

Estimation of SIR Epidemiological Model Parameters Using Computational Intelligence Methods Based on COVID-19 Data from Brazil

João Samuel Maciel de Sales

Universidade Federal do Ceará

Av. Mister Hull, s/n - Pici

joaosamuel.2001@gmail.com

Rogério Ferreira Rodrigues Júnior

Universidade Federal do Ceará

Av. Mister Hull, s/n - Pici

rger0904@hotmail.com

João Paulo do Vale Madeiro

Universidade Federal do Ceará

Av. Mister Hull, s/n - Pici

jpaulo.vale@dc.ufc.br

Ricardo Coêlho

Universidade Federal do Ceará

Av. Mister Hull, s/n - Pici

rcoelhos@dema.ufc.br

RESUMO

Este estudo investiga a aplicação de métodos de inteligência computacional (IC) para analisar a dinâmica de transmissão da COVID-19 no Brasil. Utiliza-se um modelo epidemiológico SIR (Suscetível-Infectado-Recuperado) baseado em uma equação diferencial para estimar os casos de COVID-19 no período de fevereiro de 2020 a fevereiro de 2023. Os métodos de IC empregados incluem Busca Aleatória Global, Busca Aleatória Local, Busca Harmônica, Evolução Diferencial e Otimização por Enxame de Partículas. Os resultados indicam que a Evolução Diferencial, especialmente quando aplicada bimestralmente, forneceu as estimativas de parâmetros mais precisas para o modelo.

PALAVRAS CHAVE. Inteligência computacional, Modelo SIR, COVID-19.

Tópicos: Inteligência Computacional, Metaheurísticas, Programação Matemática, Simulação.

ABSTRACT

This study investigates the application of computational intelligence (CI) methods to analyze the transmission dynamics of COVID-19 in Brazil. Utilizing a Susceptible-Infected-Recovered (SIR) epidemiological model based on differential equations, COVID-19 cases are estimated from February 2020 to February 2023. The CI methods employed in this study include Global Random Search, Local Random Search, Harmony Search, Differential Evolution, and Particle Swarm Optimization. The results indicate that Differential Evolution, particularly when applied bimonthly, yielded the most accurate parameter estimations for the model.

KEYWORDS. Computational intelligence. SIR model. COVID-19.

Topics: Computational Intelligence. Metaheuristics. Mathematical Programming. Simulation.

1. Introduction

The progression of human society brings various challenges and adversities, including concerns such as wars, predators, and climatic conditions. However, nothing instills as much fear in society as diseases, particularly viral infectious diseases, which have historically caused significant suffering and loss of life [Schecter et al., 2021]. Infectious diseases can be classified into three levels of contagion: endemics, which occur when a disease recurs in a region without a significant increase in cases; epidemics, marked by a substantial rise in the number of cases across various regions at municipal, state, and national levels; and pandemics, characterized by the rapid spread of disease around the world.

A crucial strategy to minimize virus spread involves monitoring, tracking, and estimating contagion levels, which is extremely useful for decision-making in public health crises. One common approach to predicting the dynamics of an epidemic's spread is mathematical modeling, such as epidemiological models based on differential equations. These approaches aim to model the virus's epidemiological behavior, providing a deeper understanding of how diseases impact humanity. During the COVID-19 pandemic, the importance of mathematical models was recognized for understanding disease transmission, controlling mortality rates, analyzing and projecting scenarios, and predicting the speed and spread of the disease [Schecter et al., 2021].

The SIR (Susceptible-Infected-Recovered) epidemiological model, proposed by Kermack and McKendrick [1927], is a compartmental model that divides the population into distinct groups and tracks transitions between these groups over time. According to Cecconello et al. [2021] and Schecter et al. [2021], the SIR model encompasses the main characteristics of many infectious diseases, including COVID-19. Due to its simplicity, with fewer variables, parameters, and differential equations, it requires less computational effort than other mathematical models, making it a preferred choice for studying disease spread.

For an epidemiological model to be biologically accurate, it is essential to estimate parameters related to the nature and transmission of the infectious agent. Many researchers opt for models with static parameters, meaning the parameters remain unchanged over time. Generally, these models assume that each individual can transmit and recover from the disease at a constant rate [Servidone et al., 2017].

This study aims to adjust the parameters of a SIR model using accumulated COVID-19 case data from Brazil provided by the Ministry of Health. CI methods are employed to find parameter values that minimize discrepancies between the model's predictions and observed data. By comparing model projections with actual data over time, the effectiveness of different approaches is analyzed to determine the best fit for the COVID-19 data in Brazil. Relevant computational methods and related studies are identified through searches in databases such as Web of Science, Scopus, and PubMed.

