Mann-Kendall tests to forecast critical transitions in ecology or epidemiology

Nils Thibeau--Sutre and Tristan Gamot
Joint work with Tom J.M. Van Dooren (CNRS researcher, iEES Paris)

INSTITUT de MATHÉMATIQUES de MARSEILLE

INSTITUTE OF ECOLOGY AND ENVIRONMENTAL SCIENCES OF PARIS iEES PARIS

CPHT Centre de Physique Théorique

Overview

1. Forecasting critical transitions
   - Critical transitions
   - Early warning signals
2. Trend testing
   - Kendall tau
   - Mann-Kendall tests
3. Results
   - Autocorrelation and distribution
   - Tests robustness

Critical transitions in ecology

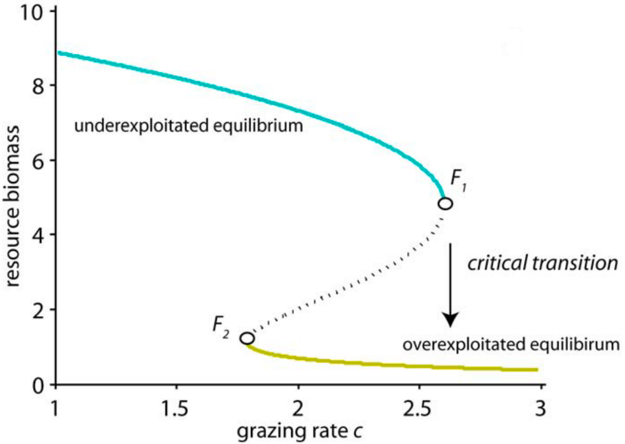


Figure: Bifurcation diagram of an ecological model of a logistically growing resource under harvesting. Resource biomass as a function of grazing rate c. Critical transitions through a fold bifurcation (F1 or F2). Solid lines represent locally stable equilibria. Figure from [Dak+12].

Critical transitions in epidemiology

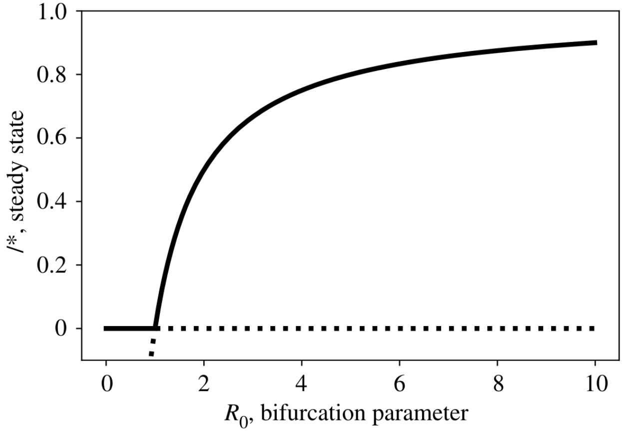


Figure: Bifurcation diagram of a compartmental SIS epidemiological model. Critical transition through a transcritical bifurcation at $R_0 = 1$. Solid lines represent locally stable equilibria. Figure from [Sou+21].

Other fields

1. Climate: abrupt shifts in ocean circulation or climate [Len11; Len+24]
2. Medicine: asthma attacks [Ven+05], epileptic seizures [MST03], psychiatric disorders [Lee+14]
3. Finance: balanced market to a financial crisis [MLS08]

Slow-fast systems

Mathematical framework to consider critical transitions [Kue11]. Consider the parametrized one-dimensional ordinary differential equation with a slowly varying one-dimensional parameter:

$$\frac{\mathrm{d}x}{\mathrm{d}t} = x' = f(x, r)$$
$$\frac{\mathrm{d}r}{\mathrm{d}t} = r' = \epsilon g(x, r)$$

where $x \in \mathbb{R}$ is the phase space variable, $r \in \mathbb{R}$ represents the one-dimensional parameter, and $0 < \epsilon \ll 1$ is a small parameter describing the **time scale separation**.

Example: the fold bifurcation

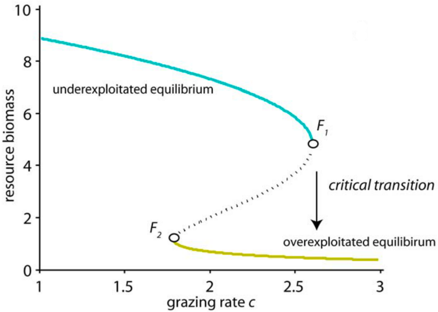


Figure: Harvest model

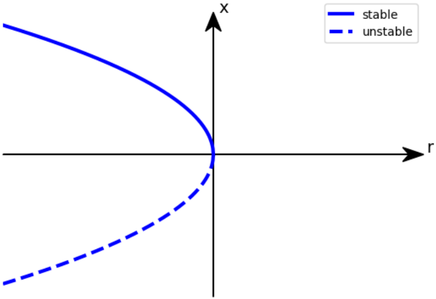


Figure: Fold bifurcation

Near a generic bifurcation, the flow is locally topologically equivalent to the flow of the normal form of the bifurcation [Kuz98].

Slow-fast systems with a local bifurcation

$$x' = f(x, r)$$
$$r' = \epsilon g(x, r)$$

Local bifurcation	Normal form
Fold	$f(x,r) = -r - x^2$
Transcritical	$f(x,r) = rx - x^2$
Pitchfork	$f(x,r) = rx + \mu x^3$ with $\mu > 0$ (subcritical case) or $\mu < 0$ (supercritical case)

Table: One-dimensional local bifurcations in continuous time, as well as their normal forms.

Near a generic bifurcation, the flow is locally topologically equivalent to the flow of the normal form of the bifurcation [Kuz98].

Parameter dynamics

In many cases it is assumed that the parameter dynamics is decoupled from phase space dynamics or even $g = 1$.

By Fenichel's theorem [Fen79], the dynamics of slow-fast models will approach those of the models where the bifurcation parameter is constant in the limit of the slow parameter approaching zero ($\epsilon \to 0$).

Parameter dynamics

In many cases it is assumed that the parameter dynamics is decoupled from phase space dynamics or even $g = 1$.

