

Deteksi Pencilan pada Model ARIMA dengan *Bayesian Information Criterion (BIC)* Termodifikasi

SELMA YULISTIANI¹, SULIADI²

^{1,2}Program Studi Statistika, Universitas Islam Bandung
email: ¹selma.yulistiani@gmail.com, ²suliadi@gmail.com

ABSTRAK

Data time series mungkin dipengaruhi oleh kejadian atau keadaan khusus seperti adanya promosi, bencana alam, dan sebagainya. Peristiwa tersebut dapat menimbulkan pengamatan yang tidak konsisten pada seriesnya yang disebut pencilan. Karena pencilan dapat membuat kesimpulan yang tidak valid, maka penting untuk melakukan prosedur dalam mendeteksi efek pencilan. Dalam deteksi pencilan terdapat salah satu tipe pencilan yaitu *additive outlier*(AO). Proses pendeteksian pencilan aditif pada model ARIMA dapat dikatakan sebagai masalah pemilihan model, dimana kandidat model mengasumsikan adanya pencilan aditif pada waktu tertentu. Dalam pemilihan model ada kriteria yang harus diperhatikan agar dapat menghasilkan model terbaik. Kriteria pemilihan model yang baik dapat menggunakan *Bayesian Information Criterion (BIC)* yang diturunkan oleh Schwarz (1978). Galeano dan Pena (2011) mengusulkan *Bayesian Information Criterion* termodifikasi untuk pemilihan model dan mendeteksi pencilan yang berpotensi. *Bayesian Information Criterion* termodifikasi untuk deteksi pencilan akan diaplikasikan pada data Outstanding Loan PT. Pegadaian Cabang Cimahi tahun 2013-2017. Sehingga diperoleh model terbaik yaitu model dengan menambahkan 2 pencilan potensial dengan model ARIMA (1,0,0), yaitu pencilan pada pengamatan ke 48, dan 58 karena memiliki nilai BICUP yang minimum yaitu 1064.95650.

Kata Kunci: ARIMA, Bayesian Information Criterion termodifikasi, Prosedur Deteksi Pencilan Potensial, additive outlier.

ABSTRACT

Time series data may be affected by special events or circumstances such as promotions, natural disasters, etc. These events can lead to inconsistent observations in the series called outliers. Because outliers can make invalid conclusions, it is important to carry out procedures in detecting outlier effects. In outlier detection there is one type of outlier, namely additive outlier (AO). The process of detecting additive outliers in the ARIMA model can be said as a model selection problem, where the candidate model assumes additive outliers at a certain time. In the selection of models there are criteria that must be considered in order to produce the best model. The good criteria for models selection can use the Bayesian Information Criterion (BIC) derived by Schwarz (1978). Galeano and Pena (2011) proposed a modified Bayesian Information Criterion for model selection and detect potential outliers. The modified Bayesian Information Criterion for outlier detection will be applied to the data Outstanding Loan PT. Pegadaian Cimahi year 2013-2017. So that the best model is obtained that the model with adding 2 potential outliers with the ARIMA model (1,0,0), that outliers at observations 48, and 58 because it has a minimum BICUP value of 1064.95650.

Keywords: ARIMA, Modified Bayesian Information Criterion, Procedure Detection Potential Outliers, additive outlier.

1. PENDAHULUAN

Data *time series* merupakan serangkaian data pengamatan yang terjadi berdasarkan indeks waktu t secara berurutan dengan interval waktu yang tetap. Dalam prakteknya, data *time series* mungkin dipengaruhi oleh kesalahan atau gangguan yang akan menimbulkan efek palsu dan menghasilkan pola yang tidak sesuai dengan pengamatan di serinya. Pengamatan yang tidak biasa ini dapat disebut pencilan. Pencilan biasanya di pengaruhi oleh peristiwa seperti

