Article:

Aknouche, A, Almohaimeed, BS and Dimitrakopoulos, S orcid.org/0000-0002-0043-180X (2021) Forecasting transaction counts with integer-valued GARCH models. Studies in Nonlinear Dynamics and Econometrics. ISSN 1081-1826

https://doi.org/10.1515/snde-2020-0095

Abdelhakim Aknouche, Bader S. Almohaimeed and Stefanos Dimitrakopoulos*

Forecasting transaction counts with integer-valued GARCH models

https://doi.org/10.1515/snde-2020-0095
Received August 13, 2020; accepted August 30, 2021; published online September 13, 2021

Abstract: Using numerous transaction data on the number of stock trades, we conduct a forecasting exercise with INGARCH models, governed by various conditional distributions; the Poisson, the linear and quadratic negative binomial, the double Poisson and the generalized Poisson. The model parameters are estimated with efficient Markov Chain Monte Carlo methods, while forecast evaluation is done by calculating point and density forecasts.

Keywords: count time series; forecasting; INGARCH models; MCMC.

1 Introduction

In recent years, there has been a surge of interest in integer-valued generalized autoregressive conditional heteroscedastic (INGARCH) models (Ahmad and Francq 2016; Aknouche, Bendjeddou, and Touche 2018; Chen et al. 2016; Christou and Fokianos 2014; Davis and Liu 2016; Doukhan, Fokianos, and Tjøstheim 2012; Ferland, Latour, and Oraichi 2006; Fokianos, Rahbek, and Tjøstheim 2009). Such processes are designed to model integer-valued series that are characterized mainly by small values and overdispersion that cannot be adequately accounted for by standard real-valued ARMA models; see also Cameron and Trivedi (2013).

In its original formulation (Heinen 2003; Rydberg and Shephard 2000), the INGARCH process had a Poisson conditional distribution with a time-varying intensity that was a linear function of its q lagged values and its p recent observations. Later, many generalizations of the Poisson INGARCH (P-INGARCH) were put forward that differed in their conditional distributions (Poisson, negative binomial, double Poisson, etc.,) and/or their specifications for the conditional mean equation (linear, exponential, threshold); see, among others, Zhu (2011, 2012a, 2012b) as well as Weiss (2018).

Quite a few studies in the literature on count time series have conducted forecasting comparisons of various models for count processes (e.g., Homburg et al. 2019, 2020 and references therein). These models include, among others, the Poisson-INAR, the Poisson-INARCH, the negative binomial INAR, the ZIP-INAR, and the binomial AR and ARCH.

However, none of these studies have focused exclusively on the INGARCH models in terms of their forecasting performance. As such, despite the large numbers of the INGARCH models that have been put

***Corresponding author**: **Stefanos Dimitrakopoulos**, Economics Division, Leeds University Business School, University of Leeds, LS2 9JT, Leeds, UK, E-mail: s.dimitrakopoulos@leeds.ac.uk
Abdelhakim Aknouche, Department of Mathematics, College of Science, Qassim University, P.O. Box 707, Buraydah 51431, Saudi Arabia; and Faculty of Mathematics, University of Science and Technology Houari Boumediene, Bab Ezzouar, Algeria, E-mail: A.Aknouche@qu.edu.sa
Bader S. Almohaimeed, Department of Mathematics, College of Science, Qassim University, P.O. Box 707, Buraydah 51431, Saudi Arabia, E-mail: bsmhiemied@qu.edu.sa

forward, the relevant literature lacks a coherent forecast comparison exercise. This paper aspires to fill this gap.

Using numerous empirical time series on the trade intensity of stocks, we evaluate the out-of-sample forecasting performance of several INGARCH models. This is our main contribution that differentiates our work from previous ones. Our set of competing INGARCH models includes those with the most popular conditional distributions, namely the Poisson, the generalized Poisson, the (linear and quadratic) negative binomial and the double Poisson.

We estimate the model parameters by efficient Bayesian methods, in particular Markov Chain Monte Carlo (MCMC). The dispersion parameters are updated using an efficient universal self-tuned sampler within Gibbs sampler, proposed by Martino et al. (2015), whilst for the GARCH parameters, the adaptive Metropolis adjusted Langevin (MALA) algorithm of Atchadé (2006) is exploited. The forecasting performance of the models is evaluated by calculating point and density forecasts.

It is also worth mentioning that Bayesian forecasting of count series has been performed, among others, by McCabe and Martin (2005), McCabe, Martin, and Harris (2011) and Berry and West (2020). McCabe and Martin (2005) considered estimates of h-step ahead predictive mass functions for INAR models, McCabe, Martin, and Harris (2011) estimated non-parametrically the forecasting distributions for the same class of models and Berry and West (2020) utilized various forecast metrics, such as, the scaled mean squared error (sMSE). For a frequentist approach, the interested reader is referred to Homburg et al. (2019, 2020).

