

Peramalan Curah Hujan di Stasiun Juanda Menggunakan Metode ARIMA *Box-Jenkins* dan *Radial Basis Function Neural Network*

Ocktalia Trisnawati dan Mike Prastuti

Departemen Statistika Bisnis, Institut Teknologi Sepuluh Nopember (ITS)

e-mail: mike_p@statistika.its.ac.id

Abstrak—Curah hujan yang tinggi tak hanya menyebabkan penundaan penerbangan, namun juga membuat pesawat terlambat mendarat karena menunggu cuaca sedikit lebih baik agar dapat dilakukan pendaratan dengan aman. Curah hujan juga mempengaruhi pendaratan dan lepas landas pesawat, terlebih lagi jika sering terjadi hujan maka akan membuat landasan pacu pesawat menjadi licin dan mengurangi jarak pandang pengelihat. Dalam upaya untuk menghindari adanya kecelakaan pesawat, maka pada penelitian ini dilakukan peramalan terhadap curah hujan di Stasiun BMKG Juanda menggunakan metode ARIMA *Box-Jenkins* dan *Radial Basis Function Neural Network* (RBFNN). Tujuan dari penelitian ini adalah untuk mendapatkan model terbaik dan hasil peramalan curah hujan di Stasiun BMKG Juanda. Data yang digunakan adalah data sekunder yang diperoleh dari BMKG Juanda Surabaya dari Januari 2019 hingga April 2021. Hasil penelitian ini menunjukkan bahwa nilai MSE model RBFNN dengan 12 *neuron input*, 3 *neuron* lapisan tersembunyi, dan 1 *neuron output* lebih kecil daripada model ARIMA ([1,9,12,35,55,58][1,1]), sehingga dapat disimpulkan bahwa metode RBFNN merupakan metode terbaik yang digunakan untuk meramalkan curah hujan di Stasiun Juanda.

Kata Kunci—ARIMA *Box-Jenkins*, Curah Hujan, Juanda, Peramalan, RBFNN.

I. PENDAHULUAN

CUACA merupakan suatu kondisi udara di suatu tempat pada waktu yang relatif singkat, yang dinyatakan dengan nilai berbagai parameter seperti suhu, tekanan udara, kecepatan angin, kelembaban udara, dan berbagai fenomena atmosfer lainnya. Cuaca merupakan aspek penting dalam aktivitas penerbangan.

Keselamatan merupakan prioritas utama dalam sarana transportasi khususnya bidang penerbangan. Keselamatan penerbangan berkaitan dengan banyak faktor diantaranya ada faktor manusia, faktor kondisi dan jenis pesawat terbang, fasilitas dan sarana bandar udara, fasilitas dan sarana telekomunikasi, dan faktor cuaca. Kemungkinan penyebab kecelakaan pesawat terbang 30% antara lain karena faktor cuaca, sedangkan 70% dari faktor manusia.

Dalam dunia penerbangan, terdapat cuaca ekstrem yang mempunyai potensi membahayakan seperti curah hujan tinggi yang disertai angin kencang, gelombang gunung, dan *turbulence* didalam awan. Curah hujan tinggi mempengaruhi pendaratan dan lepas landas pesawat, terlebih lagi jika sering terjadi hujan di bandara akan membuat landasan pacu menjadi licin dan mengurangi jarak pandang pengelihat. Hal ini bisa menyebabkan penundaan penerbangan atau pesawat terlambat mendarat, karena pesawat menunggu cuaca sedikit lebih baik sehingga dapat dilakukan pendaratan dengan aman. Apabila cuaca tak kunjung membaik, maka

akan dilakukan pendaratan di bandara terdekat atau kembali pada bandara pemberangkatan. Keberadaan awan di atmosfer pada musim hujan akan memberikan guncangan pada pesawat terbang yang disebut dengan *turbulenci*.

Badan Meteorologi Klimatologi dan Geofisika (BMKG) Juanda mempunyai tugas melaksanakan fungsi BMKG di bidang Meteorologi khususnya untuk menunjang keselamatan transportasi udara. Selain itu, BMKG menyediakan informasi cuaca harian untuk beberapa kota di Jawa Timur.

Dalam upaya untuk menghindari adanya kecelakaan pesawat, pemilihan metode yang tepat untuk menentukan kondisi cuaca seperti curah hujan adalah kegiatan yang akhir-akhir ini sering dilakukan oleh beberapa peneliti atmosfer atau cuaca. Hal ini dikarenakan banyaknya tuntutan dari berbagai pihak yang menginginkan informasi mengenai kondisi atmosfer yang lebih cepat, akurat, dan terperinci. Bahkan beberapa pihak lain menuntut tersedianya prediksi atau bahkan ramalan mengenai kondisi atmosfer dengan rentang waktu yang cukup pendek seperti waktu harian, jam, bahkan dalam waktu menit.

Metode ARIMA mengasumsikan data yang digunakan berhubungan secara linier yang sangat baik ketepatan akurasi jika digunakan untuk peramalan jangka pendek, sedangkan kurang akurat untuk peramalan jangka panjang. RBFNN merupakan salah satu metode pada jaringan saraf tiruan yang terdiri dari lapisan input, lapisan tersembunyi, dan lapisan output. RBFNN memiliki ciri khas yaitu memiliki satu lapisan tersembunyi dan mengeluarkan nilai berupa persamaan linier pada lapisan output yang membuat proses kerja RBFNN lebih cepat daripada jaringan saraf tiruan lainnya. Beberapa penelitian sebelumnya menggunakan metode ARIMA dalam meramalkan data curah hujan, dimana hasil peramalan memiliki nilai error yang cukup besar. Hal ini dikarenakan data curah hujan memiliki pola non-linier yang tidak dapat ditangkap oleh model ARIMA. Untuk mengetahui pendugaan peramalan terbaik, maka peramalan curah hujan di Stasiun Juanda akan dibandingkan dengan metode ARIMA dan RBFNN.

