RESIDUAL-BASED CUSUM BETA REGRESSION CONTROL CHART FOR MONITORING DOUBLE BOUNDED PROCESSES

CRISTINE RAUBER[1], LUIZ M. A. LIMA-FILHO[2], AND FÁBIO M. BAYER[3]

Abstract. This paper proposes a control chart useful for detecting small shifts in the mean of a double bounded process, such as fractions and proportions, in the presence of control variables. For this purpose, we consider the cumulative sum control chart applied to different residuals of the beta regression model. We conduct an extensive Monte Carlo simulation study to evaluate and compare the performance of the proposed control chart with two other control charts in the literature in terms of run length analysis. The numerical results show that the proposed control chart is more sensitive to detect changes in the process than its competitors and that the quantile residual is the most suitable residual to be used in our proposal. Finally, based on the quantile residual, we present and discuss applications to real and simulated data to show the applicability of the proposed control chart.
Keywords: Beta regression, control chart, cumulative sum, residuals, run length.

1. Introduction

Statistical process control (SPC) is a collection of techniques useful for monitoring and controlling a process (Fournier et al., 2006). Under natural variability, that is, a common variation that will always exist, the process is in statistical control. However, when the variability stems from external sources, the process is out of statistical control. Specialists desire to quickly detect shifts in the process monitoring, thus the control chart is the simplest and most used tool for this purpose (Montgomery, 2009).

The usual control chart proposed by ShewhartShewhart, 1931 is widely used to monitor independent random variables and can detect large shifts in the mean of a process (Aslam, Azam, and Jun, 2014). These charts are also known as memory-less control charts because they consider observations at a given time. However, not observing and analyzing previous observations can lead to the poor performance of the control chart. Moreover, conventional control charts require some assumptions or approximations. Alternatively, more advanced statistical methods have been proposed in the literature, such as the cumulative sum (CUSUM) (Page, 1954) and the exponentially weighted moving average (EWMA) (Roberts, 1959) control charts. CUSUM and EWMA control charts have been studied by some authors over the years. Some past works can be found in Crowder, 1989; Ewan, 1963; Gan, 1991; Hawkins, 1981; Lucas and Saccucci, 1990; Page, 1961; Woodall and Adams, 1993, while some recent developments are found in Adegoke et al., 2019; Aytaçoğlu, Driscoll, and Woodall, 2021; Haq, 2017; Park and Jun, 2015; Perry and Wang, 2020; Sanusi, Abbas, and Riaz, 2018; Xue and Qiu, 2021. Such charts quickly detect small shifts in the mean of a process and use cumulative information from the observations, thus being called memory-type control charts.

CUSUM control charts can be built using different statistics. Recently, the discussion on residual-based CUSUM control charts has received considerable attention. For example, Asadzadeh, Aghaie,

[1]Department of Mathematics and Statistics, Lancaster University. E-mail: rauberoliveira@gmail.com (corresponding author)
[2]Departamento de Estatística, Universidade Federal da Paraíba, João Pessoa, PB, Brazil.
[3]Departamento de Estatística, Universidade Federal de Santa Maria, Santa Maria, RS, Brazil.

and Niaki, 2013 monitored the Cox-Snell residuals of accelerated failure time models based on two regression-adjusted control approaches. Chen and Huang, 2014 used a residual-based CUSUM control chart to monitor syndromic data on the respiratory syndrome in Taiwan. Weiß and Testik, 2015 discussed the problem of monitoring autocorrelated, count-type discrete data by investigating the CUSUM control chart based on three different residuals. Alencar, Lee Ho, and Albarracin, 2017 and Albarracin, Alencar, and Lee Ho, 2018 evaluate the performance of the CUSUM control chart using the deviance residual, and other statistics, in the monitoring of negative binomial and negative binomial generalized autoregressive moving average processes (Benjamim, Rigby, and Stasinopoulos, 2003), respectively. Kim and Lee, 2021 introduced the residual-based CUSUM scheme in first-order Poisson integer-valued autoregressive models where the residuals are computed through squared difference estimates.

In practical situations, there is an interest in modeling and monitoring variables limited to the unit interval $(0, 1)$, such as fractions or proportions. Here, the traditional Shewhart-type control chart may not be adequate since fractions or proportions frequently follow a skew distribution, thus not holding the normality assumption. Alternatively, control charts for double bounded quality characteristics have been proposed in the literature. For example, Sant'Anna and Caten, 2012 proposed the beta control chart (BCC), where the authors assume that the variable of interest is beta distributed and the control limits are estimated using the beta cumulative distribution function. The BCC is more advantageous than the Shewhart control chart as it naturally captures the asymmetry of the quality characteristic and its control limits range between $(0, 1)$. Lima-Filho et al., 2019 proposed a new control chart to model the mean of double bounded processes in the presence of zeros and ones. The control limits of this chart are based on the inflated beta probability distribution function. As the Kumaraswamy distribution (Jones, 2009) is a good alternative to the beta distribution, Lima-Filho and Bayer, 2021 introduced a novel control chart based on the Kumaraswamy distribution to monitor environmental data limited to the unit interval. Nevertheless, production processes can also be related to external variables. For this purpose, Bayer, Tondolo, and Müller, 2018 proposed the beta regression control chart (BRCC), where the control limits are defined using the quantile function of the beta distribution, and Lima-Filho et al., 2020 provided a general framework to a recent control chart based on the inflated beta regression model to monitor the mean of environmental processes containing zeros and ones. Recently, Hwang, 2021 proposed a novel CUSUM control chart based on the deviance residual of a beta regression model to monitor the mean of a univariate process.

Although the BCC and BRCC are better alternatives than the conventional charts for double bounded data, they still are memory-less approaches and a CUSUM control chart alternative could detect more quickly a shift in the mean of a process. In this regard, the chief contribution of this paper is to propose and compare the performance of the CUSUM beta regression control chart on the residuals of the beta regression model, named CUSUM-BRCC. The CUSUM-BRCC is useful for monitoring double bounded processes where the quality characteristic is affected by control variables in which the process output may represent individual measures (e.g. efficiency score) or a ratio between continuous numbers (e.g. relative humidity). As there are different residuals for the beta regression (Espinheira, Ferrari, and Cribari-Neto, 2008; Pereira, 2019), we explore the CUSUM-BRCC based on the standardized, two types of standardized weighted, and quantile residuals. It is noteworthy that we do not consider the deviance residual because it cannot be calculated for several observations in beta regression provided that the contribution of these observations to the deviance is negative. In addition, a similar approach is already published in the literature (Hwang, 2021). We conduct an extensive Monte Carlo simulation study to evaluate and compare the performance of the proposed control chart based on different residuals in terms of run length (RL) analysis. Our Monte Carlo simulation results show that the quantile residual is the most suitable residual to be considered when controlling and monitoring processes limited to the unit interval $(0,1)$ and in the presence of external variables.

The remainder of the paper unfolds as follows. In Section 2, we present the proposed CUSUM-BRCC and the residuals considered in this paper. Section 3 presents a simulation study to evaluate and compare the performance of the proposed control chart with of the BRCC and the control chart proposed by Hwang Hwang, 2021, which we shall denote CUSUM-BRCC$_{\text{Hwang}}$, based on RL analysis. In Section 4, based on the quantile residual, we discuss and present two applications to show the applicability of the proposed control chart. Finally, some concluding remarks are presented in Section 5.

2. Residual-based CUSUM beta regression control chart

Control charts are powerful tools used to monitor a quality characteristic of interest. Usually, this graphical device consists of a center line (CL), representing the mean of the quality variable, and two other horizontal lines representing the upper (UCL) and lower (LCL) control limits. When all the sample points fall within the control limits, the process is in a state of control and no further action is needed. If an observation exceeds the desired limits, the process is assumed to be out of control and corrections are required to understand the causes and improve the process.

In this paper, we consider the monitoring of processes limited to the unit interval $(0,1)$, such as fractions or proportions. The beta distribution is frequently used to model this type of data. The beta law is very flexible and can assume a wide variety of shapes, such as symmetric, asymmetric, J-shaped, inverted J-shaped, U-shaped, and uniform. Using a parametrization that considers a location (μ) and a precision (ϕ) parameter, Ferrari and Cribari-Neto, 2004 introduce the beta density as

$$f(y;\mu,\phi)=\frac{\Gamma(\phi)}{\Gamma(\mu\phi)\Gamma((1-\mu)\phi)}y^{\mu\phi-1}(1-y)^{(1-\mu)\phi-1},\quad 0<y<1, \tag{1}$$

where $0<\mu<1$, $\phi>0$, and $\Gamma(\cdot)$ is the gamma function. The mean and variance of y are $\mathbb{E}(y)=\mu$ and $\text{Var}(y)=\mu(1-\mu)/(1+\phi)$, respectively.

The cumulative distribution function of y is given by

$$\mathcal{F}(y;\mu,\phi)=\frac{\Gamma(\phi)}{\Gamma(\mu\phi)\Gamma((1-\mu)\phi)}\int_0^y y^{\mu\phi-1}(1-y)^{(1-\mu)\phi-1}\mathrm{d}y. \tag{2}$$

Therefore, the quantile function of y is expressed as $\mathcal{F}^{-1}(u,\mu,\phi)$, where u is the desired quantile.

