

**MODEL CURAH HUJAN GUNUNGSARI LOMBOK BARAT  
DENGAN REGRESI NONPARAMETRIK CAMPURAN  
KERNEL DAN DERET FOURIER**

***RAINFALL MODEL OF GUNUNGSARI WEST LOMBOK  
USING MIXED NONPARAMETRIC  
KERNEL AND FOURIER SERIES REGRESSION***

**NAILIL IZZATI**

Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Mataram.

Jl. Majapahit No. 62, Gomong, Kec. Selaparang, Kota Mataram,

Nusa Tenggara Barat 83115, Telp. (0370) 633007,

✉email: [izzatinailil00@gmail.com](mailto:izzatinailil00@gmail.com).

**Abstrak.** Indonesia yang terletak di antara dua samudra sangat rawan terhadap kejadian iklim ekstrem berupa banjir dan kekeringan, terutama wilayah yang mempunyai pola curah hujan bertipe monsoon seperti di Lombok. Banjir di Lombok pernah terjadi di wilayah Kecamatan Gunung Sari Lombok Barat pada tahun 2021. Banjir tersebut disebabkan karena curah hujan yang tinggi, sehingga menyebabkan sungai-sungai yang ada meluap. Untuk mengurangi kerugian yang terjadi, prediksi curah hujan menjadi salah satu solusi yang dapat dilakukan. Prediksi curah hujan dapat dilakukan berdasarkan model curah hujan. Penelitian ini bertujuan menentukan model curah hujan di Gunung Sari Lombok Barat berdasarkan beberapa variabel yang mempengaruhinya. Variabel respon yang digunakan adalah curah hujan, dan sebagai variabel prediktor adalah suhu udara, kelembaban udara, dan kecepatan angin. Pemodelan menggunakan pendekatan regresi nonparametrik campuran Kernel dan Deret Fourier, karena pola data tiap variabel prediktor terhadap variabel respon sebagian mengikuti karakteristik pola Kernel dan sebagian lainnya mengikuti karakteristik pola Deret Fourier. Estimasi parameter menggunakan metode Penalized Least Square (PLS). Selain itu, pemilihan bandwidth optimal  $\Phi$  pada komponen Kernel, parameter penghalus optimal  $\lambda$  dan parameter osilasi optimal  $M$  pada komponen Deret Fourier, didasarkan pada kriteria Generalized Cross Validation (GCV). Hasil penelitian yang diperoleh adalah model terbaik terdiri dari komponen Deret Fourier dengan dua variabel dan komponen Kernel dengan satu variabel. Pada komponen Deret Fourier diperoleh parameter osilasi optimal  $M = 2$ , parameter penghalus optimal  $\lambda_1 = 3$ ,  $\lambda_2 = 1$ . Pada komponen Kernel diperoleh bandwidth optimal  $\Phi = 0,7778$ . Adapun GCV minimum yang dihasilkan sebesar 8.423. Model terbaik tersebut memiliki nilai koefisien determinasi sebesar  $R^2 = 53\%$ .

**Kata kunci:** bandwidth, curah hujan, Generalized Cross Validation, Penalized Least Square, regresi nonparametrik campuran Kernel dan Deret Fourier

**Abstract.** Indonesia, located between two oceans, is highly susceptible to extreme climate events such as floods and droughts, especially in regions with monsoon rainfall patterns like Lombok. A flood occurred in the Gunung Sari District of West Lombok in 2021 due to heavy rainfall, causing the overflowing of rivers. To mitigate such losses, rainfall prediction is one solution. Rainfall prediction can be based on rainfall models. This study aims to determine the rainfall model in Gunung Sari, West Lombok, based on several influencing variables. The response variable used is rainfall, and the predictor variables are air temperature, air humidity, and wind speed. The modeling approach used a nonparametric regression method combining Kernel and Fourier Series, as the data patterns of each predictor variable to the response variable partially follow the characteristics of Kernel patterns and partly follow the characteristics of Fourier Series patterns. Parameter estimation was done using the Penalized Least Square (PLS) method. Additionally, the optimal bandwidth  $\Phi$  in the Kernel component, optimal smoothing parameter  $\lambda$ , and optimal oscillation parameter  $M$  in the Fourier Series component were selected based on the Generalized Cross-Validation (GCV) criteria. The research results show that the best model consists of two variables in the Fourier Series component and one variable in the Kernel component. The optimal oscillation parameter obtained in the Fourier Series component is  $M = 2$ , optimal smoothing parameters  $\lambda_1=3$ ,  $\lambda_2=1$ . In the Kernel component, the optimal bandwidth is  $\Phi = 0.7778$ . The minimum GCV obtained is 8.423. The best model obtained has a coefficient of determination of  $R^2= 53\%$ .

**Key words:** *bandwidth, rainfall, Generalized Cross Validation, Penalized Least Square, nonparametric regression with mixed Kernel and fourier series*

## PENDAHULUAN

Unsur iklim yang menarik untuk dikaji di Indonesia adalah curah hujan. Hal itu disebabkan karena tidak semua wilayah Indonesia mempunyai pola hujan yang sama. Selain itu, curah hujan merupakan parameter iklim yang paling mempengaruhi pola kehidupan masyarakat (Aldrian *et al.*, 2011). Menurut Direktur Jenderal Badan Penanggulangan Bencana Daerah (BPBD) Provinsi Nusa Tenggara Barat, banjir terparah di Lombok terjadi di wilayah Lombok Barat, khususnya di Kecamatan Gunungsari pada semester II tahun 2021. Banjir yang terjadi di Gunungsari Kabupaten Lombok Barat tersebut disebabkan karena curah hujan yang tinggi, sehingga menyebabkan sungai-sungai yang ada menjadi meluap. Selain itu, di Gunungsari juga terjadi tanah longsor (Yusuf, 2011). Bencana banjir ini dapat menimbulkan dampak buruk di berbagai sektor kehidupan masyarakat. Untuk mengurangi kerugian yang terjadi, prediksi curah hujan menjadi salah satu solusi yang dapat dilakukan. Salah satu cara melakukan prediksi adalah dengan menggunakan analisis regresi.

