

Horizontal Divide and Conquer Methods for Non-Linear Time Series Forecasting Using LSTM-ANN; a Comparative Study Based on Simulated Data

Navoda Hettige*, WMND Basnayake

Department of Statistics, University of Colombo, Colombo, Sri Lanka

*Corresponding author: dininavo@gmail.com

Received July 13, 2023; Revised August 14, 2023; Accepted August 22, 2023

Abstract Time series forecasting holds a vital significance across diverse domains; however, accurately predicting the future is challenging due to the inherent complexity and non-linear nature of the data. One promising strategy for tackling non-linear time series data is the utilization of hybrid models, which have the potential to enhance forecasting accuracy. In this research paper, a comparative study is conducted, focusing on different approaches to horizontally partition data and fit Long short-term memory artificial neural network models (LSTM ANN). By using simulated data, this study effectively evaluates the efficacy of these approaches. The results demonstrate that DWT-LSTM, EEMD-LSTM, and CEEMDAN-LSTM techniques outperformed the EMD-LSTM and Threshold-LSTM models.

Keywords: complex, nonlinear, hybrid, LSTM

Cite This Article: Navoda Hettige, and WMND Basnayake, "Horizontal Divide and Conquer Methods for Non-Linear Time Series Forecasting Using LSTM-ANN; a Comparative Study Based on Simulated Data." American Journal of Applied Mathematics and Statistics, vol. 11, no. 2 (2023): 63-69. doi: 10.12691/ajams-11-2-4.

1. Introduction

Time series forecasting is the process of predicting future values of a variable based on its past observations. It has been an important area of research for many years and has numerous applications in various fields. Traditional single-model approaches such as ARIMA and ETS hold certain drawbacks in modeling non-linear time series data. Recent advancements in Machine Learning and Deep Learning have enabled the development of alternative methods such as the LSTM, which has shown promising results in forecasting non-linear data.

LSTMs are capable of simulating non-linear structural relationships, but they still have flaws that can lead to inaccurate forecasts [1]. A novel algorithm known as the "Horizontal Divide and Conquer method" has been developed to improve the accuracy of time series forecasting. The Horizontal Divide and Conquer (HDC) method involves breaking the time series into segments, applying the LSTM model to each segment independently, and obtaining the final outcome by taking the summation of those sub models. This approach can help to improve forecasting accuracy by focusing on specific patterns and dependencies within each segment rather than trying to capture all of the complexity in the entire time series at once. Instead of using a real dataset, using simulated data allows for more flexibility in the testing process and the ability to test on numerous time series.

However, despite the potential benefits of HDC methods, sufficient comparative evaluations have not been conducted up to now in order to select the optimal HDC methods to achieve increased accuracies in non-linear time series forecasting. Therefore, the problem statement for this study is to compare and evaluate the effectiveness of different HDC methods done in the literature in the context of time series forecasting using LSTM and to identify the most effective HDC method/s for a highly accurate prediction of non-linear time series data.

2. Literature Review

Time-series forecasting methods have gained significant attention in recent years, and a vast body of literature is available on the topic. Although many time series can be adequately modeled using linear methods, there are many cases where the underlying process is non-linear in the real-world scenario. As a result, more sophisticated techniques with well-established concepts including Random Forest, XGBoost, LSTM, and more have been developed by researchers and these models often provide better forecasting performance than conventional statistical methods [2]. However, dealing with nonstationary nature have proven many drawbacks even for those new approaches [1].

Many studies have been conducted in the past, that investigate the approaches in handling the nonstationary nature of time series. Incorporating exogenous variables,

using ensemble methods, developing new hybrid models that combine classical and ML-based methods, and using HDC methods are some of the highlighting approaches as such, etc [3].

There have been various research which were done on the HDC technique. The paper "Forecasting daily streamflow using hybrid ANN models" discussed the application of the divide and conquer (DAC) paradigm as a top-down black-box technique for forecasting daily streamflow. The authors used three forms of hybrid ANNs as univariate time series models: the threshold-based ANN (TANN), the cluster-based ANN (CANN), and the periodic ANN (PANN). The normal multi-layer perceptron form of ANN (MLP-ANN) was used as the baseline ANN model to compare forecasting efficiency [4]. It is found that the PANN model performed the best out of the three hybrid ANN variations tested.

