Forecasting Foreign Tourist Using Intervention Analysis On Count Time Series

Eviyana Atmanegara1*, Suhartono1, and RM Atok1

1Department of Statistics, Faculty of Mathematics, Computation, and Data Science, Institut Teknologi Sepuluh Nopember, Surabaya 60111, Indonesia

*Corresponding email: eviyana@bps.go.id

Abstract. The foreign tourist's data are count time series data that contains the discrete value. Poisson autoregressive (AR) and Negative Binomial AR are time series models used for forecasting count data. The number of foreign tourist arrivals is influenced by the series of inputs called interventions, such as the existence of bomb terror and tourism promotion. This research aims to forecast the number of foreign tourists visiting Indonesia by nationality, 2019. The number of tourist arrivals from Bahrain and Singapore represents low count data and high count data, respectively. This work employs intervention analysis on count time series model and intervention analysis on ARIMA. Intervention on Poisson AR is the best model for forecasting the number of tourist arrivals from Bahrain and Singapore to Indonesia, 2019.

Keywords: Count Time Series, Intervention, ARIMA, Foreign Tourists, Poisson AR, Negative Binomial AR.

1. Introduction
Tourism is one of the fastest growing sectors in the economy. This sector is starting to become one of the main sources of income for many countries in the world. In line with the role of tourism in the world economy, tourism also plays an important role in the Indonesian economy. The number of tourist arrivals increases. During 2018, the number of foreign tourists visiting Indonesia reached 15.81 million visits, increase 12.58 percent compared to the number of foreign tourist visits in the same period in 2017 which reached 14.04 million visits [1].

Intervention is an extraordinary event that can affect data. Box and Tiao discussed the effects of the intervention to economic and environmental problems [2]. Likewise, intervention influenced data on foreign tourists visiting Indonesia. Interventions in tourism are divided into two factors, namely internal and external factors. Promotion and various strategies that support the development of tourism are internal factors. On the opposite, external factors are out of control factors like disasters and bomb blasts.

Various literature study about forecast the number of tourist arrivals. There are some research that have been done using time series regression and ARIMA models [3-7]. Lee et al. conducted a multi-input intervention analysis to evaluate the impact of the economic crisis in Asia in 1997 and terrorist attacks such as Bali Bombing I and Bali Bombing II on the arrival of foreign tourists to Indonesia through the Soekarno-Hatta entrance [8].

The count time series data contains discrete data that is widely observed in various events. Modeling for count time series data using the generalized linear model (GLM) approach has been
increasingly studied in recent years. Liboschik et al. have developed a data processing for count time series following generalized linear models analysis [9]. Modeling count time series data using Poisson AR has already performed [10-11]. Count time series model applied to model the negative binomial distributed assumed data [12]. The Comparison of GSARIMA and SARIMA model forecasting methods for forecasting monthly data on the number of dengue fever patients in Surabaya has already done [13]. One of the time series modeling using intervention analysis is intervention analysis in the INGARCH model [14]. Furthermore, intervention analysis was also developed in the log-linear Poisson AR model [15]. Research on outliers and interventions in count time series following GLMs was carried out [16].

In this research, we applied intervention on ARIMA and intervention on count time series model to forecast monthly international visitor arrivals to Indonesia. The number of international visitor arrivals from Bahrain and Singapore represent low count data and high count data, respectively. We perform a comparison of the estimation results from the models to obtain the best model.

This paper is organized as follows. Section 2 describes the theoretical part. Section 3 explains the methodology. Section 4 reports the empirical results and discussion. At last, Section 5 shows the conclusion.

