

## Prediksi Kualitas Air Menggunakan Metode *Seasonal Autoregressive Integrated Moving Average* (SARIMA)

### *Water Quality Prediction Using Seasonal Autoregressive Integrated Moving Average (SARIMA) Method*

Ayunita Agustin<sup>1</sup>, Faisal Fajri Rahani<sup>2\*</sup>, Fitri Indra Indikawati<sup>3</sup>

Program Studi Informatika, Universitas Ahmad Dahlan, Yogyakarta, Indonesia

\*E-mail: [faisal.fajri@tif.uad.ac.id](mailto:faisal.fajri@tif.uad.ac.id)

#### Abstrak

Pelestarian air sangat diperlukan untuk mendukung terciptanya kualitas air bersih yang terbebas dari zat berbahaya yang dapat mengganggu lingkungan. Sehingga diperlukan sistem untuk memonitor kualitas air sangat untuk mengetahui kadar pencemaran yang terjadi. Sistem ini akan bekerja untuk melihat kualitas air secara realtime dengan beberapa parameter kualitas seperti pH, temperatur, dan kekeruhan air. Tujuan dilakukannya penelitian ini ialah untuk menghasilkan model prediksi dan mengetahui hasil prediksi dari sistem berbasis data mining. Metode yang digunakan untuk melakukan prediksi kualitas air menggunakan metode *Seasonal Autoregressive Integrated Moving Average* (SARIMA), karena data kualitas air diduga mengandung pola musiman. Hasil dari penelitian ini menunjukkan bahwa model SARIMA dapat diterapkan pada dataset yang digunakan dan memperoleh keakuratan dari hasil forecasting pada masing – masing data parameter yang diuji. Hasil peramalan kualitas air dengan parameter tersebut didapatkan data hasil untuk pengujian pada dataset kedalaman 30 cm dan kedalaman 60 cm untuk parameter temperature, yaitu  $MSE < 0.1$ , dan  $RMSE < 0.02$ . Untuk parameter pH,  $MSE < 0.1$ , dan  $RMSE < 0.1$ . Serta parameter turbidity, hasil  $MSE < 0.02$ , dan  $RMSE < 0.13$ . Dari hasil tersebut menunjukkan bahwa sistem ini dapat memprediksi kualitas air dengan data yang telah lalu.

**Kata kunci:** SARIMA, Data Mining, Kualitas Air, Forecasting, Time Series

#### Abstract

Water conservation is very necessary to support the creation of clean water quality that is free from harmful substances that can disturb the environment. So a system is needed to monitor water quality to determine the level of pollution that occurs. This system will work to see water quality in real time with several quality parameters such as pH, temperature, and water turbidity. The purpose of this research is to produce a predictive model and find out the prediction results of a data mining-based system. The method used to predict water quality uses the *Seasonal Autoregressive Integrated Moving Average* (SARIMA) method, because the water quality data is thought to contain seasonal patterns. The results of this study indicate that the SARIMA model can be applied to the dataset used and obtain the accuracy of the forecasting results on each of the tested parameter data. The results of water quality forecasting with this parameter are the result data for testing at a dataset of a depth of 30 cm and a depth of 60 cm for temperature parameters, namely  $MSE < 0.1$ , and  $RMSE < 0.02$ . For pH parameters,  $MSE < 0.1$ , and  $RMSE < 0.1$ . As well as the turbidity parameter, the results of  $MSE < 0.02$ , and  $RMSE < 0.13$ . From these results indicate that this system can predict water quality with past data.

**Keywords:** SARIMA, Data Mining, Water Quality, Forecasting, Time Series

Naskah diterima 9 Sept. 2022; direvisi 18 Okt. 2022; dipublikasikan 22 Okt. 2022.

JAMIKA is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.



## I. PENDAHULUAN.

Air merupakan senyawa kimia yang sangat berlimpah di alam, akan tetapi seiring bertambahnya waktu, kebutuhan akan air juga semakin bertambah khususnya untuk budidaya perikanan. Pelestarian sumber daya air ini harus menjadi prioritas utama manusia agar terciptanya kualitas air bersih yang bebas dari zat-zat berbahaya [1],[2]. Terkait dengan terciptanya kualitas air bersih, dewasa ini banyak teknologi informasi yang dapat melakukan *forecasting* (peramalan) terhadap suatu objek.

*Forecasting* adalah analisis data *time series* yang menggunakan peristiwa dalam masa lampau untuk mengetahui perkembangan kejadian dimasa depan [3]. Dalam melakukan *forecasting*, ada beberapa model yang bisa digunakan menggunakan metode statistik antara lain seperti metode Ekonometri, Regresi, *Box Jenkins*, dan *Smoothing*. Metode *Box Jenkins* yang dikenalkan oleh G.E.P Box dan G.M Jenkins merupakan salah satu model peramalan untuk data *time series*. Metode ini terdiri dari *linear* dan stasioner, serta non

stasioner. Metode *Box Jenkins linier* dan stasioner antara lain model *Moving Average (MA)*, *Autoregressive (AR)*, dan *ARMA* [4]. Sedangkan model *ARIMA* dan *SARIMA* termasuk ke dalam metode *Box Jenkins* non stasioner. Salah satu metode dari *Box Jenkins* untuk melakukan pemodelan dan peramalan musiman khususnya prediksi kualitas air adalah metode *Seasonal Autoregressive Integrated Moving Average (SARIMA)*[5]. *SARIMA* merupakan metode *time series* yang mempunyai *seasonal*, *Moving Average (MA)*, *Autoregressive (AR)*, dan *Integrated*. Bentuk dari metode *SARIMA*, yaitu (p,d,q), (P,D,Q)S [6]. Dimana ordo p ialah bagian dari *Autoregressive (AR)* yang dapat digunakan sebagai model autokorelasi yang ada pada *time series* dengan melakukan regresi kepada variabel lag sebesar p. Selain itu, untuk merepresentasikan orde *differencing* agar data yang belum stasioner menjadi stasioner dapat menggunakan orde d. Orde q, yaitu orde dari *Moving Average (MA)* yang digunakan sebagai pemodelan dari *lagged error* sebanyak q. Orde P adalah orde dari AR musiman, dan orde D merupakan orde dari *differencing* musiman, dan orde Q merupakan orde MA musiman [7] [8]. Pada penelitian ini digunakan metode *SARIMA* untuk memprediksi hasil dimasa yang akan datang berdasarkan dengan data yang lalu.

Beberapa penelitian lalu salah satunya, yaitu membahas mengenai kulaitas air untuk budidaya ikan. Masih cukup banyak masyarakat yang belum mengetahui mengenai indeks kualitas air yang baik untuk kolam budidaya ikan mereka. Sebagian dari mereka kurang memperhatikan kualitas air kolam budidaya sehingga sangat mempengaruhi pertumbuhan dan perkembangan ikan. Kualitas air yang kurang baik dapat menimbulkan banyak permasalahan seperti masalah kesehatan ikan, kematian ikan, dan lain sebagainya. Sehingga, informasi mengenai kualitas air sangat diperlukan untuk mengetahui tingkat atau kadar pencemaran yang terjadi sehingga dapat mengurangi dampak negatif yang ditimbulkan. Prediksi kualitas air ini sangat diperlukan untuk mengetahui kesesuaian nilai parameternya terhadap persyaratan optimal kualitas air yang baik terutama pada industri budidaya ikan [9].

Terdapat beberapa parameter yang dapat digunakan untuk memprediksi kualitas air khususnya air kolam seperti derajat keasaman (pH), suhu, dan kekeruhan. Persyaratan optimal kualitas air berdasarkan Peraturan Pemerintah No 82 Tahun 2001 tentang Pengelolaan Kualitas Air dan Pengendalian Pencemaran untuk standar parameter kimia dalam budidaya ikan mensyaratkan untuk nilai pH berada diantara 6-9. Kemudian untuk suhu optimal budidaya ikan pada iklim tropis berkisar antara 28°C - 32°C dan kekeruhan optimal sebesar 2,5 NTU [10] Permenkes nomor 32 tahun 2017 tentang Standar Baku Mutu Kesehatan Lingkungan, menegaskan untuk ukuran nilai pH pada air adalah 6.5-9.5 [11]. Sedangkan untuk unsur warna persyaratan fisik air antara lain: tidak berwarna, temperatur normal, rasanya tawar, tidak berbau, jernih atau tidak keruh serta tidak mengandung zat padatan [12].

