

## PEMODELAN DATA CURAH HUJAN DI KOTA LANGSA DENGAN MODEL ARIMA

Wiwin Apriani<sup>1\*</sup>, Nurhayati<sup>2</sup><sup>1</sup>Program studi Teknik Komputer, FT, Universitas Sains Cut Nyak Dhien, Langsa, Aceh<sup>2</sup>Program studi Pendidikan Matematika, FKIP, Universitas Al-Muslim, Bireun, AcehEmail korespondensi\*: [wiwina10@gmail.com](mailto:wiwina10@gmail.com)**Abstrak**

Tujuan dari penelitian untuk memberikan hasil pemodelan ARIMA pada data curah hujan di Kota Langsa tahun 2017-2021. Tahap awal dari pemodelan ARIMA adalah identifikasi stasioneritas data. Sedangkan stasioneritas dalam mean dapat dilakukan dengan plot data dan bentuk ACF. Identifikasi bentuk ACF dan PACF dari data yang sudah stasioner digunakan untuk menentukan orde model ARIMA dugaan. Tahapan selanjutnya adalah estimasi parameter untuk melihat kesesuaian model. Proses cek diagnosa dilakukan untuk mengevaluasi apakah residual model sudah memenuhi syarat white noise dan berdistribusi normal. Uji Ljung-Box adalah uji yang dapat digunakan untuk memvalidasi syarat white noise. Data curah hujan membentuk deret waktu yang stasioner. Selanjutnya dari uji kecocokan model diperoleh bahwa model MA(1) yang cocok untuk meramalkan model tersebut. Sedangkan AR(1) dan ARMA(1,1) tidak digunakan untuk meramalkan karena tidak memenuhi uji kecocokan model. Adapun model yang diperoleh dengan model MA(1) adalah sebagai berikut yaitu  $Z_t = a_t - 0,4096a_{t-1} + 521,57$ .

**Kata kunci:** Curah Hujan, ARIMA, Pemodelan**Abstract**

*The aim of this research is to provide the results of ARIMA modeling on rainfall data in Langsa City in 2017-2021. The initial stage of ARIMA modeling is the identification of data stationarity. Meanwhile, stationarity in the mean can be done with data plots and ACF forms. Identification of ACF and PACF forms from data that is already stationary is used to determine the order of the alleged ARIMA model. The next stage is parameter estimation to see the suitability of the model. The diagnostic check process is carried out to evaluate whether the residual model meets the white noise requirements and is normally distributed. The Ljung-Box test is a test that can be used to validate white noise requirements. Rainfall data forms a stationary time series. Furthermore, from the model fit test it was found that the MA(1) model was suitable for predicting the model. Meanwhile, AR(1) and ARMA(1,1) are not used to predict because they do not meet the model fit test. The model obtained with the MA(1) model is as follows, namely  $Z_t = a_t - 0,4096a_{t-1} + 521,57$ .*

**Keywords:** Rainfall, ARIMA, Modelling**Sejarah artikel**

Diterima: 18-08-2022

Direvisi: 10-09-2022

Dipublikasikan: 01-11-2022

**Article history**

Received: 18-08-2022

Revised: 10-09-2022

Published: 01-11-2022





## A. Pendahuluan

Curah hujan merupakan faktor utama yang mempengaruhi redaman hujan pada gelombang milimeter. Gelombang milimeter adalah gelombang radio yang bekerja di atas frekuensi 10 GHz. Karena redaman hujan sebanding dengan redaman spesifik dikalikan panjang link. Redaman spesifik sebanding dengan dengan curah hujan. Jika diasumsikan panjang link 1 Km maka besar redaman hujan akan bergantung pada curah hujan untuk frekuensi tertentu. Karena itu perlu ditinjau bagaimana model curah hujan yang terjadi setiap saat. Dengan mengetahui model curah hujan, maka dapat ditentukan model redaman hujan. Pemodelan curah hujan sudah banyak dilakukan dengan menggunakan uji statistik untuk menentukan validitas dari suatu model.

