



Peramalan Kualitas Air Danau Toba Melalui Citra Satelit dengan Model Peramalan Univariat

Firman Emmanuel Declarantius Parulian¹, Hasna Arifah Nur Fatih², Wimbi Uelsan Gurusinga³, Robert Kurniawan⁴

^{1,2,3}Program Studi Statistika, Politeknik Statistika STIS Jakarta, Indonesia

⁴Program Studi Komputasi Statistik, Politeknik Statistika STIS Jakarta, Indonesia

¹firmanemmanuel142@gmail.com

²arifah.hasna@gmail.com

³wimbiuel@gmail.com

⁴robertk@stis.ac.id

Corresponding author email: firmanemmanuel142@gmail.com

Abstract: The aim of this research is to carry out monitoring, compare forecasting methods, and predict the water quality of Lake Toba. Forecasting is carried out for the next 12 months which is approached by the variables Dissolved Oxygen (DO), pH, Land Surface Temperature (LST), Normalized Difference Turbidity Index (NDTI). Collecting monthly water quality data using Landsat 8 and MODIS satellite imagery. Data was taken in the period January 2014 to April 2024. The research results show that the four water quality variables are still included in the good category. The best forecasting method for each Lake Toba water quality variable for DO is Prophet, pH uses the LSTM method, and SARIMA is used to predict LST and NDTI. In terms of average RMSE, the ARIMA/SARIMA method is the best method for predicting water quality variables. Based on the forecast for the next 12 months, the DO variable tends to increase. However, LST shows a decrease, while the pH and NDTI variables show stagnant movement.

Keywords: Lake Toba, Forecasting Analysis, satellite image, ARIMA/SARIMA, Prophet, LSTM

Abstrak: Tujuan penelitian ini adalah melakukan monitoring, membandingkan metode peramalan, dan meramalkan kualitas air Danau Toba. Peramalan dilakukan selama 12 bulan ke depan yang didekati dengan variabel Dissolved Oxygen (DO), pH, Land Surface Temperature (LST), Normalized Difference Turbidity Index (NDTI). Pengambilan data bulanan kualitas air menggunakan citra satelit Landsat 8 dan MODIS. Data diambil pada periode Januari 2014 hingga April 2024. Hasil penelitian menunjukkan keempat variabel kualitas air masih termasuk dalam kategori baik. Metode peramalan yang digunakan adalah SARIMA, Prophet, dan LSTM. Metode terbaik untuk setiap variabel kualitas air Danau Toba untuk DO adalah Prophet, pH menggunakan metode LSTM, dan SARIMA digunakan untuk meramalkan LST dan NDTI. Secara rata-rata RMSE, metode ARIMA/SARIMA merupakan metode terbaik dalam meramalkan variabel-variabel kualitas air Danau Toba. Berdasarkan peramalan 12 bulan ke depan, variabel DO cenderung akan naik. Namun, LST menunjukkan adanya penurunan, sedangkan variabel pH dan NDTI menunjukkan pergerakan yang stagnan.

Kata kunci: Danau Toba, Analisis Peramalan, Citra Satelit, ARIMA/SARIMA, Prophet, LSTM

I. PENDAHULUAN

Indonesia adalah negara kepulauan dengan perairan yang cukup luas, mencapai 5,9 juta km^2 . Hal ini menjadikan Indonesia memiliki potensi sumber daya perikanan yang besar [1]. Salah satu daerah yang memiliki prospek sumber daya air yang besar adalah Danau Toba, Provinsi Sumatera Utara. Danau Toba merupakan danau terbesar di Indonesia dan Asia Tenggara dengan luas permukaan 1.124 km^2 (112.400 ha), volume danau sekitar 256,2 km^3 ($256,2 \times 10^9 m^3$) dan kedalaman mencapai 508 m. Sebagai danau terluas, Danau Toba memiliki potensi besar dalam memanfaatkan perairannya.

Potensi-potensi yang dimiliki Danau Toba antara lain adalah perikanan, bahan baku air bersih, pembangkit listrik, pariwisata, dan transportasi. Produksi perikanan Provinsi Sumatera Utara tahun 2021 sebesar 439.145 ton, meningkat 137% dari tahun 2017. Selain itu, Danau Toba makin sering dipergunakan untuk tujuan pariwisata dan perlombaan internasional [2]. Oleh karena itu, penggunaan Danau Toba sebagai penggerak perekonomian perlu dimaksimalkan. Danau Toba mampu menjadi penopang perekonomian daerah di sekitarnya dengan mengoptimalkan sumber daya alam. Baik dari

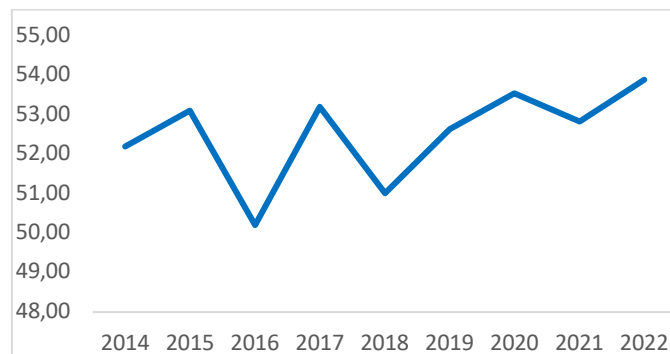


perikanan, bahan baku air bersih, pariwisata, dan potensi lain sehingga memberikan peluang besar wilayah sekitar Danau Toba untuk maju. Struktur ekonomi, pertumbuhan ekonomi regional, dan potensi yang dimiliki Danau Toba mampu mengurangi ketidakmerataan pembangunan dan kemiskinan serta menciptakan ketahanan regional dan komunitas [2]. Namun, potensi perikanan dan bahan baku air bersih sangat bergantung pada kualitas air.