The paper "Threshold Autoregression, Limit Cycles, and Cyclical Data" introduced a nonlinear time series model called threshold autoregression (TAR). The TAR model is used to capture the nonlinear dynamics of a time series by dividing the data into regions based on a threshold value. Each region is then modeled using a different linear autoregression model. The threshold value is estimated using maximum likelihood or other methods [5]. In this paper, TAR models have been applied to a variety of real data sets, including the Canadian lynx data and the rainfall-river flow data. TAR models can capture non-linearity and limit cycles, which can be important features of some time series data. However, TAR models can be more difficult to estimate and forecast than classical time series models.

The study "Wavelet-Based Nonlinear Autoregressive Neural Network to Predict Daily Reservoir Inflow" proposes a hybrid forecasting model based on wavelet transform and nonlinear autoregressive neural network (NAR-ANN) for the prediction of daily reservoir inflow in Sri Lanka [6]. The study aims to improve the accuracy of daily reservoir inflow prediction, which is critical for managing and operating hydropower plants in Sri Lanka. The DWT is applied to decompose the original time series into several sub-series, which are then fed into the ANN model to predict the daily inflow values. The authors compare the results of their proposed method with two other methods: a NAR-ANN with raw inputs, and a cluster-based modular NAR-ANN and the results show that the wavelet-based NAR-ANN outperforms the other two methods.

Norden Huang discusses the use of EMD [7] in non-linear time series analysis. The article "A review on empirical mode decomposition in forecasting time series" provides an overview of the Empirical Mode Decomposition (EMD) technique and its application in time series forecasting. In another study, a hybrid method that combines EMD and LSTM for wind speed forecasting was reported [8], that it outperformed other methods such as ARIMA and traditional ANNs.

[9], proposes an approach to improve the accuracy of forecasting using ANNs based on the EEMD decomposition technique. In this approach the original time series was decomposed into multiple sub-series, then ANNs were trained on each of the sub-series to capture the different characteristics of the data. The forecasts

obtained from the individual ANNs were then combined to obtain the final prediction. The study found that the EEMD-ANN approach outperformed traditional ANN models in terms of forecasting accuracy.

In the paper "Forecasting stock index price using the CEEMDAN-LSTM model", [10] proposes a novel method for predicting stock index prices. The method combines two popular techniques, the CEEMDAN, and the LSTM neural network. The stock index time series was first decomposed into a set of IMFs using the CEEMDAN method. The resulting IMFs represent different scales of fluctuations in the time series. The authors then trained a separate LSTM model on each IMF to capture the short-term dependencies in the data. Finally, the forecasts from each LSTM model are combined to obtain the final prediction. The authors compared the performance of their CEEMDAN-LSTM model with several other methods, including SVM, Backward Propagation (BP), the Elman network, Wavelet Neural Networks (WAV), and their mixture models combined with the CEEMDAN. The results show that the CEEMDAN-LSTM model outperformed all the other methods in terms of forecasting accuracy, as measured by the MCS test.

3. Methodology

Figure 1 provides an overview of the methodology used in this study to evaluate and compare the accuracies of non-linear time series forecasting using LSTM models with and without several HDC methods. A detailed explanation of each step is presented below.

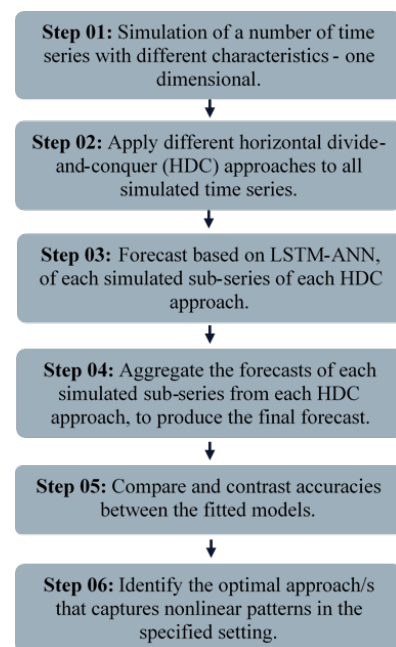


Figure 1. Research methodology

3.1. Data Simulation Procedure

Simulated data has been widely used in the field of time-series forecasting to evaluate the performance of different models. Simulated data provides a controlled environment where the characteristics of the time series

can be precisely controlled, allowing for a fair comparison of different models.

