

**PREDIKSI CURAH HUJAN KOTA TANGERANG SELATAN
DENGAN DATA IKLIM 2020, 2021, 2022**
*RAINFALL PREDICTION FOR SOUTH TANGERANG CITY
WITH CLIMATE DATA FOR 2020, 2021, 2022*

Hafidz Alamsyah Putra^{*}, Oni Dwi Saputra^{2*}, Yosafat Donni Haryanto^{3*}

Sekolah Tinggi Meteorologi Klimatologi dan Geofisika, Jl. Perhubungan I No 5
Pondok Betung, Kota Tangerang Selatan, Banten 15221

*Email: hafidzalamsyahputra@gmail.com

ABSTRAK

Perubahan cuaca dalam lingkup yang sangat besar tentu akan menjadi masalah yang harus diselesaikan dengan cepat. Curah hujan yang berubah begitu fluktuatif menyebabkan terganggunya mobilitas masyarakat. Namun, ketika masalah tersebut terjadi di kota besar, tentu akan berdampak serius. Sehingga, dilakukan uji korelasi dan regresi menggunakan bahasa pemrograman R dimana uji korelasi dilakukan untuk mengetahui keeratn hubungan antar variabel bebas dengan variabel terikat. Korelasi akan semakin kuat apabila nilai koefisien korelasi semakin mendekati 1 atau -1 dan korelasi dinyatakan sempurna apabila nilai koefisiensi korelasi sama dengan 1 atau -1 dan uji regresi dilakukan untuk mengetahui bagaimana prediksi curah hujan yang terjadi di suatu wilayah yang luas, sehingga didapatkan hasil yang mendekati ketepatan yang baik. Penelitian ini dapat dikembangkan dengan menambahkan variabel yang berperan penting dalam penurunan kualitas udara dan menggunakan perhitungan dengan ketelitian yang lebih akurat. Hasil dari peneltian ini adalah data Meteorologi suhu udara, kecepatan angin, kelembaban udara dan tekanan udara dengan metode regresi tidak akurat untuk memprediksi curah hujan, Dikarenakan nilai akurasi yang dibawah 75% dan nilai p-value pada uji Autokolerasi lebih dari 0,05 menandakan tidak ada hubungan yang signifikan dengan variabel terikat (curah hujan).

Kata Kunci : Curah Hujan, Korelasi, Regresi.

ABSTRACT

Climate Change in a very large scope will certainly be a problem that must be resolved quickly. The fluctuating rainfall has disrupted people's mobility. However, when this problem occurs in a big city, it will certainly have a serious impact. Thus, correlation and regression tests were carried out using the R programming language where correlation tests were carried out to determine the closeness of the relationship between the independent variables and the dependent variable. The correlation will be stronger if the correlation coefficient value is getting closer to 1 or -1 and the correlation is declared perfect if the correlation coefficient value is equal to 1 or -1 and a regression test is carried out to find out how the prediction of rainfall that occurs in a large area, so that the results are close to good precision. This research can be developed by adding variables that play an important role in reducing air quality and using calculations with more accurate precision rainfall, because the accuracy value is below 75% and the p-value in the autocorrelation test is more than 0.05 indicating that there is no significant relationship with the dependent variable (rainfall).

Keywords: Rainfall, Correlation, Regression.

1. PENDAHULUAN

Air merupakan salah satu hal yang sangat penting bagi kelangsungan kehidupan makhluk hidup. Dapat dikatakan bahwa air adalah kebutuhan yang sangat berharga bagi manusia. Air dapat digunakan dalam berbagai jenis manfaat dalam kehidupan seperti sebagai sarana transportasi sampai untuk pembangkit listrik (Ulama, 1989).

Curah hujan dan suhu permukaan laut memiliki tingkat variabilitas yang cukup tinggi, sehingga membutuhkan data observasi yang panjang dan sebaran spasial yang memadai (As-Syakur dan Prasetya, 2010). Indonesia terletak di daerah Khatulistiwa, sehingga semua fenomena meteorologi yang terjadi di Indonesia dipengaruhi kondisi iklim monsun seperti perbedaan pergantian musim basah dan musim kering yang jelas.

