

ARDL METHOD: Forecasting Data Jumlah Hari Terjadinya Hujan Di NTB

Ulul Azmi¹, Zilullah Nazir Hadi², Siti Soraya³

^{1,2,3}Ilmu Komputer, Universitas Bumigora, Mataram

¹1710510160@universitاسbumigora.ac.id, ²1710520197@universitاسbumigora.ac.id

³sitisorayaburhan@universitاسbumigora.ac.id

DOI 10.30812/varian.v3i2.627

INFO ARTIKEL

Riwayat Artikel:

Diterima: 21-01-2020

Disetujui: 21-03-2020

Kata Kunci:

Forecasting

ARDL

Jumlah Hari Hujan

NTB

Nusa Tenggara Barat

ABSTRAK

Abstrak: Penelitian ini berisi tentang prediksi atau *forecasting* data curah hujan di Nusa Tenggara Barat (NTB) tahun 2011, yakni khususnya jumlah hari terjadinya hujan dengan menggunakan metode *Autoregressive Distributed Lag* (ARDL). Data yang digunakan yaitu data jumlah hari terjadinya hujan per tahun di Nusa Tenggara Barat (NTB) dari 2006-2010, dengan menggunakan beberapa parameter *error* seperti *Mean Absolute Deviation* (MAD), *Mean Squared Error* (MSE), *Root Mean Squared Error* (RMSE) dan *Mean Absolute Percentage Error* (MAPE). Berdasarkan hasil simulasi data di Nusa Tenggara Barat (NTB), diperoleh prediksi jumlah hari terjadinya hujan pada tahun 2011 sebanyak 246,539 hari dengan nilai MAD 23,759, MSE 4,016, RMSE 2,004, dan MAPE 10,972. Prediksi jumlah hari terjadinya hujan ini dapat disimpulkan bahwa pada tahun 2011 terjadi peningkatan jumlah hari terjadinya hujan di Nusa Tenggara Barat (NTB).

Abstract: This study contains predictions or forecasting of rainfall data in West Nusa Tenggara (NTB) in 2011, specially the number of days of rain using the *Autoregressive Distributed Lag* (ARDL) method. The data used are data on the number of days of rain per year in West Nusa Tenggara (NTB) from 2006-2010, using several error parameters such as *Mean Absolute Deviation* (MAD), *Mean Squared Error* (MSE), *Root Mean Squared Error* (RMSE) and *Mean Absolute Percentage Error* (MAPE). Based on the results of data simulations in West Nusa Tenggara (NTB), prediction of the number of days of rain in 2011 was 246.539 days with a MAD value of 23.759, MSE 4.016, RMSE 2.004, and MAPE 10.972. This prediction of the number of days of rain can be concluded that in 2011 there was an increase in the number of days of rain in West Nusa Tenggara (NTB).

A. LATAR BELAKANG

Metode ARDL pertama kali diperkenalkan oleh Pesaran dan Shin (Pesaran & Shin, 1997) dengan pendekatan uji kointegrasi dengan pengujian *Bound Test Cointegration* (Apriyanto, 2016). Model ADL dapat disingkat dengan model ARDL. Metode ARDL merupakan salah satu bentuk metode dalam ekonometrika. Metode ini dapat mengestimasi model regresi linear dalam menganalisis hubungan jangka panjang yang melibatkan adanya uji kointegrasi diantara variabel-variabel *time series* (Elkadhi & Ben Hamida, 2014).

