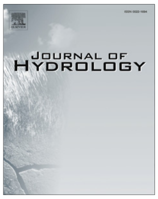
Journal of Hydrology 545 (2017) 197–211

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol



Research papers

Comparison of methods for non-stationary hydrologic frequency analysis: Case study using annual maximum daily precipitation in Taiwan

Po-Chun Chen[a], Yuan-Heng Wang[a], Gene Jiing-Yun You[a,*], Chih-Chiang Wei[b]

[a] Department of Civil Engineering, National Taiwan University, Taipei 10617, Taiwan
[b] Department of Marine Environmental Informatics, National Taiwan Ocean University, Keelung, Taiwan

ARTICLE INFO

Article history:
Received 7 April 2016
Received in revised form 12 November 2016
Accepted 2 December 2016
Available online 19 December 2016
This manuscript was handled by A. Bardossy, Editor-in-Chief, with the assistance of Felix Frances, Associate Editor

Keywords:
Non-stationarity
Identification of distribution and trends
Discrete wavelet analysis
Ensemble empirical mode decomposition
Return period

ABSTRACT

Future climatic conditions likely will not satisfy stationarity assumption. To address this concern, this study applied three methods to analyze non-stationarity in hydrologic conditions. Based on the principle of identifying distribution and trends (IDT) with time-varying moments, we employed the parametric weighted least squares (WLS) estimation in conjunction with the non-parametric discrete wavelet transform (DWT) and ensemble empirical mode decomposition (EEMD). Our aim was to evaluate the applicability of non-parameter approaches, compared with traditional parameter-based methods. In contrast to most previous studies, which analyzed the non-stationarity of first moments, we incorporated second-moment analysis. Through the estimation of long-term risk, we were able to examine the behavior of return periods under two different definitions: the reciprocal of the exceedance probability of occurrence and the expected recurrence time. The proposed framework represents an improvement over stationary frequency analysis for the design of hydraulic systems. A case study was performed using precipitation data from major climate stations in Taiwan to evaluate the non-stationarity of annual maximum daily precipitation. The results demonstrate the applicability of these three methods in the identification of non-stationarity. For most cases, no significant differences were observed with regard to the trends identified using WLS, DWT, and EEMD. According to the results, a linear model should be able to capture time-variance in either the first or second moment while parabolic trends should be used with caution due to their characteristic rapid increases. It is also observed that local variations in precipitation tend to be overemphasized by DWT and EEMD. The two definitions provided for the concept of return period allows for ambiguous interpretation. With the consideration of non-stationarity, the return period is relatively small under the definition of expected recurrence time comparing to the estimation using the reciprocal of the exceedance probability of occurrence. However, the calculation of expected recurrence time is based on the assumption of perfect knowledge of long-term risk, which involves high uncertainty. When the risk is decreasing with time, the expected recurrence time will lead to the divergence of return period and make this definition inapplicable for engineering purposes.

1. Introduction

Changes in the global climate have led to increases in the frequency and intensity of extreme hydrological events. Records of extreme weather conditions are being broken every year and the number of disasters is increasing. In facing these changes, it is crucial to utilize the information derived from hydrological time series with a sound understanding as to its interpretation. In the past, water resource engineering was based on the characterization of future events under the assumption that the future will resemble the past, and the past can be accurately represented using a sample of observations drawn based on the same physical process from which the future will be generated. In hydrologic frequency analysis, time invariance is known as stationarity, and forms the basis of many statistical methodologies. The commonly used concept, return period, is one such approach. However, climate change has raised doubts regarding this assumption. Acceleration of the changes in hydrologic cycles are now being anticipated according to theoretical understanding, predicted using climate models, and observed in hydro-meteorological data including precipitation

---

* Corresponding author.
E-mail address: genejyu@ntu.edu.tw (G.J.-Y. You).

http://dx.doi.org/10.1016/j.jhydrol.2016.12.001

and streamflow (Katz et al., 2002). This has made it necessary to revisit the concept of stationarity to verify its current relevance.

Stationarity assumes that hydrologic conditions are stationary and lacking long-term trends. Many studies have demonstrated that future climatic conditions are very likely not to satisfy this assumption (Sveinsson et al., 2003; Milly et al., 2008;). Wilks (1992) adapted stochastic weather generation models in an investigation of climate change by generating synthetic daily time series. He used observed interannual variability as an analog for climate change and constructed possible future scenarios. Strupczewski et al. (2001) applied a similar approach, using the maximum likelihood method to estimate time-dependent parameters for annual maximum flood series. Based on the simulation results of Atmosphere-Ocean General Circulation Models (AOGCMs), Katz and Brown (1992) employed extreme value theory to determine the frequency of extreme events. They performed sensitivity analysis of climate variability and suggested that extreme events are more sensitive to changes in standard deviation than to changes in the mean of distribution. As for the impact of climate change on precipitation, Fowler and Hennessy (1995) used physical and empirical arguments as well as AOGCMs to verify an increase in the frequency and intensity of extreme rainfall events around the world. They discussed the impact of climate change on the occurrence of extreme hydrological events leading to a conviction that the frequency and scale of floods or droughts will increase significantly in the future. Milly et al. (2002) applied a similar approach and also determined that climate change will intensify the global water cycle resulting in an increased risk of flooding. Efforts to verify the non-stationarity of climatic conditions has not been limited to theoretical or modeling studies but also from data observation; evidence of an increasing trend in flood peaks has also been identified in the time series of floods using statistical methods (Burn, 1998; Steel, 1998; Aliev and Vishnevskyi, 1998).