A simulated time series is not meant to replicate real-world time series, but it can be a useful tool for understanding the basic principles of non-linear dynamics and generating synthetic time series for testing or comparison purposes. The simulation process was further carried out by adding noise to the existing simulation by using different distributions in order to simulate more realistic and complex data with more variability and randomness.

3.2. Horizontal Divide and Conquer Methods

The divide-and-conquer method is a general algorithmic strategy for solving problems by breaking them down into smaller, simpler subproblems that can be solved independently, hoping that the solutions to the subproblems are relatively easier to find. The solutions to these subproblems are then combined to give the solution to the original problem. Figure 2 shown below is a basic architecture of the divide-and-conquer algorithm. This approach is often used for solving problems that have a recursive structure [11].

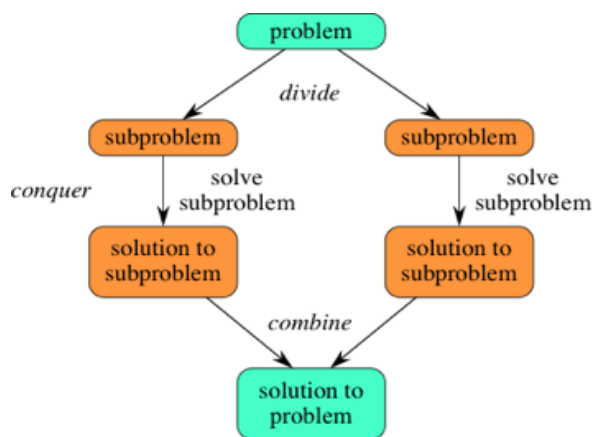


Figure 2. Divide-and-Conquer Algorithm [7]

In this study, five different HDC methods were considered, which involve splitting the time series into multiple horizontal segments.

3.2.1. Decomposing Time Series Using a Threshold Value

Using a threshold as an HDC approach in time series forecasting involves dividing a time series into two or more segments horizontally based on a specified threshold value. In this study, the threshold was selected in such a way that it divides the whole time symmetrically into two parts on the horizontal axis. One advantage of using a threshold-based approach is that it can capture the non-linear behavior of the time series. It can also handle outliers or extreme values that may be present in the data. However, the effectiveness of this approach depends on the choice of the threshold value and the number of

segments into which the time series is divided. The threshold value should be carefully selected based on the

nature of the time series data. Moreover, if the time-series does not have clearly distinct regimes, this approach may not be suitable.

3.2.2. The Discrete Wavelet Transform (DWT)

The DWT is a mathematical technique used to decompose a time series into a set of sub-series, each with a different frequency content [12].

The DWT employs a dyadic grid, where the mother wavelet is scaled by $a = 2^j$ and translated by an integer, $b = k2^j$ where k is a location index running from 1 to 2^jN (N is the number of observations) and j runs from 0 to J (J is the total number of scales).

The DWT is expressed by the following equation:

$$\Phi_{j,k}(t) = 2^{-j/2} \varphi(2^{-j}t - k) \quad (1)$$

The DWT coefficients are obtained from the following expression:

$$W_{j,k} = W(2^j, k2^j) \\ W_{j,k} = 2^{-j/2} \int f(t) \varphi(2^{-j}t - k) dt \quad (2)$$

The dyadic DWT allows for the decomposition of a series into a progression of ‘approximation’ and ‘detail’ coefficient sets under each level. The DWT supports different mother wavelets, such as ‘Harr, Daubechies, Biorthogonal, Symlet, Meyer, and Coiflets’ [13].

3.2.3. Empirical Mode Decomposition (EMD)

EMD is based on the empirical observation that any real-world signal can be decomposed into a finite number of intrinsic mode functions (IMFs), each of which has a well-defined local mean and local extrema. The IMFs are found by iteratively sifting the signal, which involves subtracting the mean of the local maxima and minima from the signal to obtain a residual and then repeating the process until the residual satisfies some stopping criteria [14].

$$x(t) = \sum c_i(t) + r_n(t) \quad (3)$$

where,

n – The number of IMFs,

$c_i(t)$ – IMFs,

$r_n(t)$ – The final residue

EMD has been applied in various fields, including biomedical signal processing, image processing, and financial time-series analysis. It has the advantage of being able to handle non-linear and non-stationary signals, which are common in real-world processes.

