Journal of Data Science 23 (1), 208–224
March 2025
DOI: 10.6339/24-JDS1124
Statistical Data Science

Dynamic Network Poisson Autoregression with Application to COVID-19 Count Data ☆

Manabu Asai[1], Amanda M. Y. Chu[2], and Mike K. P. So[3,*]

[1]*Faculty of Economics, Soka University, Hachioji, Tokyo, Japan*
[2]*Department of Social Sciences and Policy Studies, The Education University of Hong Kong, Hong Kong*
[3]*Department of Information Systems, Business Statistics and Operations Management, The Hong Kong University of Science and Technology, Hong Kong*

Abstract

There is growing interest in accommodating network structure in panel data models. We consider dynamic network Poisson autoregressive (DN-PAR) models for panel count data, enabling their use in regard to a time-varying network structure. We develop a Bayesian Markov chain Monte Carlo technique for estimating the DN-PAR model, and conduct Monte Carlo experiments to examine the properties of the posterior quantities and compare dynamic and constant network models. The Monte Carlo results indicate that the bias in the DN-PAR models is negligible, while the constant network model suffers from bias when the true network is dynamic. We also suggest an approach for extracting the time-varying network from the data. The empirical results for the count data for confirmed cases of COVID-19 in the United States indicate that the extracted dynamic network models outperform the constant network models in regard to the deviance information criterion and out-of-sample forecasting.

Keywords *Bayesian analysis; Markov chain Monte Carlo; multivariate count variables; network analysis; panel data; Poisson regression*

1 Introduction

Modeling count data have attracted a great deal of attention among researchers because of their various important applications in different fields. For example, we can evaluate financial risk through the analysis of the number of intraday transactions (Zhang et al., 2001) and the number of operational risk incidents (Panjer, 2006). When dealing with environmental risk, researchers have attempted to predict daily earthquake counts (Reasenberg and Jones, 1994). There are also examples in modeling crime data (Chen and Lee, 2016) and, more recently, analyzing the number of confirmed COVID-19 cases (Chan et al., 2021) and deaths (Chen et al., 2022). Since some count data are observed regularly in the form of time series, it is natural to study the time series structure of counts. A seminal piece of research was conducted by Fokianos et al. (2009) in this field; he developed a modeling scheme through Poisson autoregression and its accompaning

☆The authors are most grateful to Yoshihisa Baba and two anonymous reviewers for their very helpful comments and suggestions. The first author acknowledges the financial support of the Japan Society for the Promotion of Science (grant number 22KK0022). This work was partially supported by The Hong Kong University of Science and Technology research grant "Risk Analytics and Applications" (grant number SBMDF21BM07). The funding recipient was MKPS.

*Corresponding author. Email: immkpso@ust.hk.

Received June 24, 2023; Accepted March 12, 2024

statistical properties. Extensions in univariate analysis include work by Fokianos and Tjøstheim (2011), who proposed a log-linear version of the Poisson autoregression, and Chen et al. (2016), who built a dynamic count data model with over-dispersion features.

Unlike modeling multivariate continuous variables with the multivariate normal assumption, multivariate modeling of count variables is complicated, as it is nontrivial and difficult to define a high-dimensional distribution for counts. One viable approach in the multi-dimensional modeling of counts involves the use of copulas (Nikoloulopoulos and Karlis, 2009). A core idea regarding copulas involves modeling the dependence among different count variables through a copula function (Fokianos et al., 2020). Debaly and Truquet (2023) also extended this copula idea to mixed data. To incorporate correlation structures among counts, Zhang et al. (2017) considered a regression approach for relating multivariate count data. For reviews of multivariate count time series modeling, we can refer to Karlis (2016), Fokianos et al. (2022) and Fokianos (2021).

A recent class of models using the network autoregression idea put forward by Zhu et al. (2017) is the Poisson network autoregression (PNAR) developed by Armillotta and Fokianos (2021). The main innovation in the PNAR is the way in which it incorporates network information regarding the different count time series in statistical modeling. An advantage of the PNAR is its ability to explain possible dependence between different time series without using a copula function. Armillotta and Fokianos (2022) and Amillotta et al. (2022) further considered the testing of linearity and generalized linear network autoregression, respectively, in this context. In this paper, we extend the PNAR of Armillotta and Fokianos (2021) to allow for time-dependent networks and other features in count modeling.

Bayesian analysis is a common statistical inference approach to analyzing count data. For example, Frühwirth-Schnatter and Wagner (2006) studied parameter-driven models of counts in a dynamic linear model framework (West et al., 1985; West and Harrison, 2006). Ravishanker et al. (2014) considered hierarchical dynamic modeling in relation to multiple time series of counts. Aktekin et al. (2018) developed sequential Bayesian analysis using multivariate count data. Berry and West (2020) then developed efficient Bayesian computation methods to analyze many count-valued time series. West (2020) and Soyer and Zhang (2022) reviewed Bayesian modeling tools for count data and outlined challenges in their real-life applications and forecasting.

Although applications in many areas of science and business are thus evident, one particular area of focus in the literature is epidemiology. For example, Livsey et al. (2018) applied multivariate integer-valued time series models to hurricane counts. Chen et al. (2019) proposed a Markov switching integer-valued models with conditional heteroskedasticity for counts of dengue. Polwiang (2020) and Martínez-Bello et al. (2017) also looked into dengue counts and related time series patterns. Chen et al. (2021) developed a hysteric integer-valued GARCH models for disease counts. An important stream of research in epidemiology considers disease control and prevention using count data. Held et al. (2005) proposed a statistical framework for infectious disease surveillance. Wang et al. (2015) worked on the problem of influenza-like-illness. There are also many research findings regarding the modeling and prediction of COVID-19 pandemic status and severity in recent research. Some examples are Shinde et al. (2020), Chan et al. (2021), Bartolucci et al. (2021), and El-Morshedy et al. (2022), who mainly discuss the statistical modeling of the number of confirmed COVID-19 cases or deaths. One main complication in regard to data is the need to deal with the number of infectious disease cases in different regions, inducing a high-dimensional count data time series problem. In this paper, we use the COVID-19 pandemic as a case to demonstrate our proposed network count data model. Our main idea is based on some pandemic network features in the literature (So et al., 2020; Chu et al., 2021; So et al.,

2021b; Chu et al., 2023); we naturally integrate network linkage information among different regions to build a multivariate count data time series model.

In this paper, we consider a new dynamic network Poisson autoregressive model for high dimensional count data. As stated above, multivariate Poisson autoregressive models discussed in Karlis (2016), Fokianos et al. (2022) and Fokianos (2021) are useful for low dimensional count data. For high dimensional data, the number of parameters increases with an order of the square of the dimension of the data. To avoid the curse of dimensionality, we may consider a single autoregressive parameter as in the panel data analysis, by sacrificing the information of interactions of data. To recover the information using a small number of parameters, Armillotta and Fokianos (2021) developed a network Poisson autoregressive model, when the network structure is known. Our new model can accommodate dynamic network structures. Unlike Armillotta and Fokianos (2021), we introduce a technique to extract the time-varying network when the true one is unavailable. Our simulation and empirical results indicate that assuming constant correlation can lead to serious bias when the true network is dynamic.

The rest of the paper is structured as follows. Section 2 develops a new dynamic network model for Poisson autoregression, and examines the stationary condition. Section 3 introduces a Bayesian Markov chain Monte Carlo method for estimating the model, and conducts Monte Carlo experiments to demonstrate the properties of the posterior mean estimator. Section 4 provides the empirical results using the network data for the number of confirmed cases of COVID-19 in the United States. Finally, Section 5 consists of concluding remarks.