ARDL adalah metode regresi yang memasukkan *lag* dari kedua variabel dependen dan independen secara bersamaan. Dengan menggunakan model ini, peneliti dapat menganalisis hubungan jangka panjang ketika variabel-variabel penjelasnya campuran antara yang bersifat 1(1) dan 1(0). Estimator ARDL akan menghasilkan koefisien jangka panjang yang konsisten dan dapat dibuat dengan menggunakan *standard normal asymptotic theory*. Salah satu keunggulan dari pendekatan ARDL ini adalah menghasilkan estimasi yang konsisten dengan koefisien jangka panjang yang bagus tanpa peduli apakah variabel-variabel

penjelasnya atau regresornya 1(0) ataupun 1(1). Dalam kasus adanya jangka panjang yang bersifat *trend stationarity*, dengan ARDL dapat dilakukan *detrending* terhadap *series* dan memodelkan *detrended series* tersebut sebagai *distributed lag* yang stasioner (Fadhilah & Sukmana, 2017).

Model AR adalah model yang menggunakan satu atau lebih data masa lampau dari variabel dependen di antara variabel penjelas. Model DL adalah model regresi melibatkan data pada waktu sekarang dan waktu masa lampau (*lagged*) dari variabel penjelas (Gujarati & Porter, 2012). Model ARDL sangat berguna dalam ekonometrik empiris, karena membuat teori ekonomi yang bersifat statis menjadi dinamis dengan memperhitungkan peranan waktu secara eksplisit. Model ini dapat membedakan respon jangka pendek dan jangka panjang dari variabel tak bebas terhadap satu unit perubahan dalam nilai variabel penjelas (Gujarati & Porter, 2012). Keistimewaan dari model *autoregressive* dan model distribusi *lag* adalah model tersebut membuat teori statis menjadi dinamis karena model regresi yang biasanya mengabaikan pengaruh waktu. Melalui model *autoregressive* dan model distribusi *lag*, waktu ikut diperhitungkan dan panjang beda kala (*lag*) diketahui (Gujarati, 2014).

Distribution lag model adalah model regresi yang tidak hanya mencakup nilai sekarang tetapi juga nilai masa lalu (*lag*) dari variabel penjelas (X). Sedangkan *autoregressive distributed lag* adalah model yang mencakup satu atau lebih nilai masa lalu (*lag*) dari variabel terkait antara variabel penjelasnya. Model regresi yang memasukkan nilai variabel yang menjelaskan nilai masa kini atau nilai masa lalu (*lag*) dari variabel tak bebas sebagai salah satu variabel penjelas disebut *autoregressive distributed lag* (ARDL). Model ini dapat membedakan respon jangka pendek dan jangka panjang dari variabel tak bebas terhadap suatu unit perubahan dalam nilai variabel penjelas (Gujarati, 2003).

B. TINJAUAN PUSTAKA

Rumus formula apabila menggunakan 1 data:

$$X = \beta_0 + \varphi_1 X_{t-1} \quad (1)$$

Dimana:

X : stasioner tingkat level / variabel terikat

β_0 : konstanta

φ_1 : koefisien *dependent*

X_{t-1} : variabel waktu sebelumnya

$t - 1$: waktu sebelumnya

t : waktu

Rumus formula apabila menggunakan 2 data:

$$X = \beta_0 + \varphi_1 X_{t-1} + \dots + \varphi_p X_{t-p} + \beta_0 Y_t + \dots + \varphi_q Y_{t-q} \quad (2)$$

Dimana:

X : stasioner tingkat level / variabel terikat

β_0 : konstanta

φ_1 : koefisien *dependent*

X_{t-1} : variabel data pertama waktu sebelumnya

Y_{t-1} : variabel data kedua waktu sebelumnya

$t - 1$: waktu sebelumnya

t : waktu

1. Algoritma dan Flowchart

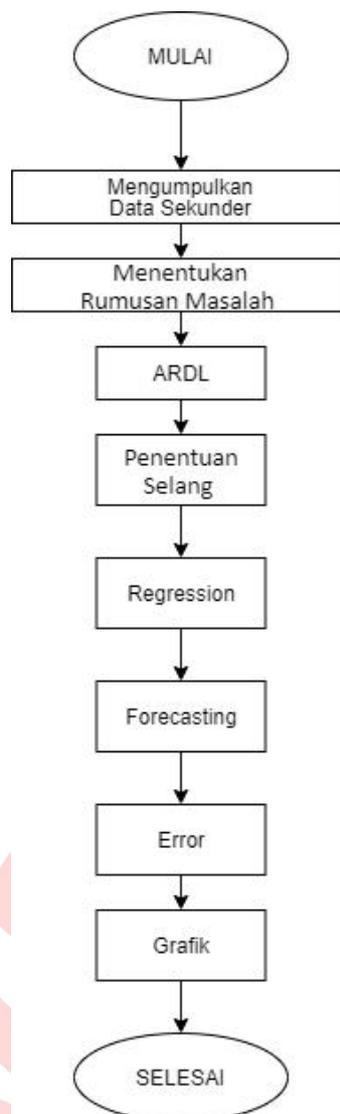
Adapun algoritmanya tersusun seperti di bawah ini:

a) Mulai

b) Mengumpulkan data penelitian berupa data sekunder yang diperoleh dari beberapa sumber, yakni data jumlah hari terjadinya hujan di NTB pada tahun 2006-2010

- c) Menentukan rumusan masalah dan tujuannya yakni memprediksi data jumlah hari terjadinya hujan pada tahun selanjutnya (2011)
- d) Menggunakan rumusan formula ARDL
- e) Uji kointegrasi dilakukan untuk melihat apakah ada hubungan jangka panjang antara variabel bebas dan variabel terikat
- f) Menentukan selang optimal, yaitu menentukan *lag* dari data
- g) Melakukan *regression* dari *Excel* untuk menentukan peramalan tahun selanjutnya
- h) Menentukan error dari data yang tersedia
- i) Membuat grafik dari data aktual dan data prediksi
- j) Selesai.

Flowchart penelitian dengan menggunakan metode ARDL dapat dilihat pada gambar di bawah ini:



Gambar 1. *Flowchart* Metode ARDL

2. Nilai Ketepatan Prediksi

Ketepatan prediksi adalah suatu hal yang penting untuk peramalan, yaitu bagaimana mengukur kesesuaian antara data yang sudah ada dengan data peramalan. Menurut Pakaja (Pakaja, Naba, & Purwanto, 2012), ada beberapa perhitungan yang biasa digunakan untuk menghitung kesalahan peramalan total. Tiga dari perhitungan yang paling terkenal adalah deviasi mutlak rerata (*Mean Absolute Deviation-MAD*), kesalahan kuadrat rerata (*Mean Squared Error-MSE*), dan kesalahan persen mutlak rerata (*Mean Absolute Percentage Error-MAPE*). Berikut uraian dari beberapa perhitungan kesalahan peramalan.

a) Rata-rata kesalahan/ *Mean Error* (ME)

Mean error (ME) merupakan rata-rata nilai kesalahan dari dari sejumlah data.

$$ME = \frac{1}{n} \sum_{i=1}^n e_i$$

ME sulit untuk menentukan kesalahan suatu data secara keseluruhan dan tepat karena terdapat nilai positif dan negatif yang akan saling mengurangi atau menambahkan nilai kesalahan.

b) Galat persentase/ *Percentage Error* (PE)

Percentage Error (PE) merupakan persentase kesalahan dari nilai aktual dengan nilai hasil peramalan.

$$PE = \left(\frac{x_i - F_t}{x_i} \right) \times 100\%$$

c) *Squared Error* (SE)

Squared Error (SE) adalah penjumlahan dari kuadrat kesalahan.

$$SE = \sum_{i=1}^n e_i^2$$

d) *Mean Absolute Deviation* (MAD)

Menurut Pakaja (Pakaja et al., 2012), metode untuk mengevaluasi metode peramalan menggunakan jumlah dari kesalahan-kesalahan yang absolut. *Mean Absolute Deviation* (MAD) mengukur ketepatan ramalan dengan merata-ratakan kesalahan dugaan (nilai absolut masing-masing kesalahan). MAD berguna ketika mengukur kesalahan ramalan dalam unit yang sama sebagai deret asli. MAD merupakan ukuran pertama kesalahan peramalan keseluruhan untuk sebuah model. Rumus untuk menghitung MAD adalah sebagai berikut:

$$MAD = \frac{\sum_{t=1}^n |X_t - F_t|}{n}$$

Dimana:

X_t : data aktual pada periode t

F_t : nilai peramalan pada periode t

n : jumlah data

e) *Mean Square Error* (MSE)

Menurut Suryaningrum (Suryaningrum & W, 2015), *Mean Squared Error* (MSE) adalah metode lain untuk mengevaluasi metode peramalan. Masing-masing kesalahan atau sisa dikuadratkan. Pendekatan ini mengatur kesalahan peramalan yang besar karena kesalahan-kesalahan itu dikuadratkan. Metode itu menghasilkan kesalahan-kesalahan sedang yang kemungkinan lebih baik untuk kesalahan kecil, tetapi kadang menghasilkan perbedaan yang besar. MSE merupakan cara kedua untuk mengukur kesalahan peramalan keseluruhan. MSE merupakan rata-rata selisih kuadrat antara nilai yang diramalkan dan yang diamati. Kekurangan penggunaan MSE adalah bahwa MSE cenderung menonjolkan deviasi yang besar karena adanya pengkuadratan. Rumus untuk menghitung MSE adalah sebagai berikut.

$$MSE = \frac{\sum_{t=1}^n (X_t - F_t)^2}{n}$$

f) Nilai tengah galat absolut/ *Mean Absolute Error* (MAE)

Mean Absolute Error (MAE) merupakan rata-rata nilai kesalahan yang bernilai mutlak positif dari jumlah data, sesuai persamaan berikut.

$$MAE = \frac{1}{n} \sum_{i=1}^n |e_i|$$

Hal tersebut bertujuan untuk mengantisipasi kesalahan atau *error* yang bernilai negatif, sehingga dapat menentukan nilai rata-rata kesalahan secara tepat.

g) *Mean Percentage Error* (MPE)

Mean Percentage Error (MPE) adalah rata-rata dari persentase kesalahan hasil peramalan.

$$MPE = \sum_{i=1}^n \frac{PE}{n}$$

Keterangan :

MPE : rata-rata persentase kesalahan

PE : persentase kesalahan pada periode ke-t

n : banyaknya data

h) *Mean Relative Error* (MRE)

$$MRE = \frac{1}{n} \sum_{i=1}^n \left| \frac{X_t - F_t}{X_t} \right| \times 100\%$$

Dimana:

X_t : data aktual pada periode t

F_t : nilai peramalan pada periode t

n : jumlah data

i) *Root mean Squared Error* (RMSE)

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (X_t - F_t)^2}{n}}$$

Keterangan:

X_t : data aktual pada periode t

F_t : nilai peramalan pada periode t

n : jumlah data

j) *Mean Absolute Percentage Error* (MAPE)

Menurut Pakaja (Pakaja et al., 2012), *Mean Absolute Percentage Error* (MAPE) dihitung dengan menggunakan kesalahan absolut pada tiap periode dibagi dengan nilai observasi yang nyata untuk periode itu. Kemudian, merata-ratakan kesalahan persentase absolut tersebut. MAPE merupakan pengukuran kesalahan yang menghitung ukuran presentase penyimpangan antara data aktual dengan data peramalan. Nilai MAPE dapat dihitung dengan persamaan:

$$MAPE = \left(\frac{1}{n} \right) \sum_{t=1}^n \left| \frac{X_t - F_t}{X_t} \right|$$

Menurut Putra (Putra, 2015), kemampuan peramalan sangat baik jika memiliki nilai MAPE kurang dari 10% dan mempunyai kemampuan peramalan yang baik jika nilai MAPE kurang dari 20%.