Let $y_1,y_2,\ldots,y_n$ be a set of independent random variables where each y_t, with $t=1,2,\ldots,n$, has density and cumulative distribution given in (1) and (2), respectively, with mean μ_t and precision ϕ_t. We shall consider the following varying precision beta regression model in our study (Simas, Barreto-Souza, and Rocha, 2010; Smithson and Verkuilen, 2006):

$$g(\mu_t)=\sum_{i=1}^{r}x_{ti}\omega_i=\eta_{1t}\qquad\text{and}\qquad h(\phi_t)=\sum_{j=1}^{s}z_{tj}\gamma_j=\eta_{2t},$$

where $\boldsymbol{\omega}=(\omega_1,\ldots,\omega_r)^\top\in\mathbb{R}^r$ and $\boldsymbol{\gamma}=(\gamma_1,\ldots,\gamma_s)^\top\in\mathbb{R}^s$ are unknown parameters vectors, $x_{1t},\ldots,x_{rt}$ and $z_{1t},\ldots,z_{st}$ are the known and fixed covariates of the mean and precision submodels, respectively, $\boldsymbol{\eta_1}=(\eta_{11},\ldots,\eta_{1n})^\top$ and $\boldsymbol{\eta_2}=(\eta_{21},\ldots,\eta_{2n})^\top$ being the mean and precision linear predictor vectors, respectively. Here, $g(\cdot)$ and $h(\cdot)$ denote the link functions that are strictly monotonic and twice differentiable such that $g:(0,1)\mapsto\mathbb{R}$ and $h:(0,\infty)\mapsto\mathbb{R}$. Thus, the mean and precision of each y_t are given, respectively, by

$$\mu_t = g^{-1}(\eta_{1t}), \tag{3}$$

$$\phi_t = h^{-1}(\eta_{2t}). \tag{4}$$

One can choose the logit, probit, loglog, and cloglog for the mean link function, while for the precision link function we have the log as a common choice.

Let $\boldsymbol{\theta} = (\boldsymbol{\omega}^\top, \gamma^\top)^\top$ be the beta regression parameter vector. The log-likelihood function of $\boldsymbol{\theta}$ is given by

$$\ell(\boldsymbol{\theta}) = \sum_{t=1}^{n} \ell_t(\mu_t, \phi_t), \tag{5}$$

where

$$\begin{aligned}\ell_t(\mu_t, \phi_t) = &\log\Gamma(\phi_t) - \log\Gamma(\mu_t\phi_t) - \log\Gamma((1-\mu_t)\phi_t) + (\mu_t\phi_t - 1)\log y_t \\ &+ [(1-\mu_t)\phi_t - 1]\log(1-y_t),\end{aligned}$$

with μ_t and ϕ_t defined in (3) and (4), respectively.

By taking first-order derivatives of (5) with respect to each element of $\boldsymbol{\theta}$, we obtain a system of equations called score vector. By setting the score vector equals zero, we obtain the maximum likelihood estimators (MLE) of the model parameters. Since the MLE do not have closed-form expressions, we maximize the log-likelihood function in (5) using numerical methods. It is noteworthy that using the MLE of $\boldsymbol{\theta}$ gives us $\hat{\eta}_{1t}$ and $\hat{\eta}_{2t}$, which are used to obtain $\hat{\mu}$ and $\hat{\phi}$. These quantities are essential in the construction of the residuals considered in the present paper. More details on inferences in beta regression models can be found in Ferrari and Cribari-Neto, 2004 and Cribari-Neto and Zeileis, 2010.

For monitoring double bounded processes with covariates, based on the beta regression model, Bayer, Tondolo, and Müller, 2018 proposed the BRCC, where the control limits are based on the quantile function and evaluated under the estimated parameters in a Shewhart fashion. The authors show that the BRCC works well for monitoring fractions and proportions and easily detects changes in the mean of the process. When the interest response variable depends on control variables, one can apply a conventional control chart to the residuals of the model as they approximately follow a normal distribution and are independently distributed if the fitted model is well specified (Montgomery, 2009). However, it is well known that CUSUM control charts are more sensitive to detect shifts than Shewhart-type charts (Lucas, 1976). In this way, we concentrate on the residual-based CUSUM beta regression control chart to model the process mean.

As discussed in Espinheira, Ferrari, and Cribari-Neto, 2008, Pereira, 2019 and Ferrari and Cribari-Neto, 2004, there are different residuals for the beta regression model. The residuals we consider are the standardized residual defined by Ferrari and Cribari-Neto, 2004, the standardized weighted 1 and 2 residuals introduced by Espinheira, Ferrari, and Cribari-Neto, 2008, and the quantile residual (Dunn and Smyth, 1996) considered in the beta regression model by Pereira, 2019. Although Anholeto, Sandoval, and Botter, 2014 proposed adjusted Pearson residuals for beta regressions, we do not consider such residuals due to the matrix formulae involved when implementing the residuals in a practical situation.

The standardized residual is given by

$$r_t^s = \frac{y_t - \hat{\mu}_t}{\sqrt{\widehat{\mathrm{Var}}(y_t)}}, \tag{6}$$

where $\widehat{\mathrm{Var}}(y_t) = \hat{\mu}_t(1-\hat{\mu}_t)/(1+\hat{\phi}_t)$ and $\hat{\mu}_t = g^{-1}(\hat{\eta}_{1t})$.

The standardized weighted 1 residual is given by

$$r_t^w = \frac{y_t^* - \hat{\mu}_t^*}{\sqrt{\hat{v}_t^*}}, \tag{7}$$

where $y_t^* = \log(y_t/(1-y_t))$, $\hat{\mu}_t^* = \hat{\mathbb{E}}(y_t^*) = \psi(\hat{\mu}_t\hat{\phi}_t) - \psi((1-\hat{\mu}_t)\hat{\phi}_t)$, $\hat{v}_t^* = \widehat{\mathrm{Var}}(y_t^*) = \psi'(\hat{\mu}_t\hat{\phi}_t) + \psi'((1-\hat{\mu}_t)\hat{\phi}_t)$, ψ and ψ' denoting the digamma and trigamma functions, respectively.

The standardized weighted 2 residual is computed by

$$r_t^{ww} = \frac{y_t^* - \hat{\mu}_t^*}{\sqrt{\hat{v}_t^*(1-v_{tt})}}, \tag{8}$$

where v_{tt} is the t-th element of $V = \widehat{W}^{1/2}X(X^\top\widehat{W}X)^{-1}X^\top\widehat{W}^{1/2}$, $X = (x_1, \ldots, x_n)^\top$ is the regressor matrix and $\widehat{W} = \mathrm{diag}(\hat{w}_1, \ldots, \hat{w}_n)$, with $\hat{w}_t = \hat{\phi}_t^2\hat{v}_t^*\left[1/\{g'(\hat{\mu}_t)\}^2\right]$.

Finally, the quantile residual is expressed as

$$r_t^q = \Phi^{-1}\{\mathcal{F}(y_t; \hat{\mu}_t, \hat{\phi}_t)\}, \tag{9}$$

where $\Phi(\cdot)$ denotes the cumulative distribution function of the standard normal distribution and $\mathcal{F}(\cdot)$ is the cumulative distribution function of the beta distribution in (2).

The deviance residual used in the CUSUM-BRCC$_{\text{Hwang}}$ is defined as

$$r_t^d = \mathrm{sign}(y_t - \hat{\mu}_t)\{2|(\ell_t(\tilde{\mu}_t, \hat{\phi}_t) - \ell_t(\hat{\mu}_t, \hat{\phi}_t))|\}^{1/2}.$$

Here, $\mathrm{sign}(\cdot)$ is the signal function, $\ell_t(\mu_t, \phi_t)$ is the contribution of the t-th observation of the log-likelihood function given in (5), $\hat{\mu}_t$ and $\hat{\phi}_t$ are the maximum likelihood estimators of μ_t and ϕ_t, respectively, and $\tilde{\mu}_t$ and $\tilde{\phi}_t$ being the estimate of μ_t and ϕ_t in the saturated model, respectively. It is noted by Ferrari and Cribari-Neto, 2004 that for large values of ϕ, $\tilde{\mu}_t \approx y_t$. Note that this residual considers the absolute value of the contribution of each observation to the deviance, which is not reasonable for defining a residual.

Based on the regression structures defined in (3) and (4) and on the residuals given in (6), (7), (8), and (9), we propose the CUSUM-BRCC using each of the residuals mentioned, thus resulting in four new control charts.

The CUSUM-BRCC statistics are given by

$$C_t^+ = \max[0, r_t - (m_0 + K) + C_{t-1}^+],$$
$$C_t^- = \max[0, (m_0 - K) - r_t + C_{t-1}^-],$$

where r_t is the t-th observation of each residual considered, m_0 is the target mean value, that is, the residual mean, K is a reference value, and $C_0^+ = C_0^- = 0$ are the starting values. The reference value is given by $K = k \times \sigma$, where σ is the standard deviation of the residual used to build the CUSUM chart, and the choice of k is related to the magnitude of the change that we want to identify, that is, $k = \frac{1}{2} \times \Delta$, where Δ is the size of the shift in standard deviation units. The decision interval of the CUSUM control chart is expressed as $H = h \times \sigma$ where h is estimated for each residual. If either C_t^+ or C_t^- exceed H, then the process is out of control.

3. Simulation study

In what follows, we shall present the results of a Monte Carlo simulation study we performed to evaluate and compare the CUSUM-BRCC with the BRCC proposed by Bayer, Tondolo, and Müller, 2018 and the CUSUM-BRCC proposed by Hwang, 2021. We considered the residuals in (6), (7), (8), and (9) to be used in the CUSUM-BRCC proposed in this paper, and $n \in \{100, 200, 500\}$. For brevity and similarity

Table 1. True parameter values for the scenarios considered in the simulation study.