II. TINJAUAN PUSTAKA

A. ARIMA *Box-Jenkins*

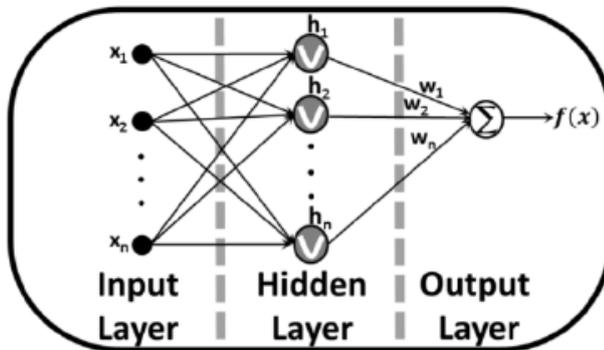
Metode *time series* adalah suatu metode peramalan untuk masa depan yang dilakukan berdasarkan nilai atau data masa lalu dari suatu variabel dan kesalahan (*error*) masa lalu. Tujuan dari metode peramalan *time series* ini adalah untuk menemukan pola data *time series* (runtun waktu) dan mengekstrapolasikan pola tersebut ke masa depan [1].

Tabel 1.
Nilai λ dan transformasinya

Nilai λ	Transformasi
-1,0	$1/Z_t$
-0,5	$1/\sqrt{Z_t}$
0,0	$\ln Z_t$
0,5	$\sqrt{Z_t}$
1,0	Z_t (Tidak perlu ditransformasi)

Tabel 2.
Struktur ACF dan PACF pada model ARIMA

Model	ACF	PACF
Autoregressive (p)	Turun eksponensial	Terpotong setelah lag-p
Moving Average (q)	Terpotong setelah lag-q	Turun eksponensial
Autoregressive Moving Average (p,q)	Turun eksponensial	Turun eksponensial



Gambar 1. Arsitektur RBFNN.

B. Stasioneritas Time Series

Stasioneritas *time series* adalah asumsi yang mendasari bahwa proses suatu deret pengamatan tidak berubah seiring dengan adanya perubahan waktu. Dengan kata lain, data *time series* dapat dikatakan stasioner adalah relatif berfluktuasi konstan disekitar *mean* [2].

Seringkali *time series* tidak stasioner dalam *mean* akibat adanya pola seiring dengan perubahan waktu. Suatu deret yang nonstasioner dapat direduksi menjadi deret yang stasioner melalui *differencing* menggunakan operator *shift* mundur (*backward shift*) yang dinotasikan Δ^d atau dapat ditulis $(1-B)^d$ dengan

$$\Delta Z_t = Z_t - Z_{t-1} \tag{1}$$

Ketika data tidak stasioner dalam *varians*, dilakukan transformasi untuk menstabilkan *varians* atau membuat *varians* menjadi homogen. Salah satu tranformasi yang bisa digunakan adalah *Power Transformation*. Secara umum, *Power Transformation* diberikan sebagai berikut.

$$T(Z_t) = \begin{cases} \frac{Z_t^\lambda - 1}{\lambda} & ; \lambda \neq 0 \\ \ln Z_t & ; \lambda = 0 \end{cases} \tag{2}$$

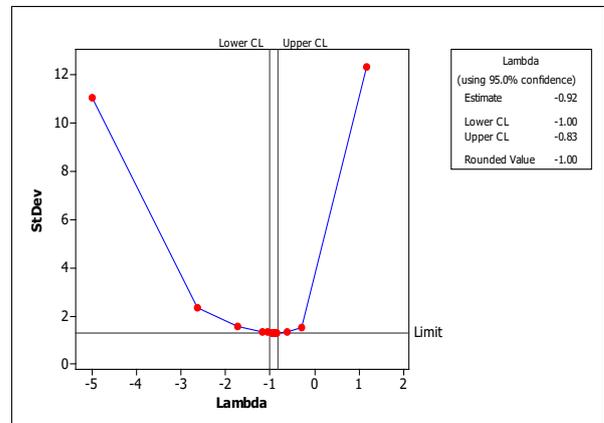
Tabel 1 adalah nilai λ yang sering digunakan dan bentuk transformasinya. Data dikatakan stasioner dalam *varians* apabila batas bawah dan batas atas dari transformasi *Box-Cox*

Tabel 3.
Variabel penelitian

Tahun	Bulan	Tanggal	Curah Hujan (mm)	
2019	Januari	1	Z_1	
	:	2	Z_2	
		:	:	:
	Desember	1	Z_{32}	
		2	Z_{33}	
	:	:	:	Z_{365}
:			:	
Januari		1	Z_{732}	
		2	Z_{733}	
2021		:	:	:
			1	Z_{822}
	April	:	Z_{851}	

Tabel 4.
Statistika deskriptif data curah hujan di Stasiun Juanda

Tahun	Mean	Variance	Minimum	Maximum
2019	5.439	227.651	0	156.600
2020	8.666	314.553	0	127.700
2021*	12.510	316.710	0	89.800



Gambar 2. Plot *Box-Cox* data curah hujan di Stasiun Juanda.

memuat nilai $\lambda=1$, sebaliknya dikatakan tidak stasioner dalam *varians* apabila batas bawah dan batas atas tidak memuat angka 1 [3].

