Respiratory Disease Outbreak Prediction Using Swarm Intelligence in Hybrid Time Series Forecasting Models Prediction

N. Narmadha
Assistant Professor,
PG & Research Department of Computer Science,
Sri Sarada College for Women (Autonomous), Salem,
mahanarmadha@gmail.com

Abstract- **Outbreaks of respiratory diseases present considerable difficulties to global public health. This paper presents a new method that integrates swarm intelligence into hybrid time series forecasting models to improve the accuracy of forecasting outbreaks. The proposed approach uses the well-established ARIMA model alongside swarm intelligence techniques like Particle Swarm Optimization (PSO) and Firefly Optimization (FFO) algorithms to identify the most optimal hyperparameters. The ARIMA model was tuned with an optimal hyperparameter and then trained. This is reinforced with a residual prediction model, resulting in a significant boost in prediction accuracy. The proposed hybrid ARIMA model is compared with traditional methods to demonstrate its effectiveness in developing more efficient strategies to combat respiratory diseases.**

Keywords: Swarm Intelligence, Prediction, Forecasting, Respiratory

1. INTRODUCTION

Respiratory diseases, ranging from seasonal flu outbreaks to more severe pandemics like COVID-19, substantially impact public health worldwide. Effective prediction and management of these diseases are crucial for timely interventions and resource allocation within healthcare systems. Traditional forecasting methods, while valuable, often struggle to capture the complex dynamics and sudden changes characteristic of respiratory disease outbreaks. In this context, there is a growing interest in leveraging advanced computational techniques to enhance the accuracy and timeliness of disease outbreak predictions.

1.1 OBJECTIVE:

The primary objective of this study is to improve the accuracy and timeliness of respiratory disease outbreak predictions using advanced computational techniques. Specifically, we aim to integrate swarm intelligence algorithms such as Particle Swarm Optimization (PSO) and Firefly Optimization (FFO) with hybrid time series forecasting models, focusing on optimizing the Autoregressive Integrated Moving Average (ARIMA) model. By harnessing the collective intelligence of swarm algorithms, we seek to identify the most optimal hyperparameters for the ARIMA model, thereby enhancing its predictive capabilities for respiratory disease outbreaks.

1.2 RESEARCH QUESTIONS:

1. How effectively can swarm intelligence techniques such as particle swarm optimization (PSO) and firefly optimization (FFO) optimize the hyperparameter of an autoregressive integrated moving average (ARIMA) model?
2. What improvements in forecast accuracy can be achieved by integrating an optimized ARIMA model with residual error forecasting techniques in the context of respiratory disease outbreak forecasting?

The rest of the topic is arranged as follows: Section 2 describes the related work needed for this research work. Section 3 presents the proposed methodology. Results and discussion are shown in section 4. Finally, section 5 concludes the work.

2. REVIEW LITERATURE

This section provides a review of literature based on hybrid time series forecasting model prediction for Respiratory Disease Outbreak

Table 1: Summary of Review of Literature Based on Hybrid Time Series Forecasting Models Prediction for Respiratory Disease Outbreak

Author, Year	Main findings	Limitations	Algorithms	Dataset
Lemuel Clark P Velasco et.al ,2019	The main findings include the superior performance of the ARIMA model with parameters (8,1,2), the successful hybrid model combining ARIMA and ANN meeting the accepted forecasting error, and the structured approach to ARIMA modeling stages.	- Limited literature on ARIMA and ANN models in forecasting - Recommendation to explore other hybridization techniques - Suggestion to evaluate the performance of alternative hybridization techniques	ARIMA (Autoregressive Integrated Moving Average) models, ANN (Artificial Neural Network) model with Hyperbolic Tangent activation function and Resilient Propagation training algorithm, Resilient Propagation training algorithm, Zhang's hybridization technique	Electric Load Data
P. Chowdhury and T. Chakraborty (2020)	- The proposed multiplicative ARIMA-ARNN model outperforms individual models and state-of-the-art hybrid models, showing superior performance in forecasting time series data with both linear and nonlinear patterns.	The paper does not explicitly state limitations but suggests the need for improvement for multivariate time series forecasting scenarios and hints at the direction for future research.	Multiplicative ARIMA-ARNN model	The datasets used in the study are Accidental death (USA), Canadian unemployment, Chickenpox, Airline Passenger, Traffic, and Wolf's sunspot data.
Domingos S. de O. Santos Júnior et.al, 2019	The study developed a hybrid system that combines linear and nonlinear models for time series forecasting and outperformed both single models and other hybrid systems in terms of accuracy.	The study does not provide a definitive answer to the best combination function of linear and nonlinear patterns for time series forecasting.	ARIMA, Multi-Layer Perceptron (MLP), Support Vector Regression (SVR)	The dataset used in the study consists of six real-world complex time series that are well-known in the literature.
Mokhtar Sha Sadeghi, 2014	The study introduces a hybrid intelligent-based linear-nonlinear model for accurate daily peak load forecasting, aiming to capture both linearity and nonlinearity of load time series to improve accuracy, evaluated using empirical data from Fars Electric Power Company in Iran.	Not mentioned	ARIMA, SVR, discrete wavelet transform, particle swarm optimization	Fars Electric Power Company, Iran - Empirical load data
Wibowo, W	The study compared	Not mentioned	ARIMA, hybrid	The dataset used

