

Optimizing COVID-19 Epidemiological Models: A Particle Swarm Approach to Parameter Estimation

Muna Afdi Muniroh¹, Sekar Sari², Dika Agustia Indrati³

^{1,3}University of PGRI Ronggolawe, ²Tamansiswa International Polytechnic of Mojokerto

Corresponding Author: Muna Afdi Muniroh, e-mail: munaafdimuniroh@gmail.com

Published: July , 2025

ABSTRACT

The SEIR (Susceptible-Exposed-Infectious-Recovered) mathematical model, represented as a system of nonlinear differential equations, has proven to be a powerful tool to describe the dynamics of the spread of infectious diseases such as COVID-19. The accuracy of the projection and understanding of this model relies heavily on the proper estimation of its parameters, such as transmission rate, incubation rate, natural birth rate, natural death rate, disease mortality rate, and recovery rate. This study focuses on the development and application of a new approach to estimate crucial parameters in the SEIR model by utilizing Particle Swarm Optimization (PSO). PSO is a metaheuristic optimization algorithm inspired by the social behavior of flocks of birds or schools of fish. PSO was chosen for its outstanding ability to find a global minimum in a complex search space, as well as its efficiency in handling nonlinear optimization problems. The advantage of PSO lies in its effective memory capacity, which allows the storage of previous best values, both individually and globally, thus accelerating convergence to the optimal solution. Through a simulation program, this study successfully identified the optimal set of parameters for the SEIR model. These estimated parameters were then carefully evaluated by comparing the model simulation outputs with available COVID-19 epidemiological data, demonstrating the model's ability to accurately replicate pandemic trends. The results of this study are expected to make a significant contribution to modelling and understanding the spread of COVID-19.

Keywords: Covid-19, SEIR Model, Particle Swarm Optimization

© The Author(s). 2021 Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0>

INTRODUCTION

The coronavirus family includes viruses that can cause disease in humans and other animals. Multiple coronaviruses infect people, causing respiratory infections and other serious illnesses including MERS and SARS. The identification of a new coronavirus, SARS-CoV-2, led to the emergence of a new illness, Covid-19. It was in late December 2019 in Wuhan City, China, when this sickness was first identified (Who, 2020). On March 2, 2020, the first cases of verified COVID-19 were reported in Indonesia; since then, the virus has spread globally. As of October 31, 2020, there have been 1,157,509 deaths and 43,341,451 confirmed positive cases of the COVID-19 pandemic. A total of 13,612 people have lost their lives and 400,483 have tested positive in Indonesia. Education, mental health, and access to vital health services are all negatively impacted by social constraints and school closures. All nations, but especially Indonesia, are facing a social and economic catastrophe on a scale never seen before because of this (UNICEF, 2020).

Mathematical models of disease spread are recognized mechanisms to determine the dynamics of disease spread. Kermack and McKendric in 1927 introduced the SIR (Susceptible-Infected-Recovered) model which is a system of nonlinear differential equations and can describe the spread of a disease in the population. In addition, there are also some individuals who are infected with the

disease but do not transmit it to other individuals. Such individuals are called exposed individuals, hence the basis for the formation of the SEIR (Susceptible, Exposed, Infected, Recovered) type epidemic model. Several researchers have conducted research by making the SEIR model a reference for constructing the Covid-19 spread model (Carcione, et al., 2020).

Parameter estimate is a technique employed to ascertain the values of parameters inside a mathematical model, facilitating the model's application to real-world scenarios (Lazzus, et al., 2020). Each nation possesses distinct Covid-19 data; parameter estimate is deemed effective if the model's result closely aligns with the actual data value. Consequently, a method is required that encompasses a search space to identify the optimal parameter value.