2 Dynamic Network Poisson Autoregressive Model

2.1 Model Specification

Denote the count variable for the ith node at time t as $y_{i,t}$ for $i = 1, \ldots, N$ and $t = 1, \ldots, T$. Adopting the network terminology, this node can be a variable, an individual, or a region in our case, as we are studying the number of confirmed COVID-19 cases or deaths in different regions. We assume there exists network linkages among the N nodes, forming a network at time t. Consider a Poisson model for network data conditional on the rate, $\lambda_{i,t}$, as:

$$P(y_{i,t} = k \mid \lambda_{i,t}) = \frac{\lambda_{i,t}^k \exp(\lambda_{i,t})}{k!}, \quad (k = 0, 1, \ldots) \tag{1}$$

$$\log \lambda_{i,t} = \gamma_0 + \boldsymbol{\gamma}_1' \boldsymbol{w}_i + \sum_{h=1}^{p} \alpha_h \log \lambda_{i,t-h} + \sum_{l=1}^{q} R_i(\boldsymbol{y}_{t-l}), \tag{2}$$

where $\boldsymbol{y}_t = (y_{1,t}, \ldots, y_{N,t})'$, $\boldsymbol{w}_t$ is an $s \times 1$ vector of exogenous variables. By definition, we obtain $E(y_{i,t}|\lambda_{i,t}) = V(y_{i,t}|\lambda_{i,t}) = \lambda_{i,t}$, hence the model accommodates not only the conditional mean but also the conditional heteroskedasticity. The term $R_i(\boldsymbol{y}_{t-l})$ represents the momentum and network effects from lth lag, defined by:

$$R_i(\boldsymbol{y}_{t-l}) = \beta_l y_{i,t-l}^* + \psi_l n_{i,t-l}^{-1} \sum_{j=1}^{N} g_{ij,t-l} y_{j,t-l}^*, \tag{3}$$

where $y_{i,t}^* = \log(1 + y_{i,t})$; $g_{ij,t}$ is the linkage variable, which takes the value of one if there is a link from node i to node j and zero otherwise; $n_{i,t} = \max(n_{i,t}^*, 1)$; and $n_{i,t}^* = \sum_{j \neq i} g_{ij,t}$ is the total number of linkages that i has. The network linkage variable $g_{ij,t}$ can be asymmetrical in directed

networks; that is, $g_{ij,t}$ may not be equal to $g_{ji,t}$. In undirected networks, $g_{ij,t} = g_{ji,t}$. On the right hand side of (3), the first term is the momentum effect, while the second term is a network effect defined by the average impact from the ith subject's neighbors. Our model reduces to the class of PNAR models in Armillotta and Fokianos (2021) if $q = 1$, $\alpha_1 = \cdots = \alpha_p = 0$ and the linkage variables are known constant, $g_{ij,t} = g_{ij}$. If we omit the network effect by setting $\psi_1 = \cdots = \psi_q = 0$, we obtain the model developed by Fokianos et al. (2020). In order to distinguish it from these models, we refer to our new model (1)-(3) as the Dynamic Network Poisson Autoregressive (DN-PAR) model. We use DN-PAR(p, q) to denote the DN-PAR model with the orders of p and q.

There are three main features in the proposed DN-PAR model. First, unlike existing network count data models in the literature, we make use of time-dependent information to formulate $R_i(\boldsymbol{y}_{t-l})$. Second, in contrast to the PNAR model, we include an additional term $\sum_{h=1}^{p} \alpha_h \log \lambda_{i,t-h}$ in the DN-PAR model. This additional term consists of the past values of the log of the conditional mean for individual i. Third, the time evolution of $\lambda_{i,t}$ in (2) is partly explained by the exogenous variable $\boldsymbol{w}_i$.

2.2 Linkage Variable

The linkage variable can be observable or unobservable. If the data are for a social network, such as Twitter, $g_{ij,t} = 1$ if the ith individual follows the jth individual at day t, and is equal to zero otherwise. When $g_{ij,t}$ is unobservable, we may consider the $\{0, 1\}$ process to be conditional on the past information of $(y_{1,t}, \ldots, y_{N,t})$. In such a case, we apply the moving-window approach from So et al. (2021a) to extract the linkage variable. Define

$$g_{ij,t} = \begin{cases} \mathbf{1}\{\zeta_t > t_{m-2}(1-\alpha/2)\} & \text{if } i \neq j \\ 0 & \text{otherwise} \end{cases}, \tag{4}$$

where

$$\zeta_t = \left| \hat{\rho}_{ij,t} \sqrt{\frac{m-2}{1-\hat{\rho}_{ij,t}^2}} \right|,$$

and $\hat{\rho}_{ij,t}$ is the sample correlation coefficient based on $(y^*_{i,t-m+1}, \ldots, y^*_{i,t})$ and $(y^*_{j,t-m+1}, \ldots, y^*_{j,t})$. The size, m $(m > 2)$, can be set to a small value if we want to capture local network structures of the $y_{i,t}$. In the empirical analysis, we choose $m = 6$ for weekly COVID-19 count data. Note that the correlation $\hat{\rho}_{ij,t}$ is statistically significant at α level if $\zeta_t > t_{m-2}(1-\alpha/2)$, where t_{df} is the quantile function of t distribution with degrees of freedom df.

To distinguish it from the model with the observable network, we refer to the DN-PAR model based on the extracted linkage variable (4) as the Extracted Network Poisson Autoregressive (EN-PAR) model.

We can consider the Constant Network Poisson Autoregressive (CN-PAR) model, using estimates of constant linkages. For this purpose, we obtain $g_{ij,T}$ $(i, j = 1, \ldots, N)$ by setting $m = T$ in equation (4), and use the values as constants for the whole sample.

2.3 Stationary Condition

To examine the stationary condition of our DN-PAR model (1)-(3), we follow the argument put forth by Debaly and Truquet (2019), who applied the convergence of the backward iterations of random maps given by Wu and Shao (2004) to multivariate Poisson autoregressive models from Fokianos et al. (2020).

Denote $|A|_v = (|a_{ij}|)_{(i,j)}$ for a square matrix A. For a case involving constant linkage, $g_{ij,t} = g_{ij}$, and Theorem 5 put forth by Debaly and Truquet (2019) implies that the stationary condition is given by:

$$\rho\left(\sum_{h=1}^{r}(|A_h|_v + |C_h|_v)\right) < 1,$$

where $\rho(A)$ is the spectral radius of a square matrix A, $r = \max(p, q)$, $A_h = \alpha_h I_N$ for $h = 1, \ldots, p$ and $A_h = O$ for $h > p$,

$$C_h = \beta_h I_N + \psi_h \begin{pmatrix} n_1^{-1} g_{11} & n_1^{-1} g_{12} & \cdots & n_1^{-1} g_{1N} \\ n_2^{-1} g_{21} & n_2^{-1} g_{22} & \cdots & n_2^{-1} g_{2N} \\ n_N^{-1} g_{N1} & n_N^{-1} g_{N2} & \cdots & n_N^{-1} g_{NN} \end{pmatrix}, \quad n_i = \max(1, \sum_{j \neq i} g_{ij}),$$

for $h = 1, \ldots, q$ and $C_h = O$ for $h > q$. See also Proposition 3 in Armillotta and Fokianos (2022).