2. Literature Review

2.1. Intervention Analysis on ARIMA Model

In general, ARIMA \((p, d, q)(P, D, Q)\) seasonal multiplicative models, generally written as follows [17]:

\[
\phi_p(B)\Phi_p(B^S)(1 - B)^d(1 - B^S)^D Y_t = \theta_q(B)\Theta_q(B^S)\alpha_t,
\]

where,

\(Y_t\) = response variable at time \(t\)

\(\phi_p(B)\) = coefficient of the AR component without a seasonal period with order \(p\)

\(\Phi_p(B^S)\) = coefficient of the AR component of the seasonal period \(S\) with order \(P\)

\(\theta_q(B)\) = coefficient of MA component without a seasonal period with order \(q\)

\(\Theta_q(B^S)\) = coefficient of the MA component of the seasonal period \(S\) with order \(Q\)

\((1 - B)^d\) = differencing without seasonal with order \(d\)

\((1 - B^S)^D\) = seasonal differencing \(S\) with order \(D\)

\(\alpha_t\) = residual white noise with mean 0 and variance \(\sigma_\alpha^2\).

The intervention model is a statistical model in the analysis group that is used to explain the effect of an event both internal and external which is expected to affect the predicted variable. There are two common types of intervention, namely step and pulse functions. Detailed explanations of intervention analysis can be found in [17]. An intervention model can be written as:

\[
Y_t = \frac{\omega_x(B)}{\delta_x(B)}B^b X_t + \frac{\theta_x(B)}{\phi_x(B)(1 - B)^d} \alpha_t.
\]

where,

\(Y_t\) = response variable at time \(t\) which shows the data is stationary

\(X_t\) = binary indicator variable that shows the existence of an intervention at time \(t\)

\(\omega_x(B) = \omega_0 - \omega_1B - \omega_2B^2 - \cdots - \omega_sB^s\)

\(\delta_x(B) = 1 - \delta_1B - \delta_2B^2 - \cdots - \delta_rB^r\)

\(\phi(B)\) = autoregressive operator

\(\theta(B)\) = moving-average operator

\(B\) = back shift operator
Eq. (2) shows the magnitude and period of intervention effect according to $b$, $s$, and $r$. The delay time is $b$, while $s$ gives information about the time which is needed for an effect of the intervention to be stable, and $r$ is the pattern of the intervention effect. A step function is an intervention type which occurs over a long term. And an intervention which occurs only at a certain time ($T$) is called a pulse intervention.

2.2. Intervention Analysis on Count Time Series Model

In general, count time series models, generally written as follows [9]:

$$g(\lambda_t) = \beta_0 + \sum_{k=1}^{p} \beta_k g(Y_{t-k}) + \sum_{l=1}^{q} \alpha_l g(\lambda_{t-l}),$$

where,

- $Y_t$ = a count time series as response variable
- $\lambda_t$ = conditional mean of the count time series, such that $E(Y_t|F_{t-1}) = \lambda_t$
- $g$: $\mathbb{R}^+ \rightarrow \mathbb{R}$ is a link function
- $\tilde{g}$: $\mathbb{N}_0 \rightarrow \mathbb{R}$ is a transformation function
- $\nu_t = g(\lambda_t)$ = the linear predictor
- $\beta_k$ = coefficient of the AR component on lagged observations with order $p$
- $\alpha_l$ = coefficient of the AR component on lagged conditional means with order $q$.

In the terminology of GLMs we call $\nu_t = g(\lambda_t)$ the linear predictor. Consider model (3), with the logarithmic link function $g(x) = \log(x)$, $\tilde{g}(x) = \log(x + 1)$. Then, we obtain a log-linear model of order $p$ and $q$ for the analysis of count time series. Indeed set $\nu_t = \log(\lambda_t)$ to obtain from (3) that

$$\nu_t = \beta_0 + \sum_{k=1}^{p} \beta_k \log(Y_{t-k} + 1) + \sum_{l=1}^{q} \alpha_l \nu_{t-l}. \quad (4)$$

Model (3) together with the Poisson assumption, i.e., $Y_t|F_{t-1} \sim Poisson(\lambda_t)$, implies that

$$P(Y_{t=y}|F_{t-1}) = \frac{\lambda_t^y \exp(-\lambda_t)}{y!}, \quad y = 0, 1, \ldots \quad (5)$$

It holds $\text{var}(Y_t|F_{t-1}) = E(Y_t|F_{t-1}) = \lambda_t$.

