Europa-Universität Viadrina Frankfurt (Oder)

Control Charts for Time Series: A Review

Sven Knoth & Wolfgang Schmid

Dep. of Statistics,
Europa–Universität Viadrina

knoth@euv-frankfurt-o.de,
schmid@euv-frankfurt-o.de

VIIth International Workshop on Intelligent Statistical Quality Control Waterloo/Canada

September 5, 2001

Outline

1. Introduction
2. Time Series Models and Standard Control Charts
3. Residual Charts
4. Modified Charts
5. Comparison Study for AR(1) Data

History

Statistical Process Control (SPC):

Shewhart (1926, 31),
Page (1954),
Roberts (1959, 66)

Time Series Analysis (TSA):

Yule/Walker (1920/30ies),
Grenander/Rosenblatt (1957),
Box/Jenkins (1970),
Engle/Bollerslev (1982, 86)

SPC for Autocorrelated Data:

Goldsmith/Whitfield (1961),
Bagshaw/Johnson (1974, 75, 77),
Nikiforov (1975, 79, 83, 84),
Rowlands (1976),
Stamboulis (1971), Vasilopoulos (1974), V/St (1978),
Alwan (1989)

Change Point Model for the Mean

$$X_t = Y_t + \sqrt{\gamma_0}\, a\, I_{\{m,m+1,\ldots\}}(t)$$

with observed process X_t and target process Y_t,

$E(Y_t) = 0$, $\gamma_0 = Var(Y_t)$,

a is known,

$I_{\{m,m+1,\ldots\}}(t)$ – indicator function

Notation:
$t < m$ – in control ($X_t = Y_t$),
$t \geq m$ – out of control ($X_t \neq Y_t$)

$P_{m,a}(\cdot)$, $E_{m,a}(\cdot)$ – probability measure, expectation for change point m and shift a

In the sequel $m = 1 \rightsquigarrow P_a(\cdot)$, $E_a(\cdot)$

$\rightarrow$ ARL (Average Run Length) $= E_a(N)$
with control chart stopping time (run length) N

(cf. steady state ARL $= \lim\limits_{m\to\infty} E_{m,a}(L - m + 1 \mid L \geq m)$)

Time Series Models

for the target process Y_t ($E(Y_t) = 0$)

1. ARMA models (with Gaussian noise)

$$Y_t = \sum_{i=1}^{p} \alpha_i \, Y_{t-i} + \varepsilon_t + \sum_{j=1}^{q} \beta_j \, \varepsilon_{t-j}$$

   with $\varepsilon_t \sim \mathcal{N}(0, \sigma^2)$

2. more general (discrete) Gaussian processes

3. GARCH processes

$$Y_t = \sigma_t \, \varepsilon_t \; , \; \sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i \, Y_{t-i}^2 + \sum_{j=1}^{q} \beta_j \, \sigma_{t-j}^2$$

Standard Control Charts

Two-sided Charts for Individual Observations

$(E(Y_t) = 0)$

1. Shewhart chart (1926,31)

$$N_s = \inf\left\{ t \in \mathbb{N} : \ |X_t| \geq c_s \sqrt{\gamma_0} \right\}$$

2. CUSUM chart (Page 1954)

$$N_c = \inf\left\{ t \in \mathbb{N} : \ \max\{-S_t^-, S_t^+\} \geq c_c \sqrt{\gamma_0} \right\}$$

with $S_t^+ \underset{t\geq 1}{=} \max\{0, S_{t-1}^+ + X_t - k\}$, $S_0^+ = 0$, S_t^- ...

3. EWMA chart (Roberts 1959)

$$N_e = \inf\left\{ t \in \mathbb{N} : \ |Z_t| \geq c_e \sqrt{\lambda/(2-\lambda)} \sqrt{\gamma_0} \right\}$$

with $Z_t \underset{t\geq 1}{=} (1-\lambda) Z_{t-1} + \lambda X_t$, $Z_0 = 0$

Effects of Autocorrelation I

Example: AR(1) data, i. e. $Y_t = \alpha\, Y_{t-1} + \varepsilon_t$,

with $\gamma_0 = 1$, $|\alpha| < 1$

If *Analyst*
falsely assumes independence
and designs related charts

for
- $a = \sqrt{\gamma_0} = 1$ ($\rightsquigarrow k = 0.5$, $\lambda = 0.1$),
- $E_0(N) = 500$ (in control)