Section 2 presents the main articles related to this study. Section 3 covers the analysis of the computational experiment results. Section 4 discusses the conclusions derived from the research conducted. Finally, Section 5 contains the acknowledgments.

2. Identification of Computational Methods

A targeted search strategy was implemented using three specific search descriptors to determine the computational methods used to estimate the parameters of the SIR epidemiological model. The first descriptor focused on the model type, incorporating terms such as "SIR" or "model". The second descriptor included the disease name, using terms such as "COVID-19" or "SARS-CoV-2". The third descriptor addressed the need for parameter estimation, using terms like ("estimation" or "estimating") and "parameters". Searches were conducted in the Web of Science, Scopus, and PubMed databases, applying these terms to the titles, abstracts, and keywords of publications. This search yielded 150 articles in Web of Science, 121 in Scopus, and 51 in PubMed.

After using the Bibliometrix package in R to remove duplicates, 157 unique articles published from January 2020 to December 2022 remained. Finally, non-English articles were excluded, resulting in 156 documents.

After reviewing 156 publications, it was found that CI methods, particularly those involving meta-heuristic optimization, were commonly used to estimate the parameters of the SIR model. These approaches are designed to find approximate solutions to complex optimization problems where finding the exact solution is either computationally impractical or requires significant time. Additionally, these methods are valuable for models that involve multiple variables, complex constraints, and high-dimensional search spaces. Figure 1 illustrates the scientific document filtering process.

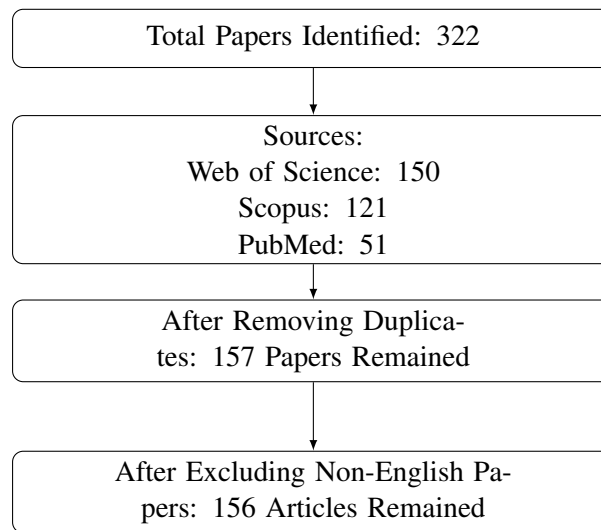


Figure 1: Flowchart of the Scientific Document Filtering Process

Table 1 highlights the CI methods used for parameter estimation, as identified in the key articles that formed the basis for developing the computational methods in this study, particularly the works of Rica and Ruz [2020] and Gopal et al. [2021].

In epidemiological models, meta-heuristic optimization techniques are essential for identifying the most effective sets of parameters. These techniques are crucial for understanding and predicting viral infectious diseases, such as COVID-19, as they play a vital role in fitting the models to observed data and identifying the parameters that best describe the disease's spread within the population.

The study by Rica and Ruz [2020] investigates the SIR epidemiological model using a differential equation to represent the cumulative number of COVID-19 cases in Santiago, Chile. Differential evolution and random search methods were applied to estimate the parameters of this dynamic model based on actual data.

The scientific paper by Singh et al. [2021] introduces an algorithm that integrates differential evolution with the Adams-Bashforth-Moulton numerical method to estimate the parameters of a SIR epidemiological model. Taking into account the effects of the nationwide lockdown, the study aims to predict confirmed COVID-19 cases in India.

Gopal et al. [2021] explored parameter estimation as an optimization challenge, employing the harmony search meta-heuristic method to derive optimal epidemiological parameters. The effectiveness of this approach was validated through ten variants applied to five COVID-19 data sets fitted to the SIR model.

The study by Wong and Juwono [2022] discusses the use of differential evolution and particle swarm optimization methods to fit curves to COVID-19 data from two Malaysian island states, Penang and Labuan. Data were sourced from official government reports and corroborated with daily case reports from October 1, 2020, to November 24, 2020.

Panigrahi et al. [2022] addressed parameter estimation as an optimization problem, using differential evolution to find optimal or near-optimal parameters for the SIR model. Their study assessed the effectiveness of this epidemiological model in the five states in India most affected by COVID-19, specifically those with the highest case numbers.

