International Conference on Time Series and Forecasting

Alexander Schnurr
Univ. Siegen
Gran Canaria, ITISE

An ordinal procedure to detect change points in the dependence structure between non-stationary time series
(Joint work with Svenja Fischer)

Outline

Starting Point

In various applications, the **co-movement** of different data sets is a desirable property. More often, it is a **source of risk**:

- If a trader holds a portfolio of various assets which are strongly 'correlated', it is more likely that several of them fall into the abyss at the same time.
- If several catchments in a basin are affected by extreme floods at the same time, severe damages can be expected for the downstream catchments by flood superposition.
- If many requests are made to the same server at the same time, it becomes overloaded.

Monitor the Strength of Dependence

- In all of these cases, it is important to monitor the strength of the interdependence between the data sets.
- This interdependence or co-movement is often modeled using the mathematical concept of correlation.
- However, sometimes it is not clear whether the time series admit second moments. Hence, it might be the case that correlation between the random variables (of which the time series consist) is not even defined.
- Another drawback of this mathematical concept is that it is well known to measure mostly *linear* dependence whilst the dependence in our applications could be far from being linear.

Our Aim

Our aim: Detect changes within the dependence structure between data sets.

Additional problem: The time series itself might not be stationary (Think of temperature under climate change).

A Simulated Toy Example

Let us start with a **simulated example** illustrating the above idea. Consider two correlated ARMA(1,1) time series, i.e.,

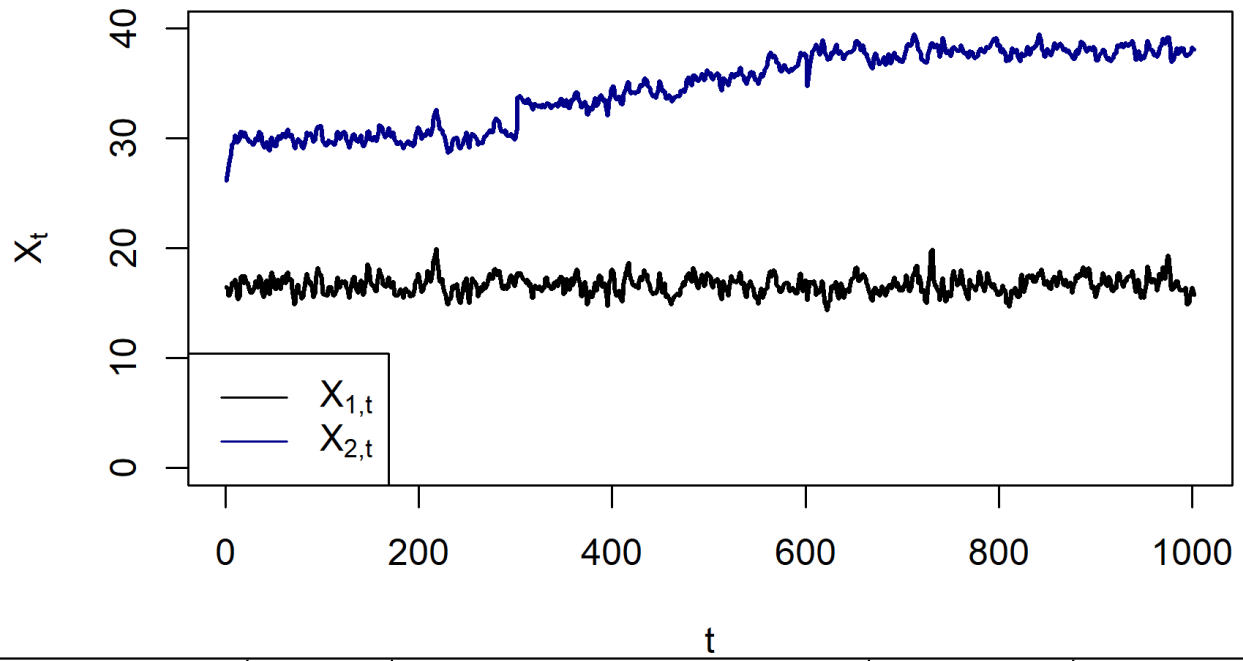
$$X_i^{(1)} = \mu_1 + \phi_1 X_{i-1}^{(1)} + \epsilon_i^{(1)} - \theta_1 \epsilon_{i-1}^{(1)}$$
$$X_i^{(2)} = \mu_2 + \phi_2 X_{i-1}^{(2)} + \epsilon_i^{(2)} - \theta_2 \epsilon_{i-1}^{(2)},$$