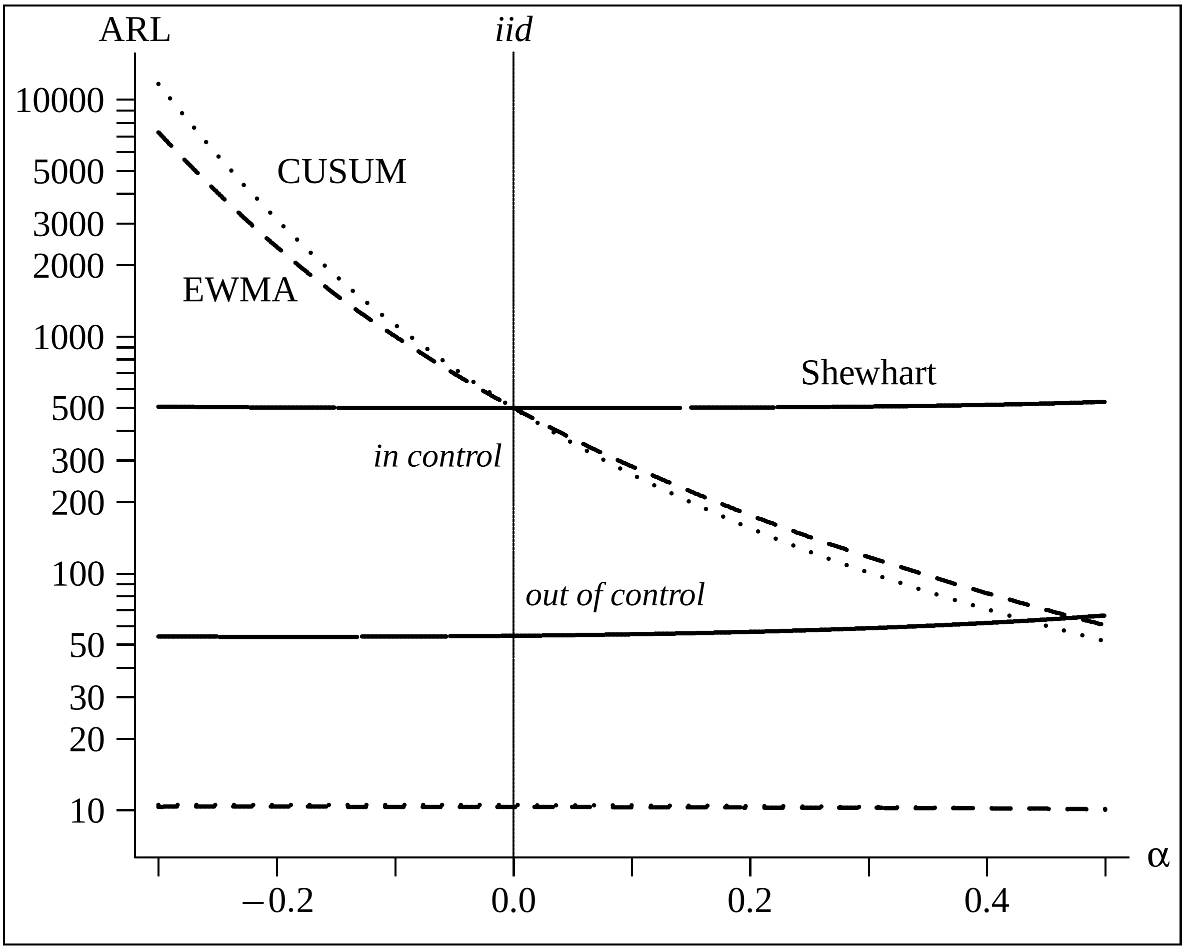
then

$c_s = 3.0902$, $c_c = 5.0707$, $c_e = 2.8143$

$E_1(N_s) = 54.6$, $E_1(N_c) = 10.5$, $E_1(N_e) = 10.3$,

but … →

Effects of Autocorrelation II



Control Charts Types for Time Series

- **Modified Charts**

control charts with modified parameters
(control limits, k, λ, variance term)

applied to

original data

- **Residual Charts**

standard control charts

applied to

transformed data
(model residuals, forecast errors etc., CUSUM/SPRT)