C. SIMULASI DATA

1. Data Jumlah Hari Terjadinya Hujan

Adapun data jumlah hari terjadinya hujan di NTB terdapat dalam Tabel 1.

Tabel 1. Data Jumlah Hari Terjadinya Hujan NTB

Tahun	Jumlah Hari
2000	0
2001	0
2002	0
2003	0
2005	0
2006	25
2007	162
2008	174
2009	174
2010	220
2011	0

Sumber: BPS Indonesia

2. Solusi dan Simulasi menggunakan Excel

Tabulasi data jumlah hari terjadinya hujan dapat dilihat pada tabel 2 di bawah ini.

Tabel 2. Tabulasi Data Jumlah Hari Terjadinya Hujan di NTB

t	Xt	Xt-1
2000	0	0
2001	0	0
2002	0	0
2003	0	0
2005	0	0
2006	25	0
2007	162	25
2008	174	162
2009	174	174
2010	220	174
2011	0	220

Tabel 3. Hasil Analisis Jumlah Hari Terjadinya Hujan di NTB

Regression Statistics	
Multiple R	0,887259
R Square	0,787229
Adjusted R Square	0,760633
Standard Error	45,78523
Observations	10

Tabel 3 di atas menunjukkan kekuatan hubungan antara model (variabel bebas) dengan variabel terikat. *Multiple R* (*R* majemuk) adalah hasil yang digunakan untuk mengukur tingkat (keeratan) hubungan linear antara variabel terikat dengan variabel bebas secara bersama-sama. Pada kasus dua variabel (satu variabel terikat dan satu variabel bebas), besaran *R* dapat bernilai positif maupun negatif ($-1 \leq R \leq 1$), tetapi untuk lebih dari dua variabel, besaran *R* selalu bernilai positif ($0 \leq R \leq 1$). Nilai *R* yang lebih besar (+ atau -) menunjukkan yang lebih kuat.

R square sering disebut dengan koefisien determinasi yang merupakan koefisien pengukur kebaikan model (*goodness of fit*) dari persamaan regresi, yaitu memberikan proporsi atau persentase variasi total dalam variabel terikat yang dijelaskan oleh variabel bebas. *Adjusted R square* adalah fungsi yang tidak pernah menurun dari banyaknya variabel bebas yang ada dalam model. *Standard error* merupakan standar eror dari estimasi variabel terikat. Jika semakin kecil standar eror ini dibandingkan standar deviasi dari permintaan, maka model regresi semakin tepat dalam memprediksi data.

Tabel 4. ANOVA Data Hari Hujan NTB

	<i>df</i>	<i>SS</i>
<i>Regression</i>	1	62048,20195
<i>Residual</i>	8	16770,29805
Total	9	78818,5
	<i>Coefficients</i>	<i>Standard Error</i>
<i>Intercept</i>	20,5414509	17,65428603
X Variabel 1	1,0272626	0,188817513

Tabel ANOVA (*analysis of variance*) digunakan untuk menguji penerimaan (*acceptability*) model dari perspektif statistik dalam bentuk analisis sumber keragaman. Anova sering juga diterjemahkan sebagai analisis ragam. *Degree of freedom* (*df*) atau derajat bebas dari total adalah $n-1$ dimana n adalah banyak observasi.

Sum of square (*ss*) atau jumlah kuadrat untuk *regression* merupakan penjumlahan kuadrat dari prediksi variabel terikat dikurangi dengan nilai rata-rata data sebenarnya. *Mean of square* (*MS*) merupakan rata-rata jumlah kuadrat. Nilai *F* ini yang dikenal dengan *f* hitung dalam pengujian hipotesis dibandingkan dengan *F* tabel. *Significance F* membandingkan antara taraf nyata dengan *F-value* atau *significance F*.

