Analysis of annual flood peak records in Mexico

C. Escalante-Sandoval & E. Garcia-Espinoza
Faculty of Engineering,
National Autonomous University of Mexico, Mexico

Abstract

Mexico suffers the occurrence of floods and droughts. Between 1980 and 2013, floods caused 2,523 fatalities and more than 26 billion dollars of economic damage. The probabilistic assessment of flood risk is made in the form of a flood frequency analysis, which requires the assumption of the stationarity of annual flood peak records. Nonparametric tests were selected to perform homogeneity analysis and to examine the presence of monotonically increasing/decreasing patterns in the data. The 44% of the samples analyzed were considered as non-stationary and the most important changes were detected during the period 1971–1980, where the maximum average annual loss of forest and jungles was experienced. The time of occurrence analysis of the 10 largest flood peaks reveals that in north-western Mexico, the highest values are concentrated in the December-January period. In Southern Mexico, the highest values occur in the October-December period, and for the rest of the country the highest values concentrate in the August-September period. When stationarity cannot be assured the results of traditional flood frequency analysis could cause fatalities and economic losses, so it is recommended that the National Water Commission take into consideration the theory of non-stationary frequency analysis.
Keywords: stationarity series, flooding, homogeneity test, trend test, flood frequency analysis.

1 Introduction

Mexico extends over a surface area of 1,964 million km^2. The 1,471 catchments identified in the country have been organized into 37 hydrological regions, which are in turn grouped into 13 Hydrological-Administrative Regions.

Mexico's rivers and streams constitute a hydrographic network of 633,000 kilometers. Seven rivers are responsible for 65% of the country's surface runoff: Balsas, Coatzacoalcos, Grijalva-Usumacinta, Pánuco, Papaloapan, Santiago, and Tonalá. The surface area of their watersheds represents 22% of the surface of Mexico. The Balsas and Santiago rivers flow into the Pacific Ocean and the other five flow into the Gulf of Mexico.

Due to its great diversity of climatic and geographical conditions, Mexico is exposed on a daily basis to severe hydro-meteorological events, which, even if they do contribute positively to the increase in water storage in dams and lakes, also cause damage to the population, to infrastructure, to services and production systems. Between 1980 and 2007, floods affected more than 8 million people and caused 2017 fatalities and more than 14.7 billion dollar of economic damage. In this period, cyclones Stan and Isidore were those that affected the largest number of people, whereas Emily, Stan and Gilbert caused the greatest economic damage. Moreover, the events occurred between 2008 and 2009 caused 150 fatalities and 2 billion dollar of economic damage.

The hurricanes that occurred in 2010, such as Alex, Karl and Mathew, mostly affected the states of Campeche, Coahuila, Chiapas, Nuevo León, Oaxaca, Puebla, Tamaulipas and Veracruz. The economic losses that they caused amounted to 7.5 billion dollar, and the death toll was 199.

In September 2013, at least 157 people died in Mexico by the simultaneous occurrence of two powerful storms, one on its east coast (Hurricane Ingrid) and another on its west coast (Tropical storm Manuel). The economic losses amounted to 1.8 billion dollar.

The damages associated with floods are increasingly more serious due to the location of irregular human settlements in areas close to rivers, the lack of enforcement of land use regulations, and deforestation in the upper portions of river basins, the effect of which is reflected in an increase in runoff and the transportation of soils and sediments towards the lower portions.

Flood damage reduction consists in adopting structural and non-structural measurements. Structural methods imply designing works such as levees, floodwalls, dams, dredging and channelization. Moreover, non-structural measures can be the relocation, flood proofing, buy-out or acquisition, and flood preparedness. In particular, the structural measurements require risk analysis of floods as design criteria. The probabilistic assessment of flood risk is made in the form of a flood frequency analysis.

The assumption of stationarity of annual flood peak records is very important in risk analysis for flood hazards. The validity of such assumption has been challenged due to the natural variability and anthropogenic induced changes in the climate system [1].