By Fenichel's theorem [Fen79], the dynamics of slow-fast models will approach those of the models where the bifurcation parameter is constant in the limit of the slow parameter approaching zero ($\epsilon \to 0$).

Therefore, we study an ordinary differential equation parametrized by a one-dimensional parameter:

$$x' = f(x, r)$$

Results hold when the bifurcation parameter changes over a sufficiently long time scale.

Stochastic dynamical systems

Add stochastic effects to capture the role of noise, leading to the following Itô stochastic differential equation:

$$\mathrm{d}x = f(x, r)\mathrm{d}t + \sqrt{h(x, r)^2}\mathrm{d}W$$

where h^2 represents the diffusion coefficient and W is a standard Brownian motion.

Stochastic dynamical systems

Add stochastic effects to capture the role of noise, leading to the following Itô stochastic differential equation:

$$\mathrm{d}x = f(x, r)\mathrm{d}t + \sqrt{h(x, r)^2}\mathrm{d}W$$

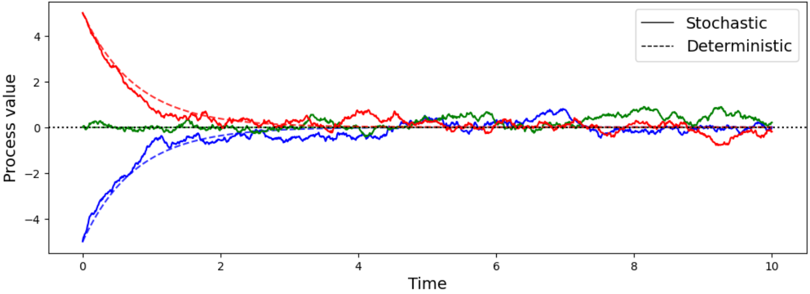
where h^2 represents the diffusion coefficient and W is a standard Brownian motion.

In the neighborhood of a locally stable equilibrium x_s, the stochastic differential equation of the deviation $\tilde{x}(t) = x(t) - x_s$ is at the leading order:

$$\mathrm{d}\tilde{x} \approx f'(x_s, r)\tilde{x}\mathrm{d}t + \sqrt{h(x_s, r)^2}\mathrm{d}W$$

Thus, the deviation is an **Ornstein–Uhlenbeck (O-U) process**.

Properties of the O-U process



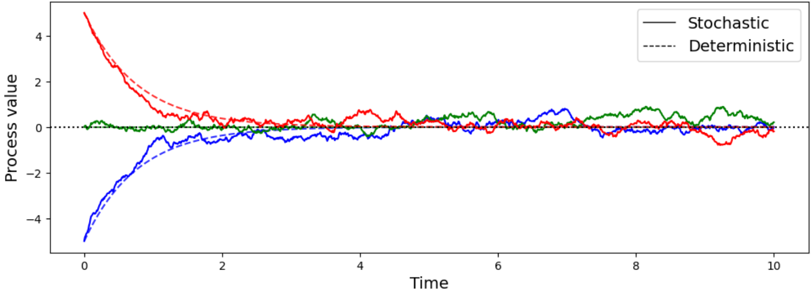
Assuming **additive** Gaussian white noise of amplitude $h(x, r) = \sigma$ we can derive the variance and autocorrelation function after a transient:

$$Var(\tilde{x}) = \frac{-\sigma^2}{2f'(x_s, r)}$$

$$\rho_{\tilde{x}}(\tau) = e^{f'(x_s, r)\tau}$$

(more generally, $f'(x_s, r)$ is the dominant eigenvalue)

Properties of the O-U process



$$Var(\tilde{x}) = \frac{-\sigma^2}{2f'(x_s, r)}$$

$$\rho_{\tilde{x}}(\tau) = e^{f'(x_s, r)\tau}$$

As $r \to 0$ (bifurcation is approached), $f'(x_s, r) \to 0^-$, **causing variance and autocorrelation to increase to infinity and one respectively.**

Early warning signals of critical transitions

More generally, comes from the fact that the rate at which the system returns to its stable equilibrium following perturbations diminishes as the critical transition gets closer, what is called **critical slowing down** [Sch+09].

Early warning signals of critical transitions

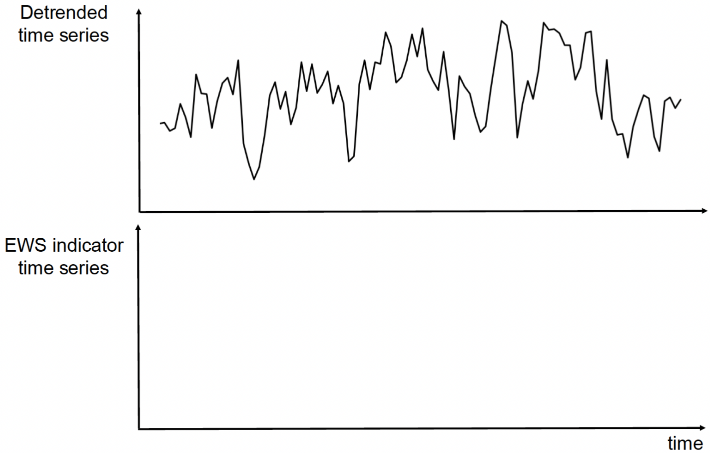
More generally, comes from the fact that the rate at which the system returns to its stable equilibrium following perturbations diminishes as the critical transition gets closer, what is called **critical slowing down** [Sch+09].

Can be used as an indicator of how far the system is away from a critical transition [Kue11]. Leads to the so-called **early warning signals** (EWS):

- autocorrelation at lag-1 [Sch+09]
- variance or standard deviation [Sch+09]
- coefficient of variation [Dak+12]
- peak power spectrum [BBA20]
- dominant eigenvalue reconstruction [Grz+23]
- and many others.

