# ENBIS 2026 — Florence, Italy

## Nonparametric Statistical Process Control for Monitoring Structural Changes in Regional Drought Dynamics

Michele Scagliarini

Department of Statistical Sciences, University of Bologna

# Outline I

1. Motivation
2. TBEA overview
3. The EWMA TBEA control chart
4. The Change Point model approach
5. The case study
6. Results
7. Performance in the out-of-control case
8. Perturbation-based sensitivity analysis
9. Main take-home messages

# Motivation

## Motivation I

In real practice it is often of interest to monitor the occurrence of negative events. In these situations, are usually relevant:

- ***T***: time between consecutive events
- ***X***: amplitude/severity of each event

Examples include:

- equipment failures
- production defects
- warranty claims
- cyber-security incidents
- environmental extremes

In these contexts, both:

- event frequency
- event severity

must be monitored simultaneously.

## Motivation I

In real practice it is often of interest to monitor the occurrence of negative events. In these situations, are usually relevant:

- *T*: time between consecutive events
- *X*: amplitude/severity of each event

Examples include:

- equipment failures
- production defects
- warranty claims
- cyber-security incidents
- **environmental extremes**

In these contexts, both:

- event frequency
- event severity

must be monitored simultaneously.

## Motivation I

In real practice it is often of interest to monitor the occurrence of negative events. In these situations, are usually relevant:

- ***T***: time between consecutive events
- ***X***: amplitude/severity of each event

Examples include:

- equipment failures
- production defects
- warranty claims
- cyber-security incidents
- **environmental extremes**

In these contexts, both:

- event frequency
- event severity

must be monitored simultaneously.

# Motivation II

## Statistical challenges

- Joint monitoring of frequency and amplitude is required
- Data are irregular in time
- Parametric assumptions are often unreliable

## Goal

This work proposes a unified distribution-free framework for monitoring **Time-Between-Event and Amplitude (TBEA)** data

# Motivation II

## Statistical challenges

- Joint monitoring of frequency and amplitude is required
- Data are irregular in time
- Parametric assumptions are often unreliable

## Goal

This work proposes a unified distribution-free framework for monitoring **Time-Between-Event and Amplitude (TBEA)** data

# Motivation III

## Key components

1. **Distribution-free EWMA control chart**
   - Sequential monitoring
   - Sensitive to gradual process shifts

3. **Nonparametric Change-Point control chart (KS-based)**
   - Retrospective analysis
   - Detects sharper detection of discrete shifts and regime transitions

# Motivation III

## Key components

1. **Distribution-free EWMA control chart**
   - Sequential monitoring
   - Sensitive to gradual process shifts

2. **Nonparametric Change-Point control chart (KS-based)**
   - Retrospective analysis
   - Detects sharper detection of discrete shifts and regime transitions

# Motivation IV

## Added value

- in addition to signalling an alarm (**i.e., identifying the detection time**), the proposed approach is able to estimate the time at which the change occurred (**i.e., the change-point time**)

## Motivation IV

### Added value

- in addition to signalling an alarm (**i.e., identifying the detection time**), the proposed approach is able to estimate the time at which the change occurred (**i.e., the change-point time**)
- identifying when a shift occurred in a TBEA statistic can aid in tracing potential root causes and enhance overall process understanding

### What's new?

# Motivation IV

## Added value

- in addition to signalling an alarm (**i.e., identifying the detection time**), the proposed approach is able to estimate the time at which the change occurred (**i.e., the change-point time**)
- identifying when a shift occurred in a TBEA statistic can aid in tracing potential root causes and enhance overall process understanding

## What's new?

# Motivation V

The EWMA TBEA and KS control charts are jointly applied to the same dataset.

**Do the two monitoring procedures provide consistent out-of-control behaviour?**

A simulation study compares the EWMA TBEA and KS control charts under controlled process shifts.

## Motivation V

The EWMA TBEA and KS control charts are jointly applied to the same dataset.

**Do the two monitoring procedures provide consistent out-of-control behaviour?**

A simulation study compares the EWMA TBEA and KS control charts under controlled process shifts.

## Motivation V

The EWMA TBEA and KS control charts are jointly applied to the same dataset.

**Do the two monitoring procedures provide consistent out-of-control behaviour?**

A simulation study compares the EWMA TBEA and KS control charts under controlled process shifts.

# Motivation VI

## How stable are the estimated change points?

# Motivation VI

## How stable are the estimated change points?

The KS chart provides no uncertainty measure.

# Motivation VI

## How stable are the estimated change points?

The KS chart provides no uncertainty measure.

↓

Perturbation-based sensitivity analysis.

# Motivation VI

## How stable are the estimated change points?

The KS chart provides no uncertainty measure.

↓

Perturbation-based sensitivity analysis.

↓

Assessment of change-point stability.

## Motivation VI

**How stable are the estimated change points?**

The KS chart provides no uncertainty measure

↓

Perturbation-based sensitivity analysis

↓

Assessment of change-point stability

↓

**Monitoring of drought characteristics in the Emilia-Romagna region (Italy)**


Michele Scagliarini
University of Bologna

# TBEA overview

## TBEA overview I

Generally, TBEA control charts use, as statistics to be monitored, some functions of the **random variables *T*** and ***X***

Most of the contributions in literature study the statistical properties of TBEA control charts under the assumption that the **random variables *T* and *X* follow some specific distributions**

However, in practice it is very difficult to identify the actual distribution of these variables. To overcome this difficulty, several **TBEA distribution-free control charts have been proposed**

# The EWMA TBEA Control Chart

## The EWMA TBEA Control Chart I

To set up a distribution-free **E**xponentially **W**eighted **M**oving **A**verage (**EWMA**) control chart for monitoring ***T*** and ***X*** simultaneously it is assumed that ***T*** and ***X*** are **continuous random variables**, both defined on [0;+∞), with **unknown distribution functions** $F_T\left(t|\theta_T\right)$ and $F_X\left(x|\theta_X\right)$ respectively, where $\theta_T$ and $\theta_X$ are known $\alpha$-quantiles.

## The EWMA TBEA Control Chart I

To set up a distribution-free **E**xponentially **W**eighted **M**oving **A**verage (**EWMA**) control chart for monitoring ***T*** and ***X*** simultaneously it is assumed that ***T*** and ***X*** are **continuous random variables**, both defined on [0;+∞), with **unknown distribution functions** $F_T\left(t|\theta_T\right)$ and $F_X\left(x|\theta_X\right)$ respectively, where $\theta_T$ and $\theta_X$ are known $\alpha$-quantiles.

Without loss of generality, it is considered that $\theta_T$ and $\theta_X$ are the median values of ***T*** and ***X***, respectively, and when the process is in-control it follows that $\theta_T = \theta_{T_0}$ and $\theta_X = \theta_{X_0}$

## The EWMA TBEA Control Chart I

To set up a distribution-free **E**xponentially **W**eighted **M**oving **A**verage (**EWMA**) control chart for monitoring ***T*** and ***X*** simultaneously it is assumed that ***T*** and ***X*** are **continuous random variables**, both defined on [0;+∞), with **unknown distribution functions** $F_T\left(t \mid \theta_T\right)$ and $F_X\left(x \mid \theta_X\right)$ respectively, where $\theta_T$ and $\theta_X$ are known $\alpha$-quantiles.

Without loss of generality, it is considered that $\theta_T$ and $\theta_X$ are the median values of ***T*** and ***X***, respectively, and when the process is in-control it follows that $\theta_T = \theta_{T_0}$ and $\theta_X = \theta_{X_0}$

can be estimated from a Phase I

# The EWMA TBEA Control Chart II

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 for $i$=1,2,…,

$$\mathrm{SX}_i = \mathrm{sign}\left(X_i - \theta_{X_0}\right),$$

where:
sign($x$) = −1 if $x$<0
and
sign($x$) = +1 if $x$>0

# The EWMA TBEA Control Chart III

It is possible to define the statistic

$$S_i = \left(\mathrm{SX}_i - \mathrm{ST}_i\right)/2 \qquad (i=1,2,\ldots),$$

## The EWMA TBEA Control Chart III

It is possible to define the statistic

$$S_i = (\mathrm{SX}_i - \mathrm{ST}_i)/2 \qquad (i=1,2,\ldots),$$

which has the following behaviour:

- ***$S_i$* =-1** when $T_i$ increases and, at the same time, $X_i$ decreases (**positive situation**);
- ***$S_i$* = 0** when both $T_i$ and $X_i$ increase or when both $T_i$ and $X_i$ decrease (**intermediate situation**).
- ***$S_i$* = +1** when $T_i$ decreases and, at the same time, $X_i$ increases (**negative situation**);

# The EWMA TBEA Control Chart III

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 ***continuousify*** method was developed (Wu et al., 2021).

# The EWMA TBEA Control Chart IV

The ***continuousify*** method consists in defining 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:

$S_i^* = Y_{i,-1}$ with $Y_{i,-1} \sim N(-1, \sigma^2)$ **if $S_i$=–1**

$S_i^* = Y_0$ with $Y_0 \sim N(0, \sigma^2)$ **if $S_i$=0**

$S_i^* = Y_{i,+1}$ with $Y_{i,+1} \sim N(1, \sigma^2)$ **if $S_i$=+1**

## The EWMA TBEA Control Chart IV

The ***continuousify*** method consists in defining 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:

$S_i^* = Y_{i,-1}$ with $Y_{i,-1} \sim N(-1, \sigma^2)$ **if $S_i$=–1**

$S_i^* = Y_0$ with $Y_0 \sim N(0, \sigma^2)$ **if $S_i$=0**

$S_i^* = Y_{i,+1}$ with $Y_{i,+1} \sim N(1, \sigma^2)$ **if $S_i$=+1**

Note that the parameter $\boldsymbol{\sigma}$ has to be fixed. However, this parameter has no impact on the performance of the control chart if it falls within the suggested range

## The EWMA TBEA Control Chart V

The upper-sided EWMA TBEA control chart uses the statistic

$$Z_i^* = \max\left(0, \lambda S_i^* + \left(1-\lambda\right) Z_{i-1}^*\right)$$

with

$$UCL = K\sqrt{\lambda\left(\sigma^2 + 0.5\right) \Big/ \left(2-\lambda\right)}$$

and initial value is $Z_0^* = 0$


Michele Scagliarini
University of Bologna

# The EWMA TBEA Control Chart V

The upper-sided EWMA TBEA control chart uses the statistic

$$Z_i^* = \max\left(0, \lambda S_i^* + \left(1-\lambda\right) Z_{i-1}^*\right)$$

with

$$UCL = K\sqrt{\lambda\left(\sigma^2 + 0.5\right)/\left(2-\lambda\right)}$$

and initial value is $Z_0^* = 0$

The optimal design parameters $\lambda \in [0, 1]$ and $K>0$ can be obtained by studying the statistical properties of the control chart by means of a Markov chain approach

# The Change Point Model Approach

## Monitoring TBEA data: a Change Point Model Approach I

### Change Point Model Approach

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 $X' = X/\mu_{X_0}$

# Monitoring TBEA data: a Change Point Model Approach II

## Change Point Model Approach

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 $X' = X/\mu_{X_0}$

Among the alternative TBEA statistics proposed in the literature, the ratio statistic $Z_R = T'/X'$ was selected because it **directly increases when negative events become both more frequent and more severe**. This property facilitates interpretation in real cases, where simultaneous changes in event frequency and magnitude are of primary concern.