The annual flood peak time series is stationary if it is free of periodicities and abrupt changes [2]. In flood frequency analysis when stationarity cannot be assured, the estimated event for design purposes could lead to catastrophic consequences in case of an underestimation and economic losses in case of an overestimation. Ouarda and El-Adlouni [3] provide a broad discussion of non-stationary frequency analysis in hydrology.

As already mentioned, flood events are responsible for numerous fatalities and large economic damage, so a detailed analysis of the flood peak records will allow a better understanding of that possible temporal change.

This study focuses on investigating the presence of temporal non stationarities in the flood peak records. To address this problem, we use annual maximum flood peak records from 205 gauging stations with a record of at least 50 years over the entire Mexican territory.

The validity of the stationarity assumption is checked in terms of both abrupt changes in the mean and variance of the flood peak distribution and monotonic trends. Some nonparametric tests were used to perform homogeneity analysis for the detection of abrupt changes in the mean and variance of the flood peak distribution, and to examine the presence of monotonically increasing/decreasing patterns in the data. For this purpose, we adopted the methodology suggested by Villarini *et al.* [4], where corrections to account for serial correlation or long-term persistence are not considered.

2 Data and methods

2.1 Data

The National Water Commission has available information from 990 gauging stations located all over Mexican territory. However, in this study we only consider 205 stations with a record of at least 50 years of annual maximum instantaneous peak discharge as shown in Figure 1.

Figure 1: Location of the 205 gauging stations included in this study.

2.2 Change-point analysis

Four statistical tests were used to test homogeneity of the mean: the Standard normal homogeneity test [5], the Buishand range test [6], the Pettitt test [7], and the Von Neumann ratio [8] were selected. Moreover, the F-test [9], the Levene

test [10], the Brown-Forsythe test [11], and the Bartlett test [12] are used to verify the assumption of homoscedasticity or homogeneity of variances. Both mean and variance homogeneity tests were assessed as follows:

The series that rejects one or none null hypothesis under the four tests at 5% significance level are considered to indicate homogeneity.

The series that reject two null hypotheses of the four tests at 5% significance level indicate series that have the inhomogeneous signal and should be critically inspected before further analysis.

When there were three or all tests rejecting the null hypothesis at 5% significance level they were considered as indicating non homogeneous conditions.

2.3 Trend analysis

The presence of monotonic trends was examined by means of two nonparametric monotonic trend tests, Mann-Kendall and Spearman rho test [13]. For a given record, the tests were performed on the entire record if series was considered homogeneous. If a change-point in mean was detected, the analysis was performed on the two subseries (before and after the change-point)

3   Results and discussion

The change-point analysis in mean and variance was performed using the statistical tests aforementioned. Out of 205 stations, 41% of the stations exhibiting a change point in mean, 45% in variance, and 41% a change-point both in mean and variance. Maps with stations showing a change-point are presented in Figures 2 and 3.
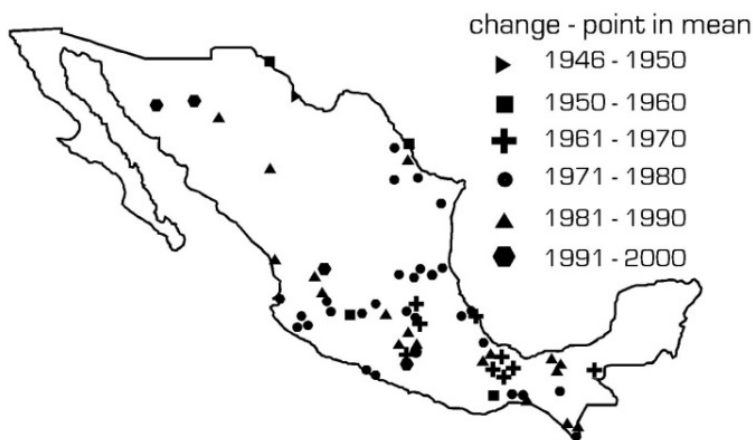
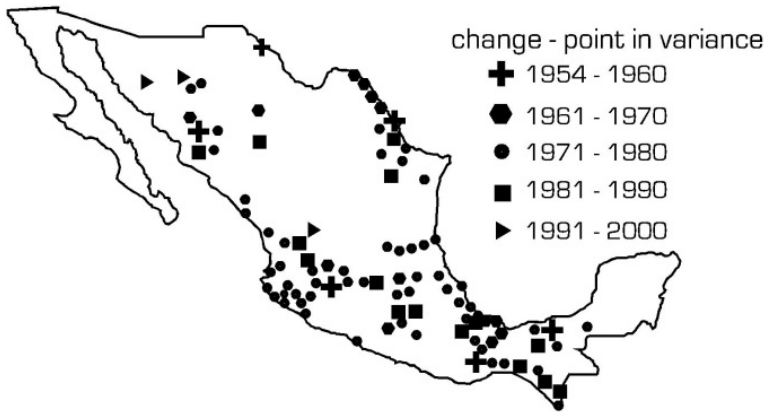

