e-ISSN: 2651-3722

HAZIRAN/JUNE 2025 CİLT/VOLUME 8 SAYI/ISSUE 1

Techno-Science

Scientific Journal Of Burdur Mehmet Akif Ersoy University

Editor

Prof. Dr. İsmail KIRBAŞ

SCIENTIFIC JOURNAL OF

BURDUR MEHMET AKİF ERSOY UNIVERSITY

e-ISSN: 2651-3722

MAKÜ
BURDUR MEHMET AKİF ERSOY ÜNİVERSİTESİ

Scientific Journal of Mehmet Akif Ersoy University
Volume 8, Number 1, 2025
e-ISSN: 2651-3722

Contents

-Research Articles

MAKÜ BBBF
BUCAK BİLGİSAYAR VE BİLİŞİM FAKÜLTESİ

Scientific Journal of Mehmet Akif Ersoy University, 2025, 8(1), Page:1-15

Research Article

Developing Demand Forecasting Models for E-Commerce: Analyzing the Impact of Time Lags on Model Performance

Alim Toprak FIRAT [1], Onur AYGÜN[1], Mustafa GÖĞEBAKAN[1], Ceren ULUS[*2], M. Fatih AKAY[2]

[1] *Department of Software Development, Innovance*

[2] *Department of Computer Engineering, Çukurova University*

ABSTRACT

Time series are an important analytical tool used in many problems today. Particularly favored in regression problems such as demand forecasting, time series enable more accurate modeling of the impact of past data on future values through various lag options. Time lag is a method used in time series analysis or machine learning models to examine the effect of past (lagged) values of a variable on current or future values. Time lag options play a crucial role, particularly in the success of demand forecasts. This study aims to develop demand forecasting models that help e-commerce businesses gain a competitive advantage by accurately predicting demand and comprehensively analyzing the impact of time delay options on forecasting performance. In this context, an interface with hyperparametric flexibility has been developed, and the effects of the lag options "Use Best N," "Use Correlation," "Use All Delays," and "Selected Delay Lag" on forecasting performance have been analyzed using demand forecasting models. Models have been created for two different months and three different products. The performance of the developed models has been evaluated using the Mean Absolute Percentage Error (MAPE) metric. The lowest MAPE value for July has been obtained with the MQRNN model developed using product A, while the lowest MAPE value for August has been obtained with the MLP model developed using product B.

Keywords: Time-lag options, E-commerce, Demand forecasting, Machine learning

1. INTRODUCTION

Time series is a type of data that represents the values of a certain variable measured at regular intervals over time and is widely used in many fields, such as finance, economics, energy, health, and meteorology. This data is typically analyzed to examine the behavior of a particular event over time, identify trends and recurring patterns, and make future predictions [1].

Time series data serve as a fundamental building block for solving many problems today, with demand forecasting being one key application. Demand forecasting is a systematic analysis process that aims to estimate future product or service demand [2]. These predictions are made by considering past data, market dynamics, consumer behavior, and other economic factors. The use of time lags is crucial for enhancing the accuracy of demand forecasts. The time lag option used in time series analysis and machine learning models is an important tool for evaluating the impact of past (lagged) values of a variable on current or future values. By accounting for the role of past data in shaping future outcomes, dynamics such as trends, seasonal patterns, and autocorrelation are captured. This allows for a better understanding of the relationship between a variable and its past values. For instance, to predict the closing prices of a stock, data from previous days are included in the model, while in the energy sector, electricity consumption data from previous weeks are included in the model for demand forecasting. Past data not only improves forecast accuracy but also provides a deeper understanding of the system's dynamic structure. Additionally, incorporating past values into the model enables the management of complexity and enhances the generalization capacity without the need for data transformation.

These options can be applied in various ways depending on the dataset and the analysis objectives. Various lag options exist in time series, with the most commonly used being Fixed Lag, Rolling Lag, and Variable Lag. However, these

***Corresponding Author Email:** f.cerenulus@gmail.com
Submitted: 27.11.2024 **Revision Requested:** 16.12.2024 **Last Revision Received:** 20.12.2024
Accepted: 24.12.2024 **Published Online:** 14.01.2025

Burdur Mehmet Akif Ersoy University Press

Cite this article as: Fırat, A. T., Aygün, O., Göğebakan, M., Ulus, C., Akay, M.F. (2025). Developing Demand Forecasting Models for E-Commerce: Analyzing the Impact of Time Lags on Model Performance. Scientific Journal of Mehmet Akif Ersoy University, 8(1): 1-15.
DOI: https://doi.org/10.70030/sjmakeu.1592024 https://dergipark.org.tr/sjmakeu

methods have certain drawbacks. The fixed lag method can overlook important factors such as seasonal changes and dynamic patterns, making it challenging to apply to data that is sensitive to changing conditions. Since the rolling lag works with continuously updated historical data, it can lead to information loss and result in high computational costs. Moreover, the loss of older data may cause long-term trends to be overlooked. On the other hand, the variable lag method can increase model complexity, complicating the optimization process, and may lead to performance issues in cases of insufficient data, thus increasing the risk of overfitting.

In this context, the Use Best N, Use Correlation, Use All Lags, and Use Selected lag options stand out by offering more flexible and optimized solutions compared to fixed, rolling, and variable lag options. These options enable more dynamic and data-specific results compared to fixed lag and other traditional methods. This enables the model to effectively adapt to dynamic changes in the data and enhances its predictive accuracy.

Accurate demand forecasting is crucial, especially for e-commerce businesses. E-commerce businesses require reliable demand forecasts to effectively manage their inventory, accurately determine the quantity of products to be supplied, and efficiently conduct cost optimizations. In this context, accurate demand forecasting emerges as a critical strategic measure, enabling businesses to utilize resources effectively, enhance customer satisfaction, and gain a competitive advantage.

The aim of this study is to develop demand forecasting models that will help e-commerce businesses effectively meet customer needs by accurately predicting demand, gain a competitive advantage in the market, and perform cost optimizations accurately. For this purpose, an interface with hyperparametric flexibility has been developed, and the effects of the lag options on forecasting performance have been analyzed using demand forecasting models.

This study is organized as follows: Section 2 includes relevant literature. Interface with hyperparameter flexibility is presented in Section 3. Section 4 presents methodology. Dataset is presented in Section 5. Development of forecast models is given in Section 6. Results and discussion are presented in Section 7. Section 8 concludes the paper.

2. LITERATURE REVIEW

Nussipova [3] developed a hybrid approach combining machine learning, representation learning, and deep learning techniques to improve the accuracy of energy demand forecasting in modern electricity transmission systems and smart grids. The study highlights notable accuracy improvements achieved through the extraction of fundamental features in energy demand forecasting. In particular, it was observed that triple losses were effective for large margin sizes and long forecast periods. The study emphasizes the role of energy demand forecasting in areas such as renewable energy integration, system management, and market analysis, and states that these technologies contribute to practical applications. Rasul [4] introduced a general-purpose model called Lag-Llama, considering the development of basic models in time series forecasting. It was noted that Lag-Llama, based on a pure decoder converter architecture, focuses on univariate probabilistic time series forecasting. The model was pre-trained on a large time series dataset collected from various domains and demonstrated superior performance in zero-shot generalization. The study further mentions that with a small fine-tuning on new datasets, Lag-Llama surpassed previous approaches and achieved state-of-the-art performance. Lag-Llama stands out as the best general-purpose model for time series forecasting, thanks to its strong generalization capabilities, and it paves the way for future advancements in this field. Peláez-Rodríguez [5] examined the performance of machine learning and deep learning methods in forecasting bike sharing demand in Madrid and Barcelona and cable car demand in Madrid. The predictor variables were divided into four groups, and 12 different regression techniques were applied. The results demonstrate that both machine learning and deep learning methods achieve high accuracy, and these approaches can be applied to other urban mobility studies. Khatun [6] investigated the effectiveness of Convolutional Neural Network (CNN) – Long Short Term Memory (LSTM) and CNN-Gated Recurrent Unit hybrid models for short- and medium-term flow forecasting in the Mahanadi River basin in India. The hybrid models achieved higher accuracy compared to standalone models, and the impact of time lags on model accuracy was analyzed. The CNN-LSTM model accurately predicted flow trends and high peak flows. Additionally, it was highlighted that upstream discharges played an important role in improving the predictions. The study also emphasized that considering all input variables with a fixed time delay resulted in effective flood prediction, even with limited

resources. Guo [7] proposed a forecasting model based on the Stacking fusion model combined with grid search to help businesses meet deadlines and reduce inventory costs. The Stacking Fusion model was constructed using "merchant - warehouse - commodity" time series data. Grid search and cross-validation techniques were used for parameter optimization, Extreme Gradient Boosting (XGBoost), Light Gradient Boosting Machine, and Categorical Boosting machine learning models were included. The forecasting performance was evaluated using metrics such as Root Mean Square Error, goodness-of-fit coefficient, recall rate, and forecast accuracy. Li [8] proposed a cascaded hybrid neural network model based on multimodal data to improve the accuracy of commodity demand projections on e-commerce platforms. Historical order information and product evaluation data were combined as the dataset. In the model, Bidirectional Gated Recurrent Unit Networks and Bidirectional LSTM Networks have been combined. The results showed that the weekly Mean Absolute Error was 0.1682 for commodity forecasts and 0.8611 for long-term demand. When the results has been evaluated, it has been observed that the model provided high accuracy in commodity demand forecasting. Li [9] presented the Spatial Graph Neural Network (SGNN) method, which increased demand forecasting accuracy by using spatial correlations in online sales data. A geographically aware graph model was created using the dataset obtained from Kaggle, and forecasting accuracy was improved with attention methods. The model outperformed traditional methods. Experimental results verified the effectiveness of the GNN-based strategy and provided valuable insights in overcoming the complexities of the online marketplace. Liu [10] studied the e-commerce agricultural products feature dataset and applied data mining. A single model-based demand forecasting model was built using the e-commerce agricultural products feature dataset. To address the issue of irrational artificial fixed parameters in machine learning models, Auto Regressive Integrated Moving Average (ARIMA), LSTM, and Random Forest (RF) models were trained. Particle Swarm Optimization (PSO) and Bayes algorithms were used to optimize the LSTM and RF models, respectively, and the prediction model was created by selecting the most appropriate parameters. Liu [11] aimed to create the feature dataset for e-commerce agricultural products. To achieve this, preprocessing tasks such as data cleaning and filling, technical tools such as web crawlers have been used. The e-commerce agricultural products feature dataset underwent cluster analysis using the K-Means technique to obtain a multi-dimensional summary of sales factors, which provided data support for the prediction model. Considering the limitations of single prediction models, ARIMA, PSO-LSTM, and other combined models were constructed using arithmetic weighting and integrated learning methods. The weights of the model were determined using various arithmetic techniques to combine the prediction results. The combined model, which calculated the weight composition using the inverse error weighted average method, gave the best result in terms of prediction performance. Liu [12] aimed to forecast demand and optimize inventory for thousands of merchants, products, and supporting warehouses in the e-commerce platform. Firstly, an ARIMA time series model has been developed for the shipment of obsolete products over time, and through continuous iteration, the optimal parameters of the model have been obtained to predict the shipment of obsolete products. The final prediction findings were then classified using K-means clustering. The old products were replaced with new ones, and the final prediction values were obtained by cosine similarity analysis after extracting the feature values of both old and new products. Lv [13] developed a model to analyze inventory management problems in e-commerce platforms. While product demand was predicted with the ARIMA model, new product classification was evaluated with the K-means method, and the effect of promotions on demand was evaluated using the LSTM model. The results revealed that LSTM outperformed ARIMA. Trisolvena [14] proposed the application of time series forecasting algorithms to determine trends in 2024 and forecast product demand on the Temu platform. LSTM, Facebook's Prophet, SARIMA, and ARIMA algorithms were used to analyze daily sales data. The analysis revealed that the Prophet model and the SARIMA algorithm outperformed ARIMA and LSTM in terms of forecast accuracy. It was predicted that if the forecasting models were applied appropriately, Temu would increase operational efficiency, improve strategic decision-making processes, and reduce expenses by optimizing inventory control. Zhang [15] aimed to develop an accurate demand forecasting scheme to help merchants better understand their commodity needs. First, a commodity type analysis was conducted to determine how products were distributed across different industries. Then, focusing on the time series data of the best-selling commodities, the sales volume of each commodity was examined. Based on the study results, the sales of similar products supplied by different merchants in various warehouses were analyzed. The Cityblock Optimized K-Means method was used to group products with similar statistical properties. Finally, ARIMA models were applied to predict sales trends for the created clusters. Zhao [16] focused on e-commerce demand forecasting using the Seasonal Auto Regressive Integrated Moving Average (SARIMA) model and the K-means clustering algorithm. Using 1996 sales data, demand changes were predicted for the next 15 days. Initially, Linear Regression, ARIMA, and SARIMA models were evaluated. Then, SARIMA was selected as the most

effective model. The K-means clustering method is applied to divide products into four distinct groups, and similar time series were integrated with cosine similarity. The findings showed that the SARIMA model accurately captured trends and seasonality, providing a reliable framework for e-commerce demand forecasting. Xu [17] presented a new method for forecasting demand during low-density periods. This method was developed by combining carefully selected proxy data with features obtained from a forecasting model based on Graph Neural Networks (GNNs). In the study, demand forecasting was formulated as a meta-learning problem. Using the relational metadata generated by GNNs and proxy data from off-density periods, the Feature-based First-Order Model-Agnostic Meta-Learning algorithm was constructed to learn feature-specific layer parameters. Theoretical analysis showed that the model accounts for domain similarities through task-specific metadata, reducing excessive risk as the number of training tasks increases. The method was evaluated with industry datasets, and the results showed that it significantly increased demand forecasting accuracy. Wang [18] proposed a time series feature-based forecasting model to forecast future product demand based on historical data. As the first step, data preparation was performed, and four regression models were trained using the first 90% of the data table: XGBoost, RF, Decision Tree (DT), and Multi-Layer Perceptron. With the second 10% of the data table, DT was selected as the best regression model. Finally, the demand for products from each merchant in each warehouse was predicted from 2023-05-16 to 2023-05-30. Using the K-means clustering technique, five classes were determined for classifying the time series generated by commodities, warehouses, and merchants. It was found that time series data belonging to the same class had similar demand characteristics.

3. INTERFACE WITH HYPERPARAMETER FLEXIBILITY

An interface with hyperparameter flexibility has been developed for forecasting models. This interface has been created using the Python programming language. The back-end utilizes TensorFlow, Keras, Pandas, Scikit-learn, Numpy, and Statsmodels libraries, while the front-end employs Tkinter, Pandastable, and Matplotlib libraries. The interface enables users to load datasets, select attribute and target variables, and partition the data for learning and testing stages. Additionally, it supports the optimization of method parameters and time delay values. Product demand forecasts are generated, and MAPE calculations are performed by comparing these forecasts with actual values. Furthermore, the interface allows for the visualization of forecast results by plotting graphs of predicted values against actual values. It has been tested using MLP and MQRNN models, and the most successful forecasting models have been identified through an evaluation of the obtained results. Interface with hyperparameter flexibility is shown in Figure 1 and 2.

The train dataset is loaded from the Get Train Set section, where the relevant variables are selected. Preferences such as how the data will be used, the proportion to be utilized, and whether intervals will be included are determined in the Customize Train Set section. Time delay options are categorized under four headings: Use All Lags includes all time delays within the specified range as independent variables in the model; Use Selected allows the user to add specific time delays to the model; Use Best N selects a certain number of delays with the highest autocorrelation with the target variable; and Use Correlation includes time delays with autocorrelation values above a specified threshold as independent variables in the model. Parameters such as the number of layers and neurons, output and activation functions, epoch, batch size, and learning rate are configured via the interface. The model is created using the Create Model button and tested with the Test Model button. Predicted values are displayed in the Values tab, while the comparison between actual and predicted values can be visualized using the Actual vs Forecasted Graph button.

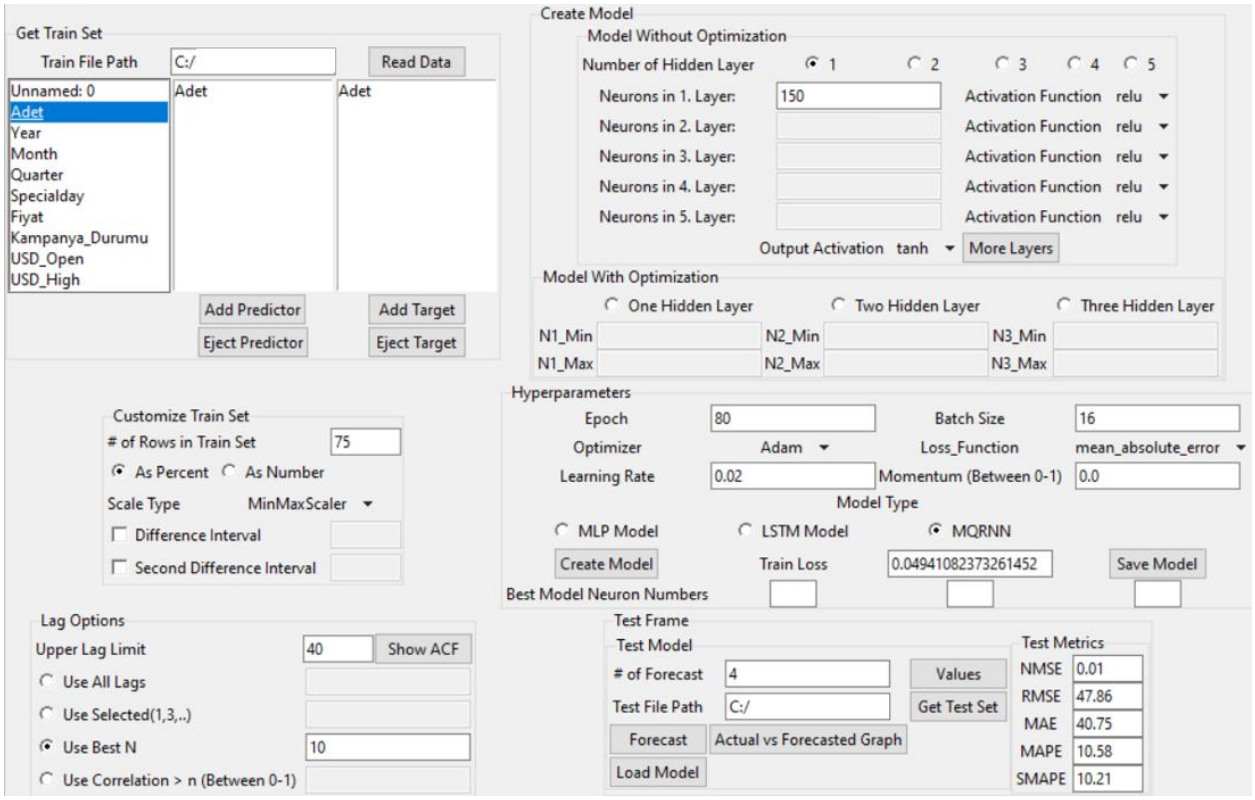

Fig 1. MQRNN screen

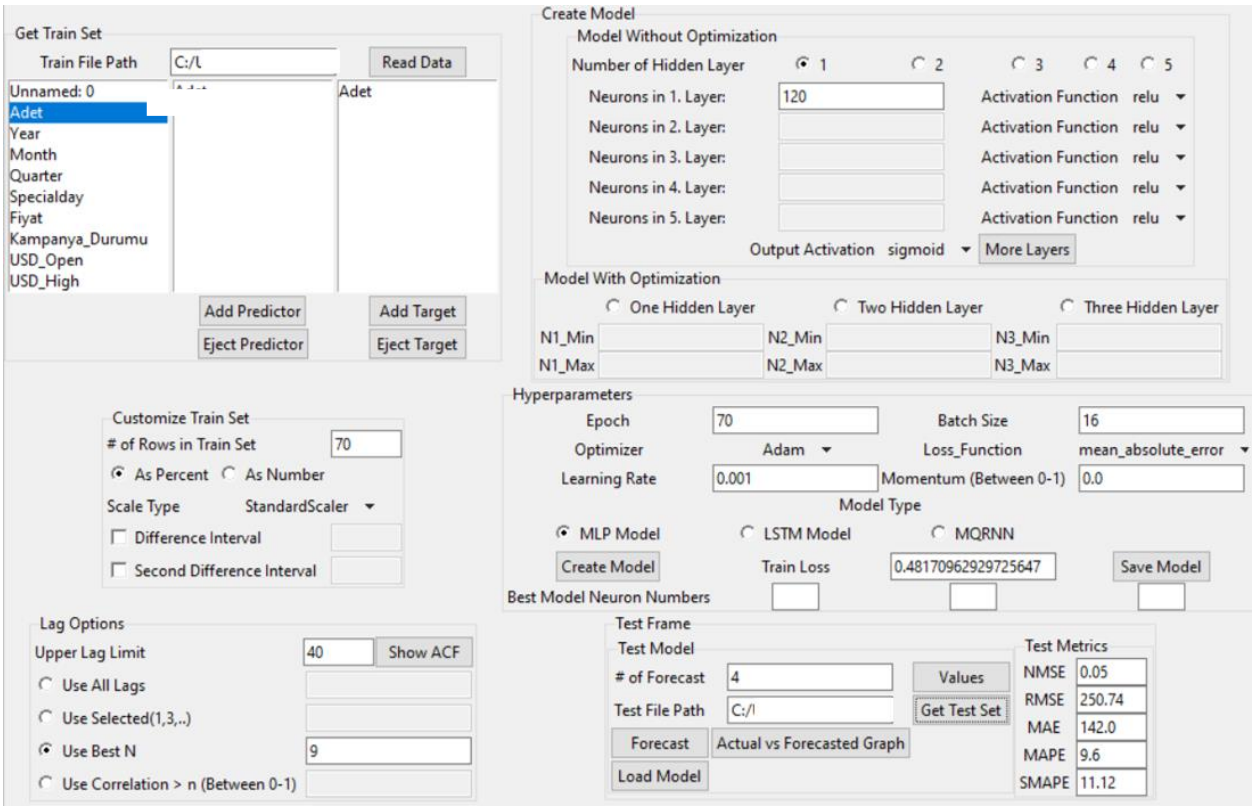

Fig 2. MLP screen

4. METHODOLOGY

4.1. Multi-Layer Perceptron

Artificial Neural Networks (ANNs) are generally divided into two fundamental categories: recurrent (feedback) networks and feedforward networks. Among these, the MLP, which follows a feedforward network architecture, is one of the most widely used types of ANN. MLP operates without making any assumptions about the structure of the output variable, the linearity of the predictor variables or output functions, or the distribution of the data. The MLP architecture comprises three primary components: the input layer (independent variables), the hidden layers (processing units), and the output layer (output variables). This structure is built by connecting multiple parallel layers of nodes with weighted connections. For instance, an MLP model might include three hidden layers containing 64, 32, and 16 neurons, respectively, as well as an output layer and an input layer with five input units [19]. MLP architecture is given in Figure 3.

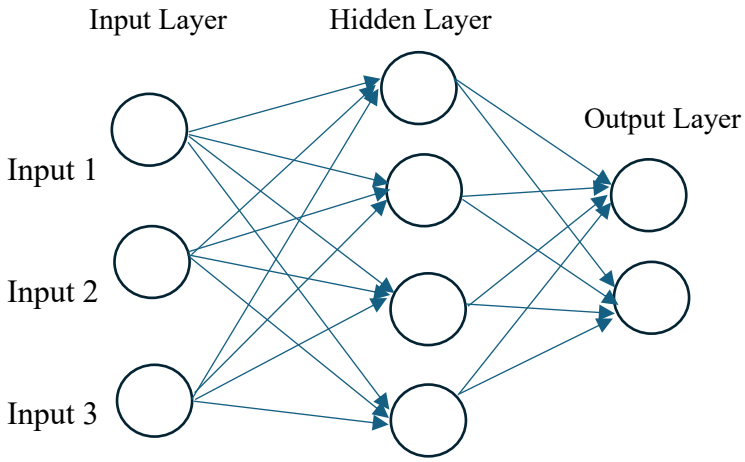

Fig 3. MLP Architecture

4.2. Multi-Horizon Quantile Recurrent Forecaster

LSTM functions as a convolutional encoder for each prediction horizon, generating the context vectors used in MQRNN. This method comprises three main components and is based on the Seq2SeqC architecture:

Encoder: The encoder creates a feature vector that represents the time series input. These feature vectors identify and extract patterns in the time series. The encoder is typically implemented using CNNs or recurrent neural networks (RNNs) with one or more layers.

Context Vector (Intermediate Vector): This vector, referred to as the context or intermediary vector, is generated by the encoder to compile the features of the input time series. The resulting vector is utilized by the decoder to generate predictions representing the time series.

Decoder: The decoder is usually implemented using an RNN or a similar model with one or more layers. It predicts values for the subsequent time steps based on the feature vector generated by the encoder. By employing various training methods, the decoder can be optimized to simulate and predict a specific percentage or distribution of the time series [20].

5. DATASET

The datasets were obtained from an Innovance customer. Daily sales data for three separate products (A, B, and C) were compiled between January 1, 2023, and August 25, 2024. Table 1 presents the characteristics and descriptions of these datasets.

Table 1. The main dataset

Attribute Name	Description
Quantity	Quantity of the products sold
Year	Year
Month	Month
Quarter	Quarter of the year
Specialday	Special days of the year
Price	Product prices during the sales period
Campaign_Status	Campaign status during the sales period
USD_Open	USD opening value
USD_High	Highest value of USD
USD_Low	Lowest value of USD
USD_Close	USD closing value
EURO_Open	Euro opening value
EURO_High	Highest value of Euro
EURO_Low	Lowest value of Euro
EURO_Close	Euro closing value
BIST_Open	BIST opening value
BIST_High	Highest value of BIST
BIST_Low	Lowest value of BIST
BIST_Close	BIST closing value

6. DEVELOPMENT OF FORECAST MODELS

In this study, the demand forecast models have been developed for products A, B and C in the FMCG sector. The models have been developed using MLP and MQRNN for the months of July and August. The effects of the lag option selection and the parameters used on the forecast performance have been also analyzed. The best value of hyperparameters has been found with grid search. The hyperparameter ranges used as a basis for developing prediction models are given in Table 2.