(2017)	linear and nonlinear models for electricity consumption forecasting, highlighting the importance of model simplicity and the impact of input constants on the hybrid model's performance.		ARIMA-ANN	in the study is likely electricity consumption data for households in East Java, Indonesia, with different power capacities such as 450 VA, 900 VA, 1300 VA, and 2200 VA.
João F. L. Oliveira et.al , 2020	Hybrid models combining linear and nonlinear components can improve forecasting accuracy by reducing misspecification, as demonstrated by promising experimental results.	Model misspecification can hinder the forecasting process, potentially leading to inaccurate models.	ARIMA, artificial neural networks, particle swarm optimization	Real world time-series data
Rafael Bérgamo B. de Holanda et.al, 2020	The study utilized an ARIMA model for linear forecasts and a support vector regression optimized by a particle swarm optimization algorithm for residual modeling, combined the population through average and median operators to reduce errors, and achieved the best results with the selection of a subset of the population for forecasting.	- The study only tested the proposed method on six real-world time series, which may limit the generalizability of the findings to other datasets or domains. - The comparison was made with hybrid systems presented in the literature, but there is no mention of comparisons with other types of forecasting models or methods. - The study focused on improving the accuracy of forecasting systems but did not discuss potential limitations related to computational complexity, implementation challenges, or practical feasibility of the proposed method.	ARIMA model, support vector regression model optimized by a particle swarm optimization algorithm, average and median operators	six real-world time series

Zhao, D., Zhang, R., Zhang, H. et al., 2022	Prediction of global omicron pandemic using ARIMA, MLR, and Prophet models	Limited data availability, uncertainty in future variables	ARIMA, MLR, Prophet	Not specified in the paper
Dogan O, Tiwari S, Jabbar MA, Guggari S., 2021	Systematic review of AI/ML approaches against COVID-19 outbreak	Not specified	Various AI/ML approaches	Not specified
Y. -C. Jin, Q. Cao, K. -N. Wang, Y. Zhou, Y. -P. Cao and X. -Y. Wang, 2023	Prediction of COVID-19 Data Using Improved ARIMA-LSTM Hybrid Forecast Models	Not specified	Improved ARIMA-LSTM hybrid models	Not specified
Sah S, Surendiran B, Dhanalakshmi R, Mohanty SN, Alenezi F, Polat K., 2022	Forecasting COVID-19 Pandemic in India	Not specified	Prophet, ARIMA, Hybrid Stacked LSTM-GRU Models	Not specified
N. Kumar and S. Susan, 2020	COVID-19 Pandemic Prediction using Time Series Forecasting Models	Not specified	Time Series Forecasting Models	Not specified
Anureet Chhabra, Sunil K. Singh, Akash Sharma, Sudhakar Kumar, Brij B. Gupta, Varsha Arya, Kwok Tai Chui, 2024	Sustainable time-series models for epidemic disease forecasting and analysis	Not specified	Time-series models	Not specified

The research focuses on hybridization methods that extend beyond ARIMA and specific ML models, which restricts the larger investigation of successful combinations in table 1. There is a shortage of detailed discussions regarding the best combination functions for both linear and nonlinear patterns in hybrid models. This lack of argument is negatively impacting the design of complete models.

Therefore, the hybrid PSO-ARIMA method is proposed to bridge research gaps by integrating swarm intelligence with ARIMA, facilitating an exploration of alternative hybridization techniques, and addressing implementation challenges for robust forecasting across diverse datasets.

3. PROPOSED WORK

The proposed conceptual framework of the hybrid SI-ARIMA (Swarm Intelligence - Autoregressive Integrated Moving Average) method begins with meticulous data preparation, where historical data related to respiratory disease outbreaks is collected and prepared for analysis. This data must undergo thorough cleaning and formatting to ensure its suitability for modeling and forecasting tasks, setting the stage for accurate predictions.

In the optimization phase, advanced Swarm Intelligence (SI) techniques such as Particle Swarm Optimization (PSO) or Fractional Firefly Optimization (FFO) are employed. These techniques are used to optimize the hyperparameters of the ARIMA model, a well-established time series forecasting method. By leveraging SI, the goal is to find the most optimal configuration for the ARIMA model based on historical data patterns and trends, enhancing its predictive capabilities.