Kota Tangerang Selatan terletak di Provinsi Banten, secara geografis (Pemerintah Provinsi Banten, 2021) $106^{\circ}38' - 106^{\circ}47'$ BT dan $06^{\circ}13'30'' - 06^{\circ}22'30''$ LS. Karakteristik curah hujan tentunya di setiap daerah berbeda-beda dipengaruhi beberapa faktor (Nieuwolt, 1977), yaitu: geografis, topografis, dan orografis. Hujan yang terjadi di wilayah tropis dipengaruhi oleh beberapa faktor salah satunya aktivitas dari manusia (Wirjohamidjojo dan Swarinto, 2007).

Dalam perkembangannya tentu prediksi curah hujan ini sangat penting bagi beberapa sektor, mengingat Kota Tangerang Selatan merupakan salah satu kota yang digunakan sebagai kota penyangga ibu kota negara. Beberapa sektor yang berkaitan antara lain : ekonomi, pemerintahan, pariwisata, dan lain-lain.

2. METODE PENELITIAN

Metode yang digunakan dalam penelitian ini adalah korelasional kuantitatif. Metode penelitian kuantitatif merupakan metode penelitian yang berlandaskan pada filsafat positivisme, digunakan untuk meneliti populasi atau sampel tertentu. Teknik pengambilan sampel dilakukan secara random, pengumpulan data menggunakan instrumen penelitian, analisis data bersifat kuantitatif/statistik dengan tujuan untuk menguji hipotesis yang telah ditetapkan (Sugiyono, 2010).

2.1 Prosedur Penelitian

Prosedur penelitian terdiri dari lokasi penelitian, data penelitian, alat yang digunakan, teknik pengumpulan data hingga teknik pengolahan data.

2.1.1 Lokasi dan waktu penelitian

Lokasi penelitian yang dipilih adalah Tangerang Selatan. Sementara waktu penelitian adalah 3 tahun terakhir, yakni antara tahun 2020 hingga 2022.

2.1.2 Data penelitian

Data yang digunakan dalam penelitian ini diantaranya:

1. Data Rata rata Curah Hujan
2. Data Suhu Udara
3. Data Kecepatan Angin
4. Data Kelembaban Udara
5. Data Tekanan Udara

Alat yang digunakan dalam penelitian ini antara lain:

1. *BMKGSoft*
BMKGSoft merupakan sebuah website dimana seluruh data observasi yang berisikan data-data Meteorologi.
2. Aplikasi Excell
Excell digunakan untuk merubah format CSV menjadi XLSX.
3. Perangkat R

Sebuah *programming language* yang berfokus pada pengolahan dan analisa data.

2.1.4 Teknik pengumpulan data

Berikut merupakan langkah-langkah pengumpulan data:

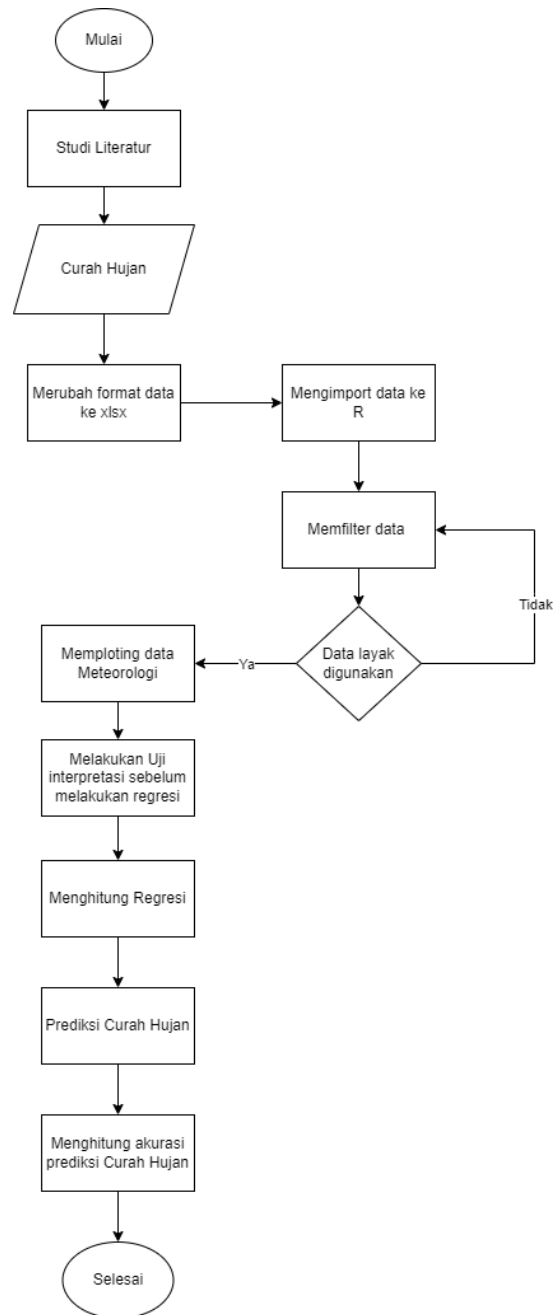
1. Data Kualitas Udara di buat dan di unduh di *BMKG Soft Website*.