Table 2. Hyperparameter ranges

Method	Hyperparameter Range
MLP	"Percentage_Of_Rows_In_Train_Set": [70 - 100] "Difference_Interval": [7] "Second_Difference_Interval": [28] "Use_All_lags": [5 - 25] "Use_Selected": [1 - 7] "Use_Best_N": [6 - 20] "Use_Correlation": [0.1 – 0.5] "Number_Of_Hidden_Layer": [1] "Neurons_In_1_Layer": [60 - 200] "Epoch": [50 - 128] Batch_Size": [16 - 32] "Learning_Rate": [0.0005 – 0.005]
MQRNN	"Percentage_Of_Rows_In_Train_Set": [70 - 100] "Difference_Interval": [3 - 6] "Second_Difference_Interval": [4] "Use_All_lags": [2 - 15] "Use_Selected": [1 - 31] "Use_Best_N": [2 - 25] "Use_Correlation": [0.1 – 0.3] "Number_Of_Hidden_Layer": [1] "Neurons_In_1_Layer": [50 - 200] "Epoch": [50 - 150] Batch_Size": [8 - 64] "Learning_Rate": [0.001 - 0.02]

7. RESULTS AND DISCUSSION

The MAPE values of the models developed for July have been presented in Table 3, while those for August have been shown in Table 4. The real and forecast values of the most successful results obtained with the developed prediction models developed for July are presented in Figures 4 to 6. The real and forecast values of the most successful results obtained with the developed prediction models developed for August are presented in Figures 7 to 9.

Table 3. MAPE values of the models developed for July

Methods	Products	Lag Options	MAPE (%)
MLP	A	Use All Lags	8.42
		Use Selected	9.39
		Use Best N	9.71
		Use Correlation	9.41
	B	Use All Lags	10.34
		Use Selected	10.96
		Use Best N	9.6
		Use Correlation	10.3
	C	Use All Lags	31.14
		Use Selected	79.28
		Use Best N	41.86
		Use Correlation	44.04
MQRNN	A	Use All Lags	8.79
		Use Selected	9.8
		Use Best N	7.02
		Use Correlation	8.94
	B	Use All Lags	12.77
		Use Selected	12.59
		Use Best N	12.35
		Use Correlation	13.56
	C	Use All Lags	36.04
		Use Selected	92.81
		Use Best N	42.69
		Use Correlation	47.96

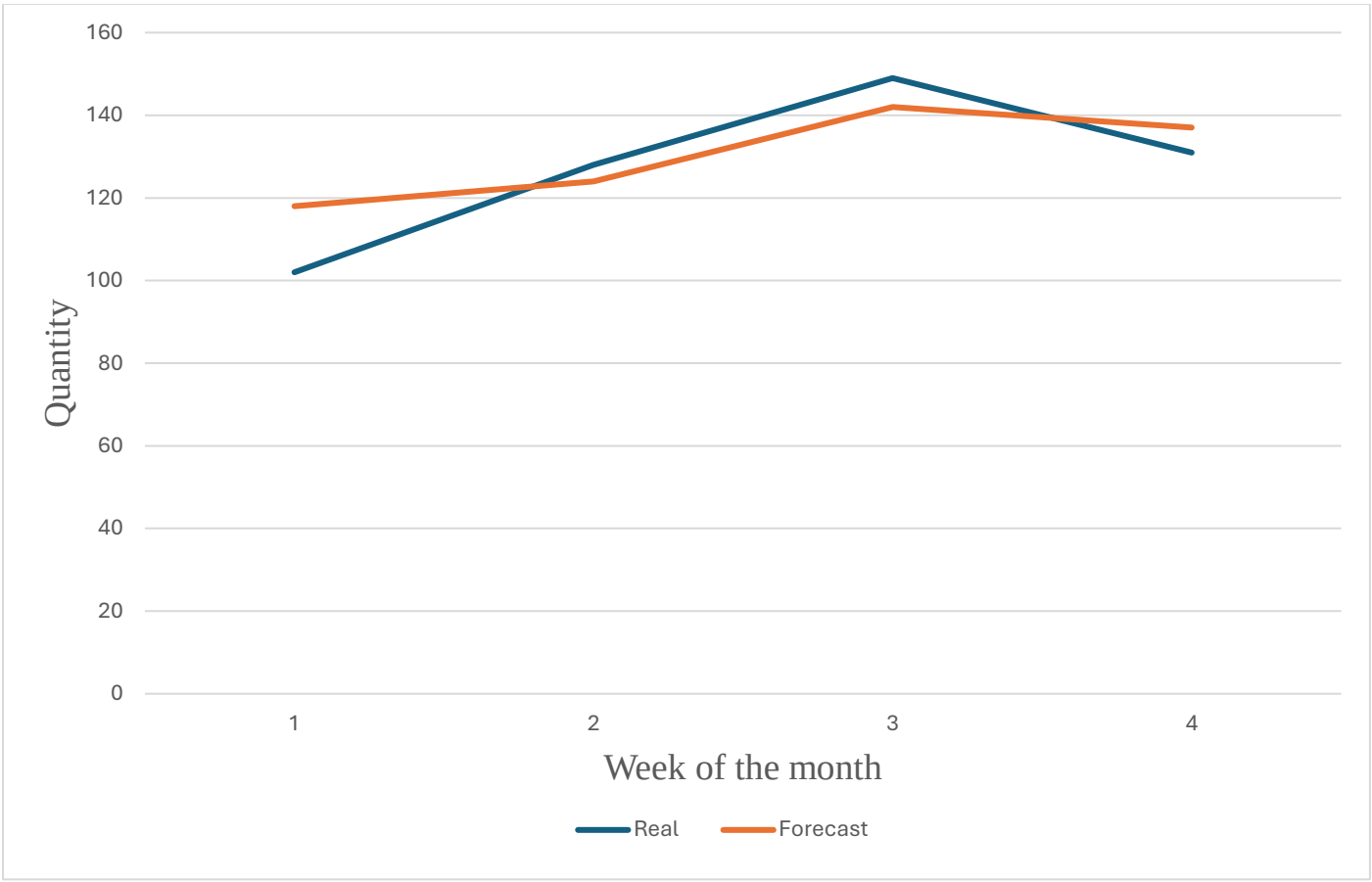

Fig 4. The real and forecast values of the model developed using MQRNN with the Use Best N for product A

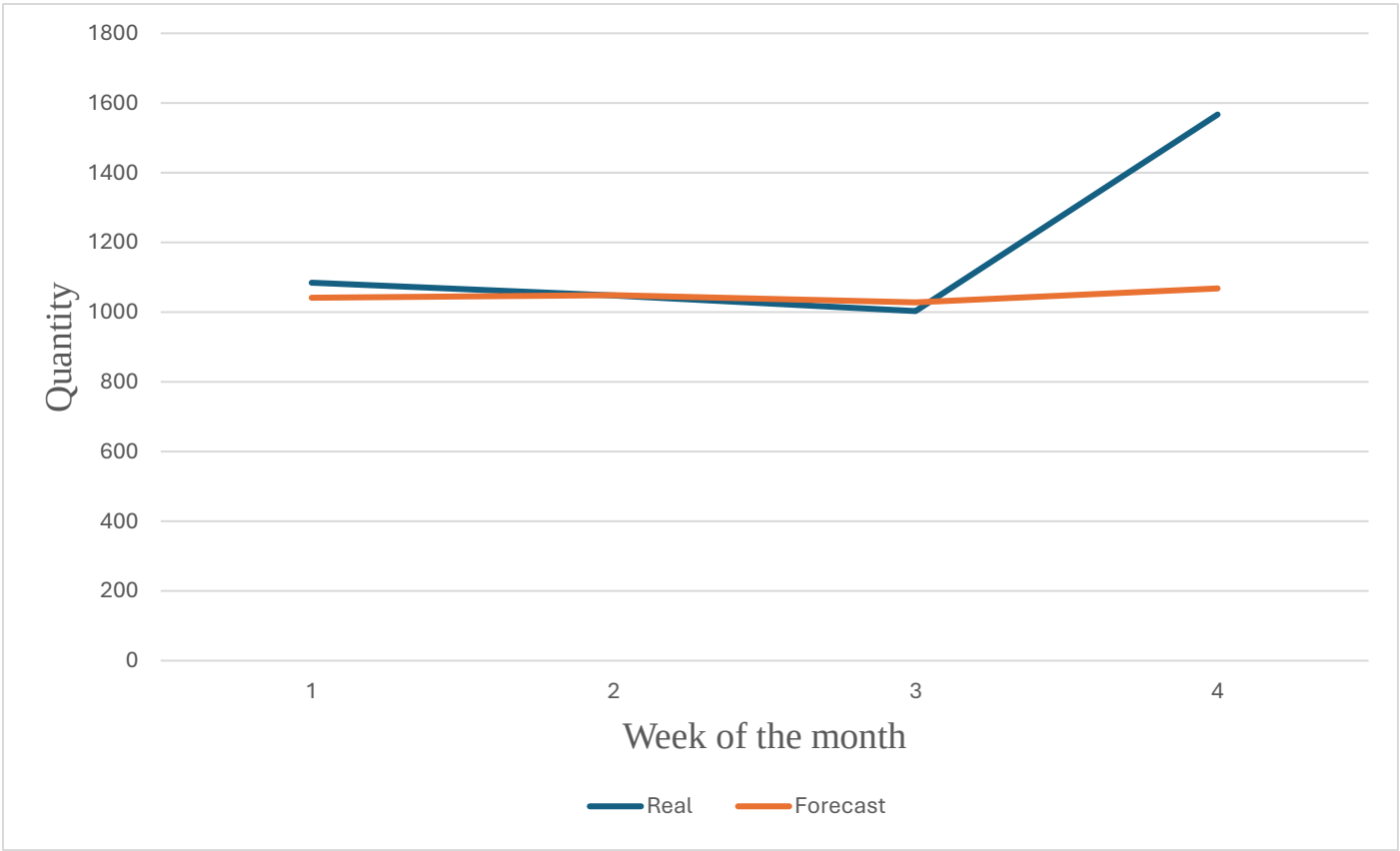

Fig 5. The real and forecast values of the model developed using MLP with the Use Best N for product B

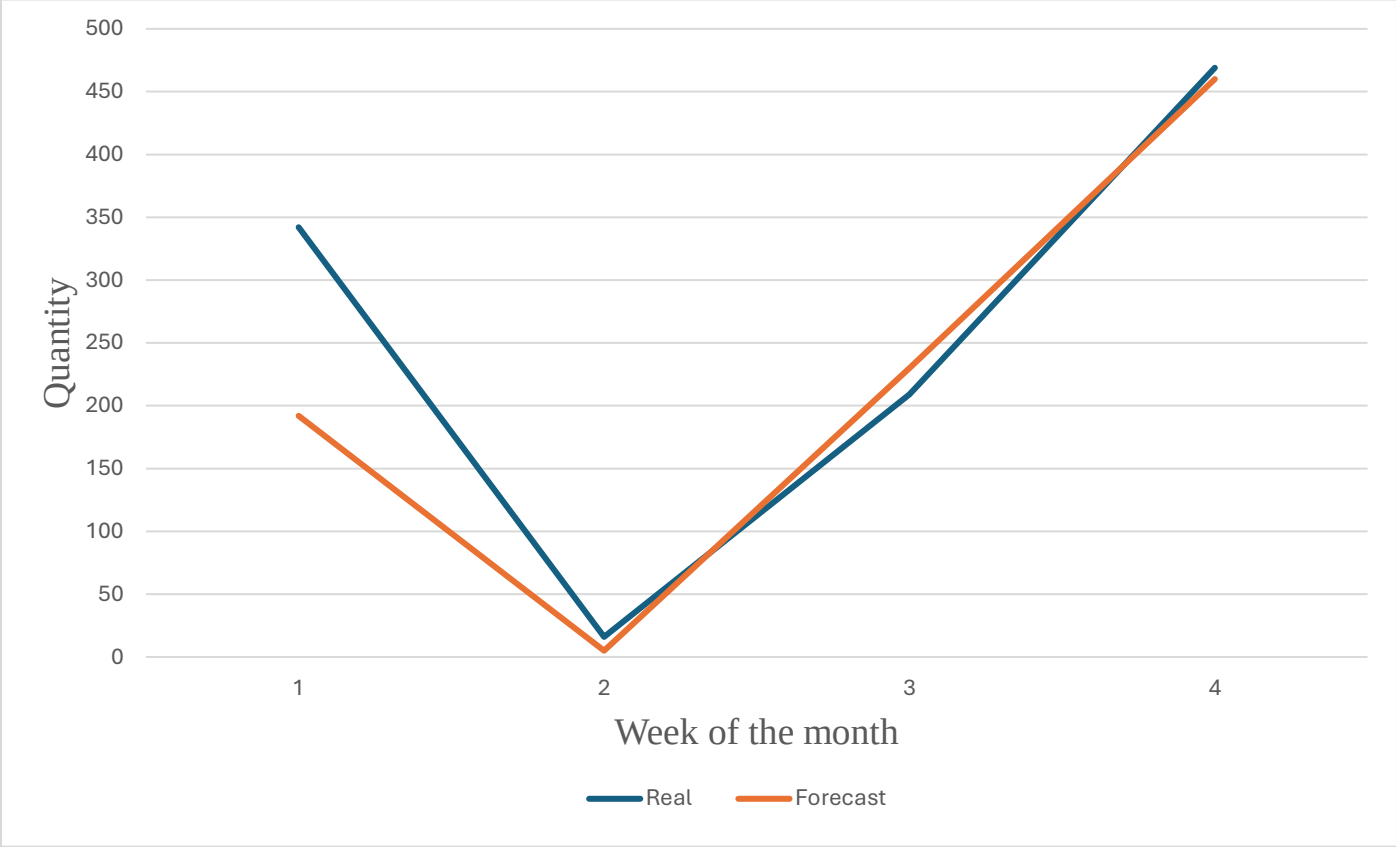

Fig 6. The real and forecast values of the model developed using MLP with the Use All Lags for product C

Upon examining the results of the models developed using MLP:

- The most successful result for Product A was achieved with the Use All Lags, yielding a MAPE of 8.42%. The Use Selected, Use Best N, and Use Correlation demonstrated similar performance.
- For Product B, the best performance has been observed with the Use Best N, achieving an error rate of 9.6%. This suggests that utilizing a specific subset of lags (Best N) improves model accuracy.
- For Product C, the Use All Lags resulted in a MAPE of 31.14%, while the worst performance has been observed with the Use Selected, which had an error rate of 79.28%.

When the results of the models developed using MQRNN have been examined:

- For Product A, the lowest MAPE value, 7.02%, has been achieved with the Use Best N.
- For Product B, the performance difference among the options was minimal. The Use Best N yielded the lowest error rate of 12.35%, while other methods produced similar results, ranging between 12.59% and 13.56%.
- For Product C, MAPE values were notably high, indicating lower model performance compared to other products. Although the best result has been achieved with the Use All Lags, with an error rate of 36.04%, the error rate remains significantly high.

Table 4. MAPE values of the models developed for August

Methods	Products	Lag Options	MAPE (%)
MLP	A	Use All Lags	16.55
		Use Selected	19.17
		Use Best N	13.9
		Use Correlation	18.1
	B	Use All Lags	6.68
		Use Selected	11.55
		Use Best N	9.84
		Use Correlation	10.28
	C	Use All Lags	11.35
		Use Selected	13.03
		Use Best N	13.88
		Use Correlation	26.32
MQRNN	A	Use All Lags	14.15
		Use Selected	18.64
		Use Best N	14.85
		Use Correlation	16.57
	B	Use All Lags	11.39
		Use Selected	16.62
		Use Best N	12.58
		Use Correlation	13.39
	C	Use All Lags	18.74
		Use Selected	22.24
		Use Best N	10.58
		Use Correlation	21.72

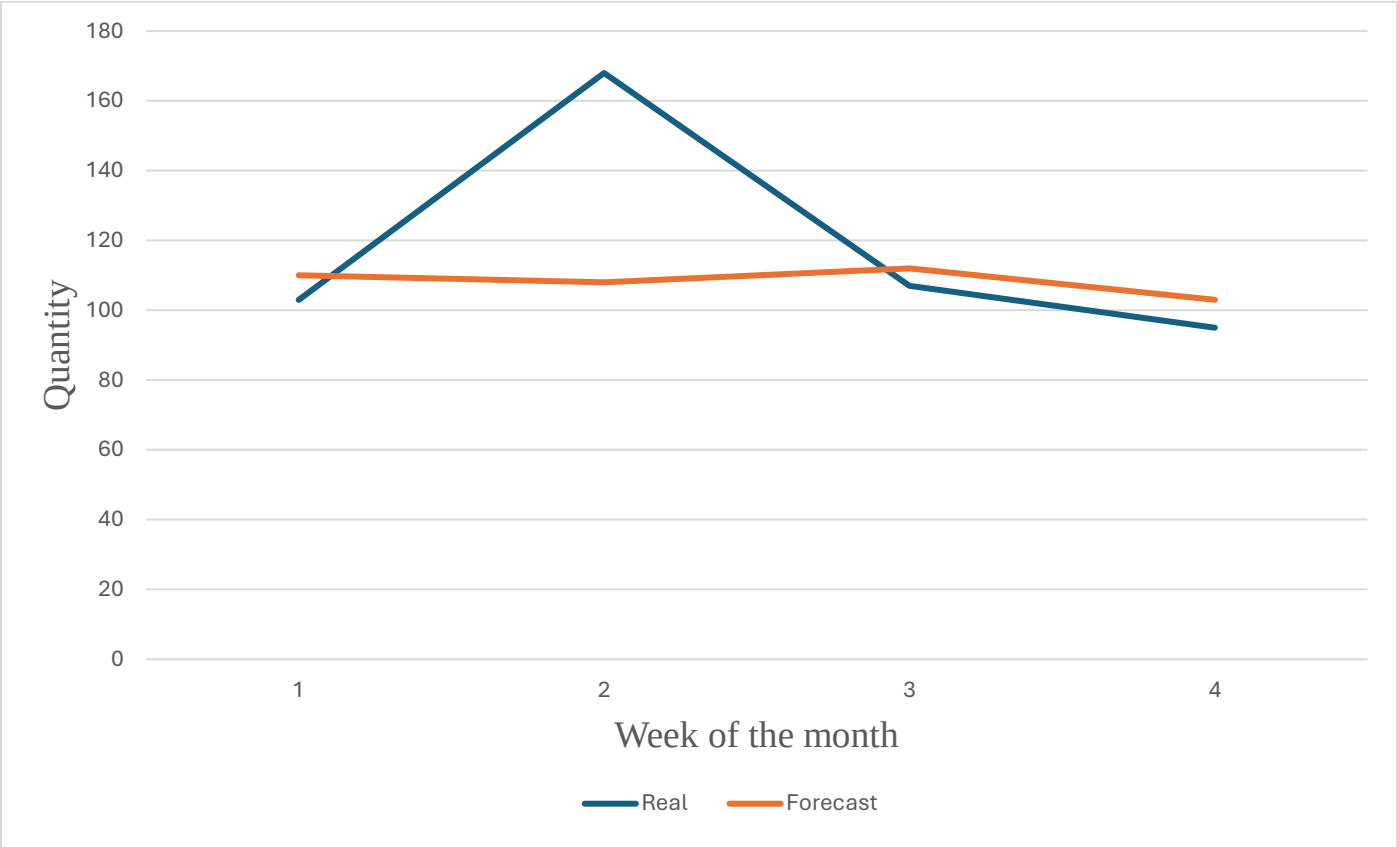

Fig 7. The real and forecast values of the model developed using MLP with the Use Best N for product A

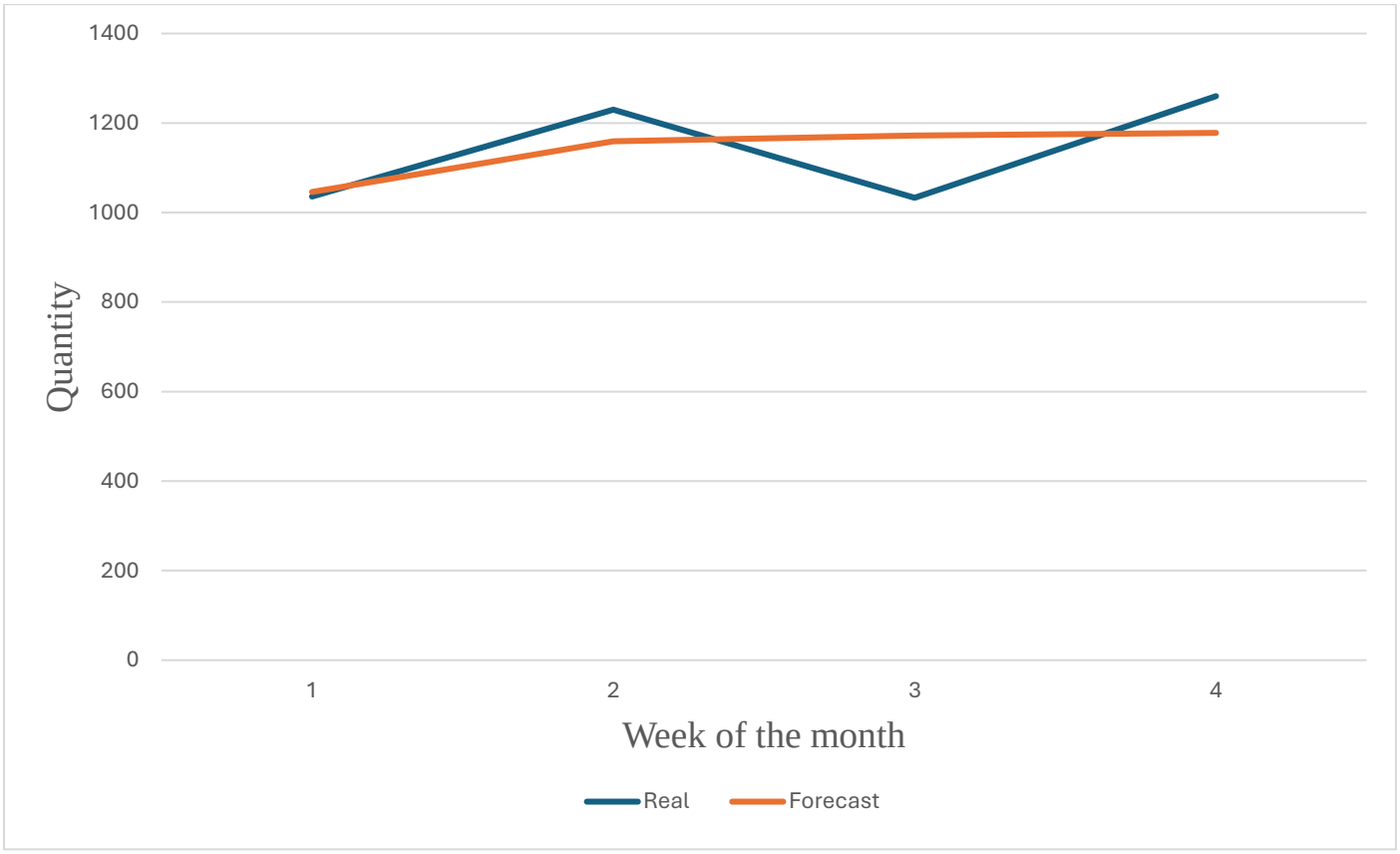

Fig 8. The real and forecast values of the model developed using MLP with the Use All Lags for product B

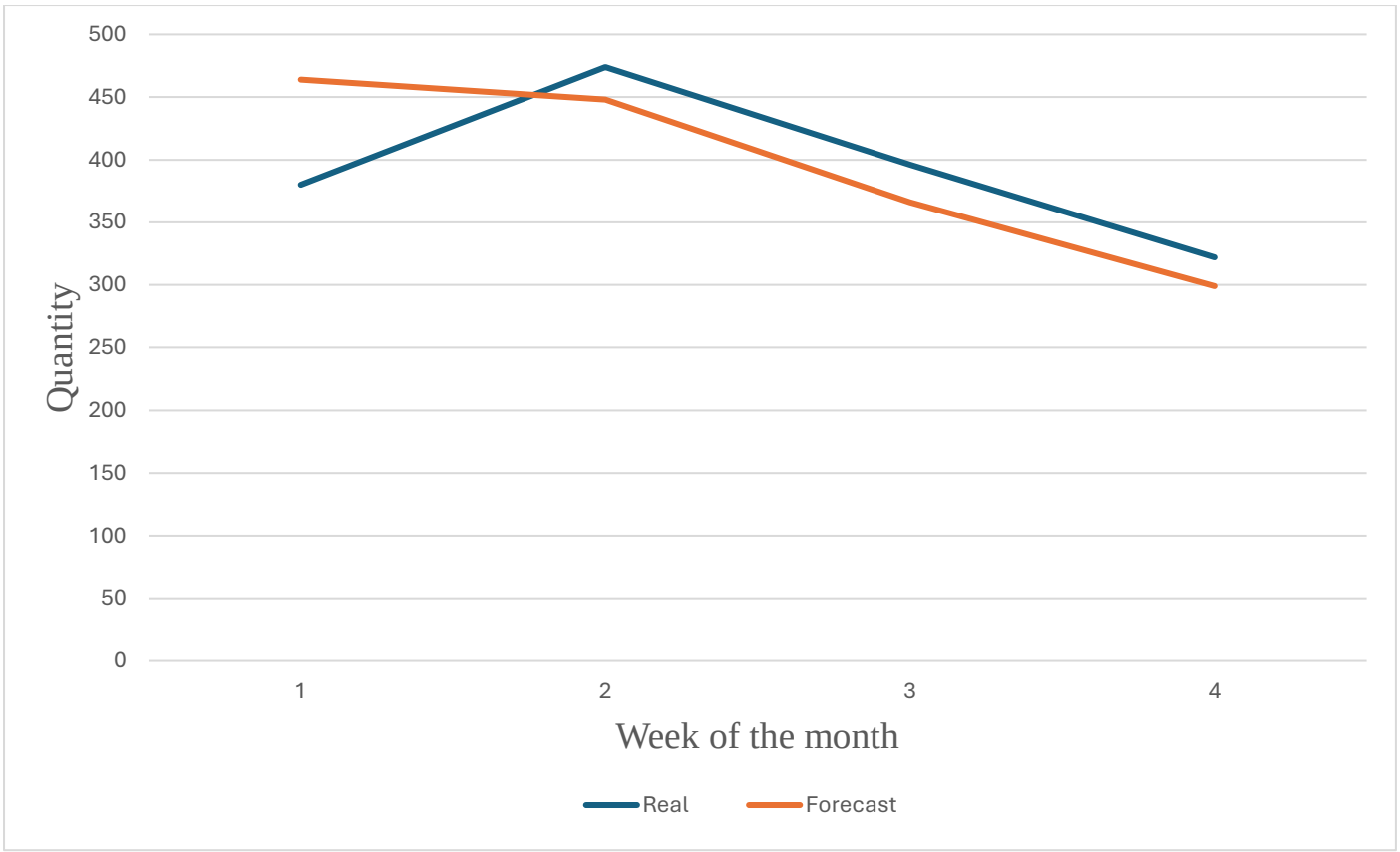

Fig 9. The real and forecast values of the model developed using MQRNN with the Use Best N for product C

When the results of the models developed with MLP have been examined:

- The lowest error rate for Product A was achieved with the Use Best N (13.9%).
- For Product B, the lowest error rate was obtained with the Use All Lags (6.68%), demonstrating very high success.
- The best performance for Product C has been observed with the Use All Lags (11.35%). However, the Use Correlation stands out with a significantly higher error rate of 26.32%, clearly separating it from the other options.