Salah satu metode statistic yang dapat digunakan untuk menganalisis pemodelan curah hujan yaitu analisis deret berkala. Analisis deret berkala merupakan proses belajar mengenai angka sebagai variabel yang berubah dalam beberapa waktu tertentu. Dalam menganalisis dan meramalkan deret berkala, banyaknya pemilihan produk berdasarkan katagori-katagori beberapa waktu tertentu dimana mampu memperlihatkan perubahan (Apriani & Nurhayati, 2020; Wahyuni & Nurhayati, 2020). Perbedaan kelas dari perbedaan produk menyebabkan variasi dalam deret berkala. Analisis deret berkala menggunakan metode pemulusan (*smoothing*). Klasifikasi metode pemulusan dibagi menjadi metode awal dari kelompok nilai tengah sederhana dari semua data masa lalu. Metode berikutnya adalah rata-rata bergerak tunggal (*single moving avarage*) dari  $n$  nilai observasi yang terakhir. Rata-rata bergerak ganda (*double moving avarage*), atau rata-rata bergerak yang akhirnya menjadi rata-rata yang berbobot tidak sama, dan dapat digunakan dalam metode peramalan yang biasanya disebut rata-rata bergerak linier (*linier moving avarage*) (Adhikari & Agrawal, 2013).

*Autoregressive Integrated Moving Average* (ARIMA) merupakan suatu model statistik yang cocok digunakan untuk meramal sejumlah variabel secara cepat, sederhana, murah, dan akurat karena hanya membutuhkan data variabel yang akan diramal. Metode ARIMA menggunakan pendekatan iteratif dalam indentifikasi terhadap suatu model yang ada. Model yang dipilih dan diuji dengan data masa lampau untuk melihat apakah model tersebut menggambarkan keadaan data secara akurat atau tidak. Model ARIMA atau model gabungan auto-regresi dengan rata-rata bergerak merupakan jenis model linier yang mampu mewakili deret waktu yang stasioner maupun non-stasioner. Pada metode peramalan dengan menggunakan metode ARIMA di mana sangat baik ketepatannya untuk peramalan jangka pendek, sedangkan untuk peramalan jangka panjang ketepatan peramalannya kurang baik. Biasanya akan cenderung *flat* (mendatar/konstan) untuk periode jangka panjang. Model ARIMA adalah model yang secara penuh mengabaikan independen variabel dalam membuat peramalan (Apriani, 2021).

Karena itu perlu dibuat model ARIMA yang dilengkapi dengan evaluasi kesesuaian model. Dalam penelitian ini akan diuraikan langkah-langkah pemodelan ARIMA pada data curah hujan. Tahapan penting dalam menentukan model ARIMA adalah identifikasi, estimasi parameter, dan cek diagnosa. Tahap Identifikasi berfungsi untuk menentukan orde model ARIMA melalui bentuk ACF (*Autocorrealation Function*) dan bentuk PACF (*Partial Autocorrelation Function*) dari data yang sudah stasioner. Estimasi parameter berfungsi untuk menduga nilai besaran konstanta dan koefisien dari model AR dan MA. Sedangkan cek diagnosa berfungsi untuk menguji kesesuaian model melalui uji residual apakah sudah memenuhi syarat white noise dan berdistribusi normal



Analogi dengan rata-rata bergerak ganda terdapat metode pemulusan ganda yang meliputi dua persamaan pemulusan eksponensial (Apriani & Nurhayati, 2020; Apriani & Hayati, 2021).

