Journal of Hydrology 519 (2014) 2040–2048

Contents lists available at ScienceDirect

Journal of Hydrology

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

A nonstationary index-flood technique for estimating extreme quantiles for annual maximum streamflow

Nicole L. O'Brien *, Donald H. Burn [1]

Department of Civil and Environmental Engineering, University of Waterloo, Waterloo, ON N2L 3G1, Canada

ARTICLE INFO

Article history:
Received 19 April 2014
Received in revised form 22 August 2014
Accepted 13 September 2014
Available online 7 October 2014
This manuscript was handled by Andras Bardossy, Editor-in-Chief, with the assistance of Attilio Castellarin, Associate Editor

Keywords:
Annual maximum flood
Nonstationarity
Pooled flood frequency analysis
Mann-Kendall test
Maximum likelihood

SUMMARY

The magnitude and timing of peak streamflow events may be affected by land-use changes along with climate change, thus leading to nonstationarity in the records. Temporal trend, along with change-points, in peak flow records can affect the accuracy of quantile estimates; therefore, these issues should not be disregarded. Commonly used techniques for pooled flood frequency analysis do not account for nonstationarity found in the data recorded for members of a region. To overcome this shortcoming, the objective of this research is to introduce a trend centered pooling approach for regionalization in which pooling groups are created based on the form of trend found in the at-site data. The approach involves the formation of regions comprised entirely of sites exhibiting either statistically significant increasing or decreasing trends. Regional parameter estimates are determined using a maximum likelihood approach, which is carried out with the assumption of second-order nonstationarity. The technique was applied to four homogenous regions all located in differing hydroclimatological Canadian regions. The uncertainty of quantile estimates calculated through the implementation of this technique was established using a balanced regional vector resampling approach. The results indicate that there is less uncertainty in quantile estimates found through the application of the trend centered pooling approach when compared to a regional stationary analysis of the same regions. The potential for overestimation/underestimation of design quantiles in the presence of significant regional nonstationarity (i.e. decreasing/increasing trends) was elucidated.

1. Introduction

The recent 2013 flooding in Alberta, Canada and the 2013 Colorado, United States floods have sparked much discussion concerning adequate planning and protection from extreme hydrological events. Concern over flooding arises due to the substantial socioeconomic risks associated with these events. Research concerning the magnitude and timing of extreme hydrological events suggests that land-use changes along with climate change may lead to nonstationarity in peak streamflow records. Stationarity can be interpreted as a form of statistical equilibrium; therefore, the statistical properties of the process in question would not be time-dependent (Hipel and McLeod, 1994). As the global population increases, along with possible increases in the mean global temperature, the need for more accurate flood frequency techniques is of paramount concern in a nonstationary environment. Intensification of the hydrologic cycle is anticipated as the mean global temperature rises, which can affect the timing and magnitude of flooding events (Trenberth et al., 2007). In colder climates, the onset of spring snow-melt runoff is expected to shift from early spring to late winter thereby decreasing total runoff amounts while warmer climates may see an increase in total rainfall (Kundzewicz et al., 2010). In addition to these projected climate-change responses, land-use changes can have marked effects on the hydrologic characteristics of a watershed. The 20th century has been characterized by intense land-use changes with respect to agricultural practices, urbanization, and forest management. These changes can cause shifts in hydrological and ecological systems and impact the rainfall–runoff relationship, thus affecting flood risk (Villarini et al., 2009a). In particular, the addition of impervious surface can lead to increase flood amplitudes and decreased time-to-peak of flooding events. Evidence of these variations in runoff has been detected in peak flow data worldwide in the form of temporal trend (Robson and Reed, 1999; Jain and Lall, 2000; Zhang et al., 2001; Burn and Hag Elnur, 2002; Hodgkins et al., 2003; Burn and Cunderlik, 2003; Burn et al., 2010;

* Corresponding author. Tel.: +1 519 888 4567x33828; fax: +1 519 888 6197.
E-mail addresses: nobrien@uwaterloo.ca (N.L. O'Brien), dhburn@uwaterloo.ca (D.H. Burn).
[1] Tel.: +1 519 888 4567x33338; fax: +1 519 888 6197.

http://dx.doi.org/10.1016/j.jhydrol.2014.09.041
0022-1694/

Schmocker-Fackel and Naef, 2010; Bormann et al., 2011; Hirsch, 2011).

Modelling nonstationarity in extreme hydrometeorological data is a topic that has received a great deal of attention in recent years (Kharin and Zwiers, 2005; Khaliq et al., 2006; El Adlouni et al., 2007; Kyselý et al., 2010; Westra et al., 2013). This may be due to the potential for systematic errors in quantile estimates in the presence of nonstationarity. For example, an increasing trend in the mean of a flow series could lead to underestimation of flood quantiles and increase the risk of failure of engineering structures. Decreasing trends could lead to overestimation of the design quantiles, resulting in increased costs for overdesigned infrastructure. There is, however, less research with respect to the implications of nonstationarity in pooled flood frequency analysis (FFA). Using pooled FFA allows for the incorporation of data from several sites, thus allowing for greater accuracy in estimating peak flows. Standard techniques used in pooled FFA were developed for independently and identically distributed (IID) data, meaning data are independent and belong to the same statistical probability distribution. In the presence of nonstationarity, the IID assumption is no longer valid and results obtained using standard methodologies may be inaccurate. There is therefore a pressing need for the development of more accurate pooled FFA techniques when time-dependence is exhibited in hydrological data.

