HYDROLOGICAL SCIENCES JOURNAL, 2018
https://doi.org/10.1080/02626667.2018.1426858

Techniques for assessing water infrastructure for nonstationary extreme events: a review

J. D. Salas[a], J. Obeysekera[b] and R.M. Vogel[c]

[a]Department of Civil and Environmental Engineering, Colorado State University, Fort Collins, Colorado, USA; [b]Hydrology & Hydraulics Bureau, South Florida Water Management District, West Palm Beach, Florida, USA; [c]Department of Civil and Environmental Engineering, Tufts University, Medford, Massachusetts, USA

ABSTRACT
Statistical and physically-based methods have been used for designing and assessing water infrastructure such as spillways and stormwater drainage systems. Traditional approaches assume that hydrological processes evolve in an environment where the hydrological cycle is stationary over time. However, in recent years, it has become increasingly evident that in many areas of the world the foregoing assumption may no longer apply, due to the effect of anthropogenic and climatic induced stressors that cause nonstationary conditions. This has attracted the attention of national and international agencies, research institutions, academia, and practicing water specialists, which has led to developing new techniques that may be useful in those cases where there is good evidence and attribution of nonstationarity. We review the various techniques proposed in the field and point out some of the challenges ahead in future developments and applications. Our review emphasizes hydrological design to protect against extreme events such as floods and low flows.

ARTICLE HISTORY
Received 13 April 2017
Accepted 10 December 2017

EDITOR
R. Woods

ASSOCIATE EDITOR
A. Viglione

KEYWORDS
nonstationarity; hydrological variability and change; hydrological extremes; anthropogenic and climate change; risk and return period; hazard functions

1 Introduction

For decades, a wide range of statistical techniques have been developed for designing hydraulic structures for protecting society and the environment against the impact of extreme events such as extreme precipitation, floods, and rising sea levels. Commonly they are based on probability distribution functions (pdf) such as the geometric, binomial, general extreme value (GEV), and log-Pearson Type III (LP3), and in some cases time series models. For example, the waiting time for the occurrence of a flood exceeding a design flood is geometric distributed and its expected value, often denoted as "return period" (expected waiting time), has been a commonly used concept in engineering practice. Likewise, the number of extreme floods exceeding the design flood magnitude of a hydraulic structure during its design life (e.g. *n* years), as well as its "hydrological reliability and risk," have been determined based on the binomial distribution. In addition, several other approaches and techniques have been developed based on risk-based design of hydraulic structures that involve additional factors such as cost of damages, vulnerability, and resilience. Furthermore, stochastic models have been used to simulate representative sets of sequences of hydrological variables (e.g. streamflow) that may occur during the planning horizon of a given structure (e.g. the risk of occurrence of 7-day, 10-year low flows during a 25-year period) and physically-based watershed models have been utilized for representing the underlying hydrological cycle and applied for assessing the performance of hydraulic systems.

The models and techniques mentioned above have generally been developed based on the assumption that hydrological events, such as annual maximum rainfall and associated flood discharges, arise from a stationary hydrological regime (i.e. the marginal distribution remains invariant with time). However, it has been reported in the literature that, because of some natural and anthropogenic factors, the hydrological cycle in many areas worldwide has been changing, causing significant changes in the temporal and spatial behavior of hydrological processes such as precipitation and streamflow and the ensuing occurrences of extreme events. Thus, a growing concern of water resources specialists has been the extent to which the traditional methods developed for stationary regimes may still be applicable, or whether newer methods are needed where there is sufficient evidence of nonstationarity (i.e. the marginal distribution changes with time). This situation has sparked worldwide interest in the

CONTACT J. D. Salas jose.salas@colostate.edu

water resources community, including planners and managers of projects, government agencies, consulting firms, research institutions, and academia, in finding new ways of taking into account the changing conditions of hydrological regimes for planning and evaluating water infrastructure.

Recently, a variety of newer concepts and techniques have emerged extending those based on stationary conditions so that they will be applicable for nonstationary hydrological conditions. The focus of this paper is to examine the recent literature on the subject, such as the application of nonstationary pdfs of the underlying hydrological variable (e.g. extreme annual floods) where the model parameters may be functions of covariates that evolve through time, along with model fitting, selection, and diagnostic techniques. Likewise, we review the development of nonstationary geometric and Poisson-binomial distributions that enable the determination of the distribution of the waiting time of the first flood exceeding the design flood and the distribution of the number of floods that exceed the design flood during the design life of the structure, respectively, and use of hazard function analysis, magnification factors, and related properties such as return period, hydrological reliability, and risk of failure, all applicable under nonstationary conditions. In addition, simple regression methods are reviewed for modeling changes in the mean, variance, and skewness, and combining such nonstationary moments with various pdfs to "update" design events, given historical data. We also discuss estimation of the Type I and II error probabilities and how they can help us to make statements about posterior probabilities needed to account for over- or underdesign during the design process. Furthermore, we review alternative metrics that may be used for assessing infrastructure investment decisions, including economic risk-based approaches, and planning under the additional uncertainty imposed by nonstationary conditions. The paper ends with a section summarizing some of the challenges involved in applying the new techniques reviewed here. The reviewed concepts and techniques that consider nonstationarity are fairly new, and time will be needed for understanding them well, gaining practical experience with them, and comparing the various alternatives that have been proposed. The current paper is aimed at assisting in this endeavor.

2 Brief review of traditional stationary techniques

The traditional techniques for designing hydraulic structures for defense against extreme events such as floods are generally based on return period and risk. They assume that extreme events arise from a stationary distribution and the occurrences of extreme events are independent and identically distributed (i.i.d.). For example, consider the case of extreme annual floods and the notation employed by Salas and Obeysekera (2014). Let us assume that the annual maximum floods (AMF) are represented by the random variable Z with cumulative distribution function (cdf) $F_Z(z, \underline{\theta})$ where $\underline{\theta}$ represents the parameter set. The typical cdfs utilized for this purpose are the GEV, the two- and three-parameter lognormal (LN2 and LN3, respectively), and LP3 (e.g. Stedinger *et al.* 1993), although other alternatives such as the generalized logistic, generalized Pareto (GP), generalized gamma, and Burr Type XII have been suggested (e.g. Koutsoyiannis 2004, Hao and Singh 2009, Papalexiou and Koutsoyiannis 2013a, 2013b, Serinaldi and Kilsby 2014). Another example is intensity–duration–frequency (IDF) curves, which are widely applied for designing structures, such as municipal storm-water drainage systems. Generally, IDF curves have been developed assuming stationarity pdfs of the precipitation process for a given duration. The preference and suitability of a given distribution depend on the way in which data are processed for analysis (e.g. annual maxima, peaks over threshold (POT), or annual minima) and the type of variable (e.g. extreme floods or precipitation). Often, we rely on guidelines from published research results and specifications provided in standards for specific regions and countries.

Consider a hydraulic structure designed using a design flood z_q, which is the flood quantile with non-exceedence probability q. Alternatively, the notation z_p and z_T are used in which $p = 1 - q$, i.e. the flood quantile with exceedence probability p, and $T = 1/p = 1/(1 - q)$ (years) is called the return period. Note it is assumed that the probability of exceeding the design flood z_q each year or the risk that a flood z may exceed z_q remains the same and is equal to p (i.e. the pdf of floods is identically distributed and stationary). The rationale behind the referred return period T is based on the waiting time x (years) in which a flood exceeding the design flood z_q will occur for the first time. That first time could be in year 1, 2, …, x, … or perhaps it will never occur. Thus the probability that a flood event (exceeding the design flood z_q) will occur for the first time in year x is $(1 - p)^{x-1} p$, or:

$$f(x) = \mathrm{P}(X = x) = (1 - p)^{x-1} p\ , \quad x = 1, 2, \ldots, \infty \tag{1}$$

which is the geometric distribution (e.g. Mood *et al.* 1974). The discrete pdf in Equation (1) (also known as a probability mass function, pmf) has an exponential decay with mean E(X) = 1/p and variance var(X) = q/p^2. Note that the expected waiting time (EWT, in years) required for the first occurrence of a flood exceeding z_q is 1/p, which has been referred to as the return period T in engineering practice (e.g. Chow *et al.* 1988, Bras 1990).

The failure of a hydraulic structure designed for a project life of n years will occur whenever the first flood exceeding z_q occurs before or during year n, i.e. the probability of that event is $\mathrm{P}(X \le n) = F_X(n)$ where $F_X(x)$ is the cdf of the geometric distribution. Therefore, the risk of failure over the n-year period is (e.g. Bras 1990):

$$R = F_X(n) = \sum_{x=1}^{n} f(x) = 1 - (1-p)^n \tag{2}$$

and the corresponding n-year reliability R_ℓ of the structure is $R_\ell = 1 - R = (1-p)^n$. Furthermore, consider the random variable Y denoting the number of floods exceeding the design flood z_q during the n-year planning period. The probabilities $\mathrm{P}(Y = y)$, $y = 0, \ldots, n$ can be obtained using the binomial probability distribution (e.g. Mood *et al.* 1974). Thus, the reliability is the probability that no floods exceeding z_q will occur in the n-year period, i.e. $\mathrm{P}(Y = 0)$, while the risk of failure over the n-year period is the probability $\mathrm{P}(Y > 0)$, i.e. $R = \mathrm{P}(Y > 0) = 1 - \mathrm{P}(Y = 0) = 1 - (1-p)^n$ as in Equation (2). In addition, it may be shown that the first two central moments of Y are E(Y) = $n\ p$ and var(Y) = $n\ p\ q$.

The foregoing concepts, definitions, and equations, related to return period, risk of failure, and reliability, have been commonly utilized in engineering practice and are widely available in books and manuals (e.g. IACWD 1982, Chow *et al.* 1988, Bras 1990, Stedinger *et al.* 1993, Viessman and Lewis 2003, CEH 2016, HFAWG 2017, Stedinger 2017, Vogel and Castellarin 2017). Depending on the type of hydraulic structure, the value of T (or the value of p) is selected from design manuals, the corresponding design flood z_q, in which $q = 1 - 1/T$ (or $q = 1 - p$), is determined from frequency analysis of the underlying flood data, i.e. by inverting the fitted distribution $F_Z(z_q, \hat{\underline{\theta}}) = q$, in which $\hat{\underline{\theta}}$ is the estimated parameter set, and the risk of failure R is obtained from Equation (2) for a specified value of n. This procedure is illustrated in Table 1 (row corresponding to EWT). Conversely, if the value of the risk of failure R is set, e.g. R = 5%, then the exceedence probability p (and return period T) is obtained from Equation (2) given the value of n, i.e. $p = 1-(1-R)^{1/n}$ and the design flood z_q is obtained by inverting the fitted model $F_Z(z_q, \hat{\underline{\theta}}) = q = 1 - p = (1 - R)^{1/n}$. This simple analysis can be extended where a more detailed examination is made that involves the associated cost of the structure, the expected damages, and the financing of the project, i.e. a risk-based (or reliability-based) approach to the design problem. In such a risk-based approach (see Section 5.5) the level of risk and/or the return period T are chosen in such a way as to maximize the expected net project benefits. Some of the foregoing concepts, definitions, and equations considering stationary processes are summarized in Table 1. In addition, the risk analysis described above may consider the effect of uncertainty (Salas *et al.* 2013). Furthermore, in engineering practice, similar approaches as above and extensions thereof, and other alternatives depending on the data available at the basin and the problem at hand, are applied for developing IDF curves (Chow *et al.* 1988), low-flow frequency analysis (Stedinger *et al.* 1993), regional analysis of extreme precipitation (e.g. Schaefer 1990), and regional flood frequency studies (e.g. Cunnane 1988).

Stochastic models have also been used for simulating possible hydrological sequences that may occur in the future, which in conjunction with predicted water demands over the planning horizon of interest can

Table 1. Alternative design methods based on return period, design quantile (return level), and risk of failure under stationarity. EWT: expected waiting time; ENE: expected number of events; DLL: design life level.

Design method	Primary parameters	Return period, T	Design quantile, z_q (return level)	Risk of failure, R, over design life, n	Probability distribution
EWT[(a)]	T	T (specified)	Solve for z_q in $F_Z(z_q, \hat{\underline{\theta}}) = 1 - 1/T$	$R = 1 - (1 - 1/T)^n$	Geometric
ENE = 1	n	$np = 1$, $p = 1/n$ $T = n$	Solve for z_q in $F_Z(z_q, \hat{\underline{\theta}}) = 1 - 1/n$	$R = 1 - (1 - 1/n)^n$	Binomial
ENE = m ($m > 1$)	n, m	$np = m$, $p = m/n$ $T = n/m$	Solve for z_q in $F_Z(z_q, \hat{\underline{\theta}}) = 1 - m/n$	$R = 1 - (1 - m/n)^n$	Binomial
DLL	R, n	$p = 1 - (1-R)^{1/n}$ $T = 1/p$	Solve for z_q in $F_Z(z_q, \hat{\underline{\theta}}) = (1-R)^{1/n}$	R (specified)	Geometric or Binomial

[(a)]Note that specifying average waiting time T as the design parameter is equivalent to specifying the exceedence probability p, since $p = 1/T$ (refer to Section 2). Then, the expressions in the columns for design quantile z_q and risk of failure R, may be written as $F_Z(z_q, \hat{\underline{\theta}}) = 1 - p$ and $R = 1 - (1 - p)^n$, respectively.

provide the needed information for estimating the statistical characteristics and pdfs of low flows, droughts of given durations, critical droughts, and water surpluses. In addition, they have been used for forecasting hydrological events for weeks and months in advance. For example, Stedinger and Crainiceanu (2001) employed an autoregressive moving average (ARMA) model after a log-normal transformation for forecasting the increasing flooding at a site of the Mississippi River. Alternative models having various features of interest for hydrological applications have been developed over the years, such as models exhibiting long-term dependence, e.g. the fractional ARMA (FARMA) model (Hosking 1984), hidden state Markov model (Thyer and Kuczera 2000), shifting mean model (Sveinsson *et al.* 2003) and the fractional Gaussian noise model (Mandelbrot 1971, Koutsoyiannis 2002), models for intermittent rainfall and streamflow processes, models with periodic components (e.g. for daily and monthly data), models where the marginal distribution of the underlying variable Z is skewed (e.g. gamma autoregressive), and multivariate models for multisite or multiple variables (e.g. Zucchini and Guttorp 1991, Efstratiadis *et al.* 2014, Sveinsson and Salas 2017). Likewise, a suite of nonparametric techniques has been developed for hydrological time series simulation and forecasting (e.g. Lall and Sharma 1996, Vogel and Shallcross 1996, Sharma *et al.* 1997, Prairie *et al.* 2006, Salas and Lee 2009, Lee *et al.* 2010). A summary of the various models generally used in practice can be found in Hao and Singh (2016), Sveinsson and Salas (2017), and Rajagopalan and Lall (2017). Furthermore, physically-based models of various degrees of sophistication have been developed in the past several decades for modeling and simulating the hydrological cycle of watersheds. A wide range of models are available, such as event-based models, which generally simulate the rainfall–runoff process for a short time period, e.g. hours, depending on the design storm and size of the basin considered. Examples of this type of model are HEC-HMS (Scharffenberg and Fleming 2010) and CASC2D (Ogden and Julien 2002). In addition, continuous time watershed models simulate the rainfall–runoff process for longer time frames, e.g. days, months, and years, and consider additional processes and interactions between the surface and subsurface storages and the stream runoff (e.g. Singh 1995, Singh and Frevert 2002). Depending on the given project, an event or a continuous time watershed model is selected. For example, for assessing the performance of a flood-related structure such as a spillway, the HEC-HMS model is often used. Likewise, for assessing the performance of a reservoir planned for water supply during the low-flow season, a continuous time model such as HSPF (Bicknell *et al.* 1997), SWAT (Srinivasan and Arnold 1994), SHETRAN (Ewen *et al.* 2002), GSSHA (Downer and Ogden 2004), and many others (Singh and Frevert 2002) may be necessary. Physically-based models are well suited for evaluating the effects of varying land use and the effect of infrastructure on the hydrological cycle that may develop in a basin over time, and those of future hydroclimatic scenarios, on the characteristics of extreme events such as floods and low flows. Examples are the model by Westra *et al.* (2014) and the physically-based and stochastic modeling framework (Efstratiadis *et al.* 2015) in which a systematic change of land use due to urbanization is considered and a stochastic generation made of precipitation and temperature. Often, a proper stochastic representation of the deterministic model error component is not included in the analysis, in which case significant systematic biases in design event quantiles can result (Farmer and Vogel 2016). Model errors can be included by adding a stochastic term in the analysis, as suggested by Efstratiadis *et al.* (2015) and Farmer and Vogel (2016) (see also Section 7).

3 New techniques in a changing environment

For several decades hydrologists have been aware of the complexity of hydrological processes and their relationships with climatic and environmental conditions (e.g. Hurst 1957). However, it was not until the last 25 years or so, that interest in the dynamics of such processes has grown exponentially, particularly due to: (a) a better understanding of large-scale atmospheric and oceanic mechanisms, and oscillations such as El Niño Southern Oscillation (ENSO) (Philander 1990), Pacific Decadal Oscillation (PDO) (Mantua and Hare 2002), and Atlantic Multidecadal Oscillation (AMO) (Enfield *et al.* 2001); (b) the global warming that the world has been experiencing, in part due to the effect of increasing emission of greenhouse gases into the atmosphere (IPCC 2013); (c) the increasing population in many regions causing major changes and stresses to the landscape, the environment, and water availability (e.g. Vörösmarty *et al.* 2000); and (d) a better understanding of the effects of such atmospheric–oceanic processes and anthropogenic factors on the hydrological cycle (e.g. IPCC 2013).