Table 1: Identified CI Methods in Key Publications

Title	Authors	Methods
Estimating SIR model parameters from data using differential evolution: an application with COVID-19 data	Rica and Ruz [2020]	Differential Evolution and Random Search
Parameter Estimation of Compartmental Epidemiological Model Using Harmony Search Algorithm and Its Variants	Gopal et al. [2021]	Harmony Search
A modified variable-order fractional SIR model to predict the spread of COVID-19 in India	Singh et al. [2021]	Differential Evolution
A Comparative Study on Parameter Estimation of COVID Epidemiological Models Using Differential Evolution Algorithm	Panigrahi et al. [2022]	Differential Evolution
Estimating Effective Reproduction Number for SIR Compartmental Model: A Stochastic Evolutionary Approach	Wong and Juwono [2022]	Differential Evolution and Particle Swarm Optimization

These methods were instrumental in estimating the parameters of the SIR model, significantly enhancing the computational outcomes of the published studies. The articles by Zhang and Geem [2019] and Al-Omoush et al. [2019] provide more information about harmony search. The study by Storn and Price [1997] proposed differential evolution, and the work by Kennedy and Eberhart [1995] identified the potential of particle swarm optimization for solving optimization problems.

3. Results of Computational Experiments

This section presents the computational results related to the evolvement of cumulative COVID-19 cases in Brazil, derived by applying a SIR-type epidemiological model. CI methods were employed to estimate the parameters of this model.

3.1. COVID-19 Data Collection

The data were gathered on daily new and accumulated COVID-19 cases in Brazil from the official Ministry of Health website. This dataset spans three years, covering 1,096 days, from February 26, 2020, to February 25, 2023.

3.2. Configuration of the SIR Model

This study evaluates a variation of the SIR epidemiological model to determine which best captures the accumulation of COVID-19 cases in the region. The SIR model investigated uses a differential equation defined as follows:

$$\begin{aligned} \frac{d}{dt}v(t) &= -\alpha \cdot v(t) + \nu \cdot \xi_0 \cdot \left(1 - \exp\left(\frac{\alpha}{\nu}v(t)\right)\right) + I_0 \\ v(t_0) &= 0. \end{aligned} \quad (1)$$

In this model, $v(t)$ represents the cumulative number of infected cases at time t , while $v(t_0)$ represents the total cases at the initial time t_0 . The parameters include α , the recovery rate, and β , the transmission rate, which together compute the parameter $\nu = \frac{\alpha}{\beta}$. Additionally, t_0 is the starting time, I_0 is the initial number of infected cases, and ξ_0 denotes the initial reproduction number.

3.3. Parameter Estimation

Estimating the parameters of the SIR model involves solving an optimization problem where the decision variables are denoted as $x = \{x_1, \dots, x_D\}$, with D representing the total number of parameters. The lower and upper bounds for these decision variables are defined according to the allowed ranges for each parameter. Table 2 shows the minimum and maximum values for each model parameter. It is important to note that these limits for the model (1) align with the values found in the study by Rica and Ruz [2020].

Table 2: Lower and Upper Bounds of Parameters

Parameter	Range
α	[0.001, 0.1]
ν	[1,000, 30,000]
t_0	[1, 50]
I_0	[0, 30]
ξ_0	[1, 7]

A standard seed was used for generating pseudorandom numbers in the methods: Global Random Search (GRS), Local Random Search (LRS), Harmony Search (HS), Differential Evolution (DE), and Particle Swarm Optimization (PSO). These methods enable the estimation of parameters for the model (1). This process is conducted through an objective function that aims to minimize the mean squared error between v_n and $v(t_n)$, expressed as:

$$\min f(v) = \sum_{n=0}^{N-1} \left(\frac{v_n}{v(t_n)} - 1 \right)^2. \quad (2)$$

In this function, N represents the total number of days in the dataset, $v(t)$ is the numerical solution of model (1) at time t using the estimated parameters, $v(t_n)$ is the calculated solution at time t_n , and v_n is the accumulated number of infected cases in the Ministry of Health's database on day t_n . Notably, the numerical solution for the model (1) was computed using a fourth-order predictor-corrector method, where the predictor is the fourth-order Adams-Bashforth method, and the corrector is the fourth-order Adams-Moulton method. This numerical method can be explored in more detail in the book by Burden et al. [2015].

Table 3 shows the initialization of control parameters for each computational method used in this study, where N_p denotes the population size, N_{max} the maximum number of iterations, HMS the harmony memory size, $HMCR$ the harmony memory consideration rate, PAR the pitch adjustment rate, BW the bandwidth, F the mutation factor for each generation, $U[0.5, 1.0]$ the continuous uniform distribution in the interval $[0.5, 1.0]$, C_r the crossover constant for each generation, w the inertia weight, c_1 the cognitive coefficient of velocity, and c_2 the social coefficient of velocity.

