Non-parametric Monitoring of Serial Dependence based on Ordinal Patterns

Christian H. Weiß

Department of Mathematics & Statistics,

Helmut Schmidt University, Hamburg

Murat Caner Testik

Department of Industrial Engineering,

Hacettepe University, Ankara

Dedicated to the memory of

Prof. Dr. Karsten Keller

(1961–2022)

HELMUT SCHMIDT UNIVERSITÄT
Universität der Bundeswehr Hamburg

MATH STAT

Monitoring of Serial Dependence

Introduction

Serial independence of observations is crucial assumption for majority of control charts (Chakraborti & Sparks, 2020).

Either monitored data themselves i. i. d. (Montgomery, 2009), or if time-series data, control charts applied to residuals, being uncorrelated if adeq. in-control model (Knoth & Schmid, 2004).

Neglecting ser. dep. severely affects performance (Alwan, 1992).

Hence, not only **monitor** for changes in mean or variance, but also **in-control assumption of serial independence**, to avoid misleading chart performance.

Several proposals to monitor for existence of serial dependencies in real-valued and continuously distributed processes.

- Large majority refers to ARMA models, some of these directly incorporate model structure into monitored statistics (types of likelihood approaches).
- Other studies easier to generalise as they monitor residuals for possible dependence, e. g.,
  - autocorrelation function (ACF) charts by Yourstone & Montgomery (1991), Atienza et al. (1997), and others;
  - periodogram chart by Beneke et al. (1988).

But aforementioned ACF and periodogram charts rely on moving-window approach, with typical window-length recommendation $w = 50$

$\Rightarrow$ **Drawback:** first collect w observations before start of monitoring, so detection delay of $\geq w$ observations.

Only individuals charts for monitoring dependence is EWMA-type monitoring of lag-1 ACF by Gardner (1983); thus later used as competitor.

Further drawback besides delay and model dependence: need model fit during Phase I, thus estimation uncertainty.

Our newly **proposed charts**

- are fully non-parametric and distribution-free,
- do not require parameter estimation but have unique design,
- used almost instantaneously at start of monitoring,
- for any real-valued and continuously distributed process (which is assumed serially independent in in-control state).

These appealing properties achieved by using so-called **ordinal patterns** for process monitoring.

Outline: various control chart proposals, analysis of out-of-control performance, application to chemical process data.

HELMUT SCHMIDT UNIVERSITÄT
Universität der Bundeswehr Hamburg

MATH
STAT

Ordinal Patterns in Time Series

Definition & Properties

Since papers Bandt & Pompe (2002), Keller et al. (2007), **ordinal patterns** (OPs) widely used for analyzing real-valued time series (Bandt, 2019).

Among others, used for hypothesis tests w.r.t. serial dependence (Weiß, 2022) $\rightarrow$ **starting point** of this research.

Idea: Map segments from (X_t) of length $m \in \mathbb{N}$ on corresponding OPs of order m (common choice: $m = 2$ or 3). Based on frequency of OPs, judge null of independence.

By contrast to ACF (default tool for dependence analyses), OP-statistics also sensitive to non-linear dependence.

Let S_m be symmetric group of order m.

$m!$ possible permutations $\pi \in S_m$ represent

$m!$ different OPs of $\boldsymbol{x} = (x_1, \ldots, x_m) \in \mathbb{R}^m$.

Rank representation $\pi = (r_1, \ldots, r_m) \in S_m$:

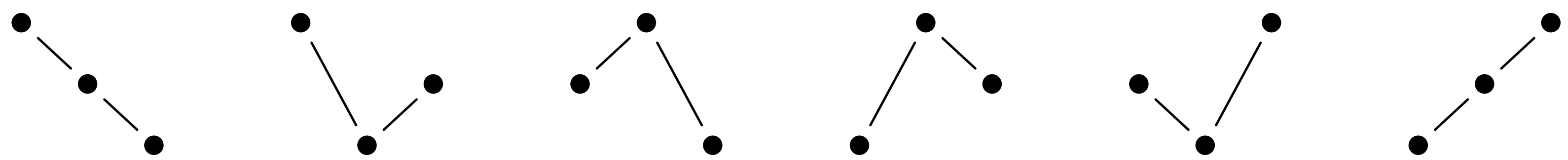
$$r_k < r_l \qquad \Leftrightarrow \qquad x_k < x_l \quad \text{or} \quad (x_k = x_l \text{ and } k < l)$$

for all $k, l \in \{1, \ldots, m\}$.