Figure 2: Stations with a change-point in mean.

Figure 3:   Stations with a change-point in variance.

The most important changes were detected during the period 1971–1980. These change-points can be related to the accelerated process of deforestation experienced in the past decades. The main causes for the disappearance in Mexico of forests are attributed to the increased area devoted to large-scale agriculture and cattle-raising, particularly in tropical and sub-tropical regions. The decadal average annual loss of forest and jungles is listed in Table 1.

Table 1: Decadal average annual loss of forest and jungles in Mexico.

Decade	Loss (ha)
1951–1960	400,000
1961–1970	600,000
1971–1980	700,000
1981–1990	400,000
1991–2000	300,000
2001–2010	200,000

The maximum average annual rate was experienced in the period 1971–1980. In the last years, Mexico has been reducing the deforestation rate applying its plans to conserve and restore forests and jungles.

After this analysis, 121 stations could be considered homogeneous in mean and variance, as shown in Figure 4. Most of the inhomogeneous stations are located at three regions well defined. The first of them is located at the central portion of Mexico, and is called "The Lerma Chapala basin". This is a strategic region for the country development. Nevertheless, the dynamics of its socioeconomic growth during the last decades has triggered severe water resources problems and conflicts as well as environmental degradation and deterioration. The basin comprises portions of 5 Mexican states: Jalisco,

Guanajuato, Michoacan, Queretaro and Mexico. The second one is located in northern Mexico, and is called "The Rio Bravo basin", and covers three US states and four Mexican states: Chihuahua, Coahuila, Nuevo Leon and Tamaulipas. In this basin, the rapid population growth and consequent economic development and land-use changes are pushing the limits of environmental sustainability and quality. During the last century, a number of large reservoirs were built on the Río Bravo and its major tributaries and extensive well fields were drilled in the basin's aquifers. The dams have greatly reduced the downstream flow of the main stem and its tributaries. The last region is located in southern Mexico, and covers portions of three Mexican States: Veracruz, Tabasco and Chiapas. This region has the biggest regional quantity of renewable water resources, a relatively low population and contribution to the GDP. The severe land-use changes in this region are attributed to the increased area devoted to large-scale agriculture and cattle-raising.

Trend analysis is performed on the entire time series if there was not statistically significant change point in mean (there were 121 stations fulfilling this requirement). From this set of stations, 4 show a negative trend and only 2 a positive trend, as shown in Figure 5. Moreover, if a change-point in mean was detected, analysis was performed on the two subseries, before and after the change-point. In this case (84 stations), 10 stations showed positive trend before change-point as shown in Figure 6, and 15 indicate negative trend after change-point as shown in Figure 7. In Most cases, the change-points rather than monotonic trends were responsible for most violations of the stationarity assumption.
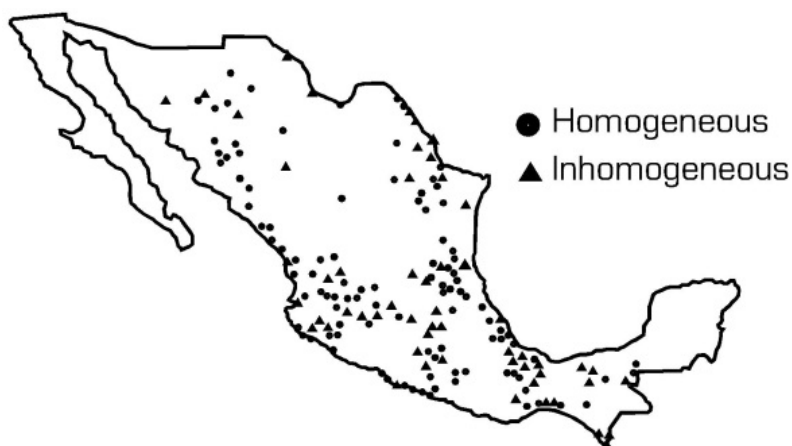