Table 3: Initialization of Control Parameters for Each Computational Method

Method	N_p	N_{max}	HMS	$HMCR$	PAR	BW	F	C_r	w	c_1	c_2
GRS	15	1,000	-	-	-	-	-	-	-	-	-
LRS	15	1,000	-	-	-	-	-	-	-	-	-
HS	-	1,000	15	0.95	0.3	0.01	-	-	-	-	-
DE	15	1,000	-	-	-	-	$U[0.5, 1.0]$	0.7	-	-	-
PSO	15	1,000	-	-	-	-	-	-	0.5	2	2

3.4. Outcomes of Numerical Simulations for COVID-19 in Brazil

Computational experiments were conducted to evaluate the effectiveness of CI methods in estimating the parameters of the SIR epidemiological model. These techniques were implemented using Python 3.11.3 within Visual Studio Code on a laptop with an AMD Ryzen 5 3500U processor, a Radeon Vega Mobile Gfx graphics card running at 2.10 GHz, and 8 GB of RAM. For the model (1), 2,450 simulations were performed using five CI methods. These methods were applied across different segmentation scenarios to analyze the model's behavior over various time intervals.

For the entire analysis period, each method generated 10 simulations, resulting in 50 simulations. For the annual segmentation, each method produced 30 simulations (10 per year over 3 years), totaling 150 simulations. In the semiannual segmentation, each method ran 60 simulations (10 per semester over 6 semesters), resulting in 300 simulations. For the four-monthly segmentation, each method generated 90 simulations (10 per period over 9 four-month periods), adding up to 450 simulations. In the quarterly segmentation, each method produced 120 simulations (10 per quarter over 12 quarters), totaling 600 simulations. Finally, for the bimonthly segmentation, each method conducted 180 simulations (10 per bimonthly period over 18 periods), amounting to 900 simulations. The Mean Squared Error (MSE) was used to evaluate the effectiveness of each method, identifying the best and worst simulations for each segmentation scenario.

Table 4: Execution Time Summary for Model (1) Across Different Segmentation Scenarios

Method	Entire Period	Annual	Semiannual	Four-Monthly	Quarterly	Bimonthly
GRS	7.3	7.7	8.0	8.3	8.6	10.0
LRS	10.4	8.0	10.3	9.3	9.7	10.7
HS	8.6	8.0	8.4	9.0	9.4	10.5
DE	20.6	23.4	28.3	33.8	36.1	46.3
PSO	7.1	8.2	8.8	9.7	10.1	10.6

Table 4 shows the execution times (in seconds) for model (1) using CI methods across various segmentation scenarios. GRS, LRS, HS, and PSO methods generally have low and consistent execution times, ranging from around 7.1 to 10.7 seconds across different segmentations. DE method, however, has significantly higher execution times, varying from 20.6 to 46.3 seconds, indicating a higher computational demand compared to the other methods. This highlights that while DE is more time-intensive, the other methods offer faster processing across all segmentation scenarios.

Two criteria were used to select the time partitions. The first criterion was based on the results from model (1) for the entire period, which showed a tendency to plateau at a specific value. This tendency was similar across all simulations, resembling a sigmoid pattern, as shown in Figure

2. This plateau significantly hampered the fit of the actual data, necessitating a subdivision of the overall period. The second criterion referred to the complexity of modeling COVID-19 using the SIR model, primarily due to the seasonality of the disease. Given that the onset of the pandemic in Brazil was in February, the decision was made to divide the entire period into bimonthly intervals. Therefore, including annual, semiannual, four-monthly, and quarterly scenarios in the study was crucial for a more detailed analysis of the computational experiments.

This subsection presents graphs and tables to identify the most suitable time segmentation for the entire analysis period and which method best fits the data. Figure 2 illustrates the best-case scenario, characterized by the lowest Mean Squared Error (MSE) achieved by each method over different time intervals for the model (1).

