EasyChair Preprint

№ 14642

Mewma Control Chart for Monitoring Clean Water Quality by PDAM Production in Surabaya City Based on Residual of Generative Adversarial Network

Raditya Widi Indarsanto and Muhammad Ahsan

EasyChair preprints are intended for rapid dissemination of research results and are integrated with the rest of EasyChair.

September 1, 2024

MEWMA Control Chart for Monitoring Clean Water Quality Production in Surabaya City Based on Residual of Generative Adversarial Network

Raditya Widi Indarsanto[1], Dr. Muhammad Ahsan, S.Si[2]

[1]Sepuluh Nopember Institute of Technology, Surabaya, Indonesia
almiradit@gmail.com
[2]Sepuluh Nopember Institute of Technology, Surabaya, Indonesia
muh.ahsan@its.ac.id

Abstract. This study investigates the water quality characteristics of pH, turbidity, and KMnO4 at the Ngagel II water treatment plant operated by Surya Sembada water treatment in Surabaya, Indonesia. Phase I and II analyses revealed that while the water quality parameters met established standards, the presence of autocorrelation compromised data reliability. To address this, a Generative Adversarial Network (GAN) model was developed and optimized to generate residual values capable of reducing autocorrelation. The performance of the GAN was evaluated using Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Mean Absolute Error (MAE) metrics. The residual series was subsequently monitored using a Moving Average Exponential Weighted Moving Average (MEWMA) control chart with a smoothing parameter λ of 0.4. Phase I analysis indicated a statistically controlled process after outlier removal. However, Phase II monitoring detected out-of-control signals, suggesting process instability. The findings demonstrate the potential of GAN-based residual analysis in mitigating autocorrelation in water quality data. Nevertheless, the complexity of GAN training and the computational resources required for optimal model development pose significant challenges.

Keywords: Forecasting, Generative Adversarial Network, Control Chart, MEWMA.

1 Introduction

Surabaya, Indonesia, faces a growing challenge in ensuring sustainable clean water access for its expanding metropolitan population. Water is a critical resource, not just for basic needs but also for public health, economic activity, and urban development. Despite efforts by Surya Sembada, localized water scarcity persists in specific areas, including Wonokromo, Wonokusumo, Wonosari, and parts of North Surabaya. The water treatment process by Surya Sembada water treatment plant in Surabaya City is currently carried out at Ngagel and Karang Pilang installations. However, in its water treatment activities, Surya Sembada water treatment has not been able to implement quality control over the water treatment process using statistical methods. Therefore, this study proposes a multivariate control chart, which handles multivariate data problems using residuals from time series-based machine learning methods [1, 2]. This approach involves predicting the water production data, assessing the residuals for autocorrelation, and addressing potential problems [3, 4]. A variety of statistical and machine learning techniques can be employed to address this issue. Traditional time series methods such as Vector Autoregression (VAR) models [5], as well as advanced machine learning algorithms including Artificial Neural Networks (ANN) [6, 7], Multioutput Least Squares Support Vector Regression (MLS-SVR) [8], XGBoost [9], and Long Short-Term Memory (LSTM) [10, 11], offer potential solutions. To improve forecasting accuracy, this research explores the Generative Adversarial Network (GAN) method, which is known for generating realistic synthetic data [12, 13]. The application of GAN in water quality forecasting, particularly in water production, offers the advantages of generating accurate synthetic or predicted data and overcoming data limitations.

This research proposes a novel framework that synergizes Generative Adversarial Networks (GANs) with Multivariate Exponentially Weighted Moving Average (MEWMA) control charts to enhance water quality monitoring and operational efficiency in water production. By effectively addressing autocorrelation inherent in water quality data, the GAN-MEWMA model aims to revolutionize traditional monitoring systems. The fidelity of GAN-generated data is paramount in ensuring the accuracy and comprehensiveness of the overall monitoring process. This study seeks to demonstrate the potential of artificial intelligence in improving forecasting capabilities and enabling proactive responses to water quality fluctuations. Ultimately, the integration of this advanced monitoring system is expected to optimize water treatment operations by facilitating timely interventions and preventing quality deterioration.

