# Using Spatial Ordinal Patterns for Non-parametric Testing of Spatial Dependence


HELMUT SCHMIDT UNIVERSITÄT
Universität der Bundeswehr Hamburg

MATH STAT

**<u>Christian H. Weiß</u>**

Department of Mathematics & Statistics

Helmut Schmidt University, Hamburg

**Hee-Young Kim**

Department of Bigdata Science

Korea University, Sejong, South Korea

HELMUT SCHMIDT UNIVERSITÄT
Universität der Bundeswehr Hamburg

MATH STAT

# Ordinal Patterns in Time Series

## Introduction

Bandt & Pompe (2002) introduced **ordinal patterns** (OPs) as complexity measures for time series characterized by "*simplicity, extremely fast calculation, robustness, and invariance with respect to nonlinear monotonous transformations*".

**Basic idea** in time-series case: map segments $\boldsymbol{X}_t = (X_t, X_{t+1}, \ldots, X_{t+m-1})$ of length $m$ from continuously distrib., real-valued process $(X_t)_{t\in\mathbb{Z}=\{\ldots,-1,0,1,\ldots\}}$ onto permutations from symmetric group $S_m$ of order $m$, where selected $\pi_t \in S_m = \{\pi^{[1]}, \ldots, \pi^{[m!]}\}$ expresses order among values in $\boldsymbol{X}_t$ in certain way (see details below).

In this way, original process $(X_t)$ (or time series $(x_t)$ thereof) transformed into (symbolic) sequence $(\pi_t)$ of OPs that reveals ordinal structure of $(X_t)$, see Keller et al. (2007).

Marginal distribution of OP series $(\pi_t)$ provides insights into dynamic structure of original process $(X_t)$.

Expressed as $m!$-dimensional probability vector $\boldsymbol{p}$ (or frequency vector $\widehat{\boldsymbol{p}}$ in case of time series data $(x_t)$), with $k$th component being $p_k = P(\pi_t = \pi^{[k]})$.

Different (equivalent) approaches to represent OP by permutation from $S_m$, see Berger et al. (2019).

We focus on **rank representation**, most intuitive approach.

Then, entries of $\pi = (r_1, \ldots, r_m) \in S_m$ interpreted as ranks within $\boldsymbol{x} = (x_1, \ldots, x_m) \in \mathbb{R}^m$, i. e.,

$$r_k < r_l \qquad \Leftrightarrow \qquad x_k < x_l \quad \text{or} \quad (x_k = x_l \text{ and } k < l)$$

for all $k, l \in \{1, \ldots, m\}$. Here, "$x_k = x_l$" if ties within $\boldsymbol{x}$.

**Example:** $(1.2, -0.7, 3.4, 1.9) \quad \mapsto \quad (2, 1, 4, 3)$,

$(1.2, -0.7, 3.4, -0.7) \quad \mapsto \quad (3, 1, 4, 2)$.

Order $m$ of OPs chosen by user.

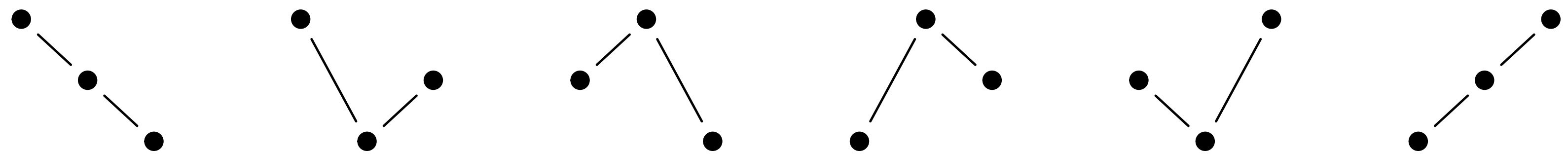
If $m = 2$, then downward OP $(2,1)$ and upward OP $(1,2)$, preserves only little information from original process.

However, range of $\pi_t$ quickly increases with $m$ as $|S_m| = m!$, so estimation of $\boldsymbol{p}$ quickly difficult in practice.

Therefore, in time series analysis,

**convenient choice** is $m = 3$ (Bandt, 2019):

$(3,2,1)$ $(3,1,2)$ $(2,3,1)$ $(1,3,2)$ $(2,1,3)$ $(1,2,3)$



$\Rightarrow$ sufficiently informative and computationally feasible.