Dalam upaya mengantisipasi terjadinya banjir dan longsor, terlebih dahulu perlu memperhatikan bentuk pola data curah hujan. Ketika data curah hujan dan masing-masing variabel prediktor yang digunakan tersebut dipolakan, pola data yang terbentuk tidak mengikuti pola data tertentu sehingga kondisi tersebut mengindikasikan adanya komponen nonparametrik. Maka pada data curah hujan dapat diterapkan model regresi nonparametrik (Fazhuddin, 2018). Model-model regresi nonparametrik multivariabel yang dikembangkan oleh para peneliti sebelumnya masih terbatas menggunakan satu jenis estimator yang sama untuk semua variabel prediktor. Hal ini disebabkan oleh anggapan bahwa pola data dari masing-masing prediktor dianggap sama. Sementara itu, pada kenyataannya sering dijumpai kasus-kasus dengan pola data yang berbeda dari masing-masing variabel prediktor. Oleh karena itu, untuk mengatasi masalah tersebut beberapa peneliti telah mengembangkan estimator kurva regresi campuran di mana masing-masing pola data dalam model regresi dihampiri dengan estimator kurva yang sesuai dengan pola data (Nisa, 2017).

Estimator dalam regresi nonparametrik yang biasanya digunakan untuk memodelkan data curah hujan adalah estimator Kernel. Estimator Kernel dalam regresi nonparametrik umumnya digunakan apabila data yang diselidiki polanya tidak mengikuti suatu bentuk tertentu dan kecenderungan memiliki jenis variasi pola yang random (Afifah, 2017). Estimator lain yang biasanya digunakan dalam memodelkan curah hujan adalah estimator Deret Fourier. Estimator Deret Fourier dalam regresi nonparametrik umumnya digunakan apabila data yang diselidiki polanya tidak diketahui dan ada kecendrungan pola berulang. Pola data musiman menunjukkan suatu gerakan yang berulang dari suatu periode ke periode berikutnya secara teratur (Afifah, 2017).

Berdasarkan uraian di atas, pokok permasalahan dalam model regresi nonparametrik ini adalah bagaimana bentuk estimator kurva regresi campuran

nonparametrik Kernel dan Deret Fourier. Pokok permasalahan selanjutnya adalah penerapan estimator kurva regresi campuran nonparametrik tersebut pada data *time series*, yaitu data curah hujan Gunungsari Lombok Barat tahun 2013-2022.

## MATERI DAN METODE

### Data Penelitian

Data yang digunakan dalam penelitian ini berupa data bulanan curah hujan, rata-rata suhu udara, rata-rata kelembaban udara, dan rata-rata kecepatan angin Lombok Barat Stasiun Hujan Gunungsari bulan Januari 2013 - Desember 2022 yang diperoleh dari Badan Meteorologi Klimatologi dan Geofisika (BMKG) Stasiun Klimatologi Kelas I Nusa Tenggara Barat. Model regresi campuran nonparametrik Kernel dan Deret Fourier diaplikasikan pada data curah hujan Kecamatan Gunungsari Kabupaten Lombok Barat tahun 2013-2022. Variabel respon yang digunakan adalah curah hujan ( $Y$ ), sedangkan variabel prediktor adalah suhu udara ( $X_1$ ), kelembaban udara ( $X_2$ ), dan kecepatan angin ( $X_3$ ). Jumlah unit observasi sebanyak 120 Bulan.

### Prosedur Kerja

Pada pemodelan regresi nonparametrik campuran Kernel dan Deret Fourier pada data curah hujan langkah analisis data meliputi tahapan, yaitu:

1. Membuat model umum regresi nonparametrik campuran Kernel dan Deret Fourier.
2. Menginput data penelitian ke dalam *software R*.
3. Membuat *scatter plot*. Dari *scatter plot* didapatkan kemungkinan model yang terbentuk.
4. Membuat model khusus. Setelah mendapatkan variabel mana yang didekati dengan Kernel dan Deret Fourier, selanjutnya dapat dilakukan pemodelan curah hujan Gunungsari Lombok Barat dengan memodelkan campuran 2 Kernel dan 1 Deret Fourier serta memodelkan 1 Kernel dan 2 Deret Fourier.

#### a. Regresi Nonparametrik

Regresi nonparametrik adalah salah satu teknik statistik yang digunakan untuk menentukan hubungan antara variabel respon dan variabel prediktor yang fungsinya tidak diketahui (Yuliati & Sihombing, 2020). secara umum model regresi nonparametrik dapat dinyatakan sebagai berikut:

$$y_i = f(x_i) + \varepsilon_i, \quad i = 1, 2, 3, \dots, n \quad (1)$$

dengan,

$y_i$  : variabel respon ke- $i$

$x_i$  : variabel prediktor ke- $i$

$f(x_i)$  : fungsi regresi nonparametrik yang tidak diketahui bentuknya

$\varepsilon_i$  : *error* yang diasumsikan identik dan independent

## b. Regresi Nonparametrik Kernel

Regresi Kernel merupakan salah satu teknik nonparametrik untuk memperkirakan ekspektasi bersyarat dari variabel acak dengan tujuan untuk menemukan hubungan *nonlinier* antara variabel acak  $x$  dan  $y$ , dan juga untuk memperoleh estimator menggunakan bobot yang sesuai (Esty, 2014). Hubungan  $y_i$  dan  $x_i$  dapat dimodelkan dalam bentuk:

$$y_i = f(\mathbf{X}) + \varepsilon_i, \quad i = 1, 2, \dots, n \quad (2)$$

Dengan *error*  $\varepsilon_i$  dan memenuhi  $E(\varepsilon_i) = 0$ . Fungsi  $f(\cdot)$  diasumsikan tidak diketahui bentuknya tetapi *smooth*. Sedangkan  $\mathbf{X}$  merupakan beberapa variabel prediktor yang bersifat acak, sedangkan  $y_i$  variabel respon. Kurva regresi  $f(\mathbf{X})$  pada dasarnya merupakan mean bersyarat, dengan bentuk umum seperti berikut dan dapat diestimasi dengan estimator Kernel (Hadijati *et al.*, 2018):

$$f(\mathbf{X}) = E(y_i | \mathbf{x}) \quad (3)$$

Ada beberapa estimator Kernel yang dapat digunakan untuk mengestimasi kurva regresi, salah satunya adalah estimator *Nadaraya-Watson*. Estimator tersebut sering digunakan untuk kasus variabel prediktor acak (Afifah, 2017).

