

# Nonstationarity of Hydrological Records and Recent Trends in Trend Analysis: A State-of-the-art Review

Mehmetcik Bayazit<sup>1</sup>

Received: 16 April 2015 / Accepted: 27 May 2015 / Published online: 30 June 2015  
© Springer International Publishing Switzerland 2015

**Abstract** Recent climate change due to global warming has given an impetus to trend analysis of hydrological time series. Climate change as well as low-frequency climate variability and human intervention in river basins violate the assumption of stationarity, which is claimed to be dead by some researchers. Detailed climate models and long hydrological records are needed to predict the future conditions in a changing world. It must be remembered, however, that all hydrological systems include a stationary element, at least in the form of a random component. A stationary model is sometimes preferable to a nonstationary one when the evolution in time of hydrological processes cannot be predicted reliably. It is attempted to generate synthetic nonstationary time series of future climates by means of a global climate model, which are then used in water resources optimization under uncertainty. The estimation of extremes (floods and low flows) is more important but also much more difficult. The statistical significance of a trend can be detected by means of statistical tests such as the nonparametric Mann-Kendall test, which must be modified when there is serial correlation, possibly by prewhitening. Long-term persistence in hydrological processes also affects the results of the test. Some authors criticized the use of significance levels in statistical tests and recommended using confidence intervals around the estimated effect size. The power of a test depends on the chosen level of significance, sample size and the accuracy of prediction of trends. In some cases, it is more important to increase the power so that errors of estimation that may lead to damages due to inadequate protection are prevented. Frequency analysis of nonstationary processes can be made by fitting a trend to the parameters of the probability distribution. Annual maxima or peaks-over-threshold series can be analyzed incorporating a trend component to the parameters. Design concepts such as return period and hydrological risk should be redefined in a changing world. Design life level is another concept that can be used in a nonstationary context. In management decisions of water structures, a risk-based approach should be used where errors that result in under-preparedness are considered as well

---

✉ Mehmetcik Bayazit  
mehmetcikb@gmail.com

<sup>1</sup> Istanbul Technical University, Istanbul, Turkey

as those resulting in over-preparedness. In a changing world, decision making in water resources management requires long-term projections of hydrological time series that include trend due to anthropogenic intervention and climate change.

**Keywords** Nonstationarity · Trend analysis · Climate change · Statistical significance · Frequency analysis · Return period · Risk · Hydrological design

## 1 Introduction

In recent years, there is an increasing trend in the number of papers concerning nonstationarity and trends in hydrological time series, published in hydrological periodicals. The main reason for this, obviously, is the change of our climate (global warming) due to increase of greenhouse gas concentrations in the atmosphere. But this is not the only reason for hydrometeorological change. Human effects in river basins such as land-cover and land-use changes, urbanization, changes in impervious surfaces and drainage network, deforestation and mining also play an important role.

The assumption of stationarity which has been made so far in water resources planning and management studies is now being challenged. In a widely cited paper by 7 authors from different institutions published in 2008 in the journal *Science* (Milly et al. 2008) it is announced that “Stationarity is dead.”

Stationarity means that hydrological variables fluctuate randomly within an unchanging envelope of variability. These variables in annual scale have time-invariant probability density functions (pdfs), whose properties can be estimated from an available record, which are then used to design and operate water resources projects. In reality, however, anthropological intervention in river basins, and more recently climate change in the world-wide scale have resulted in change and variability of hydrological variables to increase so that they are not sufficiently small to assume stationarity. Means and extremes of precipitation, evapotranspiration and streamflow are now changing.

It is very difficult, however, to predict how the climate change will affect the water cycle in a certain region. Such changes cannot be estimated with a sufficient accuracy from short hydrological records. Existing climate models are not reliable and detailed enough to project changes in runoff.

Water resources engineers are in need of hydrological information that accounts for nonstationarity in a changing world. The physics of the global land-atmosphere system should be understood more thoroughly so that better climate models can be developed that include better representation of surface- and ground-water processes. Land-cover change and land-use management effects should be included in the models. Nonstationary hydrological variables can then be modelled stochastically to describe the evolution of their pdfs in time, based on the projection of climate models driven by various climate scenarios.

This of course does not mean that “hydrologic time series are devoid of stationarity or that stationarity has only a secondary role to play” (Matalas 2012). In the future, stationarity may exist relative to an average different from that before. “It must be assessed to what degree nonstationarity affects the planning and management decisions.” It can be argued that “the announced death of stationarity is premature.”

The problems we are being faced with when dealing with nonstationarity can be summarized as follows:

1. What are the causes of nonstationarity of a hydrological time series (natural climate change, low-frequency climate variability, changes in river basin)? How can we use climate models in predicting future hydrological series? What are the changes in mean conditions and extremes? What are the effects of long-term persistence?
2. How can we decide if there is a significant trend in an observed record of a hydrological variable? How is significance defined? What are the roles of type I and type II errors in testing of a statistical hypothesis? How is a trend evaluated on physical grounds?
3. If there is a significant trend, how can we make projections for the future? Is the observed trend to continue in the future, or is it going to decrease or maybe eventually stop? What is the relation of climate models with basin-scale hydrological variables? How can the effects of land-use and land-cover change be included in the models?
4. How can we use the information about the future of hydrological variables in water resources projects? How do we estimate the pdfs, means and extremes of the variables in the future? How are the concepts such as return period and hydrological risk to be redefined in existence of nonstationarity? How can we include the effects of nonstationarity in hydrological design of water structures?

## 2 Nonstationarity of Hydrological Time Series

Stationarity implies that the pdf of a variable is independent of time. Usually it is sufficient to assume stationarity in the first and second moments of the pdf (mean, variance, covariance), which is called weak stationarity; skew and higher moments are usually not considered because they cannot be estimated accurately from a short record.