3. Residual Charts

Goal: time series data → transformation → iid data

early applications:

Davies (1964) – EWMA forecasts, CUSUM charts to forecast errors

Ermer (1980) – (in control) ARMA fit, Shewhart chart to the sequence of sums of squares of residuals

Method here: forecast (prediction) error $\hat{\varepsilon}_t = X_t - \hat{X}_t$

X_t – stationary process: $\hat{X}_t = \sum_{i=1}^{t-1} \phi_{ti} X_{t-i}$

X_t – AR(1): $\hat{X}_t \underset{t>1}{=} \alpha X_{t-1}$, $\hat{X}_1 = 0$

normalization of $\hat{\varepsilon}_t$ → standard Control Charts

Remark: starting problem, see Kramer/Schmid (1997)

AR(1)

$$\hat{\Delta}_1 = \frac{X_1}{\sqrt{\gamma_0}}\,,\ \hat{\Delta}_t \underset{t>1}{=} \frac{\sqrt{1-\alpha^2}\,\big(X_t - \alpha\,X_{t-1}\big)}{\sqrt{\gamma_0}}$$

in control: $\hat{\Delta}_t \sim \mathcal{N}(0,1)$

out of control: $\hat{\Delta}_t \sim \mathcal{N}(\mu_t,1)$ with $\mu_t = \begin{cases} a & t=1 \\ a\,\sqrt{\frac{1-\alpha}{1+\alpha}} & t>1 \end{cases}$

$\rightsquigarrow$ computation of chart characteristics as in the iid case (some slight adaptions around the change point)

CUSUM Charts

"consequential CUSUM"

Idea: CUSUM and likelihood ratio (SPRT)

Nikiforov (1975-84), Yashchin (1993), Schmid (1997)

example: AR(1)

Schmid

$$S_t^+ \underset{t>1}{=} \max \Big\{ 0, \varepsilon_t - k\,\sqrt{\gamma_0},$$

$$S_{t-1}^+ + (1-\alpha)\left(\varepsilon_t - (1-\alpha)\,k\,\sqrt{\gamma_0}\right)\Big\}\,,$$

$$S_1^+ = \max\Big\{0,\ (1-\alpha^2)\,(X_1 - k\,\sqrt{\gamma_0})\Big\}$$

Nikiforov

$$S_t^+ \underset{t>1}{=} \max\Big\{0, S_{t-1}^+ + (1-\alpha)\left(\varepsilon_t - (1-\alpha)\,k\,\sqrt{\gamma_0}\right)\Big\}\,,$$

$$S_1^+ = \max\Big\{0, (1-\alpha^2)\,(X_1 - k\,\sqrt{\gamma_0})\Big\}$$

further: Timmer/Pignatiello/Longnecker (1998), Sparks (2000), Lu/Reynolds (2001)

4. Modified Charts

Idea: Adapt the control limits (variance term inclusive) and the other chart parameters in order to get the correct in-control ARL and an "optimal" design (rules for λ, k).

$\rightsquigarrow$ more complex algorithms for the calculation of the chart characteristics (ARL, steady state ARL ... control limits)

$\rightsquigarrow$ Monte Carlo methods

exceptions: low order ARMA

- AR(1): Markov chain approach for Shewhart chart
- AR(1), ARMA(1,1): Rowlands (1976) solved integral equations for CUSUM
- ARMA: Yashchin (1993) replaced the dependent observations by a "similar" iid sequence

disadvantage: cannot transfer iid chart setup rules (λ, k)

Theoretical Results – Bounds I

- Schmid (1995), Shewhart charts, in control

Th. 1: $\{Y_t\}$ Gaussian process

$$\rightsquigarrow \; P(N_s > k) \; \geq \; P_{iid}(N_s > k) \; = \; \big(2\,\Phi(c_s) - 1\big)^k$$

$$E(N) = \sum_{k=0}^{\infty} P\big(N_s > k\big) \; \rightsquigarrow \; E(N_s) \geq E_{iid}(N)$$

Th. 2: $\{Y_t\}$ stationary Gaussian AR(1)

$\rightsquigarrow \; P(N_s > k)$ is a nondecreasing function in $|\alpha|$.

