

Comparative Analysis of ARIMA and LSTM Methods for Sea Surface Temperature Forecasting in the Sunda Strait

Analisis Perbandingan Metode ARIMA dan LSTM untuk Peramalan Suhu Permukaan Laut di Selat Sunda

Raihan Sabillah Adisecha¹, Nabil Ibni Nawawi², Naiya Dzil Izzati³, Muhammad Tsaqif Najmuddin⁴, Nawwaf Ariq Rabbani⁵, Yenni Angraini^{*6}, Adelia Putri Pangestika⁷

^{1,2,3,4,5,6,7} Statistics and Data Science Study Program, IPB University, Indonesia

Email: ¹raihanadisecha@apps.ipb.ac.id, ²nabilibninawawi@apps.ipb.ac.id, ³naiyadzilizzati@apps.ipb.ac.id, ⁴mtsaqifnajmuddin@apps.ipb.ac.id, ⁵nawwafariq@apps.ipb.ac.id, ⁶y_angraini@apps.ipb.ac.id, ⁷adelia_pangestika@apps.ipb.ac.id

^{*}Corresponding Author

Received: 18 December 2024, revised: 18 April 2025, accepted: 26 April 2025

Abstract

The Sunda Strait is an important area for Indonesia because it is the main domestic and international transportation route. As a water area, the Sunda Strait has weather conditions that are greatly influenced by sea surface temperature (SST). Crucial SST forecasting is carried out to assist maritime transportation activities. This study aims to compare the performance of the ARIMA and LSTM methods in forecasting SST in the Sunda Strait. The data used in this study are daily SST data for the Sunda Strait from August 20, 2022, to January 1, 2024. The best ARIMA model obtained in this data modeling is ARIMA(1,1,1), where this model has significant overall parameters, the smallest AIC and BIC values, and model diagnostic results that meet the assumptions. Meanwhile, in LSTM modeling, the best combination of hyperparameters obtained is a neuron of 150, an epoch of 150, and a batch size of 32, where this combination produces the lowest MSE value of 0.003799. A comparison of the performance of the ARIMA and LSTM methods is carried out by considering the MAPE values of the training data and test data. The LSTM method is superior to the ARIMA method, with a MAPE value of 0.512% for training data and 0.564 for testing data. The forecasting results using the LSTM method show a similar pattern in both training and testing data. Meanwhile, the forecasting results of the LSTM method for the following 30 periods show a fluctuating pattern throughout the day.

Keywords: ARIMA, LSTM, MAPE, Sunda Strait, temperature

Abstrak

Selat Sunda merupakan area penting untuk Indonesia karena selat ini menjadi jalur utama transportasi domestik dan internasional. Selat Sunda sebagai wilayah perairan memiliki kondisi cuaca yang sangat dipengaruhi oleh suhu permukaan laut (SPL). Peramalan SPL krusial dilakukan



untuk membantu aktivitas transportasi laut. Penelitian ini bertujuan membandingkan performa metode ARIMA dan LSTM dalam melakukan peramalan SPL di Selat Sunda. Data yang digunakan dalam penelitian ini adalah data harian SPL Selat Sunda periode 20 Agustus 2022 hingga 1 Januari 2024. Model ARIMA terbaik yang didapatkan dalam pemodelan data ini adalah ARIMA(1,1,1) di mana model ini memiliki keseluruhan parameter yang signifikan, nilai AIC dan BIC terkecil, serta hasil diagnostik model yang memenuhi asumsi. Sementara itu, dalam pemodelan LSTM, kombinasi *hyperparameter* terbaik yang didapatkan adalah *neuron* sebesar 150, *epoch* sebesar 150, dan *batch size* sebesar 32 dimana kombinasi ini menghasilkan nilai MSE terendah yaitu sebesar 0,003799. Perbandingan performa metode ARIMA dan LSTM dilakukan dengan memperhatikan nilai MAPE data latih dan data uji. Metode LSTM lebih unggul dibandingkan metode ARIMA dengan nilai MAPE data latih sebesar 0,512% dan data uji sebesar 0,564%. Hasil peramalan menggunakan metode LSTM menunjukkan pola yang serupa baik pada data latih maupun data uji. Sementara itu, hasil peramalan metode LSTM pada 30 periode ke depan menunjukkan pola yang fluktuatif di sepanjang hari.

Kata Kunci: ARIMA, LSTM, MAPE, Selat Sunda, suhu.

1. PENDAHULUAN

Indonesia merupakan negara maritim di mana 62% wilayahnya merupakan lautan. Luas laut di Indonesia mencapai sekitar 5,8 juta km². Kondisi ini menyebabkan cuaca di Indonesia sangat dipengaruhi oleh dinamika laut di sekitarnya, khususnya oleh iklim laut. Salah satu faktor yang mempengaruhi iklim laut adalah suhu permukaan laut (SPL). SPL yang tinggi membuat penguapan air laut meningkat sehingga konsentrasi uap air di atmosfer menjadi lebih tinggi [4]. Hal ini berkontribusi terhadap proses pembentukan awan dan peningkatan curah hujan. Sebaliknya, SPL yang rendah akan menurunkan tingkat penguapan. Penurunan tingkat penguapan akan menghasilkan atmosfer yang lebih kering sehingga cuaca menjadi relatif lebih panas.

Selain berpengaruh terhadap cuaca di sekitarnya, SPL juga berpengaruh terhadap aktivitas transportasi laut. Perubahan SPL dapat memicu terbentuknya pola cuaca ekstrem seperti hujan lebat, badai, dan gelombang tinggi yang berisiko mengganggu keselamatan pelayaran [19]. SPL juga mempengaruhi pola arus laut yang dapat memperlambat atau mempercepat laju kapal. Dalam kondisi tertentu, SPL memicu pembentukan kabut yang dapat menurunkan jarak pandang sehingga meningkatkan risiko kecelakaan. Selain berpengaruh terhadap cuaca di sekitarnya, SPL juga berpengaruh terhadap aktivitas transportasi laut. Perubahan SPL dapat memicu terbentuknya pola cuaca ekstrem seperti hujan lebat, badai, dan gelombang tinggi yang berisiko mengganggu keselamatan pelayaran [19]. SPL juga mempengaruhi pola arus laut yang dapat memperlambat atau mempercepat laju kapal. Dalam kondisi tertentu, SPL memicu pembentukan kabut yang dapat menurunkan jarak pandang sehingga meningkatkan risiko kecelakaan.

Dinamika SPL akan sangat berdampak pada wilayah dengan aktivitas transportasi laut yang tinggi seperti Selat Sunda. Selat Sunda merupakan selat yang menghubungkan dua wilayah dengan kepadatan penduduk dan aktivitas ekonomi yang tinggi, yaitu Pulau Jawa dan Sumatera. Hal ini membuat Selat Sunda menjadi jalur utama transportasi domestik, terutama pada lintasan penyeberangan Merak-Bakauheni. Selain itu, Selat Sunda juga termasuk salah satu dari empat alur laut kepulauan Indonesia yang ditetapkan sebagai jalur pelayaran internasional. Hal ini membuat Selat Sunda dilalui oleh 40% kapal asing yang menjalankan rute perdagangan global [18]. Tingginya aktivitas transportasi laut baik domestik maupun internasional di Selat Sunda membuat pemantauan SPL di Selat Sunda menjadi krusial untuk dilakukan. Selain pemantauan, peramalan SPL juga penting dilakukan sebagai langkah antisipatif dalam menghadapi potensi gangguan yang dapat menghambat operasional pelayaran. Salah satu metode peramalan yang dapat digunakan adalah metode *autoregressive integrated moving average* (ARIMA).

Metode ARIMA merupakan metode yang dikenal kuat dan fleksibel dalam peramalan data deret waktu. Metode ini sebelumnya digunakan oleh [17] dalam peramalan data jumlah penumpang pelabuhan bakauheni selama periode tsunami Selat Sunda. Metode ini terbukti memiliki performa peramalan yang baik dengan nilai *mean absolute percentage error* (MAPE) sebesar 9,55%. Metode lain yang dapat digunakan adalah metode berbasis pembelajaran mesin seperti *long short-term memory* (LSTM). Metode LSTM memiliki keunggulan dalam menangani ketergantungan jangka panjang pada data deret waktu. Metode ini digunakan dalam penelitian [10] untuk meramalkan data tinggi permukaan laut di Jakarta. Hasilnya, metode ini memiliki performa peramalan yang baik dengan nilai MAPE sebesar 9,04%. Performa peramalan yang baik pada metode ARIMA dan LSTM menunjukkan potensi kedua metode dalam peramalan data SPL di Selat Sunda. Oleh karena itu, penelitian ini bertujuan untuk membandingkan performa metode ARIMA dan LSTM dalam melakukan peramalan pada data SPL di Selat Sunda.