adanya promosi dan bencana alam. Karena pencilan dapat membuat kesimpulan yang dihasilkan tidak valid, maka penting untuk memiliki prosedur untuk mendeteksi efek pencilan tersebut yaitu dengan menggunakan deteksi pencilan. Deteksi pencilan dalam *time series* menjadi masalah penting karena pencilan dapat memberikan efek pada analisis. Sebagian besar metode didasarkan pada prosedur deteksi berurutan. Ketika pencilan pertama ditemukan, ukuran pencilan diprediksi, kemudian efek pencilan tersebut dibersihkan dari seri dan mencari kembali pencilan yang baru. Akan tetapi, seperti Sanchez dan Pena (2003) tunjukkan bahwa prosedur deteksi berurutan memiliki tiga kelemahan yaitu: nilai perkiraan dari parameter akan bersifat bias, dalam banyak situasi distribusi statistik uji tidak diketahui dan nilai kritis yang diperlukan untuk menerapkan *test* harus diperkirakan melalui simulasi untuk ukuran dan model sampel, serta adanya efek *masking* dan efek *swamping*. Efek *masking* yang berarti pencilan tidak terdeteksi karena adanya pengaruh data lain dan efek *swamping* yang berarti pencilan mempengaruhi data lain.

Dalam deteksi pencilan pada data *time series*, terdapat salah satu tipe pencilan yaitu *Additive Outlier* (AO). Proses pendeteksian pencilan aditif pada model ARIMA dapat dikatakan sebagai masalah pemilihan model, dimana kandidat model mengasumsikan adanya pencilan aditif pada waktu tertentu. Dalam pemilihan model ada kriteria yang harus diperhatikan ketika memilih serangkaian kandidat model. Kriteria model yang baik mampu menghasilkan model yang diharapkan dapat memprediksi observasi baru. Adapun kriteria informasi yang sering digunakan yaitu *Aikake Information Criterion* (AIC) yang diusulkan oleh Aikake (1973). Akan tetapi, AIC hanya akan bekerja optimal dalam sampel kecil. Sedangkan kriteria yang lebih baik yaitu *Bayesian Information Criterion* (BIC) yang diturunkan oleh Schwarz (1978). Kriteria Bayesian dirancang untuk memiliki tingkat kebenaran yang konsisten terhadap model.

Pedro Galeano dan Daniel Pena (2011) mengusulkan kriteria untuk pemilihan model yang diusulkan yaitu *Bayesian Information Criterion* termodifikasi yang dapat juga mendeteksi pencilan potensial. Mendeteksi pencilan berdasarkan lokasi pencilan dan mencari pencilan yang mempengaruhi data lain. Pencilan yang terdeteksi dimasukkan ke dalam model dan memilih model terbaik berdasarkan *Bayesian Information Criterion* termodifikasi yang minimum. Kemudian *Bayesian Information Criterion* termodifikasi yang dapat juga mendeteksi pencilan potensial akan diaplikasikan untuk memodelkan *Out Standing Loan* (OSL) emas logam mulia PT. Pegadaian (Persero) cabang Cimahi pada tahun 2013-2017.

Berdasarkan latar belakang yang telah diuraikan, maka perumusan masalah dalam penelitian ini sebagai berikut:

1. Bagaimana prosedur mendeteksi pencilan yang berpotensi dengan *bayesian information criterion* termodifikasi?
2. Bagaimana hasil prosedur mendeteksi pencilan yang berpotensi dengan *bayesian information criterion* termodifikasi pada data OSL emas logam mulia PT. Pegadaian cabang Cimahi tahun 2013-2017?

Selanjutnya, tujuan dalam penelitian ini diuraikan dalam pokok-pokok sebagai berikut:

1. Untuk mengetahui prosedur mendeteksi pencilan yang berpotensi dengan *bayesian information criterion* termodifikasi.

Untuk mengetahui hasil prosedur mendeteksi pencilan yang berpotensi dengan *bayesian information criterion* termodifikasi pada data OSL emas logam mulia PT. Pegadaian cabang Cimahi tahun 2013-2017.