Theoretical Results – Bounds II

1-sided EWMA: $N_e^u = \inf \left\{ t \in \mathbb{N} : Z_t > c \sqrt{Var(Z_t)} \right\}$

- Schmid/Schöne (1997)

Th. 3: $\{Y_t\}$ Gaussian process, $\gamma_v \geq 0$

$$\rightsquigarrow \; P(N_e^u > k) \;\geq\; P_{iid}(N_e^u > k)$$

- Schöne et al. (1999)

Th. 4: $\{Y_t\}$ Gaussian process with $\{\gamma_v\}$, $\{\delta_v\}$, resp.

Assume that

$$\gamma_0 > 0,\ \delta_0 > 0,\ \gamma_h \geq 0,\ \delta_h \geq 0 \quad \text{for} \quad 1 \leq h \leq k-1\,,$$

$$\gamma_i\,\delta_j \quad \geq \quad \gamma_j\,\delta_i \quad \text{for} \quad 0 \leq j < i \leq k-1 \qquad (*)$$

$$\rightsquigarrow \; P_\gamma(N_e^u > k) \;\geq\; P_\delta(N_e^u > k)$$

$(*)$ for positive $\{\gamma_v\}$, $\{\delta_v\}$: $\dfrac{\gamma_i}{\gamma_j} \geq \dfrac{\delta_i}{\delta_j}$

Further Remarks

- Influence of parameter estimation of the time series model

  Kramer/Schmid (1997) Shewhart charts

- Control charts for GARCH

  Severin/Schmid (1999) – mean, application

  Pawlak/Schmid (2001) – mean, Shewhart, theoretical

  Schipper/Schmid (2001) – scale, application

5. Comparison Study

- Target process $Y_t = \alpha Y_{t-1} + \varepsilon_t$

  with $\varepsilon_t \sim \mathcal{N}(0,1)$ and $Cov(\varepsilon_s, \varepsilon_t) = 0$ for $s \neq t$.

- cp-model $X_t = Y_t + a\sqrt{\gamma_0}$,

  where $a \in \{0, 0.5, 1.0, \ldots, 3.0\}$
  and $Var(X_t) = Var(Y_t) = \gamma_0 = 1/(1-\alpha^2)$.

- Control charts:

  1. modified EWMA, abbr. EWMAmod,
  2. residual EWMA, abbr. EWMAres,
  3. (conseq.) CUSUM (Schmid), abbr. CUSUMmod,
  4. residual CUSUM chart, abbr. CUSUMres,
  5. CUSUM with mod. control limits, abbr. CUSUMcla