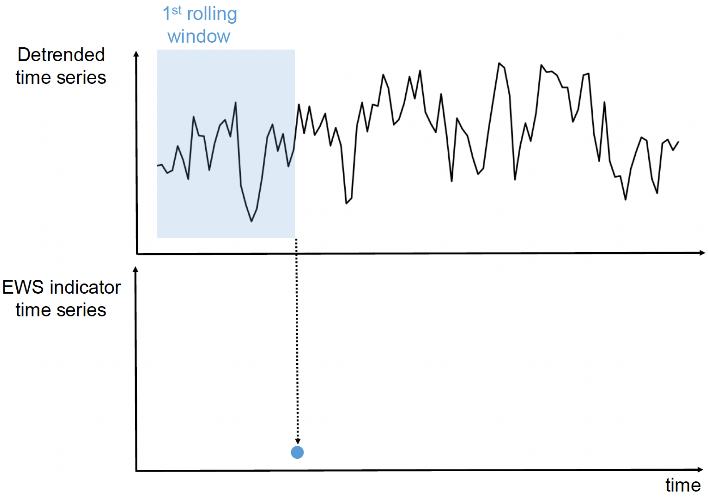
Early warning signals estimation in practice

EWS are estimated within rolling windows.



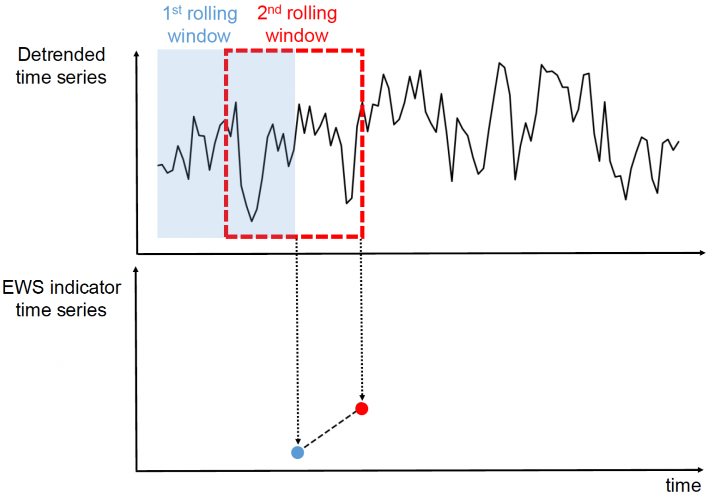
Early warning signals estimation in practice

EWS are estimated within rolling windows.



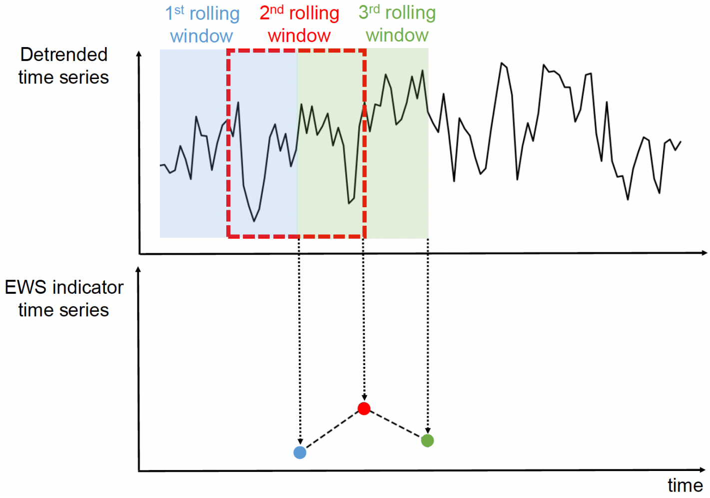
Early warning signals estimation in practice

EWS are estimated within rolling windows.



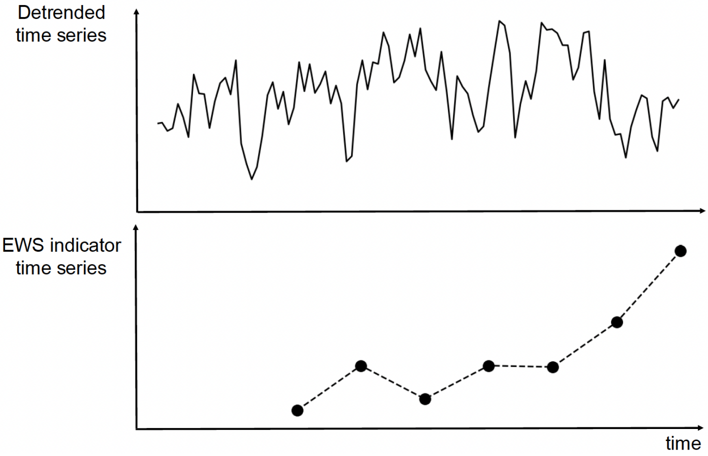
Early warning signals estimation in practice

EWS are estimated within rolling windows.



Early warning signals estimation in practice

EWS are estimated within rolling windows.



Early warning signals estimation in practice

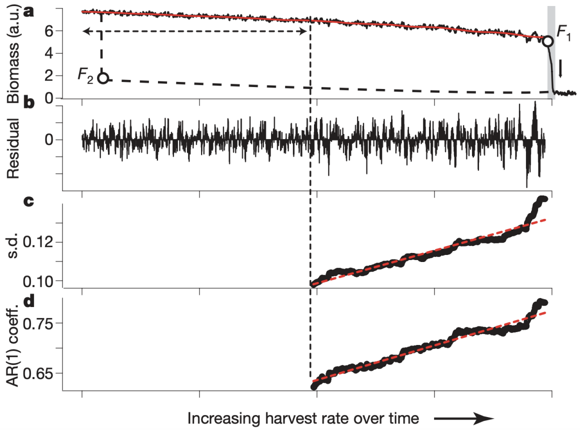


Figure: From [Sch+09]

How to assess for a significant trend in the EWS time series?

How to evaluate the significance of trends?

Method	Description	Advantages
Bootstrapping	Generate surrogate datasets with same correlation structure and proba. distrib. but model-dependent	Explicit hypothesis

How to evaluate the significance of trends?