This increased understanding, awareness, and interest led to major research activity in the field of hydrology and water resources in order to detect and identify the variability of hydrological processes at various temporal and spatial scales (e.g. Douglas *et al.* 2000, Yue *et al.* 2012, Madsen *et al.* 2014), establish the statistical significance and attribution of various types of changes (e.g. trends and abrupt shifting patterns) that have been observed in some historical records of precipitation, streamflow, and water storages (e.g. Burn and Hag Elnur 2002, Blöschl and Montanari 2010, Merz *et al.* 2012, Prosdocimi *et al.* 2015), and quantify their impacts and consequences on water resources (e.g. Kundzewicz *et al.* 2007, Adams and Peck 2008). Prior to these concerns, hydrological processes had generally been assumed to vary in a manner that the key statistics such as the mean and the variance do not change with time, i.e. a stationary process (loosely speaking). However, for the reasons outlined above, some hydrologists strongly questioned the assumption of stationarity and suggested that "*Stationarity is dead – whither water management*?" (Milly *et al.* 2008) and that alternative methods should be developed based on nonstationary concepts for more realistic design, evaluation, and planning and management of infrastructure. While the referred paper received major attention worldwide and many papers have been published along those lines, many others reacted with opposite positions and opinions, as exemplified by the titles of some of the published articles, such as: "*Stationarity: wanted dead or alive?*" (Lins and Cohn 2011), "*Comment on the announced death of stationarity*" (Matalas 2012), "*Negligent killing of scientific concepts: the stationary case*" (Koutsoyiannis and Montanari 2014), "*Modeling and mitigating natural hazards: stationarity is immortal!*" (Montanari and Koutsoyiannis 2014), and "*Stationarity is undead: uncertainty dominates the distribution of extremes*" (Serinaldi and Kilsby 2015).

As evidenced from recent review articles, there are now myriad newer concepts and methods available for flood frequency analysis (FFA) under nonstationary conditions (e.g. Petrow and Merz 2009, CEH 2013, Hall *et al.* 2014, Bayazit 2015). Despite the tremendous increased attention given to this subject, there is still no consensus or generally agreed upon set of methods for performing FFA under nonstationary conditions. Nevertheless, some government agencies are in the process of updating flood protection design guidelines to account for nonstationarity (e.g. Stedinger and Griffis 2011, CEH 2013, Madsen *et al.* 2014, Prosdocimi *et al.* 2014, HFAWG 2017). Likewise, the concern of nonstationarity has led to many studies for detection of precipitation changes over time, updating IDF curves, and comparisons thereof (e.g. Cheng and AghaKouchak 2014, Yilmaz and Perera 2014, Mondal and Mujumdar 2015, Verdon-Kidd and Kiem 2015). For example, Cheng and AghaKouchak (2014) compared results of IDF curves obtained for both stationary and nonstationary methods for some locations in the USA, and Yilmaz and Perera (2014) and Verdon-Kidd and Kiem (2015) studied IDF curves in Australia using stationary and nonstationary pdfs; their results suggested the need to update the IDF curves given by government guidelines. In addition, IDF studies were undertaken by Mondal and Mujumdar (2015), based on a gridded rainfall dataset in India and POT analysis using nonstationary GP distribution and covariates such as ENSO index and global average air temperature. Furthermore, IDF analysis has been done considering future climate scenarios where hydrological design standards for use under future climate were made using quantile mapping (QM) and equidistant quantile mapping (EQM) methods (Li *et al.* 2010), conditional on output from GCMs (global climate models) for future climate conditions (e.g. Tetra Tech 2015). While this and many other studies have been made using projections of GCMs, one must be aware of the large uncertainties, particularly in relation to extreme precipitation (e.g. Blöschl and Montanari 2010).

Some of the concepts and methods that may be useful for designing hydraulic structures to confront extreme events in nonstationary environments are reviewed in the sections below. But fundamental questions remain concerning whether or not nonstationary methods are needed in practice (e.g. Cohn and Lins 2005, Villarini *et al.* 2009, Matalas 2012, Montanari and Koutsoyiannis 2014, Serinaldi and Kilsby 2015, Silva *et al.* 2016); further, there are still disagreements on the underlying definitions, concepts, and methods, and the question of how to select an appropriate design event given evidence of nonstationarity and future uncertainty (e.g. Stedinger and Griffis 2011, Rootzen and Katz 2013, Obeysekera and Salas 2014, Read and Vogel 2015). For example, the concepts of stationarity and nonstationarity are often interpreted in different ways. If the process under study, denoted as Z_t, has the pdf $f(z;\, \underline{\theta})$, where $\underline{\theta}$ is the parameter set that remains constant through time, and $Z_1, Z_2, \ldots, Z_t$ are independent, i.e. uncorrelated, then the data series arising from it is stationary. However, if the pdf or its parameters varies with time, then Z_t is nonstationary. In addition, if Z_t has the same pdf $f(z;\, \underline{\theta})$ for all t, but Z_t is autocorrelated, say $\mathrm{cov}(Z_t, Z_{t-k}) = \sigma^2\rho(k)$, where σ^2 and $\rho(k)$ are the variance and the lag-k autocorrelation, respectively, i.e. the covariance does not depend on time t, then Z_t is stationary; but if $\mathrm{cov}(Z_t, Z_{t-k})$ is a function of t (besides being a function of k), then Z_t is nonstationary. Thus, depending on which property (e.g. moments such as the first or second moments) are constant through time, the process is denoted as first- or second-order stationary. Furthermore, if the autocorrelation function $\rho(k) \to 0$ as $k \to \infty$, then Z_t is known to have a short memory (short-term dependence), whereas if $\rho(k)$ does not converge as $k \to \infty$, then the process Z_t has long memory, or long-term dependence (e.g. Koutsoyiannis 2016). Importantly, realizations of long memory processes exhibit "changes" such as trends and shifts, so that it may appear that the underlying process is nonstationary when in fact it is not. This is even more confusing when the records are short. In this context, Koutsoyiannis (2011) pointed out that "change" does not necessarily imply "nonstationarity."

The lack of a consensus results in part from the tremendous uncertainty associated with our ability to detect, attribute, and model past trends, and the even greater uncertainty associated with our ability to predict the future evolution of hydrological processes. If an observed flood series is known to increase during some particular historical period, then the magnitude of the design flood (e.g. 100-year flood) will increase over that period, and it may not be appropriate to report a single value. In such situations in which historical trends in flood series are obvious and/or overt due to knowledge of changes in historical land use, climate and/or water infrastructure, it is imperative to provide updated estimates of design floods that reflect "those conditions." Even in situations in which there is physical knowledge of the processes that cause the hydrological nonstationarity, there is always some residual uncertainty associated with any trend detection or modeling approach. Vogel *et al.* (2013) document how one can quantify the various factors that give rise to uncertainty in our ability to detect trends in hydrological series, and Obeysekera and Salas (2014) describe how uncertainty in estimates of design events can be quantified in some cases of nonstationarity. Regardless of the level of uncertainty associated with past or future hydrological trends, due to the increased level of uncertainty associated with nonstationarity, the design of some structures under nonstationary conditions may warrant risk-based approaches (e.g. Stedinger and Crainiceanu 2001, Sivapalan and Samuel 2009, Rosner *et al.* 2014).

4 Modeling of extreme events

There are two common ways of modeling extremes of hydrological variables such as floods, precipitation, and wind. They are based on block maxima (BM) data and peaks over threshold (POT) data. Block maxima are defined as the maximum value of a given block (time period), which is typically one year but could be a season (e.g. wet period). An example of BM data is the maximum annual floods, which are the highest discharges in a calendar or water year. This approach has generally been used in actual practice even though the sample size available for modeling the data may be small because only one extreme event per year is utilized. Often there are cases where the second largest flood (or the third etc.) in a given year may outrank the annual maxima in other years, and yet, in the BM approach, those events are not considered for extreme value modeling.

The POT data use extreme events (e.g. floods) above a selected threshold (base flood in case of floods) within a given year or a season. The selection of the threshold and the POT values are important from both physical and statistical points of view. It is expected that using POT values will improve the estimation of design events because of the larger sample of extremes that can be extracted from the historical data. However, the POT data require higher frequency of the basic time series (e.g. hourly or daily) as opposed to BM data. In some cases, some annual extremes may not even be selected as POT events. The POT modeling approach provides additional flexibility in representing extreme events as compared to the BM approach but at the expense of added complexity (Lang *et al.* 1999). Unlike the BM dataset, which is well defined, the POT data depend on the base level selected. Although some criteria have been developed, standard guidelines for its application are lacking. The POT method is also known as the partial duration series (PDS) method. Both stationary and nonstationary models can be used with BM and POT data. In this section, formulation of models, model selection, and parameter estimation and uncertainty are reviewed for both types of data.

4.1 Models based on BM data

Many probability distribution functions have been used in practice for modeling BM data, such as LN, gamma, LP3, and GEV distributions. The attractive feature of the GEV distribution is the well-developed extreme value theory, which is useful for modeling extremes arising from either stationary or nonstationary regimes (e.g. Coles 2001). It leads to an extremal theorem that describes the asymptotic behavior of BM as a family of distributions, which has become increasingly popular in hydrology. Assume there are n random variables $X_1, X_2, \ldots, X_n$ which occur in a given block, and they are independent and identically distributed (i.i.d.) with marginal distribution denoted as $F_X(x)$. We are interested in the distribution of the maximum, i.e. $M_n = \max(X_1, X_2, \ldots, X_n)$. It can be shown that, regardless of the distribution $F_X(x)$, the distribution of M_n when rescaled with some appropriate constants approximates to a family of Type 1, Type 2, and Type 3 extreme value distributions.

Furthermore, it may be shown that the cdf of the three types of distributions may be expressed as (e.g. Coles 2001):

$$F_Z(z, \underline{\theta}) = \exp\left\{-\left[1+\xi\left(\frac{z-\mu}{\sigma}\right)\right]^{-1/\xi}\right\} \tag{3}$$

which is known as the generalized extreme value (GEV) distribution, in which the parameter set $\underline{\theta} = (\mu, \sigma, \xi)$ comprises the location, scale, and shape parameters, respectively. Note that $[1+(\xi/\sigma)(z-\mu)] > 0$, and if $\xi = 0$, the cdf (Equation (3)) converges to the Type 1 (Gumbel) distribution, which is unbounded, while if $\xi > 0$, the cdf is bounded on the left at $(\mu - \sigma)/\xi$ for the Type 2 (Frechet) distribution, and bounded on the right at $(\mu - \sigma)/\xi$ for the Type 3 (Weibull) distribution. In addition, the pdf of the GEV distribution can be written as:

$$f_Z(z, \underline{\theta}) = \frac{1}{\sigma}\left[1+\xi\left(\frac{z-\mu}{\sigma}\right)\right]^{-\frac{1}{\xi}-1} \exp\left\{-\left[1+\xi\left(\frac{z-\mu}{\sigma}\right)\right]^{-1/\xi}\right\} \tag{4}$$

Note that for modeling minimum values (e.g. low flows) based on the GEV, one may use the duality between the maximum and the minimum (Coles 2001).

Extensions of the stationary distributions used with BM data have been suggested in the literature for modeling nonstationary extremes. The main concept has been that extremes of hydrological variables exhibiting nonstationarity may be independent but not identically distributed. Many applications of pdfs expressed in a nonstationary framework have been made assuming that the pdf belongs to a GEV family (e.g. Coles 2001, Katz *et al.* 2002, Prosdocimi *et al.* 2015, Um *et al.* 2017), although in some cases the LN (e.g. Hecht and Vogel 2017) and LP3 (e.g. Webster and Stedinger 2017) models have been applied. We continue our discussion assuming a simplification useful for modeling, i.e. the type of cdf $F_Z(z, \underline{\theta}_t)$ remains the same over time, but the parameters vary with covariates, which could be time or explanatory variables that vary with time (refer to Section 4.3 on covariates). Further, we use the GEV with parameter set $\underline{\theta}_t = (\mu_t, \sigma_t, \xi_t)$, and the nonstationary cdf is written as (Coles 2001, Katz 2013):

$$F_Z(z, \underline{\theta}_t) = \exp\left\{-\left[1+\xi_t\left(\frac{z-\mu_t}{\sigma_t}\right)\right]^{-1/\xi_t}\right\} \tag{5}$$

where $[1+(\xi_t/\sigma_t)(z-\mu_t)] > 0$ and Equation (5) reduces to the Type 1 Gumbel distribution with varying parameters if $\xi_t = 0$. The location parameter may be a function of a covariate, for example, NIÑO3 (an index representing the El Niño phenomenon), while keeping other parameters fixed. Sometimes both the location parameter μ_t and the scale parameter σ_t are assumed functions of covariates while keeping ξ_t constant (e.g. El Adlouni *et al.* 2007, Ruggiero *et al.* 2010). Rarely, all three parameters are made as functions of covariates because it has been found in the literature that the shape parameter ξ is difficult to estimate with precision even in the stationary case. Therefore, for nonstationary models it is unrealistic to consider the shape parameter as a smooth function of time or a function of a covariate (Coles 2001). Commonly, the location parameter μ is assumed to vary with time, but if the "upper bound" of the annual maxima also increases with time, then in such cases the scale parameter σ_t may have to be modeled as a function of time (e.g. El Adlouni *et al.* 2007, Menéndez and Woodworth 2010). Some examples are $\mu_t = \mu_0 + at$ and $\log \sigma_t = \sigma_0 + bt$ (Ruggiero *et al.* 2010, Katz 2013, Salas and Obeysekera 2014).

4.2 Models based on POT data

Often observational records of BM (e.g. annual) are short, resulting in significant uncertainties in the model parameters estimated from such records. An alternative is to build models based on POT data where multiple extremes above a given threshold may be available. Using a larger sample of extremes is appealing since it may reduce the uncertainty of model parameters and design quantiles. However, the POT approach is less common and more complex than the BM approach. For example, selecting the threshold is not straightforward. Some of the early work on the theoretical basis of POT in hydrology was developed in the 1960s and 1970s (e.g. Shane and Lynn 1964, Todorovic 1970, Zelenhasic 1970, Todorovic and Yevjevich 1969, Todorovic and Woolhiser 1972, Cunnane 1973, Todorovic 1978, Cunnane 1979). Since then, additional studies have followed for developing criteria and guidelines for applying the POT models, comparing BM and POT predictions (e.g. Ashkar and Rousselle 1983, Buishand 1990, Lang *et al.* 1999), and extending the POT models into a nonstationary framework (e.g. Prosdocimi *et al.* 2015).

Assume that $Z(t)$ represents a realization of a hydrological process (e.g. streamflow, precipitation, or wind) whose extremes are of interest. Over a specified time interval $[0, t]$, e.g. a year, there are multiple peaks of the process. Consider now a threshold or base level denoted by u, and values of $Z(t)$ that are above the base level u, denoted as $Z_1, Z_2, \ldots, Z_i$, and collectively defined as the "occurrence process." Assume further that n events occur over the time interval [0, t], and the "exceedences" above the threshold, u, are identified as $Y_i = Z_i - u$, $i = 1, 2, \ldots, n$. Note that the number of exceedences n is also a random variable, which is denoted as $N(t)$ and varies from 0 to ∞. Likewise, the magnitude of exceedences Y_i is a random variable. The goal of the POT approach is to determine the distribution of the maximum of the exceedences over the arbitrary time interval [0, t]. Since both $N(t)$ and Y are random variables, the problem has been characterized as a random number of random variables (Todorovic 1970).

If we denote the event $E_n^t = [N(t) = n]$, then $\mathrm{P}[E_n^t] = P[N(t) = n]$ is the probability of n exceedences in the time interval [0,t], and $X_t = \max(Y_1, Y_2, \ldots, Y_n)$. Todorovic (1970) demonstrated that the cdf of X_t, $F_X^t(x) = P[X_t \le x]$ is given by:

$$F_X^t(x) = P[E_0^t] + \sum_{n=1}^{\infty} P\left\{ \bigcap_{j=1}^{n} [Y_j \le x] \cap E_n^t \right\} \tag{6}$$

which shows that the distribution of the maximum over the interval [0,t] is made up of the distributions of the number of exceedences and of the magnitude. In addition, if the exceedence events are characterized as a time-dependent Poisson process (Todorovic and Yevjevich 1969, Todorovic 1970), the distribution of the number of exceedences $N(t)$ can be written as:

$$\mathrm{P}[E_n^t] = \frac{\exp[-\Lambda(t)][\Lambda(t)]^n}{n!} \tag{7}$$

in which $\Lambda(t)$ is the time-varying parameter of the Poisson process. If $\Lambda(t) = \lambda t$, where λ is a constant rate, then $N(t)$ follows a homogeneous Poisson as:

$$\mathrm{P}[E_n^t] = \frac{\mathrm{e}^{-\lambda \mathrm{t}}(\lambda \mathrm{t})^{\mathrm{n}}}{n!} \tag{8}$$

Many studies have been made examining the suitability of the Poisson assumption, particularly since the mean and the variance are the same for the Poisson distribution (e.g. Cunnane 1979). Thus, alternative distributions for the occurrence process have been proposed, such as the binomial and negative binomial (e.g. Cunnane 1979, Lang *et al.* 1999, Onöz and Bayazit 2001, Bhunya *et al.* 2012, 2013). Overall, it appears that the Poisson assumption works reasonably well.

