https://doi.org/10.5194/hess-2020-576
Preprint. Discussion started: 14 November 2020

Hydrology and Earth System Sciences
Discussions
Open Access

Nonstationary weather and water extremes: a review of methods for their detection, attribution, and management

Louise J. Slater[a], Bailey Anderson[a], Marcus Buechel[a], Simon Dadson[a], Shasha Han[a], Shaun Harrigan[b], Timo Kelder[c], Katie Kowal[a], Thomas Lees[a], Tom Matthews[c], Conor Murphy[d], and Robert L. Wilby[c]

[a]School of Geography and the Environment, University of Oxford, OX1 3QY, UK
[b]European Centre for Medium-Range Weather Forecasts (ECMWF), Reading, UK
[c]Geography and Environment, Loughborough University, Loughborough, UK
[d]Maynooth University, Maynooth, County Kildare, Ireland

Correspondence: Louise J. Slater (louise.slater@ouce.ox.ac.uk)

Abstract. Hydroclimatic extremes such as intense rainfall, floods, droughts, heatwaves, and wind/storms have devastating effects each year. One of the key challenges for society is understanding how these extremes are evolving and likely to unfold beyond their historical distributions under the influence of multiple drivers such as changes in climate, land cover, and other human factors. Methods for analysing hydroclimatic extremes have advanced considerably in recent decades. Here we provide a review of the drivers, metrics and methods for the detection, attribution, prediction and projection of nonstationary hydroclimatic extremes. We discuss issues and uncertainty associated with these approaches (e.g arising from insufficient record length, spurious nonstationarities, or incomplete representation of nonstationary sources in modelling frameworks), examine empirical and simulation-based frameworks for analysis of nonstationary extremes, and identify gaps for future research.

1 Introduction: nonstationary hydroclimatic extremes

Are hydroclimatic extremes stationary or nonstationary? This question has generated much debate because of the ramifications for hazard management in a changing world. At the simplest level, a stationary process is one where the statistical properties of the distribution do not shift over time. Thus, a stationary time series (Figure 1a) would not exhibit any shift in mean, variance (Figure 1b) or shape. For hydroclimatic extremes, this implies the distribution of extreme precipitation, temperature, streamflow, or wind should merely fluctuate within a stationary envelope of variability. The assumption of stationarity has long served as the basis for statistical analysis of hazards and design of engineering structures, by defining the magnitude of events with a given frequency of occurrence, such as the probable maximum precipitation (PMP) or stationary 100-year design flood (e.g. Salas et al., 2018).

In reality, hydroclimatic extremes exhibit multiple forms of natural nonstationarity depending on the chosen timescales (e.g. Hannaford et al., 2013; Stevenson et al., 2018). Alongside 'natural' nonstationarities, the global water cycle also manifests many artificial nonstationarities induced by human activities, such as climate change, land cover change (e.g. Blum et al., 2020), water abstraction/augmentation, river regulation, and even geopolitical uncertainty (e.g. Wine, 2019) at local, regional, and global scales. Trends and step-changes in the magnitude, frequency, duration, volume, or areal extent of hydroclimatic

https://doi.org/10.5194/hess-2020-576
Preprint. Discussion started: 14 November 2020

Hydrology and Earth System Sciences
Discussions

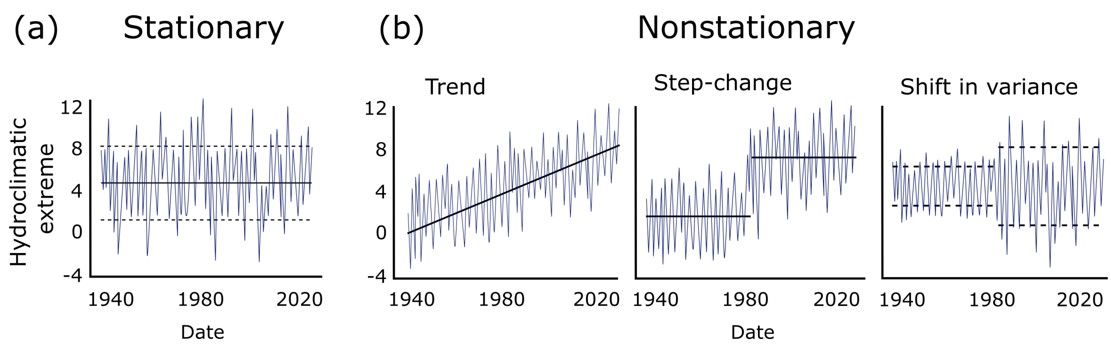

Figure 1. What is nonstationarity? Examples of (a) a stationary time series with constant mean and variance; and (b) three nonstationary time series in the form of shift in mean (trend, step-change) and shift in variance. Solid and dashed black lines represent the mean and the variance of the time series, respectively.

extremes such as intense rainfall (e.g. Sun et al., 2020a; Westra et al., 2013; Donat et al., 2016), floods (e.g. Berghuijs et al., 2019a; Do et al., 2017; Archfield et al., 2016), droughts (e.g. Andreadis and Lettenmaier, 2006; Spinoni et al., 2017), and heatwaves/heat stress (e.g. Oliver et al., 2018; Lorenz et al., 2019; Ouarda and Charron, 2018) have been widely detected, and have led to proclamations about "death" of stationarity (Milly et al., 2008). In contrast, trends in strong winds/storms are less certain and the influence of anthropogenic climate change is difficult to attribute (Shaw et al., 2016; Elsner et al., 2008; Martínez-Alvarado et al., 2018; Wohland et al., 2019). Disentangling natural and anthropogenic drivers of nonstationarity is problematic, as the two can be interlinked, and even seemingly 'natural' drivers such as climate modes may shift under the effects of anthropogenic climate change (e.g. Maher et al., 2018). Although the drivers of abrupt nonstationarities (step-changes) may be apparent (e.g. water abstraction, reservoir filling and operations), drivers of more incremental nonstationarities (trend/variance) (e.g. climate variability and change, and land cover change) may be harder to attribute and/or obfuscated by other confounding factors.

Deciding whether a time series should be treated as nonstationary for the purpose of managing extremes is one of the greatest challenges facing weather and water scientists and practitioners today. Short-term trends are widely present in variables like streamflow, which exhibit long memory (time-dependence) and periodicities. However, such trends are not necessarily indicative of nonstationarity (e.g. Koutsoyiannis, 2006; Koutsoyiannis and Montanari, 2015). Multi-decadal (e.g. 30-year) shifts may simply be temporary excursions in longer (e.g. 100-year) records, or a function of the start and end dates chosen for the trend analysis (Harrigan et al., 2018), and may not warrant application of nonstationary analysis. A growing body of literature has shown that inappropriately applying nonstationary models to short time series may have the undesired effect of increasing uncertainty; in cases where model structure and/or underlying physical drivers are uncertain, stationary models may be the preferred option for design and management of extremes (Serinaldi and Kilsby, 2015). However, it is also recognised that for engineering purposes, the concept of nonstationarity is highly valuable. Villarini et al. (2018, p. 6) contend that *"the issue is not whether observations arise from a long-term excursion from some underlying stationary process but rather whether the*

https://doi.org/10.5194/hess-2020-576
Preprint. Discussion started: 14 November 2020

Hydrology and Earth System Sciences
Discussions
Open Access

probability distribution of future (events) will resemble the distribution that is obtained from fitting a probability distribution to observations over a historical record". As such, nonstationary approaches have "functional" value when there are good reasons to suspect physically-plausible drivers of change, which in turn are somewhat predictable.

Detecting the presence, and attributing the source, of both natural and artificial nonstationarities in hydroclimatic extremes is vital for understanding and managing water resources in a changing world. Nonstationarities may have dramatic impacts for infrastructure (François et al., 2019), property, and society over a range of overlapping timescales. There are many different drivers that may alter extremes simultaneously over the short-term (e.g. influence of groundwater abstraction on streamflow) and medium- to long-term (e.g. interdecadal to multidecadal "hazard-rich" and "hazard-poor" periods driven by modes of climate variability) of relevance to societies (e.g. Rust et al., 2019). The appropriate timescales of nonstationarity, therefore, depend on the nature and dynamism of the variable of interest, and associated impacts on a given location.

Over the last two decades, hundreds of papers have addressed the detection, attribution, prediction and projection of nonstationarity in precipitation, floods, drought, heat stress, temperature, extreme winds and storms. However, there is yet no comprehensive, introductory overview of these methods across hydroclimatic extremes, or overarching discussion of key challenges that can arise. This paper offers a synthesis of methods for quantifying, attributing, and managing nonstationarity over multiple spatial and temporal scales, along with their limitations. The structure follows the logical order of steps employed in a detection, attribution and management framework (Figure 2).

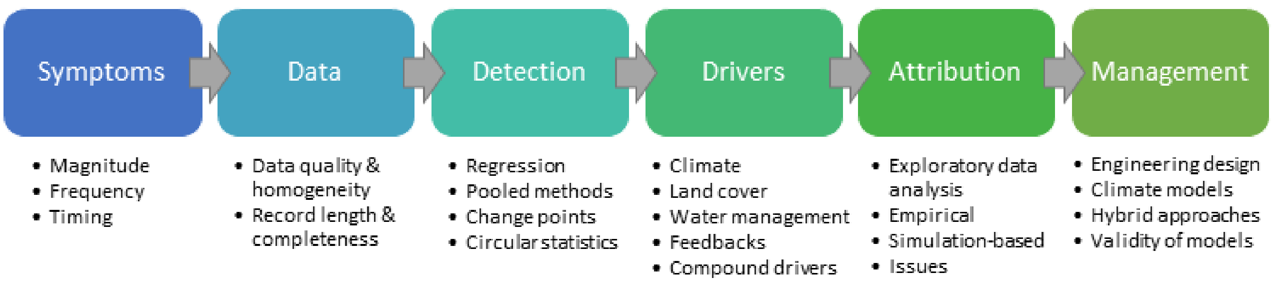

Figure 2. Workflow for the detection, attribution and management of nonstationary hydroclimatic extremes. Bullet points indicate examples discussed in the manuscript (sections 2-7).

Section 2 describes the most widely used indices for diagnosing the symptoms of nonstationarity via changes in magnitude, frequency and timing. Section 3 identifies essential data pre-requisites for detecting nonstationarity, such as homogeneity analysis. Section 4 discusses the techniques used to detect nonstationarity, including regression-based methods for gradual change, pooled approaches for analysis of rare extremes, step-change analysis for abrupt change, and other methods for discerning changing seasonality. Section 5 introduces the key drivers of nonstationarity of hydroclimatic extremes. Section 6 reviews approaches for attribution of nonstationary extremes, including both observation- and model-based approaches, and issues with attribution and engineering design under nonstationarity. Finally, Section 7 discusses approaches to manage nonstationarity via engineering design, prediction (over weekly to annual timescales) and projection (over multi-decadal timescales), including

Hydrology and Earth System Sciences
Open Access
Discussions

key limitations. The overall aim of the paper is to highlight the most significant issues and considerations when detecting, attributing, predicting and projecting nonstationarity in hydroclimatic extremes.

2 Symptoms of nonstationarity

Both natural and artificial nonstationarities in hydroclimatic extremes may be expressed through a significant shift in the mean, variance, or shape of a given time series (Figure 1). Such departures are generally diagnosed by symptoms such as a change in the magnitude (events becoming more or less extreme), frequency (events occurring more or less often than before) and timing (events occurring earlier or later in the year) of seasonal or annual extremes. There are many different ways of describing the symptoms of nonstationarity - each is discussed in turn below.

2.1 Are extreme events becoming more extreme?

Significant changes in the magnitude of extremes are relevant to society, engineers, decision-makers, and insurers alike. The magnitude or intensity of an event is generally described by estimating the percentiles of a distribution over a given period (e.g. minima or maxima within a season or year); significant changes are detected by evaluating alterations in these percentiles over time (e.g. a time series of annual maximum daily streamflow). More generally, magnitudes can also be described through the spatial extent of an event or via metrics characterising its intensity or flashiness, such as the time-to-peak.

The magnitude or intensity of precipitation (in mm, Figure 3a) can be assessed using various metrics. Many of these are part of the 'ETCCDI' indices that were proposed in 2002 by the Expert team on climate change detection and indices (ETCCDI) (see e.g. Frich et al., 2002; Zhang et al., 2011). Precipitation metrics include the maximum depth of precipitation for a given (1-, 3-, 6-, or 24-hour) duration within a given month/year; the maximum 1- or 5-day precipitation accumulation (respectively Rx1day, or Rx5day) over a month/year (Sun et al., 2020a); the percentage of a daily total that fell in the monthly maximum 1-hour precipitation (or some other period); the 90^{th}, 95^{th} or 99^{th} percentile precipitation amount (over 1-, 3-, or 6-h) during a month/year - or specifically on wet days (Moberg and Jones, 2005); or the *total* precipitation accumulated from hours exceeding specified percentiles over a month/year (e.g. Donat et al., 2013). For a complete state-of-the-art review on the 'probable maximum precipitation' (PMP) concept, see Salas et al. (2020). The PMP can be computed via hydrometeorological, statistical, grid-based and site-specific approaches using both stationary and nonstationary methods (e.g. Lee and Singh, 2020). A recent global analysis of observed rainfall data indicated that extreme precipitation has increased at about two thirds of stations – a significantly greater proportion than can be expected by chance (Sun et al., 2020a).

Percentiles of daily streamflow distribution are commonly used for floods (Figure 3b). For example, the Q_{90} (the 90^{th} percentile of the distribution, i.e. the flow that is exceeded 10% of the time, confusingly referred to as 'Q_{90}' in North America and 'Q_{10}' in the British and Irish islands), Q_{95}, Q_{99} (flow that is exceeded 3.65 days per year, on average), $Q_{99.9}$ (0.37 days per year) or AMAX (the annual maximum streamflow). When considering more extreme events than AMAX, hydroclimatic extremes are commonly expressed as 1-in-20, 1-in-50, or the 1-in-100 year events (e.g. Milly et al., 2002). Hydrograph analyses may also be used to extract metrics such as flood volume and duration (Figure 3b). Other indicators describe the flashiness of an

https://doi.org/10.5194/hess-2020-576
Preprint. Discussion started: 14 November 2020

Hydrology and Earth System Sciences
Discussions

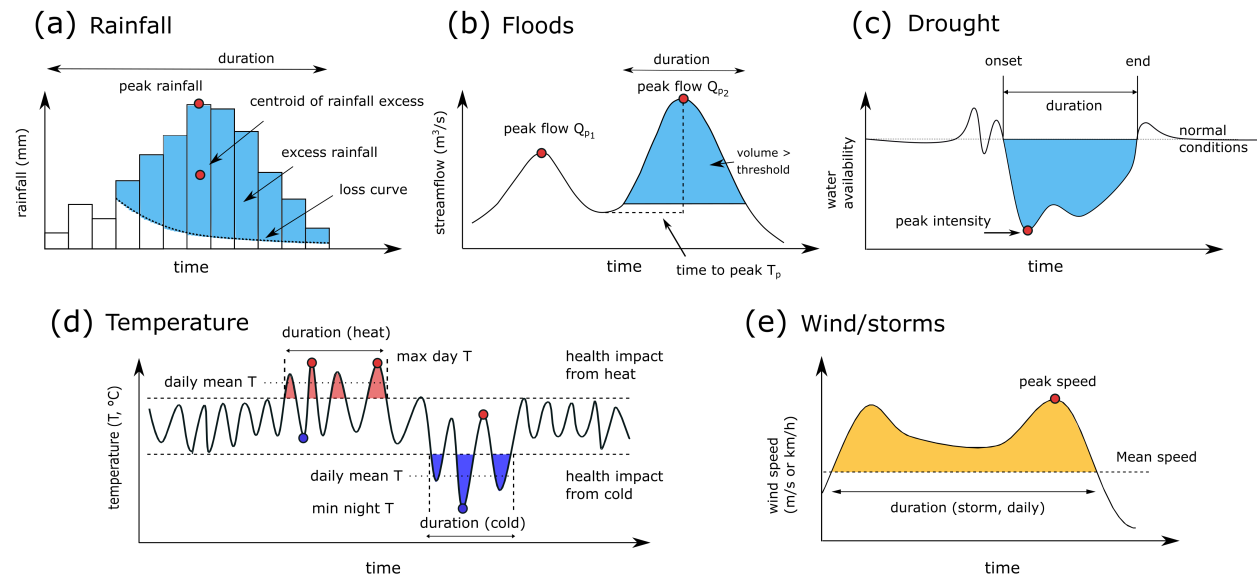

Figure 3. Metrics employed for evaluating five types of hydroclimatic extremes: (a) Precipitation hyetograph (including infiltration loss curve indicating excess rainfall); (b) Flood hydrograph; (c) Drought; (d) Temperature; (e) Wind (storms). All these variables can be described using indicators of event duration and magnitude (peak intensity).

event, such as the time-to-peak, which is defined as the total number of hours starting from the sharp rise of the hydrograph until the peak discharge. Synthetic design hydrographs (SDHs) may be used to test the sensitivity of the peak, volume and shape of the flood hydrograph to the flood generating mechanism and catchment properties (e.g. Brunner et al., 2018; Yue et al., 2002a).

The spatial extent of a flood or drought event can be described using metrics such as the number of catchments flooding simultaneously (Uhlemann et al., 2010) or the ’flood synchrony scale’ (FSS), which evaluates the largest radius around a stream gauge where more than half of the surrounding stream gauges also record flooding within the same week (Berghuijs et al., 2019a). Studies are increasingly evaluating the spatial dependence of flooding across multiple basins by considering meteorological, temporal and land surface processes leading to simultaneous flooding across varying spatial scales (e.g. Brunner et al., 2020; Kemter et al., 2020; Wilby and Quinn, 2013).

Droughts (Figure 3c) differ from other extreme weather because they develop more slowly and last longer; they are broadly defined as “a sustained period of below-normal water availability” (Tallaksen and Van Lanen, 2004), sometimes referred to as a “creeping phenomenon” (Mishra and Singh, 2010; Wilhite, 2016). These characteristics make it more difficult to assess nonstationarity as there are fewer events to compare over time plus drought onset and termination are challenging to pinpoint (Parry et al., 2016). Not all droughts are defined by aridity; and rainfall deficit alone does not imply a drought (Van Loon, 2015). Instead, a combination of factors in the hydrological cycle interact to yield below normal conditions (Figure 3c). Drought has been typically classified as: meteorological (precipitation deficit); hydrological (surface and subsurface water deficit relative to local water uses); agricultural (declining soil moisture and crop failure); or socioeconomic (failure of water resources system to meet demand) (Van Loon, 2015). Since the definition of "normal conditions" depends on spatial and temporal scales, drought anomalies are typically defined locally using composite indicators, including precipitation and temperature. Various metrics exist to measure drought stress, and they either reflect deficits in precipitation, or combined metrics of precipitation, temperature and evaporation. The World Meteorological Organisation (WMO) recommends the use of the Standardized Precipitation Index (SPI), which reflects standard deviations from normal rainfall (WMO, 2016; McKee et al., 1993). Other well-known indices that rely on monthly precipitation include the Palmer Drought Severity Index (PDSI) (Palmer, 1965), deciles (Gibbs and Maher, 1967), and the rainfall anomaly index and its modified version (RAI/mRAI) (e.g. Hänsel et al., 2016). Some metrics combine variables. PDSI accounts for soil moisture, temperature, and precipitation anomalies and therefore, reflects both sides of the water balance. The Standardized Precipitation Evapotranspiration Index (SPEI) can be interpreted similarly to SPI, but reflects both evaporative demand and precipitation inputs to a system (Vicente-Serrano et al., 2010). The spatial characteristics of drought have also become increasingly relevant, as more studies examine their areal extent over time, using metrics like spatial patterns of drought intensity mapped across livelihood zones over time (e.g., Leelaruban and Padmanabhan, 2017; Mekonen et al., 2020).

For temperature extremes (Figure 3d), studies may monitor the hottest/coldest day, warmest/coolest night, or the extreme temperature range (TXx-TNn, °C) (e.g. Donat et al., 2013). Percentiles of the distribution are also commonly assessed (Zhang et al., 2005; Kjellström et al., 2007), while some authors work with combined temperature-humidity metrics (Matthews et al., 2017; Raymond et al., 2020b; Knutson and Ploshay, 2016), which offer a more complete measure of atmospheric heat content (Pielke Sr et al., 2004; Peterson et al., 2011; Matthews, 2020), and may therefore be more closely aligned with levels of thermal stress felt by humans (Mora et al., 2017; Matthews, 2018). Other approaches include an emphasis on duration by focussing on heatwaves, defined as periods of consecutive days when heat is higher than normal (Perkins and Alexander, 2013). This very broad categorization has seen a plethora of thresholds (both absolute-value and percentile-based) and metrics (temperature and combined temperature-humidity indices) applied to heatwave studies, with the choice shaped by interests in potential impacts (Xu et al., 2016).

Changes in the magnitude of extreme wind events (Figure 3e) may also be tracked using wind speed percentiles such as the 90^{th}, 95^{th}, 98^{th} and 99^{th} seasonal or annual percentiles (e.g. Donat et al., 2011; Young and Ribal, 2019; Wang et al., 2009). Wind intensity (in metres per second) may be explicitly measured over two-minute sustained periods, or three-second gust periods (Pryor et al., 2014). Wind events may also be inferred from gradients in sea-level pressure fields (Jones et al., 2016; Matthews et al., 2016b). Winds associated with Western Hemisphere tropical cyclones are described using storm scales such as the Saffir-Simpson hurricane wind scale (SSHWS) (e.g Elsner et al., 2008; Karl et al., 2008). This classifies wind intensity into five storm categories: one (119-153 km/h), two (154-177 km/h), three (178-208 km/h), four (209-251 km/h) and five (>252 km/h). Cyclones and typhoon magnitudes are equally described using metrics such as the seasonal mean lifetime peak intensity, intensification rate, and intensification duration (Mei et al., 2015).

https://doi.org/10.5194/hess-2020-576
Preprint. Discussion started: 14 November 2020

2.2 Are extreme events occurring more frequently than before?

A second broad category of nonstationarity 'symptoms' is the frequency of events. Many metrics are used to describe changes in the frequency of hydroclimatic extremes, such as annual exceedance probabilities and counts of occurrences above or below thresholds (i.e. 'peak over threshold' (POT) approaches such as the number of days or hours exceeding a threshold in a given year). In reality, the magnitude and frequency of extremes are closely related, such that when magnitudes increase, one can also generally expect to find more peaks over a given threshold (see Figure 4). However, frequency-based metrics generally enable better detection of changes in extremes than magnitude-based metrics (e.g. Mallakpour and Villarini, 2015) because they often reflect a larger sample of data and are less prone to measurement errors. For example, while block-maxima approaches often include just one value per year/season, POT approaches count the total number of exceedances above a threshold. This fact is exploited by those using documentary evidence to evaluate flood frequency (Macdonald et al., 2006). The thresholds for detection of changes in frequency should be set high enough to describe a meaningful extreme event, yet low enough to compile an adequate sample size.

For precipitation, multiple methods exist for the selection of the most appropriate threshold (Caeiro and Gomes, 2016). Thresholds are generally chosen based on the local precipitation distribution such as the 95th, 98th, 99th or 99.5th percentile of rain over a 1-, 6-, 12-, or 24-h period (e.g. Wi et al., 2016). Percentile or fixed thresholds (such as the 10 mm or 20 mm daily total, R10mm or R20mm) are then used to count monthly/annual days with heavy precipitation exceeding or equalling these values. Alternatively, a mean residual life plot (an exploratory graphical approach) can be used to select a suitably high threshold (e.g. Coles, 2001). These thresholds can be calculated for individual years or using the entire multi-year record. For an overview of threshold selection methods, see Anagnostopoulou and Tolika (2012). An alternative approach for estimating the probability of intense rainfall events is the changing likelihood of an historical precipitation analogue (Matthews et al., 2016a).

The frequency of temperature extremes is assessed using metrics such as the percentage of the time when the daily minimum or maximum temperature is below or above a given percentile, such as the total annual count of ice/frost days (ID/FD) where the daily minimum temperature is below 0°C (e.g. Donat et al., 2013). Mwagona et al. (2018) observed changes in cold/warm night frequency at 116 stations in Northeast China. They report a decrease in cold night frequency during winter and spring, while warm night frequency increased primarily in summer. Similar metrics are used to describe changes in the frequency of heatwaves (Perkins-Kirkpatrick and Lewis, 2020) or growing season days, such as the number of days with plant heat stress (with maximum temperature exceeding 35°C) (e.g. Rivington et al., 2013). Alternatively, the accumulated frost (sum of degree days where the minimum temperature is below 0°C) (e.g. Harding et al., 2015) may be linked to crop yields during different growth phases.