Example $m = 2$: downward OP $(2, 1)$ and upward OP $(1, 2)$.

Example $m = 3$:

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



Let tth m-history equal $\boldsymbol{X}_t = (X_t, X_{t+1}, \ldots, X_{t+m-1})$, denote its OP by π_t.

Original process $(X_t)_{\mathbb{N}}$ transformed into OP-process $(\pi_t)_{\mathbb{N}}$.

Following **properties** crucial for proposed control charts:

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. in its in-control state ($\to$ exchangeability). Thus, π_t discrete uniform on S_m, i. e., PMF $P\big(\pi_t = \pi\big) = 1/m!$ for each $\pi \in S_m$ (**no Phase-I analysis** required).

HELMUT SCHMIDT UNIVERSITÄT
Universität der Bundeswehr Hamburg

MATH STAT

EWMA Control Charts based on Ordinal Patterns

Definition & Design

Denote mth-order OPs $\pi^{[1]}, \ldots, \pi^{[m!]}$ in some arrangement ("lexicographic order"), abbreviate $p_k = P(\pi_t = \pi^{[k]})$, and define PMF vector $\boldsymbol{p} = (\ldots, p_k, \ldots)$.

If $(X_t)_{\mathbb{N}}$ is i. i. d. (in-control assumption), then all $p_k = 1/m!$; abbreviate $\boldsymbol{p}_0 = (1/m!, \ldots, 1/m!)$.

For monitoring, compute estimator $\widehat{\boldsymbol{p}}_t$ of $\boldsymbol{p}$ at each time t.

Our proposal:

Define $m!$-dim. $\boldsymbol{Z}_t$ by "one-hot encoding" $Z_{t,k} = \mathbb{1}(\pi_t = \pi^{[k]})$.

EWMA estimator with smoothing parameter $\lambda \in (0;1)$:

$$\widehat{\boldsymbol{p}}_0 \;=\; \boldsymbol{p}_0, \qquad \widehat{\boldsymbol{p}}_t \;=\; \lambda\, \boldsymbol{Z}_t + (1-\lambda)\, \widehat{\boldsymbol{p}}_{t-1} \quad \text{for } t = 1, 2, \ldots$$

Based on $(\widehat{\boldsymbol{p}}_t)$, we compute appropriate statistics,

which are plotted on EWMA OP chart.

Following Bandt & Pompe (2002), Bandt (2019), Weiß (2022),

we consider:

entropy: $\widehat{H} = -\sum_{k=1}^{m!} \widehat{p}_k \ln \widehat{p}_k$,

extropy: $\widehat{H}_{\mathrm{ex}} = -\sum_{k=1}^{m!} \big(1 - \widehat{p}_k\big) \ln\big(1 - \widehat{p}_k\big)$,

distance to white noise: $\widehat{\Delta} = \sum_{k=1}^{m!} \big(\widehat{p}_k - 1/m!\big)^2$ for $m \geq 2$;

up-down balance: $\widehat{\beta} = \widehat{p}_6 - \widehat{p}_1$,

persistence: $\widehat{\tau} = \widehat{p}_6 + \widehat{p}_1 - \frac{1}{3}$,

up-down scaling: $\widehat{\delta} = \widehat{p}_4 + \widehat{p}_5 - \widehat{p}_3 - \widehat{p}_2$ for $m = 3$.

This gives three **one-sided EWMA charts**,

H-chart: plot $\widehat{H}_t$ with LCL $l_H \in (0;\ \ln m!)$;

H_{ex}-chart: plot $\widehat{H}_{\text{ex},t}$ with LCL $l_{H_{\text{ex}}} \in \big(0;\ (m!-1)\ \ln(\frac{m!}{m!-1})\big)$;

Δ-chart: plot $\widehat{\Delta}_t$ with UCL $l_\Delta > 0$;

and three **two-sided EWMA charts** (symmetric limits):

β-chart: plot $\widehat{\beta}_t$ with UCL $l_\beta > 0$ and LCL $-l_\beta$;

τ-chart: plot $\widehat{\tau}_t$ with UCL $l_\tau > 0$ and LCL $-l_\tau$;

δ-chart: plot $\widehat{\delta}_t$ with UCL $l_\delta > 0$ and LCL $-l_\delta$.

