

# Forecasting evaluation via parametric bootstrap for threshold-INARCH models

Deok Ryun Kim<sup>ab</sup>, Sun Young Hwang<sup>1,a</sup>

<sup>a</sup>Department of Statistics, Sookmyung Women's University, Korea;

<sup>b</sup>Department of Biostatistics and Data Management, International Vaccine Institute, Korea

---

## Abstract

This article is concerned with the issue of forecasting and evaluation of threshold-asymmetric volatility models for time series of count data. In particular, threshold integer-valued models with conditional Poisson and conditional negative binomial distributions are highlighted. Based on the parametric bootstrap method, some evaluation measures are discussed in terms of one-step ahead forecasting. A parametric bootstrap procedure is explained from which directional measure, magnitude measure and expected cost of misclassification are discussed to evaluate competing models. The cholera data in Bangladesh from 1988 to 2016 is analyzed as a real application.

Keywords: integer-valued time series, one-step ahead forecasting, parametric bootstrap

---

## 1. Introduction

It is obvious that one of the most important objectives in building a time series model is to forecast its future values as accurate as the model can. The better forecasting performance of the model would be the property for the better model to possess. For instance, the accurate incidence forecasting of infectious disease is critical for early prevention and for better strategy development in public health surveillance field (Cardinal *et al.*, 1999; Zhang *et al.*, 2013).

Disease incidence time series can be handled using ad hoc integer-valued time series models (instead of ARMA) since the non-negative integer-valued forecasts as well as forecast limits are obtained. The integer-valued time series models are rigorously studied and applied over past three decades. Various integer-valued models with conditional Poisson distribution (Cardinal *et al.*, 1999; Ferland *et al.*, 2006) have been studied. However, the equi-dispersion requirement in Poisson distribution is often violated in reality. The integer-valued models with negative binomial distribution for over-dispersion (Zhu, 2011; Yoon and Hwang, 2015), the model with power-law weighted Poisson distribution for under-dispersion (Weiß, 2013), and the extended model with double Poisson and generalized Poisson for both under-dispersion and over-dispersion (Zhu, 2012; Bourguignon *et al.*, 2019), the models with threshold asymmetric volatility (Wang *et al.*, 2014; Liu *et al.*, 2019; Kim *et al.*, 2019) have been introduced as alternatives to the standard Poisson model.

Due to non-linearity in most integer-valued time series models, there exist no closed-form formulas to compute multi-step ahead forecasts of nonlinear models (Tsay, 2010). Both Monte Carlo simulation and bootstrap methods are popular to compute nonlinear forecasts (Kock and Teräsvirta,

---

<sup>1</sup> Corresponding author: Department of Statistics, Sookmyung Women's University, Cheongpa-ro 47-gil 100, Seoul 04310, Korea. E-mail: shwang@sookmyung.ac.kr

2010). Taieb *et al.* (2012), presented a thorough review as well as a theoretical comparative analysis of the existing strategies for multi-step ahead forecasting.

In this paper we study a parametric bootstrap method for forecasting integer-valued-threshold autoregressive conditional heteroskedastic (INTARCH) model investigated by Kim *et al.* (2019) and evaluate forecasting performance of the model using several measures based on the parametric bootstrap. The organization of this paper is as follows. Section 2 reviews existing threshold integer-valued analogue of the ARCH models and then presents procedures for one-step ahead forecasting using parametric bootstrap. Several measures to evaluate the forecasting performance are described in Section 3. Section 4 illustrates one-step ahead forecasting and evaluation of forecasting performance with applications to actual, highly skewed, over-dispersed, and serially correlated data example of cholera disease in Bangladesh from 1988 to 2016. Concluding remarks are presented in Section 5.

## 2. One-step ahead forecast using parametric bootstrap

### 2.1. Model description

In this section, model specifications are mainly based on Kim *et al.* (2019). We define the first-order integer-valued ARCH (INARCH(1)) model (Ferland *et al.*, 2006) as

$$X_t|F_{t-1} \sim \text{Poisson}(\lambda_t), \quad \lambda_t = \alpha_0 + \alpha_1 X_{t-1}, \quad (2.1)$$

where  $F_{t-1}$  denotes the information available up to time  $t - 1$ . The parameters  $\alpha_0$  and  $\alpha_1$  are positive. The model (2.1) is extended to have a two-regime structure of the conditional mean process according to the magnitude of the observation. A threshold model based on (2.1), so called, the first-order integer-valued threshold ARCH (INTARCH(1), hereafter) model is written as (Wang *et al.*, 2014)

$$X_t|F_{t-1} \sim \text{Poisson}(\lambda_t), \quad \lambda_t = \alpha_0 + \alpha_1 X_{t-1}^{(r)} + \alpha_2 X_{t-1}^{(l)}$$