The structure of the paper is as follows. Section 2 describes the models in question and Sections 3 and 4 describe the posterior analysis. Several Monte Carlo experiments are conducted in Section 5 and in Section 6 we present the empirical results. Section 7 concludes. An Online Appendix accompanies this paper.

2 INGARCH specifications

A stochastic process $\{Y_t, t \in \mathbb{Z}\}$ is said to be an INGARCH(p, q) if its conditional distribution is given by

$$Y_t|\mathcal{F}_{t-1} \sim f_{\lambda_t} \tag{1}$$

and

$$\lambda_t = \omega + \sum_{i=1}^{q} \alpha_i Y_{t-i} + \sum_{j=1}^{p} \beta_j \lambda_{t-j}, \tag{2}$$

where $\omega > 0, \alpha_i \geq 0$ and $\beta_j \geq 0$, $\mathcal{F}_t$ is the σ-Algebra generated by $\{Y_{t-k}, k \geq 0\}$ and $f_{\lambda_t}(y_t) := f_{Y_t}(y_t|\mathcal{F}_{t-1})$ is a discrete distribution with conditional mean λ_t.

In this paper we consider the following distributions for $Y_t|\mathcal{F}_{t-1}$:

- The Poisson (P-INGARCH) (Ferland, Latour, and Oraichi 2006; Heinen 2003); $Y_t|\mathcal{F}_{t-1} \sim \mathcal{P}(\lambda_t)$.
- The double Poisson (DP-INGARCH) (Heinen 2003); $Y_t|\mathcal{F}_{t-1} \sim \mathcal{DP}(\lambda_t, \gamma)$, with $\gamma > 0$.
- The Negative binomial II (NB2-INGARCH) (Christou and Fokianos 2014; Davis and Liu 2016; Zhu 2011); $Y_t|\mathcal{F}_{t-1} \sim \mathcal{NB}\left(r_2, \frac{r_2}{r_2+\lambda_t}\right)$, with $r_2 > 0$.
- The Negative binomial I (NB1-INGARCH) (Aknouche and Francq 2021; Xu et al. 2012); $Y_t|\mathcal{F}_{t-1} \sim \mathcal{NB}\left(r_1\lambda_t, \frac{r_1}{r_1+1}\right)$, with $r_1 > 0$.
- The Generalized Poisson (GP-INGARCH) (Zhu 2012a); $Y_t|\mathcal{F}_{t-1} \sim \mathcal{GP}(\lambda_t(1-\tau), \tau)$, with $\tau > 0$.

The functional forms of these distributions along with their conditional means and conditional variances are given in Table 1. The parameters γ, υ, τ, r_1 and r_2 are usually called the dispersion parameters. As can be seen from Table 1, the conditional variance is linear in the intensity parameter λ_t for the Poisson, GP, and NB1

Table 1: Various conditional distributions for the INGARCH model.

	Notation	$f_{Y_t}(y_t \vert \mathcal{F}_{t-1})$	$E(y_t \vert \mathcal{F}_{t-1})$	$\mathrm{Var}(y_t \vert \mathcal{F}_{t-1})$
P	$\mathcal{P}(\lambda_t)$	$\mathrm{e}^{-\lambda_t}\frac{\lambda_t^{y_t}}{y_t!}$	λ_t	λ_t
GP	$\mathcal{GP}(\lambda_t(1-\tau),\tau)$	$\frac{\lambda_t(\lambda_t+\tau y_t)^{y_t-1}\mathrm{e}^{-(\lambda_t+\tau y_t)}}{y_t!}$	λ_t	$\frac{1}{(1-\tau)^2}\lambda_t$
DP	$\mathcal{DP}(\lambda_t,\gamma)$	$\gamma^{1/2}\mathrm{e}^{-\gamma\lambda_t}\frac{\mathrm{e}^{-y_t}y_t^{y_t}}{y_t!}\left(\frac{\mathrm{e}\lambda_t}{y_t}\right)^{\gamma y_t}$	λ_t	$\simeq\frac{1}{\gamma}\lambda_t$
NB1	$\mathcal{NB}\left(r_1\lambda_t,\frac{r_1}{r_1+1}\right)$	$\frac{\Gamma(y_t+r_1\lambda_t)}{y_t!\Gamma(r_1\lambda_t)}\left(\frac{r_1}{r_1+1}\right)^{r_1\lambda_t}\left(\frac{1}{r_1+1}\right)^{y_t}$	λ_t	$\left(1+\frac{1}{r_1}\right)\lambda_t$
NB2	$\mathcal{NB}\left(r_2,\frac{r_2}{r_2+\lambda_t}\right)$	$\frac{\Gamma(y_t+r_2)}{y_t!\Gamma(r_2)}\left(\frac{r_2}{r_2+\lambda_t}\right)^{r_2}\left(\frac{\lambda_t}{r_2+\lambda_t}\right)^{y_t}$	λ_t	$\lambda_t+\frac{1}{r_2}\lambda_t^2$

cases, is approximately linear for the DP case and quadratic for the NB2 case. Under

$$\sum_{i=1}^{q}\alpha_i+\sum_{j=1}^{p}\beta_j<1, \tag{3}$$

the models under consideration are stationary and ergodic with finite mean (Aknouche and Francq 2021).

