The Annals of Applied Statistics
2020, Vol. 14, No. 3, 1493–1515
https://doi.org/10.1214/20-AOAS1360

ADAPTIVE LOG-LINEAR ZERO-INFLATED GENERALIZED POISSON AUTOREGRESSIVE MODEL WITH APPLICATIONS TO CRIME COUNTS

BY XIAOFEI XU[1], YING CHEN[2], CATHY W. S. CHEN[3] AND XIANCHENG LIN[4]

[1]*Department of Mathematics, National University of Singapor*
[2]*Risk Management Institute, National University of Singapore*
[3]*Department of Statistics, Feng Chia University, chenws@mail.fcu.edu.tw*
[4]*Department of Statistics, University of California, Davis*

This research proposes a comprehensive ALG model (Adaptive Log-linear zero-inflated Generalized Poisson integer-valued GARCH) to describe the dynamics of integer-valued time series of crime incidents with the features of autocorrelation, heteroscedasticity, overdispersion and excessive number of zero observations. The proposed ALG model captures time-varying nonlinear dependence and simultaneously incorporates the impact of multiple exogenous variables in a unified modeling framework. We use an adaptive approach to automatically detect subsamples of local homogeneity at each time point of interest and estimate the time-dependent parameters through an adaptive Bayesian Markov chain Monte Carlo (MCMC) sampling scheme. A simulation study shows stable and accurate finite sample performances of the ALG model under both homogeneous and heterogeneous scenarios. When implemented with data on crime incidents in Byron, Australia, the ALG model delivers a persuasive estimation of the stochastic intensity of criminal incidents and provides insightful interpretations on both the dynamics of intensity and the impacts of temperature and demographic factors for different crime categories.

1. Introduction. The measurement of crime intensities is an important topic across multiple disciplines, varying from criminology and archaeology to economics and political science. It is also necessary to understand the dynamics of criminal incidents over time and investigate potential factors triggering high levels of crime in order to assist officers and agencies in the field at apprehending criminals and suppressing incidences of crime. There is an abundant amount of literature in this rapidly growing area; see, for example, weather and crime (Anderson et al. (2000), Cohn (1990), Ranson (2014), Mares and Moffett (2016)), crime seasonality and temporal variations (McDowall, Loftin and Pate (2012), Pereira, Andresen and Mota (2016), Yan (2004)) and dynamics modeling of crime (Chen and Lee (2017), Chen et al. (2016), Famoye and Singh (2006), Lee, Lee and Chen (2016)) among many others. This paper investigates the nonlinear dependence of several categories of crime in Byron, which is a local government area in the Northern Rivers region of New South Wales (NSW), Australia.

Various studies in the literature document the impact of exogenous factors on criminal behavior. For example, Anderson et al. (2000) find that hotter temperatures lead to increases in rapes, assaults and domestic violence across 260 U.S. cities. Hsiang, Burke and Miguel (2013) present a mean effect for a 2.3% increase in interpersonal violence for each standard deviation increase in temperature; see also Anderson (2001), Rotton and Cohn (2000) and Hird and Ruparel (2007). Moreover, Mares (2009) presents that the effect of economic change and civilizing process on interpersonal violence trends is on a long-term decline.

Received October 2019; revised April 2020.
Key words and phrases. Bayesian, integer-valued GARCH model, excess zeros, overdispersion, nonstationarity, MCMC.

Stacey, Carbone-López and Rosenfeld (2011) address the influence of a demographic change in immigration of anti-Hispanic hate crime in the United States. Famoye and Singh (2006) investigate the relationship between domestic violence and level of education, employment status and income of victims and batterers. Though insightful, the aforementioned works do not analyze the dynamics of factors' effects and/or only focus on one or a single type of factors.

Data on crime counts are discrete and integer valued, often exhibiting unique features including overdispersion, serial correlation and excess zeros and could be influenced by exogenous factors such as weather and socioeconomic status, as stated above. Studies have used Poisson regression to model count time series, where the variance and mean have to be identical, but this is not suitable for over-dispersed crime counts with a larger variance than the mean. The Poisson models are further generalized with, for example, generalized Poisson (GP), negative binomial (NB), zero-inflated Poisson (ZIP) and exponential family distributions; see Consul and Famoye (1992), Sellers and Shmueli (2010) and Zhu (2012a).

Among others, Ferland, Latour and Oraichi (2006) and Fokianos, Rahbek and Tjøstheim (2009) propose the Poisson integer-valued generalized autoregressive conditional heteroscedastic model (P-INGARCH) that allows the conditional mean (namely, the intensity), which is also the conditional variance, to be heteroscedastic and be able to handle overdispersion as well. The literature extensively develops this model to flexibly capture features in specific situations, such as nonlinearity and zero inflation; see, for example, log-linear P-INGACRH (Fokianos and Fried (2012), Fokianos and Tjøstheim (2011)) and GP/NB/ZIP/ZIGP-INGARCH models (Famoye and Singh (2006), Jazi, Jones and Lai (2012), Zhu (2011, 2012b)). Fokianos and Tjøstheim (2012) generalize the P-INGARCH type model to a nonlinear framework and study the geometric ergodicity properties. Coping with effects of exogenous covariates, studies have established the generalized linear model (Davis, Dunsmuir and Streett (2003)), linear and nonlinear INGARCHX type models where "X" refers to one or more eXogenous variables in the conditional mean equation (Agosto et al. (2016), Chen, Khamthong and Lee (2019), Chen and Lee (2017), Chen and Khamthong (2019)). Some existing works extend to bivariate or multivariate integer-valued autoregressive models; see Karlis and Pedeli (2013), Cui and Zhu (2018), Cui, Li and Zhu (2020) and Fokianos et al. (2020).

The aforementioned works derive models and estimations in a stationary framework where the coefficients are forced to be constant. The dynamics of crime counts often shift transitorily or permanently, driven by various factors such as governmental policy changes and critical events' occurrences. Lee, Lee and Chen (2016) test the existence of a structural (parameter) change in the modeling of crime counts of robbery and assault police in inner Sydney. In the existing literature of nonstationary count time series analysis, the focus is on structural break detection or estimation via a test or Bayesian approach. Fokianos and Fried (2010) develop a testing procedure for the detection of intervention effects within the framework of INGARCH models. Franke, Kirch and Kamgaing (2012) and Kang and Lee (2014) propose a CUSUM test to locate change points in a Poisson autoregressive model. Chen and Lee (2016) propose a Bayesian method to identify locations of structural transitions for time series of counts with zero inflation, while Chen et al. (2016) introduce dynamic overdispersion into the INGARCH models. Alternatively, approximating nonstationary time series via a locally stationary process with a piecewise constant or more generally time-varying coefficients is also prevalent in the time series literature; see, for example, Mercurio and Spokoiny (2004), Davis, Lee and Rodriguez-Yam (2006), Spokoiny (2009) and Chen and Li (2017). In fact, researchers have yet to examine the dynamic modeling of a local stationary integer-valued time series.

This paper proposes an Adaptive Log-linear zero-inflated Generalized Poisson INGARCHX (ALG) model to investigate the dynamics of nonstationary count time series

with unique features of autocorrelation, heteroscedasticity, overdispersion and excess zeros in a unified framework. We conduct a comprehensive analysis to the dynamics of several crimes and simultaneously investigate the impacts of multiple covariates, including temperature, seasonality, population and unemployment rate. Our study represents nonstationarity in the dynamics of crime count series with time-dependent coefficients. Instead of detecting the exact time and location of breakpoints, at each time we automatically detect a local interval of a count series over which the process is approximately stationary. We derive a local adaptive estimation procedure in the Bayesian framework based on the MCMC technique without requiring prior knowledge of the location and type of breaks. The ALG model is flexible and can take on both stationary and nonstationary count series to decipher the essential dynamic evolution. We perform simulation studies to investigate the model's finite sample performance under various scenarios and demonstrate the application of the ALG model on six types of crime incidents from January 1995 to December 2017 in Byron, NSW, Australia.

Our contributions include the following: (1) We propose an ALG model to account for a number of features in a unified framework and simultaneously incorporate the impact of multivariate exogenous covariates. In comparison, existing works consider only a part of the features and/or do so under stationarity. (2) We develop an adaptive procedure in the Bayesian framework that can automatically react to unforeseeable structural breaks, while the existing literature either assumes stationarity or requires piecewise stationarity. Our work adaptively estimates parameters and local interval of homogeneity at each time point, and both vary over time. (3) We provide an interpretable estimation of the stochastic intensity of monthly crime data and the time-dependent impacts from multivariate environmental and demographic variables.

The rest of the paper runs as follows. Section 2 describes the data on crime incidents in Byron, Australia. Section 3 presents the ALG model and derives the adaptive estimation procedure using the Bayesian MCMC sampling scheme under local homogeneity assumption. Section 4 investigates the finite sample performance of the ALG model under various scenarios. Section 5 implements real data analysis. Section 6 concludes.

2. Data. We consider six types of monthly crime datasets in Byron, NSW, Australia from January 1995 to December 2017, comprising 276 observations. The data come from the NSW Bureau of Crime Statistics and Research website (BOCSAR). The NSW Police Force organizes each dataset by type of offense, month and local government area.

The first three offense categories of "assault: nondomestic violence related assault" (labeled ASS), "malicious damage to property" (MDP) and "theft: steal from person" (TSP) are three of the major personal violence and property offenses according to BOCSAR. Among them, MDP generally denotes the intentional destruction or defacement of public, commercial and private property, including vandalism, trespassing, graffiti, illegal tipping, smashed windows or other defacing acts, where graffiti is the most common form of malicious damage. Crime statistics from the Australian Institute of Criminology indicate that malicious damage is the most commonly reported criminal offense in NSW; see Anderson et al. (2012). We also consider another three crimes of "liquor offenses" (LOS), "against justice procedures: breach bail conditions" (AJP) and "arson" (ARS). In 2012, the recorded rate of LOS in Byron ranked first out of 140 local government areas that have populations greater than 3000. The crime trend of this offense is, naturally, of particular interest for Byron. ARS is the most dangerous crime in this study.

We also obtain three environmental and demographic variables, including monthly mean maximum temperature at Byron (station number: 58198), monthly unemployment rate and quarterly population counts in NSW. All the variables are collected in the same period from the Australia Government Bureau of Meteorology. To investigate the seasonality effect of

TABLE 1
Description and summary statistics of monthly crime data, temperature and demographic variables from January 1995 *to December* 2017 (276 *observations*)

Dataset	Min	Median	Max	Variance	Variance/Mean	0*s* %
ASS	3	21	54	78	3.55	0
MDP	10	30.5	78	119.16	3.72	0
TSP	0	5	34	25.22	4.20	3.99
LOS	0	16	134	461.70	21.99	6.88
AJP	0	3	11	2.30	6.90	13.77
ARS	0	1	11	2.13	2.13	37.68
Temp (°C)	18.4	25	30.5	9.12	0.37	0
URate (%)	4.3	5.6	9.65	1.07	0.18	0
Population($\times 10^6$)	6.09	6.75	7.92	2.64×10^5	38,306	0

ASS—Assault: Nondomestic violence related assault. MDP—Malicious damage to property. TSP—Theft: Steal from person. LOS—Liquor offenses. AJP—Against justice procedures: Breach bail conditions. ARS—Arson. Temp: Monthly mean maximum temperature. URate: Monthly unemployment rate of NSW. Population: Quarterly population of NSW.

crimes, especially the hot season effect as documented in the literature (Anderson et al. (2000), McDowall, Loftin and Pate (2012), Yan (2004)), we fix a dummy variable $D_t = 1$ if the month is November, December or January, which is the summer period in Australia and equal to 0 otherwise. This dummy is to designed to examine the hot (summer) seasonality effect vs. the rest of the year.