$$X_{t-1}^{(r)} = \begin{cases} X_{t-1}, & \text{if } X_{t-1} > m_t, \\ 0, & \text{if } X_{t-1} \leq m_t, \end{cases} \quad X_{t-1}^{(l)} = \begin{cases} X_{t-1}, & \text{if } X_{t-1} \leq m_t, \\ 0, & \text{if } X_{t-1} > m_t, \end{cases} \quad (2.2)$$

where the parameters  $\alpha_0$ ,  $\alpha_1$ , and  $\alpha_2$  are positive, and the initial value  $X_0 = x_0$  is fixed. Here,  $m_t$  is a time varying threshold variable that determines the dynamic switching mechanism of the model. The dynamics of the process is governed by a two-regime scheme which are disjoint. That is,  $X_{t-1}^{(r)} X_{t-1}^{(l)} = 0$ . As to  $m_t$ , Kim *et al.* (2019) suggested to use either grand mean of the entire time series or the local mean according to the local fluctuation of the time series data.

With replacing (2.1) by a negative binomial distribution in order to accommodate over-dispersion structure in the data, the first-order integer-valued negative binomial ARCH (NB-INARCH(1)) model (Zhu, 2011; Yoon and Hwang, 2015) is defined as

$$X_t|F_{t-1} \sim \text{NB}(\lambda_t, a), \quad \lambda_t = \frac{1 - p_t}{p_t} = \alpha_0 + \alpha_1 X_{t-1}, \quad (2.3)$$

where the  $\text{NB}(\lambda, a)$  denotes the following probability mass function given by

$$P(X = k) = \frac{\Gamma(k + \frac{\lambda^{1-c}}{a})}{k! \Gamma(\frac{\lambda^{1-c}}{a})} \left( \frac{1}{1 + a\lambda^c} \right)^{\frac{\lambda^{1-c}}{a}} \left( \frac{a\lambda^c}{1 + a\lambda^c} \right)^k, \quad k = 0, 1, 2, \dots, \quad (2.4)$$

where  $\lambda > 0$ , the dispersion parameter  $a > 0$ , and  $\Gamma$  is the standard Gamma function. The index  $c$  (taking values 0 or 1) determines the form of the underlying NB distribution (Saha and Dong, 1997).

For  $c = 0$ , this particular distribution is denoted by  $NB1(\lambda, a)$  and the case of  $c = 1$  is named as  $NB2(\lambda, a)$ .

The NB-INARCH(1) can be extended to the following threshold-asymmetric model as a generalization of (2.3).

$$X_t|F_{t-1} \sim NB(\lambda_t, a), \quad \lambda_t = \frac{1-p_t}{p_t} = \alpha_0 + \alpha_1 X_{t-1}^{(r)} + \alpha_2 X_{t-1}^{(l)}, \quad (2.5)$$

where  $X_{t-1}^{(r)}$  and  $X_{t-1}^{(l)}$  are two threshold values defined by (2.2). The model (2.5) is referred to as negative binomial INTARCH(1), denoted as NB-INTARCH(1).

## 2.2. Parametric bootstrap algorithm for one-step ahead forecasting

An one-step ahead forecast  $\hat{X}_T(1)$  refers to the predicted value of  $X_{T+1}$  based on the data  $\{X_1, \dots, X_T\}$ . Here the lead time “1” denotes the forecasting horizon which is the one time ahead at the time origin  $T$ . To obtain  $\hat{X}_T(1)$ , the sample  $\{X_1, \dots, X_T\}$  is used to build the models (and estimate parameters appearing in the model) introduced earlier in Section 2.1. We simulate the conditional distribution  $X_{T+1}|F_T$  in which parameters are specified as estimated, and then obtain the forecast  $\hat{X}_T(1)$  based on simulated values. Similarly, to forecast  $\hat{X}_{T+1}(1)$ ,  $\{X_1, \dots, X_T, X_{T+1}\}$  is used to build model and estimate parameters, and finally, to forecast  $\hat{X}_{T+m-1}(1)$ ,  $\{X_1, \dots, X_T, \dots, X_{T+m-1}\}$  is used for model building and the estimation of parameters. Consequently, we obtain  $m$  1-step ahead forecasts, viz.,  $\hat{X}_{T+j-1}(1)$ ,  $j = 1, \dots, m$ . For each fixed  $j = 1, \dots, m$ , from the conditional distribution  $X_{T+j}|F_{T+j-1}$ , we generate  $M$  simulated values which are denoted as  $\{\hat{X}_{T+j-1}^{(k)}(1)\}_{k=1}^M$ . The point forecast  $\hat{X}_{T+j-1}(1)$  is taken as the sample median of  $\{\hat{X}_{T+j-1}^{(k)}(1)\}_{k=1}^M$ . Since we are dealing with integer-valued time series, we prefer to use the sample median instead of the sample mean. The  $m$  forecasts  $\{\hat{X}_T(1), \dots, \hat{X}_{T+m-1}(1)\}$  and observed  $m$  values  $\{X_{T+1}, \dots, X_{T+m}\}$  are compared to evaluate forecasting performance of the models. In the real data application discussed in Section 4, we have chosen  $M = 500$  and  $m = 300$ . We refer to, for instance, Section 4.4 of Tsay (2010) for a comprehensive treatment of parametric bootstrap methods in a broader context of time series analysis.

