EUROPEAN JOURNAL OF PURE AND APPLIED MATHEMATICS
2026, Vol. 19, Issue 1, Article Number 5439
ISSN 1307-5543 – ejpam.com
Published by New York Business Global

Estimation and Prediction of the Omicron COVID-19 Data in Indonesia

Gantina Rachmaputri[1,2,*], Nastitie[3]

[1] *Algebra Research Group, Faculty of Mathematics and Natural Sciences, Institut Teknologi Bandung, Bandung, Jawa Barat, Indonesia*
[2] *Center of Excellence in Predictive Risk, Simulation, and Modelling, Institut Teknologi Bandung, Bandung, Jawa Barat, Indonesia*
[3] *Teaching of Mathematics Master's Program, Faculty of Mathematics and Natural Sciences, Institut Teknologi Bandung, Bandung, Jawa Barat, Indonesia*

Abstract. The prediction of epidemic spread is an essential aspect of public health planning and resource allocation. This paper discusses an improved prediction method by combining the Extended Kalman Filter (EKF) as an optimal estimation technique with the Moving Average (MA) as a simple time-series prediction technique. The integration of these two methods results in the Modified Extended Kalman Filter (mEKF), which enables prediction even in the absence of actual observation data. This study applies mEKF to the SEIR model using COVID-19 Omicron case data in Indonesia. The results show that mEKF produces prediction values closer to real data compared to MA, which relies heavily on historical patterns. Therefore, mEKF provides more stable and accurate prediction performance, making it suitable for short-term epidemiological forecasting.

2020 Mathematics Subject Classifications: 92-10

Key Words and Phrases: Extended Kalman Filter, Modification of Extended Kalman Filter, Moving Average

1. Introduction

Even the most complicated natural phenomena can be simplified and solved by constructing a mathematical equation or mathematical model in which each variable represents the studied object. The complexity of the problem determines the elaborateness of the resulting mathematical model. Most simple mathematical models are linear, whereas mathematical models derived from highly complex issues will be non-linear.

Interesting natural phenomena to study include control methods, prediction methods, estimation methods, and others. This paper discusses the combination of the Extended Kalman Filter and the Moving Average method to produce a modified prediction approach.

*Corresponding author.
DOI: https://doi.org/10.29020/nybg.ejpam.v19i1.5439

Email addresses: gantina.r@itb.ac.id (G. Rachmaputri), aku.nastitie@gmail.com (Nastitie)

Kalman filter is the most optimal algorithm for recursive data processing. Optimal here means that the Kalman filter combines all the information provided. It processes all available measurements in a calculation or mathematical model, from initial and parameter values to system noise and measurement noise. Therefore, the Kalman filter will reduce potential errors [1].

Meanwhile, the word "recursive" indicates that the Kalman filter does not require the previous data to be stored and recalculated whenever new data estimates are needed. As previously stated, the system frequently present in a phenomenon is non-linear, so an estimation method for a non-linear model, the Extended Kalman filter, was developed [1].

EKF is the optimal method for discrete models, meaning it can only be estimated when real data are available. Without actual data, the EKF method cannot make estimates or predictions. In addition, the accuracy of the modified EKF (mEKF) is compared with the standard Moving Average method to evaluate its predictive performance [1].

EKF is not a prediction method but the best estimation method. Therefore, we will add a prediction method that is quite simple but widely used today, the moving average. We will use the moving averages calculation process to develop a new algorithm for EKF called MEKF. A scientific paper by Helisyah et al. [2] mentions the MEKF method. The MEKF algorithm does not, however, include a moving average. Furthermore, this study highlights how the modification enables EKF to perform multi-step predictions when actual data are unavailable.

Let us now describe the organization of this paper. Section 2 describes the methods we will employ for this paper: EKF, mEKF, and moving average. Section 3 will apply the procedures described in Section 2 to the SEIR model and COVID-19 Indonesia data. Section 4 explains the discussion and conclusion regarding the previous section's steps.

2. Methods

The Extended Kalman Filter can only estimate and predict one step due to the limited data available. If we want to make predictions beyond our actual data, we need an additional algorithm known as the Modified Extended Kalman Filter [2]. The modified, extended Kalman filter predicts beyond the actual data by generating measurement data based on previous time estimates, enabling the utilization of bounce data during the correction stage. Suppose, given the equation of a model of a non-linear system and the following measurements:

$$x_{k+1} = f(x_k, w_k, k)$$
$$z_k = h(x_k, v_k, k)$$

where f is a non-linear function that depends on the x_k state variable, h is the observation matrix, w_k is the noise system with w_k $(0, Q)$, and v_k is the measurement noise with v_k $(0, R)$.

For example, if the measurement data is as much as k, we can find the estimated data for x_{k+1} using the EKF algorithm. Next, for prediction x_{n+1} with $n > k$ and $n - k = 1$

or $n = k + 1$ difference where $n, k \in \mathbb{Z}$. Initialization of the model and measurements is carried out, namely

$$
\begin{aligned}
x_{n+1} &= f(x_n, j_n, n) \\
z_{n+1} &= h(x_n, q_n, n)
\end{aligned}
\tag{1}
$$

With j_n and q_n respectively are the noise system and the noise measurement where j_n $(0, B)$ and q_n $(0, L)$. In Equation 1, the measurement data is generated with estimation data plus noise, or can be written as $z_{n+1} = \hat{x}_{k+1} + \epsilon$.

