PERAMALAN CURAH HUJAN BULANAN KABUPATEN TANAH DATAR DENGAN MODEL *SEASONAL AUTOREGRESSIVE INTEGRATED MOVING AVERAGE* (SARIMA)

FADILA AULIA, HAZMIRA YOZZA, DODI DEVIANTO

*Program Studi S1 Matematika,
Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Andalas,
Kampus UNAND Limau Manis Padang, Indonesia.
email : fadilaulia2012@gmail.com*

Diterima 22 Juni 2019 Direvisi 6 Juli 2019 Dipublikasikan 4 Agustus 2019

Abstrak. Curah hujan merupakan salah satu faktor penting yang mempengaruhi berbagai aspek kehidupan terutama dalam bidang pertanian. Mengetahui besarnya curah hujan untuk waktu yang akan datang dapat membantu proses perencanaan manusia dalam berbagai aspek dan dapat menanggulangi berbagai permasalahan yang timbul di kemudian hari. Besarnya curah hujan untuk waktu yang akan datang dapat diprediksi dengan melakukan proses peramalan. Data curah hujan merupakan suatu data runtun waktu. Proses peramalan data runtun waktu dapat dilakukan dengan berbagai metode, salah satunya dengan menggunakan model *Seasonal Autoregressive Integrated Moving Average* (SARIMA). Pada penelitian ini dilakukan proses peramalan curah hujan bulanan kabupaten Tanah Datar dengan menggunakan data curah hujan bulanan kabupaten Tanah Datar dari bulan Januari 2013 sampai bulan Oktober 2018 dan diperoleh model SARIMA terbaik yaitu $SARIMA(0, 0, 1)(0, 1, 0)^6$. Berdasarkan hasil peramalan yang diperoleh, besarnya curah hujan kabupaten Tanah Datar untuk bulan Desember 2018 sampai bulan April 2019 konstan.

Kata Kunci: Curah Hujan, Peramalan, SARIMA

1. Pendahuluan

Hujan adalah peristiwa jatuhnya butiran-butiran air dalam bentuk cair atau padat menuju bumi. Satuan yang digunakan dalam mengukur curah hujan adalah milimeter (mm). Besarnya curah hujan 1 mm memiliki arti bahwa jumlah air hujan yang jatuh di permukaan per satuan luas (m^2) jika zat cair tersebut tidak meresap kedalam tanah atau menguap. Sehingga besarnya curah hujan 1 mm setara dengan 1 liter/ m^2 [5].

Dalam kehidupan sehari-hari, curah hujan merupakan salah satu faktor penting yang mempengaruhi berbagai aspek kehidupan manusia. Besarnya curah hujan pada suatu daerah dipengaruhi oleh banyak faktor, diantaranya letak daerah konvergensi antar tropik (DKAT), arah angin, bentuk dan arah lereng medan, jarak perjalanan

angin di atas medan datar, letak geografis daerah, dan masih banyak faktor yang lain.

Untuk daerah agraris seperti kabupaten Tanah Datar, besarnya curah hujan dapat memberi pengaruh terhadap keberhasilan pertumbuhan dan produksi tanaman dalam sektor pertanian. Besarnya curah hujan pada suatu daerah juga dapat menentukan jenis komoditi dan teknik bercocok tanam yang cocok bagi daerah tersebut. Curah hujan yang tinggi dalam waktu yang lama pada suatu daerah dapat mengakibatkan terjadinya permasalahan seperti banjir dan tanah longsor sehingga diperlukan adanya prediksi terhadap besarnya curah hujan di suatu daerah untuk periode waktu yang akan datang.

Data curah hujan merupakan suatu data runtun waktu yang proses peramalannya dapat dilakukan dengan menggunakan berbagai metode. Salah satu metode yang dapat digunakan dalam peramalan curah hujan adalah model *Seasonal Autoregressive Integrated Moving Average* (SARIMA). Model SARIMA digunakan untuk data yang mengandung pola musiman.

Pada penelitian ini akan dilakukan peramalan curah hujan bulanan kabupaten Tanah Datar model *Seasonal Autoregressive Integrated Moving Average* (SARIMA).

2. Model *Seasonal Autoregressive Integrated Moving Average* (SARIMA)

2.1. Data Runtun Waktu

Data runtun waktu (*time series*) adalah jenis data yang dikumpulkan menurut urutan waktu dalam suatu rentang waktu tertentu. Suatu data runtun waktu dapat dinotasikan dengan $\{X_t\}$ dimana t adalah waktu pengamatan mulai dari $t = 1, 2, 3, \dots, n$ [4].