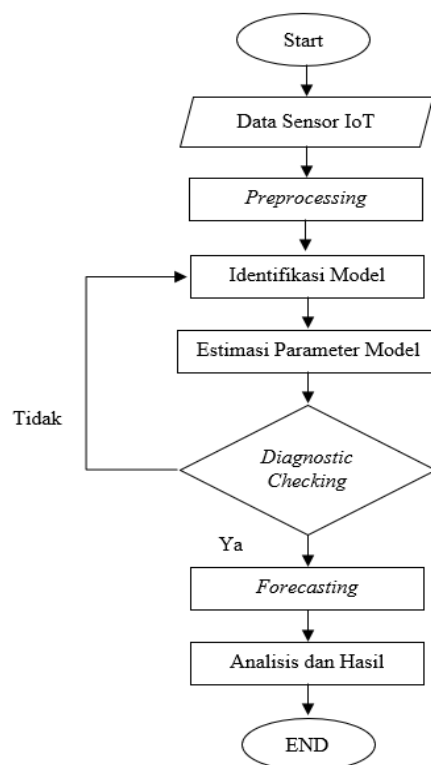
Beberapa penelitian dilakukan untuk prediksi kualitas air. Penelitian ini dilakukan untuk membuat sebuah sistem prediksi kualitas air dan untuk mendukung dilakukannya pemrosesan prediksi, yaitu dengan menerapkan algoritma klasifikasi data mining, yaitu adalah algoritma *K-Nearest Neighbor*. Data yang diperoleh berasal dari website kaggle yang mana akan digunakan untuk percobaan pada sistem yang akan bangun. Hasil pemodelan diukur menggunakan tabel *Confusion Matrix* sehingga dapat dihitung nilai akurasi dari performa model [13]. Penelitian berkaitan dengan sistem pemantauan kualitas air dari sensor berbasis IoT yang digunakan untuk memprediksi kualitas air pada kolam ikan dengan menggunakan metode *Support Vector Machine (SVM)*. Namun, hasil penelitian dengan menggunakan metode ini *errornya* masih cukup tinggi [14]. Oleh karena itu, perlu dilakukan penelitian lebih lanjut menggunakan metode yang berbeda dengan tingkat akurasi yang lebih tinggi. Penelitian lain membahas prediksi air dengan *Adaptive Neuro Fuzzy Inference System*. Kekeruhan adalah salah satu indikator mutu air, semakin tinggi tingkat kekeruhannya maka semakin rendah indeks kualitas air. Dalam makalah ini, metode *adaptive neuro fuzzy inference system (ANFIS)* digunakan untuk memprediksi tingkat kekeruhan air sebagai satu indikator kualitas air baku. Kadar pH, spektro warna dan daya hantar listrik digunakan sebagai variabel masukan. Sebanyak 587 data dikumpulkan selama bulan Januari-Juni digunakan sebagai data pelatihan dan 144 data dari pengamatan di Bulan Juli digunakan untuk memvalidasi model. Proses validasi menunjukkan bahwa selama proses pelatihan tidak terjadi *overfitting*. Penerapan model ANFIS memberikan 27 rules yang menunjukkan hubungan prediksi kekeruhan air untuk tingkata ph air, spektro warna, dan daya hantar listrik yang berbeda [15]. Penelitian lain dilakukan untuk mengetahui hasil evaluasi dari model yang dihasilkan untuk dapat memprediksi kualitas air yang dapat dikonsumsi atau tidaknya dengan menerapkan algoritma klasifikasi *data mining*, yaitu adalah algoritma *K-Nearest Neighbor*. Algoritma ini diterapkan untuk menghitung probabilitas kualitas air yang aman atau tidak untuk dikonsumsi berdasarkan data rekaman yang diambil dari lingkungan sekitar terutama di daerah padat penduduk. Kumpulan data diperoleh dari website kaggle, hasil pemodelan diukur menggunakan tabel *Confusion Matrix* untuk menghitung akurasi. Setelah diuji, model ini memiliki tingkat akurasi tertinggi 85,52% dengan nilai k (tetangga terdekat) = 3 [16].

Ruhat dan Effendi [11], menggunakan metode SARIMA untuk melakukan kajian mengenai pengaruh faktor musiman kemampuan model untuk meniru dan memprediksi aktivitas dari data debit sungai [17]. Selain metode SARIMA dalam *data mining*, terdapat juga metode *deep learning* dan *artificial neural network* dalam *machine learning*. Penerapan metode *deep learning* dengan jumlah *layer* sebanyak 200, dan perbandingan antara data *training* dan data *testing* sebesar 9:1 untuk melakukan prediksi cuaca dengan beberapa parameter seperti pengukuran kelembapan udara, kecepatan angin, tekanan udara, dan suhu udara [18]. Hasil dari penelitiannya menunjukkan bahwa nilai RMSE dari seluruh validasi parameter cuaca memperoleh hasil yang lebih baik saat digunakan metode LSTM *update*. Jumlah data yang besar merupakan salah satu karakteristik dari metode *deep learning*, sehingga apabila diterapkan pada data yang jumlahnya kecil akan kurang cocok menggunakan metode ini.

Berdasarkan permasalahan yang telah dijelaskan sebelumnya, maka penelitian yang dilakukan adalah untuk memprediksi kualitas air dalam beberapa langkah ke depan menggunakan metode SARIMA yang mana datanya diambil dari data *public*, yaitu dalam dataset [14] agar memudahkan masyarakat untuk melihat kualitas air melalui nilai parameter dan grafik yang dihasilkan. Sehingga masyarakat akan tepat dalam memprediksi kualitas air untuk keperluan sehari-hari. Metode ini digunakan karena data yang digunakan mempunyai pola musiman dan jumlah data yang digunakan masih tergolong kecil, yaitu sebanyak 9.623 dataset. Selain itu, terdapat keunggulan yang dimiliki metode SARIMA seperti fleksibel dan bisa menggambarkan berbagai karakter deret waktu yang terjadi, dalam pengujian kesesuaian model terdapat prosedur yang formal, serta interval ramalan dan prediksi telah mengikuti model. Sehingga metode ini lebih unggul dibanding dengan metode lainnya. Melalui penelitian ini nantinya bisa didapatkan hasil dengan tingkat akurasi yang lebih baik dari sebelumnya sehingga dapat membantu masyarakat untuk menentukan kesesuaian lingkungan daerah setempat untuk budidaya ikan maupun untuk mendeteksi anomali dan memperkirakan kondisi lingkungan perairan dimasa yang akan datang.

## II. METODE PENELITIAN

Metode yang digunakan pada penelitian ini adalah studi literatur dan dataset statistik, dimana data yang digunakan ialah data *public* dalam situs web [14]. Tahapan perancangan sistem yang akan dilakukan dengan pembuatan bagan alir (*flowchart*) yang memuat tahapan-tahapan proses dari sistem yang akan dibangun. Adapun *flowchart* sistem dapat dilihat pada Gambar 1.



Gambar 1. Daigram Proses Penelitian

Berdasarkan Gambar 1, maka dapat diuraikan proses *forecasting* dengan menggunakan metode SARIMA, yaitu :

1. Data Sensor IoT

Pada penelitian ini terdapat dua data yang digunakan, yaitu data pada kedalaman 30 cm dan 60 cm. Data kedalaman 30 cm memiliki tiga parameter, yaitu temperatur, pH, dan *turbidity* (kekeruhan). Sedangkan data kedalaman 60 cm hanya memiliki dua parameter, yaitu temperatur dan *turbidity* (kekeruhan). Pada penelitian ini menggunakan data sebanyak 9.623 dataset.