$$f(x) = \frac{\sum_{i=1}^n K\left(\frac{x_i - x}{\phi}\right) y_i}{\sum_{i=1}^n K\left(\frac{x_i - x}{\phi}\right)} \quad (4)$$

Dalam estimasi Kernel, bentuk Kernel dipengaruhi oleh fungsi Kernel  $K(x)$ . Ada beberapa jenis fungsi Kernel, salah satunya adalah Kernel *Gaussian* yang didefinisikan sebagai berikut (Afifah, 2017):

$$K(x) = \frac{1}{(2\pi)^{\frac{1}{2}}} e^{\left(-\frac{1}{2}x^2\right)}, \quad -\infty \leq x \leq \infty \quad (5)$$

## c. Regresi Nonparametrik Deret Fourier

Regresi Deret Fourier adalah salah satu teknik nonparametrik yang menggunakan fungsi basis sinus dan cosinus untuk mengestimasi hubungan antara  $x$  dan  $y$ . Deret Fourier adalah fungsi polinomial trigonometri yang memiliki fleksibilitas. Hubungan  $y_i$  dan  $x_i$  dapat dimodelkan dalam bentuk (Sahidah *et al.*, 2022):

$$f(x_i) = \frac{1}{2} \alpha_0 + \gamma_j x_{ij} + \sum_{k=1}^K \alpha_{jk} \cos kx_{ij} \quad (6)$$

## d. Regresi Nonparametrik Campuran Kernel dan Deret Fourier

Model campuran kurva regresi nonparametrik Kernel dan Deret Fourier adalah sebagai berikut (Afifah, 2017):

$$\mathbf{y} = \sum_{j=1}^p \hat{\mathbf{g}}_{j\phi_j}(v_{ji}) + \sum_{s=1}^q \mathbf{T}_s(\mathbf{t}_s) + \boldsymbol{\varepsilon} \quad (7)$$

$$\mathbf{y} = \mathbf{V}(\boldsymbol{\Phi})\mathbf{y} + \mathbf{X}\mathbf{a} + \boldsymbol{\varepsilon} \quad (8)$$

dengan,

$$\mathbf{V}(\Phi) = \begin{bmatrix} n^{-1} \sum_{j=1}^p W_{\Phi_{j1}}(v_{j1}) & n^{-1} \sum_{j=1}^p W_{\Phi_{j2}}(v_{j1}) & \dots & n^{-1} \sum_{j=1}^p W_{\Phi_{jn}}(v_{j1}) \\ n^{-1} \sum_{j=1}^p W_{\Phi_{j1}}(v_{j2}) & n^{-1} \sum_{j=1}^p W_{\Phi_{j2}}(v_{j2}) & \dots & n^{-1} \sum_{j=1}^p W_{\Phi_{jn}}(v_{j2}) \\ \vdots & \vdots & \ddots & \vdots \\ n^{-1} \sum_{j=1}^p W_{\Phi_{j1}}(v_{jn}) & n^{-1} \sum_{j=1}^p W_{\Phi_{j2}}(v_{jn}) & \dots & n^{-1} \sum_{j=1}^p W_{\Phi_{jn}}(v_{jn}) \end{bmatrix}$$

Entri pada matriks  $\mathbf{V}(\Phi)$  didefinisikan oleh kurva regresi Kernel *Nadaraya-Watson*  $n^{-1} \sum_{i=1}^n W_{\Phi_{ji}}(v_j) y_i$  dengan  $\Phi$  merupakan parameter *bandwidth*. Dan fungsi  $W_{\Phi_{ji}}(v_j)$  merupakan fungsi pembobot yang didefinisikan oleh  $W_{\Phi_{ji}}(v_j) = \frac{K_{\Phi_j}(v_j - v_{ji})}{n^{-1} \sum_{i=1}^n K_{\Phi_j}(v_j - v_{ji})}$  dengan  $K_{\Phi_j}(v_j - v_{ji})$  adalah fungsi Kernel  $K_{\Phi_j}(v_j - v_{ji}) = \frac{1}{\Phi_j} K\left(\frac{v_j - v_{ji}}{\Phi_j}\right)$ .

$$\mathbf{Xa} = \begin{bmatrix} t_{s1} & 1 & \cos t_{s1} & \cos 2t_{s1} & \dots & \cos M t_{s1} \\ t_{s2} & 1 & \cos t_{s2} & \cos 2t_{s2} & \dots & \cos M t_{s2} \\ t_{s3} & 1 & \cos t_{s3} & \cos 2t_{s3} & \dots & \cos M t_{s3} \\ \vdots & \vdots & \vdots & \vdots & \dots & \vdots \\ t_{sn} & 1 & \cos t_{sn} & \cos 2t_{sn} & \dots & \cos M t_{sn} \end{bmatrix} \begin{pmatrix} b_s \\ \frac{1}{2} a_{0s} \\ a_{1s} \\ \vdots \\ a_{Ms} \end{pmatrix}$$

Estimator regresi campuran nonparametrik Kernel dan Deret Fourier pada diperoleh melalui optimasi PLS sebagai berikut (Afifah, 2017):

$$\text{Min}_a \left\{ n^{-1} \|\mathbf{y} - \mathbf{V}(\Phi)\mathbf{y} - \mathbf{Xa}\|^2 + \sum_{s=1}^q \lambda_s \int_0^{\pi} \frac{2}{\pi} (T_s^2(t_{si}))^2 dt_s \right\}$$

$$\text{Min}_a \{ n^{-1} \|\mathbf{I} - \mathbf{V}(\Phi)\mathbf{y} - \mathbf{Xa}\|^2 + \mathbf{a}' \mathbf{D}(\lambda) \mathbf{a} \} \quad (9)$$

$\mathbf{a}' \mathbf{D}(\lambda) \mathbf{a}$  adalah ukuran kemulusan kurva. Dengan demikian, estimasi  $\mathbf{a}$  diperoleh melalui derivatif parsial dari Persamaan (9) terhadap  $\mathbf{a}$  diperoleh bentuk akhir sebagai berikut (Afifah, 2017):

$$\mathbf{S}(\Phi, \lambda, \mathbf{M})\mathbf{y} = \mathbf{X}(\mathbf{X}'\mathbf{X} + n\mathbf{D}(\lambda))^{-1} \mathbf{X}'(\mathbf{I} - \mathbf{V}(\Phi))\mathbf{y} \quad (10)$$

