Analisis Data Ketinggian Permukaan Air Sungai Mahakam Daerah Kutai Kartanegara Tahun 2010-2016 Menggunakan Model *Autoregressive Integrated Moving Average* (ARIMA) Dengan Efek *Outlier*
(Studi Kasus: Data Rata-rata Ketinggian Tiap Bulan Permukaan Air Sungai Mahakam, Tenggarong, Kalimantan Timur)

Analysis of Water Surface Height data of the Mahakam River in the Kutai Kartanegara Region in 2010-2016 Using the Autoregressive Integrated Moving Average (ARIMA) model With Outlier Effect
(*Case Study: Average of Water Surface Height* Data of Mahakam *River*, Tenggarong, Kalimantan Timur)

Rezky Agustianto[1], Ika Purnamasari[2], dan Suyitno[3]
[1]Laboratorium Statistika Komputasi FMIPA Universitas Mulawarman
[2]Laboratorium Statistika Ekonomi dan Bisnis FMIPA Universitas Mulawarman
[3]Laboratorium Terapan FMIPA Universitas Mulawarman
[1]E-mail: rzkyagstiantoo@gmail.com

Abstract

Measurement of water level is useful as a guide for flood events in an area. As a result of global warming it is predicted that rainfall will increase and the water level will be high, so that the chances of flooding will increase. The method often used in forecasting is the method of Autoregressive Integrated Moving Average (ARIMA). ARIMA is one of the time series forecasting methods that has been studied in depth by Box and Jenkins. ARIMA's basic concepts include, identification of models, parameter estimation, diagnostic checks and forecasting. Forecasting results with the ARIMA method are inaccurate, on data that contains outliers. The weakness of the ARIMA method can be overcome using the ARIMA method with outlier detection. The type of outlier detection in this study is additive outlier (AO). The purpose of this study was to determine the ARIMA forecasting model with an outlier effect on the average water level data of the Mahakam River in the Kutai Kartanegara Region in front of the Tenggarong Museum Building from January 2010 - December 2016. The results showed that the best forecasting model was the river Mahakam Kutai Kartanegara Region is ARIMA ([12], 1,0) with the addition of 4 outlier effects and measure of goodness is AIC with a value of 250,0776.

Keywords: ARIMA, ARIMA with outlier effect, water level height.

Pendahuluan

Peramalan merupakan suatu teknik untuk membuat suatu keputusan dimasa yang akan datang dengan memperhatikan data masa lalu maupun data saat ini. Di dalam metode statistika, peramalan memiliki dua model, yaitu model deret waktu dan model regresi. Model deret waktu berpandangan untuk memperhitungkan kemungkinan berkembangnya suatu variabel serta menilai pengaruhnya di masa yang akan datang (Aswi dan Sukarna, 2006).

Autoregressive Integrated Moving Average (ARIMA) merupakan salah satu metode peramalan model deret waktu yang telah dipelajari secara mendalam oleh Box dan Jenkins pada tahun 1976 (Aswi dan Sukarna, 2006). Adapun beberapa konsep dasar ARIMA antara lain, pengidentifikasian model, penaksiran parameter, pemeriksaan diagnostik, dan peramalan. Namun hasil peramalan yang diperoleh menjadi tidak akurat jika terdapat pencilan atau *outlier* pada residual data.

Outlier adalah nilai pengamatan yang tidak konsisten dalam data runtun waktu atau nilainya jauh berbeda dari data lainnya. Ada 4 macam jenis *outlier* yaitu *Innovation Outlier* (IO), *Additive Outlier* (AO), *Temporary Change* (TC) dan *Level Shift* (LS) (Wei, 2006). Alfi (2010) meneliti tentang masalah tinggi pasang air laut di kota Semarang dengan menggunakan metode *Seasonal Autoregressive Integrated Moving Average* (SARIMA) dan deteksi *outlier*. Dalam penelitian tersebut diperoleh kesimpulan bahwa hasil peramalan menggunakan model SARIMA dengan 7 *outlier* lebih cocok untuk memprediksi tinggi pasang air laut diperairan utara Kota Semarang tahun 2013-2014 daripada model SARIMA tanpa *outlier* (Alfi, 2014).

Air merupakan sumber daya alam yang memiliki manfaat bagi keberlangsungan hidup manusia serta makhluk hidup lainnya. Sungai merupakan tempat dan wadah serta jaringan pengaliran air dari mata air sampai ke muara. Setiap sungai memiliki beberapa Daerah Aliran Sungai (DAS). DAS yang akan dijadikan objek peneliti oleh penulis yaitu DAS Mahakam yang terletak di daerah Kutai Kartanegara. Sungai ini memiliki peran yang sangat penting bagi perkembangan kota-kota yang dilalui sungai tersebut, karena sungai ini berfungsi sebagai akhir drainase utama pusat kota, objek wisata, penyediaan bahan baku air Perusahaan Daerah Air Minum (PDAM). Kota-kota yang dilalui oleh sungai adalah kawasan yang rawan bencana, seperti banjir. Banjir adalah kejadian atau fenomena alam dimana air sungai masuk ke wilayah daratan pada waktu permukaan sungai pasang (Wahyudi, 2007). Permukaan sungai pasang diakibatkan karena kenaikan permukaan laut akibat *global warning* (Wirasatriya, 2005).