2. METODE PENELITIAN

2.1 Data

Data yang digunakan dalam penelitian ini adalah data SPL di Selat Sunda (garis lintang: -5.900 dan garis bujur: 105.600) yang dikumpulkan setiap hari dalam rentang waktu 20 Agustus 2022 hingga 1 Januari 2024. Data tersebut diambil dari sumber JAXA Earth (<https://earth.jaxa.jp/en/>) [11] yang menyediakan data satelit terkait SPL secara *real-time*. Penggunaan data ini bertujuan untuk memodelkan dan menganalisis fluktuasi SPL di kawasan Selat Sunda dalam periode 500 hari.

2.2 ARIMA

Dua komponen utama dari ARIMA, yaitu *Autoregressive* (AR) dan *Moving Average* (MA) memiliki fungsinya masing-masing. AR mengukur hubungan antara nilai pada waktu Y_t dengan nilai-nilai sebelumnya dalam *time series*, yaitu $Y_{t-1}, Y_{t-2}, \dots, Y_{t-p}$ dengan p merupakan nilai *lag*. Ini berarti, nilai saat ini dipengaruhi oleh nilai-nilai masa lalu yang disesuaikan dengan koefisien ϕ_i , yang menunjukkan seberapa besar pengaruh masing-masing nilai *lag* terhadap nilai saat ini. Lalu, MA digunakan untuk menangani kesalahan atau sisaan yang terjadi dalam prediksi model. Kesalahan ini diukur pada periode-periode sebelumnya dan digunakan untuk memperbaiki peramalan nilai di periode berikutnya. Parameter q menunjukkan jumlah *lag* kesalahan yang digunakan untuk mengoreksi prediksi.

Model ARIMA dinotasikan sebagai $ARIMA(p, d, q)$, dengan parameter:

p = Ordo dari komponen AR.

d = Ordo dari *differencing* (I) yang diperlukan untuk mencapai kondisi stasioner.

q = Ordo dari komponen MA.

Secara matematis, model ARIMA ditunjukkan oleh Persamaan 2.1.

$$Y_t = \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + e_t - \theta_1 e_{t-1} - \dots - \theta_q e_{t-q} \quad (2.1)$$

Keterangan:

Y_t = Nilai *time series* pada waktu t .

ϕ_p = Koefisien *autoregressive* pada *lag-p*.

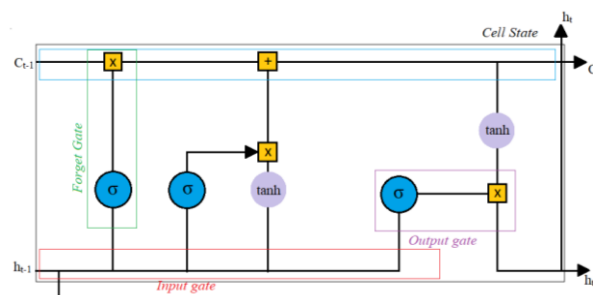
e_t = Galat sisaan (*white noise*) pada waktu t .

θ_q = Koefisien *moving average* pada *lag-q*.

Persamaan (2.1) menggambarkan model ARIMA sebagai kombinasi dari komponen AR dan MA, di mana nilai saat ini Y_t dipengaruhi oleh nilai-nilai *lag* sebelumnya dan kesalahan (sisaan) pada periode sebelumnya.

2.3 Long Short-Term Memory (LSTM)

Long Short-Term Memory (LSTM) adalah salah satu tipe dari *Recurrent Neural Network* (RNN) yang dapat mengatasi masalah *vanishing gradient* dan *exploding gradient* pada model tradisional RNN [3]. Model LSTM memiliki struktur yang disebut dengan memori sel. Dalam memori sel terdapat *gate* yang membangun memori sel tersebut. Memori sel terdiri dari beberapa *gate*, yaitu *input gate*, *output gate* dan *forget gate* yang memiliki fungsi untuk mengingat informasi sesuai dengan interval waktu yang berubah-ubah [1]. Secara umum fungsi dari *gate* tersebut memiliki peranan yang berbeda. *Input gate* berfungsi untuk mengontrol jumlah *input* yang akan memengaruhi memori. *Forget gate* berfungsi untuk mengontrol jumlah memori lama yang harus dihapus. *Output gate* berfungsi untuk mengontrol jumlah memori yang tersimpan di keadaan tersembunyi [22].



Gambar 2.1 Contoh Arsitektur LSTM

Keterangan:

x_t = nilai *input* waktu- t

h_t = nilai *output* waktu- t

c_{t-1}, c_t = *cell state* dengan waktu $(t - 1)$ dan t

Fungsi $\tanh$ dan σ secara berurutan ditunjukkan oleh Persamaan 2.2 dan 2.3.

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (2.2)$$

$$\sigma(x) = \frac{1}{1 + e^{-x}} \quad (2.3)$$

2.4 Tahapan Analisis

Tahapan analisis yang dilakukan dalam penelitian ini adalah sebagai berikut.

1. Membagi data dengan proporsi 63% data latih dan 37% data uji berdasarkan pola data. Pembagian ini mempertimbangkan kestabilan pola musiman dan kontinuitas data deret waktu, sehingga masing-masing bagian data dapat merepresentasikan karakteristik data secara menyeluruh. Proporsi ini juga bertujuan agar data uji tetap mencerminkan variabilitas pola data yang diamati dalam data latih.
2. Melakukan eksplorasi data untuk melihat pola dan karakteristik data.
3. Melakukan pemodelan ARIMA pada data latih.
 - a. Melakukan uji stasioneritas data baik dalam rata-rata maupun ragam. Uji stasioneritas data dalam rata-rata dilakukan berdasarkan plot data deret waktu, plot *Autocorrelation Function* (ACF), dan uji *Augmented Dickey-Fuller* (ADF). Apabila terjadi pelanggaran asumsi stasioneritas data dalam rata-rata, dilakukan penanganan berupa *differencing*. Sementara itu, uji stasioneritas data dalam ragam dilakukan berdasarkan nilai lambda pada plot Box-Cox di mana lambda harus bernilai 1 atau nilai 1 berada pada selang

kepercayaan lambda [23]. Apabila terjadi pelanggaran asumsi stasioneritas data dalam ragam, dilakukan penanganan berupa transformasi Box-Cox berdasarkan Persamaan 2.4.

$$y = \frac{(x^\lambda - 1)}{\lambda} \quad (2.4)$$

Di mana y merupakan nilai setelah transformasi, x merupakan nilai sebelum transformasi, dan λ merupakan parameter transformasi

- b. Melakukan model ARIMA berdasarkan plot ACF, *Partial Autocorrelation Function* (PACF), dan *Extended Autocorrelation Function* (EACF). Identifikasi model ARIMA berdasarkan plot ACF dan PACF dilakukan mengikuti Tabel 2.1. Sementara itu, identifikasi model ARIMA berdasarkan plot EACF dilakukan dengan melihat pola berbentuk segitiga nol di mana ujung "0" segitiga lancip merupakan ordo dari model ARIMA. Apabila identifikasi model ARIMA menggunakan plot ACF dan PACF sulit untuk dilakukan, maka identifikasi model ARIMA langsung dilakukan menggunakan plot EACF.

Tabel 2.1 Penentuan model ARIMA berdasarkan ACF/PACF

Model	ACF	PACF
AR(p)	Turun cepat secara eksponensial	<i>Cuts-off</i> setelah lag ke- p
MA(q)	<i>Cuts-off</i> setelah lag ke- q	Turun cepat secara eksponensial
ARMA(p, q)	Turun cepat secara eksponensial	Turun cepat secara eksponensial

- c. Melakukan pendugaan parameter model tentatif dan memilih model terbaik dengan keseluruhan parameter signifikan dan nilai *Akaike's Information Critetion* (AIC) terkecil. Perhitungan nilai AIC dilakukan berdasarkan Persamaan 2.5.

$$AIC = \ln \left(\frac{\sum_{t=1}^n (e_t)^2}{T} \right) + \frac{2p}{T} \quad (2.5)$$

Keterangan:

p = banyaknya parameter dalam model

T = banyaknya data pengamatan

$\sum_{t=1}^n (e_t)^2$ = jumlah kuadrat sisaan

- d. Melakukan uji diagnostik model. Uji diagnostik model dilakukan dengan menguji asumsi kenormalan, kebebasan sisaan, dan kehomogenan ragam sisaan model. Uji kenormalan dilakukan dengan uji Kolmogorov-Smirnov. Sementara itu, uji kebebasan sisaan dan uji kehomogenan ragam sisaan dilakukan dengan uji Ljung-Box. Selain kenormalan, kebebasan sisaan, dan kehomogenan ragam sisaan, dilakukan pula pengujian nilai tengah sisaan dengan menggunakan uji t. Apabila terjadi pelanggaran asumsi pada sisaan, dilakukan penanganan berupa transformasi data.
 - e. Melakukan *overfitting* dengan menambah ordo model tentatif terbaik sebesar satu ordo pada AR atau MA.
 - f. Memilih model terbaik antara model sebelum dan setelah *overfitting* dengan keseluruhan parameter yang signifikan, nilai AIC yang terkecil, dan hasil uji diagnostik model yang memenuhi asumsi.
4. Melakukan pemodelan LSTM pada data latih.