Optimization methods consist of exact methods and approximation methods. Exact method is a method that guarantees an optimal solution, but for large-scale problems, it cannot be easily solved using this method because it requires a very long computation time, making this method inefficient. Therefore, an approximation method such as the heuristic method is needed. The metaheuristic method is an approach that basically includes the heuristic method. The metaheuristic technique is an iterative generation process that demonstrates exploration and exploitation of the search space to obtain a solution that approximates optimality. Some examples of metaheuristic methods are GA, PSO, Ant Colony Optimization, Tabu Search, and Simulated Annealing (Talbi, 2009). Metaheuristic approaches frequently employed by academics to address parameter estimation issues in mathematical models include Genetic Algorithms (GA) and Particle Swarm Optimization (PSO), both capable of identifying a global minimum. Godio, et al., (2020) conducted parameter estimation using PSO on the SEIR type Covid-19 disease model and matched it with Covid-19 data in Italy. It is known that PSO has effective memory capabilities in storing previous best values when compared to GA. This study focuses on parameter estimation with PSO which shows the suitability of the Covid-19 model with Covid-19 data in Indonesia from March 2, 2020 to October 31, 2020.

METHOD

Model SEIR

The mathematical model of this research refers to the article (Carcione, et al., 2020) which discusses Covid-19 using a mathematical model of type . This research estimates the parameters of the model and matches them with Covid-19 data in Indonesia. It is assumed that susceptible individuals are infected with an infection rate $\frac{\beta SI}{N}$ that represents the average number of contact incidents per unit time of an infected individual to other individuals. The SEIR type Covid-19 spread model is divided into four compartments, namely S is a vulnerable individual who has the possibility of being infected, for example individuals who are in a hospital environment, individuals who often meet with strangers, do not use masks and so on, E is an exposed individual, I is a positive individual infected with Covid-19 and R is a recovered individual. The model is in the form of a nonlinear differential equation as follows.

$$\begin{aligned} S &= \Lambda - \frac{\beta SI}{N} - \mu S \\ E &= \frac{\beta SI}{N} - \delta E - \mu E \\ I &= \delta E - \gamma I - \mu I - \alpha I \\ R &= \gamma I - \mu R \end{aligned} \tag{1}$$

with initial conditions $S(0) \geq 0, E(0) \geq 0, I(0) \geq 0, R(0) \geq 0$ and $N = S + E + I + R$. The parameters in model (1) include Λ is the natural birth rate, β is the transmission rate, μ is the natural death rate, δ is the rate at which exposed people transition to infectious ones (incubation), γ is the recovery rate, and α is the disease mortality rate. The basic reproduction number ($\mathcal{R}_0$) is a key metric for comprehending the potential transmission of disease. $\mathcal{R}_0$ for model (1) is expressed as follows.

$$\mathcal{R}_0 = \frac{\beta\delta}{(\delta + \mu)(\gamma + \mu + \alpha)}$$

Objective Function

An n-dimensional system of equations $\dot{X} = (X, X_0, \theta_0)$ with $X = (x_1, x_2, \dots, x_n)^T \in \mathbb{R}^n$ is the state vector, X_0 is the initial state, dan $\theta_0 = (\theta_{01}, \theta_{02}, \dots, \theta_{0n})^T$ is the set of original parameters. When performing parameter estimation, the estimated system is $\dot{Y} = (Y, X_0, \theta)$ with $Y = (y_1, y_2, \dots, y_n)^T \in \mathbb{R}^n$ is the estimated state vector and $\theta = (\theta_1, \theta_2, \dots, \theta_n)^T$ is the set of estimated parameters (He, 2007).

Thus, the objective function can be formulated as the following RMSE optimization:

$$\text{RMSE} = \sqrt{\frac{1}{M} \sum_{k=1}^M (x_k - y_k)^2}, \text{ by determining } \theta^*,$$

with M denoted as the length of data used for parameter estimation, X_k and Y_k is the original and estimated state at time k .

Particle Swarm Optimization (PSO)

Particle swarm optimization (PSO) is an optimization technique that approximates the social behaviour of organisms, such as colonies of birds or schools of fish, in their pursuit of survival. Social behaviour indicates the actions of individuals that are influenced by the presence of others within a group. Each particle (individual) within the population (swarm) is identified by two attributes: velocity and location in PSO. The PSO method starts with the random initialization of each particle's location and velocity. The particles traverse a multidimensional search space. Each particle communicates its experience while seeking its optimal location to get the singular best position inside the swarm. Then the position and velocity of each particle is adjusted based on the best experience of each particle.

