Journal of King Saud University - Computer and Information Sciences 36 (2024) 102252

HOSTED BY

ELSEVIER

Contents lists available at ScienceDirect

Journal of King Saud University - Computer and Information Sciences

journal homepage: www.sciencedirect.com

Full length article

Enhancing stock market predictions via hybrid external trend and internal components analysis and long short term memory model

Fatene Dioubi [a,b], Negalign Wake Hundera [a,c], Huiying Xu [a,*], Xinzhong Zhu [a,d]

[a] *School of Computer Science and Technology, Zhejiang Normal University, Jinhua, China*
[b] *Zhejiang Institute of Photoelectronics, Jinhua, China*
[c] *Zhejiang Institute of Optoelectronics, Jinhua, China*
[d] *Research Institute of Hangzhou Artificial Intelligence, Zhejiang Normal University, Hangzhou, China*

ARTICLE INFO

Keywords:
Stock market predictability
Hybrid models
ETICA decomposition method
LSTM model
Financial forecasting

ABSTRACT

When it comes to financial decision-making, stock market predictability is extremely important since it offers valuable information that may guide investment strategies, risk management, and portfolio allocation overall. Traditional methods often fail to accurately predict stock prices due to their complexity and inability to handle non-linear and non-stationary patterns in market data. To address these issues, this study introduces an innovative model that combines the External Trend and Internal Components Analysis decomposition method (ETICA) with the Long Short-Term Memory (LSTM) model, aiming to enhance stock market predictions for S&P 500, NASDAQ, Dow Jones, SSE and SZSE indices. Through rigorous testing across various training data proportions and epoch settings, our findings reveal that the proposed hybrid model outperforms the single LSTM model, delivering significantly lower Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) values. This enhanced precision reduces prediction errors, underscoring the model's robustness and reliability. The superior performance of the ETICA-LSTM model highlights its potential as a powerful financial forecasting tool, promising to transform investment strategies, optimize risk management, and enhance portfolio performance.

1. Introduction

Making informed financial decisions and creating strategies that reduce risks and maximize returns require an understanding of how predictable the stock market is. In the dynamic realm of finance, being able to anticipate what might happen in the market can make a huge difference for both individual investors and big financial institutions. Delving into the background of stock market predictability unveils a lot of different ideas and studies about how and why the market moves the way it does. Drawing from enduring theories such as Fama's efficient market hypothesis (EMH) (Fama, 1995), which argues that stock prices already incorporate all known information, to newer ideas in behavioral finance, there is a wide range of viewpoints on financial markets predictability. Due to recent advancements in computing technology, predicting the stock market is now quite possible (Hoseinzade et al., 2019), the scientific community has put out several approaches to stock market forecasting in spite of Fama's theory (Bustos and Pomares-Quimbaya, 2020).

Stock market predictions have been a topic of interest for financial practitioners and policymakers for a long time, hence numerous studies have been done in this area where two primary types of analysis are conducted. The fundamental analysis is to investigate the reasons for price fluctuations, such as current industrial conditions, economic and non-economic factors, internal market factors, domestic and foreign economic conditions (Bousoño-Calzón et al., 2019). Technical analysts, on the other hand, rely on market data, like historical prices and trading volume, employing technical analysis to evaluate securities (Ahmadi et al., 2018). Recent studies have demonstrated a correlation between past and future return rates, which has prompted some to argue that

* Corresponding author.
E-mail addresses: dioubifatene@zjnu.edu.cn (F. Dioubi), nigaccna21@zjnu.edu.cn (N.W. Hundera), xhy@zjnu.edu.cn (H. Xu), zxz@zjnu.edu.cn (X. Zhu).
Peer review under responsibility of King Saud University.

ELSEVIER Production and hosting by Elsevier

https://doi.org/10.1016/j.jksuci.2024.102252
Received 12 September 2024; Received in revised form 4 November 2024; Accepted 14 November 2024
Available online 22 November 2024

previous data could be used to predict future returns if one believes that the past would repeat itself (Murphy, 1999). Consequently, technical analysis has been used in a number of studies to forecast future changes in stock values. Traditional stock prediction methods had difficulties capturing complicated market dynamics, having limited feature representation, having non-stationary data, requiring a lot of pre-processing, and having little predictive power (Ali et al., 2023). Therefore, the ability of leveraging algorithms such as support vector machines and neural networks for tackling these challenges by modeling nonlinear relationships, extracting features, adapting to market changes, handling raw data, and enhancing predictive accuracy, makes them valuable for financial forecasting and stock prediction (Kurani et al., 2023).

Recurrent neural networks (RNNs), in contrast to a basic artificial neural network, have seen significant success in the financial sector due to their excellent performance (Nabipour et al., 2020). The application of deep learning, renowned for their ability to handle nonlinear time series data without explicit input features, has significantly increased with an increasing number of articles utilizing it in recent studies (Hiransha et al., 2018). One enhanced version of the RNN technique utilized in the deep learning field is long short-term memory, this model solves the vanishing gradient problem by replacing the hidden layer units with memory cells (Bhandari et al., 2022), making it more effective for stock price prediction (Ali et al., 2023). Forecasting accuracy can be increased by using hybrid or multiple models (Hajirahimi and Khashei, 2019; Huang et al., 2021). Indeed, the hybrid LSTM model showed that in many studies by outperforming the single LSTM model (Deng et al., 2023). Even with the benefits these techniques have brought to data analysis, certain time series are still quite unstable and chaotic. This problem can be solved by isolating highly fluctuating data into lower frequency components using frequency decomposition techniques like EMD and CEEMD, which streamline the analysis (Rezaei et al., 2021).

