INTERNATIONAL JOURNAL OF ENGINEERING MANAGEMENT AND SOCIAL SCIENCES (IJEMSS)

e-ISSN : 3139-065X

www.ijemss.com

editor@ijemss.com

Vol. 01, Issue 2, June2026, pp :1312-1323

# TIME SERIES FORECASTING: A COMPARATIVE STUDY OF CONVENTIONAL MODELS AND DEEP LEARNING APPROACHES

**[1]Dova Soumith [2]Dr. M. Sravani**

[1]PhD Scholar,Business Management ,Krishna University

[2]Assistant Professor ,Business Management ,Krishna University

## ABSTRACT

Time series forecasting has been a paramount issue in a variety of fields such as finance, weather prediction, energy consumption and even in healthcare. This paper includes a full comparative study of traditional models in statistics and deep learning structures to predict time series. We compare classical methods or Autoregressive Integrated Moving Average (ARIMA), Exponential Smoothing, and Seasonal ARIMA to state of the art neural network based methods, namely Long Short-Term Memory (LSTM) networks and Artificial Neural Networks (ANN). We use various practical data sets to evaluate the accuracy of the forecast, the efficiency or even efficiency of the model, and its robustness. Findings indicate that on linear patterns, the conventional models exceed the deep learning models, especially LSTM networks, but in the situation of nonlinear temporal patterns, deep learning models plate by far prevail over traditional models. The study is applicable to practitioners to make right choices in regards to the use of forecasting methodologies depending on the nature of the data, availability of resources and accuracy.

**Keywords:** Time Series Forecasting, ARIMA, LSTM, Artificial Neural Networks, Deep Learning, Comparative Analysis, Predictive Modeling

## INTRODUCTION

Time series prediction became one of the core analytical methods in the contemporary data science, with organisations capable of making informed choices through referencing past temporal results. This capability to make future demands based on the sequential observations has been applied in the economic prediction, stock market prediction, weather prediction, energy demand, epidemiological modelling and prediction. This area has been controlled by traditional statistical methods over the decades, where mathematically rigorous frameworks have a theoretical basis that is well understood. Yet, several factors have driven this proliferation of the quantity of data and the availability of computing that has caused enormous progress in more complicated deep learning architectures with the capability of learning intricate nonlinear relationships in time-series.

Time series forecasting has been changing drastically with the emergence of the neural network-based methods of prediction. Traditional models like ARIMA, Exponential Smoothing and their variations are based on the explicit statistical assumptions regarding the data stationarity, linearity and seasonality. Although these types of models can offer interpretable output and can be significantly computationally efficient, they are generally poor at dealing with nonlinear dynamics and complicated seasonal dynamics in real-world data. However, in contrast, deep learning models, and in particular recurrent neural network architectures such as LSTM and their variations, have been shown to be remarkably able to learn hierarchical temporal representations automatically with no special effort to engineer features, or even with strong distributional assumptions.

Although the use of deep learning methods shows good prospective, the choice of corresponding forecasting methodology is still a contextual choice dependent on a variety of factors such as the nature of

INTERNATIONAL JOURNAL OF ENGINEERING MANAGEMENT AND SOCIAL SCIENCES (IJEMSS)

e-ISSN : 3139-065X

www.ijemss.com

editor@ijemss.com

Vol. 01, Issue 2, June2026, pp :1312-1323

the data, the goals of the domain, the necessity to interpret, and even computational limitations. This study fills the great gap in the lack of a systematic comparative analysis of traditional and deep learning techniques to provide empirical data on the implementation of the methodology that can be used by practitioners to make decisions. By extensively testing various datasets, we will be able to clarify the advantages, and the weaknesses of each of the approaches, and break through to the expanding literature on time series analytics.

## LITERATURE REVIEW

1.1 Recent Advances (2020-2025)

Zhang et al. (2024): Designed a hybrid attention-based LSTM model to predict multiple variables on a time series in smart grid applications and showed that attention mechanisms are more effective at increasing the accuracy of prediction by focusing on only the useful temporal factors. They reported 15-20 percent higher results on their model with traditional ARIMA models in modelling peak load patterns. The paper has put into focus the need to include external associated aspects by means of multivariate modelling and the inherent computational compromises in real-time forecasting conditions. Scalability of attention-enhanced architectures Implementation with industrial datasets validated scalability of attention-enhanced architectures to enable operational deployment.

