Chilean Journal of Statistics
Vol. 17, No. 1, June/July 2026, 88–106
DOI: 10.32372/ChJS.17-01-05

Time-Series Models
Research Article

Bias reduction for improved inference and diagnostic analysis in non-stationary Poisson INARX models with time-varying covariates

Orlando Y.E. Albarracin[1,2,*] and Airlane P. Alencar[1]
[1]Institute of Mathematics and Statistics, Universidade de São Paulo, São Paulo, Brazil
[2]Universidade Presbiteriana Mackenzie, São Paulo, Brazil

(Received: 14 May 2026 · Accepted in final form: 25 June 2026)

Abstract

Integer-valued autoregressive (INAR) models describe time series of counts, such as disease incidence and mortality data. Health-related count series exhibit time-varying means and variances that cannot be easily removed by transformations. In this article, the conditional maximum likelihood (CML) method is introduced to estimate the parameters of INAR processes with explanatory variables (INARX) in the Poisson innovation process, capturing trends, seasonality, or covariate effects in non-stationary series. A simulation study compares CML with weighted conditional least squares (WCLS) in terms of bias, root mean square error, and computational cost. A bootstrap-based bias correction is proposed to address estimator bias caused by the correlation between innovation and autocorrelation terms. Furthermore, diagnostic analysis is enhanced by introducing two new residuals that account for covariates, heteroscedasticity, and autocorrelation, alongside three bootstrap methods for parameter confidence intervals and multivariate evaluation. We recommend the faster WCLS method combined with randomized residual bootstrap for bias correction and inference. These methods are applied to model weekly respiratory hospitalizations among the elderly in São Paulo (2011–2017) using a Poisson INARX(4) model that effectively captures the serial correlation and key features of the data.

Keywords: Count time series · INAR · Covariates · Trends · Seasonality

Mathematics Subject Classification: Primary 62M10 · Secondary 62F40

1. Introduction

In recent years, the interest in developing models for integer-valued time series has increased and different approaches were motivated by regression models. Although generalized linear models (GLM) [1] are widely used as a tool to explain the mean of a response variable as a function of explanatory variables, they are not appropriate to model count time series because a key assumption made by these models is the independence of observations. However, autocorrelation is usually present in time series and may have a strong impact on parameter estimators and their variance. Several models have been proposed to make statistical inference for integer-valued time series, designed for different types of marginal distribution and autocorrelation structure.

*Corresponding author. Email: orlandoyesid.8@gmail.com (O.Y.E. Albarracin)

A popular approach is to adapt the GLM methodology by choosing a suitable distribution for count data, selecting an appropriate link function, and modeling the observations conditionally on past information of the process and/or any other auxiliary information. An overview of the main approaches can be found in [2, 3]. Some discussed models are the integer-valued GARCH model (INGARCH), where each conditional observation given the past is Poisson distributed [4], the generalized autoregressive moving average (GARMA) model proposed in [5], which was developed to extend the GLM incorporating autoregressive (AR) and moving average (MA) terms as presented in the ARMA models [6], and the generalized linear autoregressive moving average (GLARMA) model described in [7, 8], which are nonlinear state space models with a state process dependent on covariates and past observations. The observations, conditional on the state process, are independent and follow an exponential family distribution. Other extensions of the time-series modeling framework can be found in [9, 10, 11, 12, 13, 14, 15, 16, 17].

An important class of models for time series of counts is the class of integer autoregressive (INAR) models, which are based on a thinning operator and a discrete innovation process, which usually is specified as a Poisson process. This class of models was independently proposed in [18, 19]. The thinning operator makes the present response depend on the previous count. The innovation process is defined as a sequence of independent, identically distributed (IID) non-negative integer-valued random variables with constant mean and variance. The correlation structure and stationary conditions of INAR models are similar to those of the continuous-valued AR process, making INAR models more attractive and convenient for modeling count time series. Recently, in [20], two extensions of the first-order INAR process with Poisson innovations were presented for modeling stationary count time series with equidispersion, underdispersion, and overdispersion. Other proposals based on INAR models can be found in [21, 22, 23, 24].

In practice, many count time series present trend and seasonality that cannot be easily removed by transformations since any transformation would need to preserve the integer nature of the data. Also, the objective may be the estimation of trend, seasonality, and covariate effects. To extend the applicability of INAR models, several authors have included explanatory variables in these models. In [25], explanatory variables were included in an INAR(1) model, where one first set of covariates affects the parameter in the thinning operator and another set affects the mean of the innovation process. The unknown model parameters were estimated using the weighted and unweighted conditional least squares (CLS) methods and the conditional generalized method of moments (GMM). In simulation studies, the CLS estimators outperformed the GMM estimators in short time series [25]. In [26], the same set of explanatory variables was included to explain the parameters in the thinning operator and in the mean of the innovation process of an INAR(p) model. The methodology was illustrated by analyzing two real count time series and introducing a reversible jump MCMC algorithm to choose the order of the INAR model. In [27], explanatory variables were incorporated only into the parameter of the thinning operator of an INAR(1) model, the parameters were estimated by the empirical maximum likelihood (ML) method, and the asymptotic properties of the estimators were established. Simulation studies and an empirical application were used to evaluate and illustrate their approach.

In this article, we introduce the conditional maximum likelihood (CML) method for estimating the INAR(p) parameters in the Poisson innovation process —INARX(p)— to model non-stationary count processes. We compare the CML and the weighted conditional least squares (WCLS) methods through a simulation study in terms of bias and the root of the mean square error (RMSE). The WCLS method was evaluated before only for the INARX(1) model presented in [25]. The processing time for both estimation methods is also discussed. A strategy to choose the initial values is presented since the estimates are too dependent on the initial values used in the parameter estimation algorithm.

Another relevant contribution of this article is the proposal of the uncorrelated standardized and randomized residuals, which consider the inclusion of covariates, heteroscedasticity, and autocorrelation of the INARX(p) process. We present a complete residual analysis and different bootstrap methods based on the proposed and usual standardized conditional residuals, allowing the assessment of the multivariate distribution of all estimators, including the correlation between the estimators of the innovation process and the thinning parameters.

Note that, recently, in [28], a regenerative bootstrap method was developed to construct confidence intervals for the parameters of an INAR(p) process with innovations following a zero-inflated distribution without covariates. Another bootstrap scheme was discussed in [29], where a general INAR-type bootstrap procedure is proposed to discuss parametric and semi-parametric implementations of the INAR bootstrap scheme without covariates. In addition, a bias correction for the estimators of stationary INAR processes with and without parametrically specifying the distribution of the innovation was proposed in [30]. In general, the bias correction proposal works well in simulations. However, when the Poisson innovation assumptions are violated, the parametric INAR bootstrap and the asymptotic bias correction performed worse. A further contribution of this article is the proposal of a bootstrap-based bias correction method for INARX(p) models, to deal with the correlation between the Poisson innovation term and the autocorrelation terms, as identified during the simulation studies.

The present article unfolds as follows. In Section 2, we introduce the model, the estimation of its parameters, and three bootstrap procedures based on different residuals. In Section 3, we present the simulation studies and propose the bias-corrected estimators based on the bias identified in the simulations. Section 4 provides an empirical application using the time series of the weekly number of hospitalizations of elderly people in São Paulo city (Brazil) from 2011 to 2017. In Section 5, we discuss some aspects of the estimation methods and summarize the main conclusions.

2. INAR model

2.1 *The model*

Consider a non-negative integer-valued random variable X and $\{\varepsilon_t, t \in \mathbb{N}\}$ a sequence of uncorrelated non-negative integer-valued random variables with finite variance. An INAR(p) process is defined by

$$X_t = \sum_{i=1}^{p} \alpha_i \circ X_{t-i} + \varepsilon_t, \tag{1}$$

where $\{\varepsilon_t\}$ is assumed to be independent of all thinning operations $\alpha_i \circ X_{t-i}$, for $i = 1, \ldots, p$, which are in turn conditionally independent. The binomial thinning operator "$\circ$" is the main concept of INAR processes and defined by $\alpha \circ X = \sum_{i=1}^{X} Y_i \sim \text{Binomial}(X, \alpha)$, that is, Binomial$(X, \alpha)$ denotes a binomial distribution with X trials and success probability $\alpha \in [0, 1)$ and Y_i are IID Bernoulli random variables, independent of X, that are equal to zero or one with probability $1 - \alpha$ and α, respectively.