Nonstationarity can occur either as a gradual trend or a sudden shift. The emphasis here will be on trends.

### 2.1 Causes of Nonstationarity

Hydrological time series may be nonstationary for various reasons. Climate change, anthropogenic changes in river basins and low-frequency climatic variability are the main reasons for nonstationarity. Climate change due to global warming caused by greenhouse gas emissions affects the water cycle and water supply. Although most of the studies on trend analysis refer to climate change, other sources of nonstationarity may be of much greater magnitude than those that arise from climate change (Hirsch 2011).

Elements of nonstationarity should be detected, described analytically and then removed, which is not a simple procedure. Where nonstationarity arises due to deterministic processes such as land-use change or reservoir operation, it can be eliminated with a deterministic model unlike the effects of future climate change whose details are unknown.

## 2.2 Effects of Climate Variability

It is necessary to distinguish between climate variability and climate change (Stedinger and Griffis 2011). Climate variability is random variation from a long-run average distribution, whereas climate change is a trend or a shift in the long-run distribution.

For climate variability, variation in parameters may be tied to low-frequency components of oceanic-atmospheric phenomena (climate indices describing climatic patterns like NAO, ENSO, PDO, PNA, AMD, etc.). For climate change, on the other hand, it is not clear how to project it into future if we do not understand the physical mechanism causing it.

It is difficult to decide whether an observed trend will continue in future or has already reached its maximum and will stop or reverse or is nothing other than climate variability (Stedinger and Griffis 2011). Climate variability can be confused for trend when records are short, which will disappear when more data are collected (Kundzewicz and Robson 2004).

## 2.3 Effects of Long-Term Persistence

Some authors challenged the statement that stationarity is dead. Matalas (2012) argues that we cannot be sure whether a trend is indeed a trend, however long it has persisted. It is difficult to distinguish between long-term persistence and trend with the limited records that we have. Long-term persistence, called Hurst behaviour, means that wet years follow wet years and dry years follow dry years with a probability higher than that for other natural events. In the future, stationarity may exist relative to a shifting mean, an average different from that before (Matalas 2012). Long-term persistence of hydrological time series can be modelled using long memory processes such as fractional Gaussian noise (fGn) or ARIMA (Salas et al. 2012). For such processes, long-term changes are much more frequent and intense. There are large and long excursions from the mean which may be interpreted as nonstationarity. Therefore future states are much more uncertain and unpredictable (Koutsoyiannis 2013). It has been claimed that observed climate change is consistent with Hurst phenomenon, which makes predicting future states more uncertain.

## 2.4 Role of Stationarity in Nonstationary Systems

Montanari and Koutsoyiannis (2014) argue that in all hydrological systems, there are components (statistics) that are time-invariant, which may need to be interpreted on the basis of past experience and data. All hydrological models, whether deterministic or nonstationary, are affected by uncertainty and therefore include a stationary random component. Nonstationary models involve larger uncertainty additional to that of sampling because of the assumptions made about the future that may be unrealistic. A well-defined time-varying deterministic model that holds for the entire future design life of a project (30–100 years) is usually not available unless the evolution of anthropogenic influences can be predicted accurately. In the case of climate change, it is even more difficult to predict whether a trend will continue in the future, without physical understanding of the causes behind it. An unnecessary resort to nonstationarity may reduce the model's predictive capabilities to effectively plan and design mitigation policies for natural systems caused by hydroclimatic extremes, because of the increased variance of the estimates. A nonstationary model should be considered only when we know the evolution in time of hydrological parameters, so that the future projections are reliable. It is advised to make use of data from nearby sites, hydroclimatic and socioeconomic

data describing natural and anthropogenic factors and relevant documents (Serinaldi and Kilsby 2015). A simple well-understood model is generally preferable to a sophisticated one with large uncertainties. It is concluded that stationarity is still a necessary concept to make reliable predictions for engineering design. Observations of past patterns and information are key elements for a successful prediction for hydrological processes.

## 2.5 Climate Models and Trends

It has been argued that the trends of annual streamflows are consistent with model response to climate forcing (Milly et al. 2008). General circulation models (GCMs) should be improved to obtain more detailed and accurate representation of surface- and ground-water processes at the scale of a river basin. At present, GCMs provide information only on a coarse spatial and temporal scale and cannot simulate processes relevant to hydrology. Downscaling to local scale is needed for water resource systems (Salas et al. 2012).

It is important to be able to predict the future of hydrological processes with respect to trends. Use of climate models to drive hydrological models must be combined with the explanation of hydrological records to find out about the change in future conditions for various watersheds of different characteristics. A simple model with well-understood fundamentals is usually preferable to a sophisticated model whose correspondence to reality is uncertain (Lins and Cohn 2011)

## 2.6 Evidence for Trends in Hydrological Records

Although it is argued that the review of the hydrological data shows that there is not sufficient change in river flow characteristics (Galloway 2011), there are a great number of studies which report significant trends. Vogel et al. (2011), defining flood magnification and recurrence reduction, have found increasing trends in many flood flow series in the United States. Positive trends increase the magnitude of future flows and reduce their average recurrence intervals. As an example, a 100-yr. flood has become a 40-yr. flood in a decade in a certain case. Some of the largest flood magnification factors occurred in heavily urbanized regions pointing out that it is misleading to consider only the impact of climate change, without considering anthropogenic influences such as land-use, water-use, water infrastructure and regulation. Hossain (2014) points out that extreme precipitation in the U.S. has increased in magnitude over the last decades, but that the peak streamflow has not been observed to increase significantly for various reasons, probably because of the change of soil moisture and frozen ground conditions.