For time-varying linkages, we can check the stability of the structure using the top Lyapunov exponent, defined by:

$$\eta = \inf_{n \in \mathcal{N}} \left\{ \frac{1}{n} E \log ||F_n \cdots F_1|| \right\},$$

where

$$F_t = \begin{pmatrix} |A_1|_v + |C_{1,t}|_v & |A_2|_v + |C_{2,t}|_v & \cdots & |A_r|_v + |C_{r-1,t}|_v & |A_r|_v + |C_{r,t}|_v \\ I_N & O & \cdots & O & O \\ O & I_N & \cdots & O & O \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ O & O & \cdots & I_N & O \end{pmatrix},$$

and

$$C_{h,t} = \beta_h I_N + \psi_h \begin{pmatrix} n_{1,t-h}^{-1} g_{11,t-h} & n_{1,t-h}^{-1} g_{12,t-h} & \cdots & n_{1,t-h}^{-1} g_{1N,t-h} \\ n_{2,t-h}^{-1} g_{21,t-h} & n_{2,t-h}^{-1} g_{22,t-h} & \cdots & n_{2,t-h}^{-1} g_{2N,t-h} \\ n_{N,t-h}^{-1} g_{N1,t-h} & n_{N,t-h}^{-1} g_{N2,t-h} & \cdots & n_{N,t-h}^{-1} g_{NN,t-h} \end{pmatrix}$$

for $h = 1, \ldots, q$ and $C_{h,t-h} = O$ for $h > q$. Note that $g_{ii,t} = 0$ by definition. If η is strictly negative, the DN-PAR model is stationary. Note that the definition of η does not depend on the choice of the norm. Although it is difficult to obtain its value, we can determine the top Lyapounov exponent using Monte Carlo simulations of the time-varying matrices $\{F_t\}$. For $p = q = 1$, we may use the value of $|\alpha_1| + |\beta_1| + |\psi_1|$ as an approximated measure of the stationary condition. We provide a numerical example based on real data in Subsection 3.2.

2.4 Model Selection

We discuss the model selection for the new DN-PAR model. In practice, we can use the multivariate Poisson autoregressive models in Karlis (2016), Fokianos et al. (2022) and Fokianos (2021) for low dimensional cases. It is recommended to consider the network Poisson autoregressive model, which is computationally more feasible than multivariate Poisson autoregressive models for high dimensional cases. If the observed linkage variable is time-varying (constant), it is preferable to use the DN-PAR (CN-PAR) model to incorporate network information in modeling multivariate count data. When the linkage variable is unobservable, we need to consider the non-network PAR model given by setting $\psi_1 = \cdots = \psi_q = 0$ in equation (3), in addition to the

CN-PAR and EN-PAR models. We can choose promising models using the deviance information criterion (DIC), the Bayes factor, and/or the marginal likelihood in the Bayesian analysis, explained in the next section. Although we consider a parsimonious case with $p = q = 1$ in our analysis, we can select p and q using the above Bayesian model selection methods.

3 Bayesian Estimation

3.1 MCMC Algorithm

Based on recent developments of the MCMC technique, we can use a fast and efficient approach in this study: Haario et al. (2006) Delayed Rejection & Adaptive Metropolis (DRAM) algorithm to perform Bayesian estimation. The DRAM algorithm combines two ideas from the MCMC literature: adaptive Metropolis samplers (Haario et al., 1999, 2001) and delayed rejection (Tierney and Mira, 1999; Green and Mira, 2001; Mira, 2001). The adaptive Metropolis sampler is based on the idea of creating a Gaussian proposal distribution with a covariance matrix calibrated using the sample path of the MCMC chain. The basic idea of the delayed rejection is that upon rejection in a Metropolis-Hastings algorithm, instead of advancing time and retaining the same position, a higher stage move is proposed to improve the efficiency of the resulting MCMC estimators. The DRAM algorithm is useful for cases where good proposal distributions are not available and when the adaptation process has a slow start. For this reason, the DRAM algorithm can be applied to various univariate and multivariate conditional volatility models, especially when models are highly complicated.

We explain the delayed rejection and the adaptive Metropolis-Hastings algorithm used in the DRAM algorithm in Haario et al. (2006) as below. For the delayed rejection part, suppose the current position of the Markov chain is $\boldsymbol{\theta}_n = \theta$. As in a regular Metropolis-Hastings algorithm, a candidate move, $\theta_c^{(1)}$, is generated from a candidate generating density, $q_1(\theta, \cdot)$ and accepted with the usual probability $\alpha_1(\theta, \theta_c^{(1)}) = \min\left(1, \frac{N_1}{D_1}\right)$, where $N_1 = \pi(\theta_c^{(1)})q_1(\theta_c^{(1)}, \theta)$ and $D_1 = \pi(\theta)q_1(\theta, \theta_c^{(1)})$, $\pi(\theta)$ is the prior times the likelihood function evaluated at θ. When it is rejected, instead of retaining the same position, $\boldsymbol{\theta}_{n+1} = \theta$, as in a standard Metropolis-Hastings algorithm, a second stage move, $\theta_c^{(1)}$, is proposed. The second stage proposal is allowed to depend not only on the current position of the chain but also on the rejected proposal: $q_2(\theta, \theta_c^{(1)}, \cdot)$. The second stage proposal is accepted with the probability $\alpha_2(\theta, \theta_c^{(1)}, \theta_c^{(2)}) = \min\left(1, \frac{N_2}{D_2}\right)$, where $N_2 = \pi(\theta_c^{(2)})q_1(\theta_c^{(2)}, \theta_c^{(1)})q_2(\theta_c^{(2)}, \theta_c^{(1)}, \theta)[1 - \alpha_1(\theta_c^{(2)}, \theta_c^{(1)})]$ and $D_2 = \pi(\theta)q_1(\theta, \theta_c^{(1)})q_2(\theta, \theta_c^{(1)}, \theta_c^{(2)})[1 - \alpha_1(\theta, \theta_c^{(1)})]$. This process of delaying rejection can be iterated. If q_i denotes the proposal density at the i-th stage, the acceptance probability at that stage is given by $\alpha_i(\theta, \theta_c^{(1)}, \ldots, \theta_c^{(i)}) = \min\left(1, \frac{N_i}{D_i}\right)$, where

$$
\begin{aligned}
D_i &= q_i(\theta, \ldots, , \theta_c^{(i)})[q_{i-1}(\theta, \ldots, \theta_c^{(i-1)})[q_{i-2}(\theta, \ldots, \theta_c^{(i-2)}) \cdots \\
&\qquad [q_2(\theta, \theta_c^{(1)}, \theta_c^{(2)})[q_1(\theta, \theta_c^{(1)})\pi(\theta) - N_1] - N_2] - N_3] \cdots - N_{i-1}], \text{ and} \\
N_i &= \pi(\theta_c^{(i)})q_1(\theta_c^{(i)}, \theta_c^{(i-1)})q_2((\theta_c^{(i)}, \theta_c^{(i-1)}, \theta_c^{(i-2)}) \cdots q_i(\theta_c^{(i)}, \theta_c^{(i-1)}, \ldots, \theta) \\
&\qquad \times [1 - \alpha_1(\theta_c^{(i)}, \theta_c^{(i-1)})][1 - \alpha_2((\theta_c^{(i)}, \theta_c^{(i-1)}, \theta_c^{(i-2)})] \cdots [1 - \alpha_{i-1}(\theta_c^{(i)}, \theta_c^{(i-1)}, \ldots, \theta_c^{(1)})].
\end{aligned}
$$

As shown in Mira (2001), the process of delaying rejection can be interrupted at any stage. In our analysis, we attempt to move away from the current position five times at most; otherwise, we let the chain stay where it is.