5. Membandingkan nilai GCV antar semua kemungkinan kombinasi Kernel dan Deret Fourier yang terbentuk menggunakan nilai GCV minimum untuk memilih model terbaik. Regresi Kernel bergantung pada pemilihan *bandwidth*  $\Phi$ . Sementara itu, Regresi Deret Fourier bergantung pada parameter penghalus  $\lambda$  dan parameter osilasi  $M$ . *bandwidth*  $\Phi$  dan parameter penghalus  $\lambda$  bertujuan untuk mengontrol kemulusan kurva. Sedangkan parameter osilasi  $M$  merupakan banyaknya osilasi dari gelombang cosinus pada model (Nisa, 2017). Metode GCV secara umum didefinisikan sebagai berikut (Afifah, 2017)

$$GCV(\Phi, \lambda, M) = \frac{MSE(\Phi, \lambda, M)}{(n^{-1} \text{trace}(\mathbf{I} - \mathbf{Z}(\Phi, \lambda, \mathbf{M})))^2} \quad (11)$$

$MSE(\Phi, \lambda, M)$  dan  $\mathbf{Z}(\Phi, \lambda, \mathbf{M})$  dicari dengan rumus (Afifah, 2017):

$$MSE(\Phi, \lambda, M) = n^{-1} \mathbf{y}'(\mathbf{I} - \mathbf{Z}(\Phi, \lambda, \mathbf{M}))'(\mathbf{I} - \mathbf{Z}(\Phi, \lambda, \mathbf{M}))\mathbf{y}$$

$$\mathbf{Z}(\Phi, \lambda, \mathbf{M}) = (\mathbf{V}(\Phi) + \mathbf{S}(\lambda, \mathbf{M})) \text{ dengan } \mathbf{S}(\lambda, \mathbf{M}) = \mathbf{X}(n^{-1}\mathbf{X}'\mathbf{X} + \lambda\mathbf{D})^{-1}n^{-1}\mathbf{X}$$

6. Uji asumsi *error* digunakan sebagai syarat kelayakan model. Dalam syarat analisis regresi untuk kebaikan model dalam menggambarkan data sebenarnya, dibutuhkan pengecekan dalam *error*. Asumsi *error* yang harus dipenuhi adalah sebagai berikut:

#### a. Uji Normalitas

Pada model regresi harus memiliki *error* yang mengikuti distribusi normal dengan *mean* nol dan varian  $\sigma^2$ . Untuk mengetahui apakah *error* telah berdistribusi normal salah satunya dapat dilakukan pengujian *Kolmogorov-Smirnov*. Hipotesis untuk uji *Kolmogorov-Smirnov*:

$H_0$  : *error* berdistribusi normal

$H_1$  : *error* tidak berdistribusi normal

Statistik uji yang digunakan:

$$D = \text{maks}|F_0(x) - S_N(x)| \quad (12)$$

Daerah penolakan adalah apabila  $|D| > D_\alpha$  dengan nilai  $D_\alpha$  didapatkan dari tabel *Kolmogorov-Smirnov* dengan tingkat signifikan  $\alpha$  (Gujarati, 2004).

#### b. Uji Identik

Pada model regresi harus memiliki varian antar *error* sama atau tidak terjadi heteroskedastisitas. Untuk mengetahui apakah *error* telah identik salah satunya dapat dilakukan pengujian uji *Glejser*. Uji *Glejser* adalah proses meregresikan harga mutlak *error* dengan variabel prediktor ( $x$ ).

$$\varepsilon_i = y - g(x_i) \quad (13)$$

Hipotesis yang digunakan adalah sebagai berikut:

$H_0$  :  $s_1^2 = s_n^2$  (tidak terjadi heteroskedastisitas)

$H_1$  :  $s_1^2 \neq s_n^2$  (terjadi heteroskedastisitas)

Statistik uji untuk uji *Glejser*:

$$F_{hitung} = \frac{\frac{[\sum_{i=1}^m (|\hat{\varepsilon}_i - \bar{\varepsilon}|)^2]}{(K-1)}}{\frac{[\sum_{i=1}^m (|\varepsilon_i - \bar{\varepsilon}|)^2]}{n-K}} \quad (14)$$

Dengan  $n = 1, 2, \dots, n$  merupakan banyaknya data dan  $K$  merupakan banyak variabel prediktor. Daerah penolakan hipotesis adalah tolak  $H_0$  jika  $|F_{hitung}| \geq F_{tabel}(F_{\alpha; K-1; n-K})$ , hal ini berarti terjadi heteroskedastisitas atau *error* tidak identik (Gujarati, 2004).

#### c. Uji Independen

Pada model regresi harus memastikan bahwa tidak terdapat korelasi antar *error* atau autokorelasi. Untuk mengetahui apakah *error* telah independen salah satunya dapat dilakukan pengujian uji *Durbin-Watson*. Uji *Durbin-Watson*. Hipotesis untuk uji *Durbin-Watson*:

$H_0 : \rho = 0$  (tidak terdapat autokorelasi)

$H_1 : \rho \neq 0$  (terdapat autokorelasi)

Statistik uji *Durbin-Watson*:

$$DW = \frac{\sum_{t=2}^n (\varepsilon_t - \varepsilon_{t-1})^2}{\sum_{t=1}^n \varepsilon_t^2} \quad (15)$$

Kriteria pengujiannya uji Durbin-Watson adalah sebagai berikut dengan nilai  $dL$  dan  $dU$  dapat dilihat di tabel *Durbin-Watson*. Jika  $D - W < dL$  atau  $D - W > 4 - dL$ , kesimpulannya pada data tersebut terdapat autokorelasi. Jika  $dU < D - W < 4 - dU$ , kesimpulannya pada data tersebut tidak terdapat autokorelasi. Tidak ada kesimpulan jika:  $4 - dU \leq D - W \leq 4 - dL$  atau  $dL \leq D - W \leq dU$  (Wei, 2006).