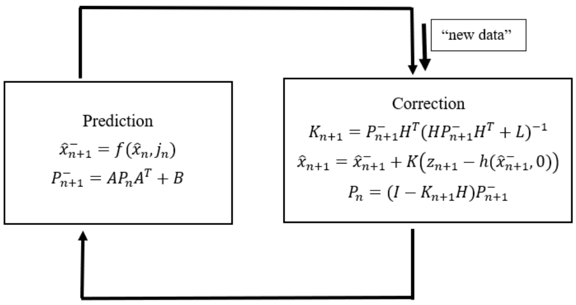

Figure 1: Algorithm of the Modified Extended Kalman Filter (mEKF), illustrating the data estimation and prediction process when actual observations are unavailable.

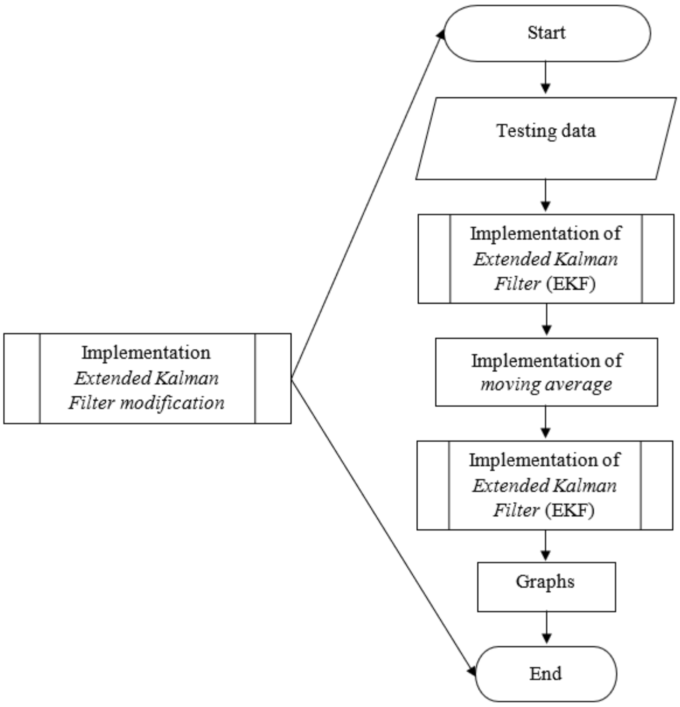

Figure 2: Flowchart of the Modified Extended Kalman Filter (mEKF) showing the integration of EKF estimation and moving average-based prediction.

This study enhances the modified EKF by incorporating data generation algorithms and additional prediction methods to improve accuracy. These methods range from the simplest to the most complex [3, 4]. The study evaluates its effectiveness by combining the modified EKF with even the simplest prediction method. Among the various prediction methods, this research utilizes a combination of the EKF and moving average methods, referred to as "modified, extended Kalman filters." The moving average plays a crucial role in determining the EKF bounce data.

In the prediction process, actual data will be separated into training and testing data. The purpose of the training data is to evaluate the COVID-19 model's implementation in the EKF method. In the meantime, testing data are utilized to compare predicted and actual data. Predictions can be made without the actual data if the predictions' results are accurate enough (close to the actual data).

The step-by-step computational procedure of the Modified Extended Kalman Filter (mEKF) is summarized in Algorithm 1, and its flowchart representation is illustrated in Figure 2.

3. Experiments

3.1. Experiment Using SEIR Model

The mathematical model used in this study is the SEIR model, reference from Annas et al. [5]. The SEIR model's purpose is to observe the effect of vaccination and the isolation period on the progression of COVID-19 infection in Indonesia. The structure of the SEIR model and the transitions between compartments are illustrated in Figure 3.

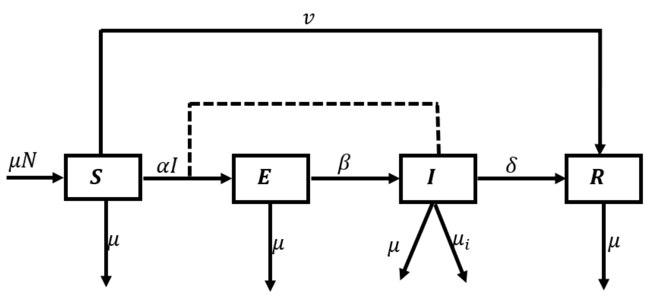

Figure 3: Flowchart of the SEIR model structure used to represent the dynamics of COVID-19 transmission, illustrating transitions between the Susceptible (S), Exposed (E), Infected (I), and Recovered (R) compartments.

The assumptions given in the SEIR model above are:

(i) In each compartment, there is a natural death of μ.

(ii) The natural birth and death rates (μ) are the same in each compartment.

(iii) Individuals in phase S who have not been vaccinated (v) have a high probability of going to phase E.

(iv) Individuals in phase S who have been vaccinated (v) will go to phase R.

(v) Individuals in phase S can go through phase E by α or directly to phase I (no parameter values are given because individuals are assumed to always go through phase E first).

(vi) In phase I, there are natural deaths (μ) and deaths from COVID-19 (μ_i).

