Journal of Advances in Mathematics and Computer Science

xx(x): ....., 20yy; Article no.JAMCS.xxxxx

ISSN: 2456-9968

(Past name: British Journal of Mathematics & Computer Science, Past ISSN: 2231-0851)

[thm]Algorithm

Evaluating the predictive performance of monthly inflation rates in Sri Lanka using the hybrid model (HB)

Abstract

Aims/ objectives:This study develops and evaluates a novel hybrid model (HB) for forecasting monthly inflation rates in Sri Lanka, a country with a unique economic context, from 1988 to 2021. By integrating the Autoregressive Integrated Moving Average (ARIMA) and Artificial Neural Networks (ANNs), the study aims to overcome the limitations of traditional linear models in capturing the nonlinear patterns often observed in Sri Lankan economic data.

Objectives: The study aims to assess the predictive accuracy of the HB model against established models, emphasizing its adaptability and robustness over a historically significant period.

Methodology: Utilizing historical data, the study compares the HB model's forecasting performance with other established models, focusing on the Mean Absolute Percentage Error (MAPE) as a key metric of predictive accuracy.

Results: The HB model demonstrates superior forecasting accuracy, with a notable reduction in MAPE to 7.10%, indicating its effectiveness in capturing the complexities of the Sri Lankan inflation trend.

Conclusion: This study contributes to the field of economic forecasting by presenting a model that not only provides more accurate predictions but also adapts to the specific economic conditions of Sri Lanka. The findings have significant implications for economic planning and policy-making, highlighting the utility of hybrid forecasting models in developing economies.

Keywords: Inflation Rate; Hybrid Models; Neural Network; Forecasting; Nonlinear Models

2010 Mathematics Subject Classification: 53C25; 83C05; 57N16

1 Introduction

Inflation serves as a crucial economic barometer, indicating the rate of increase in prices for goods and services and, consequently, the erosion of a currency's purchasing power. This economic phenomenon is of particular importance in developing economies like Sri Lanka, where the country's financial stability and growth trajectory are intricately linked to inflation trends Hajirahimi and Khashei (2022). Sri Lanka's unique economic landscape, marked by its transitional economy, presents a compelling case for detailed study, given the significant role inflation plays in both macroeconomic and microeconomic spheres Fraiha Lopes et al. (2020).

The task of accurately forecasting inflation rates extends beyond academic interests, impacting essential aspects of economic governance and financial decision-making Zhang and Qi (2005). Accurate inflation predictions are vital for policymakers in crafting monetary policies that ensure economic stability and sustainable growth. In the realm of business and investment, understanding the intricacies of inflation trends is crucial for strategic planning and risk assessment.

Traditionally, the forecasting of inflation has relied on time series models, with the Autoregressive Integrated Moving Average (ARIMA) model, as proposed by Box and Jenkins, being a prevalent choice Box and Jenkins (1976). Despite the widespread application of ARIMA models, their linear structure can be inadequate for economies exhibiting nonlinear patterns, a scenario often observed in countries like Sri Lanka McGinnity and Li (2014).

The emergence of Artificial Neural Networks (ANNs) has introduced a versatile tool capable of modeling nonlinear dynamics inherent in economic data Haykin (2009). The challenge, however, lies in integrating these two approaches—ARIMA and ANNs—to effectively handle both linear and nonlinear patterns in economic data Zhang (2003). While hybrid models, combining ARIMA and ANNs, have been explored as a solution, their application has revealed certain limitations, necessitating further research and refinement Cadenas and Rivera (2009).

Addressing these limitations, our study proposes an innovative hybrid model that merges an auto-tuned algorithm with the traditional ARIMA-ANN framework. This model is specifically designed to capture the complex dynamics of inflation in Sri Lanka, offering a theoretically robust and practically applicable forecasting tool. The performance of this model is evaluated using historical inflation data from Sri Lanka (1988–2021), implemented through the R software package and the Keras library, reflecting the latest advancements in computational tools Team (2022); Chollet et al. (2018).

Our findings illuminate the predictive strengths and broader implications of our hybrid model in understanding inflation trends in transitional economies like Sri Lanka. The study underscores the relevance and applicability of our approach to a range of stakeholders, including policymakers, business leaders, and investors. The insights gleaned from our research contribute to the existing literature by providing a novel perspective on economic forecasting in a specific context, thereby enriching the discourse on inflation prediction methodologies.