Adapun beberapa penelitian terdahulu yang telah dilakukan yaitu pada tahun 2010 Huda dkk melakukan penelitian tentang peramalan data curah hujan dengan model SARIMA dengan tambahan deteksi outlier di Kabupaten Mojokerto, hasil penelitian tersebut memberikan ramalan curah hujan yang sesuai karena mampu memprediksi data curah hujan pada tahun 2011 dengan tepat (Huda, 2010). Purwaputra juga telah melakukan penelitian tentang peramalan curah hujan untuk evaluasi pola tanam berdasarkan penanggalan Sasih di Bali dimana curah hujan di daerah Jatiluwih memiliki pola musiman dan berdasarkan curah hujan yang terjadi pola tanam kurang sesuai sebanyak 4 kali dalam 25 tahun terakhir namun kalender tanam dengan penanggalan Sasih masih layak digunakan (Purwaputra, 2015). Metode Neural Network juga pernah digunakan untuk meramalkan curah hujan di Jawa Tengah dimana metode Time Delay Neural Network lebih baik dibandingkan Feed Forward Neural Network (Fitriasari, 2013). Selain digunakan untuk meramalkan curah hujan metode ARIMA telah dilakukan untuk meramalkan penjualan pakaian di Boyolali selain dengan menggunakan metode ARIMA juga menggunakan Time Series Regression (TSR) sebagai pembanding. Pada penelitian tersebut menyimpulkan bahwa hasil peramalan dengan metode TSR lebih baik daripada ARIMA karena menghasilkan nilai RMSE out sample yang lebih kecil (Perdana, 2012).

## B. Metode Penelitian

Penelitian ini merupakan penelitian kuantitatif dengan metode deskriptif yang bertujuan menjelaskan suatu data berdasarkan karakteristik yang dimiliki. Menurut Arikunto (2006), penelitian kuantitatif, banyak menggunakan angka mulai dari pengumpulan data, menafsirkan data, serta menyajikan data tersebut sesuai hasil yang diinginkan. Penelitian ini merupakan penelitian deskriptif kuantitatif dengan tujuan untuk meramalkan dan mendeskripsikan dan memodelkan data curah hujan di Kota Langsa dengan menggunakan model ARIMA. Data yang digunakan ialah data sekunder yang diperoleh dari Badan Metrologi, Klimatologi, dan Geofisika (BMKG) Kota Langsa yang merupakan data curah hujan bulanan dengan interval waktu selama 5 tahun pada Bulan Januari 2017 hingga Desember 2021 sebanyak 60 data. Data yang diperoleh selanjutnya digunakan untuk meramalkan dan memodelkan data curah hujan dengan menggunakan model ARIMA. Tahapan penelitian yang akan dilakukan dalam penelitian ini yaitu melakukan analisis statistika deskriptif dan eksplorasi data dari curah hujan. Selanjutnya melakukan pemodelan curah hujan dengan metode ARIMA. Adapun langkah-langkah yang dilakukan pada penelitian ini yaitu menguji kestasioneran data, mendapatkan orde model pada ARIMA  $(p,d,q)$ , mengestimasi parameter model dengan metode least square, menguji parameter model dan asumsi residual, memilih model ARIMA terbaik berdasarkan kriteria RMSE.

### Model ARIMA $(p,d,q)$ ,

$p$  merupakan model auto-regresif order- $p$  yang mempunyai bentuk :

$$Y_t = \theta_0 + \theta_1 Y_{t-1} + \theta_2 Y_{t-2} + \dots + \theta_p Y_{t-p} + \varepsilon_t \quad (1)$$

Di mana :

$Y_t$  = Variabel respon (terikat) pada waktu  $t$

$Y_{t-1}, Y_{t-2}, \dots, Y_{t-p}$  = Variabel respon pada masing-masing selang waktu  $t-1, t-2, \dots, t-p$

$Y$  = variabel bebas

$\theta_0, \theta_1, \theta_2, \dots, \theta_p$  = Koefisien yang diestimasi

$\varepsilon_t$  = Galat pada saat  $t$  yang mewakili dampak variabel-variabel yang tidak dijelaskan oleh model  
 Asumsi mengenai galat adalah sama dengan asumsi model regresi standar.  $D$  merupakan banyaknya selisih yang didapat dari proses penyisihan (*differencing*) dari deret waktu yang non-stasioner menjadi deret waktu stasioner. Jika deret aslinya stasioner,  $d=0$  model ARIMA berubah menjadi model ARMA. Persamaan dari proses penyisihan (*differencing d*) yaitu  $\Delta Y_t = Y_t - Y_{t-1}$ .  $q$  merupakan model rata-rata bergerak orde ke  $q$  yang mempunyai bentuk :

$$Y_t = \mu + \varepsilon_t - \omega_1 \varepsilon_{t-1} - \omega_2 \varepsilon_{t-2} - \dots - \omega_q \varepsilon_{t-q} \tag{2}$$

di mana :