The exceedence Y_i above the threshold may have any arbitrary distribution. The simplest model suggested in the literature is the exponential distribution (Shane and Lynn 1964, Todorovic and Zelenhasic 1970):

$$F_Y(y) = 1 - \exp(-y/\tilde{\sigma}) \tag{9}$$

where $\tilde{\sigma}$ is the scale parameter. While alternative pdfs have been proposed, such as the two-parameter gamma (Zelenhasic 1970), Weibull (Ekanayake and Cruise 1993), LN (Rosbjerg 1987), and LP3 (Bačová-Mitková and Onderka 2010), the generalized Pareto (GP) distribution has become the most commonly used because of its asymptotic property (Pickands 1975, Coles 2001). The cdf of the GP distribution is:

$$F_Y(y) = 1 - (1 + \xi y/\tilde{\sigma})^{-1/\xi} \tag{10}$$

where $\tilde{\sigma}$ and ξ are the scale and shape parameters, respectively. Furthermore, if the exceedences $Y_1, Y_2, \ldots, Y_n$ are i.i.d. with cdf given by Equation (10), and are independent of the occurrence of a nonhomogeneous Poisson process given by Equation (7), then the cdf of the maximum of Equation (6) yields:

$$\begin{aligned} F_X^t(x) &= \exp[-\Lambda(t)] \\ &\quad + \sum_{n=1}^{\infty} [F_Y(x)]^n \frac{\exp[-\Lambda(t)]\,[\Lambda(t)]^n}{n!} \\ &= \exp[-\Lambda(t)\,(1 - H(x)] \end{aligned} \tag{11}$$

Equation (11) lends itself to the derivation of specific cases for stationary and nonstationary conditions. For instance, when $N(t)$ follows a homogeneous Poisson as in Equation (8) and the exceedences Y_i are exponential as in Equation (9), then the cdf of X is given by:

$$F_X^t(x) = \exp[-\lambda\, t\, \exp(-x/\tilde{\sigma})] \tag{12}$$

which is the Gumbel distribution (Cunnane 1973). In addition, if the exceedences are GP distributed, then the distribution of the maximum X is:

$$F_X^t(x) = \exp\left[-\lambda\, t\, (1 + \xi\, x/\tilde{\sigma})^{-1/\xi}\right] \tag{13}$$

which is known as the Poisson–generalized Pareto (Poisson-GP) model. Remarkably, the Poisson-GP model has the same form as a GEV with parameters (μ, σ, ξ), where $\lambda\, t = (\tilde{\sigma}/\sigma)^{-1/\xi}$ and $\tilde{\sigma} = \sigma + \xi(u - \mu)$. For further details the reader is referred to Smith (1989) and Coles (2001).

Referring to Equation (11), extension of the POT method to nonstationary conditions can be achieved in several ways. The threshold u, the occurrence process $N(t)$, and the exceedences Y_i may be considered to be nonstationary, although one must balance model complexity, the principle of parsimony, physical attribution of nonstationary components of the model(s), and data availability, while maintaining the key assumptions of POT modeling. For instance, the threshold u is often modeled using quantile regression (Koenker 2005) and this approach has been used successfully for regional POT models (e.g. Roth *et al.* 2012, 2014). Likewise, seasonality in the occurrence process may modeled by a superposition of sinusoidal cycles as (North 1980):

$$\lambda(s) = \lambda_0 \exp\left[\sum_{j=1}^{L} a_j \sin(\omega j s + b_j)\right] \tag{14}$$

where λ_0, a_j, and b_j are parameters to be estimated from data, L is the number of significant harmonics necessary to fit the periodic function, $\omega = 2\pi/t_0$, and t_0 is the length of the cycle (e.g. 365 if the time step is daily and the basic period is one year). A continuous trend may also be added to the periodic process when necessary. And time dependence in the exceedences Y may be included by using the cdf of X as (North 1980):

$$F_X^t(x) = \exp\left\{-\int_0^t [1 - F_Y^s(x)]\,\lambda(s)\,ds\right\} \tag{15}$$

Parey *et al.* (2007) modeled both the occurrences and the parameters of the GP distribution as time-dependent processes. Prosdocimi *et al.* (2015) also used the point process formulation of the POT approach and modeled the resulting GEV parameters as functions of time and other covariates, such as rainfall and temporal evolution of urbanization in a catchment. They also compared the nonstationary POT and BM models using the same covariates. Some applications of nonstationary POT models can be found in the literature (e.g. Silva *et al.* 2014, 2016, Prosdocimi *et al.* 2015, Razmi *et al.* 2017). The results show that using POT data appears to be more effective than using BM data for modeling the distribution of the maximum of hydrological extremes.

4.3 Covariates, estimation, model selection, and testing

Nonstationarity in extremes, such as floods and sea levels, may be modeled assuming that the parameters μ and σ are time dependent such as $\mu_t = \mu_0 + a\,t$ and $\log \sigma_t = \sigma_0 + b\,t$ (e.g. El Adlouni *et al.* 2007, Menéndez and Woodworth 2010, Ruggiero *et al.* 2010, Katz 2013, Salas and Obeysekera 2014). Vogel *et al.* (2011) found that a single and parsimonious exponential model may better approximate the behavior of the mean and variance of series of AMF at thousands of rivers in the USA with little evidence of any change in the coefficient of variation of flood series (refer to Section 4.4 for additional details on this issue). Importantly, the mean and standard deviation are coupled, so that assuming a constant coefficient of variation may be very plausible, even under nonstationary conditions (although it must be validated with the data.) The shape parameter is difficult to estimate reliably and is usually considered as a constant to ensure identifiability (Coles 2001, Katz 2013). Typically, the length of the underlying data (e.g. annual floods) is not long enough to estimate reliably all the parameters as time dependent (see Section 6 for additional detail).

The parameters can also be modeled as functions of exogenous variables or covariates. The covariates are external variables of phenomena, such as ENSO and AMO, that may influence the value of the event of interest (e.g. flood) at a given location. An important part of building a model is to be able to identify those covariates that are physically justified and attributed as the cause(s) of the nonstationarity (e.g. Merz *et al.* 2012). Natural climatic variability may cause abrupt shifts in hydrological regimes, which appear to be of random lengths (e.g. Kiem *et al.* 2003, Sveinsson *et al.* 2003, 2005, Akintug and Rasmussen 2005, Enfield and Cid-Serrano 2006, Rao 2009, Villarini *et al.* 2009, Park *et al.* 2011). For example, the quarter monthly annual maximum outflows of Lake Ontario and the 7-day low flows of the Parana River appear to have a shifting pattern over one or more "levels" (e.g. Sveinsson *et al.* 2005). However, if the shifts are of decadal or multi-decadal nature, typical observational records may not have a sufficient number of "shifts" to determine reliably the probabilistic characteristics of the duration and magnitude of those shifts. The AMO, which has been linked to rainfall regimes and annual streamflows (Enfield *et al.* 2001), and annual maximum floods (Rao 2009), has only a few shifts in its observational record, which started around 1850. However, proxy tree-ring data have been used to extend the records back in time, which allowed the estimation of the probability of a future shift given the length of a current AMO regime state (Enfield and Cid-Serrano 2006). Park *et al.* (2011) demonstrated the use of this approach for predicting storm surge extremes, assuming the parameters of GEV distributions depend on the AMO regime state. Likewise, other atmospheric and oceanic fluctuations, such as the Pacific Decadal Oscillation (PDO) and the El Niño Southern Oscillation (ENSO), have been known to modulate the hydrological variability of some river basins in the USA (e.g. Tootle *et al.* 2005, Nowak *et al.* 2012). One of the major challenges in dealing with shifting extremes is the estimation and projection of regime shifts. Salas and Obeysekera (2014) used a mixed model for fitting the distribution of annual floods, where data exhibited two shifting patterns associated with the AMO. In addition to climate, anthropogenic factors such as land-use changes (e.g. increasing deforestation or increasing urbanization in the basin) may influence the variability of extreme events and cause nonstationarity (e.g. Prosdocimi *et al.* 2015).

The parameter vector $\underline{\theta}_t$ of Equation (5) consists of three parameters, the location, scale, and shape parameters. In general, $\underline{\theta}_t$ may involve two to four parameters, which may be expressed as $g(\theta_t^k) = f(\text{covariates})$ where k is the parameter index, g is a link function, and f is a function that may be linear or nonlinear. A general formulation involving additive functions up to four parameters is the Generalized Additive Models in Location, Scale, and Shape (GAMLSS) suggested by Rigby and Stasinopoulos (2005) and implemented in the R software package gamlss. In the GAMLSS formulation, the distribution parameters denoted as μ, σ, ν, and τ may be modeled as linear, nonlinear, or nonparametric smoothing functions of the explanatory variables. Although the overall model is general and may also include random effects, the most useful form for hydrological applications is the semi-parametric additive model (SAM) given by:

$$g_k(\theta_t^k) = \boldsymbol{X}_k\,\boldsymbol{\beta}_k + \sum_{j=1}^{J_k} h_{jk}(x_{jk}) \tag{16}$$

where $g_k(\theta_t^k)$, k = 1, 2, 3, 4, are monotonic link functions associated with each parameter θ_t^k, $\boldsymbol{X}_k$ is a matrix of explanatory variables (i.e. covariates) of size $n \times J_k$, n is the sample size, h_{jk} is a nonparametric additive function (e.g. cubic spline or polynomial) of the explanatory variable X_{jk} evaluated at x_{jk}, J_k is the number of additive functions h, and $\boldsymbol{\beta}_k$ are the parameter vectors. It is noted that the design vector $\boldsymbol{x}_{jk}$ may be the same as or different from a design column of the matrix $\boldsymbol{X}_k$. For example, for the parameter μ, SAM in Equation (16) can be expressed as $g_1(\mu) = X_1\beta_1 + \sum_{j=1}^{J_1} h_{j1}(x_{j1})$. Similar expressions can be written for $g_2(\sigma)$, $g_3(\nu)$, and $g_4(\tau)$. Simpler models, called parametric linear, for two parameters say μ and σ, are $g_1(\mu) = X_1\,\beta_1$ and $g_2(\sigma) = X_2\,\beta_2$. The gamlss package in the R software provides an extensive suite of probability distributions whose parameters can be expressed as functions of explanatory variables (covariates) using the expressions above, in which the coefficients are estimated using a penalized maximum likelihood (ML) method. The availability of numerous forms of probability distributions in gamlss provides various options for modeling extreme value data. However, the users of software must take great care in building as parsimonious a model as possible given the concerns raised by Serinaldi and Kilsby (2015) and others.

There are several methods for estimating GEV parameters from BM data. They include the method of moments (MOM), method of maximum likelihood (ML), method of probability weighted moments (PWM) or L-moments, and Bayesian approaches (e.g. Stedinger *et al.* 1993). The ML method is based on the likelihood function, which is defined as:

$$L(\underline{\theta}_t) = \prod_{i=1}^{n} f(z_i; \underline{\theta}_t) \tag{17}$$

where $f(z_i;\, \underline{\theta}_t)$ is the selected pdf, $\underline{\theta}_t$ is the parameter set (which may be constant or varying as a function of covariates), $z_1, z_2, \ldots, z_n$ are the BM data, and n is the sample size. The approach hinges on maximizing $L(\underline{\theta}_t)$ with respect to the parameter vector $\underline{\theta}_t$. Often it is more convenient to maximize the log-likelihood function, i.e. $\ell\ (\underline{\theta}_t) = \log L(\underline{\theta}_t)$. Literature abounds using the ML method for estimation of hydrological extremes (e.g. Kottegoda and Rosso 2008).

Parameter estimation using POT data is generally based on estimating the parameters of the occurrence process $N(t)$ and those on the exceedences Y_i, as described above. For example, if $N(t)$ is Poisson and Y_i is GP distributed, then the distribution of the annual maximum is GEV with parameters that are related to the parameters of $N(t)$ and Y_i. Note that for estimating the parameters of the GP one may use the ML method based on Equation (17), as described above. However, in using POT data additional considerations are necessary. One is the selection of the threshold u for which there is no standard practice. Lang *et al.* (1999) provide a summary of the operational guidelines for POT methods and address both the threshold selection and the independence criteria. Scarrot and MacDonald (2012) also provide a review of the threshold selection methods. Bernardara *et al.* (2014) suggested a two-step approach including physical declustering and statistical optimization. The first step is intended for identification and characterization of independent events, whereas the second step is for setting a threshold for ensuring the convergence of exceedence distribution towards the commonly used GP distribution. Threshold selection in the context of regional POT has been reviewed by Roth *et al.* (2015).

Several criteria are available for verifying the independence requirement (Lang *et al.* 1999). The US Water Resources Council (USWRC 1976) specified that the exceedences of flood events must be separated by at least 5 days plus the natural logarithm of the basin area in square miles. Additionally, the intermediate flows between two consecutive peaks must drop below 75% of the lowest of the two peaks. The autocorrelation coefficient of successive peaks can also be used, although for this purpose it is a poor measure of dependence (Cunnane 1973). Lang *et al.* (1999) also suggest using the mean number of events per year as one criterion. In the case of floods, Langbein (1949) suggested the base be equal to the lowest annual flood, but in long records it should be raised such that the mean number should be around three to four, and Cunnane (1973) indicated that the mean number should higher than 1.65. Bernardara *et al.* (2014) summarize the criteria used for other variables such as extreme sea levels and rainfall. Furthermore, few quantitative criteria are available to aid in selecting the threshold in the case of the GP distribution (Coles 2001). If the threshold is too low, it is likely to violate the asymptotic assumption regarding the selection of GP for modeling exceedences. But too high a threshold will result in fewer exceedences leading to high variance of the parameters. Two methods available for threshold selection in the case of GP are the mean residual life (MRL) plot and the shape parameter stability test (Coles 2001).

Model selection and diagnostic testing are important steps in extreme value modeling (Coles 2001, Strupczewski *et al.* 2001). For nested models, the significance of the trend in the model parameters (e.g. trend in the location parameter μ) may be evaluated using the likelihood ratio test, which uses the deviance statistic $D = 2\{\ell(\mathrm{M}_r) - \ell(\mathrm{M}_s)\}$, where $\ell(\mathrm{M}_r)$ and $\ell(\mathrm{M}_s)$ are the log-likelihood functions of the fitted models M_r (e.g. a model with a linear trend in the parameter μ) and M_s (without trend in the parameter μ), respectively, and r and s ($r > s$) are the number of parameters in the models considered. The test of the validity of one model against another is based on the probability distribution of D, which is approximately chi-square distributed with $r - s$ degrees of freedom. In addition, as suggested by Katz (2013), one may apply the Akaike information criterion (AIC) to compare among competing models, where AIC (k) = $-2\ell(\mathrm{M}_r) + 2k$ for a model with k parameters. Alternatively, the Bayesian information criterion (BIC), where $\mathrm{BIC}(k) = -2\ell(\mathrm{M}_r) + k\ell(n)$, may be applied where n is the data sample size. The preferred model is that having the minimum value of AIC or BIC. The performance of the selected model is generally assessed by inspecting diagnostic plots such as the probability and quantile plots (Coles 2001). Note that for nonstationary GEV models, the quantile–quantile plot may be based on standardized residuals (Coles 2001, p. 110).

4.4 The value of regression in nonstationary flood frequency analysis

Regression approaches are extremely useful for characterizing the nonstationary behavior of extreme floods and modeling trends. Numerous benefits arise from use of regression in a nonstationary context, when a modeling approach results in residuals that are independent, homoscedastic, and normally distributed (e.g. Helsel and Hirsch 2002). Here we outline eight unique features of linear regression for nonstationary flood frequency analysis described by Serago and Vogel (2018) and Hecht (2017):

(1) A single regression can be used to estimate all conditional moments required for nonstationary flood frequency analysis, offering a parsimonious approach which can be applied to any probability distribution.
(2) Regression is useful for communicating results because of the graphical display of the available data combined with both quantitative and qualitative goodness-of-fit display of the resulting trend model.
(3) Linear regression can be used to approximate a wide class of nonlinear relationships, because monotonic nonlinear functions may be linearized using the ladder-of-powers (Mosteller and Tukey 1977) to enable estimation using ordinary least squares (OLS) regression with all its associated benefits.
(4) Unlike many nonparametric trend detection methods, regression is a method for both trend detection and modeling. Trend modeling is critical for updating design events to reflect current or future hydrological conditions.
(5) Confidence intervals can be calculated that reflect uncertainty associated with the fitted trend model as well as prediction intervals that reflect uncertainty to be expected in future predictions (e.g. Helsel and Hirsch 2002).
(6) Multivariate regression can incorporate multiple interacting covariates in modeling trends (e.g. Kwon *et al.* 2008, Delgado *et al.* 2010, López and Francés 2013, Prosdocimi *et al.* 2014, Condon *et al.* 2015, Sedano *et al.* 2017).
(7) Regression can include estimates of the likelihood of Type I and II errors, which have been shown to be important for quantifying the potential for underdesign and overdesign (Vogel *et al.* 2013). Those estimates can in turn be integrated into a risk-based decision process (Rosner *et al.* 2014).