In conclusion, this research not only builds on the foundational work in the field of economic forecasting but also integrates recent technological advancements, creating a unique contribution to the study of inflation in transitional economies. The study's significance is rooted in its context-specific approach and methodological innovation, offering new insights into the complexities of economic forecasting in environments like Sri Lanka.

2 Methodology

Augmented Dickey-Fuller (ADF) Test

The ADF test plays a pivotal role in the initial stages of time series analysis, particularly in evaluating the stationarity of the data. Developed by Dickey and Fuller (1979) and later extended by Said and Dickey (1984), the ADF test assesses the null hypothesis (H_0) that a unit root is present in a time series sample, indicating non-stationarity, against the alternative hypothesis (H_1) of the absence of a unit root, signifying stationarity.

Brock-Dechert-Scheinkman (BDS) Test)

The Brock-Dechert-Scheinkman (BDS) Test, introduced by Brock et al. (1996)., is employed for identifying nonlinearity within the chaos theory framework. Initially designed for testing independence and identical distribution (iid), the BDS test proves versatile for detecting both linear and nonlinear structures. Additionally, it is applied to residuals from a fitted model for portmanteau or misspecification tests.

BDS Test Procedure: Integral Correlation: Evaluate how frequently temporal patterns recur in the data. Define time series X_t and its m-history X_t^m. Compute the m-dimensional correlation integral, $C_{m,T}(\epsilon)$.

$$C_{m,T}(\epsilon) = \sum_{i<s} I_\epsilon(X_t^m, X_s^m) * \left\{ \frac{2}{T_m(T_m - 1)} \right\} \tag{2.1}$$

Correlation Integral Interpretation: Determine the frequency of m-histories in proximity within a hypercube. Compute the probability that any two m-dimensional points are near each other.

$$P(\mid X_t - X_s \mid < \epsilon, \mid X_{t-1} - X_{s-1} \mid < \epsilon, ..., \mid X_{t-m+1} - X_{s-m+1} \mid < \epsilon) \tag{2.2}$$

BDS Statistics: Quantify nonlinearity using BDS statistics.

$$V_{m\epsilon} = T^{1/2} * \frac{C_{m,T}(\epsilon) - C_{1,T}(\epsilon)}{S_{m,T}} \tag{2.3}$$

3 Data Set

In our research, we examine the data on Sri Lanka's monthly inflation rate from January 1988 through August 2021 TradingView (2023). Figure 01 depicts these inflation rate data.

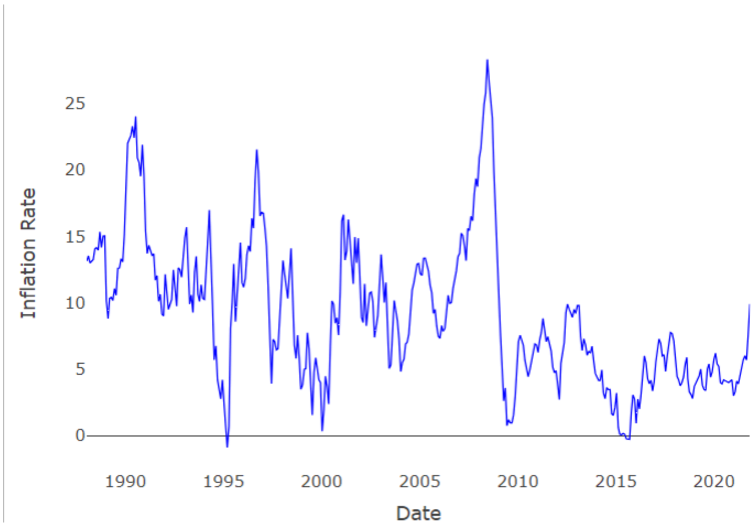

Figure 1: Monthly mean Inflation rate of Sri Lanka(1988-2021)

The dataset utilized in this study encompasses a total of 407 data points, exhibiting an average of 9.61, variance of 29.91554, a minimum value of -0.890, and a peak value of 28.310. Initial scrutiny through the Dickey-Fuller stationarity test confirmed the dataset's stationary nature. The time series representation manifests a stochastic pattern, occasionally marked by outlier points, including a notable instance in June 2008, likely influenced by the nation's political climate at that juncture. The dataset is partitioned into three segments: training data spanning from January 1988 to February 2020, instrumental in assessing modeled fits' precision; test data covering March 2020 to February 2021; and validation data ranging from March 2021 to November 2021, facilitating the evaluation of projected values' accuracy.