(vii) It does not consider the type of COVID-19 that occurs in Indonesia.

(viii) It does not consider the effectiveness of the vaccine applied in Indonesia.

The descriptions, variable values, and parameters from the preceding SEIR model are listed below.

Table 1: Definitions and initial values of variables and parameters in the SEIR model for COVID-19 in Indonesia.

Variable / parameter	Definition	Amount	Source
N	Total population	72224	-
S_0	Initial value of the vulnerable population	37538	[6]
E_0	Initial value of the latent population	13923	[6]
I_0	Initial value of the population infected	8844	[6]
R_0	Initial value of the cured individual	11919	[6]
μ	Natural birth/death rate	6.25×10^{-3}v	[5]
α	Possible change from S to E	0.62×10^{-6}	[5]
μ_i	Death rate from COVID-19	7.344×10^{-7}	[5]
δ	Possible change from I to R	0.0006667	[5]
$\frac{1}{\beta}$	Average isolation time	14	[7]
v	Vaccination rate	80%	[7]

With the above assumptions, the following mathematical model is derived:

$$\frac{dS}{dt} = \mu N - (\alpha I + \mu + v)S \tag{2}$$

$$\frac{dE}{dt} = \alpha IS - (\beta + \mu)E \tag{3}$$

$$\frac{dI}{dt} = \beta E - (\mu_i + \delta + \mu)I \tag{4}$$

$$\frac{dR}{dt} = \delta I + vS - \mu R \tag{5}$$

3.2. Experiment EKF Using Data COVID-19 from Indonesia

The COVID-19 mathematical model in Equations 2, 3, 4, 5 are continuous-time deterministic dynamic model. Using the forward finite difference method, the equation is transformed into a discrete-time dynamic system model [5, 8, 9]. The discrete form is obtained as follows:

$$\begin{aligned} S_{t+1} &= (\mu - (\alpha I_t + \mu + v)S_t)\Delta t + S_t \\ E_{t+1} &= (\alpha I_t S_t - (\beta + \mu)E_t)\Delta t + E_t \\ I_{t+1} &= (\beta E - (\mu_i + \delta + \mu)I)\Delta t + I_t \\ R_{t+1} &= (\delta I + vS - \mu R)\Delta t = R_t \end{aligned}$$

Next, linearization using the Jacobian matrix is performed [5, 8, 9].

$$J = \begin{pmatrix} J_{11} & J_{12} & 0 & J_{14} \\ 0 & J_{22} & J_{23} & 0 \\ J_{31} & J_{32} & J_{33} & J_{34} \\ 0 & 0 & 0 & J_{44} \end{pmatrix}$$

where

$$\begin{aligned} J_{11} &= -(\alpha I_t + \mu + v)\Delta t + 1 \\ J_{12} &= \alpha I_t \Delta t \\ J_{14} &= v\Delta t \\ J_{22} &= -(\beta + \mu)\Delta t + 1 \\ J_{23} &= \beta \Delta t \\ J_{31} &= -\alpha S_t \Delta t \\ J_{32} &= \alpha S_t \Delta t \\ J_{33} &= -(\mu_i + \delta + \mu)\Delta t + 1 \\ J_{34} &= \delta \Delta t \\ J_{44} &= -\mu \Delta t + 1 \end{aligned}$$

To implement the Extended Kalman Filter method, we must first create a system matrix and a measurement matrix in the row format: [5, 8, 9].

$$X_{k+1} = \phi_k X_k + W_k$$

$$Z_k = H_k X_k + V_k$$

where :

$$\phi_k = \begin{pmatrix} J & 0 \\ 0 & I \end{pmatrix}$$

$$X_k = \begin{pmatrix} S & E & I & R & \mu & \alpha & v & \beta & \mu_i & \delta \end{pmatrix}^T$$

W_k is the system noise, V_k is the measurement noise, and H_k is the measurement matrix in the following form.

$$H_{k2\times 10} = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & \ldots & 0 \\ 0 & 0 & 0 & 1 & 0 & \ldots & 0 \end{pmatrix}$$

At the prediction stage, an error covariance matrix (Q_k) is needed in the follong form:

$$Q_k = \begin{pmatrix} var(S, E, I, R)_{4\times 4} & 0 \\ 0 & I_{6\times 6} \end{pmatrix}.$$

The variance values for infection and recovery were calculated from the variance of accurate data, yielding 2.1×10^8 and 1.7×10^8, respectively. In contrast, the variance values for suspected and exposed cases, for which no actual data exist, are assumed to have greater variance, set at 2.3×10^8.

Actual data were sourced from the Indonesian COVID-19 website, spanning the period from November 27, 2021, to October 5, 2022. This dataset was divided into 282 training data points (November 27, 2021, to September 4, 2022) and 31 testing data points (September 5, 2022, to October 5, 2022). The graphical comparison between the actual infection cases and the EKF estimation results during the training period is shown in Figure 4.