We assume that, as usual in the literature, $\{\varepsilon_t\}$ is a sequence of independent Poisson distributed random variables with parameter λ and independent of previous counts $X_1, \ldots, X_{t-1}$. In this case, the mean of the X_t process is equal to $\lambda/(1 - \sum_{i=1}^{p} \alpha_i)$, and the variance is equal to the mean only when $p = 1$.

The autocorrelation at lag k for an INAR(1) process is α_1^k, which is restricted to be positive (more details can be found in [18]). The correlation structure of the INAR(p) process is identical to the corresponding AR(p) structure under stationarity and the stationarity of the process given in (1) is valid if $\sum_{i=1}^{p} \alpha_i < 1$ [31]. The parameters $\alpha_1, \ldots, \alpha_p$ are called thinning or AR parameters because they are present in the thinning operator and dictate the AR correlation structure.

The main methods used to estimate the unknown parameters of the stationary INAR(p) process are the CLS, the method of moments (MM), and the ML. In the CLS, no specific distributional assumption is required [32]. In the MM, the parameter estimation is based on the sample autocorrelations using, for example, Yule–Walker equations [33]. In the ML, the estimates are usually obtained from an iterative algorithm in which the initial values are usually obtained using the MM estimators [34]. In [35], details of these three estimation methods were provided. The ML method is more difficult to implement as the order p increases. When the innovation terms ε_t defined in (1) follow a Poisson distribution, with constant mean, the CLS and ML estimators were evaluated in [18, 31, 36]. The authors concluded that the ML estimator is more efficient in terms of both the bias and mean square error.

To make INAR models more attractive for real applications, [25, 27] included explanatory variables in an INAR(1) model. In the first article, covariates are included to model the AR parameter or the mean of the innovation process. Differently, in the second article, the authors proposed the inclusion of explanatory variables only to model the AR parameter. All model parameters were estimated by the weighted and unweighted CLS and conditional GMM in [25] and by the empirical ML method in [27]. In [26], the same set of explanatory variables was included to model both the AR parameters and the mean of the innovation process of an INAR(p) model, and a reversible jump MCMC algorithm was introduced to choose the order of the INAR model and the inclusion/exclusion of covariates. Note that it is difficult to employ reversible jump MCMC to compare different classes of models [37].

In our article, explanatory variables are included in the INAR(p) model only in the innovation process $\{\varepsilon_t\}$ defined in (1), which follows a Poisson distribution with parameter λ_t changing over time, to take into account, for example, the effects of seasonal variations and overall trends. As λ_t must be non-negative, we assume that

$$\lambda_t = \exp\left(\boldsymbol{w}_t^\top \boldsymbol{\beta}\right), \tag{2}$$

where $\boldsymbol{w}_t$ is a vector of known explanatory variables and $\boldsymbol{\beta}$ is the corresponding vector of unknown parameters. Then, a Po-INAR(p) model with explanatory variables incorporated in the mean of the innovation term, which follows a Poisson distribution, is given by expressions (1) and (2) and denoted henceforth as Po-INARX(p). In this case, the unconditional mean of X_t is given by

$$\mathrm{E}(X_t) = \sum_{j=1}^{p} \alpha_j \, \mathrm{E}(X_{t-j}) + \lambda_t. \tag{3}$$

Note that the marginal mean is decomposed into two parts: the effects of the previous means and the time-varying error mean. When $p = 1$, the unconditional mean and variance of X_t are given, respectively, by $\mathrm{E}(X_t) = \alpha \, \mathrm{E}(X_{t-1}) + \lambda_t$ and $\mathrm{Var}(X_t) = \alpha^2 \, \mathrm{Var}(X_{t-1}) + \alpha(1-\alpha)\mathrm{E}(X_{t-1}) + \lambda_t$.

Making successive substitutions, the mean of X_t can be written as

$$\mathrm{E}(X_t) = \alpha^t \mathrm{E}(X_0) + \sum_{u=0}^{t-1} \alpha^u \lambda_{t-u}.$$

When $p = 1$ and $\{\varepsilon_t\}$ follows a Poisson distribution, the marginal distribution of X_t is also Poisson provided that the initial specifications of $\mathrm{E}(X_0)$ and $\mathrm{Var}(X_0)$ are equal. When $p > 1$, the unconditional mean and variance of X_t are generally not equal [25]. The conditional moments of X_t are stated as

$$\begin{aligned} \mathrm{E}(X_t \mid X_{t-1}, \ldots, X_{t-p}) &= \sum_{j=1}^{p} \alpha_j X_{t-j} + \lambda_t, \\ \mathrm{Var}(X_t \mid X_{t-1}, \ldots, X_{t-p}) &= \sum_{j=1}^{p} \alpha_j (1-\alpha_j) X_{t-j} + \lambda_t. \end{aligned} \tag{4}$$

2.2 *Estimation*

Next, the weighted and unweighted CLS methods, which were implemented in [25] to estimate the parameters of an INAR(1) model with explanatory variables, and the CML method are generalized to estimate the parameters of the Po-INARX(p) model with time-varying λ_t in (2). In all methods, the initial values of the AR parameters (αs) are based on the partial autocorrelation function (PACF) of the standardized residuals calculated after fitting a Poisson GLM model with the same set of covariates and the logarithm as link function.

The CLS estimator of the parameter vector of the Po-INAR(p) model $\boldsymbol{\theta} = \left(\boldsymbol{\beta}^\top, \boldsymbol{\alpha}\right)^\top$ is obtained by minimizing the function $S(\boldsymbol{\theta})$ formulated as

$$S(\boldsymbol{\theta}) = \sum_{t=p+1}^{T} e_t^2, \tag{5}$$

where e_t is defined as the difference between X_t and the conditional expectation, $\mathrm{E}(X_t \mid X_{t-1}, \ldots, X_{t-p})$, defined in (4), which is the best one-step-ahead conditional predictor of X_t in the sense of minimizing the mean squared prediction error. Considering the Gauss–Newton algorithm to minimize $S(\boldsymbol{\theta})$, an iterative algorithm gives the new estimate $\boldsymbol{\theta}_{k+1}$ according to

$$\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_k - \left[\sum_{t=p+1}^{T} \frac{\partial e_t}{\partial \boldsymbol{\theta}} \frac{\partial e_t}{\partial \boldsymbol{\theta}'}\right]^{-1} \left[\sum_{t=p+1}^{T} \frac{\partial e_t}{\partial \boldsymbol{\theta}} e_t\right]\Bigg|_k,$$

where the gradient and the approximation to the Hessian matrix are evaluated at the previous estimate $\boldsymbol{\theta}_k$.

Note that the Po-INARX(p) process defined by (1) and (2) is a heteroskedastic process. Therefore, a more efficient CLS estimator that minimizes the squared standardized conditional errors is the WCLS.

Then, the WCLS estimator minimizes the expression presented as

$$S(\boldsymbol{\theta}) = \sum_{t=p+1}^{T} \frac{e_t^2}{\mathrm{Var}(e_t \mid X_{t-1}, \ldots, X_{t-p})}, \tag{6}$$

where $\mathrm{Var}(e_t \mid X_{t-1}, \ldots, X_{t-p}) = \sum_{i=1}^{p} \alpha_i(1-\alpha_i)X_{t-i} + \lambda_t$. The CLS estimate obtained by minimizing (5) can be used as the initial parameter vector to minimize the WCLS objective function stated in (6). An estimator of the asymptotic covariance matrix of $\widehat{\boldsymbol{\theta}}$ is given by

$$\mathrm{Cov}(\widehat{\boldsymbol{\theta}}) = \left[\sum_{t=p+1}^{T} c_t \frac{\partial e_t}{\partial \boldsymbol{\theta}} \frac{\partial e_t}{\partial \boldsymbol{\theta}'}\right]^{-1},$$

where $c_t = 1/\mathrm{E}(e_t^2 \mid X_{t-1}, \ldots, X_{t-p})$ and the gradient vector is $\sum_{t=p+1}^{T} c_t\, e_t\, \partial e_t / \partial \boldsymbol{\theta}$. Note that $\mathrm{E}(e_t^2 \mid X_{t-1}, \ldots, X_{t-p})) = \sum_{i=1}^{p} \alpha_i(1-\alpha_i)X_{t-i} + \lambda_t$. The asymptotic Gaussian distribution of the CLS and WCLS estimators is discussed in [38] and presented in the next theorem.