Comparison Study II

- in-control ARL 500
- EWMA

  $\lambda \in \{0.01, 0.025, 0.05, 0.1, 0.2, 0.4, 0.6, 0.8, 1\}$

  for EWMAres $\ldots \cup \{0.3, 0.5, 0.7, 0.9\}$
- CUSUM

  $k \in \{0, 0.1, 0.2, \ldots, 2.0\}$

  for CUSUMcla only $\{0, 0.25, 0.5, \ldots, 2.0\}$
- Computation methods

  EWMAmod, CUSUMcla – Monte Carlo with 10,000,000 repetitions

  others – Markov chain approximation (Brook/Evans 1972) with matrix dimension 300

Out-of-control ARLs for different Control Charts – negative Autocorrelation

α	chart	shift a 0.5	1.0	2.0	3.0
-0.6	EWMAmod	(.1)11.102	(.2)4.389	(.4)1.938	(.4,.6)1.363
	EWMAres	(.1)**10.823**	(.4)4.058	(.8)1.988	(1)1.537
	CUSUMmod	(.3)11.364	(.5)4.128	(2.4)**1.851**	(2.4)**1.314**
	CUSUMres	(.4)1.964	(.8)**4.053**	(1.4)1.997	(2.4)1.537
	CUSUMcla	(.25)11.962	(.25)4.930	(.5)2.206	(.75)1.531
-0.4	EWMAmod	(.1)15.947	(.2)5.854	(.4)2.339	(.6)**1.482**
	EWMAres	(.1)**15.738**	(.2)5.716	(.6)2.319	(.9)1.567
	CUSUMmod	(.3)16.900	(.5)5.817	(1.2)2.353	(2.4)1.500
	CUSUMres	(.4)16.571	(.7)**5.706**	(1.2)**2.318**	(1.9)1.564
	CUSUMcla	(.25)16.898	(.5)6.142	(.75)2.476	(1)1.581
-0.2	EWMAmod	(.05)21.735	(.2)7.782	(.4)2.846	(.6)1.647
	EWMAres	(.05)**21.711**	(.2)**7.675**	(.5)2.820	(.8)1.669
	CUSUMmod	(.3)23.410	(.5)7.878	(1)2.818	(1.5)**1.645**
	CUSUMres	(.3)2.945	(.6)7.786	(1.1)**2.770**	(1.7)1.646
	CUSUMcla	(.25)23.160	(.5)7.944	(.75)2.865	(1.25)1.664
0.0	EWMAmod	(.05)**28.766**	(.1)**10.333**	(.4)3.522	(.7)1.865
	EWMAres	(.05)**28.766**	(.1)**10.333**	(.4)3.522	(.7)1.865
	CUSUMmod	(.3)31.480	(.5)10.519	(1)**3.413**	(1.5)**1.792**
	CUSUMres	(.3)31.480	(.5)10.519	(1)**3.413**	(1.5)**1.792**
	CUSUMcla	(.3)31.480	(.5)10.519	(1)**3.413**	(1.5)**1.792**

Out-of-control ARLs for different Control Charts – positive Autocorrelation

α	chart	shift a 0.5	1.0	2.0	3.0
0.0	EWMAmod	(.05)**28.766**	(.1)**10.333**	(.4)3.522	(.7)1.865
	EWMAres	(.05)**28.766**	(.1)**10.333**	(.4)3.522	(.7)1.865
	CUSUMmod	(.3)31.480	(.5)[illegible].519	(1)**3.413**	(1.5)**1.792**
	CUSUMres	(.3)31.480	(.5)[illegible].519	(1)**3.413**	(1.5)**1.792**
	CUSUMcla	(.3)31.480	(.5)[illegible].519	(1)**3.413**	(1.5)**1.792**
0.2	EWMAmod	(.025)38.539	(.1)**13.590**	(.4)4.493	(.8)2.166
	EWMAres	(.025)**38.520**	(.1)13.677	(.3)4.537	(.6)2.203
	CUSUMmod	(.3)42.393	(.5)14.353	(1)4.409	(1.5)2.068
	CUSUMres	(.2)41.878	(.4)14.325	(.8)4.394	(1.3)2.038
	CUSUMcla	(.25)41.473	(.5)14.067	(1)**4.293**	(1.75)**1.988**
0.4	EWMAmod	(.025)**51.486**	(.1)**18.815**	(.2)5.977	(.8)2.552
	EWMAres	(.025)51.564	(.05)19.123	(.2)6.145	(.5)2.822
	CUSUMmod	(.3)58.273	(.5)[illegible].300	(1)6.040	(1.5)2.516
	CUSUMres	(.2)58.824	(.3)[illegible].120	(.6)6.022	(1.1)2.503
	CUSUMcla	(.25)56.499	(.5)19.422	(1.25)**5.642**	(2)**2.301**
0.6	EWMAmod	(.025)**74.066**	(.05)**27.286**	(.2)8.409	(1)3.007
	EWMAres	(.01)74.371	(.05)27.800	(.1)9.272	(.3)4.108
	CUSUMmod	(.2)84.833	(.5)31.238	(1)9.242	(1.5)3.303
	CUSUMres	(.1)83.026	(.2)[illegible].347	(.4)9.169	(.8)3.562
	CUSUMcla	(.25)[illegible]998	(.5)28.636	(1.25)**7.985**	(2)**2.860**