Untuk perkiraan tahun 2011 maupun tahun sebelumnya dapat dihitung menggunakan rumus pada persamaan (1). Hasil peramalan untuk tahun 2011 adalah sebagai berikut:

Perkiraan tahun 2011 (x) = kolom dan baris pada nilai *intercept* + kolom dan baris pada X variabel 1 * nilai jumlah pada tahun 2010

Perkiraan tahun 2011 (x) = nilai *intercept* + x variabel 1 * curah hujan 2010

$$= 20,5414509 + 1,0272626 * 220$$

$$= 246,5392$$

$$x = 246,5392$$

Hasil *forecasting* data jumlah hari terjadinya hujan di NTB dapat dilihat pada tabel di bawah ini.

Tabel 5. Hasil *Forecasting* menggunakan ARDL

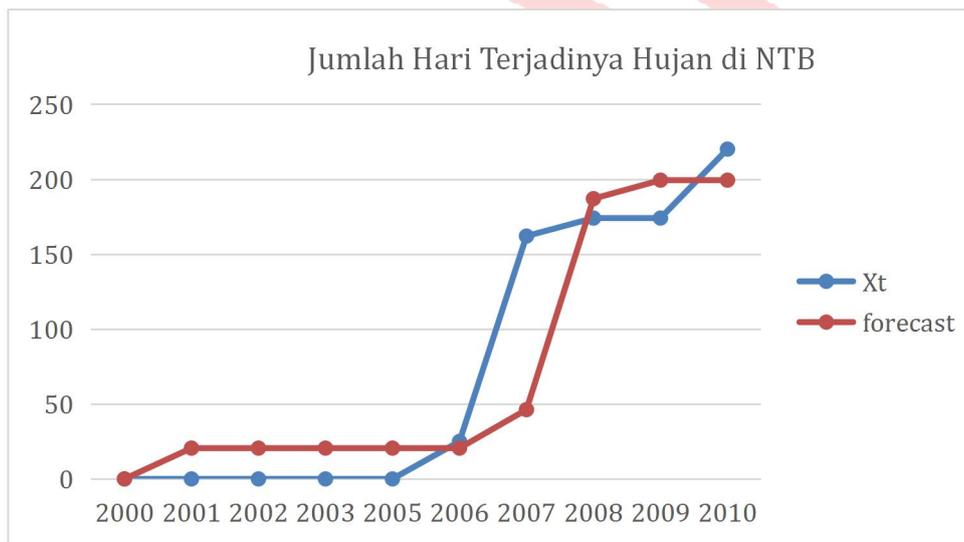
T	Xt	Xt-1	Forecast
2000	0	0	0,000
2001	0	0	20,541
2002	0	0	20,541
2003	0	0	20,541
2005	0	0	20,541
2006	25	0	20,541
2007	162	25	46,223
2008	174	162	186,958
2009	174	174	199,285
2010	220	174	199,285
2011	0	220	246,539

Kemudian, menghitung *error* dengan menggunakan rumus *error* MAD, MSE, RMSE dan MAPE seperti yang sudah dijelaskan pada pembahasan yang di atas. Perhitungan diselesaikan satu per satu sehingga menghasilkan *error* seperti pada Tabel 6.

Tabel 6. Nilai Error pada Model ARDL

t	Xt	Xt-1	Forecast	Xt-Ft	Xt-Ft	(Xt-Ft)^2	(Xt-Ft)/Xt
2000	0	0	0,000	0	0	0	0
2001	0	0	20,541	-20,5415	20,54145	4,532268	~
2002	0	0	20,541	-20,5415	20,54145	4,532268	~
2003	0	0	20,541	-20,5415	20,54145	4,532268	~
2005	0	0	20,541	-20,5415	20,54145	4,532268	~
2006	25	0	20,541	4,458549	4,458549	2,111528	0,17834197
2007	162	25	46,223	115,777	115,777	10,75997	0,71467274
2008	174	162	186,958	-12,958	12,95799	3,599721	0,07447122
2009	174	174	199,285	-25,2851	25,28514	5,028433	0,14531692
2010	220	174	199,285	20,71486	20,71486	4,551358	0,09415844
2011	0	220	246,539	-246,539	0	0	0
Total	755	535	980,998	20,541	261,359	44,180	1,207
JUMLAH TAHUN							11
MAD							23,7599364
MSE							4,01637091
RMSE							2,00408855
MAPE							0,10972375