Table 1 reports the descriptive statistics of the six categories of monthly crime offenses and three exogenous covariates. The monthly series of ASS and MDP do not have any zero observations. The crime series have 3.99%, 6.88% and 13.77% of zeros, respectively, for TSP, LOS and AJP which is considered as the medium level of zero percentage. There are 104 number counts of zeros among the 276 months (37.68%) for ARS, indicating the feature of excess zeros. The sample variance is larger than the sample mean for all the crime count data (the ratio ranges 2.12 to 21.99), implying the overdispersion feature for all the datasets. Because the exogenous variables vary in scale, the maximum value of the population, for example, is 7.65×10^5, and the maximum value of the variable unemployment rate is 4.3%; the exogenous variables are scaled to be between 0 and 1 by subtracting the minimum and dividing the range for computational convenience. As population is quarterly variable, we allocate the same value to each month within a quarter to use in the model. Figure 1 displays the time series plots, barplots and the ACF plot of two crime categories, MDP and ARS, as illustration along with three exogenous variables, respectively. The graphical demonstration of the rest four datasets refers to Figure A in the Supplementary Material (Xu et al. (2020a)). MDP reveals an increasing trend until 2009 and then a decreasing pattern thereafter, while ARS develops an increasing trend in more recent years. The serial dependence of MDP is, in general, stronger than that of ARS. The values of the monthly crime count series vary from one category to another; yet all datasets present similar complexity in data features. The empirical data features motivate the adoption of a model that can effectively handle zero observations, serial dependence, overdispersion and changing structures. The bottom panel in Figure 1 demonstrates the monthly mean maximum temperature, unemployment rate and population counts, respectively. The temperature series exhibits a strong seasonality pattern with higher values at around January and lower values at around July. The unemployment rate reveals a roughly decreasing trend until 2009 and moves up after that. Population is purely increasing over the sample period.

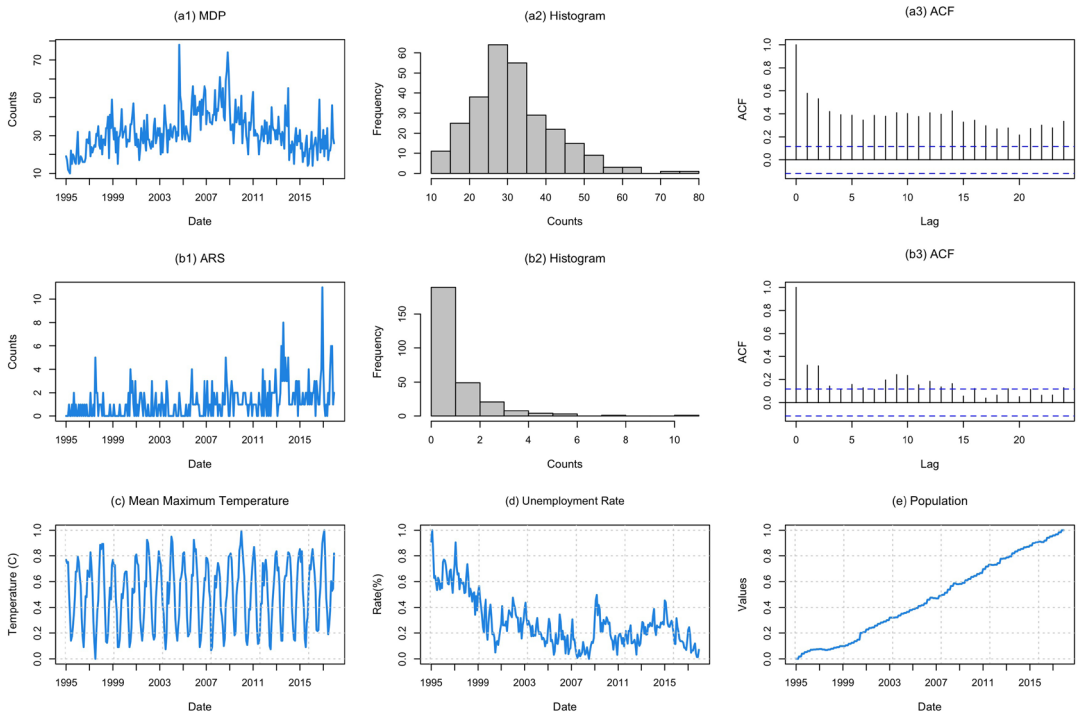

FIG. 1. *Time series plots, barplots and ACFs for MDP (upper) and ARS (middle) at Byron, and the plots of three exogenous covariates (bottom) from January* 1995 *to December* 2017, *Australia. The exogenous covariates are scaled between* 0 *and* 1.

3. ALG model. In this section we introduce the Adaptive Log-linear zero-inflated Generalized Poisson INGARCHX (ALG) model to simultaneously account for both time-varying dynamics of crime count series with multiple features, including autocorrelation, heteroscedasticity, overdispersion and excess zero observations and the impacts of exogenous factors under instability. We then introduce an adaptive Bayesian method based on the MCMC sampling scheme for parameter estimation.

3.1. *The ALG model.* We recall the definition of the ZIGP distribution (Gupta, Gupta and Tripathi (1996)). A random variable Y is said to follow the ZIGP distribution with parameters λ, φ and ρ if the probability mass function is

$$P(Y=y)=\begin{cases}\rho+(1-\rho)e^{-\lambda} & \text{if } y=0,\\ (1-\rho)\lambda(\lambda+\varphi y)^{y-1}e^{-(\lambda+\varphi y)}/y! & \text{if } y=1,2,\ldots,\\ 0 & \text{for } y>m \text{ if } \varphi<0,\end{cases}$$

where $\lambda>0$, $0\le\rho<1$, $\max(-1,-\lambda/m)<\varphi<1$ and $m(\ge 4)$ is the largest positive integer for which $\lambda+\varphi m>0$ when $\varphi<0$. If $\varphi>0$, then the distribution includes overdispersion, whereas underdispersion or no dispersion is present when $\varphi<0$ or $\varphi=0$, given that $\rho=0$. The distribution reduces to the generalized Poisson distribution when $\rho=0$ and to the Poisson distribution when $\rho=\varphi=0$.

Figure 2 displays the probability mass function of ZIGP distribution with different parameters of λ, φ and ρ. Compared to Poisson, a positive ρ significantly increases the probability of zeros, leading to a lower probability of other values to occur; a positive value of φ leads to a lower probability of smaller counts and heavier right tail. When both ρ and φ are positive, we observe a larger probability of zero and a heavier right tail.

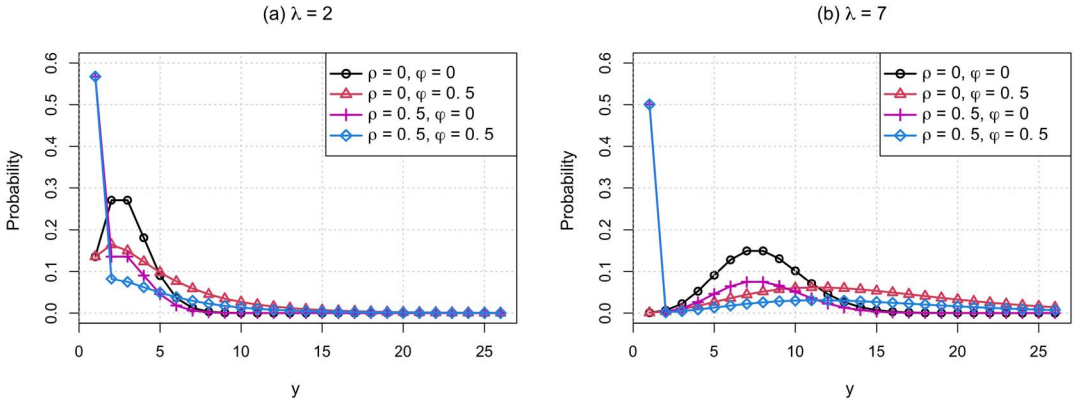

FIG. 2. *Probability mass function of the ZIGP distribution with different parameters of λ, φ and ρ.*

Under ZIGP, the conditional expectation and variance of Y are, respectively,

$$\mathrm{E}(Y) = \frac{1-\rho}{1-\varphi}\lambda \quad \text{and} \quad \mathrm{Var}(Y) = (1-\rho)\left[\frac{\rho\lambda^2}{(1-\varphi)^2} + \frac{\lambda}{(1-\varphi)^3}\right].$$

The variance of Y is greater than the mean if $0 \le \varphi < 1$. When $\rho = 0$, the variance is equal or smaller than the mean if $\varphi = 0$ or $\varphi < 0$.

Let $\{Y_t; t = 1, \ldots, n\}$ denote a count series that is conditionally ZIGP distributed with mean λ_t and log-intensity process $\mu_t = \log(\lambda_t)$, such as the number of monthly criminal offense counts. We model the dynamics of this process in terms of its own past, Y_{t-1} and q exogenous covariates $\boldsymbol{x}_t = (x_{t1}, \ldots, x_{tq})^\top$. In our study the exogenous covariates include the relevant climate and demographic factors of temperature, population, unemployment rate and a summer dummy. We define the ALG model as

$$\begin{aligned} &Y_t | (F_{t-1}^{(y)}, F_t^{(x)}) \sim \mathrm{ZIGP}(\lambda_t^*, \varphi_t, \rho_t), \\ &\lambda_t^* = \frac{1-\varphi_t}{1-\rho_t}\lambda_t, \qquad \mu_t = \log(\lambda_t), \\ &\mu_t = \omega_t + \alpha_t \log(Y_{t-1} + 1) + \beta_t \mu_{t-1} + \sum_{i=1}^{q} \gamma_{ti} x_{ti}, \end{aligned} \tag{3.1}$$

where $0 \le \rho_t < 1$, $\max(-1, -\lambda_t^*/m) < \varphi_t < 1$, $m(\ge 4)$ is again the largest positive integer for which $\lambda_t^* + \varphi_t m > 0$ when $\varphi_t < 0$, and γ_{ti} is a time-dependent real-valued parameter reflecting the effect of covariate x_{ti} for $i = 1, \ldots, q$. Here, $F_t^{(y)}$ is the σ-field generated by $\{Y_t, \ldots, Y_1, \lambda_0\}$, where λ_0 is the initial intensity and $F_t^{(x)}$ is the σ-field generated by $\{\boldsymbol{x}_t, \ldots, \boldsymbol{x}_1\}$ with $\boldsymbol{x}_j = (x_{j1}, x_{j2}, \ldots, x_{jq})$ for $j = 1, \ldots, t$, representing all available past information of exogenous variables. We set $0 \le \varphi_t < 1$ for all time indices for the overdispersion case. The unknown time-varying parameters include $(\omega_t, \alpha_t, \beta_t, \rho_t, \varphi_t, \gamma_{t1}, \ldots, \gamma_{tq})$. For notational simplicity we denote the parameter set as $\boldsymbol{\vartheta}_t = (\boldsymbol{\theta}_t, \boldsymbol{\kappa}_t, \boldsymbol{\gamma}_t)$, where $\boldsymbol{\theta}_t = (\omega_t, \alpha_t, \beta_t)$, $\boldsymbol{\kappa}_t = (\rho_t, \varphi_t)$ and $\boldsymbol{\gamma}_t = (\gamma_{t1}, \ldots, \gamma_{tq})$.

Under stationarity with all parameters constant, if $\boldsymbol{\kappa}_t$ and $\boldsymbol{\gamma}_t$ equal zero, then the ALG model reduces to the log-linear Poisson autoregression model in Fokianos and Tjøstheim (2011). In other words, that model ignores the case with excess zeros and relies on stationarity. The ALG model has time-dependent parameters $\boldsymbol{\vartheta}_t$, allowing for both smooth structural changes and abrupt breaks where no prior information of breaks is required. As it is impossible to estimate local parameters that vary from every time point, we assume at each time that there exists a subinterval over which $\boldsymbol{\vartheta}_t$ can be properly approximated by a constant. Such an

interval is called the local homogeneous interval. In other words, the ALG model does not deviate much from a stationary model with constant parameters over the local interval. As time moves, the local homogeneous interval updates at each point based on past available information. As such, the local intervals have time-dependent length, over which we can estimate the globally nonstationary yet locally stationary dynamics using an adaptive procedure. Note that Chen and Lee (2016) consider the ZIGP-INGARCH model in a nonstationary framework too, yet without exogenous variables and aiming to locate a single structural break.

3.2. *Bayesian inference under stationarity.* Under a stationarity assumption with constant parameters, we drop the subscript t of parameters for all time and have $\boldsymbol{\vartheta}_t = \boldsymbol{\vartheta} = (\boldsymbol{\theta}, \boldsymbol{\kappa}, \boldsymbol{\gamma})$. Let $\mathbf{Y}_t$ and $\mathbf{X}_t$ denote all the past count and exogenous variables' observations at time t, respectively. The conditional posterior for each parameter group is proportional to the likelihood function multiplied by the prior density of that group,

$$p(\boldsymbol{\vartheta}_\ell|\mathbf{Y}_t, \mathbf{X}_t, \boldsymbol{\vartheta}_{\neq\ell}) \propto p(\mathbf{Y}_t|\mathbf{X}_t, \boldsymbol{\vartheta}) p(\boldsymbol{\vartheta}_\ell|\boldsymbol{\vartheta}_{\neq\ell}),$$

where $\boldsymbol{\vartheta}_\ell$ denotes each parameter group with $\ell = 1, 2$ and 3 referring to $\boldsymbol{\theta}$, $\boldsymbol{\kappa}$ and $\boldsymbol{\gamma}$, respectively, $p(\boldsymbol{\vartheta}_\ell)$ is the prior density, and $\boldsymbol{\vartheta}_{\neq\ell}$ is the vector of all the model parameters except $\boldsymbol{\vartheta}_\ell$.