Theorem 1. Let the parameter space Θ be a subset of $\mathbb{R}^d$ and $\boldsymbol{\theta}$ denote a generic $d \times 1$ vector contained in Θ. Let X_t be a sequence of random variables, $\boldsymbol{w}_t \in \mathbb{R}^k$ a vector sequence, and $\mathrm{E}(X_t \mid \boldsymbol{r}_t) = m_t(\boldsymbol{r}_t; \boldsymbol{\theta}_0)$ the conditional expectation, where $\boldsymbol{r}_t = (\boldsymbol{w}_t, X_{t-1}, \boldsymbol{w}_{t-1}, \ldots, X_1, \boldsymbol{w}_1)$. Thus, letting $\boldsymbol{h}_t = (X_t, \boldsymbol{w}_t, X_{t-1}, \ldots, X_1, \boldsymbol{w}_1)$, the nonlinear least squares estimation of $\boldsymbol{\theta}_0$ such that $\mathrm{E}(X_t \mid \boldsymbol{r}_t) = m_t(\boldsymbol{r}_t, \boldsymbol{\theta}_0)$ takes $q_t(\boldsymbol{h}_t, \boldsymbol{\theta}) = (X_t - m_t(\boldsymbol{r}_t, \boldsymbol{\theta}))^2$. Then, under several assumptions stated in [38] and allowing for an estimator $\widehat{\boldsymbol{\gamma}}_T$, such that $\operatorname{plim} \widehat{\boldsymbol{\gamma}}_T = \boldsymbol{\gamma}^*$ for some $\boldsymbol{\gamma}^* \in \Gamma \subset \mathbb{R}^R$, the following results hold:

- (i) A random vector $\widehat{\boldsymbol{\theta}}_T$ exists that solves $\min_{\boldsymbol{\theta} \in \Theta} \left\{ \sum_{t=1}^{T} q_t(\boldsymbol{h}_t, \boldsymbol{\theta}; \widehat{\boldsymbol{\gamma}}_T) \right\}$ and $\widehat{\boldsymbol{\theta}}_T \xrightarrow{\mathrm{P}} \boldsymbol{\theta}_0$, where $\xrightarrow{\mathrm{P}}$ denotes convergence in probability to.
- (ii) $\sqrt{T}(\widehat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}_0) \xrightarrow{\mathrm{D}} \mathrm{N}(\mathbf{0}, \boldsymbol{A}_0^{-1})$, where $\xrightarrow{\mathrm{D}}$ denotes convergence in distribution to and $\boldsymbol{A}_0 = \lim_{T\to\infty} T^{-1} \sum_{t=1}^{T} \mathrm{E}(\boldsymbol{H}_t(\boldsymbol{\theta}_0; \boldsymbol{\gamma}^*))$.

The asymptotic variance of $\widehat{\boldsymbol{\theta}}_T$ can be estimated as $\widehat{\boldsymbol{A}}_T^{-1}/T$, where $\widehat{\boldsymbol{A}}_T$ is the consistent estimator of the expected Hessian. The proof can be found in [38, p. 2655]. The covariance $\boldsymbol{A}_0^{-1}$ corresponds to the WCLS estimator, whose weights are the conditional variance. For the unweighted CLS estimator, the asymptotic covariance takes the sandwich form $\boldsymbol{A}_0^{-1} \boldsymbol{B}_0 \boldsymbol{A}_0^{-1}$, with $\boldsymbol{B}_0$ the asymptotic variance of the gradient.

The CML estimator of the unknown parameter vector $\boldsymbol{\theta} = \left(\boldsymbol{\beta}^\top, \boldsymbol{\alpha}\right)^\top$ is obtained by maximizing the conditional log-likelihood function given by

$$l(\boldsymbol{\theta}) = \sum_{t=p+1}^{T} \log(P(X_t \mid X_{t-1}, \ldots, X_{t-p})), \tag{7}$$

where the conditional probabilities $P(X_t = x_t \mid X_{t-1} = x_{t-1}, \ldots, X_{t-p} = x_{t-p})$ can be written as the $(p+1)$-fold convolution [34]. When the innovation term ε_t stated in (1) is assumed to be Poisson, the conditional probability is expressed as

$$P(X_t = x_t \mid X_{t-1} = x_{t-1}, \ldots, X_{t-p} = x_{t-p})$$

$$= \sum_{i_1=0}^{\min\{x_{t-1},x_t\}} B(i_1, x_{t-1}, \alpha_1) \sum_{i_2=0}^{\min\{x_{t-2},x_t-i_1\}} B(i_2, x_{t-2}, \alpha_2) \cdots$$
$$\times \sum_{i_p=0}^{\min\{x_{t-p},x_t-i^*\}} B(i_p, x_{t-p}, \alpha_p) \frac{\exp(-\lambda_t)\lambda_t^{x_t-i^*-i_p}}{(x_t - i^* - i_p)!},$$

where $B(i,x,\alpha) = \binom{x}{i}\alpha^i(1-\alpha)^{x-i}$, $i^* = i_1 + \cdots + i_{p-1}$, and $x \in \mathbb{N}$, for $i = 0, 1, \ldots, x$. The CML method involves a convolution and can be difficult to implement as the order p increases [39] and for high values of observed counts due to each conditional probability term in (7), taking a long time for its computation and optimization. For $p = 1$, the conditional probability can be written as

$$P(X_t = x_t \mid X_{t-1} = x_{t-1}) = \sum_{i_1=0}^{\min\{x_{t-1},x_t\}} \binom{x_{t-1}}{i_1} \alpha_1^{i_1}(1-\alpha_1)^{x_{t-1}-i_1} \frac{\exp(-\lambda_t)\lambda_t^{x_t-i_1}}{(x_t - i_1)!}.$$

2.3 *Residuals and bootstrap procedures*

Next, we introduce three bootstrap methods to build confidence intervals for the parameters of the Po-INARX(p) models defined in (1) and (2). These bootstrap methods are based on three standardized residuals, which consider the structure of count time series with time-varying effects. The first residual proposed in this article is based on the fact that the autocorrelation structure of INAR models is similar to that of the AR process. The second residual depends on the binomial thinning operation and in this article it is generalized when the mean of the innovation process, defined in (2), changes over time. The third residual is based on the conditional moments of the process.

After computing the standardized residuals, a careful diagnostic analysis must evaluate the homoscedasticity and absence of autocorrelation of the standardized residuals. Only if the standardized residuals series behaves as a white noise process, the bootstrap can be implemented.

METHOD 1. The first bootstrap algorithm is based on the standardized residuals obtained after removing the time-varying effects, such as trend and/or seasonality, and then extracting the autocorrelation effects. Let X_t be a count time series. Our proposed algorithm consists of the following steps:

(i) Fit a Po-INARX(p) model defined by expressions stated in (1) and (2) to the data using the WCLS or CML method. Based on the estimates $\widehat{\boldsymbol{\alpha}}$ and $\widehat{\boldsymbol{\beta}}$, calculate the estimated unconditional mean of X_t, $\widehat{\mu}_t$, replacing the parameters defined in (3) by their corresponding estimates.

(ii) Compute the standardized residuals $\upsilon_t = (X_t - \widehat{\mu}_t)/\sqrt{\widehat{\mu}_t}$ to remove the time-varying effects on the data.

(iii) Remove the serial correlation of υ_t considering an autocorrelation structure of an AR(p) process, which consists of calculating the uncorrelated $r_{t,1}$ residuals established as

$$r_{t,1} = \upsilon_t - \sum_{i=1}^{p} \widehat{\alpha}_i \upsilon_{t-i}, \tag{8}$$

where $\widehat{\alpha}_i$, for $i = 1, \ldots, p$, are the estimated thinning parameters of the Po-INARX(p) model fitted in step i. Only after this step, the residuals are supposed to be permutable.

(iv) Generate bootstrap samples $\upsilon_1^+, \ldots, \upsilon_n^+$ according to $\upsilon_t^+ = \sum_{i=1}^{p} \widehat{\alpha}_i \upsilon_{t-i}^+ + r_{t,1}^+$, where $r_{t,1}^+$ are randomly and uniformly drawn with replacement from the residuals $r_{t,1}$ and the process is initialized assuming $r_{0,1}^+ = 0$. Then, generate the bootstrap observations by means of $X_t^+ = \widehat{\mu}_t + \upsilon_t^+ \sqrt{\widehat{\mu}_t}$.