- a. Melakukan normalisasi data menggunakan metode *min-max scaler*. Metode *min-max scaler* mampu memberikan hasil yang lebih baik dibandingkan dengan metode *Standard Scaler* [27] dan *Z-Score* [24]. Metode ini akan mengubah data asli menjadi suatu data baru dengan nilai yang berkisar antara 0 sampai 1 [29]. Perumusan metode *min-max scaler* dilakukan mengikuti Persamaan 2.6.

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (2.6)$$

Keterangan:

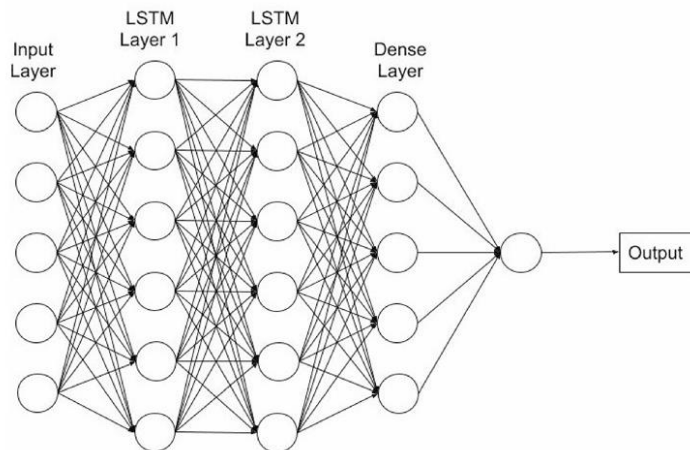
x' = hasil normalisasi

x = data asli

x_{min} = data terkecil dari dataset tersebut

x_{max} = data terbesar dari dataset tersebut

- b. Membentuk arsitektur LSTM dengan 2 *layer* utama dan 1 *dense layer* sebagaimana pada Gambar 2.2. Jumlah *layer* tersebut ditentukan secara subyektif dengan mempertimbangkan penelitian [31] yang menyatakan bahwa jumlah *layer* yang banyak tidak membuat kinerja model lebih baik dan stabil, serta membutuhkan waktu yang lebih lama.



Gambar 2.2 Ilustrasi arsitektur LSTM [7]

- c. Melakukan insialisasi *hyperparameter* dengan mengombinasikan nilai *neuron*, *epoch*, dan *batch size*. Nilai *neuron* dan *epoch* yang akan digunakan adalah 50, 100, dan 150. Sementara itu, nilai *batch size* yang akan digunakan adalah 32, 64, dan 128. *Hyperparameter* lain seperti *time step*, dan *optimizer* ditentukan secara subyektif karena mempertimbangkan waktu komputasi. *Time step* yang digunakan adalah 50, sedangkan *optimizer* yang digunakan adalah adam.
- d. Melakukan pelatihan model dan *fit model*, serta menentukan *hyperparameter* terbaik. Penentuan *hyperparameter* terbaik dilakukan dengan metode *Grid Search Cross Validation* (CV). Metode ini mengombinasikan berbagai nilai parameter yang diinginkan dan menerapkan teknik *Cross Validation* pada setiap kombinasi tersebut. Metrik evaluasi yang digunakan selama proses *Grid Search* adalah *mean squared error* (MSE). Kombinasi parameter terbaik dipilih berdasarkan nilai MSE terkecil dari

seluruh kombinasi yang diuji. Perhitungan nilai MSE dilakukan berdasarkan Persamaan 2.7.

$$MSE = \frac{\sum (y_i - \hat{y}_i)^2}{n} \quad (2.7)$$

Di mana y_i merupakan data aktual, $\hat{y}_i$ merupakan data hasil ramalan, dan n merupakan jumlah amatan. Dalam prosesnya, terdapat penyesuaian metode *Grid Search CV* yang dilakukan pada penelitian ini, yaitu pada parameter *Cross Validation*. Penelitian ini menggunakan *TimeSeriesSplit* dari pustaka *scikit-learn* sebagai metode *Cross Validation* dengan jumlah *fold* sebanyak 5. Metode ini memungkinkan segmentasi data yang sesuai untuk data deret waktu.

5. Melakukan peramalan pada data uji dan menentukan metode peramalan terbaik dengan nilai MAPE terkecil. Perhitungan nilai MAPE dilakukan berdasarkan Persamaan 2.8.

$$MAPE = \sum \left| \frac{y_i - \hat{y}_i}{\hat{y}_i} \right| \times 100\% \quad (2.8)$$

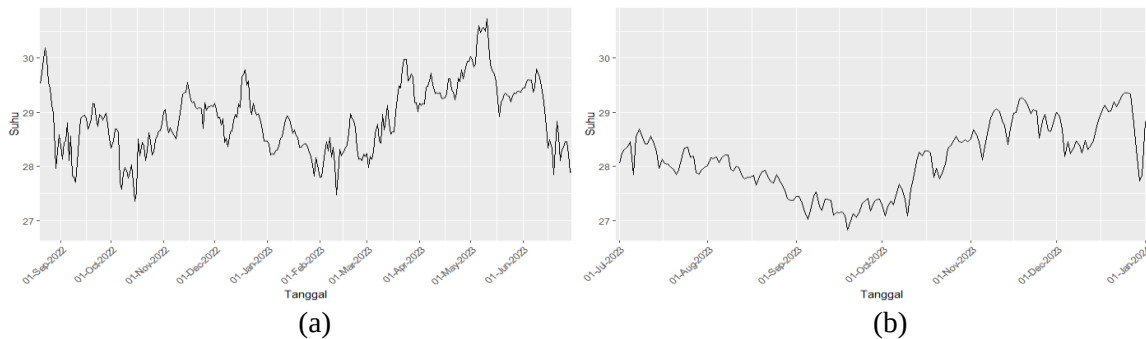
Di mana y_i merupakan data aktual dan $\hat{y}_i$ merupakan data hasil ramalan. Selain menggunakan nilai MAPE, penentuan metode terbaik juga didasarkan pada kemampuan metode dalam menggambarkan pola pada hasil *fit* data latih dengan data latih aktual dan hasil prediksi data uji dengan data uji aktual. Kesesuaian pola akan mengindikasikan bahwa metode tersebut mampu mempelajari pola data aktual dengan baik sehingga layak digunakan untuk peramalan data.

6. Melakukan peramalan sebanyak 30 periode ke depan dengan kedua metode.

3. HASIL DAN PEMBAHASAN

3.1 Pembagian Data dan Eksplorasi Data

Data pada penelitian ini dibagi dengan proporsi 63% data latih dan 37% data uji sebagaimana plot data deret waktu yang tertera pada Gambar 3.1. Plot data latih pada Gambar 3.1(a) dan plot data uji pada Gambar 3.1 (b) menggambarkan proporsi pembagian data yang baik karena kedua plot berhasil menangkap fluktuasi data yang serupa. Baik pada data latih maupun data uji, data SPL Selat Sunda cenderung menunjukkan volatilitas yang tinggi di mana data ini memiliki tren yang naik dan turun di sepanjang tahun.