Kualitas air adalah parameter krusial dalam menentukan aksesibilitas penggunaan dan pengoptimalan perikanan dan bahan baku air bersih. Akses manusia terhadap sumber daya air bersih menjadi sangat penting untuk berbagai keperluan konsumsi [3]. Hal ini yang menjadi tujuan utama *Sustainable Development Goals* (SDGs) pada goal ke-enam, yaitu menjamin ketersediaan serta pengolahan air bersih dan sanitasi yang berkelanjutan untuk semua. Kualitas air bersih sudah diatur oleh World Health Organization (WHO) dan Peraturan Kementerian Kesehatan Nomor 32 Tahun 2017 mengenai Standar Baku Mutu Kesehatan Lingkungan dan Persyaratan Kesehatan Air untuk Keperluan Higiene Sanitasi, Kolam Renang, Solus per Aqua, dan Pemandian Umum. Dalam peraturan tersebut, dipaparkan standar umum setiap parameter-parameter wajib kualitas air bersih.

Standar air bersih oleh WHO dan Peraturan Kementerian Kesehatan menjadi panduan dalam menganalisis kualitas air. Meskipun demikian, meningkatnya jumlah penduduk dan urbanisasi membawa dampak negatif pada kualitas air. Perubahan iklim dan pemanasan global yang terjadi juga menyebabkan masalah besar pada kualitas sumber daya air [4]. Secara spesifik, perubahan iklim di Danau Toba telah mempengaruhi tren peningkatan suhu dan curah hujan selama 40 puluh tahun terakhir [5].

Penghitungan kualitas air Indonesia mengikuti penghitungan WQI (*Water Quality Index*) atau IKA (Indeks Kualitas Air) yang dilakukan penyesuaian menurut Kementerian Lingkungan Hidup dan Kehutanan Republik Indonesia. IKA Indonesia pada tahun 2022 sebesar 53,88. Nilai indeks ini termasuk dalam kategori sedang. Pergerakan IKA Indonesia dari tahun 2014 hingga 2022 ditampilkan pada Gambar 1.



Gambar 1. IKA Indonesia Tahun 2014-2022

Gambar 1 menunjukkan kualitas air Indonesia cenderung fluktuatif dan memiliki kecenderungan naik. Meskipun pergerakan kualitas air Indonesia relatif membaik, perlu dilakukan pemantauan kondisi sumber daya air menggunakan citra satelit pada setiap parameter kualitas air dan setiap wilayah untuk memaksimalkan potensi perairan tiap daerah, seperti di Danau Toba. Selain itu, peramalan terhadap variabel-variabel parameter air di Danau Toba dapat menjadi pertimbangan pemerintah setempat untuk melakukan perawatan dan pencegahan. Oleh karena itu, perlu dilakukan perbandingan metode-metode peramalan yang tepat untuk meramalkan variabel-variabel kualitas air di Danau Toba.



Penelitian mengenai kualitas air di Danau Toba telah dilakukan oleh beberapa peneliti menggunakan data sekunder, web *Online Monitoring* (ONLIMO), pembangunan sensor pada pelampung, ataupun pengambilan sampel lapangan. Penelitian-penelitian terdahulu menunjukkan bahwa variabel pH, kandungan Oksigen (O_2), dan suhu cenderung baik [6]–[9].

Penelitian mengenai penggunaan metode-metode peramalan telah dilakukan untuk daerah perairan, baik sungai maupun danau. Penelitian-penelitian sebelumnya mengungkapkan bahwa penggunaan peramalan klasik seperti ARIMA dan SARIMA masih sangat relevan digunakan untuk kualitas air musiman [10]. Penelitian oleh Kogekar et al. (2021) meneliti *Biochemical Oxygen Demand* (BOD) dan *Dissolved Oxygen* (DO) Sungai Ganga menggunakan ARIMA, SARIMA, dan Prophet. Penelitian ini menemukan bahwa SARIMA dan Prophet merupakan model terbaik [11]. Shehhi dan Kaya (2021) menggunakan data citra satelit untuk meneliti klorofil dan suhu permukaan di Teluk Arab. Temuan penelitian yaitu penggunaan Neural Network lebih baik dibandingkan SARIMA [12]. Selain itu, penelitian oleh Ahmed et al. (2022) menggunakan data citra satelit untuk menganalisis kualitas air berupa *Electric Conductivity* (EC) dan *Dissolved Oxygen* (DO) di daerah aliran sungai Rawal, Pakistan menggunakan metode *deep learning* seperti RNN, CNN, MLP, dan LSTM. Penelitian ini menemukan bahwa *bi-directional LSTM* adalah metode yang paling baik [13].

Berdasarkan penelitian-penelitian terkait, kebaruan penelitian ini adalah melakukan monitoring kualitas air Danau Toba berupa pH, suhu, kekeruhan, dan kandungan oksigen yang diambil dari citra satelit dari tahun 2014 hingga 2024. Penggunaan keempat variabel dikarenakan keterbatasan ketersediaan data citra satelit yang digunakan. Selain itu, penelitian ini juga melakukan peramalan dengan menggunakan beberapa metode, baik klasik seperti ARIMA dan SARIMA ataupun *machine learning* seperti Prophet dan LSTM.

Penelitian ini bertujuan untuk memonitoring kualitas air Danau Toba dari beberapa variabel kualitas air menggunakan citra satelit. Penelitian ini juga akan membandingkan metode peramalan terbaik untuk meramalkan variabel-variabel kualitas air di Danau Toba. Selanjutnya, penelitian ini akan melakukan peramalan selama 12 bulan ke depan. Oleh karena itu, penelitian ini dapat digunakan sebagai acuan pemerintah setempat dalam mengontrol kualitas air Danau Toba.

II. TINJAUAN PUSTAKA

***Dissolved Oxygen* (DO)**

Dissolved Oxygen (DO) merupakan kandungan oksigen yang terlarut dalam air yang dapat dihasilkan oleh tanaman atau alga melalui fotosintesis [14]. DO merupakan indikator yang penting dalam mengidentifikasi kualitas air karena mencerminkan proses produksi dan konsumsi oksigen dalam air [15]. Konsentrasi DO yang terlalu rendah akan mengganggu keseimbangan ekologis. Sedangkan konsentrasi DO yang terlalu tinggi dapat menyebabkan produksi bau, asam, dan meningkatkan zat berbahaya dalam air yang akan berpengaruh pada kesehatan apabila dikonsumsi [16]. Oleh karena itu, sangat penting untuk menjaga konsentrasi DO dalam air untuk melindungi ekosistem dalam air.

pH

pH adalah variabel yang mengukur derajat keasaman atau kebasaan dalam suatu larutan. Nilai pH yang terlalu rendah atau terlalu tinggi akan mengganggu keseimbangan ekosistem akuatik. Nilai pH yang aman untuk air minum berada pada interval 6 hingga 8,5. Perubahan nilai pH dapat mempengaruhi keseimbangan ion pada organ yang dimiliki setiap organisme dalam air. Setiap



organisme dalam air memiliki kemampuan yang berbeda dalam mengatur keseimbangan asam-basa dalam tubuh [17].