3. Grafik



Gambar 2. Perbandingan Data Aktual dengan Forecasting

D. SIMPULAN DAN SARAN

Berdasarkan hasil simulasi data curah hujan, yakni berupa data jumlah hari terjadinya hujan di NTB pada tahun 2011, diperoleh kesimpulan bahwa dengan menggunakan metode *Autoregressive Distributed Lag* (ARDL), prediksi banyaknya hari terjadinya hujan di tahun 2011 adalah 247 hari dengan curah hujan sebesar 1368.424 mm dengan nilai MAD 23,75993899, MSE 4,016371118, RMSE 2,0040886, dan MAPE 10,972237527.

Selanjutnya untuk mengetahui hasil prediksi yang diperoleh baik atau tidak, dapat dilihat dari nilai MAPE-nya atau kesalahan yang menghitung ukuran persentase penyimpangan antara data aktual dengan data peramalan/*forecast* (Pakaja et al., 2012). Jika kemampuan peramalan memiliki nilai MAPE 10,97%, nilai tersebut kurang dari 20%, maka hasil prediksinya baik. Dengan begitu, bisa dikatakan hasil *forecast* dari data jumlah hari terjadinya hujan di NTB pada tahun 2011 baik, dikarenakan nilai MAPE-nya kurang dari 20%.

REFERENSI

- Apriyanto, D. (2016). *Analisis Faktor-Faktor Yang Mempengaruhi Return Indeks Harga Saham Gabungan (IHSG) Di Indonesia*. Institut Pertanian Bogor.
- Elkadhi, H., & Ben Hamida, R. (2014). The short-term effects of air pollution on health in Sfax (Tunisia): An ARDL cointegration procedure. *International Conference and Utility Exhibition on Green Energy for Sustainable Development (ICUE)*, 1–5.
- Fadhilah, N., & Sukmana, R. (2017). Pengaruh Sertifikat Bank Indoensia Syariah (SBIS), Jakarta Islamic Index (JII), Tingkat Inflasi, dan Index Harga Saham Gabungan (IHSG) terhadap Nilai Tukar: Pendekatan Autoregressive Distributed Lag (ARDL). *Jurnal Ekonomi Syariah Teori Dan Terapan*, 4(10), 833–846.
- Gujarati, D. N. (2003). *Ekonometrika Dasar*. Jakarta: Erlangga.
- Gujarati, D. N. (2014). *Econometrics by Example* (2nd ed.). Basingstoke, United Kingdom: MacMillan Education UK.
- Gujarati, D. N., & Porter, D. C. (2012). *Dasar-dasar ekonometrika* (5th ed.). Jakarta: Salemba Empat.
- Pakaja, F., Naba, A., & Purwanto. (2012). Peramalan Penjualan Mobil menggunakan Jaringan Syaraf Tiruan dan Certainty Factor. *Jurnal EECCIS*, 6(1), 23–28.
- Pesaran, M. H., & Shin, Y. (1997). An Autoregressive Distributed Lag Modelling Approach to Cointegration Analysis □. *The Symposium at the Centennial of Ragnar Frisch*, 1–33. Oslo: The Norwegian Academy of Science and Letters.
- Putra, E. A.-R. H. (2015). *Analisis curah hujan bulanan menggunakan metode exponential smoothing (studi kasus : Katulampa Bogor)*. Institut Pertanian Bogor.
- Suryaningrum, K. M., & W, S. P. (2015). Analisis dan Penerapan Metode Single Exponential Smoothing untuk Prediksi Penjualan pada Periode Tertentu (Studi Kasus: PT. Media Cemara Kreasi). *Prosiding Senatif*, (2), 259–266.