Gambar 3.1 (a) Plot Data Latih (b) Plot Data Uji

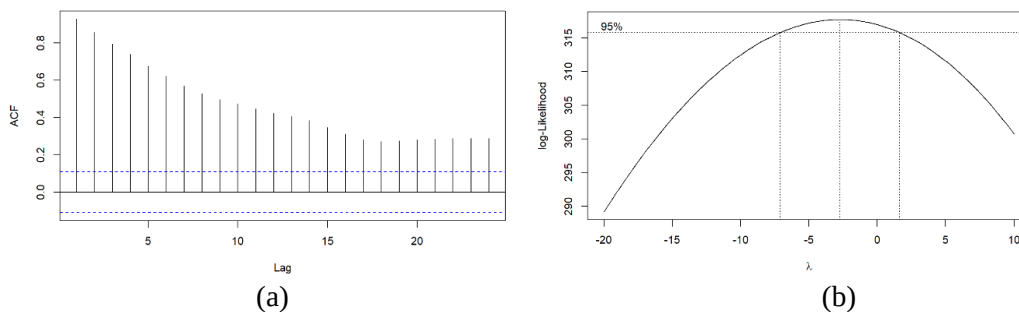
3.2 ARIMA

3.2.1 Kestasioneran Data

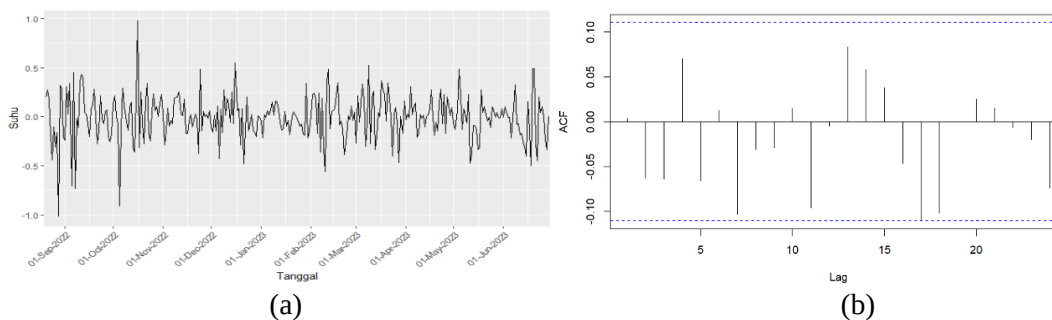
Uji kestasioneran data dalam rata-rata dilakukan secara eksploratif dengan melihat plot data deret waktu yang tertera pada Gambar 3.1(a) dan plot ACF yang tertera pada Gambar 3.2(a). Plot data

deret waktu menunjukkan adanya pola trend naik dan turun yang mengindikasikan bahwa data tidak stasioner dalam rata-rata [25]. Hasil ini diperkuat dengan pola plot ACF yang menurun secara perlahan (*tails off slowly*). Selain itu, uji kestasioneran data menggunakan uji ADF juga memperkuat kesimpulan tersebut karena memiliki p -value sebesar 0,052 yang lebih besar dari taraf signifikansi 5%. Sementara itu, uji kestasioneran data dalam ragam dilakukan berdasarkan nilai lambda pada plot Box-Cox yang tertera pada Gambar 3.2(b). Plot Box-Cox menunjukkan nilai lambda sebesar -2,727, dengan selang kepercayaan 95% yang memuat nilai 1. Hal ini menunjukkan bahwa data telah stasioner dalam ragam.

Ketidakstasioneran data harus ditangani karena pemodelan AR, MA, dan ARIMA mensyaratkan data yang stasioner. Ketidakstasioneran data dalam rata-rata selanjutnya ditangani dengan melakukan *differencing* sebanyak satu kali. Setelah dilakukan *differencing* sebanyak satu kali, terlihat bahwa data telah stasioner dalam rata-rata ditandai dengan plot data deret waktu pada Gambar 3.3(a) yang cenderung bergerak di suatu nilai tengah tertentu. Hasil ini didukung oleh plot ACF pada Gambar 3.3(b) dengan nilai ACF yang tidak signifikan pada setiap *lag*. Kestasioneran data dalam rata-rata juga diperkuat menggunakan uji ADF dengan p -value sebesar 0,01 yang kurang dari taraf nyata 5%.



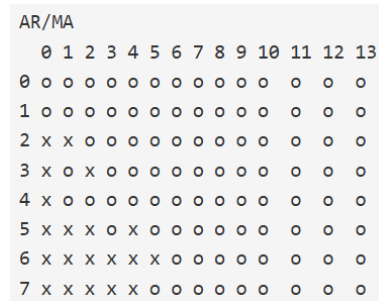
Gambar 3.2 (a) Plot ACF Data Latih (b) Plot Box-Cox Data Latih



Gambar 3.3 (a) Plot Data Latih Differencing 1 (b) Plot ACF Data Latih Differencing 1

3.2.2 Identifikasi Model dan Pendugaan Parameter

Identifikasi model menggunakan plot ACF dan PACF sulit dilakukan dalam pemodelan ARIMA [6], sehingga identifikasi model dilakukan berdasarkan plot EACF yang tertera pada Gambar 3.4. Model tentatif yang diperoleh berdasarkan hasil identifikasi model menggunakan plot EACF adalah ARIMA(0,1,1), ARIMA(1,1,1), ARIMA(2,1,2), ARIMA(3,1,3), dan ARIMA(4,1,2). Pendugaan parameter model tentatif tersebut selanjutnya ditunjukkan pada Tabel 3.1.



Gambar 3.4 Plot EACF Data Latih Differencing 1

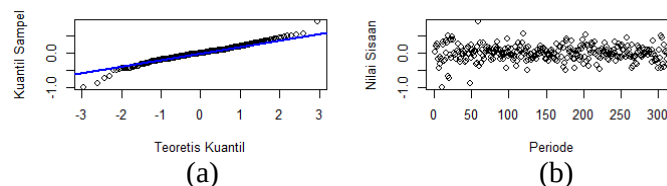
Pendugaan parameter pada Tabel 3.1 menunjukkan nilai AIC terkecil dimiliki oleh model ARIMA(2,1,2) dengan semua parameter model yang signifikan. Sementara itu, nilai BIC terkecil dimiliki oleh model ARIMA(0,1,1) dengan adanya parameter model yang tidak signifikan. Berdasarkan kedua ukuran ini, model tentatif yang dipilih adalah model ARIMA(2,1,2) karena model dengan nilai BIC terkecil memiliki parameter model yang tidak signifikan.

Tabel 3.1 Pendugaan Parameter Model Tentatif

Model	Signifikansi Parameter	AIC	BIC
ARIMA(0,1,1)	Semua parameter tidak signifikan	-27,120	-19,620
ARIMA(1,1,1)	Semua parameter tidak signifikan	-25,170	-13,920
ARIMA(2,1,2)	Semua parameter signifikan	-30,820	-12,080
ARIMA(3,1,3)	Semua parameter signifikan	-29,490	-3,240
ARIMA(4,1,2)	Semua parameter tidak signifikan	-24,050	2,190

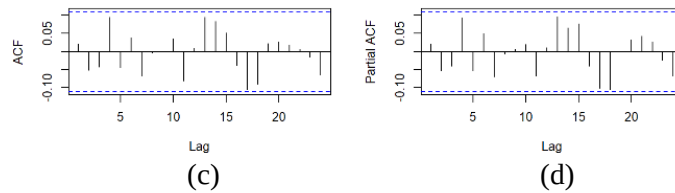
3.2.3 Uji Diagnostik Model

Uji diagnostik model dilakukan pada sisaan model ARIMA(2,1,2) baik secara eksploratif maupun dengan uji formal. Secara eksploratif, Gambar 3.5 (a) menunjukkan plot kuantil-kuantil yang cenderung tidak normal karena titik-titik yang terbentuk tidak mengikuti garis 45°. Selanjutnya, Gambar 3.5 (b) menunjukkan lebar pita sisaan yang cenderung tidak sama. Hal ini mengindikasikan bahwa ragam sisaan tidak homogen. Sementara itu, Gambar 3.5 (c) dan (d) memperlihatkan bahwa plot ACF dan PACF tidak menunjukkan adanya autokorelasi yang signifikan pada *lag-lag* awal yang mengindikasikan bahwa sisaan saling bebas. Hasil ini sejalan uji diagnostik model menggunakan uji formal yang tertera pada Tabel 3.2. Tabel 3.2 menunjukkan bahwa model ARIMA(2,1,2) yang didapat belum memenuhi asumsi normalitas sisaan dan kehomogenan ragam. Tidak terpenuhinya asumsi tersebut dalam akan menimbulkan masalah, yaitu ketidaktepatan estimasi parameter, kesalahan dalam uji hipotesis, peramalan yang kurang akurat, dan masalah pada validitas model [20].