Let $(X_t)$ be *continuously distributed* real-valued process, independent and identically distributed (i. i. d.) *under null*. Probability of ties $= 0$, so ties at most rarely in data.

Following **properties** crucial for dependence tests:

1. OPs invariant w.r.t. strictly monotonically increasing transformations of $X_t$. Thus, OPs do not depend on actual marginal distribution of $(X_t)_{\mathbb{N}}$ (**distribution-free** approach).
2. $(X_t)_{\mathbb{N}}$ is i. i. d. under null ($\rightarrow$ exchangeability). Thus, $\pi_t$ discrete uniform on $S_m$, i. e., $P\big(\pi_t = \pi\big) = 1/m!$ for each $\pi \in S_m$ (**no parameter estimation** required).

**OP-test statistics** built upon $\widehat{\boldsymbol{p}}$ computed from $\pi_1, \ldots, \pi_n$, where $\pi_t$ from $\boldsymbol{X}_t = (X_t, X_{t+1}, \ldots, X_{t+m-1})$ for $t = 1, 2, \ldots, n$.

Moving-window approach, affects asymptotics of $\sqrt{n}\,\big(\widehat{\boldsymbol{p}} - \boldsymbol{p}_0\big)$ under i. i. d.-null required, where $\boldsymbol{p}_0 = (1/m!, \ldots, 1/m!)$:

**Theorem:** (Elsinger, 2010; Weiß, 2022)

$\sqrt{n}\,\big(\widehat{\boldsymbol{p}} - \boldsymbol{p}_0\big) \;\to\; \mathrm{N}(\boldsymbol{0}, \boldsymbol{\Sigma}_m)$ with $\boldsymbol{\Sigma}_m = (\sigma_{ij})_{i,j=1,\ldots,m!}$ given by

$$\sigma_{ij} \;=\; 1/m!\,\big(\delta_{ij} - 1/m!\big) \;+\; \sum_{h=1}^{m-1} \Big(p_{ij}(h) + p_{ji}(h) - 2/m!^2\Big).$$

Afterwards, distribution of OP-test statistics by Taylor approximations ("Delta method").

**Case $m = 3$:**

$$\mathbf{P}(1) = \frac{1}{24}\begin{pmatrix} 1 & 1 & 0 & 0 & 2 & 0 \\ 0 & 0 & 1 & 1 & 0 & 2 \\ 1 & 2 & 0 & 0 & 1 & 0 \\ 2 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 2 & 0 & 1 \\ 0 & 0 & 2 & 1 & 0 & 1 \end{pmatrix}, \qquad \mathbf{P}(2) = \frac{1}{120}\begin{pmatrix} 1 & 1 & 3 & 6 & 3 & 6 \\ 3 & 3 & 4 & 3 & 4 & 3 \\ 1 & 1 & 3 & 6 & 3 & 6 \\ 3 & 3 & 4 & 3 & 4 & 3 \\ 6 & 6 & 3 & 1 & 3 & 1 \\ 6 & 6 & 3 & 1 & 3 & 1 \end{pmatrix},$$

so

$$\Sigma_3 = \frac{1}{360}\begin{pmatrix} 46 & -23 & -23 & 7 & 7 & -14 \\ -23 & 28 & 10 & -2 & -20 & 7 \\ -23 & 10 & 28 & -20 & -2 & 7 \\ 7 & -2 & -20 & 28 & 10 & -23 \\ 7 & -20 & -2 & 10 & 28 & -23 \\ -14 & 7 & 7 & -23 & -23 & 46 \end{pmatrix}.$$

HELMUT SCHMIDT UNIVERSITÄT
Universität der Bundeswehr Hamburg

MATH STAT

# Testing for Spatial Dependence in Random Fields

Approaches & Asymptotics

Real-valued and continuously distributed **spatial data** occurring in regular two-dimensional grid: $(X_t)_{t\in\mathbb{Z}^2}$.
(random field, spatial process in plane, regular lattice structure)

Data rectangles $(x_t) = (x_{t_1,t_2})$ with $0 \le t_1 \le m$ and $0 \le t_2 \le n$.

Infer dependence in $(X_t)$ via **spatial OPs** (SOPs), due to Ribeiro et al. (2012) and Bandt & Wittfeld (2023).