3.2.4. Ensemble Empirical Mode Decomposition (EEMD)

The decomposition process of EEMD starts by adding white noise to the signal and then applying the EMD algorithm to each noisy signal. The process is repeated multiple times, each time with different levels of white noise added to the signal. The IMFs extracted from each iteration are then averaged to obtain a final set of IMFs [4].

EEMD has been used in a variety of applications, including image processing, speech processing, and financial analysis. One of the main advantages of EEMD

is its ability to separate noise from the signal, making it a

useful tool for signals with a high degree of noise or interference.

3.2.5. Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN)

CEEMDAN is a signal processing method for decomposing signals into IMFs and a residual. It is an improved version of the EMD algorithm, which can handle non-stationary and non-linear signals effectively. In comparison to EMD, CEEMDAN has the ability to separate noise from the signal and it can handle the mode mixing issue which is commonly encountered in the EMD algorithm.

CEEMDAN works by decomposing the signal iteratively using the EMD algorithm and adding white noise to the signal after each iteration to prevent mode mixing [15]. The process is repeated until the desired number of IMFs and the residual are obtained. The final result of CEEMDAN is the sum of all IMFs and the residual trend, which represents the original signal.

3.2.6. The architecture of the LSTM model

It would be found that Long Short-Term Memory (LSTM) is a widely studied area in the field of ANN and ML, with a variety of applications in natural language processing, speech recognition, computer vision, and time series forecasting. LSTM is a type of recurrent neural network (RNN) that is specifically designed to handle long-term dependencies.

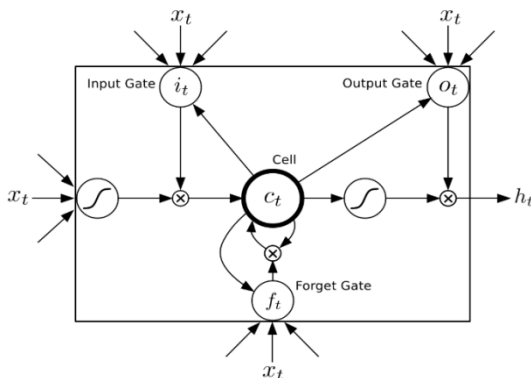


Figure 3. Standard LSTM unit [17]

The structure of a standard LSTM unit is shown in Figure 3 [16]. An LSTM network is composed of multiple LSTM cells, each of which has three gates: the input gate, the forget gate, and the output gate. The basic operations are accomplished by the input gate (i_t), the forget gate (f_t), and the output gate (O_t).

3.2.7. Hybrid HDC-LSTM Approach

This study is focused on HDC methods in the context of time series forecasting, where the time series data is divided horizontally into multiple sub-series and each sub-series is forecasted separately using a separate LSTM model. The forecasts from each segment in one simulated time series can then be combined to form the final forecast for that entire time series. The illustration of the process is shown in Figure 4.

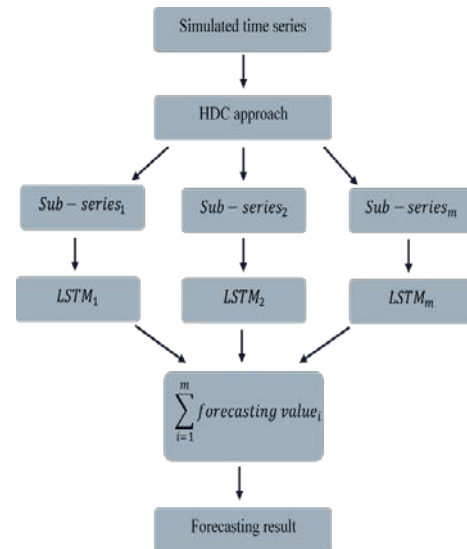


Figure 4. The architecture of the hybrid HDC-LSTM model

4. Analysis

4.1. Results of Simulations

The results of some simulated time series with different characteristics are shown below. Data simulated using the equation 4 is given in Figure 5.

$$X_t = 5 + (2\pi t) \quad (4)$$

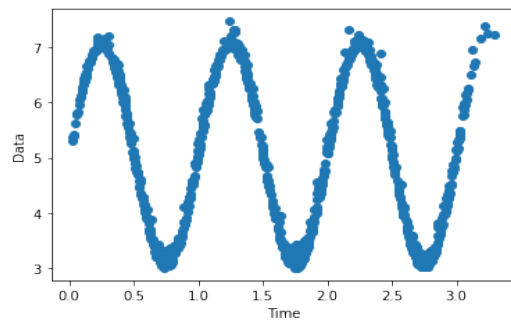


Figure 5. Simulated time series (a)