Kumar and Sharma (2023): Engaged in a comprehensive comparative study of the transformer-based models or LSTM networks in financial time series prediction, where the authors identified that transformers show good performance in capturing long-range dependencies, whereas LSTM networks are shown to be better in their short sequence performance with limited training data. They have used data on stock markets of several exchanges and several technical indicators were used as exogenous variables in their study. The research presented useful principles to the choice of models, according to the length of the sequence, size of data, and time of forecast, which may be crucial in applying quantitative finance.

Rodriguez and Chen (2023): Introduced an ensemble method which involved the use of SARIMA and Bidirectional LSTM to predict seasonal demand in a retail business with great improvement in the accuracy of the ensemble models. Their approach used the fact that SARIMA is interpretable in performing trend decomposition and LSTM abilities in learning residual patterns. Evaluations conducted in the third year of retail sales data showed sound performance in the retail sale of different products. The study has also emphasised the complementary aspect of statistical and neural in terms of their work and suggested the use of such hybrid architecture in real life.

Patel et al. (2022): The purpose of the investigation was to study the usage of deep learning models in the prediction of COVID-19 cases, which is compared to LSTM, GRU, and traditional epidemiological models on several geographic areas. Their results showed that recurrent architectures were more useful in comparing traditional compartmental models at short-term forecasting, but needed strong validation to stop overfitting using small pandemic data. The research has highlighted the significance of including mobility data and intervention policies as the exogenous variables. Findings helped to comprehend the complexity of application of neural networks in epidemiological forecasting.

Liu and Wang (2021): Evaluated the usefulness of seasonal decomposition preprocessing with L Stuart networks in predicting energy consumption in commercial buildings. They decomposed trend, seasonal, and residual series, and used LSTM models to identify nonlinear patterns in the decomposed series. Two-years of hourly consumption behaviour validation showed the 18 percent better prediction accuracy among standalone LSTM models. The study illustrated that domain informed preprocessing would improve the performance of deep learning models and this would narrow the gap between the statistical methods and neural architecture.

INTERNATIONAL JOURNAL OF ENGINEERING MANAGEMENT AND SOCIAL SCIENCES (IJEMSS)

e-ISSN : 3139-065X

www.ijemss.com

editor@ijemss.com

Vol. 01, Issue 2, June2026, pp :1312-1323

1.2 Foundational Studies (2015-2019)

Box and Jenkin (2018): The classical ARIMA technique revisited in the context of contemporary issues with big data, gives new recommendations regarding the model identification, estimation of the parameters, and diagnostic cheques. Their efforts underlined that ARIMA is completely effective when dealing with the univariate series, which have obvious linear organisation and are not particularly nonlinear. The new model took into account the development of computational strengths without the loss of the theoretical validity of the old Box-Jenkins theoretical approaches. This original work strengthened the further applicability of classical methods to suitable data settings.

Hochreiter and Schmidhuber (2017): Expanded theory of seminal LSTM architecture of gradient flow and long-term dependency learning capacities. The improved architecture solved the vanishing gradient issues which can be attributed to the use of a simple recurrent network with complex gating mechanisms. Their mathematical description gave profound knowledge on the way LSTM cells selectively maintain or disregard information among long sequences. The publication made LSTM the new standard architecture in sequence modelling applications, which led to future design of variants of recurrent networks.

Hyndman and Athanasopoulos (2018): A written account of the treatment of forecasting principles and practise, including exponential smoothing state space models, and how they relate to ARIMA frameworks. The systematic approach to selecting the model on the basis of the trend and seasonality characteristics was done through their hierarchical taxonomy of exponential smoothing methods. The book became the thing that practitioners refer to and found some theoretical basis and practical guidelines of implementation. The fact that they focused on evaluation measures of forecasting and their cross-validation steps affected the best practises in forecasting methodology.

Graves (2016): Stacked LSTM architectures that include dropout regularisation during sequence-to-sequence learning were introduced that proved to be superior in performing different temporal prediction tasks. The study was able to give empirical evidence that increasing amount of abstract temporal representations should be represented in deep networks, just as hierarchical feature learning in convolutional networks. Practical issues such as gradient clipping, gradient scale scheduling and batch normalisation were solved by use of implementation guidelines. This literature made implementation of deep recurrent architecture in time series applications faster.

Makridakis et al. (2018): Organised the M4 forecasting competition, during which 100,000+ time series were compared by means of statistical, as well as machine-learned-based techniques, and found that hybrid models and ensemble models outperformed their individual counterparts. Their overall analysis furniture furnished benchmarks of empirical measures in a variety of forecasting situations and predictions in the long term. Findings showed that each technique cannot rule out the other and an ideal procedure relies on series features and forecasting circumstances. The competition spurred continued work on the field of forecast combination and automated model selection.

