International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 04 Issue: 04 | April – 2025 DOI: 10.55041/ISJEM02631
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

Time Series Forecasting using Hybrid Holt's ESM with Damping Parameter and Neural Network Model

P R Ani Rudhran[1], Nimitha John[2]

[1]PG Student, Department of Statistics and Data Science, CHRIST (Deemed to be university), Bengaluru

[2]Assistant Professor, Department of Statistics and Data Science, CHRIST (Deemed to be university), Bengaluru

Abstract

Time series forecasting has gained much importance in various fields, and the choice of an appropriate model for forecasting is still a significant challenge. Holt's ESM with damping trend is widely used for modeling linear trends, and ANNs are effective in capturing complex nonlinear relationships. However, limitations arise when these models are implemented separately. Hybrid models in time series forecasting are used to combine strengths of different models to cover their shortcomings. In this paper, we develop a hybrid forecasting model that combines Holt's ESM with damping trend and Artificial Neural Network. Holt's ESM with damping trend captures the linearity in the time series, while the non-linearity in the residuals is captured by ANN. Results based on simulation and real-world data show that the hybrid model outperforms the component models.

Keywords: ESM, Damping trend, Artificial neural networks, Time series forecasting, Hybrid forecast

1. Introduction

Time series forecasting is a significant task in many areas, such as economics, finance, engineering, and numerous other scientific disciplines. The aim of time series analysis is to construct a model that reflects the current temporal patterns of data. It enables future accurate predictions based on historical observations. One of the biggest challenges in time series forecasting is deciding on the right modeling strategy, particularly when the process under which the data is generated has a complex nature. Statistical modeling is possible with historical data, but the challenge usually comes in establishing whether the time series has linear, nonlinear, or mixed patterns, and thus model selection becomes a prominent challenge.

Over the years, different models have been proposed to do the task of forecasting future observations of a time series. Exponential Smoothing Models (ESM) is a family of time series models widely used for the task and Artificial Neural Networks (ANN) is another such one. A particular model in the ESM family namely Holt's Exponential Smoothing Model with damping trend is used to model trends in time series data, especially in cases where the trend is expected to diminish over time. However, while this model is effective for handling linear trends, it performs relatively poorly with capturing nonlinear structures that may exist in real-world time series data.

Artificial Neural Networks (ANNs), on the other hand, are highly effective at capturing nonlinear relationships in data due to their flexible, data-driven nature. ANNs do not require any specific model form and adapt based on the features presented in the data. However, while ANNs can excel at modeling nonlinear components, they often face challenges when dealing with purely linear time series or those with complex autocorrelations, leading to mixed results when applied alone.

In this paper, we introduce a hybrid forecasting model that integrates the advantages of Holt's Exponential Smoothing Model with damping trend and Artificial Neural Networks (ANNs). The rationale behind the hybrid model is to leverage the capability of Holt's model to capture linear trend and utilize ANNs to capture the nonlinear elements of the time series. This blend of models enables a stronger solution, one that meets the demands of real-world time series, which tend to exhibit both linear and nonlinear trends. The hybrid model seeks to enhance the accuracy of forecasting by taking the best from each approach and presenting a more holistic solution to time series analysis.

The rest of the paper is organized as follows. Section 2 reviews the Holt's Exponential Smoothing Model and Artificial Neural Networks in the context of time series forecasting. In Section 3, we introduce the hybrid approach and discuss its implementation. Section 4 presents empirical results from real-world datasets, comparing the performance of the hybrid model with individual models. Finally, Section 5 concludes the study and outlines directions for future research.

2. Time Series Forecasting Models

Time series forecasting provides an estimate of future values from past observations of a variable based on patterns or trends in the data. Classical time series models including moving averages, exponential smoothing, and ARIMA are popular because they are easy to interpret and understand. These models, however, presume linearity, and there may be instances where nonlinear and complex time series data is involved. Nonlinear models, i.e., Artificial Neural Networks (ANNs), have been proposed to overcome this deficiency, with more flexibility and accuracy. Also, hybrid models that take the best from several forecasting methodologies have yielded encouraging results in improving the accuracy of forecasts.

2.1. Holt's Exponential Smoothing

Holt's Exponential Smoothing, developed by (Holt, 1957) is a development of simple exponential smoothing (SES) that can cope with time series data that includes trends. It is specifically valuable for series with changing level and trend over time. It extends SES, which smoothes the data using exponentially decreasing weights for previous observations. Holt's approach uses two primary components: the level and the trend.