Scenario	ω_0	ω_1	γ_0	γ_1
1	−3.2	2.0	3.0	1.0
2	−3.2	2.0	4.0	0.5
3	−1.0	2.0	3.0	1.5
4	−1.0	2.0	2.0	2.0
5	1.0	2.4	2.0	3.0
6	1.0	2.4	4.0	2.5

of results, we only present numerical evidence for $n = 500$ based on 5,000 Monte Carlo replications. All simulations were performed using the `R programming language` (R Core Team, 2021).

The performance of control charts is usually measured in terms of RL analysis. The average run length (ARL) is the average number of observations that must be plotted until the control chart signals. Thus, for an in-control process, this measure is known as ARL_0, whereas for an out-of-control process it is named ARL_1. The latter case means that the mean has shifted, then a smaller number of samples would allow the detection of the shift more quickly. In this work, we considered the ARL, median run length (MRL), and standard deviation run length (SDRL) for a process in control, where $\text{ARL} = 1/\alpha$, $\text{MRL} = \ln(0.5)/\ln(1-\alpha)$, and $\text{SDRL} = \sqrt{(1-\alpha)/\alpha^2}$ (Lee Ho, Fernandes, and Bourguignon, 2019; Lima-Filho et al., 2019). Here, α is the false alarm probability, that is, the probability of a single observation falling outside the control limits when the process is in control. Therefore, assuming that the process is in control and $\alpha = 0.005$, we obtain $\text{ARL}_0 = 200$, i.e. we expect an out-of-control signal every 200 samples, on average, even when the process remains in control. The nominal values of MRL and SDRL for a process with the same characteristics are 138.3, and 199.5, respectively. In order for the proposed control charts to present the same target in-control ARL, we first calibrated them using Algorithm 1 to find the optimal h for each residual. To the best of our knowledge, this calibration is not needed for the $\text{CUSUM-BRCC}_{\text{Hwang}}$, therefore we used the approximation given by Siegmund Siegmund, 1985 to obtain the optimal h for a target $\text{ARL}_0 = 200$.

To evaluate the RL when the process is out of control, we introduced a δ change in the mean regression structure of the process. In the true data generation process, we used the following beta regression model:

$$\text{logit}(\mu_t) = \delta + \omega_0 + \omega_1 x_t,$$
$$\log(\phi_t) = \gamma_0 + \gamma_1 z_t,$$

where $t = 1, \ldots, n$, δ ranging from −0.5 to 0.5 by steps of 0.1, ω_0, ω_1, γ_0, and γ_1 being the regression coefficients. Since δ is the induced change in the mean, the process is in control when $\delta = 0$. The values of x_t and z_t were obtained from a uniform distribution in the interval $(0, 1)$ and considered constant through all Monte Carlo replications.

Table 1 shows six scenarios of the parameter values with different characteristics considered in the numerical evaluation. In Scenarios 1 and 2, the mean is close to 0.1, with $\phi \in [20, 54]$ in Scenario 1 and $\phi \in [55, 90]$ in Scenario 2. In Scenarios 3 and 4, we have $\phi \in [20, 90]$ and $\phi \in [7, 54]$, respectively, and the mean is centered on the standard unit interval. Finally, in Scenarios 5 and 6, the mean is close to 0.9 with $\phi \in [7, 147]$ in Scenario 5 and $\phi \in [55, 659]$ in Scenario 6. Note that we covered a wide range of scenarios for the mean and precision of the process. The Monte Carlo simulation study is divided into two procedures and summarized by Algorithms 1 and 2.

Algorithm 1: Algorithm for estimating h in forming the proposed control charts.

(1) Define the desired probability of false alarm α (herein we used $\alpha = 0.005$);
(2) Generate n observations from a standard uniform distribution $(0, 1)$;
(3) Using covariates and parameter values, compute μ_t and ϕ_t;
(4) Generate n beta-distributed observations with parameters μ_t and ϕ_t;
(5) Fit the beta regression model with varying precision and obtain the residuals in (6), (7), (8), and (9);
(6) Obtain the CUSUM-BRCC for each residual in step 5 using H as decision interval;
(7) Plot each data point r_t together with the control limits, for $t = 1, \ldots, n$. The observation r_t that is out of the control limits interval is an out-of-control observation;
(8) Repeat steps 4 to 7 a large number of times, say 5,000;
(9) At the end of the replications, the in-control average run length is calculated, obtaining $\widehat{\text{ARL}}_0$;
(10) Repeat steps 4 to 9 for different values of H;
(11) Fit a linear regression where H is the response variable and the logarithm of the estimated ARL_0 is the control variable to obtain the exact value of H that yields the desired ARL_0 (herein we chose $\text{ARL}_0 = 200$).

Algorithm 2: Algorithm for the performance evaluation of the proposed control charts.

(1) Use the same probability of false alarm α defined in Algorithm 1;
(2) Generate n observations from a standard uniform distribution $(0, 1)$;
(3) Using covariates and parameter values, compute μ_t and ϕ_t;
(4) Generate n beta-distributed observations with parameters μ_t and ϕ_t;
(5) Fit the beta regression model with varying precision and obtain the residuals in (6), (7), (8), and (9);
(6) Obtain the CUSUM-BRCC for each residual in step 5 using the estimated value of h obtained from Algorithm 1 as decision interval;
(7) Plot each data point r_t together with the control limits, for $t = 1, \ldots, n$. The observation r_t that is out of the control limits interval is an out-of-control observation;
(8) Repeat steps 4 to 7 a large number of times, say 5,000;
(9) At the end of the replications, the average of each measure is calculated, obtaining the following Monte Carlo estimates: $\widehat{\text{ARL}}$, $\widehat{\text{MRL}}$, and $\widehat{\text{SDRL}}$;
(10) Include a shift (δ) in the linear predictor of the mean after fitting the beta model in step 5 and repeat steps 6 to 9 for different values of δ.

Tables 2 and 3 present results for $\widehat{\text{ARL}}$, $\widehat{\text{MRL}}$, and $\widehat{\text{SDRL}}$ for all scenarios. Notice that the control charts obtained similar performance when the process was in control ($\delta = 0$) and presented estimates close to their nominal values, except for $\text{CUSUM-BRCC}_{\text{Hwang}}$, which presented the worst performance among the control charts studied. We emphasize that the proposed control chart was calibrated (Algorithm 1) for each residual in order to have a target ARL_0 of 200. We observe in Scenario 1 that, when considering $\text{CUSUM-BRCC}_{r_t^w}$ (Equation (7)), we obtained $\widehat{\text{ARL}}_0 = 222.52$ when $\text{ARL}_0 = 200$ was expected, that is, in the worst scenario, there was a distortion of 11%.

When the process was out of control, the CUSUM-BRCC presented smaller values of $\widehat{\text{ARL}}_1$ than the BRCC, evidencing that the proposed control chart is more sensitive to detect changes in the mean of the variable of interest. For example, in Table 2 for Scenario 2 and $\delta = 0.1$, the $\text{CUSUM-BRCC}_{r_t^q}$ signaled

TABLE 2. Performance of the BRCC and CUSUM-BRCC considering the different residuals with $\alpha = 0.005$ for Scenarios 1, 2, and 3.