Banjir pada DAS Mahakam diprediksikan akan semakin besar yang diakibatkan oleh kenaikan muka air laut dan penurunan tanah meningkat secara konstan (Arief, 2012). Bencana yang terjadi tidak hanya banjir, kekeringan dan masuknya air laut ke aliran sungai pun akan mengganggu distribusi air bersih atau air minum daerah tersebut. Dimana, air sungai tersebut adalah bahan baku air Perusahaan Daerah Air Minum (PDAM). Oleh karena itu pada penelitian ini dilakukan untuk menganalisis data ketinggian air sungai Mahakam daerah Kutai Kartanegara dengan metode ARIMA dan efek *outlier*.

Analisis Deret Waktu

Analisis deret waktu adalah salah satu prosedur statistika yang diterapkan untuk meramalakan struktur probabilistic keadaan yang akan terjadi di masa yang akan datang dalam rangka pengambilan keputusan. Salah satu tujuan analisis deret waktu adalah meramalkan kondisi di masa yang akan datang dan untuk mengetahui hubungan antar variabel (Aswi dan Sukarna, 2006). Dalam analisis deret waktu, pengamatan sekarang (Z_t) tergantung pada satu atau beberapa pengamatan sebelumnya (Z_{t-k}).

Pemodelan deret waktu yang sering digunakan antara lain adalah metode ARIMA Box-Jenkins. Menurut Wei (2006), pemodelan data runtun waktu dengan ARIMA Box-Jenkins harus memenuhi syarat stasioneritas, yaitu nilai mean $E(Z_t) = \mu$ dan varians $Var(Z_t) = E(\mathrm{Z_t} - \mu)^2 = \sigma^2$ konstan. Uji stasioneritas data dalam mean maka digunakan Uji Dickey Fuller. Jika data tidak stasioner dalam mean maka dilakukan *differencing*. Untuk melihat dan mengatasi ketidakstasioneran dalam varian dapat digunakan transformasi Box-Cox (Aswi dan Sukarna, 2006).

$$Z_t^{\lambda} = \frac{Z_t^{(\lambda)} - 1}{\lambda}$$

Salah satu model umum runtun waktu untuk ARIMA (p,d,q) sebagai berikut :

$$\left(1 - \phi_1 B - .. - \phi_p B^p\right)(1 - B)^d Z_t = (1 - \theta_1 B - .. - \theta_q B^q) a_t$$

Model subset ARIMA merupakan bagian dari model ARIMA tergeneralisasi (Tarno, 2013). Contoh model subset ARIMA ([2,4],1,[1,10]) dapat ditulis sebagai berikut :

$$(1 - \phi_2 B^2 - \phi_4 B^4)(1 - B)Z_t = (1 - \theta_1 B - \theta_{10} B^{10}) a_t$$

atau

$$Z_t = \frac{1 - \theta_1 B - \theta_{10} B^{10}}{1 - \phi_2 B^2 - \phi_4 B^4 (1 - B)}$$

Jika ada beberapa model yang signifikan dengan semua asumsi residual terpenuhi maka dapat dipilih suatu model terbaik berdasarkan pada nilai AIC terkecil pada masing-masing model yang diverifikasi (Wei, 2006).

Efek Outlier

Outlier adalah data pengamatan yang tidak konsisten pada seriesnya. Efek kejadian tersebut dapat dihitung dengan model intervensi jika waktu dan penyebab diketahui. Ada empat macam jenis outlier yaitu *Innovation Outlier* (IO), *Additive Outlier* (AO), *Temporary Change* (TC), dan *Level Shift* (LS).

***Additive Outlier* (AO)**

Additive Outlier (AO) adalah kejadian yang mempunyai efek pada data deret waktu hanya pada satu periode saja yaitu pada pengamatan ke-T. Diasumsikan bahwa (X_t) mengikuti model ARMA (p,q), dengan Z_t adalah data pengamatan dan X_t adalah data outlier bebas. bentuk umum sebuah *Additive Outlier* (AO) dalam proses ARMA diuraikan sebagai berikut :

$$Z_t = \begin{cases} X_t & t \neq T \\ X_t + \omega, & t = T \end{cases}$$
$$= X_t + \omega I_t^{(T)}$$
$$= \frac{\theta\ (\mathrm{B})}{\phi\ (B)} \alpha_t + \omega I_t^{(T)},$$

dengan

$$I_t^{(T)} = \begin{cases} 0 & t \neq T \\ 1 & t = T \end{cases},$$

dimana $I_t^{(T)}$ adalah variabel indikator yang mewakili ada atau tidak adanya outlier pada waktu T.

