ALMA MATER STUDIORUM UNIVERSITÀ DI BOLOGNA

ARCHIVIO ISTITUZIONALE DELLA RICERCA

Alma Mater Studiorum Università di Bologna
Archivio istituzionale della ricerca

Distribution-free time between events and amplitude control charts for drought monitoring

This is the final peer-reviewed author’s accepted manuscript (postprint) of the following publication:

Published Version:
Scagliarini, M. (2025). Distribution-free time between events and amplitude control charts for drought monitoring. Springer, Cham [10.1007/978-3-031-64346-0_22].

Availability:
This version is available at: https://hdl.handle.net/11585/1007186 since: 2025-03-11

Published:

DOI: http://doi.org/10.1007/978-3-031-64346-0_22

Terms of use:

Some rights reserved. The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (https://cris.unibo.it/).
When citing, please refer to the published version.

(Article begins on next page)

14 July 2026

Distribution-Free Time Between Events and Amplitude Control Charts for Drought Monitoring

Michele Scagliarini[1[0000-0003-3209-4745]]

[1] Department of Statistical Sciences, University of Bologna, Italy
michele.scagliarini@unibo.it

Abstract. In this study, time between events and amplitude control charts are used to detect changes in the characteristics of drought events. We used non-parametric methodologies that do not require any assumption on the distribution of the phenomenon or on the observed statistics. The results indicate that the proposed methods can be valuable tools for the institutions responsible for planning drought management and mitigation measures.

Keywords: Climate Change, Control Charts, Distribution Free, Drought Index, SPEI.

1 Introduction

Droughts have severe consequences for agricultural production, the environment, the economy and social stability. To properly asses this phenomenon, it is important to objectively quantify the characteristics of drought episodes in terms of their intensity, magnitude and duration. For this purpose, the Standardized Precipitation Evapotranspiration Index, SPEI, has been proposed [5]. In summary, SPEI is a climatic drought index based on precipitation, temperature and potential evapotranspiration. It can be calculated on a range of accumulation periods, from 1 to 48 months (referred to as SPEI-1, SPEI-2, and so on) and the different timescales are used to reflect the impact of drought on different water-related sectors. Meteorological and soil moisture conditions (agriculture droughts) respond to precipitation and temperature anomalies on relatively short time scales, such as 1-6 months (SPEI-1 to SPEI-6). River flow, reservoirs, and groundwater (hydrologic drought) respond to longer-term anomalies of the order of 6 months to 12 months or longer (SPEI-6 to SPEI-12). SPEI can assume positive and negative values with the following moisture categories: values greater than 2 are considered extremely wet conditions; 1.5 to 2 very wet conditions; 1 to 1.5 moderately wet conditions; -1 to 1 normal conditions; -1.5 to -1 moderately dry conditions; -2 to -1.5 severely dry conditions, and values below -2 extremely dry conditions. SPEI data are available from the Global SPEI database SPEIbase v2.9 (https://spei.csic.es/database.html) with time scales from 1 to 48 months, spatial resolution of 0.5° lat/lon (approximately 55 km), and temporal coverage from January 1901 to December 2022.

The severe consequences of water scarcity conditions have focused attention on drought monitoring methods that can assist governments in implementing preparedness plans and mitigation measures to reduce the economic, environmental, and social impacts of drought. Among the possible monitoring methodologies, statistical process control methods have been found to be helpful in supporting effective drought management [4].

The relevant drought characteristics are severity, duration and frequency; therefore, a suitable statistical process control methodology for monitoring drought events should consider the time interval T between two occurrences and the magnitude X of each event. Time-Between-Events-and-Amplitude (TBEA) control charts have been proposed to monitor this type of phenomenon: a decrease in T and/or an increase in X may result in a negative condition that needs to be monitored and possibly detected with control charts. Most of the TBEA control charts proposed in the literature [2] assume known distribution functions for the variables T and X. However, in the majority of real situations, the distributions of these random variables are unknown or very difficult to identify.