JURNAL MATEMATIKA, STATISTIKA DAN KOMPUTASI

Raihan Sabillah Adisecha, Nabil Ibni Nawawi, Naiya Dzil Izzati, Muhammad Tsaqif Najmuddin,
Nawwaf Ariq Rabbani Yenni Angraini, Adelia Putri Pangestika



Gambar 3.5 (a) Plot kuantil-kuantil normal (b) Plot *fitted value vs residual* (c) Plot ACF (d) Plot PACF

Tabel 3.2 Hasil Uji Formal

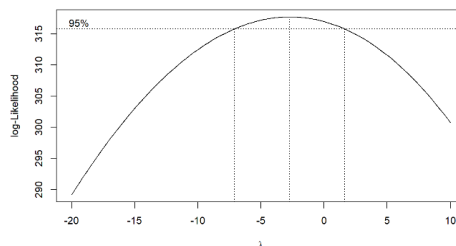
Asumsi sisaan	H_0	p -value	Kesimpulan
Normalitas sisaan	Sisaan menyebar normal	2.200×10^{-16}	Tolak H_0
Sisaan saling bebas	Sisaan saling bebas	0,966	Tak tolak H_0
Ragam sisaan homogen	Ragam sisaan homogen	0,021	Tolak H_0
Nilai tengah sisaan sama dengan nol	Nilai tengah sisaan sama dengan nol	0,690	Tak tolak H_0

3.2.4 Penanganan Tidak Terpenuhinya Asumsi Sisaan

Model ARIMA(2,1,2) yang tidak memenuhi asumsi normalitas sisaan dan kehomogenan ragam ditangani dengan melakukan transformasi data. Menurut [21], transformasi data merupakan penanganan yang lebih efektif digunakan ketika dijumpai pelanggaran asumsi normalitas dan kehomogenan ragam sisaan secara bersamaan. Dalam hal ini, transformasi data dilakukan menggunakan nilai lambda -3 karena nilai lambda -3 masih termasuk ke dalam rentang selang kepercayaan 95% pada plot Box-Cox yang tertera pada Gambar 3.6. Nilai lambda sebesar -3 membuat transformasi Box-Cox dilakukan dengan pangkat -3 sebagaimana yang tertera pada Persamaan 3.1.

$$y^* = \frac{1}{y^3} \quad (3.1)$$

Di mana y^* adalah data yang telah ditransformasikan dan y adalah data asli. Setelah data ditransformasi, hasilnya digunakan untuk memodelkan ARIMA dengan tahapan analisis yang sama dengan data aslinya.



Gambar 3.6 Plot Box-Cox Data Transformasi

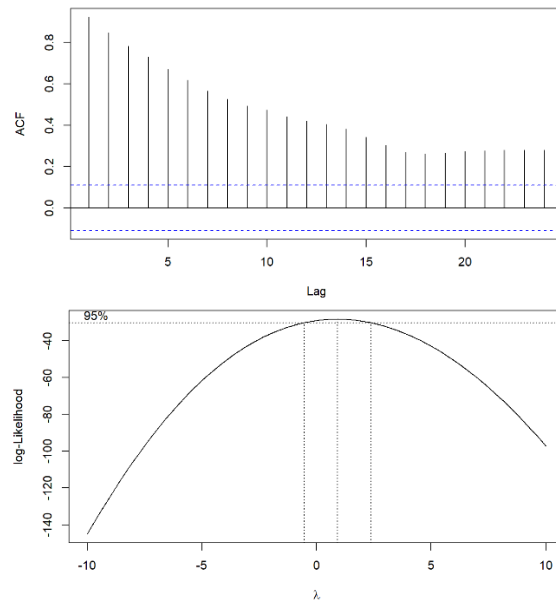
3.2.5 Kestasioneran Data Transformasi

Data hasil transformasi perlu dilihat kembali kondisi kestasionerannya dalam ratahan dan ragam. Kestasioneran dalam ratahan dapat dilihat berdasarkan plot ACF sementara kestasioneran dalam ragam dapat dilihat berdasarkan plot Box-Cox. Gambar 3.7 (a) menunjukkan bahwa plot ACF menurun secara perlahan (*tails off slowly*), yang mengindikasikan bahwa data tidak stasioner dalam

JURNAL MATEMATIKA, STATISTIKA DAN KOMPUTASI

Raihan Sabillah Adisecha, Nabil Ibni Nawawi, Naiya Dzil Izzati, Muhammad Tsaqif Najmuddin,
Nawwaf Ariq Rabbani Yenni Angraini, Adelia Putri Pangestika

rataan. Ketidakstasioneran ini juga didukung oleh hasil uji Dickey-Fuller, di mana p -value sebesar 0,060 lebih besar dari taraf signifikan 5%, sehingga hipotesis nol gagal ditolak dan data dinyatakan tidak stasioner dalam rata-rata. Sementara itu, plot Box-Cox pada Gambar 3.7 (b) menunjukkan nilai lambda optimum sebesar 0,909, dengan interval kepercayaan 95% yang memiliki batas bawah -0,505 dan batas atas 2,323. Karena interval tersebut mencakup nilai lambda sebesar satu, maka disimpulkan bahwa data bersifat stasioner dalam ragam.

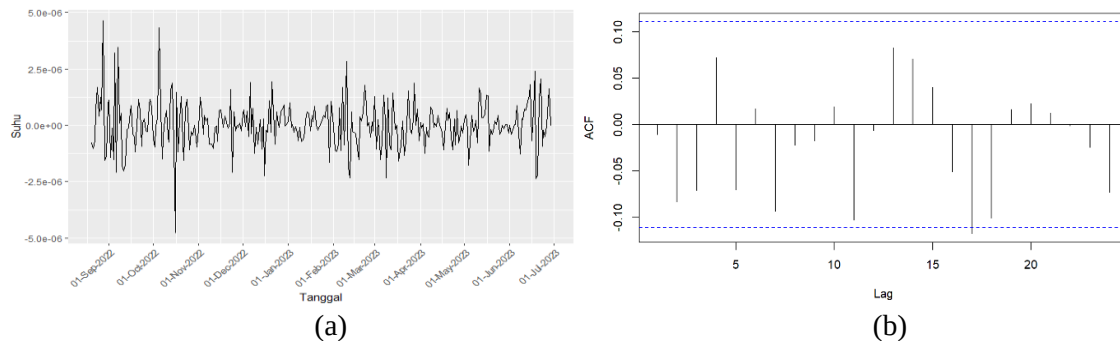


(a)

(b)

Gambar 3.7 (a) Plot ACF Data Latih Transformasi (b) Plot Box-Cox Data Latih Transformasi

Ketidakstasioneran dalam rata-rata pada data transformasi perlu ditangani dengan melakukan *differencing* sebanyak satu kali. Setelah dilakukan *differencing* sebanyak satu kali, terlihat bahwa data telah stasioner dalam rata-rata ditandai dengan plot data deret waktu pada Gambar 3.8(a) yang cenderung bergerak di suatu nilai tengah tertentu. Hasil ini didukung oleh plot ACF pada Gambar 3.8(b) dengan nilai ACF yang tidak signifikan pada setiap lag. Kestasioneran data dalam rata-rata juga diperkuat menggunakan uji ADF dengan p -value sebesar 0,01 yang kurang dari taraf nyata 5%.



(a)

(b)

Gambar 3.8 (a) Plot Data Latih Transformasi *Differencing* 1 (b) Plot ACF Data Latih Transformasi *Differencing* 1

3.2.6 Identifikasi Model dan Pendugaan Parameter Data Transformasi

Pemodelan ARIMA dilakukan dengan melihat pola segitiga nol pada plot EACF dari data latih yang telah ditransformasi dan dilakukan *differencing*. Plot EACF yang tertera pada Gambar 3.9 menunjukkan model-model tentatif yang diperoleh dari hasil identifikasi ini adalah ARIMA(0,1,1), ARIMA(1,1,1), ARIMA(2,1,2), ARIMA(3,1,3), dan ARIMA(4,1,2). Hasil pendugaan parameter kelima model tentatif ini ditunjukkan pada Tabel 3.4.