Method	Description	Advantages
Bootstrapping	Generate surrogate datasets with same correlation structure and proba. distrib. but model-dependent	Explicit hypothesis
Mann-Kendall tests	Built-in tests associated with the Mann-Kendall tau	Non-parametric Computationally faster

Mann-Kendall tests were promoted recently to evaluate the significance of the EWS trends [CGE22]. Already applied in this context [Ton+14; Geo+20; WWS20; Bos+22; Ism+22; Hel+23].

The Mann-Kendall tau coefficient

Let $(X_i)_{1 \leq i \leq n}$ be a set of random variables identically distributed.
The Mann-Kendall tau is defined as the random variable[1]:

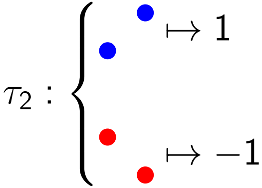
$$\tau_n := \frac{1}{\binom{n}{2}} \sum_{1 \leq i < j \leq n} \operatorname{sgn}(X_j - X_i) \tag{1}$$

For a realization $x = (x_i)_{1 \leq i \leq n}$, this defines the value:

$$\tau_n^x = \frac{1}{\binom{n}{2}} \sum_{1 \leq i < j \leq n} \operatorname{sgn}(x_j - x_i) \tag{2}$$

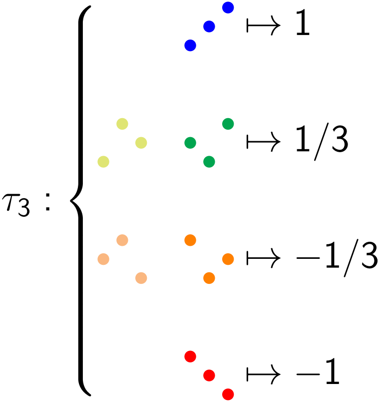
[1] Maurice G Kendall. "A new measure of rank correlation". In: *Biometrika* 30.1 (1938), pp. 81–93; Henry B Mann. "Nonparametric tests against trend". In: *Econometrica: Journal of the econometric society* (1945), pp. 245–259.

Exact values for n = 2



$$\tau_n := \frac{1}{\binom{n}{2}} \sum_{1 \le i < j \le n} \operatorname{sgn}(X_j - X_i)$$

Exact values for n = 3



$$\tau_n := \frac{1}{\binom{n}{2}} \sum_{1 \le i < j \le n} \mathrm{sgn}(X_j - X_i)$$

Independent case

Theorem [Kendall '48] - distribution convergence

Assume that $X_i \overset{iid}{\sim} X$, then :

$$\frac{\tau_n}{\sqrt{\mathbb{V}(\tau_n)}} \sim \sqrt{\frac{9n}{4}}\tau_n \xrightarrow[n\to\infty]{d} \mathcal{N}(0,1) \tag{3}$$

The exact formula for the standard deviation of the random variable is given by:

$$\sigma_n := \sqrt{\mathbb{V}(\tau_n)} = \sqrt{\frac{2}{9}\frac{(2n-5)}{n(n-1)}} \sim \sqrt{\frac{4}{9n}} \tag{4}$$

Z score - Mann-Kendall tests

Thanks to the previous theorem, we know how the data behave when they are sampled from i.i.d. random variables. For a given series, we defines its score Z^x by :

$$Z^x \simeq \tau_n^x / \sigma_n \tag{5}$$

Z score - Mann-Kendall tests

$$Z^x \simeq \tau_n^x / \sigma_n \quad (5)$$

To assess the significance of the trend, the corresponding *p-value* for the two-tailed test is:

$$p = \sqrt{\frac{2}{\pi}} \int_{|Z^x|}^{+\infty} e^{-t^2/2} \mathrm{dt} \quad (6)$$

If the p-value is small enough, the trend is quite unlikely to be caused by random sampling. This is the original Mann-Kendall test.

In practice $\mathcal{H}_0$ is understood as the following :

$\mathcal{H}_0$: There is no trend / The trend observed is due to randomness

Modified Mann-Kendall tests for autocorrelated data (1/2)

Autocorrelation in data influences the standard deviation of the Mann-Kendall tau.

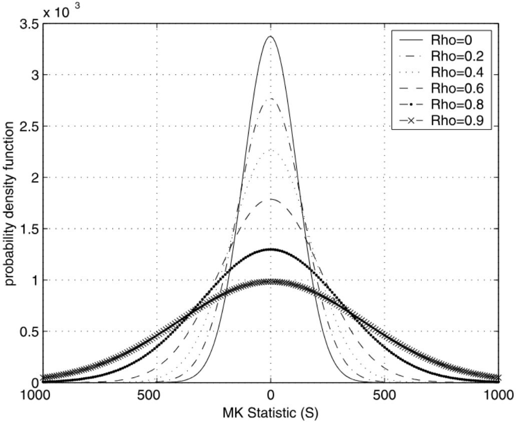


Figure: Data sampled from AR(1) processes with autocorrelation at lag-1 parameter ρ. Figure from [YW04].

Modified Mann-Kendall tests for autocorrelated data (2/2)

Two tests were elaborated to take into account autocorrelation [HR98; YW04].

The score variable is then modified:

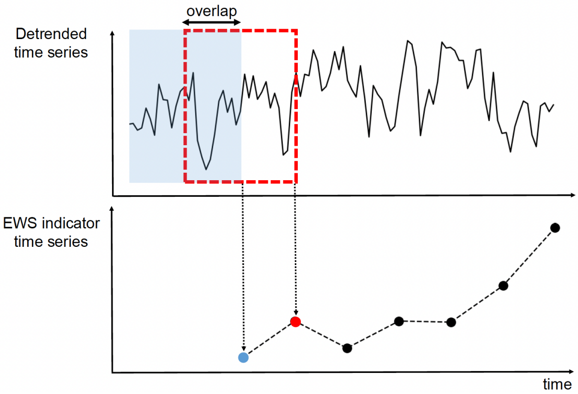
$$Z^x \simeq \tau_n^x / \sigma_n^*, \qquad \sigma_n^* = \sigma_n(1 + \rho_n^x) \tag{7}$$

Similarly, it is assumed that $Z \sim \mathcal{N}(0, 1)$ under $\mathcal{H}_0$. This gives the modified Mann-Kendall tests.