7. Melakukan prediksi dengan membandingkan data aktual dengan data prediksi sekaligus menghitung nilai ketepatan prediksi untuk mengukur keakuratan prediksi. Melihat ketepatan prediksi dapat dilakukan dengan menghitung *Mean Arctangent Absolute Percentage Error* atau biasa dikenal dengan MAAPE. MAAPE adalah salah satu cara untuk mengukur keakuratan prediksi. MAAPE digunakan jika salah satu nilai aktual sama dengan 0 karena komponen tersebut tidak akan terdefinisi. Nilai MAAPE dapat dicari menggunakan rumus (Suyudi, 2021):

$$MAAPE = \left( \frac{1}{n} \sum_{i=1}^n \arctan \left| \frac{\hat{y}_i - y_i}{y_i} \right| \right) \times 100\% \quad (16)$$

dengan,

$n$  : banyak data

$\hat{y}_i$  : data hasil prediksi pada periode ke- $i$

$y_i$  : data aktual pada periode ke- $i$

Kriteria keakuratan MAAPE adalah sebagai berikut:

**Tabel 1.** Kriteria Nilai MAAPE

Kriteria	Keputusan
< 10%	Kemampuan prediksi sangat akurat
10% - 20%	Kemampuan prediksi akurat
20% - 50%	Kemampuan prediksi cukup akurat
> 50%	Kemampuan prediksi tidak akurat

8. Menginterpretasi model dan membuat kesimpulan.

### Analisis Data

Pemilihan model regresi nonparametrik campuran Kernel dan Deret Fourier terbaik diperoleh dari penentuan *bandwidth* pada Kernel serta penentuan parameter penghalus dan parameter osilasi  $M$  pada Deret Fourier. Pemilihan *bandwidth*, parameter penghalus, dan parameter osilasi yang optimal dilakukan dengan cara memperkecil tingkat kesalahan. Semakin kecil tingkat kesalahan maka semakin baik estimasinya. Untuk mengetahui ukuran tingkat kesalahan suatu estimator dapat dilihat dari MSE (*Mean Square Error*) (Saputra, 2016). Salah satu metode yang digunakan untuk

memilih *bandwidth*, parameter penghalus, dan parameter osilasi yang optimal adalah GCV, karena GCV memiliki keterkaitan dengan MSE pada analisis regresi. Interval *bandwidth*  $\Phi$  yang digunakan dari 0-10 dan interval parameter penghalus  $\lambda$  yang digunakan dari 1-3. Pemilihan *bandwidth*, parameter penghalus, dan parameter osilasi  $M$  dilakukan secara simultan dengan *software R*.

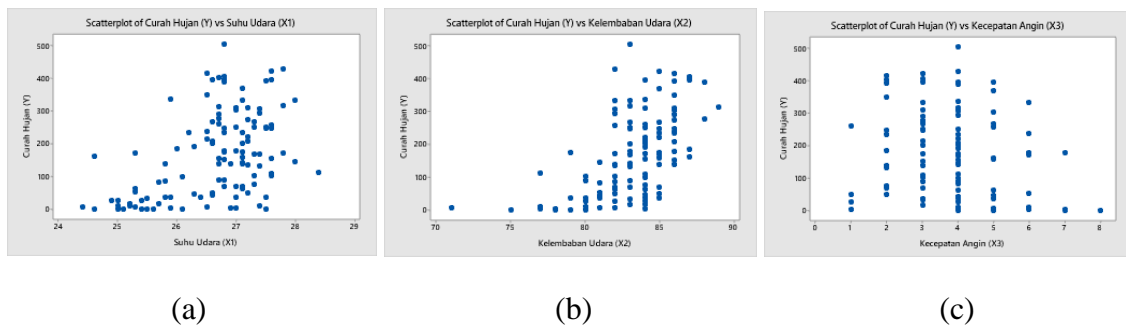
Berikut model umum regresi nonparametrik campuran Kernel dan Deret Forier. Model umum yang terbentuk seperti:

$$\begin{aligned}
 y_i &= \mu(v_{1i}, \dots, v_{ni}, t_{1i}, \dots, t_{ni}) + \varepsilon_i \\
 &= g_{1i}(v_{1i}) + \dots + g_{ni}(v_{ni}) + h_{1i}(t_{1i}) + h_{ni}(t_{ni}) + \varepsilon_i \\
 &= \frac{n^{-1} \sum_{i=1}^n K\left(\frac{v_{1i}-v}{\phi}\right) y_{1i}}{n^{-1} \sum_{i=1}^n K\left(\frac{v_{1i}-v}{\phi}\right)} + \dots + \frac{n^{-1} \sum_{i=1}^n K\left(\frac{v_{ni}-v}{\phi}\right) y_{ni}}{n^{-1} \sum_{i=1}^n K\left(\frac{v_{ni}-v}{\phi}\right)} + \\
 &\quad bt_{1i} + \frac{1}{2} a_0 + \sum_{k=1}^M a_k \cos kt_{1i} + \dots + bt_{ni} + \frac{1}{2} a_0 + \sum_{k=1}^M a_k \cos kt_{ni} \quad (12)
 \end{aligned}$$

## HASIL DAN DISKUSI

### Hasil

Sebelum proses pemodelan regresi nonparametrik campuran Kernel dan Deret Fourier pada data curah hujan, perlu dilihat pola hubungan antara variabel respon dengan masing-masing variabel prediktor dapat dilihat dari grafik *scatter plot*. Hasil *scatter plot* untuk masing-masing variabel respon dan variabel prediktor adalah sebagai berikut:



**Gambar 1.** *Scatter Plot* Curah Hujan dengan Suhu Udara, Kelembaban Udara, dan Kecepatan Angin

Gambar 1(a) menyajikan pola hubungan antara variabel respon curah hujan dengan variabel prediktor suhu udara. Hasil *plot* menggambarkan bahwa pola hubungan yang tidak mengikuti suatu bentuk pola tertentu sehingga pola tersebut dapat dimodelkan dengan regresi nonparametrik. Curah hujan dan suhu udara termasuk ke dalam kategori data musiman (Latief *et al.*, 2022). Data musiman jika dilihat pola datanya akan membentuk pola yang berulang seperti gelombang (Putri, 2022). Oleh karena itu, pada kasus ini pengaruh suhu udara terhadap curah hujan dimodelkan dengan regresi nonparametrik Deret Fourier.



Gambar 1(b) menyajikan pola hubungan dari variabel respon curah hujan dengan variabel prediktor kelembaban udara. Hasil *plot* menggambarkan bahwa pola hubungan yang tidak mengikuti suatu bentuk pola tertentu sehingga pola tersebut dapat dimodelkan dengan regresi nonparametrik. Curah hujan dan kelembaban udara termasuk ke dalam kategori data musiman (Aulia *et al.*, 2023). Data musiman ini jika dilihat pola datanya akan membentuk pola yang berulang seperti gelombang (Putri, 2022). Oleh karena itu, pada kasus ini pengaruh kelembaban udara terhadap curah hujan dimodelkan dengan regresi nonparametrik Deret Fourier.

