

Statistica Sinica Preprint No: SS-2025-0448	
Title	Double Dynamic Bounded Full-integer-valued Garch Model
Manuscript ID	SS-2025-0448
URL	http://www.stat.sinica.edu.tw/statistica/
DOI	10.5705/ss.202025.0448
Complete List of Authors	Huaping Chen, QianQian Zhao, Liugen Xue and Fukang Zhu
Corresponding Authors	Fukang Zhu
E-mails	zfk8010@163.com
Notice: Accepted author version.	

Double Dynamic Bounded Full-integer-valued GARCH Model

Huaping Chen¹, Qianqian Zhao¹, Liugen Xue¹ and Fukang Zhu²

¹ *Henan University and* ² *Jilin University*

Abstract: This article introduces a novel double dynamic trinomial GARCH model with the simultaneous modeling of the time-varying mean and variance to analyze bounded full-integer-valued time series. In this paper, we construct a trinomial autoregressive process to analyze zero-mean process with time-varying variance process, and then propose a novel double dynamic trinomial autoregressive process to analyze non-zero and time-varying mean process, besides the time-varying variance process. Furthermore, we discuss their stationarity and ergodicity, give the conditional maximum likelihood estimate, and derive their asymptotic properties. We also provide details on one-step-ahead forecasting. The performances of these estimators are compared via simulation study. Finally, we apply the proposed model to two datasets to demonstrate its favorable performance.

Key words and phrases: trinomial difference distribution, bounded full-integer-valued time series, GARCH model, time-varying variance, time-varying mean, statistical inference.

1. Introduction

There is ascending interest in the analysis and modeling of bounded full-integer-valued time series, which takes values in $\{-n, \dots, 0, \dots, n\}$, $n \in \mathbb{Z}^+$. Consequently, a variety of related models and methods have been proposed in the literature. Among these, three commonly used approaches are most frequently adopted for modeling such data. The first one is autoregressive model constructed by thinning operator, see Möller and Weiß (2020) and Chen (2025). The second one is autoregressive process constructed by a discrete distribution, see Chen et al. (2025a). The third one is the shifted method, which transforms the bounded full-integer-valued time series to take values in $\{0, \dots, 2n\}$. Then the shifted data can be analyzed by using existing nonnegative integer-valued autoregressive models as well as non-negative integer-valued ARCH or GARCH models. Among of these can be regarded as conditional-mean modeling approaches, although the last one specifies dynamic behavior for both the conditional mean and variance by allowing a common factor to vary over time.

To be specific, a binomial GARCH(1,1) model $\{Z_t, t \in \mathbb{Z}\}$ is defined by

$$Z_t | \mathcal{F}_{t-1} : \text{Bin}(n, p_t), \quad \mu_t := np_t = g_{\theta}(\mu_{t-1}, Z_{t-1}), \quad (1.1)$$

where $\mathcal{F}_t$ is the σ -field generated by $\{Z_0, Z_1, Z_2, \dots, Z_t\}$, $g_{\theta}(\mu_{t-1}, Z_{t-1})$ is a

nonnegative and twice continuously differentiable function in terms of each θ_j for given Z_{t-1} and μ_{t-1} , $\forall j = 1, 2, \dots, d$, $\boldsymbol{\theta}$ is the parameter vector involving in the link function $g_{\boldsymbol{\theta}}(\cdot, \cdot)$. The binomial ARCH-type models have been initially proposed by Weiß and Pollett (2014), explored by Ristić et al. (2016) and Chen et al. (2020). Extensions incorporating feedback mechanisms were explored by Lee and Lee (2019) and Weiß and Jahn (2024). They specify dynamic behavior for both the conditional mean ($\mu_t = np_t$) and variance ($\sigma_t^2 = np_t(1 - p_t)$) by allowing p_t to vary over time. See Chen et al. (2022), Chen et al. (2025b) and Chen et al. (2023a) for some variants.

By (1.1), the nomenclature of the Binomial ARCH or GARCH model stems from the fact that for a Binomial process $\text{Bin}(n, p_t)$, its variance and mean share a common time-varying factor p_t , thus modeling the mean implies modeling variance. Moreover, its formulation is similar to that of the classic ARCH (Engle, 1982) or GARCH model (Bollerslev, 1986). Consequently, such models are considered in the literature as a nonnegative bounded integer-valued counterpart of the ARCH model (Engle, 1982) or GARCH model (Bollerslev, 1986), see review papers Davis and Liu (2016), Weiß (2018), Davis et al. (2021), Kim et al. (2020) and Weiß (2021) for more discussions.

Nevertheless, existing binomial ARCH-type and GARCH-type models

or their extensions ignore the inherent rising and falling fluctuations of the original data, thereby losing the intrinsic information carried by positive and negative values. See Möller and Weiß (2020) and Chen (2025) for more discussion. To address this challenge, we propose a novel class of double dynamic full-integer-valued GARCH processes, in which both the conditional mean and conditional variance are jointly specified to evolve dynamically. The proposed model is constructed by using the reparameterized trinomial difference process with simultaneously time-varying mean and variance parameters. And we call it as the TDGMV model. An appealing merit of the TDGMV model lies in the separability of its mean and variance processes. This not only facilitates both the conditional mean and variance of the novel model to evolve over time, but also enables the novel model to flexibly capture the counts of upward and downward changes in the observed data. The usefulness of the TDGMV model over the ordinary nonnegative bounded integer-valued GARCH model will be illustrated in Section 5, which demonstrates that the TDGMV model outperforms existing binomial GARCH-type models and their extensions.

In fact, it is indispensable to formulate GARCH model by allowing both of its mean and variance to vary over time, rather than allowing only the conditional mean or variance to evolve in isolation. As discussed in En-

gle (1982), it makes the autoregressive model attractive by incorporating heteroskedasticity, because practical forecasters have empirically verified that predictive performance differs markedly across distinct sample periods. Especially in monetary and financial theory, portfolio allocation is determined jointly by the anticipated mean and variance of asset yields, whereby shifts in asset demand inherently correspond to co-movements of these two parts. In addition, Barreto-Souza et al. (2026) developed the integer-valued GARCH model with simultaneously time-varying mean and dispersion parameters, which outperforms existing INGARCH models in real data. Furthermore, Wendler et al. (2025) and Haq (2025) demonstrated that the CUSUM and EWMA charts constructed by joint mean and variance take lower false alarm rate, respectively.

As elaborated previously, joint mean-variance modeling has been extensively investigated for continuous-valued time series, non-negative integer-valued time series with unbounded support, and the context in statistical process control, yet relevant research remains neglected in the context of the bounded full-integer-valued time series. Although the binomial GARCH model and its extensions allow their variance to vary over time along with their mean, but this in turn is solely driven by the dynamics of the mean process. There exist two unavoidable drawbacks of this approach, where

the first one that it imposes severe restrictions on its volatility as illustrated in the real application, and the second one is that it can not describe the counts of the inherent rising and falling fluctuations for the real data. In contrast, our proposed TDGMV model not only relaxes the assumption on its volatility, but also flexibly accommodates the counts with regard to upward and downward fluctuation patterns of the data.

A prominent contribution lies in the stationarity and ergodicity properties of the TDGMV models. For the TDGV model, a special case of the TDGMV model, we follow the approach outlined in Doukhan et al. (2021) to demonstrate that the bivariate process $\{(X_t, \sigma_t^2), t \in \mathbb{Z}\}$ satisfies both the contraction condition (A1) and the stochastically small condition (A2) as specified in Doukhan et al. (2021). In contrast, $\{(X_t, \sigma_t^2), t \in \mathbb{Z}\}$ associated with the TDGMV model no longer exhibits the Markov property. Consequently, the methodology proposed in Doukhan et al. (2021) is not applicable for the TDGMV model. Here, we first account for that the triple process $\{(X_t, \mu_t, \sigma_t^2), t \in \mathbb{Z}\}$ constitutes a Bernoulli-shift sequence, while the bivariate latent process $\{(\mu_t, \sigma_t^2), t \in \mathbb{Z}\}$ is an e-chain and possesses the geometric moment contraction property. Leveraging these preliminary results, we then verify the stationarity and ergodicity of the TDGMV model by adapting the methodology outlined in Doukhan and Wintenberger (2008).

The remainder of the paper is organized as follows. Section 2 first gives a brief review of the trinomial difference distribution, explores its total variation distance and discusses its reparameterization, then proposes the TDGMV model, including its statistical property. Section 3 investigates its statistical inference, including the conditional maximum likelihood estimation and forecasting. Section 4 conducts some simulations to evaluate model performance, while Section 5 demonstrates the application of the proposed models through two illustrative examples. Concluding remarks are provided in Section 6.

2. Model Formulation and Stability Property

2.1 Preliminaries

To enhance readability, this section first introduces the trinomial difference distribution and its reparameterized form, then investigates its total variation property.