Scenario	δ	BRCC	CUSUM-BRCC					BRCC	CUSUM-BRCC					BRCC	CUSUM-BRCC				
			Hwang	r_t^{ww}	r_t^{w}	r_t^{s}	r_t^{q}		Hwang	r_t^{ww}	r_t^{w}	r_t^{s}	r_t^{q}		Hwang	r_t^{ww}	r_t^{w}	r_t^{s}	r_t^{q}
		$\widehat{\mathrm{ARL}}$						$\widehat{\mathrm{MRL}}$						$\widehat{\mathrm{SDRL}}$					
	−0.5	19.86	1.02	1.02	1.02	1.05	1.03	13.41	0.18	0.18	0.18	0.23	0.19	19.35	0.14	0.16	0.16	0.23	0.16
	−0.4	31.48	1.04	1.04	1.04	1.12	1.05	21.47	0.21	0.22	0.22	0.31	0.23	30.97	0.20	0.21	0.21	0.36	0.23
	−0.3	51.71	1.22	1.17	1.17	1.92	1.23	35.50	0.41	0.36	0.36	0.94	0.41	51.21	0.52	0.45	0.45	1.33	0.53
	−0.2	89.03	3.40	3.12	3.12	14.30	3.86	61.37	1.99	1.79	1.80	9.56	2.31	88.53	2.86	2.57	2.57	13.79	3.32
	−0.1	148.92	18.79	24.88	24.89	244.67	34.21	102.88	12.67	16.90	16.90	169.24	23.37	148.42	18.28	24.37	24.38	244.17	33.71
1	0.0	199.97	47.65	222.44	222.52	195.47	200.40	138.26	32.68	153.84	153.89	135.14	138.56	199.47	47.15	221.94	222.02	194.96	199.90
	0.1	186.73	14.10	99.82	99.89	28.86	39.40	129.09	9.42	68.84	68.89	19.65	26.96	186.23	13.59	99.32	99.39	28.35	38.90
	0.2	130.50	2.71	6.13	6.14	3.48	3.93	90.11	1.50	3.89	3.90	2.04	2.36	130.00	2.15	5.61	5.62	2.94	3.40
	0.3	77.57	1.12	1.24	1.24	1.17	1.17	53.42	0.31	0.43	0.43	0.36	0.21	77.06	0.36	0.55	0.55	0.44	0.44
	0.4	45.11	1.03	1.05	1.05	1.04	1.04	30.92	0.19	0.22	0.22	0.21	0.18	44.60	0.16	0.22	0.22	0.20	0.20
	0.5	26.45	1.01	1.03	1.03	1.02	1.02	17.99	0.16	0.19	0.19	0.18	0.19	25.95	0.12	0.16	0.16	0.15	0.15
		$\widehat{\mathrm{ARL}}$						$\widehat{\mathrm{MRL}}$						$\widehat{\mathrm{SDRL}}$					
	−0.5	11.21	1.01	1.01	1.01	1.02	1.01	7.42	0.15	0.16	0.16	0.17	0.16	10.70	0.10	0.11	0.11	0.14	0.12
	−0.4	18.62	1.02	1.02	1.02	1.03	1.02	12.55	0.17	0.18	0.18	0.20	0.18	18.11	0.13	0.14	0.14	0.17	0.15
	−0.3	34.04	1.03	1.04	1.04	1.07	1.04	23.24	0.20	0.21	0.21	0.25	0.22	33.53	0.18	0.21	0.21	0.27	0.21
	−0.2	67.04	1.33	1.36	1.36	1.98	1.44	46.12	0.50	0.52	0.52	0.99	0.59	66.54	0.67	0.70	0.70	1.40	0.80
	−0.1	133.69	8.50	12.38	12.39	42.52	15.20	92.32	5.54	8.23	8.24	29.12	10.19	133.19	7.99	11.87	11.88	42.01	14.69
2	0.0	200.53	48.41	206.27	206.27	202.40	194.52	138.65	33.21	142.63	142.63	139.94	134.49	200.03	47.91	205.77	205.77	201.90	194.02
	0.1	163.43	8.09	24.50	24.52	12.34	15.96	112.93	5.25	16.64	16.65	8.20	10.71	162.93	7.58	24.00	24.02	11.83	15.45
	0.2	88.52	1.24	1.44	1.44	1.30	1.34	61.01	0.42	0.59	0.59	0.47	0.51	88.02	0.54	0.80	0.80	0.62	0.68
	0.3	42.96	1.03	1.04	1.04	1.03	1.04	29.43	0.19	0.21	0.21	0.20	0.21	42.45	0.16	0.21	0.21	0.19	0.19
	0.4	21.49	1.01	1.02	1.02	1.02	1.02	14.55	0.16	0.18	0.18	0.17	0.17	20.99	0.11	0.14	0.14	0.13	0.13
	0.5	11.61	1.01	1.01	1.01	1.01	1.01	7.69	0.14	0.16	0.16	0.15	0.15	11.10	0.09	0.11	0.11	0.10	0.11
		$\widehat{\mathrm{ARL}}$						$\widehat{\mathrm{MRL}}$						$\widehat{\mathrm{SDRL}}$					
	−0.5	7.27	1.01	1.01	1.01	1.01	1.01	4.68	0.14	0.15	0.15	0.15	0.15	6.75	0.08	0.10	0.10	0.10	0.10
	−0.4	13.15	1.01	1.01	1.01	1.01	1.01	8.76	0.15	0.16	0.16	0.16	0.16	12.64	0.10	0.12	0.12	0.12	0.12
	−0.3	26.50	1.02	1.02	1.02	1.02	1.02	18.02	0.17	0.19	0.19	0.18	0.18	25.99	0.13	0.16	0.16	0.16	0.16
	−0.2	58.56	1.07	1.09	1.09	1.09	1.09	40.24	0.25	0.28	0.28	0.28	0.28	58.05	0.27	0.32	0.32	0.32	0.32
	−0.1	130.71	4.96	8.01	8.01	8.10	7.89	90.25	3.08	5.20	5.20	5.26	5.11	130.20	4.43	7.49	7.49	7.59	7.37
3	0.0	202.12	48.79	194.24	194.16	193.74	185.39	139.75	33.47	134.29	134.23	133.94	128.16	201.62	48.28	193.73	193.66	193.24	184.89
	0.1	132.96	5.00	8.10	8.10	8.15	7.94	91.82	3.10	5.26	5.26	5.30	5.15	132.46	4.47	7.58	7.58	7.64	7.43
	0.2	58.89	1.07	1.09	1.09	1.10	1.09	40.47	0.25	0.28	0.28	0.29	0.28	58.39	0.27	0.32	0.32	0.33	0.32
	0.3	26.75	1.02	1.02	1.02	1.02	1.02	18.20	0.17	0.18	0.18	0.19	0.18	26.25	0.13	0.16	0.16	0.16	0.16
	0.4	13.56	1.01	1.01	1.01	1.01	1.01	9.05	0.15	0.16	0.16	0.16	0.16	13.05	0.10	0.12	0.12	0.12	0.12
	0.5	7.69	1.01	1.01	1.01	1.01	1.01	4.98	0.14	0.15	0.15	0.15	0.15	7.18	0.08	0.10	0.10	0.10	0.10

TABLE 3. Performance of the BRCC and CUSUM-BRCC considering the different residuals with $\alpha = 0.005$ for Scenarios 4, 5, and 6.

Scenario	δ	BRCC	CUSUM-BRCC					BRCC	CUSUM-BRCC					BRCC	CUSUM-BRCC				
			Hwang	r_t^{ww}	r_t^{w}	r_t^{s}	r_t^{q}		Hwang	r_t^{ww}	r_t^{w}	r_t^{s}	r_t^{q}		Hwang	r_t^{ww}	r_t^{w}	r_t^{s}	r_t^{q}
		$\widehat{\text{ARL}}$						$\widehat{\text{MRL}}$						$\widehat{\text{SDRL}}$					
	−0.5	16.28	1.01	1.02	1.02	1.02	1.02	10.93	0.15	0.17	0.17	0.17	0.17	15.77	0.11	0.13	0.13	0.13	0.13
	−0.4	27.90	1.02	1.03	1.03	1.03	1.03	18.99	0.17	0.19	0.19	0.19	0.19	27.39	0.14	0.16	0.16	0.17	0.16
	−0.3	49.82	1.05	1.06	1.06	1.07	1.06	34.19	0.22	0.24	0.24	0.25	0.25	49.32	0.22	0.26	0.26	0.27	0.26
	−0.2	91.15	1.59	1.80	1.80	1.95	1.83	62.84	0.70	0.86	0.86	0.97	0.88	90.65	0.96	1.20	1.20	1.36	1.23
	−0.1	158.18	8.96	17.06	17.07	19.25	17.35	109.29	5.86	11.48	11.48	13.00	11.68	157.68	8.45	16.56	16.56	18.75	16.85
4	0.0	199.63	48.00	195.60	195.63	203.04	188.62	138.03	32.92	135.23	135.26	140.39	130.40	199.13	47.50	195.10	195.13	202.54	188.12
	0.1	156.68	14.54	34.15	34.17	32.46	32.10	108.26	9.72	23.32	23.33	22.15	21.90	156.18	14.03	33.64	33.66	31.96	31.59
	0.2	90.72	1.91	2.31	2.31	2.38	2.30	62.53	0.94	1.23	1.23	1.27	1.22	90.22	1.32	1.74	1.74	1.81	1.73
	0.3	49.71	1.05	1.07	1.07	1.07	1.07	34.11	0.23	0.25	0.25	0.26	0.25	49.21	0.23	0.27	0.27	0.28	0.27
	0.4	28.36	1.02	1.03	1.03	1.03	1.03	19.31	0.18	0.19	0.19	0.19	0.19	27.85	0.14	0.17	0.17	0.17	0.17
	0.5	17.01	1.01	1.02	1.02	1.02	1.02	11.44	0.16	0.17	0.17	0.17	0.17	16.50	0.11	0.13	0.13	0.13	0.13
		$\widehat{\text{ARL}}$						$\widehat{\text{MRL}}$						$\widehat{\text{SDRL}}$					
	−0.5	36.39	1.02	1.04	1.04	1.02	1.03	24.87	0.18	0.21	0.21	0.18	0.20	35.88	0.15	0.21	0.21	0.16	0.18
	−0.4	58.18	1.04	1.11	1.11	1.05	1.06	39.98	0.22	0.30	0.30	0.22	0.25	57.68	0.21	0.34	0.34	0.22	0.26
	−0.3	93.73	1.26	1.82	1.81	1.21	1.41	64.62	0.44	0.87	0.87	0.40	0.53	93.23	0.57	1.22	1.22	0.51	0.76
	−0.2	144.13	3.21	9.28	9.27	3.28	5.09	99.56	1.86	6.08	6.07	1.91	3.17	143.63	2.67	8.76	8.75	2.73	4.56
	−0.1	196.62	13.96	84.51	84.34	22.85	35.29	135.94	9.32	58.23	58.11	15.49	25.50	196.12	13.45	84.01	83.84	22.35	36.78
5	0.0	200.88	47.61	203.65	203.22	196.43	191.20	138.90	32.65	140.81	140.51	135.81	132.19	200.38	47.11	203.15	202.72	195.93	190.70
	0.1	152.88	27.87	49.52	49.43	377.42	65.30	105.62	18.97	33.98	33.91	261.26	44.92	152.38	27.37	49.02	48.93	376.92	64.80
	0.2	97.19	6.65	8.81	8.80	31.61	10.06	67.02	4.25	5.75	5.75	21.56	6.62	96.69	6.13	8.30	8.29	31.10	9.55
	0.3	59.42	1.78	1.96	1.95	3.03	1.98	40.84	0.84	0.97	0.97	1.73	0.99	58.92	1.18	1.37	1.37	2.48	1.39
	0.4	36.88	1.09	1.12	1.12	1.18	1.11	25.22	0.28	0.31	0.31	0.37	0.30	36.38	0.31	0.36	0.36	0.47	0.35
	0.5	24.23	1.03	1.04	1.04	1.06	1.04	16.45	0.20	0.22	0.22	0.25	0.22	23.72	0.19	0.21	0.21	0.26	0.21
		$\widehat{\text{ARL}}$						$\widehat{\text{MRL}}$						$\widehat{\text{SDRL}}$					
	−0.5	3.78	1.00	1.01	1.01	1.01	1.01	2.25	0.13	0.14	0.14	0.13	0.14	3.24	0.07	0.08	0.08	0.07	0.08
	−0.4	7.38	1.01	1.01	1.01	1.01	1.01	4.76	0.14	0.15	0.15	0.14	0.15	6.86	0.08	0.10	0.10	0.09	0.09
	−0.3	16.85	1.01	1.02	1.02	1.01	1.02	11.33	0.15	0.17	0.17	0.16	0.16	16.34	0.11	0.13	0.13	0.12	0.12
	−0.2	43.71	1.03	1.04	1.04	1.03	1.04	29.95	0.19	0.21	0.21	0.20	0.21	43.21	0.17	0.21	0.21	0.19	0.20
	−0.1	118.18	2.32	3.46	3.46	2.51	2.98	81.57	1.23	2.04	2.03	1.36	1.70	117.67	1.75	2.92	2.92	1.95	2.43
6	0.0	202.18	48.54	189.16	189.22	203.73	184.03	139.80	33.30	130.77	130.81	140.87	127.21	201.68	48.04	188.66	188.72	203.23	183.53
	0.1	101.56	2.80	3.43	3.43	4.91	3.69	70.05	1.57	2.01	2.01	3.05	2.19	101.05	2.25	2.89	2.89	4.38	3.15
	0.2	37.93	1.03	1.04	1.04	1.05	1.04	25.94	0.20	0.22	0.22	0.23	0.22	37.43	0.16	0.21	0.21	0.23	0.21
	0.3	16.02	1.01	1.02	1.02	1.02	1.02	10.75	0.16	0.17	0.17	0.18	0.17	15.51	0.11	0.13	0.13	0.14	0.13
	0.4	8.02	1.01	1.01	1.01	1.01	1.01	5.20	0.14	0.15	0.15	0.16	0.15	7.50	0.09	0.10	0.10	0.11	0.10
	0.5	4.68	1.01	1.01	1.01	1.01	1.01	2.88	0.13	0.14	0.14	0.15	0.14	4.15	0.07	0.08	0.08	0.10	0.09