The stability condition for the ALG model under the global stationarity assumption is still an open research problem. Fokianos and Tjøstheim (2011) stated the geometric ergodicity condition for $\{Y_t, \mu_t\}$ in the log-linear P-INGARCH model without exogenous variable as

$$(3.2)\quad |\beta| < 1, \qquad \alpha > 0, \qquad |\alpha + \beta| < 1 \quad \text{or} \quad |\beta| < 1, \qquad \alpha < 0, \qquad |\beta||\alpha + \beta| < 1.$$

Thus, we adopt uniform priors $p(\boldsymbol{\vartheta}_j)$ defined by indicators $I(C_j)$ for $j = 1, 2$, where C_1 and C_2 are the sets of $\boldsymbol{\theta}$ and $\boldsymbol{\kappa}$ satisfying (3.2) and $0 \le \rho, \varphi < 1$, respectively. This generates a flat prior on the parameters restricted by the indicator that is nonzero inside C_j and zero outside C_j. We also adopt a flat prior on the components of $\boldsymbol{\gamma}$. These choices of priors are not the only ones possible but are, instead, chosen to be noninformative.

The conditional posterior distributions for all $\boldsymbol{\vartheta}_\ell$ exhibit no closed-form solution. We thus adopt the adaptive MCMC method to generate the MC samples for parameter groups in their respective order from the conditional posterior distributions. We combine the random-walk Metropolis–Hastings (MH) and the independent kernel MH algorithm to draw the MCMC iterates for $\boldsymbol{\theta}$-, $\boldsymbol{\kappa}$- and $\boldsymbol{\gamma}$-groups for possibly faster convergence and better mixing. We refer to Chen and Lee (2016) for details.

We utilize the posterior mean as suggested in Chib (1995) and construct the estimate of intensity λ_t by

$$\hat{\lambda}_t = \frac{1}{N - M} \sum_{i=M+1}^{N} \lambda_t^{(i)},$$

where $\lambda_t^{(i)}$ is the ith iteration of λ_t recursively constructed through the equation of μ_t in (3.1) with each MCMC iterate of parameters denoted as $\boldsymbol{\theta}^{(i)}$, $\boldsymbol{\kappa}^{(i)}$ and $\boldsymbol{\gamma}^{(i)}$; N is the total number of iterates, and M is the number of burn-in iterates. In the following sections we set $N = 10{,}000$ for the simulation study and $N = 30{,}000$ for real data analysis. We drop the first $M = 5000$ iterations as a burn-in sample.

To test model adequacy, we apply the standardized Pearson residuals for diagnostic checking, defined as

$$R_t = \frac{Y_t - \mathrm{E}[Y_t|(F_{t-1}^{(y)}, F_t^{(x)})]}{\sqrt{\mathrm{Var}[Y_t|(F_{t-1}^{(y)}, F_t^{(x)})]}}.$$

If the model is adequate and specified correctly, then the residuals should have zero mean, unit variance and no significant serial correlations in the series of both residuals and squared residuals.

3.3. *Adaptive estimation*. The above Bayesian procedure derives from stationarity with constant parameters which is inappropriate if the dynamics are time dependent. In this section we develop an adaptive estimation procedure that detects the local interval of homogeneity in a data-driven manner in the Bayesian framework.

We assume local time homogeneity for the crime count series; that is, at a fixed time point t there exists a local interval $I_t = [t - m_t, t]$ with $1 < m_t < t$ over which all the included observations can be aptly described by a local ALG model with approximately constant parameters, that is, $\boldsymbol{\vartheta}_t \asymp$ constant. Simultaneously, we require that the modeling bias under this local parametric assumption is small, that is, the small modeling bias condition (Belomestny and Spokoiny (2007)). The length of interval I_t depends on time. In the estimation of model (3.1) at a particular time t, we assume such an I_t exists, where we can safely use the above-mentioned MCMC-based Bayesian method for parameter estimation. The estimated parameter over I_t, denoted as $\hat{\boldsymbol{\vartheta}}_t$, is called the adaptive estimator.

The question now is how to determine the homogeneous interval I_t and obtain $\hat{\boldsymbol{\vartheta}}_t$. In practice, the interval I_t is unknown and the number of possible candidates is large; for example, as many subsamples as there are past sample periods and searching among all possible interval candidates in the samples can be computationally expensive. Belomestny and Spokoiny (2007) show that an optimal choice for the interval of local homogeneity can appear via an adaptive procedure based on likelihood ratio testing. Likewise, we adopt sequential testing to detect such an interval from S candidate intervals $\mathbf{I}_t = \{I_t^{(1)}, \ldots, I_t^{(S)}\}$ to alleviate the computational burden in practice. For computational tractability, we set a common candidate interval set with $I_t^{(s)} = [t - \tau_s, t]$ for every time point t and $s = 1, \ldots, S$. The intervals are nested with increasing length, that is, $I_t^{(1)} \subset \cdots \subset I_t^{(S)}$. To each interval there exists a corresponding local Bayesian estimator, denoted by $\tilde{\boldsymbol{\vartheta}}_t^{(s)}$, which is called the weak estimator.

We present a sequential testing procedure to select the longest one from the S candidates that does not contain any breaks. The procedure starts from the shortest interval $I_t^{(1)}$ where the local homogeneity is guaranteed. We set $\hat{\boldsymbol{\vartheta}}_t^{(1)} = \tilde{\boldsymbol{\vartheta}}_t^{(1)}$ which means the weak estimator is accepted as a local homogeneous estimator. Sequentially, at each step s with $2 \le s \le S$, we test the hypothesis of local homogeneity given that at the former step the homogeneity hypothesis of interval $I_t^{(s-1)}$ has not been rejected. We define the test statistic at step s as

$$T_t^{(s)} = \bigl|L\bigl(I_t^{(s)}, \tilde{\boldsymbol{\vartheta}}_t^{(s)}\bigr) - L\bigl(I_t^{(s)}, \hat{\boldsymbol{\vartheta}}_t^{(s-1)}\bigr)\bigr|^{1/2}, \tag{3.3}$$

where $L(I_t^{(s)}, \tilde{\boldsymbol{\vartheta}}_t^{(s)})$ and $L(I_t^{(s)}, \hat{\boldsymbol{\vartheta}}_t^{(s-1)})$ are the local conditional likelihood over interval $I_t^{(s)}$ using the weak estimator to be tested for local homogeneity in the current step and using the adaptive estimator accepted in the previous step, respectively. Thus, the test statistic measures the divergence of the hypothetical ALG model and the time-varying model accepted from the previous step. Here, we define the local conditional likelihood by

$$\begin{aligned} L\bigl(I_t^{(s)}, \boldsymbol{\vartheta}\bigr) = {} & \prod_{j \in I_t^{(s)}, Y_j = 0} \bigl\{\rho + (1-\rho)e^{-\lambda_j^*}\bigr\} \\ & \cdot \prod_{j \in I_t^{(s)}, Y_j > 0} \left\{(1-\rho)\frac{\lambda_j^*(\lambda_j^* + \varphi Y_j)^{Y_j - 1}}{Y_j!} e^{[-(\lambda_j^* + \varphi Y_j)]}\right\}, \end{aligned}$$

where λ_j^* is computed recursively by

$$\lambda_j^* = \frac{1-\varphi}{1-\rho}\exp\left\{\omega + \alpha\log(Y_{j-1}+1) + \beta\log\left(\frac{1-\rho}{1-\varphi}\lambda_{j-1}^*\right) + \boldsymbol{\gamma}'\mathbf{x}_j\right\}, \quad j \in I_t^{(s)}.$$

If the divergence in (3.3) is significant, then it means the model changes more than what the sampling randomness expects, and the local homogeneity over the interval $I_t^{(s)}$ is violated. One should then terminate the procedure and select the last accepted subsample $I_t^{(s-1)}$ as the longest homogeneous interval; otherwise, one accepts the homogeneity assumption for the current interval $I_t^{(s)}$ with the updated adaptive estimator $\hat{\boldsymbol{\vartheta}}_t^{(s)} = \tilde{\boldsymbol{\vartheta}}_t^{(s)}$ and moves to the next step. The procedure continues until either a change is detected or the longest candidate interval is reached. At each time t, the algorithm runs as follows:

1. Start from the first interval $I_t^{(1)}$ with $\hat{\boldsymbol{\vartheta}}_t^{(1)} = \tilde{\boldsymbol{\vartheta}}_t^{(1)}$, where $\tilde{\boldsymbol{\vartheta}}_t^{(1)}$ is the weak estimator over $I_t^{(1)}$.
2. For $s \geq 2$, calculate the weak estimator $\tilde{\boldsymbol{\vartheta}}_t^{(s)}$ over $I_t^{(s)}$ and the statistic $T_t^{(s)}$, and:
   - if $T_t^{(s)} < \zeta_s$, then we accept $I_t^{(s)}$, that is, $\hat{\boldsymbol{\vartheta}}_t^{(s)} = \tilde{\boldsymbol{\vartheta}}_t^{(s)}$ and set $s = s+1$.
   - otherwise, we reject $I_t^{(s)}$ and stop the procedure. The interval $I_t^{(s-1)}$ is the final selection as the largest homogeneous interval at time t. We set $\hat{\boldsymbol{\vartheta}}_t = \tilde{\boldsymbol{\vartheta}}_t^{(s-1)}$.

   Here, $\zeta_2, \ldots, \zeta_S$ are certain prescribed critical values.
3. If $s < S$, then increase s by one and continue with step 2; otherwise, terminate and set $\hat{\boldsymbol{\vartheta}}_t = \hat{\boldsymbol{\vartheta}}_t^{(S)}$.

The sequence of critical value ζ_s with $s = 2, \ldots, S$ measures the significance level and plays a crucial role in the testing procedure. With a small critical value there is a higher probability to select short intervals everywhere, thus resulting in unnecessarily higher parameter uncertainty. If the critical values are too large, then it becomes easier to accept the null hypothesis of homogeneity, making it less sensitive to possible changes. Since the sampling distribution of the test statistic $T_t^{(s)}$ is unknown, even asymptotically, we calibrate the critical values via Monte Carlo experiments.

3.4. *Critical value calibration.* We generate a globally homogeneous count series from model (3.1) with constant parameters $\boldsymbol{\vartheta}_0 = (\boldsymbol{\theta}_0, \boldsymbol{\kappa}_0, \boldsymbol{\gamma}_0)$. We measure the bias of the weak estimator $\tilde{\boldsymbol{\vartheta}}_t^{(s)}$ in the homogeneous time series via the fitted log-likelihood ratio

$$D_t^{(s)} = \left|L\left(I_t^{(s)}, \tilde{\boldsymbol{\vartheta}}_t^{(s)}\right) - L\left(I_t^{(s)}, \boldsymbol{\vartheta}_0\right)\right|^{1/2}, \tag{3.4}$$

for $s = 1, \ldots, S$. Here, we numerically compute $D_t^{(s)}$ with the knowledge of $\boldsymbol{\vartheta}_0$.

Given any set of the critical values $\zeta_2, \ldots, \zeta_S$, one obtains an adaptive estimator $\hat{\boldsymbol{\vartheta}}_t^{(s)}$ at step s by employing the sequential testing procedure in Section 3.3. At time t, we measure the temporal divergence between the weak estimator and the adaptive estimator, denoted as $R_t^{(s)}$, by

$$R_t^{(s)} = \left|L\left(I_t^{(s)}, \tilde{\boldsymbol{\vartheta}}_t^{(s)}\right) - L\left(I_t^{(s)}, \hat{\boldsymbol{\vartheta}}_t^{(s)}\right)\right|^{1/2}.$$

The idea is to choose the critical values so that the adaptive estimator behaves as good as the true underlying characteristics in the artificial homogeneous situation. The distance $R_t^{(s)}$ is required to be bounded by the ideal estimation error given the true model for $s = 1, \ldots, S$:

$$E_{\boldsymbol{\vartheta}_0}\left[R_t^{(s)}\right] = E_{\boldsymbol{\vartheta}_0}\left|L\left(I_t^{(s)}, \tilde{\boldsymbol{\vartheta}}_t^{(s)}\right) - L\left(I_t^{(s)}, \hat{\boldsymbol{\vartheta}}_t^{(s)}\right)\right|^{1/2} \leq \varpi_s \bar{D}_t, \tag{3.5}$$

where $\bar{D}_t$ is the average of $\{D_t^{(s)}, s = 1, \ldots, S\}$ and ϖ_s is a hyperparameter controlling the sensitivity of the adaptive model and is, usually, selected by experience. We set ϖ_s in the form $\varpi_1 = a * \frac{1}{S-1}$ and $\varpi_s = a * \frac{s-1}{S-1}$ for $s > 1$ to reflect the increased bias as the sample size increases. Moreover, a is a constant factor, and our default choice is 1. Only the critical value is unknown in (3.5). The optimal critical value is the smallest one that makes (3.5) hold everywhere.