$m_1 \times m_2$-SOP computed from $m_1 \times m_2$-rectangle from $(x_t)$:

1. concatenate rows into vector of length $m_1 \cdot m_2$,
2. compute corresponding $(m_1 \cdot m_2)$th-order OP from $S_{m_1 \cdot m_2}$,
3. transform back into $m_1 \times m_2$-matrix in row-wise manner.

As $|S_{m_1 \cdot m_2}| = (m_1 \cdot m_2)!$ quickly unfeasibly large,

Bandt & Wittfeld (2023) recommend focus on $2 \times 2$-SOPs:

$$\mathbf{X}_t = \begin{pmatrix} X_{t_1-1,t_2-1} & X_{t_1-1,t_2} \\ X_{t_1,t_2-1} & X_{t_1,t_2} \end{pmatrix} =: \begin{pmatrix} x_1 & x_2 \\ x_3 & x_4 \end{pmatrix} \quad \mapsto \quad \begin{pmatrix} r_1 & r_2 \\ r_3 & r_4 \end{pmatrix},$$

where $(r_1, r_2, r_3, r_4) \in S_4$ is OP of $(x_1, x_2, x_3, x_4)$.

Bandt & Wittfeld (2023): further partition into **types**,

$$\mathcal{S}_1 = \left\{ \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}, \begin{pmatrix} 2 & 1 \\ 4 & 3 \end{pmatrix}, \begin{pmatrix} 2 & 4 \\ 1 & 3 \end{pmatrix}, \begin{pmatrix} 3 & 1 \\ 4 & 2 \end{pmatrix}, \begin{pmatrix} 3 & 4 \\ 1 & 2 \end{pmatrix}, \begin{pmatrix} 4 & 2 \\ 3 & 1 \end{pmatrix}, \begin{pmatrix} 4 & 3 \\ 2 & 1 \end{pmatrix} \right\},$$

$$\mathcal{S}_2 = \left\{ \begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix}, \begin{pmatrix} 2 & 1 \\ 3 & 4 \end{pmatrix}, \begin{pmatrix} 2 & 3 \\ 1 & 4 \end{pmatrix}, \begin{pmatrix} 3 & 2 \\ 4 & 1 \end{pmatrix}, \begin{pmatrix} 3 & 4 \\ 2 & 1 \end{pmatrix}, \begin{pmatrix} 4 & 1 \\ 3 & 2 \end{pmatrix}, \begin{pmatrix} 4 & 3 \\ 1 & 2 \end{pmatrix} \right\},$$

$$\mathcal{S}_3 = \left\{ \begin{pmatrix} 1 & 3 \\ 4 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 4 \\ 3 & 2 \end{pmatrix}, \begin{pmatrix} 2 & 3 \\ 4 & 1 \end{pmatrix}, \begin{pmatrix} 2 & 4 \\ 3 & 1 \end{pmatrix}, \begin{pmatrix} 3 & 1 \\ 2 & 4 \end{pmatrix}, \begin{pmatrix} 3 & 2 \\ 1 & 4 \end{pmatrix}, \begin{pmatrix} 4 & 1 \\ 2 & 3 \end{pmatrix}, \begin{pmatrix} 4 & 2 \\ 1 & 3 \end{pmatrix} \right\}.$$