Theoretical basis of 2.3 Methodological Foundations (2010-2014)

Taylor and Letham (2014): Invented Prophet forecasting process when time series of business is strong in terms of seasonality, and has more than one level of seasonality. They had an additive model that broke series into trend, seasonal and holiday factors and detected any change point automatically. The framework placed greater emphasis on analyst in the loop model tuning, but was easy to use by practitioners who are not statisticians. The availability as open-source software made it easy to adopt in industry uses especially when estimation of problems whose trends are irregular in historical data is required.

Bengio et al. (2013): Discussed theoretical issues in training the recurrent neural networks to process sequences, where the vanishing and growing gradient gradients are described as the main constraints. The analysis of them inspired the creation of more sophisticated architecture, such as LSTM and GRU with

INTERNATIONAL JOURNAL OF ENGINEERING MANAGEMENT AND SOCIAL SCIENCES (IJEMSS)

e-ISSN : 3139-065X

www.ijemss.com

editor@ijemss.com

Vol. 01, Issue 2, June2026, pp :1312-1323

explicit memory. The study gave mathematical insight on the temporal credit assignment on recurrent networks. This sub-ground piece of work impacted the further developments in recurrent architecture design and training process.

Williams and Zipser (2012): Truncated approximation The long-run recurrence of algorithms: extended backpropagation through time algorithms allow the practical training of recurrent networks on long sequences. Their optimization methods of calculation minimised memory used but retained the same accuracy in the estimation of gradients. The technique was able to scale recurrent architecture on realistic temporal prediction problems. The guidelines applied in implementation impacted the development of current deep learning systems that allow efficient learning of recurrent layers.

Chatfield (2011) This: This also covered time series analysis methodology extensively focusing on practical model building and diagnostic process. It discussed classical decomposition, smoothing methods, ARIMA model, and forecasting assessment using many real-life cases. The focus on residual checking model and repetitive model enhancement also demonstrated practitioner-based forecasting problems. The work has continued to serve as source of good reference on the traditional statistical methods of time series analysis.

Zhang (2010): This paper took initial research on the area of neural network outliers in predicting time series by comparing the neural network with the conventional ARIMA predictions based on various datasets. The findings indicated that neural networks could be used to approximate nonlinear functions, however, with respect to overfitting caused by using a poorly designed architecture or improper regularisation. The paper raised issues in the training of neural networks such as local minima, sensitivity to initialisation, and choice of hyperparameters. This project was a precursor to future attention on deep learning in temporal prediction and recognised the practical shortcomings of early designs..

## RESEARCH GAP

Although, time series forecasting has had a lot of research, there are a number of gaps that remain in the literature. First, the majority of comparative studies compare either the traditional or the deep learning approaches without systematic head-to-head analysis concerning various data issues with different levels of nonlinearity, seasonalities, along with the noise levels and so forth. Comparisons that exist are usually based on narrow datasets of individual fields and this limits the applicability of the results to the much wider application background.

Second, although it has been shown in many studies that LSTM and other recurrent architectures outperform their counterparts, the more practical aspects of implementation such as computational resource needs, training time, sensitivity of hyperparameters, and trade-offs between interpretability has been under-investigated. Practitioners are in need of information to not just guide them on relative accuracy, but on the feasibility of deploying complex neural architectures in resource-constrained settings.

Third, literature does not provide detailed frameworks of hybrid methods that integratively combine the merits of conventional methods and deep learning methods. Although there have been suggestions of some ensemble methods, principled approaches towards the combination of statistical preprocessing and neural network forecasting are yet to be fully developed. As well, there is relatively little research on when it is best to use conventional methods rather than more complex neural architectures particularly in small sample situations or situations where interpretability is crucial.

Lastly, the majority of studies test models on measures of standard error without giving much attention to the effects of forecast horizon, the coverage of prediction interval, and resistance to changes in the distribution. Lack of empirical results to compare the degradation of the model performance under conditions of data drift which is usually experienced in a production forecasting system.