When the results of the models developed with MQRNN have been examined:

- For Product A, the lowest error rate was achieved with the Use All Lags (14.15%), closely followed by the Use Best N (14.85%).
- The best performance for Product B was observed with the Use All Lags (11.39%).
- The lowest error rate for Product C was obtained with the Use Best N (10.58%).

In the forecast models developed for July, the MLP method performed better with the Use All Lags option, while the MQRNN method achieved superior results with the Use Best N option. For the models developed for August, the MLP method again demonstrated superior performance with the Use All Lags option, whereas the MQRNN method showed comparable performance between the Use All Lags and Use Best N options. These analyses show that the performance of time delay options varies based on the method employed. These results help us understand the impact of seasonal performance variations and the methods used on prediction success. The lower MAPE values for Products A and B in July suggest that this period provided more consistent results for the models, while the higher success observed for Product C in August highlights the influence of seasonal factors and product-specific dynamics on model performance. Furthermore, the better performance of MLP compared to MQRNN in August underscores the importance of considering data structure and temporal dynamics when selecting models.

8. CONCLUSION

Time series data are encountered in many different problems today, one of which is demand forecasting. Since correctly modeling the effects of past data on future demand directly influences the accuracy of forecasts, the use of time lags in demand forecasting is of great importance. Time lags help capture future trends, seasonal changes, and other dynamics based on past data. Choosing the right lag parameters enhances the model's flexibility and improves forecast performance, leading to more reliable results. Therefore, choosing the appropriate time lags is critical. In this study, demand forecasting models have been developed to help e-commerce businesses gain a competitive advantage in the market by accurately predicting demand and the effect of time lag options on forecast performance has been analyzed comprehensively. An interface with hyperparametric flexibility has been developed for the forecasting models. Models have been created for two different months and three different products, and their performance has been evaluated using the MAPE metric. The lowest MAPE value for July has been obtained with the MQRNN model developed using product A, while the lowest MAPE value for August has been obtained with the MLP model developed using product B. These findings offer valuable insights for future model development, serving as a strategic guide to improve forecast accuracy and identify the most suitable method for each product. Contrary to the [21], this study offers a comprehensive analysis of the effects and role of lag options on demand forecasting of 3 different products, A, B, C. Furthermore, the structure and functionality of the hyperparametric interface employed in the analysis are thoroughly elucidated. Future studies could enhance forecast accuracy by analyzing the performance of different time lag options on complex data structures across various sectors and integrating external factors into the models. Additionally, the effects of time lag selection and hyperparameter optimization on forecast performance can be studied in greater depth. This approach can lead to more accurate forecast results and make significant contributions to the literature.

ORCID

Alim Toprak FIRAT https://orcid.org/0000-0001-7390-7453
Onur AYGÜN https://orcid.org/0009-0008-4534-3783
Mustafa GÖĞEBAKAN https://orcid.org/0009-0006-4042-3202
Ceren ULUS https://orcid.org/0000-0003-2086-6381
M. Fatih AKAY https://orcid.org/0000-0003-0780-0679

REFERENCES

[1]. Gong, K., Peng, Y., Wang, Y., & Xu, M. (2018). Time series analysis for C2C conversion rate. Electronic Commerce Research, 18, 763-789.
[2]. Akadji, A. F., & Dewantara, R. (2024). Big Data Analysis for Product Demand Prediction in Indonesian E-commerce. West Science Information System and Technology, 2(01), 9-17.
[3]. Nussipova, F., Rysbekov, S., Abdiakhmetova, Z., & Kartbayev, A. (2024). Optimizing loss functions for improved energy demand prediction in smart power grids. International Journal of Electrical & Computer Engineering (2088-8708), 14(3).
[4]. Rasul, K., Ashok, A., Williams, A. R., Ghonia, H., Bhagwatkar, R., Khorasani, A., Darvishi Bayazi, M. J., Adamopoulos, G., Riachi, R., Hassen, N., Biloš, M., Garg, S., Schneider, A., Chapados, N., Drouin, A., Zantedeschi, V., Nevmyvaka, Y., & Rish, I. (2024). Lag-Llama: Towards foundation models for probabilistic time series forecasting. *arXiv(Preprint)*. https://doi.org/10.48550/arXiv.2310.08278
[5]. Peláez-Rodríguez, C., Pérez-Aracil, J., Fister, D., Torres-López, R., & Salcedo-Sanz, S. (2024). Bike sharing and cable car demand forecasting using machine learning and deep learning multivariate time series approaches. Expert Systems with Applications, 238, 122264.
[6]. Khatun, A., Nisha, M. N., Chatterjee, S., & Sridhar, V. (2024). A novel insight on input variable and time lag selection in daily streamflow forecasting using deep learning models. Environmental Modelling & Software, 179, 106126.
[7]. Guo, J. (2024, May). Commodity Demand Forecasting of E-Commerce Merchants Based on the Stacking Fusion Model. In 2024 6th International Conference on Communications, Information System and Computer Engineering (CISCE) (pp. 1036-1040). IEEE.
[8]. Li, C. (2024). Commodity demand forecasting based on multimodal data and recurrent neural networks for E-commerce platforms. Intelligent Systems with Applications, 22, 200364.
[9]. Li, J., Fan, L., Wang, X., Sun, T., & Zhou, M. (2024). Product Demand Prediction with Spatial Graph Neural Networks. Applied Sciences, 14(16), 6989.
[10]. Liu, J., Wu, T., Wu, J., Chen, Z., Gong, J., & Chi, H. (2024). Forecasting Analysis of Demand for Agricultural Products in E-Commerce Based on Single Forecasting Model. In Artificial Intelligence Technologies and Applications (pp. 479-489). IOS Press.
[11]. Liu, J., Wu, T., Wu, J., Chen, Z., Gong, J., & Chi, H. (2024). Combinatorial Model-Based Demand Forecast Analysis of E-Commerce Agricultural Products. In Artificial Intelligence Technologies and Applications (pp. 508-516). IOS Press.
[12]. Liu, Z., Zhao, Y., Yang, S., Ju, J., Yang, L., Li, R., ... & Lu, W. (2024). Time Series Analysis of Product Demand Forecasting and Inventory Optimization on E-commerce Platforms. Journal of Electronics and Information Science, 9(1), 49-54.
[13]. Lv, J. (2024, February). Research on Inventory Management and Demand Forecasting of E-commerce Platform Based on ARIMA and LSTM Models. In 2024 IEEE 3rd International Conference on Electrical Engineering, Big Data and Algorithms (EEBDA) (pp. 494-499). IEEE.
[14]. Trisolvena, M. N., Masruroh, M., & Ginting, Y. M. (2024). Product Demand Forecast Analysis Using Predictive Models and Time Series Forecasting Algorithms on the Temu Marketplace Platform. International Journal Software Engineering and Computer Science (IJSECS), 4(2), 430-439.
[15]. Zhang, J., Zhang, Y., Tang, F., Song, Y., Deng, Y., & He, S. (2024, May). E-commerce Retail Merchandise Based on Optimized K-means Algorithm and Multi-model Fusion Demand Forecasting Research. In Proceedings of the 2024 International Conference on Smart City and Information System (pp. 512-516).

[16]. Zhao, Y. (2024). Research on E-Commerce Retail Demand Forecasting Based on SARIMA Model and K-means Clustering Algorithm. Academic Journal of Science and Technology, 10(3), 226-231.
[17]. Xu, Z., Zhang, L., Yang, S., Etesami, R., Tong, H., Zhang, H., & Han, J. (2024). F-FOMAML: GNN-enhanced meta-learning for peak period demand forecasting with proxy data. *arXiv(Preprint)*. https://doi.org/10.48550/arXiv.2406.16221
[18]. Wang, G., Gao, J., Shao, S., & Dong, Y. (2024, February). Platform Merchant Demand Prediction Based on Decision Tree and Multi-Layer Perceptron Models. In 2024 IEEE 3rd International Conference on Electrical Engineering, Big Data and Algorithms (EEBDA) (pp. 995-1000). IEEE.
[19]. Aravazhi, A. (2021). Hybrid machine learning models for forecasting surgical case volumes at a hospital. AI, 2(4), 512-526.
[20]. Zhang, X. Y., Watkins, C., & Kuenzel, S., Multi-quantile recurrent neural network for feeder-level probabilistic energy disaggregation considering roof-top solar energy. Engineering Applications of Artificial Intelligence, 110, 104707, 2022.
[21]. Fırat, A. T., Aygün, O., Göğebakan, M., Akay, M. F., & Ulus, C. Development of machine learning based demand forecasting models for the e-commerce sector. Uluslararası Mühendislik Tasarım ve Teknoloji Dergisi, 7(1), 13-20. https://doi.org/10.70669/ijedt.1567739

MAKÜ BBBF
BUCAK BİLGİSAYAR VE BİLİŞİM FAKÜLTESİ

Scientific Journal of Mehmet Akif Ersoy University, 2025, 8(1), Page:16-34

Research Article

Numerical Solution of First and Higher Order IVPs Via a Single Continuous Block Method

Jamiu Garba[1*], Khadeejah James Audu [1], Umaru Mohammed[1], Abd'gafar Tunde Tiamiyu[1]

[1]*Department of Mathematics, Faculty of Physical Sciences, Federal University of Technology, Minna, Nigeria*

ABSTRACT

This article focuses on the development and implementation of a single continuous collocation numerical scheme for solving first and higher-order ordinary differential equations (ODEs). By employing the interpolation and collocation technique on power series as basis function, we were able to come up with a continuous scheme from which block methods for effectively solving first and higher order ODEs were derived. This is better and faster than the traditional way of developing a continuous scheme for a specific order of ODE. The method's accuracy is determined to be of order seven, establishing its consistency. Results from the implementation of our method show its applicability on nonlinear equations and application problems from first, second, third, and fourth order ODEs that are of significant implications on various fields in physics, engineering, biology and mathematics. Furthermore, the numerical results generated by the method reveal its effectiveness and accuracy, and also its superiority over some methods that exist in literature.

Keywords: Accuracy, Collocation, Continuous scheme, First order ODEs, Higher order ODEs, Interpolation

1. INTRODUCTION

Mathematical modeling entails the ability to transform issues from a particular field of application into workable mathematical expressions, whose theoretical and numerical evaluations provide understanding, solutions, and recommendations that are advantageous for the initial applications [35]. Numerous fundamental laws of nature in Physics, Chemistry, Biology, and astronomy are most effectively articulated through differential equations [34]. Their applications are widespread in mathematics, particularly in geometry, as well as in engineering, behavioural sciences, industrial mathematics, artificial intelligence. This kind of problem can be represented using either first-order or higher-order ODEs. First-order ODEs are commonly used in studying problems for example, determining the movement of an object that is ascending or descending while experiencing air resistance, and calculating the current in an electrical circuit; population expansion; radioactive decay; mixture problems; and so on. Second-order ODEs are also commonly used while analyzing vibrating systems, electromagnetism, and electrical circuits with capacitors, resistors, and inductors. Third order differential equations can also be used to solve physical problems like thin film flow, electromagnetic waves and gravity-driven flows. In general, solutions to differential equations are used to forecast the behavior of a system at a later time or in an unknown location. But there aren't many analytical ways to solve ODEs in a continuous or closed form, and nonlinear ODEs can be hard to solve or may not have a closed form of solution at all.

Given the challenges associated with solving most of these problems, numerous researchers have focused extensively on applying numerical methods to provide approximate solutions for differential equations. Numerical methods are particularly effective for addressing mathematical problems, leveraging the speed and efficiency of modern digital computers in performing arithmetic operations [42]. The process of solving these problems with high-precision digital computers typically begins with initial data, followed by the execution of suitable algorithms to produce the desired outcomes [43]. Common approaches for solving IVPs are classified into single-step methods, such as the Runge-Kutta methods [21, 43], and multistep methods [1 – 10].

Over the years, various techniques for deriving continuous linear multistep methods (LMMs) aimed at directly solving IVPs have been extensively discussed in the literature. Key techniques include collocation, interpolation, integration,

***Corresponding Author Email:** jamiu.garba@st.futminna.edu.ng
Submitted: 03.01.2025 **Revision Requested:** 06.03.2025 **Last Revision Received:** 10.03.2025
Accepted: 13.03.2025 **Published Online:** 06.05.2025

Burdur Mehmet Akif Ersoy University Press

Cite this article as: Garba, J., Audu, K. J., Mohammed, U., Taimiyu, A. T. (2025). Numerical Solution of First and Higher Order IVPs Via a Single Continuous Block Method. Scientific Journal of Mehmet Akif Ersoy University, 8(1): 16-34.
DOI: https://doi.org/10.70030/sjmakeu.1611387 https://dergipark.org.tr/sjmakeu

and the use of interpolating polynomials. Basis functions like power series, Chebyshev polynomials, Legendre polynomials, and trigonometric functions have been used for this purpose. (see [7, 10,17, 18, 22, 44, 45, 46]). Researchers have found linear multistep method very useful in solving ODEs. Some authors [7, 10, 11, 12] have develop effective methods for first order ODEs while [6, 13, 27, 35, 42, 44] have implemented linear multistep methods on second order ODEs with comparative accuracy. Other higher ODEs have been solved numerically by [2, 14, 15 31, 37, 39, 43, 47, 48, 49]. All of the linear multistep methods developed by these researchers were for a specific order of ODEs. Recently Kuboye and Adeyefa [22] developed a linear multistep method for first, second and third order ODEs.

In this research, we derive a single continuous technique from a hybrid block approach that effectively handles first, second, third, and fourth-order initial value problems. The subsequent portion of this article describes the method's derivation, while the third sets out the analysis that ensures the method's validity. In part four, a variety of numerical problems are addressed, results are displayed graphically, and the method's efficacy is demonstrated by comparing absolute errors to those of recent existing methods in literature. In section five, we offer broad conclusions on our findings.

The general form of the ODE is expressed as follows:

$$\begin{aligned} z^{(m)} &= \left(z(t), z'(t), \ldots, z^{(m-1)}(t)\right) \\ z(t_0) &= z_0, z'(t_0) = z_0', \ldots, z^{(m-1)}(t_0) = z_0^{(m-1)} \end{aligned} \tag{1}$$

2. DERIVATION OF THE METHOD

As stated in the introduction, the continuous representation of our method will be generated from one of the orthogonal basis functions, preferably the power series, due to its ease of usage. Therefore, the approximate solution to (1) is as follows:

$$Z(t) = \sum_{j=0}^{m+n-1} c_j t^j \tag{2}$$

where *m* and *n* are respectively the number of interpolation and collocation points. We derive a two-step linear multistep method using (2) with degree *m* + *n* – *1* = 10. (2) is interpolated at $t = t_n$ and the first, second and third derivatives are collocated at $t = t_{n+2}$. Furthermore, we collocate the fourth derivative at distinct points $t = t_{n+j}; \left(j = 0, \frac{1}{4}, \frac{3}{4}, 1, \frac{5}{4}, \frac{7}{4}, 2 \right)$. All these processes lead to a system of nonlinear equations in t_{n+j} with c_j's as the unknowns.

$$\left.\begin{aligned} Z(t_n) &= z_n \\ Z'(t_{n+2}) &= f_{n+2} \\ Z''(t_{n+2}) &= g_{n+2} \\ Z'''(t_{n+2}) &= p_{n+2} \\ Z^{iv}(t_{n+j}) &= g_{n+j} \end{aligned}\right\} \tag{3}$$

We solved for the unknowns using the matrix inversion method via a mathematical software – Maple 2015; and then substitute c_j's values back into (2) to obtain the continuous scheme in the form:

$$Z(t) = z_n + h\alpha'(t) f_{n+2} + h^2\alpha''(t) g_{n+2} + h^3\alpha'''(t) p_{n+2} + h^4\left(\beta_0(t) q_n + \beta_{\frac{1}{4}}(t) q_{n+\frac{1}{4}} + \beta_{\frac{3}{4}}(t) q_{n+\frac{3}{4}} + \beta_1(t) q_{n+1} + \beta_{\frac{5}{4}}(t) q_{n+\frac{5}{4}} + \beta_{\frac{7}{4}}(t) q_{n+\frac{7}{4}} + \beta_2(t) q_{n+2} \right) \tag{4}$$

Evaluating (4) at the non-interpolating points gives 6 discrete schemes that form the block method as follows:

$$z_{n+\frac{1}{4}} = z_n + \frac{1}{4}hf_{n+2} - \frac{15}{32}h^2 g_{n+2} + \frac{169}{384}h^3 p_{n+2} + \frac{1}{13005619200}h^4 \begin{pmatrix} 1837375q_n - 21942710q_{n+\frac{1}{4}} - 353393782q_{n+\frac{3}{4}} + 131048330q_{n+1} \\ -1219981490q_{n+\frac{5}{4}} - 1715628930q_{n+\frac{7}{4}} - 409914793q_{n+2} \end{pmatrix} \tag{5}$$

$$z_{n+\frac{3}{4}} = z_n + \frac{3}{4}hf_{n+2} - \frac{39}{32}h^2 g_{n+2} + \frac{129}{128}h^3 p_{n+2} + \frac{1}{4335206400}h^4 \begin{pmatrix} 1070517q_n - 8616430q_{n+\frac{1}{4}} - 163366014q_{n+\frac{3}{4}} + 67852190q_{n+1} \\ -776333082q_{n+\frac{5}{4}} - 1257048730q_{n+\frac{7}{4}} - 312696051q_{n+2} \end{pmatrix} \tag{6}$$

$$z_{n+1} = z_n + hf_{n+2} - \frac{3}{2}h^2 g_{n+2} + \frac{7}{6}h^3 p_{n+2} + \frac{1}{793800}h^4 \begin{pmatrix} 283q_n - 1880q_{n+\frac{1}{4}} - 28840q_{n+\frac{3}{4}} + 11585q_{n+1} \\ -150248q_{n+\frac{5}{4}} - 260760q_{n+\frac{7}{4}} - 66265q_{n+2} \end{pmatrix} \tag{7}$$

$$z_{n+\frac{5}{4}} = z_n + \frac{5}{4}hf_{n+2} - \frac{55}{32}h^2 g_{n+2} + \frac{485}{384}h^3 p_{n+2} + \frac{1}{2601123840}h^4 \begin{pmatrix} 1221691q_n - 7249910q_{n+\frac{1}{4}} - 88889206q_{n+\frac{3}{4}} + 28468370q_{n+1} \\ -491995826q_{n+\frac{5}{4}} - 906420930q_{n+\frac{7}{4}} - 234924589q_{n+2} \end{pmatrix} \tag{8}$$

$$z_{n+\frac{7}{4}} = z_n + \frac{7}{4}hf_{n+2} - \frac{63}{32}h^2 g_{n+2} + \frac{511}{384}h^3 p_{n+2} + \frac{1}{1857945600}h^4 \begin{pmatrix} 1082161q_n - 5977070q_{n+\frac{1}{4}} - 58808638q_{n+\frac{3}{4}} + 10672550q_{n+1} \\ -34408610q_{n+\frac{5}{4}} - 664954650q_{n+\frac{7}{4}} - 176256247q_{n+2} \end{pmatrix} \tag{9}$$

$$z_{n+2} = z_n + 2hf_{n+2} - 2h^2 g_{n+2} + \frac{4}{3}h^3 p_{n+2} + \frac{1}{99225}h^4 \begin{pmatrix} 29q_n - 160q_{n+\frac{1}{4}} - 1568q_{n+\frac{3}{4}} + 280q_{n+1} \\ -1568q_{n+\frac{5}{4}} - 17760q_{n+\frac{7}{4}} - 4712q_{n+2} \end{pmatrix} \tag{10}$$

3. ANALYSIS OF THE BLOCK METHOD

At this point, following the approaches adopted in [18, 27, 28], the order of accuracy of our method is:

$(7,7,7,7,7,7)^T$ and error constants given as

$$\left(\frac{221815417}{3515887072051200}, \frac{66674131}{390654119116800}, \frac{593}{2682408960}, \frac{184617553}{703177414410240}, \frac{150066817}{502269581721600}, \frac{251}{838252800} \right)^T.$$

We show the working of equation (10) as follows:

$$\left.\begin{aligned}
&\varepsilon_0 = 1-1=0\\
&\varepsilon_1 = 2(1)-2=0\\
&\varepsilon_2 = \frac{1}{2!}\left(2^2\right)-(2\times 2)-(-2)=0\\
&\varepsilon_3 = \frac{1}{3!}\left(2^3\right)-\frac{1}{2!}\left(2^2\times 2\right)-(-2\times 2)-\left(\frac{4}{3}\right)=0\\
&\varepsilon_4 = \frac{1}{4!}\left(2^4\right)-\frac{1}{3!}\left(2^3\times 2\right)-\frac{1}{2!}\left(-2\times 2^2\right)-\left(2\times\frac{4}{3}\right)=0\\
&\varepsilon_5 = \frac{1}{5!}\left(2^5\right)-\frac{1}{4!}\left(2^4\times 2\right)-\frac{1}{3!}\left(-2\times 2^3\right)-\frac{1}{2!}\left(2^2\times\frac{4}{3}\right)-\\
&\left(\frac{58}{99225}+\frac{64}{19845}+\frac{64}{2025}-\frac{16}{2835}+\frac{2624}{14175}+\frac{2368}{6615}+\frac{9424}{99225}\right)=0\\
&\vdots\\
&\varepsilon_{11} = \frac{1}{11!}\left(2^{11}\right)-\frac{1}{10!}\left(2^{10}\times 2\right)-\frac{1}{9!}\left(-2\times 2^9\right)-\frac{1}{8!}\left(2^8\times\frac{4}{3}\right)-\\
&\frac{1}{7!}\left(\left(\frac{1}{4}\right)^7\frac{64}{19845}+\left(\frac{3}{4}\right)^7\frac{64}{2025}-(1)^7\frac{16}{2835}+\left(\frac{5}{4}\right)^7\frac{2624}{14175}+\left(\frac{7}{4}\right)^7\frac{2368}{6615}+(2)^7\frac{9424}{99225}\right)\\
&=\frac{251}{838252800}
\end{aligned}\right\}$$

The techniques described in sections (5) through (10) can typically be represented using a matrix-based difference equation, which is detailed in the following steps.:

$$A^{(1)}Z_w = A^{(0)}Z_{w-1} + hBF_w + h^2CG_w + h^3DP_w + h^4\left[E^{(0)}Q_{w-1} + E^{(1)}Q_w\right] \tag{11}$$

where

$$\left.\begin{aligned}
&Z_w = \left(z_{n+\frac{1}{4}}, z_{n+\frac{3}{4}}, z_{n+1}, z_{n+\frac{5}{4}}, z_{n+\frac{7}{4}}, z_{n+2}\right)^T,\\
&Z_{w-1} = \left(z_{n-\frac{1}{4}}, z_{n-\frac{3}{4}}, z_{n-1}, z_{n-\frac{5}{4}}, z_{n-\frac{7}{4}}, z_{n-2}\right)^T,\\
&F_w = \left(f_{n+2}\right)^T, G_w = \left(g_{n+2}\right)^T, P_w = \left(p_{n+2}\right)^T,\\
&Q_w = \left(q_{n+\frac{1}{4}}, q_{n+\frac{3}{4}}, q_{n+1}, q_{n+\frac{5}{4}}, q_{n+\frac{7}{4}}, q_{n+2}\right)^T,\\
&Q_{w-1} = \left(q_{n-\frac{1}{4}}, q_{n-\frac{3}{4}}, q_{n-1}, q_{n-\frac{5}{4}}, q_{n-\frac{7}{4}}, q_{n-2}\right)^T
\end{aligned}\right\}$$

and the matrices $A^{(1)}, A^{(0)}, B, C, D, E^{(1)}$ and $E^{(0)}$ are matrices whose elements are given by the coefficients of the block method.

Definition 3.1. The newly developed two-step hybrid block method (5 – 10) is considered zero-stable if and only if the first characteristic polynomial $\xi(t)$ has roots that satisfy $|t| \le 1$, and for the roots $|t| \le 1$, their multiplicity does not exceed one. The characteristic function of this newly derived method is presented below:

$$\xi(\lambda) = \lambda A^{(1)} - A^{(0)} \tag{12}$$

$$\xi(\lambda) = \lambda \begin{pmatrix} 1&0&0&0&0&0\\ 0&1&0&0&0&0\\ 0&0&1&0&0&0\\ 0&0&0&1&0&0\\ 0&0&0&0&1&0\\ 0&0&0&0&0&1 \end{pmatrix} - \begin{pmatrix} 0&0&0&0&0&-1\\ 0&0&0&0&0&-1\\ 0&0&0&0&0&-1\\ 0&0&0&0&0&-1\\ 0&0&0&0&0&-1\\ 0&0&0&0&0&-1 \end{pmatrix} = (\lambda+1)\lambda^5 \tag{13}$$

the solution of which is $\lambda = (0,0,0,0,0,1)$. Hence, our method is zero-stable.

Definition 3.2. If the order of the hybrid block approach is greater than or equal to one, it is consistent. Our method is consistent as a result of order *p*=7.

Theorem 3.1. Zero stability and consistency are adequate requirements for the convergence of a linear multistep method [26, 17, 27]. The new hybrid block technique is convergent due to its zero-stability and consistency.

