Received: 22 June 2022 | Revised: 2 September 2022 | Accepted: 26 September 2022

DOI: 10.1002/qre.3215

RESEARCH ARTICLE

WILEY

Monitoring count time series: Robustness to nonlinearity when linear models are utilized

Christian H. Weiß[1] | **Murat Caner Testik**[2]

[1]Department of Mathematics and Statistics, Helmut Schmidt University, Hamburg, Germany

[2]Department of Industrial Engineering, Hacettepe University, Ankara, Turkey

Correspondence
Christian H. Weiß, Department of Mathematics and Statistics, Helmut Schmidt University, Holstenhofweg 85, 22043 Hamburg, Germany.
Email: weissc@hsu-hh.de

Abstract

Linear models are typically utilized for time series analysis as these are often simple to implement and interpret, as well as being useful in modeling many practical phenomena. Hence, most of the literature on control charts for monitoring time series also consider linearity of the data generating processes (DGP). In practice, however, a nonlinear DGP can be modeled as if it is linear, either due to overlook or illusion, when it is approximately linear. This study quantifies the effects of nonlinear DGPs, misspecified and modeled as linear, on the performance of Shewhart-type and cumulative sum (CUSUM) control charts for count time series data. Time series models for bounded and unbounded counts with several parametrizations are considered for studying the sensitivity of linear approximations to nonlinear DGPs. The Markov chain (MC) approach and simulations are used to compute the average run length (ARL) performance of the control charts. Robustness of the performance is evaluated with respect to the extent that the DGP violates linearity, with and without errors in parameter estimation. It is shown that the chart designs are, in general, robust to model misspecification when the parameters are specified. However, the CUSUM performance can be affected significantly if both the model is misspecified and the parameters are estimated. Real-world data implementations are provided to illustrate the sensitivity of the control charts' performances to the type of model and to the estimation approach.

KEYWORDS

control charts, count time series, INGARCH models, model misspecification, robustness, soft clipping, softplus

1 | INTRODUCTION

Linear models are of utmost relevance in research and practice of time series analysis.[1] For instance, when the time series data are real-valued, AutoRegressive Moving Average (ARMA) models[2] are commonly employed since these linear models are "simple, useful, and interpretable in a wide range of contexts", see p. 479 in Ref. 3. Linear models are not limited to real-valued time series but also defined for discrete-valued data[4] such as count time series (i. e., where the outcomes are

contained in the set of non-negative integers, $\mathbb{N}_0 = \{0, 1, ...\}$). With discrete-valued data, linearity is assumed with respect to the conditional mean,[3] which leads to the well-known Yule–Walker equations for the autocorrelation function (ACF).

In statistical process control implementations, the presence of autocorrelation in the time series data may have significant effects on the performance of a control chart, if ignored (see, for example, the seminal paper by Alwan[5] and the more recent surveys by Knoth and Schmid[6]; Psarakis and Papaleonida,[7] and Weiß[8]). Consequently, the literature on control charts for time series data often focus on selected model classes, and then effects of different parameter values within these classes are analyzed. In the case of real-valued time series, for example, one typically considers a data generating process (DGP) from the ARMA family,[6,7] whereas for count time series, the DGP commonly stems from either the integer-valued ARMA (INARMA) or integer-valued generalized autoregressive conditional heteroskedasticity (INGARCH) family.[8,9] As performance may be severely affected by the serial dependence of data, rules for designing control charts under ARMA, INARMA, and INGARCH assumptions have been derived by several researchers.

In practice, however, we do not know the true DGP behind the data. Therefore, methods from time series analysis have to be used to identify an appropriate model type, and then to estimate its model parameters. The latter step causes estimation uncertainty, which affects, among others, the performance of control charts being constructed based on that model fit, see Jensen et al.[10] and Psarakis et al.[11] for surveys. But also the first step, the model identification, may cause performance deteriorations, namely if the selected type of model deviates from the DGP's true model. In this research, we are concerned with time series monitoring, and our main focus regarding model uncertainty is on the actual model structure. As outlined in the first paragraph, linear time series models are quite popular among practitioners. Even if the true DGP is nonlinear, a linear model can be identified as appropriate, for example when the sample ACF roughly satisfies the Yule–Walker equations. In real world implementations, this type of model uncertainty will certainly occur in combination with additional estimation uncertainty.

In this research, we investigate to what extent the performance of Shewhart-type and cumulative sum (CUSUM) control charts depends on the particular assumption of *linear* dependence. In other words, our research question is: if a control chart is designed based on the assumption of an exactly linear DGP (as usually done in practice) but the true DGP is only *approximately linear*, how does this effect the chart's performance? It is assumed that even if a sample ACF follows a set of Yule–Walker equations in good approximation, this does not necessarily imply that the DGP is really *exactly linear*. Instead, there might be violations of linearity that are overlooked based on the available time series data. Our analyses are for the case of count time series monitoring, where we consider both unbounded counts (having full $\mathbb{N}_0$ as their range) and bounded counts (with range $\{0, \dots, n\}$ for given $n \in \mathbb{N} = \{1, 2, ...\}$). We consider the INGARCH and bounded INGARCH (BINGARCH) models as the *exactly linear* models[12,13] that are fitted to the data by the user. Note that these linear INGARCH models have already found several applications in SPC, such as in Weiß and Testik,[14] Huh et al.,[15] Rakitzis et al.,[16] Vanli et al.,[17] Ottenstreuer,[18] and Anastasopoulou and Rakitzis.[19] For the true DGP, we need a corresponding model family that allows to easily control the extent of nonlinearity. We use the softplus INGARCH (spINGARCH) and soft-clipping BINGARCH (scBINGARCH) models for this purpose, which were originally proposed as *approximately linear* counterparts of INGARCH and BINGARCH, respectively.[20,21] However, here, as stated before, our focus is not on approximating linearity, but on deviating from linearity up to an arbitrary extent, and to investigate the effect of such deviations (possibly together with estimation uncertainty) on the control charts' performance.

The paper is organized as follows. After having introduced the considered INGARCH-type models in Section 2, we provide a detailed description on the design of our robustness study in Section 3. In Section 4, we investigate the Shewhart and CUSUM control charts for count DGPs under the premise of specified parameters but illusory linearity, as well as the effect of nonlinearity on the stochastic properties of the actual DGP. In Section 5, we move forward to a performance analysis of the control charts under both model uncertainty and estimation uncertainty, and we analyze how parameter estimators are affected by violations of linearity. The practical aspects of a possible model misspecification are illustrated by two real-data examples in Section 6, one from economics and the other from service and healthcare industries. Finally, we conclude in Section 7 and outline issues for future research.

2 | LINEAR AND APPROXIMATELY LINEAR INGARCH MODELS

Despite their somewhat misleading name, see p. 74 in Ref. [4], the (B)INGARCH models possess an exactly linear structure and lead to an ARMA-like ACF, that is, satisfying a set of Yule–Walker equations. If the count process $(X_t) = (X_t)_{t\in\mathbb{Z}=\{\dots,-1,0,1,\dots\}}$ is unbounded, then the INGARCH(p, q) model of order $(p, q) \in \mathbb{N} \times \mathbb{N}_0$ is defined with respect to the

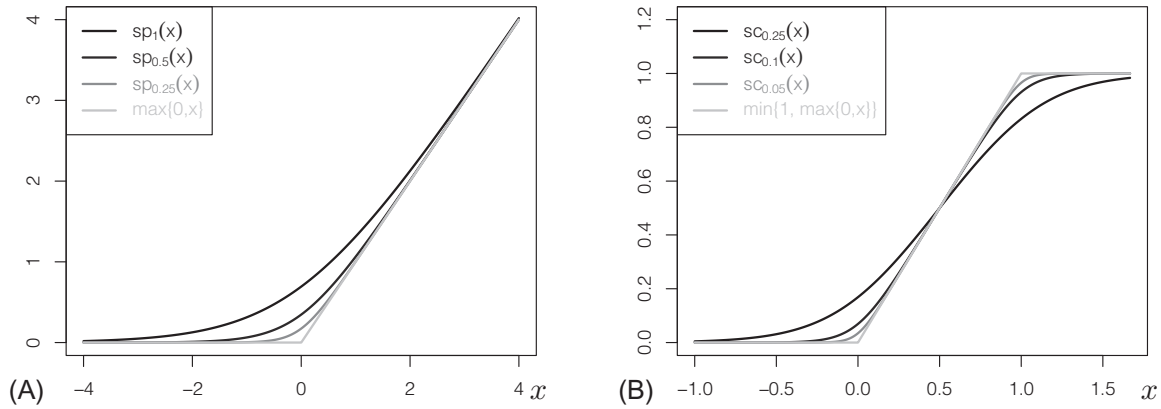

FIGURE 1 Plots of (A) softplus function and (B) soft clipping function against x

conditional mean $M_t = E[X_t|X_{t-1}, ...]$ by the recursive scheme[12]

$$M_t = a_0 + \sum_{i=1}^{p} a_i X_{t-i} + \sum_{j=1}^{q} b_j M_{t-j}, \tag{1}$$

with the constraints $a_0 > 0$ and $a_1, \dots, a_p, b_1, \dots, b_q \geq 0$ with $a_1 + \cdots + b_q < 1$. Then, the actual count X_t is emitted by a specified count distribution with mean M_t, such as the Poisson distribution $\text{Poi}(M_t)$.[12] Many further emitting distributions have been discussed in the literature, for example, the negative-binomial (NB) or the zero-inflated Poisson (ZIP) distribution.[22,23] Some SPC implementations of INGARCH models are reported in Weiß and Testik,[14] Huh et al.,[15] Rakitzis et al.,[16] Vanli et al.,[17] and Ottenstreuer.[18]

If (X_t) consists of bounded counts, then the corresponding BINGARCH(p, q) model is defined with respect to the normalized conditional mean $P_t = \frac{1}{n} M_t$, where

$$P_t = a_0 + \sum_{i=1}^{p} a_i X_{t-i}/n + \sum_{j=1}^{q} b_j P_{t-j} \tag{2}$$

with the additional constraint $a_0 + \sum_{i=1}^{p} a_i + \sum_{j=1}^{q} b_j < 1$.[13] Here, X_t is generated by a bounded-count distribution like the binomial, $\text{Bin}(n, P_t)$.