at sample 16, while the BRCC took on average 163 samples to detect an out-of-control observation. The CUSUM-BRCC$_{\text{Hwang}}$ signaled at sample 8, however, this is not an accurate estimate since the control chart presented a distorted ARL for an in-control process. The exception is for the CUSUM-BRCC$_{r_t^s}$ in Scenarios 1 and 5 for $\delta = -0.1$ and $\delta = 0.1$, respectively, which yielded an $\widehat{\text{ARL}_1}$ greater than 200. This control chart tend to be ARL-biased in the sense that some out-of-control ARL values are larger than the in-control ARL (Paulino, Morais, and Knoth, 2016).

According to the results obtained, it is important to highlight that all proposed charts and the BRCC presented a similar performance when the process was in control. However, the performance of the proposed control chart was far superior when the process was out of control. For example, in Scenario 3 and $\delta = -0.1$, the BRCC took on average 131 samples to detect a change in the process while the proposed CUSUM-BRCC, considering all residuals, took on average 8 samples to detect a change (approximately 16 times faster). Comparing the CUSUM-BRCC using the quantile residual with the CUSUM-BRCC using the other residuals when the process shifted, we note from Tables 2 and 3 that the CUSUM-BRCC$_{r_t^q}$ outperformed in Scenarios 1, 2 and 4 for $\delta = 0.1$ while in the other scenarios its performance was quite similar to that of using the other residuals. For a negative shift, the CUSUM-BRCC$_{r_t^q}$ presented better performance in Scenario 5. Although the CUSUM-BRCC using the standardized residual performed better in some scenarios, the distribution of such residual is not well approximated by the standard normal distribution Espinheira, Ferrari, and Cribari-Neto, 2008 compared to its weighted competitors, also this control chart can be ARL-biased as mentioned before.

Note that the variability of the model is directly related to the ability of the control chart to detect changes in the process. For example, in Scenario 4, where we considered a low precision ($\phi \in [7, 54]$) and $\delta = -0.1$, the CUSUM-BRCC$_{r_t^q}$ signaled every 17 samples, on average. On the other hand, when the precision in the process was high ($\phi \in [55, 659]$, Scenario 6), the same control chart took on average 3 samples to detect a change of the same magnitude in the mean.

Finally, the numerical evidence showed that regardless of the value of δ and the scenario studied, the proposed CUSUM-BRCC presented a better behavior compared to the other control charts in the literature. As the quantile residual has proved itself to be a good residual for beta regressions (Pereira, 2019) and our simulation results suggest it performs better in some scenarios and equally in others, we recommend using such residual in the proposed CUSUM-BRCC.

4. Applications

In this section, we shall present and discuss two applications to show the applicability of the proposed control chart. As the CUSUM-BRCC$_{\text{Hwang}}$ presented poor perfomance under a controlled scenario, we do not consider this control chart in our applications. We performed the applications using the quantile residual for the CUSUM-BRCC and compare it with the BRCC. The construction of the control charts follows Algorithms 1 and 2 and $\text{ARL}_0 = 200$ ($\alpha = 0.005$) for both applications.

4.1. **Application to simulated data.** In the first application, we considered simulated data to better illustrate the studied methodology. We used the following structures for the data generation process:

$$\text{logit}(\mu_t) = -3.2 + 2x_{1t} + 2x_{2t}$$
$$\log(\phi_t) = 3.0 + z_{1t} + z_{2t},$$

where $(x_1, x_2) = (z_1, z_2)$. The values of x_1 (x_2) were obtained from a uniform distribution in the unit interval (Bernoulli distribution with $p = 0.3$). We generated 1000 observations considering the process in a state of control.

Table 4. Descriptive statistics of the quantitative variables; simulated data.

Variables	min	$Q_{1/4}$	median	mean	$Q_{3/4}$	max	sd
y	0.001	0.073	0.153	0.216	0.306	0.755	0.187
x_1	0.001	0.264	0.511	0.509	0.758	0.999	0.288

Table 5. Fitted beta regression model with varying precision; simulated data.

Submodel for μ			
	Estimate	SE	p-value
Intercept	−3.2190	0.0480	< 0.0001
x_1	2.0835	0.0591	< 0.0001
x_2	2.0026	0.0341	< 0.0001
Submodel for ϕ			
	Estimate	SE	p-value
Intercept	3.2316	0.1306	< 0.0001
z_1	0.6345	0.2123	0.0028
z_2	0.8498	0.1412	< 0.0001

Some descriptive statistics about y and x_1 are shown in Table 4, namely: minimum (min), first quartile ($Q_{1/4}$), median, mean, third quartile ($Q_{3/4}$), maximum (max), and standard deviation (sd). Descriptive statistics for x_2 are not presented because it is a binary covariate. Note that 25% of the response variable (y) does not exceed the value of 0.306, and the largest value is 0.755. The mean and median are 0.216 and 0.153, respectively. For x_1, we have a minimum of 0.001 and a maximum of 0.999.

The simulated data were split into two groups, Phase I (first 500 observations) and Phase II (last 500 observations). In Phase I, we estimated the submodels parameters, and in Phase II we monitored the process. Table 5 presents the parameter estimates, standard errors (SE), and p-values for the models adjusted in Phase I. We note that all covariates were significant in both models at 5%, as expected. Before determining the control limits, we performed a diagnostic analysis of the residuals of the model fitted in Phase I. Figures 1 shows the quantile residuals and Quantile-Quantile (Q-Q) plot for the fitted model in Table 5. Figure 1a shows that the residuals are randomly distributed around zero and within three deviations from the mean. Similarly, Figure 1b suggests that there is no violation of the model assumptions as the residuals are in agreement with the 45 degree line. The next step was the calibration of both control charts to have a target ARL_0, then in Phase II (last 500 observations) we introduced a perturbation in the structure of the mean submodel of magnitude $\delta = 0.3$ to assess the power of the chart in detecting changes within a controlled scenario.