Climate change reduces water resources significantly in regions with dry subtropical climates, and also increases flood risk. More variable climate conditions and river flows are expected, causing more severe droughts and floods. Climate change affects not only available water supply but also water demand, water quality and ecosystems.

Madsen et al. (2014) obtained similar results reviewing a large number of studies concerning trend analysis of extreme precipitation and floods in Europe. Although there is some evidence of a general increase in extreme daily precipitation and short-duration rainfall, there are no clear indications of significant trends at regional level of extreme streamflow. Climate model projections indicate an increase in extreme precipitation under a future climate, which is consistent with the observed trend. Hydrological projections of peak flows and flood frequency show both positive and negative changes. A general decrease in flood magnitude

and earlier spring flood is found for catchments with snowmelt-dominated peak flows, consistent with the observed trends (Hirsch 2011).

## 2.7 Guidelines for the Future

It must be assessed to what degree nonstationarity will affect the planning and management decisions for water resources projects.

Changes in the mean conditions can be better estimated than those in the low and high tails of the pdf (low flows and floods), that are very important in water management. Considering that we do not know exactly the present hydrological risk at a site because of limited records, we know even less when we allow for historical climate variability and climate change.

A gap is highlighted between the need for considering effects of environmental change on extreme events, and the paucity of guidelines for how to incorporate them in hydrological studies. Future changes of hydrological risk should be accounted for in designing for a changing environment.

## 3 Trend Analysis of Hydrological Time Series

### 3.1 Detection of Trends

The existence of a significant trend in a time series can be detected by statistical tests. Kundzewicz and Robson (2004) reviewed the methodology for trend detection in hydrological records. The rank-based Mann-Kendall test (MK) is the most widely used nonparametric test for trend analysis (Helsel and Hirsch 1992). All statistical tests involve two kinds of errors. These are the so-called type I error (rejecting the null hypothesis  $H_0$  when it is true), and type II error (not rejecting  $H_0$  when it is false). The null hypothesis  $H_0$  here is that there is no trend in the observed series, which is tested against the alternative hypothesis  $H_1$  that there is a trend. The test is applied at a chosen level of significance  $\alpha$  (usually taken as 0.05), which is equal to the probability of the type I error. The power of the test is defined as  $1-\beta$ , where  $\beta$  is the probability of the type II error. The power of a test, which is the probability of rejecting  $H_0$  when it is false, can be determined only when the true situation is known. In a trend test, this requires the knowledge of trend magnitude. Yue et al. (2002a), and Yue and Pilon (2004) investigated the power of the MK test by Monte Carlo simulation.

### 3.2 Effect of Serial Correlation on Trend Detection

The MK test, devised for independent data, rejects the null hypothesis of no trend more often than that specified by the significance level  $\alpha$  when the data are serially correlated with a positive autocorrelation (von Storch 1995). This is because the variance of the MK test statistic increases with the magnitude of autocorrelation (Yue et al. 2002b). For this reason, trends could be detected that would not be found if the series were independent. von Storch (1995) proposed a procedure called prewhitening (PW) to eliminate this effect, assuming a lag-one autoregressive (AR(1)) model and applying the MK test to the serially independent residuals of the model. Hamed and Rao (1998) applied a different method to eliminate the effect of serial dependence, by modifying the variance of the estimate of the test statistic. Yue and Wang (2004) investigated by simulation the results of this method when there is trend and when there

is no trend. Block bootstrap is another method for removing the effect of autocorrelation on trend. The rejection rate of hypothesis of no trend approaches the nominal significance level  $\alpha$  if the block length of bootstrap samples is chosen appropriately (Önöz and Bayazit 2012).

The elimination of the effect of serial correlation on the trend is a complicated problem because of the mutual interaction between correlation and trend. Prewhitening decreases the inflation of the variance of the test statistic due to serial correlation and reduces the probability of the type I error significantly, almost to the theoretically correct value  $\alpha$ , when there is no trend. When a trend exists, however, the power of the test to detect the trend is decreased as compared with the power before PW, sometimes causing a significant trend not to be detected. There is a trade-off between the type I error and the power. Yue and Wang (2002) stated that PW should be avoided because of this reason, especially for large sample size and large magnitude of trend. Yue et al. (2002b, 2003) proposed a modified PW procedure called trend-free prewhitening (TFPW) where the magnitude of serial correlation is estimated after the trend is removed. The MK test is then applied to the series prewhitened using this serial correlation coefficient, to which the identified trend is added. TFPW has a power higher than other methods, but at a cost of detecting false trends too often (Bayazit and Önöz 2007).

Bayazit and Önöz (2004) and Zhang and Zwiers (2004) commented that there is the risk of too frequent detection of trend when it is not actually present if PW is not applied in the presence of autocorrelation, and that power is a secondary consideration. They concluded that a suitable method of eliminating the effect of serial dependence should be used. Bayazit and Önöz (2007) studied by Monte Carlo simulation under which conditions a real loss of power is caused by PW. They found that PW should be avoided when the sample size and the magnitude of trend slope are large, or when the coefficient of variation is very small. In other cases PW will prevent the false detection of a nonexistent trend, without a significant power loss in identifying a trend that exists.