Land Surface Temperature (LST)

Land Surface Temperature (LST) merupakan suhu pada permukaan bumi yang dihitung dari hasil pantulan sensor citra satelit [18]. Namun, nilai data LST yang diambil citra satelit lebih rendah dibandingkan yang dirasakan di lapangan. Hal tersebut karena citra satelit beresolusi sedang dan selalu memberikan nilai piksel rata-rata [19]. LST merupakan alat yang penting untuk analisis hidrologi. Peningkatan LST dapat menyebabkan berkurangnya pasokan air jernih. LST yang meningkat di perairan memiliki korelasi yang erat dengan penurunan vegetasi di perairan.

Normalized Difference Turbidity Index (NDTI)

Normalized Difference Turbidity Index (NDTI) merupakan indeks yang digunakan untuk mengukur tingkat kekeruhan air dengan citra satelit [20]. Kekeruhan yang dimaksud adalah kekeruhan air atau gangguan cahaya akibat material sedimen tersuspensi [21]. NDTI digunakan sebagai alat untuk mengevaluasi perubahan kualitas air. Data NDTI berada pada rentang nilai antara -1 hingga +1. Nilai NDTI yang mendekati +1 menunjukkan kekeruhan yang tinggi sehingga mengurangi kejernihan air dan penetrasi cahaya yang masuk. Nilai NDTI yang dikategorikan jernih memiliki rentang nilai dari -0,2 hingga 0. Sedangkan rentang nilai 0 hingga 0,2 dikategorikan sebagai keruh [22]. Tingkat kekeruhan air merupakan faktor penting bagi setiap perairan. Hal ini dikarenakan kekeruhan akan mempengaruhi fungsi perairan sehubungan dengan adanya beban sedimen yang tersuspensi [23]. Oleh karena itu, NDTI sebagai variabel kekeruhan menjadi penting dalam analisis kualitas air.

III. METODE PENELITIAN

Penelitian ini dibatasi pada variabel-variabel air berupa parameter fisika dan kimia non-logam. Parameter fisika yang digunakan, yaitu suhu yang didekati oleh LST dan kekeruhan yang diprosi oleh NDTI. Sedangkan parameter kimia non-logam yang digunakan adalah pH dan kandungan oksigen yang menggunakan variabel DO. Keempat variabel tersebut merupakan beberapa variabel yang digunakan untuk mengukur kualitas air.

Data diambil dari citra satelit menggunakan *Google Earth Engine* dengan memasukkan *shapefile* (shp) Kabupaten Samosir. Satuan dan satelit yang digunakan setiap variabel ditunjukkan pada Tabel 1.

Tabel 1. Keterangan satuan dan satelit setiap variabel

Variabel	Satuan	Satelit
LST	°C	MODIS
pH	-	Landsat 8
DO	mg/l	Landsat 8
NDTI	-	Landsat 8

Berdasarkan Tabel 1, penelitian menggunakan dua jenis satelit, yaitu MODIS dan Landsat 8. Data yang digunakan adalah rata-rata bulanan variabel kualitas air dari keseluruhan wilayah Danau Toba dan berupa data *time series* dari bulan Januari 2014 hingga April 2024. Adapun jumlah series pada penelitian ini sebanyak 124 *series*.



Penelitian ini melakukan peramalan (*forecasting*) pergerakan variabel kualitas air yang digunakan, yaitu LST, pH, DO, dan NDTI. Metode peramalan (*forecasting*) yang digunakan berupa model ARIMA/SARIMA, Prophet, dan LSTM (*Long-Short Term Memory*). Data dibagi menjadi 90% sebagai data *training* dan 10% sebagai data *testing*. Data testing berupa 12 *series* terakhir variabel-variabel kualitas air Danau Toba. Metode peramalan yang digunakan sebagai berikut.

1. Model ARIMA dan SARIMA

Model ARIMA merupakan salah satu pendekatan ekonometrik yang digunakan untuk pemodelan deret waktu univariat. Bentuk umum ARIMA(p, d, q) mencakup parameter bilangan bulat non-negatif. Ordo p merupakan ordo AutoRegressive (AR), ordo d merupakan jumlah perbedaan yang diperlukan untuk membuat deret waktu stasioner, dan ordo q adalah ordo Moving Average (MA). Orde pada model ARIMA diidentifikasi menggunakan fungsi autokorelasi (ACF) dan fungsi autokorelasi parsial (PACF). Model ARIMA (p,d,q) yang diperkenalkan oleh Box dan Jenkins sebagai berikut [24].

$$(1 - \phi_1 B - \dots - \phi_p B^p)(1 - B)^d y_t = c + (1 - \theta_1 B - \dots - \theta_q B^q) \varepsilon_t \quad (1)$$

Model SARIMA berasal dari model ARIMA dengan mempertimbangkan faktor-faktor musiman berdasarkan penerapan praktisnya. Dimana p adalah jumlah ordo AutoRegressive (AR) tanpa musiman, q adalah jumlah ordo Moving Average (MA) tanpa musiman, dan d adalah jumlah perbedaan tanpa musiman yang diperlukan suatu barisan untuk mencapai stasioneritas. Sedangkan (P,D,Q) menunjukkan representasi setiap bagian barisan setelah membuat perbedaan musiman. Rumus umum SARIMA(p, d, q)(P, D, Q)¹² ditunjukkan sebagai berikut.

$$\phi_p(B) \Phi_P(B^S)(1 - B)^d (1 - B^S)^D y_t = c + \theta_q(B) \Theta_Q \varepsilon_t \quad (2)$$

2. Prophet

Prophet adalah alat *open-source* atau *library* yang disediakan oleh Facebook Inc., tersedia baik dalam Python maupun R. Prophet digunakan untuk mengembangkan alat peramalan yang fleksibel yang mudah digunakan dan disesuaikan [25]. Model dasar ini menampilkan deret waktu yang dapat didekomposisi menjadi berikut [26].

$$y(t) = g(t) + s(t) + h(t) + \varepsilon_t \quad (3)$$

Dengan y(t) adalah peramalan, g(t) adalah kurva *linier piecewise* untuk pemodelan perubahan non-periodik dalam deret waktu, s(t) menunjukkan perubahan periodik, h(t) adalah efek liburan dengan jadwal tidak teratur, t mewakili akun istilah kesalahan untuk setiap perubahan yang tidak biasa tidak mengakomodasi dengan model.