2 Literature Review

2.1 Statistical Quality Control and Control Chart

Statistical quality control (SQC) is a structured methodology employed to attain and maintain desired quality levels within a product or process by systematically reducing process variation [14] It constitutes a statistical framework for the continuous monitoring, measurement, and enhancement of quality attributes [2]. Through the application of statistical tools such as control charts, hypothesis testing, and regression analysis, SQC enables the identification, quantification, and mitigation of process variability. By detecting anomalies, predicting trends, and implementing corrective actions in a timely manner, SQC optimizes operational efficiency, minimizes waste, and elevates customer satisfaction through continual quality improvement. Control charts are indispensable statistical process control tools employed in quality management to visually monitor process variation and identify anomalous shifts or trends [7]. By graphically representing process data relative to established control limits, these charts facilitate the maintenance of acceptable quality levels through the detection of special cause variation [15].

2.2 Generative Adversarial Network Test

Generative Adversarial Networks (GANs) are unsupervised learning models that generate indistinguishable synthetic data. Comprising of two components, the generator and discriminator [16]. The generator is tasked with creating synthetic data that is like original data, where previously generated data was from actual data, while the discriminator tries to differentiate between real or fake data and data generated by the generator. GANs iteratively improve their performance in data creation and differentiation. Forecasting process with a Generative Adversarial Network (GAN) involves a complex series of steps to produce synthetic data. This process starts from collecting actual data as the input vector to producing synthetic data which can be depicted in the diagram below.

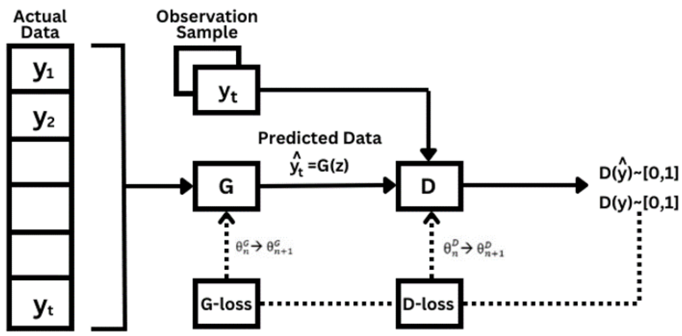


Figure 1 Generative Adversarial Network Forecasting Process

In the GAN model, synthetic data is produced from random or actual data. This is done by the generator, which uses an Artificial Neural Network (ANN) for processing. In time series forecasting, the generator applies mathematical transformations, using a weight and bias matrix for each layer and an activation function for each neuron's output to generate synthetic data.

$$\hat{y}_t = G(y_t) \tag{1}$$
$$G(y_t) = \sigma(h_{out}) \tag{2}$$
$$h_{out} = w_{out} \times h + b_{out} \tag{3}$$

where w_{out} is the weight matrix in the generator output layer, h is the value produced by generator hidden layer and b_{out} is the bias vector in the generator output layer. While discriminator in GAN also uses a simple Artificial Neural Network (ANN) model that can be represented by equations below.

$$D(y) = \sigma(w_{out} \times h + b_{out}) \tag{4}$$
$$D(\hat{y}) = \sigma(w_{out} \times h' + b_{out}) \tag{5}$$

σ is activation function, w_{out} is output layer weight, b_{out} is output layer bias, and h'

is transformation results from data produced by the generator. In GAN training, the goal is to optimize the weights and biases in the discriminator so that $D(y)$ close to 1 for the original data and $D(\hat{y})$ approaches 0 for data generated by the generator.

To optimize GAN models, select suitable generator and discriminator architectures based on task complexity and data type. Parameter tuning, like learning rate and latent vector size, is vital for stability and complexity. Utilize Batch Normalization and Residual Connections for training stability. Choose a balanced loss function for the generator and discriminator. Lastly, set batch size and epochs wisely to expedite convergence and minimize overfitting, and ensure optimal GAN model results

2.3 MEMWA Control Chart Based on Residual of Generative Adversarial Network Model

The Exponentially Weighted Moving Average (EWMA) control chart is a statistical process control (SPC) technique that assigns exponentially decreasing weights to historical observations [17]. To accommodate multivariate processes, the Multivariate EWMA (MEWMA) chart was developed [18]. By considering multiple quality characteristics simultaneously, the MEWMA chart demonstrates enhanced sensitivity to small process shifts, enabling earlier detection of anomalies. Optimal performance of the MEWMA chart is contingent upon the assumption of independent and normally distributed data. While the MEWMA chart exhibits robustness to deviations from normality, its effectiveness can be compromised by autocorrelation in the data.