Both models (1) and (2) were recently extended to a large class of approximately linear (B)INGARCH models, which cover the exactly linear models (1) and (2) as a limiting case. The original motivation of these extensions was to preserve the linearity as far as possible, but to allow for negative parameter values at the same time (and thus negative ACF values). In the present paper, our intention for using these models is different. They serve to us as a tool for experiencing the effects of deceptive linearity. They provide us a DGP, where the extent of nonlinearity is controlled by a single adjustment parameter. The basic idea behind these models is to modify recursions (1) and (2) by inserting an additional response function, which behaves nearly like a (piecewise) linear function. These are the softplus function[24]

$$\text{sp}_c(x) = c \ln(1 + \exp(x/c)) \quad \text{with adjustment parameter } c > 0 \tag{3}$$

for unbounded counts, and the soft clipping function[25]

$$\text{sc}_c(x) = c \ln\left(\frac{1 + \exp\left(\frac{x}{c}\right)}{1 + \exp\left(\frac{x-1}{c}\right)}\right) \quad \text{with adjustment parameter } c > 0 \tag{4}$$

for bounded counts. The plots in Figure 1 illustrate that $\text{sp}_c(x)$ and $\text{sc}_c(x)$ become piecewise linear for $c \to 0$: $\text{sp}_c(x) \to \max\{0, x\}$ for $c \to 0$, and $\text{sc}_c(x) \to \min\{1, \max\{0, x\}\}$ for $c \to 0$.

TABLE 1 Moment properties for models defined in Example 2.1

	$(\alpha_0, \alpha_1) = (\mathbf{0.85}, \mathbf{0.5})$				$(\alpha_0, \alpha_1) = (\mathbf{2.85}, \mathbf{-0.5})$			
c	**0**	**0.5**	**1**	**2**	**0**	**0.5**	**1**	**2**
μ	1.700	1.762	2.050	2.878	1.907	1.927	2.026	2.410
σ^2/μ	1.333	1.306	1.243	1.180	1.300	1.258	1.190	1.122
$\rho(1)$	0.500	0.484	0.441	0.389	−0.478	−0.450	−0.397	−0.328

Weiß et al.[20] suggest to modify the linear INGARCH model (1) to the spINGARCH model by defining

$$M_t = s_c\left(M_t^*\right) \quad \text{with } M_t^* = \alpha_0 + \sum_{i=1}^{p} \alpha_i X_{t-i} + \sum_{j=1}^{q} \beta_j M_{t-j}. \tag{5}$$

Here, $\alpha_1, \ldots, \alpha_p, \beta_1, \ldots, \beta_q$ might also become negative (which is useful to achieve negative ACF values), as long as $\sum_{i=1}^{p} \max\{0, \alpha_i\} + \sum_{j=1}^{q} \max\{0, \beta_j\} < 1$ and $\sum_{j=1}^{q} |\beta_j| < 1$ are satisfied.[20] The limiting case $c \to 0$ corresponds to

$$M_t = \max\left\{0, M_t^*\right\} \quad \text{with } M_t^* = \alpha_0 + \sum_{i=1}^{p} \alpha_i X_{t-i} + \sum_{j=1}^{q} \beta_j M_{t-j}, \tag{6}$$

which covers model (1) as a special case under the parameter constraints given for Equation (1).

In the case of bounded counts, Weiß and Jahn[21] extend the linear model (2) to the scBINGARCH model by

$$P_t = \mathrm{sc}_c\left(P_t^*\right) \quad \text{with } P_t^* = \alpha_0 + \sum_{i=1}^{p} \alpha_i X_{t-i}/n + \sum_{j=1}^{q} \beta_j P_{t-j}, \tag{7}$$

having the additional constraint $\alpha_0 \in (0; 1 + p + q)$. Now, the limiting case $c \to 0$ corresponds to

$$P_t = \min\left\{1, \max\left\{0, P_t^*\right\}\right\} \quad \text{with } P_t^* = \alpha_0 + \sum_{i=1}^{p} \alpha_i X_{t-i}/n + \sum_{j=1}^{q} \beta_j P_{t-j}, \tag{8}$$

which covers model (2) as a special case. Altogether, c allows to control the extent of violating linearity within the INGARCH class of models for count time series.

Example 2.1. Let us illustrate the differences between the different model types by some numerical examples referring to the mean $\mu = E[X_t]$, the variance $\sigma^2 = V[X_t]$, and the ACF $\rho(h) = Corr[X_t, X_{t-h}]$ with lag $h \in \mathbb{N}$. For the ease of presentation, we focus on the case of unbounded counts, that is, on a comparison of the exactly linear model (1), the piecewise-linear model (6), and the nonlinear softplus model (5), as well as on model order $(p, q) = (1, 0)$ ("INARCH(1)"). Further numerical results (also for bounded counts as well as for model order (1,1)) are presented later in Section 4, when discussing Table 2.

Let us begin with the parametrization $(a_0, a_1) = (\alpha_0, \alpha_1) = (0.85, 0.5)$, which corresponds to (M1a) in Section 3. Then, the exactly linear model (1) is well-defined as (a_0, a_1) satisfy the constraints $a_0 > 0, a_1 \geq 0$, and we calculate the mean μ, dispersion ratio σ^2/μ, and lag-1 ACF $\rho(1)$ as $a_0/(1 - a_1) = 1.7, 1/(1 - a_1^2) \approx 1.333$, and $a_1 = 0.5$, respectively. As α_0, α_1 are non-negative, the piecewise-linear model (6) coincides with Equation (1), thus leading to exactly the same model properties. Also the softplus model (5) is well-defined, and it becomes more and more nonlinear if c increases. As a consequence, the considered moment properties deviate from those of the linear model, see the first block in Table 1.

Next, we consider the parametrization $(a_0, a_1) = (\alpha_0, \alpha_1) = (2.85, -0.5)$, which corresponds to (M1b) in Section 3. This time, a_1 takes a negative value violating the constraint $a_1 \geq 0$, so the exactly linear model (1) does not exist at all. Models (6) and (5), however, are well-defined also for negative α_1. Here, the piecewise-linear model (6) constitutes the closest possible approximation to a truly linear model that can be achieved, whereas an exactly linear model cannot exist for $a_1 < 0$. As can be seen from the second block in Table 1, column "$c = 0$," its moment properties are close to those that one would compute under exact linearity, the latter being $a_0/(1 - a_1) = 1.9$ for the mean, $1/(1 - a_1^2) \approx 1.333$ for the dispersion ratio,

and $a_1 = -0.5$ for the lag-1 ACF. With increasing nonlinearity ($c > 0$), the moment properties of the softplus model (5) more and more deviate from those of a linear model.

To keep it simple, from now on, we shall refer to models (6) and (8) as linear models.

3 | DESCRIPTION OF ROBUSTNESS STUDY

Robustness of the control charts' performance to model misspecification is evaluated here by considering a DGP that is identified as being exactly linear when the true DGP is only *approximately linear*. The performance metric used is the average run length (ARL), that is, the average number of statistics being plotted on a control chart until the first alarm is triggered. The ARL is the most common metric for designing a control chart and for evaluating its performance, see Refs. 26, 27. Possible effects of nonlinearity are quantified with respect to different characteristics, namely stochastic properties of the DGP (X_t) itself (such as the mean μ and the ACF $\rho(h)$), as well as the design and properties of the control charts being applied to the DGP.

Remark 3.1. The true DGP might be nonlinear ($c > 0$) under normal operating conditions. When there is no change either in the true form or the parameters of the DGP, the process is said to be *in control*, otherwise it is called *out of control*. Thus, even if the process model is identified as linear ($c = 0$) when it is actually nonlinear ($c > 0$) and/or there are errors in the estimates for the parameters $\alpha_0, \ldots, \beta_q$ of models (5)–(8), the actual process is in control. However, such deviations of the specified model and/or its parameters from the true DGP may cause the actual performance of a control chart to deviate from its planned performance. This robustness study analyzes the effects of model misspecification and/or errors in parameter estimation on the control charts' ARL performance.