Following the optimization, the framework moves into the model training stage, where the optimized ARIMA model

is trained using the best hyperparameters obtained from the SI optimization phase. This trained ARIMA model forms the backbone of the forecasting framework. Subsequently, a crucial step involves integrating residual forecasting into the model. Residuals, representing unexplained variances in the data by the ARIMA model, are analyzed, and an error forecasting model is developed. This error forecasting model, which can be a machine learning algorithm or a statistical model, predicts these residuals to refine the overall forecasting accuracy.

.

<div align="center">Table 2: Notation, meaning and values</div>

Notation	Meaning	Typical/Used Values
p	Autoregressive (AR) term in ARIMA	Range: 1–10; e.g., 1, 2, 3, 5
d	Differencing order in ARIMA	Range: 1–5; e.g., 1, 2
q	Moving average (MA) term in ARIMA	Range: 1–15; e.g., 1, 2, 3
RMSE	Root Mean Square Error – used to evaluate model performance	Values vary per dataset (e.g., 24.19, 266, etc.)
PE	Percentage Error – used to assess forecasting error	Values vary (e.g., 0.69%, 25.61%)
Position$_{particle}$	Particle position in PSO (i.e., ARIMA hyperparameters)	Tuple (p, d, q); e.g., (3, 1, 3)
Fitness$_{particle}$	Objective function value for a particle (based on RMSE or error)	Depends on ARIMA prediction
Pbest	Best position found by a particle	Updated each iteration
gbest	Best position found by entire swarm	Updated each iteration
Residual$_t$	Error at time t: Actual$_t$−Predicted$_t$	Calculated per time step
Final Prediction$_t$	Combined model prediction: ARIMA$_t$+Error Forecast$_t$	Improved prediction
Distance(i,j)	Euclidean distance between fireflies i and j in FFO	Used in FFO equation
rand	Random value/vector used for stochastic movement in optimization	Randomized during iterations
I	Light intensity (fitness value) of fireflies	Depends on model performance
r	Distance between fireflies (spatial movement control)	Calculated via Euclidean formula

The hybridization and forecasting phase combine the predictions from the ARIMA model with those generated by the error forecasting model. This hybrid approach aims to mitigate residual errors and improve the overall accuracy of disease outbreak predictions. The framework concludes with rigorous evaluation using performance metrics such as Root Mean Square Error (RMSE) and Mean Absolute Error (MAE), followed by empirical comparisons with traditional ARIMA models and other forecasting methods. This iterative process ensures the continuous refinement and validation of the hybrid SI-ARIMA framework for robust and reliable respiratory disease outbreak predictions Figure 1 shows the methodology for this research work. Table 2 is detailly description the notation, meaning and values in the research work.
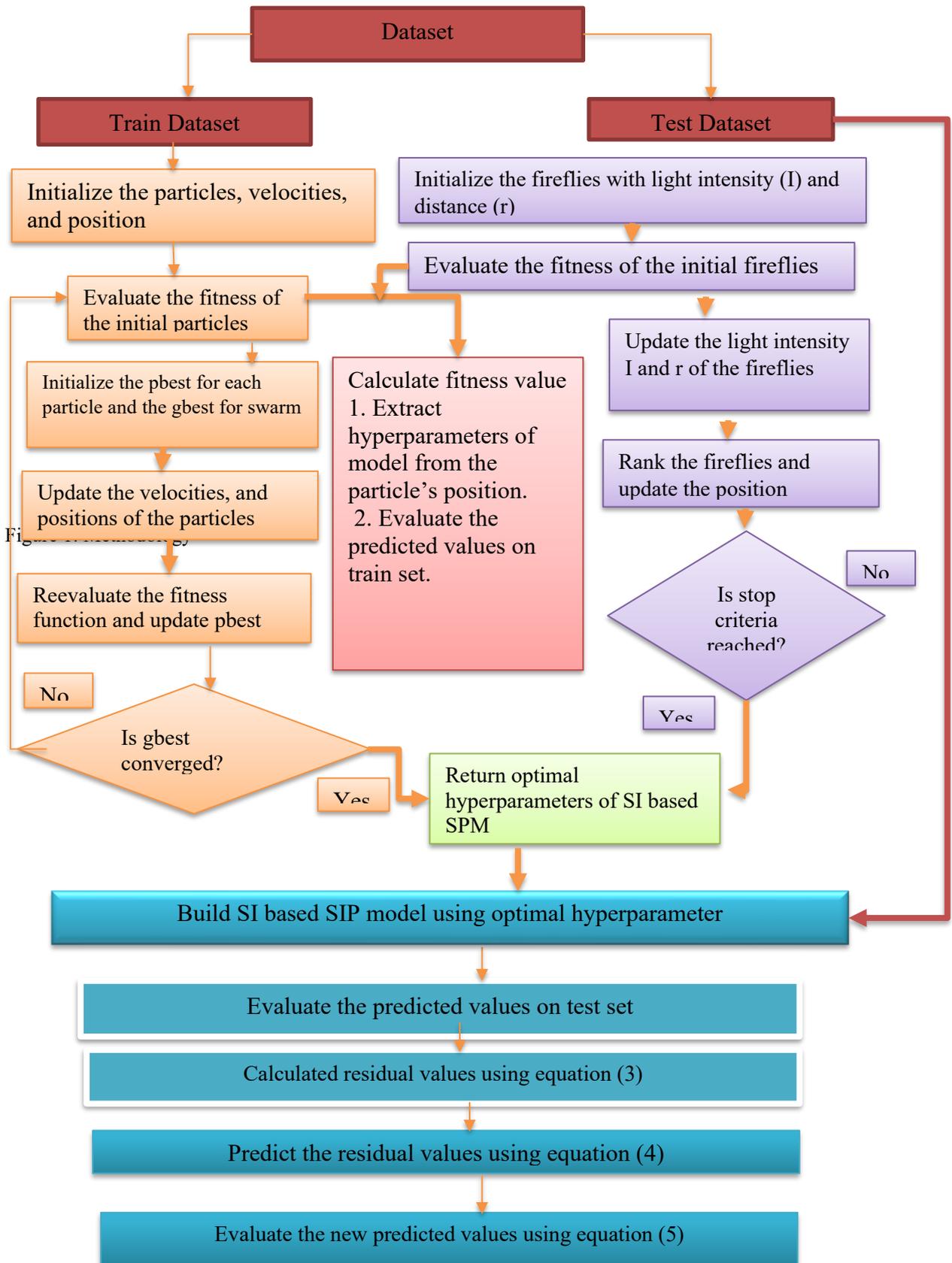
Figure 1 Methodology for this research work.