We now turn to an explanation of the adaptive Metropolis-Hastings algorithm. The basic idea is to create a Gaussian proposal distribution with a covariance matrix calibrated using the sample path of the MCMC chain. Following Haario et al. (2006), when the current position is $\boldsymbol{\theta}_n = \theta_n$, we choose the proposal distribution to be the multivariate normal distribution with mean θ_n, and covariance matrix to be defined by:

$$S_n = \begin{cases} S_0, & n \leqslant n_0 \\ s_p \mathrm{Cov}(\theta_0, \ldots, \theta_n), & n > n_0 \end{cases}$$

where S_0 is an initial covariance matrix, $s_k = 2.4^2/k$, k is the dimension of $\boldsymbol{\theta}$, and Cov creates the sample covariance matrix. In our analysis, we set $n_0 = 2,000$. Combining the two approaches, Haario et al. (2006) showed the ergodicity of the DRAM algorithm.

To obtain the posterior quantities of the parameters $\boldsymbol{\theta}$ from the posterior distribution, we implement the DRAM algorithm:

1. Initialize $\boldsymbol{\theta}$.
2. For $i = 1, \ldots, m$, generate the candidate $\theta_c^{(i)}$ via the random walk chain, and calculate the acceptance probability $\alpha_i(\theta, \theta_c^{(1)}, \ldots, \theta_c^{(i)})$ by the delayed rejection algorithm. Break this step if $\alpha_i(\theta, \theta_c^{(1)}, \ldots, \theta_c^{(i)}) = 1$.
3. Using the final candidate and the acceptance probability, accept θ_c to set $\boldsymbol{\theta} = \theta^c$ or reject θ_c to keep $\boldsymbol{\theta}$ unchanged.
4. Go to step 2.

We set $m = 100$ in our analysis. We assume that the prior distribution follows the normal distribution, $\boldsymbol{\theta} \sim N(0, 5I_k)$. As discussed in the previous section, it is hard to impose parameter restriction for stationary condition. The above algorithm indicates that the computational complexity is irrelevant to changes of (N, T, p, q). On the other hand, the computational complexity of frequentist's approach increases as p and q increase.

3.2 Simulated Data Analysis

We examine the performance of the Bayesian MCMC estimator via Monte Carlo experiments. Our data generating processes (DGPs) are based on the empirical results in the next section. We set $(N, T) = (51, 80)$ with three parameter sets: DGP1 $(\gamma_0, \beta, \psi, \alpha) = (0.6, 0.93, -0.01, 0.02)$, DGP2 $(\gamma_0, \beta, \psi, \alpha) = (0.6, 0.90, -0.01, 0.07)$, and DGP3 $(\gamma_0, \beta, \psi, \alpha) = (0.6, 0.93, -0.03, 0.02)$. DGP1 mimics the empirical result. DGP2 and DGP3 have higher persistence, since the values of the approximated measure, $|\alpha| + |\beta| + |\psi|$, for DGP1, DGP2, and DGP3 are 0.95, 0.98, and 0.98, respectively. While the difference from DGP1 comes form the value of ψ for DGP3, it comes from the difference of (α, β) for DGP2. We use the same observed linkage variables, $g_{ij,t}$ $(i, j = 1, \ldots, N)$ for $t = 1, \ldots, T$, in the next section. Under these settings, we generate data to estimate the DN-PAR(1,1) model using the given linkage variables. The number of replications is 500. For the sake of comparison, we estimate a constant network model for the same DGP. In other words, we examine the effects of assuming a constant network structure when it is time-varying. We consider the CN-PAR(1,1) model for comparison.

To examine the stationary condition for three DGPs, Figure 1 shows the values of $\eta_t = \frac{1}{t} \log ||F_t \cdots F_1||$ $(t = 1, \ldots, T)$ to approximate the top Lyapunov exponent. Although it is terminated at $t = 80$ for the available data, all η_t seem to converge to negative values, apart from zero. Hence, these parameter settings provide the stationary processes for the respective DN-PAR(1,1) models. The values of η_{80} for DGP2 and DGP3 are closer to zero than that of DGP1, since DGP2 and DGP3 have the higher persistence than DGP1.

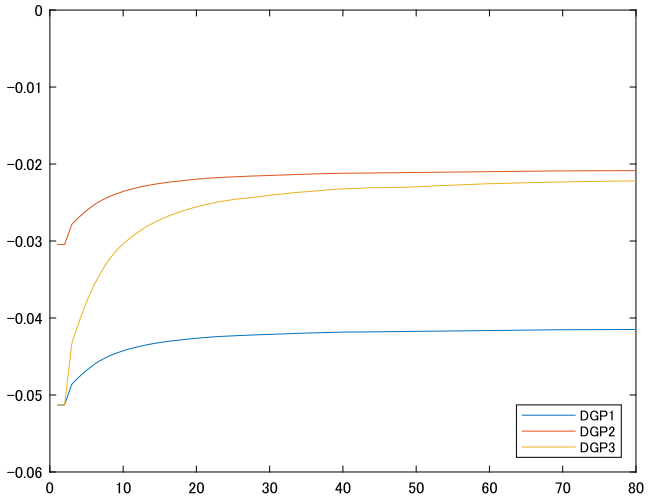

Figure 1: Top Lyapunov exponent. This shows the transition of $\eta_t = \frac{1}{t}\log||F_t \cdots F_1||$ based on the spectral radius and the extracted linkage variables.

Table 1: Monte Carlo results for the posterior means for the DN-PAR and CN-PAR Models. The DGP is generated using the time-varying linkage variables. While we estimate the DN-PAR(1,1) model using the given linkage variables, estimation of the PNAR(1,1) is based on the estimated constant linkages.

		DN-PAR(1,1)			CN-PAR(1,1)		
Parameter	True	Mean	Std. Dev.	RMSE	Mean	Std. Dev.	RMSE
DGP1							
γ	0.6	0.5996	0.0026	0.0026	0.5750	0.0046	0.0254
β	0.93	0.9297	0.0157	0.0157	1.2372	0.0117	0.3074
ψ	−0.01	−0.0100	6.6415×10^{-5}	6.6468×10^{-5}	−0.0622	0.0027	0.0523
α	0.02	0.0202	0.0157	0.0157	−0.2321	0.0125	0.2524
DGP2							
γ	0.6	0.6000	0.0003	0.0003	0.5711	0.0042	0.0292
β	0.9	0.8973	0.0130	0.0132	1.2990	0.0101	0.3991
ψ	−0.01	−0.0100	4.8158×10^{-6}	4.8302×10^{-6}	−0.0645	0.0015	0.0545
α	0.07	0.0727	0.0130	0.0132	−0.2720	0.0095	0.3421
DGP3							
γ	0.6	0.5993	0.0054	0.0054	0.5378	0.0139	0.0637
β	0.93	0.9423	0.0185	0.0222	1.2116	0.0123	0.2819
ψ	−0.03	−0.0300	0.0002	0.0002	−0.0883	0.0025	0.0583
α	0.02	0.0077	0.0184	0.0222	−0.1940	0.0123	0.2143

Table 1 shows the sample means, standard deviations, and the root mean squared errors (RMSEs) of the posterior means of the DN-PAR(1,1) and CN-PAR(1,1) models. The sample means for the DN-PAR are close to the corresponding true values. The RMSEs are close to the