For our robustness study, we always switch between two perspectives (picking up the terminology in Weiß et al.[28]):

- *Planned properties*: This corresponds to the user's perspective, who anticipates a truly linear DGP ($c = 0$, "illusory linearity") in the sense of Equations (6) and (8), which is then used for computing the DGP's stochastic properties, or for control chart design and ARL calculations.
- *Actual properties*: This corresponds to the true DGP being nonlinear (in our case, following Equation (5) or (7) with some $c > 0$). Basing computations on this true DGP, we get insight into the process' actual mean and ACF, or the designed chart's actual ARL.

In Section 4, we start our analyses by considering the case of specified parameters, that is, where the values of the model parameters $\alpha_0, \ldots, \beta_q$ for M_t^*, P_t^* in Equations (5)–(8) are given (but the adjustment parameter c varies, thus causing model uncertainty). Later in Section 5, we also study the additional effect of parameter estimation, thus the combined effect of model and estimation uncertainty. An upper-sided Shewhart individuals chart (referred to as *c-chart* for unbounded counts and *np-chart* for bounded counts) and an upper-sided *CUSUM chart* (without headstart feature) are considered since these are most relevant for low-counts processes, see Refs. 4, 8, 27. Because of the discreteness of the count processes, it is usually not possible to find a chart design such that a given target in-control ARL (the common textbook choice is $\mathrm{ARL}_0 = 370.4$) is met exactly.[29,30] Therefore, in what follows, we select the parameterization of the example DGPs in such a way that the respective Shewhart chart has a planned in-control ARL being close to the aforementioned choice of $\mathrm{ARL}_0 = 370.4$.

The following models from Equations (5)–(8) are considered within our study:

(M1) Poisson spINARCH(1) model with $M_t^* = \alpha_0 + \alpha_1 X_{t-1}$, as a representative case for AR(1)-like unbounded counts. Anticipating linearity, the mean and ACF should satisfy $\mu = \alpha_0/(1-\alpha_1)$ and $\rho(h) = \alpha_1^h$, respectively. The parametrizations are:
  (M1a) $(\alpha_0, \alpha_1) = (0.85, 0.5)$, that is, positive $\rho(1)$ and low $\mu = 1.7$;
  (M1b) $(\alpha_0, \alpha_1) = (2.85, -0.5)$, that is, negative $\rho(1)$ and low $\mu = 1.9$;
  (M1c) $(\alpha_0, \alpha_1) = (1.85, 0.5)$, that is, positive $\rho(1)$ and medium $\mu = 3.7$;
  (M1d) $(\alpha_0, \alpha_1) = (5.85, -0.5)$, that is, negative $\rho(1)$ and medium $\mu = 3.9$.

(M2) Binomial scBINARCH(1) model with $P_t^* = \alpha_0 + \alpha_1 X_{t-1}/n$, as a representative case for AR(1)-like bounded counts. Anticipating linearity, the mean and ACF should satisfy $\mu = n\,\alpha_0/(1-\alpha_1)$ and $\rho(h) = \alpha_1^h$, respectively. The parametrizations are:
  (M2a) $(n, \alpha_0, \alpha_1) = (21, 0.2, 0.5)$, that is, positive $\rho(1)$ and $\mu/n = 0.4$ (rather symmetric marginal distribution);
  (M2b) $(n, \alpha_0, \alpha_1) = (21, 0.6, -0.5)$, that is, negative $\rho(1)$ and $\mu/n = 0.4$ (rather symmetric marginal distribution);
  (M2c) $(n, \alpha_0, \alpha_1) = (20, 0.1, 0.5)$, that is, positive $\rho(1)$ and $\mu/n = 0.2$ (very asymmetric marginal distribution);
  (M2d) $(n, \alpha_0, \alpha_1) = (20, 0.309, -0.5)$, that is, negative $\rho(1)$ and $\mu/n = 0.206$ (very asymmetric marginal distribution).

(M3) Poisson spINGARCH(1,1) model with $M_t^* = \alpha_0 + \alpha_1 X_{t-1} + \beta_1 M_{t-1}$, as a representative case for counts with slowly decaying ACF. Anticipating linearity, the mean and ACF should satisfy $\mu = \alpha_0/(1-\alpha_1-\beta_1)$ and $\rho(h) = (\alpha_1 + \beta_1)^{h-1}\,\alpha_1\,(1-\beta_1(\alpha_1+\beta_1))/(1-(\alpha_1+\beta_1)^2+\alpha_1^2)$, respectively. The parametrizations are:
  (M3a) $(\alpha_0, \alpha_1, \beta_1) = (0.9, 0.25, 0.5)$, that is, positive ACF and $\mu = 3.6$;
  (M3b) $(\alpha_0, \alpha_1, \beta_1) = (2.835, -0.25, 0.5)$, that is, negative ACF and $\mu = 3.78$;
  (M3c) $(\alpha_0, \alpha_1, \beta_1) = (4.725, 0.25, -0.5)$, that is, alternating ACF and $\mu = 3.78$;
  (M3d) $(\alpha_0, \alpha_1, \beta_1) = (6.335, -0.25, -0.5)$, that is, alternating ACF and $\mu = 3.62$.

Note that the INARCH(1)-type DGPs (M1) and (M2) are CLAR(1) models (conditionally linear AR(1)) in the sense of Grunwald.[3] In particular, they constitute a Markov chain (MC) such that for the aforementioned Shewhart and CUSUM charts, an exact ARL computation is possible by the MC-approach.[31,32] Also relevant stochastic properties of the DGP are computed exactly for (M1) and (M2) by utilizing the Markov property. For DGP (M3), by contrast, stochastic properties as well as ARLs can usually only be approximated based on simulations, which, in turn, implies a high computational effort for chart design. For this reason, we did not consider further non-Markovian DGPs beyond (M3), and we limit our analysis of (M3) to the case of specified parameters. Finally, our characterizations of models (M2a–b)[1] as having a "rather symmetric marginal distribution" is motivated by the fact that $\mu/n = 0.4$ is close to 0.5, where $\text{Bin}(n, 0.5)$ is an exactly symmetric distribution (zero skewness). Analogously, models (M2c–d) having $\mu/n = 0.2$ are notably skewed to the right, thus having a "very asymmetric marginal distribution." The latter characterization also extends to models (M1a–d).

4 | EFFECTS OF NONLINEARITY FOR SPECIFIED PARAMETERS

As the first step of our robustness analyses described in Section 3, we investigate the effects of nonlinearity under the premise of specified model parameters $\alpha_0, \ldots, \beta_q$ for M_t^*, P_t^* (thus pure model uncertainty) in Equations (5)–(8). That is, the parametrizations (M1)–(M3) are assumed to be known, and we compare the actual properties (for $c > 0$) with the planned ones ($c = 0$). We start these comparisons by considering the most relevant moment properties, namely the ACF at lag $h = 1$, $\rho(1)$, and either the ordinary mean μ for unbounded counts, or the normalized mean μ/n for bounded counts. These are computed by either anticipating exact linearity (planned properties), or for the actual DGP with given c, see Table 2 for a summary (also recall Example 2.1). Note that (M1)–(M2) allow for an exact computation of these properties, whereas the tabulated values for (M3) are sample values for simulated time series of length 10^7.

A rather unique pattern can be recognized from Table 2. With increasing extent of nonlinearity (i. e., with increasing values of c), we recognize increasing values of the mean, whereas the ACF decreases in absolute extent. Thus, the generated counts tend to be larger than expected while the serial dependence structure is reduced. For the Markovian DGPs, the described patterns are more pronounced for the highly asymmetric[2] models (M1a–d), (M2c–d) than for the nearly symmetric ones (M2a–b), and more for the low-counts cases (M1a–b), (M2c–d). This appears plausible in view of the shape of the functions $\text{sp}_c(x)$ and $\text{sc}_c(x)$ in Figure 1. Note that for negative dependence parameters (models "b" and "d"), there might be a slight difference between planned and actual properties even for $c = 0$, because the model is only piecewise linear but not strictly linear in that case. An analogous behavior is observed for the non-Markovian DGPs (M3a–d), where the changes are most pronounced for (M3a) with purely positive, and least pronounced for (M3d) with purely negative dependence parameters.

Next, let us analyze the effect of nonlinearity on the in-control ARLs of the Shewhart and CUSUM charts, see Table 3. Recall that models (6) and (8) constitute the closest approximation to linearity, thus the case "$c = 0$" is interpreted as the

[1] Here, "(M2k–l)" is a short-hand notation for "(M2k), ..., (M2l)".
[2] Recall our characterizations of (a)symmetry in Section 3.