Definition 1. (Trinomial Difference Distribution, Omair et al. (2016)) A random variable Z taking values in $\{-n, \dots, -1, 0, 1, \dots, n\}$ is said to follow the trinomial difference distribution with parameters (α_1, α_2) , if its

2.1 Preliminaries

probability mass function (pmf) takes the form

$$\mathbb{P}_n(Z = k) = \sum_{j=\max(0, -k)}^{\lfloor (n-k)/2 \rfloor} \binom{n}{k+j, j} \alpha_1^{k+j} \alpha_2^j (1 - \alpha_1 - \alpha_2)^{n-k-2j}, \quad (2.2)$$

where $\binom{n}{k+j, j} = \frac{n!}{(k+j)!j!(n-k-2j)!}$, $\lfloor (n-k)/2 \rfloor$ stands for the integer part of the value $(n-k)/2$, $n \in \mathbb{Z}^+$, $0 < \alpha_i < 1$ and $\alpha_1 + \alpha_2 < 1$, $i = 1, 2$.

For simplicity, we write $Z \sim \text{TD1}(n, \alpha_1, \alpha_2)$. Specially, if $n = 1$, Z takes values -1 , 0 and 1 with probability α_2 , $1 - \alpha_1 - \alpha_2$ and α_1 , respectively. By (2.2), the moment generating function of $Z \sim \text{TD1}(n, \alpha_1, \alpha_2)$ is given as $M_Z(s) = (1 - \alpha_1 - \alpha_2 + \alpha_1 e^s + \alpha_2 e^{-s})^n$. Therefore, $E(Z) = n(\alpha_1 - \alpha_2)$, $\text{Var}(Z) = n(\alpha_1 + \alpha_2 - (\alpha_1 - \alpha_2)^2)$.

Proposition 1. *For all $i = 1, 2, \dots, n$, if $\{Z_i\}$ is an independent and identically distributed random variable sequence and $Z_1 \sim \text{TD1}(1, \alpha_1, \alpha_2)$, then*

$$(1) \sum_{i=1}^n Z_i \sim \text{TD1}(n, \alpha_1, \alpha_2).$$

$$(2) \sum_{i=1}^n |Z_i| \sim \text{Bin}(n, \alpha_1 + \alpha_2).$$

(3) *If $Z \sim \text{TD1}(n, \alpha_1, \alpha_2)$ and $Y = |Z|$, then $\mathbb{P}(Z = k, Y = l) = \mathbb{I}(|k| = l) \mathbb{P}_n(Z = k)$, if $l \geq 0$, and otherwise, $\mathbb{P}(Z = k, Y = l) = 0$. where $\mathbb{I}(|k| = l) = 1$, if $|k| = l$, else 0 .*

Proposition 2. *Let $n \in \mathbb{N}$, $\tau_1 = (\alpha_1, \alpha_2)^\top$ and $\tau_2 = (\beta_1, \beta_2)^\top$. For some*

2.2 Double Dynamic TDGARCH model

$\delta > 0$, if $\alpha_1 \geq \beta_1 > n/(2\delta) > 0$ and $\alpha_2 \geq \beta_2 > n/(2\delta) > 0$, then

$$\text{TV}(\text{TD1}(n, \alpha_1, \alpha_2), \text{TD1}(n, \beta_1, \beta_2)) \leq 1 - \exp(-\delta \|\tau_1 - \tau_2\|), \quad (2.3)$$

where $\text{TV}(Q_1, Q_2) = \sup_{A \in S} |Q_1(A) - Q_2(A)|$ denotes the total variation distance between probability measures Q_1 and Q_2 , S is the state space.

By assume $\alpha = \alpha_1 - \alpha_2$ and $\nu = \alpha_1 + \alpha_2 - \alpha^2$, Chen et al. (2023b) designed the reparameterization of $Z \sim \text{TD1}(n, \alpha_1, \alpha_2)$ and denoted it as $Z \sim \text{TD2}(n, \alpha, \nu)$, where $0 < \nu < 1$, $-1 < \alpha < 1$ and $0 < \nu + \alpha^2 < 1$. Furthermore, $|Z| \sim \text{Bin}(n, \alpha^2 + \nu)$ by using Proposition 1. Hence,

$$\text{E}(Z) = n\alpha =: \mu, \quad \text{Var}(Z) = n\nu =: \sigma^2, \quad \text{and } \text{E}|Z| = \mu^2/n + \sigma^2.$$

2.2 Double Dynamic TDGARCH model

We hereby propose a double dynamic TDGARCH model to describe time-varying mean and volatility behaviors simultaneously.

Definition 2. Let $\mathcal{F}_t$ be the σ -field generated by $\{(X_t, \mu_t, \sigma_t^2), t \in \mathbb{N}\}$ and $n \in \mathbb{N}$ be the predetermined upper limit of the range. If $X_t | \mathcal{F}_{t-1} \sim \text{TD2}(n, p_t, \nu_t)$ with

$$\mu_t := np_t = \alpha_0 + \alpha_1 X_{t-1} + \beta_1 \mu_{t-1}, \quad (2.4)$$

$$\sigma_t^2 := n\nu_t = \phi_0 + \phi_1 |X_{t-1}| + \psi_1 \sigma_{t-1}^2, \quad (2.5)$$

2.2 Double Dynamic TDGARCH model

then $\{X_t\}$ is called as the double dynamic TDGARCH model, where $\alpha_0 \neq 0$, $0 < |\alpha_1|, |\beta_1| < 1$, $\phi_0 > 0$, $0 \leq \phi_1, \psi_1 < 1$, and $\max(|\beta_1|, \psi_1) + 3 \times 2^{1-1/p} \max(|\alpha_1|, \phi_1) < 1$ for some $p > 1$.

For simplicity, we denote the new model given in Definition 2 as the TDGMV((1,1),(1,1)) model. Formally, the new TDGMV model demonstrates enhanced flexibility relative to its antecedents, i.e., those only with the conditional mean or variance to evolve in isolation. Therefore, we call the former one as the TDGM model, which has been discussed by Chen et al. (2025a). The latter one is called as the TDGV model with a zero mean. Specially, the TDGMV((1,1),(1,1)) model reduces to the TDGM(1,1) model, if $\phi_0 = 0, \phi_1 = 0$ and $\psi_1 = 0$. Conversely, if $\alpha_0 = 0, \alpha_1 = 0$ and $\beta_1 = 0$, then the TDGMV((1,1),(1,1)) model becomes the TDGV(1,1) model.

Denote $\zeta_t^{(1)} = \frac{\sigma_t^2}{2n} + \frac{\mu_t^2}{2n^2} + \frac{\mu_t}{2n}$, $\zeta_t^{(2)} = \frac{\sigma_t^2}{2n} + \frac{\mu_t^2}{2n^2} - \frac{\mu_t}{2n}$ and $\zeta_t^{(3)} = 1 - \zeta_t^{(1)} - \zeta_t^{(2)}$. Then the conditional pmf of $\{X_t\}$ takes the following form

$$\mathbb{P}_n(X_t = k | \mathcal{F}_{t-1}) = \sum_{j=m_1}^{m_2} \binom{n}{k+j, j} \left(\zeta_t^{(1)}\right)^{k+j} \left(\zeta_t^{(2)}\right)^j \left(\zeta_t^{(3)}\right)^m, \quad (2.6)$$

where $m_1 = \max(0, -k)$, $m_2 = \left[\frac{n-k}{2}\right]$, $m = n - k - 2j$ and $n \in \mathbb{Z}^+$.

Denote $\boldsymbol{\lambda}_t = (\mu_t, \sigma_t^2)^\top$, $\mathbf{Y}_t = (X_t, |X_t|)^\top$ and $\boldsymbol{\delta}_1 = (\alpha_0, \phi_0)^\top$. Then bivariate latent process $\{(\mu_t, \sigma_t^2)\}$ given in (2.4) and (2.5) is rewritten as

$$\boldsymbol{\lambda}_{t+1} = \boldsymbol{\delta}_1 + \mathbf{A}\boldsymbol{\lambda}_t + \mathbf{B}\mathbf{Y}_t \quad (2.7)$$

2.2 Double Dynamic TDGARCH model

with $\mathbf{A} = \text{diag}(\beta_1, \psi_1)$ and $\mathbf{B} = \text{diag}(\alpha_1, \phi_1)$. By Definition 2, $\rho(\mathbf{A}) < 1$.

Hence, $\boldsymbol{\lambda}_{t+1} = (\mathbf{I} + \mathbf{A} + \cdots + \mathbf{A}^{l-1})\boldsymbol{\delta}_1 + \mathbf{A}^l\boldsymbol{\lambda}_{t-l} + \sum_{k=0}^{l-1} \mathbf{A}^k \mathbf{B} \mathbf{Y}_{t-k-l}$, where

$\mathbf{I} = \text{diag}(1, 1)$. Specially, by $\rho(\mathbf{A}) < 1$, we obtain that

$$\boldsymbol{\lambda}_{t+1} = (\mathbf{I} - \mathbf{A})^{-1}\boldsymbol{\delta}_1 + \sum_{k=0}^{\infty} \mathbf{A}^k \mathbf{B} \mathbf{Y}_{t-k-1} \geq (\mathbf{I} - \mathbf{A})^{-1}\boldsymbol{\delta}_1, \forall t, \quad (2.8)$$

the equation holds if and only if $\mathbf{Y}_l = \mathbf{Y}_{l-1} = \cdots = \mathbf{Y}_1 = 0$ for $l \in \mathbb{N}$ large enough. Hence, $(\mathbf{I} - \mathbf{A})^{-1}\boldsymbol{\delta}_1$ is a reachable state.