2. LANDASAN TEORI

Analisis deret waktu (*time series*) adalah analisis yang menerangkan dan mengukur berbagai perubahan atau perkembangan data selama satu periode (Hasan, 2003). Salah satu metode analisis *time series* adalah *Autoregressive Integrated Moving Average* (ARIMA). Box dan Jenkins adalah orang yang memperkenalkan singkatan ARIMA pada tahun 1970. Dalam pemodelan ARIMA terdapat karakteristik yang harus terpenuhi yaitu kestasioneran data. Data *time series* dikatakan stasioner jika memiliki rata-rata dan varians yang konstan untuk semua t . Data dikatakan tidak stasioner dalam varians ketika plot data menunjukkan pola melebar atau pola terompot Untuk mengatasi ketidakstasioneran terhadap varians dilakukan transformasi dengan melihat plot Box-Cox. Jika *Lower CL* dan *Upper CL* mencakup 1 maka data sudah stasioner. Persamaan dari transformasi ini adalah (Wei, 2006):

$$T(Y_t) = \frac{X_t^{\lambda}-1}{\lambda} \tag{1}$$

Jika transformasi kuasa ini dihubungkan dengan bentuk transformasi stabilitas varians yang lain, maka diperoleh tabel kesetaraan seperti dibawah ini:

Tabel 1. Hubungan Nilai dengan Kesetaraan Transformasi Stabilitas Varians

λ	Transformasi
-1	$\frac{1}{X_t}$
-0,5	$\frac{1}{\sqrt{X_t}}$
0	LN X_t
0,5	$\sqrt{X_t}$
1	X_t (tidak ada transformasi)

Sedangkan data dikatakan tidak stasioner dalam rata-rata ketika data bergerak terus naik yang membentuk trend. Untuk melihat apakah data stasioner dalam rata-rata dengan melihat plot ACF dan PACF. Mengatasi ketidakstasioneran terhadap rata-rata dilakukan proses *differencing*. Persamaan dari *differencing* ini adalah:

$$X_t^d = (1 - B)^d X_t \tag{2}$$

Secara umum model ARIMA ditulis dengan ARIMA (p, d, q) yang artinya model ARIMA dengan derajat AR(p), derajat pembedaan d, dan derajat MA(q). Model *Autoregressive* (AR) merupakan suatu model persamaan regresi yang menghubungkan nilai-nilai sebelumnya dari suatu variabel *dependent* (tak bebas) dengan variabel itu sendiri. Bentuk umum model AR(p) adalah:

$$X_t = \mu + \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + e_t \tag{3}$$

Model *Moving Average* (MA) merupakan suatu model persamaan regresi yang menghubungkan nilai residual sebelumnya dari suatu variabel *dependent* (tak bebas) dengan residual. Bentuk umum model MA(q) adalah:

$$X_t = \mu + e_t - \theta_1 e_{t-1} - \dots - \theta_q e_{t-q} \tag{4}$$

Jika *d* adalah bilangan bulat non negatif, maka $\{X_t\}$ dikatakan proses ARIMA dengan $(1 - B)^d X_t$. Maka persamaan berarti bahwa model ARIMA (p,d,q) adalah:

$$X_t = \frac{\mu + \theta_q(B)e_t}{\phi_p(B)(1 - B)^d} \tag{5}$$

Model untuk musiman dalam ARIMA:

$$\Phi_P(B^s)\phi_p(B)(1 - B^s)^D(1 - B)^d X_t = \mu + \Theta_Q(B^s)\theta_q(B)e_t \tag{6}$$

Setelah mendapatkan model ARIMA sementara maka dilakukan pengujian parameter yaitu:

Hipotesis:

AR(*Autoregressive*)

$$H_0: \phi_i = 0, \text{ dimana } i=1,2,\dots,p \text{ (AR tidak signifikan dalam model)}$$