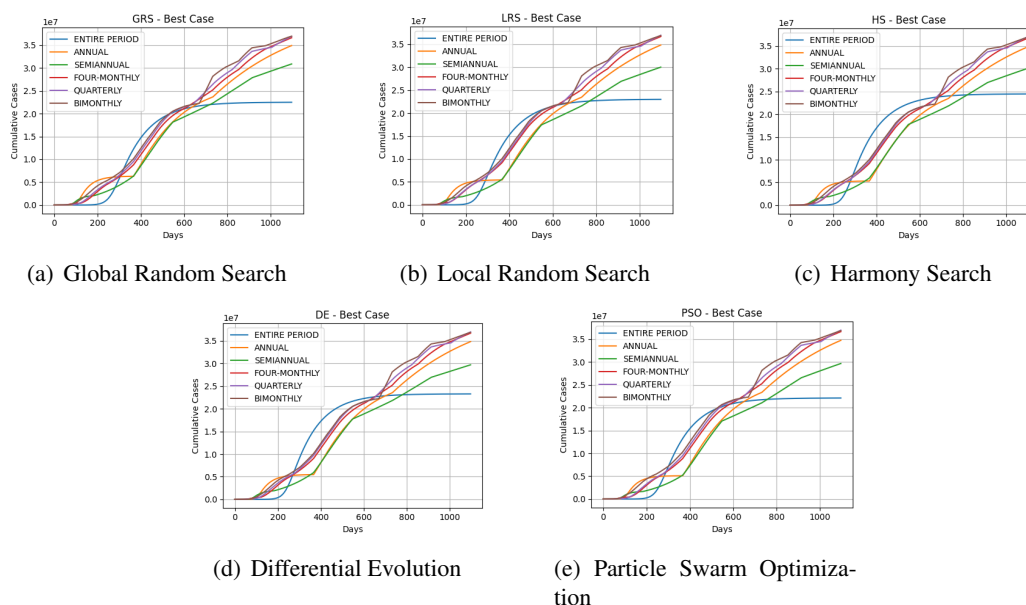


Figure 2: Best-Case Scenario Graphs for Model (1)

Based on Figure 2, it is not possible to determine which method and partition produce the best fit to the actual curve. Therefore, it was necessary to examine the MSE for each technique combined with the different time partitions. Table 5 shows that the bimonthly segmentation is the most efficient across all methods in the best-case scenario for model (1).

Table 5: Mean Squared Error in the Best-Case for Model (1)

Method	Entire Period	Annual	Semiannual	Four-Monthly	Quarterly	Bimonthly
GRS	0.266516968	0.099197410	0.106588216	0.061389543	0.033579746	0.013208146
LRS	0.254293529	0.100694161	0.121117906	0.041579122	0.031307928	0.013505218
HS	0.246877475	0.101774926	0.110728886	0.050377160	0.039986807	0.009581439
DE	0.248200825	0.098608469	0.116613295	0.046453870	0.036186256	0.010557005
PSO	0.258023967	0.102859850	0.138276290	0.045943377	0.045943377	0.019457151

This conclusion is supported by the fact that the lowest MSE is consistently achieved by this division, as evidenced by the bold column. Additionally, among all the methods analyzed,

harmony search, when associated with the bimonthly partition highlighted in blue, is the most effective approach because it presents the lowest MSE. Similar to evaluating the best-case scenario, the analysis now addresses the worst-case scenario, represented by the highest MSE for each method across different time segmentations.

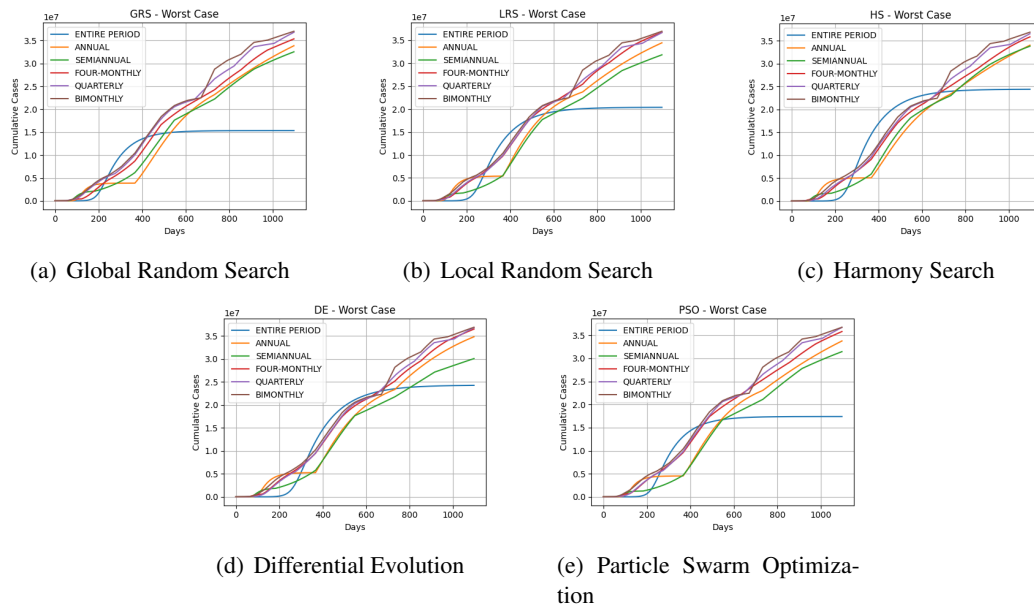


Figure 3: Worst-Case Scenario Graphs for Model (1)

Figure 3 depicts the numerical curves generated by the methods in different time subdivisions, highlighting the worst-case scenario for model (1). Similarly to what was observed in Figure 2, it is impossible to assert which method and time segmentation generates the best numerical curve in Figure 3. Table 6 shows the MSE obtained by each method in the six-time segmentations used in the worst-case scenario for model (1). The goal is to investigate whether the bimonthly partition remains the most effective and if differential evolution is the most suitable CI method for model (1), based on the observations made during the best-case scenario analysis.