By (2.4) and (2.5), the triple process $\{(X_t, \mu_t, \sigma_t^2)\}$ is a time homogeneous Markov chain. To derive its stability property, we let $f_s(\boldsymbol{\lambda})$ be the random function with $\|f_s\|_p \leq 1$ and $f_s(\boldsymbol{\lambda}) = \boldsymbol{\delta}_1 + \mathbf{A}\boldsymbol{\lambda} + \mathbf{B}F_{\boldsymbol{\lambda}}^{-1}(\mathbf{s})$, where $\boldsymbol{\lambda} = (\lambda_1, \lambda_2)^\top = (\mu, \sigma^2)^\top$, $F_{\boldsymbol{\lambda}}^{-1}(\mathbf{s}) = (F_{\mu}^{-1}(s_1), F_{\mu^2/n^2+\sigma^2/n}(s_2))$ with $F_x^{-1}(s) = \inf\{t \geq 0, F_x(t) \geq s\}$, $\mathbf{s} = (s_1, s_2) \in (0, 1)^2$. Denote

$$\mathbf{D} = \begin{pmatrix} 1 & 0 \\ (\mu + \mu_0)/n & 1 \end{pmatrix}, \forall (\mu, \mu_0)^\top \in [-n, n]^2.$$

Obviously, $\|\mathbf{D}\|_1 = 1 + |\mu + \mu_0|/n$. If $\|\mathbf{A}\|_p + 2^{1-1/p}\|\mathbf{D}\|_1\|\mathbf{B}\|_p < 1$, then

$$\max(|\beta_1|, \psi_1) + 2^{1-1/p}(1 + |\mu + \mu_0|/n) \max(|\alpha_1|, \phi_1) < 1. \quad (2.9)$$

Hence, $\{\boldsymbol{\lambda}_t\}$ is a geometric moment contraction process by Proposition 4.2.1 in Liu (2012). See supplementary material S2 for an in-depth elaboration.

Remark 1. To be precise, if $\mu + \mu_0 = 0$ ($\forall t, \mu_t = 0$ or $\mu_t = \alpha_0$), (2.9) reduces to $\max(|\beta_1|, \psi_1) + 2^{1-1/p} \max(|\alpha_1|, \phi_1) < 1$ ($\psi_1 + \phi_1 < 1$), i.e.,

2.2 Double Dynamic TDGARCH model

$\|\mathbf{A}\|_p + 3 \times 2^{1-1/p} \|\mathbf{B}\|_p < 1$, which is similar to those given in Liu (2012) and Barreto-Souza et al. (2026).

In fact, for all $(\mu, \mu_0)^\top \in [-n, n]^2$, $|\mu + \mu_0|/n \leq 2$. Thus,

$$\|\mathbf{A}\|_p + 2^{1-1/p} \|\mathbf{D}\|_1 \|\mathbf{B}\|_p \leq \|\mathbf{A}\|_p + 3 \times 2^{1-1/p} \|\mathbf{B}\|_p.$$

Therefore, if $\|\mathbf{A}\|_p + 3 \times 2^{1-1/p} \|\mathbf{B}\|_p < 1$, i.e., $\max(|\beta_1|, \psi_1) + 3 \times 2^{1-1/p} \max(|\alpha_1|, \phi_1) < 1$ for some $p > 1$, then $\|\mathbf{A}\|_p + 2^{1-1/p} \|\mathbf{D}\|_1 \|\mathbf{B}\|_p < 1$, i.e., (2.9) holds.

Assumption 1. Let $\boldsymbol{\theta} = (\alpha_1, \phi_1, \alpha_0, \beta_1, \phi_0, \psi_1)^\top$ be the parameter vector.

The parameter space Θ is a compact subset of $\mathbb{R}^6$ with $\boldsymbol{\theta} \in \Theta$, where $\boldsymbol{\theta}_0$ is an interior point of Θ , $0 < \underline{\alpha}_0 \leq |\alpha_0| \leq \bar{\alpha}_0 \leq n$, $0 < \underline{\alpha}_1 \leq |\alpha_1| \leq \bar{\alpha}_1 < 1$, $0 < \underline{\beta}_1 \leq |\beta_1| \leq \bar{\beta}_1 < 1$, $0 < \underline{\phi}_0 \leq \phi_0 \leq \bar{\phi}_0 \leq n$, $0 < \underline{\phi}_1 \leq \phi_1 \leq \bar{\phi}_1 < 1$, $0 < \underline{\psi}_1 \leq \psi_1 \leq \bar{\psi}_1 < 1$, $\max(|\beta_1|, \psi_1) + 3 \times 2^{1-1/p} \max(|\alpha_1|, \phi_1) < 1$ for some $p > 1$, all of $\underline{\alpha}_0, \underline{\alpha}_1, \underline{\beta}_1, \underline{\phi}_0, \underline{\phi}_1, \underline{\psi}_1, \bar{\alpha}_0, \bar{\alpha}_1, \bar{\beta}_1, \bar{\phi}_0, \bar{\phi}_1$ and $\bar{\psi}_1$ are finite positive constants.

Theorem 1. Let $\{X_t\}$ be the TDGMV((1,1),(1,1)) process and $\{\mathbf{Z}_t = (X_t, \mu_t, \sigma_t^2)^\top\}$. If Assumption 1 holds, then

- (1) $\{\mathbf{Z}_t\}$ exhibits a Bernoulli-shift structure.
- (2) there exists a unique stationary distribution to $\{\mathbf{Z}_t\}$.
- (3) $\{\mathbf{Z}_t\}$ is a geometrical moment contract Markov chain with unique stationary and ergodic distribution.

2.2 Double Dynamic TDGARCH model

Remark 2. For the TDGV(1,1) model, one also can establish the stationarity and ergodicity of the TDGV model by verifying the conditions (A1) and (A2) in Doukhan et al. (2021). Here, the corresponding complete technical proofs are omitted due to page limitations. See supplementary materials S2 for details.

Proposition 3. *Let $\{X_t\}$ be the TDGMV((1,1),(1,1)) process . If Assumption 1 holds, then*

$$\begin{aligned} (1) \mu_X &:= E(X_t) = \frac{\alpha_0}{1 - \alpha_1 - \beta_1}; \\ (2) \text{Var}(X_t) &= \frac{n(1 - \beta_1^2)\phi_0 + \phi_1(1 - \beta_1^2)\mu_X^2}{n(1 - \phi_1 - \psi_1)(1 - \alpha_1^2 - \beta_1^2) - \phi_1\alpha_1^2}. \end{aligned}$$

Denote $R_X(k) = \text{Cov}(X_t, X_{t-k})$ and $R_\mu(k) = \text{Cov}(\mu_t, \mu_{t-k})$, $k \geq 0$.

Similar to Corollary 1 in Chen et al. (2025a), we obtain that

$$R_X(k) = \begin{cases} \alpha_1 R_X(0) + \beta_1 R_\mu(0), & k = 1, \\ (\alpha_1 + \beta_1)^{k-1} \text{Var}(X_t), & k \geq 2, \end{cases} \quad \text{and } R_\mu(k) = (\alpha_1 + \beta_1)^k R_\mu(0), k \geq 0.$$

Therefore,

$$\rho_X(k) = \frac{R_X(k)}{R_X(0)} = \begin{cases} \alpha_1 + \beta_1 R_\mu(0)/R_X(0), & k = 1, \\ (\alpha_1 + \beta_1)^{k-1} \rho_X(1), & k \geq 2, \end{cases} \quad (2.10)$$

which implies that both negative and positive autocorrelations are allowed by the TDGMV((1,1),(1,1)) model. Additionally, the ACF presented in

2.2 Double Dynamic TDGARCH model

(2.10) has the same form as that for TDGM(1,1) model. This indicates that the variance process involved in the TDGMV model does not alter the auto-correlation structure. However, this is the case except for the value of the ACF(1), other aspects show consistency as described previous.

Remark 3. Denote $\Delta_t = 1 - \alpha_t^2 - 2\nu_t$. If $\Delta_t \geq 0, \forall t \in \mathbb{N}$, there exist two independent nonlinear binomial GARCH model denoted as $\{X_t^{(1)}\}$ and $\{X_t^{(2)}\}$, such that $\{X_t\}$ given in Definition 2 is rewritten as

$$X_t = X_t^{(1)} - X_t^{(2)} \text{ with } X_t^{(i)} | \mathcal{F}_{t-1}^{(i)} \sim \text{Bin}(n, p_t^{(i)}), i = 1, 2, \quad (2.11)$$

where $p_t^{(1)} = (1 + \alpha_t + \sqrt{\Delta_t})/2$ and $p_t^{(2)} = (1 - \alpha_t + \sqrt{\Delta_t})/2$ (or $p_t^{(1)} = (1 + \alpha_t - \sqrt{\Delta_t})/2$ and $p_t^{(2)} = (1 - \alpha_t - \sqrt{\Delta_t})/2$), $\mathcal{F}_t^{(i)}$ is the σ -field generated by $\{(X_t^{(i)}, p_t^{(i)})\}$, $\alpha_t = \mu_t/n$ and $\nu_t = \sigma_t^2/n$. This highlights that it is infeasible by using translation techniques with regard to state space, when one analyzes the complex full-integer-valued time series.