TABLE 2 Moment properties for models defined in Section 3

Poisson spINARCH(1)	**(M1a)**		**(M1b)**		**(M1c)**		**(M1d)**	
	μ	$\rho(1)$	μ	$\rho(1)$	μ	$\rho(1)$	μ	$\rho(1)$
Planned:	1.7	0.5	1.9	−0.5	3.7	0.5	3.9	−0.5
$c = 0.0$	1.700	0.500	1.907	−0.478	3.700	0.500	3.901	−0.498
0.5	1.762	0.484	1.927	−0.450	3.703	0.499	3.903	−0.494
1.0	2.050	0.441	2.026	−0.397	3.775	0.487	3.927	−0.475
2.0	2.878	0.389	2.410	−0.328	4.276	0.444	4.112	−0.418
Binomial scBINARCH(1)	**(M2a)**		**(M2b)**		**(M2c)**		**(M2d)**	
	μ/n	$\rho(1)$	μ/n	$\rho(1)$	μ/n	$\rho(1)$	μ/n	$\rho(1)$
Planned:	0.4	0.5	0.4	−0.5	0.2	0.5	0.206	−0.5
$c = 0.0$	0.400	0.500	0.400	−0.500	0.200	0.500	0.206	−0.500
0.5	0.400	0.500	0.400	−0.500	0.200	0.500	0.206	−0.498
1.0	0.400	0.500	0.400	−0.500	0.203	0.490	0.207	−0.484
2.0	0.403	0.490	0.401	−0.490	0.225	0.447	0.215	−0.431
Poisson spINGARCH(1,1)	**(M3a)**		**(M3b)**		**(M3c)**		**(M3d)**	
	μ	$\rho(1)$	μ	$\rho(1)$	μ	$\rho(1)$	μ	$\rho(1)$
Planned:	3.60	0.313	3.78	−0.219	3.78	0.219	3.62	−0.313
$c = 0.0$	3.601	0.312	3.780	−0.219	3.780	0.219	3.620	−0.313
0.5	3.604	0.312	3.781	−0.218	3.780	0.219	3.621	−0.310
1.0	3.720	0.299	3.814	−0.214	3.800	0.215	3.639	−0.292
2.0	4.530	0.263	4.150	−0.196	4.018	0.197	3.813	−0.238

TABLE 3 In-control ARLs for Shewhart charts with upper limit UCL, and for CUSUM charts with design (h, k), for DGP models defined in Section 3

Shewhart	**(M1a)**	**(M1b)**	**(M1c)**	**(M1d)**	**(M2a)**	**(M2b)**	**(M2c)**	**(M2d)**
UCL	**7**	**7**	**11**	**11**	**15**	**15**	**10**	**10**
$c = 0.0$	375.1	363.8	369.9	367.8	342.2	342.1	385.3	385.9
0.5	352.1	374.3	369.3	369.8	342.2	342.1	385.0	386.6
1.0	238.3	346.9	351.0	376.8	342.1	342.2	370.1	391.8
2.0	71.0	175.7	216.2	338.5	331.4	347.2	233.0	357.3
CUSUM	**(M1a)**	**(M1b)**	**(M1c)**	**(M1d)**	**(M2a)**	**(M2b)**	**(M2c)**	**(M2d)**
(h, k)	**(11,3)**	**(11,2)**	**(20,5)**	**(17,4)**	**(16,10)**	**(8,9)**	**(17,5)**	**(5,5)**
$c = 0.0$	375.9	390.8	369.7	377.7	276.8	259.5	260.0	255.1
0.5	350.5	289.7	369.0	365.4	276.8	259.5	259.7	254.0
1.0	220.3	112.5	344.5	286.7	276.6	258.9	245.4	239.3
2.0	48.2	26.4	169.8	90.4	265.5	231.4	122.6	135.8
	Shewhart				**CUSUM**			
	(M3a)	**(M3b)**	**(M3c)**	**(M3d)**	**(M3a)**	**(M3b)**	**(M3c)**	**(M3d)**
UCL, (h, k)	**10**	**10**	**10**	**10**	**(15,5)**	**(13,4)**	**(10,5)**	**(12,4)**
$c = 0.0$	380.5	380.9	375.1	380.2	407.9	451.1	390.2	372.7
0.5	378.8	381.3	375.2	382.3	406.7	446.3	390.1	370.5
1.0	328.2	362.9	366.8	395.0	345.4	316.4	374.3	326.8
2.0	113.3	213.2	267.4	349.7	90.4	52.0	221.0	132.3

planned properties. Accordingly, we first tried to find a chart design for $c = 0$ (i. e., the upper control limit UCL for a Shewhart chart, and upper limit h as well as reference value k for a CUSUM chart) leading to an in-control ARL close to $\text{ARL}_0 = 370.4$. This, however, was not possible for the CUSUM chart applied to some (M2)-scenarios. Thus, in the latter case, we used a different but unique target level, namely around 270. Analogously, the CUSUM ARLs for (M3a–b) are somewhat larger than ARL_0. For the CUSUM charts, the reference value k was chosen as $\lceil \mu + 1 \rceil$ for positive α_1, and $\lceil \mu \rceil$

for negative α_1, where the ceiling $\lceil\cdot\rceil$ rounds up to the next integer value. Recall that the ARLs for (M1)–(M2) are computed exactly by the MC-approach, whereas they rely on simulations for (M3) and are, thus, shown in a separate block of Table 3. To compute the ARLs for (M3), we simulated 10^6 run lengths per scenario and took the sample mean thereof.

For the obtained ARLs in Table 3, we recognize a similar pattern like for the mean μ in Table 2: with increasing c, the ARLs tend to decrease. For the Markovian DGPs (M1)–(M2), the decrease is most pronounced for (M1a–b) and (M2c–d), and least pronounced for (M2a–b). Analogously, for (M3), the Shewhart's ARL decreases most strongly for (M3a), and least strongly for (M3d). For the (M3) CUSUM charts, in turn, also the negative value for β_1 in (M3b,d) intensifies the decline in ARL. The connection between increasing μ and decreasing ARL is plausible, as increased counts cause a more frequent violation of the upper control limit. However, we do not have perfect monotonicity in all cases: the Shewhart charts under negative dependence, that is, (M1b,d), (M2b,d), and (M3b–d), first have slightly increasing ARLs. This is caused by some trade-off between the effects on mean and ACF. From Table 2, we know that increasing c leads to increasing $\rho(1)$ (a reduction in absolute extent implies an increase of a negative ACF value). Increased ACF values, in turn, are well known to cause an increase of the ARL of a Shewhart individuals chart. Note that the effect of c is generally stronger for the CUSUM than for the Shewhart charts, which follows from the CUSUM's inherent accumulation of counts and its sensitivity to already slight violations of model assumptions.

5 | EFFECTS OF NONLINEARITY UNDER PARAMETER ESTIMATION

In real-world implementations, model parameters are typically not specified but estimated from given in-control data. Thus, even if the correct model for the DGP is selected, the parameter values of the fitted model usually differ from the ones of the true model because of estimation error. If the chart design is done based on the fitted model, the actual ARL performance (computed using the developed design but for the true DGP) also deviates from the planned one. Such problems have been investigated by several authors, see Jensen et al.[10] and Psarakis et al.[11] for general surveys, and Weiß and Testik,[33] Zhang et al.,[34] and Weiß et al.[28] for studies on the particular case of autocorrelated count processes. Here, we go one step further by considering the combined effect of both estimation uncertainty and model uncertainty: the true DGP might be nonlinear according to Equation (5) or (7) with $c > 0$, but the linear models (6) and (8) ($c = 0$) are fitted to the generated data. Note that this combination of model misspecification and model fitting does not necessarily imply that results are getting worse. Instead, there might also be some trade-off effects, in the sense that the parameter values of the fitted linear model adapt to the nonlinearity in the data.

Let us first look at the effect of both model misspecification and model fitting on the parameter estimates. For the Markov models (M1)–(M2) considered in Sections 3 and 4 (recall that we omit (M3) here because of the heavy computational requirements), time series of different lengths T were generated (1000 replications per scenario) and the respective linear model ($c = 0$) was fitted to the data. In view of computational efficiency, we used moment estimates for this purpose, that is, parameters are estimated as follows:

(M1) Poisson spINARCH(1) model: $\hat{\alpha}_1 = \hat{\rho}(1)$ and $\hat{\alpha}_0 = \overline{x}\,(1 - \hat{\alpha}_1)$, where $\overline{x}$ and $\hat{\rho}(1)$ denote the sample mean and ACF, respectively.
(M2) Binomial scBINARCH(1) model: $\hat{\alpha}_1 = \hat{\rho}(1)$ and $\hat{\alpha}_0 = \frac{1}{n}\overline{x}\,(1 - \hat{\alpha}_1)$.