ARLs approximated by simulations (we use 10^5 replications).

Highly attractive feature: in-control design

for each setting (m, λ) and ARL$_0$ determined only once,

independent of distribution of $(X_t)_{\mathbb{N}}$ and without estimation,

as OPs are distribution-free.

Some in-control designs with $m = 3$ and target ARL$_0 \approx 370$:

	H-chart		H_{ex}-chart		Δ-chart		β-chart		τ-chart		δ-chart	
λ	LCL	ARL	LCL	ARL	UCL	ARL	CL	ARL	CL	ARL	CL	ARL
0.25	1.014	367.8	0.6621	369.6	0.3338	369.3	0.6437	369.7	0.4253	368.6	0.7656	368.4
0.10	1.4601	369.8	0.8405	369.9	0.1115	369.5	0.3638	369.9	0.2529	369.8	0.4876	370.4
0.05	1.6356	370.0	0.88017	369.3	0.05125	370.3	0.233	370.0	0.16775	369.4	0.3246	369.5

HELMUT SCHMIDT UNIVERSITÄT
Universität der Bundeswehr Hamburg

MATH
STAT

Performance Analyses of OP-Charts

Brief Summary

Out-of-control models:

DGP 1: AR(1) process $X_t = \alpha \cdot X_{t-1} + \epsilon_t$ with $\epsilon_t \sim \mathrm{N}(0,1)$
(linear and time-reversible process);

DGP 2: TEAR(1) process $X_t = B_t^{(\alpha)} \cdot X_{t-1} + (1-\alpha) \cdot \epsilon_t$ with
$\epsilon_t \sim \mathrm{Exp}(1)$, where i. i. d. Bernoulli $B_t^{(\alpha)}$ with $P\big(B_t^{(\alpha)} = 1\big) = \alpha$
(AR(1)-like ACF, but asymmetric in time);

DGP 3: AAR(1) process $X_t = \alpha \cdot |X_{t-1}| + \epsilon_t$ with $\epsilon_t \sim \mathrm{N}(0,1)$,

DGP 4: QAR(1) process $X_t = \alpha \cdot X_{t-1}^2 + \epsilon_t$ with $\epsilon_t \sim \mathrm{N}(0,1)$
(both with positive shocks and thus asymmetry).

Some conclusions from simulated (zero-state) ARLs:

- AR(1) with $\alpha > 0$: τ-chart superior, better if $\lambda = 0.25$;
- AR(1) with $\alpha < 0$: τ-chart superior, clearly best if $\lambda = 0.05$.
- TEAR(1) generates long-lasting rises followed by abrupt falls, so upward OP $(1, 2, 3)$ dominant. Thus, β-chart superior, improves with decreasing λ. τ-chart visibly worse.
- AAR(1) and QAR(1): If low α, β-chart with $\lambda = 0.05$ recommended. For larger α, τ-chart with $\lambda = 0.25$ superior.
- In all cases, H_{ex}- and Δ-chart not optimal but reasonable.

Note: conditional expected delays do not differ much from zero-state ARLs.

Only reasonable competitor:

EWMA-type ACF(1) chart by Gardner (1983).

Design requires fully specified in-control model!

Example: If X_t are i. i. d. N(0, 1), then

$$\left.\begin{array}{l} C_0 = 0,\;\; C_t = \lambda\, X_t\, X_{t+1} + (1-\lambda)\, C_{t-1}, \\ S_0 = 1,\;\; S_t = \lambda\, X_t^2 + (1-\lambda)\, S_{t-1}, \end{array}\right\} \Rightarrow \qquad A_t = \frac{C_t}{S_t}.$$

ACF-chart does better than τ-chart if AR(1),

reasonable as ACF designed for linear dependence.