$Y_t$  = Variabel respon (terikat) pada saat  $t$

$\mu$  = Mean konstanta proses

$\omega_1, \omega_2, \dots, \omega_q$  = Koefisien yang diestimasi

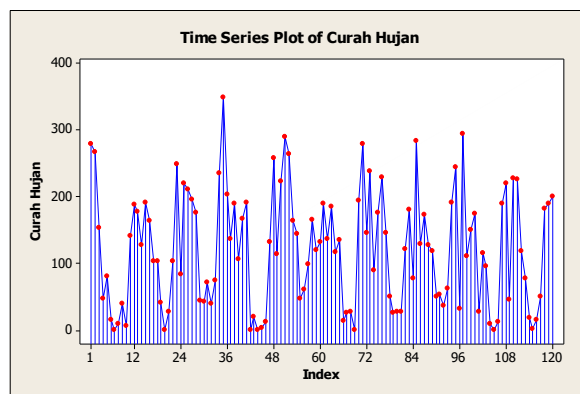
$\varepsilon_t$  = Bentuk galat yang mewakili efek variabel tak dijelaskan oleh model.

Asumsi mengenai bentuk galat adalah sama dengan asumsi model regresi standar.  $\varepsilon_{t-1}, \varepsilon_{t-2}, \dots, \varepsilon_{t-q}$  merupakan galat pada periode waktu sebelumnya yang pada saat  $t$ , nilainya menyatu dengan nilai respon  $Y_t$ .

### C. Hasil Dan Pembahasan

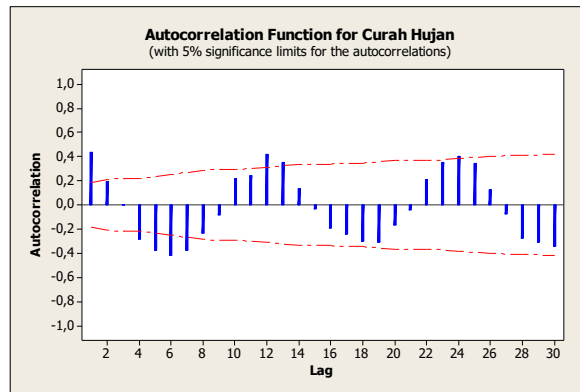
#### Uji Stasioner dan Identifikasi Model

Pada Gambar 1. terlihat pola data curah hujan di Kota Langsa Tahun 2017-2021. Pada grafik tersebut terlihat bahwa rata-rata dan variasi dari data tersebut adalah konstan, sehingga dapat disimpulkan bahwa data tersebut adalah stasioner. Berikut adalah grafik pola data curah hujan di Kota Langsa:

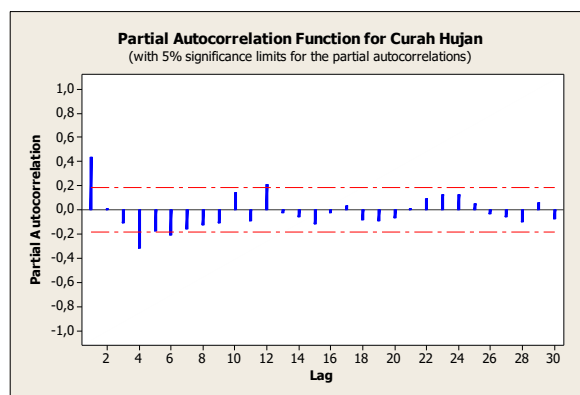


Grafik 1 Grafik Curah Hujan Tahun 2017-2021 di Kota Langsa

Dari grafik tersebut terlihat bahwa data sudah stasioner. selanjutnya dilakukan identifikasi model dengan menggunakan nilai  $f_{ak}$  dan  $f_{akp}$ . Adapun grafik  $f_{ak}$  dan  $f_{akp}$  adalah sebagai berikut:



Grafik 2 FAK Curah Hujan Tahun 2017-2021



Grafik 3 FAKP Curah Hujan Tahun 2017-2021

Selanjutnya diidentifikasi beberapa model yaitu AR(1), MA(1), dan ARMA (1,1). Kemudian dilakukan penaksiran terbaik untuk parameter model tersebut. Dari pengolahan data yang telah dilakukan dengan menggunakan software MINITAB diperoleh model AR(1) yang mempunyai bentuk  $Z_t = \varphi Z_{t-1} + a_t$  atau  $(Z_t - \bar{Z}) = \varphi(Z_{t-1} - \bar{Z}) + a_t$ . Dari hasil estimasi AR(1) diperoleh  $\varphi = 0,4506$ . Karena mean  $\bar{Z} = 123,50 > 2 SE 2(12,83)=25,66$ , maka  $\bar{Z}$  berbeda secara signifikan dengan nol. Sehingga model yang digunakan adalah model kedua. Sehingga diperoleh:

$$Z_t = 0,4506Z_{t-1} + a_t + 67,853 \tag{3}$$

Model MA(1) mempunyai bentuk  $Z_t = a_t + \theta a_{t-1}$  atau  $(Z_t - \bar{Z}) = a_t + \theta a_{t-1}$ . Dari hasil estimasi MA(1) diperoleh  $\theta = -0,3937$ . Karena mean atau  $\bar{Z} = 122,47 > 2 SE=2(10,02)=20,04$  maka  $\bar{Z}$  berada secara signifikan dengan nol. Sehingga model yang digunakan adalah model bentuk kedua. Sehingga diperoleh:

$$Z_t = a_t - 3937a_{t-1} + 122,47 \tag{4}$$

Model ARMA(1,1) mempunyai bentuk  $Z_t = \varphi Z_{t-1} + a_t + \theta a_{t-1}$  atau  $(Z_t - \bar{Z}) = \varphi(Z_{t-1} - \bar{Z}) + a_t + \theta a_{t-1}$ . Dari hasil ARMA diperoleh  $\varphi = 0,4709$  dan  $\theta = 0,0249$ . Karena mean atau  $\bar{Z} = 123,58 > 2 SE=2(13,05)=26,1$  maka  $\bar{Z}$  berada secara signifikan dengan nol. Sehingga model yang digunakan adalah model bentuk kedua. Sehingga diperoleh:

$$Z_t = 0,4709 + a_t + 0,0249a_{t-1} + 65,386 \tag{5}$$

**Verifikasi Model**



a. Uji keberatian koefisien

Model AR(1)

$$|\hat{\varphi}| = 0,4505 > 2SE = 2(0,0824)=0,1648$$

Jadi koefisien  $\varphi$  berarti berbeda secara signifikan dengan nol, hal ini berarti bahwa koefisien berarti.

Model MA(1)

$$|\hat{\theta}| = 0,3937 > 2SE = 2(0,0847)=0,1694$$

Jadi koefisien  $\theta$  berarti berbeda secara signifikan dengan nol, hal ini berarti bahwa koefisien berarti.

Model ARMA(1,1)

$$|\hat{\varphi}| = 0,4709 > 2SE = 2(0,1816)=0,3632$$

Jadi koefisien  $\varphi$  berarti berbeda secara signifikan dengan nol, hal ini berarti bahwa koefisien berarti.

$$|\hat{\theta}| = 0,0249 < 2SE = 2(0,2051)=0,4102$$

Jadi koefisien  $\theta$  berarti tidak berbeda secara signifikan dengan nol, hal ini berarti bahwa koefisien tidak berarti.

Karena model ARMA tidak memenuhi dari segi keberartian koefisien, sehingga tidak perlu lagi diikuti sertakan pada model selanjutnya.

b. Uji Kecocokan Model (*Lack of Fit*)

Uji kecocokan dilakukan untuk melihat kecocokan model dengan data. Adapun kriteria pengujian yang digunakan adalah model diterima jika P-value >  $\alpha$  dengan  $\alpha=0,01$ .

