowler and Kilsby (2003a;b)</td>
</tr>
<tr>
<td>56 stations in the English Lake District</td>
<td>1971–2000 (daily, monthly rain)</td>
<td>Increased heavy precipitation at high elevation sites; weakening of the Cumbrian rain shadow</td>
<td>Malby et al. (2007)</td>
</tr>
<tr>
<td>689 stations across UK</td>
<td>1961–2006 (daily rain)</td>
<td>Heavy precipitation events are making a greater contribution to winter totals and less in summer</td>
<td>Osborn et al. (2000); Maraun et al. (2008)</td>
</tr>
<tr>
<td>890 gauging stations</td>
<td>1870– onwards (seasonal rain, peak flows)</td>
<td>Increase in winter rainfall since 1960s; flood rich and flood poor periods but no overall trend</td>
<td>Robson (2002); Robson et al. (1998)</td>
</tr>
<tr>
<td>47 gauging stations across UK</td>
<td>1970–2002 (seasonal flow)</td>
<td>Significant increases during autumn and winter at &lt;25% sites; no clear regional pattern</td>
<td>Wade et al. (2005)</td>
</tr>
<tr>
<td>15 rivers in</td>
<td>1865–2002</td>
<td>Significant long-term increase in</td>
<td>Wilby (2006)</td>
</tr>
</tbody>
</table>
England and Wales (reconstructed seasonal flow) winter flow in only three rivers; increasing annual flows since 1970s

| Stations with snowfall >13 cm in Britain | 1861–1996 (catalogue of heavy snowfalls) | Decadal variability in the frequency of heavy snowfall peaking in the 1860s, 1870s and the 1970s; no overall trend | Wild et al. (2000) |

- **Statistical tests.** As noted, climate variability affects the detectability of a trend within a finite series of observations. The possibility of erroneous trend detection (Type I errors) are controlled by choosing an appropriate statistical test, and level of confidence. Widely used methods for extremes include (logistic) linear regression, “change point” tests, and the non-parametric Spearman rank correlation and Mann–Kendall tests (Lanzante, 1996). Alternatively, an actual trend might be overlooked because it is overshadowed by short-term climate fluctuations (Type II error). In this case, the detectability depends on the power of the test as a function of record length, magnitude of the trend and rarity of event(s). Even when presented with the same data, different (flood) indices may reveal different numbers of significant trends (e.g. Svensson et al., 2005). Detectability of trends in very rare events is particularly difficult for small sample sizes but can be improved through regional pooling of data (Frei and Schär, 2001). For most practical situations, this points to the necessity of analysing “less extreme” extremes.

- **Confounding factors.** Unerring faith in detected trends without a supporting conceptual framework can lead to invalid statistical inference (Sparks and Tryjanowski, 2005). Creeping or sudden changes in meteorological records can arise for a range of reasons such as changing instrumentation or instrument location, observing or recording practices, site characteristics, or sampling regime (Zwiers and von Storch, 2004). Some claim that the entire observational record has been contaminated to some extent by human influences on climate and hydrology (Tett et al., 2007). Discharge records may respond to a host of non-climatic influences including land cover and management, urbanisation, river regulation, water abstraction and effluent returns, or flood flows may by-pass gauging structures (Archer, 2007; Legates et al., 2005). Occasionally false trends and biases arise from the statistical method used to
divide data, such as percentile-based indices for temperature and precipitation extremes (Michaels et al., 2004; Zhang et al., 2005).

In summary, the choice of index, spatial and temporal scale of aggregation, statistical test (including significance testing), and the account taken of confounding factors, all require careful justification in a detection study. The next section briefly describes earlier work on climate change detection at river catchment scales. Then we outline a method for detecting changes in heavy precipitation using RCM output. This involves calculating when and where changes in downscaled rainfall metrics are not likely to have been entirely the result of (model-estimated) natural internal variability.