The calculation of critical values relies on the choice of hyperparameters $(\boldsymbol{\vartheta}_0, a)$ used in MC experiments which further influence the estimation accuracy. As an illustration our default choice is as follows. We randomly select one set of parameters $\boldsymbol{\vartheta}_0$ satisfying the restrictions of each parameter. We set $S = 6$ with corresponding lengths $\mathbf{I}_t = (30, 50, 70, 90, 120, 140)$ and use the setup everywhere. The choice of $\mathbf{I}_t$ is, in fact, data dependent. For example, the length of $I_t^{(1)}$ should generally increase with a larger number of unknown parameters to avoid overfitting. The nonparametric smoothing literature states that an increase in the length of intervals leads to an increase in modeling bias and a decrease in the estimators' variance (see, e.g., Härdle et al. (2012)). Under the assumption that the local homogeneous interval exists, we require the first interval $I_t^{(1)}$ to be short such that the modeling bias is small; that is, the local homogeneity is accepted by default and simultaneously shows reasonable estimation accuracy.

We perform a robustness check of the hyperparameters $(\boldsymbol{\vartheta}_0, a)$ in Section 4. We find that the adaptive Bayesian estimation is robust to the selection of hyperparameters, and the difference in estimate accuracy for possible misspecifications is not significant.

4. Simulation. In this section we conduct a simulation analysis under a known data generating process from model (3.1) with a single exogenous covariate as an illustration. We consider both homogeneous and regime-switching scenarios to investigate the finite sample performance of the ALG model. Moreover, we perform robust analysis to the choices of the hyperparameters $\boldsymbol{\vartheta}_0$ and a as well as the choice of priors in the adaptive MCMC procedure. Source code for simulation replication is available in the Supplementary Material (Xu et al. (2020b)) and online at https://github.com/Xiaofei-Xu/ALG.

4.1. *Homogeneous case.* We generate count series in a homogeneous scenario with a set of globally constant parameters $\boldsymbol{\vartheta}_0 = (\omega_0, \alpha_0, \beta_0, \rho_0, \varphi_0, \gamma_0) = (0.2, 0.9, -0.2, 0.2, 0.1, 0.5)$, satisfying the condition (3.2) and $0 \le \rho_0, \varphi_0 < 1$, initial count $Y_0 = 2$ and intensity $\lambda_0 = 5$. The covariate x_t is generated from the standard normal distribution. We consider two sample sizes of $n = 200$ and 800 and conduct an estimation from time points $t = 141$ and 301, respectively, to the end. The replication number is 500.

Figure 3 illustrates the average length of the identified local intervals and the estimated intensity λ_t, with its 95% credible interval for the case of $n = 800$. For the interval selection shown at the left plot, 1 refers to the shortest interval candidate with 30 observations, and 6 refers to the longest interval candidate with 140 observations. In the homogeneous scenario where the parameters are constant throughout the whole sample, the optimal selection of intervals should be $S = 6$, that is, the longest interval. The proposed approach yields a quite stable and reasonable choice with the values closing to 5. It also shows that the ALG model estimates the conditional expectation λ_t quite well, fitting the actual values of intensity of the count series.

We use the root mean squared error (RMSE) and the mean absolute deviation (MAD) to evaluate the estimation accuracy. For comparison, we apply the rolling window technique with a fixed window size of 140 (length of longest candidate interval) to iteratively update the parameter estimation. Table 2 reports the results, with the panel titled "ALG" referring to the ALG model and "RW140" referring to the rolling window technique. It shows that

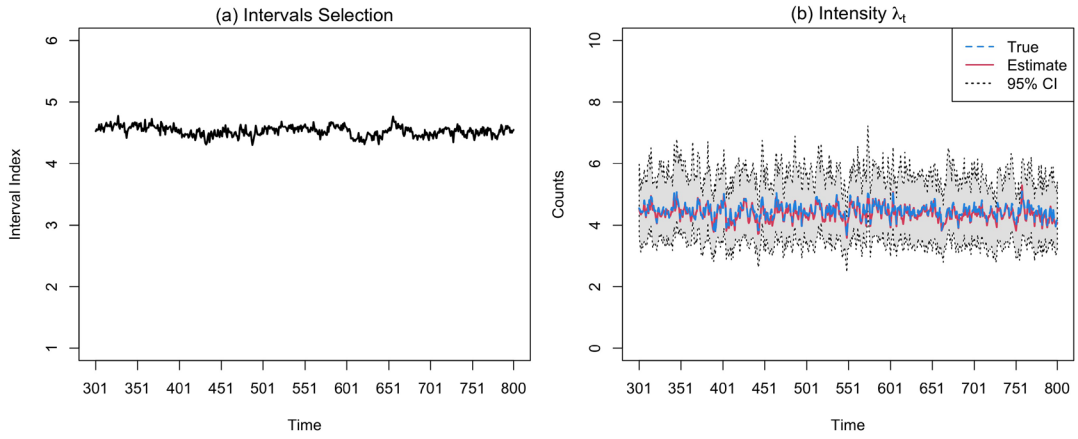

FIG. 3. *The average index of selected intervals (left) and the ALG estimate of λ_t with its corresponding 95% credible intervals (right) in $n = 800$ case from time index 301 to 800.*

the ALG estimate of parameters is accurate, with RMSE no larger than 0.266 and MAD smaller than 0.262 among all the parameters. The ALG model is comparable to the rolling window technique. This is not surprising, because RW140, known as the best model in the homogeneous scenario, always uses the longest interval.

4.2. *Heterogeneous case.* We design several nonstationary experiments with one or multiple parameters changing over time, while the other parameters remain constant. The first four experiments change a single parameter, ω_t, α_t, ρ_t and γ_t corresponding to the level, serial dependence, zero inflation and the exogenous variable effect, respectively. We denote them as RS-W where the affixed notation W refers to a particularly changing parameter. The other two designs, denoted as RS-m1 and RS-m2, shift three parameters, ω_t, ρ_t and γ_t, simultaneously. In each of the four experiments, the single or multiple changing parameters shift over their phases. Parameter shifts always happen at $t = \lfloor \frac{n}{2} \rfloor$ and $\lfloor \frac{3n}{4} \rfloor$ for all the experiments where $\lfloor z \rfloor$ denotes the maximal integer that is not larger than z and n is sample size; see Table 3 for the time-varying parameter sets. We consider two sample sizes of $n = 250$ and 800 and start the estimation from time $t = 141$ and 301, respectively, to the end. Each experiment is replicated 500 times. The exogenous covariate x_t is generated from the standard normal distribution.

To demonstrate the dynamic process, we display the average length of selected intervals for the cases RS-α and RS-m1 under $n = 800$ in Figure 4 as an illustration, and other experiments

TABLE 2

Parameter estimation accuracy for the homogeneous scenario. The adaptive estimation of ALG modeling is compared with the rolling window technique

		$n = 200$						$n = 800$					
		RW140			ALG			RW140			ALG		
	True	Estimate	RMSE	MAD	Estimate	RMSE	MAD	Estimate	RMSE	MAD	Estimate	RMSE	MAD
ω_0	0.20	0.197	0.125	0.115	0.189	0.183	0.153	0.198	0.181	0.191	0.186	0.266	0.262
α_0	0.90	0.891	0.078	0.071	0.874	0.126	0.102	0.893	0.111	0.118	0.873	0.190	0.176
β_0	−0.20	−0.200	0.066	0.060	−0.214	0.107	0.086	−0.201	0.099	0.106	−0.214	0.156	0.149
ρ_0	0.20	0.204	0.034	0.031	0.211	0.048	0.041	0.201	0.050	0.053	0.209	0.069	0.070
φ_0	0.10	0.113	0.042	0.039	0.124	0.055	0.045	0.113	0.059	0.063	0.123	0.076	0.073
γ_0	0.50	0.500	0.043	0.039	0.496	0.072	0.059	0.501	0.061	0.064	0.496	0.121	0.106

TABLE 3
Parameters in each scenario

Parameters	ω_t	α_t	β_t	γ_t	ρ_t	φ_t	ω_t	α_t	β_t	γ_t	ρ_t	φ_t
			RS-ω						RS-α			
Phase-1	**1**	0.9	−0.2	0.5	0.2	0.1	0.6	**0.4**	0.4	0.5	0.2	0.1
Phase-2	**0.2**	0.9	−0.2	0.5	0.2	0.1	0.6	**0.1**	0.4	0.5	0.2	0.1
Phase-3	**1**	0.9	−0.2	0.5	0.2	0.1	0.6	**0.4**	0.4	0.5	0.2	0.1
			RS-γ						RS-ρ			
Phase-1	0.6	0.4	0.4	**0.5**	0.2	0.1	0.6	0.4	−0.3	0.5	**0.1**	0.2
Phase-2	0.6	0.4	0.4	**0.1**	0.2	0.1	0.6	0.4	−0.3	0.5	**0.6**	0.2
Phase-3	0.6	0.4	0.4	**0.5**	0.2	0.1	0.6	0.4	−0.3	0.5	**0.1**	0.2
			RS-m1						RS-m2			
Phase-1	**0.6**	0.9	−0.2	**0.5**	**0**	0.1	**−0.2**	0.9	−0.2	**0**	**0.6**	0.1
Phase-2	**0.2**	0.9	−0.2	**0.1**	**0.6**	0.1	**0.6**	0.9	−0.2	**0.5**	**0**	0.1
Phase-3	**0.6**	0.9	−0.2	**0.5**	**0**	0.1	**−0.2**	0.9	−0.2	**0**	**0.6**	0.1

present a similar pattern. For both cases the ALG model provides a reasonable selection of local intervals. The length of the selected intervals appears large before the break, quickly drops to a very small value immediately after the break and gradually bounces back. The quick reaction to breaks is important for an accurate estimation with changing parameters. For practical reasons the data-driven approach selects the optimal local intervals only on the past information, and no "future" information is utilized. This explains the systematical lag in the plot, though with a fast update. To summarize, the ALG model selects longer intervals for the approximately homogeneous periods to utilize as much information as possible and reacts to changes efficiently so as to avoid modeling bias.

Figure 5 exhibits the true value and ALG estimate with corresponding 95% credible intervals of λ_t, ω_t, ρ_t and γ_t for the case RS-m1 under $n = 800$ as a demonstration. We see that the estimated λ_t process fits the actual values quite well and only needs a short period to adjust after the breaks happen. Specifically, the estimation of intensity is accurate at phase 2, where the occurrence possibility of zeros is large with the true parameter ρ_t being 0.6. For the individual parameters the adaptive estimates deliver similar results as λ_t. In conclusion, the ALG model provides a good estimation of intensity and time-varying parameters, with

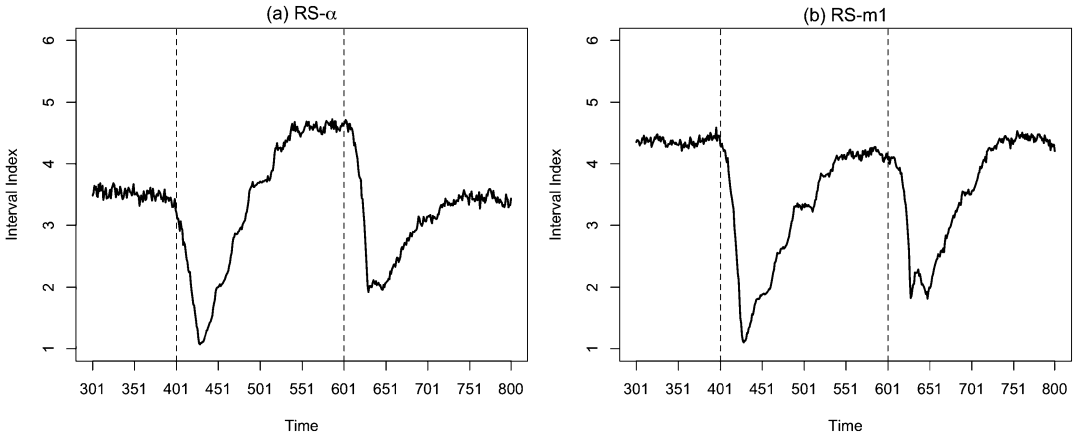

FIG. 4. *The average indices of the selected intervals for RS-α and RS-m*1 *from time index* 301 *to* 800.