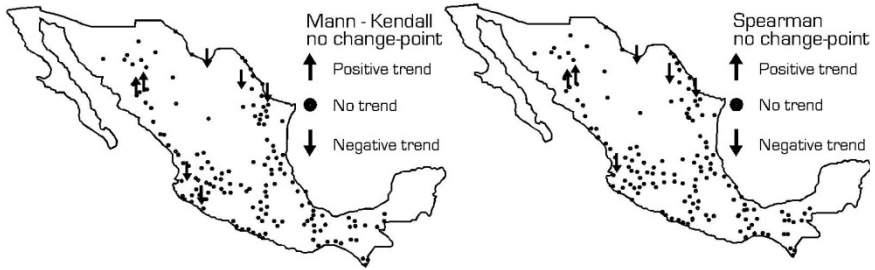Figure 4: Map showing the final results of the homogeneity tests.

Figure 5:   Results of the Mann-Kendall test and Spearman test for the series without change-point in mean.
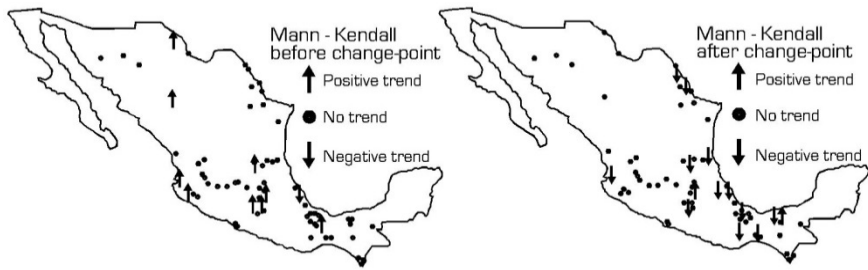

Figure 6:   Results of the Mann-Kendall test before and after the change-point in mean.
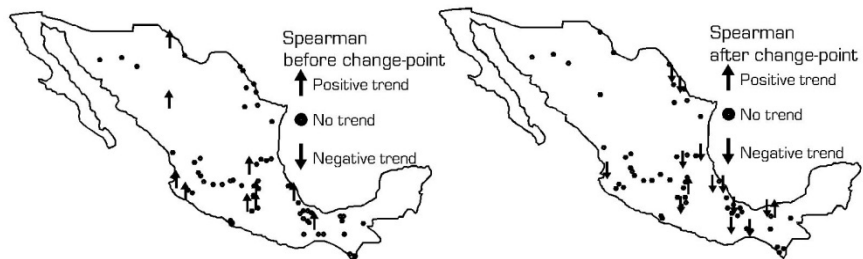

Figure 7:   Results of the Spearman test before and after the change-point in mean.

The probabilistic assessment of flood risk was made in the form of a flood frequency analysis. This analysis assumes that data are independent and identically distributed, which implies that it must meet the statistical criteria of independence, stationarity, and homogeneity (There were 115 stations fulfilling this requirement). For this purpose, unimodal and mixed distributions were used. The selection of the probability distribution that best fits the sample data is based on a goodness of-fit criterion [17]. The best fits are distributed according to the percentages showed in Table 2. The spatial distribution of these fits is presented in Figure 8.

Table 2: Percentage of occurrence of the best distribution of fit.