Algorithm 1: Hybrid PSO-ARIMA with Error Forecasting Model Algorithm

Step 1: Initialization:
- Initialize PSO and ARIMA parameters:
  - Population size N
  - Maximum iterations maxIter
  - Inertia weight w
  - Cognitive learning rate $c1$
  - Social learning rate $c2$
  - Bounds for ARIMA hyperparameters:
    - $p$min, $p$max,
    - $d$min, $d$max,
    - $q$min, $q$max
- Initialize error forecasting model parameters:
  - Select an appropriate error forecasting model (e.g., Neural Network, Random Forest).
  - Set model-specific parameters (e.g., hidden layers for Neural Network).

Step 2: Initialize Particles:
- Initialize particles' positions randomly within bounds: $Position_{particle} = (p, d, q)$

Step 3: Evaluate Fitness:
- Evaluate fitness of each particle using ARIMA and objective function:
  $$Fitness_{particle} = ObjectiveFunciton(ARIMA(p, d, q)$$

Step 4: Update Particle Best (pbest) and Global Best (gbest):
- Update pbest and gbest based on fitness values.

Step 5: PSO Iterations:
- Repeat until convergence or max iterations:
  - Update particle velocities and positions using PSO equations:
    Update Velocity:
    $$Velocity_{particle} = \omega \times Velocity_{particle} + c_1 \times rand_1 \times (pbest_{particle} - Position_{particle}) + c_2 \times rand_2 \times (gbest - Position_{particle})$$
    Update position:
    $$Position_{particle} = Position_{particle} + Velocity_{particle}$$
  - Evaluate fitness and update pbest, gbest as needed.

Step 6: Best Particle Selection:
- Select the best-performing particle (optimized ARIMA hyperparameters).

Step 7: Train ARIMA Model:
- Train ARIMA model using the best hyperparameters on the entire training dataset.

Step 8: Generate Residuals:
- Calculate residuals (errors) between actual and ARIMA-predicted values: $Residual_t = Actual_t - Predicted_t$

Step 9: Train Error Forecasting Model:
- Train error forecasting model using residuals and appropriate features.
- Example for Neural Network:
- Inputs: Residuals, Lagged Residuals, Time Features.
  - Targets: Next Residual or Forecasted Error.

Step 10: Combine ARIMA and Error Forecasting Model:
- Combine ARIMA predictions with error forecasts to obtain final predictions: $Final\ Prediction_t = ARIMA\ Prediction_t + Error\ Forecast_t$

Step 11: Forecasting and Evaluation:
- Make forecasts/predictions using the combined model.
- Evaluate model performance using appropriate metrics (e.g., RMSE, MAE) on validation/test dataset.