In contrast to conventional research focusing on frequency decomposition methods, this study delves into stock market prediction by examining the impact of the ETICA decomposition method combined with Long Short-Term Memory (LSTM) networks. While traditional LSTM models have proven effective in capturing temporal dependencies in stock market data, the non-linear and non-stationary nature of stock prices poses challenges. To address this, ETICA is used to decompose stock prices into external trends and internal components, enabling LSTM to focus on cleaner, more distinct patterns for more accurate predictions. This approach provides a nuanced understanding of market dynamics, allowing LSTM to capture complex temporal patterns more effectively. ETICA's ability to disentangle factors influencing stock prices facilitates precise modeling, which significantly reduces prediction errors. Empirical validation using historical data from the S&P 500, NASDAQ, Dow Jones, SSE and SZSE indices demonstrates that the proposed ETICA-LSTM hybrid model outperforms a single LSTM, EMD-LSTM and CEEMDAN-LSTM models, showcasing its potential for financial forecasting.

This paper is structured as follows: Section 2 presents related work. Section 3 details the data used as inputs and outlines our proposed methodology. Section 4 discussing the experimental results and performance measures. Finally, Section 5 explores potential future directions for this research.

2. Related work

By combining statistical methods with machine learning algorithms and integrating diverse data sources, hybrid models capture non-linear patterns and improve prediction reliability in the dynamic stock market environment. Many strategies exist for predicting financial time series, ranging from traditional methods to deep learning models. Of these, deep learning has drawn the most interest because of its superior performance, especially the LSTM models.

Several studies have aimed to achieve more accurate predictions by proposing various hybrid LSTM models. The LSTM-ARIMA hybrid models were developed to combine the statistical modeling capabilities of Autoregressive Integrated Moving Average (ARIMA) models with the strengths of LSTM networks in capturing intricate patterns. Temur et al. (2019) utilized ARIMA and LSTM, proposing a hybrid LSTM-ARIMA model which has shown to perform better when error rates are compared. In 2020, Kulshreshtha et al. (2020) introduced a hybrid ARIMA-LSTM model for capturing both the non-linear and linear aspects of the time series, showing that the proposed hybrid model outperforms both Prophet and the single LSTM significantly. The hybrid LSTM-ARIMA models' outperformance has been demonstrated in many cases, suggesting that combining LSTM with ARIMA improves estimation accuracy (Abdulrahman et al., 2021; Zolfaghari and Gholami, 2021; Rehman et al., 2024; Kumar et al., 2016). Nevertheless, the best decomposition technique for time series data remains to be determined. In contrast to the complex current decomposition approaches, Dave et al. (2021) proposed a simpler model for time series decomposition. Their comparative study demonstrated the superiority of the hybrid model, providing important information for policymakers. According to Albeladi et al. (2023), while ARIMA and LSTM models have proven useful in certain contexts, their effectiveness for prediction tasks can be limited by factors such as the unique characteristics of the target domain, the emergence of more advanced predictive techniques, and practical considerations related to data and resources.

Convolutional Neural Networks (CNNs), as shown by Hu (2018), have been effective in predicting time series. While CNNs are mainly used for image recognition, they have also been applied to time series forecasting. However, their performance may not always match models specifically designed for sequential data. Recent research has shown that combining CNN and LSTM models can achieve better results than using either one alone (Kumbure et al., 2022). By combining these two models and applying the CNN model's feature extraction capabilities with the LSTM time series analysis, Eapen et al. (2019) were able to predict the S&P 500 trading price. The hybrid model yielded superior results compared to SVR regression analysis. Moreover, it demonstrated better prediction accuracy than both CNN and LSTM models. Lee and Kim's proposed model architecture employed CNN and ConvLSTM to effectively train patterns across tens of thousands of time-series, serving as the market feature extractor (Lee and Kim, 2020). The outcomes show that all baseline models are outperformed by the suggested model. CNN-LSTM model was proposed by Lu et al. (2020) to predict the closing price of stocks for the next day. The features of the input data are extracted using CNN. Utilizing the retrieved feature data, LSTM is trained to forecast the stock's closing price for the following day. Among many techniques, the suggested CNN-LSTM model showed the highest predicting accuracy. Jing et al. (2021) classified investors' hidden feelings using a hybrid CNN-LSTM model. The experiment's results indicated that the proposed model surpassed the baseline classifiers in recognizing investor attitudes. Moreover, the hybrid approach demonstrated superior performance compared to both individual models and models lacking sentiment analysis, particularly in stock price prediction.

Even though these techniques have been highly beneficial for data analysis, it can still be difficult to analyze and predict some time series since they are so stochastic and variable (Rezaei et al., 2021). CEEMD and EMD algorithms have lately been used in the stock market prediction space, thanks to the advantages of sequential data decomposition into separate frequency spectra. These techniques may be helpful when used in conjunction with deep learning models like LSTM for financial time series analysis (Chen et al., 2019; Niu et al., 2020; Brandi et al., 2020; Lahmiri, 2016) because of their ability to reduce the impact of stock series' nonlinear properties (Xuan et al., 2020). Jothimani and Yadav (2019) and Jin et al. (2020) confirmed the predictive ability of hybrid models using the frequency decomposition algorithm. They also found that CEEMD might produce more precise outcomes compared to EMD when paired with other machine learning and deep learning models. The CEEMDAN-LSTM hybrid model was proposed by Cao et al. (2019), and it was compared with current hybrid models for the S&P 500 and HIS indices. The outcomes confirmed CEEMD-LSTM's superiority over CEEMD-MLP, CEEMD-SVR, and even individual LSTM and SVR models. Yan et al. (2020) proposed the CEEMD-PCA-LSTM deep learning hybrid model. The results showed that the suggested model outperformed benchmark models in terms of forecast accuracy and profitability performance.

The method proposed by Barthélemy et al. (2010), which focuses on disentangling collective trends from local dynamics, has been widely applied in fields requiring the separation of global and local factors within time series. The ability to isolate external trends from internal behaviors aligns with the approach of decomposition techniques in financial time series analysis. While existing studies have explored various hybrid models to enhance stock market predictions, the ETICA-LSTM approach proposed in this paper offers a unique method for decomposing stock prices. By isolating external trends from internal components, ETICA enables the LSTM model to focus on more structured, less noisy data, thus improving prediction accuracy. To our knowledge, this is the first study to integrate ETICA and machine learning in the financial forecasting domain, contributing a novel perspective to hybrid forecasting models. By integrating ETICA with LSTM, this study aims to build upon the success of decomposition-based approaches in stock market prediction, offering a new method for disentangling external trends from internal stock-specific factors.