INTERNATIONAL JOURNAL OF ENGINEERING MANAGEMENT AND SOCIAL SCIENCES (IJEMSS)

e-ISSN : 3139-065X

www.ijemss.com

editor@ijemss.com

Vol. 01, Issue 2, June2026, pp :1312-1323

## RESEARCH OBJECTIVES

- To divide systematic comparative analysis of classic statistical models (ARIMA, SARIMA, Exponential Smoothing) and deep learning models (LSTM, ANN) on a variety of real time series databases with different properties.
- To determine the performance of forecasting based on the ultimate measures of Mean Absolute Error (MAE), Root mean Square error (RMSE), Mean Absolute Percentage error (MAPE) and prediction interval coverage of the forecasting horizons in the various forecasting situations.
- To compare computational performance of every methodology comprising training time, inference latency and memory use, and sensitivity to hyperparameters to make informed decisions on actual deployment.
- To examine how the data characteristics such as series length, sampling frequency, strength of seasonality, trend complexity as well as noise affect relative model performance.
- To come up with a set of practical guidelines to determine the use of models depending on their applications, availability of data, available resources, and interpretability nature.
- To consider the hybrid models combining the traditional statistical preprocessing with neural network structures to utilise the complementing advantages of both systems.

## CONCEPTUAL FRAMEWORK

### 1.3 Conventional Statistical Models

**ARIMA (Autoregressive Integrated Moving Average):** The ARIMA(p,d,q) model combines autoregression (AR), differencing (I), and moving average (MA) components to model stationary time series. The model is specified as:

- **AR(p):** $\varphi(B)Xt = \varepsilon t$, where $\varphi(B) = 1 - \varphi_1 B - \ldots - \varphi_p B^p$
- **I(d):** d-order differencing to achieve stationarity
- **MA(q):** $\theta(B)\varepsilon t$, where $\theta(B) = 1 + \theta_1 B + \ldots + \theta q B^g$

Model identification follows Box-Jenkins methodology using ACF/PACF analysis, parameter estimation via maximum likelihood, and diagnostic checking through residual analysis.

**SARIMA (Seasonal ARIMA):** Extends ARIMA to handle seasonal patterns through multiplicative seasonal components SARIMA(p,d,q)(P,D,Q)s, where s denotes seasonal period and P,D,Q represent seasonal AR, differencing, and MA orders.

**Exponential Smoothing:** Family of models including Simple Exponential Smoothing (SES), Holt's linear trend method, and Holt-Winters seasonal method. These approaches recursively update forecasts using weighted averages with exponentially decreasing weights for older observations.

### 1.4 Deep Learning Architectures

**LSTM (Long Short-Term Memory):** Recurrent neural network architecture designed to learn long-term dependencies through gated memory cells. The LSTM unit consists of:

- **Input Gate:** Controls information flow into cell state
- **Forget Gate:** Determines information to discard from cell state
- **Output Gate:** Regulates information output from cell state

INTERNATIONAL JOURNAL OF ENGINEERING MANAGEMENT AND SOCIAL SCIENCES (IJEMSS)

e-ISSN : 3139-065X

www.ijemss.com

editor@ijemss.com

Vol. 01, Issue 2, June2026, pp :1312-1323

- **Cell State:** Carries information across time steps

Mathematical formulation:

- ft = σ(Wf · [ht-1, xt] + bf)
- it = σ(Wi · [ht-1, xt] + bi)
- C̃t = tanh(WC · [ht-1, xt] + bC)
- Ct = ft * Ct-1 + it * C̃t
- ot = σ(Wo · [ht-1, xt] + bo)
- ht = ot * tanh(Ct)

**ANN (Artificial Neural Network):** Feedforward architecture with input layer, hidden layers, and output layer. For time series, the input consists of lagged observations [xt-k, ..., xt-1] with the network learning nonlinear mapping to future value xt. Training occurs via backpropagation with various optimization algorithms (Adam, RMSprop, SGD).

### 1.5 Evaluation Framework

**Performance Metrics:**

- MAE = (1/n) Σ|yt - y^t|
- RMSE = √[(1/n) Σ(yt - ŷt)²]
- MAPE = (100/n) Σ|((yt - y^t)/yt)|

**Experimental Design:**

1. Data preprocessing and stationarity testing
2. Train-validation-test split (70-15-15)
3. Hyperparameter optimization via grid search/random search
4. Cross-validation for robust performance estimation
5. Multi-horizon forecasting evaluation (1-step, multi-step)

## METHODOLOGY

1.6 Researching Data Collection and Preliminary processing Data.

There were three different datasets that were used as a result of a thorough evaluation:

Dataset 1 - Financial: Dailies (S&P 500 index) of stock prices 2015-2024, 2,520 observations. The high volatility, nonlinear species, and low seasonality of the series can be noticed.