The Holt's Exponential Smoothing model can be expressed as:

$$\hat{y}_{t+1} = \ell_t + b_t$$

$\hat{y}_{t+1}$ is the forecast for the next time period $t + 1$. ℓ_t is the level component at time t and b_t is the trend component at time t.

The updating equations for the level and trend components are as follows:

$$\ell_t = \alpha y_t + (1 - \alpha)(\ell_{t-1} + b_{t-1})$$

$$b_t = \beta(\ell_t - \ell_{t-1}) + (1 - \beta)b_{t-1}$$

Here, α and β are smoothing parameters for the level and trend, respectively and y_t is the actual observed value at time t.

The model presumes that the trend is linear and is adjusted according to the new level and trend values. Holt's model is computationally effective due to its simplicity but presumes that the trend is constant, which is not always true in most applications.

2.2. The Damping Parameter

In real-world applications, the constant trend assumption may not always be valid, particularly when growth rates decline with time. For this reason, a damping parameter is added to Holt's model by (Gardner, 1985) to enable the trend to decay over time instead of persisting indefinitely. The damping parameter actually "dampens" the trend, and thus the model is more realistic where trends are anticipated to slow down or stabilize.

The modified forecast equation with a damping parameter ϕ is:

$$\hat{y}_{t+h} = \ell_t + \phi^h b_t$$

where h is the forecast horizon, and ϕ is the damping parameter ($0 \leq \phi \leq 1$). When $\phi = 1$, the model reduces to the standard Holt's Exponential Smoothing. When $\phi < 1$, the trend decays over time, gradually becoming less influential in the forecast.

The damped trend method has been shown to provide better forecasting performance for time series where trends only last for limited periods, such as in market saturation or maturing economies. (Taylor, 2003) considered the trend to be multiplicative to modify the model. The original model was further developed by (McKenzie & Gardner, 2010).

2.3. Artificial Neural Networks (ANNs)

There are situations where there is high nonlinearity in the data and the linear models because of their linear constraints can not capture the pattern in the data well. Artificial Neural Networks (ANNs) are employed in this situation because they can approximate various non linearities in the data. The principal strength of ANNs is that they can approximate any continuous function, and thus are extremely flexible and able to learn complex patterns in data. Even Recurrent Neural Networks (RNNs) are popularly used in time series forecasting as discussed by (Hewamalage, Bergmeir, & Bandara, 2021)

For forecasting time series, a common ANN structure employed is the feedforward neural network with one or more hidden layers. The network receives historical values of the time series as inputs and generates a forecast as the output. The relationship between inputs and outputs is expressed as:

$$y_t = f(Wx_t + b)$$

where: y_t is the output (forecast), W is the matrix of weights connecting the input to the hidden layer, x_t is the vector of past values of the time series, b is the bias term, f is the activation function applied to the weighted sum of inputs (typically a sigmoid or ReLU function).

The network is learned in a supervised learning algorithm, where the objective is to reduce the error in prediction between the output of the network and the actual values. This is typically performed by means of a process known as backpropagation, wherein weights are updated with an optimization algorithm (e.g., gradient descent).

ANNs are susceptible to overfitting, especially when the model is more complex than the given data, and thus regularization and judicious choice of network architecture are crucial.

2.4. Hybrid Models

In time series prediction, there is no one model that can solve for all of the complexity that real-world data will pose. Linear models, like ARIMA (AutoRegressive Integrated Moving Average), work extremely well for time series data that have linear relationships, but they do not work as well to identify more complicated nonlinear relationships. Conversely, nonlinear model structures, such as Artificial Neural Networks (ANNs), perform very well in finding complex patterns in data but could do poorly with purely linear data or if they need large amounts of data to prevent overfitting.

To get around these drawbacks, so-called hybrid models, where strengths of diverse forecast techniques are combined, have been suggested. The rationale behind hybrid modeling is that by combining linear and nonlinear models, various aspects of the underlying structure of the data—linear and nonlinear—can be described better. In applications, this method is most useful when the data has both linear trends and nonlinear features, a common situation in applied forecasting problems.

Hybrid models have been found to perform better than single models in most instances, as they harness the best of both linear and nonlinear models to enhance forecasting accuracy. Hybrid methods have been shown effective in time series forecasting by various studies. For example in (Zhang, 2003), the integration of ARIMA and ANN has been found to be effective in extracting both the linear and nonlinear patterns of time series data. In the paper by (Khashei, & Bijari, 2011), another novel hybridization of ARIMA and ANN is discussed. (Khandelwal, Adhikari, & Verma, 2015) use Discrete Wavelet Transform (DWT) as a preprocessing step to improve the existing model.(Hajirahimi & Khashei, 2019) provide an extensive review of the hybrid structures in time series forecasting.