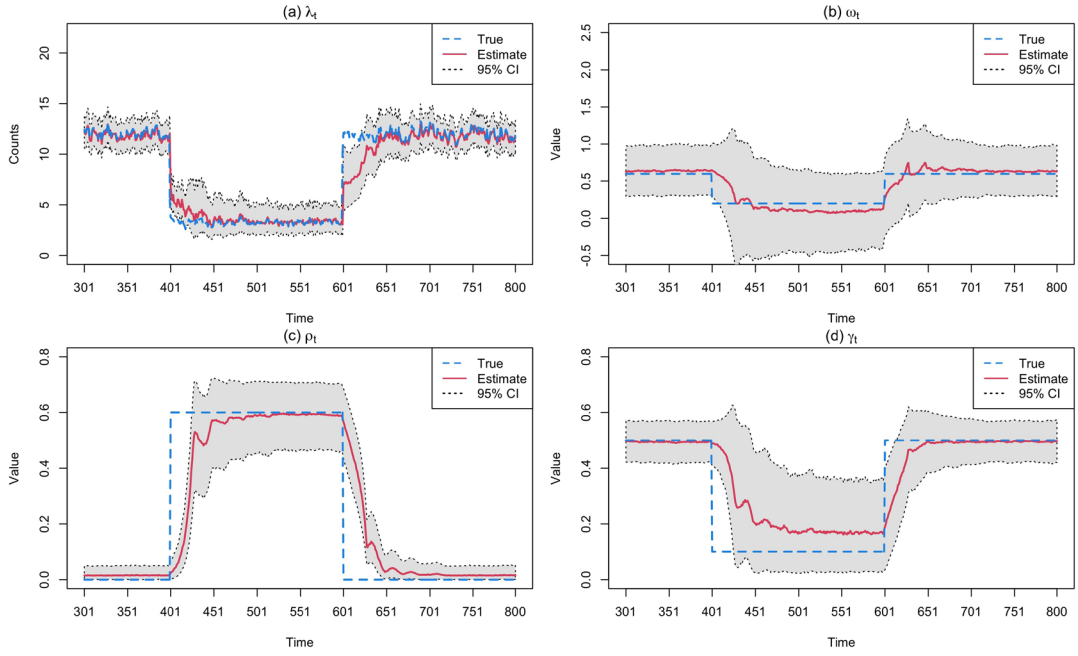

FIG. 5. *Design RS-m*1 *scenario*: *The true value*, *ALG estimate*, *and the corresponding* 95% *credible intervals of* λ_t, ω_t, ρ_t, *and* γ_t *from time index* 301 *to* 800.

the estimates being stable and close to the true value before the breaks, quickly adjusting to match the true dynamics after discovering the breaks and becoming stable again thereafter.

For estimation accuracy comparison we adopt rolling window strategies to the ALG model over the same estimation period. We consider two rolling window sizes, that is, 90 (RW90) and 140 (RW140) corresponding to the length of candidate intervals I_4 and I_6, respectively, in the ALG procedure. Table 4 summarizes RMSE and MAD of estimated intensity λ_t of the three models. The best accuracy is marked in bold. In general, the ALG model is more accurate than the rolling window technique by utilizing the adaptively selected time-varying intervals. For example, for $n = 250$, the ALG model is the best for all designs, except RS-ρ, in terms of MAD. When the sample size becomes 800, the ALG model performs better for design RS-α, RS-γ and RS-m1, with smaller RMSE and MAD. However, for RS-ρ, ALG is comparable to the rolling window technique under both sample sizes, perhaps because the breaks in λ_t are not significant though ρ_t changes over three phases, which leads to a better performance with a larger fixed window size.

TABLE 4
RMSE and MAD of the λ_t *estimation of six designs*

	$n = 250$						$n = 800$					
	RW90		RW140		ALG		RW90		RW140		ALG	
	RMSE	MAD	RMSE	MAD	RMSE	MAD	RMSE	MAD	RMSE	MAD	RMSE	MAD
RS-ω	2.890	1.589	**2.692**	1.580	2.798	**1.528**	**2.136**	**1.047**	2.181	1.099	2.149	1.053
RS-α	4.959	2.479	4.882	2.895	4.920	**2.422**	3.379	1.551	4.212	1.945	**3.257**	**1.407**
RS-ρ	1.144	0.721	**0.845**	**0.537**	1.281	0.771	1.217	0.799	**0.739**	**0.445**	0.997	0.568
RS-γ	5.304	2.588	6.515	3.390	**4.821**	**2.404**	3.599	1.758	4.735	2.105	**3.228**	**1.566**
RS-m1	**3.414**	1.799	3.752	2.236	3.470	**1.799**	2.387	1.066	2.849	1.227	**2.331**	**1.036**
RS-m2	2.483	1.127	3.125	1.662	**2.198**	**0.996**	**2.178**	0.948	2.858	1.248	2.243	**0.936**

4.3. *Robustness checking*. Local interval selection plays an important role in the adaptive estimation procedure which relies on the critical values calibrated based on the data generated with some hyperparameters $(\boldsymbol{\vartheta}_0, a)$. The default parameter set is $\boldsymbol{\vartheta}_0 = (\omega_0, \alpha_0, \beta_0, \rho_0, \varphi_0, \gamma_0) = (0.2, 0.9, -0.2, 0.2, 0.1, 0.5)$ and $a = 1$. We analyze the performance of the ALG estimation with respect to different hyperparameter choices under two experiments: homogeneous design at Section 4.1 and heterogeneous design RS-m2 at Section 4.2. We redo the estimation accuracy using different critical values obtained with misspecified parameters and various values of a.

We particularly recalibrate the critical values under six misspecified underlying parameters $\boldsymbol{\vartheta}_0^*$ with $\pm 20\%$ deviation from $\alpha_0 + \beta_0$, γ_0, and ρ_0, respectively, among others. We also study two risk factors a deviating from 1: a^* of 0.75 and 1.5. Note that the other hyperparameters still take the original true values. Moreover, we investigate the sensitivity concerning the choice of priors in the adaptive MCMC procedure. The default choice is noninformative (uniform/flat) priors for all parameters, which are not the only ones possible. To examine the sensitivity, as an illustration we consider setting beta distributions as the priors for parameter $\boldsymbol{\kappa}$ (Chen and Lee (2016)), that is, Prior 1: $\rho \sim \text{Beta}(4, 6)$, $\varphi \sim \text{Beta}(10, 90)$; Prior 2: $\rho \sim \text{Beta}(60, 40)$, $\varphi \sim \text{Uniform}(0, 1)$. We conduct an adaptive estimation to homogeneous design using alternative priors compared with the default case.

In each robustness experiment we redo the local interval selection and the adaptive estimation with the recalibrated critical values. For prior experiments, using the critical values computed under a default setup, we conduct an estimation with different priors. Table 5 reports the parameter estimation accuracy under misspecified parameters, alternative a and alternative priors $(\boldsymbol{\kappa})$. Compared to the default case with $\boldsymbol{\vartheta}_0^* = \boldsymbol{\vartheta}_0$, $a^* = 1$ and flat prior of parameters, there is no significant difference under homogeneous and RS-mult2 scenarios. It suggests that, via the adaptive and data-driven way to calibrate the critical values, the estimation accuracy is generally robust with respect to possible parameter misspecifications and alternative a. Moreover, the parameter estimation is not sensitive to the selection of priors $(\boldsymbol{\kappa})$ as well by delivering similar results to the default.

The simulation study shows stable and accurate performances of the ALG model under both stationary and nonstationary scenarios, where the ALG model provides a reasonable selection of local intervals, reacts quickly to structural breaks and is robust to the choice of hyperparameters. The adaptive Bayesian estimation is robust to the selection of hyperparameters $\boldsymbol{\vartheta}_0$ and a as well as the choice of priors in the adaptive MCMC procedure. However, the adaptive MCMC procedure is a bit time consuming, especially when the sample size and replication times become larger. The R package Template Model Builder (TMB) proposed by Kristensen et al. (2016) is designed for estimating complex nonlinear models; see, for example, Berentsen et al. (2018) and Cui, Li and Zhu (2020). It may be used to improve computational speed.

5. Real data analysis. In this section we apply the ALG model to investigate the time-dependent evolution of monthly crime count time series of the six categories in Byron, NSW, Australia, from January 1995 to December 2017 and evaluate the performance of the ALG model.

We retain the first 140 observations as initiation and start the estimation from the 141th observation dating in September 2006 and end at the last point dating in December 2017. We apply the ALG model for each crime dataset, where $\mathbf{x}_t$ includes four exogenous covariates: monthly mean maximum temperature, unemployment rate, population and a summer seasonality dummy. The detailed description of datasets is in Section 2. We use the same setup of the interval candidates and the critical values as in Section 4.1.

TABLE 5

Robustness checking: *"A*0.8*" and "A*1.2*", "R*0.8*" and "R*1.2*" and "G*0.8*" and "G*1.2*" refer to the misspecified cases where the parameters* $\alpha + \beta$, ρ, *and* γ *are, respectively, biased with* 20% *deviation from the default* $\boldsymbol{\vartheta}_0$, *and "a*0.75*" and "a*1.5*" represent the two factors* $a = 0.75$ *and* 1.5, *respectively. "Prior*1*" and "Prior*2*" refer to two alternative priors of* $\boldsymbol{\kappa}$

	RMSE							MAD						
	ω_t	α_t	β_t	ρ_t	φ_t	γ_t	λ_t	ω_t	α_t	β_t	ρ_t	φ_t	γ_t	λ_t
Hom														
Default	0.17	0.15	0.12	0.06	0.14	0.10	1.24	0.11	0.11	0.09	0.04	0.09	0.06	0.54
A0.8	0.20	0.18	0.14	0.07	0.17	0.13	1.37	0.14	0.13	0.10	0.05	0.12	0.08	0.62
A1.2	0.19	0.17	0.13	0.06	0.16	0.12	1.26	0.13	0.12	0.10	0.05	0.11	0.08	0.59
R0.8	0.16	0.15	0.12	0.05	0.13	0.09	1.22	0.11	0.10	0.09	0.04	0.09	0.06	0.53
R1.2	0.19	0.17	0.13	0.06	0.16	0.12	1.38	0.13	0.12	0.10	0.05	0.11	0.08	0.61
G0.8	0.18	0.16	0.13	0.06	0.15	0.11	1.21	0.12	0.11	0.09	0.04	0.10	0.07	0.56
G1.2	0.21	0.18	0.14	0.07	0.18	0.13	1.59	0.14	0.13	0.10	0.05	0.12	0.09	0.66
a1.75	0.18	0.16	0.13	0.06	0.15	0.11	1.19	0.12	0.11	0.09	0.04	0.10	0.07	0.55
a1.5	0.21	0.18	0.14	0.07	0.18	0.13	1.53	0.15	0.13	0.10	0.05	0.13	0.09	0.66
Prior1	0.17	0.15	0.14	0.06	0.10	0.11	1.35	0.13	0.11	0.10	0.05	0.10	0.07	0.61
Prior2	0.15	0.15	0.13	0.02	0.12	0.10	1.12	0.10	0.10	0.10	0.02	0.09	0.07	0.47
RS-m2														
Default	0.39	0.20	0.17	0.17	0.13	0.20	2.24	0.28	0.15	0.13	0.10	0.10	0.14	0.94
A0.8	0.42	0.22	0.18	0.16	0.15	0.21	2.12	0.30	0.17	0.14	0.09	0.11	0.16	0.92
A1.2	0.41	0.21	0.18	0.16	0.14	0.21	2.19	0.29	0.16	0.13	0.09	0.10	0.15	0.93
R0.8	0.39	0.19	0.17	0.17	0.13	0.20	2.31	0.27	0.15	0.13	0.10	0.10	0.14	0.94
R1.2	0.42	0.21	0.18	0.16	0.15	0.21	2.15	0.29	0.16	0.14	0.09	0.10	0.15	0.93
G0.8	0.40	0.20	0.17	0.17	0.14	0.20	2.22	0.28	0.15	0.13	0.09	0.10	0.15	0.93
G1.2	0.43	0.22	0.19	0.16	0.15	0.22	2.11	0.30	0.17	0.14	0.09	0.11	0.16	0.93
a0.75	0.41	0.20	0.17	0.17	0.14	0.21	2.27	0.28	0.15	0.13	0.10	0.10	0.15	0.95
a1.5	0.43	0.22	0.19	0.16	0.15	0.22	2.09	0.30	0.17	0.14	0.09	0.11	0.16	0.93