3. Methodology

This article's primary methodology uses the ETICA decomposition algorithms alongside with deep learning techniques. Meanwhile, it is essential to comprehend what constituting models are and how they work in order to comprehend the suggested strategy. Next, a presentation and investigation of the hybrid algorithm's design will take place.

3.1. Data

The dataset used in this study includes stock prices for the S&P 500, NASDAQ, and Dow Jones indices, covering the period from January 6, 2015, to July 11, 2022, as well as the SSE and SZSE stock indices, which span from January 6, 2015, to October 23, 2024. The data was sourced from Yahoo Finance, and from the available seven columns, only the Date and Closing Price columns were selected for further analysis. In this study, the decomposition method necessitates the following data transformation:

$$p_i(t) = \frac{P_i(t) - P_i(t-1)}{P_i(t-1)} \times 100, \tag{1}$$

here, $P_i(t)$ and $P_i(t-1)$ represent the prices at time instants t and $t-1$, respectively.

3.2. The ETICA method

In general, a set of time series $\{p_i(t)\}_{i=1,\dots,S}(t)$ is considered, where $t = 1, \dots, T$ represents the time period, and i refers to a specific stock within the market. The objective is to decompose the stock returns into components that account for broader market influences and stock-specific factors. This can be achieved by expressing the stock return $p_i(t)$ as:

$$p_i(t) = p_i^{ext}(t) + p_i^{int}(t), \tag{2}$$

The $p_i^{text}(t)$, captures the influence of overall market trends on stock i, including macroeconomic factors and broader indices. In contrast, the $p_i^{int}(t)$, reflects stock-specific factors, such as company performance or industry events. This separation helps differentiate the impact of market-wide trends from individual stock drivers.

Typically, in these techniques, it is presumed that the average of the local components equals zero. On the basis of this assumption, A technique to separate the internal dynamics has been proposed by Argollo de Menezes and Barabási (2004), from which the external components can be computed by the following equation:

$$p_i^{ext}(t) = a_i \sum_{i=1}^{S} p_i(t), \tag{3}$$

where

$$a_i = \frac{\sum_{t=1}^{T} p_i(t)}{\sum_t^T \sum_{i=1}^{S} p_i(t)}, \tag{4}$$

$$p_i^{int}(t) = p_i(t) - \frac{\sum_{t=1}^{T} p_i(t)}{\sum_t^T \sum_{i=1}^{S} p_i(t)} \sum_{i=1}^{S} p_i(t). \tag{5}$$

In specific scenarios, this method is capable of accurately forecasting outcomes. Thus, using an independent component analysis technique, the external trend and internal components analysis decomposition method was presented in Barthélemy et al. (2010).

In essence, the setting is the Arbitrage Pricing Theory (APT), where excessive α or $p_i^{int}(t)$ is the term used. The estimate of a_i is conceptually identical to the more widely used Fama–Macbeth regression techniques, which are commonly used for factor extraction. On the other hand, the ETICA methodology is a different strategy from Fama–Macbeth that offers benefits in terms of money in the APT setting. The separation approach, as outlined in Argollo de Menezes and Barabási (2004), posits that the internal component $p_i^{int}(t)$ inherently possesses a zero average. The requirement that all components of the parameter vector α must be zero stems from its implications for prices. Nevertheless, it is common for the internal contribution average to be non-zero, which leads to inaccurate findings.

To show that internal contributions are independent of the global trend and that correlations between regions mainly come from their shared dependence on this trend, the absence of the following connected correlation is imposing:

$$corr(\omega, p_i^{int}) = 0, \tag{6}$$

and

$$corr(p_i^{int}, p_j^{int}) = 0. \tag{7}$$

Following Argollo de Menezes and Barabási (2004), the possibility of having both multiplicative and additive contributions was considered with the assumption that:

$$p_i^{ext}(t) = a_i \omega(t), \tag{8}$$

thus, the decomposition approach was stated as follows:

$$p_i(t) = a_i \omega(t) + p_i^{int}(t), \tag{9}$$

with each stock responding to the overall trend using the prefactor a_i, the collective trend is represented by $\omega(t)$. The average of $\omega(t)$ is denoted by μ_ω, and its dispersion is denoted by σ_ω, so that:

$$\omega(t) = \mu_\omega + \sigma_\omega W(t). \tag{10}$$

By denoting: $P_i(t) = p_i(t) - \langle p_i \rangle$ and $G_i = p_i^{int} - \langle p_i^{int} \rangle$:

$$P_i(t) = A_i W(t) + G_i(t), \tag{11}$$

with

$$A_i = a_i \sigma_\omega. \tag{12}$$

According to Barthélemy et al. (2010), $G_i = P_i(t) - A_i W(t)$ was computed and the following was obtained:

$$\langle p_i \rangle = A_i \frac{\mu_\omega}{\sigma_\omega} + \langle p_i^{int} \rangle. \tag{13}$$

In two scenarios, $\frac{\mu_\omega}{\sigma_\omega}$ is estimated, with the first scenario assuming the absence of internal contributions, it was imposed that:

$$\frac{\mu_\omega}{\sigma_\omega} = \frac{1}{S}\sum_i \frac{\langle p_i \rangle}{A_i}, \tag{14}$$

where

$$\langle p_i \rangle = \frac{1}{T}\sum_t^T p_i(t), \tag{15}$$

Or by an alternative assumption:

$$\frac{\mu_\omega}{\sigma_\omega} = \frac{\langle p^{av}\rangle \overline{A}}{\overline{A}^2}, \tag{16}$$

where

$$\langle p^{av} \rangle = \frac{1}{S}\sum_i p_i \tag{17}$$

and

$$\overline{A} = \frac{1}{S}(A_i). \tag{18}$$

In this case μ_ω, and σ_ω can be fixed to:

$$\mu_\omega = \langle p^{av} \rangle, \tag{19}$$

and

$$\sigma_\omega = \langle W p^{av} \rangle, \tag{20}$$

where $W(t)$ is the global normalized pattern.