Data simulated with noise using the equation 5 is given in Figure 6.

$$X_t = 0.5 + \sin(0.2\pi t) * \exp(-0.05t) \quad (5)$$

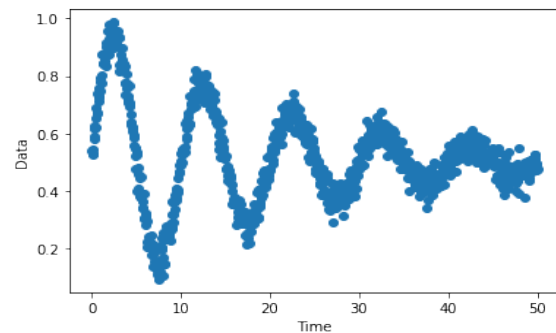


Figure 6. Simulated time series (b)

Data simulated using a random walk model with noise is given in Figure 7.

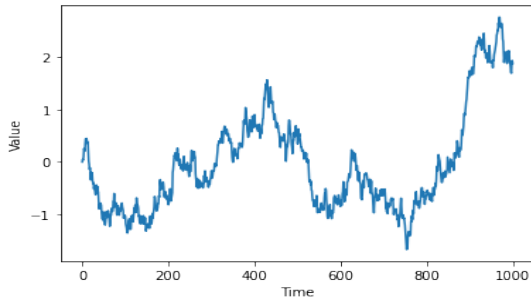


Figure 7. Simulated time series (c)

Data simulated using a random walk model with noise is given in Figure 8.

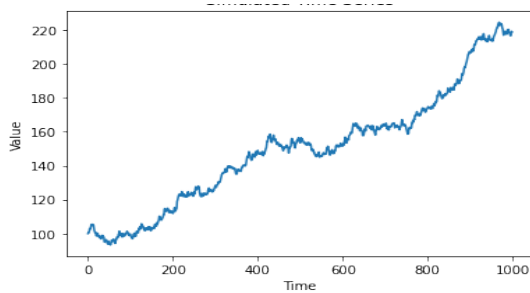


Figure 8. Simulated time series (d)

4.2. Results of Different HDC-LSTM Approach

The LSTM models were implemented using the Keras DL library and the Python 3.8 programming language. Five distinct HDC methods were used to partition each simulated time series horizontally. Before fitting an LSTM, the data set (1000 observations) was divided in to training (80%) and test sets. The mean squared error (MSE) is used as the loss function while the models are trained using the Adam optimization technique. The ideal LSTM parameters were discovered via a grid search. The final outcome of each approach was calculated by averaging the results of each sub-series in one approach. Then, the evaluation metrics MAPE were used to assess each method's forecasting performance. Moreover, the comparison of hybrid HDC-LSTM was compared with the applying only LSTM.

Table 1 summarizes the forecasting results of the best hybrid-LSTM approaches which accurately predict the non-linear behavior on the simulated time series data.