5.1. *Interpretation*. For space limitation we only present the results of three crime categories of MDP, LOS and ARS, representing low, medium and high levels of zero percentage datasets; see Figures B, C and D in the Supplementary Material (Xu et al. (2020a)) for the detailed results of ASS, TSP and AJP, respectively. Figures 6–8 illustrate the ALG estimates of parameters from September 2006 to December 2017. Specifically, as the intensity λ_t, that is, the conditional expectation of crime counts, represents the essential feature of observations, we thus display the crime counts, the adaptive estimate of λ_t and the corresponding 95% credible interval in the subplot (a). The subplots (b)–(j) display the estimates of the time-varying parameters ω_t, α_t, β_t, ρ_t, φ_t and the effect estimate of exogenous covariates $\boldsymbol{\gamma}_t$ together with their 95% credible intervals. We also display the constant estimate under the global homogeneity as a benchmark to highlight the time-varying pattern of the adaptive estimates. The ALG estimate of intensity process λ_t above all captures the dynamics of crime count time series very well. The estimated λ_t is not very dispersed, as reflected by the narrow 95% credible intervals, but does have varying dynamics. We note that closed circuit television (CCTV) cameras are located at some council facilities to reduce crime by preventing potential offenders and to assist police in the detection and prosecution of offenders. CCTV cameras and extra street lighting were installed in the Byron Bay town center from August 2015, which may explain a decreasing trend in MDP after 2015. We do observe that MDP

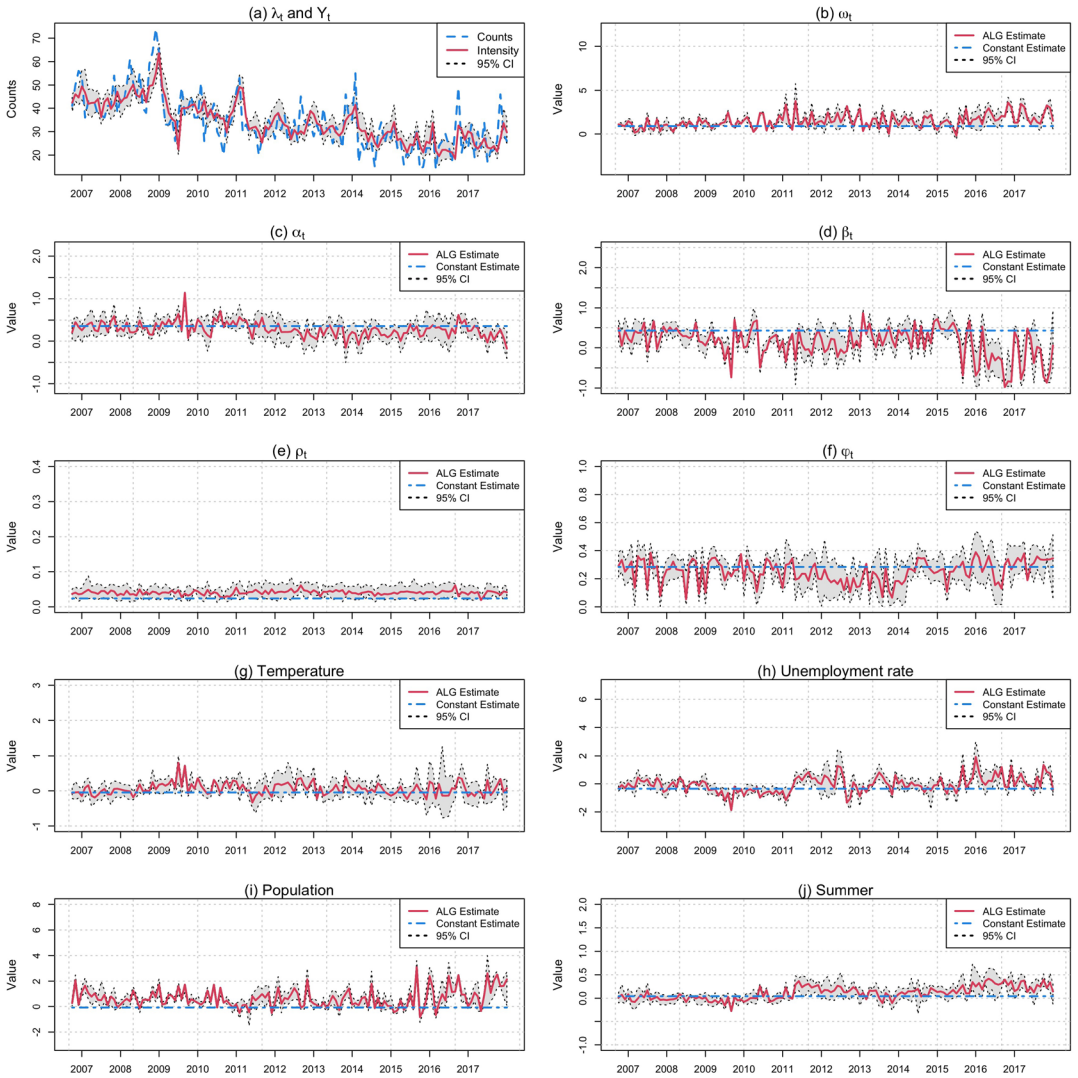

FIG. 6. *MDP: the ALG estimate and 95% CIs (credible intervals) of λ_t and the parameters. (a) indicates crime counts Y_t with dashed lines. (b)–(j) indicates the constant estimate under global homogeneity with dot-dashed lines.*

series and estimated λ_t go to lower levels after 2015 in Figure 6(a). The count of LOS exhibits a pattern of seasonality and changing level via visual inspection which is also properly reflected by estimated λ_t.

The ALG model, in general, flexibly captures the dynamics of time-varying parameters and reflects nonstationarity under different levels of zero inflation and overdispersion, while the constant estimator under global homogeneity often over- or underestimates parameters. For example, we find that α_t is relatively stable for all these three crimes, with fluctuations around zero for LOS and ARS. It indicates a stable and positive neighborhood effect for MDP. The zero-inflation parameter, ρ_t, is generally close to zero for most types of crime except ARS. This is reasonable as ARS has excessive zero observations. Specifically, ρ_t reveals a decreasing trend for ARS, which is consistent with the count feature, that is, lightly increasing intensity and smaller percentage of zeros as time goes on. The dispersion parameter φ_t for all crimes is positive with larger value for LOS than MDP and ARS, implying an overdispersion feature of these crime series. We find a sharp decrease of φ_t for LOS after

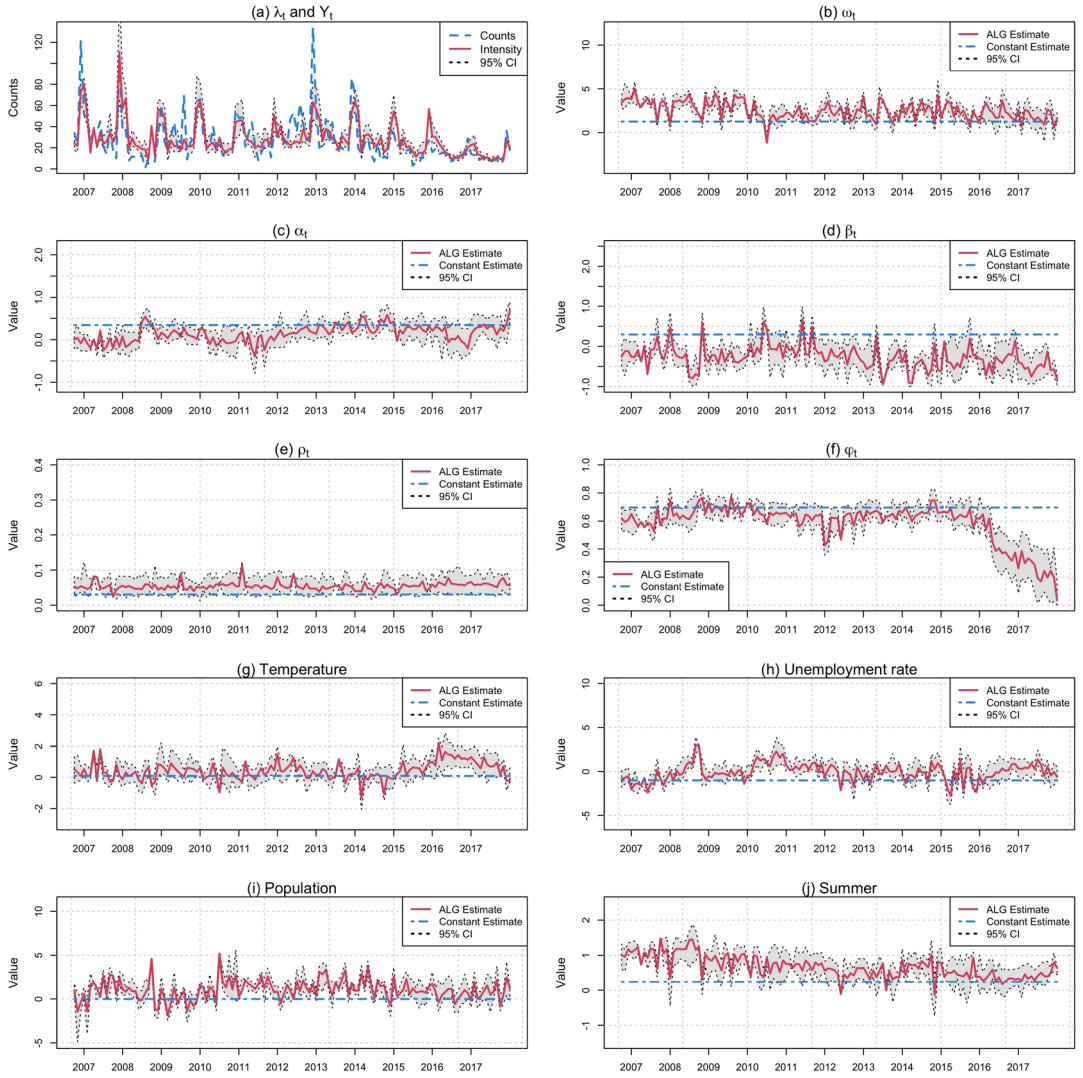

FIG. 7. *LOS: the ALG estimate and 95% CIs of λ_t and the parameters. (a) indicates crime counts Y_t with dashed lines. (b)–(j) indicates the constant estimate under global homogeneity with dot-dashed lines.*

around year 2016, which might be because of the more regular and stable pattern of LOS crime intensity between 2015–2017.

A large amount of literature has documented that higher levels of violence tend to appear under warm temperatures (Mares and Moffett (2016)) and could exhibit certain seasonality. Our study not only detects the effects of temperature and seasonality on the incidences of crimes but also investigates the evolution of temperature impacts over time. We find that LOS has a positive coefficient for the summer dummy, indicating a clear feature of higher-level crime during summer than during other seasons, and the positive temperature effect is also more significant in the recent two years (years 2016–2017). We also find a generally positive summer dummy coefficient for ASS and TSP and a positive temperature coefficient for TSP as well. This indicates that higher-level crime activities are likely to occur as a result of anomalously warm seasons/temperatures for LOS, ASS and TSP, while the other crimes like MDP and ARS are not sensitive to temperatures and hot seasons, with the coefficients for both temperature and the summer dummy close to zero or fluctuating around zero.