3 MCMC

We want to sample iteratively from the full conditional posteriors $\pi(\Delta|\text{disp},y)$ and $\pi(\text{disp}|\Delta,y)$, where $\Delta=(\omega,\alpha_1,\beta_1)'$ and disp represents the dispersion parameter, depending on the model. For Δ we used a truncated log-normal prior

$$\log(\Delta)\sim N(\mu_\Delta,\Sigma_\Delta)1_{(\alpha_1+\beta_1<1)},$$

that satisfies the stationarity condition that $\alpha_1+\beta_1<1$, whereas for the dispersion parameter we use a gamma prior

$$G(k_{\text{disp}},m_{\text{disp}}).$$

Both conditionals are intractable and therefore we use Metropolis-Hastings type algorithms. For the update of the dispersion parameter we use the Fast Universal Self-Tuned Sampler (FUSS) of Martino et al. (2015).[1] It can be used to sample efficiently from univariate distributions. It consists of four steps. In the first step, an initial set of support points of the target distribution is chosen. In the second step, unused support points drop according to some pre-defined pruning criterion (for example, optimal minimax pruning strategy). In the third step, we have the construction of the independent proposal density, tailored to the shape of the target, with some appropriate pre-defined mechanism (for example interpolation). In the final step, a Metropolis–Hastings (MH) method is used.

For the update of Δ we use the adaptive MALA of Atchadé (2006) with a truncated drift. Defining the drift as

$$\forall\ X_n, D(X)=\frac{\delta}{\max(\delta,\|\nabla\ \log\ \pi(X_n)\|)}\nabla\ \log\ \pi(X_n)$$

the general algorithm is described below, where $\pi()$ represents the posterior density and q_{σ_n,Λ_n} is the proposal, which in our case is the normal satisfying the stationarity condition. $\bar{\tau}$ is set, practically to 0.5 to achieve an acceptance rate of 50% and for numerical stability we set $\varepsilon=10^{-6}$ and $\delta=1000$. The sequence $(\gamma)_n\in N^*$ is chosen such that $\forall n,\gamma_n>0$, $\sum_n\gamma_n=+\infty$ and $\gamma_n=O(n^{-\xi})$ with $1/2<\xi\le 1$.

1 The FUSS algorithm has better mixing properties than alternative MCMC methods such as slice sampling, MALA sampling, and Hamiltonian Monte Carlo sampling and is faster. The FUSS matlab function is accessible from Martino's webpage.

4 Point and density forecasts

We conduct a recursive out-of-sample forecasting exercise in order to evaluate the predictive performance of the competing models. To this end, we compute point and density forecasts.

The conditional predictive density of the s-step ahead y_{t+s}, given the data $Y_t = (y_1, \dots, y_t)$ is given by

$$p(y_{t+s}|Y_t) = \int f(y_{t+s}|\Theta, Y_t)\mathrm{d}p(\Theta|Y_t), \tag{4}$$

where Θ denotes the model parameters.

Using Monte Carlo integration, the above expression can be approximated by

$$\hat{p}(y_{t+s}|Y_t) = \frac{1}{R}\sum_{i=1}^{R} f(y_{t+s}|\Theta^{(i)}, Y_t), \tag{5}$$

where $\Theta^{(i)}$ is the posterior draw of Θ at iteration $i = 1, \dots, R$ (after the burn-in period).

Algorithm: Adaptive MALA

Start with an initial point $X_0 \in \mathcal{U}$, a vector $\mu_0 \in \mathcal{U}$, a positive definite matrix Γ_0, $\varepsilon > 0$, a sequence of positive step sizes $(\gamma_n)_{n\geq 1}$, $\overline{\tau} \in]0,1[$ and $\sigma_0 > 0$.

Given the current value X_n and $(\mu_n, \Gamma_n, \sigma_n)$, let : $\Lambda_n = \Gamma_n + \varepsilon \mathrm{I}_d$.

Generate $Y_{n+1} \sim \mathcal{N}\left(X_n + \frac{\sigma_n^2}{2}\Lambda_n D(X_n), \sigma_n^2\Lambda_n\right)$ and generate $U \sim \mathcal{U}([0,1])$.