2. *Preprocessing*

Setelah data dikumpulkan, langkah selanjutnya, yaitu melakukan *preprocessing*. Proses ini bertujuan untuk menyiapkan data mentah yang nantinya akan dijadikan data untuk input pemodelan sehingga model yang akan dihasilkan hasilnya berkualitas [19].

3. Identifikasi Model

Dalam melakukan identifikasi model, terdapat beberapa langkah yang harus dilakukan seperti :

a. Membuat plot *time series* dan melakukan uji stasioneritas data dalam *mean* dan *variance*. Apabila data *time series* tidak stasioner terhadap *mean*, maka dapat distasionerkan dengan proses *differencing*. Jika data *time series* tidak stasioner dalam *variance*, maka dapat distasionerkan melalui transformasi *Box Cox* [20].

b. *Dickey-Fuller Test*

Selain membuat plot *time series* untuk melakukan uji stasioneritas, dapat juga dilakukan dengan menggunakan *Dickey-Fuller Test* untuk melihat nilai uji statistik apakah lebih rendah dari nilai kritis atau tidak. Serta dengan melihat *p-value* apakah nilainya lebih kecil dari 0.05 sehingga data *time series* dapat dikatakan stasioner [21].

c. Plot ACF dan PACF

Setelah data stasioner, langkah selanjutnya, yaitu melakukan identifikasi data yang berpola musiman dengan cara mengidentifikasi plot ACF dan plot PACF untuk menentukan orde pada model yang digunakan.

4. Estimasi Parameter Model

Pada tahap ini dilakukan estimasi parameter MA, AR, *seasonal*, *non seasonal*, dan melakukan uji signifikansi dari tiap parameternya. Apabila terdapat model yang parameternya tidak signifikan maka dinyatakan tidak lolos uji. Agar didapatkan model yang parameternya signifikan, maka tiap parameter yang tidak signifikan akan dieliminir.

5. *Diagnostic Checking*

Pengujian *Diagnostic Checking* pada penelitian ini, yaitu untuk mengetahui model yang digunakan tersebut apakah sudah sesuai dan layak untuk dilakukan *forecasting*. Pengujian ini dilakukan dengan cara melakukan uji *diagnostic white noise* dan uji distribusi normal menggunakan *diagnostic model* [22].

6. *Forecasting*

Melakukan peramalan (*forecasting*) menggunakan model SARIMA pada data *time series* untuk mengidentifikasi pola-pola pada data historis masa lalu untuk memperkirakan variabel di masa depan [23]. Berikut merupakan persamaan model SARIMA (p,d,q) (P,D,Q)s dengan *mean* dimana p,d,q,P,D, dan Q adalah bilangan bulat dan (s) adalah *periodicity* yang dapat dilihat pada persamaan (1).

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

Dimana :

$Z_t$	= data deret waktu yang mempunyai rata-rata $\mu$
$\phi_p(B)$	= polinomial $AR(p)$
$\theta_q(B)$	= polinomial $MA_q$
$\phi_p(B^s)$	= polinomial musiman $AR(P)$
$\theta_q(B^s)$	= polinomial musiman $MA(q)$
$(1-B^s)^D$	= operator pembeda musiman periode S
$(1-B)^d$	= operator pembeda non musiman
$\varepsilon_t$	= <i>white noise</i>

Persamaan di atas merupakan model SARIMA (p,d,q) (P,D,Q)s. Dimana p,d,q dan P,D,Q adalah bilangan bulat dan (s) adalah *periodicity* atau jangka waktu musiman. Untuk notasi  $\phi_p(B)$ ,  $\theta_q(B)$ ,  $\phi_p(B^s)$ , dan  $\theta_q(B^s)$  masing-masing merupakan persamaan polinomial dari  $AR_{(p)}$ ,  $MA_q$ ,  $AR(P)$ , dan  $MA(Q)$ .

#### 7. Analisis dan Hasil

Melakukan analisis dari hasil *forecasting* yang sudah dilakukan. Hasil *forecasting* nantinya berupa grafik prediksi kualitas air pada kolam ikan di dalam Gedung Khan Jahal Universitas Khulna untuk masa yang akan datang.

Tahap terakhir dari penelitian ini adalah implementasi sistem dan melakukan pengujian terhadap sistem. Implementasi merupakan penerapan hasil rancangan sistem ke dalam program yang akan dibuat. Program akan diimplementasikan menggunakan bahasa pemrograman *python* dengan bantuan aplikasi *jupyter notebook*. Untuk pengujian sistem menggunakan MSE dan RMSE. Pengujian ini dilakukan dengan tujuan untuk menganalisis sistem apakah sudah berjalan dengan baik atau masih membutuhkan pembaharuan agar sistem yang dibuat tersebut dapat sesuai dengan kebutuhan [24] [25]. Adapun untuk MSE dan RMSE dapat dihitung dengan persamaan (2) dan (3).

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

Dimana :

$A_t$  = nilai data asli

$F_t$  = nilai data peramalan (forecasting)

n = banyaknya data

$$RMSE = \sqrt{\sum_{t=1}^n \frac{(A_t - F_t)^2}{n}} \quad (3)$$

Dimana :

$A_t$  = nilai data asli

$F_t$  = nilai data peramalan (forecasting)

n = banyaknya data

*Mean Squared Error* (MSE) merupakan rata-rata kesalahan kuadrat antara nilai sebenarnya dengan nilai peramalan. Sedangkan *Root Mean Squared Error* (RMSE) berfungsi untuk memperoleh besaran tingkat kesalahan dari hasil peramalan [26]. Baik MSE maupun RMSE, apabila hasil yang diperoleh semakin kecil atau mendekati nilai nol maka hasil prediksi akan semakin akurat. Apabila hasil yang diperoleh sama dengan nol, maka hasil prediksi sempurna.

### III. HASIL DAN PEMBAHASAN

Pada bagian ini akan dibahas hasil dari penelitian yang terdiri dari grafik hasil prediksi, nilai prediksi, dan pengujian sistem.

#### *Data Sensor IoT*

Data yang digunakan pada penelitian ini terdapat dua data, yaitu data pada kedalaman 30 cm dan data pada kedalaman 60 cm. Dimana di dalam data tersebut terdapat beberapa parameter seperti temperatur, pH, dan *turbidity* (kekeruhan).

##### a. Data Kedalaman 30 cm

Dataset pada kedalaman 30 cm memiliki variabel *date and time*, *temperature* (°C), pH, dan *turbidity* (NTU). Masing-masing variabel memiliki jumlah data sebanyak 9623 data yang dapat dilihat pada Tabel 1.

TABEL 1  
 DATASET KEDALAMAN 30 CM

No	Date and Time	Temperatur (°C)	pH	Turbidity (NTU)
1	2020-01-15 16:00:35	20,99	7,81	197
2	2020-01-15 16:01:33	20,99	7,81	197
3	2020-01-15 16:02:33	20,98	7,81	197
4	2020-01-15 16:03:22	20,98	7,81	197
5	2020-01-15 16:04:26	20,98	7,81	197

No	Date and Time	Temperatur (°C)	pH	Turbidity (NTU)
6	2020-01-15 16:05:21	20,97	7,81	197
7	2020-01-15 16:06:20	20,97	7,81	197
8	2020-01-15 16:07:17	20,95	7,81	197
9	2020-01-15 16:08:59	20,94	7,81	197
10	2020-01-15 16:10:00	20,93	7,81	197
...	.....	.....	.....	.....
9623	2020-01-22 16:25:32	21,31	7,51	212

b. Dataset Kedalaman 60 cm

Variabel yang terdapat pada dataset di kedalaman 60 cm antara lain *date and time*, *temperature* (°C), dan *turbidity* (NTU). Jumlah data pada masing-masing variabel sama dengan dataset di kedalaman 30 cm, yaitu sebanyak 9623 dataset yang dapat dilihat pada Tabel 2.