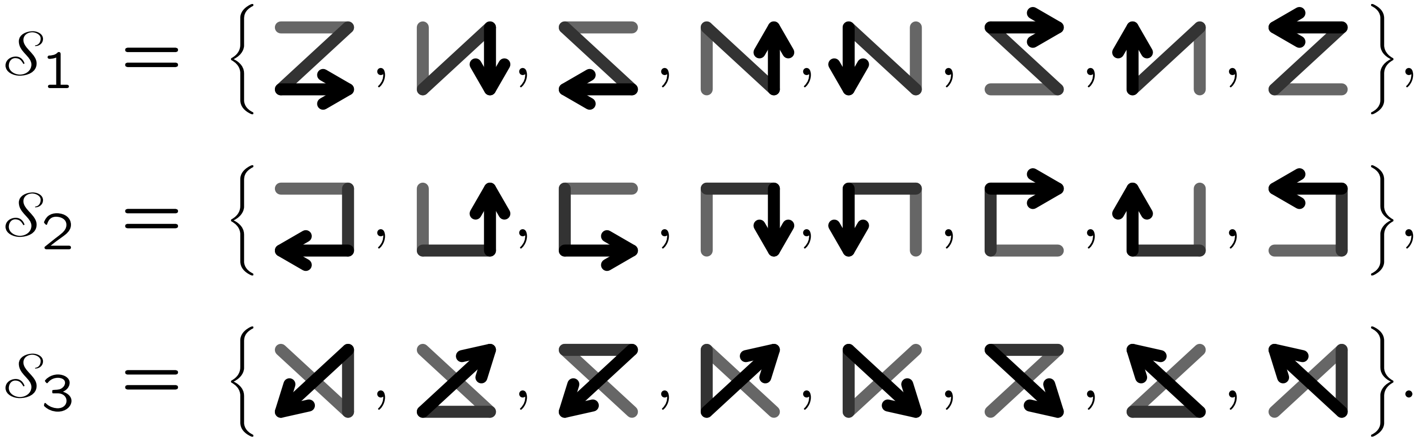
## **Visual representation** of types

$$\mathcal{S}_1 = \left\{ \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}, \begin{pmatrix} 2 & 1 \\ 4 & 3 \end{pmatrix}, \begin{pmatrix} 2 & 4 \\ 1 & 3 \end{pmatrix}, \begin{pmatrix} 3 & 1 \\ 4 & 2 \end{pmatrix}, \begin{pmatrix} 3 & 4 \\ 1 & 2 \end{pmatrix}, \begin{pmatrix} 4 & 2 \\ 3 & 1 \end{pmatrix}, \begin{pmatrix} 4 & 3 \\ 2 & 1 \end{pmatrix} \right\},$$

$$\mathcal{S}_2 = \left\{ \begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix}, \begin{pmatrix} 2 & 1 \\ 3 & 4 \end{pmatrix}, \begin{pmatrix} 2 & 3 \\ 1 & 4 \end{pmatrix}, \begin{pmatrix} 3 & 2 \\ 4 & 1 \end{pmatrix}, \begin{pmatrix} 3 & 4 \\ 2 & 1 \end{pmatrix}, \begin{pmatrix} 4 & 1 \\ 3 & 2 \end{pmatrix}, \begin{pmatrix} 4 & 3 \\ 1 & 2 \end{pmatrix} \right\},$$

$$\mathcal{S}_3 = \left\{ \begin{pmatrix} 1 & 3 \\ 4 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 4 \\ 3 & 2 \end{pmatrix}, \begin{pmatrix} 2 & 3 \\ 4 & 1 \end{pmatrix}, \begin{pmatrix} 2 & 4 \\ 3 & 1 \end{pmatrix}, \begin{pmatrix} 3 & 1 \\ 2 & 4 \end{pmatrix}, \begin{pmatrix} 3 & 2 \\ 1 & 4 \end{pmatrix}, \begin{pmatrix} 4 & 1 \\ 2 & 3 \end{pmatrix}, \begin{pmatrix} 4 & 2 \\ 1 & 3 \end{pmatrix} \right\}$$

by arrows along increasing rank:



3 types more feasible in small data than 24 SOPs.

Type 1: monotonic behaviour both along rows and columns.

Type 2: uniquely increase/decrease only along one direction.

Type 3: lowest and highest ranks on either main or antidiagonal.

type = rank number which shares diagonal with rank 4.

**Example:** arthropods data from R-package `agridat`:

$$\begin{array}{cccl} 35 & 24 & 18 & \cdots \; t_2 \\ 18 & 32 & 14 & \cdots \\ \boxed{17} & \boxed{21} & 40 & \cdots \\ \boxed{17} & \boxed{34} & 25 & \cdots \\ \vdots & \vdots & \vdots & \ddots \\ t_1 & & & \end{array} \quad \to \quad \begin{array}{ccl} \begin{pmatrix} 4 & 2 \\ 1 & 3 \end{pmatrix} & \begin{pmatrix} 3 & 2 \\ 4 & 1 \end{pmatrix} & \cdots \; t_2 \\ \begin{pmatrix} 2 & 4 \\ 1 & 3 \end{pmatrix} & \begin{pmatrix} 3 & 1 \\ 2 & 4 \end{pmatrix} & \cdots \\ \boxed{\begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}} & \begin{pmatrix} 1 & 4 \\ 3 & 2 \end{pmatrix} & \cdots \\ \vdots & \vdots & \ddots \\ t_1 & & \end{array} \quad \to \quad \begin{array}{ccl} 3 & 2 & \cdots \; t_2 \\ 1 & 3 & \cdots \\ \boxed{1} & 3 & \cdots \\ \vdots & \vdots & \ddots \\ t_1 & & \end{array}$$