## 3. Forecasting evaluation

There are many ways to evaluate the forecasting performance of a time series model. The methods of forecasting evaluation described in this section are mostly based on Section 4.3 of Tsay (2010). Directional measures and magnitude measures are discussed. A directional measure considers the future direction (up or down) implied by the model. For instance, whether next week’s cholera cases will go up and down, or keep staying is an example of directional forecasts that are of practical interest. Magnitude measures are a function of the actual and predicted values of the time series and is used to assess accuracy of the predicted values, and relative ratio of magnitude measures between models is used for the model comparison.

### 3.1. Directional measure

Tsay (2010) described the directional measure as a performance evaluation measure. By considering many zeros in the integer-valued time series as well as no directional changes due to count time series data, a  $3 \times 3$  contingency table is newly suggested which summarizes the number of “hits” or “misses” of the model in predicting upwards, downwards, and no movements of  $X_{T+j-1}$  in the forecasting

subsample. The upward, downwards, or no movements (stay) in predicted values is counted. The contingency table can be summarized as

Actual	Predicted			Total
	up (> 0)	stay (= 0)	down (< 0)	
UP (> 0)	$m_{11}$	$m_{12}$	$m_{13}$	$m_{10}$
STAY (= 0)	$m_{21}$	$m_{22}$	$m_{23}$	$m_{20}$
DOWN (< 0)	$m_{31}$	$m_{32}$	$m_{33}$	$m_{30}$
Total	$m_{01}$	$m_{02}$	$m_{03}$	$m$

where  $m$  is the total number of 1-step ahead forecasts in the forecasting subsample. The number of hits of the model in predicting upwards, no movements (stay), or downwards are counted if the signs from both predicted and actual values are positive ( $m_{11}$ ), zero ( $m_{22}$ ), or negative ( $m_{33}$ ), respectively. Otherwise, they are counted as misses such that  $m_{12}$  and  $m_{13}$  are the number of misses in predicting upward movements,  $m_{21}$  and  $m_{23}$  are the number of misses in predicting no movements, and  $m_{31}$  and  $m_{32}$  are number of misses in predicting downward movements, respectively. The larger values in  $m_{11}$ ,  $m_{22}$ , and  $m_{33}$  indicate better forecasts. The percentage of hits in each directional change and the average percentage of hits is evaluated among predictive models. In addition, the weighted average percent of hits is calculated for the evaluation because the importance of directional changes may be asymmetric in real situation. As an example, the cost for preventing infectious disease by vaccination program might be much lower than that for the response to an outbreak of the infectious disease and therefore a model has a higher proportion of hits in predicting upward movements than the one in predicting the other directional changes is deemed a better model in forecasting.

### 3.2. Magnitude Measure

Four statistics are used to measure the performance of point forecasts from each model. They are the mean squared error (MSE), mean absolute error (MAE), mean absolute scaled error (MASE), and relative mean absolute error (relMAE). The first two measures (MSE and MAE) depending on the scale of data are commonly used accuracy measures when comparing different methods applied to the same dataset, while MASE proposed by Hynman and Koehler (2006) is useful when comparing across data that have a different scale by scaling mean absolute difference between consecutive actual observations. The MASE is also alternatively useful when absolute percentage error is not computed due to division by zero observation in the data. The relMAE (Hynman and Koehler, 2006) is the ratio of the MAE between the comparator and referenced models. For one-step ahead forecasts, these measures are defined as

$$\text{MSE}(1) = \frac{1}{m} \sum_{j=0}^{m-1} [X_{T+j+1} - \hat{X}_{T+j}(1)]^2, \quad (3.1)$$

$$\text{MAE}(1) = \frac{1}{m} \sum_{j=0}^{m-1} |X_{T+j+1} - \hat{X}_{T+j}(1)|, \quad (3.2)$$

$$\text{MASE}(1) = \frac{\frac{1}{m} \sum_{j=0}^{m-1} |X_{T+j+1} - \hat{X}_{T+j}(1)|}{\frac{1}{m-1} \sum_{j=1}^{m-1} |X_{T+j+1} - X_{T+j}|}, \quad (3.3)$$

$$\text{relMAE}_{A,B}(1) = \frac{\text{MAE}_A(1)}{\text{MAE}_B(1)}, \quad (3.4)$$