TABEL 2  
DATASET KEDALAMAN 60 CM

No	Date and Time	Temperatur (°C)	Turbidity (NTU)
1	2020-01-15 16:00:35	22,54	134
2	2020-01-15 16:01:33	22,54	134
3	2020-01-15 16:02:33	22,54	134
4	2020-01-15 16:03:22	22,54	134
5	2020-01-15 16:04:26	22,54	134
6	2020-01-15 16:05:21	22,54	134
7	2020-01-15 16:06:20	22,54	134
8	2020-01-15 16:07:17	22,54	134
9	2020-01-15 16:08:59	22,54	134
10	2020-01-15 16:10:00	22,54	134
...	.....	.....	...
9623	2020-01-22 16:25:32	22,02	152

**Preprocessing**

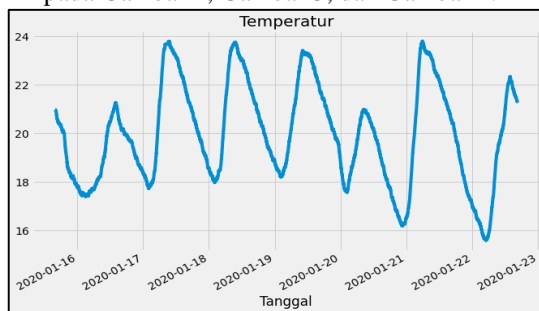
Pada tahap *preprocessing*, baik data kedalaman 30 cm maupun data kedalaman 60 cm akan dilakukan dua proses, yaitu *data cleaning* dan transformasi data. Pada proses *data cleaning* bertujuan untuk mengatasi *missing value* pada data. Selain itu, pada tahap *data cleaning* ini juga dilakukan penggantian nama kolom yang terlalu panjang dengan nama kolom yang lebih pendek dan mudah untuk dilakukan pemodelan. Sedangkan pada proses transformasi data, tipe data diubah ke dalam bentuk lain sehingga data *time series* tersebut sesuai untuk dilakukan proses *data mining*. Hasil dari *transformasi* data, yaitu tipe data Tanggal diubah menjadi tipe *datetime*, dan membuat indeks *datetime* agar proses manipulasi data menjadi lebih mudah.

**Identifikasi Model**

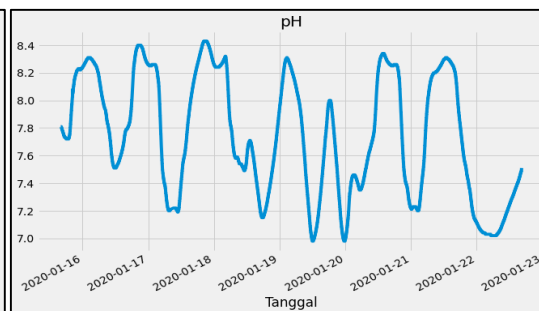
Pada tahap identifikasi model ini ada beberapa proses yang akan dilakukan seperti membuat plot *time series*, uji stasioner data, dan membuat plot ACF dan PACF.

a. Plot Time Series

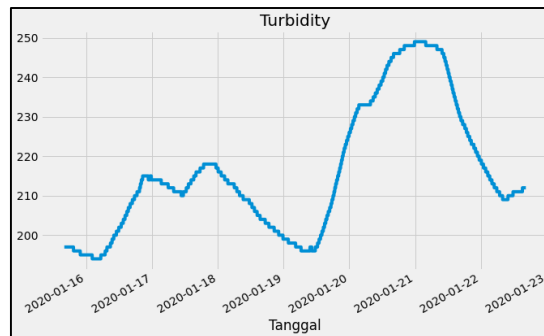
Setelah dilakukan *preprocessing*, langkah selanjutnya, yaitu melakukan *plotting* terhadap data *time series* diantaranya Temperatur, pH, dan kekeruhan (*turbidity*) pada masing-masing dataset untuk melihat pola *seasonal* pada data tersebut. Plot *time series* masing-masing parameter di kedalaman 30 cm dapat dilihat pada Gambar 2, Gambar 3, dan Gambar 4.



Gambar 2. Plot Parameter Temperatur Kedalaman 30 cm



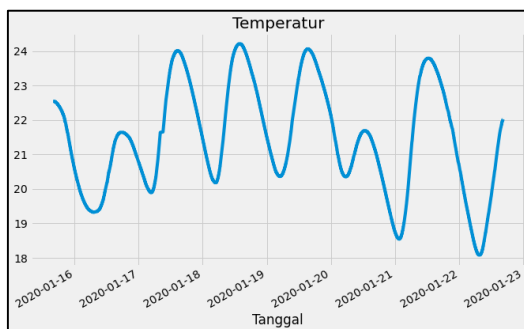
Gambar 3. Plot Parameter pH Kedalaman 30 cm



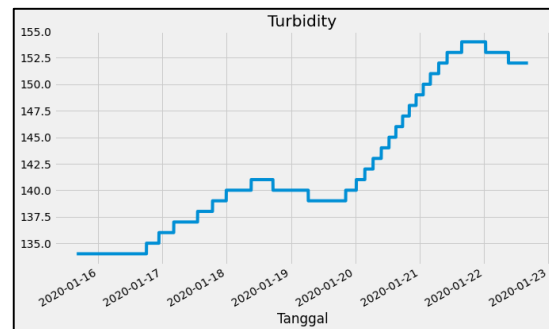
Gambar 4. Plot Parameter *Turbidity* Kedalaman 30 cm

Berdasarkan plot *time series* Gambar 2 dan Gambar 3 menunjukkan bahwa data memiliki tren dan pola *seasonal* yang kuat ditandai dengan pola berulang pada lag dan periode tertentu. Sedangkan Gambar 4 terdapat tren naik dan turun serta pola *seasonal*.

Sedangkan plot *time series* masing-masing parameter di kedalaman 60 cm dapat dilihat pada Gambar 5 dan Gambar 6. Pada Gambar 5 plot *time series* pada parameter temperatur menunjukkan bahwa data memiliki tren dan juga pola *seasonal*. Sedangkan pada Gambar 6 terlihat bahwa data memiliki tren naik serta pola *seasonal*. Berdasarkan hasil tersebut, baik data pada kedalaman 30 cm maupun 60 cm dapat disimpulkan bahwa masing-masing parameter memiliki tren dan pola *seasonal*.



Gambar 5. Plot Parameter Temperatur Kedalaman 60 cm



Gambar 6. Plot Parameter *Turbidity* Kedalaman 60 cm

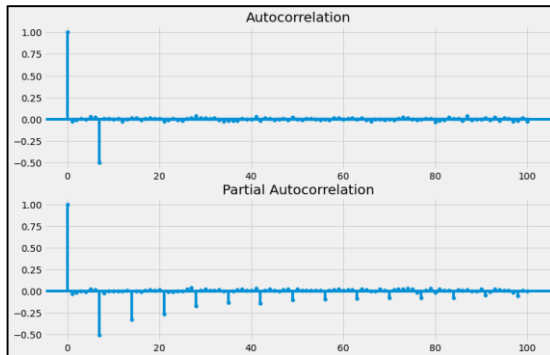
#### b. Uji Stasioner

Setelah dilakukan plot *time series* pada masing – masing parameter, langkah berikutnya adalah melakukan uji stasioner data menggunakan uji *Dickey-Fuller*. Hasil uji *Dickey-Fuller* pada data kedalaman 30 cm parameter temperatur didapatkan *p-value* 0,000031, parameter pH dengan *p-value* 0,000156, dan parameter *turbidity* (kekeruhan) *p-value* sebesar 0,695215. Berdasarkan hasil tersebut, parameter temperatur dan pH datanya telah stasioner karena *p-value* lebih kecil dari  $\alpha = 0,05$ . Sedangkan parameter *turbidity* datanya belum stasioner karena *p-value* lebih besar dari  $\alpha = 0,05$  sehingga perlu dilakukan *differencing* agar data benar-benar stasioner.