Method 2. The second proposal is a bootstrap algorithm based on the residuals discussed in [40] for Po-INAR(p) models, where the innovation term follows a Poisson distribution with constant mean. Here, these residuals are adapted to consider the time-varying parameter (λ_t) defined in (2). The steps of the algorithm are as follows:

(i) Fit a Po-INARX(p) model defined by expressions stated as (1) and (2) to the data using the WCLS or CML method.

(ii) Compute the randomized residuals

$$r_{t,2} = X_t - \sum_{i=1}^{p} \widehat{\alpha}_i \circ X_{t-i} - \widehat{\lambda}_t, \tag{9}$$

where $\widehat{\alpha}_i$, for $i = 1, \ldots, p$, are the estimated thinning parameters of the Po-INARX(p) model fitted in the previous step and $\widehat{\lambda}_t$ is calculated as in (2) based on the estimates $\widehat{\boldsymbol{\beta}}$. Note that the residuals $r_{t,2}$ depend on the thinning operator that involves the sum of Bernoulli distributed random variables.

(iii) Generate the bootstrap observations according to

$$\widehat{X}_t^+ = \sum_{i=1}^{p} \widehat{\alpha}_i \circ X_{t-i}^+ + [\widehat{\lambda}_t + r_{t,2}^+],$$

where $r_{t,2}^+$ are randomly and uniformly drawn from the residuals $r_{t,2}$ in (9). The function $[x]$ denotes a rounding operator to ensure that x is an integer value.

Method 3. The third proposed bootstrap algorithm is based on the usual standardized conditional residuals computed using the conditional moments of X_t defined in (4).

(i) Fit a Po-INARX(p) model defined by expressions (1) and (2) to the data using the WCLS or CML method.

(ii) Based on the estimates $\widehat{\boldsymbol{\alpha}}$ and $\widehat{\boldsymbol{\beta}}$ and using the conditional moments given in (4), calculate the standardized conditional residuals

$$r_{t,3} = \frac{X_t - \widehat{\mathrm{E}}(X_t \mid X_{t-1}, \ldots, X_{t-p})}{\sqrt{\widehat{\mathrm{Var}}(X_t \mid X_{t-1}, \ldots, X_{t-p})}}. \tag{10}$$

(iii) Generate the bootstrap observations according to

$$X_t^+ = \sum_{j=1}^{p} \widehat{\alpha}_j X_{t-j}^+ + \widehat{\lambda}_t + r_{t,3}^+ \sqrt{\sum_{j=1}^{p} \widehat{\alpha}_j (1 - \widehat{\alpha}_j) X_{t-j}^+ + \widehat{\lambda}_t},$$

where $r_{t,3}^+$ are randomly and uniformly drawn from the standardized conditional residuals $r_{t,3}$ defined in (10), $\widehat{\alpha}_i$, for $i = 1, \ldots, p$, are the estimated thinning parameters of the Po-INARX(p) model fitted in the first step, and $\widehat{\lambda}_t$ is calculated as in (2).

For all three methods, the generated bootstrap observations X_t^+ are rounded to the nearest integer, with negative values replaced by zero.

3. Simulation study

3.1 *Simulation scenarios*

Next, we perform two simulation studies. The first one is developed to compare the performance of the WCLS and CML methods to estimate the parameters of Po-INARX(p) models and the second simulation study to evaluate the efficiency of the three bootstrap methods detailed in the previous sections using the WCLS method and $B = 100$ bootstrap replications. Count time series are simulated following a Po-INARX(p) process with length $n = 250$ and $n = 500$ according to two scenarios.

Scenario 1. Po-INARX(1) process with the mean of the innovation term ε_t given by $\ln(\lambda_t) = \beta_0 + \beta_1 \sin(2\pi t/52.25)$, where the true values of the parameters are $\beta_0 = \ln(15) \approx 2.7081$, $\beta_1 = 0.20$ and $\alpha_1 = 0.294$.

Scenario 2. Po-INARX(4) process with the mean of the innovation term ε_t given by $\ln(\lambda_t) = \beta_0 + \beta_1 t + \beta_2 \cos(2\pi t/52.25)$, where the true values of the parameters are $\beta_0 = \ln(15) \approx 2.7081$, $\beta_1 = 0.001$, $\beta_2 = -0.151$, $\alpha_1 = 0.294$, $\alpha_2 = \alpha_3 = 0$ and $\alpha_4 = 0.111$.

The performance of the WCLS and CML methods is evaluated using the RMSE for each parameter θ given by

$$\mathrm{RMSE}(\theta) = \sqrt{\frac{1}{R}\sum_{i=1}^{R}(\widehat{\theta}_i - \theta_0)^2},$$

where θ_0 is the true value of θ, $\widehat{\theta}_i$ is its i-th estimate, and R is the number of Monte Carlo replications. The bias and RMSE are used to evaluate the bootstrap methods.

3.2 *Simulation results*

Table 1 and Table 2 present the results of the first simulation study for both scenarios. Note that the WCLS and CML methods present similar good performances. Nevertheless, both methods underestimate the α parameters and overestimate the intercept (β_0). The outputs indicate that, for both scenarios, the RMSE/$|\theta_0|$ decrease when the sample size increases. This pattern indicates improved finite-sample performance of the CML and WCLS methods as the sample size increases. The RMSE/$|\theta_0|$ are higher for the estimates of the autocorrelation parameters (αs) and lower for the intercept (β_0). The estimates are closer to the real values mainly when $n = 500$. The CML and WCLS estimators seem to follow a Gaussian distribution; see Figure 1.

Table 1. Estimates and RMSE/$|\theta_0|$ estimates for the first scenario.

		CML		WCLS	
	True values	$n = 250$	$n = 500$	$n = 250$	$n = 500$
$\widehat{\beta}_0$	2.7081	2.7108 (0.0303)	2.7104 (0.0210)	2.7193 (0.0320)	2.7144 (0.0223)
$\widehat{\beta}_1$	0.200	0.1993 (0.1280)	0.1990 (0.0962)	0.1992 (0.1284)	0.1989 (0.0963)
$\widehat{\alpha}_1$	0.294	0.2900 (0.1918)	0.2914 (0.1333)	0.2837 (0.2066)	0.2885 (0.1426)

Table 2. Estimates and RMSE/$|\theta_0|$ estimates for the second scenario.

		CML		WCLS	
	True values	$n = 250$	$n = 500$	$n = 250$	$n = 500$
$\widehat{\beta}_0$	2.7081	2.7397 (0.0469)	2.7254 (0.0335)	2.7559 (0.0493)	2.7340 (0.0350)
$\widehat{\beta}_1$	0.001	0.0010 (0.2497)	0.0010 (0.0867)	0.0010 (0.2504)	0.0010 (0.0867)
$\widehat{\beta}_2$	-0.151	-0.1501 (0.1748)	-0.1503 (0.1188)	-0.1500 (0.1742)	-0.1502 (0.1187)
$\widehat{\alpha}_1$	0.294	0.2868 (0.1954)	0.2897 (0.1396)	0.2776 (0.2063)	0.2846 (0.1493)
$\widehat{\alpha}_4$	0.111	0.0945 (0.5237)	0.1025 (0.3801)	0.0934 (0.5189)	0.1022 (0.3671)

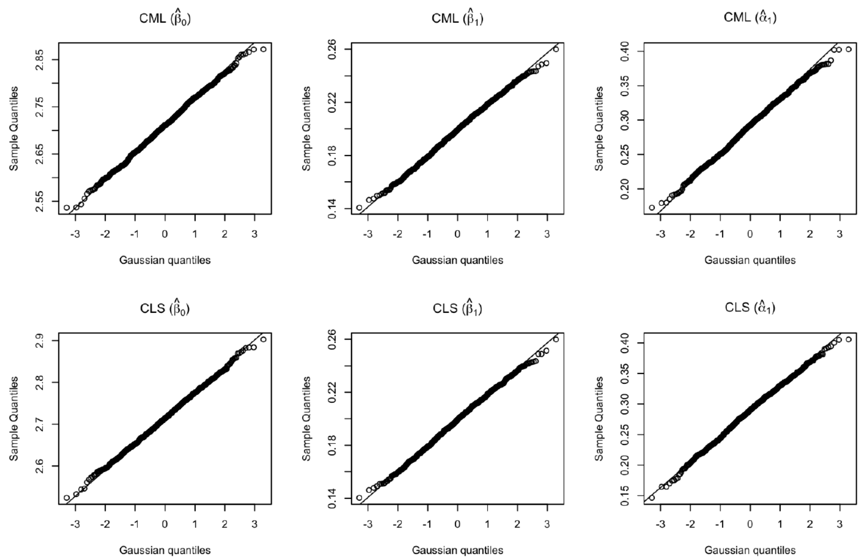

Figure 1. Theoretical quantile versus empirical quantile (Q-Q) plot for Scenario 1 and $n = 500$.