### 22.4 Detection time for changes in heavy precipitation

Detection times for climate change trends in hydroclimatic data depend on the strength of the assumed trend, the sample variance of the time series in question, and the probabilities of making errors of Type I or II. Preliminary estimates using data for river catchments in the US and UK suggest that statistically robust trends in seasonal runoff are unlikely to be found until the second half of the 21st century (Ziegler et al., 2005; Wilby, 2006).

The same statistical relationship can be inverted to estimate the strength of trend required for detection by specified time horizons. For example, Figure 22.1 shows that, on average, a 60% change in winter rainfall is needed for detection by 2025 given long-term (L) inter-annual variability over the UK catchments used by Wilby (2006), whereas the change needed for detection by 2055 is 44%. The smallest changes needed for detection in winter rainfall over the same periods are in the River Tyne catchment (54% and 39% respectively). Even smaller changes are needed for detection in annual rainfall totals by 2025 and 2055 (25% and 18% respectively).

Analysis of UK Winter and annual rainfall totals suggests that changes of ~25% will be needed for early detection (by year 2025) in the best case catchment (River Tyne). Although increases (or decreases) in seasonal totals could affect overall flood risk, it is suspected that more
useful flood-risk information can be extracted from daily precipitation indices. Furthermore, as noted before, projected trends in indices of heavy precipitation may be detectable earlier than trends in seasonal means (Hegerl et al., 2004). However, more complex flood events involving rainfall and snowmelt, or fluvial flooding and tidal surge (as in Svensson et al., 2002) are clearly not captured by “rain only” indices.

Figure 22.1 Detectable trends (% change) in seasonal mean rainfall by 2025 and 2055 for sample variances derived from short (S, 1961–1990), medium (M, 1921–1990) and long (L, 1865–1990) records. Each point represents a different river catchment; the bars show the mean detection trend by time horizon and variance estimate.

### 22.5 Case study: north-west and south-east England

Here, we provide a case study for detecting changes to extreme rainfall metrics based on a combination of regional frequency analysis (RFA) and pattern scaling. The estimation of the frequencies of extreme events is difficult as extreme events are, by definition, rare and observational records are often short. In RFA, data from the “region” are assumed to share the same frequency distribution and only differ in their magnitude (mean or median values). Therefore, RFA trades “space for time” and pools standardised data from several different sites within a “region” to fit a single frequency distribution (Hosking and Wallis, 1997).
22.5.1 Data

We employed RCM output from the PRUDENCE (Prediction of Regional scenarios and Uncertainties for Defining European Climate change risks and Effects) ensemble, which contains daily data for a range of future climate scenarios.

Table 22.3: The thirteen RCM integrations used in this study. The first part of each acronym refers to the RCM, and the second to the GCM supplying the boundary conditions. All RCM integrations are from PRUDENCE.