In this work, we use distribution-free TBEA control charts to detect changes in the characteristics of drought events. For our purposes, aware of being non-exhaustive, we consider: the distribution-free upper sided Exponentially Weighted Moving Average (EWMA) control chart proposed by [6]; furthermore, we propose to use two non-parametric change-point control charts based on the Mann-Whitney (MW) and Kolmogorov-Smirnov (KS) statistics, respectively.

The paper is structured as follows. In Section 2, the methodologies considered for our purpose are introduced. Section 3 describes the data and explains and discusses the results.

2 Methodology

Let T_i (i=1,2,…) be the time intervals between two consecutive occurrences of a given event E and let X_i be the corresponding magnitudes of the event of interest. To monitor T_i and X_i simultaneously [6] assumed that T_i and X_i are continuous random variables, both defined on [0;+∞), with unknown distribution functions $F_T\left(t\middle|\theta_T\right)$ and $F_X\left(x\middle|\theta_X\right)$ respectively, where θ_T and θ_X are known α-quantiles. Without loss of generality, it is considered that θ_T and θ_X are the median values of T_i and X_i, respectively, and when the process is in-control it follows that $\theta_T = \theta_{T_0}$ and $\theta_X = \theta_{X_0}$. The distribution-free EWMA TBEA control chart is based on the sign statistics $\mathrm{ST}_i = \mathrm{sign}\left(T_i - \theta_{T_0}\right)$ and $\mathrm{SX}_i = \mathrm{sign}\left(X_i - \theta_{X_0}\right)$, for i=1,2,…, where sign(x) = −1 if x<0 and sign(x) = +1 if x>0. Let us now define the statistic $S_i = \left(\mathrm{SX}_i - \mathrm{ST}_i\right)/2$ (i=1,2,…), which has the following behaviour: S_i = −1 when T_i increases and, at the same time, X_i decreases (positive situation); S_i = +1 when T_i decreases and, at the same time, X_i increases (negative situation); S_i = 0 when both T_i and X_i increase or when both T_i and X_i

decrease (intermediate situation). Note that S_i is a discrete random variable; therefore, it would be impossible to accurately compute the run length properties of a control chart based on this statistic. To solve this problem the authors [6] developed the "continuousify" method. They propose to define an extra parameter $\sigma \in [0.1, 0.2]$ and to convert S_i into a continuous random variable, S_i^*, defined as a mixture of 3 normal random variables $Y_{i,-1} \sim N(-1,\sigma)$, $Y_{i,0} \sim N(0,\sigma)$ and $Y_{i,+1} \sim N(1,\sigma)$, respectively. More precisely S_i^* is defined as: $S_i^* = Y_{i,-1}$ if $S_i = -1$; $S_i^* = Y_{i,0}$ if $S_i = 0$; $S_i^* = Y_{i,+1}$ if $S_i = +1$. Note that the parameter σ has to be fixed. However, the authors demonstrated that this parameter has no impact on the performance of the control chart if it falls within the suggested range. The upper-sided EWMA TBEA control chart uses the statistic $Z_i^* = \max\left(0, \lambda S_i^* + (1-\lambda) Z_{i-1}^*\right)$ with an upper asymptotic control limit given by $UCL = K\sqrt{\lambda\left(\sigma^2 + 0.5\right)/(2-\lambda)}$, where $\lambda \in [0, 1]$ and $K>0$ are the control chart parameters and the initial value is $Z_0^* = 0$. The optimal design parameters λ and K can be obtained by studying the statistical properties of the control chart by means of a Markov chain approach [6].