***Innovation Outliers* (IO)**

Innovation Outliers adalah kejadian yang efeknya mengikuti proses ARMA. Bentuk umum sebuah *Innovation Outliers* didefinisikan sebagai berikut :

$$Z_t = X_t + \frac{\theta\ (\mathrm{B})}{\phi\ (\mathrm{B})} \omega I_t^{(T)}$$
$$= \frac{\theta\ (\mathrm{B})}{\phi\ (B)} \alpha_t + \frac{\theta\ (\mathrm{B})}{\phi\ (B)} \omega I_t^{(T)}$$

$$= \frac{\theta\ (\mathrm{B})}{\phi\ (B)}(\alpha_t + \omega I_t^{(T)})$$

Dari persamaan di atas, *innovational outlier* (IO) mempengaruhi semua pengamatan $Z_{t}, _{t}+1, \cdots$, melebihi waktu T sepanjang memori sistem yang dijelaskan oleh $\theta(B)/\phi(B)$.

Secara umum, sebuah data *time series* bisa saja mengandung beberapa *outlier*, misalkan k buah *outlier* dengan tipe yang berbeda. Sehingga model umum *outlier* dapat ditulis sebagai berikut :

$$Z_t = \sum_{j=1}^{k} \omega_j \upsilon_j (B) I_t^{(T_j)} + X_t,$$

Dimana $X_t = \frac{\theta(B)}{\phi(B)} a_t$, $\upsilon_j(B)$ =1 untuk AO, dan $\upsilon_j(B) = \frac{\theta(B)}{\phi(B)}$ untuk IO pada waktu $t = T_j$.

***Level Shift* (LS)**

Suatu LS adalah kejadian yang mempengaruhi deret pada satu waktu tertentu yang memberikan suatu perubahan tiba-tiba dan permanen. Model outlier LS dinyatakan sebagai :

$$Z_t = X_t + \frac{1}{(1-B)}\ \omega I_t^{(T)}$$

***Temporary Change* (TC)**

Temporary Change (TC) adalah suatu kejadian di mana outlier menghasilkan efek awal sebesar ω pada waktu t, kemudian secara perlahan sesuai dengan besarnya δ. Model TC dapat dituliskan sebagai berikut :

$$Z_t = X_t + \frac{1}{(1-\delta B)}\ \omega I_t^{(T)}$$

Pada saat δ = 0 maka TC akan menjadi kasus *Additive Outier* (AO), sedangkan pada saat δ = 1 maka TC akan menjadi kasus *Level Shift* (LS).

Estimasi Efek *Outlier*

Estimasi *residual Additive Outliers* (AO) digunakan dalam estimasi dari ω setelah bentuk dasar AO terdeteksi. Sebuah kasus khusus dimana *outlier* dan semua parameter pada persamaan dibawah ini :

$$\pi(B) = \frac{\theta(B)}{\phi(B)} = (1 - \pi_1 B - \pi_2 B^2 - \cdots)$$

Dan mendefinisikan

$$e_t = \pi(B) Z_t$$

Adalah estimasi residual pada data yang mengandung *outlier*.

Estimasi residual dari sebuah AO dapat ditulis sebagai berikut:

$$\begin{aligned} e_t &= \pi(B) Z_t \\ &= \pi(B)\left[\frac{\theta(B)}{\phi(B)} a_t + \omega I_t^{(T)}\right] \\ &= \frac{\phi(B)}{\theta(B)}\left[\frac{\theta(B)}{\phi(B)} a_t + \omega I_t^{(T)}\right] \\ &= a_t + \frac{\theta\ (B)}{\phi(B)} \omega I_t^{(T)} \\ &= a_t + \pi(B) \omega I_t^{(T)} \end{aligned}$$

Misalkan $\widehat{\omega}$ adalah penduga dari ω,$\widehat{\omega}_{AT}$ diketahui sebagai penduga bobot dari AO. Karena $\{a_t\}$ adalah *white noise*, dari teori kuadrat terkecil didapatkan

$$\begin{aligned} AO:\ \widehat{\omega}_{AT} &= \frac{e_T - \sum_{j=1}^{n-T} \pi_j e_{T+j}}{\sum_{j=0}^{n-T} \pi_j^2} \\ &= \frac{\pi^{\circ}(F) e_T}{\tau^2} \end{aligned}$$