The position of a particle i in the search space is denoted as $X_i = [x_{i1}, x_{i2}, \dots, x_{in}]$ and the rate of change of position (velocity) for each particle i is denoted by $V_i = [v_{i1}, v_{i2}, \dots, v_{in}]$. The best position of each particle called the local best (p_{best}) is stored and denoted as $P_i = [p_{i1}, p_{i2}, \dots, p_{in}]$. The best position based on the experience of all particles in the population is called the global best (g_{best}) and is denoted by $P_g = [p_{g1}, p_{g2}, \dots, p_{gn}]$. Changes in velocity and position for each particle occur based on the following equation:

$$\begin{aligned} v_{id}(t+1) &= w_i v_{id}(t) + c_1 r_1 (p_{id}(t) - x_{id}(t)) + c_2 r_2 (p_{gd}(t) - x_{id}(t)) \\ x_{id}(t+1) &= x_{id}(t) + v_{id}(t+1) \end{aligned}$$

with d is the dimension $1 \leq d \leq n$, w is the inertial weight used to balance the local and global search, c_1 and c_2 is a positive constant, r_1 and r_2 are random numbers in range $[0,1]$. $c_1 r_1$ and $c_2 r_2$

are respectively the memory that stores the best experience of each particle and the shared information that conveys the best experience of all particles in a swarm. (Gao, 2009).

FINDING AND DISCUSSION

Based on the life expectancy in Indonesia in 2020 being 71.5, it is assumed that the parameter value $\mu = 1/71.5$. The other five parameters in model (1) are estimated using PSO. Given the initial conditions, namely $S(0)=9906$, $E(0)=355$, $I(0)=2$, $R(0)=0$. Simulation of the program with PSO involves maximum iterations for the calculation of the objective function value given as 200 and particles (position and velocity) as many as 10 according to the parameters sought. Then given $r_{1,2}$ randomly, the value of the inertial weight $w=0.729$, and the constant $c_{1,2}=1.494$. This research is a global search (swarm) of 10, then produces the best swarm with the smallest error. The PSO program was simulated using Matlab R2021a software, and the results of the objective function value (RMSE) for each swarm are shown in Table 1.

Table 1. Results for each swarm

Swarm	Λ	β	δ	γ	α	RMSE	$\mathcal{R}_0$	Time (second)
1 st	0.015	0.6122	0.1945	0.0349	0.0010	337.3470	11.4601	140.4551
2 nd	0.015	0.6129	0.1945	0.0349	0.0010	337.3471	11.4732	139.8346
3 rd	0.015	0.6129	0.1945	0.0349	0.0010	337.3471	11.4731	138.9850
4 th	10.000	1.2997	0.1000	0.0351	0.0005	350.9991	29.9362	139.3198
5 th	0.015	0.6140	0.1945	0.0349	0.0010	337.3471	11.4936	137.9372
6 th	0.015	0.6129	0.1945	0.0349	0.0010	337.3471	11.4728	138.6190
7 th	0.015	0.6129	0.1945	0.0349	0.0005	337.3471	11.4729	152.4469
8 th	10.000	0.6140	0.1942	0.0353	0.0010	337.3689	11.4952	144.8018
9 th	0.015	1.2160	0.1000	0.0185	0.0010	351.0172	31.8368	264.9265
10 th	0.015	0.6141	0.1945	0.0349	0.0010	337.3472	11.4941	283.5088

Table 1 shows that a larger computation time does not guarantee a smaller objective function value (RMSE). The largest RMSE and $\mathcal{R}_0$ values are obtained at the time of the 9th swarm, while the smallest RMSE and $\mathcal{R}_0$ values produce the best parameters at the time of the 1st swarm, namely

$$\Lambda=0.015, \quad \beta=0.6122, \quad \delta=0.1945, \quad \gamma = 0.0349, \quad \alpha = 0.001.$$

Then the best parameter and $\mu = 1/71.5$ are used in model (1), so that the value of the basic reproduction number is obtained $\mathcal{R}_0 = 11.4601$ and a visualization of the comparison of model (1) with Covid-19 data in Indonesia is shown in Figure 1.

