E-Jurnal Matematika Vol. 15(4), November 2026, pp. 294-301 ISSN: 2303-1751
DOI: https://doi.org/10.24843/MTK.2026.v15.i04.p533

ANALISIS PERBANDINGAN MODEL SARIMA DAN ARIMAX DALAM PERAMALAN KEJADIAN DBD BERDASARKAN FAKTOR IKLIM DI PROVINSI NUSA TENGGARA TIMUR

Anjelina Maria I. Goran[1§], Arista M. Tamonob[2], Jusrry R. Pahnael[3], Mira Wadu[4]

[1]Program Studi Matematika, FST, Universitas Nusa Cendana [anjelinam073@gmail.com]
[2]Program Studi Matematika, FST, Universitas Nusa Cendana [arista.tamonob@staf.undana.ac.id]
[3]Program Studi Matematika, FST, Universitas Nusa Cendana [jusrry_pahnael@staf.undana.ac.id]
[4]Program Studi Matematika, FST, Universitas Nusa Cendana [mira_wadu@staf.undana.ac.id]
[§]*Corresponding Author*

ABSTRACT

Dengue Hemorrhagic Fever (DHF) is a public health problem influenced by climate factors and exhibits seasonal patterns. This study aims to compare the SARIMA and ARIMAX models in forecasting DHF cases in East Nusa Tenggara Province. The data used consist of monthly DHF cases and climate variables, including temperature, rainfall, humidity, and wind speed from January 2015 to December 2024. The method applied is time series analysis, including model identification, parameter estimation, diagnostic checking, and accuracy evaluation using MAPE, MSE, and RMSE. The results show that the best SARIMA model is SARIMA$(1,0,0)(1,0,1)^{12}$ *with MAPE of* 81.90%, *, MSE of* 38147.42, *and RMSE of* 195.3136. *Meanwhile, the best ARIMAX model is ARIMAX*$(2,0,0)$ *with MAPE of* 57.11%, *MSE of* 259.6, *and RMSE of* 161.14. *The comparison indicates that ARIMAX provides better accuracy than SARIMA, making it more effective for forecasting DHF cases as it incorporates climate factors.*

Keywords: *SARIMA, ARIMAX, DHF, forecasting, climate factors*

1. PENDAHULUAN

Demam Berdarah Dengue (DBD) merupakan salah satu penyakit menular yang hingga saat ini masih menjadi permasalahan penting dalam bidang kesehatan masyarakat. Penyakit ini disebabkan oleh virus dengue dan ditularkkan melalui gigitan nyamuk *Aedes aegypti*. Dalam beberapa tahun terakhir, jumlah kasus DBD secara global menunjukkan peningkatan yang signifikan. Laporan World Health Organization (WHO, 2024) mencatat lebih dari 14,6 juta kasus DBD, menjadikannya angka tertinggi yang pernah dilaporkan secara global, sehingga menunjukkan bahwa DBD masih menjadi ancaman serius bagi kesehatan masyarakat dunia.

Di Indonesia, DBD juga masih menjadi masalahan kesehatan masyarakat, terutama di wilayah dengan kondisi lingkungan dan iklim yang mendukung perkembangan vektor penularnya. Provinsi NTT termasuk salah satu daerah dengan risiko tinggi penularan DBD. Berdasarkan data BPS Provinsi (BPS, 2025), yang bersumber dari Dinas Kesehatan Provinsi NTT, menunjukkan adanya fluktuasi kasus yang cukup tajam dalam empat tahun terakhir. Pada tahun 2021 tercatat sebanyak 2.543 kasus, kemudian meningkat menjadi 3.376 kasus pada tahun 2022. Jumlah tersebut menurun menjadi 2.629 kasus pada tahun 2023, namun kembali meningkat tajam hingga mencapai 3.886 kasus pada tahun 2024. Fluktuasi tersebut tidak terlepas dari berbagai penyebab, salah satunya perubahan iklim (Bone & Kaunang, 2021).

Fluktuasi kasus DBD yang berulang setiap tahun, terutama pada musim hujan, menunjukkan adanya pola musiman sehingga analisis runtun waktu (time series) menjadi pendekatan yang tepat untuk peramalan. Model SARIMA (Seasonal Autoregressive Integrated Moving Average), mampu menangkap pola musiman pada data historis, namun belum mempertimbangkan pengaruh variabel eksternal (Buntara et al., 2023) Oleh karena itu, digunakan model ARIMAX (ARIMA with Exogenous Variables) yang dapat memasukkan variabel iklim sebagai faktor eksternal dalam peramalan

kasus DBD (Nanda & Witanti, 2023).