Comparing the estimates with and without model misspecification, it turned out that their standard errors are hardly affected by model misspecification, while there might be notable effects on their bias. Therefore, let us focus our discussion on the sample means computed from the 1000 replicates per scenario. In Table 2, we analyzed the true mean and lag-1 ACF of the misspecified models. While there was an at most minor effect on the bias of the fitted models' mean, the fitted models' ACF was visibly affected. Illustrative results for sample size $T = 100$ (which is close to the sample size of the first data example discussed in Section 6.1) are shown in Table 4. There, the values labeled "spec." agree with the $\rho(1)$-values of Table 2, while those labeled "simul." are computed as the sample means across the $\rho(1)$-values of the 1000 simulated model fits. In Table 4, the effect of worsening misspecification can be seen along the rows, whereas the difference of "simul." to "spec." corresponds to the additional bias caused by parameter estimation. It becomes clear that the combination of both model fitting and misspecification leads to a deterioration of the $\rho(1)$-results, which is more pronounced for the case of positive than for negative autocorrelation. It should be noted, however, that the bias gets reduced for both with and without model misspecification if the sample size increases. For $T = 250$, for example (similar to the second data example

TABLE 4 Lag-1 ACF $\rho(1)$ for models defined in Section 3, computed either exactly for specified parameters, or as mean ACF of linear model fits to 1000 simulated time series of length $T = 100$

Poisson spINARCH(1)	**(M1a)**		**(M1b)**		**(M1c)**		**(M1d)**	
	spec.	**simul.**	**spec.**	**simul.**	**spec.**	**simul.**	**spec.**	**simul.**
Planned:	0.5		−0.5		0.5		−0.5	
$c = 0.0$	0.500	0.459	−0.478	−0.470	0.500	0.465	−0.498	−0.488
0.5	0.484	0.445	−0.450	−0.444	0.499	0.466	−0.494	−0.484
1.0	0.441	0.408	−0.397	−0.395	0.487	0.452	−0.475	−0.467
2.0	0.389	0.355	−0.328	−0.326	0.444	0.412	−0.418	−0.416
Binomial scBINARCH(1)	**(M2a)**		**(M2b)**		**(M2c)**		**(M2d)**	
	spec.	**simul.**	**spec.**	**simul.**	**spec.**	**simul.**	**spec.**	**simul.**
Planned:	0.5		−0.5		0.5		−0.5	
$c = 0.0$	0.500	0.472	−0.500	−0.493	0.500	0.468	−0.500	−0.489
0.5	0.500	0.474	−0.500	−0.489	0.500	0.470	−0.498	−0.489
1.0	0.500	0.466	−0.500	−0.491	0.490	0.460	−0.484	−0.479
2.0	0.490	0.459	−0.490	−0.486	0.447	0.413	−0.431	−0.430

of Section 6.2, but not shown here to save some space), there is an only mild bias for positive autocorrelation, and virtually no bias anymore for negative autocorrelation.

Let us now go one step further. Based on the above model fits per replication, control charts were designed according to the same rules as described in Section 4. In particular, the planned ARL (computed by the MC-approach for each simulated model fit) is chosen to be as close as possible to the target $ARL_0 = 370.4$. Then, for each simulation run, the actual ARL was computed (again by the MC-approach), that is, we have a sample of 1000 simulated ARL values per scenario (referring to the 1000 model fits). However, because of the discreteness implied by the integer design parameters UCL and h, k, respectively, these samples consist of only a few different (but repeatedly observed) ARL values. As an example, for model (M1a) with $c = 0$ and $T = 100$, we get the following actual ARLs of the c-chart:

Actual ARL	28.1	64.0	151.7	375.1	964.4	2563.3	7004.8
Frequency	1	54	352	402	146	38	7

Hence, there was once a model fit leading to the actual ARL 28.1, 54 times a model fit leading to the actual ARL 64.0, and so forth. In most cases (sample mode), we end up with a chart design leading to ARL $\approx$ 375.1, which corresponds to the first chart design in Table 3. Thus, because of discreteness, the final chart design is not affected at all by estimation error in most cases. However, in 407 cases, we end up with a chart design leading to a smaller ARL value, and in 191 cases with a larger ARL value. Since the gaps between the different ARL values are rather large, sample moments and quantiles are very sensitive to even modest changes in the frequency distribution and might, thus, be rather misleading. Therefore, we decided to analyze the simulated actual ARLs by computing the sample mode as well as the frequencies for falling below or exceeding the mode.

The obtained results for the Shewhart charts are summarized in Table 5, those for the CUSUM charts in Table 6. Let us first compare the Shewhart results in Table 5 to the first block in Table 3, that is, to the actual ARLs for specified parameters. If $c = 0$ (no model misspecification), the mode of the actual ARLs always agrees with the ARL for specified parameters. Thus, in agreement with the findings of Weiß et al.[28, p. 5], "the estimation uncertainty is completely eliminated in chart design" in the majority of simulation runs. And this happens more often for $T = 250$ than for $T = 100$, for bounded counts than for unbounded ones, and for negative than for positive ACF. Next, we consider additional misspecification ($c > 0$). While in the first block of Table 3 for specified parameters, this usually causes a decline of the actual in-control ARL, this does usually not happen for the modes under parameter estimation. So there is a (welcome!) trade-off between model fitting and model misspecification. Certainly, there are effects on the ARL performance, which also become visible in changes of the "above or below" frequencies (especially for $c = 2$). But the mode ARL is typically still close to the target ARL_0. In addition, the frequency of obtaining this mode value still increases with the sample size T. For a user, it might be relevant to have a mode frequency of, for example, more than 50 %, that is, in more than half of all cases, an appropriate

TABLE 5 Actual ARLs of Shewhart charts for fitted models (sample size T) with DGPs of Section 3: modes from 1000 replications, and frequencies of ARLs below or above mode

Shewhart		$T = 100$			$T = 250$				$T = 100$			$T = 250$		
c		Mode	Below	Above	Mode	Below	Above		Mode	Below	Above	Mode	Below	Above
0.0	**(M1a)**	375.1	407	191	375.1	320	134	**(M2a)**	342.2	209	153	342.2	85	62
0.5		352.1	385	207	352.1	268	171		342.2	198	168	342.2	64	75
1.0		238.3	158	354	238.3	36	338		342.1	212	164	342.1	83	83
2.0		413.7	432	131	413.7	311	53		331.4	184	176	331.4	60	90
0.0	**(M1b)**	363.8	91	35	363.8	21	3	**(M2b)**	342.1	50	109	342.1	0	33
0.5		374.3	98	38	374.3	13	2		342.1	54	113	342.1	2	32
1.0		346.9	47	59	346.9	4	5		342.2	42	104	342.2	2	24
2.0		570.5	407	4	570.5	308	0		347.2	47	123	347.2	4	21
0.0	**(M1c)**	369.9	420	228	369.9	320	165	**(M2c)**	385.3	339	131	385.3	186	61
0.5		369.3	428	226	369.3	297	191		385.0	309	138	385.0	196	67
1.0		351.0	384	229	351.0	274	205		370.1	284	153	370.1	184	56
2.0		216.2	163	440	462.9	453	66		233.0	85	330	233.0	6	339
0.0	**(M1d)**	367.8	237	102	367.8	135	32	**(M2d)**	385.9	94	43	385.9	8	13
0.5		369.8	242	95	369.8	116	21		386.6	78	59	386.6	11	9
1.0		376.8	254	87	376.8	123	17		391.8	60	43	391.8	12	7
2.0		338.5	169	125	338.5	68	29		357.3	28	76	357.3	4	11

TABLE 6 Actual ARLs of CUSUM charts for fitted models (sample size T) with DGPs of Section 3: modes from 1000 replications, and frequencies of ARLs below or above mode

CUSUM		$T = 100$			$T = 250$				$T = 100$			$T = 250$		
c		Mode	Below	Above	Mode	Below	Above		Mode	Below	Above	Mode	Below	Above
0.0	**(M1a)**	154.2	221	687	303.9	403	465	**(M2a)**	166.5	263	674	276.8	342	556
0.5		144.2	209	689	228.2	247	606		197.6	316	621	326.7	419	483
1.0		341.7	552	369	341.7	491	374		197.5	334	599	326.5	424	478
2.0		138.3	192	720	260.0	295	587		265.5	430	501	265.5	295	606
0.0	**(M1b)**	296.2	329	554	296.2	258	583	**(M2b)**	259.5	80	511	259.5	12	540
0.5		349.7	446	401	226.6	135	727		259.5	52	520	259.5	14	530
1.0		297.2	403	143	297.2	336	34		258.9	62	532	258.9	19	522
2.0		416.4	394	212	416.4	333	97		231.4	61	589	412.4	325	265
0.0	**(M1c)**	122.9	164	779	248.8	303	613	**(M2c)**	142.7	216	732	454.1	598	310
0.5		122.7	146	795	283.9	355	561		453.7	636	305	578.8	717	211
1.0		265.8	451	492	265.8	341	586		264.4	430	509	434.3	636	277
2.0		284.3	498	442	284.3	347	555		195.6	311	600	379.5	506	368
0.0	**(M1d)**	743.3	719	111	222.6	155	738	**(M2d)**	255.1	179	300	255.1	49	188
0.5		732.6	716	125	260.0	234	664		254.0	159	333	254.0	54	199
1.0		694.6	711	77	694.6	826	73		239.3	131	367	239.3	43	227
2.0		395.7	344	175	395.7	266	68		322.8	175	283	322.8	85	180

chart design is selected despite the model and estimation uncertainty. Looking at Table 5 from such a user's perspective, we conclude that sample sizes $T \geq 250$ are recommended for model fitting.