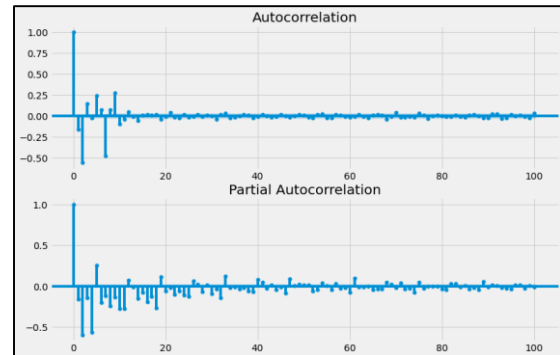
Kemudian pada data kedalaman 60 cm juga dilakukan hal yang sama seperti data kedalaman 30 cm. Hasil yang diperoleh dari uji *Dickey-Fuller*, yaitu pada parameter temperatur *p-value* sebesar 5,533976 dan pada parameter *turbidity* *p-value* sebesar 0,973980. Berdasarkan kedua hasil tersebut, maka dapat disimpulkan bahwa parameter temperatur dan *turbidity* datanya masih belum stasioner sehingga harus dilakukan *differencing* terlebih dahulu agar data stasioner.

#### c. Plot ACF dan PACF

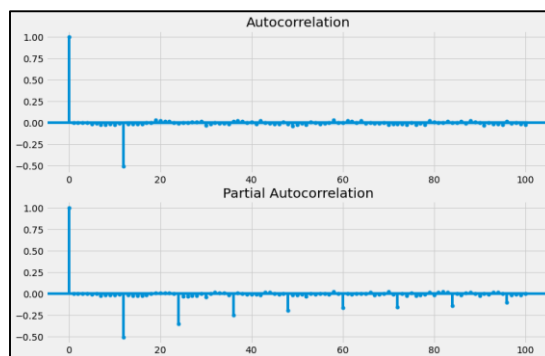
Selain menggunakan uji *Dickey-Fuller* seperti proses yang telah dilakukan sebelumnya, untuk mengetes apakah data sudah stasioner juga dapat dilakukan dengan melihat plot ACF dan PACF. Jika deret waktu stasioner, maka plot ACF/PACF akan menunjukkan penurunan korelasi yang cepat setelah sedikit jeda antar titik. Berikut merupakan plot ACF dan PACF pada data kedalaman 30 cm yang dapat dilihat pada Gambar 8, Gambar 9, dan Gambar 10.



Gambar 7. Plot ACF dan PACF Parameter Temperatur Kedalaman 30 cm

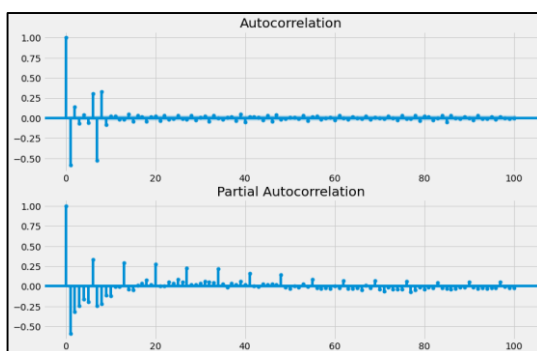


Gambar 8. Plot ACF dan PACF Parameter pH Kedalaman 30 cm

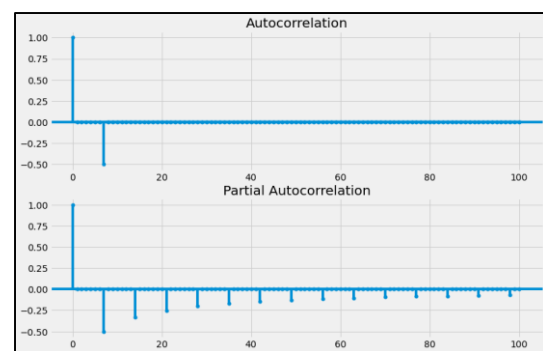


Gambar 9. Plot ACF dan PACF Parameter *Turbidity* Kedalaman 30 cm

Dari plot ACF dan PACF diatas, Gambar 8 menunjukkan bahwa nilai ACF dan PACF signifikan pada lag 1, 7 dan 14 yang menandakan data mempunyai pola *seasonal*. Lalu Gambar 9, yaitu plot ACF dan PACF parameter pH terlihat bahwa data mempunyai pola *seasonal* karena nilai ACF dan PACF signifikan pada lag ke 1, 5, dan 12. Sedangkan Gambar 10 ialah parameter *turbidity* yang hasilnya nilai ACF dan PACF signifikan pada lag 1,7, dan 14 sehingga cocok dengan model yang akan digunakan, yaitu model SARIMA.



Gambar 10. Plot ACF dan PACF Parameter Temperatur Kedalaman 60 cm



Gambar 11. Plot ACF dan PACF Parameter *Turbidity* Kedalaman 30 cm

Kemudian plot ACF dan PACF pada data kedalaman 60 cm dapat dilihat pada Gambar 10. Gambar 10 menunjukkan bahwa nilai ACF dan PACF signifikan pada lag 1, 6, dan 13. Dan pada Gambar 11 hasilnya nilai ACF dan PACF signifikan pada lag 1, 8, dan 15. Berdasarkan hasil tersebut, maka dapat disimpulkan data kedalaman 60 cm mempunyai pola *seasonal* dan dapat dilanjutkan tahapan berikutnya yaitu estimasi parameter model.



### Estimasi Parameter Model

Pada tahap estimasi parameter model akan dibuat fungsi untuk melakukan pemodelan SARIMA dengan notasi modelnya adalah SARIMA (p,d,q) (P,D,Q)s. Ketiga parameter ini digunakan untuk memperhitungkan musim, tren, dan *noise* dalam data. Dimana p adalah AR non musiman, d yaitu *differencing* non musiman, q yaitu MA non musiman. Sedangkan P ialah AR musiman, D ialah *differencing* musiman, Q ialah MA musiman, dan s adalah periode atau waktu.

TABEL 3  
 DATASET KEDALAMAN 30 CM

Parameter	Model SARIMA (p,d,q) (P,D,Q)s	AIC
Temperatur	SARIMAX (1, 1, 1)x(0, 0, 1) <sup>7</sup>	-52569.829
pH	SARIMAX (1, 0, 1)x(0, 0, 1) <sup>7</sup>	-80860.257
Turbidity	SARIMAX (1, 0, 0)x(1, 0, 1) <sup>7</sup>	-12171.691

TABEL 4  
 DATASET KEDALAMAN 60 CM

Parameter	Model SARIMA (p,d,q) (P,D,Q)s	AIC
Temperatur	SARIMAX (1, 1, 1)x(1, 0, 1) <sup>7</sup>	-79515.830
Turbidity	SARIMAX (1, 0, 0)x(0, 0, 1) <sup>7</sup>	-29499.872

Tabel 3 dan Tabel 4 merupakan model SARIMA terbaik dari masing – masing data dengan nilai AIC terkecil yang sebelumnya telah dilakukan pencarian model terbaik menggunakan fungsi *auto* SARIMA. Dari semua model yang terpilih tersebut telah memenuhi standar kriteria, yaitu *p-value* <  $\alpha = 0,05$ . Selanjutnya model tersebut dapat dilakukan *diagnostic checking* untuk mengecek keakuratan model sampai dapat dilakukan *forecasting*.

### Diagnostic Checking

Pada tahap ini, semua model terbaik dilakukan pengujian white noise dan distribusi normal yang fungsinya untuk memastikan bahwa tidak ada asumsi yang dilanggar oleh model yang digunakan. Apabila semua model yang terpilih telah *white noise* dan terdistribusi normal, maka model tersebut dapat dilanjutkan proses *forecasting*. Tabel 5 dan Tabel 6 merupakan hasil *diagnostic checking* pada masing-masing dataset dan model terbaik yang hasilnya menunjukkan bahwa semua model telah *white noise* dan terdistribusi normal. Pada Tabel 5 pengujian menggunakan MSE mendapatkan akurasi yang sempurna dan pengujian menggunakan RMSE mendapatkan nilai yang baik karena hasilnya mendekati nol. Semakin rendah nilai MSE maupun RMSE, maka hasil pengujian *forecasting* mendapatkan akurasi yang baik dan nilai *error* yang rendah.