where  $m$  is the number of 1-step ahead forecasts available in the forecasting subsample, prediction error at  $j$ ,  $X_{T+j+1} - \hat{X}_{T+j}(1)$  is a sample difference for pair  $j$ , and the difference between consecutive observations,  $X_{T+j+1} - X_{T+j}$  is the 1-lag difference of actual observed values. The model with the smallest magnitude on these measures is regarded as the best one-step ahead forecasting model. Especially, values of MASE less than one indicate better forecasts than the average naive forecast computed on the actual observed values, and conversely, it is greater than one if the forecasts are worse than the average naive forecast computed on the actual observed values in the subsample for evaluation. Comparison of model performance between two models ( $A$  and  $B$ ) can be further facilitated by the relMAE defined in (3.4) as the ratio of two MAEs, viz.,  $MAE_A(1)$  and  $MAE_B(1)$  (Hyndman and Koehler, 2006).

### 3.3. Expected cost of misclassification

It will be instructive to consider the case in public health surveillance. As vaccination is one of most cost-effective intervention that contributes to healthcare system (Andre *et al.*, 2008), it is reasonable to assume that the cost including direct treatment and indirect social loss for patients is higher than the prevention cost by vaccination program. Accordingly, the forecasting performance needs to be evaluated by considering expected cost of misclassification (ECM) in practice. The ECM is derived from the cost table and the total probability of misclassification (TPM). The cost table consists of predefined misclassification cost in off-diagonal cells and zero cost in diagonal cells. The TPM and ECM are calculated as below (Timm, 2002). Here,  $C(\cdot)$  and  $P(\cdot)$  are used to denote the cost and the probability, respectively.

Actual	Predicted		
	up	stay	down
UP	0	$C(\text{stay UP})$	$C(\text{down UP})$
STAY	$C(\text{up STAY})$	0	$C(\text{down STAY})$
DOWN	$C(\text{up DOWN})$	$C(\text{stay DOWN})$	0

Actual	Predicted			
	up	stay	down	
UP	$P(\text{up UP})$	$P(\text{stay UP})$	$P(\text{down UP})$	$P(\text{UP})$
STAY	$P(\text{up STAY})$	$P(\text{stay STAY})$	$P(\text{down STAY})$	$P(\text{STAY})$
DOWN	$P(\text{up DOWN})$	$P(\text{stay DOWN})$	$P(\text{down DOWN})$	$P(\text{DOWN})$

$$\begin{aligned} \text{TPM} = & P(\text{stay|UP})P(\text{UP}) + P(\text{down|UP})P(\text{UP}) + P(\text{up|STAY})P(\text{STAY}) + P(\text{down|STAY})P(\text{STAY}) \\ & + P(\text{up|DOWN})P(\text{DOWN}) + P(\text{stay|DOWN})P(\text{DOWN}), \end{aligned}$$

$$\begin{aligned} \text{ECM} = & C(\text{stay|UP})P(\text{stay|UP})P(\text{UP}) + C(\text{down|UP})P(\text{down|UP})P(\text{UP}) \\ & + C(\text{up|STAY})P(\text{up|STAY})P(\text{STAY}) + C(\text{down|STAY})P(\text{down|STAY})P(\text{STAY}) \\ & + C(\text{up|DOWN})P(\text{up|DOWN})P(\text{DOWN}) + C(\text{stay|DOWN})P(\text{stay|DOWN})P(\text{DOWN}), \end{aligned}$$

where  $\text{stay|UP}$  means that the observed direction is UP but is incorrectly classified as stay. Similarly,  $\text{down|UP}$ ,  $\text{up|STAY}$ ,  $\text{down|STAY}$ ,  $\text{up|DOWN}$ , and  $\text{stay|DOWN}$  can be defined.

## 4. Real data analysis: The cholera data from Matlab in Bangladesh

In this section, as a real data application, we illustrate INTARCH and NB-INTARCH models which are defined in Section 2, and evaluate forecasting performance by calculating diverse measures de-

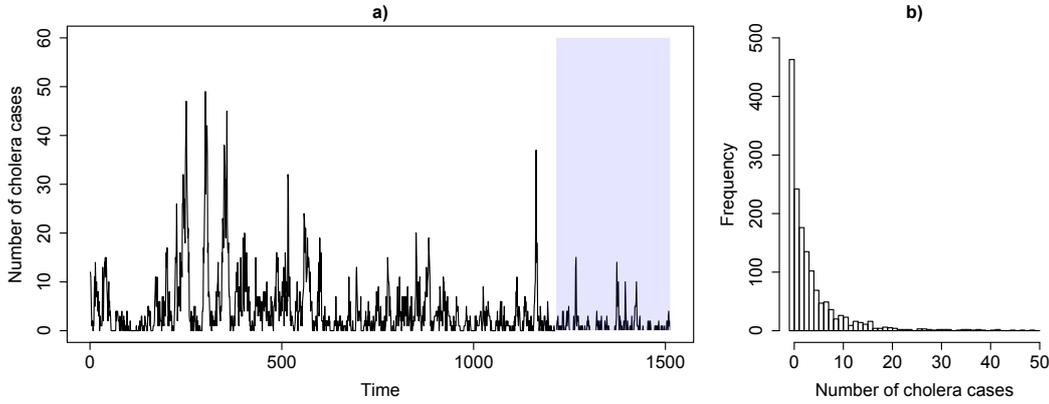


Figure 1: Weekly cholera cases series and histogram: a) entire period and 2nd-subsample for one-step ahead forecasts and evaluation (filled area, time = 1214 to 1513), b) histogram of cholera cases.