The frequency of wind extremes is often assessed in terms of the number of wind storms (e.g. Wild et al., 2015) or cyclones (e.g. Matthews et al., 2016b) in a given period. The number of events can be estimated using tracking algorithms, ranging from simple algorithms based on the mean sea-level pressure values for cyclone activity (Murray and Simmonds, 1991; Donat et al., 2011) and the exceedance over the 98th percentile wind speeds for wind storms (Leckebusch et al., 2008) to more complicated

https://doi.org/10.5194/hess-2020-576
Preprint. Discussion started: 14 November 2020

Hydrology and Earth System Sciences
Discussions

tracking of 'sting jets' (Hart et al., 2017). The frequency of extreme winds may be quantified from reanalysis data, noting that the trend magnitude and sign can be sensitive to the reanalysis product (Befort et al., 2016; Wohland et al., 2019; Torralba et al., 2017a).

POT methods are widely used to evaluate changes in the frequency of hydroclimatic events. These methods require selection of a reference threshold (a magnitude) and a period (e.g. one week) to de-cluster independent events. This is a common challenge for most hydroclimatic extremes (Figure 3). For floods, many studies apply a somewhat arbitrary streamflow threshold that is on average exceeded twice per year (e.g. Hodgkins et al., 2019; Slater and Villarini, 2017a). In practice, alternative thresholds may be equally valid, such as the number of days when water levels exceed official flood thresholds (Slater and Villarini, 2016). However, setting a lower threshold means that events are less likely to be independent and/or of practical significance. 'Clusters' of consecutive events are, thus, further de-clustered by specifying a period between events. For instance, Wi et al. (2016) separated extreme rainfall events by an interval of 7 days but other requirements have been proposed to ensure independence. For example, Lang et al. (1999, p.105) highlight that in 1976, the U.S. Water Resources Council imposed a separation between floods events of "*at least as many days as five plus the natural logarithm of square miles of basin area*", including a drop between consecutive peaks 'below 75% of the lowest of the two flood events'. The time between independent events depends on the catchment size, as a longer duration is expected in larger catchments because of the slower draw-down of hydrograph limbs. In practice, such thresholds are far too short to adequately distinguish truly independent events, given the timescales of hydroclimatic variability which are often greater than a year. Similar concerns apply when isolating successive drought events (Bell et al., 2013; Thomas et al., 2014; Parry et al., 2016). Metrics to identify drought termination and subsequently drought independence include storage deficit methods that quantify the volume of water in relation to normal water storage conditions (Thomas et al., 2014), and more generally, the return from maximum negative anomalies to above-average conditions (Parry et al., 2016). Hence, specification of the threshold and de-clustering technique make POT approaches more complicated to implement than block maxima approaches, where only one extreme per 'block' (unit of time) is selected. The more flexible selection of extremes and larger sample size of frequency-based (i.e. POT) approaches may be preferred when record length is a limiting factor, whereas simpler magnitude-based (i.e. block maxima) approaches may be preferred when longer records are available.

More severe hydroclimatic extremes (such as 1-in-10, -20, -50, -100, or -200-year events) are typically evaluated using return periods (also denoted as the 'expected waiting time'). Alternatively, annual exceedance probabilities (AEPs) define the average waiting time for an event of a given magnitude to occur, or the likelihood of attaining or exceeding a given threshold in any particular year. Other metrics have been proposed for engineering design. Reviews by Salas et al. (2018) and François et al. (2019) highlight ongoing disagreements about the utility of nonstationary methods for the design of engineering structures. As the uncertainties of nonstationary model structure may exceed that of stationary models (Serinaldi and Kilsby, 2015), specific strategies are required to manage the consequences of those uncertainties (François et al., 2019). There are thus ongoing debates about which concepts and methods are most appropriate for estimation of extremes, such as the return period, risk, reliability or Equivalent Reliability (ER), Design Life Level (DLL) or Average Design Life Level (ADLL), and Expected Number of Exceedances (ENE) (e.g. Read and Vogel, 2015; Rootzén and Katz, 2013; Yan et al., 2017; Salas and Obeysekera, 2014).

https://doi.org/10.5194/hess-2020-576
Preprint. Discussion started: 14 November 2020

Hydrology and Earth System Sciences
Open Access
Discussions

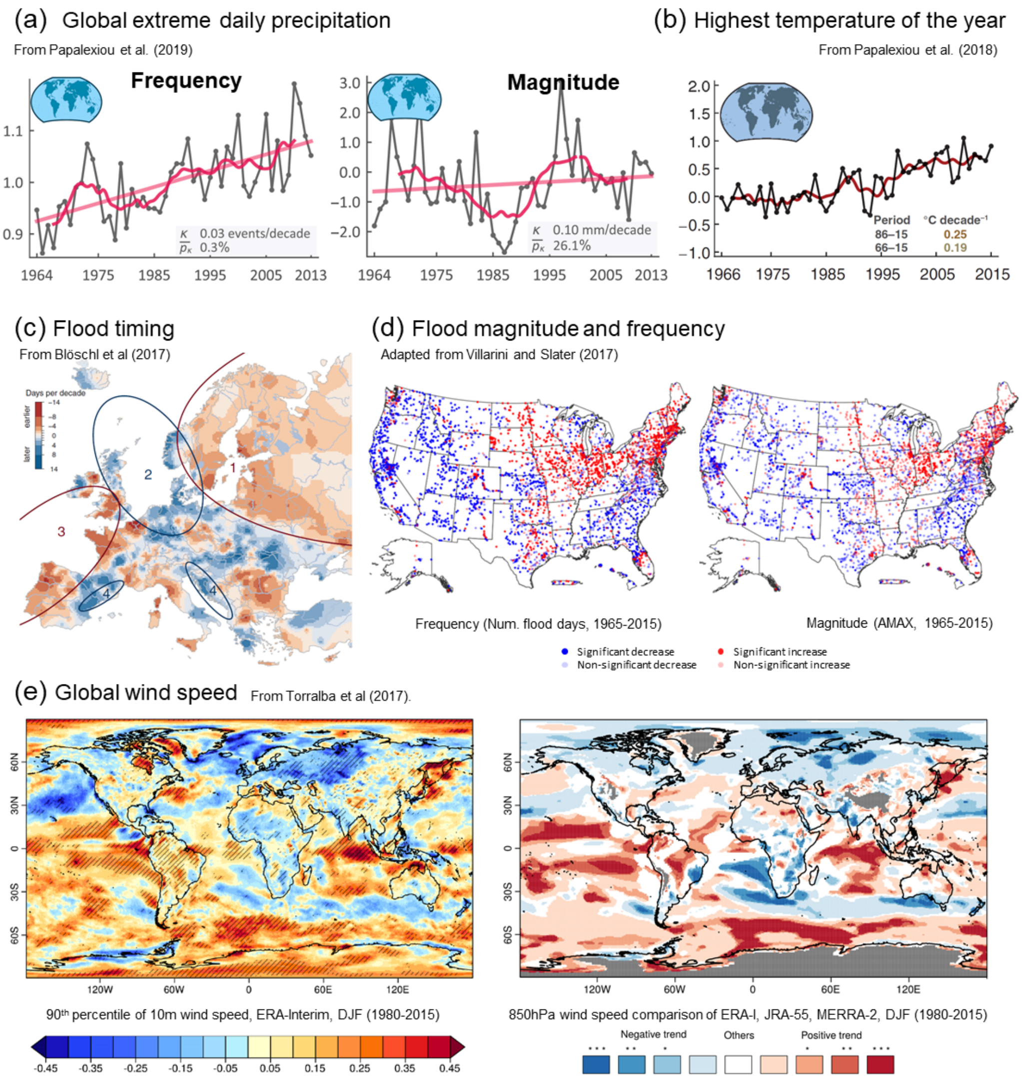

Figure 4. Examples of trends in magnitude, frequency, and timing of hydroclimatic extremes. (a) Trends in extreme daily precipitation frequency (left: events) and magnitude (right: annual mean extreme daily precipitation anomaly, mm) over the globe (Papalexiou and Montanari, 2019). (b) Trend in the highest temperature anomaly of the year (°C) over the globe (baseline period: 1970-1989); red line indicates 5-year moving average (Papalexiou et al., 2018). (c) Trends in flood timing across Europe in days/decade, 1960-2010 (Blöschl et al., 2017). (d) Trends in flood frequency (number of flood days) and magnitude (annual maxima) across the USA, 1965-2015 (significance level set to 5%, Villarini and Slater, 2017). (e) Global trends in wind speed. Left: linear trend (m/s per decade) of ERA-Interim 90th percentile of 10 m wind speed for December-January-February (DJF); significant trends hatched. Right: 850 hPA wind speed trends produced by ERA-I, JRA-55 and MERRA-2; where blues (reds) indicate level of agreement between reanalyses about negative (positive) trends (Torralba et al., 2017a).

https://doi.org/10.5194/hess-2020-576
Preprint. Discussion started: 14 November 2020

For instance, return period metrics may exhibit limitations in the case of time-correlated hydroclimatological extremes, so alternatives such as the 'equivalent return period' (*ERP*, i.e. "the period that would lead to the same probability of failure pertaining to a given return period T in the framework of classical statistics (independent case)" Volpi et al., 2015), may be preferred.

It is not just individual characteristics of weather and water extremes that can change over time but also the interdependence between different characteristics, such as frequency, magnitude, and volumes. Brunner et al. (2019) assessed future changes in flood peak-volume dependencies and found that the interdependence between variables may change more strongly that the individual variables themselves. This interdependency also applies to other variable pairs jointly of interest such as drought duration and deficit or precipitation intensity and duration. Recognising the interdependence between magnitude and frequency, many studies employ intensity-duration-frequency (IDF) metrics, which describe both the magnitude and frequency at once. It has recently been shown that generalized extreme value (GEV) distribution parameters scale robustly with event duration at the global scale ($R^2 > 0.88$), hence a universal IDF formula can be applied to estimate rainfall intensity for a continuous range of durations, including at the sub-daily scale (Courty et al., 2019). There is growing interest in the nonstationarity of IDF curves (Cheng and AghaKouchak, 2014; Ganguli and Coulibaly, 2017) as well as the implications of this nonstationarity for compound hydroclimatic extremes globally (AghaKouchak et al., 2020).

2.3 Is the timing of events changing?

Nonstationarity in the timing and seasonality of hydroclimatic extremes has been examined far less than trends in magnitude and frequency. Timing and seasonality provide information that is relevant for the management of water resources and analysis of underlying drivers of change. For instance, the start of field operations for farming may be estimated as the day of the year when 'the sum of average temperature from 1st January exceeds 200°C' (e.g. Rivington et al., 2013; Harding et al., 2015). Similarly, the start of the growing season may be measured as the first of five consecutive days with average temperature exceeding 5°C (Rivington et al., 2013). Changes in these indicators of hydro-climatic extremes may have substantial impacts, e.g. for crop yields. Additionally, changes in timing and seasonality can also affect impacts of extreme events. For example, the risk from compound tropical cyclones and heatwaves is sensitive to the seasonal cycles in tropical cyclone probability (which peaks in late summer) and extreme heat (mid summer). A greater frequency of tropical cyclones earlier in summer, or more extreme heat late in summer, would increase the risk of compounding and attendant impacts (Matthews et al., 2019).

Changes in the timing of seasonal streamflows are typically assessed using the centre of volume (CV) date (Court, 1962) or mean date of flood occurrence (mean flood day, MFD). For example, Hodgkins and Dudley (2006) assessed changes in flood timing over the conterminous USA from 1913-2002 using the winter-spring CV dates. They found that a third of stations north of 44 degrees had significantly earlier flows, likely related to changes in winter and spring air temperatures affecting winter snowpack. The MFD has been used to assess changes in streamflow timing in specific countries such as Wales (Macdonald et al., 2010) and Spain (Mediero et al., 2014). Probabilistic methods for identifying flood seasonality and their trends Cunderlik et al. (2004) have been applied in Canada (Cunderlik and Ouarda, 2009) and the northeastern United States (Collins, 2019). In Europe, an analysis of 4,262 streamflow stations in 38 countries used the date of occurrence of the highest annual flow peak

https://doi.org/10.5194/hess-2020-576
Preprint. Discussion started: 14 November 2020

Hydrology and Earth System Sciences Discussions
Open Access
EGU

to assess changes in flood timing (Blöschl et al., 2017). This showed significant changes in the seasonal timing of floods at the regional scale: in northeastern Europe, 81% of stations had shifted towards earlier floods (by 8 days per 50 years); in western Europe, 50% of stations had shifted towards earlier floods (by 15 days per 50 years); and around the North Sea, 50% of the stations had shifted towards later floods (by 8 days per 50 years), as seen in Figure 4d (Blöschl et al., 2017).

Nonstationarity in the timing of extreme precipitation has also been used to better understand the causal factors of extremes. For instance, Gu et al. (2017) examined shifts in the seasonality and spatial distribution of extreme rainfall over 728 stations in China. They found that alterations in rainfall seasonality are likely being driven by changes in the pathways of seasonal vapor flux and tropical cyclones. Others have examined shifts in the seasonality of future large-scale global precipitation and temperature extremes from model projections such as CMIP5 (e.g. Zhan et al., 2020) (see Section 7.2 for discussion of prediction and projection).

3 Essential data considerations before detecting and attributing nonstationarity

3.1 Spurious nonstationarities: the issue of data quality

Confidence in nonstationarity detection, attribution, and prediction rests on confidence in the homogeneity and quality of input data – including primary hydro-meteorological series from individual observations, accompanying meta-and qualitative information. Data quality issues tend to be prevalent with measurement of extremes. Hydroclimatologists increasingly need to be aware of data quality issues associated with homogeneity of remotely sensed data and their derivatives. For example, inaccuracies may arise from orbital drift (Weber and Wunderle, 2019), inference of precipitation from vegetation in data sparse regions (Xu et al., 2015), or changing land-surface reflectance such as snow cover over mountainous regions (Karaseva et al., 2012). Some of the most common sources of data errors and biases that reduce homogeneity/ cause nonstationarity in ground-based information over time-scales of years to decades are: site or instrument changes, biases and drift in field procedures (time of sample, preferred values), unstable rating curves and cross-sections, changes in network density/ cover, post-processing and archiving (unit changes) (Wilby et al., 2017). For example, the England and Wales Precipitation series is a specific example of spurious nonstationarity (a long-term trend towards wetter winters) arising from a combination of climate drivers (cold winters with more snowfall in the early nineteenth century) with non-standard rain gauges before the mid-1860s and snowfall under-catch, giving an apparent increase in winter precipitation (Figure 5a, Murphy et al., 2020b). Gridded products and reanalysis data are also not immune from such data quality issues and are further affected by time-space variations in raw data inputs, version updates and system drift (Sterl, 2004; Ferguson and Villarini, 2014).

Detecting spurious nonstationarities within raw data should be one of the first steps when evaluating time series that have yet to be quality-controlled. Tests for homogeneity have been proposed for uncovering such issues (see Figure 5b). Common techniques for assessing data quality range from visual inspection/expert judgment, to formal statistical tests (e.g. Chow test, Buishand range test, Pettitt test, standard normal homogeneity tests). For meteorological variables relative homogeneity tests are possible where appropriate networks of observations are available (e.g. HOMER (Mestre et al., 2013)). However, such

Hydrology and Earth System Sciences
Open Access
Discussions

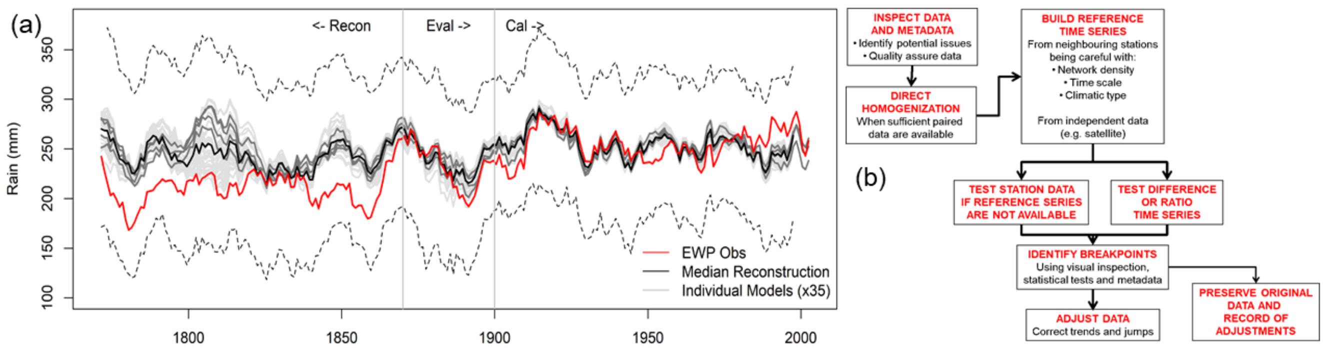

Figure 5. Homogeneity of observational records. (a) An example of spurious non-stationarity in decadal mean observed (red line) winter precipitation for England and Wales due to inhomogenous records. The ensemble median and individual reconstructions (black and grey lines) do not exhibit nonstationarity. Dashed lines indicate the 95% confidence intervals for the reconstructions; from Murphy et al. (2020b). (b) An approach to data homogenization for monthly to annual climate records, adapted from Aguilar et al. (2003).

techniques are often limited to evaluating changes in the mean rather than extremes (Peña-Angulo et al., 2020; Ribeiro et al., 2016; Yosef et al., 2019).

An important step toward detecting and attributing nonstationarity would be providing better meta-data about measurement practice and any changes in observational techniques, as well as guidance on basic quality assurance approaches and the procurement and servicing of datasets. Observations are the foundation for understanding hydroclimatic change. Unfortunately, datasets of essential variables (precipitation, evapotranspiration, discharge etc.) are typically disbursed across various global, regional, and national archives containing different variables and time scales, in varied formats. For streamflow records, changes in the rating quality are rarely noted. Large, multi-country databases such as the Global Runoff Data Centre (GRDC, https://portal.grdc.bafg.de/) are vital for providing an overview of nonstationarities at continental and global scales but do not provide information on streamflow data quality. Accordingly, there are limitations on what can be said in global studies, compared with local knowledge. Hydrologists could take the lead from climatologists and their ambitious aims to create integrated datasets of essential climate variables for understanding and detecting change (e.g. Thorne et al., 2017).

3.2 Record length and completeness

Definitions of nonstationarity depend on the information content of the data, which is inherently linked to how observations sample the true distribution, and therefore to the available record length. Given that observation networks commonly used in hydroclimate research tend to extend from the mid 20th Century (in the best cases), our ability to estimate the true distribution is limited. Consequently, early studies and guidance on what constitute ‘long records’ for trend detection are likely too short: e.g. Kundzewicz et al. (2005) claimed that 50 years of data were sufficient for trend detection. However, key modes such as variations in Atlantic sea surface temperatures (SSTs) are known to vary over periods longer than 50 years (McCarthy et al., 2015; Sutton and Dong, 2012) and to affect precipitation and temperature over multi-decadal timescales. For instance, the

https://doi.org/10.5194/hess-2020-576
Preprint. Discussion started: 14 November 2020

Hydrology and Earth System Sciences
Discussions

North Atlantic was particularly cold during the middle of the climate normal period 1961-1990 due to the Atlantic Multidecadal Oscillation/great saline anomaly (Dickson et al., 1988), with and perhaps a global "change point" in climate during the late 1960s (Baines and Folland, 2007). Hence, even 50 years of data is insufficient to robustly detect true nonstationarities because the start and end dates of records may substantially affect the sign (direction) and magnitude of trends detected, especially in records that exhibit periodicity (Harrigan et al., 2018). Ideally, the length of observation period should be appropriate for the aims of analysis, plausible physical drivers, and properties of the underlying data. Highly variable time series (such as streamflow annual maxima) require a longer period of time for a significant signal to emerge, than less variable time series based on more data (e.g. peak over threshold time series). In places where series have low signal-to-noise ratios, the time required to detect plausible trends (e.g. in precipitation, evapotranspiration, and discharge extremes) can be centuries long (e.g. Ziegler et al., 2005; Wilby, 2006). Faster detection may be possible using seasonal, rather than annual time series (e.g. Ziegler et al., 2005).

The mismatch between the temporal scales of drivers of climate variability versus the availability and quality of observations can also result in misleading conclusions. Climate is nonstationary by definition, but there is a real danger of identifying the incorrect driver of change when record length is short. For example, numerous studies have reported decreasing precipitation in Mediterranean regions since the 1960s (Longobardi and Villani, 2010; Gudmundsson and Seneviratne, 2016), with some attributing this decline and corresponding increase in drought frequency to anthropogenic forcing in the Mediterranean basin (e.g. Barkhordarian et al., 2013; Gudmundsson and Seneviratne, 2016; Hoerling et al., 2012). Vicente-Serrano et al. (2019) show that, when viewed in the context of rescued and quality assured data beginning in the mid–19th Century, these recent trends in precipitation are within the range of longer-term variability. Hence, trends depend intrinsically on the period of record and indeed on the quality of the dataset; without sufficient record length, false attribution statements may arise with potentially significant management implications.

The challenges posed by the lack of available long-term observations have prompted some to leverage advances in data rescue and historical climatology to extend discharge series back in time (e.g. O'Connor et al., 2020; Smith et al., 2017; Bonnet et al., 2020). Although such datasets lengthen the period available for analysis and better reflect ranges of variability in extremes such as drought (Murphy et al., 2020a), they are subject to limitations from changes in measurement practice, decreasing density of observations in early records, and a lack of consideration of issues such as changes in land cover and shifts in channel capacity (Slater et al., 2015).

Finally, the issue of data completeness is equally important. Gaps in extreme hydroclimatic time series affect the detection rate of significant trends. Detection rates are lower in records that have larger gaps and shorter records than those with less change (lower regression slopes) and fewer gaps, and/or when the data gaps are located towards the beginning or end of a time series (Slater and Villarini, 2017a).

In many regions, temporal and spatial data sparsity is likely to remain a key issue, hindering robust detection, attribution, and prediction of water and climate nonstationarities. Therefore different trend detection and attribution methods should be considered (e.g. using lower quantiles and peak-over-threshold methods; see Section 4), while implementing holistic 'multiple working hypotheses' approaches (Chamberlin, 1890; Harrigan et al., 2014) to avoid overlooking potential drivers of change.

https://doi.org/10.5194/hess-2020-576
Preprint. Discussion started: 14 November 2020

Hydrology and Earth System Sciences
Discussions

4 Detection of nonstationarity

Detection of nonstationarity in hydroclimatological extremes requires a sound examination of the data before applying any statistical tests, which broadly seek to detect two types of nonstationarity (see Figure 1b): monotonic change (trends) and step-changes (change points). Such changes can be considered as symptoms of nonstationarity if they represent a significant departure from normality within a long-term record. Nonstationarity may be detected either in individual time series (point-based analysis) or in larger ensembles of stations (spatially-coherent trends) (Hall et al., 2013). Regional coherence generally provides greater confidence in spatial patterns of change (Prosdocimi et al., 2019). Here, we provide an overview of existing methods employed in the fields of weather and water extremes. For a description of change-detection methods in hydrology we refer the reader to Helsel et al. (2020), and for floods to Villarini et al. (2018).

4.1 Regression-based methods for detection of incremental change

The detection of trends in hydroclimatic time series generally employs two key approaches: detection of trends in magnitude (e.g. quantiles or *block maxima*, such as the annual maxima, AMAX) or frequency-based methods (e.g. use of point process modelling frameworks to model the peaks-over-threshold (POT) series, also referred to as *partial-duration series* (PDS)) (e.g. Coles, 2001; Salas et al., 2018).

As hydroclimatic extremes typically have skewed distributions, one of the frequently used approaches for detecting monotonic trends in magnitudes is to use a non-parametric test such as the Mann-Kendall (MK) test (Mann, 1945; Kendall, 1975) which evaluates whether the central tendency or median of the distribution changes monotonically over time (see Helsel et al., 2020). The test statistic, Kendall's τ, is a rank correlation coefficient which ranges from -1 to +1 (e.g. see Figure 4c, left panel). For instance, Westra et al. (2013) evaluated trends in annual maximum daily precipitation at 8,326 precipitation stations with at least 30 years of record and found increases at approximately two-thirds of these stations. An additional test, called the Theil-Sen slope estimator, has often been used alongside MK to estimate the slope of the trend over time (Sen, 1968; Theil, 1992; Hipel and McLeod, 1994). Different versions of the MK test exist to detect seasonal and regionally-coherent trends over time (see Helsel et al. (2020) for details and examples).