Walaupun kedua model memiliki pendekatan yang berbeda, perbandingan kinerja SARIMA dan ARIMAX dalam peramalan kasus DBD di Provinsi NTT masih perlu dikaji lebih lanjut. Oleh karena itu, penelitian ini bertujuan untuk membandingkan kemampuan model SARIMA dan ARIMAX dalam meramalkan jumlah kasus DBD berdasarkan data bulanan periode Januari 2015 hingga Desember 2024. Evaluasi dilakukan menggunakan MAPE, MSE, dan RMSE. Hasil penelitian diharapkan dapat memberikan informasi mengenai model yang memiliki tingkat akurasi lebih baik sehingga dapat mendukung upaya pengamatan dan pengendalian DBD di Provinsi NTT(Junaedi et al., 2025).

2. METODE PENELITIAN

2.1 Sumber Data

Penelitian ini menggunakan data sekunder berupa data bulanan kasus DBD yang diperoleh dari Dinas Kesehaan Provinsi NTT dan data bulanan faktor iklim berupa suhu, curah hujan, kelembapan, dan kecepatan angin yang diperoleh dari BMKG Provinsi NTT periode Januari 2015-Desember 2024.

2.2 Metode Analisis Data

Langkah-langkah analisis data adalah sebagai berikut:

1. Analisis Statistik Deskriptif
2. Membuat plot
3. Membagi data training dari Januari 2015 – Desember 2022 dan data testing dari Januari 2023 – Desember 2024;
4. Melakukan pemodelan SARIMA dan ARIMAX, dengan langkah-langkah sebagai berikut:
   a. Melakukan uji stasioner data dalam mean dan varian
   b. Melakukan Transformasi Box-Cox dan proses differencing (musiman dan non-musiman) jika data tidak stasioner dalam varians dan mean
   c. Analisis ACF dan PACF musiman dan non musiman untuk menentukan orde model AR(p), AR(P), MA(q) dan MA(Q)
   d. Melakukan identifikasi parameter dan variabel eksternal
   e. Estimasi parameter dan variabel eksternal
   f. Diagnostik model
   g. Peramalan kejadian DBD dari Januari 2023 – Desember 2024.
   h. Seleksi model SARIMA dan ARIMAX terbaik
5. Membandingkan hasil peramalan dari model SARIMA dan ARIMAX
6. Menarik kesimpulan.

2.3 Runtun Waktu

Analisis runtun waktu merupakan metode statistik yang digunakan untuk memeriksa data deret yang diambil secara berurutan dalam interval waktu yang tetap misalnya bulanan dan tahunan (Khoiri, 2023). Data runtun waktu umumnya tersusun atas beberapa komponen utama yang menggambarkan pola perubahan nilai variabel dari waktu ke waktu. Komponen-komponen tersebut meliputi tren, musiman, siklus, dan acak (Hasanah, 2022).

2.4 Seasonal Autoregressive Integrated Moving Average (SARIMA)

Model SARIMA (Seasonal Autoregressive Integrated Moving Average) merupakan pengembangan dari ARIMA yang memasukkan komponen musiman ke dalam proses pemodelannya (Hasby et al., 2025).
Secara umum, bentuk model SARIMA dapat ditulis sebagai berikut :

$$\phi_p(B)\phi_P(B^s)(1-B)^d(1-B^s)^D X_t = \theta_q(B)\theta_Q(B^s)e_t \quad (1)$$

Dengan X_t : variabel deret waktu pada periode ke-t, B: backshift (operator lag), $\phi_p(B)$: parameter AR non-musiman dengan orde p, $\phi_P(B^s)$: parameter AR musiman dengan orde P, $\theta_q(B)$: parameter MA non-musiman dengan orde q, $\theta_Q(B^s)$: parameter MA musiman dengan orde Q, $(1-B)^d$: diferensiasi non-musiman ke-d, e_t : residual.

2.5 Autoregressive Integrated Moving Average with Exogenous Variable (ARIMAX)

Model ARIMAX merupakan pengembangan dari ARIMA dengan menambahkan variabel eksogen (X). Tujuannya adalah untuk meningkatkan akurasi peramalan dengan mempertimbangkan faktor-faktor luar yang berpengaruh terhadap variabel utama yang diamati (Riestiansyah et al., 2022). Untuk

menentukan model tersebut signifikan atau tidak dapat dilihat dari parameternya jika semua parameter AR dan MA signifikan dan terdapat salah satu variabel yang signifikan maka model tersebut dapat dinyatakan signifikan dan jika salah satu dari parameter AR dan MA tidak signifikan maka model tersebut tidak sigifikan (Pradana et al., 2022).

Secara umum, bentuk model ARIMAX dapat ditulis sebagai berikut :

$$\begin{aligned}(1-B)^d X_t = {} & \beta_1 X_{t-1} + \cdots + \beta_p X_{t-1} \\ & + x_1 Y_{t-1} + \cdots \\ & + x_k Y_{t-k} + e_t \\ & + \alpha_1 e_{t-1} + \cdots \\ & + \alpha_q e_{t-q}\end{aligned} \quad (2)$$

Dengan B : differencing, X_t : variabel dependen pada waktu t, β_i : koefisien Autoregressive $(i = 1, 2, 3, ...)$,Y : variabel independen, e_t : nilai galat pada waktu ke-t, α_i : parameter MA $(i = 1, 2, 3, ...)$.