Application to EWS time series

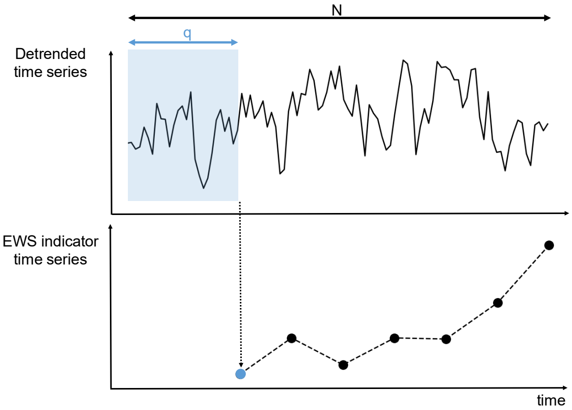
Sources of autocorrelation

- Autocorrelation in the original time series.
- Autocorrelation brought by the overlapping rolling window method to estimate EWS indicators.



Sources of autocorrelation

- Autocorrelation in the original time series.
- Autocorrelation brought by the overlapping rolling window method to estimate EWS indicators.



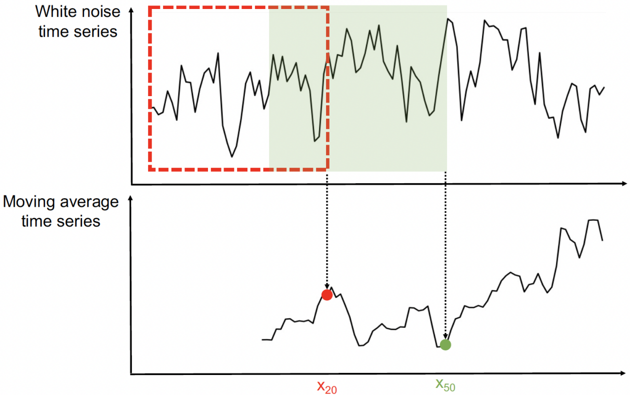
We denote by $\alpha = \frac{q}{N}$ the relative size of the rolling windows.

Case of independent original time series.

- ~~Autocorrelation in the original time series.~~
- Autocorrelation brought by the overlapping rolling window method to estimate EWS indicators.

Case of independent original time series.

- ~~Autocorrelation in the original time series.~~
- Autocorrelation brought by the overlapping rolling window method to estimate EWS indicators.



Exact calculation of variance - Moving Average

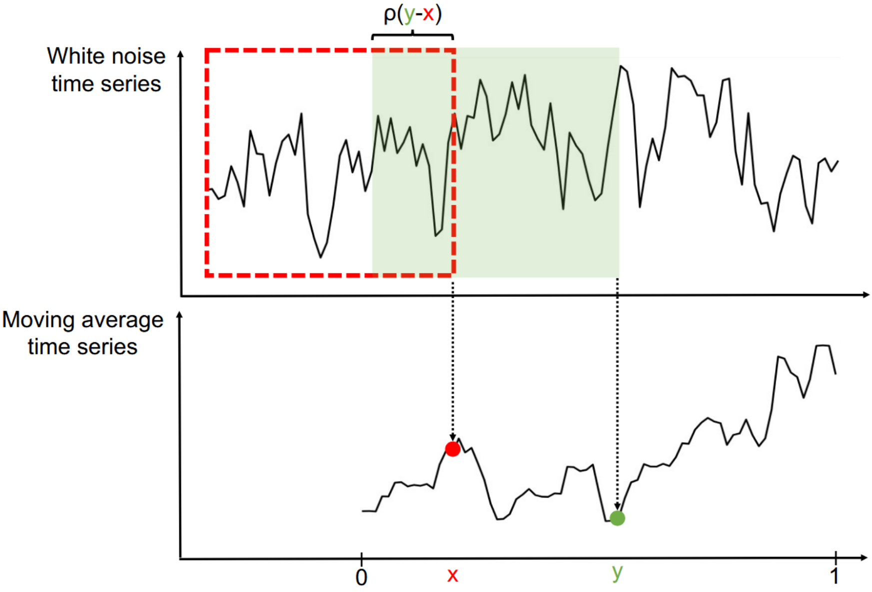
Let the X_i follow a moving-average process, then the X_i are Gaussian random variables and:

$$\mathbb{V}(\tau_n) = \frac{1}{\binom{n}{2}^2} \sum_{1 \leq i < j \leq n} \sum_{1 \leq k < l \leq n} \mathrm{sgn}(X_j - X_i)\mathrm{sgn}(X_k - X_l)$$

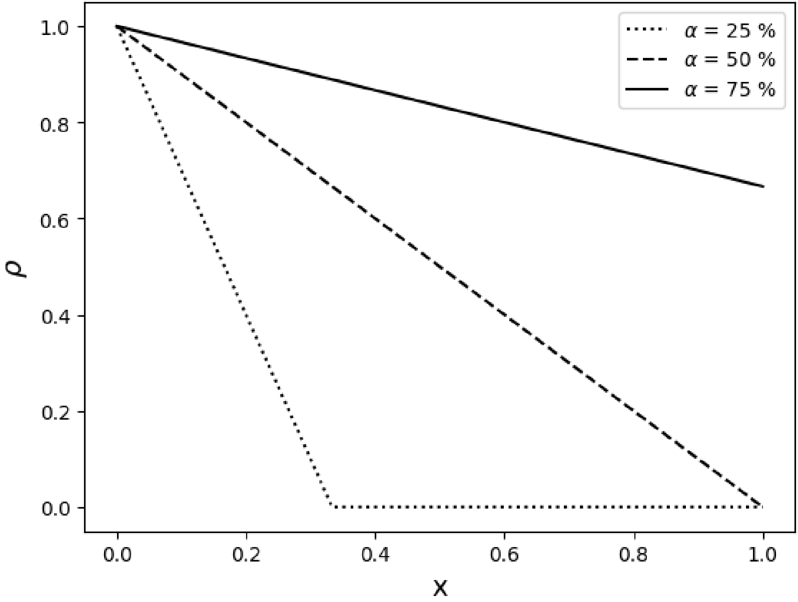
$$\mathbb{V}(\tau_n) = \frac{1}{\binom{n}{2}^2} \sum_{1 \leq i < j \leq n} \sum_{1 \leq k < l \leq n} \frac{2}{\pi} \arcsin\left(\mathrm{corr}\left(X_j - X_i, X_l - X_k\right)\right)$$

using Greiner's equality [Gre09] and where corr is the Pearson correlation coefficient.