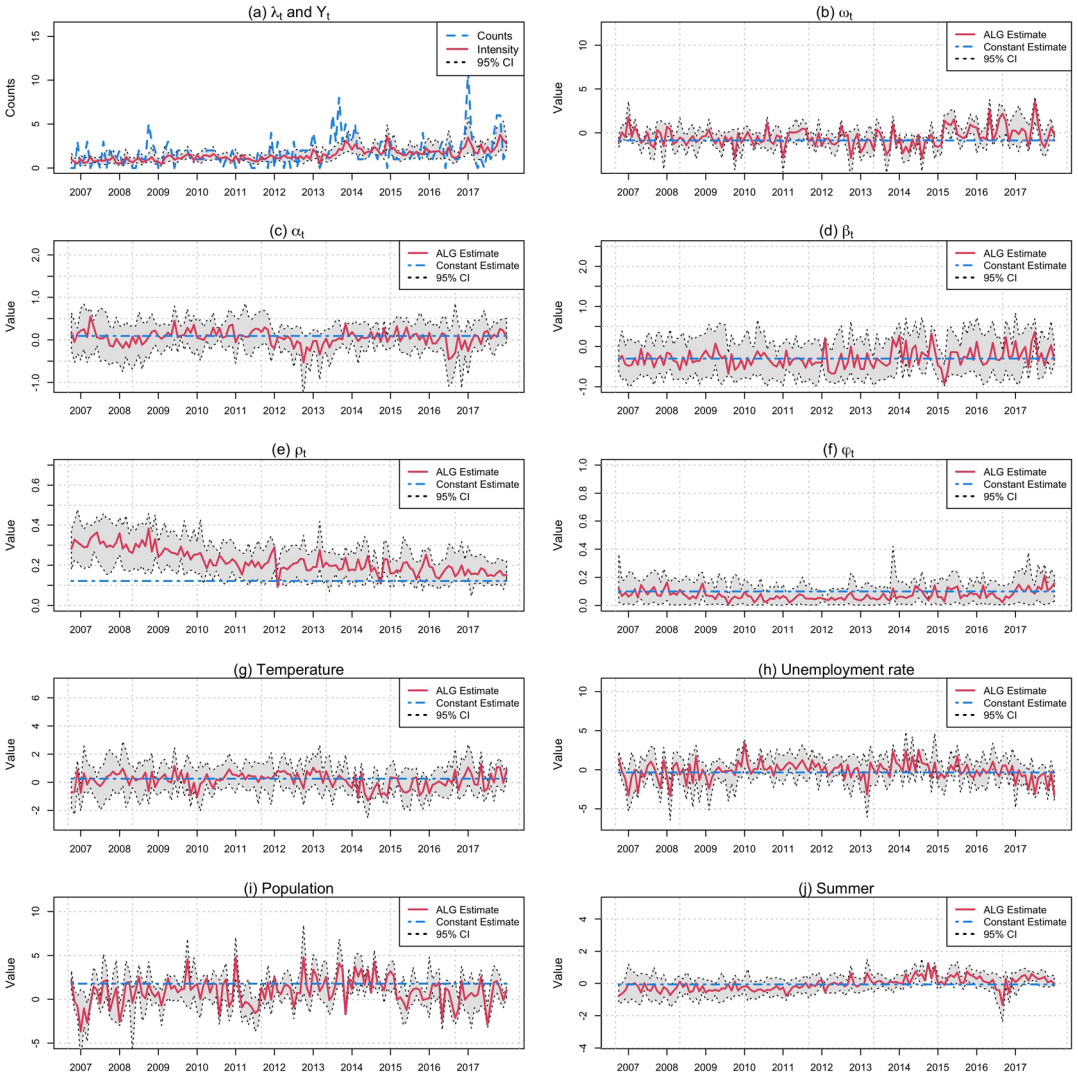

FIG. 8. *ARS: the ALG estimate and* 95% *CIs of* λ_t *and the parameters.* (*a*) *indicates crime counts* Y_t *with dashed lines.* (*b*)–(*j*) *indicates the constant estimate under global homogeneity with dot-dashed lines.*

After relaxing to local stationarity, we observe a diverse phenomenon of demographic factors. In particular, we find among all the six categories that five crime incidents (including the three illustrative cases as well as ASS and TSP) are not very sensitive to either the unemployment rate or the population amount, as the estimated coefficients are generally stable with values around 0. In other words, there is no evidence of an association between population and unemployment rate for these crime activities. However, the two produce a certain positive effect for AJP. We find a positive coefficient of the unemployment rate at around years 2011–2016 and a positive coefficient of population after year 2013 for AJP. We also observe an increasing trend of AJP criminals after year 2012. This indicates that a greater unemployment rate and population can be one reason for the increasing intensity in AJP since around 2012. The constant estimator under global homogeneity lacks any meaningful interpretation on the dynamics of exogenous factors' effects, with almost all estimates being close to zero for all the categories.

5.2. *Model selection and diagnostic checking.* We use a model selection technique to evaluate the overall impact of exogenous covariates. Specifically, we set up the dynamics of

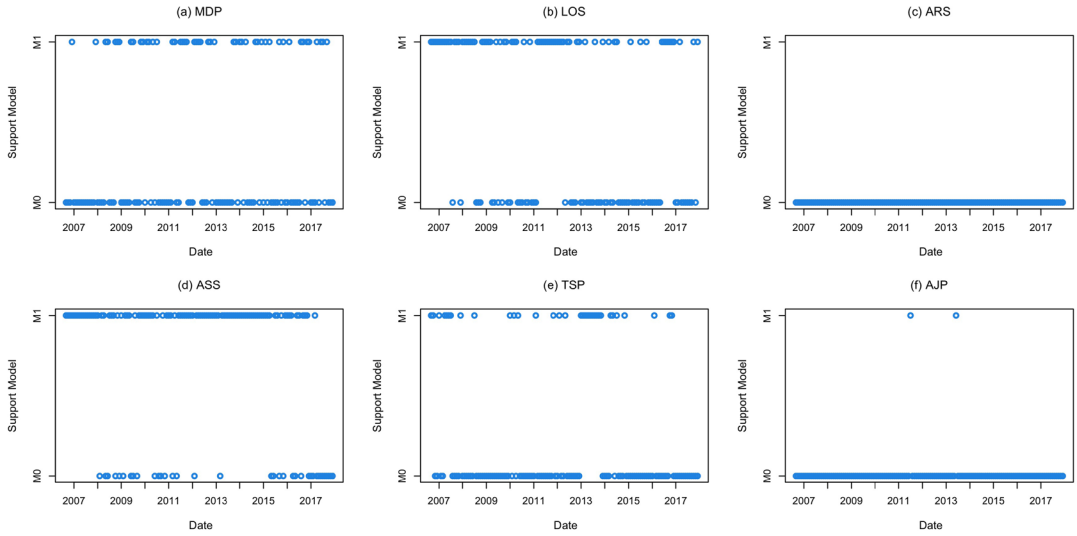

FIG. 9. *BIC criterion results*: *"M0" and "M1" refer to the model that BIC criterion favors.*

various crime series using the ALG model with (labeled M_1) and without exogenous variables (M_0), respectively. In other words, we force $\boldsymbol{\gamma}_t = 0$ in model M_0 at every time t. We apply Bayesian information criterion (BIC) to choose the better model at each time.

Let $\mathcal{D}_t$ denote the BIC difference of M_1 and M_0 at time t by

$$\mathcal{D}_t = \log(|I_t|)(k_1 - k_0) - 2\big[L(Y_{I_t}|\mathbf{x}_{I_t}, \boldsymbol{\vartheta}_t, M_1) - L(Y_{I_t}|\mathbf{x}_{I_t}, \boldsymbol{\vartheta}_t, M_0)\big],$$

where $|I_t|$ is the sample size of the detected local interval of homogeneity I_t at time t and k_0 and k_1 are the number of parameters in models M_0 and M_1, respectively. Here, $L(Y_{I_t}|\mathbf{x}_{I_t}, \boldsymbol{\vartheta}_t, M_i)$ is the log-likelihood of model M_i over interval I_t for $i = 0, 1$. Positive $\mathcal{D}_t$ indicates smaller BIC for model M_0, thus supporting M_0 over M_1 at time t. Figure 9 discloses the model selection results. As shown, the BIC criterion strongly favors M_1 at most time points for ASS and LOS, while only at some points for TSP and MDP, but always strongly supports M_0 for ARS and AJP. This indicates that the effects of exogenous covariates and seasonality, as a whole, are significant to crimes ASS and LOS but not for ARS and AJP. As the variables of unemployment rate and population are observed for the whole NSW instead of Byron, the interpretation of the test may have some limitation.

We perform diagnostic checking based on the Pearson residuals. Figure 10 exhibits the histogram and ACF of the residuals and the ACF of the squared standardized residuals for MDP, LOS and ARS, respectively. All ACFs are within the two standard error limits, except for two small correlations on the boundary for the LOS dataset. The results for other crime categories demonstrate similar performance (see Figure E in the Supplementary Material (Xu et al. (2020a)) for details). It implies that the ALG model is adequate in describing the dynamics of the crime datasets.

6. Conclusion. This research proposes a comprehensive model, named the Adaptive Log-linear zero-inflated Generalized poisson integer-valued GARCH with exogenous covariates (ALG), to capture the dynamics of count time series with various characteristics and, simultaneously, incorporates the influence of multiple exogenous variables in a unified modeling framework. The ALG model enables one to handle unforeseeable structural breaks via time-varying parameters without any predetermined information on the changing points. We adopt an adaptive approach to detect a local homogeneous interval at each time point,

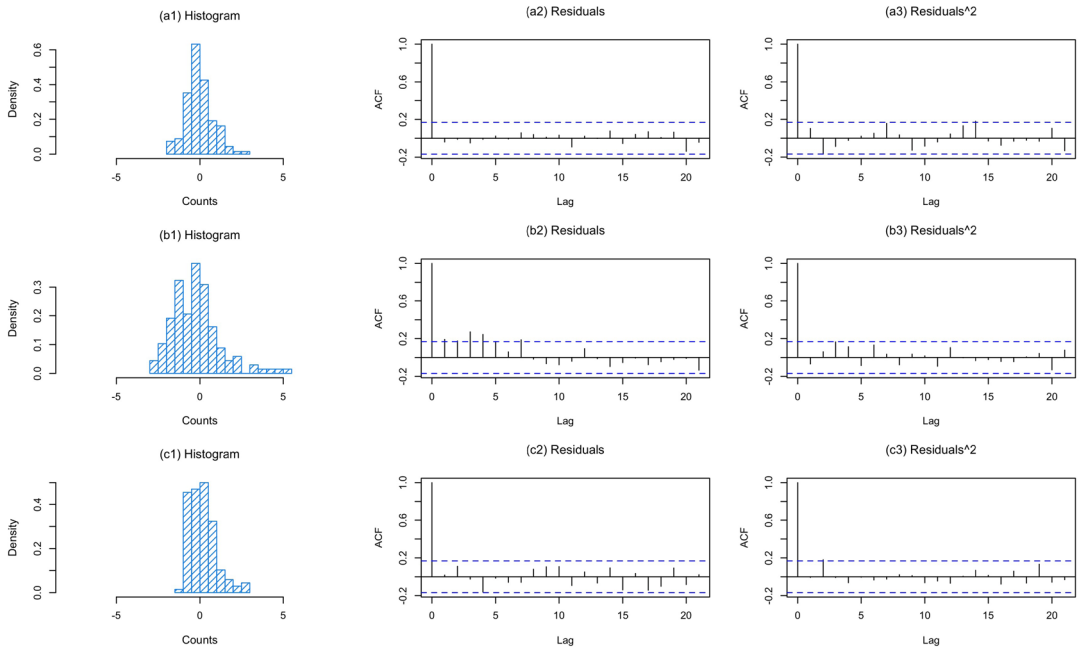

FIG. 10. *Diagnostic checking plots*: *MDP* (*upper*), *LOS* (*middle*) *and ARS* (*bottom*). *From left to right are histograms of the residuals*, *ACF of the residuals and ACF of the squared residuals*, *respectively.*

over which the time-dependent parameters are estimated through an adaptive Bayesian procedure based on the MCMC technique. A finite sample analysis shows stable and accurate performances of the ALG model under both stationary and nonstationary scenarios, where the ALG model provides a reasonable selection of local intervals, reacts quickly to structural breaks and is robust to the choice of hyperparameters. In fact, the ALG model is general and can be easily applied to other nonstationary integer-valued time series with the features of autocorrelation, heteroscedasticity, overdispersion and excessive number of zero observations.

In real data analysis on the six crime incidents categorized as "assault: nondomestic violence related assault," "malicious damage to property," "theft: steal from person," "liquor offenses," "against justice procedures: breach bail conditions" and "arson" in Byron, NSW, Australia, the ALG model not only delivers persuasive estimation accuracy of stochastic intensity but also provides insightful interpretations on the dynamics of intensity, effect of seasonality (summer) and impacts of several environment and demographic factors. It is able to flexibly capture the dynamics of changing parameters, whereas the constant model under global homogeneity often leads to overestimation or underestimation. We find that higher levels of the crime activities are likely to occur as a result of anomalously warm seasons/temperatures for nondomestic violence related assault, liquor offenses and steal from person with time-dependent dynamics, while this is not the case for the other three crimes. The effects of unemployment rate and population are not significant to these six crime incidents except for the crime of against justice procedures where a greater unemployment rate and population can be one reason for the increasing intensity of AJP since around 2012.