To analyze the stability characteristics of the developed scheme, it is applied to a standard test problem.

$$z' = \lambda z, \qquad z'' = \lambda^2 z, \qquad z''' = \lambda^3 z, \qquad z^{iv} = \lambda^4 z, \qquad \mathrm{Re}(\lambda) < 0 \tag{14}$$

to yield

$$Z_w = \xi(w) Z_{w-1}, \quad w = z\lambda \tag{15}$$

where the matrix $\xi(w)$ is given as:

$$\zeta(w) = \left(Z^{(1)} - wF - w^2G - w^3P - w^4Q^{(1)}\right)^{-1}\left(Z^{(0)} + w^4Q^{(0)}\right) \tag{16}$$

The Matrix $\zeta(w)$ has eigenvalues $(0,0,0,0,0,\lambda_6)$, and the dominant eigenvalue λ_6: $\mathbb{C} \to \mathbb{C}$ is a rational function with real coefficients given by

$$R(w) = \frac{Y(w)}{X(w)}$$

The stability region is illustrated in Figure 1, indicating that the method is A-stable.

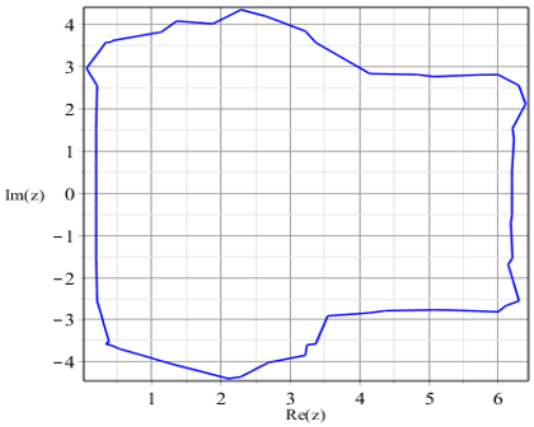

Fig 1. Region of absolute stability of the method

4. NUMERICAL EXPERIMENT

The recently developed hybrid block method is applied to both first and higher-order ODEs. This continuous approach results in a primary discrete two-step multi-derivative block method (as outlined in equations 5 – 10), along with supplementary methods that are integrated and utilized as a block method to generate approximations concurrently. $\{y_{n+1}, y_{n+2}\}$ at a block points $\{x_{n+1}, x_{n+2}\}$, $h = x_{n+1} - x_n$, $n = 0,..., \mathrm{N}-2$, on a partition $[a,b]$, where $a,b \in$ is the interval of integration, h is the constant step-size, n is a grid index and $N > 0$ is the number of steps. We obtain initial conditions at $x_{n+2}, n = 0,1,..., N-2$, using the computed values y_{n+2} over smaller intervals $[x_0, x_2],...,[x_{N-2}, x_N]$. For example, when $n = 0, [y_1, y_2]$ are acquired at the same time across the smaller interval $[x_0, x_2]$, as y_0 is established from the initial value problem, for $n = 2, [y_3, y_4]$ are also acquired simultaneously over the smaller interval $[x_3, x_4]$, as is now understood from the earlier section, and so forth. Consequently, the smaller interval $[x_n, x_{n+2}]$ does not overlap, and the solutions derived in this manner are more precise than those obtained through traditional predictor-corrector methods. Similar approach is applied to higher order ODEs considered – the first, second and third derivatives of the continuous scheme are evaluated at all points to cater for the higher derivative terms in the general higher order ODEs considered. We present the graphical solution of the problems and compare absolute errors with some existing methods in the literature.

4.1 First-Order Problems

The block method (5) to (10) is used directly on some first order problems.

Problem 1: The SIR model is an epidemiological framework that tracks the hypothetical number of individuals within a closed population who are affected by an infectious disease over time. This type of model derives its name from the interconnected equations that describe the populations of susceptible individuals *S(t)*, infected individuals *I(t)*, and recovered individuals *R(t)*. It serves as an effective and simple model for various infectious diseases, such as measles, mumps, and rubella. The model is represented by the three interconnected equations provided below.:

$$\frac{dS}{dt} = \mu(1-S) - \beta IS$$

$$\frac{dI}{dt} = \mu I - \gamma I + \beta IS$$

$$\frac{dR}{dt} = \mu R + \gamma I$$

where μ, β and γ are positive parameters. Define *z* to be:

$$z = S + I + R$$

and adding (13), (14) and (15), the following evolution equation for *z* is obtained.

$$z' = \mu(1-z)$$

Kuboye and Adeyefa [22] solved this problem with the following parameters:

$$\mu = \frac{1}{2}, \quad z(0) = \frac{1}{2}, \quad h = 0.1$$

Exact solution: $z(t) = 1 - \mu e^{-\mu t}$

Problem 2: We consider the Riccati differential equation solved in Khalsaraei *et. al.* [20]

$$z' = 1 + 2z - z^2; \quad z(0) = 0, \quad 0 \le t \le 10$$

Exact solution: $z(t) = 1 + \sqrt{2}\tanh\left(\sqrt{2}t + \frac{1}{2}\log\left(\frac{\sqrt{2}-1}{\sqrt{2}+1}\right)\right)$

Problem 3: We consider the nonlinear system of stiff chemical problem solved in Akinfenwa *et. al.* [11]

$$\left.\begin{aligned} z_1' &= \lambda z_1 + z_2^2, \quad z_1(0) = -\frac{1}{\lambda+2} \\ z_2' &= -z_2, \quad z_2(0) = 1 \end{aligned}\right\}$$

The exact solution is given as $z_1(t) = -\frac{\exp(-2t)}{\lambda+2}, \quad z_2(t) = \exp(-t)$ where $\lambda = 10000$.

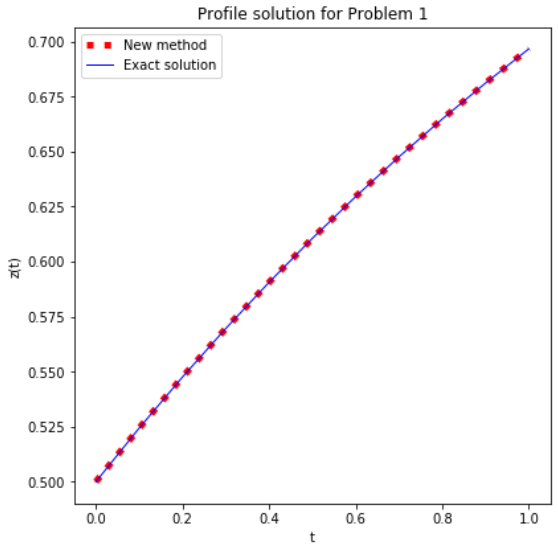

Fig 2. The profile solution for problem 1

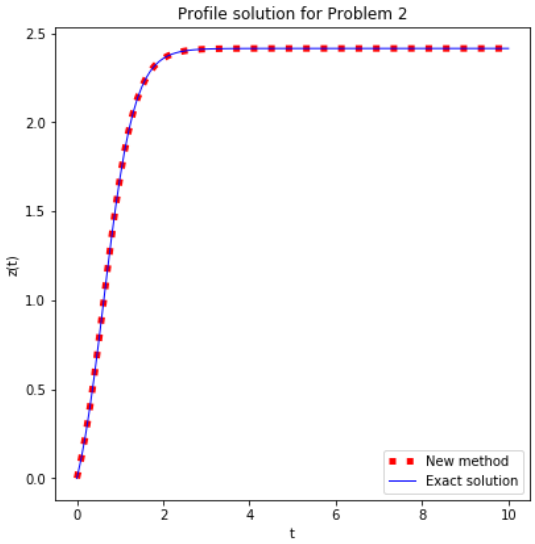

Fig 3. The profile solution for problem 2

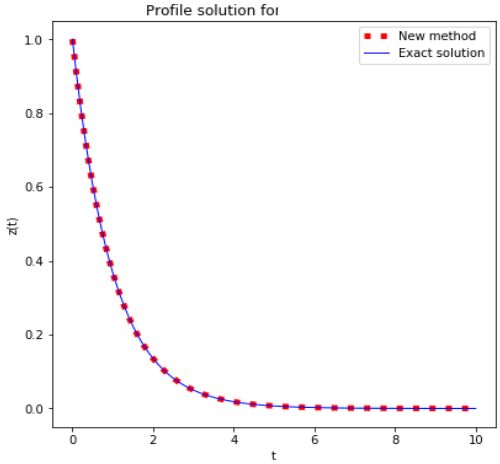

Fig 4. The profile solution for problem 3

Table 1. Error comparison for Problem 1

t	Error in [22] $(h=0.1)$	Error in [19] $(h=0.1)$	Error in [4] $(h=0.1)$	Error in New method $(h=0.1)$
0.1	3.846×10^{-13}	1.998×10^{-15}	9.104×10^{-15}	4.736×10^{-22}
0.2	7.319×10^{-13}	3.886×10^{-15}	7.105×10^{-15}	6.225×10^{-22}
0.3	1.044×10^{-12}	5.440×10^{-15}	8.882×10^{-15}	1.021×10^{-21}
0.4	1.324×10^{-12}	6.994×10^{-15}	2.121×10^{-14}	1.127×10^{-21}
0.5	1.575×10^{-12}	8.216×10^{-15}	1.368×10^{-13}	1.459×10^{-21}
0.6	1.797×10^{-12}	9.548×10^{-15}	7.983×10^{-13}	1.529×10^{-21}
0.7	1.995×10^{-12}	1.055×10^{-14}	3.699×10^{-12}	1.805×10^{-21}
0.8	2.168×10^{-12}	1.132×10^{-14}	-	1.845×10^{-21}
0.9	2.320×10^{-12}	1.221×10^{-14}	-	1.845×10^{-21}
1.0	2.452×10^{-12}	1.288×10^{-14}	-	2.086×10^{-21}

Table 2. Error comparison for Problem 2

t	Error in [19] $(h=0.05)$	Error in [4] $(h=0.1)$	Error in New method $(h=0.1)$
1	1.418×10^{-11}	9.104×10^{-15}	2.168×10^{-11}
2	7.234×10^{-13}	7.105×10^{-15}	1.052×10^{-12}
3	1.163×10^{-13}	8.882×10^{-15}	5.216×10^{-14}
4	2.132×10^{-14}	2.121×10^{-14}	3.144×10^{-15}
5	2.664×10^{-15}	1.368×10^{-13}	1.970×10^{-16}
6	4.441×10^{-16}	7.983×10^{-13}	1.233×10^{-17}
7	4.441×10^{-16}	3.699×10^{-12}	7.213×10^{-19}
8	4.441×10^{-16}	-	4.702×10^{-20}
9	4.441×10^{-16}	-	2.977×10^{-21}
10	4.441×10^{-16}	-	1.850×10^{-22}

Table 3. Error comparison for Problem 3

t	z_i	Error in [20] $(h=0.0001)$	Error in [11] $(h=0.01)$	Error in [26] $(h=0.1)$	Error in New method $(h=0.1)$
3	(z_1)	1.779×10^{-20}	2.030×10^{-19}	3.790×10^{-22}	1.972×10^{-23}
	(z_2)	2.079×10^{-20}	1.440×10^{-14}	2.998×10^{-18}	1.981×10^{-18}
5	(z_1)	2.493×10^{-19}	1.200×10^{-20}	1.60×10^{-21}	6.019×10^{-25}
	(z_2)	4.664×10^{-13}	3.210×10^{-15}	6.740×10^{-19}	4.468×10^{-19}
10	(z_1)	5.743×10^{-20}	1.110×10^{-20}	7.120×10^{-20}	5.466×10^{-29}
	(z_2)	6.346×10^{-12}	4.380×10^{-17}	9.080×10^{-21}	6.021×10^{-21}

4.2 Second-Order Problems

In order to implement our method on second order IVPs, we take the first derivative of the continuous scheme (4) and evaluated at points $t_{n+j}, \left(j = 0, \frac{1}{4}, \frac{3}{4}, 1, \frac{5}{4}, \frac{7}{4}, 2 \right)$ with the block schemes (5 – 10).

Problem 4: Consider the nonlinear second order IVP from Abdelrahim and Omar [1].

$z'' = t(z')^2,\ z(0) = 1,\ z'(0) = 0.5$

The exact solution is: $z(t) = \arctan h\left(\frac{1}{2}t\right) + 1$

Problem 5: *Cooling of a body.* Source: Kwanamu *et. al.* [24]

The temperature *z* degree of a body *t* minutes after being placed in a certain room, satisfies the differential equation

$$3\frac{d^2z}{dt} + \frac{dz}{dt} = 0,\quad z(0) = 60,\ \left.\frac{dz}{dt}\right|_{t=0} = -\frac{80}{9}$$

With the analytic solutions: $z(t) = \frac{80}{3}e^{-\frac{t}{3}} + \frac{100}{3}$

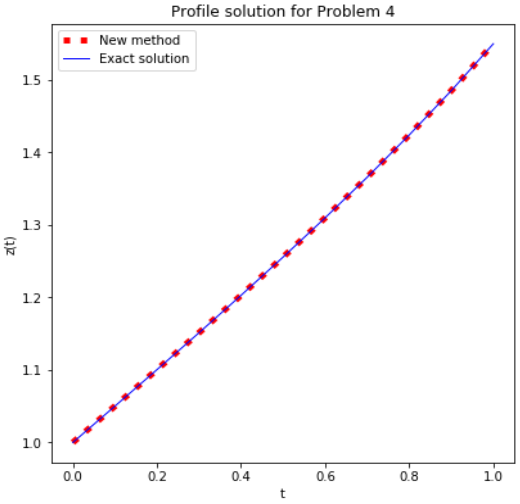

Fig 5. The profile solution for problem 4

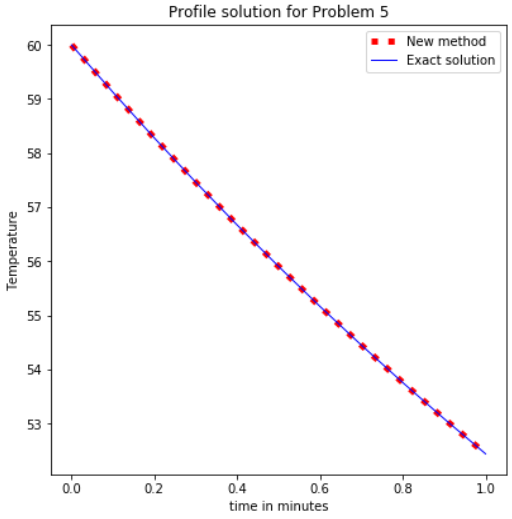

Fig 6. The profile solution for problem 5

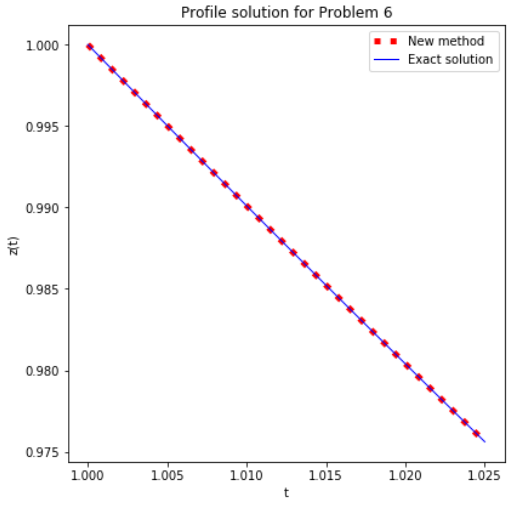

Fig 7. The profile solution for problem 6

Problem 6: Consider the nonlinear problem:

$z'' - 2z^3 = 0$ (Source: Ogunlaran & Kehinde [30])

with the following initial conditions

$z(1)=1, z'(1)=-1$ and $h=\frac{0.1}{40}$

Exact solution: $Z(t)=\frac{1}{t}$

Table 4. Error comparison for Problem 4

t	Error in [1] $\left(h=\frac{1}{30}\right)$	Error in [7] $(h=0.1)$	Error in new method $(h=0.1)$	Error in New method $\left(h=\frac{1}{30}\right)$
0.1	1.310×10^{-16}	1.957×10^{-13}	1.539×10^{-15}	6.552×10^{-20}
0.2	3.975×10^{-14}	6.040×10^{-13}	6.459×10^{-15}	2.934×10^{-19}
0.3	1.021×10^{-14}	1.262×10^{-12}	1.791×10^{-14}	7.771×10^{-19}
0.4	3.304×10^{-13}	3.715×10^{-12}	3.972×10^{-14}	1.726×10^{-18}
0.5	-	7.919×10^{-12}	8.698×10^{-14}	3.588×10^{-18}
0.6	-	1.416×10^{-11}	1.781×10^{-13}	7.353×10^{-18}
0.7	-	3.616×10^{-11}	3.987×10^{-13}	1.531×10^{-17}
0.8	-	7.473×10^{-11}	8.522×10^{-13}	3.306×10^{-17}
0.9	-	1.335×10^{-10}	2.182×10^{-12}	7.538×10^{-17}
1.0	1.293×10^{-12}	4.317×10^{-10}	5.160×10^{-12}	1.844×10^{-16}

Table 5. Error comparison for Problem 5, h = 0.1

t	Error in [38]	Error in [41]	Error in [33]	Error in [24]	Error in New method
0.1	3.550×10^{-11}	2.300×10^{-17}	7.476×10^{-06}	-	9.000×10^{-23}
0.2	4.580×10^{-11}	1.710×10^{-16}	2.939×10^{-05}	4.000×10^{-18}	3.600×10^{-22}
0.3	7.000×10^{-11}	4.370×10^{-16}	6.480×10^{-05}	9.000×10^{-18}	8.300×10^{-22}
0.4	6.500×10^{-11}	8.130×10^{-16}	1.128×10^{-05}	1.700×10^{-17}	1.450×10^{-21}
0.5	3.330×10^{-11}	1.290×10^{-15}	1.725×10^{-04}	2.600×10^{-17}	2.230×10^{-21}
0.6	4.200×10^{-11}	1.864×10^{-15}	2.431×10^{-04}	3.800×10^{-17}	3.170×10^{-21}
0.7	4.380×10^{-11}	2.525×10^{-15}	3.238×10^{-04}	5.100×10^{-17}	4.240×10^{-21}
0.8	1.070×10^{-10}	3.269×10^{-15}	4.139×10^{-04}	6.500×10^{-17}	5.410×10^{-21}
0.9	6.580×10^{-11}	4.089×10^{-15}	5.127×10^{-04}	8.100×10^{-17}	6.720×10^{-21}
1.0	1.6900×10^{-10}	4.980×10^{-15}	6.195×10^{-04}	9.700×10^{-17}	8.140×10^{-21}

Table 6. Error comparison for Problem 6

t	Exact Solution	Error in [30]	Error in New method
1.0025	0.997506234413965	3.160×10^{-17}	5.400×10^{-29}
1.0050	0.995024875621891	7.550×10^{-17}	2.480×10^{-28}
1.0075	0.992555831265509	1.190×10^{-16}	5.650×10^{-28}
1.0100	0.990099009900990	1.510×10^{-16}	1.005×10^{-27}
1.0125	0.987654320987654	1.810×10^{-16}	1.566×10^{-27}
1.0150	0.985221674876847	2.220×10^{-16}	2.241×10^{-27}
1.0175	0.982800982800983	2.630×10^{-16}	3.031×10^{-27}
1.0200	0.980392156862745	2.930×10^{-16}	3.929×10^{-27}
1.0225	0.977995110024450	3.210×10^{-16}	4.936×10^{-27}
1.0250	0.975609756097561	3.590×10^{-16}	6.044×10^{-27}

4.3 Third-Order Problems

In order to implement our method on third order IVPs, we take the first and second derivatives of the continuous scheme (4) and evaluated at points $t_{n+j}, \left(j = 0, \frac{1}{4}, \frac{3}{4}, 1, \frac{5}{4}, \frac{7}{4}, 2 \right)$ with the block schemes (5 – 10).

Problem 7: Consider the nonlinear IVP:

$$z''' = \left(2tz'' + 1\right) z'; \quad z(0) = 1, \quad z'(0) = 0.5, \quad z''(0) = 0$$

The exact solution is: $z(t) = \tan^{-1}\left(\frac{1}{2}t\right) + 1$

Source: Yakusak &. Owolanke [47]

Problem 8: Consider the nonlinear ODE

$$z''' = \frac{3}{8z^5}; \quad z(0) = 1, \quad z'(0) = \frac{1}{2}, \quad z''(0) = -\frac{1}{4}$$

Exact solution: $z(t) = \frac{1}{\sqrt{1+t}}$

Source: Adeyeye & Omar [8]

Problem 9:

$$z''' = tz'' - (tz)^2 + t\sin(t) - \cos(t) + t^2\sin^2(t); \quad z(0) = 0, \quad z'(0) = 1, \quad z''(0) = 0$$

Exact solution: $z(t) = \sin(t)$

Source: Source: Adeyeye & Omar [8]

Table 7. Error comparison for Problem 7

t	Error in [47] $(h = 0.01)$	Error in [9] $(h = 0.01)$	Error in [31] $(h = 0.01)$	Error in New method $(h = 0.01)$
0.1	9.609×10^{-12}	1.931×10^{-08}	2.043×10^{-14}	4.401×10^{-24}
0.2	7.072×10^{-10}	5.608×10^{-07}	8.371×10^{-14}	3.661×10^{-23}
0.3	6.693×10^{-09}	3.755×10^{-06}	2.813×10^{-13}	1.353×10^{-22}
0.4	3.142×10^{-08}	1.340×10^{-05}	7.667×10^{-13}	3.665×10^{-22}
0.5	1.051×10^{-07}	3.259×10^{-05}	1.853×10^{-12}	8.554×10^{-22}
0.6	2.852×10^{-07}	5.816×10^{-05}	4.163×10^{-12}	1.855×10^{-21}
0.7	6.777×10^{-07}	7.152×10^{-05}	8.965×10^{-12}	3.905×10^{-21}
0.8	1.467×10^{-06}	2.564×10^{-05}	-	8.224×10^{-21}
0.9	2.983×10^{-06}	1.709×10^{-04}	-	1.774×10^{-20}
1.0	6.189×10^{-06}	6.706×10^{-04}	-	4.003×10^{-20}

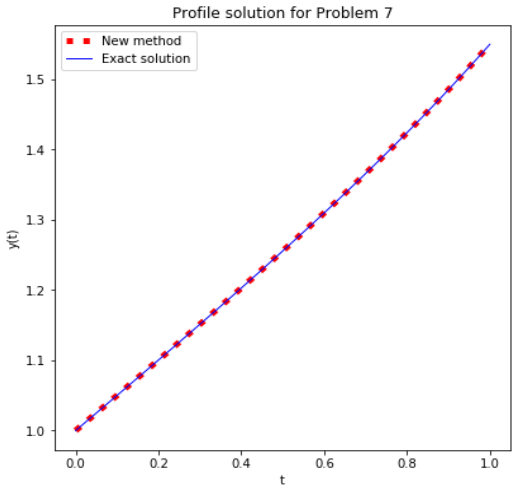

Fig 8. The profile solution for problem 7

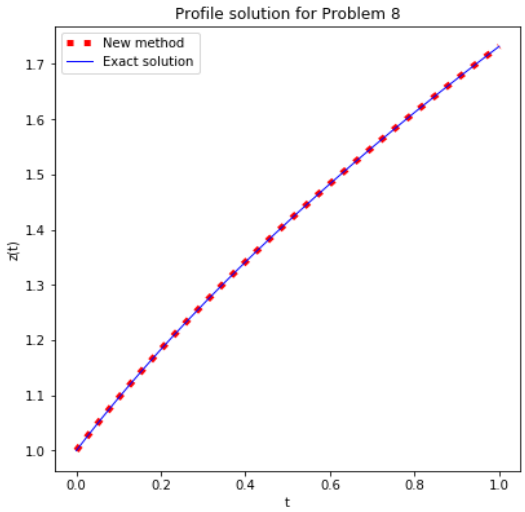

Fig 9. The profile solution for problem 8

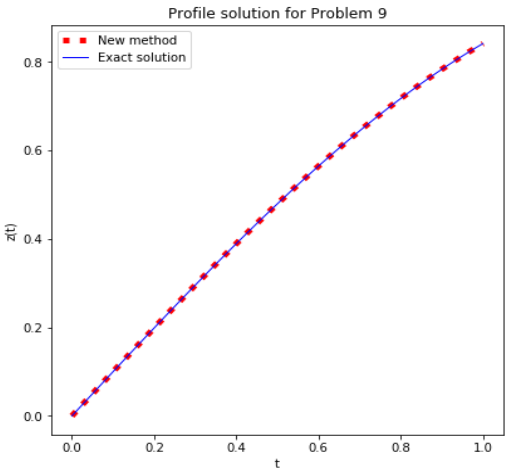

Fig 10. The profile solution for problem 9

Table 8. Comparison of errors for Problem 8

t	Exact Solution	Error in [8] $(h=0.1)$	Error in New method $(h=0.1)$	Error in New method $(h=0.01)$
0.2	1.095445115010332226913940	2.181×10^{-11}	3.890×10^{-13}	3.870×10^{-21}
0.4	1.183215956619923208513466	7.070×10^{-11}	1.812×10^{-12}	2.242×10^{-20}
0.6	1.264911064067351732799557	1.348×10^{-10}	4.673×10^{-12}	5.891×10^{-20}
0.8	1.341640786499873817845504	2.106×10^{-10}	9.038×10^{-12}	1.138×10^{-19}
1.0	1.414213562373095048801689	2.964×10^{-10}	1.490×10^{-11}	1.871×10^{-19}
1.2	1.483239697419132589742279	3.914×10^{-10}	2.225×10^{-11}	2.785×10^{-19}
1.4	1.549193338482966754071706	4.947×10^{-10}	3.106×10^{-11}	3.877×10^{-19}
1.6	1.612451549659709930473323	6.059×10^{-10}	4.128×10^{-11}	5.142×10^{-19}
1.8	1.673320053068151095956344	7.242×10^{-10}	5.290×10^{-11}	6.576×10^{-19}
2.0	1.732050807568877293527446	8.492×10^{-10}	6.586×10^{-11}	8.174×10^{-19}

Table 9. Error comparison for Problem 9

t	Exact Solution	Error in [8] $(h=0.1)$	Error in New method $(h=0.1)$	Error in New method $(h=0.01)$
0.1	0.099833416646828152307	6.661×10^{-16}	7.851×10^{-19}	3.435×10^{-27}
0.2	0.19866933079506121546	3.914×10^{-15}	3.028×10^{-18}	1.919×10^{-26}
0.3	0.29552020666133957511	1.243×10^{-14}	8.180×10^{-18}	6.472×10^{-26}
0.4	0.38941834230865049167	2.887×10^{-14}	1.779×10^{-17}	1.538×10^{-25}
0.5	0.47942553860420300027	5.601×10^{-14}	3.339×10^{-17}	3.017×10^{-25}
0.6	0.56464247339503535720	9.692×10^{-14}	5.667×10^{-17}	5.243×10^{-25}
0.7	0.64421768723769105367	1.547×10^{-13}	8.933×10^{-17}	8.383×10^{-25}
0.8	0.71735609089952276163	2.326×10^{-13}	1.333×10^{-16}	1.261×10^{-24}
0.9	0.78332690962748338846	3.346×10^{-13}	1.904×10^{-16}	1.812×10^{-24}
1.0	0.84147098480789650665	4.644×10^{-13}	2.627×10^{-16}	2.509×10^{-24}