AR/MA	0	1	2	3	4	5	6	7	8	9	10	11	12	13
0	o	o	o	o	o	o	o	o	o	o	o	o	o	o
1	x	o	o	o	o	o	o	o	o	o	o	o	o	o
2	x	x	o	o	o	o	o	o	o	o	o	o	o	o
3	x	o	x	o	o	o	o	o	o	o	o	o	o	o
4	x	x	o	o	o	o	o	o	o	o	o	o	o	o
5	x	x	x	o	x	o	o	o	o	o	o	o	o	o
6	x	x	o	x	x	x	o	o	o	o	o	o	o	o
7	x	x	o	x	x	o	o	o	o	o	o	o	o	o

Gambar 3.9 Plot EACF Data Transformasi Latih *Differencing* 1

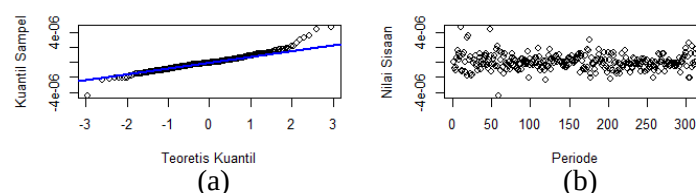
Tabel 3.4 Hasil Akurasi Pendugaan Parameter Model Tentatif

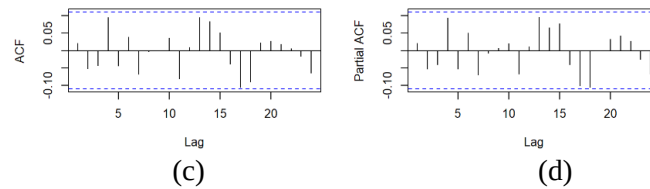
Model	Signifikansi Parameter	AIC	BIC
ARIMA(0,1,1)	Semua parameter tidak signifikan	-7761,710	-7754,210
ARIMA(1,1,1)	Semua parameter signifikan	-7766,970	-7755,720
ARIMA(2,1,2)	Semua parameter signifikan	-7765,900	-7747,150
ARIMA(3,1,3)	Semua parameter signifikan	-7764,660	-7738,420
ARIMA(4,1,2)	Hanya 3 parameter signifikan	-7764,960	-7738,720

Pendugaan parameter pada Tabel 3.4 menunjukkan bahwa model ARIMA(1,1,1) memiliki nilai AIC dan BIC yang terendah. Selain itu, semua parameter dalam model tersebut juga terbukti signifikan. Oleh karena itu, model ARIMA(1,1,1) dipilih sebagai model terbaik dari model tentatif yang akan dilanjutkan pada tahap uji diagnostik model.

3.2.7 Uji Diagnostik Sisaan

Uji diagnostik sisaan pada model ARIMA(1,1,1) dilakukan untuk memastikan bahwa sisaan menyebar normal, sisaan saling bebas, dan ragam sisaan homogen. Diagnostik model secara eksploratif ditunjukkan pada Gambar 3.10. Plot kuantil-kuantil normal pada Gambar 3.10(a) menunjukkan bahwa secara eksploratif, sisaan tidak menyebar normal, ditandai dengan titik-titik yang cenderung tidak mengikuti garis 45°. Selain itu, lebar pita sisaan pada Gambar 3.10(b) yang cenderung sama menunjukkan bahwa sisaan memiliki ragam yang homogen. Sementara itu, Gambar 3.10 (c) dan (d) memperlihatkan bahwa plot ACF dan PACF tidak menunjukkan adanya autokorelasi yang signifikan pada *lag-lag* awal yang mengindikasikan bahwa sisaan saling bebas. Kondisi ini diuji lebih lanjut menggunakan uji formal sebagaimana yang tertera pada Tabel 3.5.





Gambar 3.10 (a) Plot kuantil-kuantil normal (b) Plot *fitted value* vs *residual* (c) Plot ACF (d) Plot PACF

Tabel 3.5 menunjukkan hasil uji diagnostik sisaan menggunakan uji formal yang sejalan dengan uji diagnostik sisaan secara eksploratif. Hasil diagnostik sisaan menggunakan uji formal pada model ARIMA(1,1,1) menunjukkan bahwa metode transformasi berhasil menangani pelanggaran asumsi homoskedastisitas dengan baik. Hal ini sejalan dengan penelitian [9] dimana dalam penelitian ini penanganan heteroskedastisitas dengan metode transformasi memiliki performa yang lebih baik dibandingkan metode ARCH-GARCH. Meskipun pelanggaran asumsi homoskedastisitas dapat tertangani dengan baik, hasil uji diagnostik sisaan menggunakan uji formal pada model ARIMA(1,1,1) masih menunjukkan adanya pelanggaran asumsi normalitas sisaan. Asumsi normalitas sisaan merupakan salah satu syarat penting dalam analisis deret waktu sebagai penanda bahwa model yang digunakan telah mampu menangkap pola yang ada dalam data secara memadai. Meskipun demikian, menurut [26] pelanggaran asumsi normalitas sisaan dapat diabaikan karena jumlah data yang digunakan cukup besar.

Tabel 3.5 Hasil Uji Formal Analisis Asumsi Sisaan

Asumsi sisaan	H_0	$p\text{-value}$	Kesimpulan
Normalitas sisaan	Sisaan menyebar normal	2.200×10^{-16}	Tolak H_0
Sisaan saling bebas	Sisaan saling bebas	0,699	Tak tolak H_0
Ragam sisaan homogen	Ragam sisaan homogen	0,099	Tak tolak H_0
Nilai tengah sisaan sama dengan nol	Nilai tengah sisaan sama dengan nol	0,847	Tak tolak H_0

3.2.8 Overfitting

Overfitting model dilakukan dengan menambah ordo model tentatif terbaik sebesar 1 pada ordo AR atau MA. Berdasarkan model tentatif terbaik yang dipilih sebelumnya yaitu ARIMA(1,1,1), kandidat model *overfitting* dalam penelitian ini adalah ARIMA(2,1,1) dan ARIMA(1,1,2). Hasil evaluasi model *overfitting* yang tertera pada Tabel 3.6 menunjukkan bahwa model ARIMA(1,1,1) memiliki nilai AIC dan BIC terkecil. Semua parameter dalam model ARIMA(1,1,1) juga terbukti signifikan. Oleh karena itu, model ARIMA(1,1,1) dipilih sebagai model yang terbaik dalam pemodelan ARIMA.

Tabel 3.6 Hasil Akurasi Model *Overfitting*

Model	AIC	BIC
ARIMA(1,1,1)	-7766,970	-7755,720
ARIMA(2,1,1)	-7763,320	-7748,330
ARIMA(1,1,2)	-7763,370	-7748,370

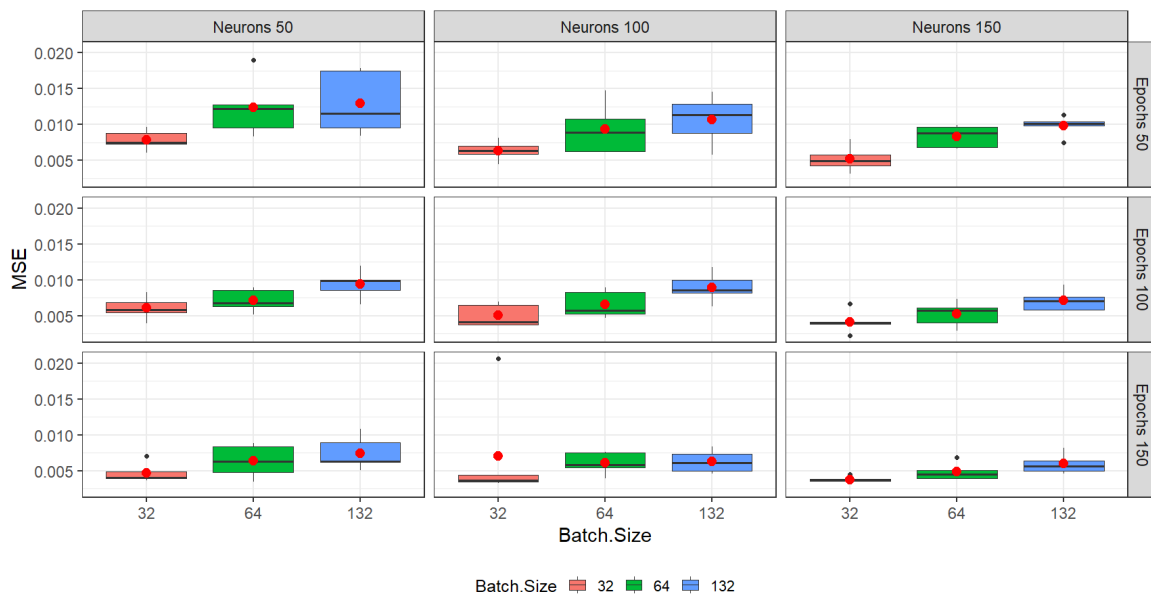
3.3 Pemodelan LSTM

Kombinasi hyperparameter yang digunakan dalam pemodelan LSTM ditentukan berdasarkan metode *Grid Search CV*. Metode ini digunakan karena telah terbukti efektivitasnya dalam menentukan hyperparameter terbaik pada penelitian [30]. Hasil *Grid Search CV* secara umum ditunjukkan oleh Gambar 3.10, di mana nilai *batch size* yang kecil dan *epoch*, serta *neuron* yang besar akan menghasilkan akurasi yang lebih baik. Gambar 3.10 juga menunjukkan bahwa kombinasi *batch size* 32, *epochs* 150, dan *neurons* 150 menjadi kombinasi terbaik karena kombinasi ini memiliki nilai rata-rata MSE yang terendah sebagaimana tertera pada Tabel 3.7.