Asymptotics of SOP frequencies under null "$(X_t)$ are i. i. d.":

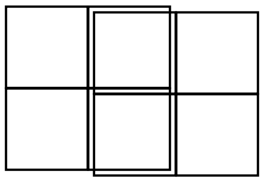
**Theorem:** (Weiß & Kim, 2024)

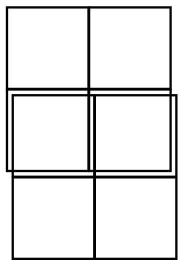
$\sqrt{mn}\,\big(\hat{\boldsymbol{p}} - \boldsymbol{p}_0\big) \ \overset{\mathrm{d}}{\to}\ \mathrm{N}\big(\mathbf{0}, \boldsymbol{\Sigma}\big)$, where $\boldsymbol{\Sigma}$ equals

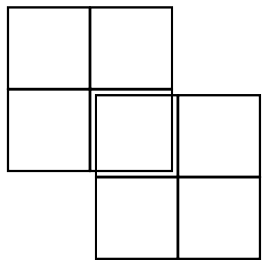
$$\mathrm{diag}(\boldsymbol{p}_0) - \boldsymbol{p}_0\,\boldsymbol{p}_0^\top \ + \ (1-\tfrac{1}{m})(1-\tfrac{1}{n})\Big(\mathbf{D}+\mathbf{D}^\top+\mathbf{A}+\mathbf{A}^\top-4\,\boldsymbol{p}_0\,\boldsymbol{p}_0^\top\Big)$$

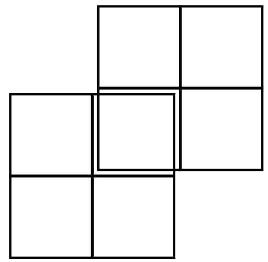
$$+\ (1-\tfrac{1}{n})\Big(\mathbf{H}+\mathbf{H}^\top-2\,\boldsymbol{p}_0\,\boldsymbol{p}_0^\top\Big)\ +\ (1-\tfrac{1}{m})\Big(\mathbf{V}+\mathbf{V}^\top-2\,\boldsymbol{p}_0\,\boldsymbol{p}_0^\top\Big).$$

Considers different overlaps of $2\times 2$-SOPs:

**H**: 

**V**: 

**D**: 

**A**: 

Closed-form expressions for $\mathbf{H}, \mathbf{V}, \mathbf{D}, \mathbf{A}$ in Weiß & Kim (2024),

e. g., entries of matrix $720 \cdot \mathbf{H}$:

| | $\pi_1$ | $\pi_2$ | $\pi_3$ | $\pi_4$ | $\pi_5$ | $\pi_6$ | $\pi_7$ | $\pi_8$ | $\pi_9$ | $\pi_{10}$ | $\pi_{11}$ | $\pi_{12}$ | $\pi_{13}$ | $\pi_{14}$ | $\pi_{15}$ | $\pi_{16}$ | $\pi_{17}$ | $\pi_{18}$ | $\pi_{19}$ | $\pi_{20}$ | $\pi_{21}$ | $\pi_{22}$ | $\pi_{23}$ | $\pi_{24}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\pi_1$ | 2 | 3 | 1 | 3 | 1 | 2 | 2 | 4 | 0 | 4 | 0 | 2 | 0 | 3 | 0 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| $\pi_2$ | 2 | 1 | 3 | 1 | 3 | 2 | 4 | 2 | 0 | 2 | 0 | 4 | 0 | 3 | 0 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| $\pi_3$ | 1 | 1 | 1 | 1 | 1 | 1 | 3 | 3 | 0 | 3 | 0 | 3 | 0 | 6 | 0 | 6 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| $\pi_4$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 0 | 3 | 0 | 4 | 0 | 2 | 0 | 2 | 4 | 2 | 3 | 1 | 3 | 1 | 2 |
| $\pi_5$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 3 | 0 | 1 | 0 | 1 | 3 | 3 | 6 | 1 | 6 | 1 | 3 |
| $\pi_6$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 2 | 0 | 2 | 0 | 2 | 2 | 4 | 3 | 3 | 3 | 3 | 4 |
| $\pi_7$ | 3 | 6 | 1 | 6 | 1 | 3 | 1 | 3 | 0 | 3 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| $\pi_8$ | 4 | 3 | 3 | 3 | 3 | 4 | 2 | 2 | 0 | 2 | 0 | 2 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| $\pi_9$ | 1 | 1 | 1 | 1 | 1 | 1 | 3 | 3 | 0 | 3 | 0 | 3 | 0 | 6 | 0 | 6 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| $\pi_{10}$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 0 | 3 | 0 | 2 | 0 | 4 | 0 | 4 | 2 | 2 | 1 | 3 | 1 | 3 | 2 |
| $\pi_{11}$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 3 | 0 | 1 | 0 | 1 | 3 | 3 | 6 | 1 | 6 | 1 | 3 |
| $\pi_{12}$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 3 | 0 | 3 | 1 | 3 | 1 | 6 | 1 | 6 | 3 |
| $\pi_{13}$ | 3 | 6 | 1 | 6 | 1 | 3 | 1 | 3 | 0 | 3 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| $\pi_{14}$ | 3 | 1 | 6 | 1 | 6 | 3 | 3 | 1 | 0 | 1 | 0 | 3 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| $\pi_{15}$ | 2 | 3 | 1 | 3 | 1 | 2 | 2 | 4 | 0 | 4 | 0 | 2 | 0 | 3 | 0 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| $\pi_{16}$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 6 | 0 | 6 | 0 | 3 | 0 | 3 | 0 | 3 | 3 | 1 | 1 | 1 | 1 | 1 | 1 |
| $\pi_{17}$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 2 | 0 | 2 | 0 | 2 | 2 | 4 | 3 | 3 | 3 | 3 | 4 |
| $\pi_{18}$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 3 | 0 | 3 | 1 | 3 | 1 | 6 | 1 | 6 | 3 |
| $\pi_{19}$ | 4 | 3 | 3 | 3 | 3 | 4 | 2 | 2 | 0 | 2 | 0 | 2 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| $\pi_{20}$ | 3 | 1 | 6 | 1 | 6 | 3 | 3 | 1 | 0 | 1 | 0 | 3 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| $\pi_{21}$ | 2 | 1 | 3 | 1 | 3 | 2 | 4 | 2 | 0 | 2 | 0 | 4 | 0 | 3 | 0 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| $\pi_{22}$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 6 | 0 | 6 | 0 | 3 | 0 | 3 | 0 | 3 | 3 | 1 | 1 | 1 | 1 | 1 | 1 |
| $\pi_{23}$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 0 | 3 | 0 | 4 | 0 | 2 | 0 | 2 | 4 | 2 | 3 | 1 | 3 | 1 | 2 |
| $\pi_{24}$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 0 | 3 | 0 | 2 | 0 | 4 | 0 | 4 | 2 | 2 | 1 | 3 | 1 | 3 | 2 |

Above Theorem simplifies if focus on **types:**

$\mathbf{D} - \boldsymbol{p}_0\boldsymbol{p}_0^\top = \mathbf{A} - \boldsymbol{p}_0\boldsymbol{p}_0^\top = \mathbf{O}$, and

$$\mathbf{H} = \mathbf{V} \;=\; \frac{1}{180}\begin{pmatrix} 21 & 20 & 19 \\ 20 & 21 & 19 \\ 19 & 19 & 22 \end{pmatrix}.$$

Hence,

$$\boldsymbol{\Sigma} \;=\; \frac{1}{9}\begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix} \;+\; \big(1 - \tfrac{1}{2m} - \tfrac{1}{2n}\big)\cdot\frac{1}{45}\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ -1 & -1 & 2 \end{pmatrix}$$

$$\to\; \frac{1}{45}\begin{pmatrix} 11 & -5 & -6 \\ -5 & 11 & -6 \\ -6 & -6 & 12 \end{pmatrix} \qquad \text{for } m, n \to \infty.$$