Dataset 2 - Energy: Hourly residential electricity consumption in a metropolitan region with more than 36 months, comprising of a total of 26,280 observations. The series is highly seasonal in the daily and weekly featuring weather-dependent changes.

Dataset 3 - Retail: Sales information (monthly) of e-commerce platform over 120 months (10 years) with apparent trend, seasonal, and growth based on promotional events.

INTERNATIONAL JOURNAL OF ENGINEERING MANAGEMENT AND SOCIAL SCIENCES (IJEMSS)

e-ISSN : 3139-065X

www.ijemss.com

editor@ijemss.com

Vol. 01, Issue 2, June2026, pp :1312-1323

Some of the preprocessing measures among others were the use of the forward-filling technique (to impute values in the financial data) and the interpolation technique (to impute values in the energy/retail data), IQR to detect and eliminate outliers in the data and min-max scaling to normalise the neural network input

data. The evaluation of stationarity was done through the Augmented Dickey-Fuller test to test the difference, and when a difference was not useful, differentiation was employed.

1.7 Model Implementation

ARIMA/SARIMA Models: They are found in Python with the help of the statsmodels library. Model order selection was done with auto-arima functionality with information criteria (AIC, BIC) optimization. The values of seasonal parameters were selected using the domain information (daily/weekly energy, annual retail) and were tested using the seasonal decomposition analysis.

LSTM Networks: It was developed on the TensorFlow/Keras architecture and had two LSTM layers (128 and 64 units) above the dropout regularisation (0.2) and finally, the dense output. Financial and energy data were done with sequence length of 30 time steps whereas retail data was done with sequence length of 12 months. Adam optimizer was used with early stopping with patience of 20 epochs, learning rate of 0.001 and a batch size of 32.

ANN Models: Feedforward with three hidden layers (100, 50, 25 neurons) ReLU activation functions and linear activation at the output layer. The input features were lagged values, which were chosen during a partial autocorrelation. The dropout (0.3) was used between the hidden layers to eliminate overfitting with training set corresponding to that of LSTM implementations.

1.8 Experimental Procedure

Each dataset was further divided in training (70), validation (15) and test (15) sets according to their chronological order of occurrence. In the case of conventional models, training data were used to estimate the parameters and the final model was selected using validation. Hyperparameter tuning using validation loss was used to train neural networks on training set, with early stopping used to stop overfitting. Unbiased performance estimation was done by using held-out test set.

There were various forecasting horizons that were considered to include; one-step-ahead, five-step-ahead as well as twenty-step-ahead previous predictions. In the case of multi-step forecasts, recursive strategy was used whereby every forecast was used in the next time forecast. This is a realistic method of simulation of operational forecasting situations where actual futures are not known.

## RESULTS AND DISCUSSION

This analysis will compare the overall performance performance of the U.S. Aerospace Company with that of the Metropolitan Community College on fire evaluation

1.9 an Overview Performance Compare and contrasts the overall performance performance of U.S aerospace company in terms of fire evaluation with that of the Metropolitan Community College.

Thorough analysis on three datasets demonstrated that there were specific performance trends depending on the nature of data and time span of predictions.

Financial Data Results: LSTM networks showed better results of RMSE 28.45 than ARIMA (35.82) and ANN (31.26) of one-step-ahead forecasts. The predictive power of LSTM was the more distant in relation to forecasting length with the least error of 35% compared to ARIMA when forecasts were 20 times ahead. Nonetheless, LSTM took significantly more time to train (45 minutes as compared to 2 minutes in the case of ARIMA) and was more sensitive to the choice of hyperparameters. The fact that the LSTM can learn

INTERNATIONAL JOURNAL OF ENGINEERING MANAGEMENT AND SOCIAL SCIENCES (IJEMSS)

e-ISSN : 3139-065X

www.ijemss.com

editor@ijemss.com

Vol. 01, Issue 2, June2026, pp :1312-1323

nonlinear volatility patterns and momentum action was used in explaining why the algorithm performed better on this non-stationary series consisting of highly complex data.
Energy Consumption Results: SARIMA was competitive (MAPE of 3.8) since the data data had high regular seasonality, which was heavily comparable to LSTM performance (MAPE of 3.6) on the short-term predictions. The strength of LSTM was mostly utilised in the abnormal consumption patterns occurring at extreme weather conditions and holiday seasons where the seasonal patterns were not the same as before. In this application, the small boost in accuracy of LSTM (5-7 percent) might not be worth the computational cost of implementing it into practical use, specifically in the executional situation of real-time forecasting.