2.1.1. Autocorrelation Function (ACF)

Autokorelasi antara X_t dengan X_{t+k} dinyatakan sebagai berikut[6]:

$$\rho_k = \text{Corr}(X_t, X_{t+k}) = \frac{\gamma_k}{\gamma_0}. \quad (2.1)$$

Fungsi Autokorelasi (*ACF*) untuk populasi biasanya tidak diketahui sehingga diduga menggunakan fungsi autokorelasi sampel. Fungsi autokorelasi untuk sampel $X_1, X_2, X_3, \dots, X_n$ yaitu [6]:

$$\widehat{\rho}_k = \frac{\sum_{t=1}^{n-k} (X_t - \bar{X})(X_{t+k} - \bar{X})}{\sum_{t=1}^n (X_t - \bar{X})^2}. \quad (2.2)$$

2.1.2. Partial Autocorrelation Function (PACF)

Secara umum Fungsi Autokorelasi Parsial (*PACF*) dapat dituliskan sebagai [6]:

$$\Phi_{kk} = \frac{\rho_k - \sum_{j=1}^{k-1} \Phi_{k-1,j} \rho_{k-j}}{1 - \sum_{j=1}^{k-1} \Phi_{k-1,j} \rho_j}, \quad k = 1, 2, \dots. \quad (2.3)$$

2.2. Pemodelan Data Runtun Waktu

2.2.1. Pola Data Runtun Waktu

Dalam melakukan pemodelan data runtun waktu perlu memperhatikan tipe atau pola data. Secara umum terdapat empat macam pola data runtun waktu, yaitu horizontal, tren, musiman, dan siklis [2].

2.2.2. Stasioneritas

Suatu data runtun waktu dikatakan stasioner terhadap nilai tengah apabila data tersebut berfluktuasi disekitar nilai tengah yang konstan dari waktu ke waktu. Dalam mengatasi data yang tidak stasioner terhadap nilai tengah dapat dilakukan proses *differencing* (pembedaan) terhadap data.

Suatu data runtun waktu dapat dikatakan stasioner terhadap ragam apabila data tersebut berfluktuasi dengan ragam yang konstan dari waktu ke waktu. Dalam mengatasi data yang tidak stasioner terhadap ragam dapat dilakukan proses transformasi *Box-Cox* [1].

2.2.3. Model Seasonal Autoregressive Integrated Moving Average (SARIMA)

Model *Seasonal Autoregressive Integrated Moving Average (SARIMA)* merupakan suatu model ARIMA yang mengandung suatu pola musiman. Suatu model ARIMA untuk data runtun waktu X_t dapat dinotasikan dengan SARIMA $(p, d, q)(P, D, Q)^S$. Secara umum bentuk persamaan model SARIMA adalah sebagai berikut [6]:

$$\Phi_p(B)\Phi_p^*(B^S)(1-B)^d(1-B^S)^D X_t = \Theta_q(B)\Theta_q^*(B^S)\varepsilon_t. \quad (2.4)$$

2.2.4. Uji Signifikansi Parameter

Misalkan θ merupakan suatu parameter yang akan diuji, maka hipotesis yang digunakan adalah

$$\begin{aligned} H_0 : \theta &= 0, \quad (\text{parameter model tidak signifikan}), \\ H_1 : \theta &\neq 0, \quad (\text{parameter model signifikan}). \end{aligned}$$

Statistik uji yang digunakan adalah

$$t_{hitung} = \frac{\hat{\theta}}{SE(\hat{\theta})}.$$

2.2.5. Kriteria Pemilihan Model Terbaik

Terdapat beberapa kriteria pemilihan model terbaik diantaranya, yaitu [6]:

- (1) *Akaike's Information Criterion* (AIC).

Bentuk umum persamaan AIC adalah sebagai berikut:

$$AIC = n \ln \hat{\sigma}_a^2 + 2M. \quad (2.5)$$

(2) *Schwartz's Criterion* (SC).

Bentuk umum persamaan SC dalam proses pemilihan model terbaik adalah sebagai berikut:

$$SC = n \ln \hat{\sigma}_a^2 + M \ln n. \quad (2.6)$$

(3) *Hannan and Quinn Criterion* (HQ).