$H_1: \phi_i \neq 0$, dimana $i=1,2,\dots,p$ (AR signifikan dalam model)

MA(Moving Average)

$H_0: \theta_i = 0$, dimana $i=1,2,\dots,q$ (MA tidak signifikan dalam model)

$H_1: \theta_i \neq 0$, dimana $i=1,2,\dots,q$ (MA signifikan dalam model)

Statistik uji:

$$t_{hitung} AR = \frac{\hat{\phi}_i}{SE(\hat{\phi}_i)} \quad (7)$$

$$t_{hitung} MA = \frac{\hat{\theta}_i}{SE(\hat{\theta}_i)} \quad (8)$$

Kriteria keputusan yang digunakan untuk menolak H_0 adalah jika $|t| > t_{\frac{\alpha}{n}, df}$, $df=n-p$ atau p -

value $< \alpha$. Kemudian diagnosis model hipotesisnya adalah:

H_0 : Model sudah memenuhi syarat cukup

H_1 : Model belum memadai

Statistik uji:

$$Q = T(T+2) \sum_{k=1}^k \frac{r_k^2}{T-k} \quad (9)$$

Kriteria pengujian:

Tolak H_0 jika $Q > \lambda^2_{(\alpha; k-p-q)}$ atau p -value $< \alpha$. Jika terdapat dua atau lebih model ARIMA maka dilakukan pemilihan kriteria model AIC yang minimum. Rumus untuk memperoleh nilai AIC (Wei,2006):

$$AIC = T \ln \hat{\sigma}^2 + 2M \quad (10)$$

Dalam data *time series* terdapat juga metode deteksi pencilan. Pencilan adalah data pengamatan yang tidak konsisten pada seriesnya. Biasanya pencilan tidak diketahui waktu penyebabnya. Efek pencilan perlu diatasi yaitu dengan deteksi pencilan. Salah satu tipe pencilan yaitu *Additive Outlier* (AO). *Additive Outlier* (AO) adalah suatu jenis pencilan yang hanya pada satu periode saja. Bentuk umum sebuah *Additive Outlier* dalam proses ARMA dapat dinyatakan sebagai berikut:

$$Z_t = \begin{cases} X_t & t \neq t^0 \\ X_t + \omega & t = t^0 \end{cases}$$

$$Z_t = X_t + \omega I_t(t^0)$$

$$Z_t = \frac{\theta(B)}{\phi(B)} e_t + \omega I_t(t^0) \quad (11)$$

dimana X_t dapat juga dinyatakan sebagai model ARIMA (4) dan (5). Jika data yang diamati mengandung campuran m pencilan aditif, maka

$$Z_t = X_t + \omega_{t_1} I_t(t_1^0) + \dots + \omega_{t_m} I_t(t_m^0)$$

Jika data mengandung campuran m pencilan aditif dan ada pengaruh musiman (5) maka:

$$\Phi_p(B^s) \phi_p(B) (1-B^s)^D (1-B)^d (Z_t - \omega_{t_1} I_t(t_1^0) + \dots + \omega_{t_m} I_t(t_m^0)) = \mu + \Theta_Q(B^s) \theta_q(B) e_t \quad (12)$$

Model diatas dinotasikan M_{τ_m} . Parameter model M_{τ_m} dapat diringkas menjadi:

$$\rho_{\tau_m} = (\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q, \Phi_1, \dots, \Phi_p, \Theta_1, \dots, \Theta_Q, \omega_{t_1}, \dots, \omega_{t_m}, \mu, \sigma),$$

dimana $p_m = p + q + P + Q + m + 1 + 1$ dan $\tau_m = t_1, \dots, t_m$. Penduga kemungkinan maksimum ρ_{τ_m} adalah $\hat{\rho}_{\tau_m}$ yang diperoleh dengan memaksimalkan fungsi kemungkinan $\mathcal{L}(\rho_{\tau_m} | Z, M_{\tau_m})$. Untuk data $Z(Z_1, \dots, Z_T)$ dengan vektor parameter dugaan $(\hat{\rho}_{\tau_m})$, banyaknya pencilan aditif (m) dan lokasi pencilan (τ_m) adalah tidak diketahui dan harus diduga dari data.