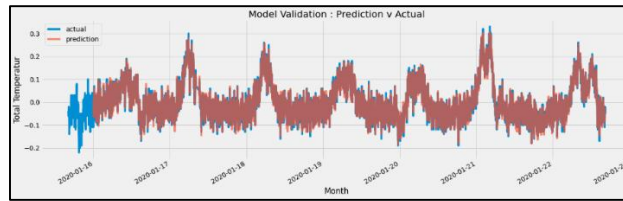
TABEL 5  
 DIAGNOSTIC CHECKING DATA KEDALAMAN 30 CM DAN 60 CM

Parameter	Model SARIMA (p,d,q) (P,D,Q)s	Uji White Noise	Uji Distribusi Normal
Data Kedalaman 30 cm			
Temperatur	SARIMAX (1, 1, 1)x(0, 0, 1) <sup>7</sup>	White Noise	Normal
pH	SARIMAX (1, 0, 1)x(0, 0, 1) <sup>7</sup>	White Noise	Normal
Turbidity	SARIMAX (1, 0, 0)x(1, 0, 1) <sup>7</sup>	White Noise	Normal
Data Kedalaman 60 cm			
Temperatur	SARIMAX (1, 1, 1)x(1, 0, 1) <sup>7</sup>	White Noise	Normal
Turbidity	SARIMAX (1, 0, 0)x(0, 0, 1) <sup>7</sup>	White Noise	Normal

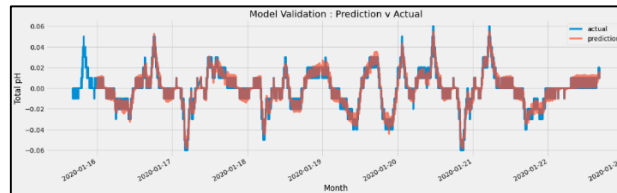
### Forecasting

Setelah dilakukan *diagnostic checking*, maka langkah selanjutnya, yaitu melakukan peramalan terhadap masing-masing model yang didapatkan dengan dilakukan perbandingan antara nilai prediksi dengan nilai nyata (aktual) dari data *time series* yang dapat membantu keakurata prakiraan.

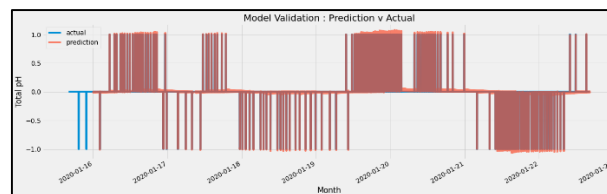
Grafik *forecasting* pada masing-masing parameter kedalaman 30 cm dapat dilihat pada Gambar 12- Gambar 14, dan parameter kedalaman 60 cm dapat dilihat pada Gambar 15 dan Gambar 16. Yang mana kelima grafik tersebut menunjukkan bahwa hasilnya selaras dengan nilai sebenarnya. Dalam hal ini berguna untuk mengukur keakuratan hasil *forecasting* yang nantinya akan dilakukan pada proses pengujian menggunakan *MSE* dan *RMSE*.



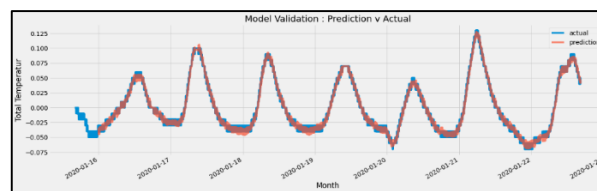
Gambar 12. Grafik Forecasting Parameter Temperatur Kedalaman 30 cm



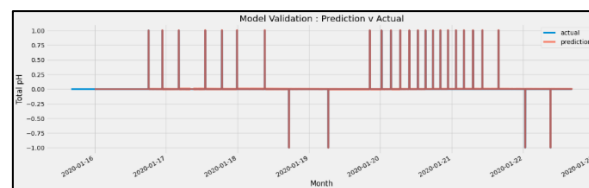
Gambar 13. Grafik Forecasting Parameter pH Kedalaman 30 cm



Gambar 14. Grafik Forecasting Parameter Turbidity Kedalaman 30 cm



Gambar 15. Grafik Forecasting Parameter Temperatur Kedalaman 60 cm



Gambar 16. Grafik Forecasting Parameter Turbidity Kedalaman 60 cm

Setelah dilakukan peramalan pada masing-masing parameter, selanjutnya dilakukan peramalan/prediksi untuk 7 hari kedepan. Pada penelitian ini, peramalan dilakukan dalam kurun waktu harian karena data yang digunakan merupakan data harian. Hasil yang diperoleh dapat dilihat pada Tabel 6 dan untuk grafik hasil peramalan 7 hari ke depan dapat dilihat pada Gambar 17-21.

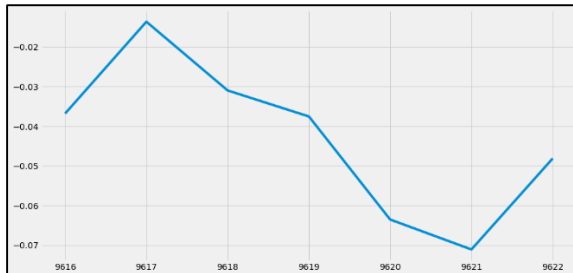
TABEL 6	
HASIL FORECASTING 7 HARI KE DEPAN	
Model SARIMA (p,d,q) (P,D,Q)s	Hasil Forecasting
Data Kedalaman 30 cm	
SARIMAX (1, 1, 1)x(0, 0, 1) <sup>7</sup>	-0,036797

Model SARIMA (p,d,q) (P,D,Q)s	Hasil <i>Forecasting</i>
	-0,013616
	-0,030972
	-0,037541
	-0,063560
	-0,071118
	-0,048151
SARIMAX (1, 0, 1)x(0, 0, 1) <sup>7</sup>	0,018377
	0,022596
	0,022596
	0,022557
	0,022537
	0,015551
	0,010446
SARIMAX (1, 0, 0)x(1, 0, 1) <sup>7</sup>	-0,006255
	-0,000971
	-0,010951
	-0,004961
	-0,009934
	-0,022124
	-0,027970
Data Kedalaman 60 cm	
SARIMAX (1, 1, 1)x(1, 0, 1) <sup>7</sup>	0,041061
	0,035475
	0,042415
	0,040671
	0,042946
	0,042538
	0,044052
SARIMAX (1, 0, 0)x(0, 0, 1) <sup>7</sup>	0,001601
	0,004659
	0,009173
	0,010771
	0,011643
	0,013243
	0,014115

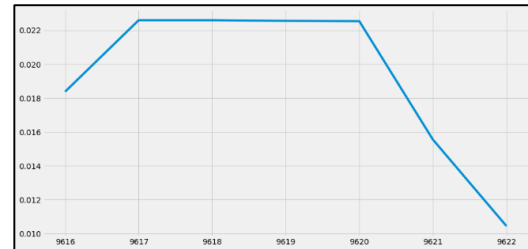
Berdasarkan Tabel 6 dan Gambar 18, parameter temperatur pada kedalaman 30 cm terlihat bahwa kualitas air mengalami kenaikan di satu hari pertama, kemudian pada hari kedua sampai hari keenam mengalami penurunan, dan pada hari ketujuh mengalami kenaikan. Prediksi tersebut menunjukkan bahwa parameter temperatur untuk 7 hari ke depan mengalami kenaikan dan penurunan yang signifikan. Kemudian pada Gambar 19, parameter pH menunjukkan kenaikan pada 3 hari pertama dan penurunan pada hari keempat sampai dengan hari ketujuh. Dari grafik tersebut dapat disimpulkan bahwa setelah hari ketujuh dimungkinkan parameter pH akan mengalami penurunan. Sedangkan Gambar 20, yaitu parameter *turbidity* mengalami kenaikan dan penurunan. Namun, pada hari kelima sampai hari ketujuh, mengalami penurunan yang signifikan. Oleh karena itu dapat disimpulkan bahwa untuk hari-hari berikutnya, parameter *turbidity* mungkin akan mengalami penurunan sehingga kualitas air akan semakin jernih.