Bentuk umum persamaan HQ adalah sebagai berikut :

$$HQ = n \ln \hat{\sigma}_a^2 + 2 M \ln (\ln(n)). \quad (2.7)$$

2.2.6. *Uji Diagnostik Model*

Setelah didapatkan model terbaik selanjutnya akan dilakukan pengujian terhadap model yang didapatkan untuk mengetahui apakah model tersebut layak digunakan dalam proses peramalan. Suatu model dikatakan layak digunakan untuk peramalan apabila residual model bersifat *white noise*, terdistribusi normal dan homoskedastisitas.

Suatu residual model dikatakan bersifat *white noise* apabila tidak terdapat autokorelasi antara residual model. uji autokorelasi dapat dilakukan dengan menggunakan uji *Q-Ljung Box* dimana hipotesis yang digunakan adalah

$$\begin{aligned} H_0 : \rho_k &= 0, \quad (\text{tidak terdapat autokorelasi residual}), \\ H_1 : \rho_k &\neq 0, \quad (\text{terdapat autokorelasi antar residual}). \end{aligned}$$

Statistik uji yang digunakan adalah

$$Q = n(n+2) \sum_{k=1}^h \frac{\rho_k^2}{n-k},$$

dengan ρ_k adalah koefisien autokorelasi sisaan pada lag k dengan $k = 1, 2, \dots, h$ dan h adalah lag maksimum. H_0 akan ditolak jika nilai $Q > \chi_{\alpha}^2(h-p-q)$ atau $p - value < \alpha$.

Untuk menguji normalitas residual dapat digunakan uji hipotesis *Jarque-Berra*. Hipotesis yang digunakan yaitu

$$\begin{aligned} H_0 : & \text{residual tidak berdistribusi normal,} \\ H_1 : & \text{residual berdistribusi normal.} \end{aligned}$$

Statistik uji yang digunakan yaitu

$$JB = \frac{n}{6} \left(S^2 + \frac{(K-3)^2}{4} \right),$$

dengan n adalah ukuran contoh, S adalah *skewness* (kemenjuluran) dan K adalah *kurtosis* (keruncingan). H_0 ditolak jika nilai probabilitas dari *Jarque Berra* kecil dari tingkat signifikansi $\alpha = 0.05$.

Uji yang dapat digunakan untuk menguji heterokedastisitas yaitu uji *White*. Hipotesis yang digunakan adalah

$$\begin{aligned} H_0 : & \text{residual tidak heterokedastisitas,} \\ H_1 : & \text{residual heterokedastisitas.} \end{aligned}$$

Statistik uji yang digunakan yaitu

$$W = n.R^2,$$

dengan n adalah banyak data observasi dan R^2 adalah koefisien determinasi. Tolak H_0 jika nilai probabilitas dari *white* ($Obs * R^2$) kecil dari tingkat signifikansi $\alpha = 0.05$ [3].

3. Metode Penelitian

Data yang digunakan adalah data curah hujan bulanan kabupaten Tanah Datar provinsi Sumatera Barat dari bulan Januari 2013 sampai bulan Oktober 2018 (dalam mm). Dalam melakukan pengolahan data digunakan bantuan *software Eviews*. Adapun langkah-langkah peramalan curah hujan bulanan kabupaten Tanah Datar adalah sebagai berikut:

- (1) Identifikasi model.
- (2) Uji Signifikansi parameter model.
- (3) Uji diagnostik model.
- (4) Peramalan.

4. Peramalan dengan Model *Seasonal Autoregressive Integrated Moving Average* (SARIMA)

4.1. Identifikasi Model

Sebelum melakukan proses identifikasi model langkah pertama yang harus dilakukan adalah melakukan uji stasioneritas data terhadap nilai tengah dan uji stasioneritas data terhadap ragam. Uji stasioneritas dapat dilakukan dengan melihat hasil plot data. Gambar 1 merupakan plot data curah hujan bulanan kabupaten Tanah Datar.

Gambar 1. Plot Data Curah Hujan Bulanan Kabupaten Tanah Datar

Berdasarkan Gambar 1 dapat dilihat bahwa besarnya curah hujan bulanan kabupaten Tanah Datar mengalami fluktuasi tetapi tidak terlihat adanya kecenderungan tren sehingga dapat dikatakan bahwa data sudah stasioner terhadap nilai tengah. Untuk lebih menguatkan kesimpulan tersebut dapat dilakukan uji ADF untuk melihat kestasioneran data terhadap nilai tengah.