Remark 4. It is worth mentioning that $\{\sigma_t^2\}$ involved in the TDGMV model is a linear function with regard to $\{|X_{t-1}|\}$, rather than $\{X_{t-1}^2\}$. Then, a natural question arises: why not use $\{X_{t-1}^2\}$ here? This arise from the following two aspects. The first reason is that $\forall t, \sigma_t^2 \in (0, n)$ because $\{\nu_t\} \in (0, 1)$. The second reason is that this design facilitates the derivation of the explicit expression of $\text{Var}(X_t)$. Therefore, we employ $\{|X_{t-1}|\}$ to

construct $\{\sigma_t^2\}$.

Remark 5. In fact, it deserves to consider that why not construct $\{\sigma_t^2\}$ by using normalization method with $\nu_t = \phi_0^* + \phi_1^* X_{t-1}^2/n^2 + \psi_1^* \nu_{t-1}$? To illustrate it, we denote $\phi_0 = n\phi_0^*$, $\psi_1 = \psi_1^*$ and $\tilde{\phi}_1 = \phi_1^* |X_{t-1}|/n$. Then $\{\sigma_t^2\}$ is rewritten as $\sigma_t^2 = \phi_0 + \tilde{\phi}_1 |X_{t-1}| + \psi_1 \sigma_{t-1}^2$. This yields a volatility process with a time-varying coefficient, which will be explored in a future research.

3. Statistical Inference

In this section, we focus on the parameter estimation and forecasting of the TDGMV model, and let $\boldsymbol{\theta} = (\boldsymbol{\theta}_1^\top, \boldsymbol{\theta}_2^\top)^\top$ with $\boldsymbol{\theta}_1 = (\alpha_0, \alpha_1, \beta_1)^\top$ and $\boldsymbol{\theta}_2 = (\phi_0, \phi_1, \psi_1)^\top$, $X_1, \dots, X_T$ be a realization of $\{X_t\}$, where $T \in \mathbb{N}$ represents sample size.

3.1 Parameter Estimation

By (2.6), the conditional log-likelihood function is written as:

$$\ell(\boldsymbol{\theta}) = \sum_{t=1}^T \log \mathbb{P}_n(X_t(\boldsymbol{\theta}) = x_t | \mathcal{F}_{t-1}), \quad (3.12)$$

where $\mathbb{P}_n(X_t(\boldsymbol{\theta}) = x_t | \mathcal{F}_{t-1})$ is given in (2.6). Denote $l_t(\boldsymbol{\theta}) = \log \mathbb{P}_n(X_t(\boldsymbol{\theta}) = x_t | \mathcal{F}_{t-1})$. The first and second derivatives of $l_t(\boldsymbol{\theta})$ with respect to $\boldsymbol{\theta}$ are expressed analytically with closed-form, see supplementary material S1.

3.1 Parameter Estimation

Lemma 1. *If Assumption 1 holds, $\{X_t\}$ is identifiable.*

Theorem 2. *Let $\{X_t\}$ be the TDGMV((1,1),(1,1)) model. If Assumption 1 holds, then, as $T \rightarrow \infty$,*

(1). *there exists an estimator $\hat{\boldsymbol{\theta}}$ such that $\hat{\boldsymbol{\theta}} \xrightarrow{a.s.} \boldsymbol{\theta}_0$;*

(2). *$\sqrt{T}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \mathbf{M}^{-1}(\boldsymbol{\theta}_0) \mathbf{V}(\boldsymbol{\theta}_0) \mathbf{M}^{-1}(\boldsymbol{\theta}_0))$,*

where $\mathbf{M}(\boldsymbol{\theta}_0) = \mathbb{E} \left(\frac{\partial^2 l_t(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top} \right)$ and $\mathbf{V}(\boldsymbol{\theta}_0) = \mathbb{E} \left(\frac{\partial l_t(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}} \frac{\partial l_t(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}^\top} \right)$.

Remark 6. In application, one often consistently estimates the asymptotic covariance of $\hat{\boldsymbol{\theta}}$ by $\hat{\mathbf{M}}_T^{-1} \hat{\mathbf{V}}_T \hat{\mathbf{M}}_T^{-1}$, where $\hat{\mathbf{M}}_T = T^{-1} \sum_{t=1}^T \partial^2 l_t(\hat{\boldsymbol{\theta}}) / \partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top$ and $\hat{\mathbf{V}}_T = T^{-1} \sum_{t=1}^T \partial l_t(\hat{\boldsymbol{\theta}}) / \partial \boldsymbol{\theta} \partial l_t(\hat{\boldsymbol{\theta}}) / \partial \boldsymbol{\theta}^\top$.

Remark 7. As is well-known, to validate the assumption of the simultaneous time-varying mean and variance, it is necessary to test the nullity of the simultaneous time-varying mean parameter α_1 and variance parameter ϕ_1 . Specially, (1) when $\alpha_1 = 0$, $\{\mu_t\}$ reduces to a constant value (α_0) if $\beta_1 = 0$, else $\mu_t = \alpha_0(1 + \beta_1 + \cdots + \beta_1^{t-1}) + \beta_1^t \mu_0 \rightarrow \alpha_0/(1 - \beta_1)$, as $t \rightarrow \infty$. (2) when $\phi_1 = 0$, $\{\sigma_t^2\}$ reduces to a constant value (ϕ_0), if $\psi_1 = 0$, else $\sigma_t^2 = \phi_0(1 + \psi_1 + \cdots + \psi_1^{t-1}) + \psi_1^t \sigma_0^2 \rightarrow \phi_0/(1 - \psi_1)$, as $t \rightarrow \infty$; (3) when $\alpha_1 = \phi_1 = 0$, both $\{\mu_t\}$ and $\{\sigma_t^2\}$ are constants ($\mu_t = \alpha_0$, $\sigma_t^2 = \phi_0$) or exhibit convergence to constant values. Hence, by examining the nullity of α_1 and ϕ_1 , one can determine whether the dataset exhibits a single dynamic structure or a double dynamic structure.

3.2 Forecasting

Lagrange multiplier (LM) test developed by Engle (1982) is a widely employed method for testing heteroscedasticity. Note that TDGV model is formally similar to the traditional GARCH model, thus, LM test is available method to test the nullity of the simultaneous time-varying variance parameter ϕ_1 . As for the nullity of the simultaneous time-varying mean parameter α_1 , Chen et al. (2025a) noted that LM test is also appropriate for the TDGM model. It is because the TDGM model is formally analogous to the integer-valued GARCH model (Ferland et al., 2006) and the null hypothesis ($H_0 : \alpha_1 = 0$) of LM test is designed to examine whether the lagged coefficients are zero. In real application, if the null hypothesis is rejected in both cases, a double dynamic model is considered.

3.2 Forecasting

In this study, we focus on the one-step-ahead forecast (Freeland and McCabe, 2004), which is often adopted in practice. By (2.6), we derive the one-step forecasting conditional distribution $\mathbb{P}_n(x|\mathcal{F}_T, \boldsymbol{\theta}) = \mathbb{P}_n(X_{T+1} = x|\mathcal{F}_T)$, where $\mathcal{F}_T$ is the σ -field generated by $\{(X_k, \mu_l, \sigma_s^2)\}$ for $k, l, s = 1, \dots, T$.

Denote $Z_T^{(1)} = (1, X_T, \mu_T)$, $Z_T^{(2)} = (1, |X_T|, \sigma_T^2)$. Then $\mu_{T+1} = Z_T^{(1)} \boldsymbol{\theta}_1$, $\sigma_{T+1}^2 = Z_T^{(2)} \boldsymbol{\theta}_2$. By Theorem 2 of Freeland and McCabe (2004), when x spans the entire state space $\{-n, \dots, -1, 0, 1, \dots, n\}$, $\mathbb{P}_n(x|\mathcal{F}_T, \hat{\boldsymbol{\theta}})$ is

asymptotically normal with mean $\mathbb{P}_n(x|\mathcal{F}_T, \boldsymbol{\theta}_0)$ and variance $T^{-1}\boldsymbol{D}\boldsymbol{\Sigma}\boldsymbol{D}^\top$, i.e.,

$$\sqrt{T} \left(\mathbb{P}_n(x|\mathcal{F}_T, \hat{\boldsymbol{\theta}}) - \mathbb{P}_n(x|\mathcal{F}_T, \boldsymbol{\theta}_0) \right) \xrightarrow{L} \mathcal{N}(\mathbf{0}, \boldsymbol{D}\boldsymbol{\Sigma}\boldsymbol{D}^\top), \quad (3.13)$$

where $\hat{\boldsymbol{\theta}}$ is the maximum likelihood estimator, $\boldsymbol{\Sigma} := \boldsymbol{M}^{-1}(\boldsymbol{\theta}_0)\boldsymbol{V}(\boldsymbol{\theta}_0)\boldsymbol{M}^{-1}(\boldsymbol{\theta}_0)$ is the covariance matrix given in Theorem 2 and $\boldsymbol{D} = (\partial\mathbb{P}_n(x|\mathcal{F}_T, \boldsymbol{\theta}_0)/\partial\boldsymbol{\theta})$. See supplementary material S1 for the closed form of $\partial\mathbb{P}_n(x|\mathcal{F}_T, \boldsymbol{\theta}_0)/\partial\boldsymbol{\theta}$.