2.6 Diagnostik Model

Tujuan dari uji diagnostik model adalah untuk memastikan bahwa model yang telah dipilih sudah sesuai dengan melakukan uji White-noise.

Uji hipotesis

H_o: $\rho_1 = \rho_2 = \cdots = \rho_k = 0$ (tidak ada autokorelasi pada residual atau model sesuai)

H_1: minimal ada satu $\rho_i \neq 0$ (ada autokorelasi pada residual atau model belum sesuai)

Satistik uji dengan uji Ljung-Box :

$$Q^* = n(n+2)\sum_{k=1}^{k}\frac{\hat{\rho}_k^2}{n-k} \quad (3)$$

Dengan Q^*: parameter berdistribusi chi-square, n : jumlah observasi, k : lag maksimum, $\hat{\rho}_k^2$: ACF residual lag ke-k.

2.7 Metode Kearutan Peramalan

2.7.1 Mean Square Error (MSE)

MSE digunakan untuk menghitung rata-rata kuadrat selisih antara nilai aktual dan hasil prediksi. Nilai MSE yang dipilih adalah nilai yang paling kecil atau mendekati nol, karena semakin rendah nilai MSE, semakin kecil tingkat kesalahan hasil peramalan yang dihasilkan oleh suatu model (Katabba, 2021). Persamaan MSE sebagai berikut :

$$MSE = \frac{1}{n}\sum_{t=1}^{n}(X_t - F_t)^2 \quad (4)$$

Dengan X_t : nilai aktual pada periode ke-t, F_t : nilai hasil peramalan pada periode ke-t, n : jumlah observasi

2.7.2 Root Mean Square Error (RMSE)

RMSE merupakan akar dari nilai MSE. Nilai RMSE menunjukkan besarnya kesalahan peramalan dalam satuan yang sama dengan data aktual. Semakin kecil nilai RMSE, semakin baik kemampuan model dalam melakukan peramalan (Zaki et al., 2023).

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(X_t - F_t)^2} \quad (5)$$

Dengan X_t : nilai aktual pada periode ke-t, F_t : nilai hasil peramalan pada periode ke-t, n : jumlah observasi.

2.7.3 Mean Absolute Percentage Error (MAPE)

MAPE digunakan untuk mengukur tingkat kesalahan peramalan dalam bentuk persentase, sehingga lebih mudah dipahami secara praktis. Persamaan MAPE dapat ditulis sebagai berikut :

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{X_t - F_t}{X_t}\right| \times 100\% \quad (6)$$

Dengan X_t : nilai aktual pada periode ke-t, F_t : nilai hasil peramalan pada periode ke-t, n : jumlah observasi.

Skala nilai error dari MAPE adalah sebagai berikut :

Tabel 1 Kriteria Nilai MAPE

MAPE	Keterangan
$< 10\%$	Prediksi sangat baik
$10 - 20\%$	Prediksi baik
$20 - 50\%$	Prediksi cukup baik/layak
$> 50\%$	Prediksi buruk

Sumber: Azzulfa & Primandari, 2026

3. HASIL DAN PEMBAHASAN

3.1 Analisis Deskriptif

Analisis deskriptif dilakukan untuk memberikan gambaran umum mengenai karakteristik data yang digunakan dalam penelitian ini.

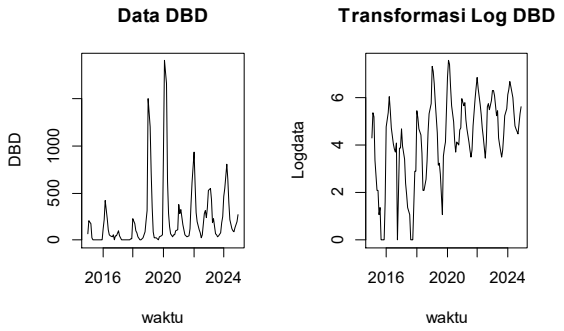
Jumlah kasus DBD di Provinsi Nusa Tenggara Timur periode 2015-2024 memiliki variasi sangat tinggi dengan rentang 0-1898 kasus perbedaan mean (217.8) dan median 98 mununjukkan data cenderung tidak simeteris akibat adanya lonjakn kasus pada periode tertentu.