Bayesian Informasi Criterion termodifikasi untuk deteksi pencilan yaitu BICUP (*Bayesian Information Criterion with Uniform Prior*) yang diusulkan oleh Galeano dan Pena (2011). Dengan kandidat model:

$$\begin{pmatrix} T_1 \\ 0 \end{pmatrix} + \begin{pmatrix} T_1 \\ 1 \end{pmatrix} + \dots + \begin{pmatrix} T_1 \\ T_1 \end{pmatrix}$$

BICUP dapat dinyatakan dalam bentuk:

$$BICUP(M_{\tau_m}) = -2\ell_{\tau_m}(\hat{\rho}_{\tau_m}) + p_m \log T + 2\log \binom{T}{M} \tag{13}$$

dimana, $\hat{\rho}_{\tau_m}$ adalah estimasi kemungkinan maksimum (MLE) dari parameter model, p_m adalah orde dari parameter model, τ_m adalah lokasi pencilan, dan T adalah ukuran data. Odds dari pendekatan BICUP untuk model deteksi pencilan yaitu:

$$o_{BICUP}(M_{\tau_m}, M_{\tau_n} | Z) = \exp\left(\frac{BICUP(M_{\tau_n}) - BICUP(M_{\tau_m})}{2}\right) \tag{14}$$

Dari odds BICUP yang dibuat matriks interaksi dengan elemen

$d_{BICUP}(r, s) = |o_{BICUP}(M_{\tau_1}^{r,s}, M_{\tau_0} | Z) - o_{BICUP}(M_{\tau_1}^r, M_{\tau_0} | Z) o_{BICUP}(M_{\tau_1}^s, M_{\tau_0} | Z)|$ (14) matriks interaksi untuk melihat pencilan mempengaruhi pengamatan lainnya. Prosedur untuk menunjukkan pencilan potensial menurut Galeano dan Pena (2011) sebagai berikut:

1. Menghitung $o_{BICUP}(M_{\tau_1}^r, M_{\tau_0} | Z)$ untuk semua model dengan lokasi pencilan $r=1, \dots, T$. Kemudian,

$$o_{BICUP}(M_{\tau_1}^r, M_{\tau_0} | Z) \geq m(o_{BICUP}) + 3 \times sd(o_{BICUP}) \tag{15}$$

dimana $m(o_{BICUP})$ adalah rata-rata dari $o_{BICUP}(M_{\tau_1}^r, M_{\tau_0} | Z)$ dan $sd(o_{BICUP})$ adalah simpangan baku dari $o_{BICUP}(M_{\tau_1}^r, M_{\tau_0} | Z)$.

2. Menghitung $d_{BICUP}(r, s)$ untuk semua model dengan dua lokasi pencilan $r, s=1, \dots, T$. Kemudian,

$$d_{BICUP}(r, s) \geq m(d_{BICUP}) + 5 \times sd(d_{BICUP}) \tag{16}$$

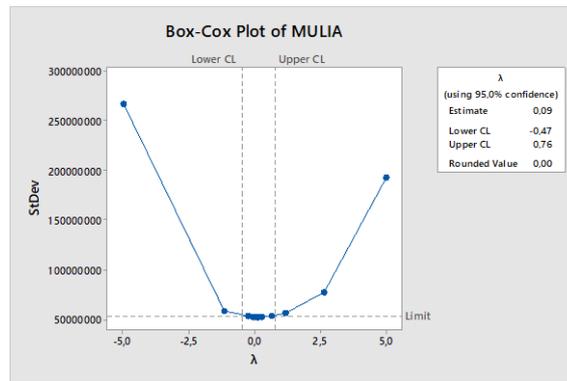
dimana $m(d_{BICUP})$ adalah rata-rata dari $d_{BICUP}(r, s)$ dan $sd(d_{BICUP})$ adalah simpangan baku dari $d_{BICUP}(r, s)$.