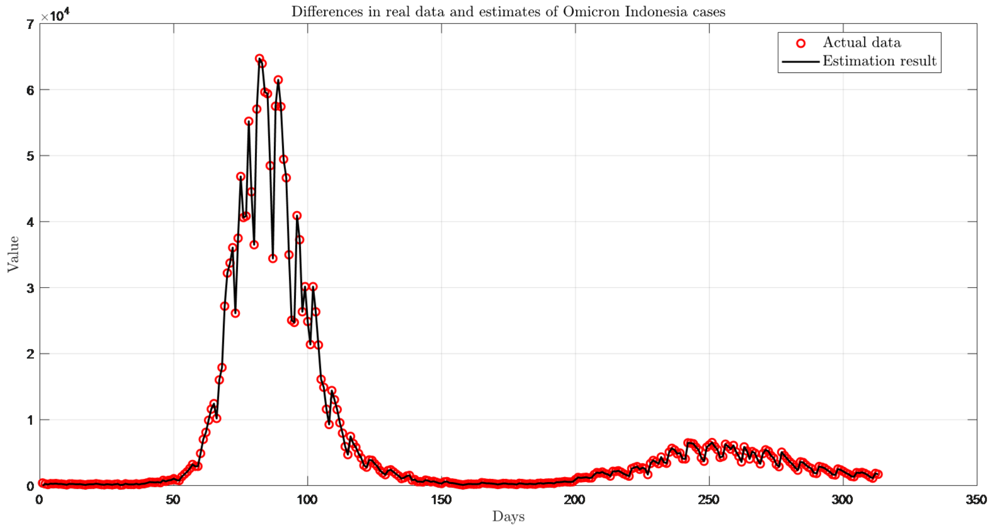

Figure 4: Comparison between actual COVID-19 Omicron case data and EKF estimated values during the training period.

The error generated from the graph is 2.3552×10^{-6}. Thus, estimates using the Extended Kalman Filter are well used in the COVID-19 model chosen for this study.

3.3. Experiment with mEKF and Moving Average using COVID-19 data from Indonesia

The Autocorrelation Function (ACF) measures the linear relationship between observations and their lagged values in a time-series. Meanwhile, the Partial Autocorrelation Function (PACF) isolates the direct effect of a given lag by controlling for the influence of intermediate lags. These functions are used to identify temporal dependency patterns and to determine suitable window length in moving average-based prediction. The Modified Extended Kalman Filter (MEKF) incorporates the Partial Autocorrelation Function (PACF) and Autocorrelation Function (ACF) to enhance its performance in analyzing time-series data. ACF provides insights into the overall relationship between current data points and their lagged counterparts, helping to identify autocorrelation patterns. On the other hand, PACF focuses on the direct influence of specific lags by isolating their impact while controlling for intermediate lags. This distinction allows MEKF to pinpoint significant data relationships, optimizing its ability to handle complex and dynamic datasets.

By utilizing PACF and ACF, MEKF improves prediction accuracy and reduces uncertainty during correction. These functions help determine relevant lags and inform the selection of optimal parameters for the algorithm. Furthermore, they help detect non-linear patterns in the data and incorporate them into the filter's predictive and corrective steps. As a result, the integration of PACF and ACF ensures that MEKF delivers more reliable and precise estimates, even when faced with imperfect or noisy data.

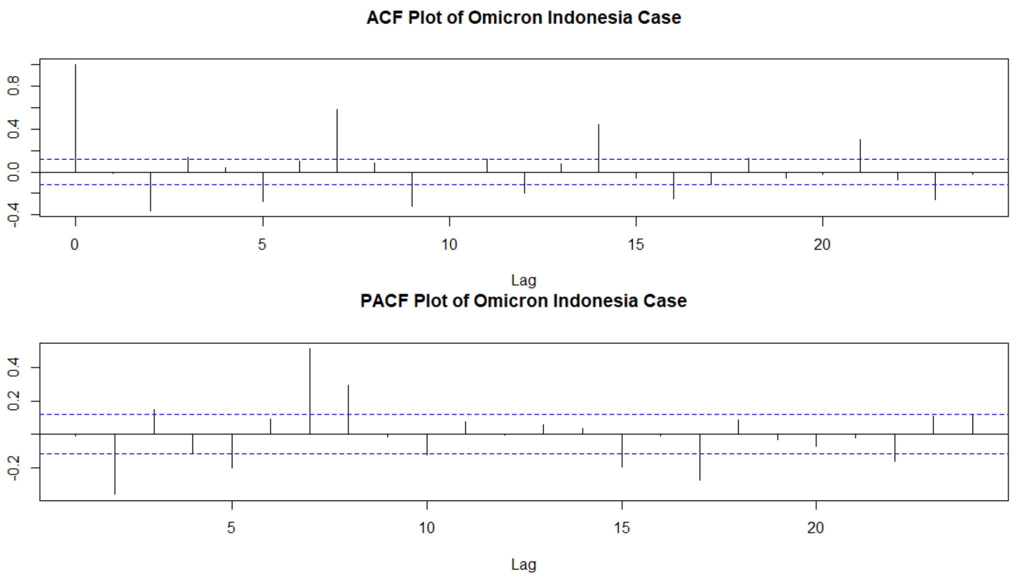

Figure 5: Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) of COVID-19 Omicron infection cases, indicating weekly periodicity.

The temporal dependency structure of the COVID-19 infection data was examined using the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF). The patterns observed in these plots, as shown in Figure 5, indicate a clear weekly cyclic behavior. This periodicity serves as the rationale for determining the window length in the Moving Average component of the mEKF method [3, 10].