The second scenario assumes that there exists no correlation between A_i's and the temporal average of $p_i^{int}(t)$'s. This parameter may be determined by calculating the slope of a linear correlation that can be found from Eq. (13). In order to address scenarios with strong correlations, both positive and negative, the following novel approach is suggested by:

$$corr(A_i, \langle p_i^{int} \rangle) = \pm 1. \tag{21}$$

This definition necessitates the presence of parameters a and b:

$$\langle p_i^{int} \rangle = aA_i + b, \tag{22}$$

By replacing $\langle p_i^{int} \rangle$ in Eq. (13), we get:

$$\langle p_i \rangle = \left(\frac{\mu_\omega}{\sigma_\omega} + a\right) A_i + b, \tag{23}$$

Without any constraint on a and $\frac{\mu_\omega}{\sigma_\omega}$, their independent separation is not feasible (obtaining $\frac{\mu_\omega}{\sigma_\omega} + a$ is possible through linear regression). Thus, when expressing that the correlation equals ± 1 and under the assumption:

$$A_i = \pm \langle p_i^{int} \rangle, \tag{24}$$

We get then

$$\frac{\mu_\omega}{\sigma_\omega} = \frac{1}{S}\sum_i \left(\frac{\langle p_i \rangle}{A_i} - 1\right)(corr = 1), \tag{25}$$

and

$$\frac{\mu_\omega}{\sigma_\omega} = \frac{1}{S}\sum_i \left(\frac{\langle p_i \rangle}{A_i} + 1\right)(corr = -1), \tag{26}$$

while considering that choosing $\mu = 1$ is thought to have no negative impacts, as discussed in Barthélemy et al. (2010).

3.3. LSTM model

Recurrent neural networks (RNNs) are designed to process sequential data using gates that retain prior inputs (Chung et al., 2014). However, they struggle with long sequences due to the vanishing (Hochreiter, 1998) and exploding gradient problems (Bengio et al., 1994). To address these issues, Long Short-Term Memory (LSTM) networks were developed, offering enhanced memory capabilities for sequence prediction (Le et al., 2019). LSTMs are particularly effective at capturing patterns over long sequences compared to traditional RNNs and feed-forward neural networks. Stock market prediction is challenging due to its non-linear and volatile nature. This study introduces a hybrid model that combines ETICA and LSTM to improve prediction accuracy. ETICA decomposes stock prices into external trends and internal factors, providing cleaner data. LSTM, with its ability to capture long-term dependencies, processes these components separately. Each LSTM cell uses forget, input, and output gates (illustrated in Figs. 1–4 and detailed in Eqs. (27)–(29)) to manage information.

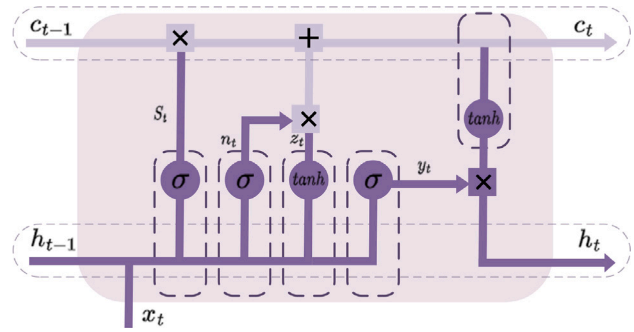

Fig. 1. LSTM model architecture.

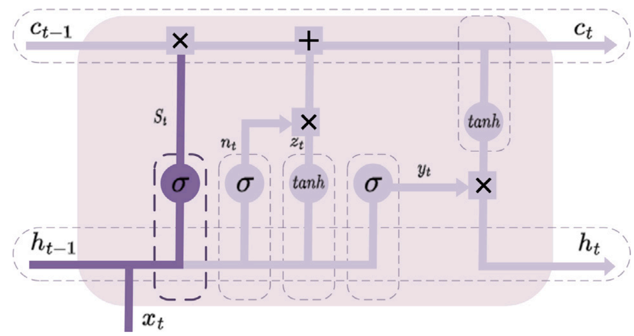

Fig. 2. Forget gate of LSTM cell.

The forget gate in an LSTM decides whether to retain or discard information from the previous time step, based on the current input and previous hidden state. It produces a value between 0 and 1, using weights D and bias g.

$$s_t = \sigma(D_s \times [h_{t-1}, x_t] + g_s). \tag{27}$$

The input gate applies a sigmoid function to the current input and previous hidden state, producing a value between 0 and 1. A tanh function is then applied, and the result is combined with the input to update the cell state.

$$\begin{aligned} n_t &= \sigma(D_n \times [h_{t-1}, x_t] + g_n), \\ z_t &= tanh(D_z \times [h_{t-1}, x_t] + g_z). \end{aligned} \tag{28}$$

The output gate applies a sigmoid to the input and previous hidden state, and a tanh to the new cell state. These are multiplied to determine the next hidden state, which, along with the updated cell state, moves to the next time step.

$$\begin{aligned} y_t &= \sigma(D_y \times [h_{t-1}, x_t] + g_y), \\ h_t &= y_t \times \tanh(c_t). \end{aligned} \tag{29}$$

LSTM's ability to retain and forget information across time steps makes it particularly suitable for the data produced by the ETICA

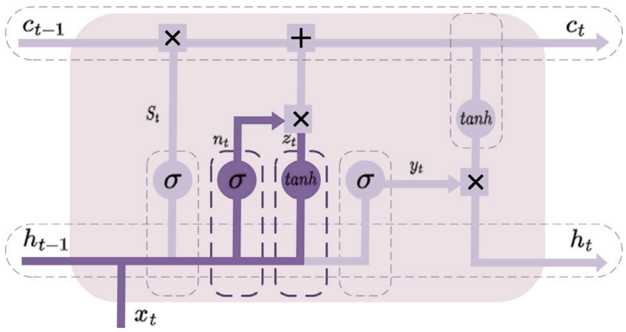

Fig. 3. Input gate of LSTM cell.