Setelah mendapatkan kandidat pencilan, maka kandidat pencilan tersebut akan dimasukkan kedalam model untuk semua $M_{\tau_m}=0, 1, \dots, m$. Model yang menunjukkan nilai BICUP yang minimum adalah model terbaik. Setelah mendapatkan kandidat pencilan, maka kandidat pencilan tersebut akan dimasukkan kedalam model untuk semua $M_{\tau_m}=0, 1, \dots, m$. Model yang menunjukkan nilai BICUP yang minimum adalah model terbaik.

3. HASIL PENELITIAN DAN PEMBAHASAN

Pemodelan ARIMA

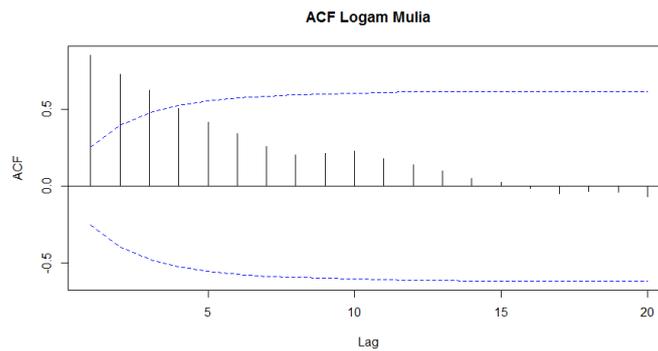
Dengan bantuan Minitab 17 dan R stationer dalam varians yaitu:



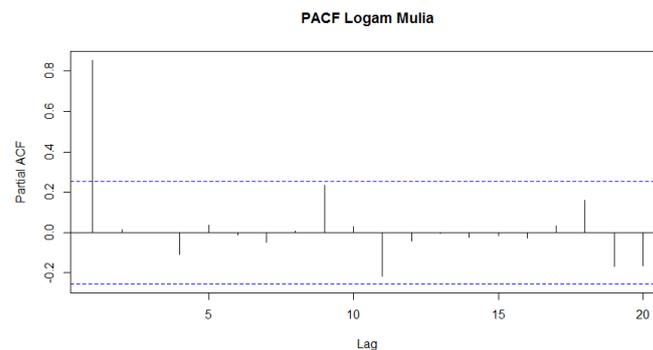
Gambar 1. Plot Box-Cox

Karena nilai *Lower CL* dan *Upper CL* tidak mencakup 1. Maka data perlu ditransformasi dengan formulasi $\sqrt{X_t}$.

Kestasioneran dalam rata-rata akan diselidiki menggunakan plot ACF dan PACF dari data yang telah ditransformasi.



Gambar 2. Plot ACF Data yang Telah Ditransformasi



Gambar 3. Plot PACF Data yang Telah Ditransformasi

Plot ACF dan plot PACF mengindikasikan bahwa data sudah stasioner dalam rata-rata. Hal ini dikarenakan plot PACF memiliki spike 1 dan turun dengan cepat setelah lag p maka data transformasi sudah stationer. Maka model untuk data tersebut yaitu ARIMA (1,0,0).

Tabel 2. Model ARIMA

Model ARIMA	Estimasi Parameter	P-value	AIC
(1,0,0)	$\hat{\phi}_1 = 0.99599$	2.2e-16	1081.49
(1,0,0) with Konstanta	$\hat{\phi}_1 = 0.84238$ $\hat{\mu} = 20800.834$	$\hat{\phi}_1 = 2.2e-16$ $\hat{\mu} = 2.2e-16$	1074.6