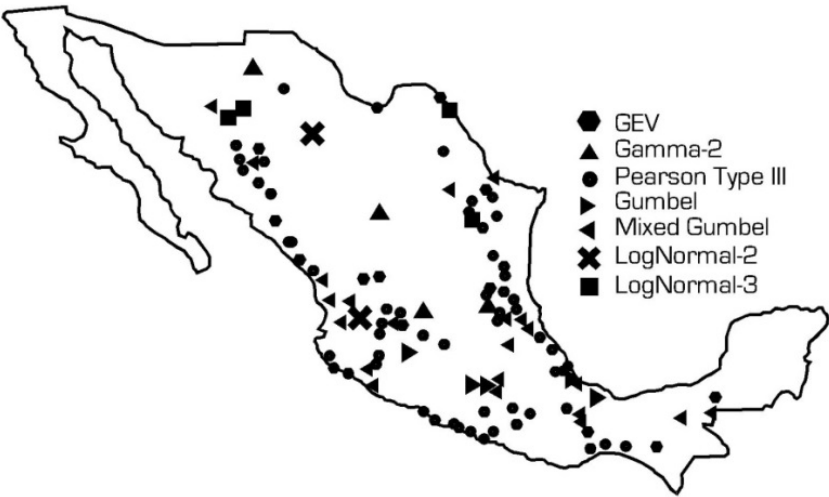| Distribution	%
Pearson Type III	55
Mixed Gumbel	21
Generalized Extreme Value	11
Gumbel	4
Gamma 2 parameters	3
Lognormal 3 parameters	3
Lognormal 2 parameters	3

Figure 8: Location of each gauging station, each associated with its best probability distribution of fit.

Now, we are able to estimate the quantiles associated to different return periods, which are required in flood risk management. However, it also was necessary to know the time of occurrence of the flood peak. To accomplish this objective, the percentage of occurrence of the 10 largest peaks from January to December was plotted as shown in Figures 9 and 10. In Northwestern Mexico, the highest values are concentrated in the December-January period. In Southern Mexico, the highest values occur in the October-December period, and for the rest of the country the highest values concentrate in the August-September period.
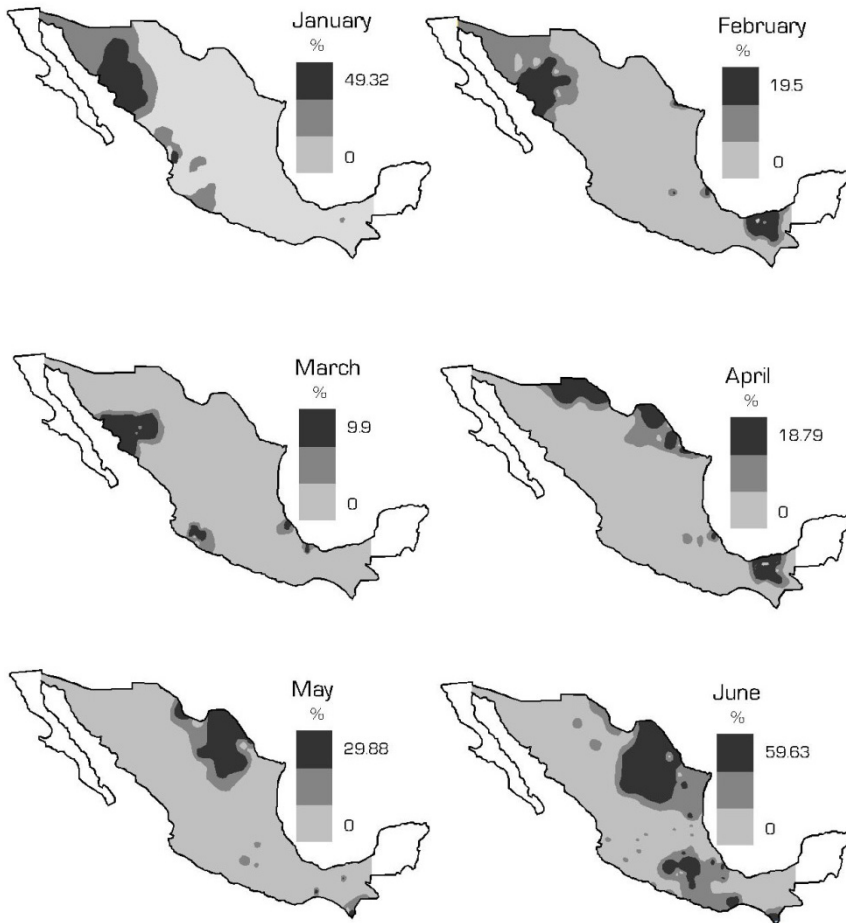