Dimana,

$\pi^*(F) = (1 - \pi_1 F - \pi_2 F^2 - .. - \pi_{n-T} F^{n-T})$, F adalah operator *forward shift* sehingga $Fe_t = e_{t-1}$, dan $\tau^2 = \sum_{j=0}^{n-T} \pi_j^2$. Varian estimator adalah

$$\begin{aligned} Var(\widehat{\omega}_{AT}) &= Var\ \left(\frac{\pi^*(F) e_T}{\tau^2}\right) \\ &= \frac{1}{\tau^4} Var[\pi^*(F) e_T] \\ &= \frac{\sigma_a^2}{\tau^2} \end{aligned}$$

Estimasi *residual* dari sebuah IO adalah sebagai berikut:

$$\begin{aligned} e_t &= \pi(B) Z_t \\ &= \pi(B)\left[\frac{\theta(B)}{\phi(B)}(a_t + \omega I_t^{(T)})\right] \\ &= \frac{\phi(B)}{\theta(B)}\left[\frac{\theta(B)}{\phi(B)}(a_t + \omega I_t^{(T)})\right] \\ &= a_t + \omega I_t^{(T)} \end{aligned}$$

Demikian pula, $\widehat{\omega}_{IT}$ merupakan estimator kuadrat terkecil ω untuk model IO.

$$IO:\ \widehat{\omega}_{IT} = e_T$$

dan

$$Var(\widehat{\omega}_{IT}) = Var(e_T) = (Var(\omega I_t^{(T)} + a_t) = \sigma_a^2$$

Dengan demikian, estimasi terbaik atas efek dari IO pada waktu T adalah residual e_T, sedangkan estimasi terbaik pengaruh AO adalah kombinasi linear dari $e_T, e_{T+1}, .., $ dan e_n dengan bobot tergantung pada struktur dari proses runtun waktu. Hal ini mudah dilihat bahwa $Var(\widehat{\omega}_{AT}) \leq Var(\widehat{\omega}_{IT}) = \sigma_a^2$ dan dalam beberapa kasus, $Var(\widehat{\omega}_{AT})$ bisa jauh lebih kecil dari σ_a^2.

Peramalan

Peramalan adalah kegiatan memperkirakan atau memprediksi apa yang akan terjadi pada masa yang akan datang dengan waktu yang relative lama. Sedangan ramalan adalah suatu situasi atau kondisi yang akan diperkirakan akan terjadi pada masa yang akan datang. Untuk memprediksi hal tersebut diperlukan data yang akurat di masa lalu, sehingga dapat dilihat prospek situasi dan kondisi di masa yang akan datang. Peramalan untuk beberapa periode yang akan datang dilakukan setelah tahap identifikasi, estimasi parameter dan pemeriksaan diagnostik yang menghasilkan model ARIMA terbaik dengan penambahan beberapa *outlier*.

Sumber Data dan Variabel Penelitian

Pada penelitian ini digunakan data sekunder yaitu data ketinggian permukaan air Sungai Mahakam Daerah Kutai Kartanegara yang telah diarsipkan oleh Dinas Pekerjaan Umum (DPU) Provinsi Kalimantan Timur. Variabel penelitian yang digunakan adalah data rata-rata ketinggian tiap bulan permukaan air Sungai Mahakam Daerah Kutai Kartanegara tepatnya didepan museum Tenggarong dari bulan Januari 2010 sampai dengan Desember 2016.

Langkah Analisis

Langkah analisis data pada penelitian ini secara umum sebagai berikut.

1. Melakukan analisis statistika deskriptif pada data
2. Melakukan identifikasi Model ARIMA dengan langkah sebagai berikut :
   a. Membuat *time series plot* untuk melihat pola data.
   b. Uji kestasioneran data, jika data belum stasioner dalam varian maka dilakukan transformasi Box-Cox, dan jika data belum stasioner dalam rata-rata maka dilakukan *differencing*. Kestasioneran dalam variansi dapat diketahui berdasarkan nilai *lamda* pada Box-Cox *plot*, dimana data stasioner dalam variansi apabila nilai *lamda* sama dengan satu.
   c. Membuat plot Fungsi Autokorelasi (FAK) dan Fungsi Autokorelasi Parsial (FAKP) dari data yang sudah stasioner.
3. Melakukan estimasi parameter model menggunakan metode *Least Square*.
4. Melakukan uji signifikansi parameter model ARIMA (*p*,*d*,*f*).
5. Melakukan uji asumsi, yang meliputi uji kenormalan residual dan uji residual *white noise*.
6. Memilih model terbaik ARIMA dengan syarat memenuhi asumsi dan memiliki nilai AIC yang terkecil.
7. Mendeteksi adanya *additive outlier* (AO).
8. Pembentukan model ARIMA dengan outlier, melakukan estimasi dan uji signifikansi parameter model ARIMA terbaik dengan penambahan outlier satu-persatu.
9. Pemilihan model ARIMA terbaik dengan dan tanpa outlier dengan menggunakan nilai AIC. Model terbaik adalah model dengan nilai AIC terkecil.
10. Peramalan ketinggian permukaan air Sungai Mahakam Daerah Kutai Kartanegara dilakukan dengan menggunakan model terbaik yang diperoleh.