To analyze the possible relationship between the estimates of $\widehat{\beta}$s (intercept and parameters of the explanatory variables) and αs (AR parameters), scatter plots of bootstrap estimates $\widehat{\beta}_i$ and $\widehat{\alpha}_j$ for $i = 0, 1, 2$ and $j = 1, 4$ were plotted for both scenarios; see Figures 2 and 3. These plots indicate that there is a relevant negative correlation between $\widehat{\beta}_0$ and $\widehat{\alpha}_1$ for both scenarios and between $\widehat{\beta}_0$ and $\widehat{\alpha}_4$ for the second scenario. Note that a larger value of β_0 is associated with a smaller value of the corresponding AR parameter, maintaining the same value of the marginal mean $\mathrm{E}(X_t)$ defined in (3). For the first scenario, which describes a Po-INARX(1) model, we have $\beta_0 = \ln(15)$, $\alpha_1 = 0.294$, and $\mathrm{E}(X_t) = \alpha_1 \mathrm{E}(X_{t-1}) + \exp(\beta_0)$, when $t = k \times 52.25$, $k = 0, 1, \ldots$. Then, $\mathrm{E}(X_t)$ oscillates around $\exp(\beta_0)/(1 - \alpha_1) = 15/(1 - 0.294) = 21.2$. For example, the estimates $\beta_0 = 2.545$ and $\alpha_1 = 0.4$, shown in Figure 2, yield the same expected value as the true parameter values.

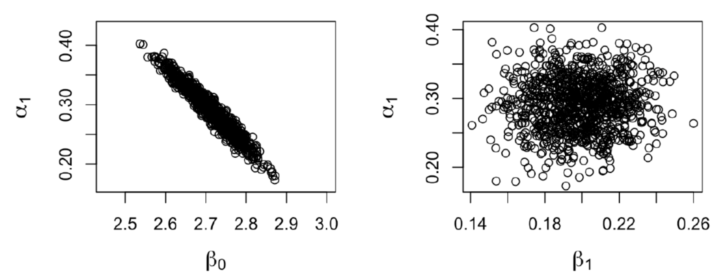

Figure 2. Correlation between $\widehat{\alpha}_1$ and $\widehat{\beta}_0$; $\widehat{\beta}_1$ using the CML method for Scenario 1.

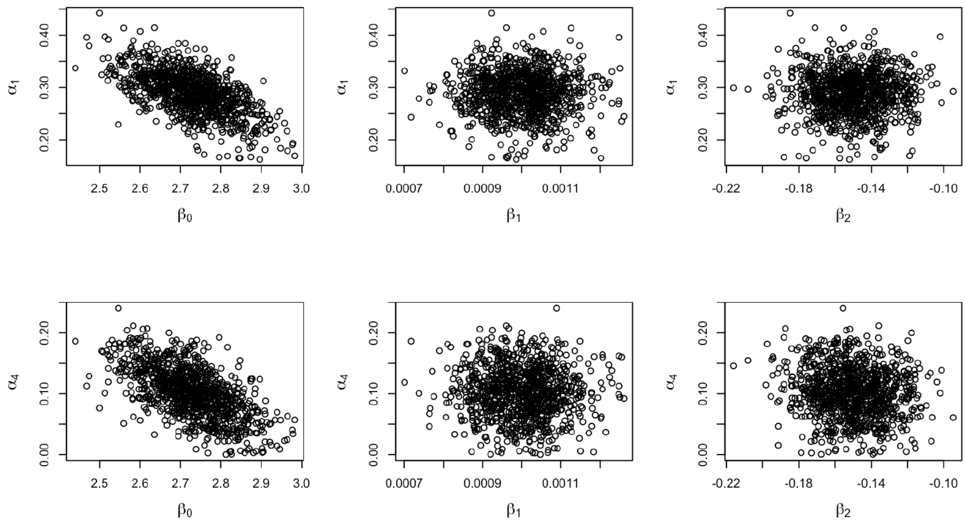

Figure 3. Correlation between $\widehat{\alpha}_1$, $\widehat{\alpha}_4$ and $\widehat{\beta}_0$, $\widehat{\beta}_1$, $\widehat{\beta}_2$ using the CML method for Scenario 2.

Tables 3 and 4 present the results of the three bootstrap methods for the first and second scenario using the WCLS method. Note that for both scenarios the bias and the RMSE decrease when the sample size increases, indicating the consistency of the WCLS method to estimate the parameters of the Po-INARX(p) models. We analyzed the true coverage probability (CP) of 95% confidence intervals. To compute a bootstrap confidence interval for an estimator the $(1-\alpha/2)$ and $\alpha/2$ quantiles were used from the bootstrap estimates. In general, for both scenarios the three bootstrap methods produce reliable confidence intervals, the CPs increase with increasing sample size reaching values close to the nominal level 0.95.

For the second scenario, the CP values generally approach the nominal 95% level for $n = 500$, whereas lower coverage is observed for some parameters when $n = 250$. This pattern is consistent with the possibility that the correlation between $\widehat{\alpha}_1$ and $\widehat{\beta}_0$, $\widehat{\beta}_1$ has a greater impact on the confidence intervals for smaller sample sizes. Among the three approaches, Method 2 generally provides the best coverage probabilities, although this improvement is not always accompanied by the smallest RMSE. As the sample size increases, the differences in coverage among the methods become less pronounced. Therefore, when the CML estimation method was implemented to estimate the parameters of the second scenario, with $n = 500$, the program took 58.888 seconds to converge while the WCLS method converged in 0.044 seconds.

Table 3. Bias, RMSE and CP for the bootstrap methods for Scenario 1.

		Method 1			Method 2			Method 3		
n		Bias	RMSE	CP (%)	Bias	RMSE	CP (%)	Bias	RMSE	CP (%)
250	β_0	0.0248	0.0927	91.4	0.0253	0.0931	91.6	0.0246	0.0931	91.3
	β_1	−0.0008	0.0272	92.2	−0.0010	0.0274	96.1	−0.0010	0.0272	91.4
	α_1	−0.0228	0.0672	91.3	−0.0235	0.0673	91.6	−0.0228	0.0674	90.8
500	β_0	0.0137	0.0611	93.9	0.0137	0.0611	94.2	0.0131	0.0610	92.9
	β_1	−0.0011	0.0192	92.6	−0.0012	0.0198	96.4	−0.0011	0.0193	92.9
	α_1	−0.0121	0.0432	92.9	−0.0122	0.0433	93.0	−0.0118	0.0431	92.8

Table 4. Bias, RMSE and CP for the bootstrap methods for Scenario 2.

n		Method 1 Bias	Method 1 RMSE	Method 1 CP (%)	Method 2 Bias	Method 2 RMSE	Method 2 CP (%)	Method 3 Bias	Method 3 RMSE	Method 3 CP (%)
250	β_0	0.0910	0.1538	86.1	0.0837	0.1435	89.8	0.0911	0.1536	85.9
	β_1	0.0001	0.0002	91.3	0.0001	0.0003	95.9	0.0001	0.0002	91.5
	β_2	−0.0016	0.0270	91.3	−0.0018	0.0270	97.3	−0.0018	0.0272	91.0
	α_1	−0.0400	0.0708	87.8	−0.0407	0.0702	88.6	−0.0402	0.0707	88.6
	α_4	−0.0289	0.0664	88.8	−0.0222	0.0574	94.6	−0.0285	0.0663	87.8
500	β_0	0.0494	0.1046	87.8	0.0508	0.1051	89.4	0.0501	0.1048	87.6
	β_1	0.0001	0.0001	92.2	0.0001	0.0001	98.3	0.0001	0.0001	92.4
	β_2	0.0001	0.0180	91.5	0.0004	0.0181	97.1	0.0001	0.0179	91.0
	α_1	−0.0197	0.0466	91.0	−0.0205	0.0469	90.8	−0.0201	0.0466	90.7
	α_4	−0.0153	0.0451	90.3	−0.0152	0.0435	92.5	−0.0154	0.0450	90.0

3.3 *Bootstrap-based bias correction*

Next, we propose to compute the bias correction, based on [30], for the estimators of the Poisson innovation parameters and the autocorrelation parameters of a Po-INARX(p) process. In both simulation scenarios, high bias values were observed using both CML and WCLS methods (see the RMSE/θ_0 values in Tables 1 and 2) mainly for series with length $n = 250$. Let $\widehat{\boldsymbol{\theta}}$ be the vector of the estimates fitting the Po-INARX(p) model computed according to the WCLS or CML estimators.