Therefore, the $100(1 - \alpha)\%$ confidence interval for $\mathbb{P}_n(x|\mathcal{F}_T, \boldsymbol{\theta}_0)$ is determined with

$$\left(\mathbb{P}_n(x|\mathcal{F}_T, \hat{\boldsymbol{\theta}}) - Z_{1-\alpha/2}\sigma/\sqrt{T}, \mathbb{P}_n(x|\mathcal{F}_T, \hat{\boldsymbol{\theta}}) + Z_{1-\alpha/2}\sigma/\sqrt{T} \right),$$

where $\sigma = \sqrt{\boldsymbol{D}\boldsymbol{\Sigma}\boldsymbol{D}^\top}$ and $Z_{1-\alpha/2}$ is the $(1 - \alpha/2)$ -lower quantile of $\mathcal{N}(0, 1)$.

4. Simulation Study

In this section, we discuss the finite-sample properties of the ML estimator under the following two scenarios.

Scenario I: Specified TDGV and TDGMV models. In the case, we first perform simulation study to assess the finite-sample properties of the ML estimator $\hat{\boldsymbol{\theta}}$ under specified TDGV and TDGMV models. Then we employ the empirical bias, empirical standard deviation (ESD) and average asymptotic standard deviation (ASD) as evaluation criteria, where the ASD

is calculated as the square roots of the diagonal elements of the inverse Hessian matrix $\mathbf{M}^{-1}(\boldsymbol{\theta}_0)\mathbf{V}(\boldsymbol{\theta}_0)\mathbf{M}^{-1}(\boldsymbol{\theta}_0)$.

Scenario II: Misspecified TDGV and TDGMV models. In this case, we first generate samples deviates from the TDGMV (or TDGV) model, i.e., misspecified model, then derive $\hat{\boldsymbol{\theta}}$ via the MLE method discussed in Section 3.1, finally summarize the empirical bias, ESD and ASD.

Without loss of generality, we assume the misspecified models as truncated Skellam GARCH model with a time-varying variance and that with jointly time-varying mean and variance, where the former one is defined as $X_t|\mathcal{F}_{t-1} \sim \text{Skellam}_n(0, \sigma_t^2)$ with $\{\sigma_t^2\}$ satisfying (2.5), and the latter one is define as $X_t|\mathcal{F}_{t-1} \sim \text{Skellam}_n(\mu_t, \sigma_t^2)$ with $\{\mu_t, \sigma_t^2\}$ satisfying (2.4) and (2.5), where $\text{Skellam}_n(\mu_t, \sigma_t^2)$ denotes as n -truncated Skellam process with

$$\mathbb{P}_n(X_t = z) = \frac{\left(\frac{\sigma_t^2 + \mu_t}{\sigma_t^2 - \mu_t}\right)^{z/2} I_{|z|}(\sqrt{\sigma_t^4 - \mu_t^2})}{\sum_{i=-n}^n \left(\frac{\sigma_t^2 + \mu_t}{\sigma_t^2 - \mu_t}\right)^{i/2} I_{|i|}(\sqrt{\sigma_t^4 - \mu_t^2})}, \quad z \in \{0, \pm 1, \dots, \pm n\},$$

where $n \in \mathbb{Z}^+$ is a fixed integer specifying the truncation boundary and

$$I_r(x) = (x/2)^r \sum_{k=0}^{\infty} \frac{(x^2/4)^k}{k! \Gamma(r+k+1)}.$$

In the simulation, we let $n = 10$ with $T = 300$ and 1000 , and generate 1000 replications per scenario. ML estimator is carried out via the `constrOptim` function in the `R stats` package. The numerical results with parameter combinations and graphical diagnostics of scenarios I and II are relegated to Supplementary Material S3, where the numerical results are

summarized in Tables T1-T5 for scenario I and T6-T10 for scenario II, and the graphical diagnostics are presented in Figures F1-F4 for scenario I and F5-F8 for scenario II, respectively.

From Tables T1-T10, it is observed that the bias of the MLE tends to zero with increasing sample size under both scenarios, alongside continuous decreases in ESD and ASD. Such trends verify the convergence of estimators to true parameters and empirically validate the asymptotic properties derived in Section 3.

From Figures F1-F2 and F5-F6, the estimated medians in both scenarios approach the true parameter values with increasing sample size. Meanwhile, the interquartile and overall ranges of the estimates shrink as sample size grows, further verifying the consistency of the maximum likelihood estimators. Furthermore, Figures F3-F4 and F7-F8 illustrate the asymptotic normality of the estimators under both scenarios. Taken together, these simulation results indicate that the ML estimate performs well for both correctly specified and misspecified models.

Remark 8. Although scenario II exhibits slightly larger ESD and ASD than those of scenario I, the differences are moderate and diminish as T grows. Moreover, the estimated medians converge to the true values and the interquartile ranges narrow steadily under both scenarios. Therefore,

the TDGMV model retains asymptotic consistency and normality under distributional misspecification.

Remark 9. In scenario II, we study distributional misspecification by adjusting only the conditional distribution while keeping the dynamic structure unchanged. Our results thus primarily evaluate the robustness of the proposed method to conditional distribution misspecification. Other forms of misspecification, such as misspecified model's order, lag structure, or joint misspecification of both distribution and dynamic structure, are not explored in this work and remain promising avenues for future research. See Aleksandrov et al. (2024) and Görtler and Henze (2000) for relevant work.

Remark 10. To assess the forecasting performance, we evaluate the calibration of the one-step-ahead predictive distributions via the randomized probability integral transform (PIT). See Figures F9-F10 for the PIT histograms in Supplementary Material S3. And they are approximately uniform under both scenarios. Furthermore, the Kolmogorov-Smirnov (KS) test and Pearson residual diagnostics (see Table T11 in Supplementary Material S3) corroborate that the proposed models exhibit robustness to misspecification of the conditional distribution in terms of prediction.

5. Application

In this section, we apply the TDGMV model to two empirical datasets derived from the field of Air Quality Index (AQI), which are available at meteorological observation stations (<https://www.aqistudy.cn/historydata/>). In what follows, we follow Ministry of Environmental Protection (2012) and consider two distinct application cases:

Case I: The daily AQI is categorized into three categories: excellent ($AQI < 50$), good ($50 \leq AQI < 100$) and pollution ($AQI \geq 100$), which are denoted as 1, 0, and -1 , respectively. Thus, the scaled AQI data take values in $\{1, 0, -1\}$ with $n = 1$.

Case II: The daily air quality is divided into six ordered grades, i.e., excellent, good, lightly polluted, moderately polluted, heavily polluted and severely polluted, which are correspondingly labeled as 0, 1, 2, 3, 4, 5. Similar classification schemes can be found in Weiß and Jahn (2024) and Liu et al. (2022).

5.1 Daily AQI Data under Case I

In this part, we consider the daily AQI data under case I for Puyang in China from 1 January 2023 to 31 December 2024. The daily AQI data takes values in $\{1, 0, -1\}$ with $n = 1$, which categorized as excellent, good

5.1 Daily AQI Data under Case I

and pollution. By computation, the sample mean and variance are -0.1833 and 0.4129 , respectively. Figure 1 presents the sample path, the Lagrange multiplier test and the corresponding ACF and PACF plots.

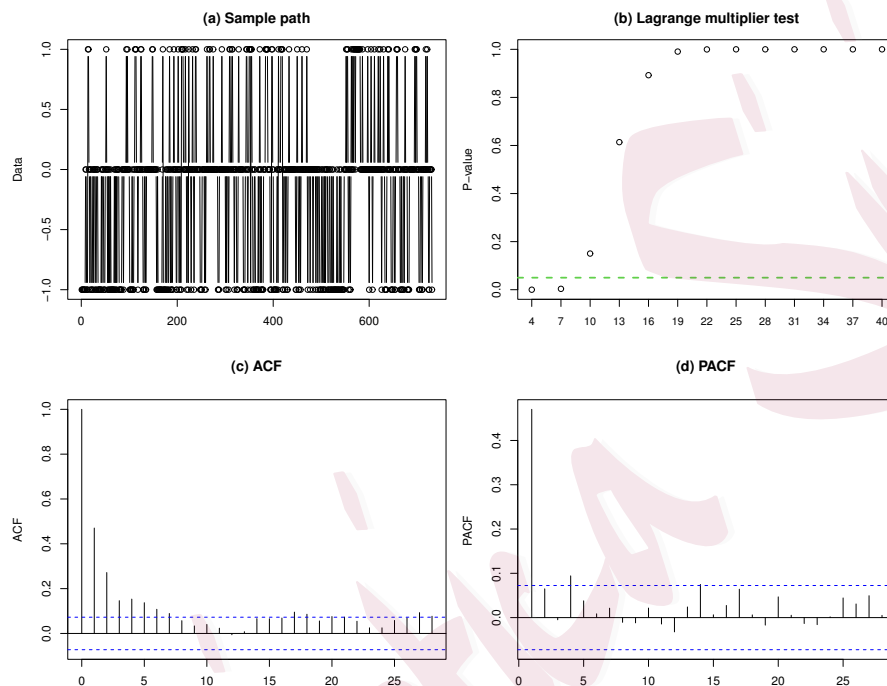


Figure 1: Path of the daily AQI data under case I.