Tabel 3.7 Kombinasi Terbaik

Batch Size	Epochs	Neurons	RMSE
32	150	150	0,003649

Kombinasi *hyperparameter* yang didapat ini cukup sejalan dengan penelitian sejenis, seperti pada penelitian [13] dengan data SPL Korea di mana penggunaan *neuron* yang besar menunjukkan peningkatan akurasi model LSTM. Selain itu, pada penelitian [5] dengan data Luas Es Laut Arktik menunjukkan penggunaan normalisasi *Min-Max* menghasilkan model yang lebih baik dengan *batch size* yang kecil, yaitu sebesar 32. Nilai *batch size* yang kecil dikatakan menghasilkan pelatihan yang lebih stabil dan dapat diandalkan. Namun, hasil ini berbeda dengan penelitian [28] dan [16] yang juga menggunakan data SPL di mana *epoch* yang optimum merupakan *epoch* yang kecil.



Gambar 3.10 Perbandingan Hasil Grid CV Setiap Kombinasi *Hyperparameter*

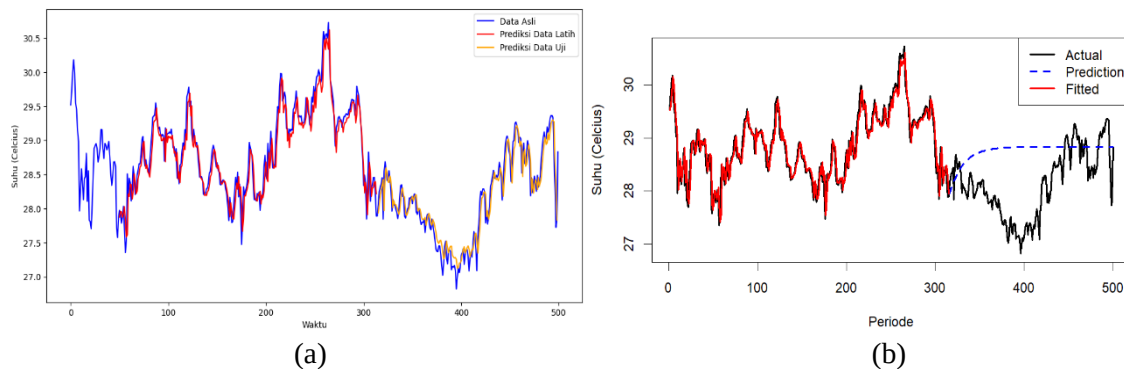
3.4 Menentukan model terbaik

Model terbaik dipilih dengan cara melakukan *fitting* model terbaik ARIMA dan LSTM pada data latih dilanjutkan dengan melakukan validasi pada data uji. Hasil *fitting* dan validasi kedua metode ditunjukkan oleh Tabel 3.8. Tabel 3.8 menunjukkan bahwa metode ARIMA dan LSTM dapat meramalkan data suhu permukaan laut selat sunda dengan sangat baik ditandai dengan nilai MAPE yang kurang dari 10. Nilai MAPE data latih dan data uji yang tidak jauh berbeda juga mengindikasikan bahwa tidak terjadi penyimpangan, baik *underfitting* maupun *overfitting* pada

masing-masing metode [2] Meskipun demikian, terlihat bahwa metode LSTM memiliki performa peramalan yang lebih baik dengan nilai MAPE yang lebih kecil. Hasil ini sejalan dengan penelitian [8] pada data kecepatan angin di Halifax. Penelitian tersebut menyimpulkan bahwa metode LSTM memiliki performa yang lebih baik dibandingkan metode ARIMA dengan nilai RMSE yang lebih kecil yaitu 3,124. Untuk melihat performa kedua metode dengan lebih detail, dilakukan analisis terhadap plot peramalan yang dihasilkan dari kedua metode sebagaimana yang tertera pada Gambar 3.11.

Tabel 3.8 Perbandingan Nilai MAPE

Metode	MAPE Data Latih	MAPE Data Uji
ARIMA	0,580	2,658
LSTM	0,512	0,564



Gambar 3.11 (a) Plot Metode LSTM (b) Plot Metode ARIMA

Gambar 3.11 menunjukkan bahwa peramalan menggunakan metode ARIMA dan LSTM menghasilkan pola yang berbeda. Pada data latih, kedua metode mampu menampilkan pola peramalan yang sesuai dengan data latih aktual. Namun, pada data uji, kedua metode menunjukkan performa peramalan yang sangat berbeda.

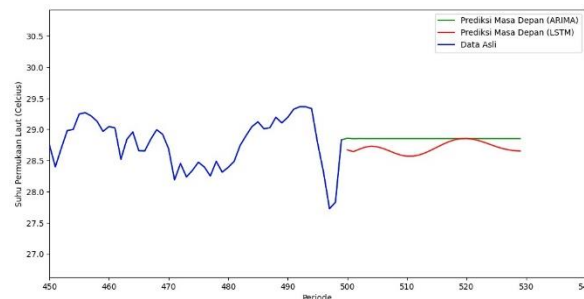
Metode ARIMA tidak mampu menggambarkan pola peramalan data uji yang sesuai dengan data uji aktual meskipun memiliki nilai MAPE yang rendah. Metode ARIMA membutuhkan data yang linier dan stasioner. Metode ini juga masih memiliki keterbatasan dalam menangkap pola nonlinier pada data [12] dan memodelkan data dengan kompleksitas yang tinggi [14]. Namun, data suhu permukaan laut Selat Sunda merupakan data yang cenderung menunjukkan pola nonlinier dengan fluktuasi yang tinggi. Akibatnya, metode ARIMA hanya akan mengambil nilai tengah pada data tanpa menangkap fluktuasi yang kompleks sehingga metode ini gagal menggambarkan pola yang sesuai dengan kondisi aktualnya.

Di sisi lain, metode LSTM mampu menampilkan pola prediksi data uji yang sesuai dengan data uji aktual. Hal ini disebabkan karena metode LSTM dirancang untuk memproses data secara berurutan dimana setiap langkah waktu dalam metode ini tidak hanya memproses input saat ini, tetapi juga mempertimbangkan *state* atau informasi dari langkah sebelumnya. Hal ini membuat model dapat mengingat informasi penting dari masa lalu yang memungkinkan model untuk mengenali pola jangka panjang dan menangkap hubungan yang kompleks (non-linier) antar data tanpa mengharuskan data bersifat stasioner [15]. Akibatnya, pola yang terbentuk dapat sesuai dengan data yang telah dipelajari sebelumnya tanpa mementingkan aspek nilai tengah pada

keseluruhan data. Oleh karena itu, metode LSTM lebih sesuai untuk digunakan dalam peramalan SPL di Selat Sunda

3.5 Peramalan SPL Selat Sunda

Peramalan SPL Selat Sunda dilakukan menggunakan metode ARIMA dan LSTM terbaik dengan periode peramalan sebanyak 30 periode kedepan. Hasil peramalan menggunakan metode ARIMA dan LSTM ditunjukkan pada Gambar 3.12. Peramalan SPL Selat Sunda menggunakan metode ARIMA menunjukkan hasil yang cenderung konstan sebagaimana yang terjadi pada proses validasi data uji. Hal ini dikarenakan metode ARIMA hanya menangkap nilai tengah pada data sehingga cenderung tidak relevan digunakan untuk data SPL Selat Sunda yang memiliki volatilitas yang tinggi. Sebaliknya, peramalan menggunakan metode LSTM cenderung menunjukkan pola yang fluktuatif. Metode LSTM cenderung mampu mengenali pola jangka panjang sehingga hasil peramalannya cenderung lebih relevan untuk data SPL Selat Sunda yang berfluktuasi sepanjang hari.



Gambar 3.12 Plot Data Ramal Metode LSTM

4. KESIMPULAN

Pemodelan data SPL Selat Sunda menggunakan metode ARIMA menghasilkan model terbaik ARIMA (1,1,1) dan dengan LSTM didapatkan kombinasi *batch size* 32, *epochs* 150, dan *neurons* 150 sebagai *hyperparameter* terbaik. Berdasarkan nilai MAPE data latih dan data uji, metode LSTM lebih unggul dibandingkan metode ARIMA dengan nilai MAPE data latih sebesar 0,512% dan data uji sebesar 0,564%. Metode LSTM mampu mengikuti pola pada validasi data uji yang tidak mampu dilakukan oleh ARIMA. Hasil peramalan 30 periode ke depan dengan ARIMA menunjukkan nilai yang konstan, sementara LSTM menunjukkan pola yang fluktuatif di sepanjang hari. Namun, SPL yang berfluktuasi mengindikasikan LSTM lebih relevan untuk digunakan.