Acknowledgments. We thank the Editor, the Associate Editor and anonymous referees for their valuable time and careful comments, which have helped improve this paper. The authors gratefully acknowledge the financial support of Singapore's Ministry of Education Academic Research Fund Tier 1 and Institute of Data Science at National University of Singapore. Cathy W. S. Chen's research is funded by the Ministry of Science and Technology, Taiwan (grant MOST 107-2118-M-035-005-MY2).

SUPPLEMENTARY MATERIAL

Supplement to "Adaptive log-linear zero-inflated generalized Poisson autoregressive model with applications to crime counts." (DOI: 10.1214/20-AOAS1360SUPPA; .pdf). Supplement on Figures A–E.

Source code for "Adaptive log-linear zero-inflated generalized Poisson autoregressive model with applications to crime counts." (DOI: 10.1214/20-AOAS1360SUPPB; .zip). R source code for the simulation design RS-m1 described in this paper and data files from this design. Other designs can run using same codes with simple modifications at data input.

REFERENCES

AGOSTO, A., CAVALIERE, G., KRISTENSEN, D. and RAHBEK, A. (2016). Modeling corporate defaults: Poisson autoregressions with exogenous covariates (PARX). *J. Empir. Finance* **38** 640–663.

ANDERSON, C. A. (2001). Heat and violence. *Curr. Dir. Psychol. Sci.* **10** 33–38. https://doi.org/10.1111/1467-8721.00109

ANDERSON, C. A., ANDERSON, K. B., DORR, N., DENEVE, K. M. and FLANAGAN, M. (2000). Temperature and aggression. In *Advances in Experimental Social Psychology* **32** 63–133.

ANDERSON, J., MORGAN, A., BOXALL, H. and LINDEMAN, K. (2012). *Effective Crime Prevention Interventions for Implementation by Local Government*. *Research and Public Policy Series* **120**. Australian Institute of Criminology, Canberra. https://aic.gov.au/publications/rpp/rpp120.

BELOMESTNY, D. and SPOKOINY, V. (2007). Spatial aggregation of local likelihood estimates with applications to classification. *Ann. Statist.* **35** 2287–2311. MR2363972 https://doi.org/10.1214/009053607000000271

BERENTSEN, G. D., BULLA, J., MARUOTTI, A. and STØVE, B. (2018). Modelling corporate defaults: A Markov-switching Poisson log-linear autoregressive model. arXiv preprint arXiv:1804.09252.

CHEN, C. W. S. and KHAMTHONG, K. (2019). Bayesian modelling of nonlinear negative binomial integer-valued GARCHX models. *Stat. Model.* https://doi.org/10.1177/1471082X19845541

CHEN, C. W. S., KHAMTHONG, K. and LEE, S. (2019). Markov switching integer-valued generalized autoregressive conditional heteroscedastic models for Dengue counts. *J. R. Stat. Soc. Ser. C. Appl. Stat.* **68** 963–983. MR4002379 https://doi.org/10.1111/rssc.12344

CHEN, C. W. S. and LEE, S. (2016). Generalized Poisson autoregressive models for time series of counts. *Comput. Statist. Data Anal.* **99** 51–67. MR3473081 https://doi.org/10.1016/j.csda.2016.01.009

CHEN, C. W. S. and LEE, S. (2017). Bayesian causality test for integer-valued time series models with applications to climate and crime data. *J. R. Stat. Soc. Ser. C. Appl. Stat.* **66** 797–814. MR3670418 https://doi.org/10.1111/rssc.12200

CHEN, C. W. S., SO, M. K. P., LI, J. C. and SRIBOONCHITTA, S. (2016). Autoregressive conditional negative binomial model applied to over-dispersed time series of counts. *Stat. Methodol.* **31** 73–90. MR3477729 https://doi.org/10.1016/j.stamet.2016.02.001

CHEN, Y. and LI, B. (2017). An adaptive functional autoregressive forecast model to predict electricity price curves. *J. Bus. Econom. Statist.* **35** 371–388. MR3663780 https://doi.org/10.1080/07350015.2015.1092976

CHIB, S. (1995). Marginal likelihood from the Gibbs output. *J. Amer. Statist. Assoc.* **90** 1313–1321. MR1379473

COHN, E. G. (1990). Weather and crime. *Br. J. Criminol.* **30** 51–64.

CONSUL, P. and FAMOYE, F. (1992). Generalized Poisson regression model. *Comm. Statist. Theory Methods* **21** 89–109.

CUI, Y., LI, Q. and ZHU, F. (2020). Flexible bivariate Poisson integer-valued GARCH model. *Ann. Inst. Statist. Math.* https://doi.org/10.1007/s10463-019-00732-4

CUI, Y. and ZHU, F. (2018). A new bivariate integer-valued GARCH model allowing for negative cross-correlation. *TEST* **27** 428–452. MR3799147 https://doi.org/10.1007/s11749-017-0552-4

DAVIS, R. A., DUNSMUIR, W. T. M. and STREETT, S. B. (2003). Observation-driven models for Poisson counts. *Biometrika* **90** 777–790. MR2024757 https://doi.org/10.1093/biomet/90.4.777

DAVIS, R. A., LEE, T. C. M. and RODRIGUEZ-YAM, G. A. (2006). Structural break estimation for nonstationary time series models. *J. Amer. Statist. Assoc.* **101** 223–239. MR2268041 https://doi.org/10.1198/016214505000000745

FAMOYE, F. and SINGH, K. P. (2006). Zero-inflated generalized Poisson regression model with an application to domestic violence data. *J. Data Sci.* **4** 117–130.

FERLAND, R., LATOUR, A. and ORAICHI, D. (2006). Integer-valued GARCH process. *J. Time Series Anal.* **27** 923–942. MR2328548 https://doi.org/10.1111/j.1467-9892.2006.00496.x

FOKIANOS, K. and FRIED, R. (2010). Interventions in INGARCH processes. *J. Time Series Anal.* **31** 210–225. MR2676152 https://doi.org/10.1111/j.1467-9892.2010.00657.x

FOKIANOS, K. and FRIED, R. (2012). Interventions in log-linear Poisson autoregression. *Stat. Model.* **12** 299–322. MR3179504 https://doi.org/10.1177/1471082X1201200401

FOKIANOS, K., RAHBEK, A. and TJØSTHEIM, D. (2009). Poisson autoregression. *J. Amer. Statist. Assoc.* **104** 1430–1439. With electronic supplementary materials available online. MR2596998 https://doi.org/10.1198/jasa.2009.tm08270

FOKIANOS, K. and TJØSTHEIM, D. (2011). Log-linear Poisson autoregression. *J. Multivariate Anal.* **102** 563–578. MR2755016 https://doi.org/10.1016/j.jmva.2010.11.002

FOKIANOS, K. and TJØSTHEIM, D. (2012). Nonlinear Poisson autoregression. *Ann. Inst. Statist. Math.* **64** 1205–1225. MR2981620 https://doi.org/10.1007/s10463-012-0351-3

FOKIANOS, K., STØVE, B., TJØSTHEIM, D. and DOUKHAN, P. (2020). Multivariate count autoregression. *Bernoulli* **26** 471–499. MR4036041 https://doi.org/10.3150/19-BEJ1132

FRANKE, J., KIRCH, C. and KAMGAING, J. T. (2012). Changepoints in times series of counts. *J. Time Series Anal.* **33** 757–770. MR2969909 https://doi.org/10.1111/j.1467-9892.2011.00778.x

GUPTA, P. L., GUPTA, R. C. and TRIPATHI, R. C. (1996). Analysis of zero-adjusted count data. *Comput. Statist. Data Anal.* **23** 207–218.

HÄRDLE, W., MÜLLER, M., SPERLICH, S. and WERWATZ, A. (2012). *Nonparametric and Semiparametric Models*. Springer Science & Business Media. MR2061786 https://doi.org/10.1007/978-3-642-17146-8

HIRD, C. and RUPAREL, C. (2007). *Seasonality in Recorded Crime*: *Preliminary Findings* **2**. Home Office London. 07.

HSIANG, S. M., BURKE, M. and MIGUEL, E. (2013). Quantifying the influence of climate on human conflict. *Science* **341** 1235367. https://doi.org/10.1126/science.1235367

JAZI, M. A., JONES, G. and LAI, C.-D. (2012). First-order integer valued AR processes with zero inflated Poisson innovations. *J. Time Series Anal.* **33** 954–963. MR2991911 https://doi.org/10.1111/j.1467-9892.2012.00809.x

KANG, J. and LEE, S. (2014). Parameter change test for Poisson autoregressive models. *Scand. J. Stat.* **41** 1136–1152. MR3277042 https://doi.org/10.1111/sjos.12088

KARLIS, D. and PEDELI, X. (2013). Flexible bivariate INAR(1) processes using copulas. *Comm. Statist. Theory Methods* **42** 723–740. MR3211946 https://doi.org/10.1080/03610926.2012.754466

KRISTENSEN, K., NIELSEN, A., BERG, C. W., SKAUG, H. and BELL, B. (2016). TMB: Automatic differentiation and Laplace approximation. *J. Stat. Softw.* **70** 1–21.

LEE, S., LEE, Y. and CHEN, C. W. S. (2016). Parameter change test for zero-inflated generalized Poisson autoregressive models. *Statistics* **50** 540–557. MR3506657 https://doi.org/10.1080/02331888.2015.1083020

MARES, D. M. (2009). Civilization, economic change, and trends in interpersonal violence in western societies. *Theor. Criminol.* **13** 419–449.

MARES, D. M. and MOFFETT, K. W. (2016). Climate change and interpersonal violence: A "global" estimate and regional inequities. *Clim. Change* **135** 297–310.

MCDOWALL, D., LOFTIN, C. and PATE, M. (2012). Seasonal cycles in crime, and their variability. *J. Quant. Criminol.* **28** 389–410.

MERCURIO, D. and SPOKOINY, V. (2004). Statistical inference for time-inhomogeneous volatility models. *Ann. Statist.* **32** 577–602. MR2060170 https://doi.org/10.1214/009053604000000102

PEREIRA, D. V., ANDRESEN, M. A. and MOTA, C. M. (2016). A temporal and spatial analysis of homicides. *J. Environ. Psychol.* **46** 116–124.

RANSON, M. (2014). Crime, weather, and climate change. *J. Environ. Econ. Manage.* **67** 274–302.

ROTTON, J. and COHN, E. G. (2000). Violence is a curvilinear function of temperature in Dallas: A replication. *J. Pers. Soc. Psychol.* **78** 1074–1081.

SELLERS, K. F. and SHMUELI, G. (2010). A flexible regression model for count data. *Ann. Appl. Stat.* **4** 943–961. MR2758428 https://doi.org/10.1214/09-AOAS306

SPOKOINY, V. (2009). Multiscale local change point detection with applications to value-at-risk. *Ann. Statist.* **37** 1405–1436. MR2509078 https://doi.org/10.1214/08-AOS612

STACEY, M., CARBONE-LÓPEZ, K. and ROSENFELD, R. (2011). Demographic change and ethnically motivated crime: The impact of immigration on anti-Hispanic hate crime in the United States. *J. Contemp. Crim. Justice* **27** 278–298.

XU, X., CHEN, Y., CHEN, C. W. S. and LIN, X. (2020a). Supplement to "Adaptive log-linear zero-inflated generalized Poisson autoregressive model with applications to crime counts." https://doi.org/10.1214/20-AOAS1360SUPPA

XU, X., CHEN, Y., CHEN, C. W. S. and LIN, X. (2020b). Source code for "Adaptive log-linear zero-inflated generalized Poisson autoregressive model with applications to crime counts." https://doi.org/10.1214/20-AOAS1360SUPPB

YAN, Y. Y. (2004). Seasonality of property crime in Hong Kong. *Br. J. Criminol*. **44** 276–283.

ZHU, F. (2011). A negative binomial integer-valued GARCH model. *J. Time Series Anal*. **32** 54–67. MR2790672 https://doi.org/10.1111/j.1467-9892.2010.00684.x

ZHU, F. (2012a). Modeling overdispersed or underdispersed count data with generalized Poisson integer-valued GARCH models. *J. Math. Anal. Appl*. **389** 58–71. MR2876481 https://doi.org/10.1016/j.jmaa.2011.11.042

ZHU, F. (2012b). Zero-inflated Poisson and negative binomial integer-valued GARCH models. *J. Statist. Plann. Inference* **142** 826–839. MR2863870 https://doi.org/10.1016/j.jspi.2011.10.002