The negative binomial distribution allows for overdispersion, where variance to be larger than the mean $\lambda_t$. Following [18], Model (3) together with the negative binomial assumption, i.e., $Y_t|F_{t-1} \sim NegBin(\lambda_t, \phi)$, implies that

$$P(Y_{t=y}|F_{t-1}) = \frac{\Gamma(\phi + y)\Gamma(\frac{\phi}{\phi + \lambda_t})}{\Gamma(\phi + y + 1)\Gamma(\phi)} \left( \frac{\phi}{\phi + \lambda_t} \right)^{\phi} \left( \frac{\lambda_t}{\phi + \lambda_t} \right)^{\phi}, \quad y = 0, 1, \ldots \quad (6)$$

In this case, $\text{var}(Y_t|F_{t-1}) = \lambda_t + \lambda_t^2/\phi$. The Poisson distribution is a limiting case of the negative binomial when $\phi \rightarrow \infty$.

Fokianos and Fried in [14,15], an intervention model can be written as:

$$g(\lambda_t) = \beta_0 + \sum_{k=1}^{p} \beta_k \tilde{g}(Y_{t-k}) + \sum_{l=1}^{q} \alpha_l g(\lambda_{t-l}) + \sum_{m=1}^{s} \omega_m \delta_{m-t} \tau_m (t \geq \tau_m), \quad (7)$$

Model interventions affecting the location by including a deterministic covariate of the form $\delta^{t-t_1} (t \geq \tau)$, where $\tau$ is the time of occurrence and the decay rate $\delta$ is a known constant. This covers various types of interventions for different choices of the constant $\delta$: a singular effect for $\delta = 0$ (spiky outlier), an exponentially decaying change in location for $\delta \in (0,1)$ (transient shift) and a permanent change of location for $\delta = 1$ (level shift).
2.3. Model Selection
The model selection is conducted using out-of-sample criteria by comparing the RMSE (Root Mean Square Error), MAD (Mean absolute deviation), and MAPE (Mean Absolute Percentage Error). RMSE, MAD, and MAPE are defined as follows [17] [21]:

\[
RMSE = \sqrt{\frac{1}{M} \sum_{i=1}^{M} (Y_{n+l} - \hat{Y}_n(l))^2}
\]

(8)

\[
MAD = \frac{1}{M} \sum_{i=1}^{M} |Y_{n+l} - \hat{Y}_n(l)|
\]

(9)

\[
MAPE = \frac{1}{M} \sum_{i=1}^{M} \left| \frac{Y_{n+l} - \hat{Y}_n(l)}{Y_{n+l}} \right| \times 100
\]

(10)

where \( M \) is the number of predictions performed, \( Z_{n+l} \) is the actual data, and \( \hat{Z}_n(l) \) is the forecast data in the next step \( l \).

3. Methodology
The data employed in this research is secondary data, namely international visitor arrivals to Indonesia, 2001-2018. Data has been published by Statistics Indonesia (BPS).

The first step is to divide the data into training and testing data. Training data is data from 2001 to 2017, 204 data. The testing data is data during 2018 as many as 12 data. In this research, we applied intervention on ARIMA models and intervention on count time series models. Intervention on the ARIMA model includes the following stages:
1. Identify stationary.
2. Transform or differencing if the data is not stationary.
3. Dividing the dataset into \( k + l \) parts, where \( k \) is the number of intervention.
4. Modeling of the first intervention.
5. Modeling of the \( m^{th} \) Intervention Model, where \( m = 2, 3, \ldots, k \).

Intervention on the count time series model includes the following stages:
1. Identify stationary.
2. Differencing if the data is not stationary.
3. Modeling using Poisson AR and Negative Binomial AR model. The model is built using intervention effects based on the results of the intervention in the ARIMA model.

Then calculate RMSE, MAD, and MAPE on out-of-sample data. Furthermore, compare RMSE, MAD, and MAPE to obtain the best model. Last, forecast monthly international visitor arrivals to Indonesia, 2019 based on the best model.