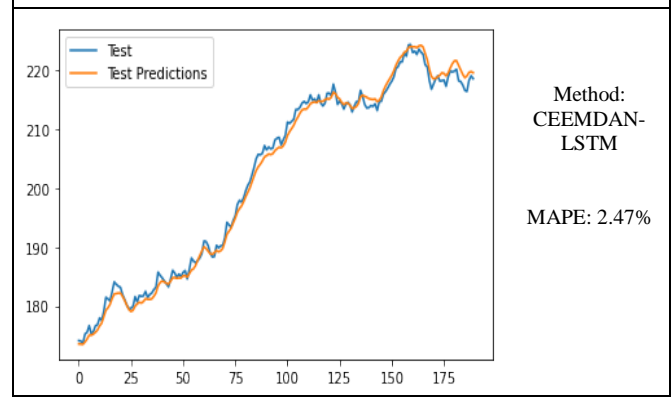
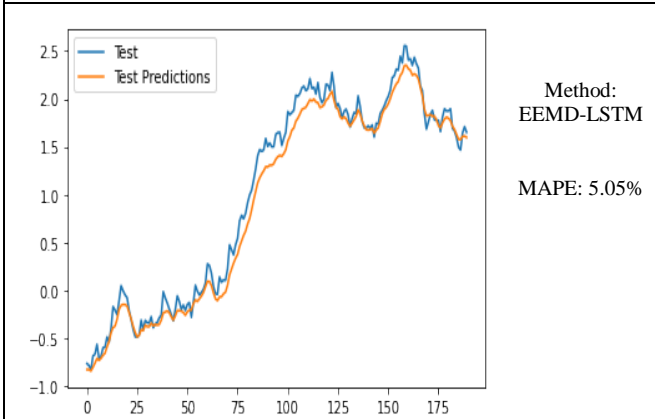
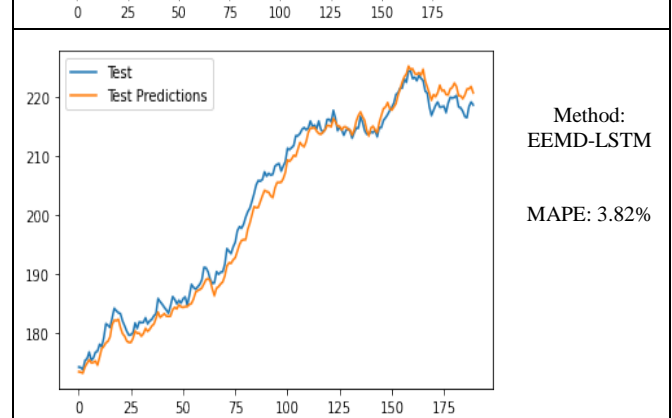
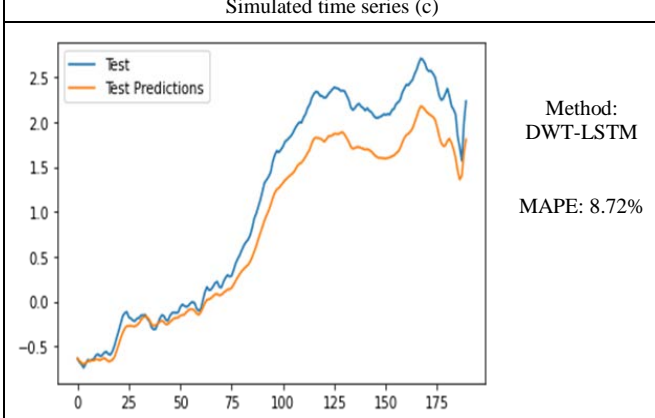
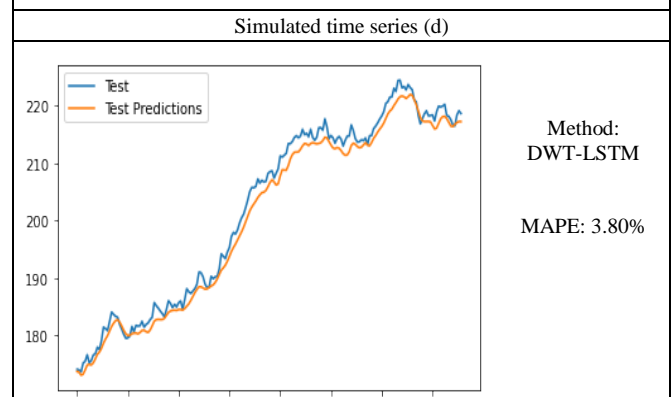
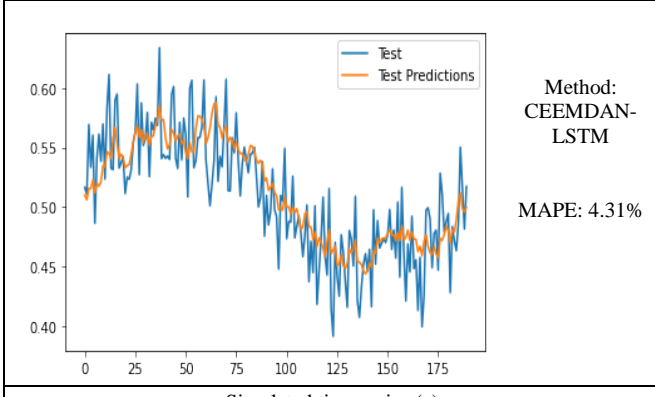
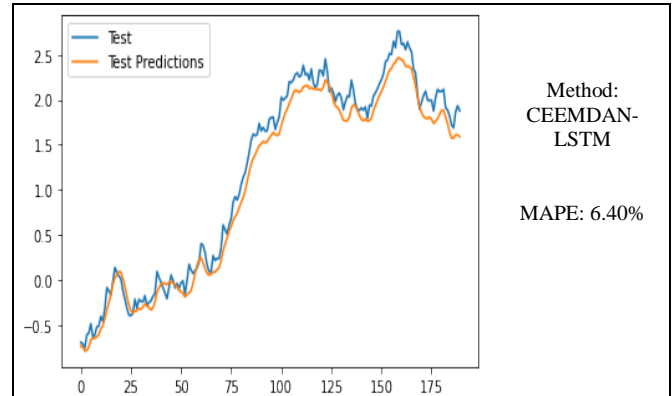
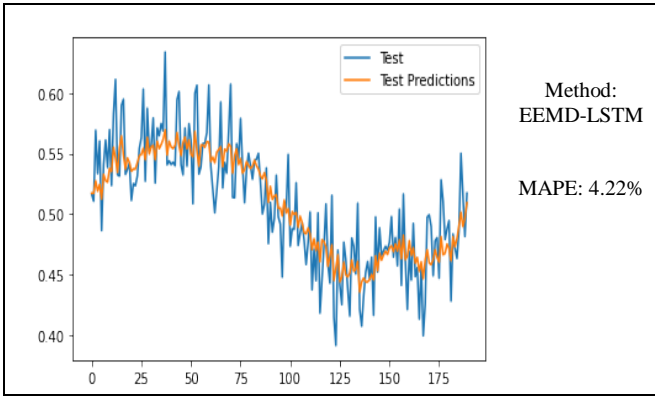
The Table 1 shows the mean absolute percentage error (MAPE) for each method. The MAPE is a measure of the accuracy of a forecast, and it is calculated by dividing the sum of the absolute errors by the sum of the actual values. A lower MAPE indicates a more accurate forecast. According to Table 1, it can be observed that predictions are so close actual values and the MAPE values are small. This is because the HDC-LSTM method is able to capture long-term dependencies in the data more accurately, which is important for forecasting non-linear time series.