(8) It is often difficult to distinguish trends from persistence and both may be present in hydrological records (e.g. Cohn and Lins 2005). Persistence generally leads to inflation in the variance of trend model coefficients leading to incorrect statistical inference for the confidence and prediction intervals and for Type I and II error probabilities. Conveniently, a method for adjusting the standard errors of regression coefficients for various forms of persistence has been suggested (Matalas and Sankarasubramanian 2003).

4.4.1 Regression model for nonstationary conditional moments

Regardless of the assumed pdf of the AMF series, which we term Z, a model of the conditional moments is needed to convert a stationary pdf into a nonstationary pdf. Perhaps the most widely used pdfs of AMF series include the LN2, LN3, LP3, and GEV. In each case there are unique properties of the associated logarithms of the AMF series $V = \ln(Z - \tau)$, in which τ denotes the lower bound of the LN3, and GEV pdfs of Z. Note that, if $\tau = 0$, Z is LN2 and $V = \ln(Z)$ is normal; when Z is LN3, $V = \ln(Z - \tau)$ is normal; if Z is LP3, then $V = \ln(Z)$ follows a Pearson Type III (P3) pdf, and when Z is GEV with negative shape parameter, $V = \ln(Z - \tau)$ follows a Gumbel pdf. Thus, it is reasonable to develop a model of the conditional moments of $V = \ln(Z - \tau)$, rather than Z, since V is generally easier to model than Z. Note that Serago and Vogel (2018) developed a more general nonstationary approach based on derived conditional moments of Z, so that this approach could be used to fit a nonstationary version of any probability distribution by using conditional moments in real space. Below we show only the results for conditional moments of V.

Consider the exponential trend model of the AMF series Z as

$$Z = \exp(\beta_o + \beta_1 t + \varepsilon) \tag{18}$$

where β_o and β_1 are model parameters, ε is a normally distributed model error with zero mean and constant variance, and t is an explanatory variable (not necessarily time) that describes the effect of some anthropogenic or climatic factors on the AMF series. Equation (18) was evaluated at thousands of rivers across the USA and the UK by Vogel *et al.* (2011, see Appendix), and Prosdocimi *et al.* (2014), respectively. Although Equation (18) only contains a single independent variable t, it can be extended to a multivariate model that includes physically meaningful covariates (e.g. Kwon *et al.* 2008, Delgado *et al.* 2010, López and Francés 2013, Prosdocimi *et al.* 2014, Condon *et al.* 2015). Model (18) reduces to a linear model by taking logarithms, i.e. $V = \beta_o + \beta_1 t + \varepsilon$, then, OLS regression may be applied. Note that the conditional mean, variance, and skewness of $V = \ln(Z - \tau)$ for the case of linear regression are given by

$$\mathrm{E}(V|t) = \mu_{V|t} = \beta_o + \beta_1 t \tag{19a}$$

$$\mathrm{var}(V|t) = \sigma^2_{V|t} = \sigma^2_V(1-\rho^2) = \sigma^2_V - \beta_1^2\sigma_t^2 \tag{19b}$$

$$\mathrm{skew}(V|t) = \gamma_{V|t} = \gamma_V - \beta_1^3\gamma_t \tag{19c}$$

where the slope coefficient is related to the product moment correlation coefficient ρ between $V = \ln(Z - \tau)$ and t, i.e. $\beta_1 = \rho\sigma_V/\sigma_t$, in which σ_V and σ_t are the standard deviations of V and t, respectively, and γ_V and γ_t are the corresponding skewness coefficients. Again, one may set $\tau = 0$ for the LN2 and P3 cases. Note that in the equations for the conditional mean, variance, and skewness, only a single additional parameter β_1 is needed, i.e. a parsimonious approach, with the added eight advantages of regression outlined above. For a model with no trend, i.e. $\beta_1 = \rho = 0$, the conditional and unconditional variances of V are equal, while if ρ increases to unity, the conditional variance of V tends toward zero. Interestingly, Read and Vogel (2015, Supplementary information) show that the simple regression model also implies a relationship between the nonstationary coefficient of variation $C_{Z|t}$ of Z and its stationary coefficient $C_Z = \sigma_Z/\mu_Z$, i.e. $C_{Z|t} = \sqrt{(C_Z^2+1)^{(1-\rho^2)} - 1}$. As expected, there is a reduction in the nonstationary value of $C_{Z|t}$ as ρ increases, which reflects improvements in our ability to predict future trends. Hecht and Vogel (2017) introduce a heteroscedastic regression approach and derive the conditional moments for a heteroscedastic regression, which enables them to develop a nonstationary LN2 model that can exhibit a trend in the coefficient of variation of Z.

4.4.2 Magnification factors and confidence intervals

Vogel *et al.* (2011) introduced a metric called the magnification factor to reflect the change in the design flood over time that results from the regression model in (18). Consider the case of a nonstationary LN model in which the nonstationary quantile function is given by $Z_p = \exp(\mu_{V|t} + w_p\sigma_{V|t}) = \exp(\beta_o + \beta_1 t \; + w_p\sqrt{\sigma_V^2 - \beta_1^2\sigma_t^2})$, where w_p is the pth quantile of a standard normal variable and again $V = \ln(Z - \tau)$ with $\tau = 0$. A flood magnification factor is defined as the ratio of the quantile functions in two different time periods t and $t + \Delta t$, which gives:

$$M = \frac{Z_p(t + \Delta t)}{Z_p(t)} = \exp(\beta_1 t) \tag{20}$$

for the cases when Z follows either an exponential or a LN2 pdf. For example, M = 1.2 with $\Delta t = 10$ corresponds to a 20% increase in all flood quantiles over a decade. Other nonstationary pdfs will lead to different expressions for the magnification factor. Vogel *et al.* (2011) and Prosdocimi *et al.* (2014) report magnification factors for hundreds of rivers in the USA and the UK, respectively. Of course, one must be very cautious when extrapolating trends with time as the independent variable, since the trend may not persist into the future (e.g. Stedinger and Griffis 2011). Therefore, it is recommended to report confidence/prediction intervals for extrapolated estimates, which tend to widen with time, thus documenting our (in)ability to predict future design events. Simple analytical formulae for such confidence and prediction intervals for estimates of the conditional mean in Equation (19a) are given by Helsel and Hirsch (2002). The reader is referred to Section 7 for further discussion on this subject.

4.4.3 Likelihood of Type I and II errors

Vogel *et al.* (2013) discuss several trend detection studies on the statistical significance of observed trends in geophysical series. Unfortunately, most such studies concentrate on the probability associated with the null hypothesis of "no trend," which we denote as α. The power of such statistical trend tests, defined as the likelihood that we might detect a trend when it exists, is denoted 1 – β, while β is the probability of missing the trend if it exists, i.e. the probability of the Type II error informs us about the likelihood of whether or not society is prepared to accommodate and respond to such trends. Another advantage of the regression-based trend model given in Equation (19a) is that analytical relationships exist for computing both the Type I and Type II error probabilities. Remarkably little attention has been given to the power of the test and the associated likelihood of Type II errors. Lettenmaier (1976), Dupont and Plummer (1998), Vogel *et al.* (2013), and Rosner *et al.* (2014) describe an analytical calculation of the Type II error probability β associated with our estimate of the slope term β_1 for a linear regression, which we describe below.

The trend test amounts to a Student's *t*-test on the estimated value of β_1 in a simple linear regression based on a sample of length N. Given the null hypothesis H_0: $\beta_1 = 0$ *versus* the one-sided alternative hypothesis H_A: $\beta_1 > 0$, one can estimate the probability of a Type I error, α, using $P(T_{N-2} \geq t)$, where T_{N-2} denotes the Student's t random variable with N –2 degrees of freedom, and $t = \hat{\beta}_1/\hat{\sigma}_{\hat{\beta}_1}$, where $\hat{\beta}_1$ is the OLS estimate of the trend slope and $\hat{\sigma}_{\hat{\beta}_1}$ is the standard deviation of the estimator. Similarly, the probability of the Type II error β corresponding to a given value of α can be estimated by $\beta = P(T_{N-2} \leq t_{1-\alpha,N-2} - \delta\sqrt{N})$ where $\delta = 1/\sqrt{(1/\rho^2) - 1}$, ρ is the correlation coefficient between $V = \ln(Z - \tau)$ and t, and $t_{1-\alpha,N-2}$ is the 1 – α quantile of the Student's t distribution with N– 2 degrees of freedom and non-exceedence probability 1 – α (Vogel *et al.* 2013). The relationship between α and β only depends on the sample size N and the correlation coefficient ρ. Note, the trend term β_1 is related to the correlation ρ as $\beta_1 = \rho\sigma_V/\sigma_t$, where σ_t is the standard deviation of t, and σ_V is the standard deviation V in Equation (19b). Rosner *et al.* (2014) document how the estimates of the Type I and II error probabilities, α and β, can be used in a risk-based framework for determining the optimal design event under nonstationary conditions.

5 Return period, risk, and reliability under nonstationarity

A traditional method for designing engineering projects when key drivers affecting the project are uncertain is based on correction factors. For example, a review and survey of guidelines for design flood and design rainfall estimation to account for climate change in Europe indicate that a few countries have guidelines in the form of correction factors (Madsen *et al.* 2014). However, in this section we focus on recently developed statistical techniques and metrics suggested in the literature for evaluating the performance of infrastructure, particularly those related to extreme events arising in nonstationary regimes. They include methods primarily derived from continuous and discrete probabilistic models, hazard functions, and economic risk-based approaches. Some of the methods reviewed here are fairly new and, as mentioned above, the assumptions are still debatable and the metrics need additional testing. Further, in some cases the methods are not established in practice, but we offer possible concepts, ideas, and paths that may be followed and developed as the need and opportunity arise.

As above, we consider the nonstationary cdf of annual floods denoted as $F_Z(z, \underline{\theta}_t)$, where parameter set $\underline{\theta}_t$ varies with covariates that evolve with time. It is also assumed that in an initial year, say year $t = 0$, a hydraulic structure has been designed (built) based on the flood quantile z_{q_0}, which has an initial return period (mean waiting time) $T_0 = 1/p_0 = 1/(1 - q_0)$, where p_0 and q_0 are the exceedence and non-exceedence probabilities of z_{q_0}, respectively. The cdf above could be defined based on BM or POT data.

5.1 Expected waiting time (EWT)

Consider the random variable X as the time where a flood exceeding the design flood z_{q_0} will occur for the first time. For example, the probability that the first flood exceeding the design flood will occur at time $x = 3$ is $(1 - p_1)(1 - p_2)p_3$. In general, the probability that the first flood exceeding z_{q_0} will occur at time x is given by (Salas and Obeysekera 2014):

$$f(x) = p_x \prod_{t=1}^{x-1}(1 - p_t), \qquad x = 1, 2, \ldots, x_{\max} \tag{21}$$

where $f(1) = p_1$ and $x_{\max}$ is the time where p_t becomes equal to 1. Equation (21) is the generalization of the geometric distribution that is applicable to nonstationary conditions, and has parameters (exceedence probabilities) varying with time. Note that if the ps are the same (stationary condition), then Equation (21) simplifies to Equation (1), the well-known geometric distribution. As indicated in Section 2, the stationary geometric distribution has an exponential decay, but that is not the case for the nonstationary geometric distribution in (21), which can take on a variety of unimodal distributional shapes (Read and Vogel 2015). Likewise, the cdf of the geometric distribution (21) becomes:

$$F_X(x) = 1 - \prod_{t=1}^{x}(1 - p_t), \qquad x = 1, 2, \ldots, x_{\max} \tag{22}$$

where $F_X(1) = p_1$ and $F_X(x_{\max}) = 1$. Mandelbaum *et al.* (2007) introduced nonhomogeneous geometric random variables, which are similar to the above and provided a convenient recursive formula for computing $f(x)$.

The nonstationary geometric distribution (Equation (21)) enables one to determine the expected waiting time (EWT) or return period, where the flood exceeding the design flood z_{q0} will occur for the first time as (Cooley 2013, Salas and Obeysekera 2014):

$$T = \mathrm{E}(X) = 1 + \sum_{x=1}^{x_{max}} \prod_{t=1}^{x}(1 - p_t) \tag{23}$$

Equation (23) gives the return period T for nonstationary conditions, which is consistent with the existing definition of return period for the stationary case. But unlike the stationary case where T is only a function of the exceedence probability p (a constant value), i.e. $T = 1/p$, in the nonstationary case T is a function of the time-varying exceedence probabilities p_t. Equation (23) has been derived relying on the geometric distribution (Equation (21)), which is applicable for nonstationary conditions. The interested reader is referred to the papers by Wigley (1988) and Olsen *et al.* (1998), who derived the return period based on concepts of the binomial distribution and mentioned the possibility of using a nonhomogeneous Poisson process. In addition, the variance of X may be determined from $\mathrm{var}(X) = E(X^2) - T^2$ in which:

$$\mathrm{E}(X^2) = \sum_{x=1}^{x_{max}} x^2 p_x \prod_{t=1}^{x-1}(1 - p_t) \tag{24}$$

Likewise, the risk of failure of the referred hydraulic structure having design life n may be determined by $R = \mathrm{P}(X \le n) = F_X(n)$ so that from Equation (22) we get (Salas and Obeysekera 2014):

$$R = 1 - \prod_{t=1}^{n}(1 - p_t) \tag{25}$$

and the reliability becomes:

$$R_\ell = \prod_{t=1}^{n}(1 - p_t) \tag{26}$$

In Section 2, the expressions for T and R and the design procedure for the stationary case were reviewed. A similar procedure can be followed in the nonstationary case, although its implementation is more involved because some of the computations must be done numerically. For example, we illustrate a procedure that may be useful for assessing the return period and risk of failure of an existing hydraulic structure that has been designed for a return period, which we denote by T_0 (e.g. $T_0 = 50$), which gives $p_0 = 1/T_0$, $q_0 = 1 - p_0$, and the design quantile z_{q0} is obtained by inverting $q_0 = F_Z(z, \hat{\underline{\theta}}_0)$, i.e. $z = z_{q_0} = F_Z^{-1}(z, \hat{\underline{\theta}}_0)$. Further, after some years of operation, additional data become available, and the concern for changes in the basin and climate suggests that the hydrological regime may be different from the past. The assessment is based on the nonstationary equations suggested above and the following steps:

(i) Fit a nonstationary flood frequency model $F_Z(z, \hat{\underline{\theta}}_t)$ using the available data.
(ii) The values of $p_t = 1 - q_t$ are obtained from $p_t = 1 - F_Z(z_{q_0}, \hat{\underline{\theta}}_t)$.
(iii) Calculate T from Equation (23) and R from Equation (25) for a specified value of the assessment time frame, say n years.
(iv) The associated uncertainties can be determined following the procedures reviewed in Section 6.

The calculations above are direct and straightforward in the sense that no iterative or numerical solutions are needed.

Alternatively, the procedure could be posed as a design problem, in which case either the risk of failure R or the return period T is fixed initially, and one wishes to compute the design flood capacity. For example, referring to Equation (23), one could set the desired value of T, e.g. T = 100, then find z_{q0} by inverting Equation (23), since $p_t = 1 - F_Z(z_{q_0}, \hat{\underline{\theta}}_t)$ (the solution will involve a numerical trial-and-error procedure). Note that z_{q0} is the design flood that the hydraulic structure would be designed for at the beginning of the time period considered. Then the risk of failure R can be determined from Equation (25). This procedure is summarized in Table 2 (row for method EWT). Conversely, one could set a value of the risk R and a design life n, e.g. R = 5% and n = 50, then z_{q0} is found by inverting Equation (25) (Table 2, row for DLL method; also see Section 5.3). In addition, as indicated in Section 5.5, an economic risk-based decision approach could be set up by considering, for example, the cost of the structure and the expected damages. Table 2 summarizes some of the foregoing concepts, definitions, and equations pertaining to nonstationary extreme processes such as floods.

5.2 Expected number of events (ENE)

A nonstationary flood regime may imply an increasing (or decreasing) frequency of extreme events in a drainage basin. Assume, as before, that the initial design quantile (return level) is z_{q0}, with a design return period T_0. The exceedence probability with respect to this initial design quantile will increase (or decrease) with time due to the nonstationarity of the probability distribution. As a result, the frequency of the occurrences of floods exceeding z_{q0} will also increase (or decrease) with time, and it would be useful for planners and managers of infrastructure to be able to quantify those frequencies for nonstationary regimes.