4. Results and Discussion
4.1. Intervention on ARIMA
4.1.1. Modeling International Visitor Arrival from Bahrain. We applied ARIMA modeling on pre-intervention data, the second Bali bombings. The 2005 Bali bombings were a series of terrorist suicide bomb and a series of car bombs and attacks that occurred on 1 October 2005, in Bali, Indonesia. Based on the analysis of data stationarity, the data is not stationary in the variance. So that the transformation of natural logarithms (ln) is provided. The mean is not stationary, either. Regular differencing order 1 and seasonal differencing 12-periods order 1 are applied for this data.
Figure 1 shows that data on Bahrainian foreign tourist visits has decreased since the explosion of the second Bali Bombings. The bomb caused the number of international visitor arrival from Bahrain decreased and the residuals from forecasting data exited $\pm 2\sigma$. Order of the intervention model is determined, $b = 3$, $s = 6$, $r = 0$, the function used is the pulse function.

After the second Bali bombings intervention model is obtained, the data will be modeled using the intervention of Sarinah Bomb, that occurred on January 2015 or at $T = 121$. Figure 2 shows that data has decreased at $T+5$ since the Sarinah bomb blast. Order of intervention models is determined, namely $b = 5$, $s = 0$, $r = 0$.

Table 1 shows the parameter estimation for Bahrain’s visitors intervention model. $P$-value is already under 5 percent; all parameter is significant.

| Parameter | Estimate | Standard Error | T value | p-value |
|-----------|----------|----------------|---------|---------|
| $\theta_1$ | 0.76     | 0.05           | 14.18   | $< 0.001$ |
| $\Phi_1$  | -0.47    | 0.07           | -6.7    | $< 0.001$ |
The final multi-input intervention on the ARIMA model for the number of tourist arrivals from Bahrain can be written as:

\[
\begin{align*}
\omega_{10} &= -2.03 \quad & 0.48 \quad & -4.24 \quad & < 0.001 \\
\omega_{11} &= 2.24 \quad & 0.47 \quad & 4.81 \quad & < 0.001 \\
\omega_{12} &= 0.94 \quad & 0.46 \quad & 2.03 \quad & 0.044 \\
\omega_{13} &= 1.88 \quad & 0.45 \quad & 4.18 \quad & < 0.001 \\
\omega_{20} &= -1.36 \quad & 0.48 \quad & -2.83 \quad & 0.005 \\
\end{align*}
\]

The final multi-input intervention on the ARIMA model for the number of tourist arrivals from Bahrain can be written as:

\[
\ln Y_t = -2.03 P_{1,t-3} - 2.24 P_{1,t-4} - 0.94 P_{1,t-5} - 1.88 P_{1,t-7} - 1.36 P_{2,t-5}
\]

\[
+ \frac{(1 + 0.76 B)}{(1 - B)(1 - B^{12})(1 + 0.47 B^{12})} a_t
\]

where \( P_{1,t} \) is the pulse function of the second Bali bombing, \( P_{2,t} \) is the pulse function of the Sarinah bombing.

4.1.2. Modeling International Visitor Arrival from Singapore. ARIMA model was performed on pre-intervention data, before closure of gambling locations in Batam, \( T = 1 \) to \( T = 53 \). The gambling locations were closed after the inauguration of National Police Chief Sutanto in June 2005. The data had constant variance, but not stationary in the mean. We applied differencing to obtain a stationary time series. Regular differencing order 1 and seasonal differencing 12-periods order 1 are applied for the data.

Figure 3. Plot of the intervention effect, closure of gambling locations.

Figure 3 shown foreign tourist visits from Singapore has decreased since the closure of gambling locations in Batam in June 2005. Based on residual, the order of intervention models is \( b = 0, s = 6, r = 0 \). After that, the intervention of sharia tourism promotion in October 2013 applied to the data. Foreign tourist visits from Singapore has increased. The intervention occurred at \( T = 154 \). It was assumed to be a pulse function. Based on the residual, the order of intervention models is \( b = 3, s = 0, r = 0 \). Table 2 shows the parameter estimation for the intervention model. \( P\)-value is under 5 percent, all parameter is significant.