Other studies also apply ordinary least squares (OLS) linear regression to estimate trends in precipitation (e.g. Figure 4a, Papalexiou and Montanari, 2019) temperature (e.g. Papalexiou et al., 2018), and flood flows (e.g. Hecht and Vogel, 2020). Practical advantages to using OLS methods include its ease of use and expression of uncertainty, graphical communication, and usability for providing decision-relevant information (e.g. Hecht and Vogel, 2020). In cases where the assumptions of OLS are not met, quantile regression (QR) may be used; instead of estimating the conditional mean of the response variable, QR considers different conditional quantiles (including the median) of a distribution. QR has been used for precipitation trends (e.g. Tan and Shao, 2017), air temperature (e.g. Barbosa et al., 2011), surface wind speed (e.g. Gilliland and Keim, 2018), and flood trends (e.g. Villarini and Slater, 2018), and is also regularly employed to investigate scaling properties between hydroclimatological variables; see section 5.1 (e.g. Wasko and Sharma, 2014).

Hydrology and Earth System Sciences
Discussions

When the empirical distribution of hydroclimatic extremes is known, many prefer to select an appropriate distribution and evaluate how the distribution parameters vary as a function of covariates such as time (e.g. Katz, 2013). For example, the non-stationary Generalized Extreme Value (GEV) or Gumbel (GU) distributions are widely used to detect trends in annual/seasonal maxima such as floods (e.g. Prosdocimi et al., 2015), precipitation (e.g. Gao et al., 2016) and wind (e.g. Hundecha et al., 2008), while the Poisson (PO, e.g. Neri et al., 2019) or negative binomial (NBO, e.g. Khouakhi et al., 2019) distributions are preferred for discrete data (e.g. counts of days over thresholds). These distributions can be fit to the data either with constant parameters (stationary case) or with the parameters expressed as a function of time (nonstationary case) (Katz, 2013). In the nonstationary case, a time covariate is employed to detect changes in the parameters of the distribution – as illustrated in Figure 6. Criteria for model selection, such as the Akaike Information Criterion (AIC) or Schwarz Bayesian Criterion (SBC - also known as the Bayesian Information Criterion, BIC) can then be used to determine whether the stationary or nonstationary model is the better fit (e.g. Figure 6a). If the nonstationary model performs better (in terms of goodness-of-fit and uncertainty) then the time series may be considered as nonstationary, pending sufficient record length (see section 3.2).

Increasingly, distributional regression modelling frameworks such as Vector Generalized Linear and Additive Models (VGLM/VGAM) or Generalized Additive Models for Location, Scale and Shape (GAMLSS) are being chosen for their flexibility in evaluating nonstationarity of hydroclimatic extremes (e.g. Serinaldi and Kilsby, 2015). These frameworks are a generalization of generalized linear models (GLMs) allowing a broader range of distributions as well as different relationships between parameters and explanatory variables (linear, nonlinear, or smooth nonparametric). GAMLSS models, for instance, can have up to four parameters μ, σ, ν, and τ, which allow modelling of the location (mean, median, mode), scale (spread: standard deviation, coefficient of variation) and shape (skewness and kurtosis) of a distribution. An example of nonstationarity detection is shown in Figure 6. Here, two nonstationary (with time-varying parameters) and two stationary (constant parameters) models are fit using both the Gamma and Weibull distributions to observed time series of instantaneous (15-minute) peak maxima (Figure 6a). In the nonstationary case, the (μ and σ) model parameters both depend linearly on the time covariate (in years). A logarithmic link function is employed to ensure the distribution parameters remain positive. The goodness-of-fit of both stationary and nonstationary models is assessed using SBC (Figure 6a). A detrended quantile-quantile- (worm) plot showing the residuals for different ranges of the explanatory variable(s) can also be used to diagnose model fit (Figure 6b). The model fit is satisfactory if the worm is relatively flat and if data points lie within the confidence intervals. In the case of the River Ouse, we find the nonstationary Gamma model is the best-fitting model, and that both the central tendency (mean: Figure 6c) and variance (Figure 6d) are increasing over time. Here, the 50-year flood (specific discharge) increased from 0.123 $m^3/s/km^2$ in 1900 to 0.183 $m^3/s/km^2$ in 2018. GAMLSS methods have been applied for different hydroclimatic extremes. For example, Bazrafshan and Hejabi (2018) developed a Nonstationary Reconnaissance Drought Index (NRDI) to assess drought nonstationarity in Iran, and found large differences between the NRDI and a traditional RDI for timeframes longer than 6 months. Sun et al. (2020b) also evaluated changes in a Nonstationary Standardised Runoff Index (NSRI) using GAMLSS over the Heihe River Basin in China. Importantly, the covariate in a GAMLSS nonstationary model is often time (e.g. Villarini et al., 2009a; Serinaldi and Kilsby, 2015), but may also include other physical drivers such as climate modes (e.g. Villarini and Serinaldi, 2012), urban or

https://doi.org/10.5194/hess-2020-576
Preprint. Discussion started: 14 November 2020

agricultural land cover (e.g. Villarini et al., 2009b; Slater and Villarini, 2017b), or other hydro-climatic variables such as dew point temperature (e.g. Lee et al., 2020).

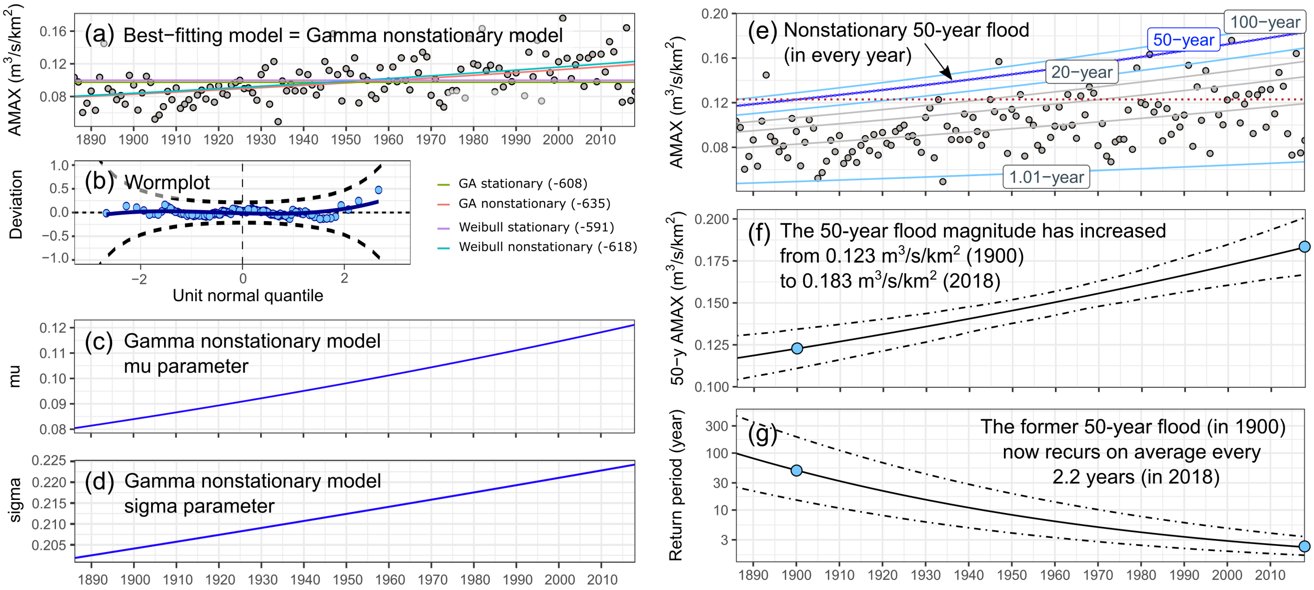

Figure 6. Example workflow for detecting trends in the magnitude and return period of extremes using distributional regression (GAMLSS). The example shows a considerable increase in flood magnitude and frequency in the River Ouse at Skelton, UK, over 130 years. (a) Two nonstationary and two stationary models are fit to the time series of 15-minute peak maxima (black circles: specific discharge, in $m^3/s/km^2$); colour lines indicate the 50% probability (centile) for each model. The best-fitting model is the nonstationary Gamma model (lowest Schwarz Bayesion Criterion, indicated in brackets). (b) The wormplot indicates satisfactory model fit (dashed lines indicate 95% confidence interval). (c) Time series of the μ parameter for the nonstationary Gamma model. (d) Time series of the σ parameter (note: the mean of the distribution is equal to μ and the variance is equal to $\sigma^2\mu^2$). (e) Centile curves for the best-fitting nonstationary model are shown for the 1st, 50th, 80th, 90th centiles (from bottom to top), 95th (20-year flood), 98th (50-year flood), and 99th (100-year flood). The dotted horizontal red line indicates the value of the 50-year flood in 1900. (f) The 50-year flood estimated every year from the nonstationary model increases over time. Confidence intervals given by dashed lines (5th and 95th). Blue circles indicate the estimated 50-year flood in 1900 and 2018. (g) The return period of the 50-year flood estimated in 1900 (with associated confidence intervals in 1900) is then estimated for every year using the time-varying μ and σ parameters. Blue circles indicate the estimated return period in 1900 (50 years) and 2018.

Finally, there is growing interest in using interpretable machine learning methods for detection of nonstationarities in weather and water extremes. For example, Prophet (Taylor and Letham, 2018) is a decomposable time series forecasting model, similar to Generalized Additive Models (GAMs) (Hastie and Tibshirani, 1987), which is increasingly popular for hydro-climatological time series modelling. For example, Papacharalampous and Tyralis (2020) used the model for forecasting mean annual discharge one year ahead, and Aguilera et al. (2019) used the approach for groundwater-level forecasting. Prophet decomposes the time series into a trend component and a seasonal or periodic component, such as annual or daily cycles. The trend com-

https://doi.org/10.5194/hess-2020-576
Preprint. Discussion started: 14 November 2020

Hydrology and Earth System Sciences
Open Access
Discussions

ponent is a piecewise linear growth model, meaning that for each partition (piece) of the time series (separated by the change points) the model fits a unique trend (varying the trend over time), and the periodic effects are modelled as a Fourier series. This approach allows users to determine the locations in time when there are significant changes.

4.2 Pooled methods for detecting changes in extremes

As noted above, trend-detection of hydroclimatic extremes is problematic when there is uncertainty arising from short record lengths or small samples. Extreme events such as the annual maximum, or 1-in-*n* (50, 100)-year events tend to be highly variable and require lengthy time series to ensure robust detection of significant nonstationarities. In cases where the sample size of observed records is insufficient, alternative methods have been proposed, ranging from "pooled" sampling to scaling approaches.

To address the issue of limited sample size over large spatial scales, one innovative approach is to pool the single largest event over an *n*-year period from multiple independent gauge-based records, effectively substituting space (large spatial sample across many gauges) for time (long temporal sample at individual gauges). In other words, by pooling the data from multiple records or datasets, the data sample is increased for greater statistical robustness. For instance, Berghuijs et al. (2017) assessed changes in 30-year floods across multiple continents by noting the date of occurrence of the single largest daily streamflow at individual gauges, and by evaluating the fraction of catchments experiencing their maximum flood at different points in time. This approach revealed temporal clustering of extreme flood occurrence at regional scales, i.e. flood-rich and flood-poor periods likely associated with hydroclimate variability.

A different and promising pooling method for detection of trends in extremes is the "UNSEEN" approach - UNprecedented Simulated Extremes using ENsembles (e.g. Van den Brink et al., 2005; Thompson et al., 2017). UNSEEN pools members of seasonal predictions or large ensemble climate models (e.g. Deser et al., 2020), using the members as multiple realisations of a plausible 'alternate reality'. In this way, historical sample sizes can be vastly increased to provide greater statistical confidence in extreme estimates. The method was further developed through an "UNSEEN-trends" approach that enables detection of nonstationary extremes (e.g. precipitation) such as 100-year events, from short (e.g. 30-year) climate model records (Kelder et al., 2020). The UNSEEN-trends approach has potential for detection of nonstationarities in a range of climate extremes, but there are some caveats. Reliability of UNSEEN-trends rests on the independence, stability, and fidelity of underlying model members, as well as on the physical plausibility of the hazard-generating mechanisms in the model world.

Most statistical tests used to detect nonstationarity in stream flow suffer from a substantial loss of power when applied to shorter time series (Yue et al., 2002b; Vogel et al., 2013; Prosdocimi et al., 2014, 2019). Pooled frequency analysis improves the estimation of rainfall events associated with long return periods in rainfall intensity-duration-frequency curves at sites where historical rainfall records are short or ungauged by compiling data from many rainfall records in a region (Requena et al., 2019). The areal model applied in Prosdocimi et al. (2019) serves a similar function, pooling regionally similar stream flow gauging stations to enhance shared trend signals. This approach can make clear the presence of trends that might otherwise remain obscured at individual sites due to short record length.

https://doi.org/10.5194/hess-2020-576
Preprint. Discussion started: 14 November 2020

Hydrology and Earth System Sciences
Open Access
Discussions

4.3 Change point analyses for detection of abrupt change

Abrupt changes in hydroclimatic extremes are often a sign that there has been significant human intervention, such as the construction of a dam or diversion, which may significantly affect a catchment's water balance. Change point (also known as step-trend) tests can thus be used either to detect such a shift, or to compare two different periods separated by a long gap (Helsel et al., 2020). Several change-point tests are available. The nonparametric Pettitt test is one of the most widely applied, and allows the user to determine the timing of the change point and its significance. One simulation study assessed the ability of the Pettitt test to detect shifts and found that the test performs best with longer records, when the step-change is located centrally in the distribution, and when the record is less variable (Mallakpour and Villarini, 2016). The performance of the Pettitt test and other similar approaches including pruned exact linear time (PELT) (Killick et al., 2012), binary segmentation (Scott and Knott, 1974), Bayesian analysis (Erdman and Emerson, 2008), wild binary segmentation (WBS) (Fryzlewicz et al., 2014), nonparametric PELT (Haynes et al., 2017), and the Mann-Whitney test (Mann and Whitney, 1947) for the estimation of abrupt changes in the mean, variance, or median of a time series was tested using both simulated and historical data with known change points (Ryberg et al., 2019). Although these methods offer potential benefits (such as the detection of multiple change points), the comparison found that the Pettitt test delivered the best combination of change point detection and minimization of false positive results. The parametric tests (PELT, binary segmentation, and Bayesian analysis) generally performed poorly at detecting known change points in peak stream flow, whilst the non-parametric tests (non-parametric PELT and Mann-Whitney) as well as the parametric WBS resulted in unacceptable false positive rates (Ryberg et al., 2019). The Mood test (Mood, 1954; Ross et al., 2011) for abrupt changes in scale was also evaluated in the same study but located only about 25% of known change points in historical data (with a relatively low false positive rate) and approximately 29% of change points in simulated data (with a relatively high false positive rate).

4.4 Circular statistics and methods for detection of shifts in timing

Circular statistics, a branch that focuses on directions, axes, or rotations, can show changes in the timing and seasonality of hydroclimatic extremes. Circular statistics have been applied to precipitation (e.g. Gu et al., 2017) and flooding (e.g. Villarini, 2016; Hall and Blöschl, 2018). The null hypothesis of circular statistics, when applied to timing, is that data are evenly distributed (uniform), with no tendency to cluster. Three seasonality types can be detected: circular uniform (no preferential direction, indicating that an event has the same probability on any day of the year); reflective symmetric (unimodal); and asymmetric (multimodal, i.e. locations with multiple generating processes, such as snowmelt and mesoscale convective systems in the case of flooding) (Villarini, 2016). Gu et al. (2017) found a significant shift in the seasonality of extreme precipitation using circular statistics, which suggested that pathways of seasonal vapor flux and tropical cyclones were likely driving these changes. A new approach based on seasonality statistics was recently developed to evaluate changes in the dominant flood-generating processes across Europe. Berghuijs et al. (2019b) use the mean date of occurrence of three processes (extreme precipitation, soil moisture excess, and snowmelt) to estimate the relative importance of each process at 3,777 European catchments with at least 20 years of flood peak timing data. They found that the relative importance of mechanisms had not changed significantly

Hydrology and Earth System Sciences
Open Access
Discussions

over 50 years. Similarly, Macdonald et al. (2010) used the mean day of the flood (MDF) to assess changes in flood timing over Wales. They used a directional statistic, computed by counting the number of days from 31st May (the date is chosen to provide a more normal distribution of British floods over the year) until the event, then converting the day of flood occurrence to an angular value (θ; a direction).

Shifts in the timing of extreme weather events can also be assessed by comparing the number of events per season between present and future climate simulations. For example, Whan et al. (2020) show that July–September could experience an increase in the number of landfalling atmospheric rivers over Norway by the end of the century. They find the seasonal differences are reduced in the Far-Future period, with an equal number of AR events in winter and summer, hence suggesting a seasonality shift.

Beyond the detection of statistically significant changes in hydroclimatic extremes (in the form of trends, step-changes, and shifts in timing), the next step is, inevitably, to attribute the most likely cause(s) of the nonstationarity.

5 Drivers of nonstationarity

The terms hydroclimatic 'mechanisms', 'agents', and 'drivers' are sometimes used interchangeably or with different intents. Here, we distinguish between these terms as follows. The expressions hydroclimatic *agents* or *mechanisms* refer to the processes generating hydroclimatic extremes. Depending on the temporal and spatial scale, precipitation-generating agents might include local convection, synoptic weather patterns or cyclones; flood-generating mechanisms might additionally include snowmelt or ice-jam release. In contrast, the expression 'nonstationary *drivers*' refers to longer-term processes which may cause significant shifts in the underlying distributions of hydroclimatic extremes, via climate or land cover change. Here we focus more on the drivers of nonstationarity than the extreme-generating mechanisms and agents. It is important to note that other (non-process related) factors may also generate spurious nonstationarities (see Section 3.1).

The drivers of nonstationarity can be broadly categorised as 'artificial'/human (e.g. land cover change, river regulation) and 'natural' (e.g. large scale climate variability), whist recognising that both are interlinked and affected by feedback processes and confounding variables. These drivers operate over a range of timescales: the annual timescales of land-surface changes (e.g. urbanisation effects on temperature, precipitation, floods and droughts) and seasonal to multidecadal timescales of climate variability (Figure 7). The effect of these nonstationary drivers is complex, and not yet fully understood: the same drivers affect extremes in different ways, depending on the site-specific conditions over different spatial and temporal scales.

5.1 Climate variability and change

Climate change is widely recognised as one of the most important multidecadal to millennial drivers of changes in hydroclimatic extremes. Potential impacts may be expressed primarily through the water cycle, as storminess and rainfall extremes are projected to increase with warming (Hartmann et al., 2013; Kharin et al., 2013; Held and Soden, 2006; Kossin, 2018). However, the association between warming and hydroclimatic extremes is not straightforward: there is not a one-to-one relationship

https://doi.org/10.5194/hess-2020-576
Preprint. Discussion started: 14 November 2020

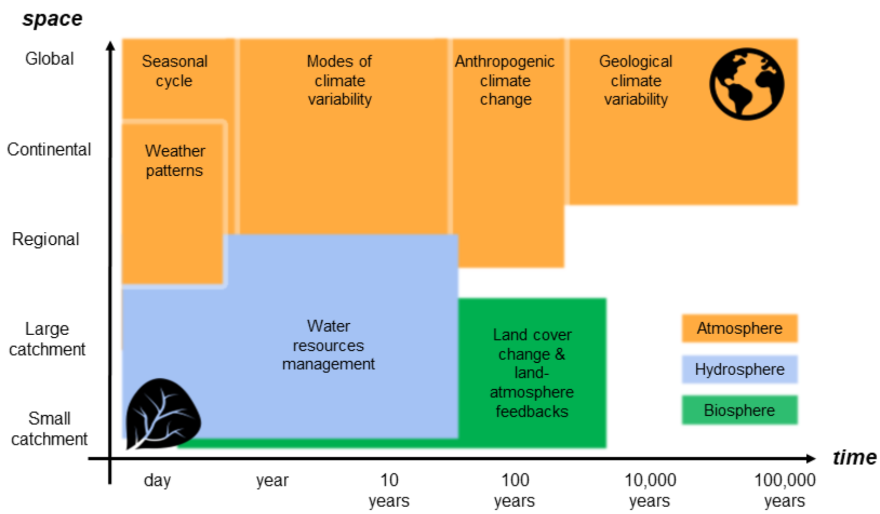

Figure 7. Drivers of nonstationary hydroclimatic extremes, by temporal and spatial scale.

between increases in temperature, precipitation and floods (e.g. Wasko and Nathan, 2019; Wasko et al., 2019), and the effects of climate change operate over different temporal and spatial scales.

To understand how warming may affect intense rainfall and floods, studies have assessed the nature of scaling relationships between extremes. As air temperature warms, the intensity of extreme precipitation is expected to increase due to the enhanced water vapour-holding capacity of warmer air, approximately following the Clausius-Clapeyron relation (7% per degree Celsius) (Trenberth et al., 2003; Ali et al., 2018). Generally, peak rainfall intensities increase with both regional and global temperature. Extreme precipitation is expected to become more intense, and weaker rainfall less intense, in all climatic regions and seasons (Wasko and Sharma, 2015). One key question is whether precipitation intensities are likely to increase proportionally more over shorter timescales. Research suggests that temperature-rainfall scaling increases over hourly timescales (Lenderink and Van Meijgaard, 2008). For example, UK summer precipitation is estimated to increase by 30-40% for short-duration extreme events (Kendon et al., 2014). As temperature increases, we can expect to see longer duration events and higher peak intensities (e.g. exceeding the 99th percentile) (Chan et al., 2016).

What does temperature scaling mean for floods and droughts? There is still little evidence that increases in heavy rainfall events at higher temperatures translate into similar increases in streamflow (Wasko and Sharma, 2017). Increased extreme rainfall does not necessarily lead to increased flooding (Blöschl et al., 2019). There are many other factors that affect flood response besides precipitation intensity, including the duration and extent of precipitation events, antecedent soil moisture conditions, catchment size, vegetation cover, and catchment imperviousness and roughness (Sharma et al., 2018). Increases in precipitation intensity are more likely to cause increases in streamflow in smaller catchments (Wasko and Sharma, 2017; Wasko and Nathan, 2019). Perhaps counter-intuitively, streamflow is likely to decrease in catchments that are experiencing significant reductions in the fraction of precipitation falling as snow (Berghuijs et al., 2014). In high-altitude, snowmelt-dominated regions

https://doi.org/10.5194/hess-2020-576
Preprint. Discussion started: 14 November 2020

such as the Hindu Kush Himalayan "water towers of Asia" (Immerzeel et al., 2020), where a large amount of water is stored as snow or ice, expected climate shifts could accelerate glacier and snowpack melting. This may in turn lead to more frequent glacial lake outburst floods, flash floods and riverine floods, thus posing potential risks to the 240 million people in the region and the 1.9 billion people living downstream (Wester et al., 2019).

Over large spatial scales, climate change also affects the characteristics (magnitude, frequency and timing) of the synoptic-scale phenomena that generate hydroclimatic extremes, such as atmospheric rivers (ARs), tropical cyclones, and atmospheric circulation patterns (e.g. Hirschboeck, 1988; Schlef et al., 2019). For example, ARs play a major role in flood occurrence in many regions of the world (Lavers et al., 2011, 2012; Paltan et al., 2017), and thus changes in their frequency and characteristics could alter the properties of future hydroclimatic extremes. In regions such as Norway where rainfall extremes are primarily driven by ARs (Whan et al., 2020), changes in the phase of precipitation are leading to increasingly rainfall-dominated, rather than snowfall-dominated regimes, which may alter flood characteristics and water resource management. Where otherwise snow would be stored in the catchment, rainfall is likely to contribute more directly to runoff, leading to more severe AR-induced floods, but less severe snowmelt-driven floods later in the season.