3. Long Short-Term Memory (LSTM)

Model LSTM adalah bagian khusus dari model RNN (*Recurrent Neural Network*) yang pertama kali diperkenalkan oleh Hochreiter dan Schmidhuber pada tahun 1997. LSTM memiliki kualitas ketergantungan jangka panjang yang unik dan membuatnya lebih presisi dibandingkan RNN konvensional mana pun [27]. Seiring bertambahnya jumlah data berurut yang ingin dilatih, RNN masih menghadapi masalah *vanishing gradient*. Bentuk LSTM yang paling dasar disebut sel memori, yang berguna untuk memanggil kembali dan mengolah output unit yang jelas dalam berbagai fase waktu. Untuk memanggil data untuk konteks temporal lagi, sel memori LSTM menggunakan langkah-langkah sel. LSTM memiliki berbagai gate yang digunakan untuk aliran dan kontrol informasi antara berbagai langkah.



Fungsi aktivasi *node* menyatakan hasil dari *node* tertentu. Fungsi aktivasi yang terutama digunakan untuk jaringan LSTM secara khusus sigmoid dan tangensial (*tanh*). Tata letak LSTM terintegrasi dengan fungsi sigmoid melalui gerbang lupakan dan gerbang input dan dengan fungsi *tanh* untuk vektor kandidat yang menggantikan vektor keadaan sel. Fungsi aktivasi LSTM ini dihitung untuk Input Gate I_t , Output gate O_t , Forget gate F_t , Candidate vector C_t , Cell state C_t , and Hidden state h_t , menggunakan rumusan sebagai berikut.

$$I_t = \text{sigmoid}(W_i[X(t), h_{t-1}] + b_i) \quad (4)$$

$$F_t = \text{sigmoid}(W_f[X(t), h_{t-1}] + b_f) \quad (5)$$

$$O_t = \text{sigmoid}(W_o[X(t), h_{t-1}] + b_o) \quad (6)$$

$$C_t = \text{tanh}(W_c[X(t), h_{t-1}] + b_c) \quad (7)$$

$$C_t = (F_t \times C_{t-1}) + (I_t \times C_t) \quad (8)$$

$$h_t = O_t \times \text{tanh}(C_t) \quad (9)$$

Keterangan:

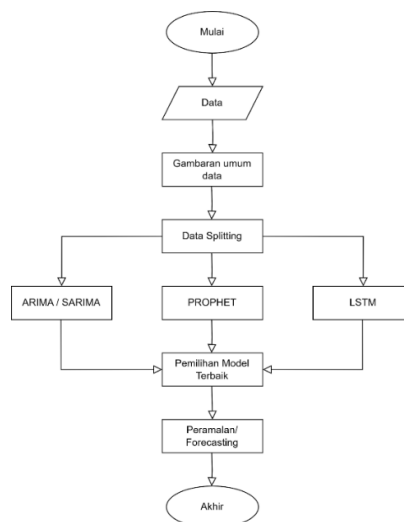
$X(t)$: *Input*

h_{t-1} : keadaan sebelumnya

W : bobot

b : *bias* di setiap *gate*

Prosedur analisis yang dilakukan yaitu mengambil data rata-rata pH, DO, LST, dan NDTI di Danau Toba pada periode Januari 2014 sampai April 2024 menggunakan citra satelit.. Selanjutnya membuat analisis deskriptif sekaligus grafik *time series* dari data yang sudah diambil. Langkah berikutnya adalah melakukan uji stasioner varians dan rata-rata dengan menggunakan *Box Cox* dan *Augmented Dicky Fuller Test* baik untuk *seasonal* ataupun *non-seasonal*. Dengan menggunakan 90% data awal sebagai data *training*, kandidat-kandidat model ARIMA/SARIMA dipilih dan model terbaiknya ditentukan berdasarkan koefisien signifikan dan nilai AIC serta SIC terkecil. Kemudian, penelitian dilanjutkan dengan pembetulan model Prophet dan LSTM menggunakan data *training*. Selanjutnya, dilakukan pemilihan model terbaik dari ketiga metode berdasarkan nilai RMSE dan MAE-nya untuk masing-masing variabel kualitas air. Dengan metode terbaik di masing-masing variabel, dilakukan peramalan selama 12 bulan ke depan. Diagram alir prosedur analisis penelitian ditunjukkan pada Gambar 2.



Gambar 2. Diagram Alir



IV. HASIL DAN PEMBAHASAN

Dalam melakukan peramalan, penelitian melakukan analisis deskriptif terlebih dahulu terhadap pergerakan setiap variabel kualitas air yang digunakan. Gambaran secara umum terkait karakteristik untuk setiap variabel yang digunakan dapat dilihat dengan statistik deskriptif. Hasil statistik deskriptif harga penutupan saham ditunjukkan pada Tabel 2.