scribed in Section 3. We consider time series of weekly cholera cases from Matlab in Bangladesh, consisting of 1,513 observations starting from 1st-week of 1988 to 52nd-week of 2016. The data prior to 1988 were excluded because there was an oral cholera vaccine trial in early 1985 that offered direct and indirect protection to people of the area (Ali *et al.*, 2013). Figure 1 shows the time series plot of entire period with the 2nd-subsample consisting of high peaks for forecasting evaluation, and histogram of weekly cholera cases. Empirical mean and variance of the data are 3.89 and 37.74, respectively, which indicates highly over-dispersed. Note that NB-INARCH and NB-INTARCH models account for over-dispersion. A histogram of the series shows they are highly skewed.

The sample autocorrelation and partial autocorrelation function (not listed here) of the series imply that the first-order integer-valued autoregressive model is a reasonable candidate model for the data.

In earlier research, Ali *et al.* (2016) reported that there exists a significantly high risk for cholera within 1 week. Thus, one-step ahead forecasting was only considered in the study. It would also be interesting to consider multi-step ahead forecast. Since there exist, however, no closed-form formulas to compute multi-step ahead forecasts of threshold (nonlinear) INARCH models, further developments in the multiple-step programming will need to be made in a future study. By following the algorithm of parametric bootstrap method for one-step ahead forecasting in Section 2, we set forecast origin  $T = 1213$ , and we obtain total 300 one-step forecasts ( $m = 300$ ) as the forecast horizon is sequentially increased. For the 1st one-step ahead forecast, forecast horizon,  $\{X_1, \dots, X_{1213}\}$  is used to build model and one-step ahead forecast,  $\hat{X}_{1213}(1)$  is calculated from estimated parameters on the forecast horizon and the data. The procedures are repeated until the last one-step ahead forecast, viz.,  $\hat{X}_{1512}(1)$ . The point forecast of  $\hat{X}_{T+j-1}(1)$  is the sample median based on the bootstrapped  $M = 500$  samples. The 300 one-step ahead forecasts derived from each model and actual data are plotted in Figure 2 from which it is noted that all models are overall well-performed.

For the threshold models, two threshold variables ( $m_t$ ) are considered, that is, grand mean and local mean. The grand mean is simply the average of all observations while the local mean is time varying which is defined by

$$m_t = [\text{average of } (X_{t-4}, X_{t-3}, X_{t-2}, X_{t-1}) + 0.5], \quad (4.1)$$

where  $[x]$  denotes the greatest integer function not exceeding  $x$  (Kim *et al.*, 2019). The choice of the order 4 moving average in (4.1) is motivated by the 28 days (which is 4 weeks) incubation period and

Table 1: Contingency table of directional changes

Actual	Predicted									Total
	INARCH			INTARCH*			INTARCH+			
	up	stay	down	up	stay	down	up	stay	down	
UP	9	57	16	9	54	19	13	51	18	82
STAY	6	113	20	6	113	20	8	102	29	139
DOWN	41	23	15	42	23	14	44	21	14	79
Total	56	193	51	57	190	53	65	174	61	300

Actual	Predicted									Total
	NB1-INARCH			NB1-INTARCH*			NB1-INTARCH+			
	up	stay	down	up	stay	down	up	stay	down	
UP	18	35	29	18	35	29	18	36	28	82
STAY	7	96	36	7	96	36	7	96	36	139
DOWN	56	9	14	56	9	14	56	9	14	79
Total	81	140	79	81	140	79	81	141	78	300

Actual	Predicted									Total
	NB2-INARCH			NB2-INTARCH*			NB2-INTARCH+			
	up	stay	down	up	stay	down	up	stay	down	
UP	16	40	26	17	41	24	14	44	24	82
STAY	11	102	26	12	99	28	9	102	28	139
DOWN	41	24	14	41	24	14	44	20	15	79
Total	68	166	66	70	164	66	67	166	67	300

Table 2: Percentage of hits

Measures	Direction of predicted values								
	INARCH			INTARCH*			INTARCH+		
	up	stay	down	up	stay	down	up	stay	down
Percent of “hits” in each direction	11%	81%	19%	11%	81%	18%	16%	73%	18%
Average percent of “hits”	46%			45%			43%		
Weighted average percent of “hits”	27%			27%			29%		