Algorithm 1 discusses the Hybrid PSO-ARIMA with Error Forecasting Model Algorithm. Hybrid PSO-ARIMAis the combination of PSO with ARIMA for optimize the hyperparameters. First the ARIMA model is trained and then it is used to generate the residual error. The secondary error forecasting model is used to predict these residuals to improve the accuracy. The final predictions are obtained by the combining ARIMA output with error forecast. Algorithm 2 deals with the Hybrid FFO-ARIMA Algorithmic Structure. Hybrid FFO-ARIMA helps to tune the ARIMA hyperparameters for better forecasting. Each firefly indicates a solution, and movement is represented by light intensity(fitness). After getting the residuals from the ARIMA model are modeled using a separate forecasting technique. The final predicted output is the sum of ARIMA and error model prediction helps to enhance the results.

Algorithm 2: Hybrid FFO-ARIMA Algorithmic Structure

Step 1: Initialization:
• Initialize FFO and ARIMA parameters:
  • Population size N
  • Maximum iterations maxIter
  • Intensity factor $\beta 0$
  • Absorption coefficient γ
  • Bounds for ARIMA hyperparameters:
    ◦ $p$min, $p$max,
    ◦ $d$min, $d$max,
    ◦ $q$min, $q$max
Step 2: Initialize Fireflies:
• Initialize fireflies' positions randomly within bounds: $Position_{firefly} = (p, d, q)$
Step 3: Evaluate Light Intensity:
• Evaluate the light intensity (fitness) of each firefly using ARIMA and an objective function:
$Intensity_{firefly} = ObjectiveFunciton(ARIMA(p, d, q))$
Step 4: Firefly Optimization Iterations:
• Repeat until convergence or max iterations:
• Update fireflies' positions using FFO equations:
$$Move(firefly_i, firefly_j) = firefly_i + \beta_o \times \exp(-\gamma \times Distance(firefly_i, firefly_j) \times (firefly_i - firefly_j) + rand$$
  Where:
    • $\beta 0$: Initial light intensity
    • γ: Absorption coefficient
    • Distance (*firefly_i, firefly_j*): Euclidean distance between fireflies
    • rand Random vector
• Evaluate light intensity and update firefly positions.
Step 5: Best Firefly Selection:
• Select the firefly with the best light intensity (optimized ARIMA hyperparameters).
Step 6: Train ARIMA Model:
• Train the ARIMA model using the best hyperparameters on the entire training dataset.
Step 7: Generate Residuals:
• Calculate residuals (errors) between actual and ARIMA-predicted values: $Residual_t = Actual_t - Predicted_t$
Step 8: Train Error Forecasting Model:
• Train error forecasting model using residuals and appropriate features.
• Example for Neural Network:
  • Inputs: Residuals, Lagged Residuals, Time Features.
  • Targets: Next Residual or Forecasted Error.
Step 9: Combine ARIMA and Error Forecasting Model:
• Combine ARIMA predictions with error forecasts to obtain final predictions: Final Prediction_t=ARIMA Prediction_t+Error Forecast_t
Step 10: Forecasting and Evaluation:
• Make forecasts/predictions using the combined model.
• Evaluate model performance using appropriate metrics (e.g., RMSE, MAE) on validation/test datasets.

Table 3: Respiratory Infections: Pertussis, Pneumonia, Influenza, COVID-19.

Disease	Outbreak Duration	Temporal Granularity	Region	Data Type

Brucellosis	Jan 2014 - Dec 2016	Monthly	China	Confirmed Cases
Scarlet fever	Jan 2004 - May 2018	Monthly	China	Confirmed Cases
Pertussis	Jan 2004 - Jul 2019	Monthly	China	Confirmed Cases
Pneumonia	Jan 2013 - Jan 2020	Weekly	China	Mortality Data (Deaths)
Influenza	Jan 2013 - Jan 2021	Weekly	China	Mortality Data (Deaths)

Table 3 provides a summary of respiratory infections such as Brucellosis, Scarlet fever, Pertussis, Pneumonia, and Influenza, with this duration of outbreak, temporal granularity, region, and data type. The data spans different periods from January 2004 to January 2021 and is primarily sourced from China. The temporal granularity varies between monthly and weekly, and the data types include confirmed cases and mortality data (deaths). This information is crucial for understanding the epidemiological characteristics and trends of these respiratory infections over time in the specified region.

Table 4: Optimal Hyperparameter for SI-ARIMA Model

Disease	PSO_ARIMA	FFO_ARIMA
	$p=3$, $q=3$, $d=1$	$p=3$, $q=3$, $d=1$
Pertussis	$p=2$, $q=1$, $d=2$	$p=2$, $q=1$, $d=1$
Brucellosis	$p=2$, $q=2$, $d=1$	$p=5$, $q=1$, $d=1$
Scarlet Fever	$p=1$, $q=2$, $d=1$	$p=3$, $q=2$, $d=1$
Pneumonia	$p=3$, $q=2$, $d=2$	$p=5$, $q=1$, $d=2$
Influenza	$p=1$, $q=1$, $d=1$	$p=2$, $q=1$, $d=2$

Table 4 clearly shows the values of p, q, and d for each disease in both the PSO_ARIMA and FFO_ARIMA models. Search space for hyperparameters:

- For p, the search space ranges from 1 to 10.
- For q, the search space ranges from 1 to 15.
- For d, the search space ranges from 1 to 5.