AOGCMs have been widely adopted as a means of representing climatic conditions in the future, and coupled AOGCMs form the basis of the vast majority of impact studies. Nonetheless, the reliability of AOGCM projections remains an issue of debate (Randall and Wielicki, 1997; Shackley et al., 1998; Henderson-Sellers and McGuffie, 1999; Petersen, 2000). Despite the fact that the simulation of AOGCMs provides a sophisticated means of predicting future climatic conditions, a degree of uncertainty always remains (Dooge et al., 1998). Even using the same emission scenario, different AOGCM climate projections often produce different results. Moreover, current AOGCMs lack the ability to reproduce current conditions on a regional or catchment scale, differing not only in quantity but sometimes also in sign. AOGCM predictions of precipitation are even less certain than those of temperature, rendering them largely inapplicable for engineering purpose (Miller and Russell, 1992; Strupczewski et al., 2001).

In water resource engineering, quality methods for non-stationary frequency analysis could prove more valuable than climate models. The non-stationarity of hydrologic events is particularly important in light of concerns over climate change. This phenomenon is commonly associated with the presence of a trend component, linear or non-linear, in the statistical characteristics of data. The presence of any trend can have a considerable effect on the interpretation of results when fitting a probability distribution to a sample of non-stationary observations.

Techniques for the probability distribution fitting of non-stationary data date back to the early 1900s by Cave and Pearson (1914). Non-stationary frequency analysis can be based on stationary frequency analysis, assuming the same probability distribution functions with a consideration only of the time variance of parameters. This simplifies problems for the estimation of parametric trends over time. Tawn (1988) proposed a multivariate extreme value method combining a parametric model with temporal dependence function. Strupczewski et al. (2001) conducted a complete investigation on the trends involved in first two moments of a probability distribution function in either linear or parabolic form. They performed maximum likelihood estimation (ML) and weighted least squares estimation (WLS) to obtain these trends, using streamflow in Poland as a case study. Their results demonstrate that the WLS method is in agreement with the maximum likelihood method in cases of normal distribution. However, in cases when the data are not normally distributed, previous researchers recommend the use of WLS, rather than ML to have a better estimation (Olsson et al., 2000).

Cox et al. (2002) employed the probability theory of extreme values to investigate changes in the maxima and return period in terms of the effects of trends in the mean level and trends in dispersion. Katz et al. (2002) used a similar approach to analyze non-stationarity in hydrologic extremes, particularly in floods. Cunderlik and Burn (2003) introduced a second-order non-stationary approach to regional flood frequency analysis. Villarini et al. (2009) proposed a framework for flood frequency analysis based on the Generalized Additive Models for Location, Scale, and Shape parameters (GAMLSS). Their aim was to investigate annual maximum flood peaks in urban river basins with respect to time as well as other covariates related to urbanization and changing climate. Their study demonstrated an increase in flood magnitudes during periods of rapid urbanization and increases in population. Although many researches have employed linear trends in parameter values to model non-stationarity, doubts remain as to whether these trends can be simplified into linear or simple functions (Hall and Tajvidi, 2000; Davison and Ramesh, 2000; Ramesh and Davison, 2002).

In addition to schemes based on the annual maximum time-varying moment, r-largest, peaks-over-threshold, and point process models are the alternatives for non-stationary analysis (Zhang et al., 2004; Rootzén and Tajvidi, 1997; McNeil and Saladin, 2000; Katz et al., 2002; Khaliq et al., 2006; Alexandrov et al., 2012). These sophisticated models have higher data requirement in additional to annual maxima due to complex theoretical structure, but make it possible to obtain accurate estimations of parameters and quantiles. Besides, incorporating teleconnection indices into frequency models as covariates also has been shown improve the statistical modeling of extreme events (Villarini et al., 2009; Cox et al., 2002; Katz et al., 2002; Vasiliades et al., 2015; Li et al., 2015). These covariates can be obtained using climate observation, reanalysis data or the outputs of global circulation models to assess the impact of climate change. Nonetheless, the computational demands of these models, even their stationary counterparts, hinder their implementation in practical engineering applications.

Artificial Neural Networks (ANNs), which were originally developed for pattern recognition in the cognitive sciences, are commonly used in hydrological forecasting. The ability of an ANN to cope with missing data and to "learn" from current forecasting cases in real time makes it an appealing alternative to conventional lumped or semi-distributed forecasting models. In some studies, ANNs have been combined with external covariate climatic indices (Wu et al., 2009; Chen et al., 2015; Gholami et al., 2015). However, this approach regards processes as a black box system, and therefore lacks explicit expressions for subsequent interpretation. Further research is required to determine the optimum ANN training period for given hydrological contexts (Dawson and Wilby, 2001).

Previous studies have used time-dependent parameters that are numerically estimated by ML and WLS (Strupczewski et al., 2001). Comparing to these parametric method, non-parametric analysis is an alternative to parametric non-stationary analysis, providing greater flexibility in the identification of trends. Wavelet analysis is