C. ACF dan PACF

Autocorretaltion Function (ACF) digunakan untuk menunjukkan hubungan linear antara Z_t dan Z_{t+k} . Suatu proses stasioner $\{Z_t\}$ mempunyai nilai $E(Z_t)=\mu$ dan $Var(Z_t) = E(Z_t - \mu)^2 = \sigma^2$, dimana nilai *mean* dan *varians* tersebut adalah konstan [2]. Persamaan dari kovarians antara Z_t dan Z_{t+k} ditunjukkan dalam Persamaan 3 sebagai berikut:

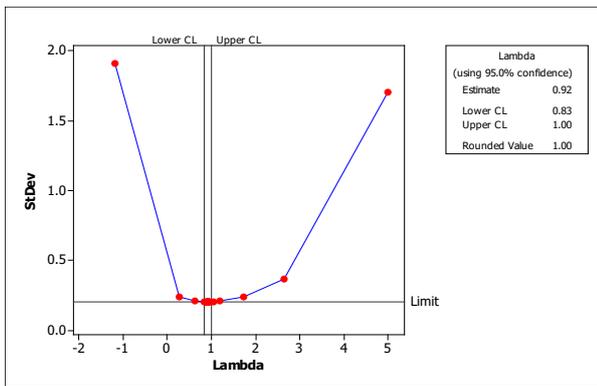
$$\gamma_k = Cov(Z_t, Z_{t+k}) = E(Z_t - \mu)(Z_{t+k} - \mu) \tag{3}$$

dan korelasi antara Z_t dan Z_{t+k} ditunjukkan dalam Persamaan 4 sebagai berikut:

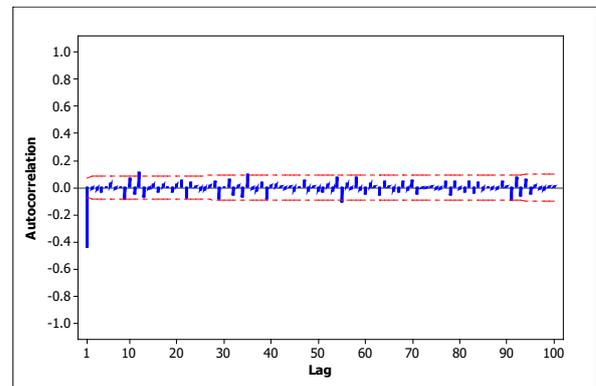
$$\rho_k = \frac{Cov(Z_t, Z_{t+k})}{\sqrt{Var(Z_t)}\sqrt{Var(Z_{t+k})}} = \frac{\gamma_k}{\gamma_0} \tag{4}$$

dimana γ_k adalah fungsi autokovarian dan ρ_k adalah fungsi autokorelasi antara Z_t dengan Z_{t+k} dari proses yang sama dan hanya terpisah oleh *lag* waktu ke-k.

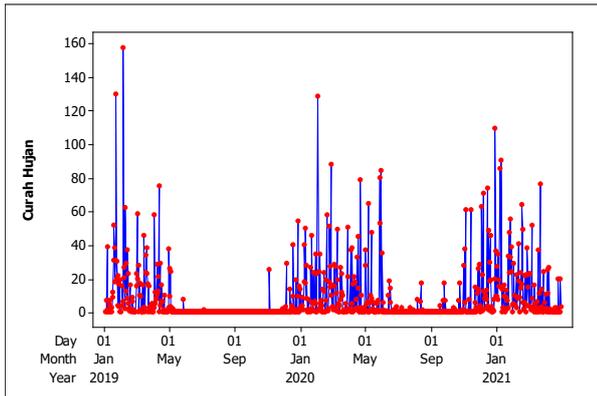
Partial Autocorrelation Function (PACF) digunakan untuk mengukur tingkat keeratan hubungan antara pasangan data Z_t dengan Z_{t+k} setelah pengaruh variabel $Z_{t+1}, Z_{t+2}, Z_{t+k-1}$ terhadap Z_{t+k} dihilangkan [2]. PACF antara Z_t dan Z_{t+k} bernilai sama dengan autokorelasi antara $(Z_t - \hat{Z}_t)$ dan



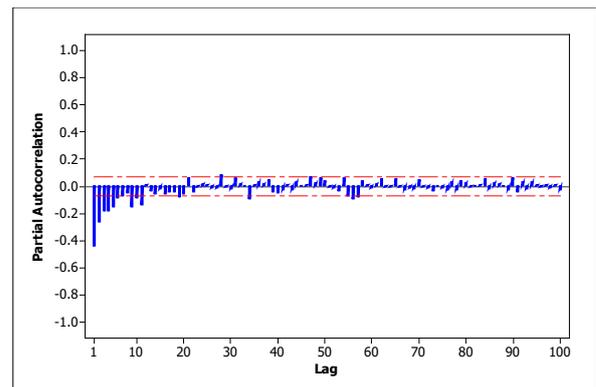
Gambar 3. Plot *box-cox* setelah transformasi.



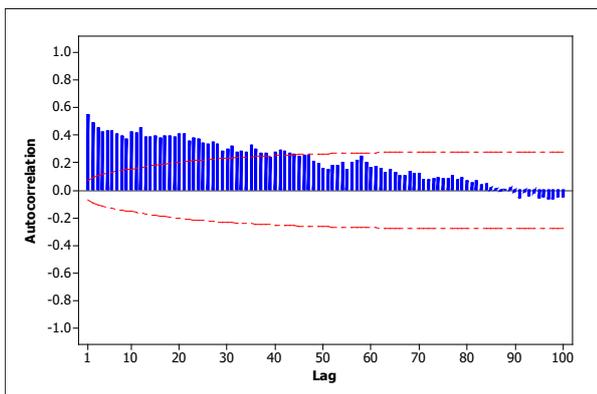
Gambar 6. Plot ACF data curah hujan setelah di *Differencing* 1.