Suhu udara relatif stabil dengan rentang $20.7℃ - 30.1℃$ serta mean $(27.18℃)$ yang mendekati median $(27.2℃)$. curah hujan memiliki variasi tinggi dengan nilai minimum 0 dan maksimum 712, sedangkan selisih mean (145.23) dan median (79.05) menunjukkna adanya nilai ekstrim. Kelembapan udara cenderung stabil dengan rentang 57-89 serta mean (78.9) yang mendekati median (78). Sementara itu, kecepatan angin memiliki variasi yang tidak terlalu besar, dengan nilai minimum 0 dan maksimum 5.1, serta mean (2.4) yang mendekati mediam (2.6).

3.2 Kestasioneran Data

1. Stasioenr dalam varians

Sebagai bagiann dari pengujian kestasioneran data, dlakukan uji stasioner dalam varians untuk memastikan penyebaran data tetap konstan sepanjang waktu sehingga memenuhi asumsi analisis deret waktu.



Gambar 1 Plot *Box-Cox* DBD di NTT

Berdasarkan Gambar 1, data asli kasus DBD menunjukkan fluktuasi yang tinggi dengan beberapa lonjakan tajam serta pola musiman pada periode tertentu, sehingga mengindikasikan varians belum konstan. Oleh karena itu, dilakukan transformasi logaritma untuk menstabilkan varians dan mengurangi pengaruh nilai ekstrem.

2. Stasioner dalam mean

Pada tahap pengujian stasioneritas dalam varians, diperoleh hasil bahwa data telah memenuhi asumsi stasioneritas dalam varians, sehingga selanjutnya dilakukan pengujian kestasioneran dalam mean menggunakan uji Augmented Dickey-Fuller (ADF). Pengujian ini bertujuan untuk memastikan bahwa rata-rata data tidak berubah secara sistematis dari waktu ke waktu sehingga data berada dalam kondisi stabil untuk pemodelan deret waktu. Hipotesis yang digunakan yaitu:

H_0: data tidak stasioner dalam mean
H_1: data stasioner dalam mean

Tabel 1 Output Augmented Dickey-Fuller (ADF)

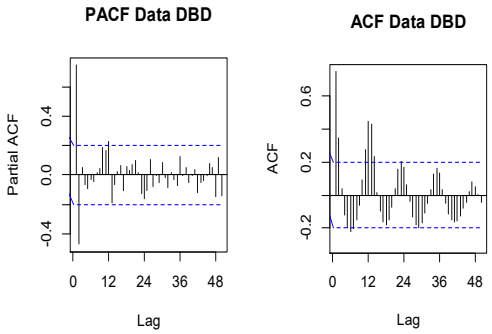
Augmented Dickey-Fuller (ADF)			
Dickey-Fuller	Lag Order	*P-value*	Keterangan
-6.3241	4	0.01	Stasioner

Sumber: data diolah (2026)

Berdasarkan Tabel 2, nilai p-value 0,01 menunjukkan bahwa hipotesis nol ditolak pada tingkat signifikansi 5%. Dengan demikian, data dapat dianggap stasioner dalam mean.

3.3 Identifikasi Model

Tahap identifikasi model dilakukan untuk menentukan kombinasi parameter model yang sesuai dengan karakteristik data deret waktu. Proses identifikasi dilakukan dengan menganalisis pola *Autocorrelation Function* (ACF) dan *Partial Autocorrelation Function* (PACF). Plot ACF digunakan untuk mengidentifikasi orde parameter *Moving Average* (MA), sedangkan plot PACF digunakan untuk menentukan orde parameter *Autoregressive* (AR). Pada model SARIMA, pola ACF dan PACF juga dianalisis untuk mengidentifikasi parameter musiman, sehingga diperoleh kombinasi model yang sesuai dalam menggambarkan pola dan dinamika data.



Gambar 2 Plot ACF dan PACF DBD di NTT

Berdasarkan Gambar 2, plot PACF data DBD, menunjukkan lag yang signifikan pada

lag 1, lag 2 dan lag 12. Hal ini menunjukkan bahwa PACF non-musiman mengalami cut off pada lag ke-2, sehingga nilai parameter AR non-musiman adalah $p = 1$ dan $p = 2$ dan PACF musiman mengalami cut off pada lag ke-12, sehingga diperoleh parameter AR adalah $P = 1$.

Plot ACF data DBD menunjukkan autokorelas signifikansi pada lag awal seperti lag 1 dan lag 2, serta pada beberapa lag musiman yaitu lag 12 dan lag 24. Pola ACF mengalami penurunan perlahan , sehingga mengindikasikan adanya komponen AR yang dominan, sedangkan parameter MA non-musiman yang mungkin adalah $q = 0$. Pada komponen musiman, ACF mengalami cut off pada lag 12 dan lag 24, sehingga diperoleh parameter MA musiman adalah $Q = 1$ dan $Q = 2$.