### 3.3 Effect of Long-Term Persistence on Trend Detection

Another phenomenon that overstates the statistical significance of observed trends is long-time persistence (LTP), which can be mistaken for trend (Sagarika et al., 2011; Dinpashoh et al. 2014). With the increase of the Hurst coefficient  $H$  above 0.5, trend test is increasingly likely to find statistical significance. It is likely to report significant trends about half the time when there is no trend for  $H=0.85-0.90$  (Cohn and Lins 2005), in this case type I error is reduced to 0.25 when AR (1) model is applied for PW, to 0.15 when fractional differencing is applied, and to 0.05 only when combination of both is applied. For northern hemisphere annual temperature data (1856–2004) that has an increasing trend of slope 0.05, MK test with no PW has a  $p$  value of  $2 \times 10^{-27}$ , whereas for AR (1) model  $p=5 \times 10^{-11}$ , for ARMA(1,1) model  $p=2 \times 10^{-4}$ , for LTP model  $p=5 \times 10^{-3}$ , and for a combination of LTP and AR(1) model  $p=0.07 > 0.05$  (Lins and Cohn 2011). The significance of trend decreases with the increase of  $p$ ;  $p$  value above 0.05 implies no statistical significance

### 3.4 Significance of Significance Levels

There have been criticisms of using significance levels in statistical tests. Nicholls (2001) questioned the adoption of a certain arbitrary value such as 0.05 for  $\alpha$ . The detection of a certain trend cannot be achieved with  $\alpha=0.05$  in case of small sample



size. If sample size is increased, on the other hand, a significant effect will eventually be found. Fixing on statistical significance can misdirect us from physically important effects. The more important aspect of the magnitude of the trend can be lost in the question of whether it is statistically significant. Nicholls (2001) recommended using confidence intervals around the effect size (trend magnitude) calculated from the sample, so that we address the question “Given these data, what is the probability that the null hypothesis is true?”. If the confidence interval includes zero, one may conclude that the evidence from the sample is not strong enough to reject  $H_0$ . Confidence intervals provide information on the magnitude of the effect of interest (trend), and it is easier to understand them.

### 3.5 Power of Trend Tests

Vogel et al. (2013) argue that little or no attention is given to the power of trend tests. The societal consequences of making a type I error is that we will prepare for a trend even when it does not exist (over-preparedness). There are situations in which society will regret for under-preparedness (type II error). Thus type II error is often more costly to society than the type I error, because it involves damages due to inadequate protection, which often exceed the cost of wasted money on infrastructure. To obtain a very low probability of under-preparedness, we must accept a fairly high probability of over-preparedness. The only way to reduce both errors is either to increase the accuracy of prediction of trends, or to wait long enough to collect more data. An extrapolation of past trends into future increases uncertainty in future trends due to model, parameter and data uncertainties.

Vogel et al. (2013) suggest that a statistical test is devised where type I error corresponds to under-preparedness (existence of trend), but it is not known how such a test can be constructed.

The power of a trend test cannot be increased to a very large value by solely increasing the sample size. It has been shown that the power can be increased by adopting an innovative model for the time series and including additional covariates (Wang et al. 2015).

### 3.6 Further Remarks on Trend Detection

For short records of about 100 years, multi-year periods of both up and down trends are frequently observed. Decisions made with respect to trend beginning and ending dates can have a dramatic effect on the final results. Koutsoyiannis (2006) thinks that with short records no deterministic interpretation about future behaviour is possible.

A sudden shift in a time series can sometimes be mistaken for a trend although there is no trend before or after the change point (Sagarika et al. 2014). Shifts should be taken into account only when they are known to have occurred at specific times, and from specific causes (Clarke 2013).

In case multi-variate trends exist, uni- and multi-variate tests may lead to the detection of different trend signals (Chebana et al. 2013). Multi-variate distribution is composed of marginal distributions for each variate taken separately and a copula for the dependence structure. Copulas and marginal distributions with time-varying parameters should be used in such cases. A multi-variate version of the MK test is developed.



## 4 Hydrological Analysis and Design in a Nonstationary World

Water resources management decisions made today will have long-term effects because the projects have life-times in the order of 50–100 years. Therefore, we must attempt to account for their impacts in the future (Thompson et al. 2013). This requires the adaptation of frequency analysis, concepts like return period and risk, and hydrological design methods to the conditions of a changing world.

### 4.1 Frequency Analysis

In frequency analysis, a probability distribution function is fitted to an observed record and then the probabilities of occurrences of certain future events are estimated. In nonstationary cases the conventional assumptions of stationarity and independence cannot be made. In a changing world parameters of the distribution of annual extreme values (maxima or minima) are varied in time to fit an observed trend which is extrapolated into the future. Nonstationarity may result in a dramatic change in the probabilities of extreme events. Villarini et al. (2009) found that a certain flood flow had a return period of 1000 years during 1950s, whereas in the present environment the return period was about 10 years. In the case of low-frequency fluctuations, climate indices such as ENSO, PDO and NAO can be directly brought into the parameters as covariates (Khaliq et al. 2006). Changing land-use pattern can also be used as covariate.

Strupczewski et al. (2001), Strupczewski and Kaczmarek (2001) and Villarini et al. (2009) presented a nonstationary approach for at-side flood frequency modeling by incorporating linear trend components into the first two parameters of the distribution. It is important to establish the significance of nonstationarity in the regional context, using long records where available, or using a suitable regional approach. Extrapolation of observed climatic trends into the future may be criticized because the future can be completely different from the past. Results from regional and global climate models can be used to assess changes in statistics of extremes. It should be remembered that the design of hydraulic structures should take into consideration hydrometeorological conditions for the whole duration of the structure. The width of the confidence interval of a quantile estimator is a warning about the risk of extrapolation (Strupczewski et al. 2001).

For the peaks-over-threshold (POT) model that analyzes all the observations exceeding a specified fixed threshold, exceedance rate and magnitude of the excesses have trends in the nonstationary case (Khaliq et al. 2006). Trend in exceedance rate corresponds to a change in the mean, and trend in magnitudes of excesses corresponds to a change in the variance (Smith 1989; Cox et al. 2002). Instead of introducing covariates into the threshold exceedance rate and parameters of the model for the threshold exceedances, Eastoe and Tawn (2009) first modeled and then removed the nonstationarity in the entire data set. Then, the extremes of the preprocessed data were modeled by using the standard approach. Kisely et al. (2010), on the other hand, employed a time-dependent threshold depending on covariates instead of using a nonhomogeneous Poisson process with a time-dependent intensity for the exceedance rate. Roth et al. (2012) used a time-varying threshold determined by quantile regression.