where $\mu_1, \mu_2 \in \mathbb{R}$, $\phi_1, \phi_2, \theta_1, \theta_2 \in (0,1)$ and $\epsilon \sim \mathcal{N}_2(\mu, \sigma)$ with

$$\Sigma = \begin{pmatrix} \sigma_1 & \rho \\ \rho & \sigma_2 \end{pmatrix}.$$

Here, we chose $\mu_1 = 5$, $\mu_2 = 6$, $\phi_1 = 0.7$, $\phi_2 = 0.8$, $\theta_1 = 0.3$, $\theta_2 = 0.2$, $\sigma_1 = \sigma_2 = 1$ to simulate $n = 1000$ data points from a short-range dependent time series with a high cross-correlation.

A Simulated Toy Example



	ρ	Drift	ρ_P	ρ_S	OPD
First part	0.9		0.786	0.837	
Second part	0.9	mild exponential	0.201	0.303	
Third part	0.6		0.592	0.601	

Our Approach: Ordinal Pattern Dependence

Roughly speaking, **ordinal pattern dependence** measures how often we encounter the same patterns at the same time considering our two data sets; the theoretical counterpart is the probability

$$\mathbb{P}(\Pi(X_1, ..., X_d) = \Pi(Y_1, ..., Y_d)).$$

for the underlying models, that is, the time series we consider.

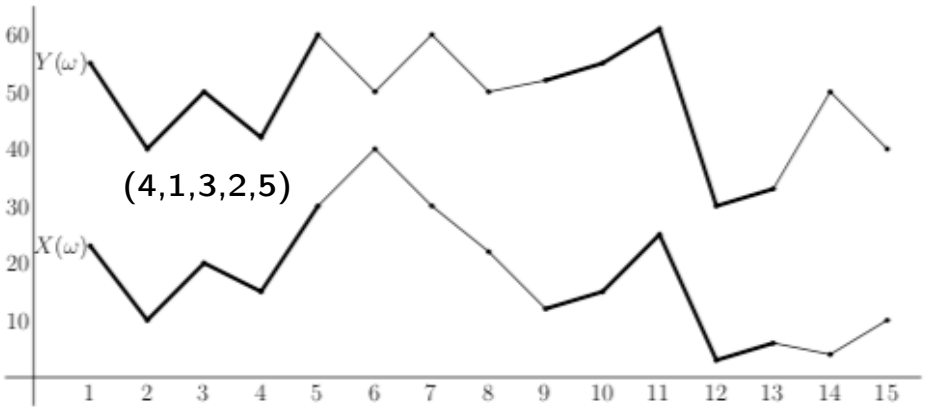


Figure: Two data sets with partially co-monotonic behavior.