Gambar 1(c) menyajikan pola hubungan variabel respon curah hujan dengan variabel prediktor kecepatan angin. Hasil *plot* menggambarkan bahwa pola hubungan yang tidak mengikuti suatu bentuk pola tertentu sehingga pola tersebut dapat dimodelkan dengan regresi nonparametrik. Curah hujan termasuk ke dalam kategori data musiman (Larief *et al.*, 2022). Sedangkan kecepatan angin termasuk ke dalam kategori data bukan musiman, dan cenderung memiliki variasi pola yang acak (Kiram, 2020). Oleh karena itu, pada kasus ini pengaruh kecepatan angin terhadap curah hujan dimodelkan dengan regresi nonparametrik Kernel.

Berdasarkan identifikasi pola yang terbentuk, terdapat tiga kemungkinan model regresi nonparametrik campuran Kernel dan Deret Fourier pada kasus ini. Berikut daftar lengkap dari tiga kemungkinan model dapat dilihat pada Tabel 2 berikut:

**Tabel 2.** Tiga Kemungkinan Model yang Terbentuk

No.	Variabel	
	Kernel	Deret Fourier
1	$X_3$	$X_1X_2$
2	$X_1X_3$	$X_2$
3	$X_2X_3$	$X_1$

Berdasarkan hasil perhitungan diperoleh nilai GCV untuk semua kemungkinan model campuran kernel dan deret fourier seperti terlihat pada Tabel 3 berikut:

**Tabel 3.** Perbandingan GCV Masing-masing Model

No.	Kernel	Deret Fourier	$M$	$\Phi_1$	$\Phi_2$	$\lambda_1$	$\lambda_2$	MSE	GCV
1	$X_3$	$X_1X_2$	1	0,7778	-	3	1	7.965	8.424
2	$X_3$	$X_1X_2$	2	0,7778	-	3	1	7.959	8.423
3	$X_3$	$X_1X_2$	3	0,7778	-	3	1	7.959	8.424
4	$X_3$	$X_1X_2$	4	7,0000	-	3	1	7.959	8.424
5	$X_3$	$X_1X_2$	5	2,3333	-	3	1	7.959	8.424
6	$X_1X_3$	$X_2$	1	0,4444	6,2222	1	-	8.520	9.043
7	$X_1X_3$	$X_2$	2	0,4444	6,2222	1	-	8.517	9.045
8	$X_1X_3$	$X_2$	3	0,4444	6,2222	1	-	8.517	9.046
9	$X_1X_3$	$X_2$	4	0,4444	6,2222	1	-	8.516	9.046
10	$X_1X_3$	$X_2$	5	0,4444	6,2222	1	-	8.516	9.046
11	$X_2X_3$	$X_1$	1	2	4,6666	3	-	9.805	10.360
12	$X_2X_3$	$X_1$	2	2	6,2222	1	-	9.796	10.354
13	$X_2X_3$	$X_1$	3	2	6,2222	1	-	9.795	10.354
14	$X_2X_3$	$X_1$	4	2	6,2222	1	-	9.794	10.353
15	$X_2X_3$	$X_1$	5	2	6,2222	1	-	9.794	10.353

Berdasarkan Tabel 3 diperoleh GCV minimum 8.423. Sehingga diperoleh  $M = 2$ ;  $\Phi = 0,7778$ ;  $\lambda_1 = 3$ ; dan  $\lambda_2 = 1$ . Serta estimasi regresi nonparametrik Deret Fourier dengan  $M = 2$  disajikan pada Tabel 4 berikut:

**Tabel 4.** Estimasi Komponen Deret Fourier

Parameter	Estimasi Deret Fourier
$b(\lambda_1)$	60,550
$a_0(\lambda_1)$	-1.806
$a_1(\lambda_1)$	-0,125
$a_2(\lambda_1)$	-0,017
$\lambda_1$	3
$b(\lambda_2)$	24,131
$a_0(\lambda_2)$	-1.806
$a_1(\lambda_2)$	4,854
$a_2(\lambda_2)$	-0,430
$\lambda_2$	1

Dengan demikian, model estimator regresi nonparametrik campuran Kernel dan Deret Fourier adalah sebagai berikut:

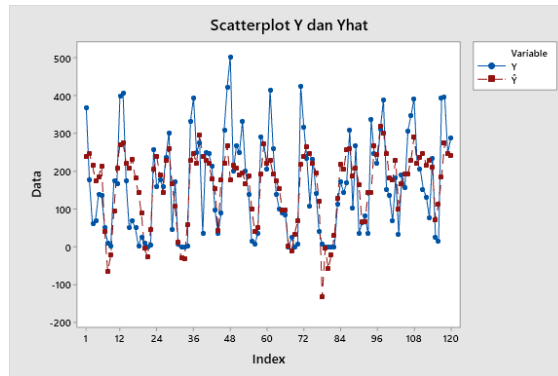
$$\hat{y}_i = \frac{\sum_{i=1}^{120} e^{\left(-\frac{1}{2}\left(\frac{v_i-v}{0,7778}\right)^2\right)} y_i}{\sum_{i=1}^{120} e^{\left(-\frac{1}{2}\left(\frac{v_i-v}{0,7778}\right)^2\right)}} +$$

$$60,550t_{1i} - 903 - 0,125 \cos t_{1i} - 0,017 \cos 2t_{1i} +$$

$$24,131t_{2i} - 903 + 4,854 \cos t_{2i} - 0,430 \cos 2t_{2i}$$

Pemodelan data curah hujan menggunakan estimator campuran kernel dan deret fourier menghasilkan nilai  $R^2$  sebesar 0,53 yang berarti 53% curah hujan Gunungsari Lombok Barat dapat dijelaskan oleh suhu udara, kelembaban udara, dan kecepatan angin, sedangkan sisanya 47% dijelaskan oleh sebab-sebab di luar model.