Bandt & Wittfeld (2023) proposed following type statistics:

$$\widehat{\tau} = \widehat{p}_1 - 1/3 \quad \text{and} \quad \widehat{\kappa} = \widehat{p}_2 - \widehat{p}_3,$$
$$\widetilde{\tau} = \widehat{p}_3 - 1/3 \quad \text{and} \quad \widetilde{\kappa} = \widehat{p}_1 - \widehat{p}_2.$$

Asymptotics under null "$(X_t)$ are i. i. d.": (Weiß & Kim, 2024)

**Corollary:** $\sqrt{mn}\,\big(\widehat{\tau}, \widehat{\kappa}\big) \overset{\text{d}}{\to} \mathrm{N}\big(\mathbf{0}, \mathbf{\Sigma}'\big)$, where $\mathbf{\Sigma}'$ equals

$$\mathbf{\Sigma}' = \frac{2}{9}\begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix} + \big(1 - \tfrac{1}{2m} - \tfrac{1}{2n}\big) \cdot \frac{1}{45}\begin{pmatrix} 1 & 1 \\ 1 & 5 \end{pmatrix} \approx \frac{1}{45}\begin{pmatrix} 11 & 1 \\ 1 & 35 \end{pmatrix}$$

$\sqrt{mn}\,\big(\widetilde{\tau}, \widetilde{\kappa}\big) \overset{\text{d}}{\to} \mathrm{N}\big(\mathbf{0}, \mathbf{\Sigma}''\big)$, where $\mathbf{\Sigma}''$ equals

$$\mathbf{\Sigma}'' = \frac{2}{9}\begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix} + \big(1 - \tfrac{1}{2m} - \tfrac{1}{2n}\big) \cdot \frac{2}{45}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \approx \frac{4}{45}\begin{pmatrix} 3 & 0 \\ 0 & 8 \end{pmatrix}.$$

HELMUT SCHMIDT
UNIVERSITÄT
Universität der Bundeswehr Hamburg

MATH
STAT

# Testing for Spatial Dependence in Random Fields

Empirical Investigations

**Performance analysis** by simulations in Weiß & Kim (2024).

**Brief summary:**

While spatial ACF superior for *linear unilateral* DGPs, e. g.,

$$X_{t_1,t_2} = \alpha_1 \cdot X_{t_1-1,t_2} + \alpha_2 \cdot X_{t_1,t_2-1} + \alpha_3 \cdot X_{t_1-1,t_2-1} + \epsilon_{t_1,t_2},$$

SOP-based tests often superior in presence of *outliers*,

for *non-linear* DGPs, such as

$$X_{t_1,t_2} = \beta_1 \cdot \epsilon^2_{t_1-1,t_2} + \beta_2 \cdot \epsilon^2_{t_1,t_2-1} + \beta_3 \cdot \epsilon^2_{t_1-1,t_2-1} + \epsilon_{t_1,t_2},$$

(. . . )

(. . . ) and also for *bilateral* spatial DGPs, among others

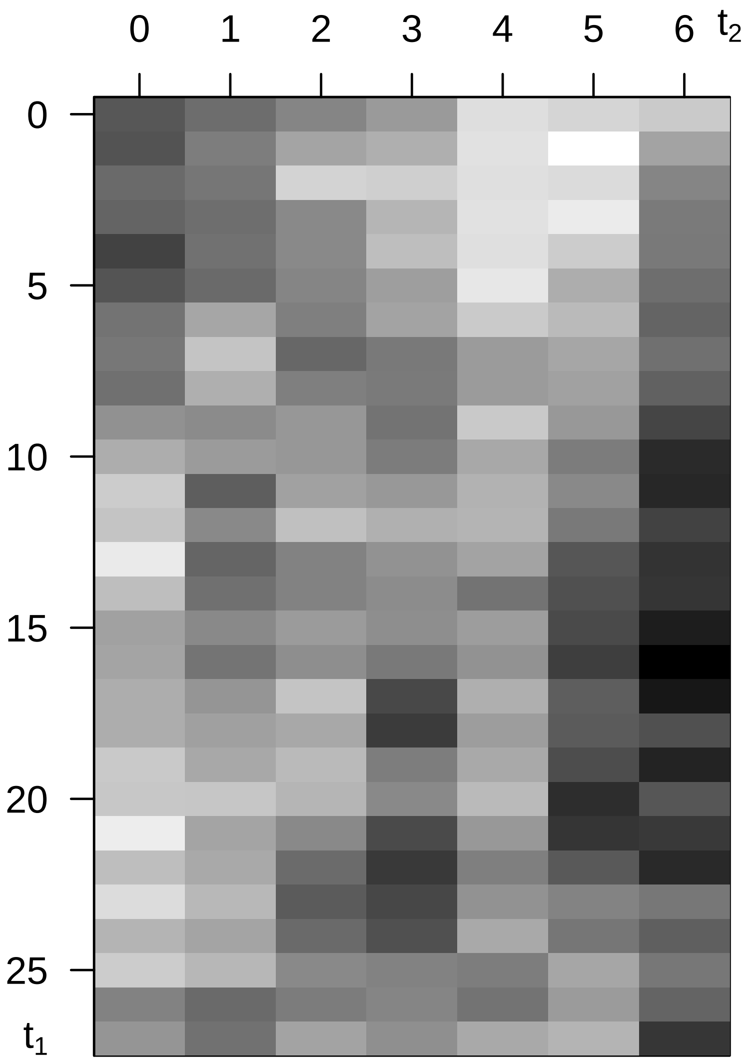
$$X_{t_1,t_2} = a_1 \cdot X_{t_1-1,t_2} + a_2 \cdot X_{t_1,t_2-1} + a_3 \cdot X_{t_1,t_2+1} + a_4 \cdot X_{t_1+1,t_2} + \epsilon_{t_1,t_2}$$