4 Models

4.1 Auto Regressive Integrated Moving Average Model (ARIMA)

The ARIMA model Box and Jenkins (1976) extends the ARMA model to address non-stationarity. In ARIMA models, a non-stationary time series is transformed into a stationary one by implementing finite differences to the data points. The mathematical representation of the ARIMA(p, d, q) model using lag polynomials is provided as follows:

$$(1-\sum_{i=0}^{p}\phi_i L^i)(1-L)^d y_t = (1+\sum_{j=0}^{q}\theta_j L^j)\epsilon_t \qquad (4.1)$$

Here, p, d, and q are integers greater than or equal to zero, representing the autoregressive, integrated, and moving average components, respectively. The integer d determines the degree of differentiation. When d is zero, the model simplifies to an ARMA(p,q) model. An ARIMA(p,0,0) corresponds to an AR(p) model, while ARIMA(0,0,q) aligns with an MA(q) model.

For our study, the ARIMA model was applied to the inflation data. The dataset was tested for stationarity using the Dickey-Fuller test, and the best-fitting ARIMA model with parameters (p, d, q) was determined based on the lowest AIC. The model was then trained using the identified parameters, and its performance was evaluated on both training and test datasets. The forecasting horizon was extended to assess the model's predictive capabilities.

4.2 Artificial Neural Networks (ANN)

The single hidden layer feed-forward neural network is a prevalent approach for time series modeling and forecasting Haykin (1999). The architecture comprises three tiers of simple processing units interconnected by acyclic links. A mathematical connection exists between the output X_t and the inputs $X_{t-1}, X_{t-2}, ..., X_{t-p}$:

$$X_t = \alpha_0 + \sum_{j=1}^{p} \alpha_j H \left(\beta_0 + \sum_{i=1}^{q} \beta_{i,j} X_{t-i} \right) + \epsilon_t \tag{4.2}$$

Here, α and β are model parameters, p is the count of input nodes, and q symbolizes the number of hidden layers. The logistic function $H(t)$, typically employed as the transfer function for the hidden layer, is given as:

$$H(X) = \frac{1}{1 + \exp(-X)} \tag{4.3}$$

Equations (4.2) and (4.3) demonstrate that the ANN model performs a non-linear function, mapping previous observations $(X_{t-1}, X_{t-2}, ..., X_{t-p})$ to future values (X_t). This can be expressed as:

$$X_t = f(X_{t-1}, X_{t-2}, ..., X_{t-p}, \boldsymbol{V}) + \epsilon_t \tag{4.4}$$

where V is a vector of all parameters, and f is determined by the network structure and the weights of connections. Therefore, the neural network corresponds to a nonlinear autoregressive model.

However, it is crucial to note that the stages of estimating the neural network model, starting from the training and testing phases, were not explicitly explained. The logistic function should not be assumed before thorough network training. This process involves optimizing the weights and biases to enhance the model's predictive capabilities. The training phase is followed by testing on independent data to assess the model's generalization performance.

4.3 Reiteration of G.Peter Zhang's Model(PZM)

G. Peter Zhang's hybrid model(PZM)Zhang (2003) harnesses the combined power of ARIMA and ANN models to enhance forecasting performance. The model initially applies an ARIMA model to scrutinize the linear components of the data, then an ANN model is developed to model the residuals generated from the ARIMA model, which encapsulate information about the data's non linearity.

By separately modeling linear and nonlinear patterns and integrating the forecasts, the hybrid model can bolster overall modeling and forecasting precision. Interestingly, the hybrid method may incorporate sub-optimal models, which paradoxically can enhance the model's utility as it's often more effective to amalgamate forecasts derived from varying information sets.