Table 6: Mean Squared Error in the Worst-Case for Model (1)

Method	Entire Period	Annual	Semiannual	Four-Monthly	Quarterly	Bimonthly
GRS	0.357857321	0.135736786	0.112087943	0.068585208	0.048006360	0.033437315
LRS	0.276212191	0.101304953	0.124240452	0.050410582	0.046198806	0.050537319
HS	0.246883991	0.112469749	0.114473053	0.057550661	0.044573376	0.013021684
DE	0.265792045	0.101655417	0.121548960	0.053578012	0.050085884	0.012154010
PSO	0.284498709	0.123094430	0.156556657	0.057970471	0.050600285	0.045943377

The bimonthly segmentation demonstrated superior performance across all methods except for LRS. The quarterly segmentation proved more appropriate for that method, as highlighted in bold. Lastly, as highlighted in blue, the DE method was the most effective when used with the bimonthly division. After analyzing the best-case and worst-case scenarios, it is necessary to consider the average-case scenario, which represents the average results obtained in the simulations for

each method applied to each subdivision of the entire period. Figure 4 depicts graphs that are very similar to those seen in Figures 2 and 3, highlighting the difficulty in determining whether the use of the bimonthly partition is the best for all methods or only for some.

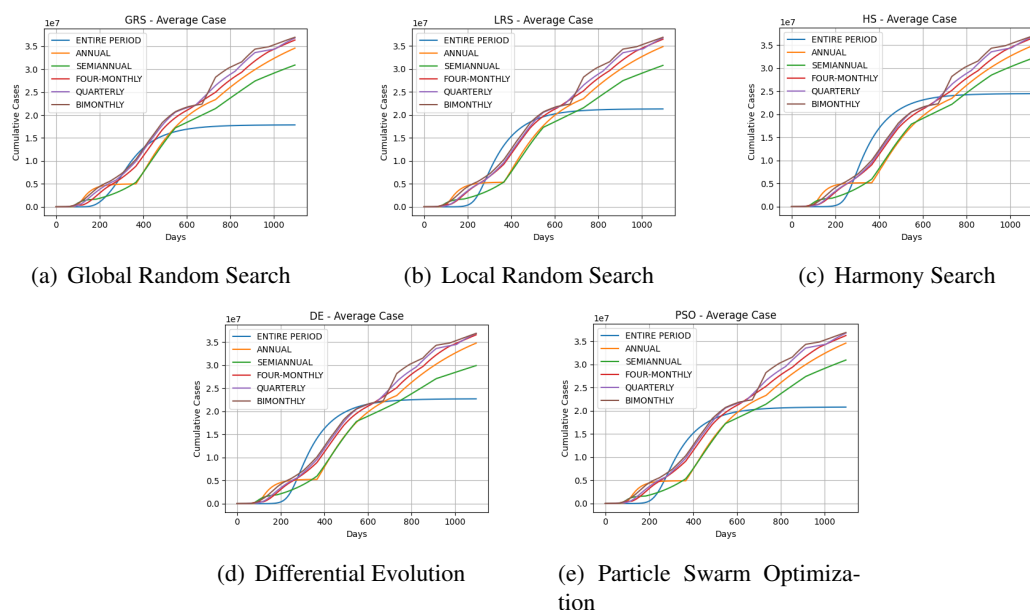


Figure 4: Average-Case Scenario Graphs for Model (1)

Table 7 reveals that the bimonthly partition resulted in the lowest MSE across all methods used for model (1), as shown in bold. Therefore, the best method combined with this segmentation is differential evolution, as highlighted in blue. This analysis yielded the same results as described in the best-case scenario.

Table 7: Mean Squared Error in the Average Case for Model (1)

Method	Entire Period	Annual	Semiannual	Four-Monthly	Quarterly	Bimonthly
GRS	0.271600554	0.102121319	0.131003477	0.057596720	0.029520947	0.021357715
LRS	0.256429543	0.099316575	0.126887076	0.048782070	0.039820244	0.012687509
HS	0.246877459	0.105428082	0.112251581	0.054895710	0.045838569	0.011802087
DE	0.247881592	0.100004170	0.114108464	0.055233506	0.042561238	0.010862145
PSO	0.255683185	0.107011377	0.133507835	0.052090577	0.034411153	0.026730220

Based on the analysis, the bimonthly partition is the best subdivision for the entire period. It stood out in all best-case and average-case scenarios, and four out of five methods in the worst-case scenario. The most effective CI method was differential evolution, which was identified as the best across all methods in the worst-case and average-case scenarios and the second-best in the best-case scenario.