Measures	Direction of predicted values								
	NB1-INARCH			NB1-INTARCH*			NB1-INTARCH+		
	up	stay	down	up	stay	down	up	stay	down
Percent of “hits” in each direction	22%	69%	22%	22%	69%	22%	22%	69%	22%
Average percent of “hits”	43%			43%			43%		
Weighted average percent of “hits”	30%			30%			30%		

Measures	Direction of predicted values								
	NB2-INARCH			NB2-INTARCH*			NB2-INTARCH+		
	up	stay	down	up	stay	down	up	stay	down
Percent of “hits” in each direction	20%	73%	22%	21%	71%	22%	17%	73%	23%
Average percent of “hits”	44%			43%			44%		
Weighted average percent of “hits”	29%			30%			29%		

The weighted average percent of “hits” is calculated with weights 0.5, 0.25, and 0.25 to each “hits” of upwards, stay, and downwards direction.

transmission period of cholera disease in the literature (Azman *et al.*, 2013; Ali *et al.*, 2016). The local constant mean  $m_t$  in (4.1) is used as a threshold variable.

In the tables and figures to follow, the grand mean and the local mean are designated respectively by the notation \* and +.

Diverse measures described in Section 3 are summarized in Tables 1 to 4 based on 300 one-step ahead forecasts in forecasting subsample. As shown in Tables 1 and 2, the models with negative binomial distribution predicted more number of “hits” in predicting upwards, while the models with Poisson distribution predicted more numbers of “hits” in predicting no movements (stay). All the