Considering a time frame of n years (e.g. the design life of the structure), let I_j, $j = 1, 2, \ldots, n$ be a series of Bernoulli random variables that denote the occurrence or absence of an extreme event exceeding z_{q0} in year j, i.e. if an annual flood exceeds z_{q0} in year j, then I_j = 1, otherwise I_j = 0. Further, if the annual floods are independent in time, then the I_js are also independent, but due to nonstationarity they are non-identically distributed. The time-varying exceedence probability is given by $p_j = \mathrm{P}[I_j = 1]$. Following Obeysekera and Salas (2016) let Y be the number of extreme events exceeding the design flood z_{q0} over the design life n, which is equal to $Y = \sum_{j=1}^{n} I_j$ and the possible values of Y are {0, 1, …, n}. It may be shown that under nonstationarity the PMF of Y is Poisson binomial (Tejada and Den Dekker 2011, Hong 2013):

Table 2. Alternative design methods based on return period, design quantile (return level), and risk under nonstationarity. EWT: expected waiting time, ENE: expected number of events, DLL: design life level, and AAR: average annual risk.

Design method	Primary parameters	Return period T, T_0, or $\bar{T}$	Design quantile z_{q0} (return level)	Risk of failure R, over design life, n	Probability distribution
EWT[(b)]	T	T (specified)	Given T, find z_{q0} in $T = \mathrm{E}(X) = 1 + \sum_{x=1}^{x_{\max}} \prod_{t=1}^{x} (1 - p_t)$ [(a)]	Equation (25)[(a)]	Nonhomogeneous geometric (NG)
ENE = 1	n	Find $p_0 = 1 - F_z(z_{q0}, \hat{\underline{\theta}}_0)$ Then $T_0 = 1/p_0$	Solve for z_{q0} in $\sum_{t=1}^{n} p_t = 1$ [(a)]	Equation (25)[(a)]	Poisson-binomial
ENE = m (m > 1)	n, m	Find $p_0 = 1 - F_z(z_{q0}, \hat{\underline{\theta}}_0)$ Then $T_0 = 1/p_0$	Solve for z_{q0} in $\sum_{t=1}^{n} p_t = m$ [(a)]	Equation (25)[(a)]	Poisson-binomial
DLL[(b)]	R, n	Find $p_0 = 1 - F_z(z_{q0}, \hat{\underline{\theta}}_0)$ Then $T_0 = 1/p_0$	Given R and n find z_{q0} in $R = 1 - \prod_{t=1}^{n} (1 - p_t)$ (25)[(a)]	R (specified)	NG or Poisson-binomial
AAR(n)	R, n	Find $\bar{p}$ in $R = 1 - (1 - \bar{p})^n$ Then $\bar{T} = 1/\bar{p}$	Find z_{q0} in $\bar{p} = (1/n) \sum_{t=1}^{n} p_t$[(a)]	R (specified)	Binomial

[(a)] $p_t = 1 - F_Z(z_{q0}, \hat{\underline{\theta}}_t)$ (refer to Section 5.1).
[(b)] If the EWT or DLL methods are used for assessing a previously designed project where the design quantile z_{q0} is known (and the corresponding values of p_0 and T0), then T can be determined from Equation (23) and R from Equation (25) without any numerical or trial and error calculations.

$$\mathrm{P}(Y=y)=\sum_{A\in S_y}\prod_{j\in A}p_j\prod_{i\in A^c}(1-p_i)\quad,\quad y=0,1,\ldots,n \tag{27}$$

where S_y is the set of all subsets of y integers that can be selected from {1, 2, 3, . . ., n}, and A^c is the complement of A with respect to {1, 2, . . ., n}, for example, if n = 3, A = [1,2] and A^c = 3, or A = [1,3] and A^c = 2, or A = [2,3] and A^c = 1, so that S_2 = {[1,2], [1,3], [2,3]}. Depending on the value of n, the number of elements in S_y, which is equal to $n!/(n - y)!y!$, can be quite large. Several approximations and an exact method for determining P(Y = y) of Equation (27) are available in the Poibin package of R-software (R Core Team 2015).

The probability of zero extreme events during the design life n gives the reliability:

$$R_\ell=\mathrm{P}(Y=0)=\prod_{j=1}^{n}(1-p_j) \tag{28}$$

and the corresponding risk of failure is (Obeysekera and Salas 2016):

$$R=1-\prod_{j=1}^{n}(1-p_j) \tag{29}$$

Note that Equations (28) and (29) are the same as (26) and (25), respectively, except they were obtained using a different approach. The expected number of extreme events, denoted ENE, over the design life n under nonstationary conditions is of interest in practical applications. The expected value and variance of Y are:

$$\mathrm{E}(Y)=\sum_{j=1}^{n}p_j \tag{30}$$

$$\mathrm{var}(Y)=\sum_{j=1}^{n}p_j(1-p_j) \tag{31}$$

Note that, under stationary conditions, i.e. p_j = p_0 for all j, so that E[Y] = np_0 and var(Y) = $np_0(1 - p_0)$, as pointed out in Section 2 for the binomial distribution. In the case of nonstationarity, the return period T_0 (consequently p_0 = 1/T_0) may be obtained by inverting Equation (30) by setting E(Y) = m, where m is the specified expected number of events (ENE) exceeding the design flood in n years. For example, the case m = 1 has been discussed by Parey *et al.* (2007), Cooley (2013), Rootzen and Katz (2013), and Obeysekera and Salas (2014), and m > 1 was considered by Obeysekera and Salas (2016). The concept here is that E(Y) = m can be used as a criterion for assessing a project. Thus, for specified values of n and m, Equation (30) can be solved for p_0, T_0, and z_{q0}, which will give the information needed for making design decisions. Refer to rows under ENE = 1 and ENE = m in Table 2 for a summary of the definitions, concepts, and equations as outlined in this section. Furthermore, one may wish to compute a tolerance interval for Y using, say, E(Y) + $c\sigma(Y)$ = r with $c \geq 0$ and $r \geq 1$, which will lead to alternative design and assessment metrics.

5.3 Design life level (DLL)

The waiting time concept has a drawback in that it requires knowledge of nonstationary behavior beyond the design life of a project, and such an extrapolation is uncertain. Rootzen and Katz (2013) suggested a risk measure that incorporates the design life n, explicitly for determining the required design quantile for the nonstationary case. They termed this measure the design life level (DLL), and it is based on the risk that the largest flood during a selected planning period n will be higher than a specified design event z_{q0}, which can be derived from the risk formula (Equation (25)). The explicit incorporation of the design life n into its definition makes it attractive in practical applications. The risk value R is specified, e.g. R = 5%, and then Equation (25) is inverted to determine the appropriate value of z_{q0} and the corresponding values of p_0 and T_0 (Table 2). Rootzen and Katz (2013) also provided an alternative measure known as Minmax Design Life Level, which is based on the yearly risk level. In this case, the design quantile is chosen such that the maximum exceedence probability in any given year during the design life is at most a specified value, say p%. This measure allows one to limit the maximum risk exposure in any given year during the project life. Two graphical displays of risk information were also suggested: a bar plot of annual risk (exceedence probability) for a given design quantile and then a constant risk plot which displays the design quantile that corresponds to a given exceedence probability (yearly risk). Both risk plots are useful communication tools for displaying time-varying risks under nonstationarity.

5.4 Hazard function analysis

The field of hazard function analysis (HFA) involves a probabilistic assessment of the "time to failure," "survival," or "return period" of an event of interest. HFA, sometimes referred to as survival analysis, is used across a wide array of fields, including epidemiology, manufacturing, medicine, actuarial statistics, reliability engineering, economics, and elsewhere. HFA is used to determine the onset or relapse of a disease in bio-statistics, the time until a person becomes unemployed (or employed) in economics, the time until a device fails in reliability engineering, and the time to death in actuarial science, among many other fields and uses (e.g. Klembaum 1996, Klein and Moeschberger 1997, Tung *et al.* 2006, Cleves *et al.* 2008, Finkelstein 2008, Kottegoda and Rosso 2008, Lawless 2011). HFA comprises a well-known set of tools for characterizing the probability distribution of the time of failure associated with a specific event or process over the course of a time period of interest. Likewise, in nonstationary cases, HFA can represent the distribution of the time of failure, as shown for natural hazards in general (Read 2015, Read and Vogel 2016a), and for floods (Read and Vogel 2016b). When the variable of interest, Z, exhibits nonstationary behavior, HFA is particularly applicable and of interest to hydrologists concerned with nonstationary processes.

The foundation of HFA is the hazard function, or failure rate function, $h(t)$ defined as the probability that a failure event occurs in the time interval $(t,\ t + \Delta t)$ (Kottegoda and Rosso 2008):

$$h(t) = \frac{f_T(t)}{1 - F_T(t)} = -\frac{1}{S_T(t)} \times \frac{\mathrm{d}S_T(t)}{\mathrm{d}t} \tag{32}$$

where $h(t)$ has units (failures/time), f_T is the pdf of the time to failure denoted here as T, F_T is the cdf, and $S_T(t) = 1 - F_T(t)$ is the survival function of T, also known as the reliability function, which represents the probability of no failure in the time interval $(0,t]$, given that failure has not yet occurred at t. From Equation (32), it follows that the cumulative hazard function $H(t)$, which represents the total number of failures over a specified time interval (Cleves 2008), is given as:

$$H(t) = \int_0^t h(s)\mathrm{d}s = \int_0^t \frac{\mathrm{d}S_T(s)}{S_T(s)}\mathrm{d}s = -\ln[S_T(t)] \tag{33}$$

Then, the survival function S_T can be rewritten from Equation (33) as:

$$S_T(t) = \exp[-\textstyle\int_0^t h(s)\mathrm{d}s] = \exp[-H(t)] \tag{34}$$

For stationary and i.i.d. processes, the hazard function $h(t)$ is constant and T associated with the design event z_{qo} follows an exponential pdf regardless of the form of $F_Z(z;\theta)$ (Gumbel 1941).

Normally, the hazard function $h(t)$ defined in Equation (32) is only constrained to be non-negative, $h(t) \geq 0$; it may be increasing or decreasing, non-monotonic, or discontinuous (Klein and Moeschberger 1997). However, Read (2015) and Read and Vogel (2016b) interpreted $h(t)$ as the changing exceedence probability associated with a particular design event of interest. They consider a time-varying model for floods, $F_Z(z,\ \underline{\theta}_t)$, in which floods are independent but not necessarily identically distributed (i./n.i.d.). They also restrict their analysis to an increasing trend in the mean and standard deviation of the AMF series, which is an excellent approximation to many actual flood series (Vogel *et al.* 2011, Prosdocimi *et al.* 2014). As indicated in previous sections, under these conditions, the exceedence probability p_t associated with the design event changes as a function of time, and thus the expected waiting time (EWT) until a flood that exceeds the design event occurs is no longer $1/p$. Further, they assume that the hazard function can be expressed and interpreted as the series of nonstationary exceedence probabilities:

$$h(t) = 1 - F_Z(z_{q_o}, \underline{\theta}_t) = p_t \tag{35}$$

where again, z_{q_0} is the design event quantile. They use several independent methods to confirm that this relationship holds in practice for many special cases. The interpretation of the function $h(t)$ in Equation (35) is the same as its definition in Equation (32); thus it still represents a failure rate, even though the hazard function is now standardized over the interval [0,1].

Using the HFA summarized in Equations (32)–(35), Read and Vogel (2016a) derive the pdf, cdf, and survival function of T corresponding to various nonstationary models of flood series Z, including nonstationary versions of the exponential, GP, and LN models. Their results show that, in all cases, the pdf of the time of failure under nonstationary conditions no longer has an exponential shape as for the stationary case, which confirms the results mentioned in Section 5.1 above. Instead, the shape of the pdf of T is complex, more symmetric than exponential, and depends upon the degree and form of the nonstationarity associated with the flood series Z, as well as the usual natural or inherent hydrological variability (Read and Vogel 2016a). In addition, their results based on the use of HFA agree entirely with their analysis summarized in Read and Vogel (2015) for the case when the AMF series Z follows a nonstationary LN model. Their findings suggest that the use of HFA provides a promising framework for summarizing our probabilistic understanding of the time of failure associated with design events under nonstationary conditions and for improving our understanding of the linkages between Z, z_{q_0}, and T.

5.5 Economic risk-based approaches

Economic risk-based decision making (RBDM) is a well-established methodology that determines appropriate levels of infrastructure based on the expected damages avoided vs. the cost of the infrastructure required (e.g.

National Research Council 2000, Tung 2005) and is now standard practice in some countries and agencies (e.g. US Federal agencies, Stakhiv 2011). RBDM approaches differ from those risk-based design-event approaches described above, because they integrate probabilistic notions of risk with economic consequences of flood hazards. RBDM can be used in place of the traditional design-event approach, which selects a particular (average) T-year event, a given risk R in a specified design lifetime n, or considers both T and R, usually specified by regulation or experience, and then designs the necessary infrastructure to protect against the hydrological event. When a design-event risk-based approach does not integrate economic consequences, a suitable risk metric that captures nonstationary conditions is Equation (25) or other alternatives, as discussed in Sections 5.2 and 5.3.

The goal of RBDM is to choose a level of infrastructure protection that minimizes the total expected annual cost (including flood damage costs) of the water infrastructure. Alternatively, RBDM may maximize the project net benefits. Thus, RBDM may lead to flood-risk mitigation measures against a flood event either larger or smaller than, say, the 100-year flood, which is a common design event considered in the traditional analyses for some projects. A common approach to performing RBDM for sequential decision problems is to use a decision tree. A decision tree is the graphical equivalent of a stochastic dynamic program (SDP), which combines a graphical representation of the overall set of alternatives and decisions, with a framework for making risk-based decisions under uncertainty. Decision trees are described in most introductory textbooks in statistical decision theory and decision sciences as well as in textbooks on water resource systems analysis (e.g. Loucks and Van Beek 2005). Below, we discuss approaches applicable for both traditional design-event risk-based and the economic RBDM suited to nonstationary conditions.

Management decisions in a nonstationary world to cope with extreme events, such as floods and low flows, are generally sequential decisions, which depend critically on the uncertainty inherent in future projections of flood (or low flow, as the case may be) scenarios and their corresponding impacts and consequences. Nevertheless, there are remarkably few examples of the application of RBDM to nonstationary water resource problems in the scientific literature and very few examples of the use of decision trees under nonstationary conditions. Fiering and Matalas (1990) provide one of the earliest examples of a sequential statistical decision process for evaluating various alternatives in the context of nonstationarity (climate change). Chao and Hobbs (1997) give a brief history of decision analysis applications to climate change; and apply SDP for evaluating breakwater adaptation under possible climate change impacts on Lake Erie. Likewise, Hobbs *et al.* (1997) applied a decision tree approach to water resources management under climate change, and recently the method has been resurrected (e.g. Gersonius *et al.* 2013, Rosner *et al.* 2014).

A critical challenge in the application of decision trees to the problem of RBDM under nonstationary conditions involves estimation of the probabilities associated with various outcomes (branches of decision tree). Hobbs *et al.* (1997) demonstrate use of a Bayesian approach to analyzing the probabilities in the decision tree for evaluating alternative adaptation strategies for climate change for the Great Lakes. Similarly, Manning *et al.* (2009) describe a Bayesian analysis consisting of aggregating predictions from a range of model predictions, such as GCMs or RCMs (regional climate models). Another approach for applying RBDM using a decision tree under nonstationary conditions would be to integrate the uncertainty inherent in projections of future trends. Such an analysis must incorporate the uncertainty associated with our ability to detect changes in flood series. Rosner *et al.* (2014) introduced a RBDM approach applicable to nonstationary conditions, using a decision tree, with the outcome probabilities based on Type I and Type II error probabilities associated with statistical trend hypothesis test outcomes. Their approach includes: a nonstationary GEV model of flood frequency, the uncertainty inherent in the trend detection process, natural hydroclimatic uncertainty, and a detailed economic analysis associated with the various infrastructure alternatives under consideration. The resulting process enables the decision maker to ask the question when enough information is available to warrant making a particular flood-management adaptation decision under nonstationary conditions.

A simple metric to describe the evolution of flood risks under nonstationary hydrological regimes considering an economic perspective has been suggested by Stedinger and Crainiceanu (2001). They assume that for any given planning period n, the risk for any threshold level can be described by the sequence of annual exceedence probabilities $p_1, p_2, \ldots, p_n$ and an average measure of such varying annual risk, AAR, is:

$$\mathrm{AAR}(n) = \bar{p} = (1/n)\,(p_1 + p_2 + \ldots + p_n) \tag{36}$$

Essentially, Equation (36) can be used in place of the risk function defined in Equation (25). In addition, the authors describe an economic RBDM approach which assumes that a damage function $\mathrm{D}(z)$ for every flood level z is known and ω is an economic discount factor that represents the time value of flood losses. If $z_1, z_2, \ldots, z_n$ are the flood levels over the planning period n, then the discounted damage function, DDF, evaluated at the present time is:

$$\mathrm{DDF}(\omega) = \mathrm{D}(z_1) + \omega\,\mathrm{D}(z_2) + \cdots + \omega^{n-1}\mathrm{D}(z_n) \quad (37)$$

The paper by Stedinger and Crainiceanu (2001) explores alternatives of the damage function $\mathrm{D}(z)$ and applications to the case of increasing flooding at a site in the Mississippi River. The simple metric based on Equation (36) is also included in Table 2 for comparison.