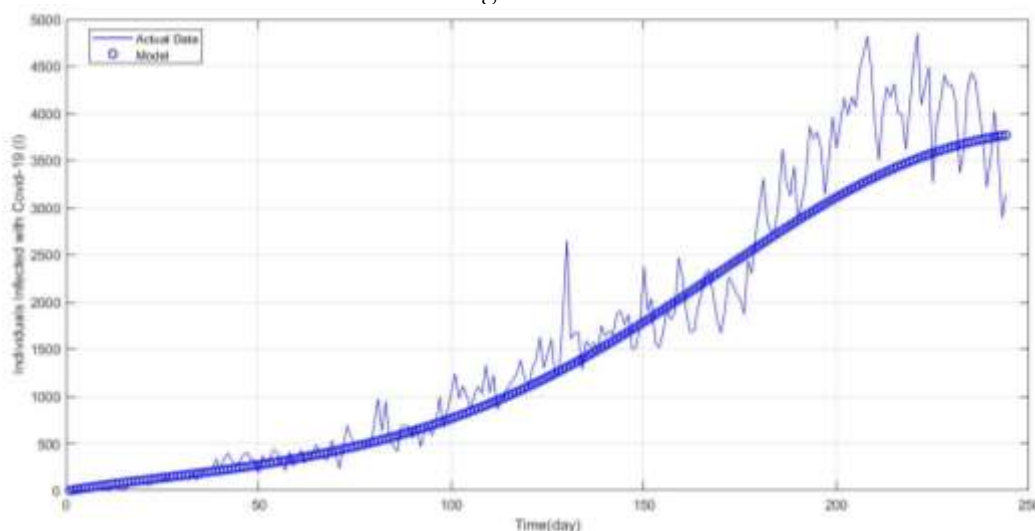


Figure 1. Comparison between actual data and model with the best PSO parameters.

Figure 1 illustrates that the model curve derived from the SEIR model parameter estimation effectively represents the initial dynamics of COVID-19 spread in Indonesia by aligning with the trend of actual data during the early period of the pandemic.

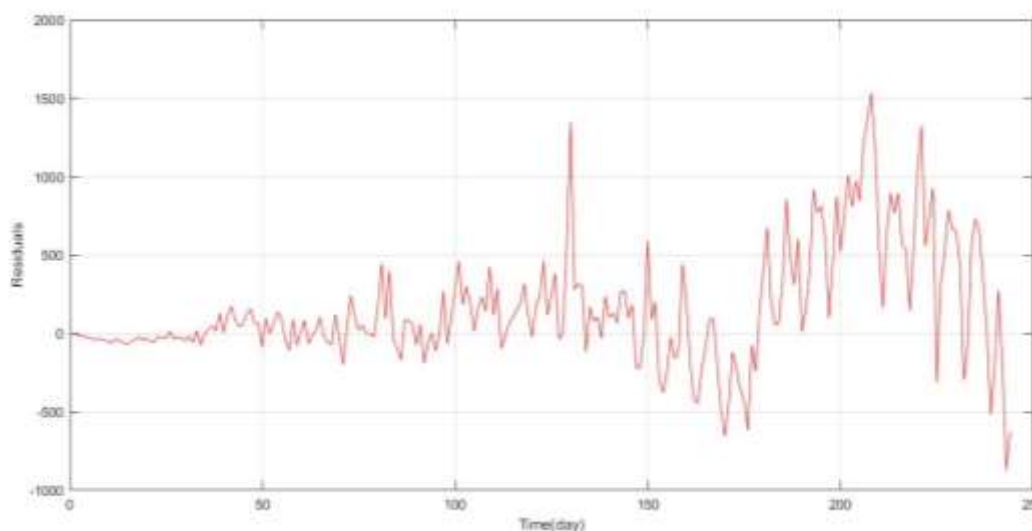


Figure 2. Fitting Errors

However, the model struggles to capture the more complex fluctuations and sharper dips in the actual data, especially in the late phase of the pandemic. This is clearly seen in Figure 2, where the residual plot shows that the residuals at the beginning of the period are close to zero (indicating a good fit), but after the peak of the epidemic, there are some large spikes in the residuals as well as a non-random pattern. These residual spikes indicate the presence of dynamics that are not fully captured by the SEIR model, most likely due to significant public health interventions such as mass vaccination programs or the implementation of strict social restrictions.

Therefore, to improve the accuracy and predictive ability of the model, future research could focus on developing a more complex SEIR model. This could include adding new parameters or compartments that explicitly represent public health interventions, such as vaccination and quarantine effects, similar to the approach discussed by Nisardi, et al. (2023).