Smith (1989) is an early reference that accounts for nonstationarity in a pooled frequency analysis. This study used IID data, which were not a function of time but of catchment characteristics. More recently, Cunderlik and Burn (2003) proposed a novel detrending approach for FFA for nonstationary data. Cunderlik and Ouarda (2006) developed a nonstationary approach to regional flood-duration-frequency (Qdf) modelling using the index-flood method. Hanel et al. (2009) propose a nonstationary index-flood (IF) equation in which the pooled growth curve and index-flood both vary with time. The authors apply their methodology to the Rhine basin where regions are determined subjectively. The IF approach develop by Hanel et al. (2009) is also used by Roth et al. (2012) but transformed for the use of a peaks-over-threshold (POT) model with nonstationary parameters. Their methodology is applied to the Netherlands, which has been determined to be a suitably homogeneous region in previous studies. Other work has focused on modelling nonstationarity in regional frequency analysis using a Bayesian framework (Leclerc and Ouarda, 2007; Renard et al., 2006).

The central focus of this paper is the development of a methodology for pooling peak streamflow data based on the form of trend detected (i.e., either increasing or decreasing trends). The standard techniques used in regional frequency analysis have been developed for data that are not temporally dependent. Therefore, if a region has members displaying nonstationarity, systematic errors may be introduced into the analysis. This becomes particularly important as more recent data are collected, resulting in an increasing number of statistically significant trends in peak annual flow data in Canadian watersheds and in catchments worldwide (Bormann et al., 2011; Burn and Cunderlik, 2003; Villarini et al., 2009b). The trend centered pooling technique involves testing applicable data for trends using the Mann–Kendall test. Sites that display temporal trend in the mean tendency are then used for the creation of homogeneous regions. This methodology is employed under the premise of second-order nonstationarity, which results in the formation of regions comprised solely of sites exhibiting significant trends, on which the remainder of the analysis is focused.

Section 2 outlines the methodology used in the analysis, including parameter estimation techniques, goodness-of-fit for selected distributions and the uncertainty associated with the use of this technique. Section 3 discusses the application of this nonstationary index-flood procedure, while Section 4 discusses the results of the analysis.

2. Methodology

This section outlines the methodology used with respect to the trend centered pooling technique. The primary goal of this section is to describe the nonstationary index-flood model, the regionalization technique implemented and the parameter estimation methodology used for this research.

2.1. Data screening and trend detection

The proposed regionalization technique makes use of statistically significant trends found in individual site data. Trend detection was carried out using the non-parametric Mann–Kendall test, which is a rank-based test commonly used for analysis of hydrometric variables. The Mann–Kendall trend statistic allows for the determination of the significance of a trend found in the data (Mann, 1945; Kendall, 1975). Significant serial correlation in at-site data can have deleterious effects on the robustness of trend detection tests; therefore autocorrelation was addressed using trend free pre-whitening (TFPW). Following Yue et al. (2002), TFPW consist of fitting a linear trend to the time series, data are then pre-whitened and the final result involves blending the monotonic trend and the pre-whitened residual series. The use of this technique allows for the removal of serial correlation while maintaining any trend present in the data.

The power of the Mann–Kendall test can also be affected by change-points or shifts in the data. Therefore, the homogeneity of the data was assessed using Bayesian Change Point Analysis (BCPA), proposed by Barry and Hartigan (1992, 1993). BCPA is a parametric test allowing for the detection of multiple change-points in a time series. Regime shift testing is an important precursor for trend detection testing as change-points in the data may lead to the detection of trends when none exist.

2.2. Index flood model

The index-flood technique (Dalrymple, 1960) provides a framework for determining design flood estimates that continues to be applied in a number of settings (Robson and Reed, 1999; Hanel et al., 2009; Roth et al., 2012; Ilorme and Griffis, 2013; Norbet et al., 2014; Wright et al., 2014). This method assumes that within an acceptably homogenous region, the flood response from all members is identically distributed aside from a site-specific scaling factor (i.e., the index flood). This model has been modified herein to include the nonstationarity of the data and can be described by the following model:

$$Q_i(F,t) = \xi_i q(F,t) \tag{1}$$

where $Q_i(F, t)$ is the time-dependent flood quantile of site i, ξ_i is the index-flood of site i, and $q(F, t)$ is the nonstationary pooled growth curve. The index-flood is taken as the mean of the at-site data. Under the assumption of second-order nonstationarity, distributional parameters can be modeled as a function of time, which can be expressed most generally as:

$$\begin{aligned} \mu_R(t) &= \mu_0 + \mu_1 t + \mu_2 t^2 + \cdots + \mu_m t^m \\ \alpha_R(t) &= \sigma_0 + \sigma_1 t + \sigma_2 t^2 + \cdots + \sigma_m t^m \\ k_R(t) &= k \end{aligned} \tag{2}$$

where $\mu_R(t)$ and $\alpha_R(t)$ are the time-dependent regional location and scale parameters, μ_i and σ_i are a collection of m regional parameters, and $k_R(t)$ is the time-invariant shape parameter. When $m = 1$, Eq. (2) reduces to a linear model. Although it is possible to model