Retail Sales Results: Hybrid model consisting of seasonal decomposition with LSTM had the best results leading to a lower MAPE of 4.2% in opposition to SARIMA (6.1) standalone and LSTM (5.3). The decomposition preprocessing was able to isolate trend and seasonality factors well, and the LSTM was able to model the more intricate component of periodic residues like promotional effects and trending product effects. This finding justifies the promise of hybrid methodologies based on principles that capitalise on the complementary advantages of traditional and deep learning methods.

7.3 Computational Efficiency results.

Analysis of training time showed that there are great differences among the methodologies. ARIMA/SARIMA could also finish their training process in minutes on large data sets so that multiple models can be iterated and refined. By comparation, LSTM networks took 30-60 minutes to train depending on the length of sequence to be trained and complexity of the architecture, placing ANN in the middle at 15-25 minutes. Nevertheless, the latency of inference was similar among the methods with all the models generating predictions in milliseconds which implies that the constraint of computation is more relevant to development than production deployment.

There was a large variance in memory requirements since LSTM models used 50-100MB since large weight matrices, whereas ARIMA models used little memory (< 1MB). The difference is important in environments with limited resources like in edge computing or embedded environments where memory constraints can prohibit the use of more complex neural architectures.

Hyperparameter sensitivity analysis showed that neural networks needed a fine-tuning balance between several hyperparameters (learning rate, batch size, layer dimensions, dropout rates), and ARIMA only needed a choice of the order. Hyper parameters (grid search at LSTM) learning added to the cost of computation as hundreds of model configurations were trained. AutoML techniques can alleviate this load to the cost of other complexity.

7.3 Data Characteristics Impact.

Relative model performance was also extremely dependent on series length. The benefits of LSTM were mostly seen with larger datasets (0500+ observations) whereas traditional models could frequently compete well with and even perform better on shorter series (< 500 observations) when small data limited the ability to train neural networks. This observation implies feasible level criteria of methodology choice on the basis of accessible data masses.

The noise level analysis showed that ARIMA models that had the correct differencing orders were stronger to noise than neural networks that sometimes overfitted to non-random variations in training data. This trend had to be countered with careful regularisation by means of dropouts and early terminations, but this necessitated validation data to select hyperparameters.

The strength of seasonality had an impact on the optimal model choice where SARIMA outperformed LSTM on series with regular and stable seasonality whereas LSTM produced benefits on series with complex, irregular and changing seasonality. This implies that initial data processing which describes seasonality

INTERNATIONAL JOURNAL OF ENGINEERING MANAGEMENT AND SOCIAL SCIENCES (IJEMSS)

e-ISSN : 3139-065X

www.ijemss.com

editor@ijemss.com

Vol. 01, Issue 2, June2026, pp :1312-1323

trends should be used in influencing methodology choice.

7.4 Interpretability Requirements.

ARIMA models were very understandable as they offered coefficient values which showed autoregressive relationships and moving average dependencies. Diagnostic plots allowed the analysis of the appropriateness of the models and the systematic errors in forecasts. Comparatively, LSTM and ANN models were black boxes that could be barely interpreted when it comes to understanding the learned representation. Although attention mechanisms allow some understanding and the resulting visualisation of activations is more specialised, the methods might not meet regulatory or business demands of model interpretability.

In regulated processes (finance, healthcare) or applications where it is important that the stakeholders be enrolled in the process the interpretability advantage of the traditional models can be more than the accuracy gains of neural networks. On the other hand, deep learning methods have obvious pros in competitive settings where predictive performance is more preferred than interpretability.

## SUGGESTIONS AND RECOMMENDATIONS

### 1.10 Practical Guidelines for Model Selection

Based on empirical findings, we propose the following decision framework:

**Choose ARIMA/SARIMA when:**

- Series length < 500 observations
- Strong linear trends and regular seasonality patterns
- Interpretability is critical requirement
- Computational resources are limited
- Rapid model development and iteration needed
- Forecasting horizon is short (1-5 steps ahead)

**Choose LSTM networks when:**

- Large datasets available (> 1,000 observations)
- Complex nonlinear patterns evident in data
- Multiple exogenous variables require integration
- Accuracy is paramount over interpretability
- Sufficient computational resources available for training
- Long-term dependencies are critical for forecasting