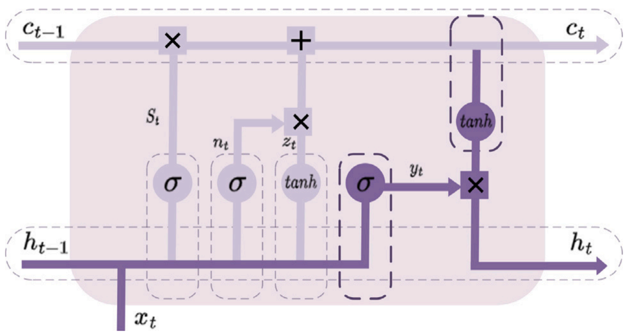

Fig. 4. Output gate of LSTM cell.

decomposition. While ETICA breaks the data into internal components and external trends, LSTM processes these decomposed elements over time, learning the temporal dependencies that exist within these components, which enhances predictive accuracy.

3.4. Hybrid model ETICA-LSTM

Algorithm 1 ETICA-LSTM model

1: Padding Null Values: any missing values are padded to ensure there are no gaps in the dataset
2: $P_i(t)$ transformation via Eq. (1) to get $p_i(t)$
3: **ETICA** Input: $p_i(t)$
4: **ETICA** Output: $p_i^{ext}(t)$, $p_i^{int}(t)$
5: $p_i^{ext}(t)$, $p_i^{int}(t)$ reverse transformation to get $P_i^{ext}(t)$ and $P_i^{int}(t)$ using Eqs. (31)–(32)
6: Normalise and prepare the $P_i(t)$, $P_i^{ext}(t)$, $P_i^{int}(t)$ for **LSTM** via Min-Max normalization function using Eq. (33)
7: Split data into training and testing sets
8: Train **LSTM** model using the selected features
9: Predict future stock prices also their internal and external components via the trained **LSTM** model
10: Evaluate the model's performance using standard metrics **RMSE** and **MAE**

In this study, a model incorporating the ETICA decomposition method alongside the LSTM model was proposed. As shown in Algorithm 1, the hybrid model first applies ETICA to decompose the stock data into distinct internal components and external trends. After ETICA decomposition, both internal components and external trends are processed separately using LSTM. This separation allows the model to capture both the external trends and internal components, ultimately improving prediction accuracy compared to using LSTM alone.

The integration of the ETICA with the LSTM aims to enhance the predictive accuracy by separating the internal components from the external trends. By merging these techniques, the model aims to improve prediction performance while offering valuable insights into the underlying factors driving the predictions. Fig. 5 outlines a hybrid model that combines ETICA and LSTM for stock price prediction. The process begins with stock price data, which is preprocessed by filling any missing values. Next, the ETICA method decomposes the data into important components, and the data is normalized to ensure consistency. The ETICA method is used to separate internal components from external trends. As proposed in Barthélemy et al. (2010), this method can be applied to rate-based data. Therefore, stock prices have been transformed into daily rates via Eq. (1). After calculating the rates for both internal and external components p_i^{int} and p_i^{ext}, we reverse the transformation to convert the data back into price format, as the LSTM model performs better with actual price values. To do this, we use the equation:

$$P_i(t) = \left[\frac{p_i^{int}(t) + p_i^{ext}(t)}{100} \times P_i(t-1)\right] + P_i(t-1). \tag{30}$$

We split $P_i(t-1)$ between the internal and external components to avoid disproportionately affecting the importance of either component:

$$P_i^{int}(t) = \left[\frac{p_i^{int}(t)}{100} \times P_i(t-1)\right] + \frac{P_i(t-1)}{2}, \tag{31}$$

and

$$P_i^{ext}(t) = \left[\frac{p_i^{ext}(t)}{100} \times P_i(t-1)\right] + \frac{P_i(t-1)}{2}. \tag{32}$$

This format allows for proper normalization, which is essential for the LSTM model to prevent gradient errors and ensure stable weight values during training. By transforming the rates back into prices, we retain the model's predictive power while ensuring consistency in the data format for both internal and external components. In conclusion, our approach of transforming data into rates for ETICA decomposition and then back into prices for LSTM prediction ensures a more accurate and stable model.

The next step is featuring scaling, which uses Eq. (33) to normalize the data and convert it into a desired range, usually 0 to 1, which ensures consistency across various ranges:

$$Pscaled = \frac{P - P_{min}}{P_{max} - P_{min}}. \tag{33}$$

Because normalization reduces the likelihood of significant gradient errors, unstable weight values in the LSTM model are avoided. All dataset can be scaled to a range between 0 and 1 by using MinMax Scaler from the sklearn preprocessing module. MinMax is suggested since it can maintain the shape of the dataset without causing any distortions. Proceeding forward, the data is divided into training and testing partitions to continue with the analysis. This is accomplished by employing a 60:40 ratio, which designates that 40% of the data are used for testing and the remaining 60% are used to train the LSTM model. From 1890 data, 1134 are used for training and 756 for testing. With a step value of 15, the data is transformed into time steps in order to get it ready for the LSTM model.

Across 200 epochs, the model training employed a batch size of 32. In order to train this configuration, the data is divided into batches of 32 samples each, and the dataset is iterated 200 times in total. Predictions are produced utilizing the training and testing datasets after training. The projected data is first converted back to its original scale using the *inversetransform ()* function in order to properly assess the predictions. To measure accuracy, one can directly compare the rescaled data against the original data. The evaluation of the model's performance employs metrics including RMSE and MAE.