4. Sensitivity Analysis

Sensitivity analysis is crucial for assessing the robustness and reliability of computational models. The objective of sensitivity analysis is to evaluate how different methods for parameter

estimation respond to variations in data segmentation periods. It is vital to assess how well CI methods capture both absolute and relative variability.

This study employs two metrics to evaluate the accuracy and stability of five CI methods. The first metric is the relative error of the standard deviation (SD), which reflects the dispersion of data points around the mean and indicates how well model predictions capture the absolute variability of the observed data. The second metric is the relative error of the coefficient of variation (CV), defined as the ratio of SD to the mean, which provides a normalized measure of variability. Table 8 presents the relative error of SD for model (1) across different segmentation periods.

Table 8: Relative Error of SD in the Average-Case for Model (1)

Method	Entire Period	Annual	Semiannual	Four-Monthly	Quarterly	Bimonthly
GRS	0.783008308	0.092265078	0.177616102	0.034998617	0.021370080	0.007475387
LRS	0.466014464	0.086486369	0.179091524	0.035925030	0.016176742	0.007493491
HS	0.264221155	0.088068956	0.141587038	0.035642723	0.013397720	0.002742238
DE	0.371194194	0.087661213	0.204483843	0.024979787	0.010798746	0.003102515
PSO	0.509568436	0.089803523	0.176126151	0.040002402	0.018614841	0.008301018

From Table 8, HS method achieves the lowest relative error, indicating superior performance in capturing short-term variability. The bimonthly segmentation performs better across methods because it allows the models to lead to lower relative errors and more accurate predictions. Table 9 shows the relative error of CV for the same five methods across different segmentation periods. The CV provides a normalized measure of variability that is especially useful when comparing datasets with different means, as it allows for a more consistent evaluation of relative variability.

Tabela 9: Relative Error of CV in the Average-Case for Model (1)

Method	Entire Period	Annual	Semiannual	Four-Monthly	Quarterly	Bimonthly
GRS	0.136602541	0.060745772	0.085147561	0.037691335	0.008565523	0.005406321
LRS	0.106797432	0.051359483	0.081014789	0.023037866	0.020342909	0.002579141
HS	0.081468191	0.056868539	0.072882903	0.029986128	0.025187235	0.007957902
DE	0.103055183	0.052035488	0.052490313	0.037870019	0.024083331	0.004976270
PSO	0.120593903	0.063882920	0.085605910	0.025171977	0.013584538	0.006130484

From Table 9, in the bimonthly segmentation, all methods achieve their lowest relative errors, suggesting that this segmentation type better captures the relative variability in the data. The LRS method shows the lowest relative error, indicating the most accurate performance for bimonthly partition. The DE method exhibited closely comparable relative errors to HS in SD and to LRS in CV. This highlights that the DE method is a robust choice that performs well across both metrics.

The results from the CV analysis reinforce the earlier findings that finer segmentations, such as bimonthly, provide more precise data points, allowing models to adjust better and capture the variability inherent in short-term periods. This finer granularity leads to lower relative errors across methods, making bimonthly segmentation particularly effective for sensitivity analysis when both absolute and relative variability are important.

5. Conclusions

In this work, the SIR model was chosen as the epidemiological model due to its ability to capture the main characteristics of COVID-19, a complex disease that profoundly impacted glo-

bal society. Although more complex models can provide more detailed representations of disease spread, the SIR model has proven to be a valuable tool for assessing the cumulative number of COVID-19 cases in Brazil. Its simplicity allowed for more accessible implementation and a more straightforward interpretation of the results.

Due to the search conducted across three databases of published works, CI methods such as GRS, LRS, HS, DE, and PSO were implemented to estimate the parameters of the SIR model. The primary objective was to find the parameters that best fit the official data on the accumulation of COVID-19 cases in Brazil, provided by the Ministry of Health, covering the period from February 26, 2020, to February 25, 2023.

After a detailed analysis of the numerical simulations conducted in this study, it became evident that the DE method stood out as the most suitable approach for fitting the SIR model to the observed data. The fourth-order predictor-corrector numerical method was employed to obtain the estimated values of the accumulated COVID-19 cases in the analyzed region. Additionally, dividing the total period into 18 bimesters proved more effective for the fit than the entire period or divisions into three years, six semesters, nine four-month periods, or 12 quarters.

The sensitivity analysis using both the relative error of SD and CV demonstrates that the choice of segmentation significantly impacts the accuracy of CI methods. The bimonthly segmentation consistently shows the lowest errors for both metrics, suggesting it is the most effective period for capturing data variability. Therefore, this study represents a significant step in understanding the dynamics of the COVID-19 pandemic by combining epidemiological models based on differential equations and CI methods.