MEWMA diagrams can be used with GAN models to monitor accuracy by focusing on residuals, the difference between original and GAN-generated data. MEWMA helps detect changes in synthetic data characteristics, providing early alerts for significant deviations. Therefore, the MEWMA equation can be formed as follows [14].

$$M_t = \lambda e_t + (1-\lambda)M_{t-1} \tag{6}$$

M_t is MEWMA value at t, e_t is residual value at t, M_{t-1} is MEWMA value at the previous time, and λ is the weighted exponential where define as $0 < \lambda < 1$ and $M_0 = 0$. Basically, λ could work well while in the interval $0.05 \leq \lambda \leq 0.25$, with $\lambda = 0.05$, $\lambda = 0.1$, $\lambda = 0.2$ can be a popular choice. The observation points plotted on the control chart are as follows.

$$T_t^2 = M_t' \textstyle\sum_{M_t}^{-1} M_t \tag{7}$$

where the covariance matrix is

$$\textstyle\sum_{M_t}^{-1} = \frac{\lambda}{(2-\lambda)}[1-(1-\lambda)^{2t}]\Sigma \tag{8}$$

Observations are said to be out of control when value $T_t^2 > H$, where H is the upper control limit (UCL) which used for MEWMA control chart based on selected lambda

(λ) and number of quality characteristics, while the lower control limit (LCL) for MEWMA control charts is equal to 0 because the T_t^2 value is always positive.

3 Methodology

3.1 Data Source

This research employed a secondary dataset comprising observational data on clean water production quality. The data was procured from the laboratory of Ngagel II Surya Sembada, encompassing the water filtration process from January 1, 2023, to October 24, 2023. The investigation was divided into two phases: Phase I (January 1, 2023 - June 30, 2023) and Phase II (July 1, 2023 - October 24, 2023). Quality assessments were conducted for 296 days post-filtration, a process designed to eliminate fine particles from the previously treated water. Three key water quality parameters were analyzed: pH, turbidity, and organic matter content (measured as KMnO4).

3.2 Research Variable

The research variables used in this research are three quality characteristics which are explained in Table below.

Table 1. Research Variables

Quality Characteristics	Description	Unit	Specification
y_1	pH	-	6.5-7.5
y_2	Turbidity	NTU	Max 5
y_3	Organic Substances (KMnO4)	Mg/L	Max 10

3.3 Analysis Steps

The research steps analysis taken in this research are as follows.

1. **Problem Definition and Data Collection**: Identify and formalize the specific water quality issues at Surya Sembada water treatment plant Ngagel II, Surabaya. Acquire comprehensive water quality data from January 1, 2023, to October 24, 2023.
2. **Autocorrelation Analysis**: Assess the presence of autocorrelation in the water quality time series using the Multivariate Cross-Correlation Function (MCCF) plot. Identify significant autocorrelation lags to inform subsequent modeling decisions.
3. **Data Partitioning**: Divide the dataset into training data (phase I: January 1, 2023, to June 30, 2023) and testing data (phase II: July 1, 2023, to October 24, 2023).
4. **GAN Modeling**: Develop a Generative Adversarial Network (GAN) model to generate synthetic water quality data. Optimize GAN hyperparameters (layers, epochs,

batch size) on the training set to achieve optimal performance. Apply the trained GAN model to generate synthetic data for the testing set.

5. **Residual Analysis**: Calculate residuals between the actual and synthetic water quality data for the training set. Assess the normality of residuals using the Shapiro-Wilk test and the presence of autocorrelation using the MCCF plot.
6. **Multivariate EWMA Control Chart**: Develop a Multivariate Exponentially Weighted Moving Average (MEWMA) control chart using the residuals from the training set. Establish control limits based on in-control assumptions. Monitor the testing set residuals using the developed control chart to detect abnormal patterns or out-of-control signals.
7. **Conclusion and Recommendations**: Findings and Implications: Summarize the key findings of the study, including the performance of the GAN model, the effectiveness of the control chart in detecting anomalies, and the identified root causes of out-of-control signals. Provide actionable recommendations for improving water quality management at Surya Sembada water treatment plant Ngagel II, Surabaya. This may involve process adjustments, data-driven decision making, or implementation of advanced monitoring systems.