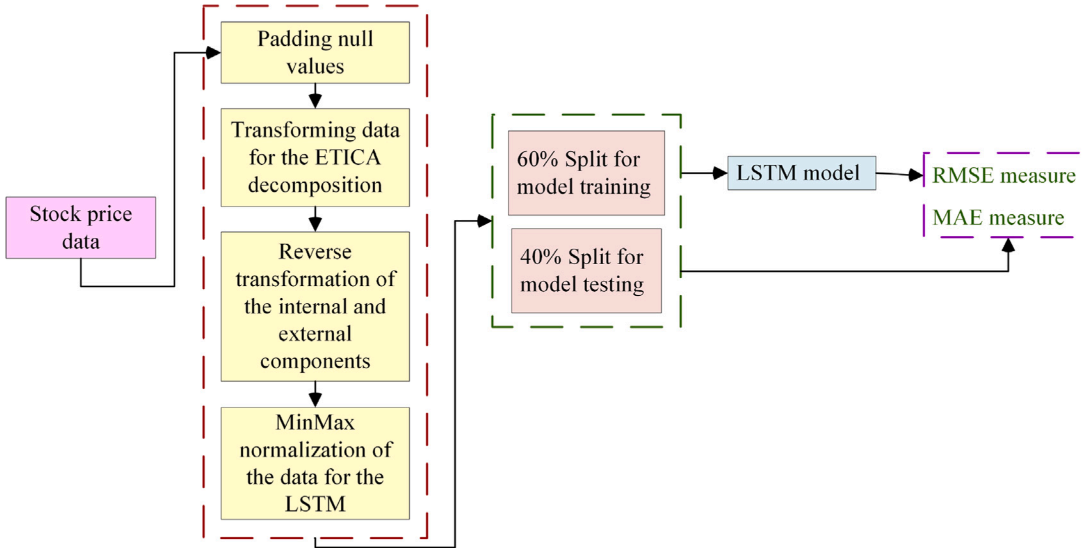

Fig. 5. Proposed model ETICA-LSTM.

Table 1
Descriptive statistics of S&P 500, NASDAQ and Dow Jones indices daily closing price.

Index	Count	Mean	Min	Max	Std
S&P 500	1891	2822.18	1829.08	4796.56	800.04
NASDAQ	1891	8365.79	4266.84	16 057.44	3338.54
Dow Jones	1891	24 208.85	15 660.18	36 799.65	5835.05
SSE	2383	3257.68	2464.36	5166.35	287.88
SZSE	2383	10 813.75	7684	17 399	2474.45

3.5. Assessment metrics

The performance of deep learning model predictions is commonly measured by loss error. It describes the gap between the actual observed values and the predicted values. RMSE and MAE are among the standard metrics employed in this study for such models.

RMSE: This metric is the square root of the mean squared errors between the actual and predicted outcomes, indicating how well the predictions align with the actual data. This measure is defined by the following expression:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(P_i - \hat{P}_i)^2}. \tag{34}$$

MAE: This metric is calculated as the average of the absolute deviations between the estimated and observed values. Its expression is:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|P_i - \hat{P}_i|. \tag{35}$$

4. Results and discussion

4.1. Data pre-processing

The dataset utilized in this study comprises stock prices for the S&P 500, NASDAQ, and Dow Jones indices, covering the period from January 6, 2015, to July 11, 2022. Additionally, it includes data for the SSE and SZSE stock indices, spanning from January 6, 2015, to October 23, 2024. To optimize the training process, the data underwent normalization via the MinMax scaling technique. During the model compilation, the loss function selected was mean squared error, and the optimizer chosen was Adam. Table 1 provides the descriptive statistics for the daily closing prices of the following stock indices: S&P 500, NASDAQ, Dow Jones, SSE and SZSE.

4.2. Results

The suggested algorithm prediction results are presented in this section. The suggested model, which is explained in detail in the section above, is a unique hybrid method that combines LSTM and ETICA decomposition. The ETICA technique is first used to split the time series into internal components and external trends, which are then predicted individually using the LSTM model. The accuracy of the model's prediction is assessed by utilizing the RMSE and MAE metrics to compare predicted values with original values. From Table 2 it can be observed that:

$$Rmse(P_i(t)) > Rmse(P_i^{int}(t)), \tag{36}$$

and

$$Rmse(P_i(t)) > Rmse(P_i^{ext}(t)). \tag{37}$$

By summing both equations:

$$2 \times Rmse(P_i(t)) > Rmse(P_i^{int}(t)) + Rmse(P_i^{ext}(t)), \tag{38}$$

so

$$Rmse(P_i(t)) > \left[Rmse(P_i^{int}(t)) + Rmse(P_i^{ext}(t))\right]/2. \tag{39}$$

Table 2 compares the performance of the ETICA-LSTM and LSTM models in predicting the S&P, NASDAQ, Dow Jones, SSE, and SZSE indices, based on RMSE and MAE. In the training phase, ETICA-LSTM outperforms LSTM for the S&P 500 index, with a lower RMSE (17.3837 vs. 22.3655) and MAE (12.2970 vs. 15.6273). Similarly, for the NASDAQ index, ETICA-LSTM shows significantly lower RMSE (46.7012 vs. 72.6584) and MAE (33.1337 vs. 50.8848), indicating better accuracy. For the Dow Jones index, ETICA-LSTM again performs better with lower RMSE (152.6741 vs. 192.6454) and MAE (106.0264 vs. 132.9207). Additionally, for the SSE index, ETICA-LSTM improves performance with a lower RMSE (31.8371 vs. 34.5248) and MAE (23.2217 vs. 24.6561). Similarly, for the SZSE index, ETICA-LSTM shows better accuracy with lower RMSE (80.11205 vs. 179.2499) and MAE (56.9475 vs. 127.1322).

In the testing phase, ETICA-LSTM consistently demonstrates superior performance. For the S&P 500, the model achieves a lower RMSE (96.3502 vs. 143.2097) and MAE (70.1420 vs. 103.5561). Likewise, for the NASDAQ index, ETICA-LSTM shows a significant reduction in RMSE (278.276 vs. 536.6437) and MAE (225.7819 vs. 393.9449). For the Dow Jones, ETICA-LSTM also outperforms LSTM with a lower RMSE

Table 2
The RMSE and MAE results for prices, internal components, and external trends for S&P500, NASDAQ, Dow Jones SSE AND SZSE indices.