But for non-linear TEAR(1), H_{ex}-, Δ-, and β-charts with $\lambda \leq 0.1$

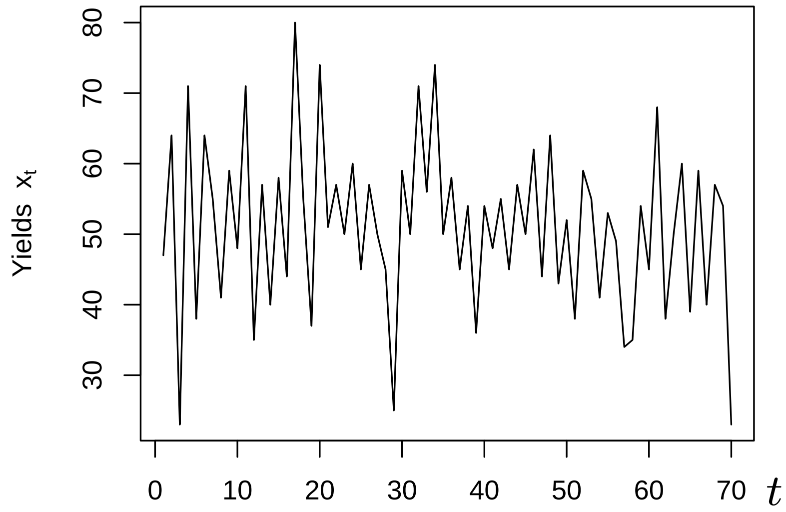
uniquely outperform any ACF-chart.

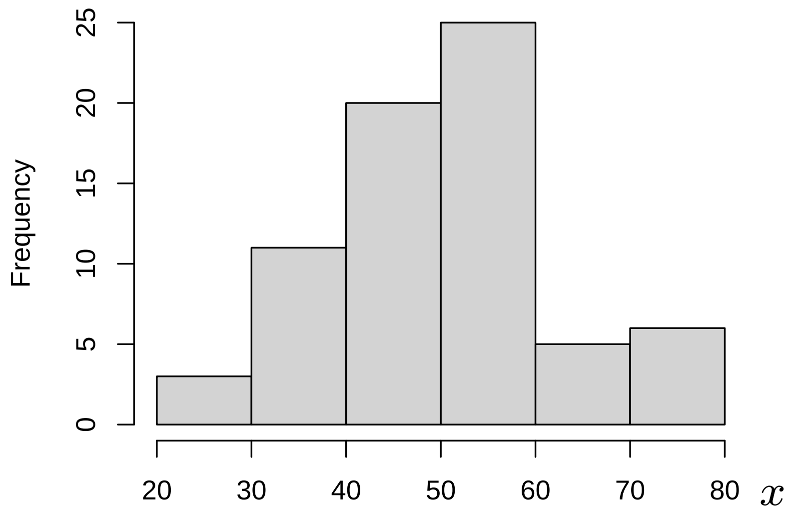
Application: Chemical Process Data

Illustration

Chemical process data $x_1, \ldots, x_{70}$, used (among others) by Beneke et al. (1988) to illustrate periodogram chart ($w = 50 \;\Rightarrow$ alarm not before x_{50}; first alarm at $t = 67$).

Data express consecutive batch yields. Negative dependence between adjacent batches reasonable as residues of high-yielding batch could reduce yield of subsequent batch.





Application: Chemical Process Data

Design of parametric charts not clear, because

non-standard distribution (see histogram of Phase-II data) and

no Phase-I data available (Beneke et al. (1988): normality).

Novel OP-charts immediately applicable to Phase-II data!

As we anticipate possibly negative dependence:

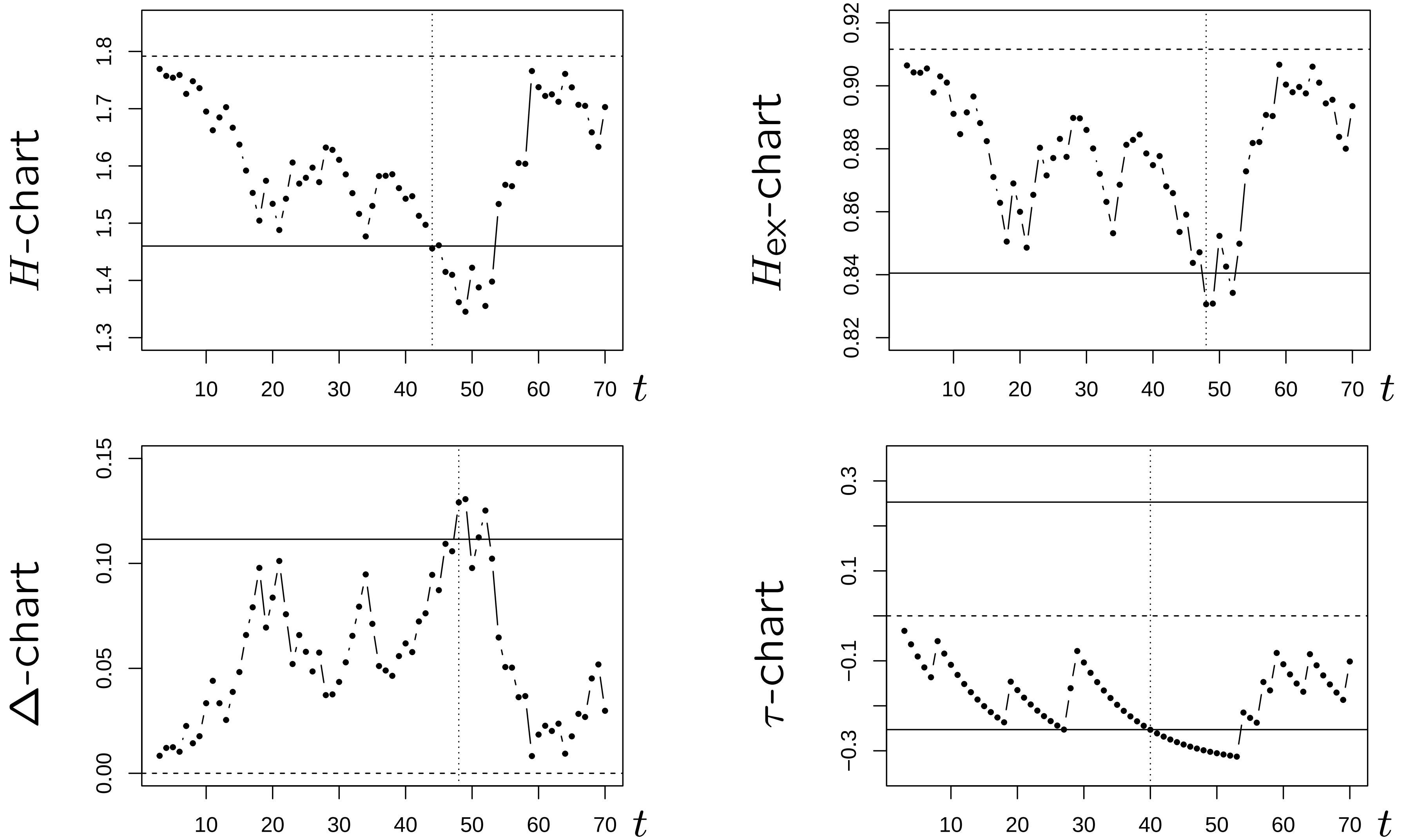
H-, H_{ex}-, Δ-, and τ-charts appropriate, esp. with $\lambda = 0.05$.

Results: If $\lambda = 0.25$, alarm by none chart.