Future estimates of hydrological extremes (floods and droughts) cannot be made accurately because of the limitations of GCMs and lack of sufficient observations. Low resolution of GCMs and their inability to simulate localized processes necessitate the

use of a statistical downscaling model together with a physically based hydrological model in order to obtain projections of variables at the river basin scale. Mondal and Mujumdar (2015) recommend that multiple GCMs are used due to lack of knowledge of the physical climate and natural variability, and then the mean of simulations is used. Nonstationarity can be incorporated into the models using physically based covariates in the parameters of the statistical model. They obtained the time of detection for the change of tail quantiles from the nonstationary projections, and compared them with observations through a detection test. Time of detection is defined as the time where there is evidence to reject the null hypothesis that the quantiles estimated from the observations and future projections are equal. Thus it is estimated for how long historically observed quantiles hold for planning and adaptation.

## 4.2 Return Period and Risk

Risk of Failure, Design Quantile (Level), Return Period and Hydrological Risk are concepts that are widely used in hydrological analysis and design. The conventional notions of these concepts are no longer valid under nonstationarity, because their values will change over time during the life-time of a project. Recently, there have been attempts to develop alternative definitions for these concepts.

Olsen et al. (1998) proposed two definitions of the return period in the nonstationary case, which are extensions of the definitions for the stationary case. In the first definition, the return period  $T$  of the event (observation of design quantile  $x_p$ , with exceedance probability  $p$ ,  $(1-p)$ -th quantile of  $X$ , e.g.,  $T$ -year flood) is defined such that the expected number of sample values larger than  $x_p$  in a sample of size  $T$  is one:

$$Tp = 1, T = 1/p \quad (1)$$

The second definition of the return period is the expected value of waiting time before the failure of probability  $p$  occurs (before the design quantile  $x_p$  is first observed). The probability that the first failure occurs in the year  $k$  is:

$$p_k = (1-p)^{k-1}p \quad (2)$$

and the expected value is:

$$E_k = \sum_{k=1}^{\infty} kp_k = 1/p \quad (3)$$

Therefore

$$T = E_k = 1/p \quad (4)$$

These expressions can easily be extended to the nonstationarity, in which case they lead to slightly different results for  $T$ . In the nonstationary case,  $p$  varies with time as  $p_t$ . The first definition gives:

$$\sum_{t=T_1}^{T_2} p_t = 1 \quad (5)$$

where  $T_1$  and  $T_2$  denote the beginning and the end of the return period  $T$ , respectively. In this definition, the return period  $T$  corresponds to the average probability of exceedance. The second definition leads to (Salas and Obeysekera (2013 and 2014):

$$T = \sum_{k=1}^{k_{\max}} k p_k \prod_{t=1}^{k-1} (1-p_t) = 1 + \sum_{k=1}^{k_{\max}} \prod_{t=1}^k (1-p_t) \tag{6}$$

where  $p_t = 1$  at  $k=k_{\max}$  (otherwise the upper limit of  $k$  is  $\infty$ ). In the case of a positive (increasing) trend, return periods of both definitions decrease with time, second definition corresponding to smaller return periods, therefore a more conservative design standard.

Design quantile to provide a specified level of protection corresponding to a certain return period can be computed from the above equations.

Hydrological risk is defined as the probability that the number of events greater than the design quantile  $x_p$  in a  $n$ -year design period (design life) is equal to or greater than one. In the stationary case risk is:

$$R_n = 1-(1-p)^n = 1-(1-1/T)^n \tag{7}$$

Extending this to nonstationarity, the following expression is obtained (Salas and Obeysekera 2014):

$$R_n = \sum_{k=1}^n p_k \prod_{t=1}^{k-1} (1-p_t) = 1-\prod_{t=1}^n (1-p_t) \tag{8}$$

Rootzen and Katz (2013) introduced a concept called Design Life Level as an extension to the nonstationarity of the concept of risk (probability of failure) in a changing climate. For a risk of  $p$  (the probability that the level is exceeded during the design life period  $T_1 - T_2$ ), the design life level is estimated as the  $(1-p)$ -th quantile of the probability distribution of the maximum event during the period  $T_1 - T_2$ . In the case of an increasing trend, design life level is higher than the design level in the year  $T_1$ . A variant of the design life level is Minimax Design Life Level, chosen such that the extreme (maximal) probability of exceedance in any 1 year in the design life period is at most  $p$ . Therefore, this concept specifies the maximal risk of failure in any 1 year during  $T_1 - T_2$ . This value is determined by first determining the  $(1-p)$ -th quantile of the distribution of  $X$  in each year, and then taking the largest of these quantiles. Risk Plot shows how the risk of exceeding a certain level varies along the design life period, whereas Constant Risk Plot gives the level that is exceeded each year with a certain probability along the design life period.

It should be remembered that the above concepts defined for the nonstationary case are meaningful only if a cause-effect mechanism can be identified for trends and shifts so that their future evolution can be predicted. Statistical methods such as trend tests and optimal model selection criteria are not sufficient to reveal physical mechanisms. Increase of model complexity will result in an increase of uncertainty, and widening of the confidence intervals for estimated pdf parameters, quantiles, return periods and risks, that represent our level of knowledge of the process (Serinaldi and Kilsby 2015).

### 4.3 Hydrological Design

In a nonstationary world, trend test results should be combined with costs and damages associated with a particular alternative to provide a rational design approach (Rosner et al.