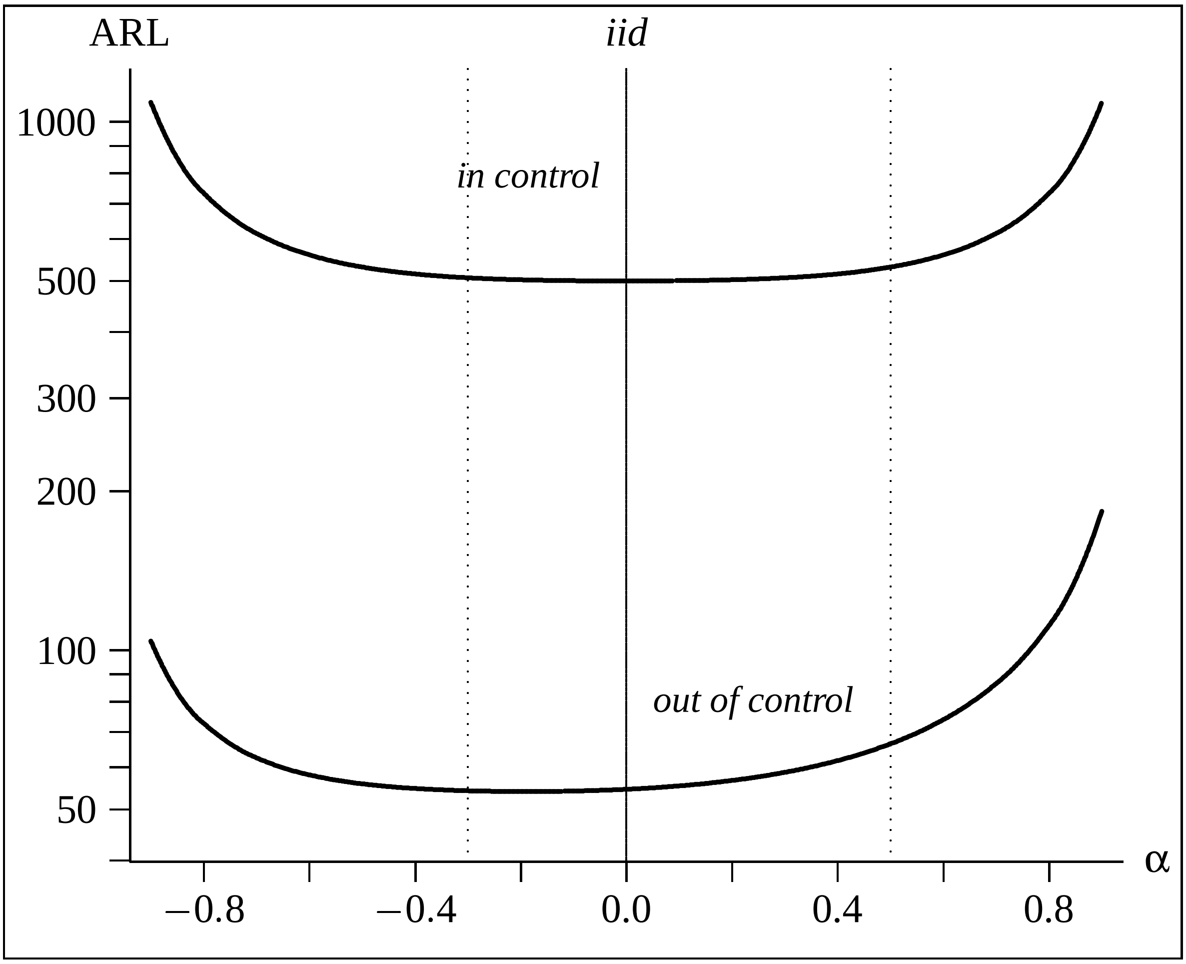
Conclusions

- Consideration of dependence is necessary and useful
- There is no obvious superior scheme (at least for AR(1))
- Residual schemes are more sophisticated, modified schemes more agreeable (to applications)
- Dependence is a more frequently investigated pattern in SPC than known
- Steady-state ARL is more suitable than the ARL (from which EWMA takes advantage) because of the starting problem for time series models
- Automatic process control (APC) might be a link between standard SPC schemes and dependence considerations (engineering literature)
- Only few results in literature on optimality in the case of dependence (cf. Lai 1998 – similar to Lorden 1971)