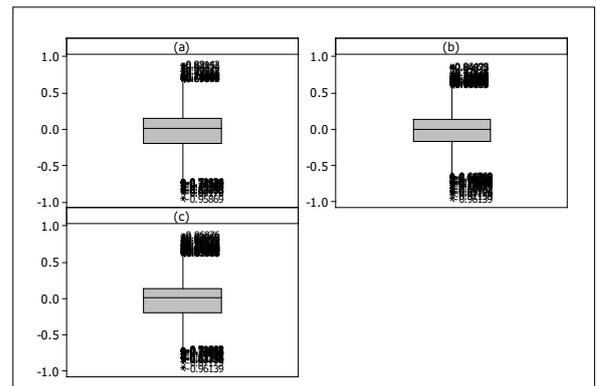
Gambar 4. *Time series plot* data curah hujan di Stasiun Juanda.



Gambar 7. Plot PACF data curah hujan setelah di *Differencing* 1.



Gambar 5. Plot ACF data curah hujan.



Gambar 8. *Boxplot* residual. (a) ARIMA ([1,12,35,58],1,[1,50]), (b) ARIMA ([1,9,12,35,55,58],1,1), (c) ARIMA (1,1,1).

$(Z_{t+k} - \hat{Z}_{t+k})$, sehingga fungsi autokorelasi parsial dapat dihitung dengan rumus sebagai berikut:

$$\hat{\phi}_{k+1,k+1} = \frac{\hat{\rho}_{k+1} - \sum_{j=1}^k \hat{\phi}_{kj} \hat{\rho}_{k+1-j}}{1 - \sum_{j=1}^k \hat{\phi}_{kj} \hat{\rho}_j} \quad (5)$$

dengan $\hat{\phi}_{k+1,k+1} = \hat{\phi}_{kj} - \hat{\phi}_{k+1,k+1} \hat{\phi}_{k,k+1-j}$ dan $j = 1, 2, \dots, k$.

D. Model ARIMA

Identifikasi model ARIMA dilakukan setelah data stationer dengan melihat pola ACF dan PACF. Pendugaan model dilakukan dengan memperhatikan hal-hal yang tertera dalam Tabel 2.

E. Pengujian Signifikansi Parameter

Pengujian signifikansi parameter dilakukan untuk mengetahui apakah hasil penaksiran parameter model ARIMA signifikan atau tidak, sehingga dapat diketahui apakah tiap variabel yang digunakan berpengaruh pada Z_t . pengujian hipotesis dilakukan dengan menggunakan uji- t

dengan hipotesis sebagai berikut [4].

Hipotesis Model AR:

$H_0: \phi_p = 0$ (parameter AR tidak signifikan terhadap model)

$H_1: \phi_p \neq 0$ (parameter AR signifikan terhadap model)

Statistik uji untuk model AR ditunjukkan pada Persamaan 6 berikut:

$$t_{hitung} = \frac{\hat{\phi}_p}{se(\hat{\phi}_p)} \quad (6)$$

Hipotesis Model MA:

$H_0: \theta_q = 0$ (parameter MA tidak signifikan terhadap model)

$H_1: \theta_q \neq 0$ (parameter MA signifikan terhadap model)

Statistik uji untuk model MA ditunjukkan pada Persamaan 7 berikut:

$$t_{hitung} = \frac{\hat{\theta}_p}{se(\hat{\theta}_p)} \quad (7)$$

dengan daerah penolakan adalah tolak H_0 jika $|t_{hitung}| > t_{\alpha/2, n-p}$ atau tolak H_0 jika $P_{value} < \alpha$ dengan α adalah taraf

Tabel 5.
Pemilihan model ARIMA terbaik

No	Model ARIMA	MSE
1	ARIMA ([1,12,35,58],1,[1,50])	39.828
2	ARIMA ([1,9,12,35,55,58],1,1)	16.902
3	ARIMA (1,1,1)	20.191

Tabel 6.
Uji linearitas

F_{hitung}	$F_{0,05;2;848}$	P_{value}
20,212	3,006	0,000

Tabel 7.
Penentuan jaringan optimum

Banyak Kelompok	MSE	
	Training	Testing
2	85473.982	9546.458
3	85462.214	9538.503
4	85510.244	9544.602
5	85484.277	9540.051
6	85499.508	9544.391

Tabel 8.
Perbandingan kebaikan model ARIMA dan RBFNN

Model	MSE
ARIMA ([1,9,12,35,55,58],1,1)	16.902
RBFNN	5.270

Tabel 9.
Ramalan curah hujan di Stasiun Juanda Periode 852-865

Periode	Nilai Ramalan	Periode	Nilai Ramalan
852	10.732	859	13.388
853	13.265	860	13.403
854	13.330	861	7.981
855	8.809	862	9.611
856	13.193	863	9.523
857	13.431	864	11.950
858	13.303	865	13.041

signifikansi, p adalah banyaknya parameter AR, q adalah banyaknya parameter MA, $se(\hat{\beta})$ adalah standar error dari nilai taksiran ϕ atau θ .