2014). The criterion of Expected Regret integrates statistical, economic and hydrological aspects of the management problem, taking into account both type I and type II error probabilities of the statistical trend test. It combines the trend detection probabilities with total expected cost, which is a sum of expected damages and infrastructure costs. The decision maker then can ask whether the economic impact of a trend is great enough even if its statistical significance is below a chosen threshold. Regret reflects the difference between benefits associated with a particular option and the benefits associated with the best option if one has perfect foresight, and is a combination of the cost of adaptation if the trend does not materialize (over-preparation cost) and the damage costs that could have been prevented if a trend that we were not expecting actually materializes (under-preparation cost). If over-preparation cost is smaller than under-preparation cost, it is recommended to decide to realize the project, i.e., to invest.

Rosner et al. (2014) applied the above concept to flood management decisions. In this context, type I error  $\alpha$  corresponds to over-preparedness, and type II error  $\beta$  to under-preparedness. Under-preparedness error signifies substantial socio-economical consequences of not adapting to a nonstationary world, which is often of equal or more relevance to society. In the study it was assumed that the mean of the logarithms of the annual flood maxima is a linear function of time. The values  $\alpha$  and  $\beta$  are inversely related to each other, their relationship depending on the sample size and correlation between flood flows and time. Obviously, one must accept a fairly high probability of under-investment in order to ensure a very low probability of over-investment. Only if the sample size or correlation is increased (larger samples or better trend detection) will the values of both  $\alpha$  and  $\beta$  decrease.

A decision tree is used to address climate change adaptation decisions made under conditions of nonstationarity and uncertainty. Situations in which a particular adaptation option is economically attractive under nonstationarity conditions, but, when stationarity is assumed, the cost of the adaptation outweighs the damages avoided, and therefore appears not to be economically viable, are those in which a risk-based decision approach is needed.

Nonstationary stochastic models are needed in risk-based water resources optimization studies. There have been attempts to generate nonstationary synthetic time series of future climates via a stochastic weather generator based on a number of GCMs (Borgomeo, 2014). Robustness of water resources management plans are then determined using simulated pdfs of decision variables such as water shortage in the future. Probabilistic information on climate uncertainty is thus introduced into the decision making process and the robustness of water resources management plans to future climate-related uncertainties are tested. Such dynamic design models with multiple plausible futures will be of great value in water resources planning and management (Galloway 2011). Probabilistic concepts can help planners to identify which sources of uncertainty are likely to have the greatest impact on long-term planning, and the degree to which they will influence the probability of undesirable outcomes (risks) under nonstationary climate conditions.

Döll et al. (2015) described Integrated Water Resources Management (IWRM) which is an iterative, evolutionary and adaptive process where problem and goal definition is modified at the beginning of each iteration, based on new information and changing external conditions learnt during the previous iteration. Under nonstationary conditions due to climate change and land-use change, scenario planning, learning from experience, development of flexible and low-regret solutions, system vulnerability assessment and decision-support methods are among the approaches used. Risk-based management reduces the vulnerability and exposure to the effects of climate change hazards by soft (institutional) and hard (infrastructural) measures.

Risk assessment for handling the uncertainty of future climate change and its impacts is made using an ensemble of climate change projections by a number of GCMs and various hydrological models to translate their results into river basin variables. Uncertainty arising from socio-economical conditions may be as large as climate-related uncertainty and must be considered in risk-based decision making. Not only ensemble means but also less likely outcomes that may have strong impacts must be taken into account.

Haguma et al. (2015) used a Bayesian dynamic model to incorporate future flow regime projections into a water resources optimization (dynamic programming) model. Historical flows were considered as existing (prior) information. Ensemble climate projections quantified the uncertainty related to the future flow regime. Trends in future flows were represented by transition probabilities between flow classes Long-term dynamic programming, where the nonstationary climate is represented explicitly by the cost-to-go function taken as the expected value of future benefit functions of all climate projections, gives a greater efficiency.

## 5 Conclusion

In view of the observed trends and long-term persistence of hydrological records, we conclude that management of water resource systems under uncertainty requires risk-based decision making with robust designs and policies (Salas et al. 2012). In engineering design we should predict future conditions with sufficient accuracy. In stationary conditions we assume that events will resemble those in the past. This is not possible in the case of nonstationarity. Change of the environment at an unprecedented pace, exposes the human settlements to natural hazard and risk to a greater extent than before (Montanari and Koutsoyiannis 2014). Relationships between nonstationary model parameters and time should not be estimated from the data using statistical methods, but need to be defined a priori based on physical mechanisms.

Although it is not possible to make deterministic forecasts on long timescales (50–100 years), predictions of a quantitative nature are possible so that we can formulate a range of potential future scenarios (Thompson et al. 2013).

Not only hydrological systems but also other environmental systems (ecological, sociological etc.) are anticipated to change as a direct or indirect result of human activity, therefore they all have time-dependent properties. These systems usually have nonlinear responses to climate and anthropogenic changes, making the extrapolation of historical data to future scenarios very difficult. Climate and land-use changes promote changes in vegetation and ecology, which will affect water resources. The coupling (feedback) of hydrological systems interacting with others must be taken into account. The dynamics of hydrological systems arising from interconnected processes lead to uncertainty and difficulty of prediction. In a changing world, decision making in water resources management projects requires long-term predictions of hydrological time series, taking into account changes in their properties. Hydrological research in the future should address these issues in exploring nonstationary water resource systems subject to human intervention. While it may never be possible to make accurate long-term predictions, short time-scale predictions can be improved (Thompson et al. 2013).