Estimasi model ARIMA yang baik digunakan yaitu model ARIMA (1,0,0) with konstanta karena memiliki nilai kriteria pemilihan AIC minimum dan p-value dari model AR(1) yaitu 0 dan konstanta 0 dengan nilai estimasi parameter $\hat{\phi}_1 = 0.84238$ dan $\hat{\mu} = 20800.834$. Menggunakan taraf signifikansi (α) 5%, maka dapat disimpulkan bahwa H_0 ditolak karena p-value < α . Diagnosis model yaitu:

Tabel 3. Hasil uji independensi residul

Ljung-Box	Df	P-value
0.0000032242	1	0.9986

Dengan nilai p-value adalah 0.986 dengan taraf signifikansi (α) 5%, maka dapat disimpulkan bahwa H_0 diterima karena p-value > α . Dari keputusan tersebut dapat disimpulkan bahwa model memenuhi syarat cukup.

Deteksi Pencilan

Maka model (M_{τ_m}) yang didapatkan dengan campuran m pencilan aditif dan ARIMA (1,0,0) yaitu :

$$\phi_1(B)(Z_t - \omega_{t_1} I_t^{(t_1)} - \dots - \omega_{t_m} I_t^{(t_m)}) = \mu + e_t$$

$$\rho_{\tau_m} = (\phi_1, \omega_{t_1}, \dots, \omega_{t_m}, \mu, \sigma)$$

$$p_m = 1 + m + 1 + 1$$

Dengan prosedur model dengan satu pencilan di lokasi r yaitu oods dari $BICUP_{BICUP}(M_{\tau_1}^r, M_{\tau_0} | Z)$ didapatkan pencilan potensial sebagai berikut:

Tabel 4. Pencilan yang Berpotensi

r	$O_{BICUP}(M_{\tau_1}^r, M_{\tau_0} Z)$	Perbandingan
48	10.73676	$10.73676 \geq 0.40748 + 3 \times 1.57439$ $10.73676 \geq 5.13065$ TRUE
58	6.20821	$6.20821 \geq 0.40748 + 3 \times 1.57439$ $6.20821 \geq 5.13065$ TRUE

Tabel 4. menunjukkan bahwa pengamatan pada waktu $t = 48$ dan $t = 58$ dengan nilai $(o_{BICUP}) = 0.40748$ dan $sd(o_{BICUP}) = 1.57439$ dapat dikatakan sebagai pencilan yang berpotensi karena nilai o_{BICUP} jauh lebih besar dari nilai ambang.

Dengan prosedur model dengan dua pencilan di lokasi r dan s yaitu matriks interaksi $d_{BICUP}(r, s)$ didapatkan pencilan potensial sebagai berikut:

Tabel 5. Pencilan yang dipengaruhi pengamatan lain

r,s	$d_{BICUP}(r, s)$	Perbandingan
48,58	175.74058	$175.74058 \geq 0.17596 + 5 \times 4.23779$ $175.74058 \geq 21.36493$ TRUE

Tabel 5. menunjukkan bahwa pengamatan pada $r = 48$ mempengaruhi pengamatan $s = 58$ dengan nilai $m(d_{BICUP}) = 0.17596$ dan nilai $sd(d_{BICUP}) = 4.23779$. Jadi dapat dikatakan bahwa pengamatan tersebut sebagai pencilan yang berpotensi karena mempengaruhi pengamatan lainnya.

Pemilihan Model Terbaik dengan Menggabungkan Pencilan Berpotensi

Dari model dengan satu pencilan dan matriks interaksi, didapatkan kandidat pencilan yang potensial yaitu $\binom{2}{m}$. Untuk pemilihan model terbaik, menggunakan perhitungan BICUP yang minimum dari kandidat model yang menggabungkan pencilan yang berpotensi sebagai berikut:

Tabel 6. Model Terbaik

Lokasi Pencilan	BICUP
0	1075.93765
48	1071.19031
58	1072.28591
48,58	1064.95650

Tabel 6 menunjukkan bahwa model yang terbaik adalah model dengan menambahkan 2 pencilan potensial yaitu pada pengamatan ke-48, dan 58, karena memiliki nilai BICUP yang minimum yaitu 1064.95650. Maka model terbaiknya adalah:

$$Z_t = X_t + \omega_{t48} I_t^{(48)} + \omega_{t58} I_t^{(58)}$$

$$(1 - \phi_1 B)Z_t = \mu + e_t + \omega_{t48} I_t^{(48)} + \omega_{t58} I_t^{(58)}$$

$$Z_t = \mu + \phi_1 Z_{t-1} + e_t + \omega_{t48} I_t^{(48)} + \omega_{t58} I_t^{(58)}$$

$$Z_t = 20690.356 + 0.87486 Z_{t-1} + e_t + 4142.438 I_t^{(48)} + 3944.007 I_t^{(58)}$$

Model terbaik tersebut dapat digunakan untuk memperkirakan data selanjutnya.

4. KESIMPULAN

Berdasarkan hasil analisis deteksi pencilan potensial dengan *Bayesian Information Criterion* termodifikasi dapat disimpulkan:

1. Dari prosedur pemilihan pencilan potensial dengan *Bayesian Information Criterion* termodifikasi maka:
 - a. Dengan menggunakan model satu pencilan yang lokasi pencilan pada pengamatan ke- r yaitu $o_{BICUP}(M_{\tau_1}^r, M_{\tau_0} | Z)$, terdeteksi bahwa pengamatan ke 48 dan 58

dikatakan sebagai potensial pencilan karena nilai $o_{BICUP}(M_{\tau_1}^r, M_{\tau_0} | Z)$ lebih besar dari 5.13065.

- b. Dengan menggunakan matriks interaksi dimana baris lokasi pencilan pada pengamatan ke-r dan kolom lokasi pencilan pada pengamatan ke-s yaitu $d_{BICUP}(r, s)$, terdeteksi bahwa pengamatan ke (48,58). Maka pengamatan tersebut dikatakan sebagai potensial pencilan yang dipengaruhi oleh pengamatan lain karena nilai $d_{BICUP}(r, s)$ lebih besar dari 21.36493.
2. Model terbaik pada data OutStanding Loan PT. Pegadaian Cabang Cimahi tahun 2013-2017 dengan menambahkan 2 pencilan potensial yaitu pencilan pada pengamatan ke 48, dan 58 karena memiliki nilai BICUP yang minimum yaitu 1064.95650. Sehingga model terbaik dengan *Bayesian Information Criterion* termodifikasi dan model ARIMA (1,0,0) sebagai berikut

$$Z_t = 20690.356 + 0.87486Z_{t-1} + e_t + 4142.438I_t^{(48)} + 3944.007I_t^{(58)}$$

DAFTAR PUSTAKA

- Box, G.E.P., Jenkins, G.M., and Reissel.G.C. 2008. *Time Series Analysis Forecasting and Control*. 4rd edition. Canada: John Wiley & Sons, Inc.
- Cryer, J.D. and Chan, K.S. (2008). *Time Series Analisis With Application In R*. New York: Springer Science+Business
- Galeano, P. and Pena, D. (2011). *Additive Outlier Detection in Seasonal ARIMA Models by a Modified Bayesian Information Criterion*. *Economic Time Series: Modeling and Seasonality*, 13:317-336.
- Hasan, Iqbal.M. (2002). *Pokok-pokok Materi Statistika 1 (Statistik Deskriptif)*. Jakarta: PT. Bumi Aksara.
- Pankratz, Alan. (1991). *Forecasting with Dynamic Regression Models*. Canada: John Wiley & Sons, Inc.
- Pena, D. (2005). *A new statistic for influence in linear regression*. *Technometrics*, 47(1):1-12
- Wei, William W S. 2006. *Time Series Analysis Univariate and Multivariate Methods Second Edition*. United States of America: Pearson Education Inc.
- Yanti, T. S. (2010). *Analisis Deret Waktu*. Bandung: Pustaka Ceria, Yayasan Pena.