## Monitoring TBEA data: a Change Point Model Approach II

### Change Point Model Approach

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 $X' = X/\mu_{X_0}$

Among the alternative TBEA statistics proposed in the literature, the ratio statistic $Z_R = T'/X'$ was selected because it **directly increases when negative events become both more frequent and more severe**. This property facilitates interpretation in real cases, where simultaneous changes in event frequency and magnitude are of primary concern.

# Monitoring TBEA data: a Change Point Model Approach III

In the **change point model** framework it is assumed that the values of the TBEA statistic $z_1, z_2, \ldots$ are generated by the random variables $Z_1, Z_2, \ldots$ with unknown distribution functions $F_1, F_2, \ldots$

A change point occurs at $\tau$ when $F_\tau \neq F_{\tau+1}$.

## Monitoring TBEA data: a Change Point Model Approach IV

Therefore, the distribution of the sequence of the TBEA statistic can be described by the following model:

$$Z_i \sim \begin{cases} F_0 \text{ if } 0 < i \le \tau_1 \\ F_1 \text{ if } \tau_1 < i \le \tau_2 \\ \ldots \\ F_k \text{ if } \tau_k < i \le \tau_{k+1} \\ \ldots \end{cases}$$

where $0 < \tau_1 < \tau_2 < \cdots < \tau_k < \cdots$ denote unknown change points.

# Monitoring TBEA data: a Change Point Model Approach IV

Therefore, the distribution of the sequence of the TBEA statistic can be described by the following model:

$$Z_i \sim \begin{cases} F_0 \text{ if } 0 < i \le \tau_1 \\ F_1 \text{ if } \tau_1 < i \le \tau_2 \\ \ldots \\ F_k \text{ if } \tau_k < i \le \tau_{k+1} \\ \ldots \end{cases}$$

- it is of interest to test $H_0 : k = 0$, versus $H_1 : k > 0$

where $0 < \tau_1 < \tau_2 < \cdots < \tau_k < \cdots$ denote unknown change points.

# Monitoring TBEA data: a Change Point Model Approach IV

Therefore, the distribution of the sequence of the TBEA statistic can be described by the following model:

$$Z_i \sim \begin{cases} F_0 \text{ if } 0 < i \le \tau_1 \\ F_1 \text{ if } \tau_1 < i \le \tau_2 \\ \ldots \\ F_k \text{ if } \tau_k < i \le \tau_{k+1} \\ \ldots \end{cases}$$

- it is of interest: to test $H_0 : k = 0$, versus $H_1 : k > 0$
- to estimate the number of change points, $k$

where $0 < \tau_1 < \tau_2 < \cdots < \tau_k < \cdots$ denote unknown change points.

## Monitoring TBEA data: a Change Point Model Approach IV

Therefore, the distribution of the sequence of the TBEA statistic can be described by the following model:

$$Z_i \sim \begin{cases} F_0 \text{ if } 0 < i \le \tau_1 \\ F_1 \text{ if } \tau_1 < i \le \tau_2 \\ \dots \\ F_k \text{ if } \tau_k < i \le \tau_{k+1} \\ \dots \end{cases}$$

- it is of interest: to test $H_0 : k = 0$, versus $H_1 : k > 0$
- to estimate the number of change points, $k$
- to estimate their locations $\tau_1, \tau_2, ..., \tau_k, ....$

where $0 < \tau_1 < \tau_2 < \cdots < \tau_k < \cdots$ denote unknown change points.


Michele Scagliarini
University of Bologna

# The Kolmogorov-Smirnov Control Chart I

The **Kolmogorov-Smirnov (KS)** control chart is designed to detect general changes to the process distribution.

# The Kolmogorov-Smirnov Control Chart I

The **Kolmogorov-Smirnov (KS)** control chart is designed to detect general changes to the process distribution.

The 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 $S_2 = (Z_{k+1}, \ldots, Z_t)$ by comparing the corresponding empirical distribution functions $\hat{F}_{S_1}(z)$ and $\hat{F}_{S_2}(z)$.

# The Kolmogorov-Smirnov Control Chart I

The **Kolmogorov-Smirnov (KS)** control chart is designed to detect general changes to the process distribution.

The 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 $S_2 = (Z_{k+1}, \ldots, Z_t)$ by comparing the corresponding empirical distribution functions $\hat{F}_{S_1}(z)$ and $\hat{F}_{S_2}(z)$.

$$D_{k,t} = \sup_z \left| \hat{F}_{S_1}(Z_i) - \hat{F}_{S_2}(Z_i) \right|$$

If $D_{k,t} > {}_D h_{k,t}$ reject $H_0$

The **KS** control chart uses a statistic defined as the maximum difference between the empirical distributions

## The Kolmogorov-Smirnov Control Chart II

If $H_0$ is rejected the **estimate $\hat{\tau}$ of the change point location** is the value of $k$ that maximizes $D_{k,t}$.

The threshold ${}_{D}h_{k,t}$, such as to guarantee the desired $ARL_0$, is determined using Monte Carlo simulation (Ross and Adams, 2012).

# The Case Study

## The Case Study I

Let's consider a generic stochastic process observed over time.

## The Case Study I

Let's consider a generic stochastic process observed over time.

Events are defined whenever the process crosses a fixed threshold.

## The Case Study I

Let's consider a generic stochastic process observed over time.

Events are defined whenever the process crosses a fixed threshold.

In this case the **adverse event** of interest is the **drought event**

## The Case Study I

Let's consider a generic stochastic process observed over time.

Events are defined whenever the process crosses a fixed threshold.

In this case the **adverse event** of interest is the **drought event**

**Event definition**

$$\text{Regional Drought Index SPEI-12}_R\left(t\right) \leq -1$$

## The Case Study II

Each event is characterized by:

- its occurrence time;
- its amplitude.

## The Case Study II

Each event is characterized by:

- its occurrence time;
- its amplitude.

From these quantities we compute the TBEA variables:

- $T_i$ = time between two successive (drought) events
- $X_i$ = event magnitude (absolute SPEI value at event time)

## The Case Study III

We considered the occurrences **194 drought events *E*** observed in the Emilia Romagna region (Italy) between December, 1901-December, 2023

The first **35 events *E*** (Phase I dataset) were used to estimate the in-control parameters for the monitoring charts. The remaining 159 events constituted the Phase II dataset.

To allow fair comparisons, all control charts were set up with the same in-control statistical properties, i.e. **ARL0=370 or $\alpha_0$=0.0027**

## The Case Study III

We considered the occurrences **194 drought events *E*** observed in the Emilia Romagna region (Italy) between December, 1901-December, 2023

The first **35 events *E*** (Phase I dataset) were used to estimate the in-control parameters for the monitoring charts. The remaining 159 events constituted the Phase II dataset.

To allow fair comparisons, all control charts were set up with the same in-control statistical properties, i.e. **ARL0=370 or $\alpha_0$=0.0027**

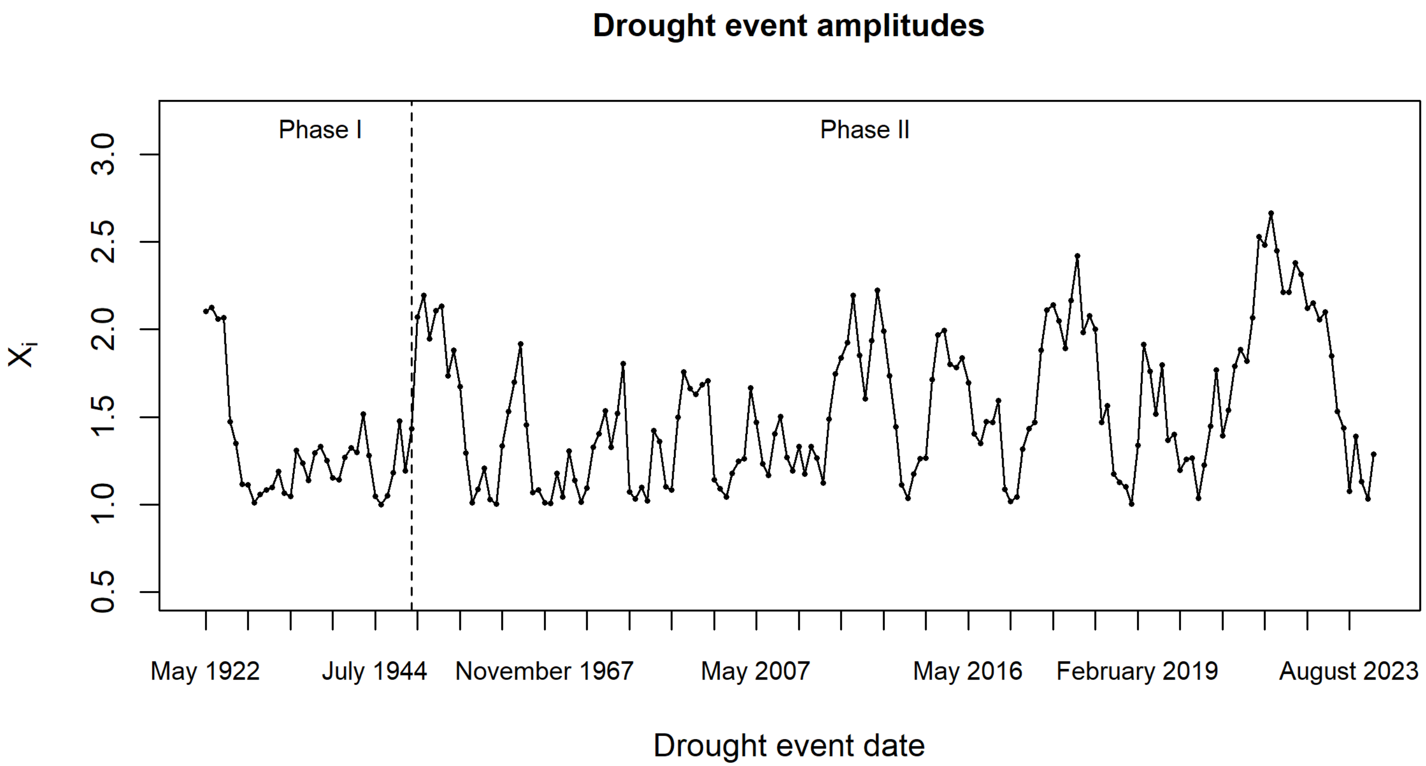
## The Case Study III

We considered the occurrences **194 drought events *E*** observed in the Emilia Romagna region (Italy) between December, 1901-December, 2023

The first **35 events *E*** (Phase I dataset) were used to estimate the in-control parameters for the monitoring charts. The remaining 159 events constituted the Phase II dataset.