Tabel 2. Statistik Deskriptif Variabel Kualitas Air

Statistik	DO	pH	LST	NDTI
Minimum	11,846	4,365	16,917	-0,134
Std. Dev	0,053	0,416	2,007	0,022
Mean	12,017	7,909	25,101	-0,0189
Median	12,027	7,939	25,286	-0,013
Maximum	12,104	8,800	31,403	0,009

Tabel 2 menunjukkan nilai minimum, standar deviasi, rata-rata, median, dan maksimum dari keempat variabel kualitas air. Berdasarkan Tabel 2, standar deviasi variabel LST lebih tinggi dibandingkan dengan variabel kualitas air lainnya sebesar 2,007. Hal ini berarti bahwa jarak simpangan antara nilai-nilai observasi dengan rata-ratanya sebesar 2,007. Standar deviasi terkecil antara keempat variabel adalah 0,022 yang berarti bahwa jarak simpangan antara nilai-nilai observasi dengan rata-ratanya sebesar 0,022. Rata-rata variabel DO adalah 12,017 mg/L yang masuk dalam kategori yang baik. Selain itu, rata-rata LST adalah 25,286 Celsius. Rata-rata NDTI Danau Toba adalah -0,013 dan termasuk dalam kategori jernih. Terakhir, pH Danau Toba memiliki rata-rata 7,909 dan termasuk dalam kategori basa. Namun, pH Danau Toba masih dalam kategori aman. Temuan penelitian ini sejalan dengan penelitian-penelitian kualitas air di Danau Toba sebelumnya, baik yang menggunakan data sekunder, web Online Monitoring (ONLIMO), pembangunan sensor pada pelampung, ataupun pengambilan sampel lapangan. Grafik pergerakan keempat variabel kualitas air Danau Toba ditunjukkan pada Gambar 3.

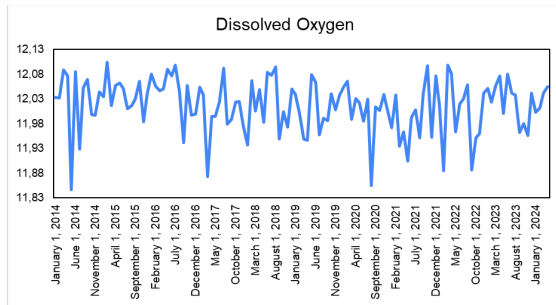
Gambar 3 menunjukkan tidak terlihatnya pola yang jelas pada keempat variabel kualitas air. Grafik memperlihatkan pola yang cukup stasioner pada variabel pH dan LST. Sedangkan variabel DO dan NDTI cukup fluktuatif. Namun, ada kecurigaan adanya seasonal baik dalam 6 bulan (per semester) atau 12 bulan (per tahun).

Dalam mengidentifikasi model ARIMA, data harus dalam keadaan stasioner dalam ragam. Pengecekan stasioneritas dalam ragam dilakukan dengan pengujian *Box-Cox*. Hasil uji membuktikan variabel LST yang memiliki λ power 2 sehingga dilakukan *Box-Cox* Transformation, yaitu kuadrat. Setelah melakukan pengecekan stasioneritas dalam ragam, dilakukan pengecekan stasioneritas dalam rata-rata menggunakan uji ADF atau *Augmented Dicky Fuller Test*. Hasil pengujian ditampilkan pada Tabel 3.

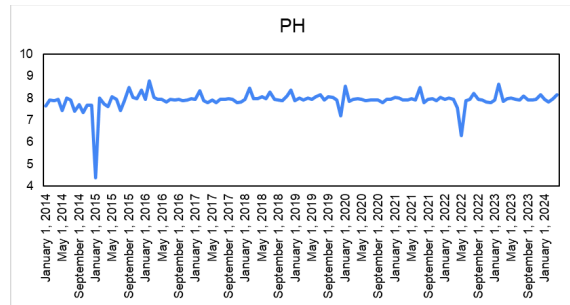
Berdasarkan hasil uji menggunakan uji ADF pada level diperoleh hasil variabel DO, NDTI, dan pH stasioner kuat pada level. Hal ini dikarenakan nilai probabilitas dari pengujian ADF kurang dari tingkat signifikansi 5% sehingga didapatkan keputusan Tolak H_0 , artinya bahwa data tidak mengandung unit root atau data telah stasioner. Sedangkan variabel LST sudah stasioner tetapi masih mengandung trend sehingga dapat dikategorikan sebagai stasioner lemah. Oleh karena itu harus



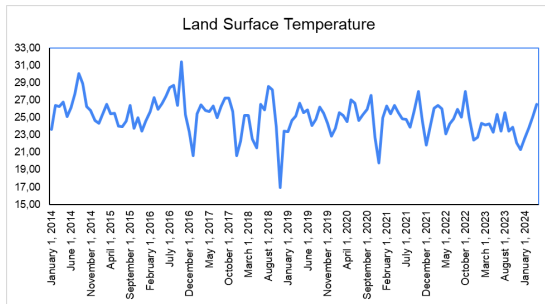
dilakukan diferensiasi. Hasil uji ADF pada *Ist difference* membuktikan bahwa variabel LST telah stasioner pada *Ist difference*. Stasioneritas pada rata-rata juga perlu diuji pada kondisi seasonal. Pengujian ADF Test untuk seasonal menghasilkan keempat variabel sudah stasioner kuat pada *level*.



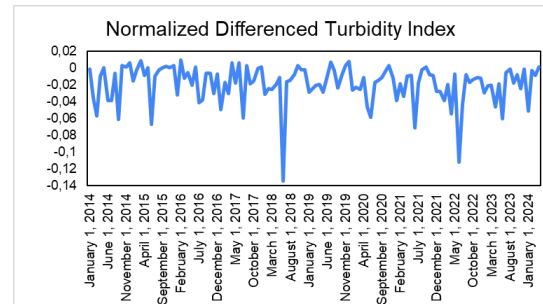
Gambar 3a.



Gambar 3b.



Gambar 3c.



Gambar 3d.