2.1.5 Teknik pengolahan data

Berikut merupakan langkah-langkah pengolahan data:

- a. Membuka website *BMKG Soft*.
- b. Memilih map Tangerang Selatan sebagai area penelitian.
- c. *Mendownload* data Meteorologi.
- d. Merubah *format file* CSV menjadi XLSX melalui excel.
- e. Membuka Aplikasi R untuk mengelolah data.
- f. Memfilter data yang rusak (data kosong).
- g. Menyatukan seluruh data menjadi satu data *frame*.
- h. Memploting Grafik data Meteorologi.
- i. Mengoding Regresi antar Variabel.
- j. Menginterpretasi Regresi.
- k. Membuat prediksi dan mengetes, serta membuat kesimpulan.

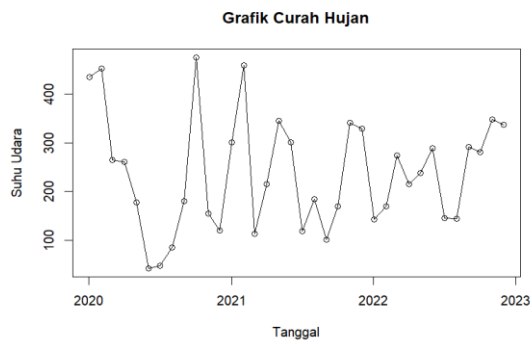
2.2 Diagram Alir Penelitian



Gambar 2.1 Diagram alir penelitian.

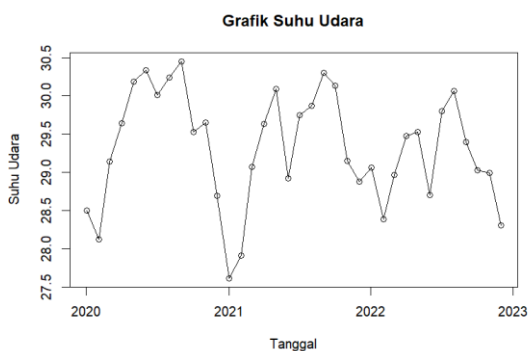
3. HASIL DAN PEMBAHASAN

3.1 Data Meteorologi



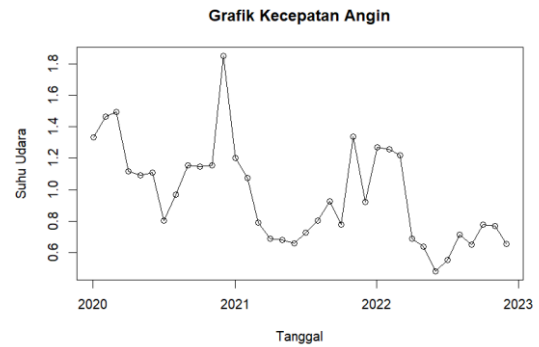
Gambar 3.1. Grafik Curah Hujan.