4.4 Fourth-Order Problems

Problem 10: Consider the oscillatory problem arising from ship dynamics:

$$z^{iv}+3z''+z\left(2+\varepsilon\cos\lambda t\right)=0;\quad z(0)=1,\ z'(0)=0,\quad z''(0)=0, z'''(0)=0$$

Where $\lambda=0$ for existence of the exact solution: $z(t)=2\cos t-\cos\left(\sqrt{2}t\right)$

Source: Familua & Omole [16]

Problem 11: Consider the nonlinear problem:

$$z^{iv}=\left(z'\right)^2-zz'''-4t^2+\mathrm{e}^t\left(1-4t+t^2\right);\quad z(0)=1,\ z'(0)=1,\quad z''(0)=3, z'''(0)=1$$

Exact solution: $z(t)=t^2+e^t$. (Source: Familua & Omole [16])

Problem 12: consider the nonlinear sinusoidal problem:

$$z^{iv}=z^2+\sin^2(t)-\cos(t)-1;\quad z(0)=-1,\ z'(0)=0,\quad z''(0)=1, z'''(0)=0$$

Exact solution: $z(t)=-\cos(t)$. Source: Tiamiyu *et al.* [43]

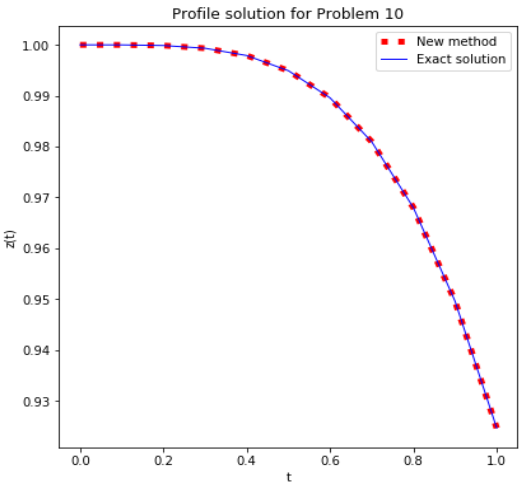

Fig 11. The profile solution for problem 10

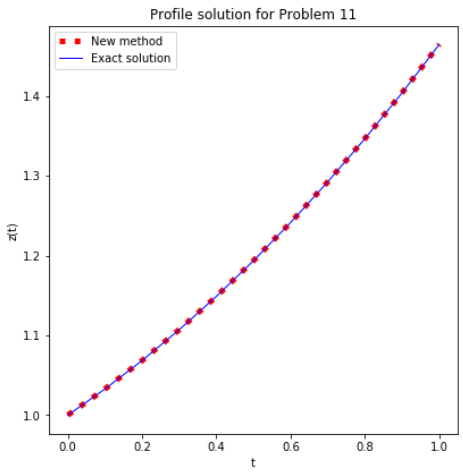

Fig 12. The profile solution for problem 11

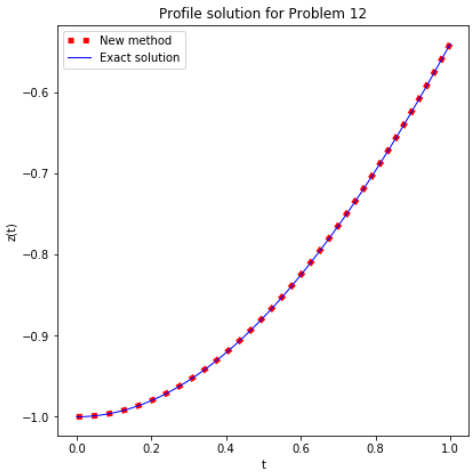

Fig 13. The profile solution for problem 12

Table 10. Error comparison for Problem 10 $(h = 0.003125)$

t	Error in [16] (Block mode)	Error in [16] (P-C mode)	Error in [45]	Error in New Method
0.003125	6.686×10^{-13}	5.686×10^{-10}	1.900×10^{-19}	3.657×10^{-36}
0.006250	1.458×10^{-11}	1.768×10^{-10}	2.300×10^{-19}	1.262×10^{-35}
0.009375	1.083×10^{-10}	5.910×10^{-09}	8.600×10^{-19}	4.233×10^{-35}
0.001250	3.918×10^{-10}	5.768×10^{-09}	1.380×10^{-18}	1.042×10^{-34}
0.015625	1.025×10^{-09}	1.100×10^{-08}	3.530×10^{-18}	2.271×10^{-34}
0.018750	2.217×10^{-09}	6.899×10^{-08}	5.310×10^{-18}	4.359×10^{-34}
0.021875	4.226×10^{-09}	4.636×10^{-08}	8.880×10^{-18}	7.728×10^{-34}
0.025000	7.358×10^{-09}	5.788×10^{-07}	3.922×10^{-17}	1.276×10^{-33}
0.028125	1.197×10^{-08}	2.246×10^{-07}	5.846×10^{-17}	2.001×10^{-33}
0.031250	1.846×10^{-08}	2.846×10^{-07}	8.477×10^{-17}	3.001×10^{-33}

Table 11. Error comparison for Problem 10 using different values of h

t	Exact solution	Error in New method $(h = 0.01)$	Error in New method $(h = 0.1)$
0.1	0.99999167499652860438	3.477×10^{-27}	4.196×10^{-18}
0.2	0.99986719911195714198	5.225×10^{-26}	1.437×10^{-17}
0.3	0.99933105226749824584	2.522×10^{-25}	4.756×10^{-17}
0.4	0.99790057330915505191	8.075×10^{-25}	1.157×10^{-16}
0.5	0.99492052670511528095	1.938×10^{-24}	2.4783×10^{-16}
0.6	0.98958301770794685383	3.941×10^{-24}	4.670×10^{-16}
0.7	0.98095229007588226219	7.137×10^{-24}	8.088×10^{-16}
0.8	0.96799382462962246366	1.186×10^{-23}	1.302×10^{-15}
0.9	0.94960705358355858551	1.844×10^{-23}	1.981×10^{-15}
1.0	0.92466091697090496135	2.718×10^{-23}	2.874×10^{-15}

Table 12. Error comparison for Problem 11

t	Error in [16]	Error in [23]	Error in New method
0.031250	1.149×10^{-12}	1.788×10^{-10}	2.000×10^{-24}
0.062500	1.885×10^{-11}	1.134×10^{-08}	9.000×10^{-24}
0.093750	9.780×10^{-11}	1.196×10^{-07}	2.300×10^{-23}
0.125000	3.166×10^{-10}	6.401×10^{-07}	4.800×10^{-23}
0.156250	7.909×10^{-10}	2.349×10^{-06}	9.000×10^{-23}
0.187500	1.676×10^{-09}	6.573×10^{-06}	1.500×10^{-22}
0.218750	3.169×10^{-09}	1.610×10^{-05}	2.3400×10^{-22}
0.250000	5.512×10^{-09}	3.501×10^{-05}	3.450×10^{-22}
0.281250	8.995×10^{-09}	6.985×10^{-05}	4.870×10^{-22}
0.312500	1.396×10^{-08}	1.245×10^{-04}	6.620×10^{-22}

Table 13. Error comparison for Problem 11 using different values of h

t	Exact solution	Error in New method $(h=0.1)$	Error in New method $(h=0.01)$
0.1	1.11517091807564762481	8.200×10^{-19}	2.430×10^{-27}
0.2	1.26140275816016983392	2.960×10^{-18}	1.891×10^{-26}
0.3	1.43985880757600310398	8.000×10^{-18}	6.317×10^{-26}
0.4	1.65182469764127031782	1.706×10^{-17}	1.486×10^{-25}
0.5	1.89872127070012814685	3.173×10^{-17}	2.881×10^{-25}
0.6	2.18211880039050897488	5.298×10^{-17}	4.947×10^{-25}
0.7	2.50375270747047652162	8.246×10^{-17}	7.811×10^{-25}
0.8	2.86554092849246760458	1.211×10^{-16}	1.160×10^{-24}
0.9	2.86554092849246760458	1.707×10^{-16}	1.645×10^{-24}
1.0	3.71828182845904523536	2.232×10^{-16}	2.251×10^{-24}

Table 14. Error comparison for Problem 12

t	Exact solution	Error in [43] $(h=0.01)$	Error in New method $(h=0.01)$
0.1	– 0.99500416527802576609556198 7804	1.400×10^{-29}	5.580×10^{-29}
0.2	–0.98006657784124163112419651 6748	8.240×10^{-28}	8.465×10^{-28}
0.3	–0.95533648912560601964231022 7568	7.825×10^{-27}	4.236×10^{-27}
0.4	–0.92106099400288508279852673 2052	3.697×10^{-26}	1.331×10^{-26}
0.5	–0.87758256189037271611628158 2604	1.209×10^{-25}	3.234×10^{-26}
0.6	–0.82533561490967829724095249 8955	3.148×10^{-25}	6.675×10^{-26}
0.7	–0.76484218728448842625585999 0192	7.021×10^{-25}	1.230×10^{-25}
0.8	–0.69670670934716542092074998 1642	1.400×10^{-24}	2.088×10^{-25}
0.9	–0.62160996827066445648471615 1407	2.564×10^{-24}	3.323×10^{-25}
1.0	–0.54030230586813971740093660 7443	4.393×10^{-24}	5.031×10^{-25}

Problem 13. And lastly, we also consider the following nearly sinusoidal problem in first-order system of equations. Source: (Akinfenwa *et. al.,* [10])

$$z_1' = -21z_1 + z_2 + 2\sin t, \qquad z_1(0) = 2$$
$$z_2' = 998z_1 - 999z_2 + 999\cos t - 999\sin t, \qquad z_2(0) = 3$$

Exact solution: $z_1(t) = 2e^{-t} + \sin t, \quad z_2(t) = 2e^{-t} + \cos t$

Table 15: Error comparison for Problem 13

	HBSDBDF [10]		New Method		
h	Maximum error	Relative error	Maximum error	Relative error	*ROC*
0.4	8.9924×10^{-7}	3.6279×10^{-7}	1.2119×10^{-11}	1.8909×10^{-10}	-
0.2	5.9042×10^{-9}	2.6294×10^{-9}	1.6035×10^{-14}	2.2570×10^{-13}	9.56
0.1	4.5695×10^{-11}	1.8848×10^{-11}	1.7159×10^{-17}	3.0416×10^{-16}	9.87
0.05	2.9376×10^{-13}	1.2826×10^{-13}	1.7479×10^{-20}	3.1990×10^{-19}	9.94

5. DISCUSSION OF RESULTS

Figures 1–3 respectively display the graphical solution of the SIR model in problem 1, the nonlinear Riccati equation in problem 2 and the nonlinear system of stiff chemical equations in problem 3. In the graphs, we plotted the results generated from our method (red boxes) and the exact solution (blue line). It is easily seen that the new method agrees with the analytical solutions to the problems. Going further, Table 1 shows the comparison of absolute errors for Problem 1. It is shown that the newly derived method has a computational advantage over the methods in ([22], [19] and [4]). Also in Table 2, we show the absolute errors of some methods and ours, for different step sizes in solving problem 2. A comparison of our method with [4] using the same step size (h = 0.1) indicates that our method outperforms the method in [4], and further comparison with the method in [19] (whose step size is smaller) also indicates the superiority of our method over the one in [19]. Table 3 presents a comparative analysis of absolute errors for problem 3 using different step sizes. We solved the nonlinear chemical stiff problem using the step size (h = 0.1) and analysis shows that our method performs better than the methods in [26] (with the same step size), [11] (whose step size is h = 0.01) and [20] (with step size h = 0.0001).

Problems 4–6 are second-order problems considered using the same block method as the first-order problems. Problem 4 is a nonlinear IVP, while Problem 5 is an application problem in the cooling of a body and Problem 6 is also a nonlinear problem. Their behavioural solutions are displayed in figures 4–6 demonstrating agreement in the numerical method that we derived and the exact solutions. To further validate the effectiveness of our method, we compare the absolute errors against those produced by existing techniques. This comparative analysis provides additional evidence supporting the superior performance of our approach. Table 4 shows the comparison of errors for problem 4. It is shown that the new method with h = 0.1 outperforms the method in [7] with the same step size and that in [1] using the same step size of $\left(h = \frac{1}{30}\right)$. Also, Table 5 shows the comparison of errors for Problem 5, which indicates the superiority of our method over those in [24, 33, 38, 41] using the same step size. And in Table 6, we show the numerical solution made by the new method, the absolute error and that in [30], it is shown that the new method is better in terms of accuracy.

The nonlinear third-order initial value problems were the next class of ODEs we looked at in this paper. The graphical solutions of problems 7–9 are displayed in Figures 7– 9. Furthermore, comparative analysis for Problem 7 in Table 7 indicates that our method is found to give better accuracy than the methods in [31], [47], and [9]. Similarly, analyses in Tables 8 and 9 also show that the newly derived method in this paper gives better accuracy than the method derived in [8] for problems 8 and 9, respectively.

Finally, a class of fourth-order ODEs is also solved by our method. Figures 10–12 display the graphical solutions for problems 10–12, respectively. Problem 10 is solved by [16] using block mode predictor–corrector (P–C) of the linear

multistep method. Also, [45] solved this particular problem in the interval [0.003125, 0.03125] with $h = 0.003125$. Table 10 shows the comparison of errors between our method, [16] and [33] as mentioned above. It has been demonstrated that the new block method provides a more accurate approximation for the application problem related to ship dynamics. We further compare our method for variable step sizes over [0, 1] using $h = 0.1$ and $h = 0.01$ and Table 11 shows that the method gives better approximation as h becomes smaller. Similarly, the nonlinear equation in problem 11 is also solved by [16] and [23] and the results in Table 12 assert that the newly derived method is superior in terms of accuracy to the methods in [16, 23]. Table 13 gives the comparison of errors in our method based on different step sizes. And in Table 14, the comparison of errors for problem 12 is made between our method and that in [43]. It is shown that as the numerical iteration progresses, our method gets closer to the exact solution than the method in [43].

Lastly, we incorporated the sinusoidal problem into a system of first-order equations, as addressed in [10]. We showcased the method's accuracy, rate of convergence *(ROC*), and strong stability characteristics. We calculated the maximum error, $\left(\max\left|z\left(t_i\right)-z_i\right|\right)$, relative error $\max_i \dfrac{\left|z\left(t_i\right)-z_i\right|}{\left|1+z\left(t_i\right)\right|}$ and compared with the method in [10], which indicates that the new method has higher accuracy than the method in [10]. Also, the $ROC = \log_2\left(\dfrac{e^{2h}}{e^{h}}\right)$, for different step sizes h, where e^h is the maximum absolute error for each h is calculated as shown in Table 15.

6. CONCLUSION

In this paper, a novel numerical method in the class of linear multistep method is developed for first order and higher order ODEs. The method is derived through the usual interpolation and collocation techniques with power series used as the basis function. Basic numerical properties as established in section 3 show that the method converges and A-stable. The method conveniently solves first, second, third and fourth order IVPs as shown in Tables 1–15. Thirteen numerical problems were considered in all, with nonlinear equation as majority and some application problems. The numerical results generated by our method show its superiority over some existing methods as compared therein. The innovative aspect of this method lies in its capability to effectively solve first, second, third, and fourth order ordinary differential equations (ODEs) without requiring distinct numerical schemes for each type. Therefore, this new approach is proposed as a practical numerical algorithm for tackling first and higher order ODEs. In our upcoming research, we plan to expand this method to address partial differential equations.

ORCID

Jamiu Garba https://orcid.org/0000-0002-2850-7730
Khadeejah Audu James https://orcid.org/0000-0002-6986-3491
Umaru Mohammed https://orcid.org/0000-0002-0777-6536
Tiamiyu Abd'gafar Tunde https://orcid.org/0000-0003-1641-7196

REFERENCES

[1] R. Abdelrahim, and Z. Omar, Direct solution of second-order ordinary differential equations using a single-step hybrid block method of order five, *Math. Comput. App*., (2016); doi:10.3390/mca21020012.
[2] Y. A. Abdullahi, Z. Omar, and J. O. Kuboye, Derivation of block predictor–block corrector method for direct solution of third order ordinary differential equations, *Global J. Pure App. Math.*, 11(1)(2016), 343-350.
[3] C. E. Abhulimen, and L. A. Ukpebor, A new class of second derivative methods for numerical integration of stiff initial value problems, *J. Nig. Math. Soc*., 38(2)(2019), 259-270.
[4] O. E. Abolarin, G. B. Ogunware, and L. S. Akinola, An efficient seven-step block method for numerical solution of SIR and growth model, *FUOYE J. Engine. Tech*., 5(1)(2020), 31–35.
[5] K. M. Abualnaja, A block procedure with linear multistep methods using Legendre polynomial for solving ODEs, *App. Math. Sci. Res*., 6(2015), 717-723.

[6] E. O. Adeyefa, Orthogonal-based hybrid block method for solving general second order initial value problems, *Italian journal of pure and applied mathematics.* 37(2017), 659-672.
[7] O. Adeyeye and Z. Omar, Maximal order block method for the solution of second order ordinary differential equations, *IAENG Int. J. App. Math.,* 46(4)(2016), 46-4-03.
[8] O. Adeyeye and Z. Omar, Solving third order ordinary differential equations using one-step block method with four equidistant generalized hybrid points, *IAENG Int. J. App. Math.*, 49(2)(2019), 42 – 56.
[9] L. O. Adoghe, B. G. Ogunware and E. O. Omole, A family of symmetric implicit higher order methods for the solution of third order initial value problems in ordinary differential equations, *Theoretical Math. App.,* 6(3)(2016), 67-84.
[10] O. A. Akinfenwa, R. I. Abdulganiy, B. I. Akinnukawe and S. A. Okunuga, Seventh order hybrid block method for solution of first order stiff systems of initial value problems, *J. Egypt. Math. Soc*., 28(34)(2020), 28-34. https://doi.org/10.1186/s42787-020-00095-3.
[11] O. A. Akinfenwa, R. I. Abdulganiy and S. A. Okunuga, A Simpsons 3/8 type block method for stiff systems of ordinary differential equation, Journal of the Nigerian Mathematical Society. 36(3)(2017), 503-514.
[12] M. O. Alabi, M. T. Raji and M. S. Olaleye, Off-grid collocation four step initial value solver for second order ordinary differential equations*, International Journal of Mathematics and Statistics Studies*, 12 (4)(2024), 1-15
[13] M. Al-Kandari, Enhanced criteria for detecting oscillations in neutral delay Emden Fowler differential equations, *Kuwait Journal of Science,* 50(4)(2023), 443 – 447.
[14] A. M. Badmus, Y. A. Yahaya, and Y. C. Pam, Adams type hybrid block methods associated with Chebyshev polynomial for the solution ordinary differential equations, *British J. Math. Comp. Sci.,* 6(1)(2015), 464-474.
[15] H. G. Debela and M. J. Kabeto, Numerical solution of fourth-order ordinary differential equations using fifth-order Runge–Kutta method, *Asian J. Sci. Tech.*, 8(2)(2017), 4332-4339.
[16] A. B. Familua and E. O. Omole, Five points mono hybrid linear multistep method for solving nth order ordinary differential equations using power series function, *Asian Res. J. Math*, 3(1)(2017), 1-17.
[17] J. Garba J. and U. Mohammed, Derivation of a new one-step numerical integrator with three intra-step points for solving first order ordinary differential equations. *Nig. J. Math. App*., 30(2020), 155 – 172.
[18] M. A. Islam, Accurate solutions of initial value problems for ordinary differential equations with the fourth order Runge-Kutta method, *J. Math. Res.,* 7(3)(2015), 41 – 53.
[19] S. H. B. Kashkaria and M. I. Syamb, Optimization of one step block method with three hybrid points for solving first-order ordinary differential equations, *Results in Physics*, 12(1)(2019), 592–596.
[20] M. M. Khalsaraei, A. Shokri, and M. Molayi, The new class of multistep multi-derivative hybrid methods for the numerical solution of chemical stiff systems of first order IVPs, *Journal of Mathematical Chemistry*, (2020). https://doi.org/10.1007/s10910-020-01160-z
[21] K. A. Koroche, Numerical solution of first order ordinary differential equation by using Runge-Kutta method, *Int. J. sys. Sci. App. Math.*, 6(1)(2021), 1-8.
[22] J. O. Kuboye and E. O. Adeyefa, New developed numerical formula for solution of first and higher order ordinary differential equations, *J. Interdisciplinary Math.* (2021). https://doi.org/10.1080/09720502.2021.1925453
[23] J. O. Kuboye, O. R. Elusakin and O. F. Quadri, Numerical algorithm for direct solution of fourth-order ordinary differential equations, *J. Nig. Soc. Phys. Sci*., 2(2020), 218-227.
[24] J. A. Kwanamu, Y. Skwame, and J. Sabo, Block hybrid method for solving higher order ordinary differential equation using power series on implicit one-step second derivative, *FUW Trends in Sci. Tech. J.*, 6(2)(2021), 576 – 582.
[25] A. I. Ma'ali, U. Mohammed, K. J. Audu, A. Yusuf and A. D. Abubakar, Extended block hybrid backward differentiation formula for second order fuzzy differential equations using Legendre polynomial as basis function, *J. Sci., Tech., Math. Edu*, 16(1)(2020), 100-111.
[26] U. Mohammed, J. Garba and M. E. Semenov, One-step second derivative block intra-step method for stiff system of ordinary differential equations *J. Nig. Math*. *Soc.,* 40(1)(2021), 47-57.
[27] U. Mohammed, O. Oyelami and M. E. Semenov, An orthogonal-based self-starting numerical integrator for second order initial and boundary value problem of ODEs, *J. Phy. Conf. Ser.* 1145(2019), 012040. Doi:10.1088/1742-6596/1145/1/012040.
[28] A. Ndanusa and F. U. Tafida, Predictor-corrector methods of high order for numerical integration of initial value problems, *Int. J. Sci. and Inno. Math. Res*, 4(2)(2016), 47-55.

[29] G. C. Nwachukwu and T. Okor Second derivative generalized backward differentiation formulae for solving stiff problems, *IAENG Int. J. App. Math*., 48(1)(2018).
[30] O. M. Ogunlaran and M. A. Kehinde, A new block integrator for second order initial value problems, *Neuroquantology*, 20(17)(2022), 1976 – 1980.
[31] B. G. Ogunware, E. O. Omole and O. O. Olanegan, Hybrid and non-hybrid implicit schemes for solving third order odes using block method as predictors, *J. Math. Theory and Modelling,* 5(3)(2015), 10-25.
[32] B. T. Olabode and A. L. Momoh, Continuous hybrid multistep methods with Legendre basis function for direct treatment of second order stiff ODEs, *Amer. J. Compt. App. Math.*, 6(2)(2016), 38-49.
[33] O. O. Olanegan, B. G. Ogunware and C. O. Alakofa, Implicit hybrid points approach for solving general second order ordinary differential equations with initial values, *J. Advan. in Math. Comp. Sci.*, 27(3)(2018), 1-14.
[34] O. O. Olanegan, B. G. Ogunware, E. O. Omole, T. S. Oyinloye and B. T. Enoch, Some variable hybrids linear multistep methods for solving first order ordinary differential equations using Taylor's series, *IOSR J. Math*. 11(2015), 08-13.
[35] Z. Omar and R. Abdelrahim, New uniform order single step hybrid block method for solving second order ordinary differential equations, *Int. J. App. Engineering Res.,* 11(4)(2016), 2402-2406.
[36] Z. Omar and O. Adeyeye, Numerical solution of first order initial value problems using a self-starting implicit two-step obrenchkoff-type block method, *J. Math. Stat.* 212(2)(2016), 127-134.
[37] Z. Omar and J. O. Kuboye, Computation of an accurate implicit block method for solving third order ordinary differential equations directly, *Global J. Pure App. Math.*, 11(1)(2015), 177-186.
[38] E. O. Omole and B. G. Ogunware, 3-point single hybrid block method (3PSHBM) for direct solution of general second order initial value problem of ordinary differential equations, *J. Sci. Res. Reports*, 20(3)(2018), 1-11.
[39] E. O. Omole and L. A. Ukpebor, "A step by step guide on derivation and analysis of a new numerical method for solving fourth-order ordinary differential equations, *J. Math. Letter*, 6(2)(2020), 13-31.
[40] G. F. Simmons*, Differential equations with applications and historical notes*, CRC Press, Taylor & Francis Group Boca Raton, London New York, 2017.
[41] Y. Skwame, J. Z. Donald, J. Sabo, T. Y. Kyagya and A. A. Bambur, The numerical applications of implicit second derivative on second order initial value problems of ordinary differential equations, *Dutse J. Pure and Appl. Sci.,* 6(4)(2020), 1-14.
[42] J. Sunday, A. A. James, M. R. Odekunle and A. O. Adesanya, Solution to free undamped and free damped motion problems in mass-spring systems, *Amer. J. Comput. App. Math.,* 6(2)(2016), 89-91.
[43] A. T. Tiamiyu, A. T. Cole and K. J. Audu, A backward differentiation formula for fourth–order initial or boundary values problems using collocation method, *Iranian J. Opt*, 13(2)(2021).
[44] P. Tumba, J. Sabo and M. Hamadina, Uniformly order eight implicit second derivative method for solving second-order stiff ordinary differential equations *Acad. Res. Pub. Group,* 4(1)(2016), 43-48.
[45] L. A. Ukpebor, E. O. Omole and L. O. Adoghe, An order four numerical scheme for fourth-order initial value problems using Lucas polynomial with application in ship dynamics, *International Journal of Mathematical Research.* 9(1)(2020), 28-41.
[46] L. A. Ukpebor, Implicit four-step approach with application to non-linear third order ordinary differential equations, *FUDMA J. Sci. (FJS),* 5(4)(2021), 406 – 412.
[47] N. S. Yakusak and A. O. Owolanke, A class of linear multi-step method for direct solution of second order initial value problems in ordinary differential equations by collocation method, *J. Advances in Math. Comp. Sci.*, 26(1)(2018), 1-11.
[48] A. T. Tiamiyu, K. I. Falade, Q. O. Rauf and S. A. Akande, A numerical technique for direct solution of special fourth-order ordinary differential equation via hybrid linear multistep method, *Cankaya University of Journal of Science and Engineering,* 18(1), (2021).
[49] K. I. Falade, Numerical solution of higher-order singular initial value problems (SIVP) by exponentially fitted collocation approximate method, *American International Journal of Research in Science, Technology, Engineering & Mathematics,* 21(1)(2018).