Berdasarkan hasil uji *Augmented Dickey Fuller* (ADF) diperoleh nilai statistik uji sebesar -5.703145 dan nilai kritis pada tabel *MacKinnon* untuk $\alpha = 5\%$ adalah

-2.904198 sehingga diperoleh kesimpulan bahwa data stasioner terhadap nilai tengah. Kemudian akan dilihat kestasioneran data terhadap ragam. Berdasarkan uji stasioneritas terhadap ragam diperoleh nilai *rounded value* (λ) sebesar 0.50 artinya data belum stasioner terhadap ragam sehingga perlu dilakukan transformasi *Box-Cox* dengan bentuk transformasi yaitu $\sqrt{X_t}$. Selanjutnya data yang telah ditransformasi diuji kembali kestasionerannya terhadap ragam dan diperoleh nilai *rounded value* (λ) sebesar 0.50 sehingga diperlukan transformasi $\sqrt{X_t}$ sekali lagi. Setelah dilakukan transformasi sebanyak dua kali diperoleh nilai λ sebesar 1 sehingga dapat dikatakan bahwa data sudah stasioner terhadap ragam.

Setelah dilakukan uji stasioneritas terhadap data langkah selanjutnya yang dilakukan adalah melakukan identifikasi model. Proses identifikasi model dilakukan dengan melihat hasil plot ACF dan PACF dari data yang telah distasionerkan.

Berdasarkan hasil plot ACF dan PACF dari data yang telah distasionerkan diperoleh bahwa plot ACF signifikan pada *lag* ke-1 sehingga kemungkinan orde $q = 1$ dan pada plot PACF signifikan pada *lag* ke-1 sehingga kemungkinan orde $p = 1$, karna tidak adanya proses *differencing* maka orde $d = 0$ sehingga diperoleh model awal yaitu $ARIMA(1,0,1)$.

Berdasarkan plot ACF terlihat adanya pola musiman sehingga perlu adanya proses *differencing* data untuk menghilangkan pola musiman tersebut. Kemudian akan dilihat plot ACF dan PACF untuk data yang sudah dilakukan proses *differencing*.

Berdasarkan plot ACF dan PACF untuk data yang sudah dilakukan proses *differencing* diperoleh bahwa plot ACF signifikan pada *lag* ke-1 artinya diperoleh orde $q = 1$ sedangkan pada plot PACF signifikan pada *lag* ke-2 sehingga kemungkinan orde $p = 2$ dan adanya proses *differencing* maka didapatkan model awal yaitu $ARIMA(2,1,1)$. Berdasarkan proses tersebut diperoleh model awal untuk data curah hujan kabupaten Tanah Datar yaitu $SARIMA(1,0,1)(2,1,1)^6$.

4.2. Uji Signifikansi Parameter Model

Berdasarkan model awal yang diperoleh selanjutnya akan ditentukan semua kemungkinan model yang sesuai dan dilakukan uji signifikansi parameter untuk setiap kemungkinan model yang diperoleh. Dari hasil uji signifikansi parameter diperoleh model-model yang signifikan yaitu $SARIMA(1,0,0)(1,1,1)^6$, $SARIMA(1,0,0)(0,1,0)^6$, $SARIMA(0,0,1)(0,1,0)^6$.

Setelah diperoleh model-model yang signifikan dari semua kemungkinan model, selanjutnya akan ditentukan model terbaik. Pemilihan model terbaik dapat dilakukan dengan melihat nilai dari *Akaike Info Criterion* (*AIC*), *Schwarz Criterion* (*SC*) dan *Hannan and Quinn Criterion* (*HQ*). Model terbaik adalah model yang memiliki nilai *Akaike Info Criterion* (*AIC*), *Schwarz Criterion* (*SC*) dan *Hannan and Quinn Criterion* (*HQ*) terkecil.

Diperoleh model terbaik yaitu $SARIMA(0,0,1)(0,1,0)^6$ dengan persamaan model:

$$(1 - B^{12}) \sqrt{X_t} = \Theta_1(B) \varepsilon_t = (1 - \theta_1 B) \varepsilon_t,$$

$$\begin{aligned}\sqrt{X_t} &= \sqrt{X_{t-12}} + \varepsilon_t - \theta_1 \varepsilon_{t-1}, \\ \sqrt{X_t} &= \sqrt{X_{t-12}} + \varepsilon_t + 0.51339 \varepsilon_{t-1}.\end{aligned}$$

4.3. Uji Diagnostik Model

Setelah diperoleh model terbaik, dilakukan uji diagnostik model untuk mengetahui apakah model terbaik layak digunakan untuk peramalan. Suatu model dikatakan layak digunakan untuk peramalan apabila residual model bersifat *white noise*, terdistribusi normal dan tidak heteroskedastisitas.