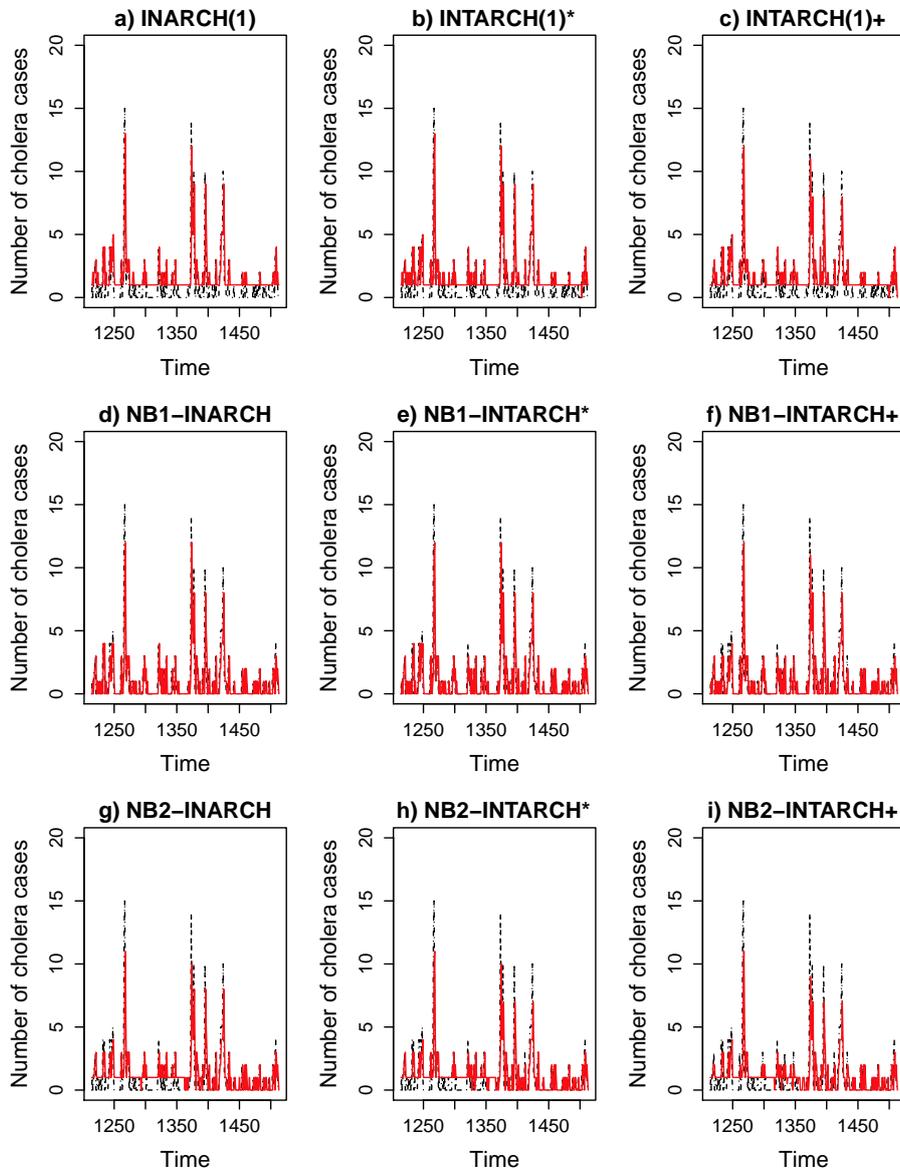


Figure 2: Actual observations and one-step ahead forecasts (sample median as predicted value): a)  $INARCH(1)$ , b)  $INTARCH(1)$  with grand mean threshold, c)  $INTARCH(1)$  with local mean threshold, d)  $NB1-INARCH(1)$ , e)  $NB1-INTARCH(1)$  with grand mean threshold, f)  $NB1-INTARCH(1)$  with local mean threshold, g)  $NB2-INARCH(1)$ , h)  $NB2-INTARCH(1)$  with grand mean threshold, i)  $NB2-INTARCH(1)$  with local mean threshold. Actual series and forecasts in dashed-black color and red color, respectively.

models predicted nearly the same number of “hits” in predicting downwards. Due to better predicting ability in no movements, the models with Poisson distribution had highest number of “hits” in predicting direction, but negative binomial model showed higher predicting ability if we consider higher

Table 3: Magnitude measures

Measures	Models		
	INARCH	INTARCH*	INTARCH+
Mean squared error (MSE)	3.36	3.40	3.25
Mean absolute error (MAE)	1.26	1.28	1.27
Mean absolute scaled error (MASE)	1.23	1.25	1.24
Relative MAE against Poisson model 1)	Ref	Ref	Ref
Relative MAE against no threshold model 2)	Ref	1.01	0.99

Measures	Models		
	NB1-INARCH	NB1-INTARCH*	NB1-INTARCH+
Mean squared error (MSE)	3.05	3.01	2.93
Mean absolute error (MAE)	0.99	0.98	0.96
Mean absolute scaled error (MASE)	0.97	0.96	0.94
Relative MAE against Poisson model 1)	0.78	0.77	0.76
Relative MAE against no threshold model 2)	Ref	0.99	0.98

Measures	Models		
	NB2-INARCH	NB2-INTARCH*	NB2-INTARCH+
Mean squared error (MSE)	2.96	2.96	2.95
Mean absolute error (MAE)	1.11	1.10	1.06
Mean absolute scaled error (MASE)	1.08	1.07	1.03
Relative MAE against Poisson model 1)	0.88	0.86	0.83
Relative MAE against no threshold model 2)	Ref	0.99	0.96

1) Relative MAE of each NB model against referenced Poisson model; 2) Relative MAE of each threshold model against referenced threshold model. Ref stands for the referenced model.

Table 4: Expected cost of misclassification

Measures	Models				
	INARCH	INTARCH*	INTARCH+	NB1-INARCH	NB1-INTARCH*
TPM	0.5433	0.5467	0.5700	0.5722	0.5733
ECM	3.3467	3.3600	3.3267	3.2733	3.2733

Measures	Models			
	NB1-INTARCH+	NB2-INARCH	NB2-INTARCH*	NB2-INTARCH+
TPM	0.5733	0.5600	0.5667	0.5633
ECM	3.2733	3.2267	3.2200	3.2933

importance of “hits” in predicting upwards over the others by giving weights, for instance, 0.5, 0.25, and 0.25 on “hits” in predicting upwards, staying, and downwards, respectively. Table 3 displays performance measures of point forecast. The models with negative binomial distribution had lower MSE, MAE, and MASE over the Poisson models. The lowest MSE was calculated in NB1-INTARCH with the local mean threshold defined in (4.1), which gives also the lowest MAE and MASE. Among the models with negative binomial distribution, the NB1-INTARCH model with local mean threshold showed slightly higher accuracy than the others. The relMAE showed negative binomial models showed better performances ( $relMAE < 1$ ) than Poisson models. Negative binomial model with local mean threshold provided better forecasting performance compared to the model without threshold. In terms of misclassification (Table 4) higher TPMs were showed in NB models due to higher misclassification rate in actual STAY direction. However, NB models and threshold models showed relatively lower misclassification rate in actual UP direction. According to the assumptions: comparable preventive healthcare cost when predicted STAY and DOWN direction, comparable treatment cost when observed stay and down direction, two times higher preventive cost when predicted UP direction than the other direction, and five time higher treatment cost when observed up direction than the other directions, we are willing to consider the following asymmetric cost table for which 6 off-diagonal

cells consist of relative cost of 2, 4, and 10 for both  $C(\text{stay}|\text{DOWN})$  and  $C(\text{down}|\text{STAY})$ , for both  $C(\text{up}|\text{STAY})$  and  $C(\text{up}|\text{DOWN})$ , and for  $C(\text{stay}|\text{UP})$  and  $C(\text{down}|\text{UP})$  respectively. With this cost table, lower ECM can be found in NB models than Poisson. Among threshold models, INTARCH+ and NB2-INATARCH\* models have lower ECM than models with no threshold.