Figure 2 presents the BRCC and CUSUM-BRCC$_{r_t^q}$ with Phase II data. The BRCC indicated only one out-of-control point below the lower limit and one out-of-control point above the upper limit while the CUSUM-BRCC$_{r_t^q}$ indicated 494 out-of-control observations (almost all the data in Phase II). It is

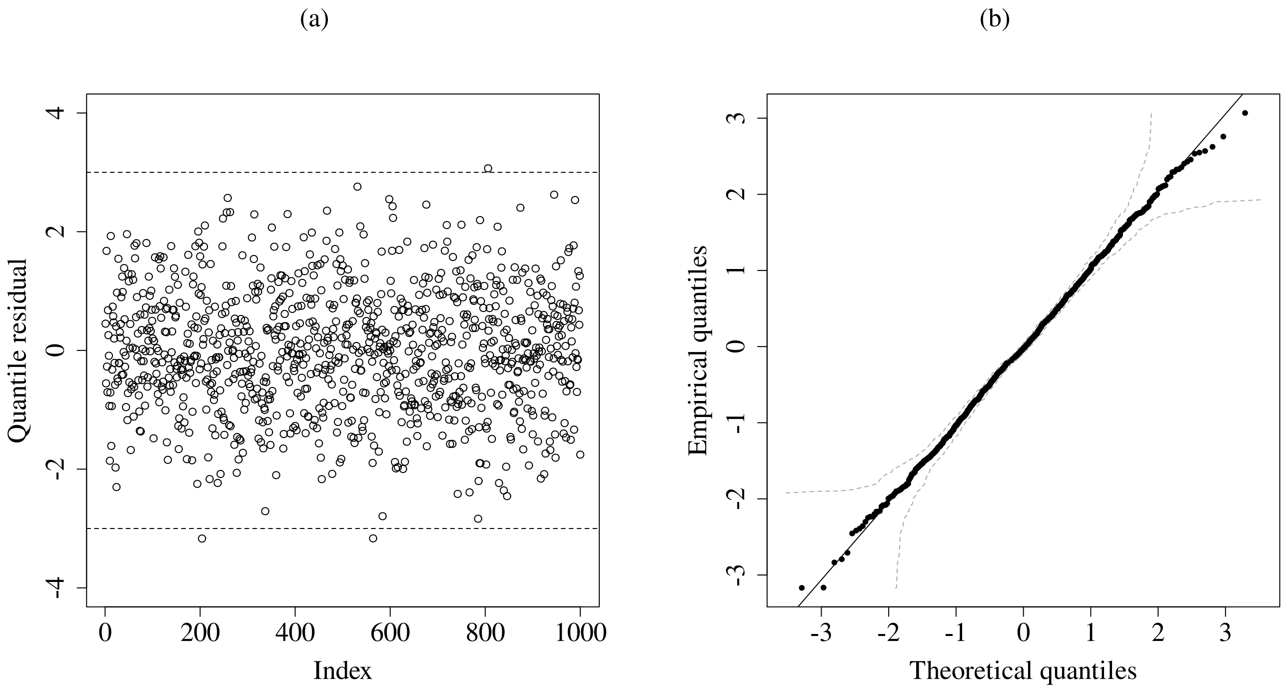

FIGURE 1. Quantile residual (a) and Q-Q plot (b); simulated data.

noteworthy that in this phase a disturbance of $\delta = 0.3$ was introduced in the process mean, that is, the process was out of control.

Figure 3 presents the performance of the BRCC and CUSUM-BRCC$_{r_t^q}$ with Phase II data. The values were obtained considering the steps of Algorithms 1 and 2. Lastly, the covariates and parameters described in Table 5 were used in this evaluation. We notice that, as evidenced in the simulation study, the CUSUM-BRCC$_{r_t^q}$ presented better results than the BRCC (smaller ARL$_1$), proving that the proposed control chart is more sensitive to detect minor changes in the manufacturing process.

4.2. **Empirical application.** In the second application, the dataset refers to the relative humidity (RH) in Australia and highlights the relevance of the proposed chart in monitoring double bounded environmental data. The RH is a ratio between continuous numbers, being the ratio of the partial pressure of water to the equilibrium vapor pressure of water, assuming values in $(0, 1)$. Due to the genesis of the beta regression model, rates and proportions usually can be well fitted by this model. Additionally, the monitoring of RH is relevant because it exerts influence on temperature, rain, and thermal sensation (Lima-Filho and Bayer, 2021).

Table 6 contains a brief description of the variables used in this analysis. The quality characteristic monitored was measured daily at 3:00pm, and the other variables were used to adjust the beta regression model for the μ and/or ϕ structures. This dataset is available in the `R` `rattle` package (Graham, 2011) from October 2010 to June 2017 for the Sydney Station in Australia.

Table 7 includes descriptive statistics of the considered variables. We observe that 25% of the RH does not exceed 43%, the largest prevalence of RH is 95%, and the mean and median values for RH are 54% and 53%, respectively. In the analyzed period, the lowest temperature was 5°C and the highest was 45.8°C.

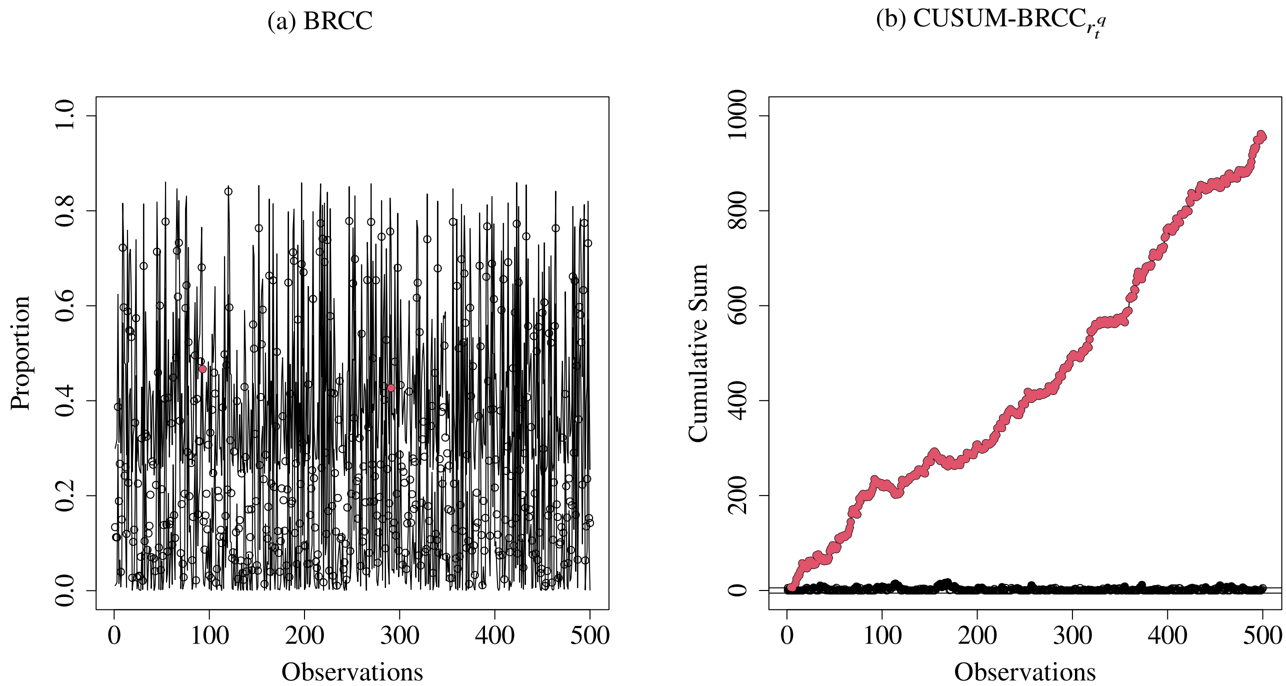

FIGURE 2. BRCC and CUSUM-BRCC$_{r_t^q}$ for simulated data with out-of-control observations highlighted in red.

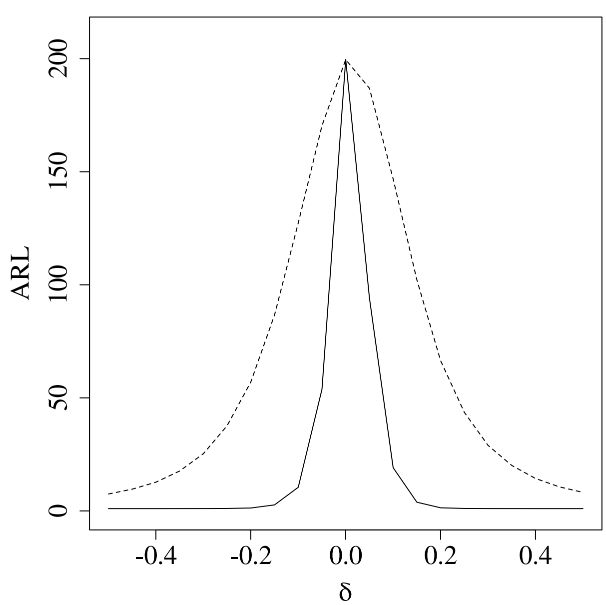

FIGURE 3. Performance of the BRCC (dashed line) and CUSUM-BRCC$_{r_t^q}$ (solid line) considering $\mathrm{ARL}_0 = 200$; simulated data.

TABLE 6. Description of the variables; RH data.

Variable	Description
RelHumid3pm	Relative humidity (%) at 3 pm
Cloud3pm	Fraction of sky obscured by cloud at 3 pm.
Evaporation	The so-called Class A pan evaporation (mm) in the 24 hours to 9 am
MaxTemp	The maximum temperature in degrees celsius
MinTemp	The minimum temperature in degrees celsius
Pressure3pm	Atmospheric pressure reduced to mean sea level at 3 pm
Rainfall	The amount of rainfall recorded for the day in mm
Sunshine	The number of hours of bright sunshine in the day

TABLE 7. Descriptive statistics of the variables; RH data.

Variables	min	$Q_{1/4}$	median	mean	$Q_{3/4}$	max	sd
RelHumid3pm	10.00	43.00	54.00	53.02	63.00	95.00	15.99
Cloud3pm	0.00	1.00	4.00	4.15	7.00	8.00	2.61
Evaporation	0.00	3.20	5.00	5.39	7.20	18.40	2.85
MaxTemp	11.70	20.20	23.30	23.45	26.40	45.80	4.48
MinTemp	5.00	11.30	15.05	15.03	18.90	27.10	4.52
Pressure3pm	994.00	1012.00	1016.00	1016.00	1021.00	1036.00	7.06
Rainfall	0.00	0.00	0.00	2.83	1.00	94.40	8.23
Sunshine	0.00	4.60	8.40	7.40	10.30	13.60	3.75

The dataset has a total of 1690 observations. We used the 845 (50%) first observations (October 2010 to December 2014) to estimate the submodels parameters (Phase I). In Phase II, we used the observations from January 2015 to June 2017 to monitor the relative humidity.