Firstly, the estimates for each bootstrap sample are centered calculating

$$\widehat{\boldsymbol{\theta}}_{\text{cen}} = \widehat{\boldsymbol{\theta}}^{*} - \text{median}(\widehat{\boldsymbol{\theta}}^{*}),$$

where $\widehat{\boldsymbol{\theta}}^{*}$ is the vector of estimates calculated for each bootstrap sample. Therefore, the corrected estimate for $\widehat{\boldsymbol{\theta}}$ is given by

$$\widehat{\boldsymbol{\theta}}_{\text{corr}} = \widehat{\boldsymbol{\theta}} - \text{mean}(\widehat{\boldsymbol{\theta}}_{\text{cen}}).$$

In Tables 5 and 6 the bias ($\widehat{\boldsymbol{\theta}} - \theta_0$) and bias correction ($\widehat{\boldsymbol{\theta}}_{\text{corr}} - \theta_0$) are presented for the estimates of the Po-INARX(p) for both scenarios using the WCLS estimator, where θ_0 is the true value of the parameter vector. In general, Method 2 presents a better performance, it manages to reduce the bias of the intercept and AR parameters.

Table 5. Bias and bias correction for the bootstrap methods for Scenario 1.

n		Bias WCLS	Bias correction Method 1	Bias correction Method 2	Bias correction Method 3
250	β_0	0.0113	0.0086	0.0031	0.0068
	β_1	−0.0008	−0.0024	−0.0038	−0.0024
	α_1	−0.0103	−0.0040	−0.0061	−0.0021
500	β_0	0.0064	0.0036	0.0035	0.0044
	β_1	−0.0011	−0.0033	−0.0023	−0.0031
	α_1	−0.0055	−0.0058	−0.0051	−0.0052

Table 6. Bias and bias correction for the bootstrap methods for Scenario 2.

n		Bias	Bias correction		
		WCLS	Method 1	Method 2	Method 3
250	β_0	0.0479	0.0489	0.0377	0.0506
	β_1	0.0000	0.0000	0.0000	0.0000
	β_2	0.0011	0.0005	0.0001	0.0003
	α_1	−0.0164	−0.0159	−0.0169	−0.0157
	α_4	−0.0176	−0.0154	−0.0092	−0.0158
500	β_0	0.0260	0.0326	0.0116	0.0330
	β_1	0.0000	0.0000	0.0000	0.0000
	β_2	0.0008	−0.0002	0.0006	−0.0005
	α_1	−0.0094	−0.0132	−0.0035	−0.0118
	α_4	−0.0088	−0.0082	−0.0086	−0.0091

4. Application with real data

4.1 *Context*

Next, the weekly number of hospital admissions due to pneumonia and influenza from January 2011 to December 2017 of people aged over 65 years in São Paulo (Brazil) is analyzed. We used a Po-INARX model where the mean of the innovation term may change over time. In Figure 4, it can be seen that the number of hospitalizations is increasing over time and it is especially higher in 2016. Also, the time series presents a seasonal pattern with higher counts in the middle of the year, which corresponds to the winter in São Paulo. A linear trend, sine and cosine functions, and a shift in 2016 are included in the model as explanatory variables in order to model deterministic trends and seasonality present in the data. Thus, in this study, the mean of the innovation term stated in (2) is given by

$$\ln(\lambda_t) = \beta_0 + \beta_1 \cos(w_{t;1}) + \beta_2 \sin(w_{t;1}) + \beta_3 \cos(w_{t;2}) + \beta_4 \sin(w_{t;2}) + \beta_5 t + \beta_6 1_t(2016), \quad (11)$$

where the week $t = 1, \ldots, 366$, the frequencies are $w_{t;1} = 2\pi t/52.25$, $w_{t;2} = 2\pi t/187.5$, and 1_t is the indicator function if observation t-th occurs in 2016, for $t \in \{261, \ldots, 312\}$. These frequencies were identified using the periodogram of the series [41] and they correspond to periods of one year and almost three years and a half.

4.2 *Modeling*

Firstly, a GLM model was fitted to the hospitalization series with a Poisson error distribution, the logarithm as link function, and the seasonal components and linear trend considered in the Po-INARX model formulated in (11). However, the first autocorrelations of the deviance residuals were significant at 5% based on the Ljung–Box test [42] with a p-value < 0.001. Thus, a Po-INARX(p) model was fitted to the data with $p = 1, 2, 3, 4$. Considering $p = 4$, the residual autocorrelations were not significant. The hospitalizations and the corresponding fitted values —the unconditional mean defined in (3)— using a Po-INARX(4) model and the time-varying mean of the Poisson innovation error are shown in Figure 4. The predicted values are higher than the innovation mean because the unconditional mean at instant t is defined as the weighted sum of weighted terms from previous instants plus the innovation mean as stated in (3).

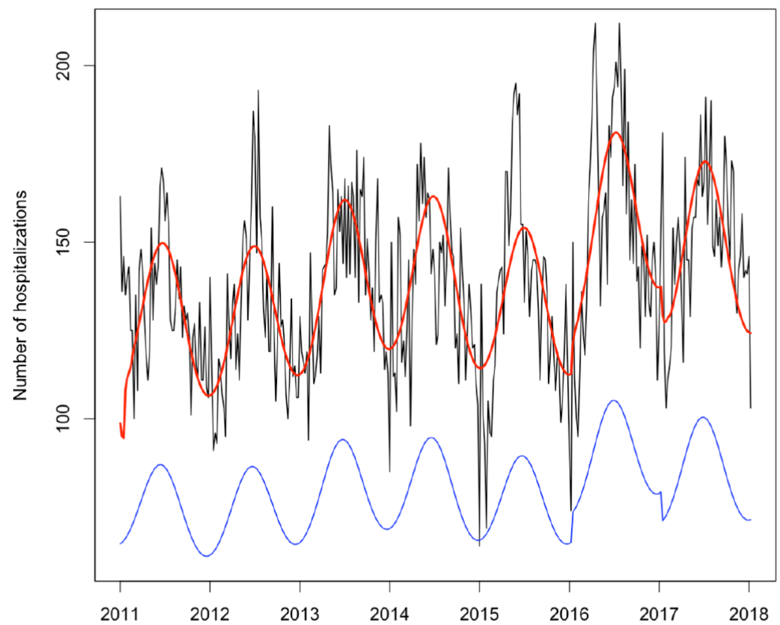

Figure 4. Weekly number of hospitalizations (black line), predicted values (red line), and the mean of the innovation process λ_t (blue line) using the Po-INARX(4) model and medical data from São Paulo, Brazil.

The observed series presents trend, seasonality, and an increase in 2016 that were well captured by the estimated λ_t defined in (11). All estimates are presented in Table 7 with the standard errors (SE) obtained using the three bootstrap procedures proposed in this article and the usual asymptotic SE.

Table 7. Estimates and corresponding asymptotic and bootstrap standard errors (SE); final Po-INARX(4) model using the WCLS method and medical data from São Paulo, Brazil.

Parameter	Estimate	SE (asymptotic)	SE1	SE2	SE3
Intercept (β_0)	4.260	0.110	0.117	0.120	0.115
$\cos(2\pi t/52.25)$ (β_1)	-0.166	0.014	0.017	0.018	0.016
$\sin(2\pi t/52.25)$ (β_2)	0.040	0.016	0.019	0.020	0.019
$\cos(2\pi t/187.5)$ (β_3)	0.016	0.017	0.020	0.022	0.019
$\sin(2\pi t/187.5)$ (β_4)	-0.042	0.015	0.018	0.021	0.018
t ($100\beta_5$)	0.025	< 0.001	< 0.001	< 0.001	< 0.001
Year-2016 (β_6)	0.117	0.034	0.041	0.044	0.041
α_1	0.294	0.043	0.050	0.052	0.050
α_2	0.045	0.044	0.039	0.039	0.039
α_4	0.111	0.041	0.047	0.047	0.047

All the parameters are significant at 5%, except β_3 and α_2 (p-values > 0.05). The estimate of α_3 was too small ($< 10^{-10}$) and was not included. Note that the standard errors obtained in the bootstrap procedures were, in general, similar and a little bit larger than those obtained by the asymptotic Gaussian distribution. The residual autocorrelation function (ACF) and the quantile–quantile plot of the randomized residuals defined in (9) are presented in Figure 5. It can be seen that the residuals are a white noise process, seem to be normally distributed and are homoscedastic (p-value > 0.05, the White test). Then, all the model assumptions seem to be valid. The residuals stated in (8) and (10) presented similar results and were omitted for brevity.