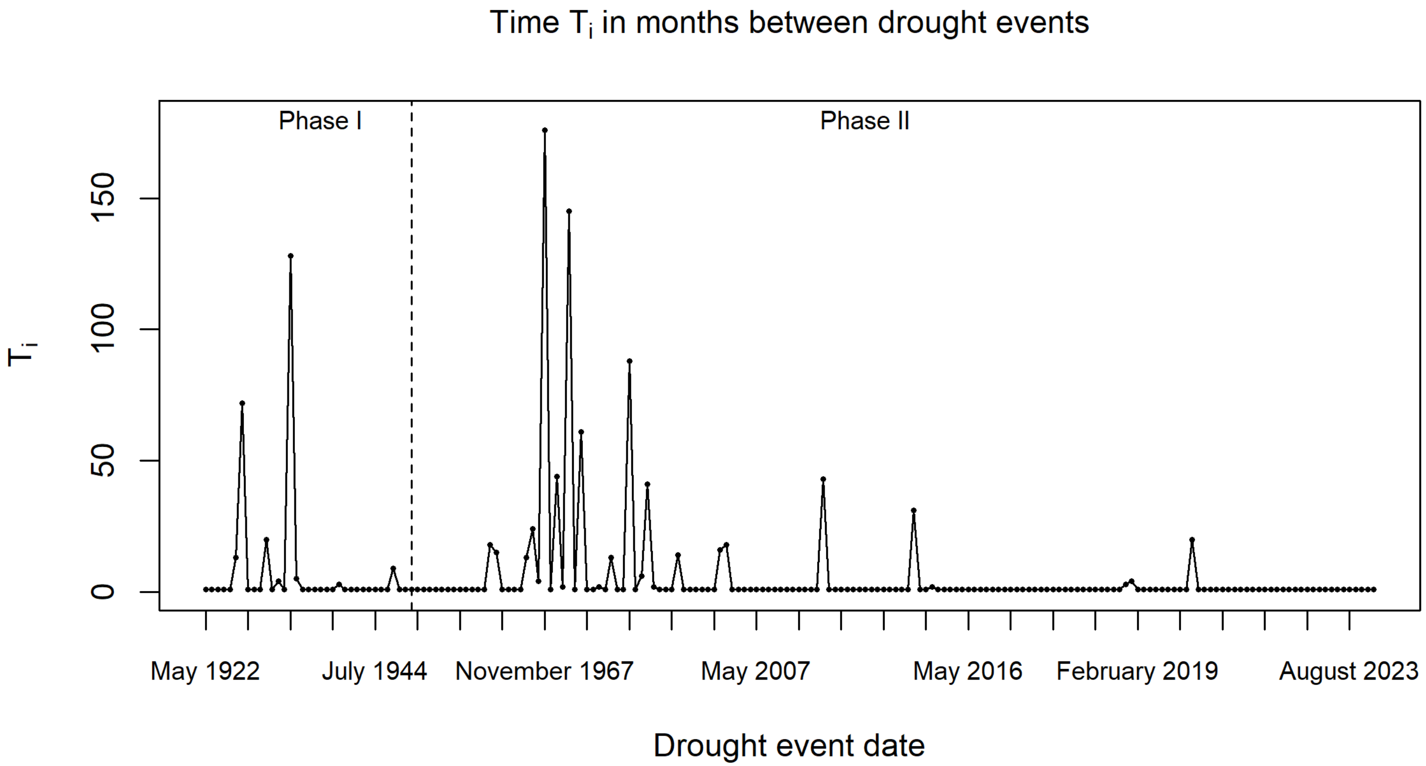
To allow fair comparisons, all control charts were set up with the same in-control statistical properties, i.e. **ARL0=370 or $\alpha_0$=0.0027**

# The Case Study V




Amplitudes $X_i$ of the events $E$

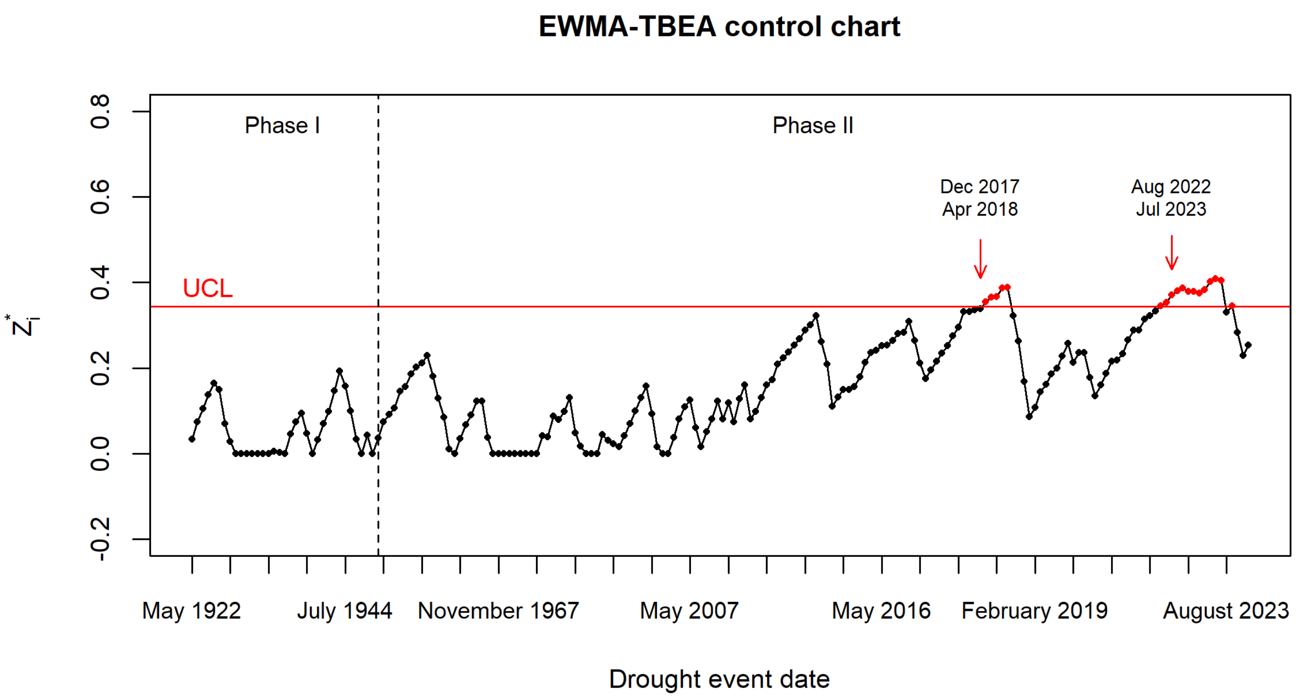
# The Case Study VI




Time $T_i$ in months between events $E$

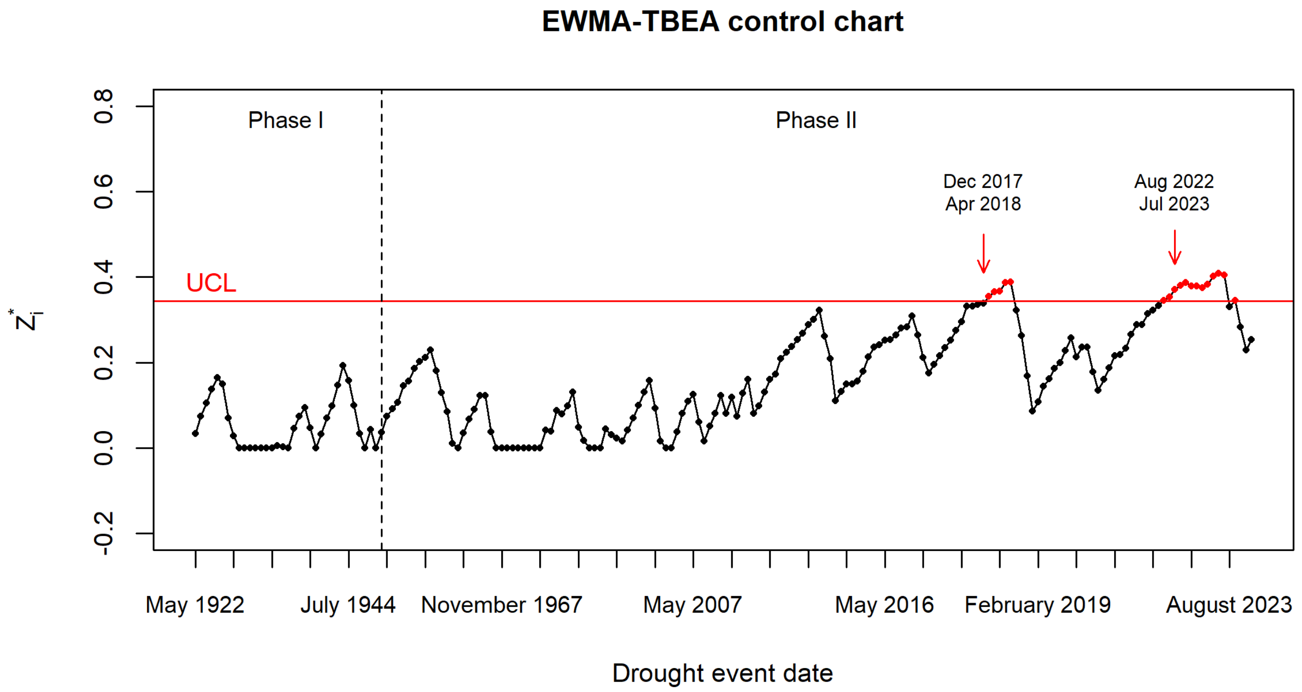
# Results

## EWMA TBEA control chart



# Results I

## EWMA TBEA control chart

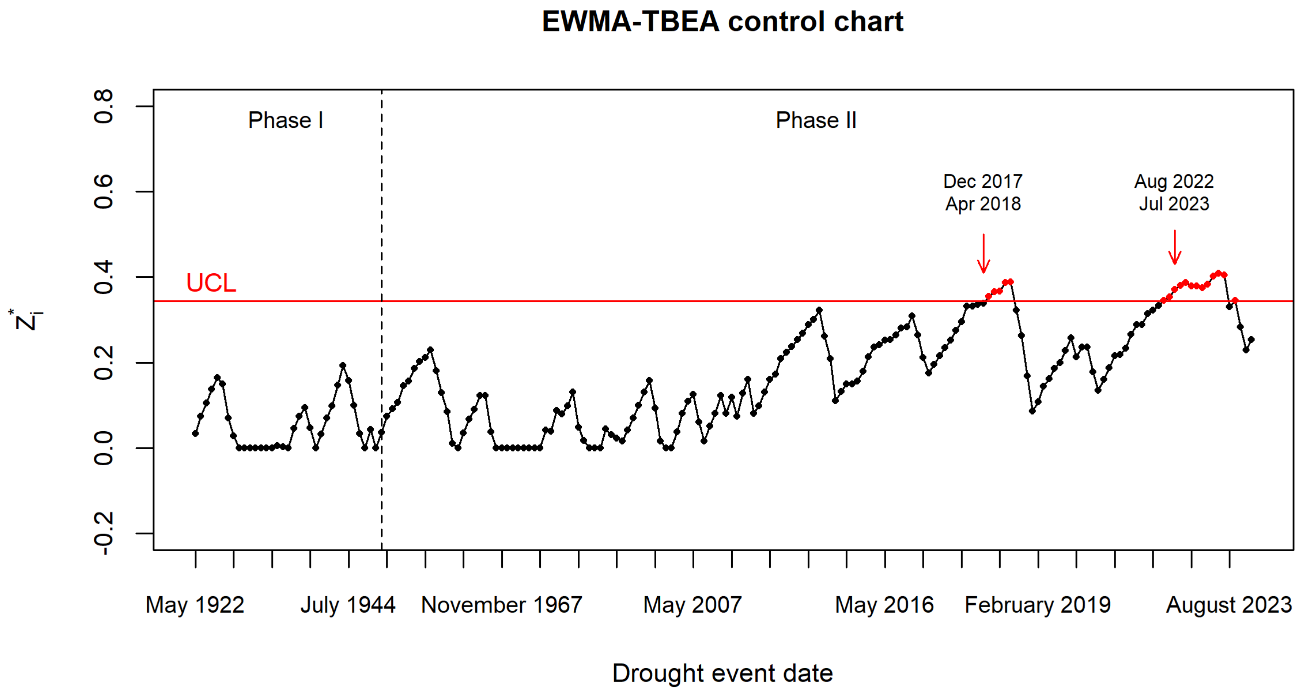



Two major sequences of out-of-control signals are observed:

- the first between December 2017 and August 2018,

# Results I

## EWMA TBEA control chart

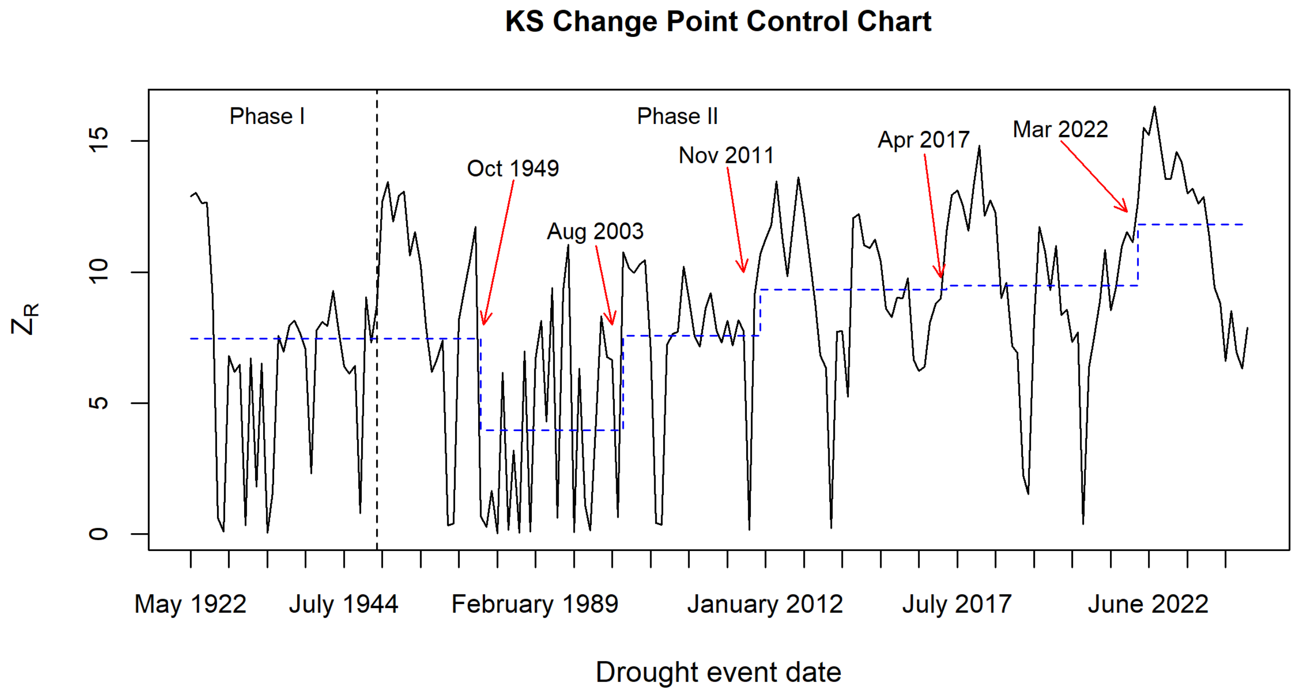



Two major sequences of out-of-control signals are observed:

- the first between December 2017 and August 2018,
- the second between August 2022 and July 2023.