4 Result and Discussion

4.1 Autocorrelation Analysis

To use control charts for analysis, observation data must be independent. If autocorrelation is detected, it must be addressed. Prior to autocorrelation checking, the three quality characteristics are standardized due to their different units. Autocorrelation is checked using the MCCF plot for each quality characteristic of water, as shown in Table 2.

Table 2. MCCF Result of Quality Characteristics

Variable /Lag	0	1	2	3	4	5	6	7	8	9
pH	++-	++.	++.	++.	++.	++.	++.	++.	++.	++.
Turbidity	+++	+++	+++	+++	+++	++.	++.	++.	++.	++.
KMnO4	-++	..+	-.+	-.+	-.+	..+	..+	..+	..+	...

Table 2. MCCF Result of Quality Characteristics (continue)

Variable / Lag	10	11	12	13	14	15	16	17	18	19
pH	++.	++.	++.	++.	++.	++.	++.	++.	++.	++.
Turbidity	++.	++.	++.	++.	++.	++.	++.	++.	++.	++.
KMnO4	...	...	...	...	...	...	...	...	...	...

Table 2 reveals a pronounced and persistent positive autocorrelation for pH and turbidity across nearly all lags, suggesting a strong temporal dependency in these parameters. This indicates that fluctuations in pH and turbidity tend to be correlated with their preceding values over extended periods. In contrast, KMnO4 exhibits a more erratic and inconsistent autocorrelation pattern, characterized by substantial fluctuations and a lack of discernible regularity. Collectively, these findings imply a higher degree of stationarity for pH and turbidity compared to KMnO4. The presence of autocorrelation in the clean water process data is attributable to the continuous nature of the production process. To address the challenges posed by autocorrelation in the observational data, the application of Generative Adversarial Network (GAN) modeling is proposed.

4.2 Generative Adversarial Network's Modelling

GAN-based ANN is employed to address autocorrelation in clean water production. Optimal GAN architecture and hyperparameters are determined through hyperparameter tuning to minimize model residuals. These optimized parameters will be applied to the modeling process in phase II.

1. Phase I GAN's Modelling

Phase I of the GAN modeling process utilized clean water data spanning January 1 to June 30, 2023. The network architecture and hyperparameters employed are detailed in Table 3.

Table 3. Architecture & Hyper-Parameter

Component	*Hyper-Parameter*
Generator Layer 1	Number of Neuron = 200 Activation = LeakyReLu (Alpha 0.01)
Generator Layer 2	Number of Neuron = 600 Activation = LeakyReLu (Alpha 0.01)
Generator Layer 3	Number of Neuron = 3 Activation = Linear
Discriminator Layer 1	Number of Neuron = 600 Activation = LeakyReLu (Alpha 0.01)
Discriminator Layer 2	Number of Neuron = 200 Activation = LeakyReLu (Alpha 0.01)
Discriminator Layer 3	Number of Neuron = 1 Activation = Sigmoid
Discriminator Compile	Loss Function = Binary Cross Entropy

	Optimizer = Adam (Lr 0.0001, Beta 0.3)
Compile of Model	Optimizer = Adam (Learning Rate = 0.0001, Beta = 0.3)

Model training was conducted using the Adam optimizer with a learning rate of 0.0001 and parameter β. A range of epochs were explored during training, and resulting loss values for each epoch are tabulated in Table 4. Training metrics were captured and stored within a history variable for subsequent analysis.

Table 4 Loss Values for Generator and Discriminator

Epoch	Generator Loss	Discriminator Loss
600	0.3751	0.9538
650	0.3748	0.9467
700	0.3729	0,9443

To optimize GAN model performance, a hyperparameter tuning process was conducted, focusing on the epoch value. From three candidate epochs, the model trained for 700 epochs exhibited the lowest loss values, suggesting superior convergence and predictive accuracy. This outcome indicates the generator's effectiveness in producing highly realistic synthetic data, thereby challenging the discriminator's ability to distinguish between genuine and artificial samples. Model evaluation was conducted using Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Mean Absolute Error (MAE) metrics. These metrics quantified the discrepancy between predicted and actual values for the three quality characteristics generated by the GAN model. The resulting evaluation results are summarized in Table 5.