As mentioned in Section 1, to simultaneously monitor the time T between an event E and its amplitude X, the majority of the studies proposed parametric methods. In this framework T and X are assumed to be continuous random variables, both defined on $[0;+\infty)$, with known distribution functions $F_T(t|\theta_T)$ and $F_X(x|\theta_X)$ respectively, where θ_T and θ_X are the corresponding vector of parameters. Let $\mu_T=E(T)$, $\mu_X=E(X)$, $\sigma_T=SD(T)$ and $\sigma_X=SD(X)$ be the expectation and standard deviation of T and X, respectively. As usual, when the process is in control it follows that $\mu_T = \mu_{T_0}$, $\mu_X = \mu_{X_0}$, $\sigma_T = \sigma_{T_0}$ and $\sigma_X = \sigma_{X_0}$. Since the variables T and X can have very different scales, it is appropriate to define and use the normalized variables $T' = T/\mu_{T_0}$ and $X' = X/\mu_{X_0}$. Therefore, for TBEA monitoring [2] defined several dedicated statistics $Z=Z(T', X')$, functions of T' and X', satisfying the following two properties: Z increases if either T' decreases or X' increases; Z decreases if either T' increases or X' decreases. The difficulty with this approach lies in the fact that the distributions of T and X are assumed to be known. Furthermore, it is difficult to derive the distribution of the Z statistics needed to design the control chart.

Here we propose to monitor the Z statistics with non-parametric change-point control charts. In the Change Point Model (CPM) framework, it is assumed that the values of the TBEA statistic z_1, z_2,... are generated by the random variables Z_1, Z_2,... with unknown distribution functions F_1, F_2,... , respectively. A change point occurs at instant τ when $F_\tau \neq F_{\tau+1}$ and it is usually assumed that the observations are independent and identically distributed between every pair of change points. Therefore, the distribution of the sequence can be described by the following model: $Z_i \sim F_0$ if $0 < i \le \tau_1$, $Z_i \sim F_1$ if $\tau_1 < i \le \tau_2, \ldots,$ $Z_i \sim F_k$ if $\tau_k < i \le m$, where $0 < i < \tau_1 < \tau_2 < \cdots < \tau_k < m$ denote k unknown change points. Within this framework, it is of interest to test H_0: $k=0$, versus

H_1: $k \neq 0$, to estimate the number of change points, k, and to estimate their locations $\hat{\tau}_1, \hat{\tau}_2, ..., \hat{\tau}_k$. For testing the aforementioned hypothesis system, [1] proposed a change point control chart, designed to detect changes in the location of the process distribution, based on the Mann-Whitney (MW) U-statistic: $U_{k,n} = \sum_{i=1}^{k} \sum_{j=k+1}^{n} D_{ij}$, $1 \leq k \leq n-1$, where $D_{ij} = \text{sign}\left(Z_i - Z_j\right)$ with $D_{ij} = 1$ if $Z_i > Z_j$, $D_{ij}=0$ if $Z_i = Z_j$, $D_{ij}= -1$ if $Z_i > Z_j$. The variance of $U_{k,n}$ depends on the split point k, so for this reason it is suitably standardized, thus obtaining the statistic $T_{k,n}$. Therefore, the test for the presence of a change point and the estimate of its time of occurrence are given by maximizing $T_{k,n}$ over k: $T_{\max,n} = \max_{1 \leq k \leq n-1} \left|T_{k,n}\right|$. To detect arbitrary changes in the process distribution, [3] proposed a control chart based on the Kolmogorov-Smirnov (KS) statistic. This control chart tests for a change point immediately following any observation z_k by partitioning the observations into two samples, $S_1=(z_1,\ldots,z_k)$ and $S_2=(z_{k+1},\ldots,z_t)$, and subsequently comparing the corresponding empirical distribution functions $\hat{F}_{S_1}\left(z\right)$ and $\hat{F}_{S_2}\left(z\right)$. The test statistic is based on the maximum difference between the empirical distributions $D_{k,t} = \sup_z \left|\hat{F}_{S_1}\left(z_i\right) - \hat{F}_{S_2}\left(z_i\right)\right|$. Further details can be found in [1] and [3].