4.4 The Proposed ARIMA-ANN Hybrid Model (HB)

In accordance with Zhang et al. (2008), it is suggested that a time series is a mixture of a linear component and a non-linear component. Therefore, we can represent a time series X at time t as:

$$X_t = L_t + N_t \tag{4.5}$$

Here, L denotes the linear part of the time series X at time t, while N stands for the non-linear part. In the preliminary application of the ARIMA model to X, the non-linear part of X can be estimated as follows:

$$N_t = X_t - L_t^{'} \tag{4.6}$$

Where $L_t^{'}$ is the estimated linear part of time series X at time t. In this study, 12 data points ahead were chosen as our test data length, and an "Akaike Information Criteria" based auto-tuned ARIMA model was used to forecast ten data points ahead using training data lengths ranging from 12 to 386, leveraging the power of Algorithm 2. The initial ten data points were omitted as the ARIMA model requires some data to make its first prediction. The advantage of this method is that it increases the length of training data, ensuring successful estimation of N_t. As pointed out by Faraway and Chatfield **?**, the accuracy of ANN models is heavily dependent on training data length. N_t is estimated at each step as follows:

$$N_t = [X_{t_1}, X_{t_2}....X_{t_{12}}] - [L_{t_1}, L_{t_2}....L_{t_{12}}] \tag{4.7}$$

Let X_t-L_t=ϵ_t and denote N_t as:

$$N_{352\times 10} = \begin{bmatrix} \epsilon_{1,1} & \epsilon_{1,2} & \cdots & \epsilon_{1,h} \\ \epsilon_{2,1} & \epsilon_{2,2} & \cdots & \epsilon_{2,h} \\ \vdots & \vdots & \ddots & \vdots \\ \epsilon_{L,1} & \epsilon_{L,2} & \cdots & \epsilon_{L,h} \end{bmatrix} \tag{4.8}$$

We intend to model N_t based on univariate ANN modeling techniques and transform $N_{386\times 12}$ to $N_{1\times 4632}$:

$$N_t^{'} = (N_t)^T \tag{4.9}$$

By treating $N_t^{'}$ as a univariate time series, we can model the non-linear component of X_t. Finally, we can modify [equation 6] as follows:

$$X_t = L_t^{'} + N_t^{'} + e_t \tag{4.10}$$

Here, e_t represents the error term generated from estimating N_t using ANN modeling processes. The objective of making an accurate forecast for this task (Algorithm 3) is to minimize e_t.

4.5 Error Calculation Methods

Let n denote the number of fitted points, Y_t represent the actual value of the response variable Y at time t, and $\hat{Y}_t$ correspond to the predicted value of Y_t.

4.5.1 Mean Absolute Percentage Error (MAPE)

The Mean Absolute Percentage Error (MAPE) can be computed using the formula provided below.

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{|Y_t - \hat{Y}_t|}{Y_t}.$$

MAPE is the most commonly used measure of the forecast error and it works best in the absence of an extremes values in the data set.

5 Simulation Results

5.1 Pseudo Code for Optimal $ARIMA$ Order

Algorithm 1: Pseudo Code for Time Series Stationarity Check and Optimal $ARIMA$ Order Selection

Input: X, p_max, d_max, q_max
Output: Optimum $ARIMA$ Order

- Perform ADF test on time series data, rejecting the null hypothesis if stationarity is observed.
- If stationarity is confirmed, proceed to optimal $ARIMA$ order selection.
- Otherwise, apply necessary transformations to achieve stationarity.
- For $p \longleftarrow$ 0 $to\ p\_max$
- For $d \longleftarrow$ 0 $to\ d\_max$
- For $q \longleftarrow$ 0 $to\ q\_max$
- Try:
  model $\leftarrow$ fit(arima(p,d,q))
- current_AIC $\leftarrow$ $AIC(model)$
- If $current\_AIC < AIC$
- $AIC \leftarrow Optimum\ ARIMA\ Order$
- optimum_model $\leftarrow$ model
- Cat: $error$ = function(e)
- End If
- End For
- **return** $Optimum\ ARIMA\ Order(p, d, q)$

The algorithm detailed here delineates the procedure for determining the ideal order of an ARIMA model, a critical component of time series forecasting. The ARIMA model necessitates three parameters: p (autoregressive part's order), d (degree of first differencing involved), and q (moving average part's order). The algorithm launches a sequence of loops for each parameter, within the bounds of 0 to their respective maximum limits (p_{max}, d_{max}, and q_{max}). It then attempts to fit the ARIMA model with each parameter combination. The Akaike Information Criterion (AIC) for each model is computed and contrasted with the best (lowest) AIC noted so far. If a model results in a lower AIC, it supplants the current optimal model. Ultimately, the algorithm returns the parameters of the ARIMA model that produced the lowest AIC, delivering the optimal order for the ARIMA model for the provided time series data.