Berdasarkan pola ACF dan PACF, diperoleh model ARIMA dan SARIMA yang mungkin adalah ARIMA$(1,0,0)$,ARIMA $(2,0,0)$,ARIMA $(1,0,1)$, ARIMA $(2,0,1)$, ARIMA$(1,0,2)$,ARIMA$(2,0,2)$,ARIMA$(0,0,1)$,SARIMA$(1,0,0)(1,0,0)^{12}$,SARIMA $(2,0,0)(1,0,0)^{12}$,SARIMA$(1,0,1)(1,0,0)^{12}$,SA RIMA$(2,0,1)(1,0,0)^{12}$,SARIMA $(1,0,0)(0,0,1)^{12}$,SARIMA$(2,0,0)(0,0,1)^{12}$,SA RIMA$(1,0,1)(0,0,1)^{12}$,SARIMA $(2,0,1)(0,0,1)^{12}$,SARIMA$(1,0,2)(1,0,1)^{12}$,SA RIMA$(2,0,2)(1,0,1)^{12}$,SARIMA $(0,0,1)(0,0,1)^{12}$,SARIMA$(1,0,0)(2,0,0)^{12}$,SA RIMA$(1,0,0)(1,0,1)^{12}$SARIMA $(1,0,0)(2,0,1)^{12}$,SARIMA$(2,0,0)(1,0,1)^{12}$,SA RIMA$(2,0,0)(1,0,2)^{12}$,SARIMA $(2,0,0)(0,0,2)^{12}$,SARIMA$(2,0,1)(1,0,1)^{12}$,SA RIMA$(1,0,1)(1,0,1)^{12}$.

3.4 Estimasi dan Signifikan Parameter Model SARIMA dan ARIMAX

1. Model SARIMA

Tahap selanjutnya adalah melakukan estimasi dan pengujian signifikansi parameter. Parameter yang tidak signifikan menunjukkan bahwa model SARIMA dan ARIMAX yang dihasilkan juga tidak signifikan.

Model SARIMA yang diuji pada taraf nyata α=5%, terdapat 7 model signifikan dan 12 model tidak signifikan. Kesembilan belas model tersebut dibentuk berdasarkan pola ACF, PACF, serta prinsip kesederhanaan model agar diperoleh model yang mampu mewakili karakteristik data tanpa terlalu kompleks. Sementara model lainnya dinyatakan tidak signifikan karena terdapat parameter yang tidak berkontribusi nyata terhadap model. Hasil ini menunjukkan bahwa model dengan struktur yang lebih sederhana cenderung menghasilkan parameter yang lebih stabil dan signifikan, sedangkan penambahan parameter berlebih dapat menyebabkan *overparameterization*. Oleh karena itu, ketujuh model signifikan tersebut layak digunakan pada tahap analisis selanjutnya.

2. Model ARIMAX

Selain menggunakan model SARIMA, penelitian ini juga menggunakan model ARIMAX dengan variabel eksternal berupa suhu, curah hujan, kelembapan, dan kecepatan angin untuk mengetahui pengaruh faktor iklim terhadap jumlah kasus DBD serta meningkatkan akurasi peramalan. Model dinyatakan signifikan apabila seluruh parameter AR dan MA signifikan serta terdapat minimal satu variabel eksogen yang signifikan.

7 model ARIMAX yang diuji pada taraf nyata α=5%, terdapat 5 model signifikan dan 2 model tidak signifikan. Model-model tersebut dibentuk berdasarkan pola ACF, PACF, serta mempertimbangkan kesederhanaan model dan variabel eksogen. Model signifikan yaitu ARIMAX (1,0,0), ARIMAX (2,0,0), ARIMAX (1,0,1), ARIMAX (1,0,2), dan ARIMAX (0,0,1), sedangkan ARIMAX (2,0,1) dan ARIMAX (2,0,2) dinyatakan tidak signifikan karena masih terdapat parameter AR atau MA yang tidak signifikan. Hasil ini menunjukkan bahwa model yang lebih sederhana cenderung memiliki parameter yang lebih stabil dan signifikan dibandingkan model yang lebih kompleks. Selain itu, variabel kelembapan udara (%) secara konsisten memberikan pengaruh signifikan terhadap jumlah kasus DBD, sehingga menjadi variabel meteorologi yang paling dominan dalam model. Oleh karena itu, kelima model signifikan tersebut layak digunakan pada tahap analisis selanjutnya.

3.5 Diagnostik Model

Setelah diperoleh model SARIMA dan ARIMAX yang signifikan, dilakukan uji diagnostik residual menggunakan uji Ljung-Box untuk mengetahui apakah residual bersifat white noise.

Hipotesis pengujian Ljung-Box :

H_0: Residual bersifat white noise

H_1: Residual tidak bersifat white noise

Taraf signifikan : $\alpha = 0.05$

Kriteria keputusan : Tolak H_0 jika $Q^* > \chi^2(\alpha, df)$ atau p-value $< \alpha$, dimana $df = k -$

p, dengan k =banyaknya lag yang diuji dan p =banyaknya parameter.