The effects of climate variability on hydroclimatic extremes are well recognised. Hazard-rich and hazard-poor periods in the historical record tend to be driven by the spatial and temporal periodicities of multiple, sometimes overlapping, climate modes. For instance, during a positive phase of the North Atlantic Oscillation (NAO), the jet stream tends to shift northwards, generating a greater frequency of heavy rain events over the British and Irish islands (Hannaford and Marsh, 2008). The NAO, together with the Scandinavian pattern (SCA), is also the main driver of European windstorm variability (Walz et al., 2018). El Nino Southern Oscillation (ENSO) phases have different effects on flooding and droughts depending on locality. In Central America, for instance, the west coast experiences increased likelihood of droughts during ENSO, while the east coast has increased flood risk (Cid-Serrano et al., 2015; Enfield and Mayer, 1997; Aguilar et al., 2005). Moreover, El Niño events have been observed to drive episodes of extreme heat in Southeast Asia (Thirumalai et al., 2017), whilst also increasing the lifetime and strength of tropical cyclones in the Western North Pacific (Camargo and Sobel, 2005). Temperatures extremes in Northern Europe are associated with atmospheric blocking situations, both during winter cold spells (Sillmann et al., 2011) and summer heatwaves (Schaller et al., 2018). However, the interlinkages between climate variability and change are still not fully understood for atmospheric blocking and remain challenging to disentangle (e.g. Woollings and Blackburn, 2012; O'Reilly et al., 2019).

5.2 Land cover changes

Land cover modulates the response of local, regional, and remote regions to shifts in climate variability and change. Land is both a source and sink of greenhouse gases: changes in land use such as expansion of areas under agriculture and forestry may alter hydroclimatic extremes through a combination of both biophysical effects (e.g. photosynthesis, respiration, drying, greening) and greenhouse gas feedbacks (see Chapter 2 of IPCC (2019)).

The effect of urbanisation on hydroclimatic extremes is perhaps one of the most well documented land cover changes, but there are still many unknowns. Cities alter the local atmosphere through the urban heat island (UHI) effect which increases the

https://doi.org/10.5194/hess-2020-576
Preprint. Discussion started: 14 November 2020

Hydrology and Earth System Sciences
Open Access
EGU
Discussions

mean annual surface air temperature within cities relative to surrounding rural areas. The magnitude and diurnal amplitude of the UHI effect varies between cities (Ward et al., 2016) lifting night temperatures more than daytime temperatures (Hausfather et al., 2013). For example, in Kuala Lumpur, Malaysia, hourly intensities of extreme rainfall have increased by ~35% in the last 30 years due to the UHI creating an unstable atmosphere (i.e. altering the local atmospheric vertical structure, Li et al., 2019). The effects of these local land surface changes are compounded by the thermodynamic and dynamical effects of climate change. Local changes in extreme temperature are driven over short/medium spatial and temporal scales by changes in the land surface such as the UHI effect and large-scale irrigation, (e.g. Mahmood et al., 2014), but over multi-decadal timescales they are principally caused by large-scale shifts in greenhouse gases modifying the global mean temperature.

Urbanisation effects on flooding and drought are also widely acknowledged but not yet fully understood. Prosdocimi et al. (2015) evaluated the impacts of urban land cover in a paired-catchment study in the UK using both block maxima and POT approaches. They found a significant effect of urbanization on high flows in all seasons, with the strongest effect in summer. A recent analysis of 280 stream gauges in the United States found that annual maximum floods increase by 3.3% on average for every 1% increase in impervious land cover, using panel regression (e.g. Blum et al., 2020). Further, the vertical structure of cities can alter precipitation and flood extremes. When hurricane Harvey hit Houston in August 2017, enhanced rainfall was produced by the storm system's drag induced by increased surface roughness (Zhang et al., 2018). Another recent study examined the effect of urbanization on long term persistence in the stream flow records of 22 catchments in the North-eastern United States using scaling exponents (Jovanovic et al., 2016). They found evidence that streamflow responds more quickly to rainfall in urbanized catchments than in less urbanized counterparts.

Afforestation/deforestation affect both floods and drought, but not always in a single way. For example, afforestation can increase streamflow magnitude by ditching but reduce flood peaks by canopy interception, evapotranspiration and drier antecedent conditions (e.g. Birkinshaw et al., 2014; Soulsby et al., 2017). The effects of changes in vegetated land are believed to be far more pronounced for low flows than high flows (e.g. Birkinshaw et al., 2014; Bathurst et al., 2020; Do et al., 2017; Vicente-Serrano et al., 2019) although there have been very few large-sample studies using observational records. To understand the influence of land cover changes on catchment hydrology, numerical modelling is often used to assess potential impacts and provide evidence for ambitious land cover changes related to policy before implementation. Some use theoretical land cover changes (e.g. Gao et al., 2018; Iacob et al., 2017); others use scenario-based land cover changes (e.g. Harrison et al., 2019). The response of hydroclimatic extremes to land cover change varies over both spatial and temporal scales. At the very fine scale, the hydraulic structure of the soil may be changing due to reorganisation of macro- and micropores associated with land management practices just within a single parcel of land. At continental and multi-decadal scales, changes in the magnitude, intensity and pathway of storms may lead to widespread changes in runoff. These effects may not be distinguishable over short periods in individual catchment data so long-term monitoring is required to understand their impact on catchment hydrology over a range of spatial and temporal scales (Dadson et al., 2017).

Major land cover changes may also have remote teleconnection effects on hydroclimatic extremes. For instance, it has been shown that tropical deforestation alters precipitation patterns not only locally, but also in the mid- and high latitudes. Deforestation of the Amazon and Central Africa may reduce precipitation in the US Midwest; while deforestation of Southeast

https://doi.org/10.5194/hess-2020-576
Preprint. Discussion started: 14 November 2020

Hydrology and Earth System Sciences
Discussions

Asia has been shown to affect China (Avissar and Werth, 2005). In northern India, Saeed et al. (2009) found that irrigation suppresses the development of monsoon-driving land-sea temperature gradients and, thereby, exerts a first-order control on precipitation in central India and the Bay of Bengal.

5.3 Water resources management and geomorphological change

With the growing demand for food and increased productivity following the mechanisation of agriculture, large swathes of land have been subject to arterial and land drainage. Such installations can take various forms. Their impact on hydrological response is poorly understood but they may alter catchment morphology, soil and groundwater hydrological response, depending on the extent of the catchment affected and the extent of works completed on river channels. The nature of hydrological nonstationarities associated with arterial and field drainage has been debated: some argue that changes should not be detectable in discharge as abrupt shifts, while others have identified change points in specific components of the hydrological regime (Harrigan et al., 2014). Numerous studies show that arterial and field drainage can affect flood peaks (increasing magnitude and reducing time to peak, e.g.Wilcock and Wilcock (1995), Bhattarai and O'Connor (2004)). Conversely, understanding of the impacts of drainage on low flows and drought remains poor, as does our understanding of how other components of runoff response are affected due to inadequate monitoring of sub-surface hydrological processes.

Changes in river channel morphology and conveyance capacity are also poorly-understood drivers of changes in flooding. In the Mississippi River for example, construction of wing dikes, navigational structures, and levees have contributed significantly to increases in flood levels (Pinter et al., 2008). Decreases in river conveyance may significantly affect the frequency of overbank flooding and can be identified from trends in flood stages (Pinter et al., 2006) as well as from stream gauge transect data (James, 1991; Smelser and Schmidt, 1998). It is now possible to estimate the relative effect of hydrologic and geomorphic drivers of changing flood risk (Slater et al., 2015) and even to estimate the influence of tropical cyclones on conveyance and flood risk (Li et al., 2020).

Water management can also generate significant nonstationarities. Dam construction increased dramatically over the last century and is likely to continue to continue apace in developing regions. Grill et al. (2015) highlight that on a global basis, 48% of river volume is moderately to severely impacted by either flow regulation, river fragmentation (i.e. diminished connectivity within river systems), or both as a result of dam construction. The impacts on trends in river flow have been debated. Lorenzo-Lacruz et al. (2012) examined trends in Iberian river flows over the period 1945-2005, finding that river regulation by dams was more likely to affect the magnitude of trends rather the direction of change,with important seasonal difference. Examining trends in floods at a global scale, Do et al. (2017) found that the presence of dams did not have a large impact on trend results in all catchments. Rather, catchment size and local context was deemed to be most important in determining response. Abstractions and discharges from watercourses can also drive nonstationarities, with an important task in detection being to ensure naturalised flow regimes are used.

Hydrology and Earth System Sciences
Open Access
Discussions

5.4 Land-atmosphere feedbacks

Nonstationarity can also arise when non-linear dynamical systems exhibit multiple metastable states. Much debate has arisen in the literature surrounding the flawed diagnosis of nonstationarity in systems that have metastable states due to internal feedbacks. The difficulty is particularly acute when records are short and understanding of the system is poor. Examples of metastable states imposed by the climate system include the persistence of periods of high levels in Lake Victoria (Sutcliffe and Parks, 1999), now thought likely to be driven by the Indian Ocean Dipole (Schott et al., 2009). Important examples of land-atmosphere feedbacks internal to hydrological systems include the role of rainfall recycling in Amazon dieback (Wenzel et al., 2014). Feedbacks are also thought to have amplified climate variability in the Sahel and in the 20th century western US dust bowl (Berg et al., 2016). More recently, the role of soil moisture storage in controlling land-surface feedbacks was highlighted through enhancement of heatwaves and associated severe drought conditions in Central Europe (e.g. Seneviratne et al., 2010; Kornhuber et al., 2020). Of key importance is the potential for human water use to affect the water balance in regions where hydrological systems are close to sensitive thresholds (Gleick and Palaniappan, 2010). Taken together, these examples demonstrate that internal feedbacks in hydrological systems can amplify nonstationary drivers of change. Moreover, some of these feedbacks exhibit behaviour which suggests that they introduce tipping points in the Earth system, thus providing not only a response to external changes in the hydrological system, but also that they can be a driver of nonstationarity itself (Lenton et al., 2008). Analyses which address the mechanisms behind such feedbacks will be required in order to account for their effects in the future.

Most of the time there are many drivers affecting hydroclimatic extremes simultaneously, across overlapping multiple temporal and spatial scales (Figure 7). This makes it all the more challenging to discern their individual effects.

5.5 Compound drivers: difficulties understanding nonstationarity in compound risk and consecutive disasters

Interest in compound hydrological extremes is growing rapidly, not least because these events can deliver particularly severe societal impacts (de Ruiter et al., 2019; Raymond et al., 2020a; Zscheischler et al., 2020). Although rarity and complexity pose challenges for their assessment, physical reasoning and empirical methods have highlighted that they can be sensitive to anthropogenic climate warming (AghaKouchak et al., 2020) as well as modes of climate variability (Hillier et al., 2020). It is also possible that novel compound hazards emerge as the climate continues to change (Raymond et al., 2020a; Matthews et al., 2019), causing potentially devastating consequences for communities taken by surprise (Masys et al., 2016). With so much at stake, losses may be too substantial to learn from novel hazards one they have emerged (Masys et al., 2016; Sornette and Ouillon, 2012). Instead, modelling and simulation can be used to explore the so far unseen events. As the name suggests, the climate modelling framework in Section 4.2 may be particularly well suited to this task, but so too are resampling and stochastic weather generation techniques that provide a quantitative underpinning to storyline approaches (Woo, 2019; Shepherd et al., 2018; Matthews et al., 2019).

The role of spatial and temporal correlations in hydrological hazards like drought is important for probabilistic calculations of risk. For example, simultaneous crop failure in Russia, South West Australia and South West China caused a major spike

https://doi.org/10.5194/hess-2020-576
Preprint. Discussion started: 14 November 2020

in grain prices in 2011–12 (e.g. Gaupp et al., 2020). Crop production losses can be underestimated by a factor of three if spatially-correlated risks are ignored, but losses can be reduced by international cooperation to pool risks (e.g. Gaupp et al., 2020). Coincident hydrological extremes at the global scale are driven by climate modes such as ENSO, the Pacific Decadal Oscillation, and the Atlantic Multi-decadal Oscillation (De Luca et al., 2020).

6 Attribution of nonstationarity in hydroclimatic extremes

The term 'attribution' implies understanding the causes or driving processes of change. In the climate sphere, it is often understood in a strict sense as quantifying the influence of humans on the climate or by demonstrating that climate change is consistent with climate model predictions rather than alternative causal explanations. However, beyond climate science, the field of attribution is still relatively new. In hydroclimatology, the concept of attribution is used more broadly to disentangle the wide range of drivers of hydroclimatic extremes, which include climate variability, land cover change, and land-atmosphere feedbacks (section 5). Attribution is typically preceded by exploratory data analysis (EDA) to understand the underlying data (section 6.1). Attribution methods can then broadly be divided into empirical (section 6.2) and simulation-based (section 6.3) approaches. Empirical approaches use statistical techniques to relate the causes to detected changes in the observed record. Simulation-based approaches use model simulations to explain drivers of changes in climatic extremes. In both cases, attribution requires developing multiple working hypotheses based on a process-based understanding of potential drivers of change (Harrigan et al., 2014). There is considerable overlap between the methods employed for attribution (in the past), prediction (over short future time scales) and projection (over multi-decadal timescales).

6.1 Exploratory data analysis (EDA)

Exploratory analyses play an essential role in understanding historical changes. Often, such analyses involve visual exploration of the time series and of the statistical test results for many gauges to identify common patterns or outliers. At the most basic level, temporal trend lines (e.g. Figure 6f) from multiple stations can be plotted on the same graph to identify any record that diverge from the coherence of the pooled analysis. Similarly, trend persistence plots may be created for multiple stations to identify suspect data where nonstationarities may be present. For example, Noone et al. (2016) computed the MK Z-score over different time periods, starting with the full record, then shortening the start year one year at time to a minimum length required for statistical robustness (such as 30 years). Plotting persistence lines for multiple stations allowed them to flag stations with possibly spurious nonstationarities. Wavelet power spectra can also be plotted for multiple stations to assess whether a common periodicity could be detected across multiple sites. For instance, Rust et al. (2019) used continuous wavelet transforms to identify multi-annual periodicities of the North Atlantic Oscillation in UK groundwater records.

6.2 Empirical attribution approaches

Statistical attribution of hydroclimatic nonstationarity is typically regression-based. Statistical methods require less computational power than physical model-based approaches, but cannot account for complex feedbacks between processes. Addition-

https://doi.org/10.5194/hess-2020-576
Preprint. Discussion started: 14 November 2020

Hydrology and Earth System Sciences
Discussions

ally, statistical approaches are unable to distinguish between correlation and causation, even when a predictor is physically plausible. The first step in attributing the drivers of a nonstationary process is to detect the presence of nonstationarity in the time series. This can be done by applying a trend plus significance test (section 4.1); or by fitting a distributional regression model with constant parameters (stationary case), then again with time-varying parameters (see Figure 6a)), and evaluating the superior description of the data. If the stationary model fit is better (e.g. lower AIC or SBC), then it is not recommended to proceed further. If the time-varying model is better, and if there are physically-plausible reasons to suspect nonstationarity, then attribution may proceed. However, it is important to note that small differences in AIC and SBC are not always meaningful (Burnham and Anderson, 2004), and visual assessment is generally recommended. If a time series is deemed nonstationary, regression can be used to determine potential drivers of change by introducing additional predictors that are representative of the drivers (e.g. changing land cover, climate, or reservoir indices). Regression model coefficients can be used to quantify the effect of covariates (e.g. Prosdocimi et al., 2015). 'Soft' statistical attribution approaches involve fitting a range of models with different sets of covariates in a given location, and selecting the best-fitting model (i.e. most important covariates). For example, Slater and Villarini (2017b) fit a range of models at 290 stream gauges using plausible combinations of five predictors – precipitation, antecedent wetness, temperature, agriculture, and population density – to identify the dominant driving processes in different seasons.

Panel regression techniques are increasingly popular because they can be used to leverage temporal and spatial variation to isolate a causal effect, separate from other drivers of change (Blum et al., 2020). These methods pool both dynamic (e.g. land cover) and static (e.g. physical catchment characteristics) data across time and space. They are able to identify generalized relationships between drivers and hydrologic response (i.e. reliable model coefficients), and are particularly powerful in regional and large sample analyses (Bassiouni et al., 2016; Steinschneider et al., 2013). Examples of recent applications of panel regression include estimating low-flow response to rainfall in ungauged basins (Bassiouni et al., 2016), estimating the effects of urbanization on annual runoff coefficients (Steinschneider et al., 2013) and peak flows in the United States (Blum et al., 2020) and in Belgium (De Niel and Willems, 2019). Others have applied panel regression to analyse the effects of forest cover change and other socioeconomic factors on the frequency of large floods in developing regions (Ferreira and Ghimire, 2012) or to examine the effects of deforestation and agricultural development on streamflow in Brazil (Levy et al., 2018).

6.3 Simulation-based attribution of nonstationary extremes

Unlike statistical methods, where models are "fit" to observations, climate model experiments are physically-based: they solve series of mathematical equations representing various Earth-system processes. The strength of these methods relies on their use of physics to accurately describe real-world feedbacks and recognise important physical limits, for example, the cap on extreme humid heat in the tropics linked to deep convection (Sherwood and Huber, 2010; Zhang and Fueglistaler, 2020). Here we provide a brief overview but we point the reader to Hegerl et al. (2010) and Stocker et al. (2013) for details.

In attribution studies, climate models can be run under different scenarios to identify the role of anthropogenic forcing in contributing to the likelihood of an observed extreme. For example, comparing the probability of an extreme event simulated by a model for the current climate (p1) with its probability in a modelled pre-industrial scenario (p0), indicates how much

https://doi.org/10.5194/hess-2020-576
Preprint. Discussion started: 14 November 2020

Hydrology and Earth System Sciences
Discussions

"more likely" (p1/p0) the event has become relative to a "counter-factual" world in which the pre-industrial climate persisted (see Figure 8a-b). The probability ratio framework is also used to identify the "fraction of attributable risk" (1-p0/p1) – the portion of the event's probability that can be assigned to anthropogenic modification of the climate (Fischer and Knutti, 2015). Changes in event magnitude can also be examined using the same framework. Wang et al. (2018) applied such methods to the exceptional rainfall caused by Hurricane Harvey in 2017. The analysis indicated that, in a counter-factual environment of no trends in SSTs or tropospheric variables, the storm would have delivered almost 30% less precipitation. Such methods, which blend climate change scenarios (in this case, no trends since 1980) with numerical weather prediction models can also be used to assess in detail how the characteristics of observed extremes may differ in counter-factual worlds (hotter or colder than present). Lackmann (2015) demonstrated this for Hurricane Sandy, noting that climate warming since pre-industrial levels caused a small northward shift and intensification of the storm's track; further intensification and northward displacement could be expected in a counterfactual climate even warmer than observed in 2012. Statistical comparisons, such as the Anderson-Darling (Figure 8c) or Kolmogorov-Smirnov tests, can formally confirm any significant differences between historical observations and future simulated extremes.

Event attribution has emerged as a major field of research in the last decade, aiming to assess whether specific extreme events can be ascribed to human-induced climate change (Stott et al., 2016). Climate models can be used to estimate the influence of climate warming on the likelihood or intensity of individual events. For example, Diffenbaugh et al. (2017) compare four different attribution approaches used to evaluate the influence of climate warming on the hottest monthly and daily events at the global scale. The first two methods estimate the contribution of the observed historical trend to the magnitude of the event (Figure 8d), or to the probability of the event magnitude. The second two methods estimate the probability of the observed trend in the historical climate forcing (Figure 8e); or the probability of the event magnitude in the historical and pre-industrial climate forcing. For a review of the statistical methods employed for event attribution in climate science, see Naveau et al. (2020).

6.4 Issues with attribution: complexity, confounding variables, and undetected drivers

Attribution encounters different challenges depending on the variables being considered. In the case of hydrological extremes such as floods or droughts, there is substantial complexity in disentangling multiple drivers, which may have a confounding influence. One of the principal obstacles in hydrology is the lack of data on the drivers of nonstationarity at appropriate temporal and spatial scales. Important drivers of hydrological change (such as arterial drainage, land use changes, or adjustments in the conveyance capacity of river channels) may be overlooked if the initial attribution framework is too narrow – i.e. does not consider the possibility of multiple plausible drivers. Hydrologists seeking to quantify the influence of land cover signals are faced with a conundrum: how to disentangle multiple driving factors in big datasets that have not been assembled for such a purpose. There are many national hydrological reference networks (also known as benchmark networks) that are minimally affected by anthropogenic influences and designed to enable detection of global change signals (e.g. Whitfield et al., 2012; Harrigan et al., 2018). However, these data are not well-suited to detect drivers such as land cover changes: there are currently no benchmark networks for detection and attribution of the effects of land cover changes on streamflow at any scale. Hy-

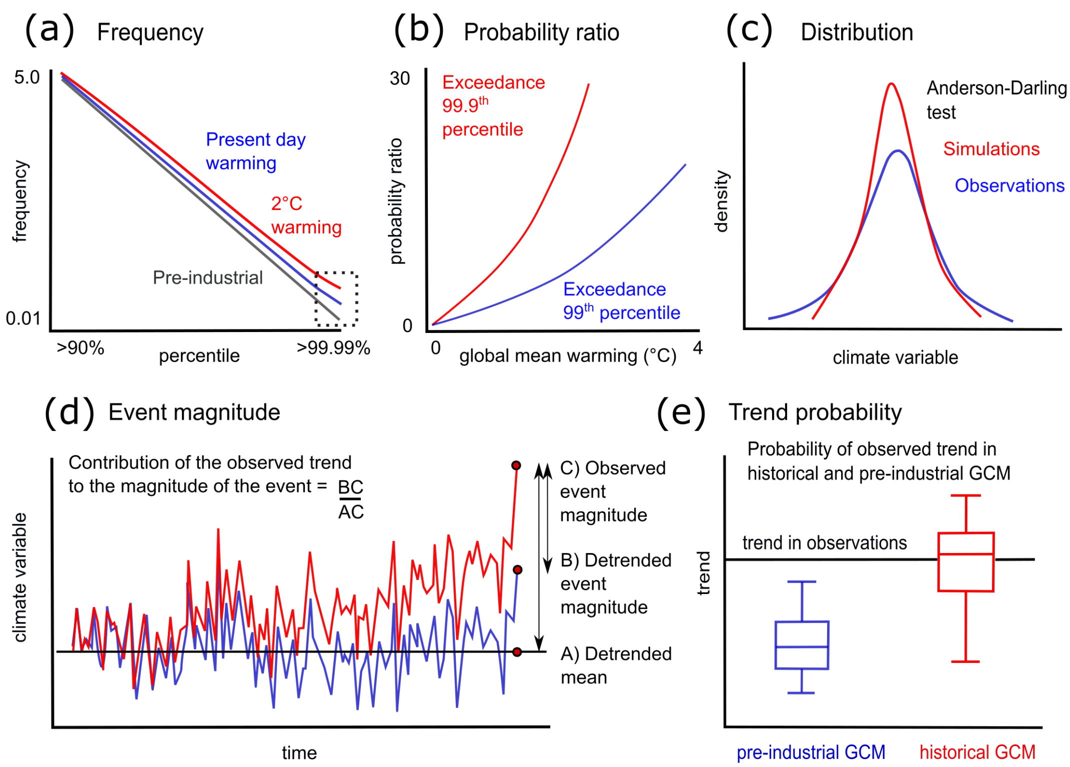

Figure 8. Attribution techniques for exploring nonstationarity with climate model simulations and observations. (a) Frequency analyses: example of daily precipitation simulations for pre-industrial conditions, present day and future warming scenarios. (b) Probability ratio: example of exceeding $99^{th}/99.9^{th}$ percentile of pre-industrial temperature at a given warming level relative to pre-industrial conditions, averaged across land. (c) Comparing distributions: here, using the Anderson-Darling test to evaluate differences between observed and simulated interannual variability. (d) Event magnitude: estimating the contribution of an observed trend to an event magnitude. (e) Trend probability: estimating the probability of an observed trend in a historical climate model simulation. Panels a-b are simplifications of the general patterns shown in Fischer and Knutti (2015); panels c-e adapted from Diffenbaugh et al. (2017).

drologists seeking to attribute such effects may find spurious relationships, or that the data are affected by other confounding variables such as upstream dams or abstraction. Because of the multiple confounding variables that affect streamflow, a broader attribution framework, with multiple working hypotheses, is needed.

Complexity is reduced for attribution of extreme high temperatures because they scale strongly with global mean warming (Buzan and Huber, 2020), particularly over land (Matthews et al., 2017). However, some confounding influences remain even for heatwaves because of their sensitivity to land use change. For example, any modification leading to surface drying (e.g. cessation of irrigation) may partition more radiant energy into the sensible heat (at the expense of latent heat), thereby amplifying high temperature extremes (Miralles et al., 2014; Peterson et al., 2011). This principle of trading off between sensible and latent heat is often exploited in reverse by urban designers trying to engineer cool cities (Coccolo et al., 2018), but it may

https://doi.org/10.5194/hess-2020-576
Preprint. Discussion started: 14 November 2020

Hydrology and Earth System Sciences
Discussions

confound local-scale attribution studies focusing on temperature extremes. The total sensible and latent heat content of the air is, therefore, recommended for studies of extreme heat events (Pielke Sr et al., 2004; Matthews, 2020), not least because it is more tightly coupled to physiological heat stress (Matthews, 2018).