3. Proposed hybrid framework

While existing research on hybrid models has explored various combinations of time series models, the specific combination of Holt's ESM with damping and ANN has not been extensively studied. We aim to explore the potential of combining these two models, focusing on their ability to model both the linear and non-linear components of time series data. The use of the damping parameter in Holt's ESM offers a unique advantage in capturing trends that slow down over time, which is often observed in real-world data such as economic indicators, stock market returns, and product demand.

The proposed hybrid model combines the strengths of Holt's ESM (for modeling the linear components of the time series) with ANNs (for capturing non-linear residual patterns). The general workflow is as follows:

1. **Step 1**: Use Holt's ESM (with damping) to forecast the time series and calculate the residuals (the difference between the actual values and the forecasted values).

$$e_t = y_t - \hat{y}_t$$

The actual value at time t is represented by y_t and e_t is the residual at corresponding time.$\hat{y}_t$ is the Holt's ESM (with damping) forecast at time t.

2. **Step 2**: Use the residuals from Holt's ESM (with damping) as the input to an ANN. The ANN is trained to learn the non-linear patterns in the residuals and provide additional forecasts to correct for these patterns.

3. **Step 3**: The final forecast is obtained by adding the residual forecasts from the ANN to the base Holt's ESM (with damping) forecast:

$$\hat{y}_t^{hybrid} = \hat{y}_t^{Holt} + \hat{e}_t^{ANN}$$

Here $\hat{y}_t^{hybrid}$ is the final hybrid forecast, $\hat{y}_t^{Holt}$ is the forecast from Holt's ESM (with damping) and $\hat{e}_t^{ANN}$ is the forecast from the ANN model applied to the residuals.

By combining these two models, we aim to leverage the strengths of both: Holt's ESM (with damping)for capturing the linear and trend components and ANN for modeling the non-linear residual patterns that remain unexplained by Holt's ESM.

4. Empirical Results

In this section, several simulated time series are used to demonstrate the effectiveness of the proposed hybrid model. A real-world dataset that exhibits a damping trend is further considered for checking the effectiveness of the proposed model.

4.1. Simulation study

In order to set up the simulation, it is necessary to understand the effect of each parameter on the generated time series. The alpha **(α)** parameter determines the influence of the most recent observation on the forecasted value. Therefore smaller values of alpha are considered to give more weightage to older observations and thus not deviate from the trend. The beta **(β)** parameter, especially in the model with a damping trend, determines how quickly the trend reaches its threshold. Smaller values of beta are considered in the simulation to slow the time reaching threshold. The phi value (ϕ) controls the damping of the trend, a value equal to one results in the model being Holt's ESM without damping trend. Therefore, values closer to one are considered in the simulation.

The sample size (n) ranged from 36 to 100, with increments of 2. The alpha (α) and beta (β) values varied between 0.1 and 0.3, with increments of 0.05. The damping parameter (ϕ) was adjusted between 0.95 and 0.98 in increments of 0.01, with an initial level (L_0) of 150, initial trend (T_0) of 5, and an error standard deviation (error_sd) of 1.

Based on the varying parameters, 3300 different time series were simulated for understanding the performance of the proposed hybrid model. In the simulation, the proposed hybrid framework's performance is observed in metrics like RMSE, MAE, and MAPE. The performance of the hybrid model is compared with **Holt's ESM**, **Holt's ESM with damping (DHM)**, and **Artificial Neural Networks (ANN)**. These parameters are used to generate the time series data using a custom function simulate_data(), which models both the level and trend components while incorporating noise.

4.2. Real-time data analysis

The dataset used in this study comprises the annual growth rate of the population of India from the year 1961 to 2023. This dataset, sourced from the World Bank's World Development Indicators, includes two variables - the year of observation, ranging from 1961 to 2023, and the annual percentage growth rate of the population in India for each corresponding year. The data used for modeling is the cumulative growth calculated from the yearly growth rate.
It has 63 observations representing each year. The final 12 observations are used to evaluate the model. The below figure shows the time series plot of considered data.