From Figure 1, the considered AQI data appears stationary. The Lagrange multiplier test displayed in Figure 1(b) is significant at the 5% level for several low-order lags, indicating the presence of ARCH effects. This result is in line with the time-varying volatility. In addition, the ACF and PACF plots in Figures 1 (c) and (d) indicate that a low-order dynamic

5.1 Daily AQI Data under Case I

structure is sufficient.

To explore the evolutionary behavior of the daily AQI data, we use the TDGM(1,1), TDGV(1,1) and TDGMV((1,1),(1,1)) models to fit data by using MLE. To further illustrate the outperform of TDGMV model, we adopt a two-step strategy. To begin with, we shift the AQI data with +1 such that the shifted data take values in $\{0, 1, 2\}$. Thus, this creates applicable conditions for adopting existing nonnegative bounded integer-valued GARCH models. Second, we use the binomial GARCH(1,1) (Lee and Lee, 2019), beta-binomial GARCH(1,1) (Chen et al., 2022) and Conway-Maxwell-Poisson-Binomial (CMB)GARCH(1,1) (Chen et al., 2025b) models with linear, logit and soft-clipping transformations to fit the shifted data.

For convenience, we call the binomial GARCH(1,1) model with linear, logit and soft-clipping transformation as BGAR, logit-BGAR and Sc-BGAR, respectively. Similarly, BBGAR, logit-BBGAR and Sc-BBGAR correspond to beta-binomial GARCH(1,1) models with the three aforementioned transformations. Meanwhile, CMBGAR, logit-CMBGAR and Sc-CMBGAR are defined in a similar way. To assess their performance, we compare the estimated standard errors (SE), the negative log-likelihood ($-\log\text{-lik}$) and the Akaike information criterion (AIC). The fitted results, including $-\log\text{-lik}$ and AIC, are reported in Table 1.

5.1 Daily AQI Data under Case I

Table 1: Fitted results and SE in grey for daily AQI data under case I.

Model	estimates					−log-lik	AIC
BGAR(1,1)	$\hat{\omega}$	$\hat{\alpha}_1$	$\hat{\beta}$				
	0.1717	0.1425	0.4383				634.2923 1274.5847
	0.0324	0.0966	0.0428				
logit-BGAR(1,1)	$\hat{\omega}$	$\hat{\alpha}_1$	$\hat{\beta}$				
	−0.6018	0.5002	0.4999				645.1812 1296.3625
	0.1123	0.1076	0.0892				
Sc-BGAR(1,1)	$\hat{\omega}$	$\hat{\alpha}_1$	$\hat{\beta}$				
	−0.0885	0.2902	0.8101				634.5403 1275.0806
	0.0663	0.1858	0.0837				
BBGAR(1,1)	$\hat{\omega}$	$\hat{\alpha}_1$	$\hat{\beta}$	$\hat{\phi}$			
	0.3454	0.1418	0.4364	0.0001			
	0.0788	0.1194	0.0561	0.0009			
logit-BBGAR(1,1)	$\hat{\omega}$	$\hat{\alpha}_1$	$\hat{\beta}$	$\hat{\phi}$			
	−1.1161	0.0824	0.9175	0.0001			
	0.1203	0.1058	0.0980	0.0000			
Sc-BBGAR(1,1)	$\hat{\omega}$	$\hat{\alpha}_1$	$\hat{\beta}$	$\hat{\phi}$			
	−0.0885	0.2902	0.8101	0.0001			
	0.0663	0.1858	0.0837	0.0000			
CMBAR(1)	$\hat{r}_0$	$\hat{r}_1$	$\hat{\nu}$				
	0.1788	0.4999	1.8029				611.3038 1228.6076
	0.0208	0.0425	0.1299				
logit-CMBAR(1)	$\hat{r}_0$	$\hat{r}_1$	$\hat{\nu}$				
	−1.7279	2.8668	1.9483				608.3666 1222.7331
	0.1362	0.2463	0.1355				
Sc-CMBAR(1)	$\hat{r}_0$	$\hat{r}_1$	$\hat{\nu}$				
	−0.4501	1.1556	1.9417				609.6632 1225.326
	0.0609	0.0861	0.1363				
TDGV(1,1)	$\hat{\phi}_0$	$\hat{\phi}_1$	$\hat{\psi}_1$				
	0.3006	0.2450	0.3893				901.1297 1822.0385
	0.1451	0.0764	0.2100				
TDGM(1,1)	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\beta}_1$	$\hat{\nu}$			
	0.0031	−0.2862	0.9274	0.1311			
	0.0030	0.0251	0.0346	0.0071			
TDGMV((1,1),(1,1))	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\beta}_1$	$\hat{\phi}_0$	$\hat{\phi}_1$	$\hat{\psi}_1$	
	−0.0725	0.4307	0.0993	0.0104	0.0124	0.9487	603.2496 1218.4993
	0.0216	0.0361	0.0824	0.0066	0.0073	0.0237	

5.1 Daily AQI Data under Case I

Table 1 shows that the TDGMV((1,1),(1,1)) model takes the smallest $-\log\text{-lik}$ and AIC. To check its adequacy, we compute the mean and variance of its fitted standardized Pearson residuals are approximately -0.0331 and 1.0380 , respectively. These show that the TDGMV model effectively captures the inherent variability in the analyzed data.

Additionally, we present the ACF and PACF of the fitted standardized Pearson residuals in Figure 2, which reveals that the TDGMV((1,1),(1,1)) model successfully captures the correlated patterns presenting in the analyzed data, further affirming its suitability.

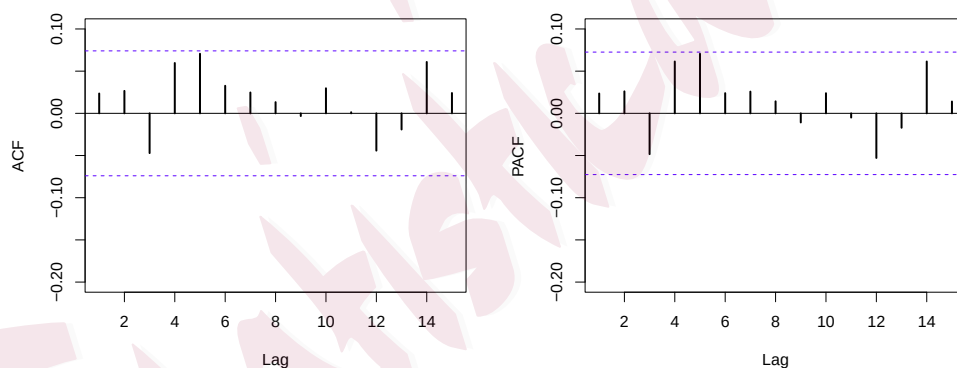


Figure 2: Pearson residual analysis: (a). ACF (b). PACF.

Finally, we give the one-step ahead forecasting probability and its lower and upper bounds of 95% confidence interval in Table 2.

5.2 Weekly AQI Data under Case II

Table 2: Results of the one-step ahead forecasting probability.

	-1	0	1
Lower bound	0.1812	0.7150	0.1003
True probabilty	0.1823	0.7162	0.1015
Upper bound	0.1835	0.7173	0.1026

From Table 2, the one-step ahead forecasting probability is unimodal and asymmetric shape. This is consistent with (3.13).

5.2 Weekly AQI Data under Case II

Herein, we aim to compare the weekly AQI between Xinxiang and Zhengzhou from December 31, 2018 to December 29, 2024 under case II, in which the daily AQI forms an ordinal categorical sequence (denoted as $\{Y_{it}\}, i = 1, 2$) on $\{0, 1, 2, 3, 4, 5\}$, where $\{Y_{1t}\}$ and $\{Y_{2t}\}$ denote the daily AQI series of Xinxiang and Zhengzhou.

In subsequent, we define a trinary time series $\{Y_t\}$, i.e., $\forall t, Y_t = 1$ if $Y_{1t} > Y_{2t}$, $Y_t = 0$ if $Y_{1t} = Y_{2t}$, and $Y_t = -1$ if $Y_{1t} < Y_{2t}$. We then compute the weekly number of days (denoted as $\{X_t\}$ with $X_t = \sum_{j=t}^{t+6} Y_j, t = 1, \dots, 313$), which records the number of weekly days where the air quality in Xinxiang outperforms that in Zhengzhou. Thus, $\{X_t\}$ takes value in $\{-7, \dots, 0, \dots, 7\}$. By calculation, the sample mean and variance of $\{X_t\}$ are -0.1917 and 2.8606 , respectively. To further examine the air quality of

5.2 Weekly AQI Data under Case II

the two cities, we give some plots of $\{X_t\}$ in Figure 3, which shows that the series seems stationary, exhibits volatility clustering and contains significant low-order ARCH effects. The ACF and PACF plots further suggest that $p = 1$ is appropriate.