Figure 9:   Percentage of occurrence of the 10 largest flood peaks during the January–June period.
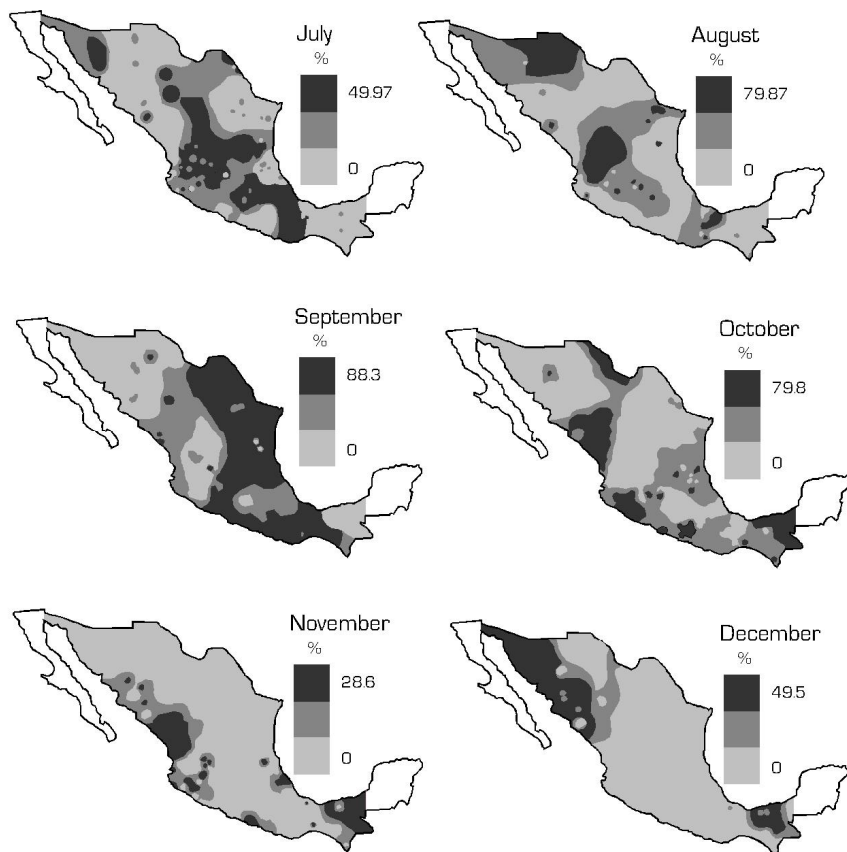
Figure 10: Percentage of occurrence of the 10 largest flood peaks during the July–December period.

4 Concluding remarks

Mexico is periodically subject to meteorological vents that cause serious damage in large portions of its territory, and it is expected they will be even more frequent and intense due to the effects associated with climate change.

As the drainage basin scale, consideration of flood risk plays an important role in planning of water infrastructure project. The probabilistic assessment of flood risk is made in the form of a flood frequency analysis, which requires the assumption of stationarity of annual flood peak records.

Four nonparametric tests were used to perform homogeneity analysis for the detection of abrupt changes in the mean and another four for the variance. Moreover, two additional nonparametric tests were applied to examine the

presence of monotonically increasing/decreasing patterns in the data. Out of 205 stations, 41% of the stations exhibit a change point in mean, 45% in variance, and 41% a change-point both in mean and variance.

The 44% of the analyzed samples were considered non stationaries and the most important changes are detected during the period 1971-1980. In the vast majority of the cases, violations on stationarity assumption are associated with change-points rather than monotonic trends. These non stationarities could be associated with anthropogenic effects, such as the population growth, unordered urban settlements, the deforestation and degradation of catchments.