The results of the study showed that the HDC-LSTM methods are effective in improving the accuracy of non-linear time series forecasting. The results of the study showed that the HDC-LSTM methods are effective in forecasting non-linear time-series in terms of accuracy.

Therefore, the hybrid HDC-LSTM methods can be a beneficial strategy for handling non-linear time series data.

Table 1. The Forecasting Performance of Different HDC Methods on Simulated Time Series

Simulated time series (a)	
	Method: DWT-LSTM MAPE: 11.47%
	Method: EEMD-LSTM MAPE: 10.42%
	Method: CEEDMAN-LSTM MAPE: 10.11%
Simulated time series (b)	
	Method: DWT-LSTM MAPE: 5.56%



5. Conclusion & Future Prospects

5.1. Conclusion

This study presented a comparative analysis of HDC methods for non-linear time series forecasting using LSTM. The study employed simulated data to compare the performance of HDC-LSTM methods with that of the traditional LSTM model on non-linear time series data. The results showed that HDC-LSTM methods outperformed traditional LSTM models in terms of forecasting accuracy, particularly when dealing with highly non-linear time series data.

The study also investigated the impact of different horizontal partitioning strategies on the performance of HDC-LSTM methods. The results showed that the DWT-LSTM, EMD-LSTM, and CEEMDAN-LSTM performed better than the other partitioning strategies used.

The results of this study demonstrate the potential of HDC-LSTM methods for non-linear time series forecasting. This method can be particularly useful when dealing with long and complex time series in providing accurate and efficient forecasts for various applications, including financial forecasting, weather forecasting, and traffic forecasting, among others.

5.2. Future Directions

This study offers valuable insights into the performance of HDC-LSTM methods for non-linear time series forecasting. However, it also highlights the potential for further research in this field.

While this study provides valuable insights into the performance of HDC-LSTM methods for non-linear time series forecasting, it is worth noting that the analysis was limited to simulated data. Therefore, an intriguing avenue for future research lies in examining the effectiveness of HDC-LSTM methods on real-world time series data. By conducting experiments on authentic datasets from different fields, we can assess the generalizability and reliability of HDC-LSTM approaches in real-life scenarios.