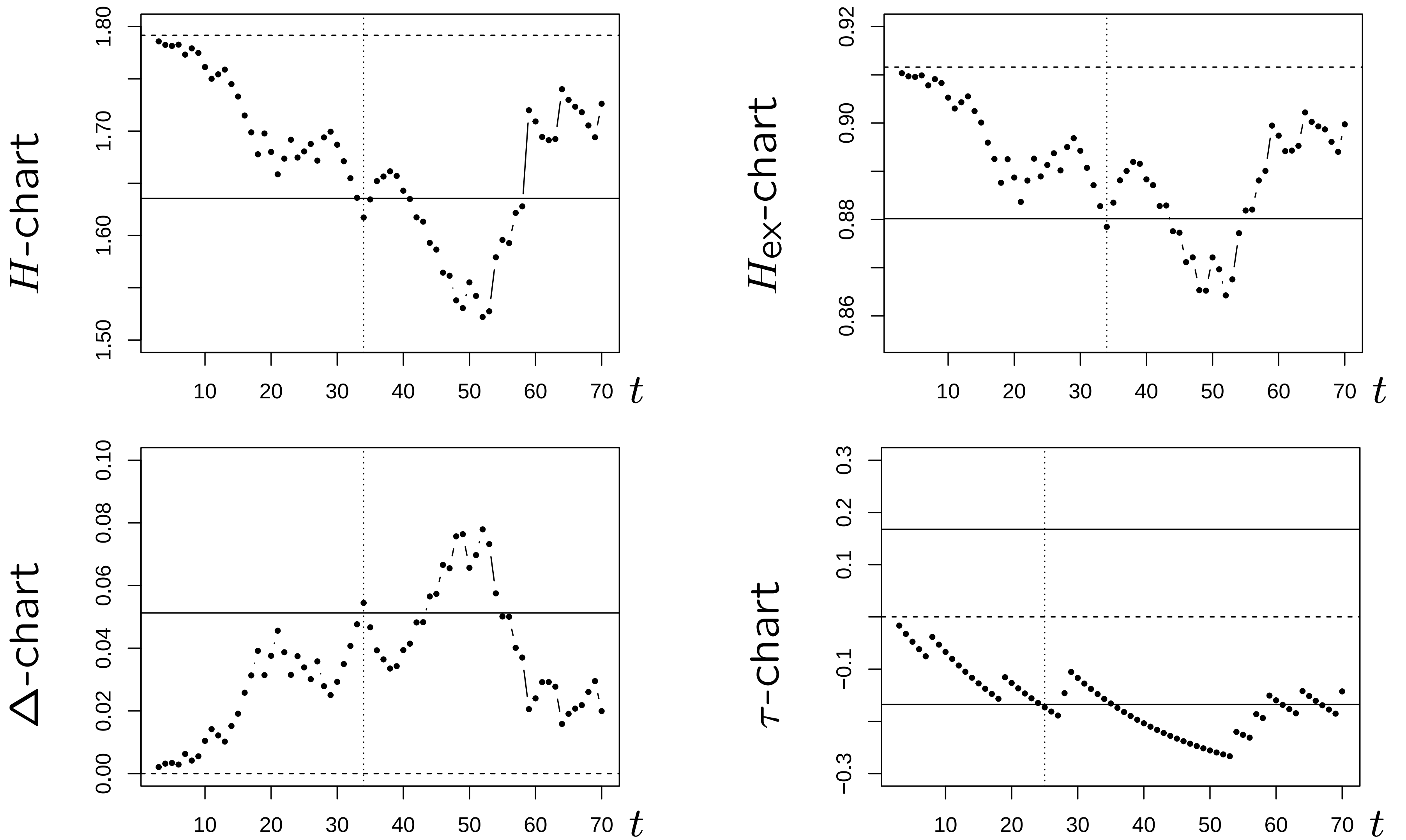
Charts get more sensitive with decreasing λ: ...

Application: Chemical Process Data

Novel **EWMA OP-charts** with $\lambda = 0.10$:



Novel **EWMA OP-charts** with $\lambda = 0.05$:



Conclusions

Chakraborti & Sparks (2020, pp. 137–138):

"Most statistical process monitoring begins with an assumed model (. . .) and further assumptions about the components of the model. (. . .) two of the important and common assumptions [are] normality and independence. (. . .) Violations of one or more of the assumptions might render the decisions invalid and hence useless even though there would seem nothing wrong in terms of crunching the numbers."

⇒ These issues solved by **novel EWMA OP-charts**, which monitor assumption of serial independence by distribution-free approach!

Thank You

for Your Interest!

Christian H. Weiß

Department of Mathematics & Statistics

Helmut Schmidt University, Hamburg

weissc@hsu-hh.de

Literature

Alwan (1992) Effects of autocorrelation on . . . *CSTM* **21**, 1025–1049.

Atienza et al. (1997) ARL properties of a sample . . . *CAIE* **33**, 733–736.

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

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

Beneke et al. (1988) Spectral analysis in quality . . . *Technom.* **30**, 63–70.

Chakraborti & Sparks (2020) SPM. . . In *Dist.-Free Meth. . .* , 137–155.

Gardner (1983) Automatic monitoring of forecast errors. *J Forec* **2**, 1–21.

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

Knoth & Schmid (2004) Control charts for time. . . In *Frontiers 7*, 210–236.

Montgomery (2009) *Introd. Statistical Quality Control*. 6th ed., Wiley.

Weiß (2021) Non-parametric tests . . . *Chaos* **32**, 093107.

Yourstone & Montg. (1991) Detection of process . . . *QREI* **7**, 133–140.