Exact calculation of variance - Moving Average



Exact calculation of variance - Moving Average



Exact calculation of variance - General case

For $n \in \mathbb{N}$, let $(X_i^{(n)})_{1 \leq i \leq n}$ be Gaussian random variables. We assume the following:

Assumption: correlation function

$\exists \rho : [0,1] \to [-1,1]$ such that, $\forall x, y \in]0,1[$,

$$\lim_{n \to \infty} \mathrm{corr}(X_{\lceil nx \rceil}^{(n)}, X_{\lceil ny \rceil}^{(n)}) = \rho\left(|x-y|\right) \ , \tag{8}$$

Exact calculation of variance - General case

Assumption: correlation function

$\exists \rho : [0,1] \to [-1,1]$ such that, $\forall x, y \in]0,1[$,

$$\lim_{n\to\infty} \mathrm{corr}(X^{(n)}_{\lceil nx \rceil}, X^{(n)}_{\lceil ny \rceil}) = \rho\left(|x-y|\right) , \tag{8}$$

Under the previous assumption, we have an explicit formula[2]:

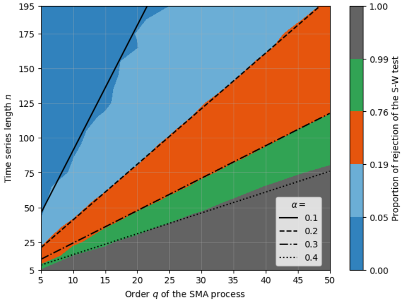
Lemma: renormalization

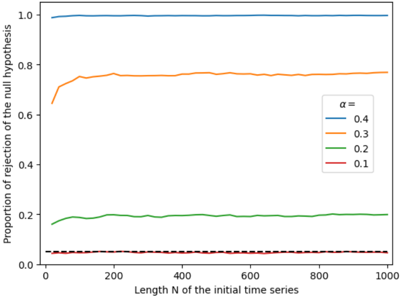
$$\lim_{n\to\infty} \mathbb{V}(\tau_n) = \frac{16}{\pi} \int_0^1 (1-z) \int_0^z \int_0^y f(x,y,z) \mathrm{d}x \mathrm{d}y \mathrm{d}z \, , \tag{9}$$

where $f(x,y,z) = \lim_{n\to\infty} \mathrm{corr}(X^{(n)}_{\lceil nx \rceil} - X^{(n)}_1, X^{(n)}_{\lceil nz \rceil} - X^{(n)}_{\lceil ny \rceil})$

[2] Tristan Gamot, Nils Thibeau–Sutre, and Tom J.M. Van Dooren. "On the Gaussian distribution of the Mann-Kendall tau statistic in the case of autocorrelated data". unpublished.

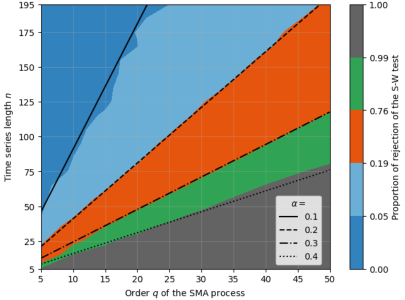
On the Gaussian distribution of the Mann-Kendall tau

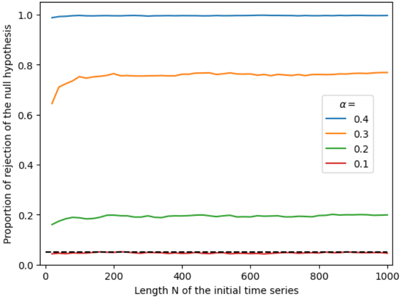


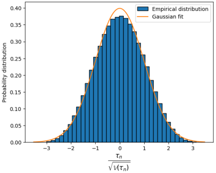


Provide scaling laws to determine whether the Gaussian approximation remains valid for finite-length time series generated by stationary AR(1) and SMA processes.

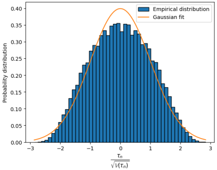
On the Gaussian distribution of the Mann-Kendall tau