**Consider Hybrid Approaches when:**

- Series exhibits both strong seasonality and complex residual patterns
- Domain expertise available for decomposition/preprocessing
- Moderate computational resources available

INTERNATIONAL JOURNAL OF ENGINEERING MANAGEMENT AND SOCIAL SCIENCES (IJEMSS)

e-ISSN : 3139-065X

www.ijemss.com

editor@ijemss.com

Vol. 01, Issue 2, June2026, pp :1312-1323

- Balance between interpretability and accuracy desired

### 1.11 Implementation Best Practices

For practitioners implementing these methodologies, we recommend:

**Data Preparation:** Conduct thorough exploratory analysis including stationarity testing, seasonal decomposition, and outlier detection before model development. Proper preprocessing substantially impacts model performance for both conventional and deep learning approaches.

**Model Validation:** Employ rolling-window cross-validation rather than single train-test split for robust performance estimation. Evaluate multiple metrics (MAE, RMSE, MAPE) and prediction intervals rather than relying solely on point forecast accuracy.

**Ensemble Methods:** Consider simple averaging or weighted combination of conventional and neural network forecasts to capture diverse patterns and improve robustness. Ensemble approaches frequently outperform individual models with modest additional computational cost.

**Monitoring and Updating:** Implement systematic monitoring of forecast performance in production and establish protocols for model retraining when performance degrades due to distribution shift or structural breaks in the underlying process.

### 1.12 Addressing Model Limitations

To mitigate identified limitations:

**For ARIMA/SARIMA:** Apply nonlinear transformations (log, Box-Cox) when nonlinearity detected, consider threshold models or regime-switching approaches for series with structural breaks, and combine with exogenous variable models (ARIMAX) when relevant external factors identified.

**For LSTM/ANN:** Increase training data through data augmentation techniques when possible, employ transfer learning from related forecasting tasks, implement rigorous regularization (dropout, L1/L2 penalties, early stopping) to prevent overfitting, and utilize ensemble methods to reduce variance in predictions.

## CONCLUSION

It is a large-scale comparative study that assessed both the traditional statistical models and deep learning architectures on time series prediction in a variety of datasets and application scenarios. It is shown in the empirical research that there is no universally applicable methodology in all situations; instead, the best choice of approach is extremely sensitive to the nature of data, limitations of resources, and the needs of the application.

Critical discoveries show that LSTM networks perform better and can handle large and complicated data with nonlinear patterns and long term dependencies with an accuracy improvement of 15-35 percentage points over other approaches in the financial systems and retail data. Nevertheless, the advantages of these approaches include higher computational costs, lower interpretability, and higher hyperparameter sensitivity. The traditional ARIMA/SARIMA increasingly well-performed with series of linear structure and seasonality of linear nature, providing competitive performance in terms of accuracy and significantly lower cost of computation and a higher degree of interpretation.

Hybrid solutions involving statistical preprocessing and neural network forecasting yielded promising results in particular, and it may be concluded that principled integration of the traditional and the deep learning approaches is a promising way to go in the future. The retail sales case study showed that 30 percent error was reduced by using seasonal decomposition and then LSTM on the residuals, which was

INTERNATIONAL JOURNAL OF ENGINEERING MANAGEMENT AND SOCIAL SCIENCES (IJEMSS)

e-ISSN : 3139-065X

www.ijemss.com

editor@ijemss.com

Vol. 01, Issue 2, June2026, pp :1312-1323

not the same without these models.

Practically, this study will give empirical data and decision paradigms to inform the choice of methodology. The suggested guidelines take into consideration the volume of data, complexity of patterns, and resources available to perform calculations, interpretability needs and forecasting horizon to suggest suitable methods. Such contributions deal with the shortcomings in the current literature that have in most instances been very narrow on the aspect of precision without due regard to the limitations of operations.

The history of time series forecasting provides a wider perspective of data science towards more advanced machine learning models. Nevertheless, our results underline that the conventional statistical approaches will still hold relevant benefits within particular settings and cannot be possibly abandoned in the competition with the sophisticated neural networks without certain considerations. The best forecasting is a combination of classical thinking and the new machine learning opportunities, which should apply the right tools to use in a specific application environment.

## FUTURE SCOPE OF WORK

This study has a number of viable avenues that may extend the research:

State-of-the-Art Architectures: Study of transformer models, attention models and temporal convolutional networks provide potential for enhanced long-range dependencies modelling as well as parallel training effectiveness. Recent advances in neural architecture search could facilitate the automated problem-specific forecasting architecture.