A promising direction for future work is the implementation of machine learning methods for analyzing COVID-19 data. These techniques can be used to predict the spread of the disease, identify risk factors, and enhance the understanding of transmission patterns. Additionally, comparing the machine learning methods with the approaches used in this study will provide valuable insights into the effectiveness of different computational approaches in modeling and predicting epidemics.

In summary, there is a vast, unexplored territory at the intersection of epidemiology, computational intelligence, and machine learning. Future studies have the potential to investigate these areas and enhance the ability to understand, predict, and respond to viral infectious disease epidemics, providing significant benefits for public health and global security.

6. Acknowledgements

The following Brazilian research agencies have supported this work: Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) — Proc: 140788/2024-7 and Fundação Cearense de Apoio ao Desenvolvimento Científico e Tecnológico (FUNCAP) — Proc: BMD-0008-00739.01.03/24.

References

- Al-Omoush, A. A., Alsewari, A. A., Alamri, H. S., and Zamli, K. Z. (2019). Comprehensive Review of the Development of the Harmony Search Algorithm and its Applications. *IEEE Access*, 7: 14233–14245.
- Burden, R. L., Faires, J. D., and Burden, A. M. (2015). *Numerical analysis*. Cengage Learning, Boston.
- Cecconello, M. S., Mizukoshi, M. T., and Lodwick, W. A. (2021). Interval nonlinear initial-valued problem using constraint intervals: Theory and an application to the Sars-Cov-2 outbreak. *Information Sciences*, 577:871–882. ISSN 0020-0255. URL <https://www.sciencedirect.com/science/article/pii/S0020025521008513>.

- Gopal, K., Lee, L. S., and Seow, H.-V. (2021). Parameter Estimation of Compartmental Epidemiological Model Using Harmony Search Algorithm and Its Variants. *Applied Sciences*, 11(3). ISSN 2076-3417. URL <https://www.mdpi.com/2076-3417/11/3/1138>.
- Kennedy, J. and Eberhart, R. (1995). Particle swarm optimization. In *Proceedings of ICNN'95-international conference on neural networks*, volume 4, p. 1942–1948, Perth. IEEE.
- Kermack, W. O. and McKendrick, A. G. (1927). A contribution to the mathematical theory of epidemics. *Proceedings of the royal society of london. Series A, Containing papers of a mathematical and physical character*, 115(772):700–721.
- Panigrahi, S. S., Muthukumar, A. J., Thangavelu, S., Jeyakumar, G., and Shunmuga Velayutham, C. (2022). *A Comparative Study on Parameter Estimation of COVID Epidemiological Models Using Differential Evolution Algorithm*, p. 241–263. Springer Nature Singapore, Singapore. ISBN 978-981-16-8082-3. URL https://doi.org/10.1007/978-981-16-8082-3_9.
- Rica, S. and Ruz, G. A. (2020). Estimating SIR model parameters from data using differential evolution: an application with COVID-19 data. In *2020 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*, p. 1–6, Chile. Institute of Electrical and Electronics Engineers Inc.
- Schecter, S., Matos, V., Marchesin, D., and López-Flores, M. M. (2021). *Equações diferenciais e modelos epidemiológicos*. Instituto de Matemática Pura e Aplicada, Brasil. ISBN 978-65-89124-42-9.
- Servidone, G., Ternes, S., Rossi, M. M., and Nogueira, M. F. (2017). Fuzzy theory use in deterministic epidemiological models. In *Proceeding Series of the Brazilian Society of Computational and Applied Mathematics*, Brazil. Brazilian Society of Computational and Applied Mathematics. URL <https://proceedings.sbmec.org.br/sbmec/article/viewFile/1214/1226>.
- Singh, A., Mehra, M., and Gulyani, S. (2021). A modified variable-order fractional SIR model to predict the spread of COVID-19 in India. *Mathematical Methods in the Applied Sciences*, 46.
- Storn, R. and Price, K. (1997). Differential evolution-a simple and efficient heuristic for global optimization over continuous spaces. *Journal of global optimization*, 11(4):341.
- Wong, W. K. and Juwono, F. H. (2022). Estimating Effective Reproduction Number for SIR Compartmental Model: A Stochastic Evolutionary Approach. *Journal of Social Computing*, 3(2): 182–189.
- Zhang, T. and Geem, Z. W. (2019). Review of harmony search with respect to algorithm structure. *Swarm and Evolutionary Computation*, 48:31–43. ISSN 2210-6502. URL <https://www.sciencedirect.com/science/article/pii/S2210650218303791>.