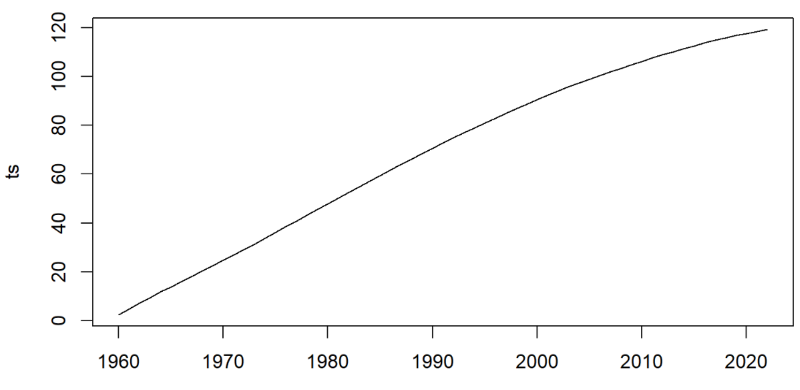

Fig. 1. Time series plot of cumulative population growth rate of India

We use cumulative population growth specifically to observe the damping effect in the time series. The results of forecasting should be transformed again in order to get the actual population growth rates for each year. For example, the population growth rate of 2020 is derived from subtracting the forecasted cumulative population growth rate of 2019 from that of 2020.

International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 04 Issue: 04 | April – 2025 DOI: 10.55041/ISJEM02631
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

4.3. Results

The results of the simulation were visualized through the RMSE, MAE, and MAPE curves for each model, where the x-axis represented the sample size.