Pada gambar 3.1, menunjukkan grafik Curah Hujan Tangerang Selatan tahun 2020 – 2022. Rata rata Curah Hujan Tangerang Selatan tertinggi pada Oktober 2020 yaitu 475,5 dan terendah pada tahun Juni 2020 yaitu 42,2.



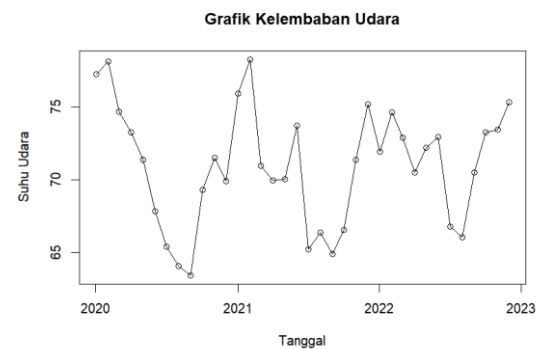
Gambar 3.2. Grafik Suhu Udara.

Pada gambar 3.2, menunjukkan grafik Suhu Udara Tangerang Selatan tahun 2020 – 2022 *stationary*. Rata rata Suhu Udara Tangerang Selatan tertinggi pada September 2020 yaitu 30,45 dan terendah pada tahun Januari 2021 yaitu 27,60..



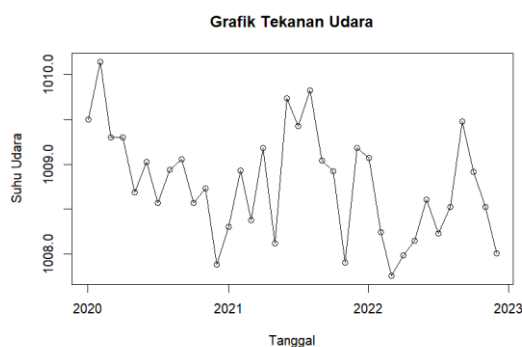
Gambar 3.3. Grafik Kecepatan Angin.

Pada gambar 3.3, menunjukkan grafik Kecepatan Angin Tangerang Selatan tahun 2020 – 2022 *stationary*. Rata rata Kecepatan Angin Tangerang Selatan tertinggi pada Desember 2020 yaitu 1,85 dan terendah pada tahun Juni 2022 yaitu 0,48.



Gambar 3.4. Grafik Kelembaban Udara.

Pada gambar 3.4, menunjukkan grafik Kelembaban Udara Tangerang Selatan tahun 2020 – 2022 *stationary*. Rata rata Kelembaban Udara Tangerang Selatan tertinggi pada Februari 2021 yaitu 78,23 dan terendah pada tahun September 2020 yaitu 63,46.

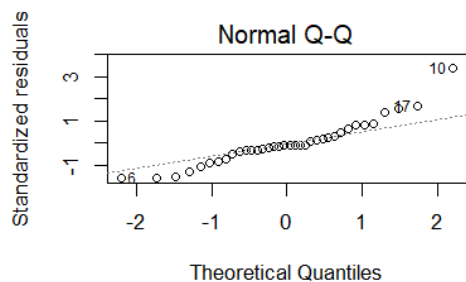


Gambar 3.5. Grafik Tekanan Udara.

Pada gambar 3,5, menunjukkan grafik Tekanan Udara Tangerang Selatan tahun 2020 – 2022 *stationary*. Rata rata Tekanan Udara Tangerang Selatan tertinggi pada Februari 2020 yaitu 1010,14 dan terendah pada tahun Maret 2022 yaitu 1007,75.

3.2 Uji Normalitas

Uji normalitas merupakan sebuah metode untuk menentukan apakah residu terdistribusi secara normal atau tidak.



Gambar 3.6. Plot Normalitas.

Pada Interpretasi uji Normalitas hal yang penting dilihat pada gambar Normal Q-Q menunjukkan point-point data berada di sekitar garis normal yang berarti data residual terdistribusi dengan normal sehingga dapat dikatakan model regresi memenuhi asumsi Normalitas.