Table 5 Evaluation Metrics

Epoch	$\overline{MSE}$	$\overline{RMSE}$	$\overline{MAE}$
600	0.0067	0.0822	0.0633
650	**0.0054**	**0.0738**	**0.0556**
700	0.0071	0.0846	0.0637

The GAN model trained for 650 epochs exhibited the most promising performance, demonstrating superior predictive accuracy. The model achieved a MSE of 0.0054, indicating minimal variance between predicted and actual values. Correspondingly, RMSE of 0.0738 suggests a relatively low average prediction error. Further, the MAE of 0.0556 quantifies the average absolute deviation between predicted and observed values. Collectively, these metrics underscore the model's strong predictive capabilities. As visualized in Figure 1, a comparison of predicted and actual values provides empirical evidence supporting the model's reliability.

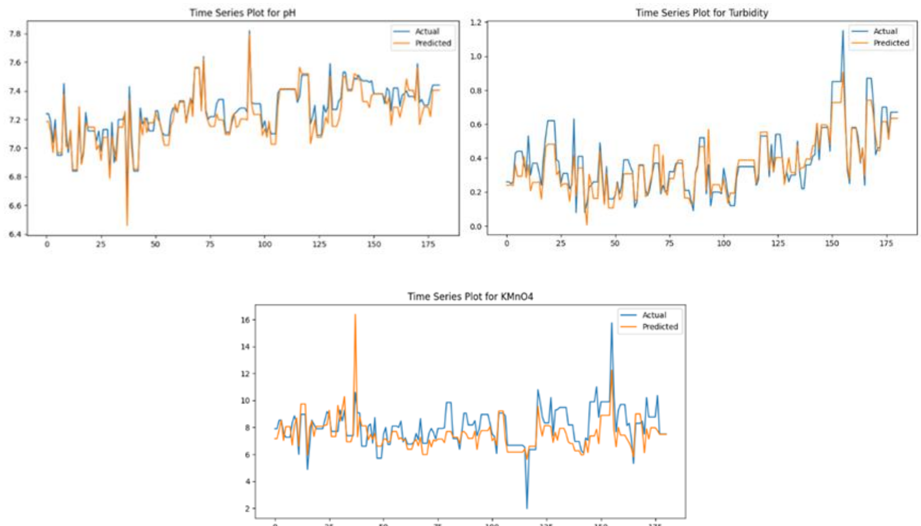


Figure 2 Time Series Plot of Actual Data and Predicted Data in Phase I

Figures 2(a) and 1(b) show that the data pattern of the predicted values for the quality characteristics of pH and turbidity from GAN modeling has a pattern that is similar or follows the actual data values for phase I. Meanwhile, the quality characteristics of KMnO4 can be seen in Figure 1(c) has a predicted data pattern that tends to be on average or in the middle with actual data. This proves that predictions by modeling using GANs can be used because the patterns produced from predicted data tend to have the same patterns as actual data.

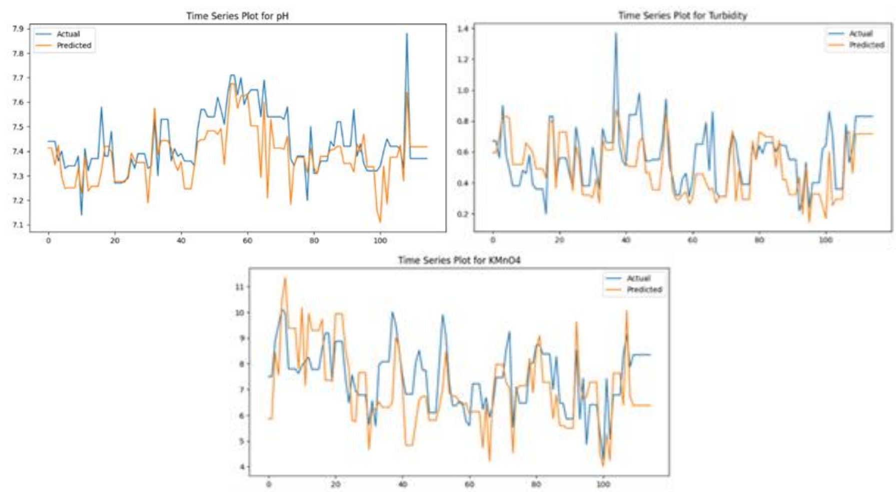


Figure 3 Time Series Plot of Actual Data and Predicted Data in Phase II

2. Phase II GAN's Modelling

Phase II GAN modeling was conducted using identical architectural and hyperparameter settings as employed in phase I. The model was trained on clean water data spanning from July 1 to October 24, 2023. Subsequent to training, predicted values were generated and compared to corresponding actual data points, as visualized in Figure 3. An analysis of Figure 3 reveals that the predicted values for all three quality characteristics exhibit a pattern that deviates significantly from the observed trends in the phase II actual data. Nevertheless, the predicted data points are distributed within a range proximate to the mean of the actual data. These findings suggest that the GAN model's predictive capabilities during phase II were suboptimal in capturing the specific nuances and fluctuations of the actual data. Consequently, the generated data displayed markedly different patterns compared to the ground truth. To quantitatively assess these pattern discrepancies, a control chart analysis will be performed.