From the total of 205 stations, 115 did not present statistically significant change-points and trend. We describe its distribution in terms of its cumulative distribution function. The 55% of the samples were best fitted by the Pearson Type III distribution, followed by the Mixed Gumbel distribution (21%) and the Generalized Extreme Value Distribution (11%).

The time of occurrence analysis of the 10 largest flood peaks reveals that in north-western Mexico, the highest values are concentrated in the December–January period. In Southern Mexico, the highest values occur in the October–December period, and for the rest of the country the highest values concentrate in the August–September period.

When stationarity cannot be assured, the results of traditional flood frequency analysis produce representations that might cause fatalities and economic losses, so it is recommended that Mexican authorities take in consideration the non-stationary frequency analysis theory.

With this kind of analysis, authorities can be able to substantially improve the emergency plans for rivers that may cause damage by overflowing. As a complement to this, the reforestation program in the upper parts of river basins and eroded areas must be strengthen. This will contribute to reducing possible floods in the lower parts of river basins.

References

[1]   P. Milly, J. Betancourt, M. Falkenmark,R. Hirsh, Z. Kundzewicz, D. Lettenmaier and R. Stouffer, Stationarity Is Dead: Whiter Water Management? *Science*, **319**, pp. 573-574, 2008.

[2]   N. Matalas, Stochastic Hydrology in the Context of Climate Change, Climate Change, **37**, pp. 89-101, 1997.

[3]   T. Ouarda and S. El-Adlouni, Bayesian Non-stationary Frequency Analysis of Hydrological Variables, *Journal of the American Water Resources Association*, **47(3)**, pp. 496-505, 2011.

[4]   G. Villarini, J. Smith, M. Baeck, and W. Krajewski, Examining Flood Frequency Distributions in the Midwest US, Journal of the American Water Resources Association, **47(3)**, pp. 447-463, 2011.

[5]   H. Alexandersson, A Homogeneity Test Applied to Precipitation Data, *Journal of Climate*, **6**, pp. 661-675, 1986.

[6]   T. Buishand, Some Methods for Testing the Homogeneity of Rainfall Record, *Journal of Hydrology*, **58**, pp. 11-27, 1982.

[7]   A. Pettitt, A Non-parametric Approach to the Change-point Problem, *Journal of Royal Statistical Association*. Series C. **28(2)**, pp. 126-135, 1979.

[8]   J. Von Neumann, Distribution of the ratio of the mean square successive difference to the variance. *Annals Math. Stat.*, **13**, pp. 367–395, 1941.

[9]   D. Rees, *Essential Statistics* (4th Edition), Chapman & Hall/CRC, 2001.

[10] H. Levene, Contributions to Probability and Statistics*: Essays in Honor of Harold Hotelling. Stanford University Press. 1960.*

[11] M. Brown and A. Forsythe, Robust Tests for Equality of Variances, *Journal of the American Statistical Association,* **69**, pp. 364–367, 1974.

[12] G. Snedecor and W. Cochran, *Statistical Methods,* (8th Edition), Iowa State University Press. 1989.

[13] Z. Kundzewics and A. Robson, Change Detection in Hydrological Records – A Review of the Methodology. *Hydrological Sciences Journal*, **49(1)**, pp. 7-19, 2004.

[14] E. Douglas, R. Vogel, and C. Kroll, Trends in Floods and Low Flows in the United States: Impact of Serial Correlation. *Journal of Hydrology*, **240**, pp. 90-105, 2000.

[15] S. Yue, P. Pilon and G. Cavadias, Power of the Mann-Kendall and Spearman's Rho Test for Detecting Monotonic Trends in Hydrological Series. *Journal of Hydrology*, **259**, 254-271, 2002.

[16] S. Yue and C. Wang, Applicability of Prewhitening to Eliminate the Influence of Serial Correlation on the Mann-Kendall Test. *Water Resources Research*, **38(6)**, 4-1-4-7, 2002.

[17] G. Kite, Frequency and Risk Analyses in Hydrology. Water Resources Publications. Fort Collins, 1988.