1. Model SARIMA

Diagnostik model dilakukan untuk memastikan model SARIMA layak digunakan dalam peramalan dengan memeriksa apakah residual bersifat acak (*white noise*) dan tidak mengandung autokorelasi. Model yang baik tidak menyisakan pola pada residual.

Uji Ljung-Box dilakukan pada lag 12, 24, 36, dan 48 untuk mengetahui apakah residual model bersifat *white-noise*. Model memenuhi asumsi *white-noise* apabila seluruh *p-value* > 0,05. Model SARIMA (1,0,0)(1,0,0), SARIMA (2,0,1)(0,0,1), dan SARIMA (1,0,0)(2,0,1) tidak memenuhi asumsi karena pada lag 48 masing-masing memiliki *p-value* 0,008601; 0,02064; dan 0,04924 (<0,05). Model SARIMA (1,0,0)(0,0,1) dan SARIMA (0,0,1)(0,0,1) juga tidak *white-noise* karena seluruh *p-value* <0,05. Sementara itu, model SARIMA (1,0,0)(2,0,0) dan SARIMA (1,0,0)(1,0,1) memiliki seluruh *p-value* >0,05 sehingga residual bersifat *white-noise* dan model layak digunakan. Oleh karena itu, kedua model tersebut dipertimbangkan sebagai model terbaik pada tahap selanjutnya.

2. Model ARIMAX

Uji diagnostik dilakukan untuk memastikan model ARIMAX layak digunakan dalam peramalan dengan memeriksa apakah residual bersifat acak, tidak mengandung autokorelasi, dan variabel eksternal memberikan pengaruh yang baik terhadap model.

Uji Ljung-Box pada model ARIMAX dilakukan pada lag 12, 24, 36, dan 48 untuk mengetahui apakah residual bersifat *white noise*. Pemilihan lag tersebut didasarkan pada data bulanan, di mana lag 12 merepresentasikan satu tahun hingga lag 48 merepresentasikan empat tahun. Model ARIMAX(1,0,0), ARIMAX (2,0,0), ARIMAX (1,0,1), ARIMAX (1,0,2), dan ARIMAX(2,0,2) memiliki seluruh *p-value* > 0.05, seperti ARIMAX(1,0,0) dengan *p-value* 0,196; 0,4801; 0,6922; dan 0,5186, sehingga residual tidak mengandung autokorelasi dan model layak digunakan. Sementara itu, ARIMAX (0,0,1) memiliki *p-value* $5{,}548\times10^{-7}$;0,0001485; 0,00133; dan 0,001464 (<0,05), sehingga residual masih mengandung autokorelasi dan model belum layak digunakan. Dengan demikian, sebagian besar model ARIMAX telah memenuhi asumsi residual *white noise* dan layak digunakan untuk peramalan.

3.6 Pemilihan Model Terbaik

Pemilihan model terbaik dilakukan dengan membandingkan hasil nila MAPE.

1. Model SARIMA

Pemilihan model terbaik dilakukan dengan membandingkan hasil nila MAPE (Mean Absolute Percentage Error).

Tabel 3 Nilai MAPE Model SARIMA

Model SARIMA	MAPE
SARIMA(1,0,0)(2,0,0)	82.50%
SARIMA(1,0,0)(1,0,1)	81.90%

Sumber: data diolah (2026)

Berdasarkan Tabel 3, dilakukan perbandingan akurasi antara model SARIMA (1,0,0)(2,0,0) dan SARIMA(1,0,0)(1,0,1) yang telah memenuhi asumsi residual *white-noise*. Model SARIMA(1,0,0)(2,0,0) menghasilkan nilai MAPE sebesar 82,50%, sedangkan SARIMA(1,0,0)(1,0,1) sebesar 81,90%. Karena memiliki nilai MAPE lebih kecil, model SARIMA(1,0,0)(1,0,1) dinilai lebih baik dalam melakukan peramalan.

2. Model ARIMAX

Pemilihan model terbaik dilakukan dengan membandingkan hasil nila MAPE (Mean Absolute Percentage Error).

Tabel 4 Nilai MAPE Model ARIMAX

Model ARIMAX	MAPE
ARIMAX(1,0,0)	64.91%
ARIMAX(2,0,0)	57.11%
ARIMAX(1,0,1)	77.68%
ARIMAX(1,0,2)	83.04%

Sumber: data diolah (2026)

Berdasarkan Tabel 4, setiap model ARIMAX menghasilkan tingkat akurasi peramalan yang berbeda pada periode Januari 2023–Desember 2024. Model ARIMAX(2,0,0) memiliki nilai MAPE paling rendah yaitu 57,11%, sehingga menjadi model terbaik karena hasil prediksinya paling mendekati data aktual. Model ARIMAX(1,0,0) memiliki MAPE sebesar 64,91%, sedangkan ARIMAX (1,0,1) dan ARIMAX(1,0,2) memiliki MAPE lebih tinggi yaitu 77,68% dan 83,04%. Hal ini menunjukkan bahwa semakin kecil nilai MAPE,

maka semakin baik kemampuan model dalam melakukan peramalan.