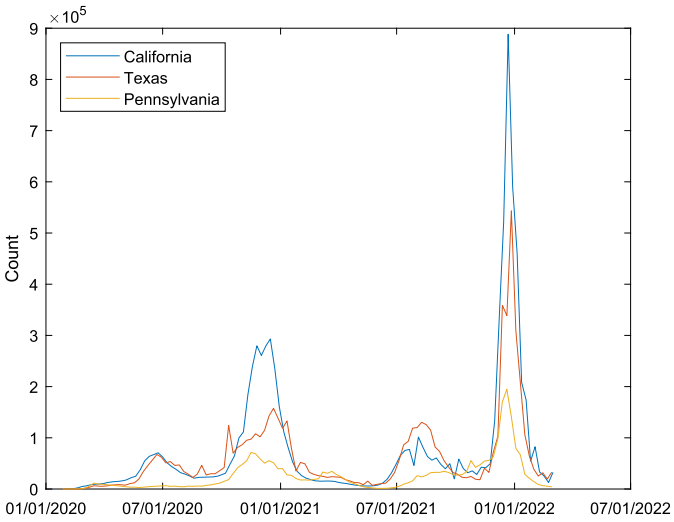

Figure 2: Confirmed cases for three states.

standard deviations, implying that bias is negligible except for DGP3. For DGP3 the sample mean of the estimates of β has small upward bias, while that of α has the small downward bias. A comparison with DGP1 and DGP2 implies that the size of ψ affects the bias in (α, β) when $|\alpha| + |\beta| + |\psi|$ is close to one. The bias will disappear, as the sample size T increases.

In contrast, the estimates for the CN-PAR model suffer from bias, which stems from the model assumption of a constant network when the DGP accommodates the dynamic linkage variables. The sample mean of the estimated posterior means for γ features upward bias. For the remaining parameters, the sample means of the estimates are far from the true values.

The results indicate that our Bayesian MCMC technique is satisfactory under $(N, T) = (51, 80)$. The misspecification for the network structure causes severe bias when estimating posterior means.

The statistical inference of high-dimensional models for count data is challenging as there may involve a large number of parameters in the models. We study the use of the network linkages as auxiliary information to formulate the DN-PAR model to capture potential dependence among the high-dimensional count observations. In practice, we need to use the DN-PAR model when the observed linkage variables are time-varying. In this case, assuming constant network model causes a serious bias. As discussed in the previous section, we need to consider the non-network PAR, CN-PAR, and EN-PAR models when the network is unobservable, in order to select a preferable model. For estimating the DN-PAR model, the Bayesian MCMC method is satisfactory. As p and q increase, the computational complexity in the frequentist's approach increases substantially. On the contrary, the computational complexity is still acceptable in our MCMC technique when p and q increase.

4 Empirical Analysis

Denote $y_{i,t}$ as the number of confirmed COVID-19 cases for i-th state in the United States at week t for the period starting from the fifth week of January 2020 to the fourth week of March 2022, with 113 observations for $N = 51$ states, including the District of Columbia. Figure 2 shows the confirmed cases for three states: California, Texas, and Pennsylvania. We can observe a pattern

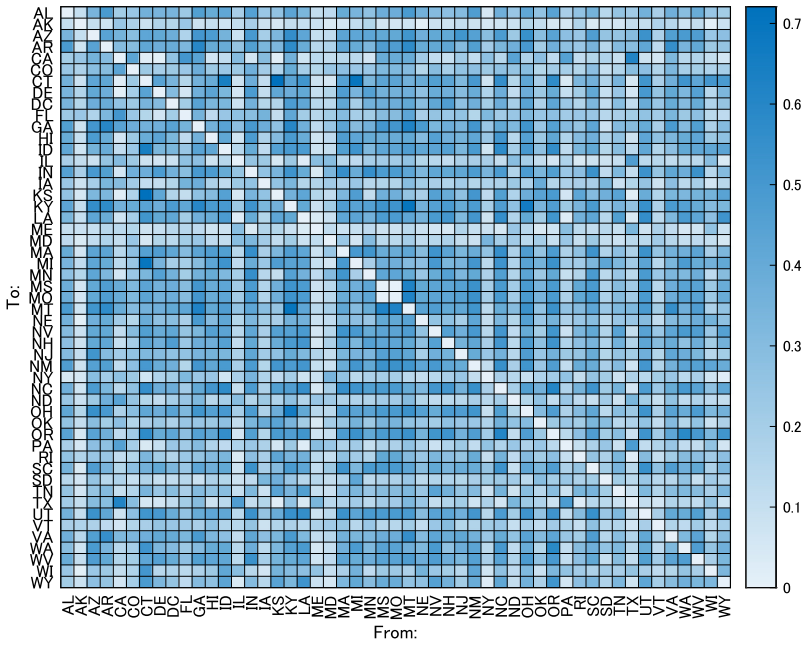

Figure 3: Sample mean for the dynamic network. The heatmap shows the values of $T^{-1}\sum_{t=1}^{T} g_{ij,t}$ for $i, j = 1, \ldots, N$.

in the three states, and can find similar patterns across the 51 states. We use three types of linkage variables: (a) the dynamic linkage provided by the method of So et al. (2021b) based on changes in the count; (b) the dynamic linkage extracted by the approach of So et al. (2021a) as explained in Subsection 2.2; and (c) the constant linkage given in Subsection 3.2. The dynamic linkage in (a) is based on links between states specified by the level of 'co-movement' of newly confirmed COVID-19 cases. Following So et al. (2021b), the correlations between the change in the square root of the counts, $\sqrt{y_{i,t}} - \sqrt{y_{i,t-1}}$, of two states are used to define the dynamic linkage $g_{ij,t}$. For (b), we use the moving-window correlation of $y^*_{i,t} = \log(y_{i,t} + 1)$ instead of the change in the square root of counts to form the dynamic linkage. It is interesting to compare the effects of different linkages when modeling the number of confirmed COVID-19 cases and the forecasting performance. Furthermore, we can consider (c), defined in Subsection 3.2 with the full-sample correlations of $y^*_{i,t}$, as a substitute for the PNAR model. Figure 3 is a heatmap that shows the values of $T^{-1}\sum_{t=1}^{T} g_{ij,t}$ $(i, j = 1, \ldots, N)$ for the dynamic linkage variables in (a). For example, there is a clear edge between California and Texas, but there is no edge between California and Delaware.

There are two main areas of focus in the empirical analysis: investigating how well the three count models with network linkages fit the number of confirmed COVID-19 cases, and their forecasting performance. We consider forecasting in the last 30 weeks, using the rolling window with size $T = 80$. The first three observations are used for initial values. We examine the in-sample and forecasting performances for our DN-PAR(1,1) model.

The competitive models are as follows:

(i) DN-PAR(1,1) model with $\boldsymbol{\gamma}_1 = 0$ in equation (2);
(ii) EN-PAR(1,1) model with $\boldsymbol{\gamma}_1 = 0$;
(iii) CN-PAR(1,1) model with $\boldsymbol{\gamma}_1 = 0$;

Table 2: Posterior Quantities for Network PAR Models. The values in brackets show the 95% credible interval, calculated using the 2.5th and 97.5th percentiles of the simulated draws. Ave. INEF is the average of inefficiency factors.