5.2 Pseudo Code for Proposed Artificial Neural Network (ANN) Model

Algorithm 2: Pseudo Code for Proposed ANN Model

Input: D-Data set, H-test data length, L-learning rate, Network structure
Output: Trained model

- Initialize all weights and biases for the artificial neural network.
- For each training tuple X in D:
  - For each input layer unit j:
    - $O_j = I_j;$
    - $I_j = \Sigma_i W_{i,j} O_i + \Theta_j$
    - $O_j = \frac{1}{1+e^{-1}j};$
  - For each unit j in the output layer:
    - $E_j = O_j(1 - O_j)(T_j - O_j)$
  - For each unit j in the hidden layers, from the last to the first hidden layer:
    - $E_j = O_j(1 - O_j) \sum E_k W_{j,k}$
  - For each weight $W_{i,j}$ in the network:
    - $\Delta W_{i,j} = L(E_j O_i);$
    - $W_{i,j} = Wi, j + \Delta W_i;$
  - For each bias $\Theta_j$in the network:
    - $\Delta\Theta_j = L(E_j);$
    - $\Theta_j = \Theta_j + \Delta\Theta_j;$
- End For
- Return the trained model.

The pseudo-code outlines the process for training the proposed Artificial Neural Network (ANN) model. It involves initializing weights and biases, performing forward and backward passes, adjusting parameters based on the learning rate, and updating the model iteratively. The resulting trained ANN model can be used for time series forecasting tasks.

5.3 Pseudo code for proposed hybrid model

Algorithm 3: Pseudo Code for Hybrid Model

Input: D-Data set, p_max, d_max, q_max , H-test data length, L-the learning rate, K-Initial index train $ARIMA$ model to, Network structure
Output: Trained model

- Perform ADF test on time series data, rejecting the null hypothesis if stationarity is observed.
- If stationarity is confirmed, proceed to optimal $ARIMA$ order selection using the previously provided pseudo-code.
- Otherwise, apply necessary transformations to achieve stationarity.
- Initialize all weights and biases for the artificial neural network.
- For each training tuple X in D:
  - For each input layer unit j:
    - $O_j = I_j$;
    - $I_j = \Sigma_i W_{i,j} O_i + \Theta_j$
    - $O_j = \frac{1}{1+e^{-1}j}$;
  - For each unit j in the output layer:
    - $E_j = O_j(1 - O_j)(T_j - O_j)$
  - For each unit j in the hidden layers, from the last to the first hidden layer:
    - $E_j = O_j(1 - O_j) \sum E_k W_{j,k}$
  - For each weight $W_{i,j}$ in the network:
    - $\Delta W_{i,j} = L(E_j O_i)$;
    - $W_{i,j} = Wi, j + \Delta W_i$;
  - For each bias $\Theta_j$in the network:
    - $\Delta\Theta_j = L(E_j)$;
    - $\Theta_j = \Theta_j + \Delta\Theta_j$;
- End For
- Return the trained model.

The proposed hybrid model iteratively applies the ARIMA model with the optimal order selected based on stationarity checks and simultaneously employs an artificial neural network for further refinement, ensuring a robust and accurate forecasting model. Algorithm 2 shows that the pseudo code for proposed model. Hear H is the test data length and K is the initial index of univareate data set and integer H needs minimum values to train the $ARIMA$ model. In this case we selected 10 data points to train the $ARIMA$ model.

5.4 Augmented Dickey-Fuller Test

The Augmented Dickey-Fuller test was performed to assess the stationarity of the inflation training data. The test yielded a test statistic of -4.8804 with a lag order of 7 and a p-value of 0.01. The alternative hypothesis suggests stationary.

```
data:  inflation_data 1988-2021
Dickey-Fuller = -4.8804, Lag order = 7, p-value = 0.01
alternative hypothesis: stationary
```

The small p-value (0.01) provides evidence against the null hypothesis and data set is stationary.

5.5 BDS Test Results

m(Embedding dimension) = 2, 3, 4
Epsilon for close points = 2.7348, 5.4695, 8.2043, 10.9390

Table 1: Results of BDS test (Applying BDS test for Data set)

m	[2.7348]	[5.4695]	[8.2043]	[10.939]
[2]	p <2.2e-16	p <2.2e-16	p <2.2e-16	p <2.2e-16
[3]	p <2.2e-16	p <2.2e-16	p <2.2e-16	p <2.2e-16
[4]	p <2.2e-16	p <2.2e-16	p <2.2e-16	p <2.2e-16

Table 1 shows that results of BDS test. According to those results our data set is not a indentation identical distribution (iid). Because P-values of close points are close to zero in first four Embedding dimension and we can reject null hypothesis.