Our conclusions differ if comparing the CUSUM ARLs in Table 6 to the second block in Table 3. Except for bounded counts with negative ACF, the mode ARL usually deviates considerably from the ARL with specified parameters (although it does not necessarily get worse than in the specified case). In addition, there are much larger frequencies for falling above or below the mode. This additional variability can be explained by the additional flexibility in chart design (two design

TABLE 7 Actual ARLs of Shewhart charts for fitted models (sample size T), where true DGPs correspond to (M1a–b) from Section 3 with additional model-order misspecification (if $\beta_1 \neq 0$): modes from 1000 replications, and frequencies of ARLs below or above mode

(M1a)								(M1b)						
Shewhart		$T = 100$			$T = 250$				$T = 100$			$T = 250$		
c	β_1	Mode	Below	Above	Mode	Below	Above	β_1	Mode	Below	Above	Mode	Below	Above
0.0	0.0	375.1	407	191	375.1	320	134	0.0	363.8	91	35	363.8	21	3
	0.1	376.7	400	192	376.7	324	129	0.1	363.4	81	38	363.4	12	1
	−0.1	369.2	394	187	369.2	279	120	−0.1	355.6	96	62	355.6	16	2
0.5	0.0	352.1	385	207	352.1	268	171	0.0	374.3	98	38	374.3	13	2
	0.1	349.7	387	242	349.7	255	147	0.1	367.7	70	43	367.7	18	0
	−0.1	349.7	380	227	349.7	257	113	−0.1	373.4	85	48	373.4	30	1
1.0	0.0	238.3	158	354	238.3	36	338	0.0	346.9	47	59	346.9	4	5
	0.1	224.0	136	390	224.0	37	388	0.1	331.8	43	65	331.8	2	7
	−0.1	247.3	159	319	247.3	49	290	−0.1	357.9	86	50	357.9	12	9
2.0	0.0	413.7	432	131	413.7	311	53	0.0	570.5	407	4	570.5	308	0
	0.1	340.0	338	204	340.0	210	130	0.1	519.6	304	6	519.6	235	0
	−0.1	192.0	64	476	192.0	6	501	−0.1	620.7	461	5	620.7	470	0

parameters h, k instead of a single one), that is, the strong (and useful) effect of discreteness for the Shewhart charts is fading for the CUSUM charts.

In a final simulation experiment, we investigated the situation where model misspecification is not solely caused by illusory linearity, but by an additional type of erroneous model assumption, namely a misspecified model order (another type of additional misspecification, an erroneous distributional assumption, is considered later in Section 6.2). As the starting point, we selected the Poisson spINARCH(1) models (M1a–b) from Section 3 and allowed the true DGP to include an additional feedback term (like in models (M3)). The possible values of β_1 are ± 0.1 (whereas $\beta_1 = 0$ corresponds to an (M1)-type DGP), and the parameters (α_0, α_1) are chosen such that the planned mean and lag-1 ACF remain constant, that is, as 1.7 and 0.5 for (M1a), and 1.9 and −0.5 for (M1b). Then, like before, a linear Poisson spINARCH(1) model was fitted to the simulated time series, and the fitted model was used for chart design. Table 7 summarizes the results obtained for the Shewhart chart (where the values for $\beta_1 = 0$ are taken from Table 5). It can be observed that for low extent of nonlinearity (such as $c \leq 0.5$), the Shewhart's actual ARL performance is quite robust against the erroneous model order, but the effect becomes more and more pronounced for increasing c. For $c = 2$ and model (M1a) with positive ACF, both deviations $\beta_1 = \pm 0.1$ lead to a reduced mode ARL, with a stronger effect for $\beta_1 = -0.1$ (corresponding to a "shorter memory" than planned). For $c = 2$ and model (M1b) with negative ACF, by contrast, only $\beta_1 = 0.1$ (again going along with a shorter memory) causes a reduced mode ARL, whereas $\beta_1 = -0.1$ leads to an increased value thereof.

Next, let us turn to the corresponding CUSUM ARLs in Table 8. Recalling the discussion of Table 6, we already noted that the effect of disregarded nonlinearity on the CUSUM's actual ARLs is clearly stronger than in the Shewhart case (both regarding the mode ARLs and the frequencies for falling above or below the mode). The results in Table 8 show that this effect is further intensified if there is an additional misspecification of the model order. However, this additional effect is relatively moderate especially for model (M1b) with negative ACF and if $c \leq 1$. Together with the effects of parameter estimation and disregarded nonlinearity, we generally observe a high variability in the actual ARLs (much higher than for the Shewhart charts), as expressed by the frequencies for deviating from the mode. Thus, for applications of the CUSUM chart in practice, design adjustments might be advisable, a topic being further discussed in Section 6.1.

6 | DATA APPLICATIONS

6.1 | Monitoring strike counts

As the first real-world data application, we consider an economic time series, $x_1, \ldots, x_T$ where $T = 108$, of monthly strike counts (work stoppages of ≥ 1000 workers in the United States) for the period 1994–2002, see Figure 2 for a plot of the data and their sample ACF. These strike counts were often discussed in the literature, among others by Weiß and Testik[14]

TABLE 8 Actual ARLs of CUSUM charts for fitted models (sample size T), where true DGPs correspond to (M1a–b) from Section 3 with additional model-order misspecification (if $\beta_1 \neq 0$): modes from 1000 replications, and frequencies of ARLs below or above mode

(M1a)								(M1b)						
CUSUM		$T = 100$			$T = 250$				$T = 100$			$T = 250$		
c	β_1	Mode	Below	Above	Mode	Below	Above	β_1	Mode	Below	Above	Mode	Below	Above
0.0	0.0	154.2	221	687	303.9	403	465	0.0	296.2	329	554	296.2	258	583
	0.1	148.5	225	685	228.2	281	594	0.1	249.1	193	673	338.1	242	580
	−0.1	203.2	286	626	258.5	248	598	−0.1	262.9	331	541	262.9	261	590
0.5	0.0	144.2	209	689	228.2	247	606	0.0	349.7	446	401	226.6	135	727
	0.1	170.8	310	602	257.7	368	504	0.1	348.9	415	420	247.6	124	736
	−0.1	115.5	119	783	307.6	353	505	−0.1	342.6	509	312	208.6	155	713
1.0	0.0	341.7	552	369	341.7	491	374	0.0	297.2	403	143	297.2	336	34
	0.1	393.6	623	289	301.0	455	409	0.1	292.5	365	152	292.5	281	42
	−0.1	377.1	542	374	377.1	508	363	−0.1	297.0	432	130	297.0	381	38
2.0	0.0	138.3	192	720	260.0	295	587	0.0	416.4	394	212	416.4	333	97
	0.1	399.6	636	293	158.1	119	795	0.1	409.0	315	262	409.0	273	122
	−0.1	216.4	329	575	216.4	195	678	−0.1	190.2	22	553	421.5	451	65

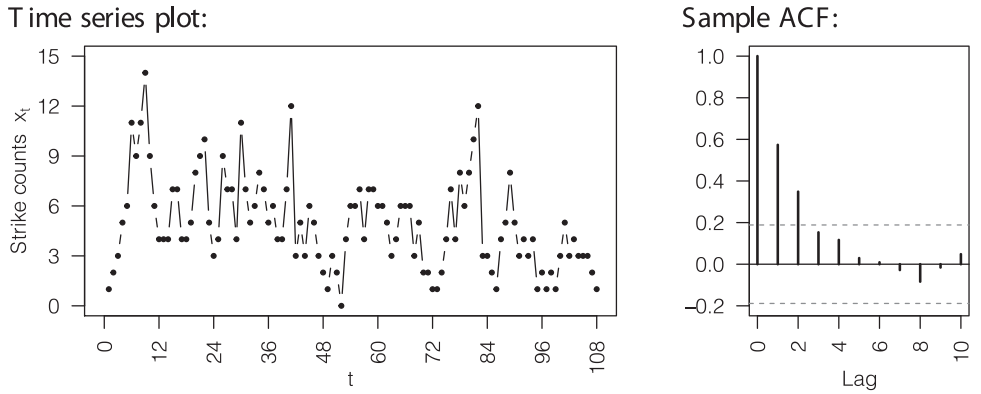

FIGURE 2 Plots of strike-counts time series and their sample ACF, see Section 6.1

in an SPC context. In Weiß[4, Example 4.1.8], it was shown that the ordinary (linear) Poisson INARCH(1) model (i. e., model (1) with $p = 1$ and $q = 0$) provides an excellent fit to the data, going along with the apparent AR(1)-like autocorrelation structure in Figure 2. Thus, a practitioner being interested in monitoring the further development of this economic process would proceed as discussed in Section 5. First, as for any "(M1)-type model," the sample mean $\overline{x} \approx 4.944$ and ACF $\hat{\rho}(1) \approx 0.573$ are used to compute the moment estimates $(\hat{\alpha}_0, \hat{\alpha}_1) \approx (2.109, 0.573)$. Then, the fitted model is used for chart design, leading to the c-chart with UCL = 14 or to the CUSUM chart with $(h, k) = (36, 6)$, where the planned in-control ARL performances are ARL ≈ 404.3 and ARL ≈ 366.8, respectively. Note that all counts in Figure 2 are ≤ UCL such that a retrospective application of the c-chart does not lead to an alarm. This time, certainly, we do not know the true model behind the DGP, as it was the case in our simulation study in Section 5. However, an analogous sensitivity analysis regarding the possible effects of parameter estimation and model misspecification might be done as follows.