Define $\alpha(X_n, Y_{n+1}) = \min\left(1, \frac{\pi(Y_{n+1})q_{\sigma_n,\Lambda_n}(X_n \mid Y_{n+1})}{\pi(X_n)q_{\sigma_n,\Lambda_n}(Y_{n+1} \mid X_n)}\right)$. If $U \leq \alpha(X_n, Y_{n+1})$, set $X_{n+1} = Y_{n+1}$. Otherwise, set $X_{n+1} = X_n$.

Set :

$$\begin{aligned}
\mu_{n+1} &= \mu_n + \gamma_n\left(X_{n+1} - \mu_n\right)\\
\Gamma_{n+1} &= \Gamma_n + \gamma_n\left((X_{n+1} - \mu_n)(X_{n+1} - \mu_n)^\top - \Gamma_n\right)\\
\sigma_{n+1} &= \sigma_n + \gamma_n\left(\alpha(X_n, Y_{n+1}) - \overline{\tau}\right).
\end{aligned}$$

The conditional predictive likelihood of y_{t+s} is the conditional predictive density of y_{t+s} evaluated at the observed y^o_{t+s}, namely, $p(y_{t+s} = y^o_{t+s}|Y_t)$. A usual metric for the evaluation of the density forecasts is the log predictive score (LPS) (Geweke and Amisano 2011)

$$\mathrm{LPS} = \sum_{t=t_0}^{T-s} \log\, p(y_{t+s} = y^o_{t+s}|Y_t), \tag{6}$$

where $t = t_0 + 1, \dots, T - s$ is the evaluation period. The higher the LPS value, the better the (out-of-sample) forecasting power of the model.

We also calculated s-step ahead point forecasts. As a metric for the evaluation of point forecasts we used the scaled mean squared error (sMSE); see Berry and West (2020). For a particular observation, y^o_{t+s}, the scale squared error is defined as

$$\mathrm{sSE}_t(s) = (y^o_{t+s} - E(y_{t+s}))^2/\bar{y}_{1:t} \tag{7}$$

where $E(y_{t+s})$ is the s-step ahead predictive mean. The average of this metric over all days t is the scaled mean squared error for the forecast horizon s.

The lower the sMSE value, the better the (out-of-sample) forecasting power of the model. In our analysis, $s = 1, 4$ and 8.

5 Monte Carlo experiments

To assess the performance of the proposed Bayesian methodology we simulated various INGARCH series. Throughout our simulations, we generated $T = 100$, $T = 500$ and $T = 1000$ data points from all models with various sets of real values of the parameters. We run the samplers for 10,000 iterations after discarding the initial 10,000 cycles (burn-in period).

For the INGARCH parameter $\Delta = (\omega, \alpha_1, \beta_1)'$ we used a truncated log-normal prior

$$\log(\Delta) \sim N(\mu_\Delta, \Sigma_\Delta)1_{(\alpha_1+\beta_1<1)},$$

where $\mu_\Delta = (1, \log(0.1), \log(0.8))'$ and $\Sigma_\Delta = \mathrm{diag}(10, 1, 1)$. For the dispersion parameters, we used the following gamma prior

$$G(5, 0.1).$$

To monitor the performance of our sampling algorithms, we estimated the inefficiency factor (IF); see Chib (2001). To monitor any lack of convergence, we also computed the Convergence Diagnostics (CD) statistic of Geweke (1992).

From the Tables (Tables 2–6) below we see that the estimated values are close to their true values and the mixing of the algorithms are satisfactory. No convergence problems were detected.

Table 2: Simulated data for the P-INGARCH.

True values	Mean	Stdev	IF	CD
		T = 100		
$\omega = 1$	0.805	0.309	14.902	−0.468
$\alpha_1 = 0.7$	0.706	0.076	16.937	−1.528
$\beta_1 = 0.2$	0.192	0.076	18.641	1.437
$\omega = 2$	2.308	0.501	56.115	−1.759
$\alpha_1 = 0.3$	0.155	0.091	96.098	−1.748
$\beta_1 = 0.6$	0.697	0.099	85.752	2.030
		T = 500		
$\omega = 1$	1.018	0.193	12.865	−0.708
$\alpha_1 = 0.7$	0.641	0.041	15.541	−1.106
$\beta_1 = 0.2$	0.248	0.046	16.795	1.026
$\omega = 2$	2.108	0.418	22.098	1.298
$\alpha_1 = 0.3$	0.286	0.035	21.815	1.650
$\beta_1 = 0.6$	0.612	0.044	24.141	−1.804
		T = 1000		
$\omega = 1$	1.145	0.157	13.694	−1.389
$\alpha_1 = 0.7$	0.671	0.028	13.265	−1.94
$\beta_1 = 0.2$	0.207	0.033	13.471	1.949
$\omega = 2$	2.333	0.382	19.613	−0.442
$\alpha_1 = 0.3$	0.271	0.027	32.791	−0.387
$\beta_1 = 0.6$	0.614	0.036	34.447	0.443

Table 3: Simulated data for the NB1-INGARCH.