DAFTAR PUSTAKA

- [1] Alim, M., N., 2023. Pemodelan Time Series Data Saham LQ45 dengan Algoritma LSTM, RNN, dan Arima, PRISMA: *Prosiding Sem-Nas Matematika*, Vol. 6 thn 2024, 694-701. Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Semarang, Semarang.
- [2] Amalia, DW, 2024. Peramalan Jumlah Incident Information Technology PT XYZ Menggunakan Artificial Neural Network (ANN). *Journal of Emerging Information Systems and Business Intelligence*. Vol. 5, No. 3, 20–27.
- [3] Arissinta, I., O., Sulistiyowati, I., D Kurnianto, D. & Kharisudin, I., 2022. Pemodelan Time Series untuk Peramalan Web Traffic Menggunakan Algoritma Arima, LSTM, dan GRU, PRISMA: *Prosiding Sem-Nas Matematika*, Vol. 5 thn 2022, 693–700. Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Semarang, Semarang.
- [4] Batubara, M., P., N., Zainuri, M., Kunarso, K. & Puryajati, A., D., 2022, Diagnosa Perilaku

- MJO Aktif pada saat La Nina Kuat (2011/2012) di Perairan Tropis. *Buletin Oseanografi Marina*. Vol. 11, No. 2, 193–205.
- [5] Christofi, V., P., Wang, X., Christofi, V., P. & Wang, X., 2024. Optimizing the LSTM Deep Learning Model for Arctic Sea Ice Melting Prediction. *Atmospheric and Climate Sciences*. Vol. 14, No. 4, 429–449.
 - [6] Cryer, J., D. & Chan, K., S. 2008. *Time Series Analysis with Applications in R*, Second Edition. Springer Science & Business Media., New York.
 - [7] Dhakal, D., Gautam, Y. & Bhattarai, A., 2021. Exploring a deep LSTM neural network to forecast daily PM2.5 concentration using meteorological parameters in Kathmandu Valley, Nepal, Air Quality. *Atmosphere and Health*, Vol. 14, 83–96.
 - [8] Elsaraiti, M. & Merabet, A., 2021. A Comparative Analysis of the ARIMA dan LSTM Predictive Models and Their Effectiveness for Predicting Wind Speed. *Energies*, Vol. 14, No. 20, 1-16.
 - [9] Gestel, T., V., Suykens, J., A., K., Baestaens, D., E., Lambrechts, A., Lanckriet, G. & Vandaele, B., 2001. Financial Time Series Prediction Using Least Squares Support Vector Machines Within the Evidence Framework. *IEEE Transactions on Neural Networks*, Vol. 12, No. 4, 809–821.
 - [10] Hilal, Y., N., Nainggolan, G., D., A., Syahputri, S., H. & Kartiasih, F., 2024. Comparison of ARIMA and LSTM methods in predicting Jakarta sea level. *Jurnal Ilmu Dan Teknologi Kelautan Tropis*, Vol. 16, No. 2, 163–178.
 - [11] JAXA, JAXA Earth-graphy / Space Technology Directorate, (2024).
 - [12] Jeong, K., J. & Shin, Y., M., 2022. Time-Series Anomaly Detection with Implicit Neural Representation. *arXiv preprint arXiv: 2201.11950*.
 - [13] Kim, M., Yang, H. & Kim, J., 2020. Sea Surface Temperature and High Water Temperature Occurrence Prediction Using a Long Short-Term Memory Model. *Remote Sensing*, Vol. 12, No. 21, 3654.
 - [14] Kundu, A., Sahu, A., Serpedin, E. & Davis, K., 2020. A3D: Attention-based auto-encoder Anomaly Detector for False Data Injection Attacks. *Electric Power Systems Research*, Vol. 189, No. 106795, 1-7.
 - [15] Lin, S. & Feng, Y., 2022. Research on stock price prediction based on orthogonal Gaussian basis function expansion and Pearson correlation coefficient weighted lstm neural network, *Advances in Computer. Signals and Systems*, Vol. 6, No. 5, 23–30.
 - [16] Lin, Y. & Zhong, G. A., 2021. Multi-Channel LSTM Model for Sea Surface Temperature Prediction. *Journal of Physics: Conference Series*, Vol. 1880, 012029.
 - [17] Mahkya, D Angraini, Fitriawati, Siahaan, R., M., 2020. Pemodelan dan prediksi jumlah penumpang pelabuhan bakauheni selama periode tsunami Selat Sunda menggunakan autoregressive integrated moving average. *Journal of Science and Applicative Technology*. Vol. 4, No. 1, 32-37.
 - [18] Mawaddah, S., N., Prakoso, L., Y. & Sudiarso, A., 2021. Aspek Geostrategi pada Selat Sunda dalam Mencapai Visi Poros Maritim Dunia. *Jurnal Strategi Pertahanan Laut*, Vol. 7, No. 3, 185-194.
 - [19] Mulyani, A., S., 2020. Antisipasi Terjadinya Pemanasan Global dengan Deteksi Dini Suhu Permukaan Air Menggunakan Data Satelit. *e-Journal CENTECH*, Vol. 2, No. 1, 22-29.
 - [20] Nastiti, E., Damayanti, T., Madina, S., A., 2023. Dampak Pelanggaran Asumsi Klasik Terhadap Estimasi Model Ekonometrika. *Jurnal Pijar Studi Manajemen Dan Bisnis*, Vol. 2, No. 1, 566–577.
 - [21] Ng, V., K., Y. & Cribbie, R., A., 2018. The Gamma Generalized Linear Model, Log Transformation, and the Robust Yuen-Welch Test for Analyzing Group Means with Skewed

- and Heteroscedastic Data. *Communications in Statistics – Simulation and Computation*, Vol. 48, No. 8, 2269-2286.
- [22] Nurhambali, M., R., Angraini, Y. & Fitrianto, A., 2024. Implementation of Long Short-Term Memory for Gold Prices Forecasting. *Malaysian Journal of Mathematical Sciences*, Vol. 18, No. 2, 394-422.
- [23] Pamungkas, M., B. & Wibowo, A., 2018. Aplikasi Metode ARIMA Box-Jenkins untuk Meramalkan Kasus DBD di Provinsi Jawa Timur. *The Indonesian Journal of Public Health*, Vol. 13, No. 2, 181-194.
- [24] Pranolo, A., Setyaputri, F., U., Paramarta, A., K., I., Triono, A., P., P., Fadhilla, A., F., Akbari, A., K., G., Utama, A., B., P., Wibawa, A., P., Uriu, W., 2024. Enhanced Multivariate Time Series Analysis Using LSTM: A Comparative Study of Min-Max and Z-Score Normalization Techniques. *ILKOM Jurnal Ilmiah*, Vol. 16, No. 2, 210–220.
- [25] Roza, A., Violita, E., S. & Aktivani, S., 2022. Study of Inflation using Stationary Test with Augmented Dickey Fuller & Phillips-Peron Unit Root Test (Case in Bukittinggi City Inflation for 2014-2019). *Eksakta: Berkala Ilmiah Bidang MIPA*, Vol. 23, No. 2, 106-116.
- [26] Schaffer, A., L., Dobbins, T., A. & Pearson, S., A. 2021. Interrupted Time Series Analysis using Autoregressive Integrated Moving Average (ARIMA) Models: a Guide for Evaluating Large-Scale Health Interventions. *BMC Medical Research Methodology*, Vol. 21, 1-12.
- [27] Switrayana, I., N., Hammad, R., Irfan, P., Sujaka, T., T., Nasri, M., H., 2025. Comparative Analysis of Stock Price Prediction Using Deep Learning with Data Scaling Method, *JTIM : Jurnal Teknologi Informasi Dan Multimedia*, Vol. 7, No. 1, 78–90.
- [28] Taylor, J., Feng, M., 2022. A deep learning model for forecasting global monthly mean sea surface temperature anomalies. *Frontiers in Climate*, Vol. 4, 932932.
- [29] Wiranda, L. & Sadikin, M., 2019. Penerapan Long Short-Term Memory pada Data Time Series untuk Memprediksi Penjualan Produk PT. Metiska Farma. *Jurnal Nasional Pendidikan Teknik Informatika: JANAPATI*, Vol. 8, No. 3, 184-196.
- [30] Yanti, F., Sari, B., N., Defiyanti, S., 2024. Implementasi Algoritma Lstm Pada Peramalan Stok Obat (Studi Kasus: Puskesmas Beber). *Jurnal Mahasiswa Teknik Informatika*, Vol. 8, No. 4, 6082–6089.
- [31] Zhang, Q., Wang, H., Dong, J., Zhong, G., Sun, X., 2017. Prediction of Sea Surface Temperature Using Long Short-Term Memory. *IEEE Geoscience and Remote Sensing Letters*, Vol. 14, No. 10, 1745–1749.