Gambar 3. Grafik Variabel Kualitas Air Danau Toba Tahun 2014-2024

(3a) *Dissolced Oxygen* (3b) pH (3c) *Land Surface Temperature* (3d) *Normaized Differenced Turbidity Index*

Tabel 3. Nilai Uji *Unit Root* dengan Metode Uji ADF Non-Seasonal Pada Tingkat *Level* dan *Ist Difference*

Indikator	Prob. ADF	Prob. Trend	Keterangan
<i>Level</i>			
DO	0,0000*	0,8686	Stasioner Kuat
pH	0,0000*	0,1371	Stasioner Kuat
LST	0,0000*	0,0112*	Stasioner Lemah
NDTI	0,0000*	0,9725	Stasioner Kuat
<i>Ist Difference</i>			
LST	0,0000*	0,8779	Stasioner Kuat

Selanjutnya, peneliti melakukan estimasi model ARIMA dan SARIMA dalam memodelkan dan meramalkan data. Berdasarkan pemilihan model ARIMA/SARIMA melalui AIC dan SIC terkecil, model terbaik untuk variabel DO adalah SARIMA(1,0,1)⁶, pH adalah ARIMA(2,0,2), LST adalah SARIMA(3,1,0)(2,0,1)⁶, dan NDTI adalah SARIMA(2,0,1)⁶. Model SARIMA terbaik pada penelitian ini menggunakan seasonal 6 bulan. Sementara itu, belum ada model prediksi yang solid untuk seasonal 12 bulan. Sejalan dengan penelitian Jerves yang menyatakan bahwa pembentukan model



yang disesuaikan dengan musim akan lebih relevan digunakan untuk analisis kualitas air [28]. Hal ini terjadi karena Indonesia merupakan negara tropis yang memiliki 2 musim yang dapat mempengaruhi kualitas air.

Setelah mendapatkan model ARIMA/SARIMA terbaik, dilakukan peramalan data testing untuk metode Prophet dan LSTM. Perbandingan metode ARIMA/SARIMA terbaik, Prophet, dan LSTM berdasarkan RMSE (*Root Mean Squared Error*) dan MAE (*Mean Absolute Error*) ditunjukkan pada Tabel 4.

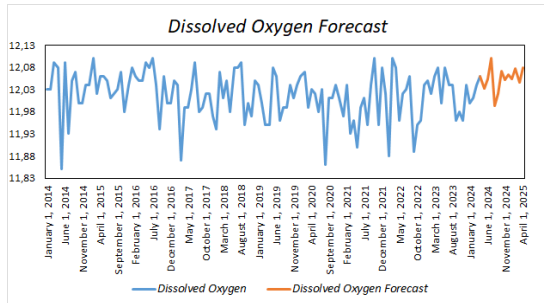
Berdasarkan Tabel 4, metode terbaik untuk meramalkan data testing variabel-variabel kualitas air bervariasi. Berdasarkan RMSE dan MAE, metode SARIMA mampu meramalkan LST dan NDTI lebih baik dibandingkan metode lain. Sedangkan metode Prophet menjadi metode terbaik dalam meramalkan DO. Lalu, metode LSTM mampu meramalkan variabel pH dengan baik. Secara rata-rata RMSE, metode ARIMA/SARIMA merupakan metode terbaik untuk meramalkan keempat variabel kualitas air.

Hasil penentuan metode terbaik variabel DO sejalan dengan penelitian Kogekar et al. (2021) dan Mehreen Ahmed et al. (2022), yaitu penggunaan metode machine learning seperti Prophet dalam meramalkan variabel DO. Namun, penelitian ini berbeda dengan penelitian oleh Shehhi dan Kaya (2021) karena LST lebih cocok untuk diramalkan menggunakan SARIMA. Hasil penelitian ini juga membenarkan temuan Wahid dan Arunbabu, 2022, bahwa ARIMA dan SARIMA masih sangat relevan digunakan untuk kualitas air. Hasil peramalan keempat variabel kualitas air 12 bulan ke depan berdasarkan metode terbaik yang ditunjukkan pada Gambar 4.

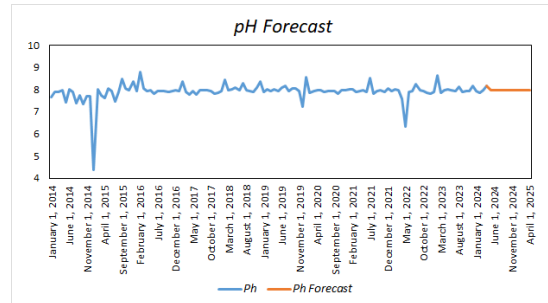
Gambar 4 menunjukkan peramalan keempat variabel kualitas air selama 12 bulan ke depan, yaitu dari bulan Mei 2024 hingga April 2025. Keempat variabel kualitas air dihasilkan dari pemilihan metode terbaik di tahapan sebelumnya. Dari keempat hasil peramalan 12 bulan ke depan, variabel DO cenderung akan naik. Namun, LST menunjukkan adanya penurunan, sedangkan variabel pH dan NDTI menunjukkan pergerakan yang stagnan.

Tabel 4. Perbandingan Metode Terbaik Data *Testing*

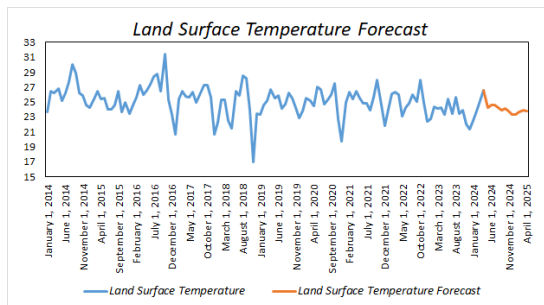
Variabel	Metrik	ARIMA/SARIMA	Prophet	LSTM
DO	RMSE	0,041410	0,034966	0,037104
	MAE	0,034902	0,029723	0,032032
pH	RMSE	0,175316	0,212455	0,105476
	MAE	0,145439	0,170091	0,081516
LST	RMSE	1,396366	1,506766	1,683863
	MAE	1,204851	1,158076	1,461292
NDTI	RMSE	0,016220	0,017027	0,019783
	MAE	0,013548	0,014453	0,017249
<i>Avg. RMSE</i>		0,407328	0,442804	0,461557



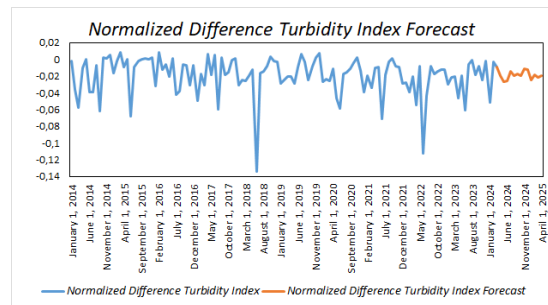
Gambar 4a.



Gambar 4b.



Gambar 4c.



Gambar 4d.