5.6 Predicted values for test data

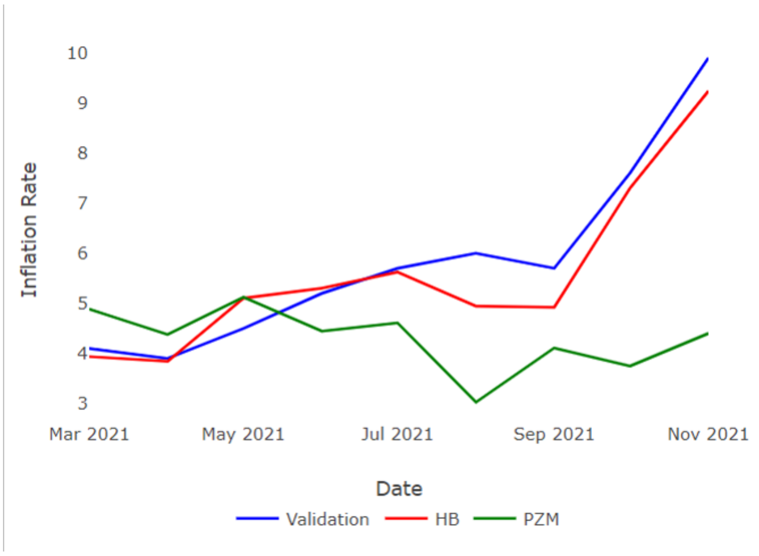

Figure 2: Predicted data comparison with Test Data

The displayed plot(Figure 02) is a multi-line chart representing the Inflation Rate predictions made by four different models – Test, HB and PZM, over a period from March 2020 to February 2021. Each line, distinguished by a unique color, represents a specific model's prediction trajectory over the specified period. By observing the line trends, one can gain insights into the behavior of each model. For instance, a line consistently above others indicates a model with a higher inflation rate prediction. Conversely, a jagged line with noticeable peaks and troughs suggests volatile model predictions with significant fluctuations over time. Thus, this plot offers a comprehensive visual depiction of the models' performances, allowing for an easy comparison and identification of patterns in their respective predictions.

Table 2: Test data comparison of predicted data(Applying HB model to test data and PZM models)

Date	Test	HB	PZM
Mar-2020	5.4	5.44	5.96
Apr-2020	5.2	5.15	5.29
May-2020	4	3.97	4.22
Jun-2020	3.9	3.98	3.50
Jul-2020	4.2	4.04	4.38
Aug-2020	4.1	4.08	4.16
Sep-2020	4	3.86	4.29
Oct-2020	4	4.07	4.19
Nov-2020	4.1	4.06	3.23
Dec-/2020	4.2	4.15	4.41
Jan-2021	3	2.90	3.43
Feb-2021	3.3	3.29	3.53
MAPE	-	1.60	7.83

The Table 02 presents the predicted inflation rates(test data) by HB and PZM,across a time frame from March 2020 to February 2021. Each row corresponds to a specific month, while each column signifies the inflation rate predicted by a particular model for that month. It allows for a thorough examination of the individual model's performance over time by comparing the data vertically within each column.

Most notably, the table includes the Mean Absolute Percentage Error (MAPE) at the end, a key indicator of the accuracy of the model's predictions. The MAPE for HB, and PZM, is provided as 1.60% and 7.83%, respectively. These MAPE values serve as benchmarks for model comparison, with lower values signifying higher predictive accuracy. For instance, with a MAPE of 1.60%, the HB model exhibits the most accurate predictions among the two models in this case.

In summary, this table not only provides a granular view of each model's monthly predictions but also quantitatively compares their performance via the MAPE metric. This approach, in conjunction with the visual comparison offered by the line graph, facilitates a comprehensive evaluation of the models' inflation rate predictions.