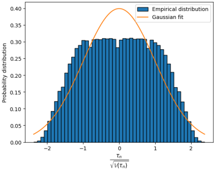




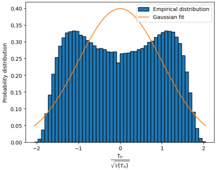
(a) $\alpha = 0.1$



(b) $\alpha = 0.2$

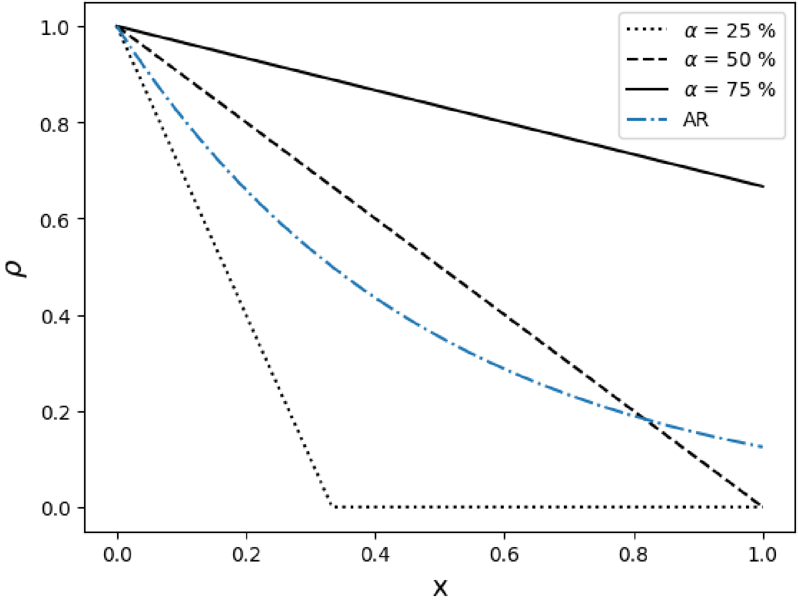


(c) $\alpha = 0.3$



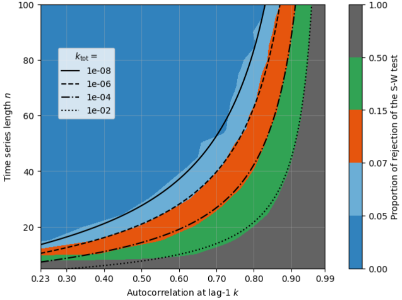
(d) $\alpha = 0.4$

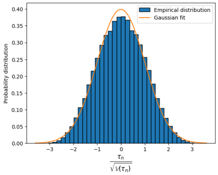
Exact calculation of variance - AR(1)



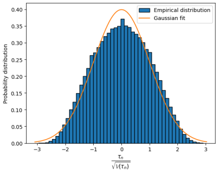
On the Gaussian distribution of the Mann-Kendall tau

$$k_{\text{tot}} = k^{(n-1)}$$

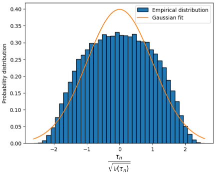




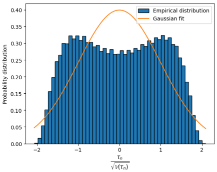
(a) $k_{\text{tot}} = 10^{-8}$



(b) $k_{\text{tot}} = 10^{-4}$



(c) $k_{\text{tot}} = 10^{-2}$

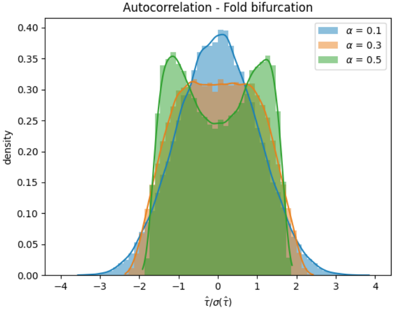


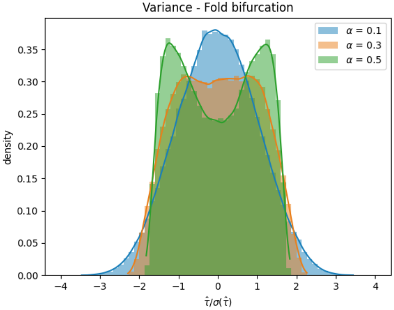
(d) $k_{\text{tot}} = 0.7$

What if the original time series are not independent?

- Autocorrelation in the original time series.
- Autocorrelation brought by the overlapping rolling window method to estimate EWS indicators.

Distributions are strongly influenced by the relative size of the rolling windows





Minimal sensitivity to the bifurcation type, the EWS indicator, the type and intensity of noise, and the sample size.

The distribution is principally governed by the use of rolling windows.

Hypothesized distributions of the Mann-Kendall tests

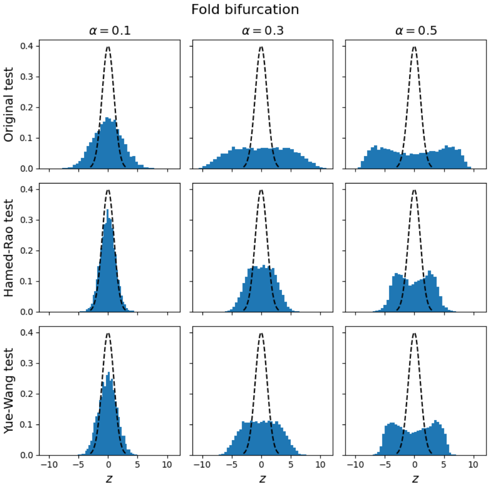


Figure: Comparison of the empirical distribution of the test statistics z of the original test (first row) the Hamed and Rao test (second row) and the Yue and Wang test (third row) with the corresponding hypothesized distributions. The z statistics are the Mann-Kendall tau renormalized differently for each test.

Inflated type I error rates

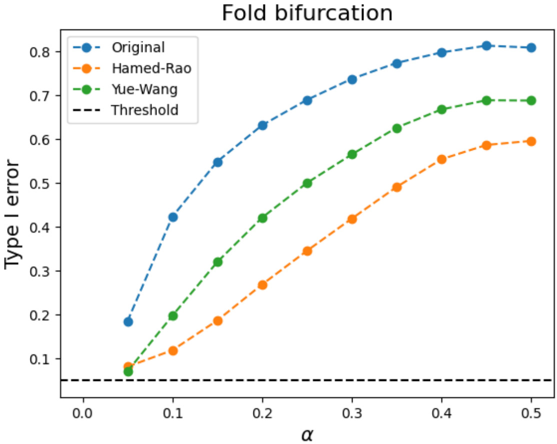
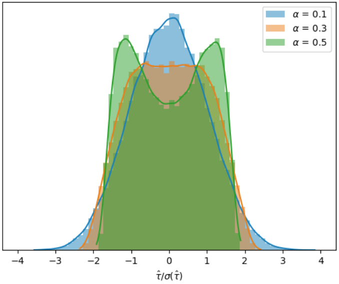
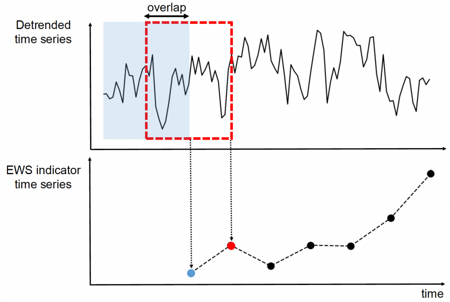


Figure: Rejection rates of the null hypothesis of no trend at the statistical significance level of 5%. The black line is the significance threshold of 5%. Each original time series is of length $N = 100$ and is sampled from the Fold normal form.