4.3 Control Chart Assumption Testing

The residuals of the GAN-modeled water quality data from Phase I observations are assessed for autocorrelation using MCCF plots prior to multivariate normality testing, a prerequisite for control chart implementation.

Table 6. MCCF Result Based on GAN Phase I Residuals

Variable / Lag	0	1	2	3	4	5	6	7	8	9
pH	+-+	+-+	+..	...	...	...	...	...	...	...
Turbidity	-+.	.+.	.+.	...	...	...	..+	...	...	...
KMnO4	+.+	+.+	...	...	...	...	...	...	...	...

Table 6. Result Based on GAN Phase I Residuals (continue)

Variable / Lag	10	11	12	13	14	15	16	17	18	19
pH	...	.-.	...	..+	..+	...	...	...	...	...
Turbidity	...	...	..-	..-	-+-	.+.	...	...	...	...
KMnO4	...	+..	...	...	...	...	...	...	...	...

Analysis of Table 6 indicates that GAN modeling effectively mitigates autocorrelation present in Phase I observational data for the examined quality characteristics. The pH variable exhibited pronounced cyclical patterns, evidenced by significant and strong autocorrelation at initial and multiple higher lags. Turbidity data displayed higher variability and a more random structure, although significant autocorrelation was detected at specific lags. KMnO4 demonstrated strong autocorrelation at early lags, transitioning to a more stable pattern at subsequent lags. Subsequent to autocorrelation assessments of GAN model residuals, a Shapiro-Wilk test was employed to evaluate multivariate normality. Results are tabulated in Table 7.

Table 7 Normal Multivariate Test for Residuals Phase I

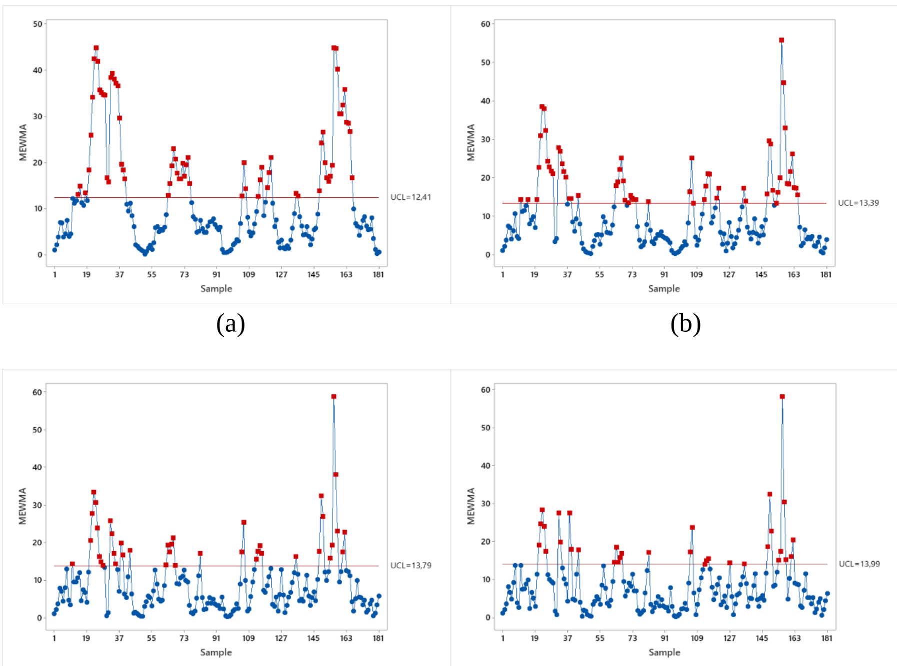
Shapiro Wilk	P-Value	Decision
0.86278	< 2.2e-16	Not normally distributed

Based on the multivariate normal distribution test reveals that the clean water quality data doesn't follow a multivariate normal distribution, as the p-value is less than the 5% alpha level. Despite this, the research can proceed with the MEWMA control chart for process control, as it's robust to normality assumptions and is based on GAN model residuals.