MAKÜ BBBF
BUCAK BİLGİSAYAR VE BİLİŞİM FAKÜLTESİ

Scientific Journal of Mehmet Akif Ersoy University, 2025, 8(1), Page:35-46

Research Article

Segmentation of Histopathological Images with LinkNet Model Supported by Vgg16 Backbone

Furkan ATLAN[1*]

[1]*Department of Management Information Systems, Burdur Mehmet Akif Ersoy University, Burdur, Türkiye*

ABSTRACT

Nuclei segmentation in histopathological images is crucial for the processing and analysis of medical images. Manual segmentation of nuclei images is challenging due to subjective errors by experts and image noise. Before the use of artificial intelligence in medical image analysis, segmentation tasks were performed with common classical methods such as thresholding and watershed. The development of deep learning has led to the emergence of models specifically designed for segmentation tasks. In this study, LinkNet model supported with Vgg16 backbone is proposed for segmenting histopathological images in CryoNuSeg dataset created for nucleus segmentation. After a small number of images are multiplied with data augmentation, feature maps are generated using the Vgg16 model integrated into the encoder of the LinkNet architecture. The results obtained in this study, with F1 Score, Intersection over Union (IoU), and Aggregated Jaccard Index (AJI) values of 0.8447, 0.7312, and 0.7312 respectively, demonstrate superior performance compared to recent studies utilizing the same dataset.

Keywords: Medical image segmentation, CryoNuSeg, Backbone, Vgg16, LinkNet

1. INTRODUCTION

Histopathology is the microscopic examination of tissue samples to detect structural abnormalities and pathological changes, serving as a cornerstone of disease diagnosis, particularly in oncology. In clinical practice, pathologists analyze stained tissue sections—typically prepared with Hematoxylin and Eosin (H&E)—to identify disruptions in tissue architecture, variations in nuclear morphology, and the presence of inflammatory or malignant cells. This process, although vital for determining disease stage and guiding treatment, is labor-intensive and subject to inter-observer variability. Recent advances in digital pathology and computational image analysis have transformed histopathology into a data-rich domain, allowing whole-slide images (WSIs) to be assessed algorithmically. In this context, deep learning-based approaches, often referred to as computational pathology (CPATH), have shown great potential in automating diagnostic workflows, improving reproducibility, and addressing the global shortage of expert pathologists [1].

Digital pathology enables the extraction of information from stained and digitized tissue samples obtained from patients. This information is shared and managed among experts, thereby providing benefits such as allowing remote specialists to interpret these images or utilizing samples for scientific research [2]. Additionally, digitized images can also be employed for computer-aided quantitative image analysis [3]. Performing image analysis in digital pathology supported by artificial intelligence raises expectations of significantly improving clinical applications [4].

Examination of tissue sections stained with H&E provides valuable insights into cells and their functions [5]. This is because H&E-stained tissue images play a crucial role in diagnosing various cancer types, including breast, prostate, and liver cancers. Factors such as shape, type, morphology, density, and quantity of nuclei are fundamental components in the evaluation of H&E-stained tissue images [6].

Nuclei segmentation is biologically crucial, as information extracted from tissue images enables observations regarding cell cycles and mutations in cancer-related proteins, thereby facilitating the advancement of research. However, challenges exist in nuclei segmentation due to factors such as noise in images, overlapping of cells, and complications arising during manual preparation processes [7-8].

***Corresponding Author Email:** fatlan@mehmetakif.edu.tr
Submitted: 16.03.2025 **Revision Requested:** 07.04.2025 **Last Revision Received:** 02.05.2025
Accepted: 06.05.2025 **Published Online:** 22.05.2025

Burdur Mehmet Akif Ersoy University Press

Cite this article as: Atlan, F. (2025). Segmentation of Histopathological Images with LinkNet Model Supported by Vgg16 Backbone. Scientific Journal of Mehmet Akif Ersoy University, 8(1): 35-46.
DOI: https://doi.org/10.70030/sjmakeu.1658832

https://dergipark.org.tr/sjmakeu

Manual segmentation is considerably costly due to factors such as the necessity for clinical expertise, the time-consuming nature of the task, and susceptibility to human error. Consequently, automated nuclei segmentation methods have been developed to reduce workload and establish models with the highest possible accuracy.

The initial models developed for automated nuclei segmentation mainly consisted of watershed segmentation, morphological operations, and thresholding methods. However, these classical approaches had several disadvantages, including the requirement for manual parameter tuning, limited generalizability across multiple organs and tissue types, and reduced performance in the presence of noise [9]. Following the availability of large annotated datasets and the success of deep learning models such as Convolutional Neural Networks (CNNs), the use of classical methods in automated nuclei segmentation tasks has progressively declined. Instead, models employing Fully Convolutional Networks (FCNs) have become more favorable [10]. FCN architectures typically consist of encoder-decoder blocks, which often exhibit symmetrical structures. Specifically, within medical image segmentation, the most popular FCN-based architecture is the U-Net model [11]. Upon recognition of the significance of nuclei segmentation, state-of-the-art models and their modified versions have emerged in this field. Examples of such advanced models include U-Net++ [12] and 3D U-Net [13].

Accurate decoding of features learned by the encoder at higher resolutions in the decoder is essential for automated nuclei segmentation. For this reason, proven deep learning architectures can be used as "backbones" in encoder sections instead of relying solely on the layers of the original model [14-16].

The objective of this study is to automate nuclei segmentation by utilizing the Vgg16 model [17] as the backbone structure of LinkNet [18], a Fully Convolutional Network (FCN) architecture specifically developed for segmentation tasks. The choice of the LinkNet model instead of the widely adopted U-Net architecture, which is extensively used in medical image segmentation, stems from the hypothesis that employing state-of-the-art architectures as backbones within segmentation models—rather than relying on their original internal layers—may yield superior performance. Here, the primary focus is placed upon the potential improvement in segmentation accuracy achievable by integrating more effective models within the encoder section. In other words, the emphasis is on evaluating the influence of the backbone model rather than the segmentation architecture itself. Consequently, the LinkNet model was preferred over the conventional U-Net.

This study is organized as follows: Section 2 presents the relevant literature. The dataset is described in Section 3. Methodology is provided in Section 4. System details are discussed in Section 5. Section 6 presents the results of the study. Section 7 concludes the paper.

2. LITERATURE REVIEW

Deep learning-based models play a crucial role in nuclei segmentation. Table 1 provides an overview of the literature on nuclei segmentation.

Table 1. Literature review on nuclei segmentation in histopathological images

Author(s)	Methodology	Dataset	Metrics			
			DCS	F1-Score	IoU	AJI
[19]	Enhanced U-Net with deeper encoder layers	Data Science Bowl	-	-	0.567	-
[20]	Residual + Bottleneck + Attention-based Decoder	Kumar [21]	-	0.811	0.685	-
[22]	Multi-task, 3-branch U-Net with region enhancement	ConSep	0.854	-	-	0.561
		CPM17	0.884	-	-	0.712
[23]	U-Net with ResNet encoder	PanNuke	-	0.841	-	0.740
[24]	Vision Transformer-based model with multi-scale encoding	GCNS	0.725	-	-	-
[25]	Dense conv + normalization + local & global feature fusion	PanNuke	0.865	-	0.844	-
		ConSep	0.844	-	0.823	-
[26]	U-Net with VGG16 encoder	MoNuSeg	-	0.845	0.693	-

In this study, a LinkNet segmentation model supported by a VGG16 backbone is proposed for nuclear segmentation, in which the VGG16 model serves as the encoder. The primary distinction of the proposed method from other conventional encoder-based models lies in the preprocessing of images before being fed into the model. Here, the proposed approach is not only a model but a comprehensive framework that incorporates image preprocessing prior to segmentation. The effectiveness of the proposed method was evaluated on histopathological images from the CryoNuSeg dataset, yielding successful results.

3. DATASET

3.1. CryoNuSeg Dataset

The CryoNuSeg dataset [27] consists of 30 histopathological images stained with H&E, representing ten different organs, each with a resolution of 512×512 pixels. Figure 1 shows nuclei images of different organs in the CryoNuSeg dataset.

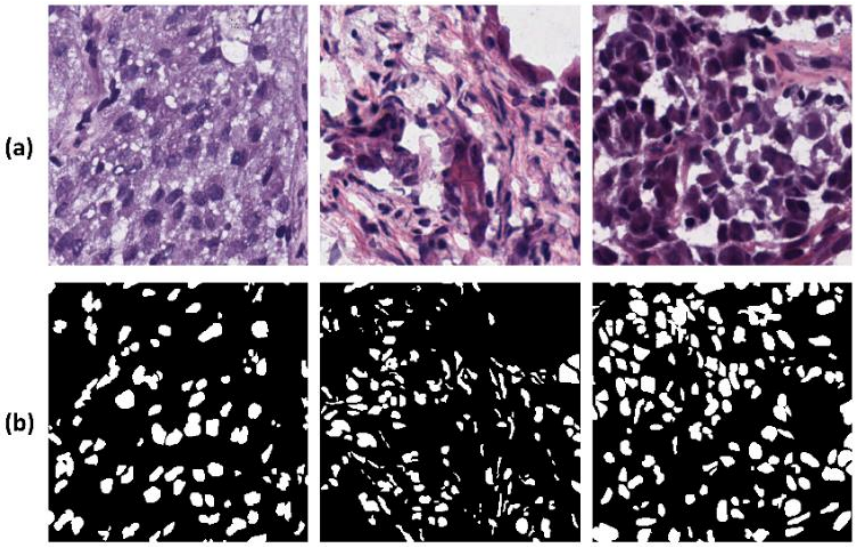

Fig 1. Examples from the CryoNuSeg dataset (a) Organ images, (b) Segmentation masks (From left to right: adrenal gland, pancreas, and skin images.)

Table 2 presents the descriptive statistics of the dataset, while Figure 1 illustrates example images from the CryoNuSeg dataset.

Table 2. Descriptive Statistics of the CryoNuSeg Dataset

Organ	Number of Images	Number of Nuclei
Adrenal gland	3	344
Thyroid gland	3	464
Pleura (Lung membrane)	3	515
Lymph node	3	1308
Testis	3	793
Skin	3	436
Thymus	3	1646
Pancreas	3	548
Mediastinum	3	1349
Larynx	3	641
Total	**30**	**8044**

3.2. CPM-17 Dataset

The CPM-17 [28] dataset, provided as part of the MICCAI 2017 Digital Pathology Challenge, comprises a total of 64 histopathological images—32 for training and 32 for testing—each with a resolution of 500×500 pixels, including 7,570

annotated nuclei. In this study, the proposed model, initially trained on the CryoNuSeg dataset, was evaluated on the 32 test images from the CPM-17 dataset.

4. METHODOLOGY

4.1 Data Preprocessing

Data augmentation was applied to the 30 images in the dataset. Data augmentation is a common regularization technique in deep learning [29]. When the number of training images is limited, increasing the dataset size enhances data diversity and improves model learning, making it a preferred preprocessing approach [30].
Flipping and cropping data augmentation techniques were applied to 25 training images. In the flipping technique, images were mirrored along the x-axis, y-axis, and both axes. In the cropping technique, images were randomly cropped within a range of pixel values. As a result of applying these two augmentation techniques, resulting in a total of 1,000 augmented training images.

4.2 LinkNet

LinkNet is a semantic segmentation model consisting of a total of eight blocks, including four encoder and four decoder blocks. On the encoder side, a convolution operation is first performed using a 7×7 matrix, followed by max pooling over a 3×3 region. Batch normalization [31] is applied between each convolutional layer, followed by the nonlinear ReLU activation function [32]. The convolution process in the encoder begins with 64 feature maps, and through downsampling, the number of feature maps doubles at each stage, reaching 512 in the final encoder block. In the decoder part, the 512 feature maps undergo upsampling to restore the resolution, and at each stage, the number of feature maps is reduced by half. By the time the process reaches the first block, the segmentation is completed with 64 feature maps, matching the initial configuration.

4.3 Vgg16

VGG16 is a deep learning model consisting of five convolutional blocks, each containing a max pooling layer. The first two convolutional blocks include two convolutional layers followed by a max pooling layer, while the remaining three convolutional blocks contain three convolutional layers and a max pooling layer [33]. The first convolutional block utilizes 64 filters, and the number of filters doubles at each subsequent block until the fifth block, where 512 filters are maintained, as in the previous block. Finally, the model is completed with three fully connected layers, followed by the softmax activation function [34].

4.4 Metrics

The performance of the proposed model was evaluated using the Dice Coefficient Score (DCS), F1-score, Intersection over Union (IoU), and Aggregated Jaccard Index (AJI) metrics. These metrics are widely utilized in segmentation studies to assess the effectiveness of the segmentation process by quantifying the similarity between predicted and ground truth masks.
The mathematical formulations of the DCS, F1-score, IoU, and AJI metrics are presented in Equations 1–4, respectively.

$$\mathrm{DCS} = \frac{2*|t \cap \hat{t}|}{|t+\hat{t}|} \tag{1}$$

In Equation 1, t represents the ground truth segmented nucleus, while t̂ denotes the nucleus obtained as a result of the prediction.

$$F1\ \mathrm{Score} = 2 * \frac{\mathrm{Recall}*\mathrm{Precision}}{\mathrm{Recall}+\mathrm{Precision}} \tag{2}$$

$$\mathrm{IoU} = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}+\mathrm{FN}} \tag{3}$$

In Equation 3, TP refers to True Positive, FP represents False Positive, and FN denotes False Negative values.

$$AJI = \frac{\sum_{i=1}^{nQ} |G_i \cap S(G_i)|}{\sum_{i=1}^{nQ} |G_i \cup S(G_i)| + \sum_{k \in K} |S_k|} \tag{4}$$

In Equation 4, $n\varrho$ represents the number of segmented nuclei, G_i denotes the set of segmented nuclei, G_i refers to the set of matched fragmented nuclei, and S_k corresponds to the set of fragmented nuclei that do not match any segmented nuclei [35,8]. The AJI metric is a crucial measure for evaluating segmentation performance, as it is more robust in penalizing incorrect segmentation predictions compared to other metrics, making it a reliable indicator of segmentation accuracy [8].

5. DETAILS OF THE SYSTEM

In the proposed LinkNet model supported by a VGG16 backbone for nuclear segmentation, the convolutional blocks in the encoder part of the LinkNet segmentation model have been replaced with the blocks from the VGG16 model. Figure 2 presents the architecture of the developed model in this study.

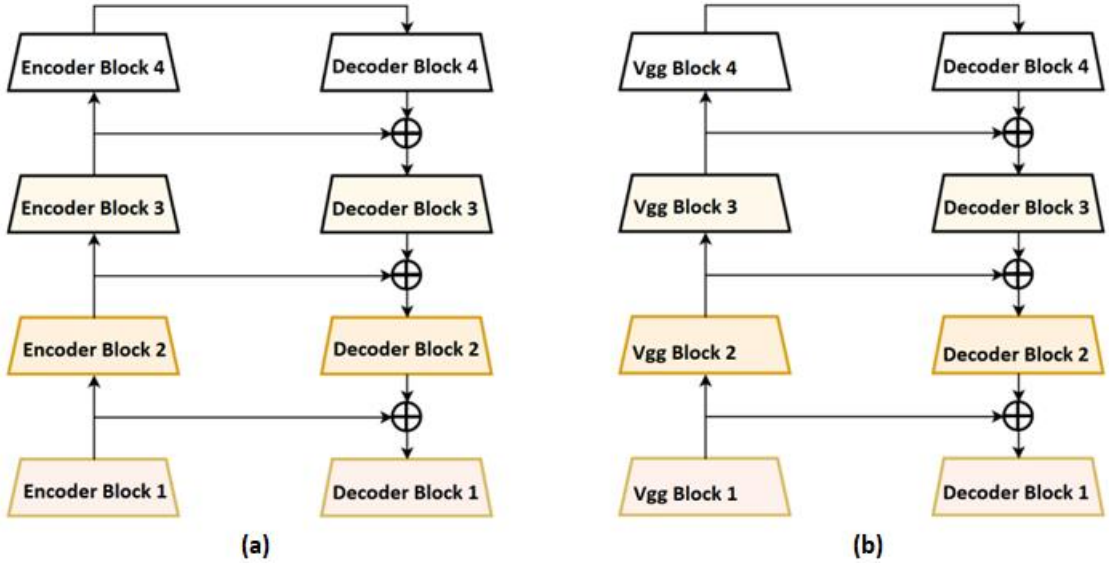

Fig 2. LinkNet Model Architecture Supported by a VGG16 Backbone (a) Original LinkNet Model, (b) LinkNet Model Supported by a VGG16 Backbone.

The proposed segmentation model is an enhanced version of the original LinkNet architecture, incorporating a VGG16 backbone in place of the standard LinkNet encoder. The original LinkNet model consists of four encoder blocks and four decoder blocks, where the encoder progressively extracts hierarchical features from the input image through convolutional operations, batch normalization, and activation functions. In this process, the spatial resolution of the feature maps is reduced while the number of feature channels increases. The decoder section, in turn, restores the resolution through upsampling, utilizing skip connections that retain fine-grained spatial details to improve segmentation accuracy.

In the modified model, the encoder component of LinkNet is replaced with the convolutional blocks of VGG16 to enhance feature extraction capabilities. Specifically, the first encoder block of LinkNet is substituted with VGG16 Block 1, while VGG16 Block 2 replaces the second encoder block. The third and fourth encoder blocks of LinkNet are replaced with the combined structure of VGG16 Blocks 3 and 4, and the final encoder block is substituted with VGG16 Block 5. The integration of VGG16 as the feature extractor allows for deeper and more refined hierarchical feature representation, enabling the model to better capture both low-level textures and high-level semantic structures within an image. The decoder section of the model retains the original LinkNet structure, where each decoder block progressively upsamples the extracted features and refines the segmentation mask, leveraging skip connections for improved spatial detail retention.

The main advantage of using a VGG16-backed LinkNet model is its improved capacity to capture complex structural features, which is particularly beneficial for nuclear segmentation tasks. The deep convolutional layers of VGG16 facilitate robust feature representation, contributing to improved segmentation accuracy. Additionally, the model benefits from pre-trained VGG16 weights, which improve generalization when applied to diverse datasets. The

presence of skip connections ensures that the detailed features extracted in the earlier layers of VGG16 are effectively incorporated during the upsampling process, leading to more precise boundary delineation in segmented images.
By integrating VGG16 into the LinkNet framework, the proposed model achieves a balance between computational efficiency and segmentation accuracy. The hierarchical structure of VGG16 provides strong feature extraction capabilities, while the efficient upsampling mechanism of LinkNet maintains the model's ability to generate high-resolution segmentation masks. This combination renders the proposed approach well-suited to complex segmentation tasks, such as nuclei segmentation in histopathological images, where fine structural details play a crucial role in analysis and interpretation.

6. RESULTS AND DISCUSSION

6.1. Parameter Settings

Experimental studies were performed with NVIDIA RTX 4070 8GB GPU resource. Tensorflow version 2.15.0 was preferred in the study and was run using 0.0001 learning rate and Adam optimizer for 100 epochs on Jupyter Notebook.

6.2 Experimental Results

Table 3 presents the literature results obtained using the CryoNuSeg dataset, along with the results achieved by the proposed model in this study. The best-performing results for each performance metric are highlighted in bold.
Additionally, Figure 3 provides a visual representation of test images processed using the proposed model, demonstrating its segmentation performance.

Upon examining the comparative results in Table 3, it is evident that studies conducted using the CryoNuSeg dataset have employed state-of-the-art deep learning models in the encoder components of segmentation architectures [36-37].

The key distinction of the proposed method from similar and previous studies lies in its focus on the encoder model rather than the segmentation model itself it focuses on the encoder model instead of focusing on the segmentation model. This is because the results in Table 3 show that studies using the U-Net model developed for medical image segmentation are predominant. In this study, after a simple data augmentation preprocessing, the LinkNet model supported by the Vgg16 backbone achieved the highest AJI score in the literature. Achieving the highest score with the AJI metric, which is quite ruthless in penalizing incorrectly matched or incorrectly predicted kernels in segmentation tasks, demonstrates the effectiveness of the proposed model in nuclei segmentation.

Table 3. Results of Studies with CryoNuSeg Dataset

Model	DS	F1 Score	IoU	AJI
Two-Stage U-Net [36]	0.803	-	-	0.525
Nested U-Net backed by EfficientNet backbone [35]	0.929	-	-	0.604
CNN encoder-based U-Net [38]	0.815	-	-	0.541
EfficientNetv2 and the Attention Module [37]	**0.941**	-	-	0.605
U-Net based on Recurrent Neural Networks [39]	0.820	-	0.697	-
DONSeg [40]	0.672	-	-	0.441
UN-SAM [41]	0.804	0.807	0.652	-
MDLA-Unet [42]	0.807	-	-	-
Light-weight multi-scale attention [43]	0.810	-	0.685	-
LinkNet supported by Vgg16 backbone	0.840	**0.845**	**0.731**	**0.731**

The LinkNet model supported by the proposed VGG16 backbone for nuclei segmentation demonstrated remarkable performance, surpassing the results of multiple previous studies, including those conducted by the creators of the CryoNuSeg dataset [29, 38]. This superior performance can be attributed to several critical factors that merit further discussion.

First and foremost, the integration of a robust, multi-parameter deep learning architecture like VGG16 in the encoder component of the LinkNet model proved to be highly effective. VGG16's hierarchical structure, with its ability to capture both low-level features in initial blocks and high-level semantic information in deeper layers, enabled comprehensive feature extraction from histopathological images. This dual capability is particularly crucial for nuclei segmentation, where both textural details and structural context significantly influence segmentation accuracy. The experimental results, as evidenced by the highest AJI score (0.731) among comparable studies, validate the efficacy of this architectural decision.

A key methodological contribution of this work was the emphasis on enhancing the encoder component rather than focusing exclusively on modifying the segmentation architecture itself. Most previous studies utilizing the CryoNuSeg dataset have primarily concentrated on adapting the U-Net architecture, which has been the de facto standard for medical image segmentation. Our approach diverged from this trend by implementing LinkNet with a VGG16 backbone, demonstrating that alternative segmentation architectures can achieve superior results when paired with appropriate feature extractors. This finding suggests that the choice of encoder may have a more substantial impact on segmentation performance than the underlying segmentation framework, especially for complex tasks like nuclei segmentation in histopathological images.

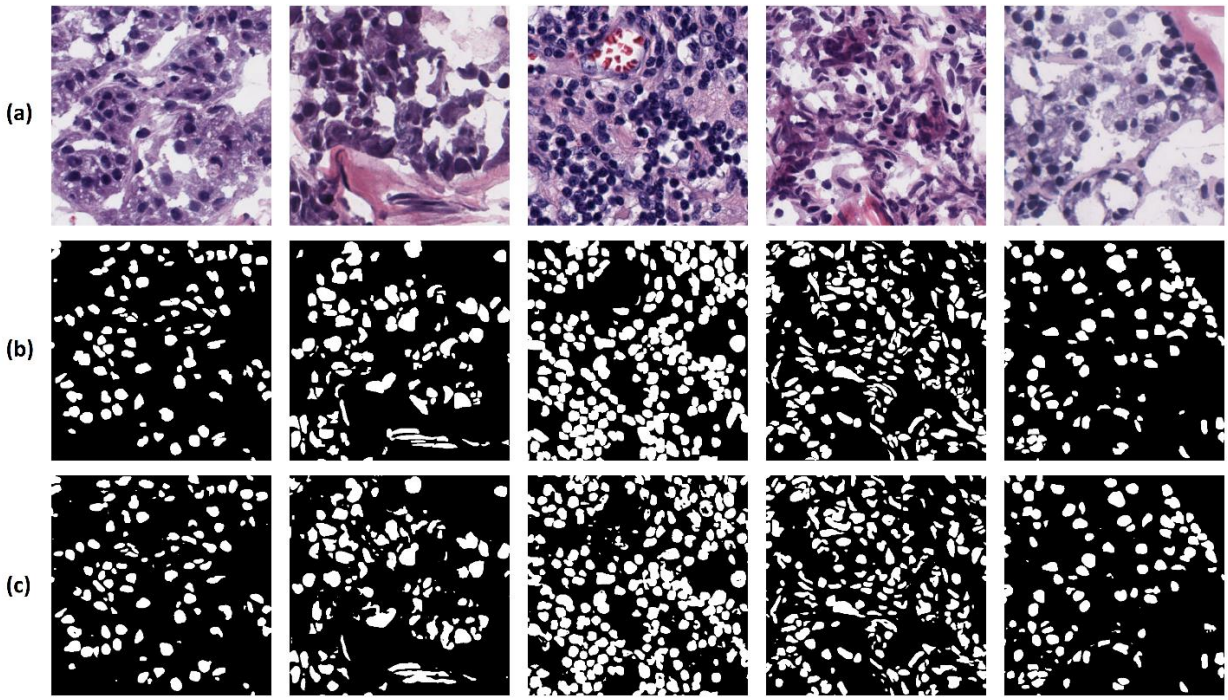

Fig 3. Running the CryoNuSeg dataset with the proposed model (a) Original histopathological input images, b) Ground truth segmentation masks annotated by experts, c) Predicted segmentation masks generated by the proposed LinkNet model supported by VGG16 backbone

The data preprocessing stage, particularly the application of data augmentation techniques, played a vital role in the model's success. Despite the relatively small size of the CryoNuSeg dataset (only 30 original images), our augmentation strategy expanded the training set to 1,000 images, significantly enhancing data diversity. This approach mitigated potential overfitting issues that often plague deep learning models trained on limited datasets. The combination of flipping and cropping techniques introduced variations in orientation and scale, enabling the model to learn more robust and generalizable features. This comprehensive preprocessing approach contrasts with some previous studies that utilized more limited augmentation techniques or none at all.