Table 8 shows the parameter estimates, SE, and p-values for the fitted beta regression model with varying precision. We considered the logit and log link functions in the mean and precision submodels, respectively. Considering the covariates that were statistically significant at the significance level of 5%, two of them were significant in both submodels, namely: minimum temperature (MinTemp) and atmospheric pressure reduced to mean sea level (Pressure3pm).

As in the previous application, we performed a diagnostic analysis of the residuals to check if the beta model is a good fit to the data. Figure 4 shows the residuals and Q-Q plot for the model fitted to RH data. In Figure 4a, the residuals are randomly distributed around zero and within three deviations from the mean. Figure 4b displays theoretical quantiles against the empirical quantiles of the residuals. There is no evidence of violation of the model assumptions as the residuals are mostly on the 45 degree line, indicating that this model is a good fit to the data.

Figure 5 shows the BRCC and CUSUM-BRCC$_{r_t^q}$ with Phase II data. Considering the BRCC for monitoring relative humidity, the control chart indicated no more than ten out-of-control points below the lower limit and three points exceeded the upper limit. Differently, the CUSUM-BRCC$_{r_t^q}$ triggered 62

Table 8. Parameter estimates, standard errors (SE), and p-values for the fitted beta regression model with varying precision; RH data.

Submodel for μ			
	Estimates	SE	p-value
Intercept	−20.6930	2.6299	< 0.0001
Cloud3pm	0.0657	0.0062	< 0.0001
Evaporation	−0.0509	0.0066	< 0.0001
MaxTemp	−0.0729	0.0062	< 0.0001
MinTemp	0.1161	0.0060	< 0.0001
Pressure3pm	0.0204	0.0025	< 0.0001
Rainfall	0.0121	0.0026	< 0.0001
Submodel for ϕ			
	Estimates	SE	p-value
Intercept	−60.4090	7.6854	< 0.0001
MinTemp	0.0284	0.0121	0.0190
Sunshine	0.0765	0.0129	< 0.0001
Pressure3pm	0.0615	0.0075	< 0.0001

(a) (b)

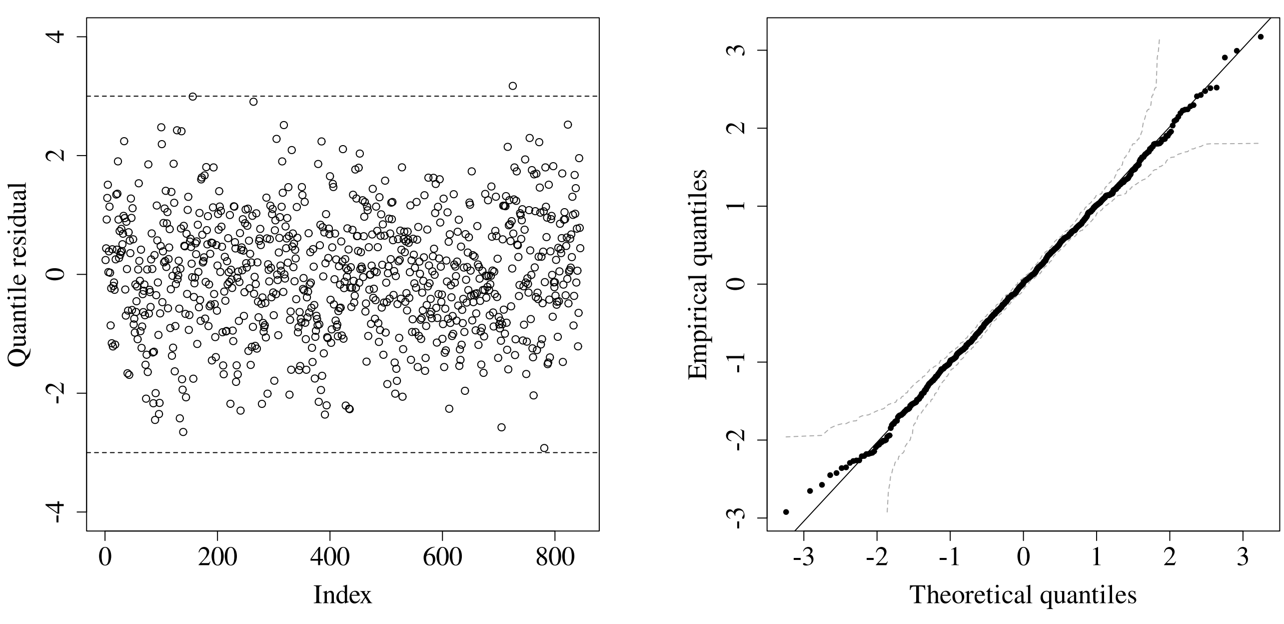

Figure 4. Quantile residual (a) and Q-Q plot (b); RH data.

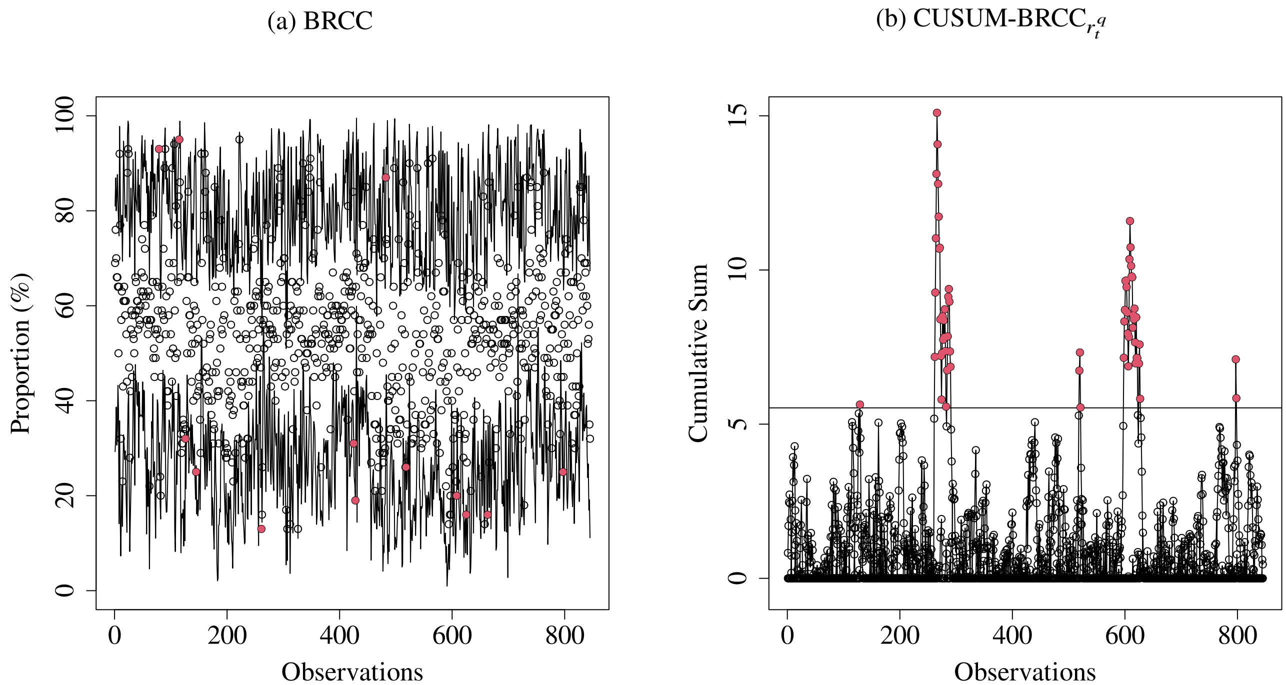

FIGURE 5. BRCC and CUSUM-BRCC$_{r_t^q}$ for the monitoring of relative humidity in Australia with out-of-control observations highlighted in red.

out-of-control points. These results reinforce the characteristic of the CUSUM control chart's power to detect changes in the quality characteristic of interest.

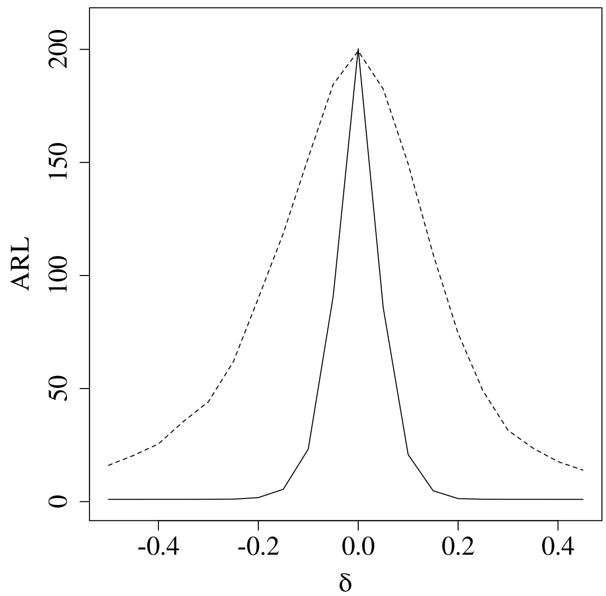

Figure 6. Performance of the BRCC (dashed line) and CUSUM-BRCC$_{r_t^q}$ (solid line) considering $\text{ARL}_0 = 200$; RH data.