4.4 MEWMA Control Chart Based on Residual of GAN Models

The Multivariate Exponentially Weighted Moving Average (MEWMA) control chart extends the univariate EWMA to monitor multiple quality characteristics simultaneously. It excels at detecting small process shifts and is sensitive to changes in the process mean. Control limits are determined based on Average Run Length (ARL) criteria, with a target ARL of 370 achieved through appropriate weighting (λ).

1. MEWMA Control Chart Based on Residual of GAN Models Phase I



(a) (b)

(c) (d)

Figure 4 MEWMA Control Chart Based on Residual of GAN Phase I
(a) λ = 0.1, (b) λ = 0.2, (c) λ = 0.3, (d) λ = 0.4

Figure 4 indicates optimal MEWMA chart performance at λ = 0.4 with UCL = 13.99. However, 33 out-of-control points were detected, suggesting process instability in the GAN model during phase I. To address this, residual values were centered to zero, as depicted in the subsequent MEWMA chart.

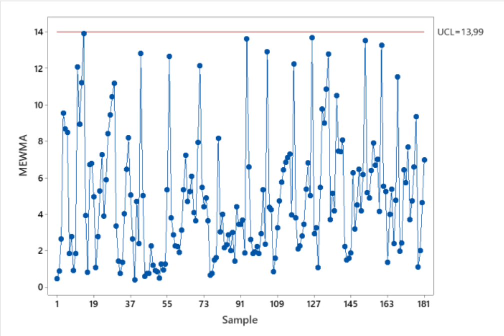


Figure 6 MEWMA Control Chart Based on Residual of GAN Phase I (In Control)

Figure 6 depicts an updated MEWMA control chart for Phase I GAN model residuals (UCL=13.99, λ=0.4). No out-of-control signals were detected post-intervention (removal of the highest outlier). The Phase I water production GAN model is now statistically in-control and prepared for Phase II monitoring.

2. MEWMA Control Chart Based on Residual of GAN Models Phase II

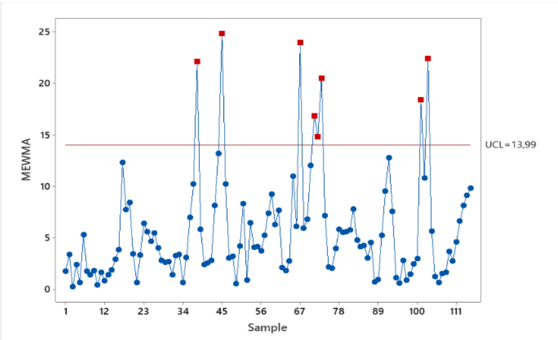


Figure 7 MEWMA Control Chart Based on Residual of GAN Phase II

Figure 7 depicts a Phase II Multivariate Exponentially Weighted Moving Average (MEWMA) control chart with λ = 0.4 and UCL = 13.99. The presence of eight data points exceeding the upper control limit indicates a statistically significant deviation from the process mean in clean water production. This suggests the process is out of control. To identify the specific factors contributing to this process shift, a detailed analysis of residuals from the Phase I Generative Adversarial Network (GAN) model is warranted. These residuals, representing the discrepancies between observed and predicted clean water quality characteristics, can provide valuable insights into the underlying process dynamics. By examining combinations of these residual patterns, it is possible to pinpoint potential root causes of the process excursion. Such knowledge is essential for implementing corrective actions to restore the process to a state of statistical control and ensure the consistent delivery of high-quality clean water to customers.

5 Conclusion

This study introduces a novel approach to water quality monitoring by integrating Generative Adversarial Network (GAN) residuals into a Multivariate Exponentially Weighted Moving Average (MEWMA) control chart. Analysis of Surya Sembada's clean water production process revealed that current pH, turbidity, and KMnO4 levels are within specified limits, indicative of a controlled process. The proposed GAN model effectively mitigated autocorrelation in residuals and exhibited superior predictive performance, as quantified by lower MSE, RMSE, and MAE compared to alternative methods. While the MEWMA chart effectively detected process deviations in Phase II, the computational demands of GAN training underscore the need for careful consideration of computational efficiency and robustness to process shifts for practical applications. Future research should explore strategies to optimize computational performance and enhance the model's sensitivity to subtle process variations. Also, some robust approaches can be applied to the control chart.