The final multi-input intervention on the ARIMA model for the number of tourist arrivals from Singapore can be written as:
\[ Y_t = -57267.00 \ p_{1,t} - 48219.40 \ p_{1,t-1} - 46583.90 \ p_{1,t-2} - 32110.30 \ p_{1,t-3} \]
\[ - 34541.50 \ p_{1,t-4} - 33813.30 \ p_{1,t-5} - 52154.40 \ p_{1,t-6} + 21343.50 \ p_{2,t-3} \]
\[ + \frac{(1 - 0.66 B - 0.16 B^{13} - 0.16 B^{25})(1 - 0.49 B^{12})}{(1 - B)(1 - B^{12})(1 + 0.18 B^{12})} a_t \]

where \( p_{1,t} \) is the pulse function of the closure of gambling locations in Batam, \( p_{2,t} \) is the pulse function of the sharia tourism promotion in Indonesia.

### Table 2. Estimation For Singapore Visitors Intervention Model on ARIMA

| Parameter | Estimate | Standard Error | T value | p-value |
|-----------|----------|----------------|---------|---------|
| \( \theta_1 \) | 0.66 | 0.06 | 10.84 | < 0.001 |
| \( \theta_2 \) | 0.16 | 0.06 | 2.6 | 0.010 |
| \( \theta_3 \) | 0.16 | 0.06 | 2.64 | 0.009 |
| \( \Theta_1 \) | 0.49 | 0.07 | 6.71 | < 0.001 |
| \( \Theta_0 \) | -0.18 | 0.08 | -2.16 | 0.032 |
| \( \omega_{10} \) | -57267.00 | 10000.50 | -5.73 | < 0.001 |
| \( \omega_{11} \) | 48219.40 | 10268.60 | 4.70 | < 0.001 |
| \( \omega_{12} \) | 46583.90 | 9983.40 | 4.67 | < 0.001 |
| \( \omega_{13} \) | 32110.30 | 9969.00 | 3.22 | < 0.001 |
| \( \omega_{14} \) | 34541.50 | 9980.70 | 3.46 | < 0.001 |
| \( \omega_{15} \) | 33813.30 | 10015.70 | 3.38 | < 0.001 |
| \( \omega_{16} \) | 52154.40 | 9657.30 | 5.40 | < 0.001 |
| \( \omega_{20} \) | 21343.50 | 9196.80 | 2.32 | 0.022 |

### 4.2. Intervention on Count Time Series Model

#### 4.2.1. Modeling International Visitor Arrival from Bahrain

Regular differencing order 1 and seasonal differencing 12-periods order 1 are applied for this data. On the count time series model, the data must be positive. Based on differencing results, the smallest value is \(-214\), so that all data is added with constants \( c = 215 \).

We applied Poisson AR and Negative Binominal AR model. And used the same intervention order as in the ARIMA. Six tentative count time series models have considered, namely Poisson-AR(1,1), NB-AR(1,1), Poisson-AR([1,2,3],[1,12]), NB-AR([1,2,3],[1,12]), Poisson-AR([1,12],1), NB-AR([1,12],1). Table 3 shows the parameter estimation for the intervention model on poisson-AR([1,12],1), NB-AR([1,12],1). On Negative Binomial AR, degree of overdispersion is 0.056.