Hasil Penelitian dan Pembahasan

1. Statistika Deskriptif

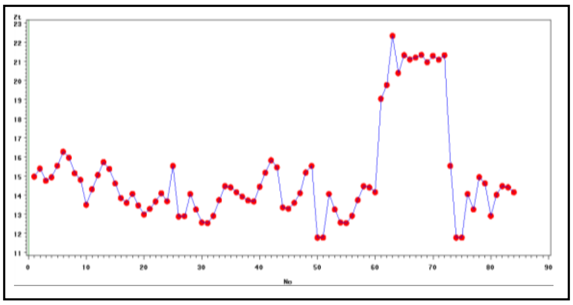
Analisis deskriptif digunakan untuk mengetahui gambaran umum dari data tersebut. Gambaran data yang dimaksud adalah seberapa besar nilai rata-rata, sebaran data, nilai maksimum, nilai minimum, serta jumlah data tinggi permukaan air sungai tiap bulannya yang digunakan dalam penelitian ini. Adapun nilai-nilai deksriptif dapat dilihat pada Tabel 1.

Tabel 1. Analisis Deskriptif Data

Banyak Data	Minimum	Maximum	Mean	Standar Deviasi
84	11,77	22,34	15,02	2,63

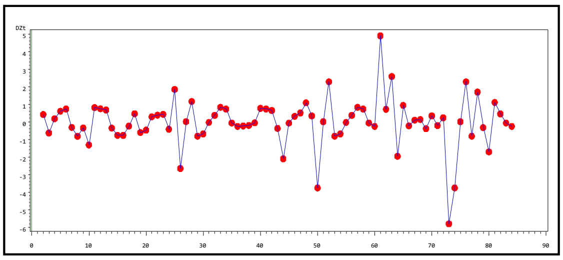
2. Identifikasi Model ARIMA

Langkah awal dalam melakukan identifikasi model ARIMA adalah dengan memeriksa kestasioneran data menggunakan *time series plot*. Adapun hasil dari *time series plot* dapat dilihat pada gambar 1 :



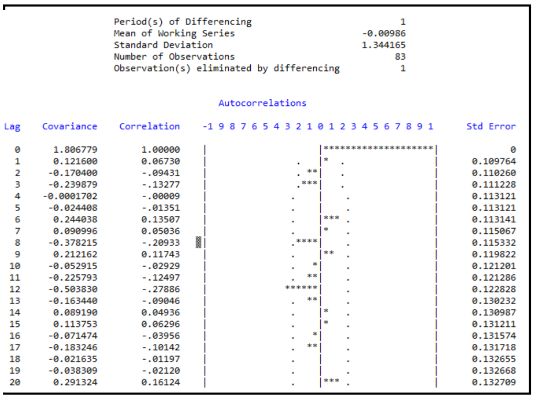
Gambar 1. *Time series plot* Data

Pada gambar 1 terlihat bahwa data dikatakan belum stasioner dalam rata-rata sehingga dilakukan *differencing* pada data. Adapun hasil *time series plot* setelah *differencing* dapat dilihat pada gambar 2 :



Gambar 2. *Time series plot* setelah *differencing*

Berdasarkan hasil pada gambar 2 dapat terlihat bahwa data dapat dikatakan stasioner dalam rata-rata. Setelah asumsi stasioneritas terpenuhi, maka selanjutnya dapat menduga model runtun waktu dengan melihat plot ACF dan plot PACF.

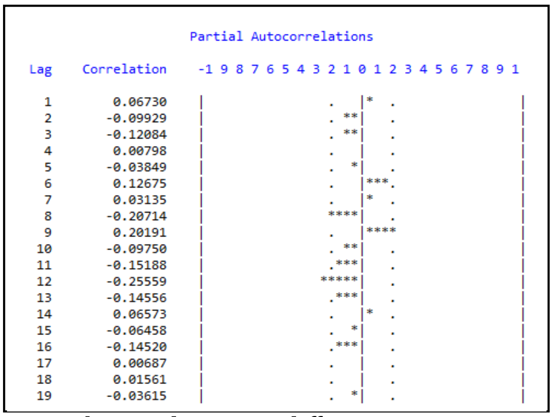


Period(s) of Differencing	1
Mean of Working Series	-0.00986
Standard Deviation	1.344165
Number of Observations	83
Observation(s) eliminated by differencing	1