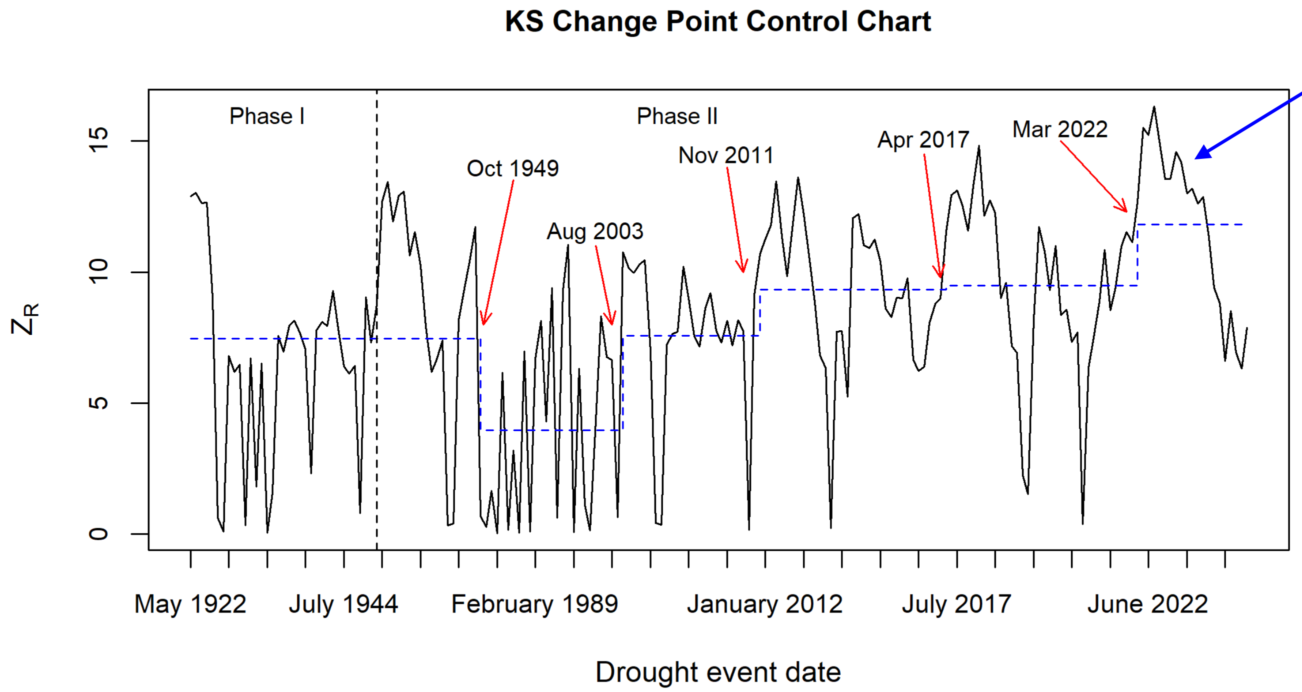
# Results II

## Kolmogorov-Smirnov control chart



# Results II

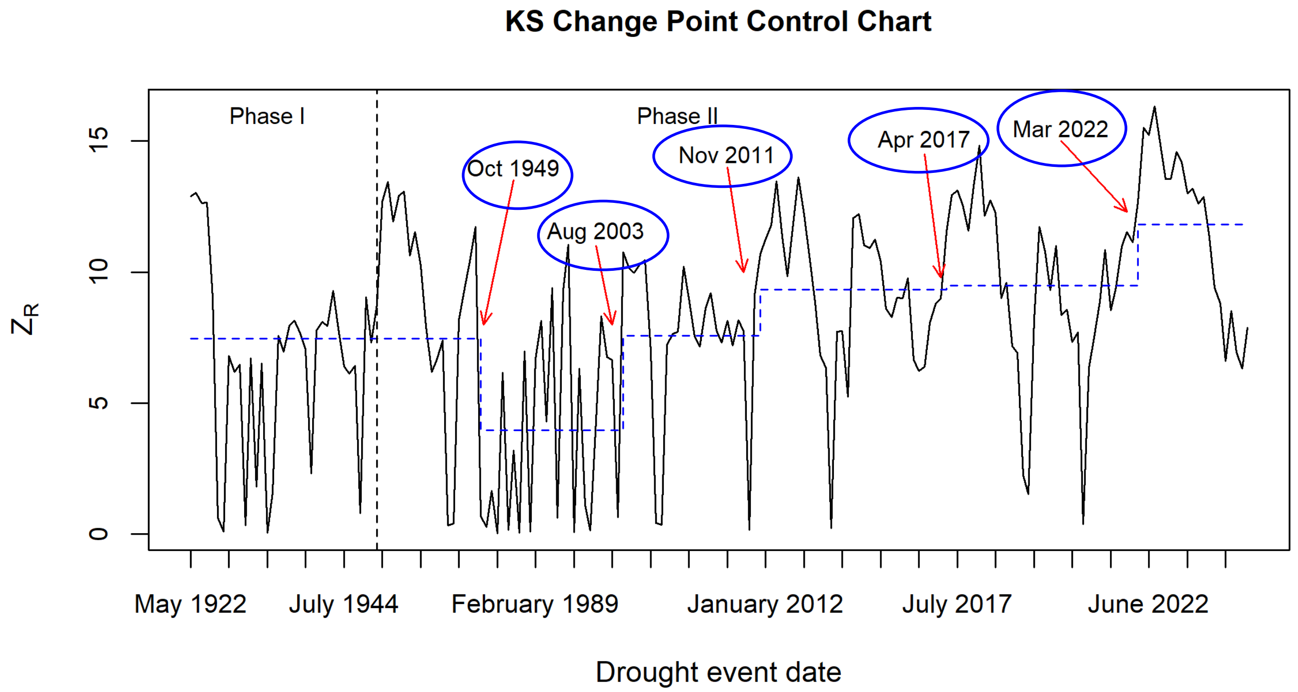
## Kolmogorov-Smirnov control chart




- The values of the $Z_R$ statistic

# Results II

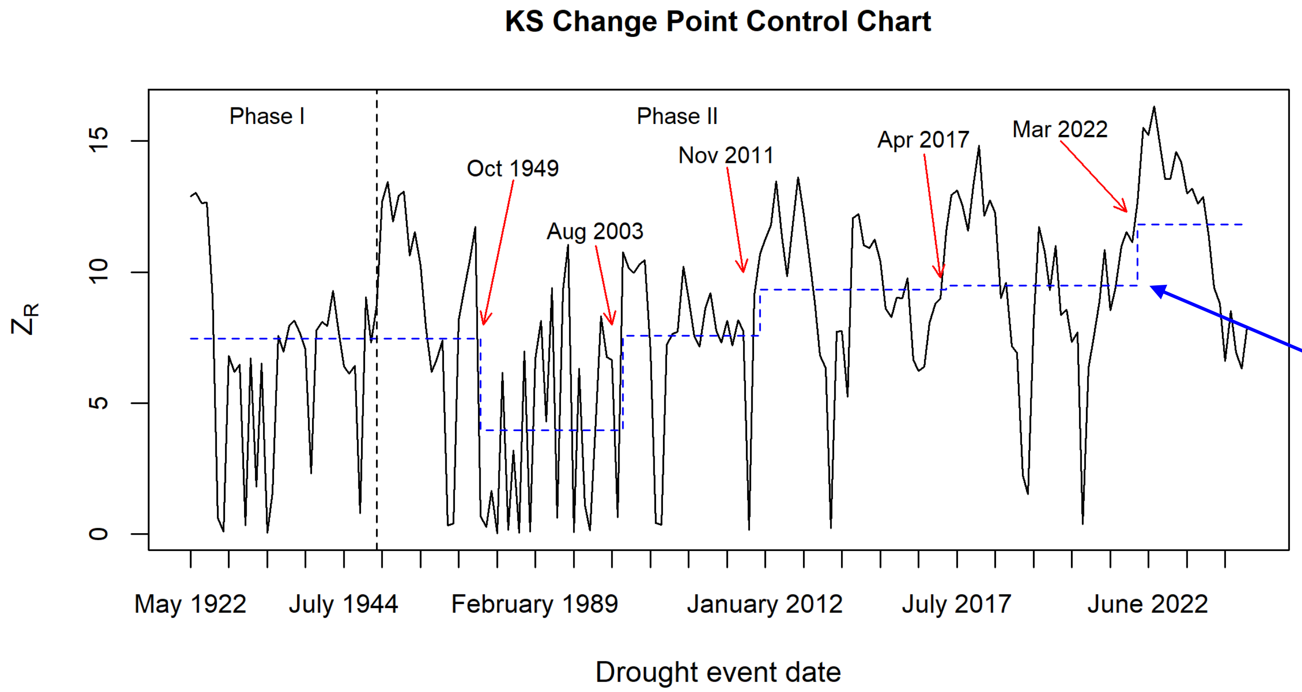
## Kolmogorov-Smirnov control chart




- The values of the $Z_R$ statistic
- The estimated change points

# Results II

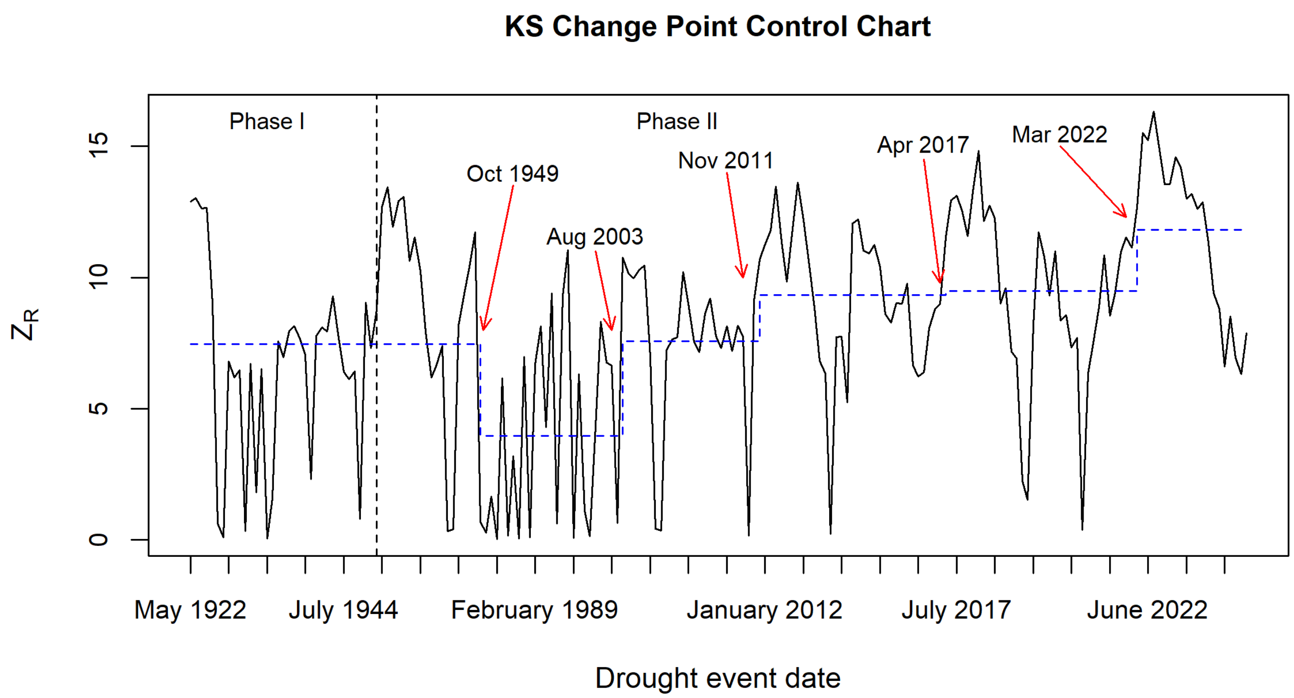
## Kolmogorov-Smirnov control chart




- The values of the $Z_R$ statistic
- The estimated change points
- The averages of the $Z_R$ statistic between each pair of change points (dashed line)

# Results III

## Kolmogorov-Smirnov control chart

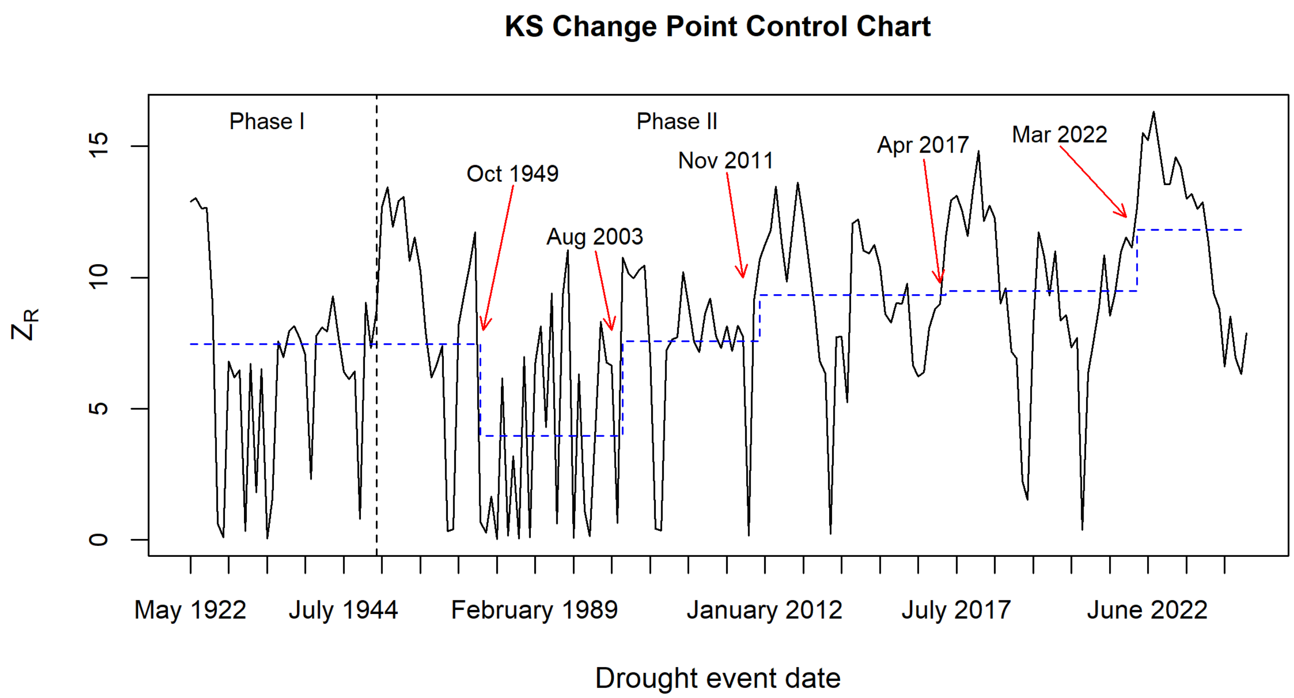



Five change points were identified:

- October 1949
- August 2003
- November 2011
- April 2017
- March 2022

# Results IV

## Kolmogorov-Smirnov control chart



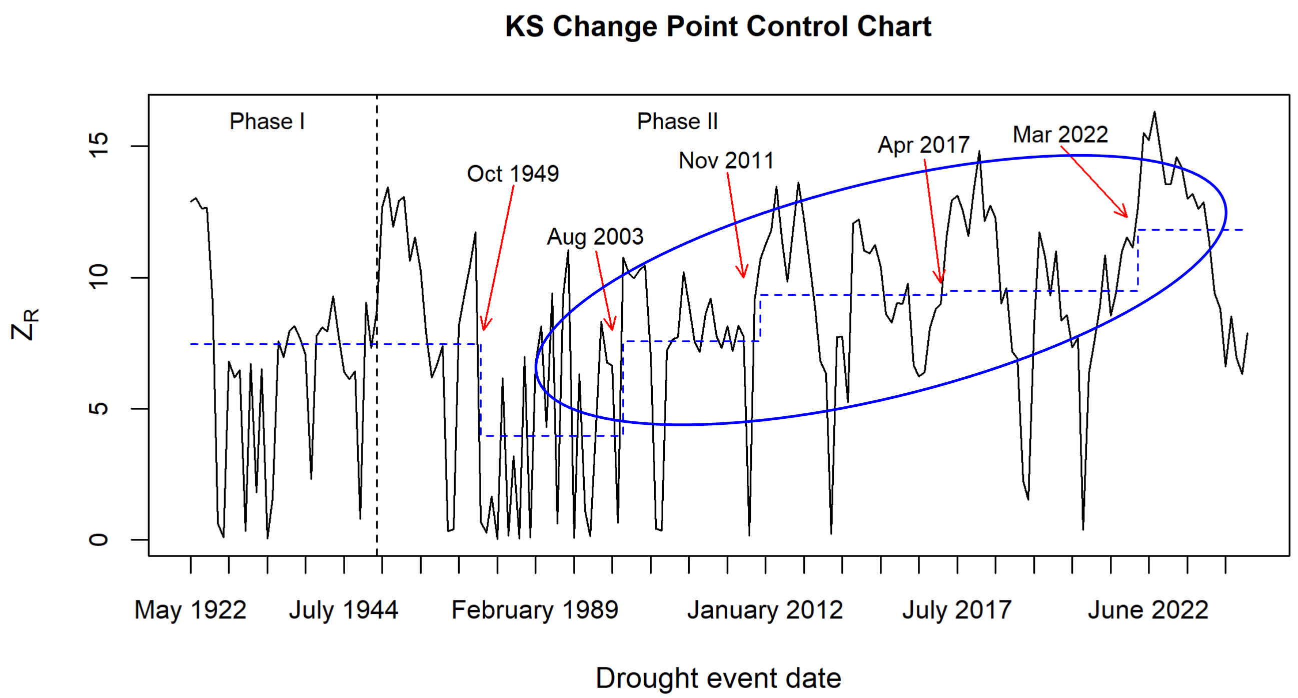

Five change points were identified:

- October 1949
- August 2003
- November 2011
- April 2017
- March 2022

Since the KS chart monitors the statistic $Z_R$, which jointly accounts for time between drought events and their magnitudes, **the detected change points correspond to significant shifts in drought characteristics**

# Results V

## Kolmogorov-Smirnov control chart




Five change points were identified:

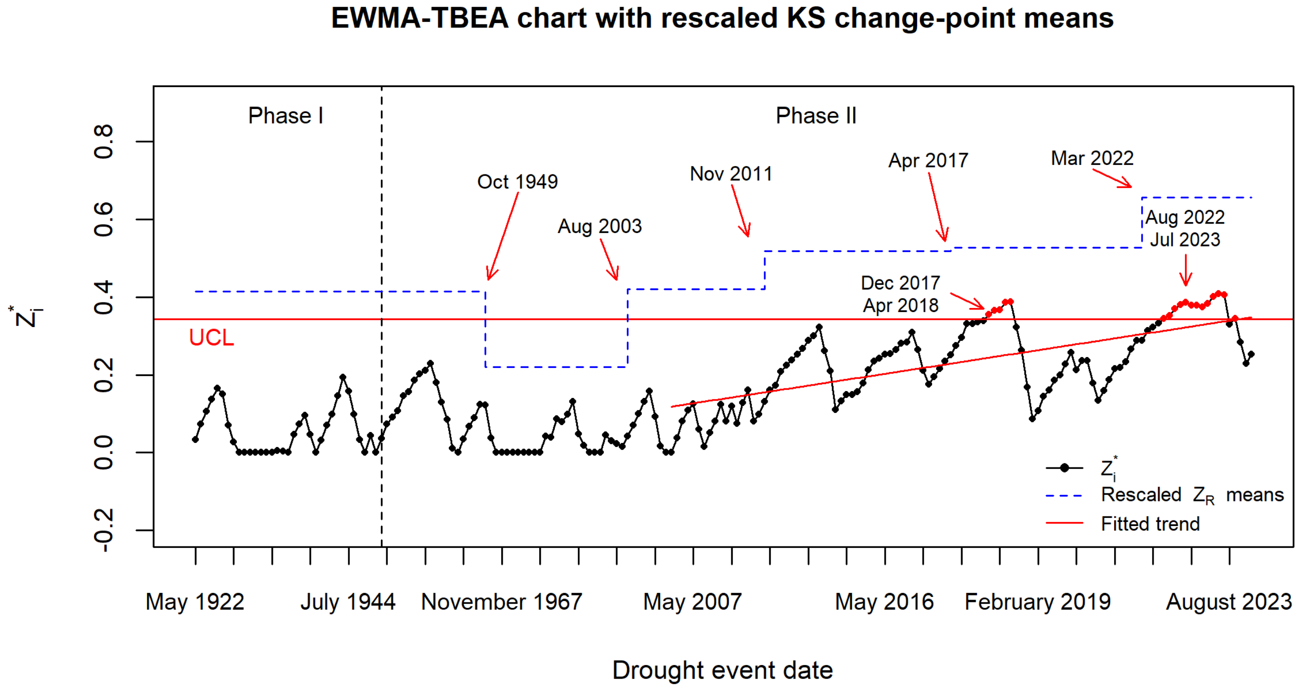
- October 1949
- August 2003
- November 2011
- April 2017
- March 2022

In particular, **the post-2003** period exhibits a pronounced upward shift in the mean of the $Z_R$ statistic, indicating a deterioration in drought conditions characterized by increasingly frequent and severe events

# Results VI

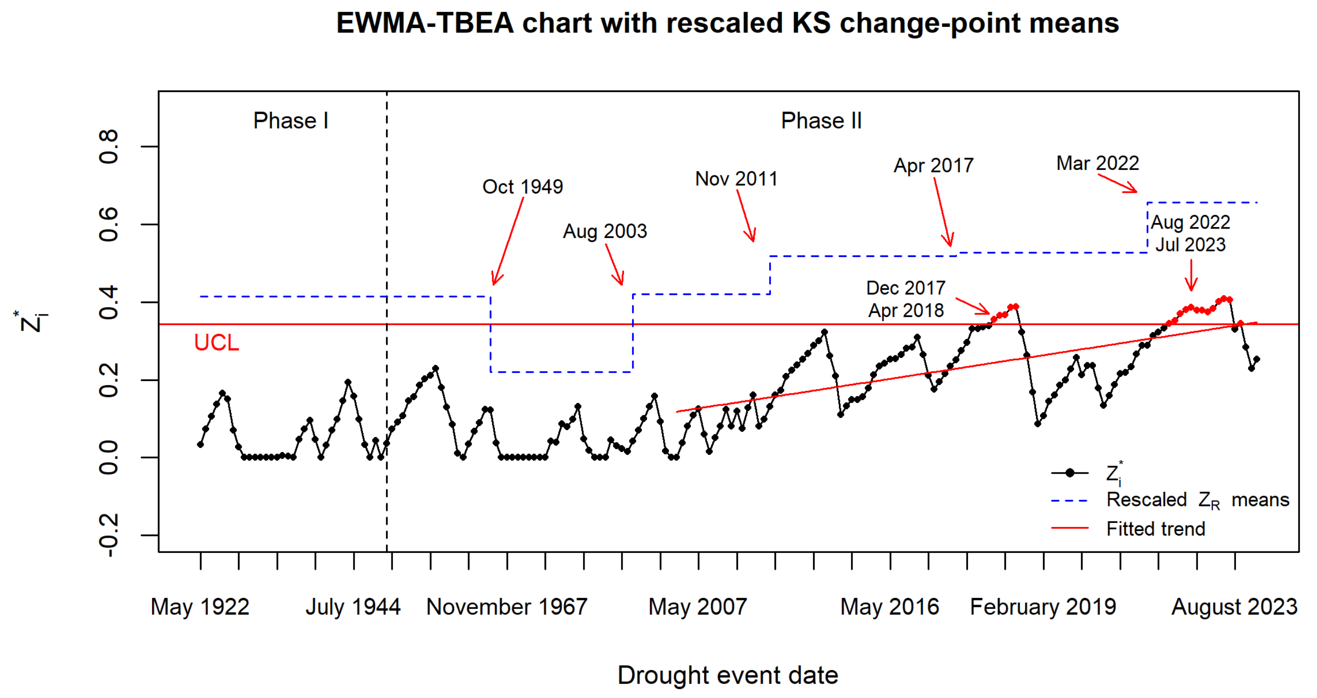
**Consistency between the results** can be highlighted if the averages (appropriately rescaled) of the $Z_R$ statistic between each pair of change points are superimposed on the EWMA chart

# Results VI



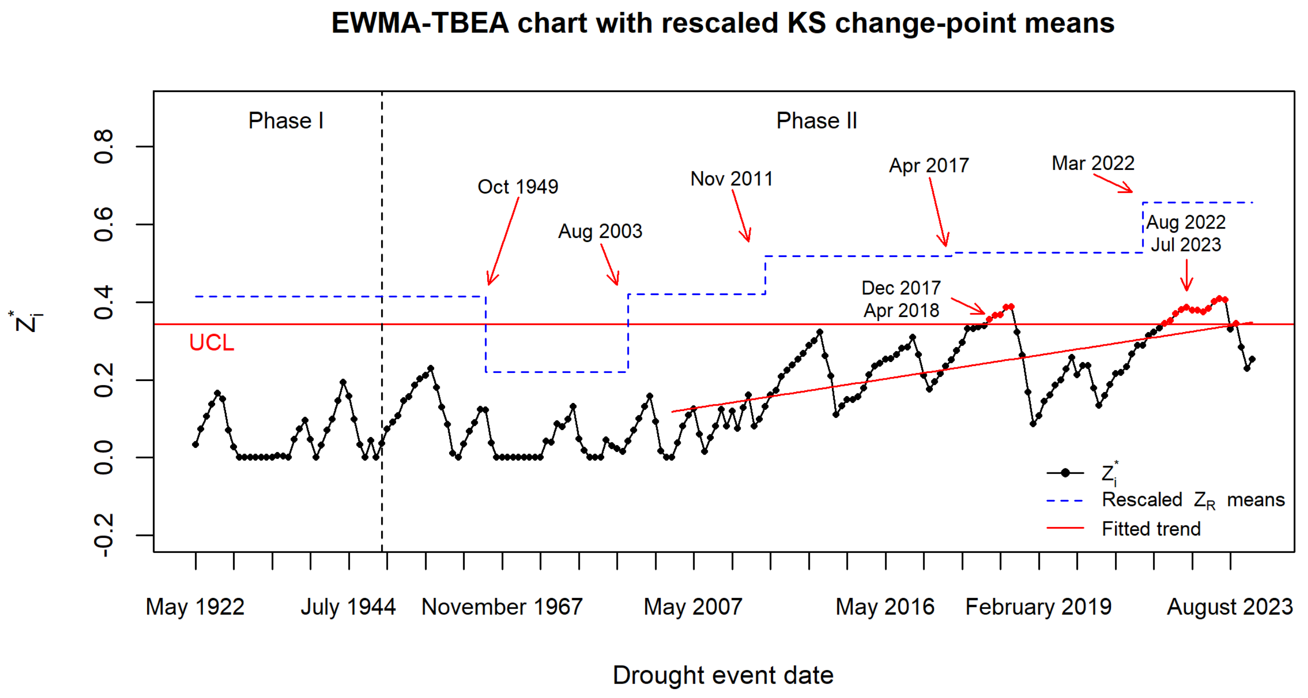

**Consistency between the results can be highlighted if the averages (appropriately rescaled) of the $Z_R$ statistic between each pair of change points are superimposed on the EWMA chart**