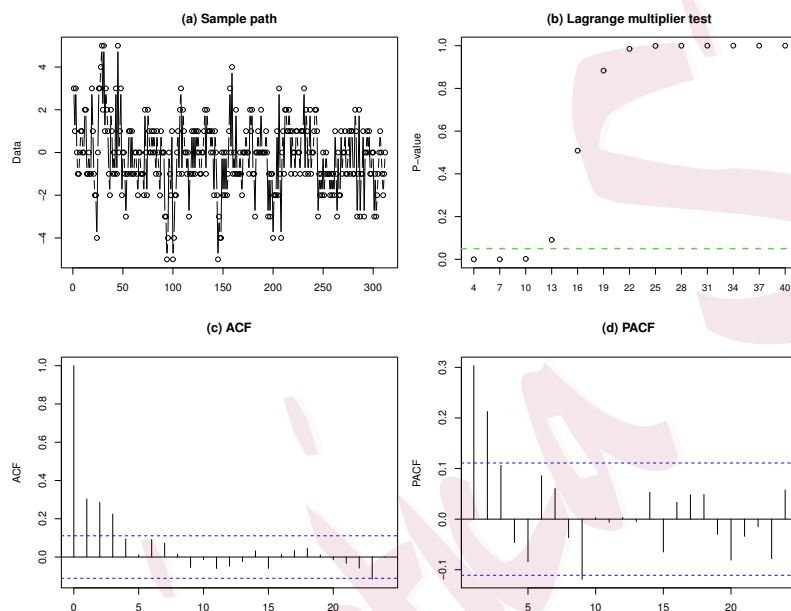


Figure 3: Some plots of the weekly AQI data.

In what follows, we use TDGM(1,1), TDGV(1,1) and TDGMV((1,1),(1,1)) to fit $\{X_t\}$, and then shift $\{X_t\}$ with +7 such that the shifted data takes a nonnegative finite range. Then we use BGAR(1,1) model, logit-BGAR(1,1) model and Sc-BGAR(1,1) model, BBGAR(1,1) model, logit-BBGAR(1,1) model and Sc-BBGAR(1,1) model, CMBGAR(1,1) model, logit-CMBGAR(1,1)

5.2 Weekly AQI Data under Case II

model and Sc-CMBGAR(1,1) model with $c = 0.01$ to fit the shifted data.

The fitted results, including $-\log\text{-lik}$ and AIC, are given in Table 3.

Table 3: Fitted results and SE in grey of the weekly AQI data.

Model	estimates					$-\log\text{-lik}$	AIC
BGAR(1,1)	$\hat{\omega}$	$\hat{\alpha}_1$	$\hat{\beta}$				
	0.1188	0.4891	0.2655			591.5366	1189.0733
	0.0436	0.1147	0.0571				
logit-BGAR(1,1)	$\hat{\omega}$	$\hat{\alpha}_1$	$\hat{\beta}$				
	-0.5440	0.4937	0.0755			591.4792	1188.9583
	0.1185	0.1151	0.0166				
Sc-BGAR(1,1)	$\hat{\omega}$	$\hat{\alpha}_1$	$\hat{\beta}$				
	-0.2608	0.5484	0.4947			594.5059	1195.0119
	0.0920	0.1984	0.0874				
BBGAR(1,1)	$\hat{\omega}$	$\hat{\alpha}_1$	$\hat{\beta}$	$\hat{\phi}$			
	1.6775	0.4882	0.2642	0.0001		591.6009	1191.202
	0.6172	0.1155	0.0570	0.0017			
logit-BBGAR(1,1)	$\hat{\omega}$	$\hat{\alpha}_1$	$\hat{\beta}$	$\hat{\phi}$			
	-0.5310	0.5039	0.0736	0.0001		591.0852	1188.1704
	0.1186	0.1121	0.0167	0.0000			
Sc-BBGAR(1,1)	$\hat{\omega}$	$\hat{\alpha}_1$	$\hat{\beta}$	$\hat{\phi}$			
	-0.2608	0.5484	0.4947	0.0001		594.5669	1195.1337
	0.0921	0.1985	0.0875	0.0000			
CMBAR(1,1)	$\hat{r}_0$	$\hat{r}_1$	$\hat{\nu}$				
	0.2741	0.4242	1.3841			589.8256	1185.651
	0.0398	0.0786	0.1142				
logit-CMBAR(1,1)	$\hat{r}_0$	$\hat{r}_1$	$\hat{\nu}$				
	-0.9035	1.6966	1.3803			589.9469	1185.8938
	0.1635	0.3226	0.1137				
Sc-CMBAR(1,1)	$\hat{r}_0$	$\hat{r}_1$	$\hat{\nu}$				
	-0.0872	0.6668	1.3763			590.1822	1186.3643
	0.0651	0.1252	0.1132				
TDGV(1,1)	$\hat{\phi}_0$	$\hat{\phi}_1$	$\hat{\psi}_1$				
	0.7067	0.5055	0.4945			593.1798	1192.3597
	0.5597	0.1863	0.2373				
TDGM(1,1)	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\beta}_1$	$\hat{\nu}$			
	-0.0565	0.2578	0.4966	0.3387		579.1394	1166.2788
	0.0662	0.0542	0.1256	0.0277			
TDGMV((1,1),(1,1))	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\beta}_1$	$\hat{\phi}_0$	$\hat{\phi}_1$	$\hat{\psi}_1$	
	-0.0624	0.2440	0.5001	0.3195	0.3005	0.7082	574.6348
	0.0465	0.0522	0.1092	0.2537	0.1369	0.1427	1161.2697

5.2 Weekly AQI Data under Case II

From Table 3, the TDGMV((1,1),(1,1)) model takes the smallest $-\log$ -lik and AIC. Then we calculate its fitted standardized Pearson residuals, whose mean and variance are 0.0028 and 1.0065, respectively. This confirms that the TDGMV((1,1),(1,1)) model adequately captures the underlying data volatility. Furthermore, we present the ACF and PACF in Figure 4 of the fitted standardized Pearson residuals, which reveals that there exists no significant residual correlation. It demonstrates that the TDGMV((1,1),(1,1)) model fully captures the dependence structure within the dataset.

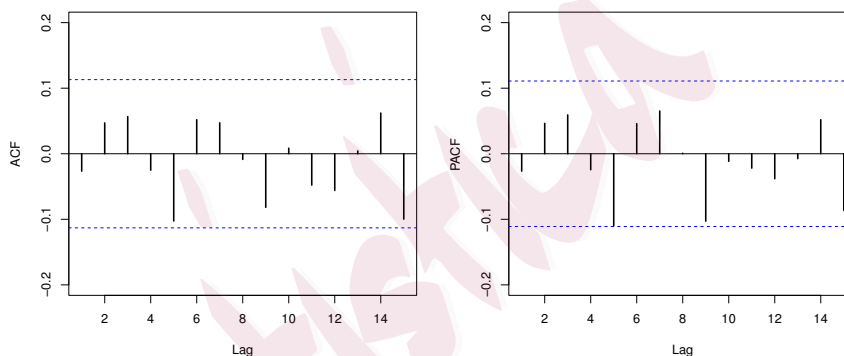


Figure 4: Pearson residual analysis: (a). ACF (b). PACF.

Finally, we present the one-step-ahead forecasting distribution of weekly AQI data in Figure 5, along with its 95% confidence intervals, which confirms that the true distribution lies entirely within the 95% confidence

5.2 Weekly AQI Data under Case II

bounds. This consistency with the conclusions from Section 3.2 strengthens the overall validation of the TDGMV model.

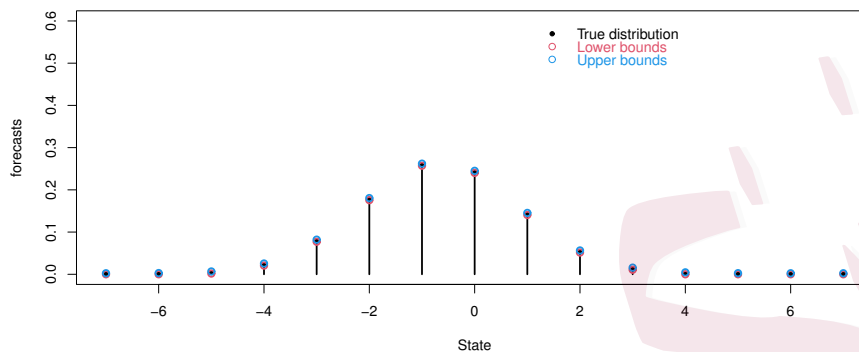


Figure 5: Results of the one-step ahead forecasting probability.

Remark 11. As discussed in Weiß and Jahn (2024), “But except for these extreme cases, the ScBINARCH(1) model with low c (such as $c = 0.01$) does very well in imitating linearity.” Therefore, we take $c = 0.01$ for the compared model with a soft-clipping transformation in real application.

Remark 12. Analogous to the binomial GARCH-type models, the TDGMV model can also be specified with soft-clipping or logit transformations, which may yield improved performance relative to the linear specification. However, closed-form conditional moments are unavailable for such nonlinear formulations, and establishing stability properties becomes substantially more difficult than for the linear TDGMV model. We therefore restrict our

attention to the linear-transformation TDGMV model for a preliminary investigation of its double-dynamic structure, and leave more comprehensive exploration for future research.