3.7 Perbandingan Model SARIMA dan ARIMAX

Dari hasil analisis, model terbaik yang diperoleh adalah SARIMA(1,0,0)(1,0,1) dan ARIMAX (2,0,0). Kedua model dapat digunakan untuk meramalkan kasus DBD di Provinsi Nusa Tenggara Timur, namun model dengan nilai MSE, MAPE, dan RMSE paling kecil dipilih karena memberikan tingkat kesalahan peramalan yang lebih rendah.

Tabel 5 Tabel Perbandingan Model

Model ARIMAX	MAPE	MSE	RMSE
SARIMA (1,0,0)(1,0,1)	81.90%	382.42	195.31
ARIMAX(2,0,0)	57.11%	259.6	161.14

Sumber: data diolah (2026)

Berdasarkan Tabel 5, dapat dilihat bahwa model yang memiliki tingkat kesalahan paling kecil model ARIMAX(2,0,0) dengan nilai MAPE 57.11%, MSE 259.6 dan RMSE 161.14.

3.8 Peramalan Model ARIMAX

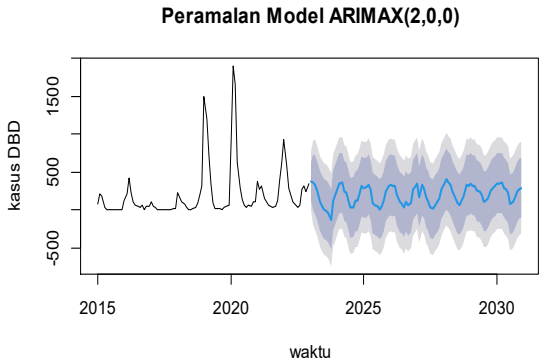
Hasil peramalan faktor iklim yang mempengaruhi kejadian DBD di Provinsi NTT pada periode Januari 2025 −Desember 2026 dengan menggunakan model ARIMAX(2,0,0) dapat dilihat pada tabel berikut.

Tabel 6 Hasil Peramalan Model ARIMAX(2,0,0)

Periode	Nilai Peramalan
Januari 2025	294.2912
Februari 2025	299.9201
Maret 2025	335.9456
.	.
:	:
:	:
November 2030	275.5648
Desember 2030	295.4798

Sumber: data diolah (2026)

Pada Tabel 6, hasil peramalan kasus DBD periode 2025–2030 menunjukkan pola fluktuasi musiman yang berulang. Tahun 2025 diperkirakan mencapai puncak pada Maret sebesar 335,9456 kasus dan terendah pada Agustus sebesar 11,3877 kasus. Puncak tertinggi selama periode peramalan terjadi pada Februari 2028 sebesar 408,0929 kasus. Sementara itu, tahun 2029–2030 menunjukkan pola yang lebih stabil dengan kasus tertinggi masing-masing sebesar 350,0771 kasus pada Januari 2029 dan 356,8511 kasus pada Maret 2030. Secara umum, kasus DBD cenderung meningkat pada awal dan akhir tahun serta menurun pada pertengahan tahun, sehingga menunjukkan adanya pola musiman.



Gambar 3. Plot Peramalan Kasus DBD dengan model ARIMAX(2,0,0)

Pada Gambar 3, hasil peramalan menggunakan model ARIMAX(2,0,0) untuk periode 2023–2030 menunjukkan bahwa jumlah kasus DBD tetap mengalami fluktuasi dari tahun ke tahun. Garis hitam menunjukkan data historis dengan beberapa lonjakan besar, terutama sekitar tahun 2019–2020, sedangkan garis biru menunjukkan hasil peramalan yang masih memperlihatkan pola musiman. Secara keseluruhan, model ARIMAX mampu mengikuti pola historis kasus DBD dan menunjukkan bahwa hingga tahun 2030 kasus diperkirakan tetap berfluktuasi tanpa tren peningkatan atau penurunan yang drastis, dengan pengaruh faktor iklim yang masih berperan terhadap perubahan jumlah kasus DBD.