Parameter	DN-PAR(1,1)	EN-PAR(1,1)	CN-PAR(1,1)
γ	0.6087	0.6214	0.7023
	[0.6046, 0.6129]	[0.6169, 0.6261]	[0.6968, 0.7078]
β	0.9308	0.9317	0.9348
	[0.9299, 0.9318]	[0.9307, 0.9326]	[0.9338, 0.9357]
ψ	−0.0094	−0.0084	−0.0251
	[−0.0098, −0.0091]	[−0.0088, −0.0081]	[−0.0259, −0.0244]
α	0.0196	0.0166	0.0200
	[0.0187, 0.0206]	[0.0157, 0.0175]	[0.01909, 0.0209]
Ave. INEF	3.5133	3.5081	3.5124

(iv) DN-PAR(1,1)-I. DN-PAR(1,1) model with $\gamma_0 = 0$ and with individual effects, $\boldsymbol{\gamma}_1' \boldsymbol{w}_i$, where $\boldsymbol{w}_i = (0, \ldots, 0, 1, 0, \ldots, 0)'$ is an $N \times 1$ unit vector with 1 in the ith element;
(v) EN-PAR(1,1)-I. EN-PAR(1,1) model with individual effects;
(vi) CN-PAR(1,1)-I. CN-PAR(1,1) model with individual effects.

Note that the EN-PAR and CN-PAR models are defined in Subsections 2.2 and 3.2, respectively. We set $m = 6$ for the EN-PAR model, to model the correlation under six months. The three models, DN-PAR, EN-PAR, and CN-PAR, correspond to the three types of network linkage variables we adopt. The major difference between {(i),(ii),(iii)} and {(iv),(v),(vi)} stems from the individual effects, since $\boldsymbol{\gamma}_1' \boldsymbol{w}_i = \gamma_{1i}$. Compared with (i) and (iv), models (ii) and (v) respectively accommodate the extracted linkage variables. Similarly, models (iii) and (vi) are based on the constant network.

Table 2 shows the posterior quantities for the DN-PAR, EN-PAR, and CN-PAR models. We first check the result for the DN-PAR(1,1) model. All 95% credible intervals are apart from zero. The posterior mean for $\alpha + \beta$ is 0.950, implying there is a high level of persistence in $\log \lambda_{i,t}$. For the network effect, $\psi n_{i,t-1}^{-1} \sum_{j=1}^{N} g_{ij,t-1} y_{j,t-1}^*$, the posterior mean for ψ is negative, implying that an increase in the confirmed cases in other states decreases the number one-step-ahead. Compared to the value of γ for the DN-PAR model, the posterior means for the EN-PAR and CN-PAR models are greater in the magnitude. The posterior means of $|\alpha| + |\beta| + |\psi|$ for the DN-PAR, EN-PAR, and CN-PAR models are 0.9598, 0.9567, and 0.9799, respectively. While DN-PAR and EN-PAR model has a similar value, CN-PAR model has a higher persistence as in the Monte Carlo experiment. Although the extracted network structure mimics the network structure based on So et al. (2021b), there are still non-negligible difference in the posteriors. The averages of the inefficiency factors are similar in the three models.

Table 3 presents the posterior quantities for the DN-PAR, EN-PAR, and CN-PAR models with individual effects. We first check the results for the DN-PAR(1,1)-I model. All 95% credible intervals are apart from zero. Individual effects for California and Texas are similar. On the other hand, the individual effect of Pennsylvania is smaller. The value of $\alpha + \beta$ is 0.891, which is smaller than the value for the DN-PAR model. The posterior mean for the network effect is negative. For the EN-PAR(1,1)-I and CN-PAR(1,1)-I models, the individual effects are similar to those of the corresponding values for the DN-PAR(1,1)-I model. The posterior mean for the network

Table 3: Posterior quantities for the network PAR models with individual effects. Although $\boldsymbol{\gamma}_1$ has 51 elements, we omitted remaining results to save space. The values in brackets show the 95% credible interval, calculated using the 2.5th and 97.5th percentiles of the simulated draws. Ave. INEF is the average of inefficiency factors.

Parameter	DN-PAR(1,1)-I	EN-PAR(1,1)-I	CN-PAR(1,1)-I
$\gamma_{1,5}$ (CA)	1.3073	1.29959	1.3139
	[1.2715, 1.3480]	[1.2637, 1.3407]	[1.2761, 1.3576]
$\gamma_{1,39}$ (PA)	1.1237	1.1047	1.0474
	[1.0775, 1.1788]	[1.0584, 1.1599]	[1.0005, 1.1044]
$\gamma_{1,44}$ (TX)	1.2990	1.2857	1.2916
	[1.2566, 1.3464]	[1.2432, 1.3336]	[1.2483, 1.3425]
β	0.8746	0.8694	0.8294
	[0.8658, 0.8830]	[0.8605, 0.8781]	[0.8207, 0.8385]
ψ	−0.0018	0.0069	0.0826
	[−0.0055, −0.0029]	[0.0039, 0.0110]	[0.0758, 0.0869]
α	0.0162	0.0156	−0.0040
	[0.0056, 0.0268]	[0.0050, 0.0261]	[−0.0140, 0.0063]
Ave. INEF	3.1447	3.1448	3.1448

Table 4: DIC. The entries are divided by 10^6.

Model	DIC	Std.Dev.	Min	Max
DN-PAR(1,1)	−5.19255	0.0000066	−5.19255	−5.19250
EN-PAR(1,1)	−5.19385*	0.0000067	−5.19386	−5.19381
CN-PAR(1,1)	−5.18666	0.0000069	−5.18667	−5.18662
DN-PAR(1,1)-I	−5.15911	0.0394659	−5.25794	−5.11031
EN-PAR(1,1)-I	−5.15160	0.0393267	−5.25017	−5.10308
CN-PAR(1,1)-I	−5.07890	0.0393080	−5.17760	−5.03007

effect is positive in the two models. The posterior mean of α for the EN-PAR-I model is similar to that of the DN-PAR-I model. On the other hand, the 95% credible interval for α contains zero for the CN-PAR-I model, implying there is misspecification involved in the CN-PAR-I model.

Table 4 provides the results for the DIC. Among the six models, the EN-PAR model has the smallest DIC, while the CN-PAR-I model has the largest DIC. The data prefer the dynamic network structure over the constant network, and favor the model without individual effects. The in-sample results in Tables 2, 3, and 4 show that the EN-PAR(1,1) model, having the lowest DIC, captures the dynamic of the number of confirmed COVID-19 cases better than the other models. On the other hands, the two models with constant network linkages, CN-PAR(1,1) and CN-PAR(1,1)-I, are the least preferable, according to their highest DIC. The above in-sample results indicate the superiority of using dynamic network linkages when modeling the count data. We also investigate below the models' forecasting performance.

For the last 30 weeks, we obtain the one-step-ahead forecasts of $y_{i,t}$ for $t = T+1, \ldots, T+30$. For example, we use the data up to week T to forecast $y_{i,T+1}$. Since $E(y_{i,t}|\lambda_{i,t}) = \lambda_{i,t}$, we can use

Table 5: Root mean squared forecast error.