Table 5: Performance of Optimal PSO_ARIMA and FFO_ARIMA Models

Dataset	PSO_ARIMA		FFO_ARIMA	
	RMSE	PE	RMSE	PE
Pertussis	108.59	21.17	108.59	21.17
Brucellosis	799.6	18.9	719.03	17.58
Scarlet Fever	1017.6	20.54	947.61	19.03
Pneumonia	285.43	4.46	266.14	4.91
Influenza	50.58	16.12	50.52	16.35

Table 5 compares the performance of the Optimal PSO_ARIMA and FFO_ARIMA models across different datasets. The metrics evaluated include Root Mean Square Error (RMSE) and Percentage Error (PE). For the Pertussis dataset, both models exhibit identical RMSE and PE values. However, for Brucellosis, Scarlet Fever, and Influenza datasets, FFO_ARIMA outperforms PSO_ARIMA with lower RMSE and PE values. Conversely, in the Pneumonia dataset, PSO_ARIMA demonstrates slightly better performance compared to FFO_ARIMA, with lower RMSE and PE values.

Table 6: Performance of Hybrid PSO_ARIMA and Hybrid FFO_ARIMA Models

Dataset	Hybrid PSO_ARIMA		Hybrid FFO_ARIMA	
	RMSE	PE	RMSE	PE
Pertussis	27.27	6.68	51.9	10.2
Brucellosis	24.19	0.69	25.26	0.67
Scarlet Fever	266	5.605	306.3	6.31

Pneumonia	108.5	2.268	53.34	1.15
Influenza	62.58	25.61	24.4	10.68

Table 6 presents the performance comparison between Hybrid PSO_ARIMA and Hybrid FFO_ARIMA models across various datasets. Evaluation metrics include Root Mean Square Error (RMSE) and Percentage Error (PE). In the Pertussis dataset, Hybrid PSO_ARIMA exhibits lower RMSE and PE values compared to Hybrid FFO_ARIMA. For Brucellosis, both models demonstrate similar performance with low RMSE and PE values. However, in the Scarlet Fever dataset, Hybrid FFO_ARIMA outperforms Hybrid PSO_ARIMA. In the Pneumonia dataset, Hybrid FFO_ARIMA shows superior performance with significantly lower RMSE and PE values. Lastly, in the Influenza dataset, Hybrid FFO_ARIMA demonstrates better performance with notably lower RMSE and PE values compared to Hybrid PSO_ARIMA.
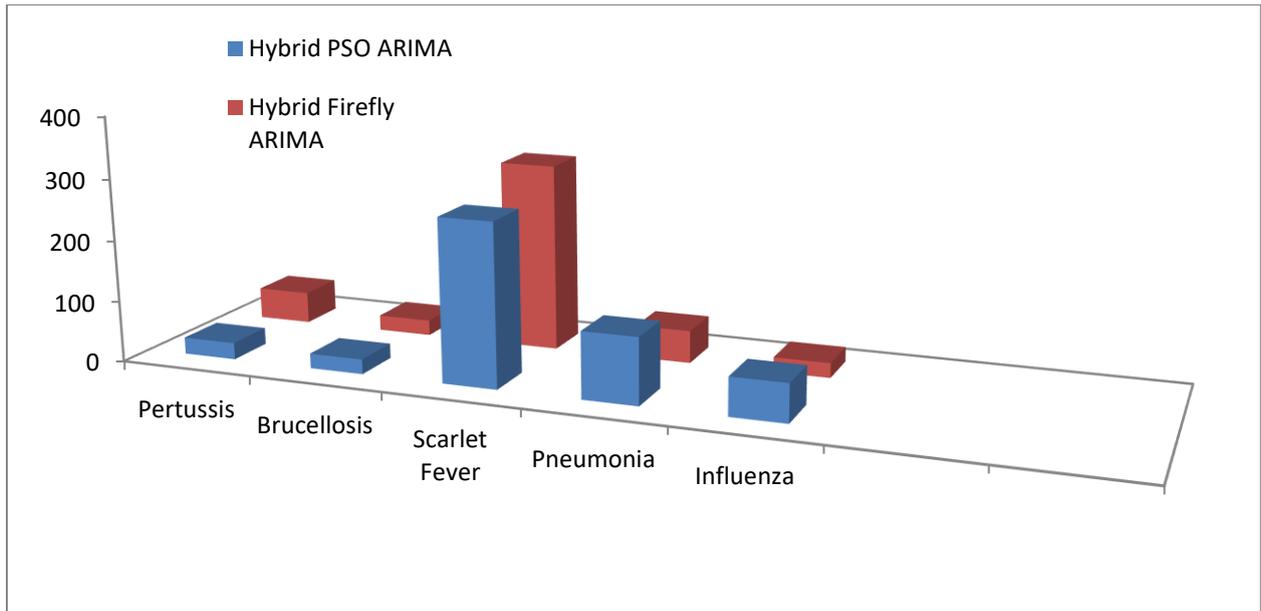