Earlier Reviews and Surveys

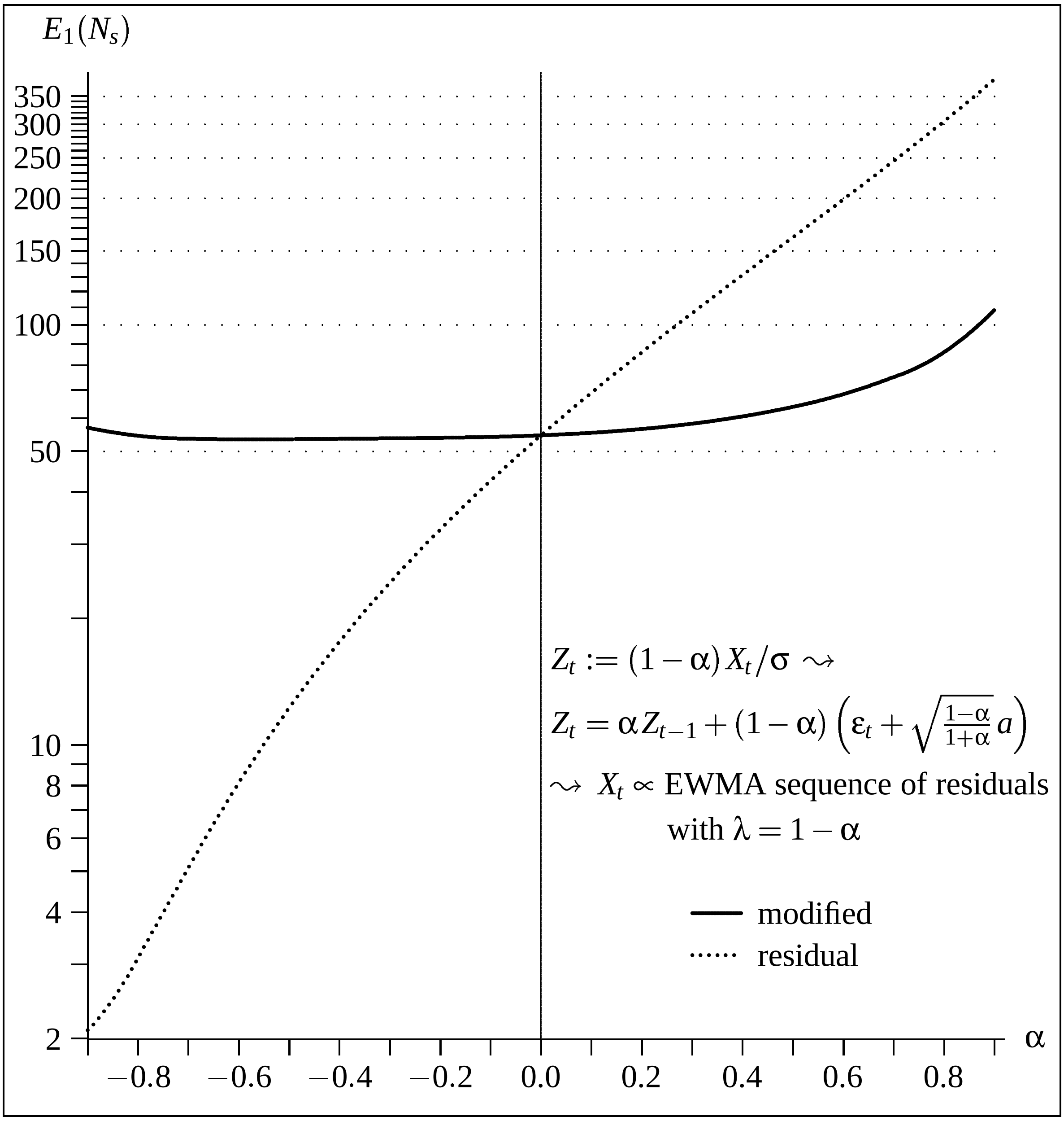
Kligene/Tel'ksnis (1982),
Basseville (1988),
Rowlands/Wetherill (1991),
Montgomery/Mastrangelo (1991),
Basseville/Nikiforov (1993)