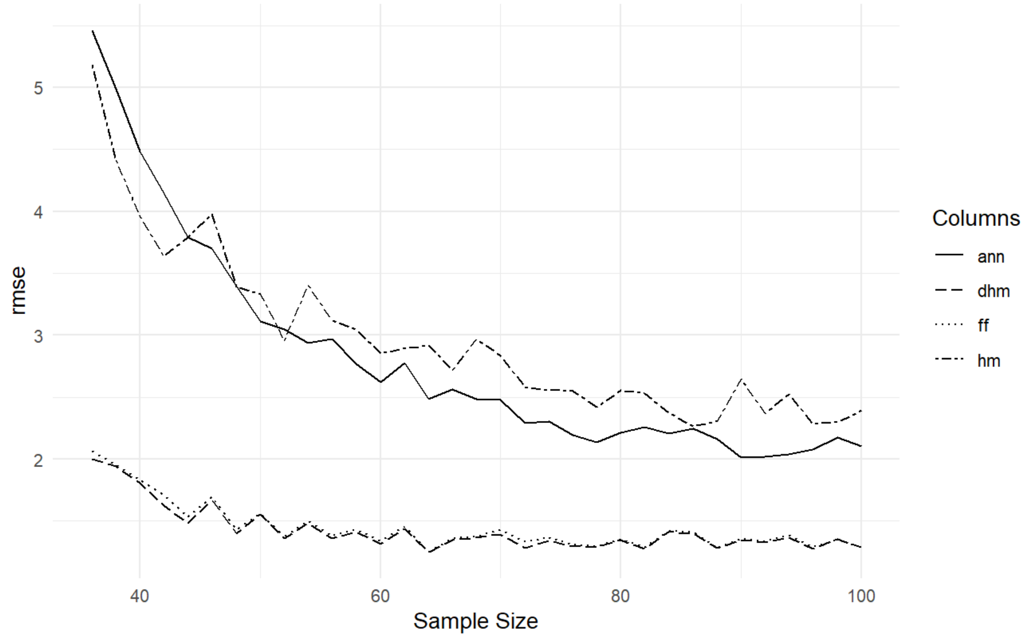

Fig. 2. RMSE values for different models

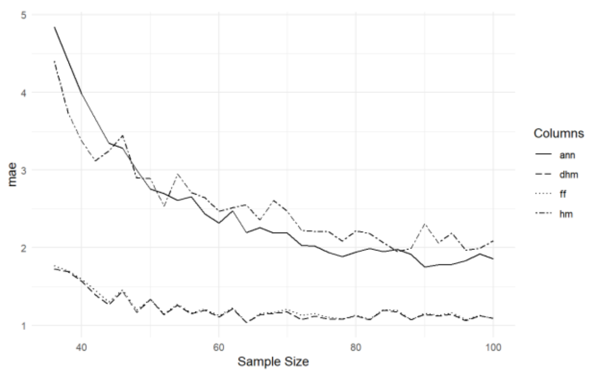

Fig. 3. MAE values for different models

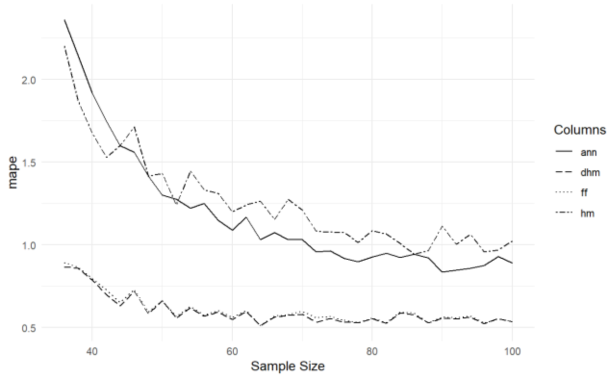

Fig. 4. MAPE values for different models

As shown in the figures, **Holt's ESM with damping** (DHM) and the **hybrid model (HM)** demonstrated nearly identical performance across all metrics, indicating that the integration of ANN with Holt's ESM with damping did not lead to significant improvements over the standalone DHM model. In contrast, **Holt's ESM without damping** (Holt's ESM) consistently performed the worst, followed by **ANN**, which showed higher errors and variability across the sample sizes. This confirms that the hybrid model provides competitive performance compared to traditional approaches, with the added flexibility of nonlinear pattern detection from the ANN, but with a slight improvement in the case of the damping trend model.

For the data analysis, a table containing the RMSE, MAE, and MAPE values for each model was formulated.

International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 04 Issue: 04 | April – 2025 DOI: 10.55041/ISJEM02631
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

Model	RMSE	MAE	MAPE
HM	1.9272163	1.4780727	1.2650187
DHM	0.8629191	0.6426342	0.5492760
ANN	5.8334145	5.1715322	4.4657145
Hybrid	0.8595685	0.6405405	0.5475264

Table 1. Evaluation metrics for data analysis

The findings, as seen above in the table, indicate that the hybrid model (HM) is ever so slightly better than Holt's ESM with damping (DHM), both of which report the lowest errors in all the given measures. Precisely, RMSE (Root Mean Square Error), MAE (Mean Absolute Error), and MAPE (Mean Absolute Percentage Error) for the hybrid model were similar to DHM but with slightly improved performance in error minimization. However, ANN did not do well, registering much larger error values, particularly in MAPE. This indicates that although the hybrid model has a minor advantage over Holt's ESM with damping, both perform far better than ANN for actual forecasting.

5. Conclusion

In this paper, we introduced a hybrid time series forecasting model combining Holt's Exponential Smoothing with damping (DHM) and Artificial Neural Networks (ANNs). The hybrid strategy utilized the linear pattern-capturing strength of Holt's model as well as the nonlinear pattern-identifying ability of ANNs, presenting a holistic solution to time series forecasting data containing linear and nonlinear attributes. Results through experiments illustrated how the hybrid model surpassed DHM and ANN as standalone techniques using lower error statistics in all measures of RMSE, MAE, and MAPE.

Competitive performance came from the hybrid model with small gains compared to DHM using forecasting precision as a criterion. The capacity of capturing intricate nonlinear patterns in residuals from Holt's ESM with damping also boosted the robustness of the model. In comparison, although ANN was promising to model nonlinear relations, its consistency was lower and much worse than the hybrid and DHM models. These results underscore the benefit of blending linear and nonlinear models for time series prediction, especially in actual data sets where there are linear trends and nonlinear residuals. The technique yields a more precise and dependable means of forecasting, complementing the drawbacks of standard approaches and neural networks alone.

ISSN: 2583-6129
DOI: 10.55041/ISJEM02631

References

Gardner Jr, E. S. (1985). Exponential smoothing: The state of the art. *Journal of Forecasting, 4(1), 1-28.*

Hajirahimi, Z., & Khashei, M. (2019). Hybrid structures in time series modeling and forecasting: A review. *Engineering Applications of Artificial Intelligence, 86, 83-106.*

Hewamalage, H., Bergmeir, C., & Bandara, K. (2021). Recurrent neural networks for time series forecasting: Current status and future directions. *International Journal of Forecasting, 37(1), 388-427.*

Holt, C. E. (1957). Forecasting seasonals and trends by exponentially weighted averages (ONR Memorandum No. 52). *Carnegie Institute of Technology, Pittsburgh USA, 10.*

Khandelwal, I., Adhikari, R., & Verma, G. (2015). Time series forecasting using hybrid ARIMA and ANN models based on DWT decomposition. *Procedia Computer Science, 48, 173-179.*

Khashei, M., & Bijari, M. (2011). A novel hybridization of artificial neural networks and ARIMA models for time series forecasting. *Applied Soft Computing, 11(2), 2664-2675.*

McKenzie, E., & Gardner Jr, E. S. (2010). Damped trend exponential smoothing: A modeling viewpoint. *International Journal of Forecasting, 26(4), 661-665.*

Taylor, J. W. (2003). Exponential smoothing with a damped multiplicative trend. *International Journal of Forecasting, 19(4), 715-725.*

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