## 5. Concluding remarks

In this paper, we have discussed the parametric bootstrap method of forecasting threshold INARCH models with conditional Poisson and negative binomial distributions. Overall, the negative binomial threshold ARCH (NB-INTARCH) model with local threshold variable showed relatively better performances compared to the others. Since the threshold models have certain advantages in integer valued time series, it will be of interest to make applications to time series of infectious disease incidence in public health surveillance. Further, in order to accommodate seasonal circulation, sanitation improvement, vaccination, and climate changes, intervention analysis in the context of INTARCH would make an interesting future study. This is now under investigation and will be addressed elsewhere.

## Acknowledgements

We thank the two reviewers for constructive comments which led to a substantial improvement in the revised version. The real data analysis in this paper is based on data collected and shared by the International Centre for Diarrhoeal Disease Research, Bangladesh (icddr,b) from an original study which was conducted with support from the governments of Bangladesh, Canada, Sweden and UK. SY Hwang's work was supported by a grant from the National Research Foundation of Korea (NRF-2018R1A2B2004157).

## References

- Ali M, Debes AK, Luquero FJ, Kim DR, Park JY, and Digilio L (2016). Potential for controlling cholera using a ring vaccination strategy: re-analysis of data from a cluster-randomized clinical trial, *PLOS Medicine*, **13**, e1002120.
- Ali M, Kim DR, Yunus M, and Emch M (2013). Time series analysis of cholera in Matlab, Bangladesh, during 1988–2001, *Journal of Health, Population, and Nutrition*, **31**, 11–19.
- Andre FE, Booy R, Bock HL, Clemens J, Datta SK, John TJ, and Schmitt HJ (2008). Vaccination greatly reduces disease, disability, death and inequity worldwide, *Bulletin of the World Health Organization*, **86**, 140–146.
- Azman AS, Rudolph KE, Cummings DAT, and Lessler J (2013). The incubation period of cholera: a systematic review, *Journal of Infection*, **66**, 432–438.
- Bourguignon M, Rodrigues J, and Santos-Neto M (2019). Extended Poisson INAR(1) processes with equidispersion, underdispersion and overdispersion, *Journal of Applied Statistics*, **46**, 101–118.
- Cardinal M, Roy R, and Lambert J (1999). On the application of integer-valued time series models for the analysis of disease incidence, *Statistics in Medicine*, **18**, 2025–2039.
- Ferland R, Latour A, and Oraichi D (2006). Integer-valued GARCH process, *Journal of Time Series Analysis*, **27**, 923–942.
- Hyndman RJ and Koehler AB (2006). Another look at measures of forecast accuracy, *International Journal of Forecasting*, **22**, 679–688.
- Kim DR, Yoon JE, and Hwang SY (2019). Threshold-asymmetric volatility models for integer-valued time series, *Communications for Statistical Applications and Methods*, **26**, 295–304.

- Kock AB and Teräsvirta T (2010). *Forecasting with nonlinear time series models* (CREATES Research Papers), Department of Economics and Business Economics, Aarhus University, 2010-01.
- Liu M, Li Q, and Zhu F (2019). Threshold negative binomial autoregressive model, *Statistics*, **53**, 1–25.
- Saha A and Dong D (1997). Estimating nested count data models, *Oxford Bulletin of Economics and Statistics*, **59**, 423–430.
- Taieb SB, Bontempi G, Atiya AF, and Sorjamaa A (2012). A review and comparison of strategies for multi-step ahead time series forecasting based on the NN5 forecasting competition, *Expert Systems with Applications*, **39**, 7067–7083.
- Timm NH (2002). *Applied Multivariate Analysis* (2nd ed), Springer, New York.
- Tsay RS (2010). *Analysis of Financial Time Series* (3rd Ed), John Wiley & Sons, Hoboken.
- Wang C, Liu H, Yao JF, Davis RA, and Li WK (2014). Self-excited threshold Poisson autoregression, *Journal of the American Statistical Association*, **109**, 777–787.
- Weiß CH (2010). The INARCH(1) model for overdispersed time series of counts, *Communications in Statistics: Simulation and Computation*, **39**, 1269–1291.
- Yoon JE and Hwang SY (2015). Integer-valued GARCH models for count time series: case study, *Korean Journal of Applied Statistics*, **28**, 115–122.
- Zhang X, Liu Y, Yang M, *et al.* (2013). Comparative study of four time series methods in forecasting typhoid fever incidence in China, *PLoS ONE*, **8**, e63116.
- Zhu F (2011). A negative binomial integer-valued GARCH model, *Journal of Time Series Analysis*, **32**, 54–67.
- Zhu F (2012). Modeling overdispersed or underdispersed count data with generalized Poisson integer-valued GARCH models, *Journal of Mathematical Analysis and Applications*, **389**, 58–71.

Received July 8, 2019; Revised December 30, 2019; Accepted January 3, 2020