The challenge of nonstationarity does not end with attribution because it is only after attribution that we can consider the management implications then approaches to prediction and projection.

7 Management of future nonstationary extremes

Estimation of future nonstationary extremes is sensitive to methodological choices such as the influence of different bias correction methods, and philosophical questions about whether we can (or should) predict nonstationarity into the future given inherent uncertainties. Here we describe methods for factoring nonstationarity into design estimation and present the current difficulties and emerging methods in predicting and projecting hydroclimatic extremes under a nonstationary future.

7.1 Adjusting for nonstationarity in engineering design

It is generally accepted that assuming stationarity in the hydrologic variables used for long-lived engineering designs is no longer tenable (Milly et al., 2008). However, there are differences in opinion about how environmental change information can be used in detailed engineering designs. For instance, a decade ago, some claimed that climate models were not yet 'ready for prime time' (Kundzewicz and Stakhiv, 2010). The reasoning was that outputs from climate models and downscaling procedures were too coarse and uncertain for use in site-specific adaptations to infrastructure. However, this was premised on a 'predict-then-act' approach to nonstationarity, where the goal was to characterize, or even constrain, major sources of uncertainty for decision-makers (Clark et al., 2016).

An alternative view is that nonstationary hydroclimatic information can be applied in smarter ways by stress-testing the performance of a design or adaptation decision via risk and reliability metrics (e.g. Brown and Wilby, 2012). Such frameworks promote early engagement with the decision-maker in the design process, to identify key system vulnerabilities, performance criteria and trade-offs between management goals (Poff et al., 2016). Furthermore, so-called 'scenario-neutral' methods can establish when and where national safety margins for nonstationary hydroclimatic conditions might be inadequate (Prudhomme et al., 2010; Broderick et al., 2019). When combined with storylines that describe more elaborate scenarios of change – such as forest die-back and dust on snowpack under hotter-drier conditions shifting the volume and timing of runoff – more holistic capture of nonstationarity in adaptation planning is then feasible (e.g. Yates et al., 2015).

A few agencies have issued specific guidance for incorporating climate change into detailed engineering designs (e.g. Asian Development Bank, 2020; Environment Agency, 2016; International Hydropower Association and others, 2019; United States Army Corps of Engineers, 2019). Favoured techniques include: setting lower bound change estimates by extrapolating historic rates of change (such as for sea level rise); using changes in primary design variables (such as annual daily maximum precipitation amount) from regional climate projections as an adjustment factor (or climate change allowance) to uplift baseline series; or rescaling the parameters of extreme value distributions to reflect multi-decadal climate variability. For a review of these

Hydrology and Earth System Sciences
Open Access
Discussions

procedures and other robust decision-making approaches in the water sector see Wilby and Murphy (2019). Approaches for incorporating uncertainty into flood design include: climate factors (i.e. adjusting peak flow estimates); the 'prudent approach' (i.e. selecting a larger return period based on the precautionary principle); and robustness-based decision methods (i.e. which may be suited to a large range of plausible futures) (François et al., 2019). Other sectors are recognizing the need to adopt new standards as well, for example, to improve the thermal performance of buildings (e.g. Lomas and Giridharan, 2012).

7.2 Model-based prediction and projection

In a world of nonstationary hydroclimatic extremes, society must adapt and become increasingly resilient to their negative impacts. A critical tool in the Disaster Risk Reduction (DRR) effort is operational early-warning systems. These are used over short timescales (sub-daily to several months forecast horizons) and are based on numerical weather prediction (NWP) or seasonal prediction models. Seasonal predictions start from an observed state of the climate system and provide a picture of the probability that the coming season will be wetter, drier, warmer or colder than usual. A comprehensive overview is provided by ECMWF (2020b). Seasonal predictions are used for many extremes (e.g. Vitart and Robertson, 2018), including precipitation (e.g. Shukla et al., 2019), temperature (e.g. Gubler et al., 2020), floods (e.g. Ren et al., 2019; Slater and Villarini, 2018), droughts (e.g. Hao et al., 2018; Yuan et al., 2017) and wind/storms (e.g. Befort et al., 2019; Torralba et al., 2017b). For hydrological extremes, several early warning systems are now operational at global and continental scales (Emerton et al., 2016) (e.g. EFAS, GloFAS, E-HYPE, WW-HYPE, GLOFFIS) and are being used for decision-making during major events. For example, GloFAS was used to provide operational flood bulletins to civil protection and humanitarian agencies in Mozambique, Zimbabwe and Malawi in the wake of Tropical Cyclones Idai and Kenneth in 2019 (Emerton et al., 2020). For a comprehensive review of hydrometeorological ensemble forecasting, we recommend Duan et al. (2019).

Over years to decades ahead, climate model projections (such as those of the Climate Model Intercomparison Project, CMIP) are used to understand how hydroclimatic extremes may develop according to different scenarios. They are typically initialized once with perturbations to the initial conditions and with stochastic physics. Climate model projections are less constrained to observed conditions than seasonal prediction systems that are for example initialized every month. They can be downscaled and bias-corrected, but they cannot be evaluated against observations in the same manner as seasonal and annual climate model predictions. Similarly, projections are employed to better assess the range of future extremes of precipitation and temperature (e.g. Li et al., 2016; Bao et al., 2017; Niu et al., 2018; Wu et al., 2020b), floods (e.g. Villarini and Zhang, 2020; Shkolnik et al., 2018), droughts (e.g. Yuan et al., 2019; Wu et al., 2020a) and wind/storms (e.g. Krishnan and Bhaskaran, 2020; Emanuel and Center, 2020). An overview of global climate projections (CMIP) and regional climate projections (CORDEX) is provided by ECMWF (2020a).

Despite the advantages of climate models for understanding nonstationary extremes, they also have several non-trivial limitations. Their immense computational demand means that ensemble experiments may take years to complete. Their skill depends on the ability to represent climate dynamics accurately, and so most climate models have notable biases when projecting nonstationarities decades ahead. For instance, several CMIP5 GCMs with 6-hourly fields exhibit 'implausible' projections (McSweeney et al., 2015). Care should be taken when using climate projections, as their biases (and changes over time of the biases) may considerably distort the multi-model means of water cycle components (Liepert and Lo, 2013). Another issue with using climate models to understand nonstationarity is their tendency to "drift" over time, which means they develop progressive changes beyond natural variability. This behaviour strongly depends on both the model and climate variable being analysed. For example, precipitation drift contributes to less than 10% of the historical trend in most of the CMIP5 models, whereas drift in steric sea level might contribute up to 30-60% of the historical trend (Gupta et al., 2013).

Ensemble modelling approaches enhance confidence in our understanding of nonstationarities by better quantifying the uncertainty arising from models versus the uncertainty arising from internal variability. Sample size is increased for individual climate models by running many model members (e.g. 51 members in ECMWF's SEAS5; Johnson et al., 2019); or via multi-model ensembles (such as the North-American Multimodel Ensemble (NMME, Kirtman et al., 2014) or the China Multi-Model Ensemble Prediction System (CMME, Ren et al., 2019)); or from single model initial-condition large ensembles (SMILEs), which quantify the effect of internal variability (Deser et al., 2020). For instance, Maher et al. (2020) showed that surface temperature trends in near-term (15 to 30 year) projections are dominated by internal variability, using six SMILEs and the CMIP5 output. Post-processing methods such as the Bayesian Ensemble Uncertainty Processor (BEUP) have been developed to quantify the predictive uncertainty associated with both ensemble climate predictions and hydrologic data (Reggiani et al., 2009). In BEUP approaches, ensemble weather predictions are used as inputs to hydrologic models and the resulting ensemble streamflow predictions are then post-processed by the BEUP, delivering a full probability distribution of the expected flow. The climate uncertainty is represented by ensembles of prediction and the aggregated hydrologic uncertainty is quantified and then added after post-processing of BEUP (Han and Coulibaly, 2019).

7.3 Hybrid statistical-dynamical prediction and projection

Hybrid statistical-dynamical approaches offer several advantages over purely statistical or purely dynamical methods for predicting and projecting nonstationary extremes. A hybrid model uses climate predictions that are sometimes bias-corrected or merged using techniques such as regression or Bayesian model averaging. Hybrid approaches are particularly useful for hydroclimatic extremes where dynamical models have limited skill (e.g. rainfall predictions). They have been applied to a wide range of weather and water extremes including precipitation (Wang et al., 2012; Fernando et al., 2019), temperature (Strazzo et al., 2019), hurricanes (Kim and Webster, 2010; Vecchi et al., 2011), floods (Slater and Villarini, 2018) and droughts (see Hao et al. (2018) for a review). For example, Kim and Webster (2010) developed extended-range seasonal hurricane forecasts using North Atlantic SST anomalies and vertical wind shear hindcasts from the ECMWF Seasonal Forecasting System 3. Lee et al. (2020) used projected surface air temperature (SAT), specific humidity and surface pressure from two GCMs, four RCMs, and two Representative Concentration Pathway (RCP) climate change scenarios as covariates to compute dew-point temperature (DPT). They then employed the SAT and DPT to project nonstationary precipitation peaks over threshold (POT) using a Generalized Pareto distribution. This indirect approach produced more consistent projections of precipitation across a range of climate models than the raw (direct) precipitation projections. Finally, Slater and Villarini (2018) used forecasts of seasonal precipitation and temperature from the NMME alongside antecedent climate conditions, agriculture, and population density, to generate enhanced predictions of seasonal flooding. In this approach, the regression model relating extremes (e.g.

https://doi.org/10.5194/hess-2020-576
Preprint. Discussion started: 14 November 2020

Hydrology and Earth System Sciences Discussions

flood indices) to covariates (e.g. precipitation) generates future predictions/projections of extremes from outlooks of seasonal forecasting models or climate models.

There is considerable scope for developing enhanced statistical-dynamical prediction using sophisticated postprocessing techniques such as ensemble model output statistics (EMOS). For instance, Bayesian joint probability (BJP) modeling, an EMOS-type method, employs a joint probability distribution to characterize the relationship between the raw GCM ensemble mean and observations. BJP has been found to be superior to traditional quantile mapping approaches when dealing with bias, reliability and coherence (Zhao et al., 2017).

7.4 Validity of models for evaluation of nonstationarity

Process-based models are often used in climate science and hydrological impact studies precisely to cope with the uncertainty and nonlinear feedbacks that arise because of nonstationarity but are not always able to solve those problems. Incorporating nonstationarity in the calibration of model parameters and physics remains an issue. Models are used in situations where nonstationarity is present, sometimes with the express purpose of understanding the nonstationary process (e.g. Crooks and Kay, 2015). Models that assume stationarity or are heavily reliant on calibration (with fixed parameters) are less credible at representing nonstationarity. However, this distinction between statistical models (that assume data on past system behaviour offers insight into future behaviour) and physically-based models (which incorporate fundamental physical, biological, and chemical principles) is unhelpful. In practice, there are many physical principles (e.g. conservation of mass, energy, momentum) encoded in or relied upon by statistical models, whereas there are many data-specific parameters hard wired into physical models (Hrachowitz and Clark, 2017). As always, the key is to understand all of the steps at which assumptions of nonstationarity are made because these determine model realism when applied under nonstationary conditions (Kirchner, 2009).

More generally, models used for attribution and prediction suffer from incomplete representation of all sources of nonstationarity. For instance, hydrodynamic models used for inundation modelling may be inaccurate for projecting flood extents because they consider the land surface or river channel properties as fixed, by assuming bankfull flooding occurs once every two years (employing a two-year flood recurrence interval) (e.g. Kettner et al., 2018). One promising approach is to begin to tackle nonstationarity within an Earth System Modelling perspective that includes interactions between atmosphere, oceans, land, biosphere, and human activities. Such an interdisciplinary approach is expected to accelerate scientific progress (Harrigan et al., 2020). However, there are still obstacles in the way of a more holistic framework. For instance, grid resolution will inevitably determine which features may be included (e.g. river catchments) or excluded (e.g. field drains).

8 Conclusions and recommendations

The presence of nonstationarity in weather and water extremes has long been controversial, due to lack of clarity about the timescales required for identifying nonstationarity; spurious identification of nonstationarities arising from data inhomogeneity; poor or incomplete information about the drivers of change; and the mistaking of short-term excursions for long-term change. Caution is undoubtedly necessary when attempting to detect and project nonstationarity (Serinaldi et al., 2018). However,

https://doi.org/10.5194/hess-2020-576
Preprint. Discussion started: 14 November 2020

Hydrology and Earth System Sciences
Discussions

recognising the presence of an abrupt temporal departure from normality in the form of a trend, step-change, shift in variance, or more complex associations between variables may have profound utility for the appropriate management of hydroclimatic hazards and associated resources. We argue thus in favour of a "functional" nonstationarity (Villarini et al., 2018), defined as shifts in the probability distribution of a given dataset rather than an entire population. This approach recognises the value in using the concept of nonstationarity as a practical tool for diagnosis, prediction and management of changing extremes. Hence, it is important to acknowledge the reasons for employing nonstationary methods. Inaccurate detection of a trend or step-change (Type I errors) may be just as hazardous as not detecting a significant shift where it does in fact exist (Type II error). Such type II errors may lead to societal under-preparedness and infrastructure maladaptation based on a flawed assumption of stationarity (e.g. Vogel et al., 2013).

This review outlines a 'toolkit' to guide investigations into the detection, attribution, and management of nonstationary extremes (Figure 2). Recognising nonstationary symptoms (section 2) is essential for developing risk management strategies under global change. Such shifts mean society must adapt both over the short term through early warning systems, and over the long term through nature-based and engineered solutions to risk management. However, methods used for hazard risk management should only factor nonstationarity into their design after thorough exploratory data analysis (section 3), due consideration of the uncertainties inherent in nonstationary techniques (Serinaldi and Kilsby, 2015), and evaluation of the transferability and robustness of model parameters from the calibration period to a nonstationary future (Broderick et al., 2016; Fowler et al., 2016, 2020).

When faced with incomplete knowledge of hazard-generating processes, 'multiple working hypotheses' frameworks must be developed to ensure robust detection of drivers (Chamberlin, 1890; Clark et al., 2011; Harrigan et al., 2014) (sections 4-5). Rather than seeking a single solution or 'cure' for extremes, the precautionary principle requires a more inclusive investigation of possible underlying causes. Deliberately perturbing one component of a physically-based model remains currently the only approach to control drivers with confidence, and search a response in system behaviour.

Attribution remains an ongoing challenge (section 6) as the shortness of observational records frequently precludes robust detection of nonstationarities. Pooled approaches – both temporal (e.g. Thompson et al., 2017; Kelder et al., 2020; Massey et al., 2015) and spatial (e.g. Prosdocimi et al., 2019; Blum et al., 2020) – hold considerable promise for enhanced detection and attribution of the different types of nonstationarity and their driving processes, by increasing the sample size required to detect significant change. At individual sites, ensemble approaches will be indispensable to distinguish between periodicities and temporal trends, as well as the inter-dependencies of the driving variables. Over regional scales, large-sample approaches are also one of the solutions to understanding the more general properties of both environmental hazards and their drivers (e.g. Addor et al., 2020).

A final step is in the management of future extremes (section 7). Beyond consideration of single extremes, there are fundamental issues with the management, prediction, and projection of compound extremes. Just as the lack of observations is a key challenge for individual extremes, the difficulty is even greater when it comes to understanding the intersection, sequencing and joint distribution of very rare events, which by definition are limited in the observed record. Pooled or ensemble approaches are also likely to be one of the key ways forward to tackle some of the major outstanding research directions –

https://doi.org/10.5194/hess-2020-576
Preprint. Discussion started: 14 November 2020

Hydrology and Earth System Sciences
Open Access
Discussions

such as future changes; effects of land-atmosphere interactions and feedbacks; their impacts; land cover effects; and cascading events (AghaKouchak et al., 2020). However, as climate models become more complex, there is no guarantee that nonstationary signals may emerge more clearly than before. Dynamical changes in atmospheric circulation are highly uncertain, and postprocessing of large multi-model ensembles is required to help extract a signal from the noise (e.g. Smith et al., 2020). The 'promise' of reduced uncertainty with better science is also not guaranteed: statistical postprocessing and the most advanced techniques may still be inconsequential if a process is poorly captured within the modelled world.

Lastly, alongside sophisticated detection and attribution techniques, there remains a need for practical tools for managing future change. Methods range from simple look-up tables, factors, and storylines, through to user-driven design of environmental decision support systems (EDSS; e.g. McIntosh et al., 2011; Zulkafli et al., 2017). Alongside the more quantitative methods, storylines may allow us to imagine futures with multiple nonstationary drivers, and to frame narratives in terms that are meaningful for decision making (Shepherd et al., 2018). Developing intelligible ways to translate the science, including "low regret" options for adaptation, remains a key area for research.

Competing interests. No conflict of interest

Acknowledgements. This review stems from a virtual hydroclimatology workshop held between the universities of Oxford, Loughborough, Maynooth and the ECMWF during the Covid summer of 2020, and seeks to provide a general introduction to the concept, methods, and debates associated with evaluating nonstationarity in weather and water extremes. We welcome all formal or informal comments and suggestions for improvement from the community.

https://doi.org/10.5194/hess-2020-576
Preprint. Discussion started: 14 November 2020

Hydrology and Earth System Sciences
Discussions

References

Addor, N., Do, H. X., Alvarez-Garreton, C., Coxon, G., Fowler, K., and Mendoza, P. A.: Large-sample hydrology: recent progress, guidelines for new datasets and grand challenges, Hydrological Sciences Journal, 65, 712–725, 2020.

AghaKouchak, A., Chiang, F., Huning, L. S., Love, C. A., Mallakpour, I., Mazdiyasni, O., Moftakhari, H., Papalexiou, S. M., Ragno, E., and Sadegh, M.: Climate Extremes and Compound Hazards in a Warming World, Annual Review of Earth and Planetary Sciences, 48, 519–548, https://doi.org/10.1146/annurev-earth-071719-055228, 2020.

Aguilar, E., Auer, I., Brunet, M., Peterson, T. C., and Wieringa, J.: Guidance on metadata and homogenization, Wmo Td, 1186, 1–53, 2003.

Aguilar, E., Peterson, T., Obando, P. R., Frutos, R., Retana, J. A., Solera, M., Soley, J., García, I. G., Araujo, R., Santos, A. R., et al.: Changes in precipitation and temperature extremes in Central America and northern South America, 1961–2003, Journal of Geophysical Research: Atmospheres, 110, 2005.

Aguilera, H., Guardiola-Albert, C., Naranjo-Fernández, N., and Kohfahl, C.: Towards flexible groundwater-level prediction for adaptive water management: using Facebook's Prophet forecasting approach, Hydrological Sciences Journal, 64, 1504–1518, https://doi.org/10.1080/02626667.2019.1651933, 2019.

Ali, H., Fowler, H. J., and Mishra, V.: Global observational evidence of strong linkage between dew point temperature and precipitation extremes, Geophysical Research Letters, 45, 12–320, 2018.

Anagnostopoulou, C. and Tolika, K.: Extreme precipitation in Europe: statistical threshold selection based on climatological criteria, Theoretical and Applied Climatology, 107, 479–489, 2012.

Andreadis, K. M. and Lettenmaier, D. P.: Trends in 20th century drought over the continental United States, Geophysical Research Letters, 33, 2006.

Archfield, S. A., Hirsch, R. M., Viglione, A., and Blöschl, G.: Fragmented patterns of flood change across the United States, Geophysical research letters, 43, 10–232, 2016.

Asian Development Bank: Climate change adjustments for detailed engineering design of roads: Experience from Viet Nam. Knowledge Product, http://dx.doi.org/10.22617/TIM200148-2, 2020.

Avissar, R. and Werth, D.: Global hydroclimatological teleconnections resulting from tropical deforestation, Journal of Hydrometeorology, 6, 134–145, 2005.

Baines, P. G. and Folland, C. K.: Evidence for a rapid global climate shift across the late 1960s, Journal of Climate, 20, 2721–2744, 2007.

Bao, J., Sherwood, S. C., Alexander, L. V., and Evans, J. P.: Future increases in extreme precipitation exceed observed scaling rates, Nature Climate Change, 7, 128–132, 2017.

Barbosa, S., Scotto, M., Alonso, A., and Lasaponara, R.: Summarising changes in air temperature over Central Europe by quantile regression and clustering., Natural Hazards & Earth System Sciences, 11, 2011.

Barkhordarian, A., von Storch, H., and Bhend, J.: The expectation of future precipitation change over the Mediterranean region is different from what we observe, Climate Dynamics, 40, 225–244, 2013.

Bassiouni, M., Vogel, R. M., and Archfield, S. A.: Panel regressions to estimate low-flow response to rainfall variability in ungaged basins, Water Resources Research, 52, 9470–9494, 2016.

Bathurst, J. C., Fahey, B., Iroumé, A., and Jones, J.: Forests and floods: using field evidence to reconcile analysis methods, Hydrological Processes, 2020.

Bazrafshan, J. and Hejabi, S.: A non-stationary reconnaissance drought index (NRDI) for drought monitoring in a changing climate, Water Resources Management, 32, 2611–2624, 2018.

Befort, D. J., Wild, S., Kruschke, T., Ulbrich, U., and Leckebusch, G. C.: Different long-term trends of extra-tropical cyclones and windstorms in ERA-20C and NOAA-20CR reanalyses, Atmospheric Science Letters, 17, 586–595, https://doi.org/10.1002/asl.694, 2016.

Befort, D. J., Wild, S., Knight, J. R., Lockwood, J. F., Thornton, H. E., Hermanson, L., Bett, P. E., Weisheimer, A., and Leckebusch, G. C.: Seasonal forecast skill for extratropical cyclones and windstorms, Quarterly Journal of the Royal Meteorological Society, 145, 92–104, 2019.

Bell, V., Davies, H., Kay, A., Marsh, T., Brookshaw, A., and Jenkins, A.: Developing a large-scale water-balance approach to seasonal forecasting: application to the 2012 drought in Britain, Hydrological Processes, 27, 3003–3012, 2013.

Berg, A., Findell, K., Lintner, B., Giannini, A., Seneviratne, S. I., Van Den Hurk, B., Lorenz, R., Pitman, A., Hagemann, S., Meier, A., et al.: Land–atmosphere feedbacks amplify aridity increase over land under global warming, Nature Climate Change, 6, 869–874, 2016.

Berghuijs, W., Woods, R., and Hrachowitz, M.: A precipitation shift from snow towards rain leads to a decrease in streamflow, Nature Climate Change, 4, 583–586, 2014.

Berghuijs, W. R., Aalbers, E. E., Larsen, J. R., Trancoso, R., and Woods, R. A.: Recent changes in extreme floods across multiple continents, Environmental Research Letters, 12, 114 035, 2017.

Berghuijs, W. R., Allen, S. T., Harrigan, S., and Kirchner, J. W.: Growing spatial scales of synchronous river flooding in Europe, Geophysical Research Letters, 46, 1423–1428, 2019a.

Berghuijs, W. R., Harrigan, S., Molnar, P., Slater, L. J., and Kirchner, J. W.: The relative importance of different flood-generating mechanisms across Europe, Water Resources Research, 55, 4582–4593, 2019b.

Bhattarai, K. and O'Connor, K.: The effects over time of an arterial drainage scheme on the rainfall-runoff transformation in the Brosna catchment, Physics and Chemistry of the Earth, Parts A/B/C, 29, 787–794, 2004.

Birkinshaw, S. J., Bathurst, J. C., and Robinson, M.: 45 years of non-stationary hydrology over a forest plantation growth cycle, Coalburn catchment, Northern England, Journal of Hydrology, 519, 559–573, 2014.

Blöschl, G., Hall, J., Parajka, J., Perdigão, R. A., Merz, B., Arheimer, B., Aronica, G. T., Bilibashi, A., Bonacci, O., Borga, M., et al.: Changing climate shifts timing of European floods, Science, 357, 588–590, https://doi.org/10.1126/science.aan2506, 2017.

Blöschl, G., Hall, J., Viglione, A., Perdigão, R. A., Parajka, J., Merz, B., Lun, D., Arheimer, B., Aronica, G. T., Bilibashi, A., et al.: Changing climate both increases and decreases European river floods, Nature, 573, 108–111, 2019.

Blum, A. G., Ferraro, P. J., Archfield, S. A., and Ryberg, K. R.: Causal effect of impervious cover on annual flood magnitude for the United States, Geophysical Research Letters, 47, no–no, 2020.