Metrics	Index	Prices		Internal components		External trend	
		Train	Test	Train	Test	Train	Test
RMSE	S&P 500	22.3655	143.2097	13.0655	87.57098	21.70192	105.1294
	NASDAQ	72.6584	536.6437	34.9628	362.2687	58.4395	194.2832
	Dow Jones	192.6454	457.5253	122.0281	454.3025	183.3201	427.1806
	SSE	34.5248	49.5867	24.8098	31.9203	38.8644	52.9558
	SZSE	179.2499	203.6735	85.1974	99.3579	75.0267	93.1630
MAE	S&P 500	15.6273	103.5561	9.4493	63.6573	15.1447	76.6267
	NASDAQ	50.8848	393.9449	25.8360	303.8989	40.4314	147.6648
	Dow Jones	132.9207	350.1138	85.8132	342.9603	126.2396	302.0947
	SSE	24.6561	30.5344	18.6487	22.3041	27.7947	35.0905
	SZSE	127.1322	134.6148	60.7412	66.4097	53.1538	60.4220

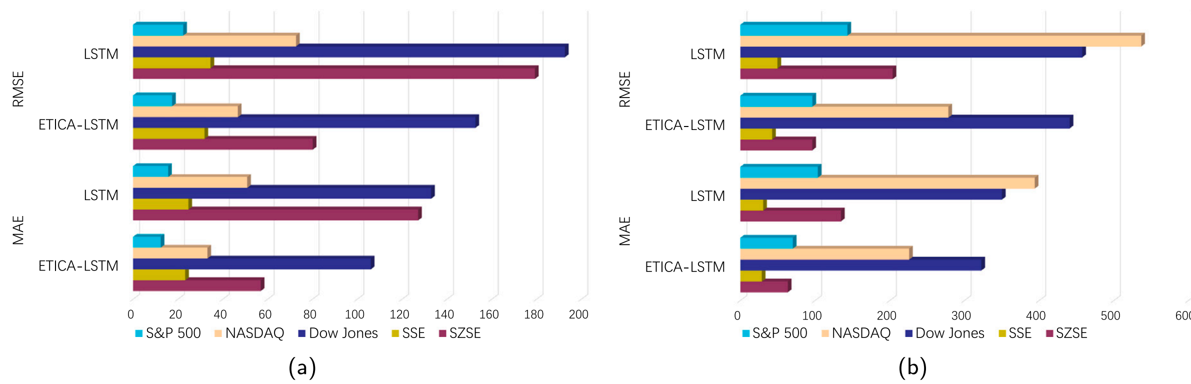

Fig. 6. Training and testing RMSE and MAE values.

(440.7416 vs. 457.5253) and MAE (322.5275 vs. 350.1138). Furthermore, for the SSE index, ETICA-LSTM shows better test results with a lower RMSE (42.43805 vs. 49.5867) and MAE (28.6973 vs. 30.5344). For the SZSE index, ETICA-LSTM again improves prediction accuracy, with a lower RMSE (96.26045 vs. 203.6735) and MAE (63.41585 vs. 134.6148). Overall, ETICA-LSTM consistently demonstrates superior accuracy and predictive performance compared to the LSTM model across all five indices in both the training and testing phases.

Fig. 6 compare the performance of the ETICA-LSTM and LSTM models in predicting the S&P 500, NASDAQ, Dow Jones, SSE and SZSE indices based on RMSE and MAE values during the training and testing phases. In the training phase (Fig. 6(a)), ETICA-LSTM consistently outperforms LSTM across all indices, showing lower RMSE and MAE values, indicating better accuracy and fit to the training data. Similarly, in the testing phase (Fig. 6(b)), ETICA-LSTM again outperforms LSTM with significantly lower RMSE and MAE values, demonstrating superior generalization to new, unseen data. This consistent performance across both phases highlights that the ETICA-LSTM model not only fits the training data better but also performs more accurately on the testing data, making it a more reliable model for predicting the S&P 500, NASDAQ, Dow Jones, SSE and SZSE indices compared to the standard LSTM model.

In both training and testing phases, the proposed model exhibits lower RMSE and MAE values based on the S&P 500, NASDAQ, Dow Jones, SSE, and SZSE indices, signifying that it generally produces smaller prediction errors compared to the LSTM model. A lower RMSE indicates that the proposed model has fewer large errors, minimizing the magnitude of significant deviations from the actual values. Meanwhile, a lower MAE signifies that the proposed model has fewer overall errors, consistently providing accurate predictions. This dual reduction in error metrics highlights the superior predictive accuracy and reliability of the proposed model.

The robustness and generalization capacity of the model can be evaluated by testing it with different proportions of training data and varying the number of epochs. A resilient model should maintain its performance despite changes in the training arrangement. In this study, the RMSE and MAE were assessed for epochs set to 200, 100, 75, 50, and 25. The results are presented in Table 3 for scenarios where 60% of the data was used for training, and in Table 4 for scenarios where 80% of the data was used for training using S&P 500 index. As shown in these tables, the RMSE and MAE values consistently remain lower on average when the decomposition method is integrated with the LSTM model. Regardless of the variations in epochs and training data proportions, the LSTM model demonstrates stable and continuous performance. This consistency confirms the model's efficacy and dependability for prediction, highlighting its strong generalization capability, well-chosen initial training parameters, and stability in handling various data patterns. Notably, the comparison of several performance metrics clearly indicates that the proposed ETICA-LSTM model is more precise than the single LSTM model. Specifically, this study reveals that the ETICA-LSTM algorithm significantly improves the effectiveness of predicting S&P 500, NASDAQ and Dow Jones price indices.

Figs. 7–8 display the actual and predicted values of various stock indices S&P 500, NASDAQ, Dow Jones, SSE, and SZSE using the ETICA-LSTM model. In these figures, black lines represent actual values, pink lines denote predictions on training data, and cyan lines indicate predictions on testing data. Across all indices, the ETICA-LSTM model demonstrates strong predictive accuracy, closely aligning with actual values and effectively capturing overall trends. For the S&P 500, NASDAQ, Dow Jones, SSE and SZSE price indices, the model shows a high degree of accuracy, tracking both upward and downward movements with minimal error, even in testing phases. Similarly, for the SSE and SZSE indices, the model effectively captures both the trend direction and general volatility, though minor discrepancies appear in periods of high volatility, particularly during sharp declines or rebounds. These results highlight the ETICA-LSTM model's robust performance across different stock markets, illustrating its capability to forecast both trend patterns and short-term fluctuations in market behavior.