Model	RMSFE	Ranking
DN-PAR(1,1)	198437.6*	1
EN-PAR(1,1)	201657.1	2
CN-PAR(1,1)	207009.0	5
DN-PAR(1,1)-I	202689.8	3
EN-PAR(1,1)-I	203027.9	4
CN-PAR(1,1)-I	210129.1	6

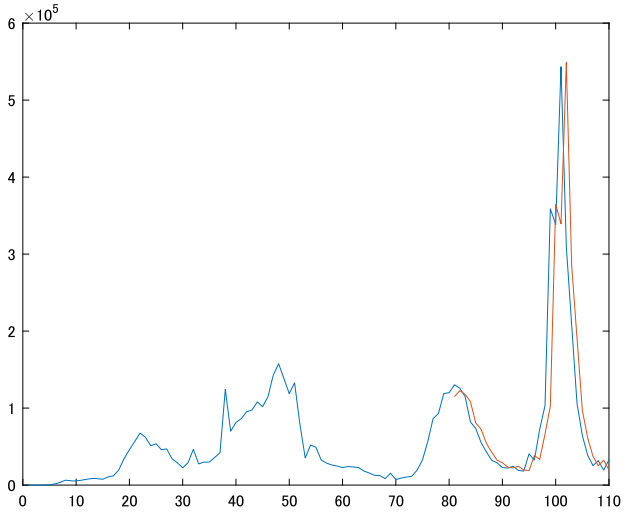

Figure 4: One-step-ahead forecasts for confirmed cases in Texas.

$\hat{\lambda}_{i,T+1}$, an estimate of $\lambda_{i,t}$, as a forecast of $y_{i,T+1}$, where

$$\hat{\lambda}_{i,T+1} = \frac{1}{R}\sum_{r=1}^{R}\exp\left(\gamma_0^{(r)} + \boldsymbol{\gamma}_1^{(r)\prime}\boldsymbol{w}_i + \alpha^{(r)}\log\lambda_{i,T} + \beta^{(r)}y_{i,T}^{*} + \psi^{(r)}n_{i,T}^{-1}\sum_{j=1}^{N}g_{ij,T}y_{j,T}^{*}\right),$$

with the number of the DRAM draws, R. We set $R = 1000$ for the out-of-sample forecasts. Table 5 presents the root mean squared forecast errors (RMSFEs) for the six models, along with their rankings. The DN-PAR has the smallest RMSFE, followed by the EN-PAR model. The constant network models have the largest RMSFE. Figure 4 shows the one-step-ahead forecasts for Texas ($i = 44$). The forecasts capture the fluctuations in the observed values, with a one-step delay, based on the nature of the model.

The empirical results from the DIC and the out-of-sample forecast indicate that the models with dynamic network information embedded are preferred over the models with constant network information. This finding implies that extracted linkage variables are useful when the true network structure is unknown.

5 Conclusion

In this paper, we develop a DN-PAR model that can accommodate dynamic network structures when multivariate count data are involved. When the relevant network between counts is unobservable, we suggest extracting time-varying network information to build the DN-PAR model. We introduce a DRAM algorithm for estimating the new model. The Monte Carlo results show that the bias of the estimated posterior means is negligible, and that the constant network model produces bias in the estimates. We apply the DN-PAR model with both dynamic network and constant network linkages when modeling the number of confirmed COVID-19 cases in the United States. The empirical results show that the models with dynamic network linkages have smaller DIC and RMSFE than the model with constant network linkage. Our DN-PAR methodology showcases the importance of integrating dynamic network linkage information when modeling multivariate count data.

There are several important extensions for the DN-PAR model. First, we can incorporate other statistical properties of the count data, such as, over-dispersion in the modeling by using alternative distributions to Poisson. Second, we may accommodate threshold autoregression to capture asymmetric network linkage effects. Third, it may be useful to develop a frequentist approach for estimating the model. The above extensions warrant further research on multivariate count data in the future.

Supplementary Material

Programming code can be found at https://github.com/ManabuAsai/Dynamic_network_poisson.

Acknowledgement

The authors are most grateful to Yoshihisa Baba and two anonymous reviewers for their very helpful comments and suggestions.

Funding

The first author acknowledges the financial support of the Japan Society for the Promotion of Science (grant number 22KK0022). This work was partially supported by The Hong Kong University of Science and Technology research grant "Risk Analytics and Applications" (grant number SBMDF21BM07). The funding recipient was MKPS.

References

Aktekin T, Polson N, Soyer R (2018). Sequential Bayesian analysis of multivariate count data. *Bayesian Analysis*, 13: 385–409.

Amillotta M, Fokianos K, Krikidis I (2022). Generalized linear models network autoregression. In: *Network Science: Proceedings of the 7th International Winter Conference, NetSci-X 2022* (P Ribeiro, F Silva, JF Mendes, R Laureano, eds.), In: *Lecture Notes in Computer Science*, 112–125. Network Science Society, Springer, Porto, Portugal.

Armillotta M, Fokianos K (2021). Poisson network autoregression. arXiv preprint: https://arxiv.org/abs/2104.06296.

Armillotta M, Fokianos K (2022). Testing linearity for network autoregressive models. arXiv preprint: https://arxiv.org/abs/2202.03852.

Bartolucci F, Pennoni F, Mira A (2021). A multivariate statistical approach to predict Covid-19 count data with epidemiological interpretation and uncertainty quantification. *Statistics in Medicine*, 40(24): 5351–5372. https://doi.org/10.1002/sim.9129

Berry LR, West M (2020). Bayesian forecasting of many count-valued time series. *Journal of Business & Economic Statistics*, 38(4): 872–887. https://doi.org/10.1080/07350015.2019.1604372

Chan S, Chu J, Zhang Y, Nadarajah S (2021). Count regression models for COVID-19. *Physica A: Statistical Mechanics and its Applications*, 563: 125460. https://doi.org/10.1016/j.physa.2020.125460

Chen CWS, Khamthong K, Lee S (2019). Markov switching integer-valued generalized autoregressive conditional heteroscedastic models for dengue counts. *Journal of the Royal Statistical Society. Series C. Applied Statistics*, 68(4): 963–983. https://doi.org/10.1111/rssc.12344

Chen CWS, Lee S (2016). Generalized Poisson autoregressive models for time series of counts. *Computational Statistics & Data Analysis*, 99: 51–67. https://doi.org/10.1016/j.csda.2016.01.009

Chen CWS, Lee S, Khamthong K (2021). Bayesian inference of nonlinear hysteretic integer-valued garch models for disease counts. *Computational Statistics*, 36(1): 261–281. https://doi.org/10.1007/s00180-020-01018-7

Chen CWS, So MKP, Li JC, Sriboonchitta S (2016). Autoregressive conditional negative binomial model applied to over-dispersed time series of counts. *Statistical Methodology*, 31: 73–90. https://doi.org/10.1016/j.stamet.2016.02.001

Chen CWS, So MKP, Liu FC (2022). Assessing government policies' impact on the COVID-19 pandemic and elderly deaths in East Asia. *Epidemiology and Infection*, 150: e161.

Chu AMY, Chan TWC, So MKP, Wong WK (2021). Dynamic network analysis of COVID-19 with a latent pandemic space model. *International Journal of Environmental Research and Public Health*, 18(6): 3195. https://doi.org/10.3390/ijerph18063195

Chu AMY, Chong ACY, Lai NHT, Tiwari A, So MKP (2023). Enhancing the Predictive Power of Google Trends Data Through Network Analysis: Infodemiology Study of COVID-19. *JMIR Public Health and Surveillance*, 9(1): e42446. https://doi.org/10.2196/42446

Debaly ZM, Truquet L (2019). Stationarity and moment properties of some multivariate count autoregressions. arXiv preprint: https://arxiv.org/abs/1909.11392.