### Table 3. Parameter Estimation For The Intervention Model on Poisson-AR([1,12],1), NB-AR([1,12],1)

| Parameter | Estimate | Standard Error | T value | p-value |
|-----------|----------|----------------|---------|---------|
| Intercept | 5.26     | 5.26           | 0.16    | 0.68    | 32.42 | 7.78 | < 0.001* | < 0.001* |
| \( \beta_1 \) | -0.55    | -0.55          | 0.02    | 0.07    | -29.71 | -7.52 | < 0.001* | < 0.001* |
| \( \beta_{12} \) | -0.06    | -0.06          | 0.01    | 0.04    | -6.68  | -1.45 | < 0.001* | 0.150 |
| \( \alpha_1 \) | 0.63     | 0.63           | 0.03    | 0.10    | 24.00  | 6.24  | < 0.001* | < 0.001* |
| \( \omega_{10} \) | -5.05    | -5.05          | 1.00    | 1.03    | -5.05  | -4.91 | < 0.001* | < 0.001* |
| \( \omega_{11} \) | 0.80     | 0.80           | 0.65    | 0.91    | 1.22   | 0.87  | 0.225 | 0.384 |
| \( \omega_{12} \) | 0.08     | 0.08           | 0.07    | 0.26    | 1.17   | 0.32  | 0.242 | 0.746 |
\( \omega_{13} \), -0.24
\( \omega_{20} \), -0.09

\[ \begin{array}{cccccc}
\omega_{13} & -0.24 & -0.24 & 0.06 & 0.22 & -3.84 & -1.10 < \alpha = 5% \n\omega_{20} & -0.09 & -0.09 & 0.06 & 0.20 & -1.66 & -0.46 0.100 0.645
\end{array} \]

The final multi-input intervention on count time series model for the number of tourist arrivals from Bahrain can be written as:
\[
\log(\lambda_t) = 5.26 - 0.55Y_{t-1} - 0.06Y_{t-12} + 0.63\lambda_{t-1} - 5.05P_{1,t-3} + 0.8 P_{1,t-4} + 0.08 P_{1,t-5} - 0.24 P_{1,t-7} - 0.09 P_{2,t-5}
\]

where \( P_{1,t} \) is pulse function of the second Bali bombing, \( P_{2,t} \) is pulse function of the Sarinah bombing.

4.2.2. Modeling International Visitor Arrival from Singapore. Differencing regular order 1 and the differencing seasonal 12 order 1 is applied for the data. The smallest value is −63222, to make positive data, all data are added with constants \( c = 63223 \).

We considered six tentative models, namely Poisson-\( AR(1,2,12) \), \( NB-AR(1,2,12) \), Poisson-\( AR(2,1,12,13,25) \), \( NB-AR(2,1,12,13,25) \), Poisson-\( AR(2,1,12) \), \( NB-AR(2,1,12) \). Table 4 shows the parameter estimation for the intervention model on Poisson-\( AR(2,1,12) \), \( NB-AR(2,1,12) \), degree of overdispersion is 0.0697.

| Parameter | Estimate | Standard Error | T value | p-value |
|-----------|----------|----------------|---------|---------|
| \( \beta_2 \) | 18.99 | 0.027 | 1.71 | 11.10 | < 0.001* |
| \( \alpha_1 \) | -0.49 | 0.002 | 0.15 | -72.24 | -1.05 | < 0.001* |
| \( \alpha_2 \) | -0.21 | 0.10 | -145.76 | -3.20 | < 0.001* | 0.002* |
| \( \omega_{10} \) | -1.05 | 0.28 | -166.13 | -3.76 | < 0.001* | 0.035* |
| \( \omega_{11} \) | -0.36 | 0.35 | -64.23 | -0.10 | < 0.001* | 0.303 |
| \( \omega_{12} \) | -0.05 | 0.29 | -11.02 | -0.17 | < 0.001* | 0.864 |
| \( \omega_{13} \) | 0.12 | 0.29 | 29.27 | 0.42 | < 0.001* | 0.676 |
| \( \omega_{14} \) | 0.04 | 0.29 | 10.08 | 0.15 | < 0.001* | 0.880 |
| \( \omega_{15} \) | 0.24 | 0.29 | 59.28 | 0.82 | < 0.001* | 0.415 |
| \( \omega_{16} \) | -0.89 | 0.34 | -151.03 | -2.65 | < 0.001* | 0.009* |
| \( \omega_{20} \) | 0.46 | 0.23 | 160.99 | 2.02 | < 0.001* | 0.045* |