The strike counts were also analyzed by Weiß et al.[20] Although these authors confirmed the adequacy of the linear Poisson INARCH(1) model, they also experimented with (mildly) nonlinear spINARCH(1) models and showed that these alternative models perform similarly well. This serves as an inspiration for us to study the possible actual properties by using a different estimation method, namely maximum likelihood (ML) estimation, and by fitting Poisson spINARCH(1) models with several values of the adjustment parameter c. The results are summarized in Table 9. In addition to the respective parameter estimates, also the maximal value of the (conditional) log-likelihood function ℓ is shown, as this is commonly used as the base for model selection (for so-called "information criteria," see Burnham and Anderson[35]). It can be seen that the estimates clearly vary across the different candidate models (including a notable deviation between the moment and ML estimates for $c = 0$), whereas the maximized ℓ-value remains rather stable. In fact, considering the

TABLE 9 Strike counts data: parameter estimates with maximized log-likelihood ("max. ℓ"), properties of fitted models, and in-control ARLs of Shewhart chart with UCL = 14 or CUSUM chart with $(h, k) = (36, 6)$ for respective model fit

Poisson spINARCH(1)	Parameter estimation			Model properties			ARLs	
	$\hat{\alpha}_0$	$\hat{\alpha}_1$	max. ℓ	μ	σ^2/μ	$\rho(1)$	Shewh.	CUSUM
Moment fit:	2.109	0.573	–	4.944	1.490	0.573	404.3	366.8
ML, $c = 0.0$	1.811	0.636	−230.1	4.981	1.681	0.636	263.6	218.2
0.5	1.804	0.638	−230.1	4.981	1.684	0.637	261.8	216.8
1.0	1.728	0.650	−230.2	4.992	1.701	0.642	249.2	206.3
2.0	1.244	0.703	−230.6	5.027	1.721	0.647	226.5	190.3

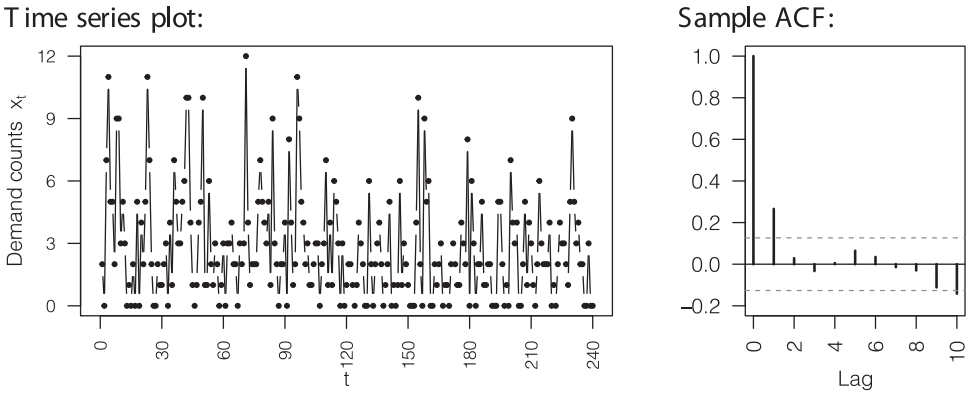

FIGURE 3 Plots of demand-counts time series and their sample ACF, see Section 6.2

common rules-of-thumb for interpreting information criteria[35, Section 2.6], we have to treat any of the fitted models as a viable candidate.

Let us now compare their properties (and those implied for the aforementioned control charts). Comparing the values for the mean μ, dispersion ratio σ^2/μ, and ACF $\rho(1)$ (the corresponding sample values are 4.944, 1.587, and 0.573, respectively), we note larger values for the ML-fits than for the moment fit, and further increases among the ML-fits with increasing c. For the control charts, in turn, the opposite behavior is observed, that is, decreasing ARLs from row to row in Table 9. This is plausible as the charts were designed based on the moment fit, but the ML-fits show more and more deviation for increasing c. All in all, we recognize that the charts' performances are rather sensitive to the precise type of model and estimation approach. This is due to the relatively small sample size of $T = 108$ (which also does not satisfy our sample-size recommendation derived from Table 5) and goes along with our findings in Section 5 for sample size $T = 100$. Thus, the chart designs and ARL values should be taken with caution, and it is recommended to revise them from time to time once further in-control observations from the monitored process became available. In addition, one may also think of further design adjustments derived from some kind of "worst-case considerations" in the following way: the chart design is revised such that the target value is approximately met (or exceeded) for most or all of the candidate models. For example, if increasing UCL to 15, then the Shewhart ARLs in Table 9 increase to 758.1, 449.2, 445.4, 418.0, and 364.7, respectively, that is, in none of the scenarios, the in-control ARL is much below the target $\text{ARL}_0 = 370.4$ anymore. Analogously, one could choose $h \geq 47$ for the CUSUM chart to meet the target for more and more of the ML-fits. Thus, although we cannot know the true model behind the DGP, we may use the spINARCH(1) model as a tool for experiencing the effects of nonlinearity on the chart's performance and adjust the chart design accordingly.

6.2 | Monitoring demand counts

Our second real-world data application is from the fields of service and healthcare industries. The $T = 240$ counts $x_1, \ldots, x_T$ express the daily demand for RBC O+ transfusion blood bags in a large regional hospital in southeastern Wisconsin (period June, 2009 to January, 2010), see Alwan and Weiß[36] for further details. The ACF plot in Figure 3 again shows an AR(1)-like decay such that the ordinary (linear) Poisson INARCH(1) model might appear reasonable to the practitioner. From

TABLE 10 Demand counts data: parameter estimates with maximized log-likelihood ("max. ℓ"), properties of fitted models, and in-control ARLs of Shewhart chart with UCL = 14 or CUSUM chart with $(h, k) = (22, 4)$ for respective model fit

Neg. bin. spINARCH(1)	Parameter estimation				Model properties			ARLs	
	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\pi}$	max. ℓ	μ	σ^2/μ	$\rho(1)$	Shewh.	CUSUM
Moment fit:	2.050	0.266	0.441	—	2.792	2.439	0.266	451.6	397.7
ML, $c = 0.0$	2.092	0.251	0.424	−507.8	2.792	2.516	0.251	394.0	388.6
0.5	2.084	0.252	0.424	−507.8	2.791	2.517	0.251	394.3	387.8
1.0	1.962	0.269	0.425	−507.7	2.789	2.520	0.256	393.9	376.7
2.0	1.239	0.336	0.426	−507.5	2.789	2.530	0.266	389.0	350.0

$\overline{x} \approx 2.792$ and $\hat{\rho}(1) \approx 0.266$, one computes the moment estimates $(\hat{\alpha}_0, \hat{\alpha}_1) \approx (2.050, 0.266)$, which lead to the c-chart design UCL = 8 (planned ARL ≈ 284.2) in analogy to Section 6.1. This time, however, a retrospective application of the c-chart causes several alarms, see Figure 3, clearly indicating that the fitted Poisson INARCH(1) model is not adequate for the data. In fact, the data exhibit much more overdispersion (sample value 2.439 for σ^2/μ) than covered by the model (model value 1.076 for σ^2/μ). Thus, for the demand-counts data, we are concerned with another type of model uncertainty, namely a misspecified distribution. But as we see from the retrospective c-chart application, this issue was easily uncovered, and as a consequence, the practitioner will now turn to an INARCH(1)-type model allowing for additional dispersion.

There are many distributions for modeling overdispersed counts, see Weiß,[4] with one of the most common choices being the NB distribution. Therefore, instead of using the Poi(M_t)-distribution for emitting the count X_t at time t, an obvious modification is to use NB(M_t $\frac{\pi}{1-\pi}$, π) instead, also see the discussion in Section 2, which has mean M_t but variance $\frac{1}{\pi} \cdot M_t$. So $\pi \in (0; 1)$ serves as an additional dispersion parameter, where the limit $\pi \to 1$ leads to Poi(M_t). The next steps are analogous to Section 6.1, that is, we start with moment estimation, design the control charts, and then do sensitivity analyses based on the ML-fitted NB-spINARCH(1) models. The results are summarized in Table 10. Compared to Table 9, we have considerably less estimation uncertainty this time if comparing the moment and ML fits for $c = 0$. This is plausible as $T = 240$ is a much larger sample size, being close to the recommendation derived from Table 5. But also the model uncertainty with respect to possible nonlinearity (as imitated by the ML fits for $c > 0$) is clearly reduced. The Shewhart ARLs reach the target $\text{ARL}_0 = 370.4$ without exception, and for the CUSUM ARLs, we have a mild violation only for the strong nonlinearity $c = 2.0$. Thus, if judged as being necessary at all, a practitioner might conclude on adjusting the CUSUM's limit to $h = 23$ (corresponding to ARL = 398.5 for $c = 2.0$) for achieving robustness against possibly neglected nonlinearity. This choice, however, would lead to ARL = 446.1 for $c = 0.0$, which notably exceeds the target $\text{ARL}_0 = 370.4$.

7 | CONCLUSIONS

In this study, we investigated the sensitivity of the performance of Shewhart-type and CUSUM control charts on the assumption of linear dependence of count-type time series when the DGP is in fact nonlinear. For the studied models and parametrizations, a general conclusion for the cases of specified model parameters is that the Shewhart-type and CUSUM control charts can be considered to be robust to model misspecification such that the planned in-control ARL performances do not deviate much from the actual ARL, when a linear model is utilized for modeling a nonlinear DGP. There are some exceptions to this conclusion when the extent of violation of linearity is moderate-to-high ($c \geq 1$), for the models with low counts (M1a–b), with very asymmetrical marginal distributions (M2c–d), and with the slowly decaying ACF (M3a–d). The effect of disregarded nonlinearity is stronger for the CUSUM than for the Shewhart-type control charts, since the CUSUM control chart is sensitive to even slight violations of model assumptions due to the accumulation of counts.