True values	Mean	Stdev	IF	CD
		T = 100		
$\omega = 0.1$	0.128	0.074	16.646	−0.580
$\alpha_1 = 0.7$	0.544	0.142	15.02	1.154
$\beta_1 = 0.2$	0.403	0.134	18.165	−0.240
$r_1 = 4$	3.097	0.555	2.658	0.196
$\omega = 1$	1.794	0.664	22.683	−2.606
$\alpha_1 = 0.3$	0.407	0.104	34.131	−0.451
$\beta_1 = 0.6$	0.450	0.118	30.431	1.614
$r_1 = 8$	7.312	1.609	1.061	−1.095
		T = 500		
$\omega = 0.1$	0.070	0.016	12.073	−0.687
$\alpha_1 = 0.7$	0.664	0.068	12.626	−1.011
$\beta_1 = 0.2$	0.225	0.061	12.15	1.375
$r_1 = 4$	3.147	1.596	1.208	1.005
$\omega = 1$	1.193	0.272	22.033	0.504
$\alpha_1 = 0.3$	0.294	0.037	21.711	1.349
$\beta_1 = 0.6$	0.549	0.054	29.378	−0.970
$r_1 = 8$	6.574	2.090	1.030	0.361
		T = 1000		
$\omega = 0.1$	0.097	0.014	13.324	−0.280
$\alpha_1 = 0.7$	0.693	0.047	11.469	−0.829
$\beta_1 = 0.2$	0.183	0.043	12.429	−0.023
$r_1 = 4$	3.735	0.792	1.2687	−0.472
$\omega = 1$	1.292	0.215	16.301	−2.680
$\alpha_1 = 0.3$	0.303	0.027	17.708	−2.263
$\beta_1 = 0.6$	0.522	0.043	18.519	2.818
$r_1 = 8$	7.101	2.375	1.022	−0.147

6 Empirical analysis

6.1 Data

Our empirical data consist of four time series that record the number of trades for four stocks (Glatfelter Company (GLT), Wausau Paper Corporation (WPP), Empire District Electric Company (EDE), Ericsson B). For the first three stocks (GLT, WPP, EDE) we monitor the number of stock transactions in 5-min intervals between 9:45 AM and 4:00 PM. Each of these three series has $T = 2925$ observations and the time period is from January 3, 2005 to February 18, 2005. For the last stock (Ericsson B) the time series is of length $T = 460$ and records the number of transactions per minute between 9:35 AM and 17:14 PM on 2 July 2002. Plots of the time series and histograms are given in Figures 1 and 2. The data are strongly overdispersed. The estimation results are presented in the Online Appendix.

6.2 Forecasting results

For our out-of-sample forecasting exercise, the evaluation period consists of the last 100 data points. The summary of the forecasting results is presented in Tables 7 (density forecasts) and 8 (point forecasts). The detailed forecasting results are reported in the Online Appendix.

Table 4: Simulated data for the NB2-INGARCH.

True values	Mean	Stdev	IF	CD
		T = 100		
$\omega = 1$	1.433	0.53323	14.28	0.626
$\alpha_1 = 0.7$	0.614	0.114	20.843	0.407
$\beta_1 = 0.2$	0.242	0.106	23.353	−0.867
$r_2 = 8$	7.324	0.960	1.076	2.487
$\omega = 2$	1.642	0.645	16.53	−2.307
$\alpha_1 = 0.4$	0.399	0.0937	19.577	−1.047
$\beta_1 = 0.4$	0.474	0.104	20.481	2.686
$r_2 = 4$	3.296	0.640	1.077	−2.484
		T = 500		
$\omega = 1$	0.970	0.189	13.018	−0.879
$\alpha_1 = 0.7$	0.652	0.047	15.223	−1.039
$\beta_1 = 0.2$	0.244	0.050	16.47	0.946
$r_2 = 8$	9.013	1.244	1	1.391
$\omega = 2$	1.959	0.414	14.726	−1.188
$\alpha_1 = 0.4$	0.422	0.047	13.091	−1.556
$\beta_1 = 0.4$	0.392	0.062	13.63	1.896
$r_2 = 4$	3.911	0.301	1.486	0.877
		T = 1000		
$\omega = 1$	1.276	0.161	11.897	−1.524
$\alpha_1 = 0.7$	0.692	0.031	12.667	−1.624
$\beta_1 = 0.2$	0.176	0.033	12.447	2.585
$r_2 = 8$	8.093	0.711	1	−0.906
$\omega = 2$	2.050	0.318	15.3	−0.411
$\alpha_1 = 0.4$	0.380	0.032	12.198	−1.050
$\beta_1 = 0.4$	0.413	0.049	15.073	0.761
$r_2 = 4$	4.036	0.201	3.393	0.890

Table 5: Simulated data for the DP-INGARCH.