Secondly, this study only considered LSTM-based HDC methods. Future research could investigate the effectiveness of other types of ANN-based HDC methods, such as Convolutional Neural Networks (CNNs), Multi-Layer Perceptron (MLP), and other types of RNNs rather than LSTM, *etc.*, for non-linear time series forecasting.

Thirdly, the study will focus only on one-dimensional data, future results can be done to find the applicability of the results to higher-dimensional time series.

At last, while this study focused on horizontal partitioning, future research could investigate the potential benefits of combining horizontal and vertical partitioning strategies for time series forecasting.

Future research can focus on developing hybrid models forecasting methods to further improve the accuracy of

non-linear time series, which could have significant implications for a wide range of applications.

References

- [1] Basnayake, W.M.N.D., Attygalle, M. D. T., Liyanage-hansen, L., & Nandalal, K. D. W. (2019). *Modified 1D Multilevel DWT Segmented ANN Algorithm to Reduce Edge Distortion*. 7(1), 25–31.
- [2] Moroff, N. U., Kurt, E., & Kamphues, J. (2021). Machine Learning and Statistics: A Study for assessing innovative Demand Forecasting Models. *Procedia Computer Science*, 180, 40–49.
- [3] Brownlee, J. (2020). *How to Develop LSTM Models for Time Series Forecasting*. Machine Learning Mastery. <https://machinelearningmastery.com/how-to-develop-lstm-models-for-time-series-forecasting/>
- [4] Wang, W., Gelder, P. H. A. J. M. V., Vrijling, J. K., & Ma, J. (2006). Forecasting daily streamflow using hybrid ANN models. *Journal of Hydrology*, 324(1–4), 383–399.
- [5] Tong, H., & Lim, K. S. (1980). Threshold Autoregression, Limit Cycles and Cyclical Data. *Journal of the Royal Statistical Society: Series B (Methodological)*, 42(3), 245–268.
- [6] Basnayake, W.M.N.D (2017). *Wavelet Based Nonlinear Autoregressive Neural Network to Predict Daily Wavelet Based Nonlinear Autoregressive Neural Network to Predict Daily Reservoir Inflow*. March 2018.
- [7] Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Snin, H. H., Zheng, Q., Yen, N. C., Tung, C. C., & Liu, H. H. (1998). The empirical mode decomposition and the Hubert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 454(1971), 903–995.
- [8] Liu, H., Chen, C., Tian, H. Q., & Li, Y. F. (2012). A hybrid model for wind speed prediction using empirical mode decomposition and artificial neural networks. *Renewable Energy*, 48, 545–556.
- [9] Wang, W. chuan, Chau, K. wing, Qiu, L., & Chen, Y. bo. (2015). Improving forecasting accuracy of medium and long-term runoff using artificial neural network based on EEMD decomposition. *Environmental Research*, 139, 46–54.
- [10] Lin, Y., Yan, Y., Xu, J., Liao, Y., & Ma, F. (2021). Forecasting stock index price using the CEEMDAN-LSTM model. *North American Journal of Economics and Finance*, 57, 101421.
- [11] Le, J. (2018). *Divide and Conquer Algorithms*. *Divide-and-Conquer Algorithm | by James Le | Data Notes*. Data Notes. <https://data-notes.co/divide-and-conquer-algorithms-b135681d08fc>
- [12] Mallat, S. G. (1989). A Theory for Multiresolution Signal Decomposition: The Wavelet Representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(7), 674–693.
- [13] Rhif, M., Abbes, A. Ben, Farah, I. R., Martínez, B., & Sang, Y. (2019). Wavelet transform application for/in non-stationary time-series analysis: A review. In *Applied Sciences (Switzerland)* (Vol. 9, Issue 7, p. 1345). Multidisciplinary Digital Publishing Institute.
- [14] Zeiler, A., Faltermeier, R., Keck, I. R., Tomé, A. M., Puntonet, C. G., & Lang, E. W. (2010). Empirical mode decomposition - An introduction. *Proceedings of the International Joint Conference on Neural Networks*.
- [15] Torres, M. E., Colominas, M. A., Schlotthauer, G., & Flandrin, P. (2011). A complete ensemble empirical mode decomposition with adaptive noise. *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, 4144–4147.
- [16] Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*, 9(8), 1735–1780.
- [17] Gall, R. (2018). *What is LSTM?* | Packt Hub. Packtpub. <https://hub.packtpub.com/what-is-lstm/>