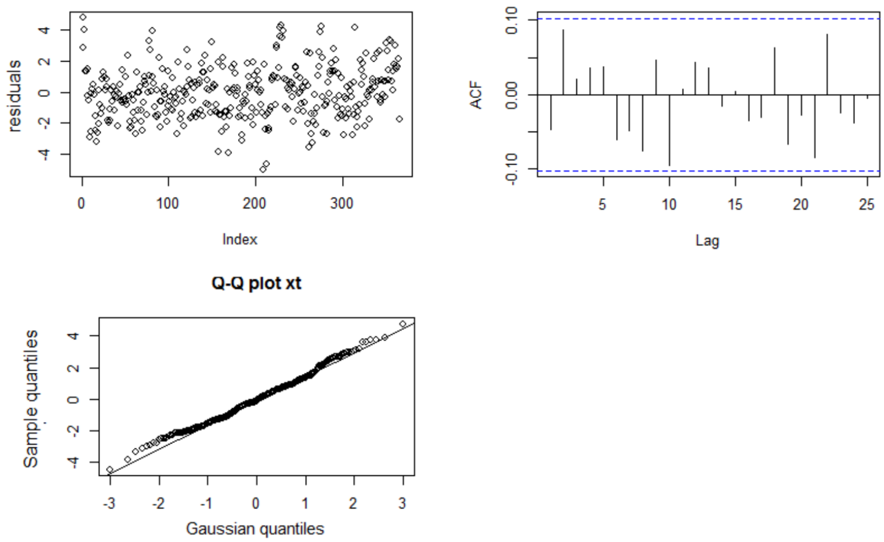

Figure 5. Randomized residuals defined in (9), corresponding correlogram, and Q-Q plot for the Po-INARX(4) model with medical data from São Paulo, Brazil.

5. Conclusions

In this article, we proposed an INAR(p) model with explanatory variables included in the mean of the Poisson innovation process to model non-stationary count time series, estimating covariate effects and modeling trends and seasonality (Po-INARX(p)).

The WCLS and CML methods were introduced to estimate the parameters of the proposed Po-INARX(p) model. Initial values for the AR parameters of the Po-INARX(p) model were proposed based on the partial autocorrelation function of the standardized residuals, calculated after removing the effects of the covariates using a Poisson GLM with the same link function considered in the Po-INARX(p) model. Wrong initial values for the AR parameters may lead to estimators with larger bias, mainly for the intercept and AR parameters. The simulation results suggest that the CML and WCLS estimators have decreasing RMSE and approximately Gaussian distributions as the sample size increases. Both methods showed similar RMSE performance. However, in both simulation scenarios, the AR parameters were underestimated and the intercept was overestimated, regardless of the estimation method, but the time-varying marginal mean and the effects of the covariates were well estimated.

A bootstrap-based bias correction was proposed to correct the bias of the estimators. Generally, Method 2 provides the lowest bias for the intercept and, in general, for the AR parameters. In simulation studies the bias correction proposal works well. An R routine was developed to estimate the parameters of the Po-INARX(p) model.

It is important to point out that the CML method took a long time to converge due to the sum of several terms, particularly for large count data. Thus, we suggest the use of the WCLS method in high-order INARX(p) models. The bootstrap-based bias correction method yielded satisfactory results for both estimation methods. Nevertheless, it is advisable to utilize WCLS due to the higher computational cost associated with CML.

During the implementation of the WCLS method, the convergence of the algorithm deserves some attention since too small values for some parameters can cause problems with the invertibility of a matrix (the implementation of the generalized inverse of the matrix does not solve the problem because the RMSE increases with this method).

The three bootstrap algorithms showed similar performances in both scenarios. The bootstrap standard errors are a little higher than the asymptotic standard errors. However, in the first scenario and considering Method 2, where the residuals depend on the thinning operator, we obtained a slightly better performance in terms of bias. The bootstrap estimates also revealed a relevant negative correlation between the intercept ($\widehat{\beta}_0$) and the thinning estimators ($\widehat{\alpha}_1$, $\widehat{\alpha}_4$), which keeps the marginal mean unchanged and helps to explain the CP of the confidence intervals for smaller sample sizes.

The inclusion of the same set of covariates to model the AR parameters and the mean of the innovation process may lead to confounding effects of the explanatory covariates. The inclusion of covariates only in the mean of the innovation process seems to be sufficient to explain the trends and seasonality of the observed time series since the predicted values capture the main features of the number of weekly admissions in Figure 4.

Future research building on this study could focus on other time-series models and count-related distributions [43, 44].

Statements

Acknowledgments

The authors thank the reviewers for their valuable comments, which helped improve the presentation and content of this article.

Author contributions

Conceptualization: O.Y.E. Albarracin, A.P. Alencar; Data curation: O.Y.E. Albarracin, A.P. Alencar; Investigation: O.Y.E. Albarracin, A.P. Alencar; Methodology: O.Y.E. Albarracin, A.P. Alencar; Software: O.Y.E. Albarracin, A.P. Alencar; Validation: O.Y.E. Albarracin, A.P. Alencar; Writing – original draft: O.Y.E. Albarracin, A.P. Alencar; Writing – review & editing: O.Y.E. Albarracin, A.P. Alencar. All authors have read and agreed to the published version of the article.

Conflicts of interest

The authors declare no conflict of interest.

Data and code availability

The code used to reproduce the analyses and results presented in this article is publicly available in the GitHub repository github.com/yesid017/Bootstrapping-Poisson-INARX. This repository contains all scripts, data processing workflows, and documentation required to ensure full reproducibility of the study's findings.

Declaration on the use of artificial intelligence (AI) technologies

The authors declare that no generative AI was used in the preparation of this article.

Funding

This research was funded by the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), Brazil, with number 2018/04654-9.

Open access statement

References

[1] McCullagh, P. and Nelder, J.A., 1989. Generalized Linear Models. Chapman and Hall, London, UK. DOI: 10.1201/9780203753736.

[2] Davis, R.A., Holan, S.H., Lund, R., and Ravishanker, N., 2016. Handbook of Discrete-Valued Time Series. CRC Press, Boca Raton, FL, US. DOI: 10.1201/b19485.

[3] Kedem, B. and Fokianos, K., 2002. Regression Models for Time Series Analysis. Wiley, Hoboken, NJ, US. DOI: 10.1002/0471266981.

[4] Ferland, R., Latour, A., and Oraichi, D., 2006. Integer-valued GARCH process. Journal of Time Series Analysis, 27, 923–942. DOI: 10.1111/j.1467-9892.2006.00496.x.

[5] Benjamin, M.A., Rigby, R.A., and Stasinopoulos, D.M., 2003. Generalized autoregressive moving average models. Journal of the American Statistical Association, 98, 214–223. DOI: 10.1198/016214503388619238.

[6] Box, G.E.P., Jenkins, G.M., Reinsel, G.C., and Ljung, G.M., 2015. Time Series Analysis: Forecasting and Control. Wiley, Hoboken, NJ, US. URL: www.wiley.com/en-us/Time+Series+Analysis.

[7] Davis, R.A., Dunsmuir, W.T., and Wang, Y., 1999. Modeling time series of count data. In Ghosh, S. (Ed.). Asymptotics, Nonparametrics, and Time Series, 158, pp. 63–114. CRC Press, Boca Raton, FL, US. URL: www.routledge.com/Asymptotics-Nonparametrics-and-Time-Series/Ghosh/p/book/9780367399924.

[8] Davis, R.A., Dunsmuir, W.T., and Streett, S.B., 2003. Observation-driven models for Poisson counts. Biometrika, 90, 777–790. DOI: 10.1093/biomet/90.4.777.

[9] Doukhan, P., Leucht, A., and Neumann, M.H., 2022. Mixing properties of non-stationary INGARCH(1,1) processes. Bernoulli, 28, 663–688. DOI: 10.3150/21-BEJ1362.