It is particularly noteworthy that our model achieved the highest score on the AJI, a metric known for its stringent evaluation of segmentation accuracy by severely penalizing incorrectly matched or predicted nuclei. AJI provides a more rigorous assessment of instance-level segmentation performance compared to pixel-level metrics like Dice Score

or IoU. Our model's superior performance on this challenging metric underscores its effectiveness in correctly identifying individual nuclei boundaries, a critical requirement for practical applications in histopathological analysis and clinical decision support.

The symmetric two-stage architecture of segmentation models, comprising encoder and decoder components, allows for an interesting analysis of feature learning dynamics. In our approach, the VGG16 backbone in the encoder fulfilled a dual function: extracting low-level textural features in the initial blocks and high-level contextual features in the deeper blocks. This hierarchical feature extraction, when coupled with LinkNet's efficient decoder pathway, facilitated precise boundary delineation of nuclei in histopathological images. The skip connections between encoder and decoder blocks further enhanced the model's ability to preserve fine spatial details while incorporating contextual information, resulting in more accurate segmentation masks.

Despite these achievements, there remain several limitations and opportunities for future research. First, although our model showed superior performance on the CryoNuSeg dataset, its generalizability to other histopathological datasets with different staining protocols or tissue types requires further investigation. Second, the computational complexity of the VGG16 backbone, while justified by its performance benefits, may pose challenges for deployment in resource-constrained environments.

Future research directions could explore several promising avenues. The evaluation of alternative combinations of segmentation frameworks and backbone architectures could yield further improvements or identify optimal pairings for specific histopathological applications. More recent architectures such as EfficientNet, Vision Transformers, or hybrid models could potentially enhance feature extraction while reducing computational demands. Additionally, hyperparameter optimization of the backbone models, particularly learning rates and regularization strategies, could further refine segmentation performance. Advanced data augmentation techniques, such as style transfer or adversarial training, might further improve model robustness to variations in staining and imaging conditions commonly encountered in clinical settings.

Another promising direction involves the incorporation of attention mechanisms specifically designed for histopathological image analysis, which could enhance the model's focus on relevant nuclear structures while suppressing background noise. Additionally, exploring multi-task learning approaches that simultaneously perform nuclei segmentation and classification could provide more comprehensive analytical capabilities for pathological assessment.

Table 4 shows the evaluation results of the proposed method on the CPM-17 dataset. In order to fully evaluate the model success, the results of recent works conducted with the CPM-17 dataset are also included.

Table 4. Comparision recent works with CPM-17 dataset

Model	DCS	F1 Score	IoU	AJI
Mask2Former [44]	-	0.782	-	0.602
Rtmdet [45]	-	0.775	-	0.607
CACS [46]	0.751	-	-	0.546
Micro-Net [47]	**0.857**	-	-	0.661
HistoNeXt [48]	0.826	-	-	0.625
LinkNet supported by Vgg16 backbone	0.809	**0.813**	**0.687**	**0.681**

The evaluation of the proposed LinkNet supported by Vgg16 backbone model on the CPM-17 dataset offers critical insights into its generalizability across different histopathological domains. Although the model was exclusively trained on the CryoNuSeg dataset, it achieved highly competitive results on the CPM-17 test set, with an F1 Score of 0.813 and an AJI score of 0.681. These findings are particularly significant given the inherent differences between the two datasets in terms of tissue types, staining variability, and nuclei morphology. The consistency of performance across distinct datasets without additional fine-tuning indicates that the feature representations learned by the model are

robust and transferable, underscoring its potential applicability in real-world clinical settings where training data may be limited or heterogeneous.

This generalization capability highlights a major strength of the proposed approach, namely its architectural design that emphasizes rich and hierarchical feature extraction through the VGG16 backbone. While many existing models exhibit high accuracy within the scope of the training domain, they often fail to maintain performance when exposed to unseen data distributions. In contrast, our model demonstrates resilience against domain shift, suggesting that its encoder effectively captures nucleus-invariant characteristics that are preserved across datasets. This level of adaptability not only validates the architectural decisions made but also reinforces the relevance of encoder-focused design for histopathological segmentation tasks requiring strong cross-dataset generalization.

7. CONCLUSION

This study demonstrates that the LinkNet model supported by a VGG16 backbone provides an effective solution for nuclei segmentation in histopathological images. By prioritizing enhancements to the encoder component and implementing comprehensive data preprocessing, our approach achieved state-of-the-art performance on the CryoNuSeg dataset.

The success of this methodology highlights the importance of feature extraction capabilities in segmentation tasks and suggests that judicious selection of backbone architectures may be as critical as the choice of segmentation framework itself. These findings contribute valuable insights to the ongoing development of automated tools for histopathological image analysis, with potential applications in cancer diagnosis, prognosis, and personalized treatment planning.

ACKNOWLEDGMENT

The author gratefully acknowledges that a preliminary version of this paper appeared as an abstract in the IMISC2024 Conference Abstract Book.

ORCID

Furkan ATLAN https://orcid.org/0000-0003-1602-1941

REFERENCES

[1]. Van der Laak, J., Litjens, G., & Ciompi, F. (2021). Deep learning in histopathology: the path to the clinic. Nature medicine, 27(5), 775-784. https://doi.org/10.1038/s41591-021-01343-4

[2]. Baidoshvili, A., Bucur, A., van Leeuwen, J., van der Laak, J., Kluin, P., & van Diest, P. J. (2018). Evaluating the benefits of digital pathology implementation: time savings in laboratory logistics. Histopathology, 73(5), 784-794. https://doi.org/10.1111/his.13691

[3]. Zheng, Y., Li, J., Shi, J., Xie, F., Huai, J., Cao, M., & Jiang, Z. (2023). Kernel attention transformer for histopathology whole slide image analysis and assistant cancer diagnosis. *IEEE Transactions on Medical Imaging*, *42*(9), 2726-2739. https://doi.org/10.1109/TMI.2023.3264781

[4]. Barisoni, L., Lafata, K. J., Hewitt, S. M., Madabhushi, A., & Balis, U. G. (2020). Digital pathology and computational image analysis in nephropathology. *Nature Reviews Nephrology*, 16(11), 669-685. https://doi.org/10.1038/s41581-020-0321-6

[5]. Martos, O., Hoque, M. Z., Keskinarkaus, A., Kemi, N., Näpänkangas, J., Eskuri, M., ... & Seppänen, T. (2023). Optimized detection and segmentation of nuclei in gastric cancer images using stain normalization and blurred artifact removal. *Pathology-Research and Practice*, 248, 154694. https://doi.org/10.1016/j.prp.2023.154694

[6]. Yildirim, Z., Samet, R., Hancer, E., Nemati, N., & Mali, M. T. (2023). Gland Segmentation in H&E Histopathological Images using U-Net with Attention Module. *In 2023 Twelfth International Conference on Image Processing Theory, Tools and Applications (IPTA)*, 1-6. https://doi.org/10.1109/IPTA59101.2023.10320014

[7]. Irshad, H., Veillard, A., Roux, L., & Racoceanu, D. (2013). Methods for nuclei detection, segmentation, and classification in digital histopathology: a review—current status and future potential. *IEEE reviews in biomedical engineering*, 7, 97-114. https://doi.org/10.1109/RBME.2013.2295804

[8]. Hancer, E., Traore, M., Samet, R., Yıldırım, Z., & Nemati, N. (2023). An imbalance-aware nuclei segmentation methodology for H&E stained histopathology images. *Biomedical Signal Processing and Control*, 83, 104720. https://doi.org/10.1016/j.bspc.2023.104720

[9]. Verma, R., Kumar, N., Patil, A., Kurian, N. C., Rane, S., Graham, S., ... & Sethi, A. (2021). MoNuSAC2020: A multi-organ nuclei segmentation and classification challenge. *IEEE Transactions on Medical Imaging*, 40(12), 3413-3423. https://doi.org/10.1109/TMI.2021.3085712

[10]. Long, J., Shelhamer, E., & Darrell, T. (2015). Fully convolutional networks for semantic segmentation. *In Proceedings of the IEEE conference on computer vision and pattern recognition*, 3431-3440.

[11]. Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. *In Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference*, 234-241. https://doi.org/10.1007/978-3-319-24574-4_28

[12]. Zhou, Z., Rahman Siddiquee, M. M., Tajbakhsh, N., & Liang, J. (2018). Unet++: A nested u-net architecture for medical image segmentation. *In Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: 4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018, Held in Conjunction with MICCAI 2018*, 3-11. https://doi.org/10.1007/978-3-030-00889-5_1

[13]. Çiçek, Ö., Abdulkadir, A., Lienkamp, S. S., Brox, T., & Ronneberger, O. (2016). 3D U-Net: learning dense volumetric segmentation from sparse annotation. *In Medical Image Computing and Computer-Assisted Intervention–MICCAI 2016: 19th International Conference*, 424-432. https://doi.org/10.1007/978-3-319-46723-8_49

[14]. Abedalla, A., Abdullah, M., Al-Ayyoub, M., & Benkhelifa, E. (2020). The 2ST-UNet for pneumothorax segmentation in chest X-Rays using ResNet34 as a backbone for U-Net. *arXiv preprint arXiv:2009.02805*.

[15]. Sharma, N., & Gupta, S. (2023). Semantic Segmentation of Gastrointestinal Tract using UNet Model with ResNet 18 Backbone. *In 2023 International Conference in Advances in Power, Signal, and Information Technology (APSIT)*, 226-230. https://doi.org/10.1109/APSIT58554.2023.10201739

[16]. Das, N., & Das, S. (2024). Attention-UNet architectures with pretrained backbones for multi-class cardiac MR image segmentation. *Current Problems in Cardiology*, 49(1), 102129. https://doi.org/10.1016/j.cpcardiol.2023.102129

[17]. Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.

[18]. Chaurasia, A., & Culurciello, E. (2017). Linknet: Exploiting encoder representations for efficient semantic segmentation. *In 2017 IEEE visual communications and image processing (VCIP)*, 1-4. https://doi.org/10.1109/VCIP.2017.8305148

[19]. Long, F. (2020). Microscopy cell nuclei segmentation with enhanced U-Net. *BMC bioinformatics*, 21(1), 8. https://doi.org/10.1186/s12859-019-3332-1

[20]. Lal, S., Das, D., Alabhya, K., Kanfade, A., Kumar, A., & Kini, J. (2021). NucleiSegNet: Robust deep learning architecture for the nuclei segmentation of liver cancer histopathology images. *Computers in Biology and Medicine*, 128, 104075. https://doi.org/10.1016/j.compbiomed.2020.104075

[21]. Kumar, N., Verma, R., Anand, D., Zhou, Y., Onder, O. F., Tsougenis, E., ... & Sethi, A. (2019). A multi-organ nucleus segmentation challenge. *IEEE transactions on medical imaging*, 39(5), 1380-1391. https://doi.org/10.1109/TMI.2019.2947628

[22]. Qin, J., He, Y., Zhou, Y., Zhao, J., & Ding, B. (2022). REU-Net: Region-enhanced nuclei segmentation network. *Computers in Biology and Medicine*, 146, 105546. https://doi.org/10.1016/j.compbiomed.2022.105546

[23]. Ahmad, I., Xia, Y., Cui, H., & Islam, Z. U. (2023). DAN-NucNet: A dual attention based framework for nuclei segmentation in cancer histology images under wild clinical conditions. *Expert Systems with Applications*, 213, 118945. https://doi.org/10.1016/j.eswa.2022.118945

[24]. Li, Z., Tang, Z., Hu, J., Wang, X., Jia, D., & Zhang, Y. (2023). Nst: a nuclei segmentation method based on transformer for gastrointestinal cancer pathological images. *Biomedical Signal Processing and Control*, 84, 104785. https://doi.org/10.1016/j.bspc.2023.104785

[25]. Goceri, E. (2024). Nuclei segmentation using attention aware and adversarial networks. *Neurocomputing*, 579, 127445. https://doi.org/10.1016/j.neucom.2024.127445

[26]. Hasan, M. J., Ahmad, W. S. H. M. W., Fauzi, M. F. A., & Abas, F. S. (2024). Hybrid Deep Learning Architectures for Histological Image Segmentation. *In 2024 International Conference on Artificial Intelligence in Information and Communication (ICAIIC)*, 75-80. https://doi.org/10.1109/ICAIIC60209.2024.10463355

[27]. Hernández-García, A., & König, P. (2018). Data augmentation instead of explicit regularization. *arXiv preprint arXiv:1806.03852*.

[28]. Vu, Q. D., Graham, S., Kurc, T., To, M. N. N., Shaban, M., Qaiser, T., ... & Farahani, K. (2019). Methods for segmentation and classification of digital microscopy tissue images. *Frontiers in bioengineering and biotechnology*, *7*, 53. https://doi.org/10.3389/fbioe.2019.00053

[29]. Mahbod, A., Schaefer, G., Bancher, B., Löw, C., Dorffner, G., Ecker, R., & Ellinger, I. (2021). CryoNuSeg: A dataset for nuclei instance segmentation of cryosectioned H&E-stained histological images. *Computers in biology and medicine*, 132, 104349. https://doi.org/10.1016/j.compbiomed.2021.104349

[30]. Zhong, Z., Zheng, L., Kang, G., Li, S., & Yang, Y. (2020). Random erasing data augmentation. *In Proceedings of the AAAI conference on artificial intelligence*, 34(7), 13001-13008. https://doi.org/10.1609/aaai.v34i07.7000

[31]. Ioffe, S., & Szegedy, C. (2015, June). Batch normalization: Accelerating deep network training by reducing internal covariate shift. *In International conference on machine learning*, 448-456.

[32]. Nair, V., & Hinton, G. E. (2010). Rectified linear units improve restricted boltzmann machines. *In Proceedings of the 27th international conference on machine learning (ICML-10)*, 807-814.

[33]. Nijaguna, G. S., Babu, J. A., Parameshachari, B. D., de Prado, R. P., & Frnda, J. (2023). Quantum Fruit Fly algorithm and ResNet50-VGG16 for medical diagnosis. *Applied Soft Computing*, 136, 110055. https://doi.org/10.1016/j.asoc.2023.110055

[34]. Sun, J., Yang, F., Cheng, J., Wang, S., & Fu, L. (2024). Nondestructive identification of soybean protein in minced chicken meat based on hyperspectral imaging and VGG16-SVM. *Journal of Food Composition and Analysis*, 125, 105713. https://doi.org/10.1016/j.jfca.2023.105713

[35]. Zhang, J., Li, C., Kosov, S., Grzegorzek, M., Shirahama, K., Jiang, T., ... & Li, H. (2021). LCU-Net: A novel low-cost U-Net for environmental microorganism image segmentation. *Pattern Recognition*, 115, 107885. https://doi.org/10.1016/j.patcog.2021.107885

[36]. Le Dinh, T., Lee, S. H., Kwon, S. G., & Kwon, K. R. (2022). Cell nuclei segmentation in cryonuseg dataset using nested unet with efficientnet encoder. *In 2022 International Conference on Electronics, Information, and Communication (ICEIC)*, 1-4. https://doi.org/10.1109/ICEIC54506.2022.9748537

[37]. Liao, L., & Zhang, Z. (2023). Nuclear Segmentation Based on Recurrent Iteration and Fusion Attention Mechanism. *In 2023 3rd International Conference on Electronic Information Engineering and Computer Science (EIECS)*, 365-371. https://doi.org/10.1109/EIECS59936.2023.10435439

[38]. Mahbod, A., Schaefer, G., Dorffner, G., Hatamikia, S., Ecker, R., & Ellinger, I. (2022). A dual decoder u-net-based model for nuclei instance segmentation in hematoxylin and eosin-stained histological images. *Frontiers in Medicine*, 9, 978146. https://doi.org/10.3389/fmed.2022.978146

[39]. Kadaskar, M., & Patil, N. (2023). Nuclei Segmentation using EfficientNetV2 and Convolutional Block Attention Module. *In 2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, 1-5. https://doi.org/10.1109/ICCCNT56998.2023.10307210

[40]. Wang, Z., Zhang, Y., Wang, Y., Cai, L., & Zhang, Y. (2024). Dynamic Pseudo Label Optimization in Point-Supervised Nuclei Segmentation. *arXiv preprint arXiv:2406.16427*.

[41]. Chen, Z., Xu, Q., Liu, X., & Yuan, Y. (2024). UN-SAM: Universal Prompt-Free Segmentation for Generalized Nuclei Images. *arXiv preprint arXiv:2402.16663*.

[42]. Sun, X., Li, S., Chen, Y., Chen, J., Geng, H., Sun, K., ... & Zhang, H. (2025). Multiple Differential Convolution and Local-Variation Attention UNet: Nucleus Semantic Segmentation Based on Multiple Differential Convolution and Local-Variation Attention. *Electronics*, 14(6), 1058. https://doi.org/10.3390/electronics14061058

[43]. Zhang, X., Xu, J., He, D., Wang, K., & Wang, L. (2025). Lightweight multi-scale attention group fusion structure for nuclei segmentation. *The Journal of Supercomputing*, 81(1), 199. https://doi.org/10.1007/s11227-024-06710-9

[44]. Cheng, B., Misra, I., Schwing, A. G., Kirillov, A., & Girdhar, R. (2022). Masked-attention mask transformer for universal image segmentation. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 1290-1299).

[45]. Lyu, C., Zhang, W., Huang, H., Zhou, Y., Wang, Y., Liu, Y., ... & Chen, K. (2022). Rtmdet: An empirical study of designing real-time object detectors. *arXiv preprint arXiv:2212.07784*. arXiv preprint arXiv:2212.07784 2022

[46]. Ruan, H. (2024, December). A Color-Aware Unsupervised Segmentation Network for Nuclei in Histopathology Images. In *2024 4th International Conference on Electronic Information Engineering and Computer Communication (EIECC)* (pp. 627-631). IEEE. https://doi.org/10.1109/EIECC64539.2024.10929140

[47]. Raza, S. E. A., Cheung, L., Shaban, M., Graham, S., Epstein, D., Plantaris, S., ... & Rajpoot, N. M. (2019). Micro-Net: A unified model for segmentation of various objects in microscopy images. *Medical image analysis*, *52*, 160-173. https://doi.org/10.1016/j.media.2018.12.003

[48]. Chen, J., Wang, R., Dong, W., He, H., & Wang, S. (2025). HistoNeXt: dual-mechanism feature pyramid network for cell nuclear segmentation and classification. *BMC Medical Imaging*, *25*(1), 9. https://doi.org/10.1186/s12880-025-01550-2

MAKÜ BBBF
BUCAK BİLGİSAYAR VE BİLİŞİM FAKÜLTESİ

Scientific Journal of Mehmet Akif Ersoy University, 2025, 8(1), Page:47-56

Research Article

Comparative Analysis of Förster Resonance Energy Transfer (FRET) in Spherical and Planar Geometries

Onur İNAN[1*]

[1]*Department of Biomedical Device Technology, Bucak Emin Gülmez Technical Sciences Vocational School, Burdur Mehmet Akif Ersoy University, Burdur, Türkiye*

ABSTRACT

This research explores the influence of geometric configuration on Förster Resonance Energy Transfer (FRET) efficiency. Specifically, it compares spherical arrangements (relevant to structures like nanoparticles) with planar arrangements (found in systems like cell membranes). A key goal is to clarify the interplay between FRET efficiency, inter-molecular distances, and the characteristic Förster distance. By employing both mathematical models and visual representations, the study seeks to provide a detailed understanding of how FRET operates under these distinct geometric constraints. The findings are intended to be broadly applicable, offering valuable insights for the design and analysis of FRET-based experimental work across diverse scientific disciplines.

Keywords: FRET (Förster Resonance Energy Transfer), Geometric Configuration, Molecular Distance, Simulation

1. INTRODUCTION

Förster Resonance Energy Transfer (FRET), also known in some contexts as Fluorescence Resonance Energy Transfer, is a non-radiative process through which energy is transferred between two fluorophores: a donor and an acceptor. This transfer occurs when the donor molecule, in an excited state, passes its energy to a nearby acceptor molecule without emitting a photon. The mechanism relies on dipole-dipole interactions, which set it apart from traditional radiative energy transfer methods. One of the most distinctive characteristics of FRET is its extreme sensitivity to the distance between the donor and acceptor molecules [1-3].

The efficiency of energy transfer decreases rapidly as the separation distance increases, with the process being most effective when the molecules are within a range of approximately 1 to 10 nanometers. This relationship is described mathematically by the Förster equation, which shows that the transfer efficiency is inversely proportional to the sixth power of the distance between the donor and acceptor [4, 5]. Because of this strong dependence on distance, FRET is often called a "spectroscopic ruler," as it allows researchers to measure nanoscale distances with exceptional precision. The ability of FRET to reveal molecular interactions and spatial arrangements has made it an indispensable tool in fields such as biophysics, molecular biology, and chemistry. Researchers use this technique to investigate dynamic processes, including protein-protein interactions, structural changes in biomolecules, and the assembly of molecular complexes [6, 7].

FRET efficiency, denoted as E, is a quantitative measure representing the proportion of energy transferred from the donor molecule to the acceptor molecule. This efficiency, ranging from 0 (indicating no transfer) to 1 (representing complete transfer), is fundamentally governed by the distance, r, separating the donor and acceptor, and the Förster distance, R_o, a characteristic parameter unique to each donor-acceptor pair [8, 9].

The relationship is mathematically expressed as: $E = 1 / (1 + (r/R_o)^6)$. The Förster distance, R_o, signifies the specific separation at which the FRET efficiency reaches 50%. Several factors influence its value: the extent of spectral overlap between the donor's emission and the acceptor's absorption, the relative orientation of the donor and acceptor transition dipoles, and the refractive index of the intervening medium. The Förster distance is calculated using the equation (1). [10]

***Corresponding Author Email:** oinan@mehmetakif.edu.tr
Submitted: 26.03.2025 **Revision Requested:** 23.04.2025 **Last Revision Received:** 21.05.2025
Accepted: 21.05.2025 **Published Online:** 23.05.2025

Burdur Mehmet Akif Ersoy University Press

Cite this article as: Inan, O. (2025). Comparative Analysis of Förster Resonance Energy Transfer (FRET) In Spherical and Planar Geometries. Scientific Journal of Mehmet Akif Ersoy University, 8(1): 47-56.
DOI: https://doi.org/10.70030/sjmakeu.1665692

https://dergipark.org.tr/sjmakeu

$$R_0^{\,6} = \frac{9000(\ln 10)\kappa^2\phi D}{128\pi^5 N_A n^4} J(\lambda) \quad (1)$$

Here, κ^2 is the orientation factor, reflecting the relative orientation of the donor emission and acceptor absorption dipoles, typically assumed to be 2/3 for randomly oriented molecules but ranging from 0 to 4. Φ_D represents the quantum yield of the donor in the absence of the acceptor. *n* is the refractive index of the medium, and N_A is Avogadro's number. The spectral overlap integral, J(λ), quantifies the degree of overlap between the donor's normalized fluorescence intensity, $F_D(\lambda)$, and the acceptor's molar extinction coefficient, $\varepsilon_A(\lambda)$, and is calculated by equation (2) [11, 12].

$$J(\lambda) = \int_0^\infty F_D(\lambda) \in_A(\lambda)\lambda^4 d\lambda \quad (2)$$

Because of its sensitivity to distance, FRET has found widespread applications in various fields (Table 1).

Table 1. Applications of FRET in Various Fields [2, 13-16]

Field	Applications of FRET
Biophysics	Studying protein folding, protein-protein interactions, conformational changes in biomolecules, and membrane dynamics.
Biochemistry	Investigating enzyme kinetics, receptor-ligand binding, and DNA/RNA interactions.
Cell Biology	Monitoring intracellular signaling pathways, visualizing molecular distributions, and tracking cellular processes.
Nanotechnology	Characterizing nanomaterials, developing biosensors, and creating nanoscale devices.
Materials Science	Studying polymer blends and self-assembly processes.

While Förster Resonance Energy Transfer (FRET) is widely used in various applications, comparative analyses of how different geometric configurations—specifically spherical versus planar arrangements—affect FRET efficiency are limited. Previous studies have often addressed FRET in spherical or planar contexts, but a direct comparison of these geometries has not been sufficiently explored. This study aims to provide further insights into this area and to understand the geometric influences on FRET efficiency.

This study investigated the intricacies of FRET efficiency. In particular, we examined how the two main geometrical configurations affect it. We focus on two main geometrical configurations in FRET analysis. The research covers two different geometric configurations for FRET analysis. The first configuration examines spherical geometry representing molecular arrangements in various biological and synthetic systems. This geometric framework accommodates donor and acceptor molecules on the spherical surface or distributed within the volumetric domain. Such arrangements are particularly important when studying biological vesicles, micelle structures, spherical nanoparticles, and specific protein complexes that exhibit spherical symmetry. The second geometric configuration involves planar arrangements in which molecular interactions occur along two-dimensional surfaces. This configuration is particularly important in biological systems, especially in studying cellular membranes, surface-fixed molecular assemblies, and layered material architectures. The planar geometry provides a fundamental framework for understanding FRET behavior in systems constrained to two dimensions. This research integrates rigorous mathematical formulations with comprehensive visual representations to develop a detailed understanding of FRET behavior under varying geometric constraints. The findings aim to advance the theoretical framework underlying FRET-based experimental design and data interpretation across diverse scientific applications. This understanding proves particularly valuable for researchers developing new experimental protocols or interpreting complex FRET data in various molecular systems.

2. MATHEMATICAL MODELING: FRAMEWORK AND APPROACH

This section outlines the mathematical underpinnings of our analysis, detailing the methods employed to compute distances and, subsequently, FRET efficiencies within both spherical and planar geometric configurations. Furthermore, it provides a brief overview of the parameter sensitivity analysis conducted to assess the influence of the Förster distance (R_o) on the overall FRET process.

2.1. Spherical Geometry: A Curved Landscape

FRET) efficiency highly depends on the distance, r, between the donor and the acceptor. Accurately determining this distance is critical in calculating FRET efficiency. Using a spherical coordinate system allows for geometrically accurate position determination for distance calculations. In the spherical coordinate system, the distance of the donor molecule from the center is represented by r_d. In the simplest case, if both donor and acceptor are located on the sphere's surface, r_d is directly equal to the sphere's radius. However, if the donor molecule is located at a different distance from the center, r_d can also express this distance. Similarly, we define the distance of the acceptor from the center by r_a. Besides distances, angles must also be taken into account. The polar angle θ measures the angular separation between the donor and the acceptor, and this angle is determined when viewed from the center of the sphere. The angle θ varies between 0 and π radians (180 degrees). Another angle, the azimuthal angle ϕ, defines the projection of the donor and acceptor positions on the "equator" (xy-plane). This angle is measured from the positive x-axis and ranges from 0 to 2π radians (360 degrees). These angles and distances are necessary for an accurate calculation of FRET efficiency, and these calculations play a critical role in determining the efficiency of energy transfer.