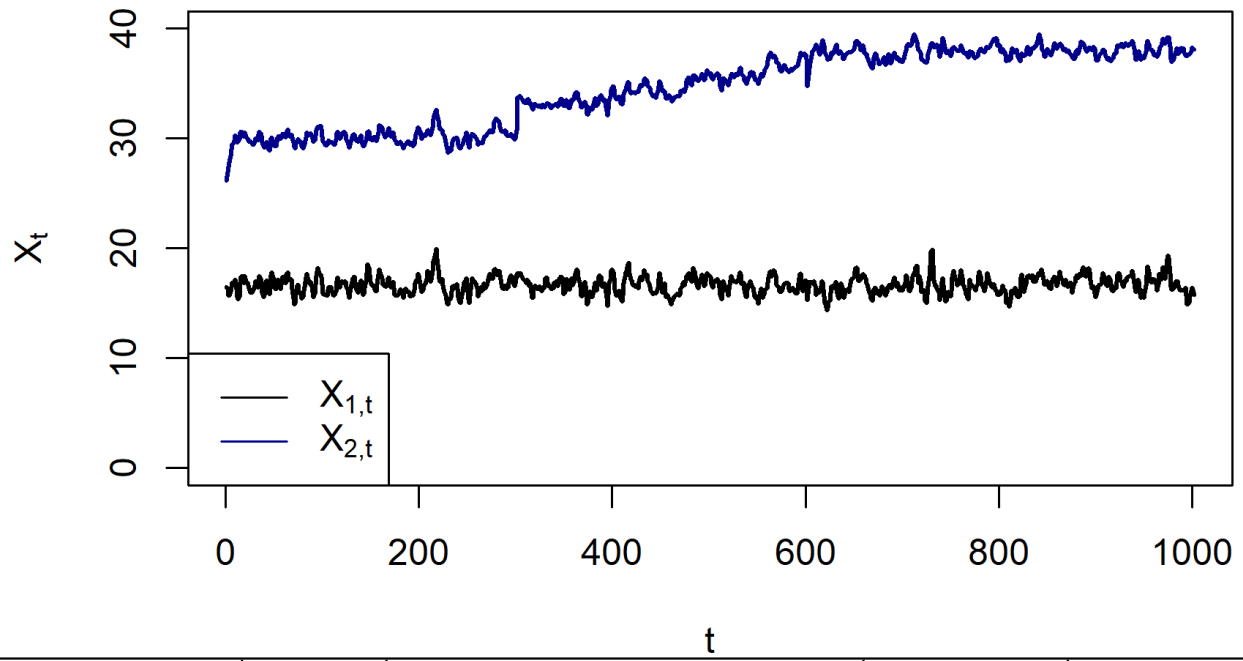
Advantages

In fact, one subtracts the probability for the hypothetical case of independence and divides by some normalizing factor in order to obtain values between zero and one in the case of positive dependence.

Advantages:

- We do not need to assume the existence of second moments
- Time series do not necessarily have to be stationary
- The method is robust against small noise and/or measurement errors
- The intuitive concept makes it easy to apply in practice
- The algorithms are quick.

Back to Our Simulated Example



	ρ	Drift	ρ_P	ρ_S	OPD
First part	0.9		0.786	0.837	0.366
Second part	0.9	mild exponential	0.201	0.303	0.300
Third part	0.6		0.592	0.601	0.105

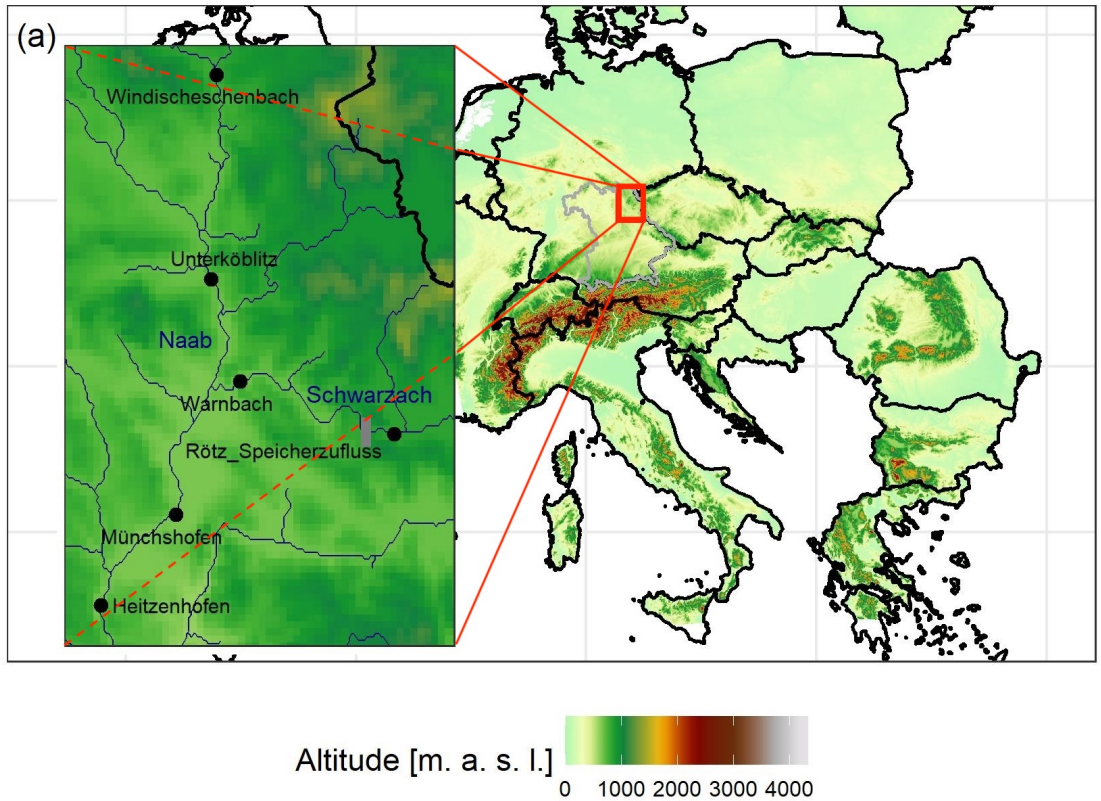
Hydrological Data

We consider six runoff data sets of catchments of the mesoscale in Bavaria, Southern Germany.