Debaly ZM, Truquet L (2023). Multivariate time series models for mixed data. *Bernoulli*, 29(1): 669–695. https://doi.org/10.3150/22-BEJ1474

El-Morshedy M, Altun E, Eliwa M (2022). A new statistical approach to model the counts of novel coronavirus cases. *Mathematical Sciences*, 16: 37–50. https://doi.org/10.1007/s40096-021-00390-9

Fokianos K (2021). Multivariate count time series modelling. *Econometrics and Statistics*. https://doi.org/10.1016/j.ecosta.2021.11.006

Fokianos K, Fried R, Kharin Y, Voloshko V (2022). Statistical analysis of multivariate discrete-valued time series. *Journal of Multivariate Analysis*, 188: 104805. https://doi.org/10.1016/j.jmva.2021.104805

Fokianos K, Rahbek A, Tjøstheim D A Tjøstheim D R (2009). Poisson autoregression. *Journal of the American Statistical Association*, 104(488): 1430–1439. https://doi.org/10.1198/jasa.2009.tm08270

Fokianos K, Støve B, Tjøstheim D, Doukhan P (2020). Multivariate count autoregression. *Bernoulli*, 26(1): 471–499. https://doi.org/10.3150/19-BEJ1132

Fokianos K, Tjøstheim D (2011). Log-linear Poisson autoregression. *Journal of Multivariate Analysis*, 102(3): 563–578. https://doi.org/10.1016/j.jmva.2010.11.002

Frühwirth-Schnatter S, Wagner H (2006). Auxiliary mixture sampling for parameter-driven models of time series of counts with applications to state space modelling. *Biometrika*, 93(4): 827–841. https://doi.org/10.1093/biomet/93.4.827

Green PJ, Mira A (2001). Delayed rejection in reversible jump Metropolis-Hastings. *Biometrika*, 88: 1035–1053. https://doi.org/10.1093/biomet/88.4.1035

Haario H, Laine M, Mira A, Saksman E (2006). Dram: Efficient adaptive MCMC. *Statistics and Computing*, 16: 339–354. https://doi.org/10.1007/s11222-006-9438-0

Haario H, Saksman E, Tamminen J (1999). Adaptive proposal distribution for random walk Metropolis algorithm. *Computational Statistics*, 14: 375–395. https://doi.org/10.1007/s001800050022

Haario H, Saksman E, Tamminen J (2001). An adaptive Metropolis algorithm. *Bernoulli*, 7: 223–242. https://doi.org/10.2307/3318737

Held L, Höhle M, Hofmann M (2005). A statistical framework for the analysis of multivariate infectious disease surveillance counts. *Statistical Modelling*, 5(3): 187–199. https://doi.org/10.1191/1471082X05st098oa

Karlis D (2016). Models for multivariate count time series. In: Davis RA, Holan SH, Lund R, Ravishanker N, eds., *Handbook of Discrete-Valued Time Series*, 407–424.

Livsey J, Lund R, Kechagias S, Pipiras V (2018). Multivariate integer-valued time series with flexible autocovariances and their application to major hurricane counts. *Annals of Applied Statistics*, 12(1): 408–431. https://doi.org/10.1214/17-AOAS1098

Martínez-Bello DA, López-Quílez A, Torres-Prieto A (2017). Bayesian dynamic modeling of time series of dengue disease case counts. *PLoS Neglected Tropical Diseases*, 11(7): e0005696. https://doi.org/10.1371/journal.pntd.0005696

Mira A (2001). On Metropolis-Hastings algorithms with delayed rejection. *Metron*, 59: 231–241.

Nikoloulopoulos AK, Karlis D (2009). Modeling multivariate count data using copulas. *Communications in Statistics. Simulation and Computation*, 39(1): 172–187. https://doi.org/10.1080/03610910903391262

Panjer HH (2006). *Operational Risk: Modeling Analytics*. John Wiley & Sons, Hoboken, New Jersey.

Polwiang S (2020). The time series seasonal patterns of dengue fever and associated weather variables in Bangkok (2003-2017). *BMC Infectious Diseases*, 20: 208. https://doi.org/10.1186/s12879-020-4902-6

Ravishanker N, Serhiyenko V, Willig MR (2014). Hierarchical dynamic models for multivariate times series of counts. *Statistics and its Interface*, 7(4): 559–570. https://doi.org/10.4310/SII.2014.v7.n4.a11

Reasenberg P, Jones L (1994). Earthquake aftershocks: Update. *Science*, 265(5176): 1251–1252. https://doi.org/10.1126/science.265.5176.1251

Shinde GR, Kalamkar AB, Mahalle PN, Dey N, Chaki J, Hassanien AE (2020). Forecasting models for coronavirus disease (COVID-19): A survey of the state-of-the-art. *SN Computer Science*, 1: 197. https://doi.org/10.1007/s42979-020-00209-9

So MKP, Chan LSH, Chu AMY (2021a). Financial network connectedness and systemic risk during the COVID-19 pandemic. *Asia-Pacific Financial Markets*, 28(4): 649–665. https://doi.org/10.1007/s10690-021-09340-w

So MKP, Chu AMY, Tiwari A, Chan JNL (2021b). On topological properties of COVID-19: Predicting and assessing pandemic risk with network statistics. *Scientific Reports*, 11(1): 1–14. https://doi.org/10.1038/s41598-020-79139-8

So MKP, Tiwari A, Chu AMY, Tsang JTY, Chan JNL (2020). Visualizing COVID-19 pandemic risk through network connectedness. *International Journal of Infectious Diseases*, 96: 558–561. https://doi.org/10.1016/j.ijid.2020.05.011

Soyer R, Zhang D (2022). Bayesian modeling of multivariate time series of counts. *Wiley Interdisciplinary Reviews: Computational Statistics*, 14(6): e1559. https://doi.org/10.1002/wics.1559

Tierney L, Mira A (1999). Some adaptive Monte Carlo methods for Bayesian inference. *Statistics in Medicine*, 18: 2507–2515. https://doi.org/10.1002/(SICI)1097-0258(19990915/30)18:17/18<2507::AID-SIM272>3.0.CO;2-J

Wang Z, Chakraborty P, Mekaru SR, Brownstein JS, Ye J, Ramakrishnan N (2015). Dynamic Poisson autoregression for influenza-like-illness case count prediction. In: *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD'15* (Cao L, Zhang C, eds.), 1285–1294. Association for Computing Machinery, Association for Computing Machinery, Sydney, NSW, Australia.

West M (2020). Bayesian forecasting of multivariate time series: Scalability, structure uncertainty and decisions. *Annals of the Institute of Statistical Mathematics*, 72: 1–31. https://doi.org/10.1007/s10463-019-00741-3

West M, Harrison J (2006). *Bayesian Forecasting and Dynamic Models*. Springer, New York, NY.

West M, Harrison PJ, Migon HS (1985). Dynamic generalized linear models and Bayesian forecasting. *Journal of the American Statistical Association*, 80(389): 73–83. https://doi.org/10.1080/01621459.1985.10477131

Wu WB, Shao X (2004). Limit theorems for iterated random functions. *Journal of Applied Probability*, 41(2): 425–436. https://doi.org/10.1239/jap/1082999076

Zhang MY, Russell JR, Tsay RS (2001). A nonlinear autoregressive conditional duration model with applications to financial transaction data. *Journal of Econometrics*, 104(1): 179–207. https://doi.org/10.1016/S0304-4076(01)00063-X

Zhang Y, Zhou H, Zhou J, Sun W (2017). Regression models for multivariate count data. *Journal of Computational and Graphical Statistics*, 26(1): 1–13. https://doi.org/10.1080/10618600.2016.1154063

Zhu X, Pan R, Li G, Liu Y, Wang H (2017). Network vector autoregression. *The Annals of Statistics*, 45(3): 1096–1123. https://doi.org/10.1214/16-AOS1476