Uji autokorelasi dapat digunakan untuk menentukan apakah residual model bersifat *white noise*. Suatu residual model dikatakan bersifat *white noise* apabila tidak terjadi autokorelasi. Salah satu uji yang dapat digunakan untuk menguji autokorelasi adalah uji *Q-Ljung Box*. Nilai uji *Q-Ljung Box* dapat dilihat dari nilai probabilitas Q-Stat pada korelogram residual model. Jika nilai probabilitas Q-Stat kurang dari $\alpha = 0.05$ maka dapat dikatakan terjadi autokorelasi antar residual model.

Dari korelogram residual model diperoleh nilai probabilitas Q-Stat dari residual model terbaik lebih dari $\alpha = 0.05$ maka dapat dikatakan tidak terjadi autokorelasi antar residual model sehingga model terbaik bersifat *white noise*.

Untuk mengetahui residual model terbaik terdistribusi normal atau tidak dapat dilihat dari histogram normalitas residual model terbaik. Suatu residual model dikatakan terdistribusi normal jika nilai probabilitas residual model lebih besar dari $\alpha = 0.05$. Dari histogram normalitas residual model terbaik diperoleh nilai probabilitas lebih besar dari $\alpha = 0.05$ maka dapat dikatakan residual model terdistribusi normal.

Heteroskedastisitas dari residual model dapat dilihat dari korelogram kuadrat residual model. Suatu residual model dikatakan tidak heteroskedastisitas jika nilai probabilitas kuadrat residual model lebih besar dari $\alpha = 0.05$. Dari korelogram kuadrat residual model diperoleh nilai probabilitas kuadrat residual model lebih besar dari $\alpha = 0.05$ sehingga dapat dikatakan bahwa residual model tidak heteroskedastisitas.

Berdasarkan ketiga uji tersebut dapat disimpulkan bahwa model SARIMA $(0, 0, 1)(0, 1, 0)^6$ layak digunakan untuk peramalan.

4.4. Peramalan

Setelah dilakukan uji diagnostik model, selanjutnya dilakukan peramalan terhadap data dengan menggunakan data yang telah distasionerkan. Diperoleh hasil peramalan curah hujan bulanan kabupaten Tanah Datar.

5. Kesimpulan

Berdasarkan hasil pengolahan data dapat disimpulkan bahwa:

- (1) Model SARIMA terbaik untuk peramalan curah hujan bulanan kabupaten

Tabel 1. Data Hasil Peramalan

Bulan	Hasil Peramalan
Nov 2018	187.8736
Des 2018	173.9631
Jan 2019	173.9631
Feb 2019	173.9631
Mar 2019	173.9631
Apr 2019	173.9631

Tanah Datar adalah $SARIMA(0, 0, 1)(0, 1, 0)^6$ dengan model persamaan:

$$\sqrt{X_t} = \sqrt{X_{t-12}} + \varepsilon_t + 0.51339 \varepsilon_{t-1}.$$

- (2) Berdasarkan hasil peramalan diperoleh curah hujan bulanan kabupaten Tanah Datar dari bulan Desember 2018 sampai April 2019 konstan.

Daftar Pustaka

- [1] Devianto, Dodi, dkk. 2018. *Time Series of Rainfall Model with Markov Switching Autoregressive*. IEEE : Page: 202 – 207
- [2] Hanke, J.E., and Wichern, D.W. 2005. *Business Forecasting*. Prentice Hall, New York
- [3] Iswahyuli, Hazmira Yozza dan Dodi Devianto. 2018. Peramalan Curah Hujan Bulanan Desa Sungai Ipuh Solok Selatan dengan Model Autoregressive Integrated Moving Average. *Jurnal Matematika UNAND*. Vol. **VII**(1): 85 – 92
- [4] Rosadi, Dedi. 2006. *Pengantar Analisis Runtun Waktu*. Penerbit Universitas Gadjah Mada, Yogyakarta
- [5] Tjasyono, Bayong. 1999. *Klimatologi Umum*. Penerbit ITB, Bandung
- [6] Wei. W. W. S. 2006. *Time Series Analysis: Univariate and Multivariate Methods*. 2nd ed. Pearson, New York