5.7 Froecasted values for validation data

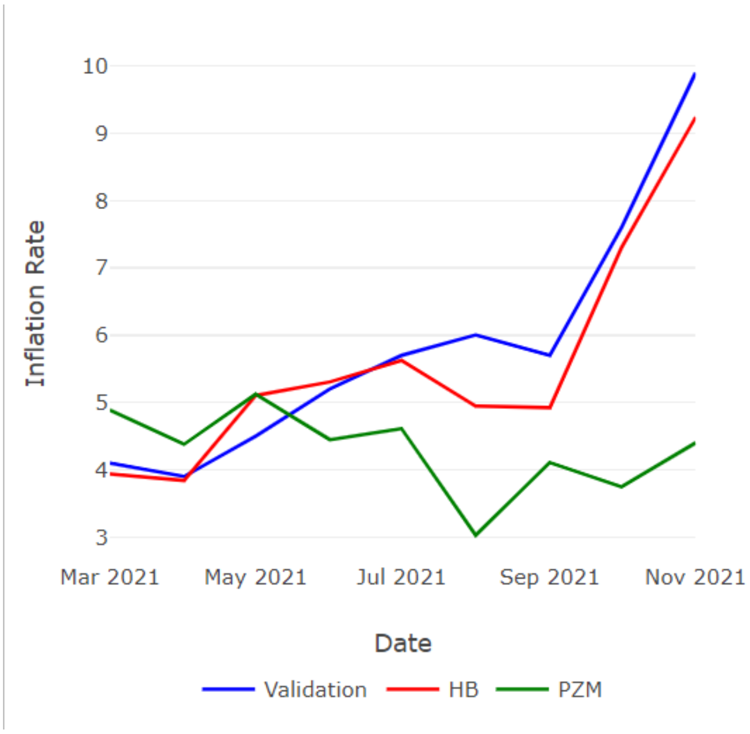

Figure 3: Validation Data with Forecasted values

The Figure 03 depicted here showcases the inflation rate forecasts from March 2021 to November 2021 using two different models: HB and PZM, . The blue line represents the validation data, which serves as the actual observed data against which the forecasts are validated. The other lines - red for HB, green for PZM, the forecasts made by each respective model. Over the given time period, it is evident that all two models exhibit varying degrees of deviation from the validation data. For instance, the PZM model (Green line) demonstrates a higher deviation in later months, whereas the HB model (Red line) appears to track relatively closer to the actual data.

Table 3: Validation data comparison of forecasted data(Appling Traned models to validation data PZM and HB)

Date	Validation	HB	PZM
Mar-2021	4.1	3.93	4.88
Apr-2021	3.9	3.84	4.37
May-2021	4.5	5.10	5.12
June-2021	5.2	5.30	4.44
Jul-2021	5.7	5.62	4.60
Aug-2021	6	4.94	3.02
Sep-2021	5.7	4.92	4.10
Oct-2021	7.6	7.30	3.74
Nov-2021	9.9	9.23	4.40
MAPE	-	7.10	27.17

This table presents a comparison of two forecasting models (HB, PZM) against the actual validation data for a period from March 2021 to November 2021. For each model, the Mean Absolute Percentage Error (MAPE) is also provided, which serves as an indicator of the model's accuracy.

The HB model has the lowest MAPE at 7.10, which indicates that on average, the model's predictions are approximately 7.10% off from the actual values. The PZM model's MAPE is 27.17, suggesting that its predictions deviate by about 27.17% from the actual values. These MAPE values can be useful for determining which model provides the most accurate predictions for the given data. Based on this table, the HB model appears to provide the most accurate forecasts, with the lowest average deviation from the actual validation data.