Any of the methods described in previous sections may be extended to incorporate economics into the design procedure under nonstationarity, and such extensions are recommended. However, integration of costs and benefits over the design life under nonstationarity may not be straightforward, as remarked by Rootzen and Katz (2013), and adds another element of difficulty into the analysis.

6 Impact of uncertainty and nonstationarity on design events

It is well known that, when using stationary pdf models, physically-based models, and stochastic models, the uncertainty of design quantiles (e.g. flood return levels) or of any risk metric of a given project is attributed to the uncertainty of the model parameters because they must be estimated based on a limited sample, often small datasets. Likewise, there is an additional uncertainty associated with our lack of knowledge of the true model of the system under consideration. In the case of nonstationary models, the problem is the same, but even more complex and challenging because of the larger number of parameters involved and the alternative ways of including nonstationarity into the analysis. In this regard, while for many models determining parameter uncertainty has been well advanced at least for stationary models, a relevant problem is to what extent the quantification of those uncertainties (e.g. uncertainty of the 100-year flood) is utilized in practice. This is even more so for project assessment under nonstationary conditions because of the many uncertainties involved. The combined impact of parameter uncertainty and model uncertainty is a challenging subject that needs additional research. While many pdfs have been utilized in practice for fitting the distribution of hydrological extremes, the review of literature suggests that pdfs of annual maximums, particularly those arising from POT data, points towards the GEV distribution because of the well-defined extreme value theory. Therefore, one can take advantage of the tools of model selection, comparison, and testing, which for nested models are well established, and include stationary and nonstationary pdfs.

As indicated above, in modeling stationary and nonstationary extreme events an issue of concern is the parameter uncertainty of the selected pdf and the design metrics involved. Since the size of the typical dataset is generally too small to estimate the parameters reliably, one must be cautious in selecting a model with an appropriate number of parameters. For example, even for stationary models, estimating the shape parameter is not reliable. It is for this reason that in some countries (e.g. in the USA) maps have been developed to estimate regional skewness, which combined with at-site skewness gives a more reliable skew value to be used in estimation of the model parameters (HFAWG 2017). Various approaches have been proposed to estimate the variance (uncertainty) of a design quantile under nonstationary conditions (e.g. Coles 2001). They generally fall into three categories: (a) the delta method, (b) bootstrapping, and (c) profile likelihood. If the parameters are estimated using the method of maximum likelihood (ML), the three approaches may be used for constructing confidence intervals associated with estimates of a design event quantile. The delta method is based on the large sample properties of ML estimators. The bootstrap method is based on standardized data, which are then resampled with replacement and used to fit an ensemble of nonstationary models. The profile likelihood method, which is based on the log-likelihood function, is usually more accurate than the other two methods, and may be used to compute confidence intervals for design events. However, even if the profile likelihood method is more accurate, it is more complex in cases of nonstationary models because commonly they involve a bigger number of parameters than the stationary models. Obeysekera and Salas (2014) provide the details of all three methods and illustrate their applications with data of extreme hydrological events.

Another increasingly common approach to uncertainty analysis of design event quantiles involves the use of Monte Carlo methods for repeated sampling of design event quantiles obtained from physically-based models. For example, the generalized likelihood uncertainty estimation (GLUE) method is a common approach used in this category of uncertainty methods, although it has been widely criticized (e.g. Stedinger *et al.* 2008) when it is used with an "informal" likelihood function. Generally, there is now a widespread application of deterministic models for use in deriving design event quantiles and for obtaining uncertainty intervals associated with such quantiles, because such deterministic models can often accommodate physical changes in watershed behavior. An example in this direction has been suggested by Efstratiadis *et al.* (2015), based on a deterministic modeling framework with stochastic residual errors. Often a proper stochastic representation of the deterministic model error component is not included in the analysis, in which case significant systematic biases in design event quantiles can result (Farmer and Vogel 2016). The application of deterministic models without accounting properly for the model residual uncertainty introduces substantial distributional bias into important statistics computed from simulated responses. Farmer and Vogel (2016) apply a complex distributed rainfall–runoff model to hundreds of river basins across the USA to document that this bias is particularly severe at the distributional extremes corresponding to floods and droughts. Even though a deterministic model may generate unbiased simulations overall, differential bias in various important statistics derived from such simulations still arises from ignoring the model residuals. Some of this bias can be mitigated by considering a stochastic reintroduction of model residuals into simulated responses. Exactly how to do so requires additional exploration and may depend on the particular application and the simulation model developed. Farmer and Vogel (2016) document that the stochastic use, as an *ex post facto* solution to the problem of distributional bias, may be an attractive approach. Importantly, many of the components of a rigorous uncertainty analysis can be redirected for use in the stochastic development and implementation of environmental simulation models in operational studies (Vogel 2017). Additional bias in design events will also result from not including other known sources of uncertainty into model output, such as uncertainty in climatic inputs and model parameters. Consideration of uncertainty arising due to climatic input series that drive such models, in addition to uncertainty due to model parameter estimation as well as reintroduction of model errors into simulation output, must all be considered when performing a proper uncertainty analysis based on physically based simulation models (Vogel 2017). Such analyses and developments may be particularly important within the context of the additional uncertainty corresponding to the impacts of climate change.

Uncertainty, which may be even more important than parameter uncertainty, arises from the lack of knowledge of the nonstationary evolution of the model (e.g. a pdf) in the future. It is not advisable to simply extrapolate the historical trends into the future. Such an extrapolation should be based on additional information regarding plausible evidence-based futures (e.g. anthropogenic and climate variability and change investigations), which is why it is advocated to use physically-based covariates for modeling future nonstationary conditions. However, even when physically-based covariates are considered, the extrapolation could prove to be inaccurate due to lack of knowledge of future conditions. For example, in cases where the nonstationarity of floods is due to urbanization, the built-out condition implies an upper limit on future trends in floods, and this limit must be considered in any plausible analysis. Often it may be reasonable to use the historical trends due to urbanization to simply update a flood frequency analysis to reflect current urbanization conditions, which may remain roughly constant for the near-term or for a short period into the foreseeable future. Similarly, there are significant uncertainties associated with future climates and the use of the current suite of GCMs because such models are unreliable, particularly for extreme precipitation events (e.g. Blöschl and Montanari 2010, Kundzewicz and Stakhiv 2010). Quantification of the predictive uncertainty of future hydroclimatic conditions will remain a considerable challenge in the further development of nonstationary methods for extreme events in the future.

7 Further remarks: caveats and complications

As stated above, projections made using any statistical or deterministic model inherently involve uncertainties. For example, the design flood computed using a pdf such as GEV or LP3, fitted to historical records at a particular location, has uncertainties that are typically expressed in the form of confidence intervals. This is so for both stationary and nonstationary conditions; however, in the case of nonstationarity, additional challenges arise due to predictive uncertainties of future trajectories of the underlying anthropogenic and hydroclimatic processes and their impact and consequences on society and the environment.

First, the perceived nonstationarity may not necessarily be caused by anthropogenic or climate change effects. For example, even though trend tests and change point analysis may suggest statistically significant changes in the hydrological regime, such inhomogeneity in the time series may be due to stationary processes, such as low-frequency components of the atmospheric and oceanic system, or to the effect of persistence (e.g. Sveinsson *et al.* 2003, Cohn and Lins 2005, Koutsoyiannis 2011, Koutsoyiannis and Montanari 2014, Sveinsson and Salas 2017). For this reason, Serinaldi and Kilsby (2015) recommend that nonstationary frequency analysis should not be based solely on the at-site time series, but rather require additional information and detailed exploratory data analysis. Regardless of the perceived nonstationarity, whether due to ongoing changes of the physical characteristics of the basin (e.g. urbanization or deforestation), climatic or human influences, or other effects, such effects should be explored and considered. Through such an investigation, a well-defined deterministic mechanism that explains the nonstationary behavior must be identified. Attribution of change is not an easy task (Merz *et al.* 2012), but it may be worth pursuing in some detail, depending on the importance of the project. While a proper model building that includes nonstationarity should in principle be well justified both physically and statistically, one must be aware of the limitations involved (e.g. cost, resources, time, political factors, inertia of institutions, and others) when planning infrastructure.

Matalas (2012), Serinaldi and Kilsby (2015), and others have recommended that stationarity should always remain the default assumption. Silva *et al.* (2016) compared the results obtained by applying alternative flood frequency models considering climate-related covariates, and found that "*the difference in estimates and uncertainty of the design life level (DLL), obtained using stationary and nonstationary models was very small,*" for the particular example considered. In situations when nonstationary behavior is apparent, it is even more important to include the effect of uncertainty in planning and management decisions (Stakhiv 2011). Vogel *et al.* (2013) and Rosner *et al.* (2014) recommend considering outcomes relating to both stationary and nonstationary conditions, in an integrated probabilistic framework known as RBDM (see Section 5.5).

In any effort to account for nonstationarity due to climate change, numerous additional complications arise. Historical changes have been documented by observations and proxy data, but future changes are generally studied by GCM projections and scenarios. However, it is well known that climate projections, particularly for precipitation extremes, are extremely uncertain in most regions of the world (e.g. Blöschl and Montanari 2010, Kundzewicz and Stakhiv 2010). Future projections from climate models may include multiple realizations (ensembles) that project the plausible futures of, say, the precipitation regime. However, there is no accepted standard practice for using alternative futures of climate over the design life to determine an appropriate design flood (or low flows and drought) level for the project. The development of nonstationary methods for pdfs and IDF curves using uncertain climate projections are subjects where research is urgently needed, especially in regions where impacts of climate change are known to be considerable, as well as in areas lacking pertinent data. In this regard, it may be appropriate to refer to Blöschl and Montanari (2010), and as emphasized by Merz *et al.* (2012), who propose distinguishing hard and soft facts in some of the changes arising from the climate models, such as air temperature changes, which are robust and can be considered as a "hard fact," while changes in precipitation, where the uncertainty increases as the scale decreases, must be considered as "soft facts."

Serinaldi and Kilsby (2015) and Obeysekera and Salas (2016) have identified additional challenges in applying nonstationary methods. They include the uncertainty of model structure, and the lack of clear understanding of the actual probabilistic meaning of the concepts of stationary and nonstationary return periods and risks. In fact, the real meaning of return period T, a typical jargon commonly utilized in actual hydrological practice under stationarity, is often misunderstood and does not always reflect or communicate future flood risk over the particular project planning horizon. Read and Vogel (2015, see Table 1 and associated discussion) provide numerous reasons for considering the concepts of system reliability or risk of failure over a planning horizon rather than using return period, under both stationary and nonstationary conditions. An important reason for favoring risk of failure R over the use of return period T is that calculation of R depends on planning horizon n, whereas calculation of T does not. Under stationarity, there is a simple relationship between T and the risk of failure R, for a given planning horizon, but that is not the case with the newer methods developed that are applicable for nonstationary conditions. Therefore, one must make an effort to understand the underlying concepts and assumptions, etc. for their proper applications. Likewise, there are several alternative measures for assessing projects when considering risk and reliability under nonstationarity, but formal comparisons of such applications are still lacking.

In addition, there is a need for developing assessment criteria where more explicit consideration of uncertainty is taken into account. There are many existing classes of water infrastructure worldwide, such as flood walls and spillways, that have been designed under the concepts of a stationary flood regime. For example, the exceedence probability p of the design flood may have been set at, say, $p = 0.002$ (or 500 years of return period), i.e. there is 0.2% chance that a flood will exceed the design flood in any given year. Then a relevant question may be, what is the expected risk of failure of the structure and its uncertainty, say in the next 25 years, given that nonstationary conditions of the flood regime have been detected based on a careful investigation? If the flood regime continues to be stationary, then one can answer the questions using Equation (2), which gives a risk of about R = 5% and its uncertainty (standard deviation) can be determined by a procedure described in the literature (e.g. Salas *et al.* 2013). However, calculating the expected risk, and particularly its uncertainty, is more complex under nonstationary conditions. Then some of the risk measures discussed in Section 5 above could be extended to consider the underlying uncertainties.

Regardless of the models, the methods that are being developed to take into consideration nonstationary conditions are more complex than those that have typically been used assuming stationary conditions. Generally, models are needed, but the equations are more involved, the number of model parameters increases, there is less experience with such models, the data are generally insufficient, and the sources of uncertainties are many. Perhaps some of the difficulties may be overcome by the development of software and databases, data sharing and accessibility (although in some countries the needed data are not easily available, e.g. hydro-meteorological data), and the development of training programs and courses to transfer and disseminate needed knowledge. It is hoped and expected that many of these difficulties will become more manageable as more experience and knowledge are developed. As indicated above, the availability of computational tools and software (e.g. Guilleland and Katz 2011) will alleviate some of the problems, but a word of caution is that they could easily be misused if basic concepts, assumptions, and limitations are not considered.

Lastly, given the various uncertainties involved in planning and management of infrastructure in a nonstationary world, as mentioned above, it may be suggested that the planning horizons may have to be shortened, e.g. from 50 years life to 25 years, and at the same time flexible design and construction of structures are required to better enable extensions, modifications, and retrofitting, as needed and at reasonable cost to society.

Acknowledgements

The first author acknowledges the continuing support of the Department of Civil and Environmental Engineering of Colorado State University. The authors wish to acknowledge the detailed revision of the reviewers. Their suggestions and those of the Associate Editor, Alberto Viglione, improved the depth and breadth of this review paper.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

The second author was partly supported by the National Science Foundation [Grant no. EAR-1204762]. The third author is indebted to the US Army Corps of Engineers, Institute for Water Resources (IWR), for their financial support of this research project.

Disclaimer

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. The views expressed represent those of the author and do not necessarily reflect the views or policies of IWR.

References

Adams, R.M. and Peck, D.E., 2008. Effects of climate change on drought frequency: impacts and mitigation opportunities. *In*: Chapter 7, A. Dinar and A. Garrido, eds. *Mountains, valleys, and flood plains: managing water resources in a time of climate change*. London: Routledge Publishing.

Akintug, B. and Rasmussen, P.F., 2005. A Markov switching model for annual hydrologic time series. *Water Resources Research*, 41 (9). doi:10.1029/2004WR003605

Ashkar, F. and Rousselle, J., 1983. The effect of certain restrictions imposed on the interarrival times of flood events on the Poisson distribution used for modeling flood counts. *Water Resources Research*, 19 (2), 481–485. doi:10.1029/WR019i002p00481

Bačová-Mitková, V. and Onderka, M., 2010. Analysis of extreme hydrological Events on THE danube using the Peak Over Threshold method. *Journal of Hydrology and Hydromechanics*, 58 (2), 88–181. doi:10.2478/v10098-010-0009-x

Bayazit, M., 2015. Nonstationarity of hydrological records and recent trends in trend analysis: A state-of-the-art review. *Environmental Processes*, 2 (3), 527–542. doi:10.1007/s40710-015-0081-7

Bernardara, P., *et al.*, 2014. A two-step framework for over-threshold modelling of environmental extremes. *Natural Hazards Earth Systems Sciences*, 14, 635–647. doi:10.5194/nhess-14-635-2014

Bhunya, P.K., *et al.*, 2012. Flood analysis using generalized logistic models in partial duration series. *Journal of Hydrology*, 420-421, 59–71. doi:10.1016/j.jhydrol.2011.11.037

Bhunya, P.K., *et al.*, 2013. Flood analysis using negative binomial and generalized Pareto models in partial duration series (PDS). *Journal of Hydrology*, 497, 121–132. doi:10.1016/j.jhydrol.2013.05.047

Bicknell, B.R., *et al*., 1997. *Hydrological simulation program - Fortran. User's manual for version 11*. Athens, GA: U.S. Environmental Protection Agency, National Exposure Research Laboratory, EPA/600/R-97/080. 755 p.

Blöschl, G. and Montanari, A., 2010. Climate change impacts – throwing the dice? *Hydrologic Process*, 24, 374–381.

Bras, R.L., 1990. *Hydrology: an introduction to hydrologic science*. Reading, MA: Addison-Wesley, 643 p.

Buishand, T.A., 1990. Bias and variance of quantile estimates from a partial duration series. *Journal of Hydrology*, 120, 35–49. doi:10.1016/0022-1694(90)90140-S

Burn, D.H. and Hag Elnur, M.A., 2002. Detection of hydrologic trends and variability. *Journal of Hydrology*, 255, 107–122. doi:10.1016/S0022-1694(01)00514-5

CEH (Centre for Ecology and Hydrology), 2013. *A review of applied methods in Europe for flood-frequency analysis in a changing environment*. H. Madsen, D. Lawrence, M. Lang, M. Martinkova, and T.R. Kjeldsen, eds. Wallingford, UK: Centre for Ecology and Hydrology.

CEH (Centre for Ecology and Hydrology), 2016. *Flood estimation handbook*. Wallingford, UK: Center for Ecology and Hydrology.

Chao, P.T. and Hobbs, B.F., 1997. Decision analysis of shoreline protection under climate change uncertainty. *Water Resources Research*, 33 (4), 817–829. doi:10.1029/96WR03496

Cheng, L. and AghaKouchak, A., 2014. Nonstationary precipitation intensity-duration-frequency curves for infrastructure design in a changing climate. *Scientific Reports*, 4, 7093. doi:10.1038/srep0793

Chow, V.T., Maidment, D.R., and Mays, L.W., 1988. *Applied hydrology*. New York: McGraw Hill, 572 p.