The training data are then used to implement both mEKF and the Moving Average method, allowing a direct comparison of their estimation performance.

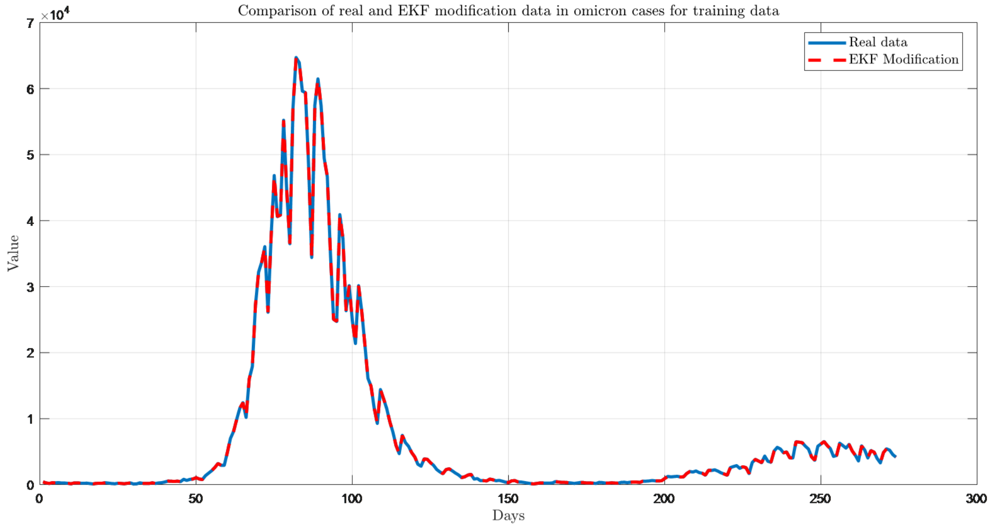

Figure 6: Comparison of actual and mEKF-estimated infection cases during the training period.

Implementing training data with mEKF for infection cases yields an error of 0.0038731 (Figure 6). The error associated with implementing training data using moving averages in Figure 7 is 0.195916

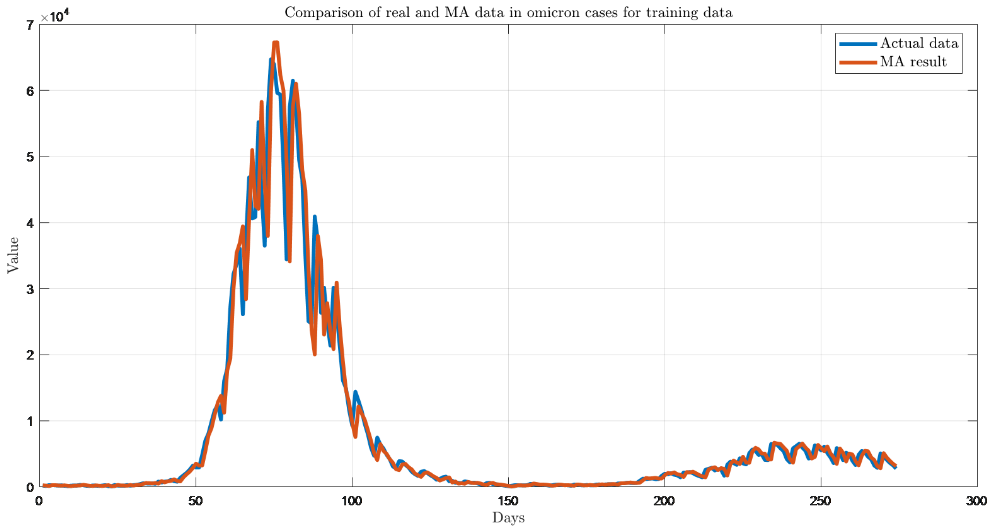

Figure 7: Prediction results of infection cases using the Moving Average method during the training period.

Based on the error value of how training data was utilized with mEKF and MA, it was determined that mEKF produced better results. The subsequent implementation of testing data using mEKF and MA was conducted from September 5 to October 5, 2022, the results can be seen in Figure 8.

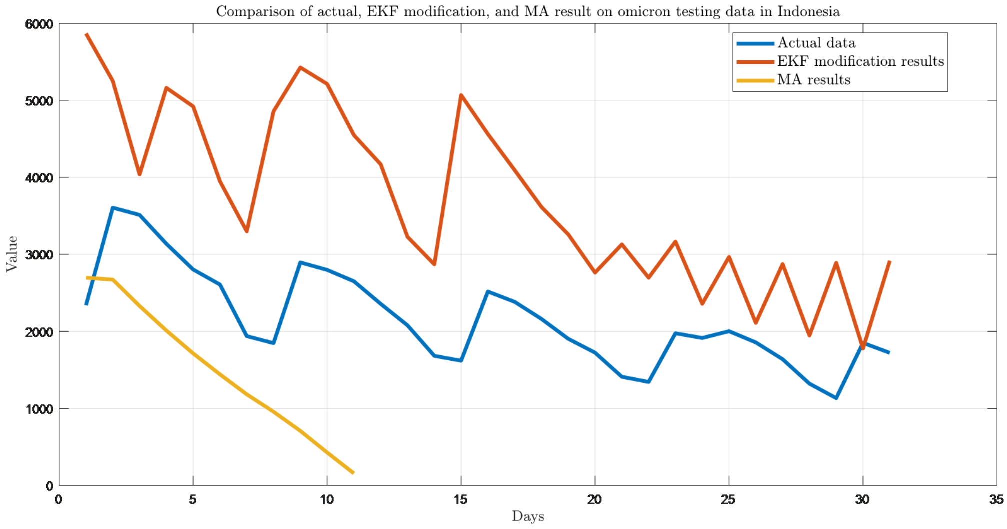

Figure 8: Comparison of MA and mEKF predictions for the next 30 days based on COVID-19 Omicron data.