A new Mann-Kendall test?

- Possible to adapt the variance for $\alpha \leq 0.1$ (Gaussian approximation).
- Possible to reduce the window overlap but minimal effects on type I error rates.
- Similar effects are observed for other rank correlation coefficients (such as the Spearman correlation coefficient).





Take-home messages

- **Mann-Kendall tests are not suitable to assess the significance of trends when testing for critical transitions** because of the type of autocorrelation brought by the rolling window method to estimate EWS indicators.
- Better use alternative tests that make no assumption on the Mann-Kendall tau distribution, or use methods to forecast critical transitions that do not use it to quantify trends.

Merci pour votre attention !

Remerciements:

(a) Tom Van Dooren

(b) Jean-René Chazottes

(c) Mickael Kopp

References I

[BBA20] Thomas M Bury, Chris T Bauch, and Madhur Anand. “Detecting and distinguishing tipping points using spectral early warning signals”. In: *Journal of the Royal Society Interface* 17.170 (2020), p. 20200482.

[Bos+22] Fionneke M Bos et al. “Anticipating manic and depressive transitions in patients with bipolar disorder using early warning signals”. In: *International Journal of Bipolar Disorders* 10.1 (2022). Article 12.

[CGE22] Shiyang Chen, Amin Ghadami, and Bogdan I Epureanu. “Practical guide to using Kendall’s τ in the context of forecasting critical transitions”. In: *Royal Society Open Science* 9.7 (2022), p. 211346.

[Dak+12] Vasilis Dakos et al. “Methods for detecting early warnings of critical transitions in time series illustrated using simulated ecological data”. In: *PloS one* 7.7 (2012), e41010.

References II

[Fen79] Neil Fenichel. “Geometric singular perturbation theory for ordinary differential equations”. In: *Journal of differential equations* 31.1 (1979), pp. 53–98.

[Geo+20] Sandip V George et al. “Early warning signals indicate a critical transition in Betelgeuse”. In: *Astronomy & Astrophysics* 640 (2020). Article L21.

[Gre09] R Greiner. “Über das fehlersystem der kollektivmasslehre”. In: *Zeitschift für Mathematik und Physik* 57 (1909), pp. 121–158.

[Grz+23] Florian Grziwotz et al. “Anticipating the occurrence and type of critical transitions”. In: *Science Advances* 9.1 (2023), eabq4558.

[GTV] Tristan Gamot, Nils Thibeau–Sutre, and Tom J.M. Van Dooren. “On the Gaussian distribution of the Mann-Kendall tau statistic in the case of autocorrelated data”. unpublished.

References III

[Hel+23] Marieke A Helmich et al. "Detecting impending symptom transitions using early-warning signals in individuals receiving treatment for depression". In: *Clinical Psychological Science* 11.6 (2023), pp. 994–1010.

[HR98] Khaled H Hamed and A Ramachandra Rao. "A modified Mann-Kendall trend test for autocorrelated data". In: *Journal of hydrology* 204.1-4 (1998), pp. 182–196.

[Ism+22] Mohd Sabri Ismail et al. "Early warning signals of financial crises using persistent homology". In: *Physica A: Statistical Mechanics and Its Applications* 586 (2022). Article 126459.

[Ken38] Maurice G Kendall. "A new measure of rank correlation". In: *Biometrika* 30.1 (1938), pp. 81–93.

References IV

[Kue11] Christian Kuehn. "A mathematical framework for critical transitions: Bifurcations, fast–slow systems and stochastic dynamics". In: *Physica D: Nonlinear Phenomena* 240.12 (2011), pp. 1020–1035.

[Kuz98] Yuri Kuznetsov. *Elements of applied bifurcation theory*. Vol. 112. Springer, 1998.

[Lee+14] Ingrid A van de Leemput et al. "Critical slowing down as early warning for the onset and termination of depression". In: *Proceedings of the National Academy of Sciences* 111.1 (2014), pp. 87–92.

[Len+24] Timothy M Lenton et al. "Remotely sensing potential climate change tipping points across scales". In: *nature communications* 15.1 (2024), p. 343.

[Len11] Timothy M Lenton. "Early warning of climate tipping points". In: *Nature climate change* 1.4 (2011), pp. 201–209.

References V

[Man45] Henry B Mann. “Nonparametric tests against trend”. In: *Econometrica: Journal of the econometric society* (1945), pp. 245–259.

[MLS08] Robert M May, Simon A Levin, and George Sugihara. “Ecology for bankers”. In: *Nature* 451.7181 (2008), pp. 893–894.

[MST03] Patrick E McSharry, Leonard A Smith, and Lionel Tarassenko. “Prediction of epileptic seizures: are nonlinear methods relevant?” In: *Nature medicine* 9.3 (2003), pp. 241–242.

[Sch+09] Marten Scheffer et al. “Early-warning signals for critical transitions”. In: *Nature* 461.7260 (2009), pp. 53–59.

[Sou+21] Emma Southall et al. “Early warning signals of infectious disease transitions: a review”. In: *Journal of the Royal Society Interface* 18.182 (2021), p. 20210555.

References VI

[Ton+14] Ji-Long Tong et al. "Early warning signals of abrupt temperature change in different regions of China over the past 50 years". In: *Chinese Physics B* 23.4 (2014). Article 049201.

[Ven+05] Jose G Venegas et al. "Self-organized patchiness in asthma as a prelude to catastrophic shifts". In: *Nature* 434.7034 (2005), pp. 777–782.

[WWS20] Chenghao Wang, Zhi-Hua Wang, and Linda Sun. "Early-warning signals for critical temperature transitions". In: *Geophysical Research Letters* 47.14 (2020). Article e2020GL088503.

[YW04] Sheng Yue and ChunYuan Wang. "The Mann-Kendall test modified by effective sample size to detect trend in serially correlated hydrological series". In: *Water resources management* 18.3 (2004), pp. 201–218.