Model AR(1). Berikut ini merupakan hasil pengujian model AR(1):

Tabel 2 Nilai Q Box-Pierce Model AR(1) untuk Data Curah Hujan

LAG	P-Value	Kesimpulan
12	0,000	Model Tidak Sesuai
24	0,000	Model Tidak Sesuai
36	0,000	Model Tidak Sesuai
48	0,000	Model Tidak Sesuai

Selanjutnya dilakukan pengujian kecocokan model MA(1) sebagai berikut:

Tabel 3 Nilai Q Box-Pierce Model MA(1) untuk Data Curah Hujan

LAG	P-Value	Kesimpulan
12	0,015	Model Sesuai
24	0,019	Model Sesuai
36	0,025	Model Sesuai
48	0,029	Model Sesuai

Berdasarkan uraian diatas,maka model yang paling cocok dengan data adalah MA(1). Maka model peramalan untuk curah hujan adalah:

$$Z_t = a_t - 3937a_{t-1} + 122,47, \text{ Dengan } a_t \sim N(0; 29192,72) \tag{6}$$

Setelah didapat model peramalan yang sesuai dengan data curah hujan, maka selanjutnya diperoleh nilai ramalan untuk beberapa periode kedepan.



#### D. Simpulan

Dari hasil penelitian yang telah dilakukan dapat disimpulkan bahwa data curah hujan di Kota Langsa tahun 2017-2021 membentuk deret waktu yang stasioner. Hal ini terlihat dari nilai varian data tidak dipengaruhi oleh waktu atau nilai variannya konstan (tetap). Selanjutnya dilakukan uji kecocokan model AR(1) dan MA(1), dan diperoleh hasil bahwa model MA(1) yang cocok untuk meramalkan model tersebut untuk beberapa periode kedepannya. Adapun model yang diperoleh:  $Z_t = a_t - 0,4096a_{t-1} + 521,57$  Dengan  $a_t \sim N(0; 29192,72)$ .

#### E. Daftar Pustaka

- Apriani, W. (2020). Perbandingan Metode MA dan LS pada Transaksi Harian Penyetoran Tunai Bank BRI Sungai Liput. *Asimetris: Jurnal Pendidikan Matematika dan Sains*, 1(1), 37-42.
- Apriani, W. (2020). Efektivitas Blended Learning Berbantuan SPSS terhadap Tingkat Pemahaman Mahasiswa pada Mata Kuliah Statistik. *Asimetris: Jurnal Pendidikan Matematika dan Sains*, 1(1), 12-16.
- Apriani, W., Hayati, R. (2021). Metode ARIMA untuk Memodelkan Volume Produksi Kelapa Sawit pada PT. Socfindo di Kabupaten Aceh Tamiang. *Jurnal Absis: Jurnal pendidikan Matematika dan Matematika* 3(2), 309 – 319.
- Fithriasari, K., Iriawan, N., Ulama, B.S.S., Sutikno. (2013). On The Multivariate Time Series Rainfall Modeling Using Time Series Delay Neural Network. *International Journal of Applied Mathematics and Statistics*TM, 44 (14), 193-201.
- Huda, A. M., Choiruddin, A., Budianto, O., Sutikno. (2010). Peramalan Data Curah Hujan dengan Seasonal Autoregressive Integrated Moving Average (SARIMA) dengan Deteksi Outlier sebagai Upaya Optimalisasi Produksi Pertanian di Kabupaten Mojokerto. Surabaya: ITS.
- Purwawaputra, I. M. (2015). Peramalan Curah Hujan untuk Evaluasi Pola Tanam Berdasarkan Penanggalan Sasih di Bali. Surabaya: ITS.
- Wahyuni, R. (2020). Penerapan Model Time Series Untuk Meramalkan Nilai Un Matematika Di Man 4 Bireuen Tahun 2019. *Asimetris: Jurnal Pendidikan Matematika dan Sains*, 1(2), 70-76