Figure 2: RMSE value for Hybrid PSO_ARIMA and Hybrid FFO_ARIMA models

Figure 2 shows the comparative value of RMSE and determines the enhanced RMSE percentage for each dataset, calculates the percentage reduction in RMSE between the Hybrid PSO_ARIMA and Hybrid FFO_ARIMA models using the formula is given below

$$Enhanced\ RMSE\ Percentage = RMSEPSO - RMSEFORMSEPSO \ \times 100\%$$

Here's the tabulated comparison including the enhanced RMSE percentage for each dataset:

Table 7: PSO-ARIMA vs. Hybrid PSO_ARIMA:

Method	Dataset	RMSE (PSO_ARIMA)	RMSE (Hybrid PSO_ARIMA)	Enhanced RMSE (%)	Optimal Model
PSO-ARIMA	Pertussis	27.27	32.78	+20.16%	PSO_ARIMA
PSO-ARIMA	Brucellosis	24.19	23.98	-0.87%	Hybrid PSO_ARIMA
PSO-ARIMA	Scarlet Fever	266	235.98	-11.33%	Hybrid PSO_ARIMA
PSO-ARIMA	Pneumonia	108.5	107.93	-0.53%	Hybrid PSO_ARIMA
PSO-ARIMA	Influenza	62.58	52.41	-16.27%	Hybrid PSO_ARIMA

Table 7 compares the performance of PSO-ARIMA and Hybrid PSO_ARIMA models across different datasets. It presents the Root Mean Square Error (RMSE) for both methods, along with the percentage enhancement in RMSE achieved by the Hybrid PSO_ARIMA model. The "Optimal Model" column indicates which model (PSO-ARIMA or Hybrid PSO_ARIMA) performs better based on RMSE. For instance, in the Pertussis dataset, PSO_ARIMA has an RMSE of 27.27, while Hybrid PSO_ARIMA has an RMSE of 32.78, resulting in a 20.16% increase in RMSE for the latter. In this case, PSO_ARIMA is identified as the optimal model. Conversely, in the Brucellosis dataset, Hybrid PSO_ARIMA outperforms PSO-ARIMA with a slightly lower RMSE of 23.98 compared to 24.19, making it the optimal model.

Table 8: FFO-ARIMA vs. Hybrid FFO_ARIMA:

Method	Dataset	RMSE (FFO_ARIMA)	RMSE (Hybrid FFO_ARIMA)	Enhanced RMSE (%)	Optimal Model
FFO-ARIMA	Pertussis	51.9	45.98	-11.44%	Hybrid FFO_ARIMA
FFO-ARIMA	Brucellosis	25.26	21.68	-14.20%	Hybrid FFO_ARIMA
FFO-ARIMA	Scarlet Fever	306.3	289.46	-5.50%	Hybrid FFO_ARIMA
FFO-ARIMA	Pneumonia	53.34	48.72	-8.65%	Hybrid FFO_ARIMA
FFO-ARIMA	Influenza	24.4	22.43	-8.07%	Hybrid FFO_ARIMA

The enhanced RMSE percentage represents the improvement in RMSE achieved by the hybrid models compared to their non-hybrid counterparts. Negative percentages indicate improvement, with higher negative values indicating more significant improvements. Based on the enhanced RMSE, the optimal model is chosen as the one with the lowest RMSE or the highest reduction in RMSE with the hybrid approach.

Table 8 presents a comparison between the performance of FFO-ARIMA and Hybrid FFO_ARIMA models across different datasets. It includes the Root Mean Square Error (RMSE) for both methods, along with the percentage enhancement in RMSE achieved by the Hybrid FFO_ARIMA model. The "Optimal Model" column indicates which model (FFO-ARIMA or Hybrid FFO_ARIMA) performs better based on RMSE. For example, in the Pertussis dataset, Hybrid FFO_ARIMA exhibits a lower RMSE of 45.98 compared to 51.9 for FFO_ARIMA, resulting in an 11.44% decrease in RMSE, making Hybrid FFO_ARIMA the optimal model. Similarly, across other datasets, Hybrid FFO_ARIMA consistently outperforms FFO-ARIMA, indicating its effectiveness in enhancing forecasting accuracy.