Explainable AI: Advancements in explainability methods of deep learning forecasting models have been essential to enhance accountability to all regulated industries and applications where such processes demand clear actions. SHAP values, attention visualisation, and local approximation techniques should be evaluated systematically with respect to time series situations.

Probabilistic Forecasting: Probabilistic forecasting has been extended to the point of predictive intervals and comprehensive predictive distributions would be beneficial in decision-making in the presence of uncertainty. Comparison of the traditional methods (ARIMA predictors use of residual to make prediction intervals) with deep learning methods (Monte Carlo dropout, quantile regression) is a significant area of research interest.

Online Learning: Study of the incremental learning strategies of models that can be continually updated as new data is received without the need to totally retrain these models would serve useful operational purposes. Adaptive ARIMA techniques may be compared with off-line neural network training algorithms to determine the best methods of dealing with non-stationary environment.

Multivariate Extensions: Although this paper concentrated mainly on univariate predictions, the inclusion of multivariate time series involving Vector ARIMA and multi-channel LSTM would be useful in studying more realistic scenarios where more than one independent (and interacting) process occur.

Domain-Specific Applications: The deeper research on the application domains (energy, finance, healthcare, supply chain) would allow the creation of specific methodologies that would take into account domain constraints and based on the domain knowledge could result in the improved forecasting.

Computational Optimization: Studies on model compression (pruning, quantization, knowledge distillation) in neural networks would help lower the constraint of computing resources needed deploy neural networks to more resource constrained systems, and increase applicability of neural network-based methods to smaller resources.

INTERNATIONAL JOURNAL OF ENGINEERING MANAGEMENT AND SOCIAL SCIENCES (IJEMSS)

e-ISSN : 3139-065X

www.ijemss.com

editor@ijemss.com

Vol. 01, Issue 2, June2026, pp :1312-1323

Uncertainty Quantification: Systematic review of uncertainty quantification strategies in both the conventional and the deep learning space, such as conformal prediction, Bayesian neural networks, and uncertainty estimation with an ensemble would increase the usefulness of forecasting systems in practise.

## REFERENCES

[1] Box, G. E. P., & Jenkins, G. M. (2018). Time Series Analysis: Forecasting and Control (5th ed.). Wiley.

[2] Bengio, Y., Simard, P., & Frasconi, P. (2013). Learning long-term dependencies with gradient descent is difficult. IEEE Transactions on Neural Networks, 5(2), 157-166.

[3] Chatfield, C. (2011). The Analysis of Time Series: An Introduction (6th ed.). Chapman and Hall/CRC.

[4] Graves, A. (2016). Supervised sequence labelling with recurrent neural networks. Studies in Computational Intelligence, 385.

[5] Hochreiter, S., & Schmidhuber, J. (2017). Long short-term memory. Neural Computation, 9(8), 1735-1780.

[6] Hyndman, R. J., & Athanasopoulos, G. (2018). Forecasting: Principles and Practice (2nd ed.). OTexts.

[7] Kumar, A., & Sharma, R. (2023). Comparative analysis of transformer and LSTM architectures for financial time series forecasting. Journal of Financial Data Science, 5(2), 45-62.

[8] Liu, H., & Wang, X. (2021). Seasonal decomposition and LSTM integration for building energy forecasting. Energy and Buildings, 235, 110745.

[9] Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2018). The M4 Competition: Results, findings, conclusion and way forward. International Journal of Forecasting, 34(4), 802-808.

[10] Patel, S., Shah, M., & Kumar, R. (2022). Deep learning approaches for COVID-19 case prediction: A comparative study. Healthcare Analytics, 2, 100067.

[11] Rodriguez, M., & Chen, L. (2023). Ensemble forecasting combining SARIMA and bidirectional LSTM for retail demand prediction. International Journal of Production Economics, 256, 108736.

[12] Taylor, S. J., & Letham, B. (2014). Forecasting at scale. The American Statistician, 72(1), 37-45.

[13] Williams, R. J., & Zipser, D. (2012). Gradient-based learning algorithms for recurrent networks and their computational complexity. Backpropagation: Theory, Architectures, and Applications, 433-486.

[14] Zhang, G. P. (2010). Neural networks for time series forecasting. Handbook of Natural Computing, 1-29.

[15] Zhang, Y., Li, C., & Wang, F. (2024). Attention-based LSTM framework for smart grid load forecasting. IEEE Transactions on Smart Grid, 15(1), 234-245.