Figure 6 shows the performance of the BRCC and CUSUM-BRCC$_{r_t^q}$. The construction of the control charts followed the same steps of Application 1. The control charts obtained similar performance when the process was in control ($\delta = 0$); however their performance differed when the process was out of control. The CUSUM-BRCC$_{r_t^q}$ presented smaller values of $\widehat{\text{ARL}}_1$ than its counterpart, evidencing that the proposed control chart is more sensitive to trigger a signal in the quality characteristic. For example, considering $\delta = 0.1$, the BRCC took on average 149 samples to detect a change in the process while the CUSUM-BRCC$_{r_t^q}$ took on average 20 samples to detect a change of the same magnitude. In a nutshell, in the presence of control variables, this analysis shows that the proposed CUSUM-BRCC is useful to monitor quality characteristics in the interval $(0, 1)$ in practical situations.

5. Concluding remarks

In this paper, we developed a new control chart for monitoring double bounded quality characteristics in the presence of control variables (covariates). For this purpose, we proposed the residual-based CUSUM beta regression control chart considering different residuals of the beta distribution. This control chart has the advantage of accumulating information from the past as well as being more sensitive to detect changes in the mean of a process. We conducted a Monte Carlo simulation study to evaluate and compare the performance of the proposed control chart with two competing control charts in the literature. The numerical results evidenced the superiority of the proposed control chart, presenting values of $\widehat{\text{ARL}}_0$, $\widehat{\text{MRL}}_0$, and $\widehat{\text{SDRL}}_0$ close to their nominal values when the process was in control, and smaller $\widehat{\text{ARL}}_1$ for an out-of-control process. We also presented and discussed applications to real and simulated data that showed the practical importance of our proposal. Finally, we suggest the use of the CUSUM beta regression control chart with the quantile residual when the objective is to monitor double bounded quality characteristics in the presence of control variables and detect small shifts in the mean of the process.

Acknowledgements

The authors gratefully acknowledge partial financial support from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Brazil.

References

Adegoke, N. A., A. N. H. Smith, M. J. Anderson, R. A. Sanusi, and M. D. M. Pawley (2019). Efficient homogeneously weighted moving average chart for monitoring process mean using an auxiliary variable. *IEEE Access* 7, 94021–94032.

Albarracin, O. Y. E., A. P. Alencar, and L. Lee Ho (2018). CUSUM chart to monitor autocorrelated counts using negative binomial GARMA model. *Statistical Methods in Medical Research* 27, (9), 2859–2871.

Alencar, A. P., L. Lee Ho, and O. Y. E. Albarracin (2017). CUSUM control charts to monitor series of negative binomial count data. *Statistical Methods in Medical Research* 26, (4), 1925–1935.

Anholeto, T., M. C. Sandoval, and D. A. Botter (2014). Adjusted Pearson residuals in beta regression models. *Journal of Statistical Computation and Simulation* 84, (5), 999–1014.

Asadzadeh, S., A. Aghaie, and S. T. A. Niaki (2013). AFT regression-adjusted monitoring of reliability data in cascade processes. *Quality & Quantity* 47, (6), 3349–3362.

Aslam, M., M. Azam, and C.-H. Jun (2014). A new exponentially weighted moving average sign chart using repetitive sampling. *Journal of Process Control* 24, (7), 1149–1153.

Aytaçoğlu, B., A. R. Driscoll, and W. H. Woodall (2021). Controlling the conditional false alarm rate for the MEWMA control chart. *Journal of Quality Technology*, 1–16.

Bayer, F. M., C. M. Tondolo, and F. M. Müller (2018). Beta regression control chart for monitoring fractions and proportions. *Computers & Industrial Engineering* 119, 416–426.

Benjamim, M. A., R. A. Rigby, and M. Stasinopoulos (2003). Generalized autoregressive moving average models. *Journal of the American Statistical Association* 98, (461), 214–223.

Chen, H. and C. Huang (2014). The use of a CUSUM residual chart to monitor respiratory syndromic data. *IIE Transactions* 46, (8), 790–797.

Cribari-Neto, F. and A. Zeileis (2010). Beta regression in R. *Journal of Statistical Software* 34, 1–24.

Crowder, S. V. (1989). Design of exponentially weighted moving average schemes. *Journal of Quality technology* 21, (3), 155–162.

Dunn, P. K. and G. K. Smyth (1996). Randomized quantile residuals. *Journal of Computational and Graphical Statistics* 5, (3), 236–244.

Espinheira, P. L., S. L. P. Ferrari, and F. Cribari-Neto (2008). On beta regression residuals. *Journal of Applied Statistics* 35, (4), 407–419.

Ewan, W. D. (1963). When and how to use CUSUM charts. *Technometrics* 5, (1), 1–22.

Ferrari, S. L. P. and F. Cribari-Neto (2004). Beta regression for modelling rates and proportions. *Journal of Applied Statistics* 31, (7), 799–815.

Fournier, B., N. Rupin, M. Bigerelle, D. Najjar, and A. Iost (2006). Application of the generalized lambda distributions in a statistical process control methodology. *Journal of Process Control* 16, (10), 1087–1098.

Gan, F. F. (1991). An optimal design of CUSUM quality control charts. *Journal of Quality Technology* 23, (4), 279–286.

Graham, J. W. (2011). *Data Mining with Rattle and R: The art of excavating data for knowledge discovery*. Springer Science & Business Media.

Haq, A. (2017). New synthetic CUSUM and synthetic EWMA control charts for monitoring the process mean using auxiliary information. *Quality and Reliability Engineering International* 33, (7), 1549–1565.

Hawkins, D. M. (1981). A CUSUM for a scale parameter. *Journal of Quality Technology* 13, (4), 228–231.

Hwang, W-Y. (2021). Deviance residual-based control charts for monitoring the beta-distributed processes. *Quality and Reliability Engineering International*.

Jones, M. C. (2009). Kumaraswamy distribution: A beta-type distribution with some tractability advantages. *Statistical Methodology* 6, (1), 70–81.

Kim, H. and S. Lee (2021). On residual CUSUM statistic for PINAR(1) model in statistical design and diagnostic of control chart. *Communications in Statistics-Simulation and Computation* 50, (5), 1290–1314.

Lee Ho, L., F. H. Fernandes, and M. P. Bourguignon (2019). Control charts to monitor rates and proportions. *Quality and Reliability Engineering International* 35, (1), 74–83.

Lima-Filho, L. M. A. and F. M. Bayer (2021). Kumaraswamy control chart for monitoring double bounded environmental data. *Communications in Statistics-Simulation and Computation* 50, (9), 2513–2528.

Lima-Filho, L. M. A., T. L. Pereira, T. C. Souza, and F. M. Bayer (2019). Inflated beta control chart for monitoring double bounded processes. *Computers & Industrial Engineering* 136, 265–276.

Lima-Filho, L. M. A., T. L. Pereira, T. C. Souza, and F. M. Bayer (2020). Process monitoring using inflated beta regression control chart. *PLOS ONE* 15, (7), 1–20.

Lucas, J. M. (1976). The design and use of V-mask control schemes. *Journal of Quality Technology* 8, (1), 1–12.

Lucas, J. M. and M. S. Saccucci (1990). Exponentially weighted moving average control schemes: properties and enhancements. *Technometrics* 32, (1), 1–12.

Montgomery, D. C. (2009). *Introduction to Statistical Quality Control*. John Wiley & Sons.

Page, E. S. (1954). Continuous inspection schemes. *Biometrika* 41, (1/2), 100–115.

Page, E. S. (1961). Cumulative sum charts. *Technometrics* 3, (1), 1–9.

Park, J. and C.-H. Jun (2015). A new multivariate EWMA control chart via multiple testing. *Journal of Process Control* 26, 51–55.

Paulino, S., M. C. Morais, and S. Knoth (2016). An ARL-unbiased c-chart. *Quality and Reliability Engineering International* 32, (8), 2847–2858.

Pereira, G. H. A. (2019). On quantile residuals in beta regression. *Communications in Statistics-Simulation and Computation* 48, (1), 302–316.

Perry, M. B. and Z. Wang (2020). A distribution-free joint monitoring scheme for location and scale using individual observations. *Journal of Quality Technology*, 1–18.

R Core Team (2021). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria.

Roberts, S. W. (1959). Control chart tests based on geometric moving averages. *Technometrics* 1, (3), 239–250.

Sant'Anna, A. M. O. and C. S. ten Caten (2012). Beta control charts for monitoring fraction data. *Expert systems with applications* 39, (11), 10236–10243.

Sanusi, R. A., N. Abbas, and M. Riaz (2018). On efficient CUSUM-type location control charts using auxiliary information. *Quality Technology & Quantitative Management* 15, (1), 87–105.

Shewhart, W. A. (1931). *Economic Control of Quality of Manufactured Product*. Macmillan And Co Ltd, London.

Siegmund, David (1985). *Sequential Analysis: Tests and Confidence Intervals*. Springer Science & Business Media.

Simas, A. B., W. Barreto-Souza, and A. V. Rocha (2010). Improved estimators for a general class of beta regression models. *Computational Statistics & Data Analysis* 54, (2), 348–366.

Smithson, M. and J. Verkuilen (2006). A better lemon squeezer? Maximum-likelihood regression with beta-distributed dependent variables. *Psychological Methods* 11, (1), 54.

Weiß, C. H. and M. C. Testik (2015). Residuals-based CUSUM charts for Poisson INAR(1) processes. *Journal of Quality Technology* 47, (1), 30–42.

Woodall, W. H. and B. M. Adams (1993). The statistical design of CUSUM charts. *Quality Engineering* 5, (4), 559–570.

Xue, L. and P. Qiu (2021). A nonparametric CUSUM chart for monitoring multivariate serially correlated processes. *Journal of Quality Technology* 53, (4), 396–409.