In November 1975, a **large reservoir** was build at the Schwarzach river, called the Eixendorfer See, located directly downstream of the Rötz gauge. It serves flood protection, low water elevation (water regulation), hydropower generation and recreation and with a storage of 19.30 million m^3 it is among the largest reservoirs in Southern Germany.

We consider the monthly maximum discharges of the period from 1959–2018, which often serve as basis for flood frequency analyses. Due to their temporal resolution, monthly maximum discharges are usually considered to be short-range dependent (cf. Wu 2009).

A Map


(a)
Windischeschenbach
Unterköblitz
Naab
Warnbach
Schwarzach
Rötz_Speicherzufluss
Münchshofen
Heitzenhofen
Altitude [m. a. s. l.]
0 1000 2000 3000 4000

Step I: Preparation and First Analysis

The available discharge data sets are checked for consistency and trimmed to a joint time period between 1959 and 2018 ($n = 730$ data points each).

For each data series, the Wilcoxon test for short-range dependent data is applied to test whether there is a significant change-point in the mean. For the gauges Münchshofen and Heitzenhofen in the Naab river as well as Warnbach in the Schwarzach river, such a significant change-point (5% significant level) was detected in the period of 1975 till 1977.

Therefore, all gauges located downstream of the reservoir can be considered as non-stationary. Hence, the classical tests for structural breaks in the dependence structure cannot be applied.

Step II: Ordinal Pattern Dependence of the Whole Series

Ordinal pattern dependence of all six gauges for the time period 1959-2018. Since the OPD is symmetric, only the upper half of the table is shown.

Gauge	**Win.**	**Unt.**	**Rötz**	**War.**	**Mün.**	**Heit.**
Windischeschenbach	1.000	0.326	0.151	0.170	0.250	0.221
Unterköblitz		1.000	0.179	0.249	0.507	0.462
Rötz			1.000	0.278	0.209	0.205
Warnbach				1.000	0.317	0.343
Münchshofen					1.000	0.670
Heitzenhofen						1.000

The results reveal that there is a high dependence between the gauges of the main river, decreasing with distance.

Step III: Dependence Changes Over Time

Ordinal pattern dependence of all six gauges for the time period 1959-1975 (upper table) and 1976-2018 (lower table).

Gauge	**Win.**	**Unt.**	**Rötz**	**War.**	**Mün.**	**Heit.**
Windischeschenbach	1.000	0.222	0.095	0.115	0.179	0.179
Unterköblitz		1.000	0.164	0.240	0.525	0.486
Rötz			1.000	0.390	0.181	0.187
Warnbach				1.000	0.289	0.317
Münchshofen					1.000	0.746
Heitzenhofen						1.000

Gauge	**Win.**	**Unt.**	**Rötz**	**War.**	**Mün.**	**Heit.**
Windischeschenbach	1.000	0.372	0.175	0.191	0.273	0.241
Unterköblitz		1.000	0.191	0.259	0.517	0.465
Rötz			1.000	0.246	0.224	0.205
Warnbach				1.000	0.337	0.356
Münchshofen					1.000	0.655
Heitzenhofen						1.000

Some Results

- Our method shows how the new reservoir has **changed the dependence** between the discharge data sets.

Some Results

- Our method shows how the new reservoir has **changed the dependence** between the discharge data sets.
- In a similar way it could be used on the other frameworks described in the introduction.

Some Results

- Our method shows how the new reservoir has **changed the dependence** between the discharge data sets.
- In a similar way it could be used on the other frameworks described in the introduction.
- The method is robust, easy to implement and has the advantage that the concept behind it has a simple intuitive meaning.

Outlook

Further directions:

Outlook

Further directions:

- Proof proper limit results

Outlook

Further directions:

- Proof proper limit results
- Rely on ordinal pattern stationarity

Outlook

Further directions:

- Proof proper limit results
- Rely on ordinal pattern stationarity
- Consider patterns with ties

Thank you for your attention!