The use of hydrological records for this purpose is compulsory but also problematic. Long-term data sets that address both hydrological variability and variability of other environmental subsystems are needed. Some approaches that can be employed are hydrological reconstruction (empirical studies along long time-scales of hydrological change, using long-term data

records as an empirical baseline to assess contemporary changes), and comparative hydrology (studies of a large number of watersheds to identify and test hypotheses about hydrological trends at the catchment scale).

In this review it is attempted to compile and discuss pessimistic and optimistic approaches to dealing with nonstationarity in hydrological time series. It has been argued that stationarity is dead, future states are highly uncertain and cannot be predicted with sufficient accuracy by existing climate models to allow the design of water resource projects reliably in a changing world. With the short hydrological records that are available, it is very difficult to interpret the trends and project them into the future. Even in nonstationary cases, however, there are time-invariant components, and the observation of past patterns and data can supply us with necessary information to predict the future, at least qualitatively, so that the effects of nonstationarity can be included in hydrological design. Simulations by a number of climate models, although highly uncertain but still potentially useful, can be used to make robust risk-based decisions.

Synthetic series for specified time periods in the future are generated, sampling uncertainty of natural variability due to climate change in a probabilistic way, and are used to estimate the frequency distributions of decision variables under alternative water management strategies, showing the extent of system's ability to meet the requirements. Uncertainties of future demand due to population growth and development in river basins can also be considered in a similar way, leading to risk-based decision making.

Hydrological reconstruction along long time-scales, comparative hydrology over catchments and using data to improve the existing models are among the methods that can be used to evaluate the implications of future changes on the behaviour of water resource systems.

In conclusion, to deal with nonstationarity, we must adopt a dynamic design approach that is effective in meeting multiple plausible future scenarios. Decision makers should be provided with information about new multi-disciplinary paradigms in developing robust and sustainable solutions that are risk-based, flexible and adaptive.

## References

- Bayazit M, Önöz B (2007) To prewhiten or not to prewhiten in trend analysis? *Hydrol Sci J* 52(4):611–624
- Bayazit M, Önöz B (2004) Comment on “Application of prewhitening to eliminate the influence of serial correlation on the Mann-Kendall test” by S. Yue and C.Y. Wang”. *Water Resour. Res.*, 40:W08801
- Borgomeo E, Hall JW, Fung F, Watts G, Colquhoun K, Lambert C (2014) Risk-based water resources planning: Incorporating probabilistic nonstationary climate uncertainties. *Water Resour. Res.*, 50: WR01558
- Chebana F, Ouarda TBMJ, Duong TC (2013) Testing for multivariate trends in hydrologic frequency analysis. *J Hydrol* 486:519–530
- Clarke RT (2013) How should trends in hydrological extremes be estimated? *Water Resour Res* 49, WRC20485
- Cohn TA, Lins HF (2005) Nature's style: naturally trendy. *Geophys Res Lett* 32, L23402
- Cox DR, Isham SV, Northrop PJ (2002) Floods: some probabilistic and statistical approaches. *Phil Trans R Soc Lond A* 360:1389–1408
- Dinpashoh Y, Mirabbasi R, Jhajharia D, Abdaneh R, Mostafaeipour A (2014) Effect of short-term and long-term persistence on identification of temporal trends. *J Hydrol Eng ASCE* 19(3):617–625
- Döll P, Jimenez-Cisneros B, Oki T, Amell HW, Benito G, Cogley JG, Jiang T, Kundzewicz ZW, Mwakilila S, Hijima A (2015) Integrating risks of climate change into water management. *J Hydrol Sci* 60(1):4–17
- Eastoe EF, Tawn JA (2009) Modelling non-stationary extremes with application to surface level ozone. *Appl Statist* 58(1):25–45
- Galloway GE (2011) If stationarity is dead, what do we do now? *J Amer Water Res Assoc* 47(3):563–570