3.3 Uji Asumsi Multikolinieritas

Uji Asumis Multikolinieritas adalah sebuah metode untuk menentukan apakah ada korelasi erat antara variabel independen dengan model regresi linier berganda, apabila ada korelasi tinggi antara variabel independen hubungan dengan variabel independen serta variabel dependen terganggu. Hal tersebut dapat dilihat dengan *package* “car” dari R dengan berfokus melihat nilai VIF nya.

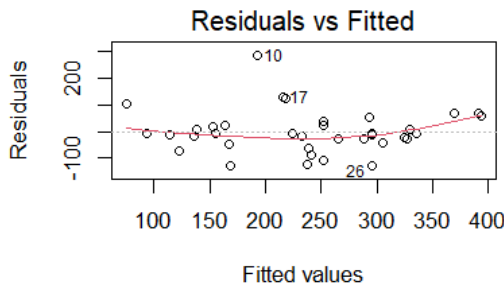
Tabel 3.1. Uji Asumsi Multikolinieritas.

Variabel Independen	Vif()
Suhu Udara	4,01
Kecepatan Angin	1,101
Tekanan Udara	1,04
Kelembaban Udara	3,95

Nilai VIF untuk Variabel Kecepatan angin, Tekanan Udara, Suhu Udara dan Kelembaban Udara berada dibawah 10 yang berarti tidak terjadi multikolinieritas sehingga model regresi memenuhi asumsi Multikolinieritas.

3.4 Uji Asumsi Heteroskedastisitas

Uji Asumis Heteroskedastisitas adalah sebuah metode untuk mengukur apakah terjadi ketidaksamaan varian residual (Heteroskedastisitas) dari data-data pengukuran dengan melihat sebaran pada grafik Residual vs *Fitted*. Model regresi yang baik adalah model ketika grafik tidak mengandung pola-pola seperti berkumpul ditengah, memperbesar, menyempit maupun memperkecil, dan lainnya.



Gambar 3.7. Plot Residual vs Fitted

Pada interpretasi Residuals vs *Fitted* menunjukkan bahwa data tersebar dan tidak membuat satu pola tertentu dan tidak terjadi perbedaan varians residual dan pengujian model memenuhi uji asumsi heteroskedastisitas.

3.5 Uji Asumsi Autokolerasi

Uji Asumis Autokolerasi adalah sebuah metode untuk mengetahui apakah ada kolerasi dari data terdahulu kepada data yang baru, uji dilakukan dengan menggunakan R dengan *function* *dwtest* () dalam *package* ("*lmtest*").

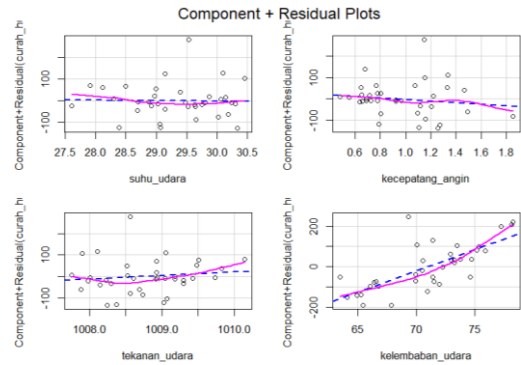
Tabel 3.2. Uji Asumsi Autokolerasi

Durbin-Watson test	
DW	1,7042
P-Value	0,09

Hasil uji diatas menunjukkan nilai Dw test p value lebih dari 0,05 yang berarti tidak terjadi korelasi antar data sebelumnya dengan data terbaru dan terdapat asumsi autokorelasi yang terpenuhi antar data dalam model yang sudah dibuat.

3.6 Uji Asumsi Linieritas

Uji Asumsi Linieritas adalah sebuah metode untuk Menguji hubungan linier antara variabel terikat dan independen untuk memenuhi asumsi adanya hubungan linier antar variabel.



Gambar 3.8. Plot Component +residual

Hasil menunjukkan bahwa dua garis dengan warna ungu dan biru berada pada posisi yang hampir berdempetan, artinya variabel independen memiliki linieritas dengan variabel terikat sehingga dapat dikatakan asumsi linieritas model regresi terpenuhi.