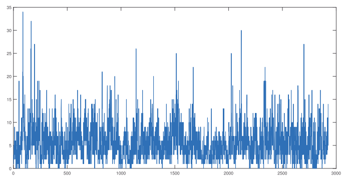
True values	Mean	Stdev	IF	CD
		T = 100		
$\omega = 4$	5.165	1.323	23.868	−0.992
$\alpha_1 = 0.2$	0.119	0.067	25.212	−0.333
$\beta_1 = 0.5$	0.218	0.093	23.824	1.594
$\gamma = 1$	0.827	0.124	1.0251	−1.370
$\omega = 1$	1.062	0.442	16.755	0.431
$\alpha_1 = 0.6$	0.518	0.099	19.194	−1.037
$\beta_1 = 0.2$	0.304	0.117	27.208	−0.086
$\gamma = 0.2$	0.274	0.040	1.036	−0.799
		T = 500		
$\omega = 4$	4.823	0.945	68.29	0.326
$\alpha_1 = 0.2$	0.158	0.040	67.554	1.534
$\beta_1 = 0.5$	0.343	0.105	88.009	−0.689
$\gamma = 1$	1.508	0.095	1.006	−0.678

Table 5: (continued)

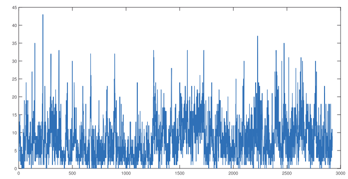
True values	Mean	Stdev	IF	CD
		T = 500		
$\omega = 1$	1.116	0.180	13.548	−1.563
$\alpha_1 = 0.6$	0.471	0.049	13.935	−1.740
$\beta_1 = 0.2$	0.181	0.064	13.467	2.154
$\gamma = 0.2$	0.292	0.018	1	0.234
		T = 1000		
$\omega = 4$	3.994	0.681	79.012	−1.254
$\alpha_1 = 0.2$	0.178	0.026	26.58	−0.866
$\beta_1 = 0.5$	0.408	0.077	60.079	1.264
$\gamma = 1$	1.391	0.063	1.0026	−1.114
$\omega = 1$	1.167	0.126	11.768	−1.294
$\alpha_1 = 0.6$	0.541	0.034	12.64	−1.270
$\beta_1 = 0.2$	0.133	0.041	11.994	1.598
$\gamma = 0.2$	0.271	0.012	1.670	−0.093

Table 6: Simulated data for the GP-INGARCH.

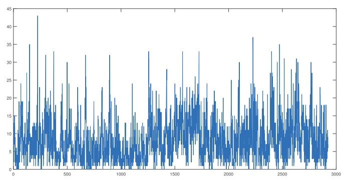
True values	Mean	Stdev	IF	CD
		T = 100		
$\omega = 0.8$	1.0958	0.0577	8.9458	−1.8958
$\alpha_1 = 0.3$	0.4189	0.0636	12.723	2.0518
$\beta_1 = 0.4$	0.3059	0.0342	14.943	−2.5536
$\tau = 0.4$	0.508	0.267	60.38	0.2622
$\omega = 2$	2.4604	0.1664	1.3222	−3.128
$\alpha_1 = 0.6$	0.7107	0.0384	1.2769	−3.7252
$\beta_1 = 0.2$	0.3594	0.0343	1.2593	3.5576
$\tau = 1$	1.436	9.158	46.951	0.1369
		T = 500		
$\omega = 0.8$	1.0072	0.0327	1.9565	12.628
$\alpha_1 = 0.3$	0.3558	0.0126	2.9502	−5.2772
$\beta_1 = 0.4$	0.4539	0.0070	18.675	0.4067
$\tau = 0.4$	0.4559	0.355	51.84	−0.0146
$\omega = 2$	2.2753	0.0187	2.1335	−3.3031
$\alpha_1 = 0.6$	0.6420	0.0111	2.5379	−1.8379
$\beta_1 = 0.2$	0.2078	0.0067	1.6509	2.6845
$\tau = 1$	1.199	9.06	71.57	0.0022
		T = 1000		
$\omega = 0.8$	0.8358	0.0079	4.1953	7.3745
$\alpha_1 = 0.3$	0.3211	0.0073	12.278	−0.7614
$\beta_1 = 0.4$	0.4080	0.0071	15.162	−0.5835
$\tau = 0.4$	0.424	0.379	70.542	−2.6647
$\omega = 2$	2.0748	0.0691	2.0951	6.3428
$\alpha_1 = 0.6$	0.6108	0.1066	2.16	−7.4312
$\beta_1 = 0.2$	0.1591	0.0576	1.9985	4.8223
$\tau = 1$	1.1458	9.443	10.41	−1.3431



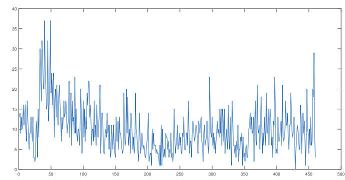
(a) GLT.