F. Cek Diagnosa Residual

White noise merupakan asumsi dimana gangguan pada residual telah dihilangkan. Jika residual model telah memenuhi asumsi white noise, berarti residual tersebut saling independen dan identik. Pemeriksaan asumsi residual white noise dilakukan dengan Uji Ljung-Box dengan hipotesis sebagai berikut [2].

$$H_0: \rho_1 = \rho_2 = \dots = \rho_K = 0$$

H_1 : minimal ada satu $\rho_j \neq 0$, dengan $j=1,2, \dots, K$

Statistik uji yang digunakan sebagai berikut:

$$Q = n(n + 2) \sum_{k=1}^K (n - k)^{-1} \hat{\rho}_k^2 \tag{8}$$

keterangan:

n = banyaknya pengamatan

$\hat{\rho}_k$ = autokorelasi lag ke- k

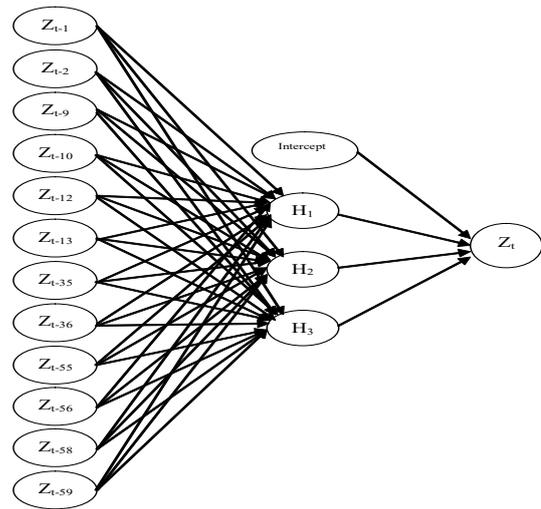
K = jumlah lag maksimum

H_0 akan ditolak jika $Q > \chi^2_{\alpha,(K-m)}$ atau tolak H_0 jika $P_{value} < \alpha$.

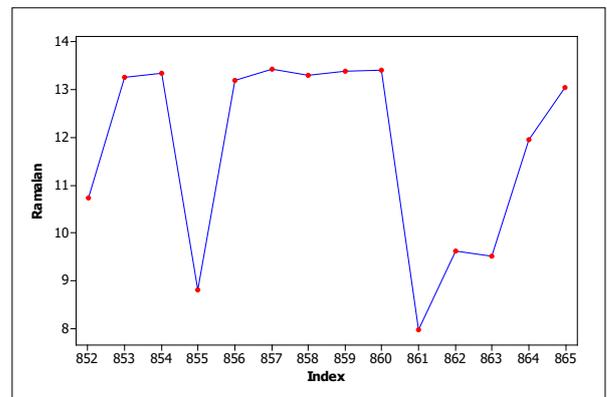
Selanjutnya akan dilakukan pemeriksaan diagnostik dengan pengujian asumsi residual berdistribusi normal dengan Uji Kolmogorov-Smirnov dengan hipotesis sebagai berikut:

$$H_0 : F_n(x) = F_0(x)$$

$$H_1 : F_n(x) \neq F_0(x)$$



Gambar 9. Arsitektur jaringan RBFNN.



Gambar 10. Time series plot nilai ramalan model RBFNN.

Statistik uji yang digunakan adalah sebagai berikut:

$$D = \max|F(x) - F_0(x)| \tag{9}$$

keterangan:

$F(x)$ = nilai peluang kumulatif residual

$F_0(x)$ = nilai peluang kumulatif distribusi normal

H_0 akan ditolak jika nilai $D > D_{(1-\alpha),n}$ atau tolak H_0 jika

$P_{value} < \alpha$, dimana n sebagai derajat bebasnya.

G. Pemilihan Model Terbaik

Pemilihan model terbaik dilakukan dengan membandingkan nilai Mean Square Error (MSE). Nilai MSE digunakan untuk mengukur ketepatan nilai dugaan model yang dinyatakan dalam rata-rata kuadrat kesalahan [2]. Perhitungan nilai MSE dilakukan dengan rumus sebagai berikut:

$$MSE = \frac{1}{n} \sum_{t=1}^n |Z_t - \hat{Z}_t|^2 \tag{10}$$

H. Uji Linearitas

Uji linearitas dapat dilakukan dengan Uji Terasvitra. Uji Terasvitra merupakan pengujian linearitas tipe LM (Lagrange Multiplier). Prosedur pengujian linieritas dengan Uji Terasvitra adalah sebagai berikut:

- 1) Regresikan Z_t pada 1 dan Z_{t-k} , $k=1,2,\dots,p$. Sehingga diperoleh model linear

$$Z_t = f_t - \hat{\epsilon}_t \tag{11}$$

- 2) Meregresikan $\hat{\epsilon}_t$ pada 1, $Z_{t-1}, Z_{t-2}, \dots, Z_{t-p}$ dan kemudian menghitung jumlah kuadrat nilai residual $SSR_1 = \sum \hat{v}_t^2$.
- 3) Menghitung nilai statistik uji sebagai berikut:

$$F_{hitung} = \frac{(SSR_0 - SSR_1)/m}{SSR_1/(n-p-1-m)} \quad (12)$$

Keterangan:

p = prediktor awal

m = prediktor tambahan ($m=2$)

n = jumlah pengamatan

\hat{v}_t = residual ke- t

SSR = Sum Square Residual

I. Radial Basis Function Neural Network

Radial Basis Function Neural Network (RBFNN) memiliki 3 layer, yaitu *input layer*, *hidden layer*, dan *output layer* [5]. Arsitektur jaringan RBFNN dijelaskan pada Gambar 1. *Output* di rumuskan sebagai berikut:

$$\hat{y} = \sum_{j=1}^m w_j h_j(x) + w_0 \quad (13)$$

fungsi aktivasi dengan RBFNN dinyatakan sebagai berikut:

$$h_j(x) = \exp\left(-\sum_{i=1}^k \left(\frac{\|x_i - c_j\|^2}{2\sigma_j^2}\right)\right) \quad (14)$$

dimana,

$$\sigma_j = \sqrt{\frac{1}{p} \sum_{p=1}^p \sigma_{jp}^2}$$

w_j = parameter pada node ke- j

h_j = fungsi aktivasi pada node ke- j

c_j = vektor *center* pada node ke- j

σ_j = standar deviasi pada node ke- j

x = vektor *input*

III. METODOLOGI PENELITIAN

A. Sumber Data

Data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh dari BMKG Juanda Surabaya www.dataonline.bmkg.go.id dengan unit penelitian curah hujan harian mulai Januari 2019 hingga April 2021 sebanyak 851 data. Data curah hujan ini akan dibagi menjadi dua, yaitu data *in-sample* dan data *out-sample*. Data *in-sample* dimulai dari Januari 2019 hingga Desember 2020 (731 data) dan data *out-sample* dimulai dari Januari 2021 hingga April 2021 (120 data).

B. Variabel Penelitian

Variabel yang digunakan pada penelitian ini akan ditampilkan pada Tabel 3.

C. Langkah Analisis

Langkah-langkah yang digunakan dalam menganalisis data, yaitu langkah analisis dengan metode ARIMA *Box-Jenkins* dan metode RBFNN akan dijelaskan sebagai berikut: melakukan *pre-processing* data curah hujan di Stasiun Juanda, mendapatkan karakteristik data curah hujan di Stasiun BMKG Juanda menggunakan statistika deskriptif. Selanjutnya melakukan peramalan curah hujan menggunakan

metode ARIMA *Box-Jenkins*, berikut adalah langkah-langkah yang akan dilakukan:

- 1) Membuat *time series* plot pada data *in-sampel* untuk melakukan identifikasi pola *time series* pada data curah hujan
- 2) Melakukan identifikasi stasioneritas data, yaitu stasioner terhadap varian dan stasioner terhadap *mean*. Jika tidak stasioner terhadap varian maka akan dilakukan transformasi *Box-Cox*, dan jika tidak stasioner terhadap *mean* akan dilakukan *differencing*
- 3) Membuat plot ACF dan PACF untuk melihat pola data
- 4) Identifikasi dan pendugaan model ARIMA berdasarkan pada plot ACF dan PACF.
- 5) Estimasi parameter dan pengujian signifikansi parameter
- 6) Pengujian asumsi residual *white noise* dan berdistribusi normal pada model yang terbentuk
- 7) Mendapatkan model peramalan terbaik menggunakan kriteria MSE.

Melakukan peramalan curah hujan menggunakan metode RBFNN, berikut adalah langkah-langkah yang akan dilakukan:

- 1) Menentukan variabel *input* dengan melihat plot PACF dari data *training*. Variabel input ditentukan oleh *lag-lag* yang nilai PACF-nya signifikan
- 2) Menentukan nilai pusat dengan menggunakan metode kluster *k-means* dan menghitung nilai varian dari hasil pengelompokan
- 3) Menentukan jaringan yang optimum dengan jaringan yang memiliki nilai MSE terkecil

Terakhir yaitu memilih model peramalan terbaik berdasarkan model ARIMA dan RBFNN menggunakan nilai MSE.

IV. ANALISIS DAN PEMBAHASAN

A. Statistika Deskriptif Data Curah Hujan di Stasiun Juanda

Data yang digunakan merupakan data harian yang kemudian akan dilakukan *pre-processing* data dengan imputasi terhadap data kosong, dari 851 data ada sekitar 52 data yang kosong dan penambahan nilai 1 pada semua data agar tidak ada data yang bernilai nol. Berdasarkan data curah hujan di Stasiun Juanda diperoleh hasil statistika deskriptif yang dijelaskan pada Tabel 4.

Tabel 4 menjelaskan bahwa Tahun 2019 curah hujan maksimum sebanyak 156.600 mm (03 Februari) dengan keragaman data sebesar 227.651. Tahun 2020 curah hujan maksimum sebanyak 127.700 mm (01 Februari) dengan keragaman data sebesar 314.553. Tahun 2021 curah hujan maksimum sebanyak 89.800 mm (08 Januari) dengan keragaman data sebesar 316.710. Pada Tahun 2019 tidak hujan terjadi 252 hari dari 365 hari, Tahun 2020 tidak hujan terjadi 168 hari dari 366 hari, dan tahun 2021 tidak hujan terjadi 27 hari dari 120 hari.

B. Pemodelan Data Curah Hujan di Stasiun Juanda Menggunakan ARIMA Box-Jenkins

Tahap awal pendugaan model ARIMA adalah memeriksa stasioneritas data terhadap *varians* dan *mean* karena model ARIMA mengasumsikan bahwa data yang digunakan stasioner dalam *varians* dan *mean*. Lebih jelasnya akan dilakukan estimasi parameter menggunakan *box-cox* (λ)

dapat dilihat pada Gambar 2.

Gambar 2 menjelaskan hasil plot *box-cox* data curah hujan di Stasiun Juanda memiliki nilai λ tidak sama dengan 1, serta batas atas dan batas bawah tidak melewati nilai 1 sehingga dapat disimpulkan data curah hujan di Stasiun Juanda belum stasioner dalam *varians*. Nilai λ yang diperoleh sebesar -1,00 sehingga data akan di transformasi dengan $T(Z_t) = \frac{1}{Z_t}$. Hasil transformasi *box-cox* dapat dilihat pada Gambar 3.