References

1. Khusna, H., Mashuri, M., Suhartono, Prastyo, D.D., Lee, M.H., Ahsan, M.: Residual-based maximum MCUSUM control chart for joint monitoring the mean and variability of multivariate autocorrelated processes. Prod Manuf Res. 7, 364–394 (2019).
2. Khusna, H., Mashuri, M., Suhartono, S., Prastyo, D.D., Ahsan, M.: Multioutput least square SVR based multivariate EWMA control chart: The performance evaluation and application. Cogent Eng. (2018). https://doi.org/10.1080/23311916.2018.1531456.
3. Prastyo, D.D., Khusna, H., Mashuri, M., Suhartono, S., Ahsan, M.: Multivariate CUSUM control chart based on the residuals of multioutput least squares SVR for monitoring water quality. MJS. 38, 73–83 (2019).
4. Sulistiawanti, N., Ahsan, M., Khusna, H.: Multivariate Exponentially Weighted Moving Average (MEWMA) and Multivariate Exponentially Weighted Moving Variance (MEWMV) Chart Based on Residual XGBoost Regression for Monitoring Water Quality. Engineering Letters. 31, (2023).
5. Veerendra, G.T.N., Kumaravel, B., Rao, P.K.R., Dey, S., Manoj, A.V.P.: Forecasting models for surface water quality using predictive analytics. Environ Dev Sustain. 26, 15931–15951 (2024).
6. Palani, S., Liong, S.-Y., Tkalich, P.: An ANN application for water quality forecasting. Mar Pollut Bull. 56, 1586–1597 (2008).

7. won Seo, I., Yun, S.H., Choi, S.Y.: Forecasting water quality parameters by ANN model using pre-processing technique at the downstream of Cheongpyeong Dam. Procedia Eng. 154, 1110–1115 (2016).
8. Khusna, H., Mashuri, M., Suhartono, S., Prastyo, D.D., Ahsan, M.: Multioutput least square SVR-based multivariate EWMA control chart: The performance evaluation and application. Cogent Eng. 5, 1531456 (2018).
9. Yusri, H.I.H., Hassan, S.L.M., Halim, I.S.A., Abdullah, N.E.: Water quality classification using SVM and XGBoost method. In: 2022 IEEE 13th Control and System Graduate Research Colloquium (ICSGRC). pp. 231–236. IEEE (2022).
10. Wang, Y., Zhou, J., Chen, K., Wang, Y., Liu, L.: Water quality prediction method based on LSTM neural network. In: 2017 12th international conference on intelligent systems and knowledge engineering (ISKE). pp. 1–5. IEEE (2017).
11. Tejoyadav, M., Nayak, R., Pati, U.C.: Multivariate water quality forecasting of river ganga using var-lstm based hybrid model. In: 2022 IEEE 19th India Council International Conference (INDICON). pp. 1–6. IEEE (2022).
12. Kumar, A., Alsadoon, A., Prasad, P.W.C., Abdullah, S., Rashid, T.A., Pham, D.T.H., Nguyen, T.Q.V.: Generative adversarial network (GAN) and enhanced root mean square error (ERMSE): deep learning for stock price movement prediction. Multimed Tools Appl. 1–19 (2022).
13. Qi, S., Chen, J., Chen, P., Wen, P., Niu, X., Xu, L.: An efficient GAN-based predictive framework for multivariate time series anomaly prediction in cloud data centers. J Supercomput. 80, 1268–1293 (2024).
14. Montgomery, D.C.: Introduction to statistical quality control. John wiley & sons (2019).
15. Walpole, R.E., Myers, R.H., Myers, S.L., Ye, K.: Probability and statistics for engineers and scientists. Macmillan New York (1993).
16. Sharma, P., Kumar, M., Sharma, H.K., Biju, S.M.: Generative adversarial networks (GANs): Introduction, Taxonomy, Variants, Limitations, and Applications. Multimed Tools Appl. 1–48 (2024).
17. Roberts, S.W.: Control chart tests based on geometric moving averages. Technometrics. 42, 97–101 (2000).
18. Lowry, C.A., Woodall, W.H., Champ, C.W., Rigdon, S.E.: A multivariate exponentially weighted moving average control chart. Technometrics. 34, 46–53 (1992).