Using the values of r_d, r_a, θ and ϕ, a special version of the cosine law adapted to spherical coordinates can calculate the distance (r) between donor and acceptor. This gives the precise separation needed for FRET calculations.

This adaptation is crucial because it accounts for the inherent curvature of the spherical geometry.

$$r = \sqrt{r_d^{\,2} + r_a^{\,2} - 2r_d r_a \cos\gamma} \tag{3}$$

Where γ represents the central angle between the donor and acceptor position vectors, it is calculated using equation (2).

$$\cos\gamma = \cos\theta_d \cos\theta_a + \sin\theta_d \sin\theta_a \cos(\phi_d - \phi_a) \tag{4}$$

In the simplified case where both donor and acceptor are on the surface of a sphere with radius R, and only concerned with the angle between them, the distance equation simplifies to equation (3). If only given a single angle, equation (5) becomes equation (3).

$$r = \sqrt{2R^2(1-\cos\gamma)} = 2R\sin(\gamma/2) \tag{5}$$

Once the distance, r, is determined, the FRET efficiency, E, is calculated using the standard Förster equation.

2.2. Planar Geometry

Planar geometry deals with donor and acceptor molecules on a flat, two-dimensional plane. This applies to systems such as cell membranes or surface-fixed molecules. The distance, *r*, between the donor and acceptor is calculated using the Pythagorean theorem in Cartesian coordinates (equation 6). Here, *x*, *y*, and *z* represent the Cartesian coordinates for the donor and acceptor molecules.

$$r = \sqrt{(x_d - x_a)^2 + (y_d - y_a)^2 + (z_d - z_a)^2} \tag{6}$$

The FRET efficiency is calculated using the same Förster equation as the spherical geometry.

2.3. Parameter Sensitivity Analysis

A parameter sensitivity analysis was performed to understand the impact of the Förster distance R_o on FRET efficiency:

1. **Varying R_o**: Systematically changing the value of R_o within a relevant range (e.g., from 5 nm to 15 nm).
2. **Calculating E**: For each R_o value, we calculate the FRET efficiency, E, as a function of distance, r, using the appropriate distance equation for the chosen geometry (spherical or planar).

3. **Plotting E vs. r**: Plotting E versus r for different values of R_0 allows one to visualize how the efficiency-distance relationship changes with R_0.
4. **Logarithmic Scale Analysis**: Plotting the relationship between distance and FRET efficiency on a logarithmic scale helps highlight the efficiency behavior at very short and very long distances and better visualize the steepness of efficiency declines.
5. **Orientation Factor (κ^2)**: While often assumed to be 2/3 (for dynamic random averaging), it's important to acknowledge that the orientation factor can significantly impact R_0 and thus FRET efficiency. A brief discussion of κ^2 and its potential influence could be included, although a full analysis of κ^2 is beyond the scope of this basic model.

This systematic approach allows us to quantitatively assess the sensitivity of FRET efficiency to changes in the Förster distance and to understand the implications for interpreting FRET measurements.

3. RESULTS AND DISCUSSION

This section presents and discusses the results of the FRET efficiency calculations for both spherical and planar geometries, highlighting the key differences and the underlying reasons for these differences.

3.1. Spherical Geometry

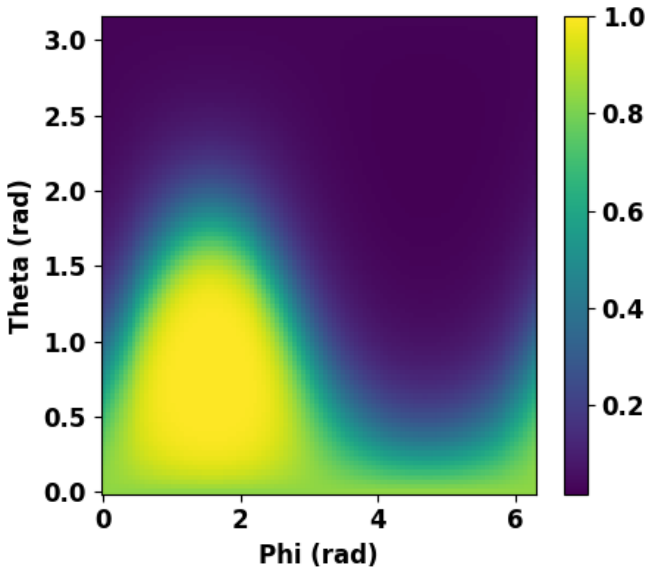

Fig 1. FRET Efficiency (2D Heat Map)

The 2D heat map (Figure 1) visualizes FRET efficiency as a function of the azimuthal angle (ϕ) and the polar angle (θ). Crucially, maximum FRET efficiency doesn't occur uniformly but is localized to specific combinations of ϕ and θ. This indicates a strong angular dependence. As θ increases (moving from the "pole" towards the "equator" of the sphere), the distance, r, between the donor and acceptor generally increases, leading to a decrease in FRET efficiency. The color gradient of the heat map (e.g., from yellow for high efficiency to purple for low efficiency) directly reflects this distance-dependent change.

The 3D surface plot (Figure 2) represents the same data's three-dimensional representation. The "wave-like" pattern is a direct visual consequence of the angular dependence of the distance, r, and consequently, the FRET efficiency, *E*. The symmetry observed along the azimuthal axis (ϕ) is expected, especially when r_d and r_a are constant, as rotating the system around the z-axis (changing ϕ) doesn't change the *relative* distance between donor and acceptor if their θ values are fixed.

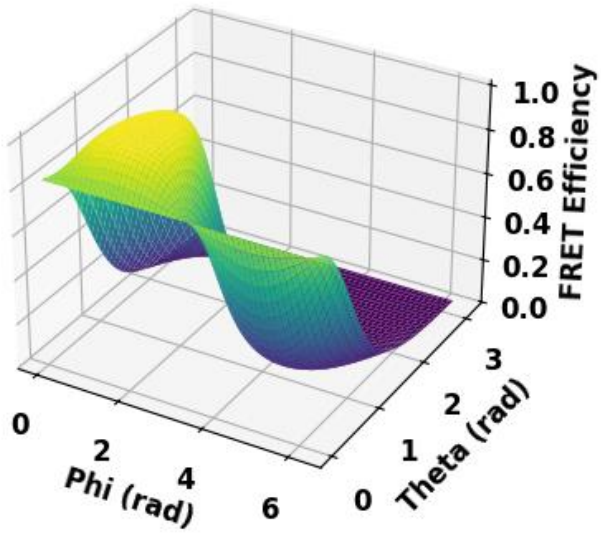

Fig 2. FRET Efficiency (3D Surface Plot)

By taking a cross-section of the 3D surface plot at a fixed ϕ (e.g., ϕ = π), we obtain a 2D plot of FRET efficiency versus θ (Figure 3). This plot demonstrates the monotonic decrease in FRET efficiency as θ increases. This monotonic decrease is a direct consequence of the inverse sixth-power relationship between distance and FRET efficiency in the Förster equation. As θ increases, *r* increases, and *E* decreases rapidly.

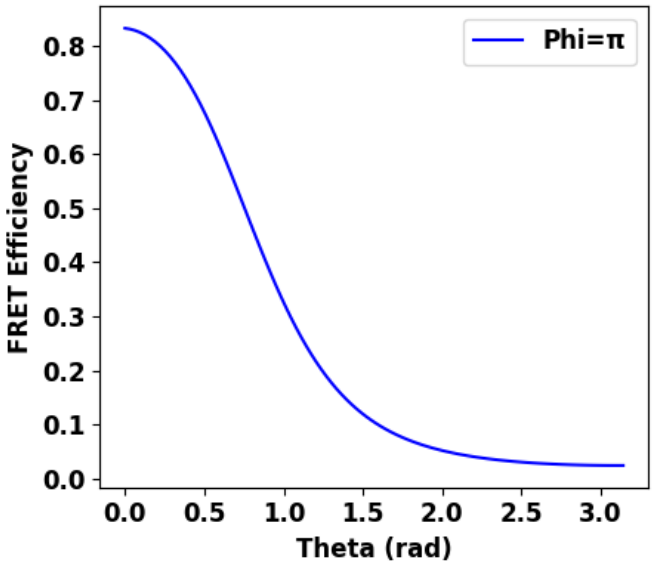

Fig 3. FRET Efficiency vs. θ (Cross-section)

The curvature inherent in spherical geometry is the primary reason for the observed FRET efficiency patterns. Even small angular separation (θ) changes can significantly change the distance, *r*, between the donor and acceptor. This makes FRET efficiency highly sensitive to the angular configuration in spherical systems. The symmetry along the azimuthal angle (ϕ), when r_d and r_a are constant, simplifies the analysis somewhat, but the fundamental angular dependence remains.

3.2. Planar Geometry

The 2D heat map (Figure 4) for planar geometry shows FRET efficiency as a function of the donor and acceptor positions on the X-Y plane. The maximum FRET efficiency is observed at the plane's center (assuming the acceptor is

fixed at the origin), where the donor and acceptor are closest. As we move radially outward from the center, the distance, r, increases, and the FRET efficiency decreases. The color gradient reflects this radial decrease.

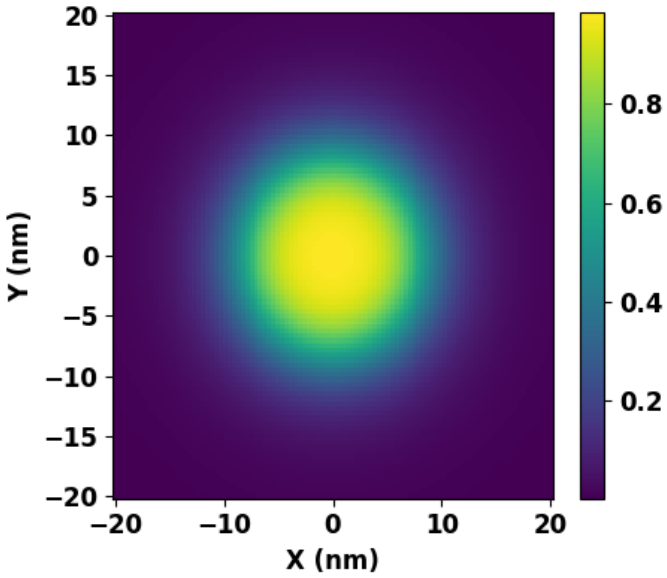

Fig 4. FRET Efficiency (2D Heat Map)

The 3D surface plot (Figure 5) shows a peak at the center, corresponding to the minimum donor-acceptor distance. The efficiency decreases symmetrically in all directions, reflecting the isotropic nature of the plane. The spherical geometry has no wave-like patterns because there's no angular dependence; the efficiency depends solely on the radial distance.

A cross-section of the 3D surface plot (Figure 6) along the X-axis (with Y fixed, e.g., Y = 0) shows the FRET efficiency as a function of the X-coordinate. The efficiency is highest at X = 0 (where the donor is closest to the acceptor) and decreases symmetrically as the donor moves away from the center along the X-axis. This symmetry is a direct consequence of the linear distance relationship in planar geometry.

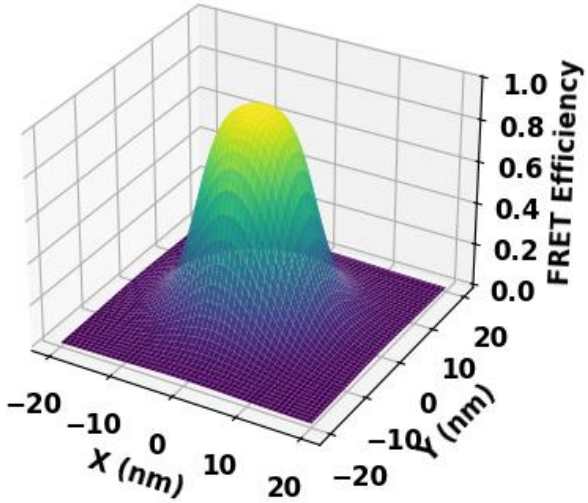

Fig 5. FRET Efficiency (3D Surface Plot)

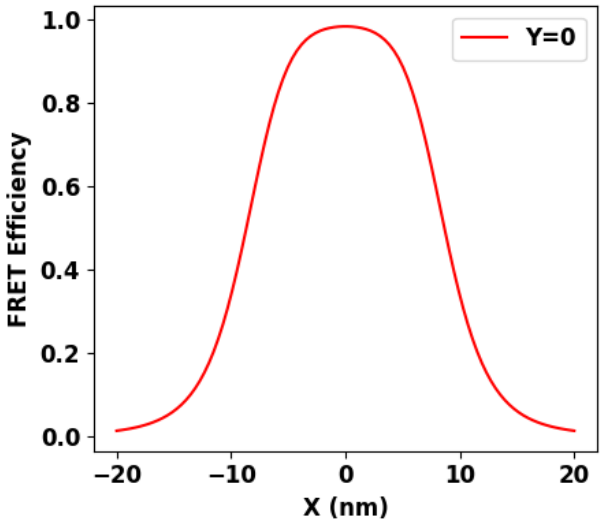

Fig 6. FRET Efficiency vs. X (Cross-section)

The absence of curvature in planar geometry is the key factor influencing FRET efficiency. The distance between the donor and acceptor is a simple, linear function of their Cartesian coordinates, leading to symmetric and predictable efficiency patterns. There are no angular dependencies; the FRET efficiency is solely a function of the Cartesian distance, r.

3.3. Parameter Analysis

Figure 7 shows how FRET efficiency varies depending on the distance between donor and acceptor. In Figure 7.a, the Y axis shows the FRET efficiency, and the X axis shows the distance (in nm). In this graph, the blue line represents R_0=8 nm, the orange line R_0=10 nm, and the green line R_0=12 nm. The blue line starts at 1.0000 and decreases to approximately 0.99975 at 20 nm, while the orange line decreases from 1.0000 to 0.99985, and the green line decreases from 1.0000 to 0.99995. This shows that the FRET efficiency of lower R_0 values decreases more rapidly at shorter distances. Figure 7.b presents the relationship between distance and FRET efficiency on a logarithmic scale. The Y-axis again shows FRET efficiency, while the X-axis shows distance on a logarithmic scale. This graph's blue, orange, and green lines represent R_0=8 nm, R_0=10 nm, and R_0=12 nm, respectively. The blue line shows a sharp decrease at about 15 nm starting at 1.0000, while the orange line shows a similar but less pronounced decrease. The green line shows less change with increasing distance, allowing the efficiency to remain more constant. It emphasizes that FRET efficiency decreases more rapidly at shorter distances and that higher R_0 values tend to keep the efficiency more constant.

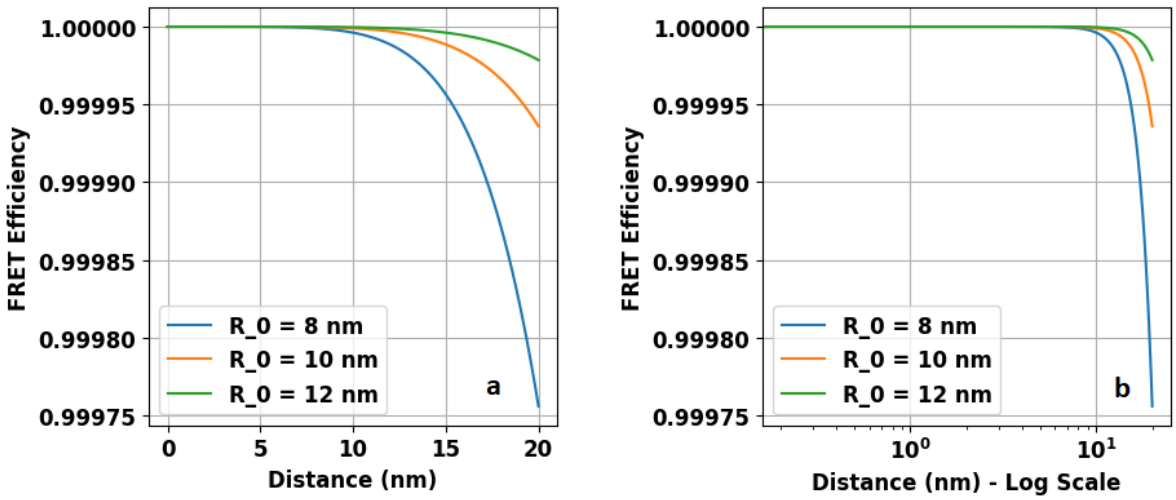

Fig 7. FRET efficiency versus distance

3.4. Comparison of Spherical and Planar Geometries

The following table summarizes the key differences between spherical and planar geometries:

Table 1. Comparison of Spherical and Planar Geometries

Aspect	Spherical Geometry	Planar Geometry
Distance Calculation	Depends on angular separation (θ, and potentially ϕ) and curvature. The relationship between distance and angle is non-linear.	Depends linearly on Cartesian coordinates (x, y, and potentially z). The relationship between distance and coordinates is linear.
Efficiency Patterns	Exhibits angular dependence. Efficiency is not uniform and shows a wave-like pattern in 3D plots. Symmetry may exist along the azimuthal angle (ϕ) under certain conditions.	Exhibits radial symmetry. Efficiency decreases uniformly as distance from the acceptor increases. No angular dependence.
Sensitivity	Highly sensitive to small changes in angular separation, especially near the poles.	Uniform sensitivity across the plane. Changes in distance result in predictable changes in efficiency.
Complexity	More complex due to the non-linear relationship between distance and angle. Requires careful consideration of angular coordinates.	Simpler due to the linear relationship between distance and coordinates.
Applications	Vesicles, micelles, spherical nanoparticles, and protein complexes with curved surfaces.	Cell membranes, surface-immobilized molecules, layered materials.

This study quantitatively demonstrates that FRET efficiency behaves nonlinearly in spherical systems due to angular dependence, while in planar systems, there is a predictable relationship with radial distance. Saini et al. [18] emphasize that the assumptions of Förster theory are limited in large conjugated structures, metal nanoparticles, and polymer systems. Both studies recognize the critical importance of the Förster distance parameter: The current research examines the practical implications of its optimization, while Saidi's analysis shows that this parameter can be affected by geometry and environmental factors. Non-spherical structures and orientation dynamics indicate that the theoretical models in both studies require revision. In conclusion, geometric design and theoretical model selection in FRET-based systems should be optimized to achieve consistent results from nanomaterials engineering to biophysical applications.

Oliden-Sánchez et al. [19] show that organic chromophores (DMAN, rhodamine 123, Nile Blue) entrapped in a rigid aluminophosphate matrix provide emission in the entire visible spectrum by energy transfer via FRET. The one-dimensional extra-wide channels of the IFO-type zeolitic structure allow the confinement of bulky dyes, while the fixation of the D-A pairs in the rigid matrix facilitates the control of geometrical parameters. While this study investigates its optimization as geometry-dependent, this work shows how it can be used in practice with optimized D-A pairs in a rigid matrix. In conclusion, this study models the fundamental effects of geometry on FRET, while this work presents a feasible way to improve FRET efficiency through material design. The two approaches can complement the development of FRET-based materials in controlled geometries.

Grzedowsk et al. [20] have demonstrated how DNA nanocube structures can manipulate FRET signals in a geometrically optimized system with molecular level control on the gold surface. DNA nanocubes' self-organizing monolayer (SAM) allows the inter-fluorophore distance to be fixed with the nanocube dimensions, minimizing the angular dependence problem in this study. This rigid structure allows for a controlled decrease or increase of the FRET signal during target DNA hybridization, similar to the predictable FRET behavior in planar systems. It emphasizes the critical role of geometry in FRET-based biosensor design with different dimensions. The geometrical optimization obtained at the end of this study may prove indispensable for understanding fundamental FRET dynamics and developing practical biosensor applications.

4. CONCLUSION

This study systematically explores the interplay between geometry and FRET efficiency, revealing critical insights into energy transfer dynamics in spherical and planar systems. The key conclusions are as follows:

Geometric Influence on FRET Efficiency: In spherical geometries, FRET efficiency exhibits strong angular dependence due to the nonlinear relationship between donor-acceptor distance and polar angle (θ). The sphere's curvature amplifies sensitivity to angular displacements, with efficiency declining sharply as θ increases (e.g., from the pole to the equator). The observed wave-like patterns in 3D plots and azimuthal symmetry underscore the complexity introduced by spherical curvature. FRET efficiency depends solely on radial distance in planar geometries, resulting in radially symmetric patterns. The absence of angular dependence simplifies the relationship, with efficiency decreasing predictably as donors move away from the acceptor. This linearity contrasts starkly with the nonlinear behavior in spherical systems.

Role of Förster Distance (R_o): The parameter analysis highlights R_o as a critical determinant of FRET efficiency. Smaller R_o values (e.g., 8 nm) lead to a rapid decline in efficiency at shorter distances, while larger R_o values (e.g., 12 nm) extend the effective energy transfer range. This emphasizes the importance of selecting donor-acceptor pairs with R_o tailored to the system's spatial constraints.

Sensitivity and Applications: Spherical systems (e.g., vesicles, nanoparticles) demand careful consideration of angular configurations, as small positional changes near the poles significantly alter efficiency. This sensitivity may limit FRET reliability in highly curved environments unless spatial constraints are well-characterized.

Planar systems (e.g., cell membranes) offer greater predictability due to radial symmetry, making them ideal for quantitative studies where distance is the primary variable.

Practical Implications: These findings underscore the necessity of accounting for geometry when designing FRET-based experiments or interpreting data. For instance, spherical geometries may require advanced modeling to disentangle angular effects, while planar systems benefit from straightforward distance calibration. Additionally, the R_o-dependent efficiency profiles suggest that optimizing the donor-acceptor pair's Förster radius can enhance resolution in targeted applications, such as biosensing or molecular imaging.

By bridging theoretical models with geometric realities, this work provides a framework for improving the accuracy of FRET measurements across diverse nanoscale and biological systems. Future studies could extend these principles to hybrid geometries or dynamic systems where curvature and donor-acceptor mobility evolve over time.

ORCID

Onur INAN https://orcid.org/0000-0003-4573-7025

REFERENCES

[1]. Clegg, R. M. (2009). Förster resonance energy transfer—FRET what is it, why do it, and how it's done. Laboratory techniques in biochemistry and molecular biology, 33, 1-57.
[2]. Kaur, A., Kaur, P., & Ahuja, S. (2020). Förster resonance energy transfer (FRET) and applications thereof. Analytical Methods, 12(46), 5532-5550.
[3]. Szabó, Á., Szöllősi, J., & Nagy, P. (2022). Principles of resonance energy transfer. Current protocols, 2(12), e625.
[4]. Chen, G. (2005). Nanoscale energy transport and conversion: a parallel treatment of electrons, molecules, phonons, and photons. Oxford University Press.
[5]. Metz, S., & Marian, C. M. (2025). Computational Approach to Phosphor-Sensitized Fluorescence Based on Monomer Transition Densities. Journal of Chemical Theory and Computation, 21(5), 2569-2581.

[6]. Nüesch, M., Ivanović, M. T., Nettels, D., Best, R. B., & Schuler, B. (2025). Accuracy of distance distributions and dynamics from single-molecule FRET. Biophysical Journal.
[7]. Loidolt-Krüger, M. (2025). Perspective: fluorescence lifetime imaging and single-molecule spectroscopy for studying biological condensates. Methods in Microscopy, (0).
[8]. Berney, C., & Danuser, G. (2003). FRET or no FRET: a quantitative comparison. Biophysical journal, 84(6), 3992-4010.
[9]. Shrestha, D., Jenei, A., Nagy, P., Vereb, G., & Szöllősi, J. (2015). Understanding FRET as a research tool for cellular studies. International journal of molecular sciences, 16(4), 6718-6756.
[10]. Zhou, M., Zhang, K., Li, X., Ge, Y., Zhang, W., Lu, P., & Hao, X. (2024). Improved Exciton Diffusion through Modulating Förster Resonance Energy Transfer for Efficient Organic Solar Cells. Solar RRL, 8(13), 2400136.
[11]. Wong, K. F., Bagchi, B., & Rossky, P. J. (2004). Distance and orientation dependence of excitation transfer rates in conjugated systems: beyond the Förster theory. The Journal of Physical Chemistry A, 108(27), 5752-5763.
[12]. Patterson, G. H., Piston, D. W., & Barisas, B. G. (2000). Förster distances between green fluorescent protein pairs. Analytical biochemistry, 284(2), 438-440.
[13]. Fang, C., Huang, Y., & Zhao, Y. (2023). Review of FRET biosensing and its application in biomolecular detection. American journal of translational research, 15(2), 694.
[14]. Gopal, A. R., Joy, F., Dutta, V., Devasia, J., Dateer, R., & Nizam, A. (2024). Carbon dot-based fluorescence resonance energy transfer (FRET) systems for biomedical, sensing, and imaging applications. Particle & Particle Systems Characterization, 41(1), 2300072.
[15]. Liu, C. (2024, January). Application of FRET and TBET in bioimaging and biosensors. In Third International Conference on Biological Engineering and Medical Science (ICBioMed2023) (Vol. 12924, pp. 346-353). SPIE.
[16]. Kaur, A., & Dhakal, S. (2020). Recent applications of FRET-based multiplexed techniques. TrAC Trends in Analytical Chemistry, 123, 115777.
[17]. Medintz, I. L., & Hildebrandt, N. (Eds.). (2013). FRET-Förster resonance energy transfer: from theory to applications. John Wiley & Sons.
[18]. Saini, S., Srinivas, G., & Bagchi, B. (2009). Distance and orientation dependence of excitation energy transfer: from molecular systems to metal nanoparticles. The Journal of Physical Chemistry B, 113(7), 1817-1832.
[19]. Oliden-Sánchez, A., Sola-Llano, R., Pérez-Pariente, J., Gómez-Hortigüela, L., & Martínez-Martínez, V. (2024). Exploiting the photophysical features of DMAN template in ITQ-51 zeotype in the search for FRET energy transfer. Physical Chemistry Chemical Physics, 26(2), 1225-1233.
[20]. Grzedowski, A. J., Jun, D., Mahey, A., Zhou, G. C., Fernandez, R., & Bizzotto, D. (2024). Engineering DNA Nanocube SAM Scaffolds for FRET-Based Biosensing: Interfacial Characterization and Sensor Demonstration. Journal of the American Chemical Society, 146(46), 31560-31573.