(b) EDE.

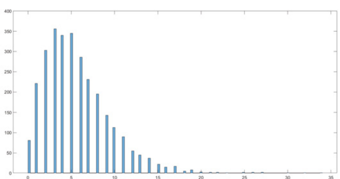


(c) WPP.

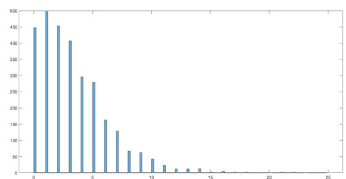


(d) Ericsson B

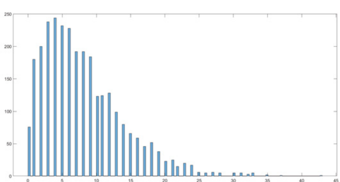
Figure 1: Empirical results: Time series plots for the four financial series.



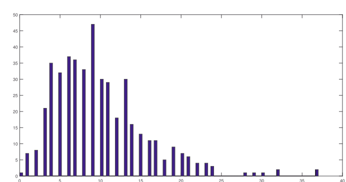
(a) GLT.



(b) EDE.



(c) WPP.



(d) Ericsson B

Figure 2: Empirical results: Histograms for the four financial time series.

Table 7: Summary table for the LPS results.

Data	$\boldsymbol{s = 1}$	$\boldsymbol{s = 4}$	$\boldsymbol{s = 8}$
GLT	NB2-INGARCH	NB2-INGARCH	NB2-INGARCH
WPP	NB2-INGARCH	NB2-INGARCH	NB2-INGARCH
EDE	NB1-INGARCH	NB2-INGARCH	NB2-INGARCH
Ericsson B	NB1-INGARCH	NB1-INGARCH	NB1-INGARCH

Table 8: Summary table for the sMSE results.

Data	$\boldsymbol{s = 1}$	$\boldsymbol{s = 4}$	$\boldsymbol{s = 8}$
GLT	NB2-INGARCH	NB2-INGARCH	NB2-INGARCH
WPP	NB1-INGARCH	NB1-INGARCH	NB1-INGARCH
EDE	NB1-INGARCH	NB1-INGARCH	NB1-INGARCH
Ericsson B	NB2-INGARCH	NB2-INGARCH	NB2-INGARCH

From Table 7 we can see that the NB2-INGARCH model is dominant for the first three data sets, producing better density forecasts than the other completing INGARCH specifications across all forecast horizons. The NB1-INGARCH yielded the best forecasting results for the Ericsson B data set only. The third-best model is the DP-INGARCH (see Online Appendix). The GP-INGARCH produced the worst density forecasting results.

From Table 8, the results indicate that both the NB1-INGARCH and NB2-INGARCH models produce better point forecasts than the P-INGARCH, the DP-INGARCH and the GP-INGARCH models. In most of the datasets, the DP-INGARCH did better than the P-INGARCH and the GP-INGARCH (Online Appendix).

7 Conclusions

We conducted a Bayesian forecasting exercise using INGARCH models with various conditional distributions. Our empirical application concerned the number of stock trades. We found that the NB2-INGARCH model is superior, in terms of density forecasts, to other competing models in predicting transaction counts, whereas the NB1-INGARCH and NB2-INGARCH models seem to dominate in point forecasting.

The set of existing forecasting models considered here could be extended to include additional models such as the COM-poisson INGARCH model (Zhu 2012c), the zero-inflated versions of it, the log-linear INGARCH model and the Softplus INGARCH model (Weiss et al. 2022). Such an extension will be examined in a future paper.

Acknowledgement: All computations have been performed in the Advanced Research Computing (ARC) environment at the University of Leeds.
Author contribution: All the authors have accepted responsibility for the entire content of this submitted manuscript and approved submission.
Research funding: None declared.
Conflict of interest statement: The authors declare no conflicts of interest regarding this article.