Autocorrelations

Lag	Covariance	Correlation	Std Error
0	1.806779	1.00000	0
1	0.121600	0.06730	0.109764
2	-0.170400	-.09431	0.110260
3	-0.239879	-.13277	0.111228
4	-0.0001702	-.00009	0.113121
5	-0.024408	-.01351	0.113121
6	0.244038	0.13507	0.113141
7	0.090996	0.05036	0.115067
8	-0.378215	-.20933	0.115332
9	0.212162	0.11743	0.119822
10	-0.052915	-.02929	0.121201
11	-0.225793	-.12497	0.121286
12	-0.503830	-.27886	0.122828
13	-0.163440	-.09046	0.130232
14	0.089190	0.04936	0.130987
15	0.113753	0.06296	0.131211
16	-0.071474	-.03956	0.131574
17	-0.183246	-.10142	0.131718
18	-0.021635	-.01197	0.132655
19	-0.038309	-.02120	0.132668
20	0.291324	0.16124	0.132709

Gambar 3. Plot ACF *differencing* pertama



Partial Autocorrelations

Lag	Correlation
1	0.06730
2	-0.09929
3	-0.12084
4	0.00798
5	-0.03849
6	0.12675
7	0.03135
8	-0.20714
9	0.20191
10	-0.09750
11	-0.15188
12	-0.25559
13	-0.14556
14	0.06573
15	-0.06458
16	-0.14520
17	0.00687
18	0.01561
19	-0.03615

Gambar 4. Plot PACF *differencing* pertama

Pada gambar 3 dapat dilihat bahwa plot ACF hanya keluar pada lag ke-12, diperoleh orde untuk MA adalah 12. Sementara untuk plot PACF yang keluar signifikan adalah pada lag ke-8, 9 dan 12, diperoleh orde untuk AR adalah 8, 9, 12. Sehingga model dugaan ARIMA yang diperoleh dapat dilihat pada tabel 2.

Tabel 2. Model Dugaan ARIMA

No.	Model ARIMA	No.	Model ARIMA
1	ARIMA ([8],1,0)	8	ARIMA (8,1,0)
2	ARIMA ([9],1,0)	9	ARIMA (9,1,0)
3	ARIMA ([12],1,0)	10	ARIMA (12,1,0)
4	ARIMA (0,1,[12])	11	ARIMA (0,1,12)
5	ARIMA ([8],1,[12])	12	ARIMA (8,1,12)
6	ARIMA ([9],1,[12])	13	ARIMA (9,1,12)
7	ARIMA ([12],1,[12])	14	ARIMA (12,1,12)

3. Estimasi Parameter Model

Estimasi parameter model ini dilakukan untuk menguji apakah koefisien parameter pada model signifikan atau tidak. Pendugaan model dilakukan berdasarkan lag yang keluar pada plot ACF dan PACF. Dari 14 pendugaan model diperoleh hanya parameter model dugaan ARIMA ([8],1,0), ([12],1,0) dan (0,1,2) yang memenuhi karena semua parameternya signifikan.

4. Pemeriksaan Diagnostik Model ARIMA

Setelah diperoleh model yang signifikan dari pendugaan model yaitu ARIMA ([8],1,0), ([12],1,0) dan (0,1,[12]) maka dilakukan pemeriksaan diagnostik. Pemeriksaan diagnostik terdiri dari dua uji, yaitu uji kesesuaian model dan uji kenormalan residual. Adapun hasil pemeriksaan diagnostik dapat dilihat pada tabel 3.

Tabel 3. Pemeriksaan Diagnostik

Model	*White Noise*	Normalitas
ARIMA ([8],1,0)	Tidak	Tidak
ARIMA ([12],1,0)	Ya	Tidak
ARIMA (0,1,[12])	Ya	Tidak

Berdasarkan tabel 3 dapat disimpulkan bahwa residual model untuk semua model tidak memenuhi normal sehingga asumsi normalitas residual model belum terpenuhi. Jadi, *diagnostic checking model* ketiga model memiliki kesimpulan yang sama bahwa ketiga model belum bisa memodelkan data pengamatan. Menurut hasil ukuran kebaikan model dapat dilihat pada nilai AIC, model **ARIMA ([12],1,0)** memiliki nilai AIC terkecil dan bentuk model ARIMA ([12],1,0) dengan melihat nilai estimasi parameternya sebagai berikut:

Diketahui parameter AR (12) pada lag-12 (ϕ_{12}) sebesar -0,47068, maka model ARIMA ([12],1,0) menjadi

$$Z_t = Z_{t-1} - 0{,}47068 Z_{t-12} + a_t$$

5. Pendeteksian *Outlier* pada Model ARIMA

Memperoleh Residual Model dari *Output* Model ARIMA Pilihan

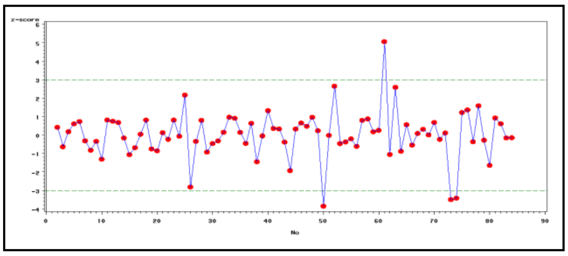
Dengan diperoleh model pilihan yaitu ARIMA ([12],1,0) dihasilkan residual model dengan rumusan berikut:

$$Z_t = Z_{t-1} - 0{,}47068 Z_{t-12} + a_t$$
$$a_t = Z_t - (Z_{t-1} - 0{,}47068 Z_{t-12})$$
$$a_t = Z_t - Z_{t-1} + 0{,}47068 Z_{t-12}$$

Menghitung Nilai *z-score* dari Residual Model (a_t) yang Terbentuk

Residual model (a_t) digunakan untuk menghitung nilai *z-score* (z_t) untuk mendeteksi adanya *outliers* dalam residual model. Data residual model akan terdeteksi sebagai Outlier jika nilai mutlak *z-score* lebih besar dari tiga.

Untuk mendeteksi outlier digunakan nilai z-score yaitu nilai standardisasi residual model. Outlier diperoleh ketika nilai mutlak z-score lebih besar dari 3. Dari gambar 5, diperoleh outlier pada data z-score yaitu *Additive Outlier* (AO) pada pengamatan 50, 61, 73, dan 74. Selanjutnya dilakukan estimasi parameter model dengan penambahan *outlier* satu persatu.



Gambar 5. *Time series plot* dari nilai *z-score*

6. Estimasi Parameter Model ARIMA dengan Efek *Outlier*

Estimasi parameter model ini dilakukan untuk menguji apakah koefisien parameter pada model signifikan atau tidak. Pengujian hipotesis dilakukan untuk model ARIMA dengan penambahan *outlier* satu persatu dan setelah itu dilakukan pengujian signifikansi parameter model. Adapun hasilnya dapat dilihat pada tabel 4.

Tabel 4. Uji Signifikansi Parameter Model dengan Penambahan *Outlier*

ARIMA ([12],1,0) dengan penambahan *outlier* residual data ke-	Signifikansi Parameter
50	Ya
61	Ya
73	Tidak
74	Ya
50,61	Ya
50,61,73	Tidak
50,61,73,74	Ya

7. Pemeriksaan Diagnostik Model ARIMA dengan Efek *Outlier*

Setelah diperoleh model yang signifikan dari model ARIMA ([12],1,0) dengan penambahan *outlier* satu persatu selanjutnya dilakukan pemeriksaan diagnostik. Pemeriksaan diagnostik terdiri dari dua uji, yaitu uji kesesuaian model dan uji kenormalan residual. Adapun hasil pemeriksaan diagnostik dapat dilihat pada tabel 5:

Tabel 5. Pemeriksaan Diagnostik Model ARIMA dengan Efek *Outlier*

ARIMA ([12],1,0) dengan penambahan *outlier* residual data ke-	*White Noise*	Normalitas
50	Ya	Tidak
61	Ya	Tidak
73	Ya	Tidak
74	Ya	Tidak
50,61	Tidak	Tidak
50,61,73	Tidak	Tidak
50,61,73,74	Tidak	Ya

Dari pengujian signifikansi parameter, dihasilkan parameter model untuk kelima model signifikan karena *p-value* kurang dari 5%. Pada pengujian *white noise*, residual model untuk semua model *white noise* karena semua *p-value* lebih besar dari 5%. Untuk pengujian asumsi normalitas residual model, hanya residual model untuk model ARIMA dengan penambahan 4 efek *outlier* yang memenuhi normal atau asumsi normalitas residual model terpenuhi karena *p-value* lebih besar dari 5%.