Bonnet, R., Boé, J., and Habets, F.: Influence of multidecadal variability on high and low flows: the case of the Seine basin, Hydrology and Earth System Sciences, 24, 1611–1631, 2020.

Broderick, C., Matthews, T., Wilby, R. L., Bastola, S., and Murphy, C.: Transferability of hydrological models and ensemble averaging methods between contrasting climatic periods, Water Resources Research, 52, 8343–8373, 2016.

Broderick, C., Murphy, C., Wilby, R. L., Matthews, T., Prudhomme, C., and Adamson, M.: Using a scenario-neutral framework to avoid potential maladaptation to future flood risk, Water resources research, 55, 1079–1104, 2019.

Brown, C. and Wilby, R. L.: An alternate approach to assessing climate risks, Eos, Transactions American Geophysical Union, 93, 401–402, 2012.

Hydrology and Earth System Sciences
Open Access
Discussions

Brunner, M. I., Furrer, R., Sikorska, A. E., Viviroli, D., Seibert, J., and Favre, A.-C.: Synthetic design hydrographs for ungauged catchments: a comparison of regionalization methods, Stochastic Environmental Research and Risk Assessment, 32, 1993–2023, 2018.

Brunner, M. I., Hingray, B., Zappa, M., and Favre, A.-C.: Future Trends in the Interdependence Between Flood Peaks and Volumes: Hydro-Climatological Drivers and Uncertainty, Water Resources Research, 55, 4745–4759, 2019.

Brunner, M. I., Gilleland, E., Wood, A., Swain, D. L., and Clark, M.: Spatial dependence of floods shaped by spatiotemporal variations in meteorological and land-surface processes, Geophysical Research Letters, p. e2020GL088000, 2020.

Burnham, K. P. and Anderson, D. R.: Multimodel inference: understanding AIC and BIC in model selection, Sociological methods & research, 33, 261–304, 2004.

Buzan, J. R. and Huber, M.: Moist heat stress on a hotter Earth, Annual Review of Earth and Planetary Sciences, 48, 2020.

Caeiro, F. and Gomes, M. I.: Threshold selection in extreme value analysis, Extreme value modeling and risk analysis: Methods and applications, 1, 69–86, 2016.

Camargo, S. J. and Sobel, A. H.: Western North Pacific tropical cyclone intensity and ENSO, Journal of Climate, 18, 2996–3006, 2005.

Chamberlin, T. C.: The method of multiple working hypotheses, Science, 15, 92–96, 1890.

Chan, S., Kendon, E., Roberts, N., Fowler, H., and Blenkinsop, S.: The characteristics of summer sub-hourly rainfall over the southern UK in a high-resolution convective permitting model, Environmental Research Letters, 11, 094 024, 2016.

Cheng, L. and AghaKouchak, A.: Nonstationary precipitation intensity-duration-frequency curves for infrastructure design in a changing climate, Scientific reports, 4, 7093, https://doi.org/https://doi.org/10.1038/srep07093, 2014.

Cid-Serrano, L., Ramírez, S. M., Alfaro, E. J., and Enfield, D. B.: Analysis of the Latin American west coast rainfall predictability using an ENSO index, Atmósfera, 28, 191–203, 2015.

Clark, M. P., Kavetski, D., and Fenicia, F.: Pursuing the method of multiple working hypotheses for hydrological modeling, Water Resources Research, 47, 2011.

Clark, M. P., Wilby, R. L., Gutmann, E. D., Vano, J. A., Gangopadhyay, S., Wood, A. W., Fowler, H. J., Prudhomme, C., Arnold, J. R., and Brekke, L. D.: Characterizing uncertainty of the hydrologic impacts of climate change, Current Climate Change Reports, 2, 55–64, 2016.

Coccolo, S., Kämpf, J., Mauree, D., and Scartezzini, J.-L.: Cooling potential of greening in the urban environment, a step further towards practice, Sustainable cities and society, 38, 543–559, 2018.

Coles, S.: An Introduction to Statistical Modeling of Extreme Values, Springer Series in Statistics, Springer-Verlag, London, //www.springer.com/gp/book/9781852334598, 2001.

Collins, M. J.: River flood seasonality in the Northeast United States: Characterization and trends, Hydrological Processes, 33, 687–698, 2019.

Court, A.: Measures of streamflow timing, Journal of Geophysical Research, 67, 4335–4339, 1962.

Courty, L. G., Wilby, R. L., Hillier, J. K., and Slater, L. J.: Intensity-duration-frequency curves at the global scale, Environmental Research Letters, 14, 084 045, 2019.

Crooks, S. and Kay, A.: Simulation of river flow in the Thames over 120 years: Evidence of change in rainfall-runoff response?, Journal of Hydrology: Regional Studies, 4, 172–195, 2015.

Cunderlik, J. M. and Ouarda, T. B.: Trends in the timing and magnitude of floods in Canada, Journal of hydrology, 375, 471–480, 2009.

Cunderlik, J. M., Ouarda, T. B., and Bobée, B.: Determination of flood seasonality from hydrological records/Détermination de la saisonnalité des crues à partir de séries hydrologiques, Hydrological Sciences Journal, 49, 2004.

https://doi.org/10.5194/hess-2020-576
Preprint. Discussion started: 14 November 2020

Hydrology and Earth System Sciences Discussions

Dadson, S. J., Hall, J. W., Murgatroyd, A., Acreman, M., Bates, P., Beven, K., Heathwaite, L., Holden, J., Holman, I. P., Lane, S. N., et al.: A restatement of the natural science evidence concerning catchment-based 'natural' flood management in the UK, Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 473, 20160 706, 2017.

De Luca, P., Messori, G., Wilby, R. L., Mazzoleni, M., and Di Baldassarre, G.: Concurrent wet and dry hydrological extremes at the global scale, Earth System Dynamics, 11, 251–266, 2020.

De Niel, J. and Willems, P.: Climate or land cover variations: what is driving observed changes in river peak flows? A data-based attribution study., Hydrology & Earth System Sciences, 23, 2019.

de Ruiter, M. C., Couasnon, A., van den Homberg, M. J., Daniell, J. E., Gill, J. C., and Ward, P. J.: Why we can no longer ignore consecutive disasters, Earth's Future, 2019.

Deser, C., Lehner, F., Rodgers, K., Ault, T., Delworth, T., DiNezio, P., Fiore, A., Frankignoul, C., Fyfe, J., Horton, D., et al.: Insights from Earth system model initial-condition large ensembles and future prospects, Nature Climate Change, pp. 1–10, 2020.

Dickson, R. R., Meincke, J., Malmberg, S.-A., and Lee, A. J.: The "great salinity anomaly" in the northern North Atlantic 1968–1982, Progress in Oceanography, 20, 103–151, 1988.

Diffenbaugh, N. S., Singh, D., Mankin, J. S., Horton, D. E., Swain, D. L., Touma, D., Charland, A., Liu, Y., Haugen, M., Tsiang, M., et al.: Quantifying the influence of global warming on unprecedented extreme climate events, Proceedings of the National Academy of Sciences, 114, 4881–4886, 2017.

Do, H. X., Westra, S., and Leonard, M.: A global-scale investigation of trends in annual maximum streamflow, Journal of hydrology, 552, 28–43, 2017.

Donat, M., Renggli, D., Wild, S., Alexander, L., Leckebusch, G., and Ulbrich, U.: Reanalysis suggests long-term upward trends in European storminess since 1871, Geophysical Research Letters, 38, 2011.

Donat, M., Alexander, L., Yang, H., Durre, I., Vose, R., Dunn, R., Willett, K., Aguilar, E., Brunet, M., Caesar, J., et al.: Updated analyses of temperature and precipitation extreme indices since the beginning of the twentieth century: The HadEX2 dataset, Journal of Geophysical Research: Atmospheres, 118, 2098–2118, 2013.

Donat, M. G., Lowry, A. L., Alexander, L. V., O'Gorman, P. A., and Maher, N.: More extreme precipitation in the world's dry and wet regions, Nature Climate Change, 6, 508, https://doi.org/https://doi.org/10.1038/nclimate2941, 2016.

Duan, Q., Pappenberger, F., Wood, A., Cloke, H. L., and Schaake, J.: Handbook of Hydrometeorological Ensemble Forecasting, Springer, 2019.

ECMWF: C3S Climate projections, https://confluence.ecmwf.int/display/CKB/C3S+Climate+projections, 2020a.

ECMWF: Seasonal forecasts and the Copernicus Climate Change Service, https://confluence.ecmwf.int/display/CKB/C3S+Seasonal+Forecasts, 2020b.

Elsner, J. B., Kossin, J. P., and Jagger, T. H.: The increasing intensity of the strongest tropical cyclones, Nature, 455, 92–95, 2008.

Emanuel, K. and Center, L.: Response of Global Tropical Cyclone Activity to Increasing CO2: Results from Downscaling CMIP6 Models, Journal of Climate, pp. 1–54, 2020.

Emerton, R., Cloke, H., Ficchi, A., Hawker, L., de Wit, S., Speight, L., Prudhomme, C., Rundell, P., West, R., Neal, J., et al.: Emergency flood bulletins for Cyclones Idai and Kenneth: A critical evaluation of the use of global flood forecasts for international humanitarian preparedness and response, International Journal of Disaster Risk Reduction, 50, 101 811, 2020.

Emerton, R. E., Stephens, E. M., Pappenberger, F., Pagano, T. C., Weerts, A. H., Wood, A. W., Salamon, P., Brown, J. D., Hjerdt, N., Donnelly, C., et al.: Continental and global scale flood forecasting systems, Wiley Interdisciplinary Reviews: Water, 3, 391–418, 2016.

https://doi.org/10.5194/hess-2020-576
Preprint. Discussion started: 14 November 2020

Hydrology and Earth System Sciences
Open Access
Discussions

Enfield, D. B. and Mayer, D. A.: Tropical Atlantic sea surface temperature variability and its relation to El Niño-Southern Oscillation, Journal of Geophysical Research: Oceans, 102, 929–945, 1997.

Environment Agency, U.: Adapting to Climate Change: Advice for Flood and Coastal Erosion Risk Management Authorities, https://www.gov.uk/government/publications/adapting-to-climate-change-for-risk-management-authorities, 2016.

Erdman, C. and Emerson, J. W.: A fast Bayesian change point analysis for the segmentation of microarray data, Bioinformatics, 24, 2143–2148, 2008.

Ferguson, C. R. and Villarini, G.: An evaluation of the statistical homogeneity of the Twentieth Century Reanalysis, Climate dynamics, 42, 2841–2866, 2014.

Fernando, D. N., Chakraborty, S., Fu, R., and Mace, R. E.: A process-based statistical seasonal prediction of May–July rainfall anomalies over Texas and the Southern Great Plains of the United States, Climate Services, 16, 100 133, 2019.

Ferreira, S. and Ghimire, R.: Forest cover, socioeconomics, and reported flood frequency in developing countries, Water resources research, 48, 2012.

Fischer, E. M. and Knutti, R.: Anthropogenic contribution to global occurrence of heavy-precipitation and high-temperature extremes, Nature Climate Change, 5, 560–564, 2015.

Fowler, K., Knoben, W., Peel, M., Peterson, T., Ryu, D., Saft, M., Seo, K.-W., and Western, A.: Many commonly used rainfall-runoff models lack long, slow dynamics: implications for runoff projections, Water Resources Research, 56, e2019WR025 286, 2020.

Fowler, K. J., Peel, M. C., Western, A. W., Zhang, L., and Peterson, T. J.: Simulating runoff under changing climatic conditions: Revisiting an apparent deficiency of conceptual rainfall-runoff models, Water Resources Research, 52, 1820–1846, 2016.

François, B., Schlef, K., Wi, S., and Brown, C.: Design considerations for riverine floods in a changing climate–a review, Journal of Hydrology, 574, 557–573, 2019.

Frich, P., Alexander, L. V., Della-Marta, P., Gleason, B., Haylock, M., Tank, A. K., and Peterson, T.: Observed coherent changes in climatic extremes during the second half of the twentieth century, Climate research, 19, 193–212, 2002.

Fryzlewicz, P. et al.: Wild binary segmentation for multiple change-point detection, The Annals of Statistics, 42, 2243–2281, 2014.

Ganguli, P. and Coulibaly, P.: Does nonstationarity in rainfall requires nonstationary intensity-duration-frequency curves?, Hydrology and Earth System Sciences, 21, 6461–6483, 2017.

Gao, J., Kirkby, M., and Holden, J.: The effect of interactions between rainfall patterns and land-cover change on flood peaks in upland peatlands, Journal of Hydrology, 567, 546–559, 2018.

Gao, M., Mo, D., and Wu, X.: Nonstationary modeling of extreme precipitation in China, Atmospheric Research, 182, 1–9, 2016.

Gaupp, F., Hall, J., Hochrainer-Stigler, S., and Dadson, S.: Changing risks of simultaneous global breadbasket failure, Nature Climate Change, 10, 54–57, 2020.

Gibbs, W. and Maher, J.: Rainfall deciles as drought indicators, bureau of meteorology bulletin no. 48, Commonwealth of Australia, Melbourne, 29, 1967.

Gilliland, J. M. and Keim, B. D.: Surface wind speed: trend and climatology of Brazil from 1980–2014, International Journal of Climatology, 38, 1060–1073, 2018.

Gleick, P. H. and Palaniappan, M.: Peak water limits to freshwater withdrawal and use, Proceedings of the National Academy of Sciences, 107, 11 155–11 162, 2010.

Grill, G., Lehner, B., Lumsdon, A. E., MacDonald, G. K., Zarfl, C., and Liermann, C. R.: An index-based framework for assessing patterns and trends in river fragmentation and flow regulation by global dams at multiple scales, Environmental Research Letters, 10, 015 001, 2015.

Gu, X., Zhang, Q., Singh, V. P., and Shi, P.: Nonstationarity in timing of extreme precipitation across China and impact of tropical cyclones, Global and Planetary Change, 149, 153–165, 2017.

Gubler, S., Sedlmeier, K., Bhend, J., Avalos, G., Coelho, C., Escajadillo, Y., Jacques-Coper, M., Martinez, R., Schwierz, C., De Skansi, M., et al.: Assessment of ECMWF SEAS5 seasonal forecast performance over South America, Weather and Forecasting, 35, 561–584, 2020.

Gudmundsson, L. and Seneviratne, S. I.: Anthropogenic climate change affects meteorological drought risk in Europe, Environmental Research Letters, 11, 044 005, 2016.

Gupta, A. S., Jourdain, N. C., Brown, J. N., and Monselesan, D.: Climate drift in the CMIP5 models, Journal of Climate, 26, 8597–8615, 2013.

Hall, J. and Blöschl, G.: Spatial patterns and characteristics of flood seasonality in Europe., Hydrology & Earth System Sciences, 22, 2018.

Hall, J., Arheimer, B., Borga, M., Brázdil, R., Claps, P., Kiss, A., Kjeldsen, T. R., Kriaučuniene, J., Kundzewicz, Z. W., Lang, M., et al.: Understanding flood regime changes in Europe: a state of the art assessment, Hydrology and Earth System Sciences Discussions, 10, 15 525–15 624, 2013.

Han, S. and Coulibaly, P.: Probabilistic flood forecasting using hydrologic uncertainty processor with ensemble weather forecasts, Journal of Hydrometeorology, 20, 1379–1398, 2019.

Hannaford, J. and Marsh, T. J.: High-flow and flood trends in a network of undisturbed catchments in the UK, International Journal of Climatology: A Journal of the Royal Meteorological Society, 28, 1325–1338, 2008.

Hannaford, J., Buys, G., Stahl, K., and Tallaksen, L.: The influence of decadal-scale variability on trends in long European streamflow records, Hydrology and Earth System Sciences, 17, 2717–2733, 2013.

Hänsel, S., Schucknecht, A., and Matschullat, J.: The Modified Rainfall Anomaly Index (mRAI)—is this an alternative to the Standardised Precipitation Index (SPI) in evaluating future extreme precipitation characteristics?, Theoretical and applied climatology, 123, 827–844, 2016.

Hao, Z., Singh, V. P., and Xia, Y.: Seasonal drought prediction: advances, challenges, and future prospects, Reviews of Geophysics, 56, 108–141, 2018.

Harding, A., Rivington, M., Mineter, M., and Tett, S.: Agro-meteorological indices and climate model uncertainty over the UK, Climatic Change, 128, 113–126, 2015.

Harrigan, S., Murphy, C., Hall, J., Wilby, R., and Sweeney, J.: Attribution of detected changes in streamflow using multiple working hypotheses, Hydrology and Earth System Sciences, 18, 1935, 2014.

Harrigan, S., Hannaford, J., Muchan, K., and Marsh, T. J.: Designation and trend analysis of the updated UK Benchmark Network of river flow stations: The UKBN2 dataset, Hydrology Research, 49, 552–567, 2018.

Harrigan, S., Cloke, H., and Pappenberger, F.: Innovating global hydrological prediction through an Earth system approach, WMO Bulletin, 69, 2020.

Harrison, P. A., Dunford, R. W., Holman, I. P., Cojocaru, G., Madsen, M. S., Chen, P.-Y., Pedde, S., and Sandars, D.: Differences between low-end and high-end climate change impacts in Europe across multiple sectors, Regional Environmental Change, 19, 695–709, 2019.

Hart, N. C., Gray, S. L., and Clark, P. A.: Sting-jet windstorms over the North Atlantic: climatology and contribution to extreme wind risk, Journal of Climate, 30, 5455–5471, 2017.

Hartmann, D., Klein Tank, A., Rusticucci, M., Alexander, L., Bronnimann, S., Charabi, Y., Dentener, F., Dlugokencky, E., Easterling, D., Kaplan, A., Soden, B., Thorne, P., Wild, M., and Zhai, P.: Observations: Atmosphere and Surface, book section 2, p. 159–254, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, https://doi.org/10.1017/CBO9781107415324.008, www.climatechange2013.org, 2013.

Hastie, T. and Tibshirani, R.: Generalized additive models: some applications, Journal of the American Statistical Association, 82, 371–386, 1987.

Hausfather, Z., Menne, M. J., Williams, C. N., Masters, T., Broberg, R., and Jones, D.: Quantifying the effect of urbanization on US Historical Climatology Network temperature records, Journal of Geophysical Research: Atmospheres, 118, 481–494, 2013.

Haynes, K., Fearnhead, P., and Eckley, I. A.: A computationally efficient nonparametric approach for changepoint detection, Statistics and Computing, 27, 1293–1305, 2017.

Hecht, J. S. and Vogel, R. M.: Updating urban design floods for changes in central tendency and variability using regression, Advances in Water Resources, 136, 103 484, 2020.

Hegerl, G. C., Hoegh-Guldberg, O., Casassa, G., Hoerling, M. P., Kovats, R., Parmesan, C., Pierce, D. W., Stott, P. A., et al.: Good practice guidance paper on detection and attribution related to anthropogenic climate change, in: Meeting report of the intergovernmental panel on climate change expert meeting on detection and attribution of anthropogenic climate change, IPCC Working Group I Technical Support Unit, University of Bern, Bern, Switzerland, 2010.

Held, I. M. and Soden, B. J.: Robust responses of the hydrological cycle to global warming, Journal of climate, 19, 5686–5699, 2006.

Helsel, D., Hirsch, R., Ryberg, K., Archfield, S., and Gilroy, E.: Statistical methods in water resources: U.S. Geological Survey Techniques and Methods, Elsevier, https://doi.org/10.3133/tm4a3, supersedes USGS Techniques of Water-Resources Investigations, book 4, chapter A3, version 1.1., 2020.

Hillier, J. K., Matthews, T., Wilby, R. L., and Murphy, C.: Multi-hazard dependencies can increase or decrease risk, Nature Climate Change, pp. 1–4, 2020.

Hipel, K. W. and McLeod, A. I.: Time series modelling of water resources and environmental systems, Elsevier, 1994.

Hirschboeck, K. K.: Flood hydroclimatology, Flood geomorphology, 27, 49, 1988.

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

Hodgkins, G. A. and Dudley, R. W.: Changes in the timing of winter–spring streamflows in eastern North America, 1913–2002, Geophysical Research Letters, 33, 2006.

Hoerling, M., Eischeid, J., Perlwitz, J., Quan, X., Zhang, T., and Pegion, P.: On the increased frequency of Mediterranean drought, Journal of climate, 25, 2146–2161, 2012.

Hrachowitz, M. and Clark, M. P.: HESS Opinions: The complementary merits of competing modelling philosophies in hydrology, Hydrol. Earth Syst. Sci, 21, 3953–3973, 2017.

Hundecha, Y., St-Hilaire, A., Ouarda, T., El Adlouni, S., and Gachon, P.: A nonstationary extreme value analysis for the assessment of changes in extreme annual wind speed over the Gulf of St. Lawrence, Canada, Journal of applied meteorology and climatology, 47, 2745–2759, 2008.

Iacob, O., Brown, I., and Rowan, J.: Natural flood management, land use and climate change trade-offs: the case of Tarland catchment, Scotland, Hydrological Sciences Journal, 62, 1931–1948, 2017.

Immerzeel, W. W., Lutz, A., Andrade, M., Bahl, A., Biemans, H., Bolch, T., Hyde, S., Brumby, S., Davies, B., Elmore, A., et al.: Importance and vulnerability of the world's water towers, Nature, 577, 364–369, 2020.

International Hydropower Association and others: Hydropower Sector Climate Resilience Guide, https://www.hydropower.org/publications/hydropower-sector-climate-resilience-guide, 2019.

IPCC: Summary for Policymakers. In: Climate Change and Land: an IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems [P.R. Shukla, J. Skea, E. Calvo Buendia, V. Masson-Delmotte, H.-O. Pörtner, D. C. Roberts, P. Zhai, R. Slade, S. Connors, R. van Diemen, M. Ferrat, E. Haughey, S. Luz, S. Neogi, M. Pathak, J. Petzold, J. Portugal Pereira, P. Vyas, E. Huntley, K. Kissick, M. Belkacemi, J. Malley, (eds.)] in press, 2019.

James, L. A.: Incision and morphologic evolution of an alluvial channel recovering from hydraulic mining sediment, Geological Society of America Bulletin, 103, 723–736, 1991.

Johnson, S. J., Stockdale, T. N., Ferranti, L., Balmaseda, M. A., Molteni, F., Magnusson, L., Tietsche, S., Decremer, D., Weisheimer, A., Balsamo, G., et al.: SEAS5: The new ECMWF seasonal forecast system, Geoscientific Model Development, 12, 2019.

Jones, P. D., Harpham, C., and Lister, D.: Long-term trends in gale days and storminess for the Falkland Islands, International Journal of Climatology, 36, 1413–1427, 2016.

Jovanovic, T., Mejía, A., Gall, H., and Gironás, J.: Effect of urbanization on the long-term persistence of streamflow records, Physica A: Statistical Mechanics and its Applications, 447, 208–221, 2016.

Karaseva, M. O., Prakash, S., and Gairola, R.: Validation of high-resolution TRMM-3B43 precipitation product using rain gauge measurements over Kyrgyzstan, Theoretical and Applied Climatology, 108, 147–157, 2012.

Karl, T. R., Meehl, G. A., Miller, C. D., Hassol, S. J., Waple, A. M., and Murray, W. L.: Weather and climate extremes in a changing climate, Tech. rep., US Climate Change Science Program, 2008.

Katz, R. W.: Statistical methods for nonstationary extremes, in: Extremes in a Changing Climate, pp. 15–37, Springer, 2013.

Kelder, T., Muller, M., Slater, L., Marjoribanks, T., Wilby, R. L., Prudhomme, C., Bohlinger, P., Ferranti, L., and Nipen, T.: Using UNSEEN trends to detect decadal changes in 100-year precipitation extremes, npj Climate and Atmospheric Science (in press), pp. 1–13, https://doi.org/10.1038/s41612-020-00149-4, 2020.

Kemter, M., Merz, B., Marwan, N., Vorogushyn, S., and Blöschl, G.: Joint trends in flood magnitudes and spatial extents across Europe, Geophysical Research Letters, 47, e2020GL087 464, 2020.

Kendall, M.: Rank Correlation Methods; Charles Griffin: London, 1975.

Kendon, E. J., Roberts, N. M., Fowler, H. J., Roberts, M. J., Chan, S. C., and Senior, C. A.: Heavier summer downpours with climate change revealed by weather forecast resolution model, Nature Climate Change, 4, 570–576, 2014.