* significant parameter with \( \alpha = 5% \)

The final multi-input intervention on count time series model for the number of tourist arrivals from Singapore can be written as:
\[
\log(\lambda_t) = 18.99 - 0.02Y_{t-2} - 0.49\lambda_{t-1} - 0.21\lambda_{t-2} - 1.05P_{1,t} - 1.03 P_{1,t-1} - 0.05 P_{1,t-3} + 0.12 P_{1,t-4} + 0.04 P_{1,t-5} + 0.24 P_{1,t-6} + 0.89 P_{1,t-7} + 0.46 P_{2,t-3}
\]

where \( P_{1,t} \) is pulse function of the closure of gambling locations in Batam, \( P_{2,t} \) is pulse function of the sharia tourism promotion in Indonesia.
4.3. Selection of The Best Model

RMSE, MAD, and MAPE on out-of-sample data based on 3 models above have been calculated. We compared the value to obtain the best model. On Bahrain data, the intervention functions were pulse function of the second Bali bombing and pulse function of the Sarinah bombing. And the best intervention on count time series model were Poisson-AR([1,12],1) and NB-AR([1,2,3],[1,12]). On Singapore data, the best model were Poisson-AR(2,[1,12]) and NB-AR(2,[1,12,13,25]). And The intervention functions were pulse function of the closure of gambling locations in Batam and pulse function of the sharia tourism promotion in Indonesia.

Summary of RMSE, MAD, and MAPE (%) shown in Figure 4. It was concluded that intervention on Poisson AR is the best model to forecast international visitor from Bahrain and Singapore.

Figure 4. RMSE, MAD, and MAPE comparison (a) Bahrain’s Tourist Data and (b) Singapore’s Tourist Data.

4.4. Forecast

The best model to forecast international visitor from Bahrain and Singapore is an intervention model on Poisson-AR([1,12],1) and Poisson-AR(2,[1,12]), respectively. The forecast provides count data. Forecast of international visitors from Bahrain and Singapore, 2019 can be seen in Figure 5.

Figure 5. Time Series Plot of Forecast International Visitors, 2019 (a) Bahrain’s Tourist and (b) Singapore’s Tourist.

5. Conclusion

This research performed an analysis of intervention on ARIMA and count time series model to forecast monthly international visitor arrivals from Bahrain and Singapore to Indonesia. We used the same intervention order as in the ARIMA. Based on RMSE, MAD, and MAPE on out-of-sample data,
it was concluded that intervention on Poisson AR is the best model to forecast international visitor from Bahrain and Singapore. The best intervention model used to forecast international visitors from Bahrain is intervention model on Poisson-AR([[1,12],1]). The intervention functions were pulse function of the second Bali bombing and pulse function of the Sarinah bombing. Intervention model on Poisson-AR(2,[[1,12]]) was used to forecast international visitors from Singapore, 2019. The intervention functions were pulse function of the closure of gambling locations in Batam and pulse function of the sharia tourism promotion in Indonesia.