It was concluded that MEKF could predict for 30 days and predict actual data trends, whereas MA could only predict for 11 days. Moreover, it is incapable of predicting the trends in actual data. The next step is to iterate every 11 days to assess the suitability of the predicted mEKF and MA results, as shown in the Table 2, Table 3 and Table 4.

Table 2: Comparison between actual infection data, Moving Average predictions, and mEKF predictions for the period 5–15 September 2022.

Date	Actual data	Prediciton of MA	Prediction of EKF modification
5 Sept 2022	2340	2698	3130
6 Sept 2022	3607	2673	2698
7 Sept 2022	3513	2331	3167
8 Sept 2022	3138	2012	2359
9 Sept 2022	2804	1715	2968
10 Sept 2022	2609	1443	2111
11 Sept 2022	1939	1183	2875
12 Sept 2022	1848	957	1947
13 Sept 2022	2896	709	2891
14 Sept 2022	2799	428	1778
15 Sept 2022	2651	156	2919

Table 3: Comparison between actual infection data, Moving Average predictions, and mEKF predictions for the period 16–26 September 2022.

Date	Actual data	Prediction of MA	Prediciton of EKF modification
16 Sept 2022	2358	2582	2697
17 Sept 2022	2079	2550	2582
18 Sept 2022	1683	2542	2682
19 Sept 2022	1620	2628	2578
20 Sept 2022	2518	2739	2788
21 Sept 2022	2384	2716	2682
22 Sept 2022	2162	2705	2773
23 Sept 2022	1904	2712	2665
24 Sept 2022	1724	2731	2790
25 Sept 2022	1411	2757	2661
26 Sept 2022	1344	2787	2820

Table 4: Comparison between actual infection data, Moving Average predictions, and mEKF predictions for the period 27 September–7 October 2022.

Date	Actual data	Prediciton of MA	Prediction of EKF modification
27 Sept 2022	1976	1305	1372
28 Sept 2022	1915	1132	1305
29 Sept 2022	2003	953	1209
30 Sept 2022	1857	780	1151
1 Oct 2022	1639	619	1073
2 Oct 2022	1322	461	1043
3 Oct 2022	1134	326	980
4 Oct 2022	1851	180	991
5 Oct 2022	1722	19	927
6 Oct 2022	1831	-	936
7 Oct 2022	1501	-	874

The third iteration demonstrates that MA cannot predict infection cases until 7 October 2022, whereas mEKF can. To further demonstrate the superiority of mEKF over MA, a numerical comparison table will be provided in Table 5.

Table 5: Statistical comparison of real data, mEKF predictions, and Moving Average predictions during three prediction intervals.

Date	Method	Error	Mean	Variance	Minimal	Maximal
5 - 15	Actual data	-	2740.364	316188.055	1848	3607
Sept	MA	0.487	1481.842	778759.485	155.633	2697.714
2022	EKF modification	0.199	2621.873	240355.846	1777.700	3167
16 - 26	Actual data	-	1926.091	162387.891	1344	2518
Sept	MA	0.451	2676.635	7470.444	2541.152	2786.847
2022	EKF modification	0.462	2701.264	6778.665	2577.500	2819.400
17 Sept -	Actual data	-	1713.222	90941.444	1134	2003
7 Oct	MA	0.637	641.207	191154.874	18.912	1304.571
2022	EKF modification	0.336	1116.289	23560.389	926.990	1372.100

The numerical comparison table above shows that the mEKF results are better than those of the MA method. To further strengthen this conclusion, both methods were evaluated using a different set of testing data collected between September 12 and October 13, 2022. The comparison of prediction performance during the testing period is shown in Figure 9. The mEKF method is able to maintain stable and consistent forecasts throughout the 30-day horizon, whereas the MA method begins to diverge after approximately 11 days, indicating its limited capability for longer-term prediction. This finding is also reflected in the magnitude of the prediction deviations, where the mEKF remains closer to the actual infection data across the forecasting interval, while the MA method gradually drifts away from the true trend. These results suggest that the mEKF provides a more robust and reliable short-term forecasting approach compared to the MA method.