[10] Silva, R.B. and Barreto-Souza, W., 2019. Flexible and robust mixed Poisson INGARCH models. Journal of Time Series Analysis, 40, 788–814. DOI: 10.1111/jtsa.12459.

[11] Albarracin, O.Y.E., Alencar, A.P., and Ho, L.L., 2019. Generalized autoregressive and moving average models: multicollinearity, interpretation and a new modified model. Journal of Statistical Computation and Simulation, 89, 1819–1840. DOI: 10.1080/00949655.2019.1599892.

[12] Douc, R., Doukhan, P., and Moulines, E., 2013. Ergodicity of observation-driven time series models and consistency of the maximum likelihood estimator. Stochastic Processes and their Applications, 123, 2620–2647. DOI: 10.1016/j.spa.2013.04.010.

[13] Fahrmeir, L. and Tutz, G., 2013. Multivariate Statistical Modelling Based on Generalized Linear Models. Springer, New York, NY, US. DOI: 10.1007/978-1-4757-3454-6.

[14] Fokianos, K., 2012. Count time series models. In Handbook of Statistics, Vol. 30, pp. 315–347. Elsevier, Amsterdam, Netherlands. DOI: 10.1016/B978-0-444-53858-1.00012-0.

[15] Sözen, Ç., 2025. Short-horizon volatility spike forecasting via integrated functional and topological representations. Chilean Journal of Statistics, 16, 178–216. DOI: 10.32372/ChJS.16-02-03.

[16] Davis, R.A., Dunsmuir, W.T., and Wang, Y., 2000. On autocorrelation in a Poisson regression model. Biometrika, 87, 491–505. DOI: 10.1093/biomet/87.3.491.

[17] Cruz-Torres, C. and Ruggeri, F., 2025. On the use of the concentration function to compare predictive distributions in ARMA models. Chilean Journal of Statistics, 16, 217–233. DOI: 10.32372/ChJS.16-02-04.

[18] Al-Osh, M. and Alzaid, A.A., 1987. First-order integer-valued autoregressive (INAR(1)) process. Journal of Time Series Analysis, 8, 261–275. DOI: 10.1111/j.1467-9892.1987.tb00438.x.

[19] McKenzie, E., 1988. Some ARMA models for dependent sequences of Poisson counts. Advances in Applied Probability, 20, 822–835. DOI: 10.2307/1427362.

[20] Bourguignon, M., Rodrigues, J., and Santos-Neto, M., 2019. Extended Poisson INAR(1) processes with equidispersion, underdispersion and overdispersion. Journal of Applied Statistics, 46, 101–118. DOI: 10.1080/02664763.2018.1458216.

[21] Eliwa, M.S., Altun, E., El-Dawoody, M., and El-Morshedy, M., 2020. A new three-parameter discrete distribution with associated INAR(1) process and applications. IEEE Access, 8, 91150–91162. DOI: 10.1109/ACCESS.2020.2993593.

[22] Bourguignon, M., Vasconcellos, K.L.P., Reisen, V.A., and Ispány, M., 2016. A Poisson INAR(1) process with a seasonal structure. Journal of Statistical Computation and Simulation, 86, 373–387. DOI: 10.1080/00949655.2015.1015127.

[23] Barreto-Souza, W., 2015. Zero-modified geometric INAR(1) process for modelling count time series with deflation or inflation of zeros. Journal of Time Series Analysis, 36, 839–852. DOI: 10.1111/jtsa.12131.

[24] Weiß, C.H., 2010. The INARCH(1) model for overdispersed time series of counts. Communications in Statistics: Simulation and Computation, 39, 1269–1291. DOI: 10.1080/03610918.2010.490317.

[25] Brännäs, K., 1995. Explanatory variables in the AR(1) count data model. Umeå Economic Studies, 381, Department of Economics, Umeå University, Umeå, Sweden. URL: www.usbe.umu.se/digitalAssets/39/39103_ues381.pdf.

[26] Enciso-Mora, V., Neal, P., and Subba Rao, T., 2009. Integer valued AR processes with explanatory variables. Sankhyā B, 71, 248–263. URL: www.jstor.org/stable/41343031.

[27] Ding, X. and Wang, D., 2016. Empirical likelihood inference for INAR(1) model with explanatory variables. Journal of the Korean Statistical Society, 45, 623–632. DOI: 10.1016/j.jkss.2016.05.004.

[28] Bertail, P., Garay, A.M., Medina, F.L., and Jales, I.C.S., 2024. A maximum likelihood and regenerative bootstrap approach for estimation and forecasting of INAR(p) processes with zero-inflated innovations. Statistics, 58, 336–363. DOI: 10.1080/02331888.2024.2344670.

[29] Jentsch, C. and Weiß, C.H., 2019. Bootstrapping INAR models. Bernoulli, 25, 2359–2408. DOI: 10.3150/18-BEJ1057.

[30] Weiß, C.H. and Jentsch, C., 2019. Bootstrap-based bias corrections for INAR count time series. Journal of Statistical Computation and Simulation, 89, 1248–1264. DOI: 10.1080/00949655.2019.1576179.

[31] Du, J.G. and Li, Y., 1991. The integer-valued autoregressive (INAR(p)) model. Journal of Time Series Analysis, 12, 129–142. DOI: 10.1111/j.1467-9892.1991.tb00073.x.

[32] Klimko, L.A. and Nelson, P.I., 1978. On conditional least squares estimation for stochastic processes. The Annals of Statistics, 6, 629–642. DOI: 10.1214/aos/1176344207.

[33] Hansen, L., 1982. Large sample properties of generalized method of moments estimators. Econometrica, 50, 1029–1054. DOI: 10.2307/1912775.

[34] Bu, R., McCabe, B., and Hadri, K., 2008. Maximum likelihood estimation of higher-order integer-valued autoregressive processes. Journal of Time Series Analysis, 29, 973–994. DOI: 10.1111/j.1467-9892.2008.00590.x.

[35] Bu, R., 2006. Essays in Financial Econometrics and Time Series Analysis. Ph.D. thesis, University of Liverpool, Liverpool, UK. URL: livrepository.liverpool.ac.uk/3174790/1/433051.pdf.

[36] Ronning, G. and Jung, R.C., 1992. Estimation of a first order autoregressive process with Poisson marginals for count data. In L. Fahrmeir, B. Francis, R. Gilchrist, and G. Tutz (Eds.). Advances in GLIM and Statistical Modelling, pp. 188–194. Springer, New York, NY. DOI: 10.1007/978-1-4612-2952-0_29.

[37] Brooks, S.P., Giudici, P., and Roberts, G.O., 2003. Efficient construction of reversible jump Markov chain Monte Carlo proposal distributions. Journal of the Royal Statistical Society B, 65, 3–39. URL: www.jstor.org/stable/3088825.

[38] Wooldridge, J.M., 1994. Estimation and inference for dependent processes. Handbook of Econometrics, 4, 2639–2738. DOI: 10.1016/S1573-4412(05)80014-5.

[39] Pedeli, X., Davison, A.C., and Fokianos, K., 2015. Likelihood estimation for the INAR(p) model by saddlepoint approximation. Journal of the American Statistical Association, 110, 1229–1238. DOI: 10.1080/01621459.2014.983230.

[40] Bisaglia, L. and Gerolimetto, M., 2018. Estimation and forecasting in INAR(p) models using sieve bootstrap. Research Paper Series No. 06/WP/2018, Department of Economics, University Ca' Foscari of Venice, Venice, Italy. DOI: 10.2139/ssrn.3141801.

[41] Shumway, R.H. and Stoffer, D.S., 2006. Time Series Analysis and Its Applications: With R Examples. Springer, New York, NY, US. DOI: 10.1007/0-387-36276-2.

[42] Ljung, G.M. and Box, G.E.P., 1978. On a measure of lack of fit in time series models. Biometrika, 65, 297–303. DOI: 10.1093/biomet/65.2.297.

[43] Sardar, I., Akbar, M.A., Leiva, V., Alsanad, A., and Mishra, P., 2023. Machine learning and automatic ARIMA/Prophet models-based forecasting of COVID-19: Methodology, evaluation, and case study in SAARC countries. Stochastic Environmental Research and Risk Assessment, 37, 345–359. DOI: 10.1007/s00477-022-02307-x.

[44] Saulo, H., Leão, J., Leiva, V., and Aykroyd, R.G., 2019. Birnbaum–Saunders autoregressive conditional duration models applied to high-frequency financial data. Statistical Papers, 60, 1605–1629. DOI: 10.1007/s00362-017-0888-6.