Gambar 4. Grafik Peramalan 12 Bulan Variabel Kualitas Air Danau Toba

(4a) Dissolved Oxygen (4b) pH (4c) Land Surface Temperature (4d) Normalized Differenced Turbidity Index

Penelitian ini melakukan evaluasi model melalui uji normalitas residual dan uji residual *White Noise* atau uji independensi untuk mengetahui kecocokan model. Pada uji normalitas residual, residual tidak berdistribusi normal dengan *p-value* kurang dari tingkat signifikansi 5%. Namun, distribusi residual didekati menggunakan CLT (*Central Limit Theorem*) dikarenakan jumlah observasi yang cukup banyak (124 series). Uji residual *White Noise* dilakukan dengan analisis korelogram ACF dan PACF pada pemodelan ARIMA/SARIMA. Berdasarkan uji tersebut, *residual prob.* yang dihasilkan keempat model variabel kualitas air pada pemodelan ARIMA/SARIMA sudah lebih dari 0,05 dengan tingkat signifikansi 5%. Berdasarkan hal tersebut, keempat residual model variabel kualitas air sudah dapat dikatakan *White Noise*, random, atau tidak terdapat korelasi antar residual.

V. KESIMPULAN DAN SARAN

Kesimpulan

Kualitas air Danau Toba berdasarkan variabel DO, pH, LST, dan NDTI cenderung baik dari periode Januari 2014 hingga April 2024. Pergerakan setiap variabel cenderung stasioner dan belum menunjukkan adanya trend naik atau turun. Hasil metode peramalan terbaik untuk setiap variabel kualitas air Danau Toba untuk DO adalah Prophet, pH menggunakan metode LSTM, dan SARIMA digunakan untuk meramalkan LST dan NDTI. Secara rata-rata RMSE, metode ARIMA/SARIMA merupakan metode terbaik dalam meramalkan variabel-variabel kualitas air. Model SARIMA yang digunakan dalam penelitian ini adalah dengan seasonal 6 bulan yang diasumsikan dipengaruhi oleh iklim Indonesia yang tropis. Sementara itu, belum ada model prediksi yang solid untuk *seasonal* 12 bulan. Berdasarkan peramalan 12 bulan ke depan, variabel DO cenderung akan naik. Namun, LST



menunjukkan adanya penurunan, sedangkan variabel pH dan NDTI menunjukkan pergerakan yang stagnan.

Saran

Temuan dalam penelitian dapat digunakan sebagai dasar keputusan pemerintah setempat dalam mengontrol kualitas air di Danau Toba. Selain itu, hasil penelitian ini juga dapat dijadikan acuan dalam melakukan peramalan variabel-variabel kualitas air.

Berkaitan dengan keterbatasan penelitian, yaitu variabel kualitas air yang terbatas pada dua variabel dari parameter fisika dan dua variabel dari parameter kimia non-logam, maka disarankan bagi penelitian selanjutnya dapat mempertimbangkan variabel-variabel lain yang mempengaruhi kualitas air di Danau Toba seperti kandungan klorofil, *Biochemical Oxygen Demand* (BOD), Total Suspended Solids (TSS), dan lain sebagainya. Selain itu, metode terbaik ARIMA/SARIMA yang digunakan belum memenuhi asumsi normalitas, sehingga perlu dipertimbangkan penggabungan citra satelit dan metode spasial (*geotagging*) dalam memetakan kualitas air di Danau Toba.

REFERENSI

1. W. A. Siringoringo, I. Gumilar, A. Nurhayati, and A. A. H. Suryana, “Productivity Analysis of Fish Farming in Floating Net Cages in Lake Toba (Case Study in Pangururan Subdistrict, Samosir District, Indonesia),” *Asian J. Fish. Aquat. Res.*, vol. 21, no. 5, pp. 40–48, 2023, doi: 10.9734/ajfar/2023/v21i5554.
2. I. Widianingsih *et al.*, “Sport Tourism, Regional Development, and Urban Resilience: A Focus on Regional Economic Development in Lake Toba District, North Sumatra, Indonesia,” *Sustain.*, vol. 15, no. 7, 2023, doi: 10.3390/su15075960.
3. M. Najafzadeh and S. Basirian, “Evaluation of River Water Quality Index Using Remote Sensing and Artificial Intelligence Models,” *Remote Sens.*, vol. 15, no. 9, 2023, doi: 10.3390/rs15092359.
4. N. Thai-Nghe, N. Thanh-Hai, and N. C. Ngon, “Deep learning approach for forecasting water quality in IoT systems,” *Int. J. Adv. Comput. Sci. Appl.*, vol. 11, no. 8, pp. 686–693, 2020, doi: 10.14569/IJACSA.2020.0110883.
5. H. Irwandi, M. S. Rosid, and T. Mart, “Effects of Climate change on temperature and precipitation in the Lake Toba region, Indonesia, based on ERA5-land data with quantile mapping bias correction,” *Sci. Rep.*, vol. 13, no. 1, pp. 1–11, 2023, doi: 10.1038/s41598-023-29592-y.
6. R. F. Rahmat, Athmanathan, M. F. Syahputra, and M. S. Lydia, “Real time monitoring system for water pollution in Lake Toba,” *2016 Int. Conf. Informatics Comput. ICIC 2016*, no. Icic, pp. 383–388, 2017, doi: 10.1109/IAC.2016.7905749.
7. W. Silaban and M. V. Silalahi, “Analisis Kualitas Air Di Perairan Danau Toba Kecamatan Pangururan, Kabupaten Samosir,” *JST (Jurnal Sains dan Teknol.*, vol. 10, no. 2, pp. 299–307, 2021, doi: 10.23887/jstundiksha.v10i2.39500.
8. A. A. Damayanti, H. D. Wahjono, and A. D. Santoso, “Pemantauan Kualitas Air Secara Online dan Analisis Status Mutu Air di Danau Toba, Sumatera Utara,” *J. Sumberd. Alam dan Lingkung.*, vol. 9, no. 3, pp. 113–120, 2022, doi: 10.21776/ub.jsal.2022.009.03.4.
9. G. A. L. Raja, R. Retno, and S. Sitompul, “Studi Kualitas Air Di Perairan Danau Toba Kecamatan Ajibata Kabupaten Toba,” *ARMADA J. Penelit. Multidisiplin*, vol. 1, no. 7, pp. 640–650, 2023, doi: 10.55681/armada.v1i7.657.
10. A. Abdul Wahid and E. Arunbabu, “Forecasting water quality using seasonal ARIMA model by integrating in-situ measurements and remote sensing techniques in Krishnagiri reservoir, India,” *Water Pract. Technol.*, vol. 17, no. 5, pp. 1230–1252, 2022, doi: 10.2166/wpt.2022.046.
11. A. P. Kogekar, R. Nayak, and U. C. Pati, “Forecasting of Water Quality for the River Ganga using Univariate Time-series Models,” *2021 8th Int. Conf. Smart Comput. Commun. Artif. Intell. AI Driven Appl. a Smart World, ICSCC 2021*, pp. 52–57, 2021, doi: 10.1109/ICSCC51209.2021.9528216.
12. M. R. Al Shehhi and A. Kaya, “Time series and neural network to forecast water quality parameters using satellite data,” *Cont. Shelf Res.*, vol. 231, no. July 2020, p. 104612, 2021, doi: 10.1016/j.csr.2021.104612.
13. M. Ahmed, R. Mumtaz, Z. Anwar, A. Shaukat, O. Arif, and F. Shafait, “A Multi-Step Approach for