CONCLUSIONS

Based on the results and discussion, it can be concluded that parameter estimation using Particle Swarm Optimization (PSO) can obtain parameters that are quite suitable for the model. The best parameter values from the simulation results are

$$\Lambda=0.015, \quad \beta=0.6122, \quad \delta=0.1945, \quad \gamma = 0.0349, \quad \alpha = 0.001.$$

with an objective function value of 337.3470.

REFERENCES

- Annas, S., Pratama, M. I., Rifandi, M., Sanusi, W., & Side, S. (2020). Stability Analysis and Numerical Simulation of SEIR Model for Pandemic COVID-19 Spread in Indonesia. *Chaos, Solitons and Fractals*. Vol 139.
- Carcione, J. M., Santos, J. E., Claudio, B., & Bang, J. (2020). A Simulation of a COVID-19 Epidemic, Based on a Deterministic SEIR. Model. *Frontiers in Public Health*. 8:230.
- CDC. (2020). Symptoms of Coronavirus Disease <https://www.cdc.gov/coronavirus/2019-ncov/symptoms-testing/symptoms.html>.
- Chen, M., Li, M., Hao, Y., Liu, Z., Hu, L., & Wang, L. (2020). The Introduction of Population Migration to SEIAR for COVID-19 epidemic modeling with an Efficient Intervention Strategy. *Information Fusion*. 64:252-258.

- Gao, F., Lee, J. Z., Tong, Li. H., & Lu, X. (2009). Parameter Estimation for Chaotic System with Initial Random Noises by Particle Swarm Optimization. *Chaos, Solitons and Fractals*. 42:1286-1291.
- Godio, A., Pace, F., & Vergnano, A. (2020). SEIR Modeling of The Italian Epidemic of SARS-CoV-2 using Computational Swarm Intelligence. *International Journal of Environmental Research and Public Health*.
- He, Q., Wang, L., & Liu, B. (2007). Parameter Estimation for Chaotic Systems by Particle Swarm Optimization. *Chaos, Solitons and Fractals*. 34:654-661.
- Kermack, W. O., & McKendrick, A. G. (1927). A Contribution of the Mathematical Theory of Epidemics. *The Royal Society*. 115:700-721.
- Lazzus, J. A., Vega-Jorquera, P., Lopez-Caraballo, Palma-Chilla, C. H., L., & Salfate, I. (2020) Parameter Estimation of a Generalized Lotka–Volterra System Using a Modified PSO Algorithm. *Applied Soft Computing Journal*. Vol 96.
- Lazzus, J. A., Rivera, M., & López-Caraballo, C.H. (2016). Parameter Estimation of Lorenz Chaotic System Using a Hybrid Swarm Intelligence Algorithm. *Physics Letters A*. 380:1164-1171.
- Nisardi, M. R., Sulma, Hukmah, Suriani, & Toaha, S. (2023). Mathematical Model of COVID-19 with Quarantine and Vaccination. *Jurnal Matematika, Statistika, dan Komputasi*. 19: 266-285.
- Peng, L., Yang, W., Zhang, D., Zhuge, C., & Hong, L. (2020). Epidemic Analysis of COVID-19 in China by Dynamical Modeling. *Natural Public Health Emergency Collection*.
- Shi, Y., & Eberhart, R. C. (2001). Comparing Inertia Weights and Constriction Factors in Particle Swarm Optimization. *Proceedings of the 2000 Congress on Evolutionary Computation*. 1:84-88.
- Talbi, E. (2009). *Metaheuristics from Design to Implementation*. John Wiley & Sons.
- UNICEF. (2020). Covid 19: What You Should Know and How to Protect Yourself. <https://www.unicef.org/indonesia/id/coronavirus>.
- WHO. (2020). Coronavirus Disease (COVID-19) Advice for the Public. <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/advice-for-public>.
- WHO. (2020). Pertanyaan dan Jawaban Terkait Coronavirus <https://www.who.int/indonesia/news/novel-coronavirus/qa-for-public>.
- WHO. (2020). Naming the coronavirus disease (COVID-19) and the virus that causes it. [https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/naming-the-coronavirus-disease-\(covid-2019\)-and-the-virus-that-causes-it](https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/naming-the-coronavirus-disease-(covid-2019)-and-the-virus-that-causes-it).