and

$$X_{t_1,t_2} = b_1 \cdot \epsilon^2_{t_1-1,t_2-1} + b_2 \cdot \epsilon_{t_1+1,t_2-1} + b_3 \cdot \epsilon^2_{t_1+1,t_2+1} + b_4 \cdot \epsilon_{t_1-1,t_2+1} + \epsilon_{t_1,t_2}.$$

$\tilde{\tau}$-test performs particularly well!

Weiß & Kim (2024): two **agricultural data examples**, e. g., **yield of barley** (in kg) in $28 \times 7$-grid ($m = 27, n = 6$) from uniformity trial experiment (Kempton & Howes, 1981). Yield as deviations in 0.01 kg-units from mean yield 2.63 kg.




Except $\tilde{\kappa}$, all SOP-tests significant spatial dependence.

Type 3 too rare (7.4%), types 1 and 2 too frequent (46.9% resp. 45.7%).

Those SOPs too frequent, where maximal ranks within columns.

**Explanation:**
"sowing, harvesting and all intermediate farming practices were carried out column by column, and this could produce intra-column correlations" (Kempton & Howes, 1981)

# Conclusions & Outlook

OPs are well-interpretable, robust, and flexibly adapted to different types of dependence.

If data continuously distributed, we get non-parametric tests.

**Work in progress & future research:**

- SOP-based hypothesis tests, which use a refined definition of types;
- control charts based on SOPs and types, in analogy to Weiß & Testik (2023);
- SOPs based on "generalized OPs", where ties are explicitly accounted for, in analogy to Weiß & Schnurr (2023).

# Thank You

# for Your Interest!

**Christian H. Weiß**

Department of Mathematics & Statistics

Helmut Schmidt University, Hamburg

weissc@hsu-hh.de

# Literature



Bandt (2019) Small order patterns . . . *Entropy* **21**, 613.

Bandt & Pompe (2002) Permutation entropy . . . *Phys Rev L* **88**, 174102.

Bandt & Wittfeld (2023) Two new parameters . . . *Chaos* **33**, 043124.

Berger et al. (2019) Teaching ordinal patterns . . . *Entropy* **21**, 1023.

Elsinger (2010) Independence tests . . . *Working paper* **165**, Öst. Nat.bank.

Keller et al. (2007) Time series from . . . *Stoch Dyn* **7**, 247–272.

Kempton & Howes (1981) The use of neighbouring . . . *JRSS-C* **30**, 59–70.

Ribeiro et al. (2012) Complexity-entropy . . . *PLoS ONE* **7**, e40689.

Weiß (2022) Non-parametric tests . . . *Chaos* **32**, 093107.

Weiß & Kim (2024) Using spatial ordinal patterns . . . *Spat Stat* **59**, 100800.

Weiß & Schnurr (2023) Generalized OPs . . . *J Nonpar Stat*, in press.

Weiß & Testik (2023) Non-param. control ch. . . *Technomet* **65**, 340–350.