Sehingga, hanya model ARIMA dengan penambahan 4 efek *outlier* yang memiliki kesimpulan terbaik. Menurut hasil ukuran kebaikan model AIC, model ARIMA pilihan jatuh pada model ARIMA ([12],1,0) dengan penambahan 4 efek *outlier* yang mana nilai AIC dapat dilihat pada tabel 6 dan dapat dirumuskan sebagai berikut:

ARIMA ([12],1,0) dengan penambahan 4 efek *outlier*

$$Z_t = Z_{t-1} + \phi_1 Z_{t-12} - \phi_1 Z_{t-13} + \mu_1 O_{50} + \mu_2 O_{61} + \mu_3 O_{73} + \mu_4 O_{74} + a_t$$

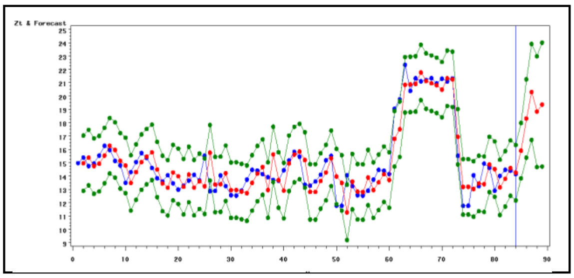
$$Z_t = Z_{t-1} - 0{,}64476\ Z_{t-12} + 0{,}64476\ Z_{t-13} - 1{,}68594\ O_{50} + 2{,}86808\ O_{61} - 3{,}06982\ O_{73} - 3{,}093080\ O_{74} + a_t$$

Tabel 6. Nilai AIC dengan Penambahan Efek *Outlier*

ARIMA ([12],1,0) dengan penambahan *outlier* residual data ke-	AIC
50	268,0072
61	265,9802
73	276,9507
74	269,6414
50,61	261,2135
50,61,73	260,6557
50,61,73,74	250,0776

8. Peramalan

Hasil ramalan data ketinggian permukaan air sungai Mahakam model ARIMA ([12],1,0) dengan penambahan 4 efek *outlier* di daerah Kutai Kartanegara pada bulan januari 2010 sampai 2017 dapat dilihat pada gambar 6 atau tabel 7.



Gambar 6. Grafik Hasil Ramalan Data Tinggi Sungai Mahakam Daerah Kutai Kartanegara Model ARIMA ([12],1,0) dengan penambahan 4 *outlier* bertipe AO

Pada gambar 6 dapat dilihat bahwa hasil ramalan mendekati nilai aktual bahkan beberapa nilainya memotong data aktual sehingga dapat disimpulkan bahwa model ARIMA ([12],1,0) dengan penambahan 4 efek *outlier* lebih cocok untuk memprediksi tinggi air sungai Mahakam Daerah Kutai Kartanegara Tahun 2010-2017.

Tabel 7. Hasil *Forecast* Data Ketinggian Air Sungai Mahakam Bulan Januari 2017 – Mar 2017

No	Data Bulanan	*Forecast*	Batas Bawah	Batas Atas
85	Jan-17	15,9075	13,8297	17,9853
86	Feb-17	18,3148	15,3764	21,2533
87	Mar-17	20,3051	16,7062	23,9039

Kesimpulan

Berdasarkan analisis dan pembahasan yang telah dilakukan, maka dapat disimpulkan bahwa model terbaik yang didapat adalah ARIMA ([12],1,0) penambahan 4 efek *outlie*r dengan nilai AIC yaitu sebesar 250,0776 dan model yang terbentuk adalah sebagai berikut:

$$Z_t = Z_{t-1} - 0{,}64476\, Z_{t-12} + 0{,}64476\, Z_{t-13} - 1{,}68594\, O_{50} + 2{,}86808\, O_{61} - 3{,}06982\, O_{73} - 3{,}093080\, O_{74} + a_t$$

hasil peramalan menggunakan model ARIMA ([12],1,0) dengan penambahan 4 efek *outlie*r cocok untuk memprediksi ketinggian air sungai Mahakam daerah Kutai Kartanegara. Hasil peramalan pada data bulan januari – mar 2017 yaitu, 15,90, 18,31, 20,30.

Daftar Pustaka

Alfi F., Dwi I., dan Suparti. (2014). Prediksi Tinggi Pasang Air Laut Di Kota Semarang Dengan Menggunakan Metode SARIMA Dan Deteksi Outlier. *Jurnal Gaussian*. 3 (3).

Arief L.N., Purnama B.S., & Trias A. (2012). Pemetaan Risiko Bencana Banjir Rob Kota Semarang. *The 1st Conference on Geospatial Information Science and Engineering*. Teknik Geodesi UNDIP. Semarang.

Aswi dan Sukarna. (2006). *Analisis Deret Waktu: Teori dan Aplikasi*. Makassar: Andira Publisher.

Tarno. (2013). Kombinasi Prosedur Pemodelan Subset Arima dan Deteksi Outlier untuk Prediksi Data Runtun Waktu. *Prosiding Seminar Nasional Statistika UNDIP 2013*. Semarang.

Wahyudi. (2007). Tingkat Pengaruh Elevasi Pasang Laut Terhadap Banjir dan Rob di Kawasan Kaligawe Semarang. Riptek Vol, 1 (1), 27-34.

Wei W.W.S. (2006). *Time Series Analysis: Univariat dan Multivariate Methods Second Edition*. Addison Wesley Boston: Pearson Education, Inc.