3 Data, results and discussion

Among the possible timescales for the SPEI, the 12-month accumulation period was chosen because the SPEI-12 is a reliable measure of hydrological drought. The SPEI-12 data have been downloaded from the Global SPEI database for the coordinates of Bologna (44.4949° N, 11.3426° E), where observations are available from December, 1901 to December, 2022 (1453 observations).

We considered the occurrences of SPEI-12≤-1, corresponding to drought conditions, as the events (E) of interest. Since the TBEA methodology assumes that the magnitudes of the events are defined on [0; +∞), the absolute values of SPEI-12 at the dates of occurrence have been considered as the corresponding magnitudes X_i. Given the monthly frequency of the SPEI-12, the time T_i between two events is computed in months. The available dataset contains 226 events E (values of SPEI-12≤-1). The first 45 events have been used for performing the Phase I and the remaining 181 represent the monitored data. To compute the values of ST_i, SX_i, S_i and S_i^* the in-control median values have been estimated from Phase I data: $\hat{\theta}_{T_0} = 1$ and $\hat{\theta}_{X_0} = 1.2177$. Following the suggestions of the authors [6], we set σ=0.125 and to obtain a control chart with an in-control ARL equal to ARL_0=370, we select λ=0.07 and K=2.515. Figure 1 shows the results obtained with the EWMA TBEA control chart.

As far as the change point control charts is concerned, among the possible choices for the statistics Z to be monitored, we considered the ratio $Z_R = T'/X'$. To obtain the normalized variables T' and X' the in-control means have been estimated from Phase

I data: $\hat{\mu}_{T_0}$ =11.7333, $\hat{\mu}_{X_0}$ =1.3385. To allow fair comparisons, the MW and KS control charts were set up with the same in-control statistical properties as the EWMA TBEA control chart (ARL_0=370). The results are shown in Figure 2, which reports the estimated change points, the values of the Z_R statistic, and the averages of the observed statistics between each pair of change points (dashed line).