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Institution/Model Origin and References</th>
<th>RCM</th>
<th>GCM</th>
</tr>
</thead>
<tbody>
<tr>
<td>HADH</td>
<td>Hadley Centre, UK Met Office, Exeter; Regional model at the Hadley Centre. Model: Jones et al. (2004a)</td>
<td>HadRM3P</td>
<td>HadAM3P</td>
</tr>
<tr>
<td>HIRE</td>
<td>Hadley Centre, UK Met Office, Exeter; Regional model at the Hadley Centre. Model: Jones et al. (2004a)</td>
<td>HadRM3P</td>
<td>HadAM3P</td>
</tr>
<tr>
<td>RCAH</td>
<td>Swedish Meteorological and Hydrological Institute, Stockholm; Rossby Centre Atmosphere Ocean Model. Model: Jones et al. (2004b), Meier et al. (2003), Dössen et al. (2002), Räisänen et al. (2004).</td>
<td>RCOAO</td>
<td>HadAM3H</td>
</tr>
<tr>
<td>CHRMH</td>
<td>Swiss Federal Institute of Technology (ETH), Zurich; Climate High-Resolution Model. Model: Lüthi et al. (1996), Vidale et al. (2003).</td>
<td>CHRM</td>
<td>HadAM3H</td>
</tr>
<tr>
<td>CLMH</td>
<td>GKSS, Institute for Coastal Research, Geesthacht, Germany; Climate version of “Lokalmodell” of German Weather Service. Model: Steppeler et al. (2003).</td>
<td>CLM</td>
<td>HadAM3H</td>
</tr>
<tr>
<td>REMOH</td>
<td>Max Planck Institute for Meteorology, Hamburg, Germany; Dynamical core from “Europamodell” of German Weather Service, Parameterisations from ECHAM4. Model: Jacob (2001), Roeckner et al. (1996).</td>
<td>REMO</td>
<td>HadAM3H</td>
</tr>
<tr>
<td>PROMH</td>
<td>Universidad Complutense de Madrid, Spain; Climate version of PROMES model. Model: Castro et al. (1993), Arribas et al. (2003).</td>
<td>PROMES</td>
<td>HadAM3H</td>
</tr>
<tr>
<td>REGH</td>
<td>Abdus Salam International Centre for Theoretical Physics, Italy (ICTP); Dynamical core from MM5, Parameterisations from CCM3. Model: Giorgi et al. (1993a,b, 2000), Pal et al. (2000).</td>
<td>RegCM</td>
<td>HadAM3H</td>
</tr>
<tr>
<td>RACH</td>
<td>The Royal Netherlands Meteorological Institute (KNMI), Netherlands; Dynamical core from HIRLAM, Parameterisations from ECMWF physics. Model: Tiedtke (1989, 1993), Lenderink et al. (2003).</td>
<td>RACMO2</td>
<td>HadAM3H</td>
</tr>
<tr>
<td>METH</td>
<td>Norwegian Meteorological Institute; Version of HIRHAM. Model: Christensen et al. (2001), Hanssen-Bauer et al. (2003).</td>
<td>MetNo</td>
<td>HadAM3H</td>
</tr>
</tbody>
</table>
of climatic variables for control (1961–1990) and future (2071–2100) time slices (Christensen et al., 2007). We used 13 RCM integrations from the PRUDENCE ensemble. All experiments yield daily precipitation totals for control (1961–1990) and future (2071–2100) time periods (Christensen et al., 2007) under the IPCC SRES A2 emissions scenario (Nakićenović et al., 2000) (Table 22.3). Nine of the RCM experiments were performed by nesting within the atmosphere-only high-resolution global climate model (GCM) HadAM3H of the UK Hadley Centre. One RCM, HadRM3P, was nested within HadAM3P, a more recent version of the same atmosphere-only GCM; but HadRM3H and HadRM3P can be considered as essentially the same model for Europe (Moberg and Jones, 2004). The variable resolution global atmospheric model, Arpège, is nested directly within HadCM3. Additionally, two RCM integrations, HIRHAM and RCAO, are driven by lateral boundary conditions from two separate integrations of the ECHAM4/OPYC3 coupled ocean-atmosphere GCM.

All the RCMs operate with grid spacing of ~0.5° longitude by ~0.5° latitude (~50 km spatial resolution) over a European domain and data were re-gridded to a regular 0.5° x 0.5° grid to allow direct comparison between models. Suffixes E and H denote RCMs driven by ECHAM4/OPYC3 and HadAM3H/P/HadCM3 GCMs respectively. More details of the experimental design of the PRUDENCE integrations can be found in Jacob et al. (2007).

A dataset of comparable spatial scale to the RCM outputs was produced by aggregating a daily observed precipitation 5 km by 5 km grid produced by the UK Meteorological Office (Perry and Hollis, 2005a,b) to the regular 0.5° x 0.5° grid. This was achieved by taking a daily average across the 5 km boxes contained within each 0.5° x 0.5° grid cell for each day of 1961–1990.