Effects of (auto)correlation III



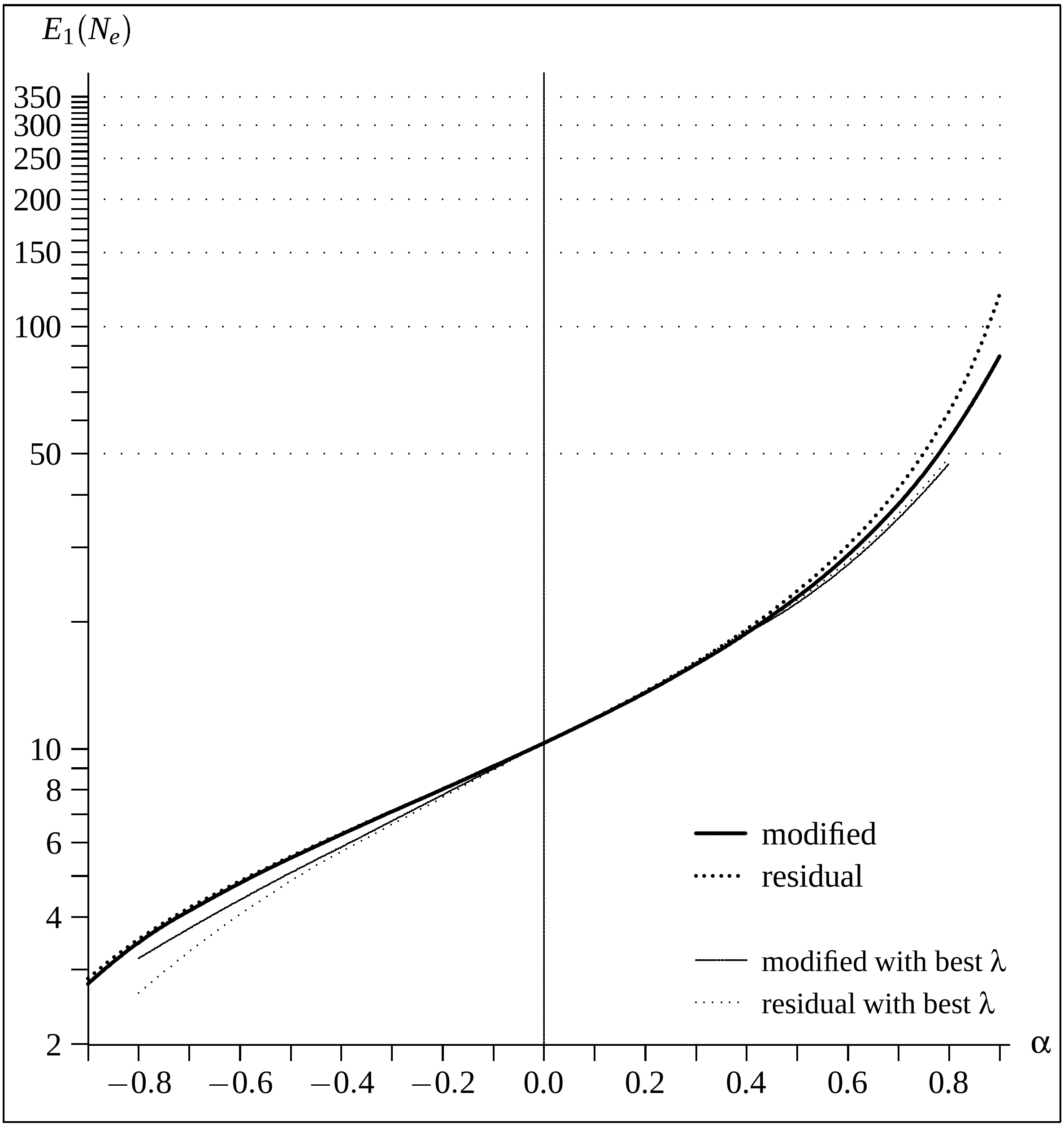
Residual vs. Modified Shewhart Chart

$E_0(N_s) = 500$, $a = 1$,
out-of-control ARL $E_1(N_s)$ vs. autocorrelation coefficient α



Residual vs. Modified EWMA Chart

$\lambda = 0.1$, $E_0(N_e) = 500$, $a = 1$,
out-of-control ARL $E_1(N_e)$ vs. autocorrelation coefficient α



Timmer/Pignatiello/Longnecker (1998)

Different cp-Model

$$X_t = \delta_t + \alpha\, X_{t-1} + \varepsilon_t \quad ,$$

$$\delta_t = \begin{cases} \delta_0 & , t < m \\ \delta_A & , t \geq m \end{cases} ,$$

$$E(X_t) = \begin{cases} \dfrac{\delta_0}{1-\alpha} & , t < m \\ \dfrac{\delta_A}{1-\alpha} - \dfrac{\alpha^{t-m+1}}{1-\alpha}(\delta_A - \delta_0) & , t \geq m \end{cases}$$