Cleves, M.A., *et al*., 2008. *An introduction to survival analysis using stata*. Stata Press, Texas.

Cohn, T. and Lins, H.F., 2005. Nature's style: naturally trendy. *Geophysical Research Letters*, 32, L23402. doi:10.1029/2005GL024476

Coles, S., 2001. *An introduction to statistical modeling of extreme values*. London: Springer-Verlag, 208 p.

Condon, L.E., Gangopadhyay, S., and Pruitt, T., 2015. Climate change and non-stationary flood risk for the Upper Truckee River Basin. *Hydrology and Earth System Sciences*, 19, 159–175. doi:10.5194/hess-19-159-2015

Cooley, D., 2013. Return periods and return levels under climate change. *In*: Chapter 4, A. AghaKouchak, *et al*., eds. *Extremes in a changing climate: detection, analysis and uncertainty*. Dordrecht: Springer Science + Business media.

Cunnane, C., 1973. A particular comparison of annual maxima and partial duration series methods of flood frequency prediction. *Journal of Hydrology*, 18, 257–271. doi:10.1016/0022-1694(73)90051-6

Cunnane, C., 1979. A note on the Poisson assumption in partial duration series models, *Water Resources Research*, 15, 489–494. doi:10.1029/WR015i002p00489

Cunnane, C., 1988. Methods and merits of regional flood frequency analysis. *Journal of Hydrology*, 100 (1–3), 269–290. doi:10.1016/0022-1694(88)90188-6

Delgado, J.M., Apel, H., and Merz, B., 2010. Flood trends and variability in the Mekong River. *Hydrology and Earth System Sciences*, 14, 407–418. doi:10.5194/hess-14-407-2010

Douglas, E.M., Vogel, R.M., and Kroll, C.N., 2000. Trends in floods and low flows in the United States: impact of spatial correlation. *Journal of Hydrology*, 240 (1–2), 90–105. doi:10.1016/S0022-1694(00)00336-X

Downer, C.W. and Ogden, F.L., 2004. GSSHA: Model To Simulate Diverse Stream Flow Producing Processes. *Journal of Hydrologic Engineering*, 9 (3), 161–174. doi:10.1061/(ASCE)1084-0699(2004)9:3(161)

Dupont, W.D. and Plummer, W.D., 1998. Power and sample size calculations for studies involving linear regression. *Controlled Clinical Trials*, 19, 589–601. doi:10.1016/S0197-2456(98)00037-3

Efstratiadis, A., *et al*., 2014. A multivariate stochastic model for the generation of synthetic time series at multiple time scales reproducing long-term persistence. *Environmental Modelling & Software*, 62, 139–152. doi:10.1016/j.envsoft.2014.08.017

Efstratiadis, A., Nalbantis, I., and Koutsoyiannis, D., 2015. Hydrological modelling of temporally-varying catchments: facets of change and the value of information. *Hydrological Sciences Journal*, 60 (7–8), 1438–1461. doi:10.1080/02626667.2014.982123

Ekanayake, S.T. and Cruise, J.F., 1993. Comparisons of Weibull- and exponential-based partial duration stochastic flood models. *Stochastic Hydrology and Hydraulics*, 7 (4), 283–297. doi:10.1007/BF01581616

El Adlouni, S., *et al*., 2007. Generalized maximum likelihood estimators for the nonstationary generalized extreme value model. *Water Resources Research*, 43 (3), 1–13. doi:10.1029/2005WR004545

Enfield, D.B. and Cid-Serrano, L., 2006. Projecting the risk of future climate shifts. *International Journal of Climatology*, 26 (7), 885–895. doi:10.1002/joc.1293

Enfield, D.B., Mestas-Nuñez, A.M., and Trimble, P.J., 2001. The Atlantic multidecadal oscillation and its relation to rainfall and river flows in the Continental U.S. *Geophysical Research Letters*, 28 (10), 2077–2080. doi:10.1029/2000GL012745

Ewen, J., *et al*., 2002. SHETRAN: physically-based distributed river basin modelling system. *In*: Ch. 3, V.P. Singh and D.K. Frevert, eds. *mathematical models of small watershed hydrology and applications*. Highlands Ranch, CO: WRP.

Farmer, W.H. and Vogel, R.M., 2016. On the deterministic and stochastic use of hydrologic models. *Water Resources Research*, 52. doi:10.1002/2016WR019129

Fiering, M.B. and Matalas, N.C., 1990. Decision-making under uncertainty. *In*: P.E. Waggoner, ed. *Climate change and U.S. water resources*. New York: John Wiley.

Finkelstein, M., 2008. *Failure rate modelling for reliability and risk*. Berlin: Springer Science & Business Media.

Gersonius, B., *et al*., 2013. Climate change uncertainty: building flexibility into water and flood risk infrastructure. *Climatic Change*, 116, 411–423. doi:10.1007/s10584-012-0494-5

Gilleland, E. and Katz, R.W., 2011. New software to analyze how extremes change over time. *Eos, Transactions American Geophysical Union*, 92 (2), 13–14. doi:10.1029/2011EO020001

Gumbel, E.J., 1941. The return period of flood flows. *The Annals of Mathematical Statistics*, 12 (2), 163–190. doi:10.1214/aoms/1177731747

Hall, J., *et al*., 2014. Understanding flood regime changes in Europe: a state-of-the-art assessment. *Hydrology and Earth*

System Sciences, 18, 2735–2772. doi:10.5194/hess-18-2735-2014

Hao, Z. and Singh, V.P., 2009. Entropy-based parameter estimation for extended BurrXII distribution. *Stochastic Environmental Research and Risk Assessment*, 23 (8), 1113–1122. doi:10.1007/s00477-008-0286-7

Hao, Z. and Singh, V.P., 2016. Review of dependence modeling in hydrology and water resources. *Progress in Physical Geography*, 40 (4), 549–578. doi:10.1177/0309133316632460

Hecht, J.S. .2017. Updating design flood estimates using regression. Chapter 1 in Making multi-stakeholder water resources decisions with limited streamflow information. Ph.D. Dissertation, Department of Civil and Environmental Engineering, Tufts University.

Helsel, D.R. and Hirsch, R.M., 2002. *Statistical methods in water resources. US Geological Survey, Techniques of water-resources investigations*. Book 4, Chapter A3-12. Reston, VA: US Geological Survey.

HFAWG (Hydrologic Frequency Analysis Working Group), 2017. *US national flood frequency guidelines*. Bulletin 17C. Reston, VA: US Geological Survey.

Hobbs, B.F., Chao, P.T., and Venkatesh, B.N., 1997. Using decision analysis to include climate change in water resources decision making. *Climatic Change*, 37, 177–202. doi:10.1023/A:1005376622183

Hong, Y., 2013. On computing the distribution function for the poisson binomial distribution. *Computational Statistics & Data Analysis*, 59, 41–51. doi:10.1016/j.csda.2012.10.006

Hosking, J.R.M., 1984. Modeling persistence in hydrological time series using fractional differencing. *Water Resources Research*, 20 (12), 1898–1908. doi:10.1029/WR020i012p01898

Hurst, H.E., 1957. A suggested statistical model of some time series which occur in nature. *Nature*, 180, 494. doi:10.1038/180494a0

IACWD (Interagency Committee on Water Data), 1982. *Guidelines for determining flood flow frequency*. Bulletin 17B. Reston, VA: Office of Water Data Coordination, U.S. Geological Survey.

IPCC (Intergovernmental Panel on Climate Change), 2013. *Climate change 2013: the physical science basis. Contribution of Working Group I to the Fifth Assessment Report of the IPCC*. T.F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.). Cambridge, UK and New York, NY, USA: Cambridge University Press, 1535 p. doi:10.1017/CBO9781107415324

Katz, R.W., 2013. Statistical methods for nonstationary extremes. *In*: A. AghaKouchak, *et al.*, eds. *Extremes in a changing climate: detection, analysis and uncertainty*. Dordrecht: Springer Science + Business media.

Katz, R.W., Parlange, M.B., and Naveau, P., 2002. Statistics of extremes in hydrology. *Advances in Water Resources*, 25, 1287–1304. doi:10.1016/S0309-1708(02)00056-8

Kiem, A.S., Franks, S.W., and Kuczera, G., 2003. Multi-decadal variability of flood risk. *Geophysical Research Letters*, 30 (2), 1035. doi:10.1029/2002GL015992

Klein, J.P. and Moeschberger, M.L., 1997. *Survival analysis: techniques for censored and truncated data*. Berlin: Springer Science & Business Media.

Klembaum, D.G., 1996. *Survival analysis: a self learning text*. New York: Springer.

Koenker, R., 2005. *Quantile regression*. Cambridge: Cambridge University Press.

Kottegoda, N.T. and Rosso, R., 2008. *Applied statistics for civil and environmental engineers*. Malden, MA: Blackwell.

Koutsoyiannis, D., 2002. The Hurst phenomenon and fractional Gaussian noise made easy. *Hydrological Sciences Journal*, 47 (4), 573–595. doi:10.1080/02626660209492961

Koutsoyiannis, D., 2004. Statistics of extremes and estimation of extreme rainfall: II. Empirical investigation of long rainfall records. *Hydrological Sciences Journal*, 49 (4), 590–610. doi:10.1623/hysj.49.4.591.54424

Koutsoyiannis, D., 2011. Hurst-Kolmogorov dynamics and uncertainty1. *JAWRA Journal of the American Water Resources Association*, 47 (3), 481–495. doi:10.1111/j.1752-1688.2011.00543.x

Koutsoyiannis, D., 2016. Generic and parsimonious stochastic modelling for hydrology and beyond. *Hydrological Sciences Journal*, 61 (2), 225–244. doi:10.1080/02626667.2015.1016950

Koutsoyiannis, D. and Montanari, A., 2014. Negligent killing of scientific concepts: the stationarity case. *Hydrological Sciences Journal*. doi:10.1080/02626667.2014.959959

Kundzewicz, Z.W., *et al.*, 2007. Freshwater resources and their management. *In*: M.L. Parry, *et al.*, eds. *Climate change 2007: impacts, adaptation and vulnerability. contribution of working group II to the fourth assessment report of the intergovernmental panel on climate change*. Cambridge, UK and New York, NY, USA: Cambridge University Press, 173–210. Retrieved 2009-05-20.

Kundzewicz, Z.W. and Stakhiv, E., 2010. Are climate models "ready for prime time" in water resources management applications, or is more research needed? *Hydrological Sciences Journal*, 55 (7), 1085–1089. doi:10.1080/02626667.2010.513211

Kwon, H., Brown, C., and Lall, U., 2008. Climate informed flood frequency analysis and prediction in Montana using hierarchical Bayesian modeling. *Geophysical Research Letters*, 35. doi:10.1029/2007GL032220

Lall, U. and Sharma, A., 1996. A nearest neighbor bootstrap for resampling hydrologic time series. *Water Resources Research*, 32 (3), 679–693. doi:10.1029/95WR02966

Lang, M., Ouarda, T.B.M.J., and Bobée, B., 1999. Towards operational guidelines for over-threshold modeling. *Journal of Hydrology*, 225, 103–117. doi:10.1016/S0022-1694(99)00167-5

Langbein, W.B., 1949. Annual floods and the partial-duration flood series. *Transactions, American Geophysical Union*, 30 (6), 879. doi:10.1029/TR030i006p00879

Lawless, J.F., 2011. *Statistical models and methods for lifetime data*. Hoboken, NJ: John Wiley & Sons, 362.

Lee, T.S., Salas, J.D., and Prairie, J., 2010. An enhanced nonparametric streamflow disaggregation model with genetic algorithm. *Water Resources Research*, 46 (8), 1–14. doi:10.1029/2009WR007761

Lettenmaier, D.P., 1976. Detection of trends in water quality data from records with dependent observations. *Water Resources Research*, 12, 1037–1046. doi:10.1029/WR012i005p01037

Li, H., Sheffield, J., and Wood, E.F., 2010. Bias correction of monthly precipitation and temperature fields from Intergovernmental Panel on Climate Change AR4 models using equidistant quantile matching. *Journal of*

Geophysical Research: Atmospheres, 115, D10101. doi:10.1029/2009JD012882

Lins, H.F., and Cohn, T.A., 2011. Stationarity: wanted dead or alive?. *Journal of The American Water Resources Association*, 47 (3), 475–480. doi:10.1111/j.1752-1688.2011.00542.x

López, J. and Francés, F., 2013. Non-stationary flood frequency analysis in continental spanish rivers, using climate and reservoir indices as external covariates. *Hydrology and Earth System Sciences*, 17, 3189–3203. doi:10.5194/hess-17-3189-2013

Loucks, D.P. and Van Beek, E., 2005. *Water resource systems planning and management – an introduction to methods, models and applications*. Place de Fontenoy F-75352 Paris: UNESCO Publishing, WL–Delft Hydraulics, 680.

Madsen, H., *et al.*, 2014. Review of trend analysis and climate change projections of extreme precipitation and floods in Europe. *Journal of Hydrology*, 519, 3634–3650. doi:10.1016/j.jhydrol.2014.11.003

Mandelbaum, M., Hlynka, M., and Brill, P.H., 2007. Nonhomogeneous geometric distributions with relations to birth and death processes. *In*: *Sociedad de Estadistica e Investigacion Operativa*. New York: Springer-Verlag, 281–296. doi:10.1007/s11750-007-0018-z

Mandelbrot, B.B., 1971. A fast fractional Gaussian noise generator. *Water Resources Research*, 7 (3), 543–553. doi:10.1029/WR007i003p00543

Manning, L.J., *et al.*, 2009. Using probabilistic climate change information from a multimodel ensemble for water resources assessment. *Water Resources Research*, 45 (11), 1–13. doi:10.1029/2007WR006674

Mantua, N.J., and Hare, S.R., 2002. The pacific decadal oscillation. *Journal of Oceanography*, 58 (1), 35–44. doi:10.1023/A:1015820616384

Matalas, N.C., 2012. Comment on the announced death of stationarity. *Journal of Water Resources Planning and Management*, 138, 311–312. doi:10.1061/(ASCE)WR.1943-5452.0000215

Matalas, N.C. and Sankarasubramanian, A., 2003. Effect of persistence on trend detection via regression. *Water Resources Research*, 39 (12), 1342. doi:10.1029/2003WR002292

Menéndez, M. and Woodworth, P.L., 2010. Changes in extreme high water levels based on a quasi-global tide-gauge data set. *Journal of Geophysical Research*, 115, C10011. doi:10.1029/2009JC005997

Merz, B., *et al.*, 2012. More efforts and scientific rigour are needed to attribute trends in flood time series. *Hydrology and Earth Systems Science*, 16, 1379–1387. doi:10.5194/hess-16-1379-2012

Milly, P.C.D., *et al.*, 2008. Stationarity is dead: whither water management? *Science*, 319, 573–574. doi:10.1126/science.1151915

Mondal, A. and Mujumdar, P.P., 2015. Modeling non-stationarity in intensity, duration and frequency of extreme rainfall over India. *Journal of Hydrology*, 521, 217–231. doi:10.1016/j.jhydrol.2014.11.071

Montanari, A. and Koutsoyiannis, D., 2014. Modeling and mitigating natural hazards: stationarity is immortal!. *Water Resources Research*, 50 (12), 9748–9756. doi:10.1002/2014WR016092

Mood, A., Graybill, F., and Boes, D.C., 1974. *Introduction to the theory of statistics*. 3rd ed. New York: McGraw Hill.

Mosteller, F. and Tukey, J.W., 1977. *Data analysis and regression*. Menlo Park, CA: Addison-Wesley, 588.

National Research Council, 2000. *Risk analysis and uncertainty in flood damage reduction studies*. Washington, DC: The National Academies Press.

North, M., 1980. Time-dependent model of floods. *Journal of the Hydraulics Division*, 106 (HY5). ASCE, 649–665.

Nowak, K., Hoerling, M., Rajagopalan, B., and Zagona, E., 2012. Colorado river basin hydroclimatic variability. *Journal of Climate, AMS*, 25, 4389–4403. doi:10.1175/JCLI-D-11-00406.1

Obeysekera, J. and Salas, J.D., 2014. Quantifying the uncertainty of design floods under nonstationary conditions. *Journal of Hydrologic Engineering*, 19 (7), 1438–1446. doi:10.1061/(ASCE)HE.1943-5584.0000931

Obeysekera, J. and Salas, J.D., 2016. Frequency of recurrent extremes under nonstationarity. *Journal of Hydrologic Engineering*, 21, 04016005. ASCE. ISSN 1084-0699. doi:10.1061/(ASCE)HE.1943-5584.0001339

Ogden, F.L. and Julien, P.Y., 2002. CASC2D: a two-dimensional, physically-based, hortonian hydrologic model. *In*: V.P. Singh and D.K. Frevert, eds. *Ch.4 Mathematical models of small watershed hydrology and applications*. Littleton, CO: WRP, 972 p.