22.5.2 Detection indices

Seasonal maximum (SM) series of 1, 5 and 10 day precipitation totals were extracted for each grid cell, for each RCM time-slice (the control period, 1961–1990, and future period, 2071–2100), and for observations. These SM series were standardized by their median (Rmed; following
Fowler et al., 2005) and the standardised SM were then pooled for each of the nine UK rainfall regions delineated by Wigley et al. (1984) (Figure 22.2). The homogeneity of these regions for extreme precipitation was confirmed previously by Fowler and Kilsby (2003a).

![Figure 22.2 The RCM regular 0.5° by 0.5° grid and the nine coherent rainfall regions. The regions are: North Scotland (NS), East Scotland (ES), South Scotland (SS), Northern Ireland (NI), Northwest England (NWE), Northeast England (NEE), Central and Eastern England (CEE), Southeast England (SEE) and Southwest England (SWE) (from Fowler et al., 2007).]

Next, a Generalised Extreme Value (GEV) distribution was fitted to each pooled SM sample using Maximum Likelihood Estimation (MLE) and return values of precipitation intensities with average recurrence of 10 and 50 years were estimated. The estimates were then rescaled using the regional average Rmed from the original SM dataset. We also calculate the associated error by the 95% confidence intervals on the return value estimates via the delta method using three different estimates of variance to test the sensitivity: (i) RCM pooled sample variance; (ii) observed variance, 1961–1990; (iii) observed variance, 1958–2002 (full length of record). Note that the use of the full record of observed variance gives a
more conservative estimate of confidence intervals compared to the RCM pooled sample variance.
A further step is needed to estimate transient changes in the return values of precipitation intensities for each RCM between the two time-slices, 1961–1990 and 2071–2100. We applied a conventional pattern scaling approach (Mitchell, 2003). The technique assumes that regional changes in extreme precipitation (or any climate variable) will occur in proportion to the projected change in global mean temperature, in this case, from the two GCMs providing lateral boundary conditions for the 13 PRUDENCE RCMs (i.e. HadCM3 and ECHAM4). Pattern scaling was based on the change in global mean temperature predicted for 30-year time slices centred on the years 1975, 2025, 2055 and 2085. Therefore, designating the global mean temperature as \( T_y \) where \( y \) is the central year (1975, 2025, 2055 and 2085 respectively), the scale factor for each GCM time slice may be written as:

\[
SF_y = \frac{T_y - T_{\text{Con}}}{T_{\text{Fut}} - T_{\text{Con}}},
\]

(22.1)

where \( T_{\text{Con}} \) and \( T_{\text{Fut}} \) indicate the global mean temperature for the Control (1961–1990) and Future (2071–2100) time slices respectively. For intervening years the scale factors were linearly interpolated to provide transient scale factors. Therefore, for 1990 to 2100, \( SF \) varies from zero to one. The transient scale factors show that global mean temperature change will accelerate over the next 100 years. An alternative to pattern scaling on global mean temperature change would be to simply apply linear scaling between 1990 and 2100; this would imply earlier detection times. We use the more conservative non-linear pattern-scaling to estimate return values of precipitation intensities and the associated confidence intervals for each year between 1990 and 2100.

22.5.3 Detection times

We define a detectible increase in extreme precipitation, \( D_x \), as the point (year) at which we would reject (at the \( \alpha = 0.05 \) or 95% significance level) the null hypothesis that the return level estimated for the 1961–1990 period, \( \mu_1 \), and the return level estimated for a year \( x \) (where \( x > \)
1990), \( \mu_x \), are equal in favour of the alternative hypothesis that \( \mu_x \) is greater (lower) than \( \mu_c \). This statistical test is based on the signal to noise ratio and provides a distribution that is approximately normally distributed with a mean of zero and a standard deviation of 1; \( N(0,1) \) (see Equation 22.2). We then use a two-tailed Student’s \( t \)-test (assuming that the trend can go down as well as up) to estimate the point at which the return levels are shown to be from a significantly different population at the \( \alpha = 0.05 \) level, i.e. where \( D_x \geq 1.96 \):

\[
D_x = \frac{\mu_x - \mu_c}{\sqrt{\sigma_f^2 + \sigma_c^2}} \geq 1.96 \tag{22.2}
\]