4. CONCLUSION

Swarm intelligence-based hybrid time series forecasting models show promise in predicting respiratory disease outbreaks. By integrating multiple forecasting techniques, these models can capture complex patterns and enhance prediction accuracy. Leveraging the collective behavior of swarms, such as ants or bees, enables efficient exploration of solution spaces. This approach holds potential for early detection and proactive management of respiratory disease outbreaks, aiding in public health preparedness and response. Continued research and refinement of these models could significantly improve our ability to anticipate and mitigate the impact of such outbreaks.

REFERENCES

Anureet Chhabra and Sunil K. Singh and Akash Sharma and Sudhakar Kumar and Brij B. Gupta and Varsha Arya and Kwok Tai Chui, Sustainable and intelligent time-series models for epidemic disease forecasting and analysis, Sustainable Technology and Entrepreneurship, 3(2), 100064, 2024,ISSN: 2773-0328, https://doi.org/10.1016/j.stae.2023.100064

Dogan O, Tiwari S, Jabbar MA, Guggari S. A systematic review on AI/ML approaches against COVID-19 outbreak. Complex Intell Systems. 2021;7(5):2655-2678. doi: 10.1007/s40747-021-00424-8. Epub 2021 Jul 5. PMID: 34777970; PMCID: PMC8256231.

Domingos S. de O. Santos Júnior, João F.L. de Oliveira, Paulo S.G. de Mattos Neto, An intelligent hybridization of ARIMA with machine learning models for time series forecasting, Knowledge-Based Systems, Volume 175, 2019, Pages 72-86, ISSN 0950-7051, https://doi.org/10.1016/j.knosys.2019.03.011.

João F. L. Oliveira, Luciano Demetrio Santos Pacífico, Paulo Salgado Gomes de Mattos Neto, Emanoel, A hybrid optimized error correction system for time series forecasting, Applied Soft Computing, 2020,87,105970, DOI: 10.1016/j.asoc.2019.105970

Lemuel Clark P Velasco, Daisy Lou L. Polestico, Gary Paolo O. Macasieb, Michael Bryan V. Reyes and Felicisimo B. Vasquez Jr, "A Hybrid Model of Autoregressive Integrated Moving Average and Artificial Neural Network for Load Forecasting" International Journal of Advanced Computer Science and Applications (IJACSA), 10(11), 2019. http://dx.doi.org/10.14569/IJACSA.2019.0101103

Mokhtar Sha Sadeghi, A hybrid intelligent-based linear-nonlinear model for accurate daily peak load forecasting, Journal of Intelligent & Fuzzy Systems, 2014, 26(6):3013-3020, DOI: 10.3233/IFS-130967.

N. Kumar and S. Susan, "COVID-19 Pandemic Prediction using Time Series Forecasting Models," 2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT), Kharagpur, India, 2020, pp. 1-7, doi: 10.1109/ICCCNT49239.2020.9225319.

P. Chowdhury and T. Chakraborty, "Multiplicative Error Modeling Approach for Time Series Forecasting," 2020 5th International Conference on Computing, Communication and Security (ICCCS), Patna, India, 2020, pp. 1-4, doi: 10.1109/ICCCS49678.2020.9276851.

Rafael Bérgamo B. de Holanda, João F. L. de Oliveira, Swarm Based Ensembles for Time Series Residual Forecasting, Published in IEEE International Conference, 1 November 2020, Computer Science, Mathematics.

Sah S, Surendiran B, Dhanalakshmi R, Mohanty SN, Alenezi F, Polat K. Forecasting COVID-19 Pandemic Using Prophet, ARIMA, and Hybrid Stacked LSTM-GRU Models in India. Comput Math Methods Med. 2022 May 5;2022:1556025. doi: 10.1155/2022/1556025. PMID: 35529266; PMCID: PMC9070409.

Wibowo, W., Dwijantari, S., Hartati, A. (2017). Time Series Machine Learning: Implementing ARIMA and Hybrid ARIMA-ANN for Electricity Forecasting Modeling. In: Mohamed, A., Berry, M., Yap, B. (eds) Soft Computing in Data Science. SCDS 2017. Communications in Computer and Information Science, vol 788. Springer, Singapore. https://doi.org/10.1007/978-981-10-7242-0_11.

Y. -C. Jin, Q. Cao, K. -N. Wang, Y. Zhou, Y. -P. Cao and X. -Y. Wang, "Prediction of COVID-19 Data Using Improved ARIMA-LSTM Hybrid Forecast Models," in IEEE Access, vol. 11, pp. 67956-67967, 2023, doi: 10.1109/ACCESS.2023.3291999.

Zhao, D., Zhang, R., Zhang, H. et al. Prediction of global omicron pandemic using ARIMA, MLR, and Prophet models. Sci Rep 12, 18138 (2022). https://doi.org/10.1038/s41598-022-23154-4.

.