On the other hand, when the effects of errors in parameter estimation are considered together with the model misspecification, it is also observed that the performance does not necessarily get worse than the ones of parameters specified cases, since the parameter values of the fitted linear model may adapt to the nonlinearity (and possibly deviating model order) in the data. While increasing nonlinearity causes a decline of the actual in-control ARL for Shewhart-type charts when the parameters are specified, this usually does not happen for the mode ARL under parameter estimation so that one can benefit from the trade-off between model fitting and model misspecification with a good chance. Consequently, as the mode ARL is typically close to the target ARL_0 (and an increased sample size usually leads an increased frequency

of the mode ARL), we consider Shewhart-type control charts to be robust, for the models and parametrizations considered in the study, when the model is misspecified to be exactly linear and parameters are estimated. We recommend a sample size of ≥250 for achieving the desired Shewhart chart's performance even under model misspecification.

However, we have differing conclusions for the CUSUM control charts since the mode ARLs usually deviate considerably from the ones of specified parameters cases, and since there is a much higher variability among the actual ARLs. Therefore, the CUSUM designs are not considered to be robust in performance when both parameters are estimated and model is misspecified. Generally, regarding the issues with the CUSUM chart or being faced with insufficient Phase-I data, we recommend a regular update of the model fit and chart design once further in-control observations are available. In particular, as illustrated by our data examples, we suggest to use the spINGARCH or scBINGARCH models (or other suitable nonlinear models) as a tool for experiencing the effects of possible nonlinearity on the charts' performances, and to adjust the chart designs accordingly.

As possible directions for future research, the effect of (possibly misspecified) model fitting could be analyzed for control charts with a randomized decision rule such as used by Paulino et al.[29] and Morais et al.[30] Furthermore, analogous robustness analyses could be done for other types of INGARCH model (and nonlinear extensions thereof), such as the zero-one-inflated bounded Poisson autoregressive (ZOBPAR) model considered by Liu et al.[37]

ACKNOWLEDGMENT

The authors thank the referee for useful comments on an earlier draft of this article.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Christian H. Weiß https://orcid.org/0000-0001-8739-6631
Murat Caner Testik https://orcid.org/0000-0003-2389-4759

REFERENCES

1. Paolella MS. *Linear Models and Time-Series Analysis: Regression, ANOVA, ARMA and GARCH*. John Wiley & Sons Ltd.; 2019.
2. Holan SH, Lund R, Davis G. The ARMA alphabet soup: a tour of ARMA model variants. *Stat Surv*. 2010;4:232-274.
3. Grunwald G, Hyndman RJ, Tedesco L, Tweedie RL. Non-Gaussian conditional linear AR(1) models. *Aust N Z J Stat*. 2000;42(4):479-495.
4. Weiß CH. *An Introduction to Discrete-Valued Time Series*. John Wiley & Sons, Inc.; 2018.
5. Alwan LC. Effects of autocorrelation on control chart performance. *Commun Stat — Theory Methods*. 1992;21(4):1025-1049.
6. Knoth S, Schmid W. Control charts for time series: a review. In: Lenz H-J, Wilrich P-T, eds. *Frontiers in Statistical Quality Control*. Vol 7. Physica-Verlag; 2004:210-236.
7. Psarakis S, Papaleonida GEA. SPC procedures for monitoring autocorrelated processes. *Qual Technol Quant Manag*. 2007;4(4):501-540.
8. Weiß CH. SPC methods for time-dependent processes of counts — a literature review. *Cogent Math*. 2015;2(1):1111116.
9. Lee S, Kim B. Recent progress in parameter change test for integer-valued time series models. *J Korean Stat Soc*. 2021;50(3):730-755.
10. Jensen WA, Jones-Farmer LA, Champ CW, Woodall WH. Effects of parameter estimation on control chart properties: a literature review. *J Qual Technol*. 2006;32(4):395-409.
11. Psarakis S, Vyniou AK, Castagliola P. Some recent developments on the effects of parameter estimation on control charts. *Qual Reliab Eng Int*. 2014;30(8):1113-1129.
12. Ferland R, Latour A, Oraichi D. Integer-valued GARCH processes. *J Time Ser Anal*. 2006;27(6):923-942.
13. Ristić MM, Weiß CH, Janjić AD. A binomial integer-valued ARCH model. *Int J Biostat*. 2016;12(2):20150051.
14. Weiß CH, Testik MC. Detection of abrupt changes in count data time series: cumulative sum derivations for INARCH(1) models. *J Qual Technol*. 2012;44(3):249-264.
15. Huh J, Kim H, Lee S. Monitoring parameter shift with Poisson integer-valued GARCH models. *J Stat Comput Simul*. 2017:87(9):1754-1766.
16. Rakitzis AC, Weiß CH, Castagliola P. Control charts for monitoring correlated Poisson counts with an excessive number of zeros. *Qual Reliab Eng Int*. 2017;33(2):413-430.
17. Vanli OA, Giroux R, Ozguven EE, Pignatiello Jr,. J.J. Monitoring of count data time series: cumulative sum change detection in Poisson integer valued GARCH models. *Qual Eng*. 2019;31(3):439-452.
18. Ottenstreuer S. The Shiryaev–Roberts control chart for Markovian count time series. *Qual Reliab Eng Int*. 2022;38(3):1207-1225.
19. Anastasopoulou M, Rakitzis AC. One- and two-sided monitoring schemes for BINARCH(1) processes. *Appl Stoch Models Bus Ind*. 2022;38(2):262-286.
20. Weiß CH, Zhu F, Hoshiyar A. Softplus INGARCH models. *Stat Sinica*. 2022;32(2):1099-1120.
21. Weiß CH, Jahn M. Soft-clipping INGARCH models for time series of bounded counts. *Stat Model*. Forthcoming, 2022. https://doi.org/10.1177/1471082X221121223

22. Zhu F. A negative binomial integer-valued GARCH model. *J Time Ser Anal*. 2011;32(1):54-67.
23. Zhu F. Zero-inflated Poisson and negative binomial integer-valued GARCH models. *J Stat Plan Inference*. 2012;142(4):826-839.
24. Mei H, Eisner J. The neural Hawkes process: a neurally self-modulating multivariate point process. In: von Luxburg et al. eds. *Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS'17)*. Curran Associates Inc; 2017:6757-6767.
25. Klimek MD, Perelstein M. Neural network-based approach to phase space integration. *SciPost Phys*. 2020;9(4):053.
26. Knoth S. The art of evaluating monitoring schemes — how to measure the performance of control charts? In: Lenz H-J, Wilrich P-T, eds. *Frontiers in Statistical Quality Control*. Vol 8. Physica-Verlag; 2006:74-99.
27. Montgomery DC. *Introduction to Statistical Quality Control*. 6th ed. John Wiley & Sons, Inc.; 2009.
28. Weiß CH, Testik MC, Homburg A. On the design of Shewhart control charts for count time series under estimation uncertainty. *Comput Ind Eng*. 2021;157:107331.
29. Paulino S, Morais MC, Knoth S. On ARL-unbiased c-charts for INAR(1) Poisson counts. *Stat Pap*. 2019;60(4):1021-1038.
30. Morais MC, Knoth S, Cruz CJ, Weiß, CH. ARL-unbiased CUSUM schemes to monitor binomial counts. In: Knoth S, Schmid W, eds. *Frontiers in Statistical Quality Control*. Vol 13. Physica-Verlag; 2021:77-98.
31. Brook D, Evans DA. An approach to the probability distribution of CUSUM run length. *Biometrika*. 1972;59(3):539-549.
32. Weiß CH. The Markov chain approach for performance evaluation of control charts — a tutorial. In: Werther SP, ed. *Process Control: Problems, Techniques and Applications*. Nova Science Publishers, Inc.; 2011:205-228.
33. Weiß CH, Testik MC. The Poisson INAR(1) CUSUM chart under overdispersion and estimation error. *IIE Trans*. 2011;43(11):805-818.
34. Zhang M, Nie G, He Z, & Hou X. The Poisson INAR(1) one-sided EWMA chart with estimated parameters. *Int J Prod Res*. 2014;52(18):5415-5431.
35. Burnham KP, Anderson DR. *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. 2nd ed. Springer-Verlag; 2002.
36. Alwan LC, Weiß CH. INAR implementation of newsvendor model for serially dependent demand counts. *Int J Prod Res*. 2017;55(4):1085-1099.
37. Liu M, Zhu F, Zhu K. Modeling normalcy-dominant ordinal time series: an application to air quality level. *J Time Ser Anal*. 2022;43(3):460-478.

AUTHOR BIOGRAPHIES

Christian H. Weiß is a Professor at the Department of Mathematics and Statistics at the Helmut Schmidt University in Hamburg, Germany. His research areas include time series analysis, statistical quality control, and computational statistics.

Murat C. Testik is a Professor and Department Chair in the Industrial Engineering Department at Hacettepe University. He has a PhD degree in Industrial Engineering from Arizona State University with a major in quality engineering. His research interests and publications are mainly in the area of quality engineering and data mining for quality and process improvement. Testik is the past President of the European Network for Business and Industrial Statistics. He is a past Editor-In-Chief of Quality Engineering.

How to cite this article: Weiß CH, Testik MC. Monitoring count time series: Robustness to nonlinearity when linear models are utilized. *Qual Reliab Eng Int*. 2022;38:4356–4371. https://doi.org/10.1002/qre.3215