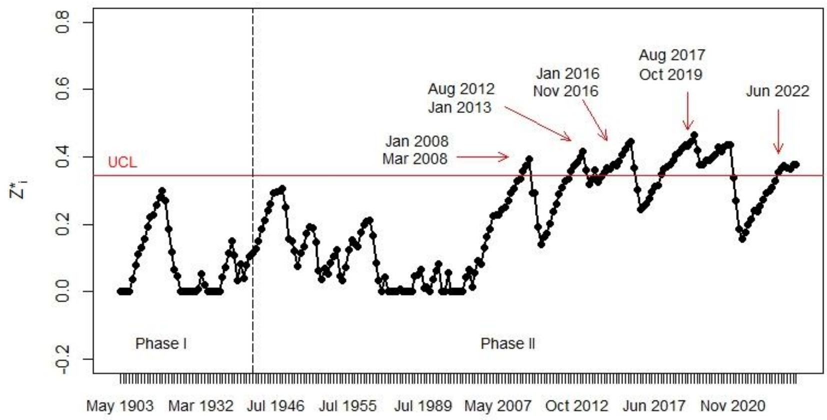

Fig. 1. Time Between Event and Amplitude EWMA control chart

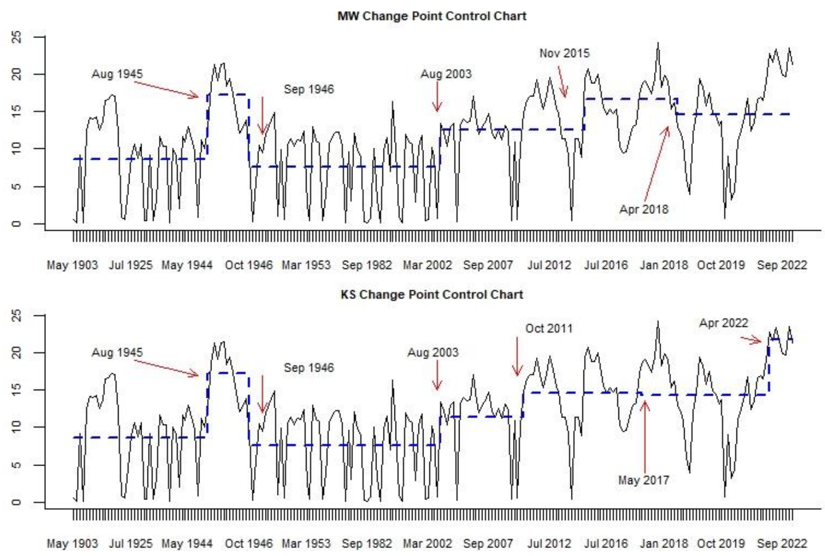

Fig. 2. Mann-Whitney (upper panel) and Kolmogorov-Smirnov (lower panel) control charts

By examining the results, it can be seen that for Phase I data no alarms are reported; therefore, this dataset is in statistical-control and the estimated medians $\hat{\theta}_{T_0} = 1$, $\hat{\theta}_{X_0} = 1.2177$ and averages $\hat{\mu}_{T_0}$ =11.7333, $\hat{\mu}_{X_0}$ =1.3385 can be considered as reliable estimates of the corresponding unknown parameters and used for the monitoring phase. Before commenting further on the results, it should be remembered that we have used different methodologies and therefore we do not expect identical results. However, it can be noted that the control charts considered agree in establishing that the drought evolved for the worse during our study period: since August 2003 (MW and KS control charts, Figure 2) or August 2012 (EWMA control chart, Figure 1), the TBEA statistics show significant increases. Furthermore, it is worth noting that the examined methodologies provide results that can be used in a complementary way. As far as the EWMA chart is concerned, it is possible to perform a post-signal analysis to highlight out-of-control events that are consecutive with respect to the original time scale. For example, note the long period of temporally consecutive out-of-control Z_i^* values between August 2017 and October 2019. To discuss the results obtained with the change point control charts, it is worth remembering that we are monitoring $Z_R = T'/X'$, which simultaneously takes into account the time between drought events and their magnitudes. This means that by detecting change points in the sequence of this TBEA statistic, it is possible to identify significant changes in drought characteristics. For example, consider the MW control chart. The observed increases in the location of the TBEA statistic in August 2003 and November 2015 indicate an increase in drought severity.

In summary, since drought mitigation measures vary according to drought severity and duration, TBEA control charts can provide valuable input to institutions responsible for planning suitable drought management and designing resilience policies.

References

1. Hawkins, D.M., Deng, Q.: A Nonparametric Change-Point Control Chart. Journal of Quality Technology, 42(2), 165-173 (2010).
2. Rahali, D., Castagliola, P., Taleb, H., Khoo, M.B.C.: Evaluation of Shewhart Time-Between-Events-and-Amplitude Control Charts for Several Distributions. Quality Engineering, 31(2), 240-254 (2019).
3. Ross, G.J., Adams, N.M.: Two Nonparametric Control Charts for Detecting Arbitrary Distribution Changes. Journal of Quality Technology, 44(12), 102-116 (2012).
4. Townsley, J., Chimka, J.R: A Control Chart for Severity Index to Detect Drought Compare Favourably to Logistic Regression. Hydrology Reseragh, 42(1),1-9 (2011).
5. Vicente-Serrano, S. M., Beguería, S., López-Moreno, J. I.: A Multiscalar Drought Index Sensitive to Global Warming: The Standardized Precipitation Evapotranspiration Index. Journal of Climate, 23(7), 1696–1718 (2010).
6. Wu, S., Castagliola, P., Celano G.: A Distribution-Free EWMA Control Chart for Monitoring Time-Between-Events-and-Amplitude Data. Journal of Applied Statistics, 48(3), 434-454 (2021).