Kettner, A. J., Cohen, S., Overeem, I., Fekete, B. M., Brakenridge, G. R., and Syvitski, J. P.: Estimating Change in Flooding for the 21st Century Under a Conservative RCP Forcing: A Global Hydrological Modeling Assessment, book section 9, pp. 157–167, Wiley Online Library, https://doi.org/10.1002/9781119217886.ch9, 2018.

Kharin, V. V., Zwiers, F., Zhang, X., and Wehner, M.: Changes in temperature and precipitation extremes in the CMIP5 ensemble, Climatic change, 119, 345–357, 2013.

Khouakhi, A., Villarini, G., Zhang, W., and Slater, L. J.: Seasonal predictability of high sea level frequency using ENSO patterns along the US West Coast, Advances in Water Resources, 131, 103 377, 2019.

Killick, R., Fearnhead, P., and Eckley, I. A.: Optimal detection of changepoints with a linear computational cost, Journal of the American Statistical Association, 107, 1590–1598, 2012.

https://doi.org/10.5194/hess-2020-576
Preprint. Discussion started: 14 November 2020

Kim, H.-M. and Webster, P. J.: Extended-range seasonal hurricane forecasts for the North Atlantic with a hybrid dynamical-statistical model, Geophysical research letters, 37, 2010.

Kirchner, J. W.: Catchments as simple dynamical systems: Catchment characterization, rainfall-runoff modeling, and doing hydrology backward, Water Resources Research, 45, 2009.

Kirtman, B. P., Min, D., Infanti, J. M., Kinter III, J. L., Paolino, D. A., Zhang, Q., Van Den Dool, H., Saha, S., Mendez, M. P., Becker, E., et al.: The North American multimodel ensemble: phase-1 seasonal-to-interannual prediction; phase-2 toward developing intraseasonal prediction, Bulletin of the American Meteorological Society, 95, 585–601, 2014.

Kjellström, E., Bärring, L., Jacob, D., Jones, R., Lenderink, G., and Schär, C.: Modelling daily temperature extremes: recent climate and future changes over Europe, Climatic Change, 81, 249–265, 2007.

Knutson, T. R. and Ploshay, J. J.: Detection of anthropogenic influence on a summertime heat stress index, Climatic Change, 138, 25–39, 2016.

Kornhuber, K., Coumou, D., Vogel, E., Lesk, C., Donges, J. F., Lehmann, J., and Horton, R. M.: Amplified Rossby waves enhance risk of concurrent heatwaves in major breadbasket regions, Nature Climate Change, 10, 48–53, 2020.

Kossin, J. P.: A global slowdown of tropical-cyclone translation speed, Nature, 558, 104–107, 2018.

Koutsoyiannis, D.: Nonstationarity versus scaling in hydrology, Journal of Hydrology, 324, 239–254, 2006.

Koutsoyiannis, D. and Montanari, A.: Negligent killing of scientific concepts: the stationarity case, Hydrological Sciences Journal, 60, 1174–1183, 2015.

Krishnan, A. and Bhaskaran, P. K.: Skill assessment of global climate model wind speed from CMIP5 and CMIP6 and evaluation of projections for the Bay of Bengal, Climate Dynamics, pp. 1–21, 2020.

Kundzewicz, Z. W. and Stakhiv, E. Z.: Are climate models “ready for prime time” in water resources management applications, or is more research needed?, Hydrological Sciences Journal–Journal des Sciences Hydrologiques, 55, 1085–1089, 2010.

Kundzewicz, Z. W., Graczyk, D., Maurer, T., Pińskwar, I., Radziejewski, M., Svensson, C., and Szwed, M.: Trend detection in river flow series: 1. Annual maximum flow/Détection de tendance dans des séries de débit fluvial: 1. Débit maximum annuel, Hydrological Sciences Journal, 50, 2005.

Lackmann, G. M.: Hurricane Sandy before 1900 and after 2100, Bulletin of the American Meteorological Society, 96, 547–560, 2015.

Lang, M., Ouarda, T., and Bobée, B.: Towards operational guidelines for over-threshold modeling, Journal of hydrology, 225, 103–117, 1999.

Lavers, D. A., Allan, R. P., Wood, E. F., Villarini, G., Brayshaw, D. J., and Wade, A. J.: Winter floods in Britain are connected to atmospheric rivers, Geophysical Research Letters, 38, 2011.

Lavers, D. A., Villarini, G., Allan, R. P., Wood, E. F., and Wade, A. J.: The detection of atmospheric rivers in atmospheric reanalyses and their links to British winter floods and the large-scale climatic circulation, Journal of Geophysical Research: Atmospheres, 117, 2012.

Leckebusch, G. C., Renggli, D., and Ulbrich, U.: Development and application of an objective storm severity measure for the Northeast Atlantic region, Meteorologische Zeitschrift, 17, 575–587, 2008.

Lee, K. and Singh, V. P.: Analysis of Uncertainty and Non-stationarity in Probable Maximum Precipitation in Brazos River Basin, Journal of Hydrology, p. 125526, 2020.

Lee, O., Sim, I., and Kim, S.: Application of the non-stationary peak-over-threshold methods for deriving rainfall extremes from temperature projections, Journal of Hydrology, 585, 124 318, 2020.

Leelaruban, N. and Padmanabhan, G.: Drought occurrences and their characteristics across selected spatial scales in the Contiguous United States, Geosciences, 7, 59, 2017.

Lenderink, G. and Van Meijgaard, E.: Increase in hourly precipitation extremes beyond expectations from temperature changes, Nature Geoscience, 1, 511–514, 2008.

Lenton, T. M., Held, H., Kriegler, E., Hall, J. W., Lucht, W., Rahmstorf, S., and Schellnhuber, H. J.: Tipping elements in the Earth's climate system, Proceedings of the national Academy of Sciences, 105, 1786–1793, 2008.

Levy, M., Lopes, A., Cohn, A., Larsen, L., and Thompson, S.: Land use change increases streamflow across the arc of deforestation in Brazil, Geophysical Research Letters, 45, 3520–3530, 2018.

Li, W., Jiang, Z., Xu, J., and Li, L.: Extreme precipitation indices over China in CMIP5 models. Part II: probabilistic projection, Journal of Climate, 29, 8989–9004, 2016.

Li, Y., Fowler, H. J., Argüeso, D., Blenkinsop, S., Evans, J. P., Lenderink, G., Yan, X., Guerreiro, S. B., Lewis, E., and Li, X.-F.: Strong intensification of hourly rainfall extremes by urbanization, Geophysical Research Letters, p. e2020GL088758, 2019.

Li, Y., Wright, D. B., and Byrne, P. K.: The Influence of Tropical Cyclones on the Evolution of River Conveyance Capacity in Puerto Rico, Water Resources Research, p. e2020WR027971, 2020.

Liepert, B. G. and Lo, F.: CMIP5 update of 'Inter-model variability and biases of the global water cycle in CMIP3 coupled climate models', Environmental Research Letters, 8, 029 401, 2013.

Lomas, K. J. and Giridharan, R.: Thermal comfort standards, measured internal temperatures and thermal resilience to climate change of free-running buildings: A case-study of hospital wards, Building and Environment, 55, 57–72, 2012.

Longobardi, A. and Villani, P.: Trend analysis of annual and seasonal rainfall time series in the Mediterranean area, International journal of Climatology, 30, 1538–1546, 2010.

Lorenz, R., Stalhandske, Z., and Fischer, E. M.: Detection of a climate change signal in extreme heat, heat stress, and cold in Europe from observations, Geophysical Research Letters, 46, 8363–8374, 2019.

Lorenzo-Lacruz, J., Vicente-Serrano, S. M., López-Moreno, J. I., Morán-Tejeda, E., and Zabalza, J.: Recent trends in Iberian streamflows (1945–2005), Journal of Hydrology, 414, 463–475, 2012.

Macdonald, N., Werritty, A., Black, A., and McEwen, L.: Historical and pooled flood frequency analysis for the River Tay at Perth, Scotland, Area, 38, 34–46, 2006.

Macdonald, N., Phillips, I. D., and Mayle, G.: Spatial and temporal variability of flood seasonality in Wales, Hydrological processes, 24, 1806–1820, 2010.

Maher, N., Matei, D., Milinski, S., and Marotzke, J.: ENSO change in climate projections: Forced response or internal variability?, Geophysical Research Letters, 45, 11–390, 2018.

Maher, N., Lehner, F., and Marotzke, J.: Quantifying the role of internal variability in the temperature we expect to observe in the coming decades, Environmental Research Letters, 15, 054 014, 2020.

Mahmood, R., Pielke Sr, R. A., Hubbard, K. G., Niyogi, D., Dirmeyer, P. A., McAlpine, C., Carleton, A. M., Hale, R., Gameda, S., Beltrán-Przekurat, A., et al.: Land cover changes and their biogeophysical effects on climate, International journal of climatology, 34, 929–953, 2014.

Mallakpour, I. and Villarini, G.: The changing nature of flooding across the central United States, Nature Climate Change, 5, 250–254, 2015.

Mallakpour, I. and Villarini, G.: A simulation study to examine the sensitivity of the Pettitt test to detect abrupt changes in mean, Hydrological Sciences Journal, 61, 245–254, 2016.

Mann, H. B.: Nonparametric tests against trend, Econometrica: Journal of the econometric society, pp. 245–259, 1945.

Mann, H. B. and Whitney, D. R.: On a test of whether one of two random variables is stochastically larger than the other, The annals of mathematical statistics, pp. 50–60, 1947.

Martínez-Alvarado, O., Gray, S. L., Hart, N. C., Clark, P. A., Hodges, K., and Roberts, M. J.: Increased wind risk from sting-jet windstorms with climate change, Environmental research letters, 13, 044 002, 2018.

Massey, N., Jones, R., Otto, F., Aina, T., Wilson, S., Murphy, J., Hassell, D., Yamazaki, Y., and Allen, M.: weather@ home—development and validation of a very large ensemble modelling system for probabilistic event attribution, Quarterly Journal of the Royal Meteorological Society, 141, 1528–1545, 2015.

Masys, A. J., Yee, E., and Vallerand, A.: 'Black Swans','Dragon Kings' and Beyond: Towards Predictability and Suppression of Extreme All-Hazards Events Through Modeling and Simulation, in: Applications of Systems Thinking and Soft Operations Research in Managing Complexity, pp. 131–141, Springer, 2016.

Matthews, T.: Humid heat and climate change, Progress in Physical Geography: Earth and Environment, 42, 391–405, 2018.

Matthews, T.: Death valley: world-beating temperatures, no sweat, Weather, https://rmets.onlinelibrary.wiley.com/doi/full/10.1002/wea.3858, 2020.

Matthews, T., Mullan, D., Wilby, R. L., Broderick, C., and Murphy, C.: Past and future climate change in the context of memorable seasonal extremes, Climate Risk Management, 11, 37–52, 2016a.

Matthews, T., Murphy, C., Wilby, R. L., and Harrigan, S.: A cyclone climatology of the British-Irish Isles 1871–2012, International Journal of Climatology, 36, 1299–1312, 2016b.

Matthews, T., Wilby, R. L., and Murphy, C.: An emerging tropical cyclone–deadly heat compound hazard, Nature Climate Change, 9, 602–606, 2019.

Matthews, T. K., Wilby, R. L., and Murphy, C.: Communicating the deadly consequences of global warming for human heat stress, Proceedings of the National Academy of Sciences, 114, 3861–3866, 2017.

McCarthy, G. D., Gleeson, E., and Walsh, S.: The influence of ocean variations on the climate of Ireland, Weather, 70, 242–245, 2015.

McIntosh, B. S., Ascough II, J. C., Twery, M., Chew, J., Elmahdi, A., Haase, D., Harou, J. J., Hepting, D., Cuddy, S., Jakeman, A. J., et al.: Environmental decision support systems (EDSS) development–Challenges and best practices, Environmental Modelling & Software, 26, 1389–1402, 2011.

McKee, T. B., Doesken, N. J., Kleist, J., et al.: The relationship of drought frequency and duration to time scales, in: Proceedings of the 8th Conference on Applied Climatology, 22, pp. 179–183, Boston, 1993.

McSweeney, C., Jones, R., Lee, R. W., and Rowell, D.: Selecting CMIP5 GCMs for downscaling over multiple regions, Climate Dynamics, 44, 3237–3260, 2015.

Mediero, L., Santillán, D., Garrote, L., and Granados, A.: Detection and attribution of trends in magnitude, frequency and timing of floods in Spain, Journal of Hydrology, 517, 1072–1088, 2014.

Mei, W., Xie, S.-P., Primeau, F., McWilliams, J. C., and Pasquero, C.: Northwestern Pacific typhoon intensity controlled by changes in ocean temperatures, Science Advances, 1, e1500 014, 2015.

Mekonen, A. A., Berlie, A. B., and Ferede, M. B.: Spatial and temporal drought incidence analysis in the northeastern highlands of Ethiopia, Geoenvironmental Disasters, 7, 1–17, 2020.

Mestre, O., Domonkos, P., Picard, F., Auer, I., Robin, S., Lebarbier, E., Böhm, R., Aguilar, E., Guijarro, J. A., Vertacnik, G., et al.: HOMER: a homogenization software–methods and applications, 2013.

Milly, P. C., Betancourt, J., Falkenmark, M., Hirsch, R. M., Kundzewicz, Z. W., Lettenmaier, D. P., and Stouffer, R. J.: Stationarity is dead: Whither water management?, Science, 319, 573–574, 2008.

Milly, P. C. D., Wetherald, R. T., Dunne, K., and Delworth, T. L.: Increasing risk of great floods in a changing climate, Nature, 415, 514–517, 2002.

Miralles, D. G., Teuling, A. J., Van Heerwaarden, C. C., and De Arellano, J. V.-G.: Mega-heatwave temperatures due to combined soil desiccation and atmospheric heat accumulation, Nature geoscience, 7, 345–349, 2014.

Mishra, A. K. and Singh, V. P.: A review of drought concepts, Journal of hydrology, 391, 202–216, 2010.

Moberg, A. and Jones, P. D.: Trends in indices for extremes in daily temperature and precipitation in central and western Europe, 1901–99, International Journal of Climatology: A Journal of the Royal Meteorological Society, 25, 1149–1171, 2005.

Mood, A. M.: On the asymptotic efficiency of certain nonparametric two-sample tests, The Annals of Mathematical Statistics, pp. 514–522, 1954.

Mora, C., Dousset, B., Caldwell, I. R., Powell, F. E., Geronimo, R. C., Bielecki, C. R., Counsell, C. W., Dietrich, B. S., Johnston, E. T., Louis, L. V., et al.: Global risk of deadly heat, Nature climate change, 7, 501–506, 2017.

Murphy, C., Wilby, R. L., Matthews, T., Horvath, C., Crampsie, A., Ludlow, F., Noone, S., Brannigan, J., Hannaford, J., McLeman, R., et al.: The forgotten drought of 1765–1768: Reconstructing and re-evaluating historical droughts in the British and Irish Isles, International Journal of Climatology, 2020a.

Murphy, C., Wilby, R. L., Matthews, T. K., Thorne, P., Broderick, C., Fealy, R., Hall, J., Harrigan, S., Jones, P., McCarthy, G., et al.: Multi-century trends to wetter winters and drier summers in the England and Wales precipitation series explained by observational and sampling bias in early records, International Journal of Climatology, 40, 610–619, 2020b.

Murray, R. J. and Simmonds, I.: A numerical scheme for tracking cyclone centres from digital data. Part I: Development and operation of the scheme, Aust. Meteor. Mag, 39, 155–166, 1991.

Mwagona, P. C., Yao, Y., Shan, Y., Yu, H., and Zhang, Y.: Trend and Abrupt Regime Shift of Temperature Extreme in Northeast China, 1957–2015, Advances in Meteorology, 2018, 2018.

Naveau, P., Hannart, A., and Ribes, A.: Statistical methods for extreme event attribution in climate science, Annual Review of Statistics and Its Application, 7, 89–110, 2020.

Neri, A., Villarini, G., Slater, L. J., and Napolitano, F.: On the statistical attribution of the frequency of flood events across the US Midwest, Advances in Water Resources, 2019.

Niu, X., Wang, S., Tang, J., Lee, D.-K., Gutowski, W., Dairaku, K., McGregor, J., Katzfey, J., Gao, X., Wu, J., et al.: Ensemble evaluation and projection of climate extremes in China using RMIP models, International Journal of Climatology, 38, 2039–2055, 2018.

Noone, S., Murphy, C., Coll, J., Matthews, T., Mullan, D., Wilby, R. L., and Walsh, S.: Homogenization and analysis of an expanded long-term monthly rainfall network for the Island of Ireland (1850–2010), International Journal of Climatology, 36, 2837–2853, 2016.

O'Connor, P., Murphy, C., Matthews, T., and Wilby, R.: Reconstructed monthly river flows for Irish catchments 1766-2010, Geosciences Data Journal, 2020.

Oliver, E. C., Donat, M. G., Burrows, M. T., Moore, P. J., Smale, D. A., Alexander, L. V., Benthuysen, J. A., Feng, M., Gupta, A. S., Hobday, A. J., et al.: Longer and more frequent marine heatwaves over the past century, Nature communications, 9, 1–12, 2018.

Ouarda, T. B. and Charron, C.: Nonstationary Temperature-Duration-Frequency curves, Scientific reports, 8, 1–8, 2018.

O'Reilly, C. H., Zanna, L., and Woollings, T.: Assessing External and Internal Sources of Atlantic Multidecadal Variability Using Models, Proxy Data, and Early Instrumental Indices, Journal of Climate, 32, 7727–7745, 2019.

Palmer, W. C.: Meteorological drought, Research paper no. 45, US Weather Bureau, Washington, DC, 58, 1965.

Paltan, H., Waliser, D., Lim, W. H., Guan, B., Yamazaki, D., Pant, R., and Dadson, S.: Global floods and water availability driven by atmospheric rivers, Geophysical Research Letters, 44, 10–387, 2017.

Papacharalampous, G. and Tyralis, H.: Hydrological time series forecasting using simple combinations: Big data testing and investigations on one-year ahead river flow predictability, Journal of Hydrology, 590, 125 205, https://doi.org/10.1016/j.jhydrol.2020.125205, 2020.

Papalexiou, S. M. and Montanari, A.: Global and regional increase of precipitation extremes under global warming, Water Resources Research, 55, 4901–4914, 2019.

Papalexiou, S. M., AghaKouchak, A., Trenberth, K. E., and Foufoula-Georgiou, E.: Global, regional, and megacity trends in the highest temperature of the year: Diagnostics and evidence for accelerating trends, Earth's future, 6, 71–79, 2018.

Parry, S., Prudhomme, C., Wilby, R. L., and Wood, P. J.: Drought termination: Concept and characterisation, Progress in Physical Geography, 40, 743–767, 2016.

Peña-Angulo, D., Vicente-Serrano, S. M., Domínguez-Castro, F., Murphy, C., Reig, F., Tramblay, Y., Trigo, R. M., Luna, M. Y., Turco, M., Noguera, I., et al.: Long-term precipitation in Southwestern Europe reveals no clear trend attributable to anthropogenic forcing, Environmental Research Letters, 2020.

Perkins, S. E. and Alexander, L. V.: On the measurement of heat waves, Journal of Climate, 26, 4500–4517, 2013.

Perkins-Kirkpatrick, S. and Lewis, S.: Increasing trends in regional heatwaves, Nature Communications, 11, 1–8, 2020.

Peterson, T. C., Willett, K. M., and Thorne, P. W.: Observed changes in surface atmospheric energy over land, Geophysical Research Letters, 38, 2011.

Pielke Sr, R. A., Davey, C., and Morgan, J.: Assessing "global warming" with surface heat content, Eos, Transactions American Geophysical Union, 85, 210–211, 2004.

Pinter, N., Ickes, B. S., Wlosinski, J. H., and Van der Ploeg, R. R.: Trends in flood stages: Contrasting results from the Mississippi and Rhine River systems, Journal of Hydrology, 331, 554–566, 2006.

Pinter, N., Jemberie, A. A., Remo, J. W., Heine, R. A., and Ickes, B. S.: Flood trends and river engineering on the Mississippi River system, Geophysical Research Letters, 35, 2008.

Poff, N. L., Brown, C. M., Grantham, T. E., Matthews, J. H., Palmer, M. A., Spence, C. M., Wilby, R. L., Haasnoot, M., Mendoza, G. F., Dominique, K. C., et al.: Sustainable water management under future uncertainty with eco-engineering decision scaling, Nature Climate Change, 6, 25–34, 2016.

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

Prosdocimi, I., Kjeldsen, T., and Miller, J.: Detection and attribution of urbanization effect on flood extremes using nonstationary flood-frequency models, Water Resources Research, 51, 4244–4262, 2015.

Prosdocimi, I., Dupont, E., Augustin, N. H., Kjeldsen, T. R., Simpson, D. P., and Smith, T. R.: Areal models for spatially coherent trend detection: the case of British peak river flows, Geophysical Research Letters, 46, 13 054–13 061, 2019.

Prudhomme, C., Wilby, R. L., Crooks, S., Kay, A. L., and Reynard, N. S.: Scenario-neutral approach to climate change impact studies: application to flood risk, Journal of Hydrology, 390, 198–209, 2010.

Pryor, S., Conrick, R., Miller, C., Tytell, J., and Barthelmie, R.: Intense and extreme wind speeds observed by anemometer and seismic networks: An eastern US case study, Journal of Applied Meteorology and Climatology, 53, 2417–2429, 2014.

Raymond, C., Horton, R. M., Zscheischler, J., Martius, O., AghaKouchak, A., Balch, J., Bowen, S. G., Camargo, S. J., Hess, J., Kornhuber, K., et al.: Understanding and managing connected extreme events, Nature climate change, pp. 1–11, 2020a.

Raymond, C., Matthews, T., and Horton, R. M.: The emergence of heat and humidity too severe for human tolerance, Science Advances, 6, eaaw1838, 2020b.

Read, L. K. and Vogel, R. M.: Reliability, return periods, and risk under nonstationarity, Water Resources Research, 51, 6381–6398, 2015.

Reggiani, P., Renner, M., Weerts, A., and Van Gelder, P.: Uncertainty assessment via Bayesian revision of ensemble streamflow predictions in the operational river Rhine forecasting system, Water Resources Research, 45, 2009.

Ren, H.-L., Wu, Y., Bao, Q., Ma, J., Liu, C., Wan, J., Li, Q., Wu, X., Liu, Y., Tian, B., et al.: The China multi-model ensemble prediction system and its application to flood-season prediction in 2018, Journal of Meteorological Research, 33, 540–552, 2019.

Requena, A. I., Burn, D. H., and Coulibaly, P.: Estimates of gridded relative changes in 24-h extreme rainfall intensities based on pooled frequency analysis, Journal of Hydrology, 577, 123 940, 2019.

Ribeiro, S., Caineta, J., and Costa, A. C.: Review and discussion of homogenisation methods for climate data, Physics and Chemistry of the Earth, Parts A/B/C, 94, 167–179, 2016.

Rivington, M., Matthews, K. B., Buchan, K., Miller, D., Bellocchi, G., and Russell, G.: Climate change impacts and adaptation scope for agriculture indicated by agro-meteorological metrics, Agricultural Systems, 114, 15–31, 2013.

Rootzén, H. and Katz, R. W.: Design life level: quantifying risk in a changing climate, Water Resources Research, 49, 5964–5972, 2013.

Ross, G. J., Tasoulis, D. K., and Adams, N. M.: Nonparametric monitoring of data streams for changes in location and scale, Technometrics, 53, 379–389, 2011.

Rust, W., Holman, I., Bloomfield, J., Cuthbert, M., and Corstanje, R.: Understanding the potential of climate teleconnections to project future groundwater drought, 2019.

Ryberg, K. R., Hodgkins, G. A., and Dudley, R. W.: Change points in annual peak streamflows: Method comparisons and historical change points in the United States, Journal of Hydrology, p. 124307, 2019.

Saeed, F., Hagemann, S., and Jacob, D.: Impact of irrigation on the South Asian summer monsoon, Geophysical Research Letters, 36, 2009.

Salas, J., Obeysekera, J., and Vogel, R.: Techniques for assessing water infrastructure for nonstationary extreme events: A review, Hydrological Sciences Journal, 63, 325–352, 2018.

Salas, J. D. and Obeysekera, J.: Revisiting the concepts of return period and risk for nonstationary hydrologic extreme events, Journal of Hydrologic Engineering, 19, 554–568, 2014.

Salas, J. D., Anderson, M. L., Papalexiou, S. M., and Frances, F.: PMP and Climate Variability and Change: A Review, Journal of Hydrologic Engineering, 25, 03120 002, 2020.