# Results VII



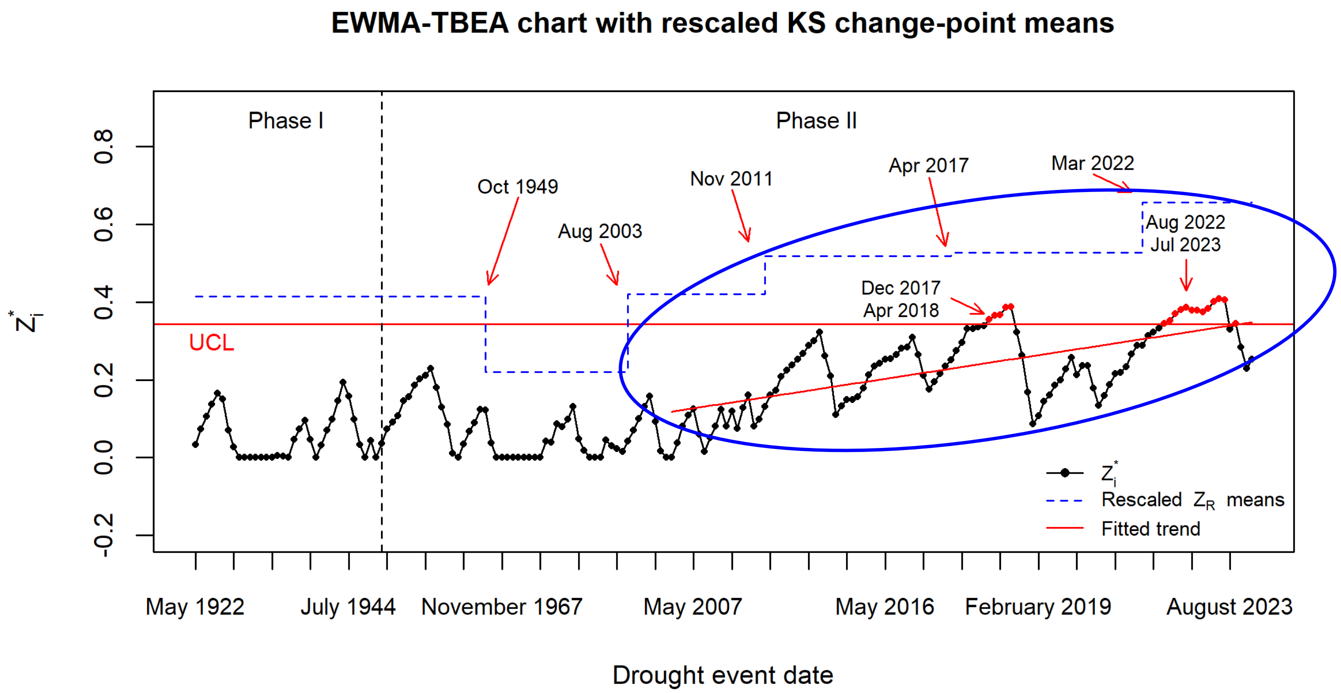

This visualization makes it possible to assess how the smoothed EWMA statistic aligns with the discrete regime shifts identified by the KS control chart

# Results VII



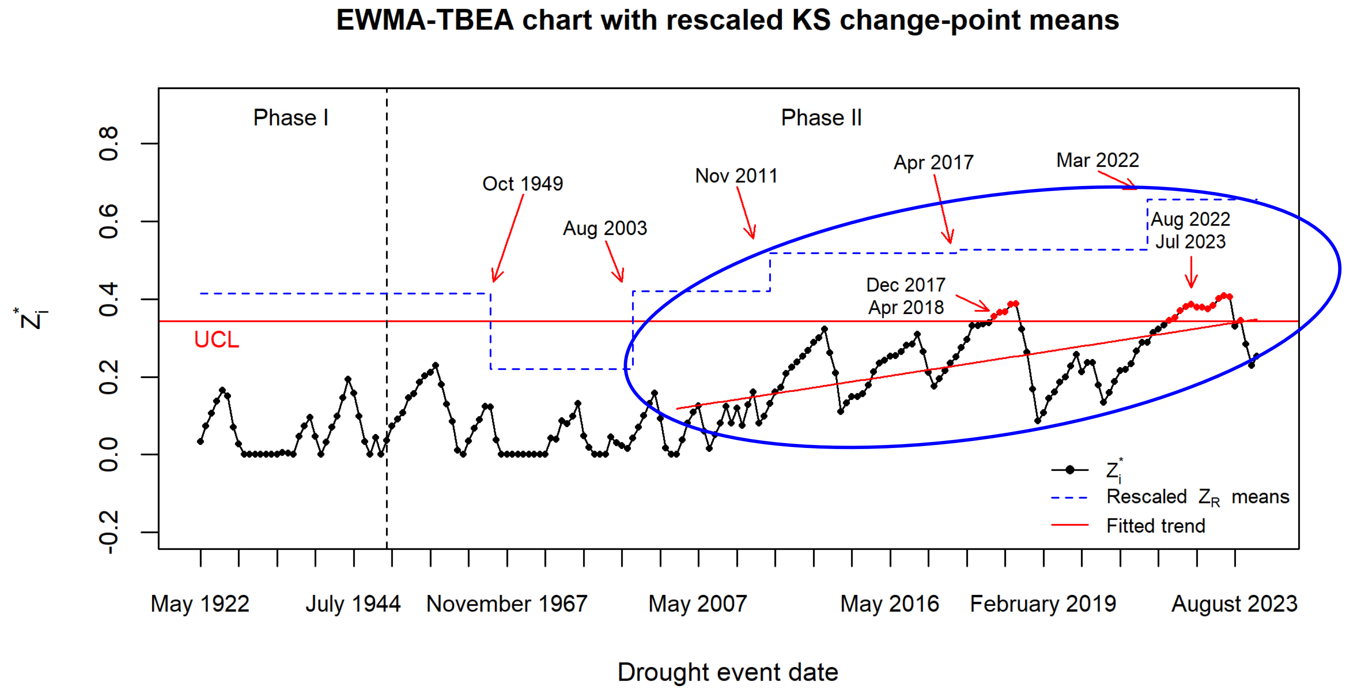

The EWMA statistic begins to deviate from its stationary baseline and trend upward around August 2003, coinciding with the first increasing change point detected by the KS chart

# Results VII




The EWMA statistic begins to deviate from its stationary baseline and trend upward around August 2003, coinciding with the first increasing change point detected by the KS chart

# Results VIII

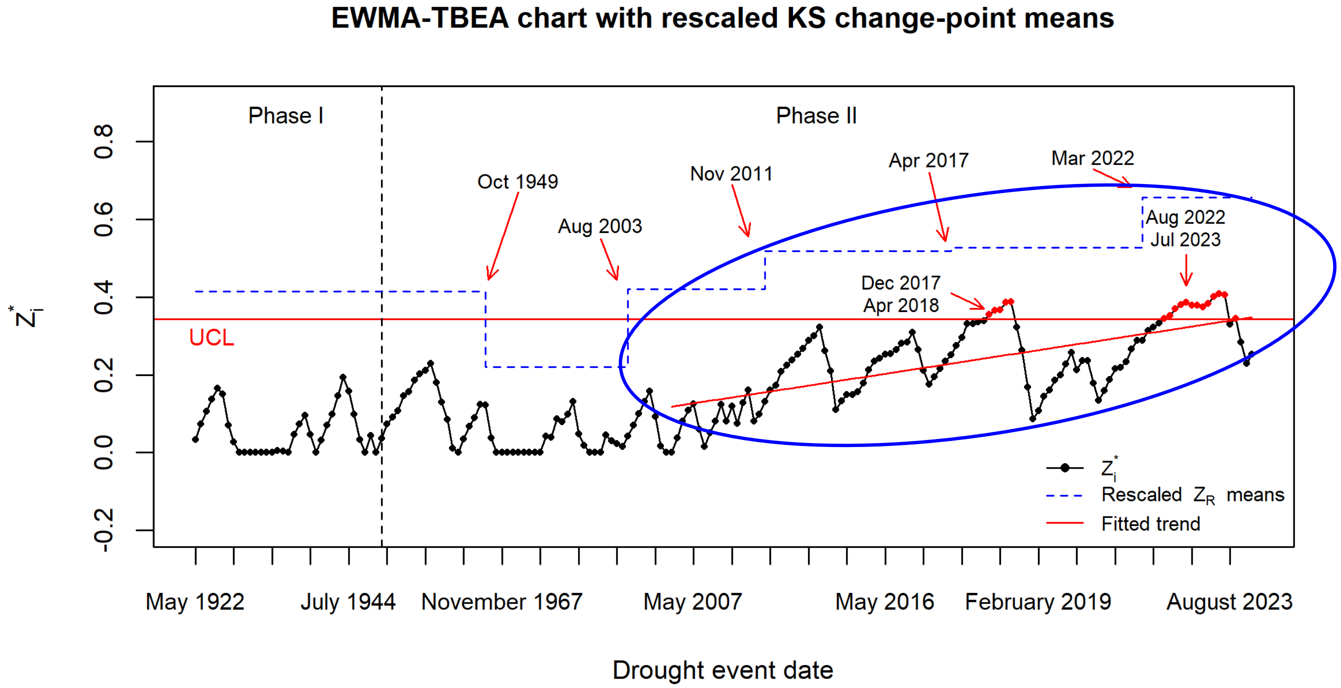



**These results show consistency between the control charts**

But does this consistency also extend to the performance of the methodologies used?

The EWMA statistic begins to deviate from its stationary baseline and trend upward around August 2003, coinciding with the first increasing change point detected by the KS chart

# Results VIII




**These results show consistency between the control charts**

**But does this consistency also extend to the performance of the methodologies used?**

The EWMA statistic begins to deviate from its stationary baseline and trend upward around August 2003, coinciding with the first increasing change point detected by the KS chart

# Performance in the out-of-control case

## Performance in the out-of-control case I

The TBEA observations were not simulated directly but were induced by the underlying process.

The simulation experiment is based on a $10^6$ replications

## Performance in the out-of-control case II

Phase I: 300 observations

Events identification

Phase-1 was performed on 35 events $E$ identified within 281 SPEI-12$_R$ observations, normally distributed, characterized by a mean $\mu_{SPEI-12_R}$ and standard deviation $\sigma_{SPEI-12_R}$

## Performance in the out-of-control case II

Phase I: 300 observations

Events identification

we generated a stationary Gaussian process representing the in-control condition.

## Performance in the out-of-control case II

Phase I: 300 observations

Events identification

Compute $T_i$ and $X_i$

EWMA chart

KS chart

The event extraction procedure was then applied exactly as in the real application

## Performance in the out-of-control case II

Phase I: 300 observations

▼

Events identification

▼

Compute $T_i$ and $X_i$

▼

EWMA chart

KS chart

The event extraction procedure was then applied exactly as in the real application

## Performance in the out-of-control case II

Phase I: 300 observations

↓

Events identification

↓

Compute $T_i$ and $X_i$

↓

EWMA chart

KS chart

The event extraction procedure was then applied exactly as in the real application

# Performance in the out-of-control case III

Phase II (1000 observations)

Mean shift $\delta$=(-0.5, -1, -1.5)

## Performance in the out-of-control case III

Phase II (1000 observations)

Mean shift $\delta$=(-0.5, -1, -1.5)

we introduced a mean shift of magnitude $\delta$, while keeping the variance unchanged

$ARL_1$

Missed alarms

# Performance in the out-of-control case III

Phase II (1000 observations)

Mean shift $\delta$=(-0.5, -1, -1.5)

ARL1

Missed alarms

out-of-control average run length (ARL1)

If, at the end of the $n$ observations, a control chart did not detect any alarm, this was classified as a missed alarm.

## Performance in the out-of-control case III

Phase II (1000 observations)

Mean shift $\delta$=(-0.5, -1, -1.5)

$ARL_1$

Missed alarms

out-of-control average run length ($ARL_1$)

If, at the end of the *n* observations, a control chart did not detect any alarm, this was classified as a missed alarm.

## Performance in the out-of-control case V

| | EWMA | KS | EWMA | KS | EWMA | KS |
|---|---|---|---|---|---|---|
| %missAL | 0.6153 | 2.2993 | 0.0000 | 0.0074 | 0.0000 | 0.0019 |
| $ARL_1$ | 26.3424 | 35.0557 | 10.2356 | 13.5661 | 7.5477 | 12.0029 |
| SD | 27.5813 | 35.8646 | 5.4972 | 6.4698 | 2.9384 | 6.2964 |
| | Results for $\delta$=-0.5 | | Results for $\delta$=-1 | | Results for $\delta$=-1.5 | |

For each control chart and each simulated scenario, the following quantities are reported:

- **%missAL**, the percentage of missed alarms;
- **$ARL_1$**, the estimated average length;
- **SDRL**, the standard deviation of the RL.