- Optically Active and Inactive Water Quality Parameter Estimation Using Deep Learning and Remote Sensing,” 2022, doi: <https://doi.org/10.3390/w14132112>.
14. Z. Wang, Q. Wang, Z. Liu, and T. Wu, “A deep learning interpretable model for river dissolved oxygen multi-step and interval prediction based on multi-source data fusion,” *J. Hydrol.*, vol. 629, no. January, p. 130637, 2024, doi: 10.1016/j.jhydrol.2024.130637.
 15. C. Xu, P. Luo, P. Wu, C. Song, and X. Chen, “Detection of periodicity, aperiodicity, and corresponding driving factors of river dissolved oxygen based on high-frequency measurements,” *J. Hydrol.*, vol. 609, no. February, p. 127711, 2022, doi: 10.1016/j.jhydrol.2022.127711.
 16. M. G. Hutchins, Y. Qu, and M. B. Charlton, “Successful modelling of river dissolved oxygen dynamics requires knowledge of stream channel environments,” *J. Hydrol.*, vol. 603, no. PB, p. 126991, 2021, doi: 10.1016/j.jhydrol.2021.126991.
 17. L. Cui *et al.*, “A comprehensive review of the effects of salinity, dissolved organic carbon, pH, and temperature on copper biotoxicity: Implications for setting the copper marine water quality criteria,” *Sci. Total Environ.*, vol. 912, no. September 2023, p. 169587, 2024, doi: 10.1016/j.scitotenv.2023.169587.
 18. S. Singh, P. Kumar, R. Parijat, B. Gonengcil, and A. Rai, “Establishing the relationship between land use land cover, normalized difference vegetation index and land surface temperature: A case of Lower Son River Basin, India,” *Geogr. Sustain.*, vol. 5, no. 2, pp. 265–275, 2024, doi: 10.1016/j.geosus.2023.11.006.
 19. A. H. Tania, M. Y. Gazi, and M. B. Mia, “Evaluation of water quantity–quality, floodplain landuse, and land surface temperature (LST) of Turag River in Bangladesh: an integrated approach of geospatial, field, and laboratory analyses,” *SN Appl. Sci.*, vol. 3, no. 1, pp. 1–18, 2021, doi: 10.1007/s42452-020-04011-3.
 20. R. Singh, V. Saritha, and C. B. Pande, “Monitoring of wetland turbidity using multi-temporal Landsat-8 and Landsat-9 satellite imagery in the Bisalpur wetland, Rajasthan, India,” *Environ. Res.*, vol. 241, no. April 2023, p. 117638, 2024, doi: 10.1016/j.envres.2023.117638.
 21. S. Bid and G. Siddique, “Identification of seasonal variation of water turbidity using NDTI method in Panchet Hill Dam, India,” *Model. Earth Syst. Environ.*, vol. 5, no. 4, pp. 1179–1200, 2019, doi: 10.1007/s40808-019-00609-8.
 22. N. Das, R. Bhattacharjee, A. Choubey, A. K. Agnihotri, A. Ohri, and S. Gaur, “Analysing the change in water quality parameters along river Ganga at Varanasi, Mirzapur and Ghazipur using Sentinel-2 and Landsat-8 satellite data during pre-lockdown, lockdown and post-lockdown associated with COVID-19,” *J. Earth Syst. Sci.*, vol. 131, no. 2, 2022, doi: 10.1007/s12040-022-01825-0.
 23. M. C. Singh, J. Singh, and K. Sur, “Spatio–temporal water quality assessment of Chohal and Damsal dams located in Kandi region of Punjab, India by geospatial technique and on-site investigation,” *Environ. Earth Sci.*, vol. 83, no. 2, pp. 1–18, 2024, doi: 10.1007/s12665-023-11354-8.
 24. H. R. Makridakis S, Wheelwright SC, “1 / the Forecasting Perspective,” *Forecast. methods Appl.*, pp. 1–632, 1997.
 25. L. Menculini *et al.*, “Comparing Prophet and Deep Learning to ARIMA in Forecasting Wholesale Food Prices,” *Forecasting*, vol. 3, no. 3, pp. 644–662, 2021, doi: 10.3390/forecast3030040.
 26. G. Battineni, N. Chintalapudi, and F. Amenta, “Forecasting of COVID-19 epidemic size in four high hitting nations (USA, Brazil, India and Russia) by Fb-Prophet machine learning model,” *Appl. Comput. Informatics*, 2020, doi: 10.1108/aci-09-2020-0059.
 27. U. Khalil *et al.*, “Developing a Spatiotemporal Model to Forecast Land Surface Temperature: A Way Forward for Better Town Planning,” *Sustain.*, vol. 14, no. 19, 2022, doi: 10.3390/su141911873.
 28. R. Jerves-Cobo *et al.*, “Biological water quality in tropical rivers during dry and rainy seasons: A model-based analysis,” *Ecol. Indic.*, vol. 108, no. September 2019, p. 105769, 2020, doi: 10.1016/j.ecolind.2019.105769.