Dalam penelitian ini pengujian normalitas data menggunakan uji *Kolmogrov-Smirnov*. Berdasarkan hasil yang diperoleh, nilai *Kolmogrov-Smirnov* ( $D$ ) sebesar 0,063 < nilai tabel *Kolmogrov-Smirnov* sebesar 0,123 sehingga menjadi kesimpulan terima  $H_0$ . Yang berarti residual memenuhi asumsi mengikuti distribusi normal. Untuk pengujian identik dalam penelitian ini menggunakan uji *Glejser*. Hasil pengujian bahwa jika  $|F_{hitung}|$  sebesar 4,73 dan  $F_{tabel}(F_{0,05;3-1;120-3})$  diperoleh 3,07, sehingga  $|F_{hitung}| > F_{tabel}$  maka terjadi heterokedastisitas karena tolak  $H_0$  yang berarti ragam residual tidak homogen. Oleh karena itu, dapat disimpulkan bahwa pengujian asumsi identik tidak terpenuhi. Dan untuk pengujian independen menggunakan uji *Durbin-Watson*. Hasil pengujian nilai tabel *Durbin-Watson* dengan jumlah sampel 120 dan jumlah variabel prediktor 3 yaitu nilai  $dL = 1,613$  dan  $dU = 1,736$ . Karena nilai *Durbin-Watson* tidak berada diantara  $dL$  dan  $dU$  atau nilai *Durbin-Watson* <  $dL$  maka tolak  $H_0$  sehingga dapat disimpulkan bahwa terdapat autokorelasi antar residual yang artinya asumsi independen tidak terpenuhi. Dari model yang diperoleh, dapat digunakan untuk memprediksi curah hujan.



**Gambar 2.** Perbandingan Data Aktual dan Data Prediksi

Perbandingan berdasarkan Gambar 2 nilai data aktual terhadap nilai data prediksi menunjukkan bahwa hasil prediksi tidak mendekati data aktual. Untuk menghitung nilai ketepatan prediksi dapat dilakukan dengan menggunakan nilai *Mean Arctangent Absolute Percentage Error* (MAAPE). Hasil MAAPE melalui perhitungan diperoleh sebesar 60%. Oleh karena itu, kesimpulan berdasarkan kriteria keputusan MAAPE menyatakan bahwa kemampuan prediksi menggunakan estimator regresi nonparametrik campuran Kernel dan Deret Fourier pada kasus ini adalah tidak akurat. Hal tersebut diakibatkan karena uji identik dan uji independen tidak terpenuhi artinya terjadi heteroskedastisitas (*error* tidak identik) dan terdapat autokorelasi (*error* tidak independen). Pada penelitian ini dihasilkan *error* yang masih berpola. Artinya masih adanya unsur yang tidak dimasukkan ke dalam model atau masih terdapat variabel yang signifikan mempengaruhi variabel respon seperti unsur waktu, tempat, dll. Mengingat bahwa penelitian ini merupakan penelitian *time series* namun tidak memasukkan unsur waktu ke dalam model.

### Diskusi

Model untuk pola data yang mengikuti Kernel (kecepatan angin) tidak dapat diinterpretasi karena tidak memiliki koefisien dan karena model regresi nonparametrik Kernel tidak menggunakan asumsi tertentu mengenai bentuk kurva regresi maupun distribusi galat. Model ini hanya berfokus pada mengestimasi kurva regresi dengan menggunakan fungsi densitas Kernel. Namun pola data yang mengikuti Deret Fourier (suhu udara dan kelembaban udara) dapat diinterpretasi. Berdasarkan Tabel 5.8 terlihat nilai  $M$  optimal adalah 2. Hal ini dikarenakan nilai  $M = 2$  menghasilkan nilai GCV yang lebih kecil dibandingkan dengan  $M = 1$ ,  $M = 3$ ,  $M = 4$ , dan  $M = 5$ . Jumlah nilai  $M$  sebenarnya bisa lebih dari 5, namun di sini peneliti membatasi menggunakan hanya pada 5 osilasi  $M$ . Karena jika semakin besar akan berdampak pada jumlah parameter yang diestimasi dan model yang dihasilkan tidak sederhana. Berikut ini adalah model curah hujan masing-masing kelompok data tersebut beserta interpretasi:

1. Jika  $v$  dan  $t_2$  dianggap konstan, maka pengaruh suhu udara terhadap curah hujan ( $Y$ ) adalah:

$$\hat{y}_i = 60,550t_{1i} - 903 - 0,125 \cos t_{1i} - 0,017 \cos 2t_{1i}$$

Hasil estimasi menggambarkan bahwa terdapat 2 osilasi di mana osilasi pertama dan kedua curah hujan akan menurun dengan *trend* negatif yang fluktuatif mengikuti gelombang cosinus.

2. Jika  $v$  dan  $t_1$  dianggap konstan, maka pengaruh kelembaban udara terhadap curah hujan ( $Y$ ) adalah:

$$\hat{y}_i = 24,131t_{2i} - 903 + 4,854 \cos t_{2i} - 0,430 \cos 2t_{2i}$$

Hasil estimasi menggambarkan bahwa terdapat 2 osilasi di mana pada osilasi pertama curah hujan akan naik dengan *trend* positif yang fluktuatif mengikuti gelombang cosinus dan pada osilasi kedua curah hujan akan menurun dengan *trend* negatif yang fluktuatif mengikuti gelombang cosinus.

Dalam regresi Deret Fourier, koefisien parameter menggambarkan kontribusi dari setiap gelombang dalam menjelaskan variasi data. Koefisien ini menunjukkan seberapa besar amplitudo (jarak terjauh gelombang) dan fase (jumlah gelombang yang terbentuk) dari setiap gelombang yang terlibat dalam model regresi. Semakin besar koefisien, semakin besar pula kontribusi gelombang tersebut dalam menjelaskan variasi data. Namun, interpretasi koefisien parameter dalam regresi Deret Fourier tidaklah mudah karena koefisien tersebut tidak memiliki makna yang intuitif seperti pada regresi linier. Oleh karena itu, interpretasi koefisien parameter dalam regresi Deret Fourier lebih bersifat teknis dan memerlukan pemahaman yang mendalam.