# Performance in the out-of-control case VI

| | EWMA | KS | EWMA | KS | EWMA | KS |
|---|---|---|---|---|---|---|
| **%missAL** | 0.6153 | 2.2993 | 0.0000 | 0.0074 | 0.0000 | 0.0019 |
| **$ARL_1$** | 26.3424 | 35.0557 | 10.2356 | 13.5661 | 7.5477 | 12.0029 |
| **SD** | 27.5813 | 35.8646 | 5.4972 | 6.4698 | 2.9384 | 6.2964 |
| | Results for **$\delta$=-0.5** | | Results for **$\delta$=-1** | | Results for **$\delta$=-1.5** | |

For each control chart and each simulated scenario, the following quantities are reported:

- **%missAL**, the percentage of missed alarms;
- **$ARL_1$**, the estimated average length;
- **SDRL**, the standard deviation of the RL.

# Performance in the out-of-control case VI

| | EWMA | KS | EWMA | KS | EWMA | KS |
|---|---|---|---|---|---|---|
| **%missAL** | 0.6153 | 2.2993 | 0.0000 | 0.0074 | 0.0000 | 0.0019 |
| **$ARL_1$** | 26.3424 | 35.0557 | 10.2356 | 13.5661 | 7.5477 | 12.0029 |
| **SD** | 27.5813 | 35.8646 | 5.4972 | 6.4698 | 2.9384 | 6.2964 |
| | Results for $\delta$=-0.5 | | Results for $\delta$=-1 | | Results for $\delta$=-1.5 | |

The two control charts exhibit comparable performance:

- the EWMA chart shows slightly lower percentages of missed alarms and shorter $ARL_1$ values than the KS chart
- but the overall differences are limited

# Performance in the out-of-control case VI

| | EWMA | KS | EWMA | KS | EWMA | KS |
|---|---|---|---|---|---|---|
| **%missAL** | 0.6153 | 2.2993 | 0.0000 | 0.0074 | 0.0000 | 0.0019 |
| **ARL1** | 26.3424 | 35.0557 | 10.2356 | 13.5661 | 7.5477 | 12.0029 |
| **SD** | 27.5813 | 35.8646 | 5.4972 | 6.4698 | 2.9384 | 6.2964 |
| | Results for $\delta$=-0.5 | | Results for $\delta$=-1 | | Results for $\delta$=-1.5 | |

The two control charts exhibit comparable performance:

- the EWMA chart shows slightly lower percentages of missed alarms and shorter ARL1 values than the KS chart
- but the overall differences are limited

# Performance in the out-of-control case VI

| | EWMA | KS | EWMA | KS | EWMA | KS |
|---|---|---|---|---|---|---|
| **%missAL** | 0.6153 | 2.2993 | 0.0000 | 0.0074 | 0.0000 | 0.0019 |
| **ARL1** | 26.3424 | 35.0557 | 10.2356 | 13.5661 | 7.5477 | 12.0029 |
| **SD** | 27.5813 | 35.8646 | 5.4972 | 6.4698 | 2.9384 | 6.2964 |
| | Results for **$\delta$=-0.5** | | Results for **$\delta$=-1** | | Results for **$\delta$=-1.5** | |

The two control charts exhibit comparable performance:

- the EWMA chart shows slightly lower percentages of missed alarms and shorter ARL1 values than the KS chart
- but the overall differences are limited

# Performance in the out-of-control case VI

| | EWMA | KS | EWMA | KS | EWMA | KS |
|---|---|---|---|---|---|---|
| **%missAL** | 0.6153 | 2.2993 | 0.0000 | 0.0074 | 0.0000 | 0.0019 |
| **ARL1** | 26.3424 | 35.0557 | 10.2356 | 13.5661 | 7.5477 | 12.0029 |
| **SD** | 27.5813 | 35.8646 | 5.4972 | 6.4698 | 2.9384 | 6.2964 |
| | Results for $\delta$=-0.5 | | Results for $\delta$=-1 | | Results for $\delta$=-1.5 | |

The two control charts exhibit comparable performance:

- the EWMA chart shows slightly lower percentages of missed alarms and shorter $ARL_1$ values than the KS chart
- but the overall differences are limited

# Performance in the out-of-control case VI

| | EWMA | KS | EWMA | KS | EWMA | KS |
|---|---|---|---|---|---|---|
| **%missAL** | 0.6153 | 2.2993 | 0.0000 | 0.0074 | 0.0000 | 0.0019 |
| **$ARL_1$** | 26.3424 | 35.0557 | 10.2356 | 13.5661 | 7.5477 | 12.0029 |
| **SD** | 27.5813 | 35.8646 | 5.4972 | 6.4698 | 2.9384 | 6.2964 |
| | Results for **$\delta$=-0.5** | | Results for **$\delta$=-1** | | Results for **$\delta$=-1.5** | |

The two control charts exhibit comparable performance:

- the EWMA chart shows slightly lower percentages of missed alarms and shorter $ARL_1$ values than the KS chart
- but the overall differences are limited

# Perturbation-based sensitivity analysis

# Perturbation-based sensitivity analysis I

## Objective

Assess the stability of the estimated change-point locations

## Procedure

- generate K=1000 perturbed realizations of the observed TBEA statistic
- adding Gaussian perturbations with standard deviation equal to 5% of the empirical SD
- re-estimate the change points for each realization.

# Perturbation-based sensitivity analysis I

## Objective

Assess the stability of the estimated change-point locations

## Procedure

- generate K=1000 perturbed realizations of the observed TBEA statistic

- adding Gaussian perturbations with standard deviation equal to 5% of the empirical SD

- re-estimate the change points for each realization.

# Perturbation-based sensitivity analysis I

## Objective

Assess the stability of the estimated change-point locations

## Procedure

- generate K=1000 perturbed realizations of the observed TBEA statistic

- adding Gaussian perturbations with standard deviation equal to 5% of the empirical standard deviation of the observed series

- re-estimate the change points for each realization.

# Perturbation-based sensitivity analysis I

## Objective

Assess the stability of the estimated change-point locations

## Procedure

- generate K=1000 perturbed realizations of the observed TBEA statistic

↓

- adding Gaussian perturbations with standard deviation equal to 5% of the empirical standard deviation of the observed series

↓

- re-estimate the change points for each realization.

# Perturbation-based sensitivity analysis II

## Stability measure

For each estimated change point $\hat{\tau}$ the **Stability Frequency (SF)** is the proportion of perturbed datasets that relocate the change point within a tolerance window of ±6 events.

$$\mathrm{SF}\left(\hat{\tau}\right)=\frac{1}{K}\sum_{k=1}^{K} I\left(\left|\hat{\tau}^{(k)}-\hat{\tau}\right|\leq 6\right)$$

# Perturbation-based sensitivity analysis III

## Stability assessment of the detected change points

| Original CP position | Stability Frequency (SF) | Stability Assessment |
|---|---|---|
| 53 | 0.648 | moderate |
| 79 | 0.981 | very high |
| 104 | 0.985 | very high |
| 138 | 0.543 | moderate |
| 173 | 0.997 | very high |

Three change points are consistently recovered across the perturbed data sets whereas two change points show greater sensibility to small perturbations

Overall, the sensitivity analysis supports the robustness of then main structural changes identified by the KS monitoring procedure

# Perturbation-based sensitivity analysis III

## Stability assessment of the detected change points

| Original CP position | Stability Frequency (SF) | Stability Assessment |
|---|---|---|
| 53 | 0.648 | moderate |
| 79 | **0.981** | **very high** |
| 104 | **0.985** | **very high** |
| 138 | 0.543 | moderate |
| 173 | **0.997** | **very high** |

Three change points are consistently recovered across the perturbed data sets whereas two change points show greater sensibility to small perturbations

Overall, the sensitivity analysis supports the robustness of then main structural changes identified by the KS monitoring procedure

## Perturbation-based sensitivity analysis III

### Stability assessment of the detected change points

| Original CP position | Stability Frequency (SF) | Stability Assessment |
|---|---|---|
| 53 | 0.648 | moderate |
| 79 | **0.981** | **very high** |
| 104 | **0.985** | **very high** |
| 138 | 0.543 | moderate |
| 173 | **0.997** | **very high** |

Three change points are consistently recovered across the perturbed data sets whereas two change points show greater sensibility to small perturbations

Overall, the sensitivity analysis supports the robustness of then main structural changes identified by the KS monitoring procedure

# Main take-home messages

## Main take-home messages I

The EWMA TBEA and KS control charts provide **complementary information** on the behaviour of the monitored process.

- EWMA TBEA: effective for detecting persistent small shifts and long-term process evolution.
- KS Change-Point: effective for identifying abrupt distributional changes and estimating regime transitions.

Together, they provide a more complete characterization of structural process changes.

The framework is applicable to any process characterized by event occurrence times and event magnitudes.

## Main take-home messages I

The EWMA TBEA and KS control charts provide **complementary information** on the behaviour of the monitored process.

- EWMA TBEA: effective for detecting persistent small shifts and long-term process evolution.
- KS Change-Point: effective for identifying abrupt distributional changes and estimating regime transitions.

Together, they provide a more complete characterization of structural process changes.

The framework is applicable to any process characterized by event occurrence times and event magnitudes.

## Main take-home messages I

The EWMA TBEA and KS control charts provide **complementary information** on the behaviour of the monitored process.

- EWMA TBEA: effective for detecting persistent small shifts and long-term process evolution.
- KS Change-Point: effective for identifying abrupt distributional changes and estimating regime transitions.

Together, they provide a more complete characterization of structural process changes.

The framework is applicable to any process characterized by event occurrence times and event magnitudes.

## Main take-home messages I

The EWMA TBEA and KS control charts provide **complementary information** on the behaviour of the monitored process.

- EWMA TBEA: effective for detecting persistent small shifts and long-term process evolution.
- KS Change-Point: effective for identifying abrupt distributional changes and estimating regime transitions.

Together, they provide a more complete characterization of structural process changes.

The framework is applicable to any process characterized by event occurrence times and event magnitudes.

# Many thanks for your attention

Contacts:
michele.scagliarini@unibo.it

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

enbis European Network for Business and Industrial Statistics

ENBIS-26 Conference

# References I

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).

Wu, S., Castagliola, P., Celano G.: A Distribution-Free EWMA Control Chart for Monitor-ing Time-Between-Events-and-Amplitude Data. Journal of Applied Statistics, 48(3), 434-454 (2021).

Ross, G.J., Adams, N.M.: Two Nonparametric Control Charts for Detecting Arbitrary Dis-tribution Changes. Journal of Quality Technology, 44(12), 102-116 (2012).

Pascale, S. Ragone, F.: Widespread Multi-Year Droughts in Italy: Identification and Causes of Development. International Journal of Climatology, 45(8), 1-19 (2025).