3.7 Uji Regresi Berganda

Uji regresi berganda memiliki tujuan untuk mengetahui seberapa besar pengaruh variabel bebas terhadap variabel terikat. Uji regresi ini dilakukan setelah mengetahui hubungan variabel bebas dan variabel terikat melalui uji korelasi. Hasil uji regresi disajikan pada gambar 3.9.

```

Residuals:
    Min       1Q   Median       3Q      Max
-126.743  -30.517   -8.963    26.788   283.296

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -18040.363  24539.250  -0.735  0.46777
suhu_udara    -1.496     39.317  -0.038  0.96988
kecepatang_angin -37.810    48.429  -0.781  0.44088
tekanan_udara  16.714    24.596  0.680  0.50183
kelembaban_udara 21.101     7.124  2.962  0.00582 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 85.21 on 31 degrees of freedom
Multiple R-squared:  0.5242,    Adjusted R-squared:  0.4628
F-statistic: 8.539 on 4 and 31 DF,  p-value: 9.12e-05
    
```

4	287,78
5	238,85
6	168,8

3.9 Mencari Selisih antara data prediksi dengan data curah hujan

Setelah diprediksi, kami akan mencoba membandingkan data tersebut dengan data nyata pada Januari 2020 hingga Juni 2020 menggunakan *package* "dplyr". Berikut tabel selisihnya.

Tabel 3.4. Tabel Prediksi vs Curah Hujan.

Date	Prediksi	CH	Selisih
Januari 2020	368,82	435,0	66,17
Februari 2020	393,67	451,9	58,22
Maret 2020	305,01	264,6	40,41
April 2020	287,78	261,0	26,78
Mei 2020	238,85	178,1	60,75
Juni 2020	168,8	42,2	126,69

Gambar 3.9. Hasil Regresi Berganda

Dari hasil uji regresi didapatkan persamaan model regresi berganda Curah Hujan = **-18040,36-1,496 (Suhu Udara) - 37,81 (Kecepatan Angin) + 16,71 (Tekanan Udara) + 21,1 (Kelembaban Udara).**

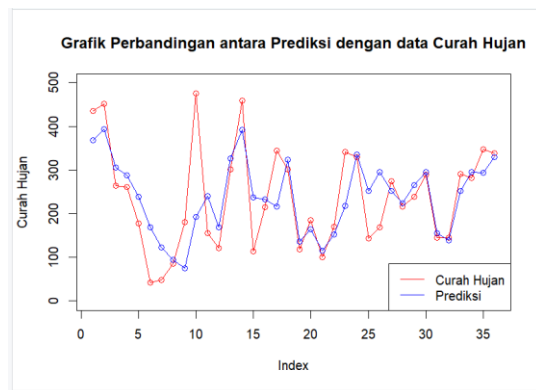
Melihat dari hasil pengujian diatas dapat dinyatakan bahwa pengaruh kelembaban udara terhadap curah hujan adalah signifikan dengan nilai p-value pada kolom Pr (>|t|) lebih kecil atau sama dengan dari 0.05.

3.8 Prediksi Curah Hujan

Untuk memprediksi curah hujan, penelitian akan menggunakan sebuah *function* "predict" dari R dengan menggunakan model dari regresi. Fungsi tersebut akan mencoba memprediksi berapa curah hujan yang terjadi pada awal Januari 2020 hingga Juni 2020 menggunakan data variabel independen (Suhu udara, kecepatan angin, tekanan udara, dan kelembaban udara). Regresi akan mencoba untuk memprediksi. Berikut Tabel Prediksi :

Tabel 3.3. Tabel Prediksi.