4. KESIMPULAN DAN SARAN

Kesimpulan tentang peningkatan ketepatan peramalan berlaku berdasarkan hasil pengujian pada data penelitian, bukan berarti semua faktor iklim terbukti berpengaruh signifikan secara statistik.Hasil tersebut menunjukkan bahwa penambahan variabel eksternal berupa suhu, curah hujan, kelembapan, dan kecepatan angin dapat meningkatkan ketepatan peramalan jumlah kasus DBD. Meskipun demikian, tidak semua variabel iklim menunjukkan pengaruh yang signifikan pada setiap model ARIMAX yang diestimasi. Dengan demikian, model ARIMAX terpilih sebagai model terbaik dalam

penelitian ini karena menghasilkan tingkat kesalahan peramalan yang lebih rendah dibandingkan model SARIMA

Persamaan umum model ARIMAX(2,0,0) yang diperoleh adalah:

$$X_t = 0.522946X_{t-1} + 0.278096X_{t-2} - 67.266424X_{t-3} + 1.077526X_{t-4} + 0.133300Y_{t-1} + 15.423433Y_{t-2} + et - 0.138209et - 1$$

Penelitian selanjutnya disarankan untuk menambahkan variabel lain yang dapat memengaruhinya agar hasil peramalan semakin akurat. Selain itu, dapat dilakukan perbandingan dengan metode lain untuk melihat konsistensi hasil.

DAFTAR PUSTAKA

Azzulfa, A. N., & Primandari, A. H. (2026). Penerapan Metode Seasonal ARIMA dalam Peramalan Konsumsi Listrik di PT PLN (Persero) Area Magelang. *Emerging Statistics and Data Science Journal*, 4(1), 819-830.

Bone, T., & Kaunang, W. P. J. (2021). Hubungan Antara Curah Hujan, Suhu Udara Dan Kelembaban Dengan Kejadian Demam Berdarah Dengue Di Kota Manado Tahun 2015-2020. *Jurnal KESMAS*, 10(5), 36-45.

BPS, 2025. *Statistik Daerah Kota Kupang 2025 - Badan Pusat Statistik Kota Kupang*.

Buntara, M. S., Napitupulu, H., & Gusriani, N. (2023). Pemrograman Python Untuk Peramalan Data Deret Waktu Menggunakan Metode Seasonal Autoregressive Integrated Moving Average (SARIMA). *In Search*, 22(2), 354–362. https://doi.org/10.37278/insearch.v22i2.774

Hasanah, U. (2022). Analisis model Autoregressive Integrated Moving Average (ARIMA) Box–Jenkins pada penderita diabetes. Skripsi. Malang: Universitas Islam Negeri Maulana Malik Ibrahim Malang.

Hasby Kuswanto, Pradita Eko Prasetyo Utomo, Ulfa Khaira, & Akhiyar Waladi. (2025). Prediksi Nilai Ekspor Migas Indonesia Menggunakan Metode SARIMA dan LSTM. *SATESI: Jurnal Sains Teknologi dan Sistem Informasi*, 5(1), 69–79. https://doi.org/10.54259/satesi.v5i1.4103

Junaedi, L., Damastuti, N., & Widodo, A. (2025). Penerapan Metode Seasonal ARIMA (SARIMA) untuk Peramalan Penjualan Barang dengan Pola Musiman Tahunan. *JISEM Jurnal Program Studi Informatika Universitas Katolik Widya Mandala Surabaya*, 01(01), 38-48.

Katabba, Y. I. (2021). Metode Seasonal Autoregressive Integrated Moving Average (SARIMA) untuk Memprediksi Jumlah Penumpang Kereta Api di Pulau Sumatera. Skripsi. Jambi: Universitas Jambi. Skripsi. Jambi: Universitas Jambi.

Khoiri, H. A. (2023). *Analisis deret waktu univariat: Teori dan pengolahan data.* Madiun Jawa Timur: UNIPMA PRESS.

Nanda Tria Lestari, & Witanti, A. (2023). Analisis Prediksi Kasus DBD Berdasarkan Faktor Cuaca Dengan Multivariat ARIMA. *PETIR*, 16(2), 228–236. https://doi.org/10.33322/petir.v16i2.2117

Pradana, D. A. P., Mahananto, F., & Djunaidy, A. (2022). Sistem Peramalan Menggunakan Autoregressive Integrated Moving Average with Exogenous Variables (ARIMAX) Untuk Harga Minyak Sawit Indonesia. *Jurnal Teknik ITS*, 11(2), A97–A102. https://doi.org/10.12962/j23373539.v11i2.86373

Riestiansyah, F., Damayanti, D., Reswara, M., & Susetyoko, R. (2022). Perbandingan metode ARIMA dan ARIMAX dalam Memprediksi Jumlah Wisatawan Nusantara di Pulau Bali. *Jurnal Infomedia,* 7(2), 58-62. https://doi.org/10.30811/jim.v7i2.3336

World Health Organization., 2024. *Dengue*.

Zaki, A., Wahyuni, M. S., Irwan, I., & Rahman, A. (2023). Peramalan Jumlah Penderita Demam Berdarah Dengue Menggunakan Metode Seasonal- ARIMA. *ARRUS Journal of Mathematics and Applied Science*, 3(2), 65–74. https://doi.org/10.35877/mathscience2143