References

Ahmad, A., and C. Francq. 2016. "Poisson QMLE of Count Time Series Models." *Journal of Time Series Analysis* 37 (3): 291–314.
Aknouche, A., S. Bendjeddou, and N. Touche. 2018. "Negative Binomial Quasi-Likelihood Inference for General Integer-Valued Time Series Models." *Journal of Time Series Analysis* 39 (2): 192–211.
Aknouche, A., and C. Francq. 2021. "Count and Duration Time Series with Equal Conditional Stochastic and Mean Orders." *Econometric Theory* 37 (2): 248–80.
Atchadé, Y. F. 2006. "An Adaptive Version for the Metropolis Adjusted Langevin Algorithm with a Truncated Drift." *Methodology and Computing in Applied Probability* 8: 235–54.
Berry, L., and M. West. 2020. "Bayesian Forecasting of Many Count-Valued Time Series." *Journal of Business & Economic Statistics* 38 (4): 872–87.
Cameron, A., and P. Trivedi. 2013. *Regression Analysis of Count Data*. Cambridge: University Press.
Chen, C. W. S., M. So, J. C. Li, and S. Sriboonchitta. 2016. "Autoregressive Conditional Negative Binomial Model Applied to Over-dispersed Time Series of Counts." *Statistical Methodology* 31: 73–90.
Chib, S. 2001. "Markov Chain Monte Carlo Methods: Computation and Inference." In *Handbook of Econometrics*, edited by J. J. Heckman, and E. V. Leamer, 3569–649. Amsterdam: North-Holland.
Christou, V., and K. Fokianos. 2014. "Quasi-Likelihood Inference for Negative Binomial Time Series Models." *Journal of Time Series Analysis* 35 (1): 55–78.
Davis, R. A., and H. Liu. 2016. "Theory and Inference for a Class of Observation-Driven Models with Application to Time Series of Counts." *Statistica Sinica* 26 (4): 1673–707.
Doukhan, P., K. Fokianos, and D. Tjøstheim. 2012. "On Weak Dependence Conditions for Poisson Autoregressions." *Statistics & Probability Letters* 82 (5): 942–8.
Ferland, R., A. Latour, and D. Oraichi. 2006. "Integer-Valued GARCH Process." *Journal of Time Series Analysis* 27 (6): 923–42.
Fokianos, K., A. Rahbek, and D. Tjøstheim. 2009. "Poisson Autoregression." *Journal of the American Statistical Association* 140 (488): 1430–9.

Geweke, J. 1992. "Evaluating the Accuracy of Sampling-Based Approaches to the Calculation of Posterior Moments." In *Bayesian Statistics 4*, edited by J. M. Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith, 641–9. Oxford: Clarendon Press.

Geweke, J., and G. Amisano. 2011. "Hierarchical Markov Normal Mixture Models with Applications to Financial Asset Returns." *Journal of Applied Econometrics* 26 (1): 1–29.

Heinen, A. 2003. *Modelling Time Series Count Data: An Autoregressive Conditional Poisson Model*. Also available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1117187.

Homburg, A., C. H. Weiss, L. C. Alwan, G. Frahm, and R. Göb. 2019. "Evaluating Approximate Point Forecasting of Count Processes." *Econometrics* 7 (3).

Homburg, A., C. H. Weiss, L. C. Alwan, G. Frahm, and R. Göb. 2020. "A Performance Analysis of Prediction Intervals for Count Time Series." *Journal of Forecasting* 40 (4): 603–25.

McCabe, B. P. M., and G. M. Martin. 2005. "Bayesian Predictions of Low Count Time Series." *International Journal of Forecasting* 21 (2): 315–30.

McCabe, B. P. M., G. M. Martin, and D. Harris. 2011. "Efficient Probabilistic Forecasts for Counts." *Journal of the Royal Statistical Society: Series B* 73 (2): 253–72.

Martino, L., H. Yang, D. Luengo, J. Kanniainen, and J. Corande. 2015. "A Fast Universal Self-Tuned Sampler within Gibbs Sampling." *Digital Signal Processing* 47: 68–83.

Rydberg, T. H., and N. Shephard. 2000. *BIN Models for Trade-by-Trade Data. Modelling the Number of Trades in a Fixed Interval of Time*. Also available at https://ideas.repec.org/p/ecm/wc2000/0740.html.

Weiss, C. H. 2018. *An Introduction to Discrete-Valued Time Series*. Hoboken: Wiley.

Weiss, C. H., F. Zhu, and A. Hoshiyar. 2022. "Softplus INGARCH Models." *Statistica Sinica*. (forthcoming).

Xu, H-Y., M. Xie, T. N. Goh, and X. Fu. 2012. "A Model for Integer-Valued Time Series with Conditional Overdispersion." *Computational Statistics & Data Analysis* 56 (12): 4229–42.

Zhu, F. 2011. "A Negative Binomial Integer-Valued GARCH Model." *Journal of Time Series Analysis* 32 (1): 54–67.

Zhu, F. 2012a. "Modeling Overdispersed or Underdispersed Count Data with Generalized Poisson Integer-Valued GARCH Models." *Journal of Mathematical Analysis and its Applications* 389 (1): 58–71.

Zhu, F. 2012b. "Zero-Inflated Poisson and Negative Binomial Integer-Valued GARCH Models." *Journal of Statistical Planning and Inference* 142 (4): 826–39.

Zhu, F. 2012c. "Modeling Time Series of Counts with COM-Poisson INGARCH Models." *Mathematical and Computer Modelling* 56 (9–10): 191–203.

Supplementary Material: The online version of this article offers supplementary material (https://doi.org/10.1515/snde-2020-0095).