Gambar 3 menunjukkan hasil plot *box-cox* data curah hujan di Stasiun Juanda setelah transformasi diketahui nilai *rounded value* (λ) sebesar 1, serta batas atas dan batas bawah telah melewati nilai 1 sehingga dapat disimpulkan bahwa data telah stasioner dalam *varians*.

Selanjutnya akan dilakukan pemeriksaan stasioneritas dalam *mean* yang dapat dilihat dari *time series plot* dan Plot ACF. *Time series plot* data curah hujan di Stasiun Juanda dapat dilihat pada Gambar 4.

Gambar 4 menunjukkan *time series plot* data curah hujan di Stasiun Juanda tidak stasioner dalam *mean*, yang terlihat dari fluktuasi data tidak stabil, tetapi akan dilakukan pengecekan melalui *plot* ACF agar lebih jelas. Plot ACF dapat dilihat pada Gambar 5.

Gambar 5 menunjukkan bahwa pada plot ACF turun lambat sehingga dapat dikatakan data curah hujan di Stasiun Juanda belum stasioner dalam *mean*. Data yang tidak stasioner dalam *mean* perlu dilakukan *differencing* ($d=1$) agar stasioner dalam *mean*. Plot ACF dan PACF hasil *differencing* data dapat dilihat pada Gambar 6 dan Gambar 7

Berdasarkan Gambar 6 dapat diketahui bahwa plot ACF *cut off* pada lag ke-1 sehingga dapat disimpulkan data sudah stasioner dalam *mean*. Setelah data sudah stasioner dalam *varians* dan *mean*, maka langkah selanjutnya adalah mengidentifikasi orde model ARIMA.

Berdasarkan plot ACF dan PACF pada Gambar 6 dan Gambar 7 didapatkan 33 dugaan model ARIMA yang selanjutnya akan dilakukan pengujian signifikansi parameter dan diperoleh 15 model yang signifikan. Dari 15 model yang signifikan didapatkan 3 model yang memenuhi residual *white noise* tetapi tidak berdistribusi normal, yaitu ARIMA ([1,12,35,58],1, [1,50]), ARIMA ([1,9,12,35,55,58],1,1), dan ARIMA (1,1,1) Oleh karena itu, akan dilakukan cek *outlier* untuk melihat penyebab residual tidak berdistribusi normal. Hasil cek *outlier* dapat dilihat pada Gambar 8

Gambar 8 menunjukkan bahwa *boxplot* residual dari ketiga model memiliki banyak data yang *outlier*, hal ini menyebabkan residual tidak berdistribusi normal.

Pemilihan model ARIMA terbaik menggunakan kriteria MSE dengan melihat nilai MSE terkecil yang diperoleh dari data *out-sample*. Hasil dari pemilihan model ARIMA terbaik dapat dilihat pada Tabel 5.

Tabel 5 menunjukkan hasil pemilihan model terbaik berdasarkan kriteria kebaikan model yang digunakan, yaitu MSE. Model ARIMA ([1,9,12,35,55,58],1,1) menghasilkan nilai MSE terkecil.

Persamaan matematis yang dibangun oleh model ARIMA ([1,9,12,35,55,58],1,1) dapat dilihat pada persamaan berikut:

$$Z_t = 1,146 Z_{t-1} - 0,146 Z_{t-2} - 0,074 Z_{t-9} + 0,074 Z_{t-10} + 0,084 Z_{t-12} - 0,084 Z_{t-13} + 0,082 Z_{t-35} - 0,082 Z_{t-36}$$

$$-0,081 Z_{t-55} + 0,081 Z_{t-56} + 0,122 Z_{t-58} - 0,122 Z_{t-59} + a_t - 0,878 a_{t-1} \tag{15}$$

C. Pengujian Linearitas

Uji linearitas digunakan untuk mengetahui apakah data terdapat pola linear atau non-linear. Hasil uji terasvirta data curah hujan di Stasiun Juanda dapat dilihat pada Tabel 6.

Tabel 6 menunjukkan hasil uji terasvirta data curah hujan di Stasiun Juanda didapatkan nilai F_{hitung} sebesar 20,212 lebih besar dari $F_{0,05;2;848}$ sebesar 3,006 dan diperkuat dengan P_{value} (0,000) lebih kecil dari α (0,05) maka diputuskan tolak H_0 yang artinya data curah hujan di Stasiun Juanda berpola non-linear.

D. Pemodelan Curah Hujan di Stasiun Juanda Menggunakan RBFNN

Langkah pertama untuk membentuk model adalah menentukan matrik *input* yang terdiri dari variabel-variabel *input*. Variabel *input* ditentukan oleh nilai *partial autokorelasi* yang signifikan. Pada pemilihan model terbaik ARIMA didapatkan nilai *partial autokorelasi* ada 12 lag yang signifikan, berarti ada 12 variabel input dalam model RBFNN. Variabel input yang akan digunakan, yaitu $Z_{t-1}, Z_{t-2}, Z_{t-9}, Z_{t-10}, Z_{t-12}, Z_{t-13}, Z_{t-35}, Z_{t-35}, Z_{t-55}, Z_{t-56}, Z_{t-58},$ dan Z_{t-59} . Dalam penelitian ini terdapat 12 buah neuron pada lapisan input, karena banyak variabel *input* sama dengan banyaknya *neuron*.