## KESIMPULAN

Berdasarkan analisis yang telah dilakukan dapat diambil kesimpulan sebagai berikut:

1. *Bandwidth*  $\Phi$  serta parameter penghalus  $\lambda$  dan parameter osilasi  $M$  optimal sehingga memperoleh model terbaik diperoleh  $\Phi = 0,7778$ ;  $M = 2$ ;  $\lambda_1 = 3$ ; dan  $\lambda_2 = 1$  di mana komponen kernel terdiri dari 1 variabel dan komponen deret fourier terdiri dari 2 variabel yang ditentukan berdasarkan nilai GCV minimum.
2. Model regresi nonparametrik campuran kernel dan deret fourier diterapkan pada data curah hujan di Gunung Sari Kabupaten Lombok Barat tahun 2013-2022 dengan variabel responnya adalah curah hujan ( $Y$ ), sedangkan variabel prediktornya adalah suhu udara ( $t_1$ ), kelembaban udara ( $t_2$ ), dan kecepatan angin ( $v$ ). Model estimator regresi nonparametrik campuran kernel dan deret fourier adalah sebagai berikut dan pemodelan ini menghasilkan  $R^2$  sebesar 53%.

$$\hat{y}_i = \frac{\sum_{i=1}^{120} e^{\left(-\frac{1}{2}\left(\frac{v_i-v}{0,7778}\right)^2\right)} y_i}{\sum_{i=1}^{120} e^{\left(-\frac{1}{2}\left(\frac{v_i-v}{0,7778}\right)^2\right)}} + 60,550t_{1i} - 903 - 0,125 \cos t_{1i} - 0,017 \cos 2t_{1i} + 24,131t_{2i} - 903 + 4,854 \cos t_{2i} - 0,430 \cos 2t_{2i}$$

## UCAPAN TERIMAKASIH

Penghargaan dan terima kasih yang setulus-tulusnya penulis sampaikan kepada Allah SWT selalu melimpahkan rahmat, kesehatan, karunia, dan keberkahan di dunia dan di akhirat atas budi baik yang telah diberikan, Ayahanda tercinta Nasrulloh S.Pd dan Ibunda tercinta Wahyu Andriani yang telah mencurahkan segenap cinta dan kasih sayang serta perhatian moril maupun materil kepada penulis. Penghargaan dan terima kasih juga penulis berikan kepada Ibu Mustika Hadijati, S.Si., M.Si. selaku Pembimbing I, dan Ibu Nuzla Af'idatur Robbaniyyah, S.Si., M.Sc. selaku Pembimbing II atas bimbingannya selama penulisan jurnal ini.

## DAFTAR PUSTAKA

- Afifah, N., 2017, Estimator Campuran Kernel dan Deret Fourier dalam Regresi Nonparametrik (Studi Kasus: Pemodelan Persentase Pernduduk Miskin di Indonesia), Tesis, Program Magister Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Institut Teknologi Sepuluh Nopember, Surabaya.
- Aldrian, E., Karmini, M., dan Budiman, 2011, Adaptasi dan Mitigasi Perubahan Iklim di Indonesia, Jakarta: Pusat Perubahan Iklim dan Kualitas Udara Kedepujian Bidang Klimatologi Badan Meteorologi, Klimatologi, dan Geofisika (BMKG).
- Aulia, N., Sumarjaya, I.W., dan Srinadi, I.G.A.M., 2023, Peramalan Curah Hujan di Bandara Internasional I Gusti Ngurah Rai, Menggunakan Metode Fungsi Transfer, E-Jurnal Matematik, 12(3):209-215.
- Esty, 2014, Regresi Kernel dengan Metode Nadaraya Watson, Skripsi, Program Studi Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Negeri Yogyakarta: Yogyakarta.
- Fazhuddin, M.I., 2018, Model Regresi Nonparamtrik Smoothing Spline pada Data Curah Hujan di Jawa Tengah, Skripsi, Program Studi Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Sebelas Maret, Surakarta.
- Gujarati, D.N., 2004, Basic Econometric 4th edition, New York: The Mc Gra Hill Companies.
- Hadijati, M., Komalasari, D., Irwansyah., 2018, Model Statistical Downscaling Nonparametrik pada Simulasi Data Curah Hujan Harian Pos Jurang Malang Daerah Aliran Sungai Jangkok, Eigen Mathematics Journal, 1(2): 43-47.
- Kiram, M.R., 2020, Peramalan Kecepatan Angin di Kabupaten Banyuwangi dengan Menggunakan Metode Hybrid Ensemble Empirical Mode Decomposition-Autoregressive Integrated Moving Average (EEMD-ARIMA), Tesis, Program Magister Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Institut Teknologi Sepuluh Nopember, Surabaya.
- Latief, N.H., Ani, N., dan Setiawan, I., 2022, Peramalan Curah Hujan di Kota Makassar dengan Menggunakan Metode SARIMAX, Forum Teori dan Aplikasi Statistika, 22(1):55-63.

- Nisa, K., 2017, Model Regresi Semiparametrik Campuran Spline Truncated dan Deret Fourier (Studi Kasus : Angka Harapan Hidup Provinsi Jawa Timur), Tesis, Program Magister Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Institut Teknologi Sepuluh Nopember, Surabaya.
- Putri, H.M., 2022, Penerapan Metode Dekomposisi dalam Peramalan Data Deret Waktu (Time Series), Skripsi, Program Studi Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung, Bandar Lampung.
- Sahidah, Kuzairi, dan Mardianto, M.F.F., 2022, Estimator Deret Fourier dalam Regresi Nonparametrik dengan Penalti untuk Perencanaan Penjualan Produk Musiman, Zeta-Math Journal, 7(2):69-78.
- Suyudi, A.H., 2021, Prediksi Lokasi, Kekuatan, dan Waktu Gempa Bumi di Wilayah Sulawesi Menggunakan Model Semi-Marcov, Skripsi, Fakultas Sains dan Teknologi UIN Syarif Hidayatullah, Jakarta.
- Wei, W.W., 2006, Time Series Analisis: Univariate and Multivariate, United States: Pearson Education Inc.
- Yuliati, I.F., dan Sihombing, P.R., 2020, Pemodelan Fertilitas di Indonesia Tahun 2017 Menggunakan Pendekatan Regresi Nonparametrik Kernel dan Spline, Jurnal Statistika dan Aplikasinya (JSA), 4(1):48-60.
- Yusuf, M., 2021, BPBD NTB Sebut Banjir Terparah Terjadi di Lombok Barat (<https://m.antaranews.com>), Diakses Jam 16:49 WITA, Tanggal 20/10/2022.