Schaller, N., Sillmann, J., Anstey, J., Fischer, E. M., Grams, C. M., and Russo, S.: Influence of blocking on Northern European and Western Russian heatwaves in large climate model ensembles, Environmental Research Letters, 13, 054 015, 2018.

Schlef, K. E., Moradkhani, H., and Lall, U.: Atmospheric circulation patterns associated with extreme United States floods identified via machine learning, Scientific reports, 9, 1–12, 2019.

Schott, F. A., Xie, S.-P., and McCreary Jr, J. P.: Indian Ocean circulation and climate variability, Reviews of Geophysics, 47, 2009.

Scott, A. J. and Knott, M.: A cluster analysis method for grouping means in the analysis of variance, Biometrics, pp. 507–512, 1974.

Sen, P. K.: Estimates of the regression coefficient based on Kendall's tau, Journal of the American statistical association, 63, 1379–1389, 1968.

Seneviratne, S. I., Corti, T., Davin, E. L., Hirschi, M., Jaeger, E. B., Lehner, I., Orlowsky, B., and Teuling, A. J.: Investigating soil moisture–climate interactions in a changing climate: A review, Earth-Science Reviews, 99, 125–161, 2010.

Serinaldi, F. and Kilsby, C. G.: Stationarity is undead: Uncertainty dominates the distribution of extremes, Advances in Water Resources, 77, 17–36, 2015.

Serinaldi, F., Kilsby, C. G., and Lombardo, F.: Untenable nonstationarity: An assessment of the fitness for purpose of trend tests in hydrology, Advances in Water Resources, 111, 132–155, 2018.

Sharma, A., Wasko, C., and Lettenmaier, D. P.: If precipitation extremes are increasing, why aren't floods?, Water Resources Research, 54, 8545–8551, 2018.

Shaw, T., Baldwin, M., Barnes, E. A., Caballero, R., Garfinkel, C., Hwang, Y.-T., Li, C., O'Gorman, P., Rivière, G., Simpson, I., et al.: Storm track processes and the opposing influences of climate change, Nature Geoscience, 9, 656–664, 2016.

Shepherd, T. G., Boyd, E., Calel, R. A., Chapman, S. C., Dessai, S., Dima-West, I. M., Fowler, H. J., James, R., Maraun, D., Martius, O., et al.: Storylines: an alternative approach to representing uncertainty in physical aspects of climate change, Climatic change, 151, 555–571, 2018.

Sherwood, S. C. and Huber, M.: An adaptability limit to climate change due to heat stress, Proceedings of the National Academy of Sciences, 107, 9552–9555, 2010.

Shkolnik, I., Pavlova, T., Efimov, S., and Zhuravlev, S.: Future changes in peak river flows across northern Eurasia as inferred from an ensemble of regional climate projections under the IPCC RCP8. 5 scenario, Climate dynamics, 50, 215–230, 2018.

Shukla, S., Roberts, J., Hoell, A., Funk, C. C., Robertson, F., and Kirtman, B.: Assessing North American multimodel ensemble (NMME) seasonal forecast skill to assist in the early warning of anomalous hydrometeorological events over East Africa, Climate Dynamics, 53, 7411–7427, 2019.

Sillmann, J., Croci-Maspoli, M., Kallache, M., and Katz, R. W.: Extreme cold winter temperatures in Europe under the influence of North Atlantic atmospheric blocking, Journal of Climate, 24, 5899–5913, 2011.

Slater, L. and Villarini, G.: On the impact of gaps on trend detection in extreme streamflow time series, International Journal of Climatology, 37, 3976–3983, 2017a.

Slater, L. J. and Villarini, G.: Recent trends in US flood risk, Geophysical Research Letters, 43, 12–428, 2016.

Slater, L. J. and Villarini, G.: Evaluating the drivers of seasonal streamflow in the US Midwest, Water, 9, 695, 2017b.

Slater, L. J. and Villarini, G.: Enhancing the predictability of seasonal streamflow with a statistical-dynamical approach, Geophysical Research Letters, 45, 6504–6513, 2018.

Slater, L. J., Singer, M. B., and Kirchner, J. W.: Hydrologic versus geomorphic drivers of trends in flood hazard, Geophysical Research Letters, 42, 370–376, 2015.

Smelser, M. and Schmidt, J.: An assessment methodology for determining historical changes in mountain streams. USDA Department of Agriculture Forest Service, Rocky Mountain Research Station, Tech. rep., General Technical report RMRS-GTS-6. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station. 29 p, 1998.

Smith, D. M., Scaife, A. A., Eade, R., Athanasiadis, P., Bellucci, A., Bethke, I., Bilbao, R., Borchert, L., Caron, L.-P., Counillon, F., et al.: North Atlantic climate far more predictable than models imply, Nature, 583, 796–800, 2020.

Smith, K., Hannaford, J., Bloomfield, J., McCarthy, M., Parry, S., Barker, L., Svensson, C., Tanguy, M., Marchant, B., McKenzie, A., et al.: The hydroclimatology of UK droughts: evidence from newly recovered and reconstructed datasets from the late 19 th century to present, AGUFM, 2017, H11O–02, 2017.

Sornette, D. and Ouillon, G.: Dragon-kings: mechanisms, statistical methods and empirical evidence, The European Physical Journal Special Topics, 205, 1–26, 2012.

Soulsby, C., Dick, J., Scheliga, B., and Tetzlaff, D.: Taming the flood-How far can we go with trees?, Hydrological Processes, 31, 3122–3126, 2017.

Spinoni, J., Naumann, G., and Vogt, J. V.: Pan-European seasonal trends and recent changes of drought frequency and severity, Global and Planetary Change, 148, 113–130, 2017.

Steinschneider, S., Yang, Y.-C. E., and Brown, C.: Panel regression techniques for identifying impacts of anthropogenic landscape change on hydrologic response, Water Resources Research, 49, 7874–7886, 2013.

Sterl, A.: On the (in) homogeneity of reanalysis products, Journal of Climate, 17, 3866–3873, 2004.

Stevenson, S., Overpeck, J. T., Fasullo, J., Coats, S., Parsons, L., Otto-Bliesner, B., Ault, T., Loope, G., and Cole, J.: Climate variability, volcanic forcing, and last millennium hydroclimate extremes, Journal of Climate, 31, 4309–4327, 2018.

Stocker, T. F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, V., Midgley, P. M., et al.: Climate change 2013: The physical science basis, Contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change, 1535, 2013.

Stott, P. A., Christidis, N., Otto, F. E., Sun, Y., Vanderlinden, J.-P., van Oldenborgh, G. J., Vautard, R., von Storch, H., Walton, P., Yiou, P., et al.: Attribution of extreme weather and climate-related events, Wiley Interdisciplinary Reviews: Climate Change, 7, 23–41, 2016.

Strazzo, S., Collins, D. C., Schepen, A., Wang, Q., Becker, E., and Jia, L.: Application of a hybrid statistical–dynamical system to seasonal prediction of North American temperature and precipitation, Monthly Weather Review, 147, 607–625, 2019.

Sun, Q., Zhang, X., Zwiers, F., Westra, S., and Alexander, L. V.: A global, continental and regional analysis of changes in extreme precipitation, Journal of Climate, pp. 1–52, 2020a.

Sun, X., Li, Z., and Tian, Q.: Assessment of hydrological drought based on nonstationary runoff data, Hydrology Research, 2020b.

Sutcliffe, J. and Parks, Y.: The hydrology of the Nile IAHS Special Publication. no. 5 IAHS Press, Institute of hydrology, Wallingford, Oxfordshire, 1999.

Sutton, R. T. and Dong, B.: Atlantic Ocean influence on a shift in European climate in the 1990s, Nature Geoscience, 5, 788–792, 2012.

Tallaksen, L. M. and Van Lanen, H. A.: Hydrological drought: processes and estimation methods for streamflow and groundwater, vol. 48, Elsevier, 2004.

Tan, X. and Shao, D.: Precipitation trends and teleconnections identified using quantile regressions over Xinjiang, China, International Journal of Climatology, 37, 1510–1525, 2017.

Taylor, S. J. and Letham, B.: Forecasting at Scale, The American Statistician, 72, 37–45, https://doi.org/10.1080/00031305.2017.1380080, https://doi.org/10.1080/00031305.2017.1380080, 2018.

Theil, H.: A rank-invariant method of linear and polynomial regression analysis, in: Henri Theil's contributions to economics and econometrics, pp. 345–381, Springer, 1992.

Thirumalai, K., DiNezio, P. N., Okumura, Y., and Deser, C.: Extreme temperatures in Southeast Asia caused by El Niño and worsened by global warming, Nature communications, 8, 1–8, 2017.

Thomas, A. C., Reager, J. T., Famiglietti, J. S., and Rodell, M.: A GRACE-based water storage deficit approach for hydrological drought characterization, Geophysical Research Letters, 41, 1537–1545, 2014.

Thompson, V., Dunstone, N. J., Scaife, A. A., Smith, D. M., Slingo, J. M., Brown, S., and Belcher, S. E.: High risk of unprecedented UK rainfall in the current climate, Nature communications, 8, 1–6, 2017.

Thorne, P., Allan, R. J., Ashcroft, L., Brohan, P., Dunn, R. H., Menne, M., Pearce, P., Picas, J., Willett, K., Benoy, M., et al.: Toward an integrated set of surface meteorological observations for climate science and applications, Bulletin of the American Meteorological Society, 98, 2689–2702, 2017.

Torralba, V., Doblas-Reyes, F. J., and Gonzalez-Reviriego, N.: Uncertainty in recent near-surface wind speed trends: a global reanalysis intercomparison, Environmental Research Letters, 12, 114 019, 2017a.

Torralba, V., Doblas-Reyes, F. J., MacLeod, D., Christel, I., and Davis, M.: Seasonal climate prediction: A new source of information for the management of wind energy resources, Journal of Applied Meteorology and Climatology, 56, 1231–1247, 2017b.

Trenberth, K. E., Dai, A., Rasmussen, R. M., and Parsons, D. B.: The changing character of precipitation, Bulletin of the American Meteorological Society, 84, 1205–1218, 2003.

Uhlemann, S., Thieken, A., and Merz, B.: A consistent set of trans-basin floods in Germany between 1952-2002, Hydrology and Earth System Sciences, 14, 1277–1295, 2010.

United States Army Corps of Engineers: Incorporating sea level change in civil works ER 1110-2-8162. Sea Level Change Calculator, https://www.usace.army.mil/corpsclimate/Climate_Preparedness_and_Resilience/App_Flood_Risk_Reduct_Sandy_Rebuild/SL_change_curve_calc/, 2019.

Van den Brink, H., Können, G., Opsteegh, J., Van Oldenborgh, G., and Burgers, G.: Estimating return periods of extreme events from ECMWF seasonal forecast ensembles, International Journal of Climatology: A Journal of the Royal Meteorological Society, 25, 1345–1354, 2005.

Van Loon, A. F.: Hydrological drought explained, Wiley Interdisciplinary Reviews: Water, 2, 359–392, 2015.

Vecchi, G. A., Zhao, M., Wang, H., Villarini, G., Rosati, A., Kumar, A., Held, I. M., and Gudgel, R.: Statistical–dynamical predictions of seasonal North Atlantic hurricane activity, Monthly Weather Review, 139, 1070–1082, 2011.

Vicente-Serrano, S. M., Beguería, S., and López-Moreno, J. I.: A multiscalar drought index sensitive to global warming: the standardized precipitation evapotranspiration index, Journal of climate, 23, 1696–1718, 2010.

Vicente-Serrano, S. M., Peña-Gallardo, M., Hannaford, J., Murphy, C., Lorenzo-Lacruz, J., Dominguez-Castro, F., López-Moreno, J. I., Beguería, S., Noguera, I., Harrigan, S., et al.: Climate, irrigation, and land cover change explain streamflow trends in countries bordering the northeast Atlantic, Geophysical Research Letters, 46, 10 821–10 833, 2019.

Villarini, G.: On the seasonality of flooding across the continental United States, Advances in Water Resources, 87, 80–91, 2016.

Villarini, G. and Serinaldi, F.: Development of statistical models for at-site probabilistic seasonal rainfall forecast, International Journal of Climatology, 32, 2197–2212, 2012.

Villarini, G. and Slater, L.: Climatology of flooding in the United States, 2017.

Villarini, G. and Slater, L. J.: Examination of changes in annual maximum gauge height in the continental United States using quantile regression, Journal of Hydrologic Engineering, 23, 06017 010, 2018.

Villarini, G. and Zhang, W.: Projected changes in flooding: a continental US perspective, Annals of the New York Academy of Sciences, 2020.

Villarini, G., Serinaldi, F., Smith, J. A., and Krajewski, W. F.: On the stationarity of annual flood peaks in the continental United States during the 20th century, Water Resources Research, 45, 2009a.

Villarini, G., Smith, J. A., Serinaldi, F., Bales, J., Bates, P. D., and Krajewski, W. F.: Flood frequency analysis for nonstationary annual peak records in an urban drainage basin, Advances in Water Resources, 32, 1255–1266, 2009b.

https://doi.org/10.5194/hess-2020-576
Preprint. Discussion started: 14 November 2020

Hydrology and Earth System Sciences
Discussions

Villarini, G., Taylor, S., Wobus, C., Vogel, R., Hecht, J., White, K., Baker, B., Gilroy, K., Olsen, R. J., and Raff, D.: Floods and nonstationarity - A review, USACE Climate preparedness and resilience, CWTS 2018-01, 2018.

Vitart, F. and Robertson, A. W.: The sub-seasonal to seasonal prediction project (S2S) and the prediction of extreme events, npj Climate and Atmospheric Science, 1, 1–7, 2018.

Vogel, R., Rosner, A., and Kirshen, P.: Brief Communication: Likelihood of societal preparedness for global change: trend detection., Natural Hazards & Earth System Sciences, 13, 2013.

Volpi, E., Fiori, A., Grimaldi, S., Lombardo, F., and Koutsoyiannis, D.: One hundred years of return period: Strengths and limitations, Water Resources Research, 51, 8570–8585, 2015.

Walz, M. A., Befort, D. J., Kirchner-Bossi, N. O., Ulbrich, U., and Leckebusch, G. C.: Modelling serial clustering and inter-annual variability of European winter windstorms based on large-scale drivers, International Journal of Climatology, 38, 3044–3057, https://doi.org/10.1002/joc.5481, 2018.

Wang, Q., Schepen, A., and Robertson, D. E.: Merging seasonal rainfall forecasts from multiple statistical models through Bayesian model averaging, Journal of Climate, 25, 5524–5537, 2012.

Wang, S. S., Zhao, L., Yoon, J.-H., Klotzbach, P., and Gillies, R. R.: Quantitative attribution of climate effects on Hurricane Harvey's extreme rainfall in Texas, Environmental Research Letters, 13, 054 014, 2018.

Wang, X. L., Zwiers, F. W., Swail, V. R., and Feng, Y.: Trends and variability of storminess in the Northeast Atlantic region, 1874–2007, Climate Dynamics, 33, 1179, 2009.

Ward, K., Lauf, S., Kleinschmit, B., and Endlicher, W.: Heat waves and urban heat islands in Europe: A review of relevant drivers, Science of the Total Environment, 569, 527–539, 2016.

Wasko, C. and Nathan, R.: Influence of changes in rainfall and soil moisture on trends in flooding, Journal of Hydrology, 575, 432–441, 2019.

Wasko, C. and Sharma, A.: Quantile regression for investigating scaling of extreme precipitation with temperature, Water Resources Research, 50, 3608–3614, 2014.

Wasko, C. and Sharma, A.: Steeper temporal distribution of rain intensity at higher temperatures within Australian storms, Nature Geoscience, 8, 527–529, 2015.

Wasko, C. and Sharma, A.: Global assessment of flood and storm extremes with increased temperatures, Scientific reports, 7, 1–8, 2017.

Wasko, C., Sharma, A., and Lettenmaier, D. P.: Increases in temperature do not translate to increased flooding, Nature Communications, 10, 1–3, 2019.

Weber, H. and Wunderle, S.: Drifting Effects of NOAA Satellites on Long-Term Active Fire Records of Europe, Remote sensing, 11, 467, 2019.

Wenzel, S., Cox, P. M., Eyring, V., and Friedlingstein, P.: Emergent constraints on climate-carbon cycle feedbacks in the CMIP5 Earth system models, Journal of Geophysical Research: Biogeosciences, 119, 794–807, 2014.

Wester, P., Mishra, A., Mukherji, A., and Shrestha, A. B.: The Hindu Kush Himalaya assessment: mountains, climate change, sustainability and people, Springer Nature, 2019.

Westra, S., Alexander, L. V., and Zwiers, F. W.: Global increasing trends in annual maximum daily precipitation, Journal of Climate, 26, 3904–3918, https://doi.org/https://doi.org/10.1175/JCLI-D-12-00502.1, 2013.

Whan, K., Sillmann, J., Schaller, N., and Haarsma, R.: Future changes in atmospheric rivers and extreme precipitation in Norway, Climate Dynamics, 54, 2071–2084, 2020.

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

Wi, S., Valdés, J. B., Steinschneider, S., and Kim, T.-W.: Non-stationary frequency analysis of extreme precipitation in South Korea using peaks-over-threshold and annual maxima, Stochastic environmental research and risk assessment, 30, 583–606, 2016.

Wilby, R.: When and where might climate change be detectable in UK river flows?, Geophysical Research Letters, 33, 2006.

Wilby, R. and Murphy, C.: Decision-making by water managers despite climate uncertainty, in: The Oxford Handbook of Planning for Climate Change Hazards, 2019.

Wilby, R. L. and Quinn, N. W.: Reconstructing multi-decadal variations in fluvial flood risk using atmospheric circulation patterns, Journal of Hydrology, 487, 109–121, 2013.

Wilby, R. L., Clifford, N. J., De Luca, P., Harrigan, S., Hillier, J. K., Hodgkins, R., Johnson, M. F., Matthews, T. K., Murphy, C., Noone, S. J., et al.: The 'dirty dozen'of freshwater science: detecting then reconciling hydrological data biases and errors, Wiley Interdisciplinary Reviews: Water, 4, e1209, 2017.

Wilcock, D. and Wilcock, F.: Modelling the hydrological impacts of channelization on streamflow characteristics in a Northern Ireland catchment, IAHS Publications-Series of Proceedings and Reports-Intern Assoc Hydrological Sciences, 231, 41–48, 1995.

Wild, S., Befort, D. J., and Leckebusch, G. C.: Was the extreme storm season in winter 2013/14 over the North Atlantic and the United Kingdom triggered by changes in the West Pacific warm pool?, Bulletin of the American Meteorological Society, 96, S29–S34, 2015.

Wilhite, D. A.: Droughts: a global assesment, Routledge, 2016.

Wine, M. L.: Under non-stationarity securitization contributes to uncertainty and Tragedy of the Commons, Journal of Hydrology, 568, 716–721, 2019.

WMO: Handbook of Drought Indicators and Indices, Geneva: World Meteorological Organization (WMO) and Global Water Partnership (GWP), 2016.

Wohland, J., Omrani, N.-E., Witthaut, D., and Keenlyside, N. S.: Inconsistent wind speed trends in current twentieth century reanalyses, Journal of Geophysical Research: Atmospheres, 124, 1931–1940, 2019.

Woo, G.: Downward Counterfactual Search for Extreme Events, Frontiers in Earth Science, 7, 340, 2019.

Woollings, T. and Blackburn, M.: The North Atlantic jet stream under climate change and its relation to the NAO and EA patterns, Journal of Climate, 25, 886–902, 2012.

Wu, C., Yeh, P. J.-F., Chen, Y.-Y., Hu, B. X., and Huang, G.: Future Precipitation-Driven Meteorological Drought Changes in the CMIP5 Multimodel Ensembles under 1.5° C and 2° C Global Warming, Journal of Hydrometeorology, 21, 2177–2196, 2020a.

Wu, J., Han, Z., Xu, Y., Zhou, B., and Gao, X.: Changes in Extreme Climate Events in China Under 1.5° C–4° C Global Warming Targets: Projections Using an Ensemble of Regional Climate Model Simulations, Journal of Geophysical Research: Atmospheres, 125, e2019JD031 057, 2020b.

Xu, S., Wu, C., Wang, L., Gonsamo, A., Shen, Y., and Niu, Z.: A new satellite-based monthly precipitation downscaling algorithm with non-stationary relationship between precipitation and land surface characteristics, Remote sensing of environment, 162, 119–140, 2015.

Xu, Z., FitzGerald, G., Guo, Y., Jalaludin, B., and Tong, S.: Impact of heatwave on mortality under different heatwave definitions: a systematic review and meta-analysis, Environment international, 89, 193–203, 2016.

Yan, L., Xiong, L., Guo, S., Xu, C.-Y., Xia, J., and Du, T.: Comparison of four nonstationary hydrologic design methods for changing environment, Journal of Hydrology, 551, 132–150, 2017.

Yates, D. N., Miller, K. A., Wilby, R. L., and Kaatz, L.: Decision-centric adaptation appraisal for water management across Colorado's continental divide, Climate Risk Management, 10, 35–50, 2015.

Yosef, Y., Aguilar, E., and Alpert, P.: Changes in extreme temperature and precipitation indices: Using an innovative daily homogenized database in Israel, International Journal of Climatology, 39, 5022–5045, 2019.

Young, I. R. and Ribal, A.: Multiplatform evaluation of global trends in wind speed and wave height, Science, 364, 548–552, https://doi.org/https://doi.org/10.1126/science.aav9527, 2019.

Yuan, X., Zhang, M., Wang, L., and Zhou, T.: Understanding and seasonal forecasting of hydrological drought in the Anthropocene, Hydrology and Earth System Sciences, 21, 5477, 2017.

Yuan, X., Wang, L., Wu, P., Ji, P., Sheffield, J., and Zhang, M.: Anthropogenic shift towards higher risk of flash drought over China, Nature communications, 10, 1–8, 2019.

Yue, S., Ouarda, T. B., Bobée, B., Legendre, P., and Bruneau, P.: Approach for describing statistical properties of flood hydrograph, Journal of hydrologic engineering, 7, 147–153, 2002a.

Yue, S., Pilon, P., and Cavadias, G.: Power of the Mann–Kendall and Spearman's rho tests for detecting monotonic trends in hydrological series, Journal of hydrology, 259, 254–271, 2002b.

Zhan, W., He, X., Sheffield, J., and Wood, E. F.: Projected seasonal changes in large-scale global precipitation and temperature extremes based on the CMIP5 ensemble, Journal of Climate, 33, 5651–5671, 2020.

Zhang, W., Villarini, G., Vecchi, G. A., and Smith, J. A.: Urbanization exacerbated the rainfall and flooding caused by hurricane Harvey in Houston, Nature, 563, 384–388, 2018.

Zhang, X., Hegerl, G., Zwiers, F. W., and Kenyon, J.: Avoiding inhomogeneity in percentile-based indices of temperature extremes, Journal of Climate, 18, 1641–1651, 2005.

Zhang, X., Alexander, L., Hegerl, G. C., Jones, P., Tank, A. K., Peterson, T. C., Trewin, B., and Zwiers, F. W.: Indices for monitoring changes in extremes based on daily temperature and precipitation data, Wiley Interdisciplinary Reviews: Climate Change, 2, 851–870, 2011.

Zhang, Y. and Fueglistaler, S.: How tropical convection couples high moist static energy over land and ocean, Geophysical Research Letters, 47, e2019GL086 387, 2020.

Zhao, T., Bennett, J. C., Wang, Q., Schepen, A., Wood, A. W., Robertson, D. E., and Ramos, M.-H.: How suitable is quantile mapping for postprocessing GCM precipitation forecasts?, Journal of Climate, 30, 3185–3196, 2017.

Ziegler, A. D., Maurer, E. P., Sheffield, J., Nijssen, B., Wood, E. F., and Lettenmaier, D. P.: Detection time for plausible changes in annual precipitation, evapotranspiration, and streamflow in three Mississippi River sub-basins, Climatic Change, 72, 17–36, 2005.

Zscheischler, J., Martius, O., Westra, S., Bevacqua, E., Raymond, C., Horton, R. M., van den Hurk, B., AghaKouchak, A., Jézéquel, A., Mahecha, M. D., et al.: A typology of compound weather and climate events, Nature reviews earth & environment, pp. 1–15, 2020.

Zulkafli, Z., Perez, K., Vitolo, C., Buytaert, W., Karpouzoglou, T., Dewulf, A., De Bievre, B., Clark, J., Hannah, D. M., and Shaheed, S.: User-driven design of decision support systems for polycentric environmental resources management, Environmental Modelling & Software, 88, 58–73, 2017.