Penentuan jaringan optimum ditentukan oleh variabel *input* yang kemudian akan dikelompokkan dengan metode *k-means cluster*. Dari semua kelompok yang dicobakan, akan dipilih satu kelompok terbaik dengan melihat kriteria kebaikan model, yaitu dengan MSE (*Mean Square Error*) terkecil. Hasil penentuan jaringan optimum dapat dilihat pada Tabel 7

Tabel 7 menunjukkan hasil penentuan jaringan optimum yang terpilih adalah 3 kelompok karena memiliki nilai MSE terkecil. Sehingga model RBFNN yang digunakan dapat dibentuk dengan arsitektur 12 *neuron* pada lapisan input, 3 *neuron* pada lapisan tersembunyi, dan 1 *neuron* pada lapisan *output*. Arsitektur jaringan RBFNN dapat dilihat pada Gambar 9.

Pembobot yang digunakan pada jaringan RBFNN untuk curah hujan di Stasiun Juanda adalah sebagai berikut:

$$c = \begin{pmatrix} 0.032 & 3.924 & 10.903 \\ 0.055 & 3.759 & 10.948 \\ 0.028 & 3.780 & 10.739 \\ 0.013 & 4.152 & 10.574 \\ 0.020 & 1.106 & 11.166 \\ 0.029 & 0.445 & 11.255 \\ 0.034 & 0.495 & 11.446 \\ 0.038 & 0.997 & 11.258 \\ 0.059 & 0.549 & 11.227 \\ 0.022 & 1.380 & 11.108 \\ 0.042 & 1.401 & 11.067 \\ 0.060 & 0.820 & 11.205 \end{pmatrix}$$

$$\sum (0.337 \quad 8.560 \quad 28.583)$$

$$w_0 = 13.435, w_1 = -1.497, w_2 = -10.746, w_3 = -12.352$$

Berdasarkan arsitektur jaringan RBFNN curah hujan di Stasiun Juanda pada Gambar 9, maka model yang terbentuk seperti pada persamaan (15).

$$\hat{Z}_t = 13.435 - 1.497 h_1(x) - 10.746 h_2(x) - 12.352 h_3(x) \quad (16)$$

dengan nilai $h_j(x)$ sebagai berikut:

$$h_j(x) = \exp \left\{ - \left[\frac{(x_1 - c_{1j})^2}{2\sigma_j^2} + \frac{(x_2 - c_{2j})^2}{2\sigma_j^2} + \dots + \frac{(x_{12} - c_{12j})^2}{2\sigma_j^2} \right] \right\}$$

E. Perbandingan Model ARIMA dan RBFNN

Untuk membandingkan nilai ramalan model ARIMA dan RBFNN yang telah didapatkan, akan digunakan MSE sebagai kriteria kebaikan model. Nilai MSE model ARIMA dan RBFNN dapat dilihat pada Tabel 8.

Tabel 8 menunjukkan bahwa model ARIMA memiliki nilai MSE sebesar 16,902 sedangkan model RBFNN memiliki nilai MSE sebesar 5,270. Berdasarkan hasil tersebut diketahui Model RBFNN memiliki nilai MSE lebih kecil daripada model ARIMA, sehingga dapat disimpulkan bahwa model RBFNN lebih baik untuk meramalkan curah hujan di Stasiun Juanda.

F. Peramalan Curah Hujan di Stasiun Juanda

Model terbaik yang digunakan untuk peramalan curah hujan di Stasiun Juanda adalah model RBFNN. Hasil peramalan untuk periode 852 hingga 865 dapat dilihat pada Tabel 9.

Gambar 10 menjelaskan nilai ramalan yang dihasilkan model RBFNN periode 852 hingga 865, pola data bergerak fluktuatif. Curah hujan tertinggi berada pada periode ke-857 sebesar 13.841 mm dan curah hujan terendah berada pada

periode ke-861 sebesar 7.981 mm.

V. KESIMPULAN

Berdasarkan hasil analisis dan pembahasan peramalan curah hujan di Stasiun Juanda dapat ditarik kesimpulan sebagai berikut: Model ARIMA terbaik untuk meramalkan curah hujan di Stasiun Juanda adalah ARIMA ([1,9,12,35,55,58],1,1). Model *Radial Basis Function* (RBFNN) yang optimum adalah model dengan 12 *neuron* pada input, 3 *neuron* pada lapisan tersembunyi, dan 1 *neuron* pada lapisan output.

Berdasarkan perbandingan kebaikan model dengan nilai MSE pada model ARIMA dan RBFNN, dapat disimpulkan bahwa model RBFNN lebih baik dalam meramalkan curah hujan di Stasiun Juanda karena memiliki nilai MSE lebih kecil dibandingkan model ARIMA.

Peramalan curah hujan periode 852 hingga 865 memiliki pola data yang bergerak fluktuatif. Curah hujan tertinggi berada pada periode ke-857 sebesar 13.841 mm dan curah hujan terendah berada pada periode ke-861 sebesar 7.981 mm.

DAFTAR PUSTAKA

- [1] S. Makridakis, S. C. Wheelwright, and V. E. McGee, *Metode dan Aplikasi Peramalan*. Jakarta: Erlangga, 1999.
- [2] W. Wei, *Time Series Analysis Univariate and Multivariate*. London: Pearson education, 2006.
- [3] J. . Cryer and K. Chen, *Time Series Analysis With Applications in R*, 2nd ed. New York: Springer, 2008.
- [4] B. Bowerman, *Forecasting and Time Series : An Applied Approach*, 3rd ed. California: Duxbury Press, 1993.
- [5] L. V Fausett, *Fundamentals of Neural Networks (Architectures, Algorithms, and Applications)*. Mumbai: Pearson education, 2006.