where \( \mu_x \) is the pattern-scaled return level for year \( x \), \( \mu_c \) is the estimated return level for the control period (1961–1990), \( \sigma_f^2 \) is the variance in the return level estimate for the RCM generated future period (2071–2100) and \( \sigma_c^2 \) is the variance in the return level estimate for RCM generated control period (1961–1990). Note that the observed variance, \( 2\sigma_o^2 \), for 1961–1990 and 1958–2002 respectively, was substituted for the summed variance from the RCM control and future integrations, \( \sigma_f^2 + \sigma_c^2 \), in Equation (22.2) to establish the sensitivity of the test to assumed variance (where the delta confidence intervals on the return level estimates were also calculated using the same assumed variance estimate, either RCM generated or observed for 1961–1990 and 1958–2002 respectively, as previously stated).

This test defines the “detection year”, \( D_x \), as the first year at which \( D_x \geq 1.96 \) (the transient return level estimate for year \( x \) is significantly different to the return level estimate for the control period (1961–1990) at the \( \alpha = 0.05 \) level).

We apply the principle of equal weighting across different RCMs to produce probability distributions of estimated \( D_x \) for change to the return levels of seasonal extreme precipitation in the 9 UK rainfall regions. This is reasonable because weighting RCMs in the PRUDENCE ensemble by their skill at reproducing observed extremes did not make any practical difference since most are driven by lateral boundary conditions from the same GCM, HadAM3H/CM3 (Fowler and Ekström, 2009; Manning et al., in revision). However, we examine the sensitivity
of $D_s$ to the assumed variance in extremes and to the $\alpha$ level used in the detection test. The analysis is truncated at 2100 as we cannot assume that the pattern scaling relationship holds true after this time. Furthermore, although the analysis includes summer precipitation extremes, it should be noted that the present generation of RCMs do not adequately reproduce observed extremes during this season (Fowler and Ekström, 2009). Throughout the study we define the point beyond which the probability of detection is more likely than not as when more than 50% of the RCMs detect a significant change in extreme precipitation.

22.5.4 Results

Cumulative frequency distributions (cdfs) of detection time by region and season are presented for 1 and 10 day precipitation totals at the 10-year return level. Figure 22.3 shows the detection times for significant change (at the $\alpha = 0.05$ level) to Winter extreme precipitation. In these plots, the vertical black dotted line indicates the year 2050 and the horizontal red dashed line indicates the point at which the probability of detection is 0.5 (i.e. the chance of detection is 50%). Beyond this point, change to these metrics is more likely than not to be detectible. Three individual cumulative frequency distributions are shown on each plot: the green cdf shows the detection time, $D_s$, using RCM internal variability (i.e. using Equation 22.2); the red cdf shows the same but for observed variability for 1961–1990 (i.e. substituting this into Equation 22.2); the black cdf shows the same but for the full record of observed variability from 1958–2002 (i.e. substituting this into Equation 22.2).

The detection time, $D_s$, is sensitive to (i) the assumed variance used to calculate confidence intervals on the return level estimates (95% confidence intervals are calculated in all cases) and in the detection test, and (ii) the $\alpha$ significance level used in testing “detectability” (note that the standard used is the $\alpha = 0.05$ level). We find that there are no consistent patterns (across regions) of earlier or later detection time for the different assumed variance estimates. For the two case study regions, use of the observed variability for 1961–1990 yielded longer detection times. Across all regions, for the Winter 1 day 10-year return level, the mean detection time is 2045 for RCM-estimated variability, 2055 for
observed variability 1961–1990, but 2042 for the longer observed variability 1958–2002. Note that even with regional pooling of data, confidence intervals for the return period estimate are wide due to the relatively large uncertainty (variability) in the estimate. In the case study
Figure 22.3 Detection year for significant change ($\alpha = 0.05$) in the return levels of maximum 1-day (upper panels) and 10-day (lower panels) winter precipitation totals with 10-year return period in NWE (left hand column) and SEE (right hand column). Data used to estimate natural variability were: observed 1958–2002 (red lines); RCM 1961–1990 (green lines); observed 1961–1990 (black lines).

regions, use of the observed variability available from the 5km gridded dataset (for 1958–2002) provides the shortest detection times for 1-day totals. But for the 10-day totals the RCM-estimated natural internal variability enables earlier detection. In these cases, low variance of seasonal maxima within the RCM integration (and/or outliers in the longer observed record) decreases (increases) the width of confidence intervals and thus produces shorter (longer) detection times.