Nomer	Prediksi
1	368,82
2	393,67
3	305,01



Gambar 3.10. Grafik Perbandingan antara prediksi dengan data Curah Hujan

Dari hasil grafik antara hasil prediksi dengan data Curah Hujan menunjukkan bahwa data prediksi hampir segaris dengan data curah hujan sesungguhnya

3.10 Uji Akurasi

Untuk menguji akurasi, maka harus dicari nilai dari MAE, MSE, RMSE, dan R-Squared dimana:

1. **MAE** (*Mean Absolute Error*) mewakili perbedaan antara nilai asli dan prediksi yang diekstraksi dengan rata-rata perbedaan absolut pada kumpulan data.
2. **MSE** (*Mean Squared Error*) merepresentasikan perbedaan antara nilai asli dan prediksi yang diekstraksi dengan mengkuadratkan perbedaan rata-rata pada kumpulan data.
3. **RMSE** (*Root Mean Squared Error*) Apakah tingkat kesalahan dengan akar kuadrat dari MSE.
4. **R-squared** (*Coefficient of determination*) mewakili koefisien seberapa baik nilai cocok dibandingkan dengan nilai aslinya. Nilai dari 0 sampai 1 diartikan sebagai persentase. Semakin tinggi nilainya, semakin baik modelnya.

Berikut tabel dari nilai akurasi menggunakan *package* "caret".

Tabel 3.5. Tabel Akurasi.

MAE	MSE	RMSE	R-square
55,65	6252,9	79,07	0,52

Dapat dilihat pada nilai *r-squared* menunjukkan bahwa nilainya adalah 0,52 atau 52% yang dimana menunjukkan akurasi tidak terlalu bagus.

4. KESIMPULAN

Penelitian ini dilakukan dikarenakan timbulnya kesadaran akan meningkatnya curah hujan yang terjadi di Tangerang Selatan pada akhir tahun 2022. Penelitian ini memberikan beberapa kesimpulan yang penting :

- a. Data Suhu udara, kecepatan angin, kelembaban udara, dan tekanan udara tidak cocok digunakan untuk memprediksi curah hujan, dikarenakan tingkat akurasi yang dibawah 75%
- b. Dalam uji Autokolerasi, nilai p-value lebih dari 0,05 yang menandakan tidak terjadi korelasi antar data sebelumnya dengan data terbaru.

REFERENSI

- [1] Fridayana Yudiatmaja (2013). Analisis Regresi dengan menggunakan Aplikasi Komputer Statistik. Jakarta: Kompas Gramedia.
- [2] Masruroh, Kemal Faruoq Maulidi. (2020) "Perbandingan metode regresi linear dan neural network backpropagation dalam prediksi nilai ujian nasional siswa SMP menggunakan Software R.

- [3] Kurniawan, Robert. (2016) "Analisa Regresi: Dasar dan penerapan dengan R".
- [4] Wahab, Abdul. (2017). Pengembangan Modul Pembelajaran Literasi Statistika (Analisis Regresi Linier Sederhana dengan R). S2 thesis, Universitas Negeri Makassar.
- [5] T.I Wasalaine, M. W. TALAKULA, Y. ALESNUSSA. (2014). "Model Regresi ridge untuk mengatasi model regresi linier berganda yang mengandung multikolinieritas". Jurnal matematika .
- [6] Handayani, Lestari, and Muhammad Adri. "Penerapan JST (*Backpropagation*) untuk Prediksi curah hujan (Studi kasus: Kota Pekanbaru). Seminar Nasional Teknologi Informasi Komunikasi dan Industri. 2015.
- [7] Fadholi, Akhmad. "Persamaan regresi prediksi curah hujan bulanan menggunakan data suhu dan kelembapan udara di Ternate." *Statistika* 13.1 (2013).
- [8] Susilokarti, Dyah, et al. "Identifikasi perubahan iklim berdasarkan data curah hujan di wilayah selatan Jatiluhur Kabupaten Subang, Jawa Barat." *Agritech* 35.1 (2015): 98-105.
- [9] Rahayu, Nofiana Dian, Bandi Sasmito, and Nurhadi Bashit. "Analisis pengaruh fenomena indian ocean dipole (IOD) terhadap curah hujan di pulau Jawa." *Jurnal Geodesi Undip* 7.1 (2018): 57-67.
- [10] Mulyono, Dedi. "Analisis karakteristik curah hujan di wilayah Kabupaten Garut Selatan." *Jurnal Konstruksi* 12.1 (2014).