Olsen, J.R., Lambert, J.H., and Haimes, Y.Y., 1998. Risk of extreme events under nonstationary conditions. *Risk Analysis*, 18 (4), 497–510. doi:10.1111/risk.1998.18.issue-4

Önöz, B., and Bayazit, M., 2001. Effect of the occurrence process of the peaks over threshold on the flood estimates, *Journal of Hydrology*, 244, 86–96. doi:10.1016/S0022-1694(01)00330-4

Papalexiou, S.M. and Koutsoyiannis, D., 2013a. How extreme is extreme? an assessment of daily rainfall distribution tails. *Hydrology and Earth System Sciences*, 17, 851–862. doi:10.5194/hess-17-851-2013

Papalexiou, S.M. and Koutsoyiannis, D., 2013b. Battle of extreme value distributions: a global survey on extreme daily rainfall. *Water Resources Research*, 49, 187–201. doi:10.1029/2012WR012557

Parey, S., *et al.*, 2007. Trends and climate evolution: statistical approach for very high temperatures in France. *Climatic Change*, 81 (3–4), 331–352. doi:10.1007/s10584-006-9116-4

Park, J., *et al.*, 2011. Storm surge projections and implications for water management in South Florida. *Climatic Change*, 107, 109–128. doi:10.1007/s10584-011-0079-8

Petrow, T. and Merz, B., 2009. Trends in flood magnitude, frequency and seasonality in Germany in the period 1951-2002. *Journal of Hydrology*, 371, 129–141. doi:10.1016/j.jhydrol.2009.03.024

Philander, S.G.H., 1990. *El Niño, La Niña and the Southern oscillation*. San Diego, CA: Academic Press, 289 p.

Pickands, J., 1975. Statistical inference using extreme order statistics. *The Annals of Statistics*, 3, 119–131. doi:10.1214/aos/1176343003

Prairie, J., *et al.*, 2006. Modified K-NN model for stochastic streamflow simulation. *Journal of Hydrologic Engineering*, 11 (4), 371–378. doi:10.1061/(ASCE)1084-0699(2006)11:4(371)

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

Prosdocimi, I., Kjeldsen, T.R., and Svensson, C., 2014. Non-stationarity in annual and seasonal series of peak flow and precipitation in the UK. *Natural Hazards Earth Systems Sciences*, 14, 1125–1144. doi:10.5194/nhess-14-1125-2014

R Core Team, 2015. *R:a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Available from:: http://www.R-project.org/

Rajagopalan, B. and Lall, U., 2017. Stochastic streamflow simulation and forecasting. *In*: Chapter 128, V.P. Singh, ed. *Handbook of applied hydrology*. 2nd ed. New York: McGraw Hill.

Rao, D.V., 2009. Discussion of Log-Pearson type 3 distribution and its application in flood frequency analysis. II: parameter estimation methods by V.W. Griffis and J.R. Stedinger. *Journal of Hydrologic Engineering*, 12 (5), 492–500. ASCE.

Razmi, A., Golian, S., and Zahmatkesh, Z., 2017. Non-stationary frequency analysis of extreme water level: application of annual maximum series and peak-over threshold approaches. *Water Resources Management*, 31 (7), 2065. doi:10.1007/s11269-017-1619-4

Read, L.K., 2015. *Planning and communicating risk for non-stationary natural hazards*. Ph.D. Dissertation. Tufts University

Read, L.K. and Vogel, R.M., 2015. Reliability, return periods, and risk under nonstationarity. *Water Resources Research*, 51, 6381–6398. doi:10.1002/2015WR017089

Read, L.K. and Vogel, R.M., 2016a. Hazard function analysis for flood planning under nonstationarity. *Water Resources Research*, 52, 4116–4131. doi:10.1002/2015WR018370

Read, L.K. and Vogel, R.M., 2016b. Hazard function theory for nonstationary natural hazards. *Nhessd*, 3, 6883–6915.

Rigby, R.A. and Stasinopoulos, D.M., 2005. Generalized additive models for location, scale and shape (with discussion). *Applications Statistical*, 54 (part 3), 507–554.

Rootzen, H. and Katz, R.W., 2013. Design life level: quantifying risk in a changing climate. *Water Resources Research*, 49, 5964–5972. doi:10.1002/wrcr.20425

Rosbjerg, D., 1987. Partial duration series with lognormal distributed peak values. *In*: V.P. Singh, ed. *Hydrological frequency modeling*. Massachusetts: Reidel, 117–129.

Rosner, A., Vogel, R.M., and Kirshen, P.H., 2014. A risk-based approach to flood management decisions in a non-stationary world. *Water Resources Research*, 50, 1928–1942. doi:10.1002/2013WR014561

Roth, M., *et al.*, 2012. A regional peaks-over-threshold model in a nonstationary climate. *Water Resources Research*, 48, WRO12214. doi:10.1029/2012WR012214

Roth, M., et al., 2014. Projections of precipitation extremes based on a regional, non-stationary peaks-over-threshold approach: a case study for the Netherlands and north-western Germany. *Weather and Climate Extremes*, 4, 1–10. doi:10.1016/j.wace.2014.01.001

Ruggiero, P., Komar, P.D., and Allan, J.C., 2010. Increasing wave heights and extreme value projections: the wave climate of the U.S. Pacific Northwest. *Coastal Engineering*, 57 (5), 539–552. doi:10.1016/j.coastaleng.2009.12.005

Salas, J.D., *et al.*, 2013. Quantifying the uncertainty of return period and risk in hydrologic design. *Journal of Hydrologic Engineering*, 18 (5), 518–526. Published online 3/17/2012, ASCE. doi:10.1061/(ASCE)HE.1943-5584.0000613

Salas, J.D. and Lee, T.S., 2009. Nonparametric simulation of single site seasonal streamflows. *ASCE Journal of Hydrologic Engineering*, 15 (4), 284–296. April.

Salas, J.D. and Obeysekera, J., 2014. Revisiting the concepts of return period and risk for nonstationary hydrologic extreme events. *ASCE Journal of Hydrologic Engineering*, 19 (3), 554–568. doi:10.1061/(ASCE)HE.1943-5584.0000820

Scarrot, C. and MacDonald, A., 2012. A review of extreme value threshold estimation and uncertainty quantification. *Revstat*, 10, 33–60.

Schaefer, M.G., 1990. Regional analyses of precipitation annual maxima in Washington State. *Water Resources Research*, 26 (1), 119–131. doi:10.1029/WR026i001p00119

Scharffenberg, W.A. and Fleming, M.J., 2010. *Hydrologic modeling system HEC-HMS*. Davis, CA: User's Manual, U.S. Army Corps of Engineers, Hydrologic Engineering Center.

Sedano, K.R., *et al.*, 2017. Predicting the frequency and magnitude of annual floods in the Upper Cauca River based on climatic and reservoir operation indices. *In*: *ASCE-EWRI World Environmental and Water Resources Congress*, 21–27 May Sacramento, CA.

Serago, J. and Vogel, R.M., 2018. Parsimonious nonstationary flood frequency analysis. Advances in Water Resources, 112, 1–16.

Serinaldi, F. and Kilsby, C.G., 2014. Rainfall extremes: toward reconciliation after the battle of distributions. *Water Resources Research*, 50, 336–352. doi:10.1002/wrcr.v50.1

Serinaldi, F. and Kilsby, C.G., 2015. Stationarity is undead: uncertainty dominates the distribution of extremes. *Advances in Water Resources*, 77, 17–36. doi:10.1016/j.advwatres.2014.12.013

Shane, R.M. and Lynn, W.R., 1964. Mathematical model for flood risk evaluation. *Journal of Hydraulics Division*, ASCE 90 (HY6). 1–20.

Sharma, A., Tarboton, D.G., and Lall, U., 1997. Streamflow simulation: a nonparametric approach. *Water Resources Research*, 33 (2), 291–308. doi:10.1029/96WR02839

Silva, A.T., Naghettini, M., and Portela, M.M., 2016. On some aspects of peaks-over-threshold modeling of floods under nonstationarity using climate covariates. *Stochastic Environmental Research and Risk Assessment*, 30, 207–224. doi:10.1007/s00477-015-1072-y

Silva, A.T., Portela, M.M., and Naghettini, M., 2014. On peaks-over-threshold modeling of floods with zero-inflated Poisson arrivals under stationarity and nonstationarity. *Stochastic Environmental Research and Risk Assessment*, 28, 1587–1599. doi:10.1007/s00477-013-0813-z

Singh, V.P., 1995. *Computer models of watershed hydrology*. Highlands Ranch, CO: WRP, 1130.

Singh, V.P. and Frevert, D.R., 2002. *Mathematical models of large watershed hydrology*. Highlands Ranch, CO: WRP, LLC, 891.

Sivapalan, M. and Samuel, J.M., 2009. Transcending limitations of stationarity and the return period: process-based approach to flood estimation and risk assessment. *Hydrological Processes*, 23, 1671–1675. doi:10.1002/hyp.v23:11

Smith, R.L., 1989. Extreme value analysis of environmental time series: an application to trend detection in ground-level ozone. *Statistical Science*, 4 (4), 367–377. doi: 10.1214/ss/1177012400. https://projecteuclid.org/euclid.ss/1177012400

Srinivasan, R. and Arnold, J.G., 1994. Integration of a basin-scale water quality model with GIS. *Journal of the American Water Resources Association*, 30 (3), 453–462. doi:10.1111/jawr.1994.30.issue-3

Stakhiv, E.Z., 2011. Pragmatic approaches for water management under climate change uncertainty. *JAWRA Journal of the American Water Resources Association*, 47 (6), 1183–1196. doi:10.1111/j.1752-1688.2011.00589.x

Stedinger, J. and Crainiceanu, C., 2001. Climate variability and flood-risk management. *Risk-Based Decision Making in Water Resources*, IX, 77–86. doi:10.1061/40577(306)7

Stedinger, J.R., *et al.*, 2008. Appraisal of the generalized likelihood uncertainty estimation (GLUE) method. *Water Resources Research*, 44, WOOB06. doi:10.1029/2008WR006822

Stedinger, J.R. and Griffis, V.W., 2011. Getting from here to where? flood frequency analysis and climate. *JAWRA Journal of the American Water Resources Association*, 47 (3), 506–513. doi:10.1111/j.1752-1688.2011.00545.x

Stedinger, J.R., 2017. Flood Frequency analysis. *In*: Chapter 76, V.P. Singh, ed. *Handbook of Applied Hydrology*. 2nd ed. New York: McGraw Hill.

Stedinger, J.R., Vogel, R.M., and Foufoula-Georgiou, E., 1993. Frequency analysis of extreme events. *In*: Ch.18, D. R. Maidment, ed. *Handbook of hydrology*. New York: McGraw-Hill.

Strupczewski, W.G., Singh, V.P., and Feluch, W., 2001. Non-stationary approach to at-site flood frequency modelling I. Maximum likelihood estimation. *Journal of Hydrology*, 248 (1–4), 123–142. doi:10.1016/S0022-1694(01)00397-3

Sveinsson, O.G.B., *et al.*, 2003. Modeling the dynamics of long term variability of hydroclimatic processes. *Journal of Hydrometeorology*, 4, 489–505. doi:10.1175/1525-7541(2003)004<0489:MTDOLV>2.0.CO;2

Sveinsson, O.G.B., Salas, J.D., and Boes, D.C., 2005. Prediction of extreme events in hydrologic processes that exhibit abrupt shifting patterns. ASCE. *Journal of Hydrologic Engineering*, 10 (4), 315–326. doi:10.1061/(ASCE)1084-0699(2005)10:4(315)

Sveinsson, O.G.B. and Salas, J.D., 2017. Time series analysis and models. *In*: Chapter 18, V.P. Singh, ed. *Handbook of applied hydrology*. 2nd ed. N.Y: McGraw Hill.

Tejada, A. and Den Dekker, A.J., 2011. The role of poisson's binomial distribution in the analysis of TEM images. *Ultramicroscopy*, 111 (11), 1553–1556. doi:10.1016/j.ultramic.2011.08.010

Tetra Tech, 2015. *Hydrologic design standards under future climate for Grand Rapids, Michigan*. Prepared for the City of Grand Rapids, June. Research Triangle Park, NC: Tetra Tech.

Thyer, M. and Kuczera, G., 2000. Modeling long-term persistence in hydroclimatic time series using a hidden state Markov model. *Water Resources Research*, 36 (11), 3301–3310. doi:10.1029/2000WR900157

Todorovic, P., 1970. On some problems involving random number of random variables. *The Annals of Mathematical Statistics*, 41 (3), 1059–1063. doi:10.1214/aoms/1177696981

Todorovic, P., 1978. Stochastic models of floods. *Water Resources Research*, 14 (2), 345–356. doi:10.1029/WR014i002p00345

Todorovic, P. and Woolhiser, D.A., 1972. On the time when the extreme flood occurs. *Water Resources Research*, 8, 1433–1438. doi:10.1029/WR008i006p01433

Todorovic, P. and Yevjevich, V., 1969. *Stochastic process of precipitation*. Hydrology Papers No. 35. Fort Collins: Colorado State University.

Tootle, G.A., Piechota, T.C., and Singh, A., 2005. Coupled oceanic-atmospheric variability and US streamflow. *Water Resources Research*, 41, 12. doi: 10.1029/2005WR004381

Tung, Y.-K., 2005. Flood defense systems design by risk-based approaches. *Water International*, 30 (1), 50–57. doi:10.1080/02508060508691836

Tung, Y.-K., Yen, B.C., and Melching, C.S., 2006. *Hydrosystems engineering reliability assessment and risk analysis*. New York: McGraw-Hill.

Um, M.-J., *et al.*, 2017. Modeling nonstationary extreme value distributions with nonlinear functions: an application using multiple precipitation projections for U.S. cities. *Journal of Hydrology*, 552, 396–406. doi:10.1016/j.jhydrol.2017.07.007

USWRC, 1976. *Guidelines for determining flood flow frequency*. United States Water Resources Council, Bulletin 17, Washington DC: Hydrology Committee, 73 p.

Verdon-Kidd, D.C. and Kiem, A.S., 2015. Regime shifts in annual maximum rainfall across Australia – implications for intensity-frequency-duration (IFD) relationships. *Hydrology and Earth System Sciences*, 19, 4735–4746. doi:10.5194/hess-19-4735-2015

Viessman Jr., W. and Lewis, G.L., 2003. *Introduction to hydrology*. 5th ed. Upper Saddle River, NJ: Prentice Hall, Pearson Education, Inc.

Villarini, G., *et al.*, 2009. On the stationarity of annual flood peaks in the continental United States during the 20th century. *Water Resources Research*, 45 (8), 17. doi:10.1029/2008WR007645

Vogel, R.M., 2017. Stochastic watershed models for hydrologic risk management. *Water Security*, 1 (First Issue, Under Review), 28–35. doi:10.1016/j.wasec.2017.06.001

Vogel, R.M. and Castellarin, A., 2017. Risk, reliability, return periods, and hydrologic design. *In*: Chapter 78, V.P. Singh, ed. *Handbook of applied hydrology*. 2nd ed. New York: McGraw Hill.

Vogel, R.M., Rosner, A., and Kirshen, P.H., 2013. Likelihood of societal preparedness for global change – trend detection. *Natural Hazards and Earth System Science*, Brief Communication, 13, 1–6. doi:10.5194/nhess-13-1-2013

Vogel, R.M. and Shallcross, A.L., 1996. The moving blocks bootstrap versus parametric time series models. *Water Resources Research*, 32 (6), 1875–1882. doi:10.1029/96WR00928

Vogel, R.M., Yaindl, C., and Walter, M., 2011. Nonstationarity: flood magnification and recurrence reduction factors in the United States. *JAWRA Journal of the American Water Resources Association*, 47 (3), 464–474. doi:10.1111/j.1752-1688.2011.00541.x

Vörösmarty, C.J., *et al.*, 2000. Global water resources: vulnerability from climate change and population growth. *Science*, 289, 284–288. doi:10.1126/science.289.5477.284

Webster, V.L. and Stedinger, J.R., 2017. Flood frequency analysis in the United States. *In*: R. Teegavarapu, J.D. Salas, and J.R. Stedinger, eds. *Statistical analysis of hydrological variables: methods and applications*. New York: ASCE.

Westra, S., *et al*., 2014. A strategy for diagnosing and interpreting hydrological model nonstationarity. *Water Resources Research*, 50, 5090–5113. doi:10.1002/2013WR014719

Wigley, T.M.L., 1988. The effect of changing climate on the frequency of absolute extreme events. *Climate Monitoring*, 17, 44–55.

Yilmaz, A.G. and Perera, B.J.C., 2014. Extreme rainfall nonstationarity investigation and intensity-frequency-duration relationship. ASCE. *Journal of Hydrologic Engineering*, 19 (6), 1160–1172. doi:10.1061/(ASCE)HE.1943-5584.0000878

Yue, S., Kundzewicz, Z.K., and Wang, L., 2012. Detection of changes. *In*: Chapter 22, Z.W. Kundzewicz, ed. *Changes in flood risk in Europe*. Wallinford, UK: International Association of Hydrological Sciences, IAHS Special Publication 10, 97–120.

Zelenhasic, E., 1970. *Theoretical probability distributions for flood peaks*. Fort Collins: Colorado State University, Hydrology Papers No. 42.

Zucchini, W. and Guttorp, P., 1991. A hidden Markov model for space-time precipitation. *Water Resources Research*, 27 (8), 1917–1923. doi:10.1029/91WR01403