References
[1] Badan Pusat Statistik (BPS). (2019). Berita Resmi Statistik Perkembangan Pariwisata dan Transportasi Nasional, Desember 2018. Jakarta: Badan Pusat Statistik (BPS).
[2] Box GEP, Tiao GC (1975). “Intervention Analysis with Applications to Economic and Environmental Problems.” Journal of the American Statistical Association, 70(349), 70–79. doi:10.2307/2285379.
[3] Cho, V. (2001). Tourism Forecasting and Its Relationship with Leading Economic Indicators. Hospitality and Tourism Research, 25(4): 399-420.
[4] Rukini, R., Arini, P.S., and Nawangsih, E. (2015). Peramalan Jumlah Kunjungan Wisatawan Mancanegara (Wisman) Ke Bali Tahun 2019: Metode ARIMA. Jurnal Ekonomi Kuantitatif Terapan, [S.I.], nov. 2015.
[5] Ma, E., Liu, Y., Li, J., and Chen, S. (2016). Anticipating Chinese tourists arrivals in Australia: A time series analysis. Tourism Management Perspectives, 17, 50-58. doi.org/10.1016/j.tmp.2015.12.004.
[6] Loganathan N., and Yahaya I. (2010). Forecasting International Tourism Demand in Malaysia using Box-Jenkins Sarima Applications. South Asian Journal of Tourism and Heritage, 3(2): 50-60.
[7] Hassani, H., Silva, E.S., Antonakakis, N., Filis, G., and Gupta, R. (2017), Forecasting Accuracy Evaluation of Tourist Arrivals, Elsevier : Annals of Tourism Research, 63, 112-127.
[8] Lee, M.H., Suhartono, and Sanugi, B. (2010). Multi Input Intervention Model for Evaluating the Impact of the Asian Crisis and Terrorist Attacks on Tourist Arrivals. MATEMATIKA, 26(1), 83–106.
[9] Liboschik, T., Fokianos, K., and Fried, R. (2017) tcount: An R package for Analysis of Count Time Series Following Generalized Linearized Models. Journal of Statistical Software, 82(5), 1–51, http://dx.doi.org/10.18637/jss.v082.i05.
[10] Fokianos, K. (2011). “Some Recent Progress in Count Time Series.” Statistics, 45(1), 49–58. doi:10.1080/02331888.2010.541250.
[11] Fokianos, K. and Tjostheim, D. (2011). Log-Linear Poisson Autoregression. Journal of Multivariate Analysis, 102(3), 563–578. doi:10.1016/j.jmaa.2010.11.002.
[12] Christou, V. and Fokianos, K. (2014). Quasi-Likelihood Inference for Negative Binomial Time Series Models. Journal of Time Series Analysis, 35(1), 55–78.
[13] Asrirawan, Kuswanto, H., and Suhartono. (2014). Model Generalized Seasonal Autoregressive Integrated Moving Average (GSARIMA) untuk Peramalan Jumlah Penderita Demam Berdarah Dengue (DBD) di Kota Surabaya. Seminar nasional Sains dan Pendidikan MIPA (SEMIRATA) 2014. Bogor: FMIPA-IPB.
[14] Fokianos, K. and Fried, R. (2010). Interventions in INGARCH Processes. Journal of Time Series Analysis, 31(3), 210–225. doi: 10.1111/j.1467-9892.2010.00657.x.
[15] Fokianos, K. and Fried, R. (2012). Interventions in Log-Linear Poisson Autoregression. Statistical Modelling, 12(4), 299–322. doi:10.1177/1471082x1201200401.
[16] Fried, R., Liboschik, T., Elsaied, H., Kitromilidou, S., and Fokianos, K. (2014). On Outliers and Interventions in Count Time Series Following GLMs. Austrian Journal of Statistics, 43(3), 181–193. doi:10.17713/ajs.v43i3.30.
[17] Wei, W.W. (2006). Time Series Analysis Univariate and Multivariate Methods. 2nd edition. USA: Pearson Education Inc.

[18] Christou, V. and Fokianos, K. (2013). On Count Time Series Prediction. Journal of Statistical Computation and Simulation, 85(2), 357–373. doi:10.1080/00949655.2013.823612.

[19] Gneiting, T., Balabdaoui, F., and Raftery, A.E. (2007) Probabilistic Forecasts, Calibration and Sharpness. Journal of the Royal Statistical Society B, 69 (2), 243–268. doi:10.1111/j.1467-9868.2007.00587.x.

[20] Czado C, Gneiting T, Held L (2009). “Predictive Model Assessment for Count Data.” Biometrics, 65(4), 1254–1261. doi:10.1111/j.1541-0420.2009.01191.x.

[21] Gooijer, J.G.D., and Hyndman, R.J. (2006). 25 Years of Time Series Forecasting. International Journal of Forecasting, 22, 443 – 473.