- Haguma D, Leconte K, Krau S, Cote P, Brissette P (2015) Water resources optimization method in the context of climate change. *J. Water Resour. Planning and Manag.*, ASCE, 141: 04014051–1.
- Hamed KH, Rao AR (1998) A modified Mann-Kendall trend test for autocorrelated data. *J Hydrol* 204:182–196
- Helsel DR, Hirsch RM (1992) *Statistical methods in water resources*. Elsevier, Amsterdam
- Hirsch RM (2011) A perspective on nonstationarity and water management. *J Amer Water Resour Assoc* 47(3): 436–446
- Hossain F (2014) Paradox of peak flows in a changing climate. *J. Hydrol. Eng.* ASCE, 19 (9): 02514001
- Khaliq MN, Ouarda TBMJ, Ondo J-C, Gachon P, Bobee B (2006) Frequency analysis of a sequence of dependent and/or non-stationary hydro-meteorological observations: a review. *J Hydrol* 329: 534–552
- Kisely J, Picek J, Beranova F (2010) Estimating extremes in climate change simulations using the peak-over-threshold method with a non-stationary threshold. *Glob Planet Chang* 72:53–68
- Koutsoyiannis D (2006) Nonstationarity versus scaling in hydrology. *J Hydrol* 324:239–254
- Koutsoyiannis D (2013) Hydrology and change. *Hydrol Sci J* 58(6):1177–1197
- Kundzewicz ZW, Robson AJ (2004) Change detection in hydrological records—a review of the methodology. *Hydrol Sci J* 49(1):7–19
- Lins HF, Cohn TA (2011) Stationarity: wanted dead or alive? *J Amer Water Resour Assoc* 47(3):475–480
- Madsen H, Lawrence D, Lang M, Martinkova M, Kjeldsen TR (2014) Review of trend analysis and climate change projections of extreme precipitation and floods in Europe. *J Hydrol* 519:3614–3650
- Matalas NC (2012) Comment on the announced death of stationarity. *J Water Resour Planning and Manag ASCE* 138:311–312
- Milly PCD, Betancourt J, Falkenmark M, Hirsch FM, Kundzewicz ZW, Lettenmaier DR, Stouffer RJ (2008) Stationarity is dead: whither water management? *Science* 319:573–574
- Mondal A, Mujumdar PP (2015) Return levels of hydrologic droughts under climate change. *Adv in Water Resour* 75:67–79
- Montanari A, Koutsoyiannis D (2014) Modeling and mitigating natural hazards: stationarity is immortal! *Water Resour Res* 50, WR016092
- Nicholls N (2001) The insignificance of significance testing. *Bull Am Meteorol Soc* 81(5):981–986
- Olsen JR, Lambert JH, Haines YV (1998) Risk of extreme events under nonstationary conditions. *Risk Anal* 18(4):497–510
- Önöz B, Bayazit M (2012) Block bootstrap for Mann-Kendall test of serially dependent data. *Hydrol Proc* 26: 3552–3560
- Rootzen H, Katz RW (2013) Design life level: quantifying risk in a changing climate. *Water Resour Res* 49: 5964–5972
- Rosner A, Vogel RM, Kirschen PH (2014) A risk-based approach to flood management decisions in a nonstationary world. *Water Resour Res* 50, WR014561
- Roth M, Buishand TA, Jangbloed G, Klein Tank AMG, van Zanten JH (2012) A regional peaks-over-threshold model in a nonstationary climate. *Water Resour. Res.*, 48: WR012214.
- Sagarika S, Kalva A, Ahmad S (2014) Evaluating the effect of persistence on long-term trends and analyzing step changes in streamflows of the continental United States. *J Hydrol* 517:36–53
- Salas JD, Obeysekera J (2013) Return period and risk for nonstationary hydrologic extreme events. *World Envir. and Water Resour. Cong. 2013: Showcasing the Future*, ASCE: 1213–1223
- Salas JD, Obeysekera J (2014) Revisiting the concepts of return period and risk for nonstationary hydrologic extreme events. *J Hydrol Eng ASCE* 19(3):554–568
- Salas JD, Rajagopalan BR, Saito L, Brown C (2012) Special section on climate change and water resources: climate nonstationarity and water resources management. *J Water Resour Planning and Manag ASCE* 138: 385–388
- Serinaldi F, Kilsby CG (2015) Stationarity is undead: uncertainty dominates the distribution of extremes. *Adv in Water Resour* 77:17–36
- Smith RL (1989) Extreme value analysis of environmental times series: An application to trend detection in ground-level ozone. *Statis. Sci.*, 367–393
- Stedinger JR, Griffis VW (2011) Getting from here to where: flood frequency analysis and climate. *J Amer Water Resour Assoc* 47(3):506–513
- Strupczewski WG, Kaczmarek Z (2001) Non-stationary approach to at-site flood frequency modelling II weighted least square estimation. *J Hydrol* 248:143–151
- Strupczewski WG, Singh VP, Feluch W (2001) Non-stationary approach to at-site flood frequency modelling I maximum likelihood estimation. *J Hydrol* 248:123–142
- Thompson SE, Sivapalan M, Harman CJ, Srinivasan V, Hipsey MR, Reed P, Montanari A, Blöschl E (2013) Developing predictive insight into changing water systems: use-inspired hydrologic science for the anthropocene. *Hydrol Earth Syst Sci* 17:5013–5039



- Villarini G, Smith JA, Serinaldi F, Bales J, Bates PD, Krajewski WF (2009) Flood frequency analysis for nonstationary annual peak records in an urban drainage basin. *Adv in Water Resour* 32:1255–1266
- Vogel RM, Yaindl C, Walter M (2011) Nonstationarity: flood magnification and recurrence reduction factors in the United States. *J Amer Water Resour Assoc* 47(3):464–474
- Vogel RM, Rosner A, Kirschen PA (2013) Brief communication: likelihood of societal preparedness for global change: trend detection. *Nat Hazards Earth Syst Sci* 13:1773–1778
- von Storch H (1995) Misuses of statistical analysis in climate research. *Analysis of Climate Variability: Applications of Statistical Techniques*, ed H.v.Storch, A.Navarra, Springer-Verlag, New York: 11–26
- Wang Y-G, Wang SSJ, Dunlop J (2015) Statistical modelling and power analysis for detecting trends in total suspended sediment loads. *J Hydrol* 520:439–447
- Yue S, Pilon P (2004) A comparison of the power of the t-test, Mann-Kendall and bootstrap tests for trend detection. *Hydrol Sci J* 49(1):21–37
- Yue S, Wang CY (2002) Applicability of prewhitening to eliminate the influence of serial-correlation on the Mann-Kendall test. *Water Resour. Res.*, 38(6): 4-1-4-7
- Yue S, Wang C (2004) The Mann-Kendall test modified by effective sample size to detect trend in serially correlated hydrological series. *Water Resour Manag* 18:201–218
- Yue S, Pilon P, Cavadias G (2002a) Power of the Mann-Kendall and Spearman's rho tests for detecting monotonic trends in hydrological series. *J Hydrol* 259:254–271
- Yue S, Pilon P, Phinney B, Cavadias G (2002b) The influence of autocorrelation on the ability to detect trend in hydrological series. *Hydrol Processes* 16:1807–1829
- Yue S, Pilon P, Phinney B (2003) Canadian streamflow trend detection: impacts of serial and cross-correlation. *Hydrol Sci J* 48(1):51–63
- Zhang X, Zwiers FN (2004) Comment on “Application of prewhitening to eliminate the influence of serial correlation on the Mann-Kendall test” by S.Yue and C.Y. Wang. *Water Resour. Res.*, 40: W08801