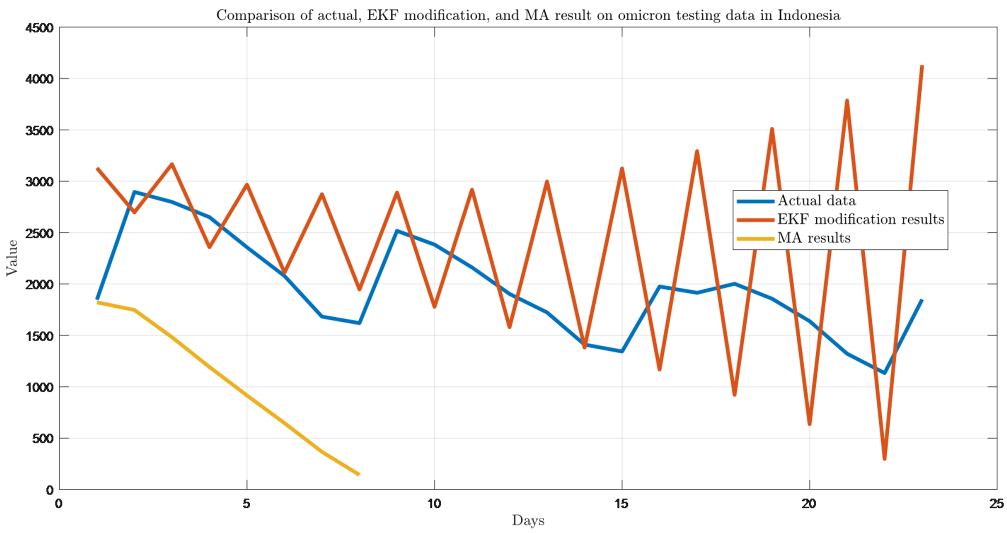

Figure 9: Comparison of mEKF and MA predictions during the testing period to evaluate short-term prediction accuracy.

Evidently, the MA method can only produce reliable predictions for the first eight days of the testing period. This limitation arises because the MA method depends directly on recent observed values, making it sensitive to short-term fluctuations and unable to maintain stable forecasts over longer horizons. Therefore, in the next stage, both MA and mEKF will be evaluated over the same eight-day prediction window to ensure a fair and consistent comparison of performance.

The prediction results for the testing phase are summarized in Tables 6–9, which present the actual infection data alongside the MA and mEKF predictions across several testing intervals. Overall, the mEKF predictions remain closer to the observed data and retain stability over time, while the MA predictions tend to follow short-term fluctuations and diverge as the forecasting horizon increases. This pattern demonstrates that the mEKF is better suited to capture underlying transmission dynamics rather than merely reflecting recent case counts. A summary of performance across all test intervals is provided in Table 10, which further confirms that the mEKF achieves more accurate and stable predictions than the MA method, particularly in short-term forecasting scenarios.

Table 6: Comparison of actual infection data, Moving Average predictions, and mEKF predictions for the period Sept 12 - 19, 2022.

Date	Actual data	Prediction of MA	Prediction of EKF Modification
Sept 12, 2022	1848	1822	3130
Sept 13, 2022	2896	1748	2698
Sept 14, 2022	2799	1483	3167
Sept 15, 2022	2651	1193	2359
Sept 16, 2022	2358	916	2968
Sept 17, 2022	2079	647	2111
Sept 18, 2022	1683	367	2875
Sept 19, 2022	1620	143	1947

Table 7: Comparison of actual infection data, Moving Average predictions, and mEKF predictions for the period Sept 20 - 27, 2022.

Date	Actual data	Prediction of MA	Prediction of EKF Modification
Sept 20, 2022	2518	1588	1647
Sept 21, 2022	2384	1401	1588
Sept 22, 2022	2162	1201	1468
Sept 23, 2022	1904	994	1415
Sept 24, 2022	1724	799	1299
Sept 25, 2022	1411	616	1280
Sept 26, 2022	1344	464	1188
Sept 27, 2022	1976	298	1222

Table 8: Comparison of actual infection data, Moving Average predictions, and mEKF predictions for the period Sept 28 - Oct 5, 2022.

Date	Actual data	Prediction of MA	Prediction of EKF Modification
Sept 28, 2022	1915	1899	1305
Sept 29, 2022	2003	1830	1899
30 Sept, 2022	1857	1782	1151
Oct 1, 2022	1639	1765	1861
Oct 2, 2022	1322	1770	1043
Oct 3, 2022	1134	1822	1881
Oct 4, 2022	1851	1890	991
Oct 5, 2022	1722	1888	1958

Table 9: Comparison of actual infection data, Moving Average predictions, and mEKF predictions for the period Oct 6 - 13, 2022.

Date	Actual data	Prediction of MA	Prediction of EKF Modification
Oct 6, 2022	1831	1695	1834
Oct 7, 2022	1501	1651	1695
Oct 8, 2022	1325	1621	1809
Oct 9, 2022	999	1619	1672
Oct 10, 2022	1195	1661	1834
Oct 11, 2022	2077	1736	1722
Oct 12, 2022	2028	1720	1934
Oct 13, 2022	1830	1719	1702

Table 10: Comparison of errors and statistical value from real data, EKF modification, and MA of COVID-19 infection in Indonesia on Sept 12 - Oct 13, 2022

Date	Method	Error	Mean	Variance	Minimal	Maximal
Sept 12 - 19, 2022	Actual data	-	2241.750	257124.5	1620	2896
	MA	0.553	1039.875	394258.411	143	1822
	EKF modification	0.273	2565.875	217253.554	1947	3167
Sept 20 - 27, 2022	Actual data	-	1927.875	180518.982	1344	2518
	MA	0.539	919.694	209071.771	298	1588
	EKF modification	0.262	1388.375	28764.268	1188	1647
Sep 28 - Oct 5, 2022	Actual data	-	1680.375	92835.411	1134	2003
	MA	0.158	1828.917	2968.652	1765	1899
	EKF modification	0.295	1511.125	181601.839	991	1958
Oct 6 - 13, 2022	Actual data	-	1598.250	161085.929	999	2077
	MA	0.223	1677.297	2127.822	1619	1736
	EKF modification	0.249	1775.250	8363.643	1672	1934

The numerical comparison and prediction iterations demonstrate that the mEKF method provides more accurate and stable forecasts than the MA method in predicting

COVID-19 infection cases in Indonesia. This improvement reflects the ability of mEKF to incorporate the underlying transmission dynamics of the SEIR model, allowing it to maintain prediction reliability even when the data exhibit fluctuations.