6. Conclusion and Discussion

To address the modeling of bounded full-integer-valued time series, this paper introduced a novel double dynamic autoregressive process, along its special cases. We established the stationarity and ergodicity of the model, investigated its statistical properties, discussed the MLE, including the asymptotic property of the ML estimator. Additionally, we conducted some simulation studies to validate the finite-sample performance and applied the novel model on two datasets to highlight its outperformance.

There also exist some important directions call for prompt investigation. First, the AQI data consistently exhibit a significant prevalence of values at the lower end of the scale, particularly zero. Consequently, an inflated trinomial difference autoregressive process warrants serious consideration for modeling such data, see Kang et al. (2023) and Liu et al. (2022). Second, it is an important issue to test the nullity of α_1 and ϕ_1 for the TDGMV((1,1),(1,1)) model. To resolve this question, one can follow the approach in Jiang et al. (2020) and Liu et al. (2022), and then construct

a Lagrange multiplier test. Third, it is worth noting that the likelihood function given in (3.12) involves a convolution and two time-varying latent processes, both of which impair the efficiency of the ML estimator. This issue is fairly prevalent, see Fokianos et al. (2009). Pedeli et al. (2015) and Zhang et al. (2019) have confirmed that the saddlepoint approximation provides near-exact likelihood inference. Hence, it is worthwhile to consider the MLE of the TDGMV model via saddlepoint approximation.

Supplementary Materials

The online supplementary materials provide detailed results for the first and second derivatives of $l_t(\boldsymbol{\theta})$ with respect to $\boldsymbol{\theta}$ of Section 3.1 (S1), the proofs of the Propositions, Lemma 1, Remark 2 and Theorems (S2), detailed results for the simulation study of Section 4 (S3) and some auxiliary results (S4).

Acknowledgements

The authors thank the Editor-in-Chief and the two anonymous referees for the valuable comments and suggestions that result in a substantial improvement of this paper. This work is supported by National Nature Science Foundation of China (Nos. 12671320, 12471252, 12271206), Natural Science Foundation of Jilin Province (No. 20250102003JC) and Science and

REFERENCES

Technology Research Program of Henan Province (No. 262102320106).

References

- Aleksandrov, B., Weiß, C.H., Nik, S., Faymonville, M. and Jentsch, C. (2024). Modelling and diagnostic tests for Poisson and negative-binomial count time series. *Metrika* **87**, 843–887.
- Barreto-Souza, W., Piancastelli, L. S., Fokianos, K. and Ombao, H. (2026). Time-varying dispersion integer-valued GARCH models. *Journal of Time Series Analysis* **47**, 839–853.
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* **31**, 307–327.
- Chen, H. (2025). A novel bounded $\mathbb{Z}$ -valued autoregressive model with its application on crime data. *Journal of Statistical Computation and Simulation* **95**, 408–425.
- Chen, H., Han, Z. and Zhu F. (2025a). A trinomial difference autoregressive process for the bounded $\mathbb{Z}$ -valued time series. *Journal of Time Series Analysis* **46**, 152–180.
- Chen, H., Li, Q. and Zhu F. (2020). Two classes of dynamic binomial integer-valued ARCH models. *Brazilian Journal of Probability and Statistics* **34**, 685–711.
- Chen, H., Li, Q. and Zhu F. (2022). A new class of integer-valued GARCH models for time series of bounded counts with extra-binomial variation. *AStA Advances in Statistical Analysis* **106**, 243–270.
- Chen, H., Li, Q. and Zhu, F. (2023a). A covariate-driven beta-binomial integer-valued GARCH model for bounded counts with an application. *Metrika* **86**, 805–826.

REFERENCES

- Chen, H., Wang, H. and Zhang, Y. (2025b). Modelling bounded integer-valued time series of counts with a novel class of Conway-Maxwell-Poisson-Binomial ARCH models. *Statistics* **59**, 819-844.
- Chen, H., Zhang, J. and Zhu, F. (2023b). A trinomial difference autoregressive model and its applications. *Stat* **12**, e596.
- Davis, R.A., Fokianos, K., Holan, S.H., Joe, H., Livsey, J., Lund, R., Pipiras, V. and Ravishanker, N. (2021). Count time series: a methodological review. *Journal of the American Statistical Association* **116**, 1533–1547.
- Davis, R. A. and Liu, H. (2016). Theory and inference for a class of observation-driven models with application to time series of counts. *Statistica Sinica* **26**, 1673–1707.
- Doukhan, P., Naushad, M. and Michael, H. (2021). Mixing properties of integer-valued GARCH processes. *ALEA* **18**, 401–420.
- Doukhan, P. and Wintenberger, O. (2008). Weakly dependent chains with infinite memory, *Stochastic Processes and their Applications* **118**, 1997–2013.
- Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* **50**, 987–1008.
- Ferland, R., Latour, A. and Oraichi, D. (2006). Integer-valued GARCH process. *Journal of Time Series Analysis* **27**, 923–942.
- Fokianos, K., Rahbek, A. and Tjøstheim, D. (2009). Poisson autoregression. *Journal of the*

REFERENCES

- American Statistical Association* **104**, 1430–1439.
- Freeland, R. K. and McCabe, B. P. (2004). Forecasting discrete valued low count time series. *International Journal of Forecasting* **20**, 427–434.
- Gürtler, N. and Henze, N. (2000). Recent and classical goodness-of-fit tests for the Poisson distribution. *Journal of Statistical Planning and Inference* **90**, 207–225.
- Haq, A. (2025). Design of the EWMA charts for simultaneously monitoring mean and variance using the conditional false alarm rate. *Quality Engineering* **37**, 552–559.
- Jiang, F., Li, D. and Zhu, K. (2020). Non-standard inference for augmented double autoregressive models with null volatility coefficients. *Journal of Econometrics* **215**, 165–183.
- Kang, Y., Wang, S., Wang, D. and Zhu, F. (2023). Analysis of zero-and-one inflated bounded count time series with applications to climate and crime data. *TEST* **32**, 34–73.
- Kim, H. Y., Weiß, C. H. and Möller, T. A. (2020). Models for autoregressive processes of bounded counts: How different are they? *Computational Statistics* **35**, 1715–1736.
- Lee, Y. and Lee, S. (2019). CUSUM test for general nonlinear integer-valued GARCH models: comparison study. *Annals of the Institute of Statistical Mathematics* **71**, 1033–1057.
- Liu, H. (2012). *Some models for time series of counts*. Dissertations, Columbia University. <https://core.ac.uk/download/pdf/27293903.pdf>.
- Liu, M., Zhu, F. and Zhu, K. (2022). Modeling normalcy-dominant ordinal time series: An application to air quality level. *Journal of Time Series Analysis* **43**, 460–478.

REFERENCES

- Ministry of Environmental Protection. (2012). Technical regulation on ambient air quality index (on trial) (HJ 633-2012). *China Environmental Science Press*, Beijing.
- Möller, T. A. and Weiß, C.H. (2020). Generalized discrete autoregressive moving-average models. *Applied Stochastic Models in Business and Industry* **36**, 641–659.
- Omar, A., Alzaid, A. and Odhah, O. (2016). A trinomial difference distribution. *Revista Colombiana de Estadística* **39**, 1–15.
- Pedeli, X., Davison, A. C. and Fokianos, K. (2015). Likelihood estimation for the INAR (p) model by saddlepoint approximation. *Journal of the American Statistical Association* **110**, 1229–1238.
- Ristić, M. M., Weiß, C. H. and Janjić, A. D. (2016). A binomial integer-valued ARCH model. *International Journal of Biostatistics* **12**, 20150051.
- Weiß, C. H. (2018). *An introduction to discrete-valued time series*. Chichester: John Wiley & Sons.
- Weiß, C. H. (2021). Stationary count time series models. *WIREs Computational Statistics* **13**, e1502,
- Weiß, C. H. and Jahn, M. (2024). Soft-clipping INGARCH models for time series of bounded Counts. *Statistical Modelling* **24**, 295–319.
- Weiß, C. H. and Pollett, P. K. (2014). Binomial autoregressive processes with density-dependent thinning. *Journal of Time Series Analysis* **35**, 115–132.

REFERENCES

- Wendler, A., Beruvides, M. G. and Tercero-Gómez, V. G. (2025). Cautious parameter learning for a joint mean and variance monitoring CUSUM scheme with guaranteed in-control performance. *Quality Engineering* **37**, 523–540.
- Zhang, W., Bravington, M. V. and Fewster, R. M. (2019). Fast likelihood-based inference for latent count models using the saddlepoint approximation. *Biometrics* **75**, 723–733.

Huaping Chen

School of Mathematics and Statistics, Henan University, Kaifeng 475004, China

E-mail: chenhp0107@163.com

Qianqian Zhao

School of Mathematics and Statistics, Henan University, Kaifeng 475004, China

E-mail: zjq11282025@163.com

Liugen Xue

School of Mathematics and Statistics, Henan University, Kaifeng 475004, China

E-mail: lgxue@bjut.edu.cn

Fukang Zhu

School of Mathematics, Jilin University, Changchun 130032, China

E-mail: zfk8010@163.com (Corresponding author)

ORCID ID: 0000-0002-8808-8179