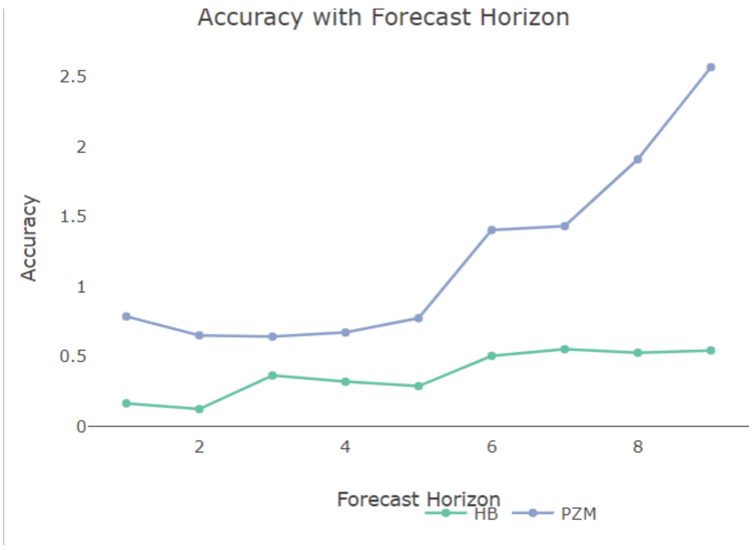

Figure 4: RMSE Change with forecast horizon -Validation data

The RMSE plot further substantiates the assessment of model performance by visualizing the variation of error as a function of the forecast horizon. As illustrated in the plot, the RMSE of all models tends to increase with forecast length, indicating a reduced predictive accuracy over time. The PZM model (Blue line) shows the highest increase, corroborating the comparatively higher MAPE in the earlier table. Conversely, the HB model (Green line) presents a more steady error increase, affirming its superior forecast precision. This graphical representation thus provides an intuitive means to compare forecast model robustness over an expanding forecast horizon.

6 Conclusion and Discussion

This study undertook a comparative evaluation of two distinct forecasting models: the newly proposed Hybrid Model (HB) and the existing PZM model. Utilizing real-world validation data, the study's primary focus was on assessing the models' performance through the Mean Absolute Percentage Error (MAPE), a widely recognized metric in forecasting accuracy.

The HB model, central to our research, exhibited superior performance by recording the lowest MAPE at 7.10%. This figure indicates a minimal deviation from the actual validation data, underscoring the model's accuracy and reliability in forecasting inflation rates. In contrast, the PZM model manifested a higher MAPE of 27.17%, indicating a larger deviation from the actual values.

It is crucial to acknowledge that while the HB model showed promising results in this specific context, the efficacy of forecasting models can be contingent on various factors inherent to the data set, such as noise levels, seasonality, and underlying trend patterns. This observation highlights the importance of choosing a model that aligns with the specific characteristics and requirements of the data at hand.

Our findings lay the groundwork for future research endeavors aimed at enhancing forecasting accuracy. Prospective studies could explore the integration of advanced methodologies, including machine learning or deep learning techniques, and the amalgamation of multiple forecasting models to achieve greater predictive precision. Furthermore, additional validation of the HB model across diverse data sets would be instrumental in affirming its robustness and applicability in various economic contexts.

In conclusion, the research contributes significantly to the field of economic forecasting, particularly in the context of developing economies like Sri Lanka. The efficacy of the HB model in accurately predicting inflation rates represents a notable advancement in forecasting methodologies. This study not only enriches the academic discourse in economic forecasting but also provides practical insights for policymakers and economists in data-driven decision-making processes.

References

Box, G. E. and Jenkins, G. M. (1976). *Time series analysis: forecasting and control*. Holden-Day.

Brock, W. A., Dechert, W. D., and LeBaron, B. (1996). The bds statistic as a test for the adequacy of a linear time series model. *Journal of Time Series Analysis*, 17:215–235.

Cadenas, E. and Rivera, W. (2009). Wind speed forecasting in three different regions of mexico, using a hybrid arima-ann model. *Renewable Energy*, 34:2732–2738.

Chollet, F. et al. (2018). Keras: The python deep learning library.

Fraiha Lopes, R. L., Fraiha, S. G. C., Lima, V. D., Gomes, H. S., and Cavalcante, G. P. S. (2020). Hybrid arima and neural network modelling applied to telecommunications in urban environments in the amazon region. *International Journal of Antennas and Propagation*.

Hajirahimi, Z. and Khashei, M. (2022). A novel parallel hybrid model based on series hybrid models of arima and ann models. *Neural Processing Letters*.

Haykin, S. (1999). *Neural Networks: A Comprehensive Foundation*. Prentice Hall.

Haykin, S. (2009). *Neural Networks and learning machines*. Pearson Education, 3 edition.

McGinnity, S. M. S. C. T. M. and Li, Y. (2014). A hybrid forecasting approach using arima models and self-organising fuzzy neural networks for capital markets. *The 2013 International Joint Conference on Neural Networks*.

Team, R. C. (2022). R: A language and environment for statistical computing.

TradingView (2023). Economics-lkiryy.

Zhang, G., Patuwo, B. E., and Hu, M. Y. (2008). Forecasting with artificial neural networks: The state of the art. *International Journal of Forecasting*, 14:35–62.

Zhang, G. P. (2003). Time series forecasting using a hybrid arima and neural network model. *Neurocomputing*, 50:159–175.

Zhang, G. P. and Qi, M. (2005). Are arima neural network hybrids better than single models. *International Joint Conference on Neural Networks*.