Disisi lain, data pada kedalaman 60 cm parameter temperatur yang terlihat pada Gambar 21 menunjukkan bahwa kualitas air mengalami penurunan dan kenaikan. Pada hari pertama, parameter temperatur mengalami penurunan, kemudian dihari kedua mengalami kenaikan yang lumayan signifikan. Selanjutnya untuk hari ketiga sampai hari ketujuh mengalami penurunan dan kenaikan. Berdasarkan prediksi tersebut dapat

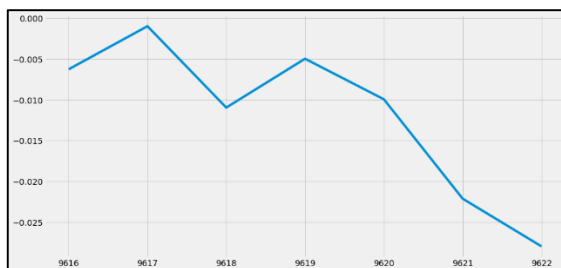
disimpulkan bahwa parameter temperatur untuk dalam jangka waktu 7 hari ke depan mengalami penurunan disertai dengan kenaikan temperatur air.



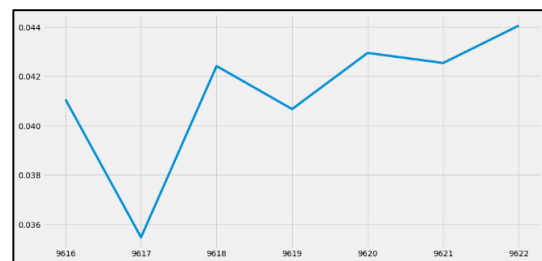
Gambar 17. Grafik Forecasting 7 Hari Parameter Temperatur Kedalaman 30 cm



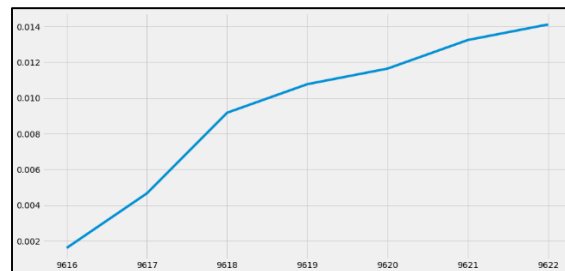
Gambar 18. Grafik Forecasting 7 Hari Parameter pH Kedalaman 30 cm



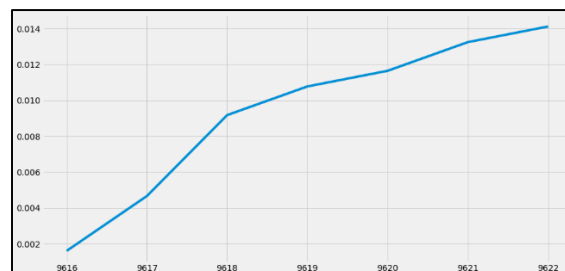
Gambar 19. Grafik Forecasting 7 Hari Parameter Turbidity Kedalaman 30 cm



Gambar 20. Grafik Forecasting 7 Hari Parameter Temperatur Kedalaman 60 cm



Gambar 21. Grafik Forecasting 7 Hari Parameter Turbidity Kedalaman 60 cm



Gambar 22. Grafik Forecasting 7 Hari Parameter Turbidity Kedalaman 60 cm

Sedangkan pada Gambar 22, parameter *turbidity* cenderung mengalami kenaikan dari hari pertama hingga hari ketujuh. Sehingga dari hasil tersebut dapat disimpulkan bahwa untuk prediksi 7 hari ke depan, parameter *turbidity* mengalami kenaikan secara signifikan yang mengakibatkan kadar kekeruhan air tinggi pada kedalaman 60 cm.

### Analisis dan Hasil

Tahap terakhir pada penelitian ini adalah analisis hasil yang di dalamnya berisi pengujian sistem dengan menggunakan MSE dan RMSE yang dapat dilihat pada Tabel 7. Berdasarkan Tabel 7, parameter temperatur

pada kedalaman 30 cm memperoleh nilai MSE 0,0 dan RMSE 0,02. Dan untuk parameter pH memperoleh nilai MSE 0,0 serta RMSE 0,0. Sedangkan pada parameter *Turbidity*, memperoleh nilai MSE sebesar 0,02 dan RMSE 0,13.

Pada data kedalaman 60 cm, parameter temperatur memperoleh nilai MSE sebesar 0,0 dan RMSE 0,0. Sedangkan parameter *turbidity* memperoleh nilai MSE sebesar 0,0 dan RMSE 0,05. Berdasarkan hasil tersebut, maka dapat disimpulkan bahwa semua parameter baik dari kedalaman 30 cm maupun 60 cm memperoleh hasil yang akurat untuk prediksi kualitas air 7 hari ke depan karena menghasilkan nilai *error* yang sangat rendah.

TABEL 7  
 HASIL PERHITUNGAN EROR

Parameter	Model SARIMA (p,d,q) (P,D,Q)s	MSE	RMSE
Data Kedalaman 30 cm			
Temperatur	SARIMAX (1, 1, 1)x(0, 0, 1) <sup>7</sup>	0,0	0,02
pH	SARIMAX (1, 0, 1)x(0, 0, 1) <sup>7</sup>	0,0	0,0
Turbidity	SARIMAX (1, 0, 0)x(1, 0, 1) <sup>7</sup>	0,02	0,13
Data Kedalaman 60 cm			
Temperatur	SARIMAX (1, 1, 1)x(1, 0, 1) <sup>7</sup>	0,0	0,0
Turbidity	SARIMAX (1, 0, 0)x(0, 0, 1) <sup>7</sup>	0,0	0,05

#### IV. KESIMPULAN

Berdasarkan hasil yang telah diperoleh, maka dapat disimpulkan bahwa model terbaik yang dihasilkan dari masing-masing parameter memperoleh nilai *error* yang sangat rendah sehingga hasilnya akurat dalam memprediksi parameter- parameter kualitas air 7 hari ke depan. Hasil peramalan kualitas air dengan parameter tersebut didapatkan data hasil untuk pengujian pada dataset kedalaman 30 cm untuk parameter temperature, yaitu MSE sebesar 0.0, dan RMSE sebesar 0.02. Untuk parameter pH, MSE sebesar 0.0, dan RMSE sebesar 0.0. Serta parameter turbidity, hasil MSE sebesar 0.02, dan RMSE sebesar 0.13. Sedangkan pada dataset kedalaman 60 cm, hasil dari parameter temperature, yaitu MSE sebesar 0.0, dan RMSE sebesar 0.0. Dan yang terakhir parameter turbidity, nilai MSE sebesar 0.0, dan RMSE sebesar 0.05. Selain itu, hasil dari penelitian ini juga menunjukkan bahwa metode SARIMA dalam penerapannya lebih baik dibandingkan dengan metode SVM pada penelitian terdahulu yang dilakukan oleh Nahid, et.al karena metode SARIMA menghasilkan nilai *error* yang lebih rendah dibandingkan metode SVM. Pada penelitian selanjutnya dapat dipadukan dengan menggunakan metode yang lain agar dapat mengetahui hasil terbaik dari perbandingan beberapa metode yang dilakukan