Records prior to the early nineteenth century are thought to be less representative of the present climate regime (see for example, Marsh et al., 2007) and so were not used to characterize the variance statistic. Even so, it is recognized that long-term rainfall and reconstructed runoff records can help to quantify multi-decadal variability without anthropogenic climate change (see Jones et al., 2006). For example, a 200-year monthly precipitation series for the English Lake District varied by -17% to +11% of long-term average between the driest (1850s) and wettest (1920s) decades (Barker et al., 2004).

For the two exemplar regions, future changes in some winter precipitation indices could be detectable relative to 1958–2002 conditions as early as the 2020s and 2030s (Table 22.4). The earliest detection year (2027) was found for 10-day winter precipitation totals in SEE, although even earlier detection may be possible in south-west England (not shown). Detection years were earlier for 1- than 10-day precipitation totals in NWE; but the opposite applies in SEE. However, detection is earliest for 10-year return period winter precipitation indices, and earlier than for spring or autumn. At 50-year return periods it is unlikely that changes will be detectible by 2050, and for some regions changes are not even detectible by the end of the 21st century (not shown). Across all regions and seasons, the earliest detection times were generally found in Winter, for longer duration (10-day) precipitation events, with shorter (10-year) return periods, based on estimates of variability taken from RCM output. The earliest mean detection times for
variance estimates based on observations (1958–2002) were also found for Winter 10-year return period events but were broadly similar for 1- and 10-day totals.

Table 22.4 Detection years for significant changes in 1- and 10-day seasonal precipitation totals with 10-year return period, using variance estimates based on observations for the period 1958–2002. Note that summer was excluded because of the low confidence in RCM projections of regional rainfall. See Fowler and Wilby (2009) for other regions.

<table>
<thead>
<tr>
<th>Season</th>
<th>1-day</th>
<th>10-day</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dec-Feb</td>
<td>NWE</td>
<td>SEE</td>
</tr>
<tr>
<td>Mar-May</td>
<td>2047</td>
<td>2083</td>
</tr>
<tr>
<td>Sep-Oct</td>
<td>2039</td>
<td>2071</td>
</tr>
</tbody>
</table>

22.6 Concluding remarks

Following extensive flooding in England during Summer 2007, budgets for flood risk management were significantly increased. This raises the questions as to when and where to prioritise future investment in flood management measures? This chapter has summarised the latest scientific evidence for the detection and attribution of extreme events to climate change. Although attribution of changes in precipitation or flood risk is not yet possible at regional scales, techniques are being developed for detection of trends in these indices at river catchment scales, and for estimating the time-horizons for formal detection. We set the scene by reviewing evidence of changing flood risk in the observational records of the UK and by identifying some of the main factors that confound detection. We then demonstrated a method for estimating detection times for changes in seasonal precipitation extremes projected by the EU PRUDENCE regional climate model (RCM) ensemble. We showed that for selected regions and extreme precipitation indices, the climate change signal(s) in the PRUDENCE projections could be detectable as early as the 2020s. The next step is to broaden the analysis to a national scale, and to investigate a larger suite of extreme indicators. By identifying potential “hotspots” of emerging flood risk, a more targeted approach to monitoring and investment might then be feasible.
22.7 Acknowledgements

This research was supported by the Environment Agency and by a NERC fellowship award to Dr Hayley Fowler. The views expressed are those of the authors and are not necessarily indicative of the position held by the Environment Agency.

REFERENCES