4. Discussion and Conclusion

This section presents the comparison between the Modified Extended Kalman Filter (mEKF) and the Moving Average (MA) method. The initial step involves selecting a mathematical model. This study uses the SEIR model [5] and COVID-19 Omicron case data from Indonesia. To implement EKF, the continuous deterministic dynamic model is converted into a discrete form and linearized using Jacobians. The suitability between the model structure, parameter values, and real data is reflected in the small estimation error, indicating that the selected SEIR model appropriately represents the observed infection dynamics.

The next step involves separating the dataset into training and testing subsets. The training data are used to compare the estimation and prediction performance of mEKF and MA. The results show that mEKF achieves lower prediction error than MA. Meanwhile, MA predictions depend heavily on historical observations based on ACF and PACF characteristics, causing MA to follow the recent trend of the data: decreasing when recent values decrease and increasing when they increase. In contrast, mEKF maintains the dynamic structure of the disease transmission model, leading to more stable forecasts even when actual data are unavailable. The testing results further show that mEKF can generate predictions up to 30 days ahead, whereas MA is limited to approximately 11 days. Repeated experiments with different testing periods reinforce the conclusion that mEKF provides more consistent predictive accuracy over time.

The predictive capability of mEKF provides valuable insight for short-term epidemic response planning. In practical applications, 30-day ahead forecasts can support public health authorities in adjusting hospital bed availability, testing strategies, and vaccine distribution. Unlike purely statistical predictors, mEKF maintains internal consistency with the SEIR transmission model, ensuring that the resulting predictions remain epidemiologically interpretable. This interpretability is essential when predictions are used to guide public health interventions and mobility control policies [11].

Other forecasting approaches, such as ARIMA and deep learning-based time-series models (e.g., LSTM), have also been widely applied to COVID-19 prediction [10]. However, these data-driven models do not explicitly incorporate epidemiological transmission mechanisms, which may reduce interpretability when used for policy decision-making. Hybrid approaches that combine SEIR models with filtering techniques have been shown to improve predictive performance while preserving epidemiological structure [12]. In this context, mEKF provides a balance between prediction accuracy and model interpretability, making it suitable for short-term infectious disease forecasting [11].

In conclusion, the results demonstrate that the Modified Extended Kalman Filter (mEKF) is more effective than the Moving Average method for predicting COVID-19 Omicron case trends in Indonesia. The mEKF approach provides stable, accurate, and interpretable short-term forecasts, making it a promising method for real-time epidemic forecasting and public health decision support.

References

[1] P.S. Maybeck. *Stochastic Models, Estimation, and Control.* Academic Press, Inc., San Diego, California, 1999.

[2] H.N. Fadhilah, E. Apriliani, and D.K. Arif. Prediksi Penyebaran Covid-19 di Indonesia dan Jawa Timur dengan Metode Extended Kalman Filter. *LIMITS*, 18(1):1–13, 2021.

[3] L.L. Frank, Xie, and D. Popa. *Optimal and Robust Estimation with an Introduction to Stochastic Control Theory, Second ed.* CRC Press, Boca Raton, FL, 2008.

[4] J.E. Hanke and D. Wichern. *Business Forecasting, 9th Ed.* Pearson Education Limited, London, UK, 2014.

[5] S. Annas et al. Stability analysis and numerical simulation of seir model for pandemic covid-19 spread in indonesia. *Chaos, Solitons and Fractals*, 139:110072, 2020.

[6] COVID-19 Response Acceleration Task Force. Peta Sebaran. Website, May 15 2022. www.covid19.go.id.

[7] Nastitie. Estimations and Short-Term Predictions Cases of COVID-19: Study Cases in Indonesia. Master's thesis, Institut Teknologi Bandung, 2022.

[8] D.M. Causon and C.G. Mingham. *Introductory Finite Difference Methods for PDEs.* Ventus Publishing Aps., Erie, Colorado, 2010.

[9] G. Chen. *Series in Approximations and Decompositions – Vol. 2: Approximate Kalman Filtering.* World Scientific, London, UK, 1994.

[10] Z. Yang et al. Covid-19 epidemic prediction using time-series deep learning models. *Scientific Reports*, 13:14522, 2023.

[11] A. Acheampong. Predictive accuracy of extended kalman filter in epidemiological forecasting. *Journal of Infectious Disease Modelling*, 7:1–12, 2023.

[12] T. Tran et al. Hybrid seir and kalman filtering framework for covid-19 forecasting. *Applied Mathematical Modelling*, 105:515–529, 2022.