#### DAFTAR PUSTAKA

- [1] S. R. Fauzia and S. H. Suseno, "Resirkulasi Air Untuk Optimalisasi Kualitas Air Budidaya Ikan Nila Nirwana (*Oreochromis niloticus*)," *Jurnal Pusat Inovasi Masyarakat (PIM)*, vol. 2, no. 5, Art. no. 5, Jul. 2020.
- [2] D. Azhari and A. M. Tomaso, "Kajian Kualitas Air dan Pertumbuhan Ikan Nila (*Oreochromis niloticus*) yang Dibudidayakan dengan Sistem Akuaponik," *Akuatika Indonesia*, vol. 3, no. 2, Art. no. 2, Sep. 2018, doi: 10.24198/jaki.v3i2.23392.
- [3] S. Hanifi, X. Liu, Z. Lin, and S. Lotfian, "A Critical Review of Wind Power Forecasting Methods— Past, Present and Future," *Energies*, vol. 13, no. 15, Art. no. 15, Jan. 2020, doi: 10.3390/en13153764.
- [4] T. Nyoni, "Box-Jenkins ARIMA approach to predicting net FDI inflows in Zimbabwe," Jul. 2018, Accessed: Sep. 19, 2022. [Online]. Available: <https://mpr.ub.uni-muenchen.de/87737/>
- [5] "Perbandingan Metode Seasonal Autoregressive Integrated Moving Average (SARIMA) dengan Support Vector Regression (SVR) dalam Memprediksi Jumlah Kunjungan Wisatawan Mancanegara ke Bali | Jurnal Varian", Accessed: Sep. 19, 2022. [Online]. Available: <https://journal.universitاسbumigora.ac.id/index.php/Varian/article/view/668>
- [6] S. S. Aruan, "Perbandingan Metode Arima dan Sarima Dalam Peramalan Penjualan Kelapa," *JAMI: Jurnal Ahli Muda Indonesia*, vol. 2, no. 2, Art. no. 2, Dec. 2021, doi: 10.46510/jami.v2i2.82.
- [7] N. S. R. Ahmad, S. Martha, and N. Imro'ah, "PREDIKSI PRODUKSI KELAPA SAWIT DI PTPN XIII DENGAN ADDITIVE OUTLIER PADA MODEL SEASONAL AUTOREGRESSIVE INTEGRATED MOVING AVERAGE (SARIMA)," *Bimaster : Buletin Ilmiah Matematika, Statistika dan Terapannya*, vol. 8, no. 4, Art. no. 4, Oct. 2019, doi: 10.26418/bbimst.v8i4.36562.

- [8] K. He, L. Ji, C. W. D. Wu, and K. F. G. Tso, "Using SARIMA–CNN–LSTM approach to forecast daily tourism demand," *Journal of Hospitality and Tourism Management*, vol. 49, pp. 25–33, Dec. 2021, doi: 10.1016/j.jhtm.2021.08.022.
- [9] B. Y. Yohannes, S. W. Utomo, and H. Agustina, "Kajian Kualitas Air Sungai dan Upaya Pengendalian Pencemaran Air.," *IJEEM - Indonesian Journal of Environmental Education and Management*, vol. 4, no. 2, Art. no. 2, Jul. 2019, doi: 10.21009/IJEEM.042.05.
- [10] W. H. Siegers and Y. Prayitno, "PENGARUH KUALITAS AIR TERHADAP PERTUMBUHAN IKAN NILA NIRWANA (*Oreochromis sp.*) PADA TAMBAK PAYAU," vol. 3, no. 11, p. 10, 2019.
- [11] P. E. Wulandari, O. R. Pinontoan, and H. B. Boky, "KUALITAS AIR SUMUR BERDASARKAN PARAMETER FLUORIDA DAN PARAMETER PH DI KELURAHAN SUMOMPO KECAMATAN TUMINTING KOTA MANADO," *KESMAS*, vol. 8, no. 6, Art. no. 6, 2019, Accessed: Oct. 18, 2022. [Online]. Available: <https://ejournal.unsrat.ac.id/index.php/kesmas/article/view/25337>
- [12] H. Harianti and N. Nurasia, "ANALISIS WARNA, SUHU, pH DAN SALINITAS AIR SUMUR BOR DI KOTA PALOPO," *Prosiding*, vol. 2, no. 1, Art. no. 1, May 2016, Accessed: Oct. 18, 2022. [Online]. Available: <http://journal.uncp.ac.id/index.php/proceeding/article/view/520>
- [13] H. Said, N. H. Matondang, and H. N. Irmanda, "Perancangan Sistem Prediksi Kualitas Air Yang Dapat Dikonsumsi Dengan Menerapkan Algoritma K-Nearest Neighbor," *Prosiding Seminar Nasional Mahasiswa Bidang Ilmu Komputer dan Aplikasinya*, vol. 3, no. 1, Art. no. 1, Oct. 2022.
- [14] A. I. Arafat, T. Akter, Md. F. Ahammed, Md. Y. Ali, and A.-A. Nahid, "A dataset for internet of things based fish farm monitoring and notification system," *Data in Brief*, vol. 33, p. 106457, Dec. 2020, doi: 10.1016/j.dib.2020.106457.
- [15] Y. Muharni, K. Kulsum, and A. Denisa, "Prediksi Kualitas Air Baku dengan Pendekatan Adaptive Neuro Fuzzy Inference System," Mar. 2018, Accessed: Oct. 18, 2022. [Online]. Available: <http://publikasiilmiah.ums.ac.id/handle/11617/9832>
- [16] H. Said, N. H. Matondang, and H. N. Irmanda, "Penerapan Algoritma K-Nearest Neighbor Untuk Memprediksi Kualitas Air Yang Dapat Dikonsumsi," *Techno.Com*, vol. 21, no. 2, Art. no. 2, May 2022, doi: 10.33633/tc.v21i2.5901.
- [17] D. Ruhiat and A. Effendi, "PENGARUH FAKTOR MUSIMAN PADA PEMODELAN DERET WAKTU UNTUK PERAMALAN DEBIT SUNGAI DENGAN METODE SARIMA," *Teorema: Teori dan Riset Matematika*, vol. 2, no. 2, Art. no. 2, Mar. 2018, doi: 10.25157/teorema.v2i2.1075.
- [18] E. Supriyadi, "PREDIKSI PARAMETER CUACA MENGGUNAKAN DEEP LEARNING LONG-SHORT TERM MEMORY (LSTM)," *Jurnal Meteorologi dan Geofisika*, vol. 21, no. 2, Art. no. 2, Jan. 2021, doi: 10.31172/jmg.v21i2.619.
- [19] S. K. Agus Sasmito Aribowo, "Text-Preprocessing Model Youtube Comments in Indonesian," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 4, no. 4, pp. 648–654, Aug. 2020.
- [20] E. T. McGonigle, R. Killick, and M. A. Nunes, "Modelling time-varying first and second-order structure of time series via wavelets and differencing," *Electronic Journal of Statistics*, vol. 16, no. 2, pp. 4398–4448, Jan. 2022, doi: 10.1214/22-EJS2044.
- [21] M. Tiyani and A. I. Achmad, "Uji Stasioneritas Dickey Fuller Generalized Least Square pada Data Inflasi Indonesia dari Bulan Januari 2006 Hingga Bulan Mei 2020," *Prosiding Statistika*, vol. 6, no. 2, Art. no. 2, Aug. 2020, doi: 10.29313/.v6i2.22587.
- [22] E. A. Damanhuri, Y. I. Siregar, and E. Elfizar, "Penerapan Model Berbasis Artificial Neural Network Untuk Memprediksi Kualitas Air di Sungai Subayang Kabupaten Kampar," *Jurnal Ilmu Lingkungan*, vol. 14, no. 1, Art. no. 1, Mar. 2020, doi: 10.31258/jil.14.1.p.18-28.
- [23] R. Rajan, S. Rajest, and B. Singh, "Spatial Data Mining Methods Databases and Statistics Point of Views," in *Innovations in Information and Communication Technology Series*, 2021. doi: 10.46532/978-81-950008-7-6\_010.
- [24] G. N. Ayuni and D. Fitriana, "Penerapan Metode Regresi Linear Untuk Prediksi Penjualan Properti pada PT XYZ," *Jurnal Telematika*, vol. 14, no. 2, Art. no. 2, 2019.
- [25] A. Khumaidi, R. Raafi'udin, and I. P. Solihin, "Pengujian Algoritma Long Short Term Memory untuk Prediksi Kualitas Udara dan Suhu Kota Bandung," *Jurnal Telematika*, vol. 15, no. 1, Art. no. 1, Dec. 2020.
- [26] Susanto, "Model Perbandingan Metode Prediksi Jumlah Penjualan Produk Aplikasi HRIS Dengan Algoritma Forecasting Time Series Perusahaan SaaS," *Jurnal Sisfokom (Sistem Informasi dan Komputer)*, [Online]. Available: <http://jurnal.atmaluhur.ac.id/index.php/sisfokom/article/view/1387>