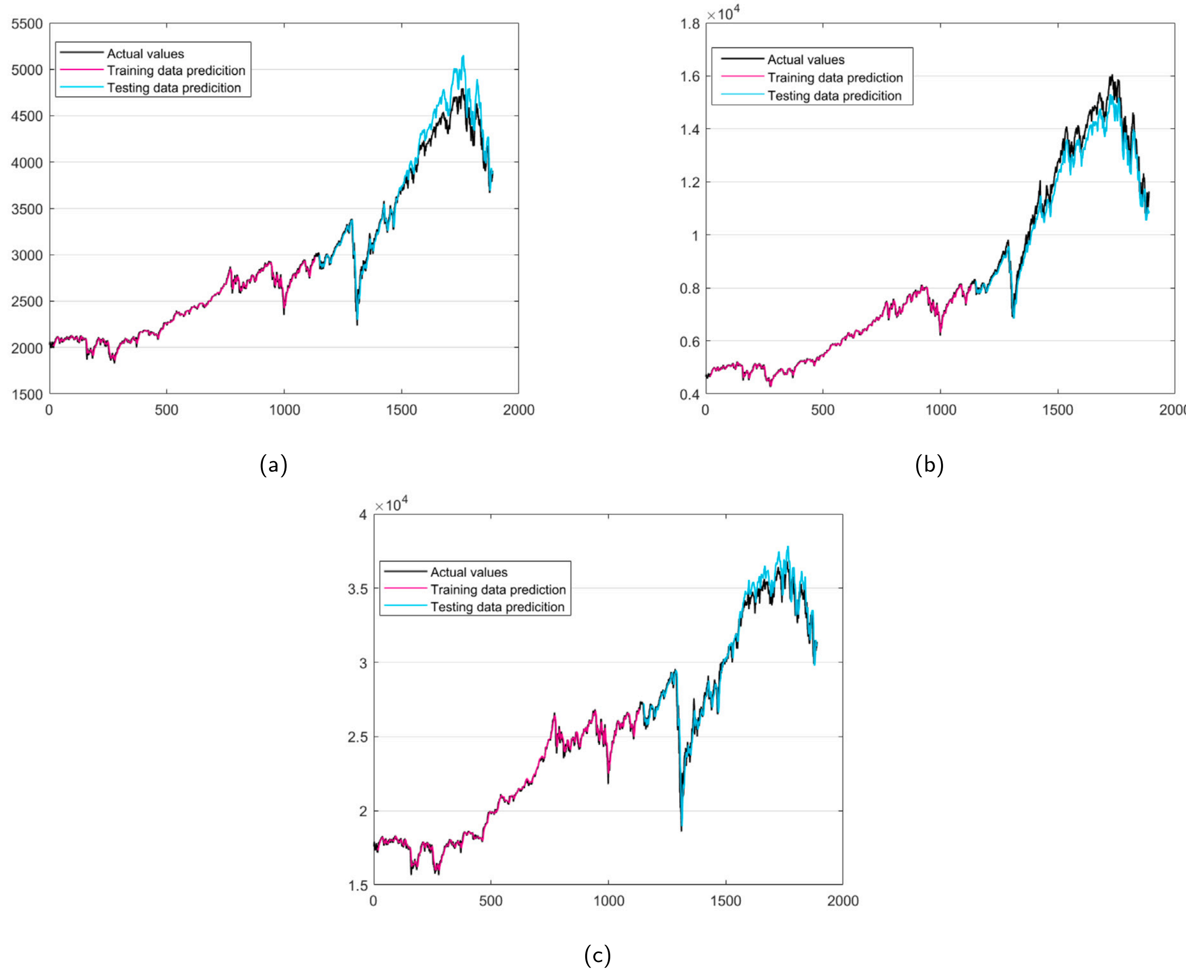

Fig. 7. Actual and predicted values of S&P 500, NASDAQ and Dow Jones indices.

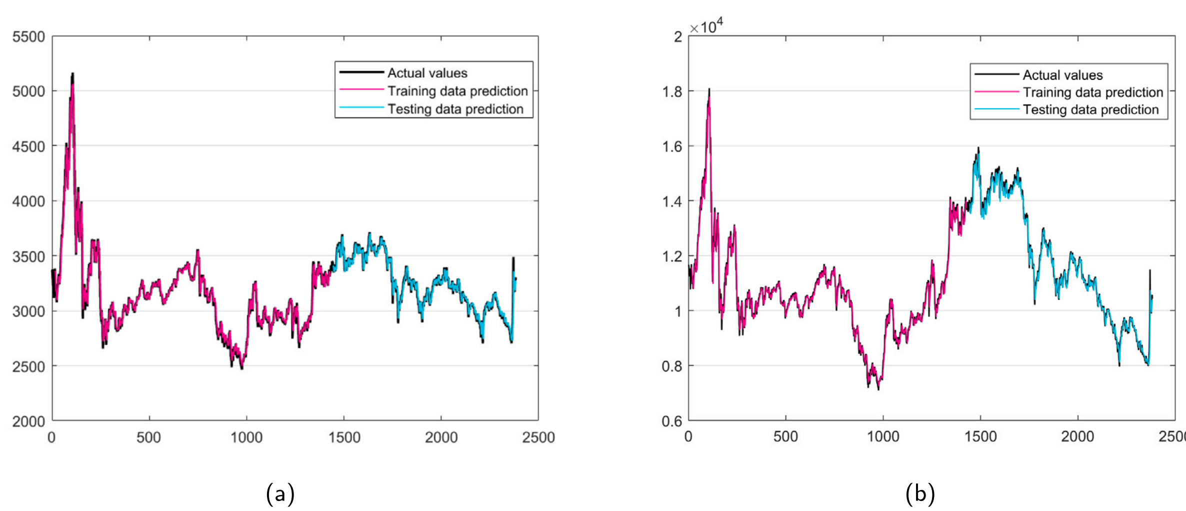

Fig. 8. Actual and predicted values of SSE and SZSE indices.

The results presented in Table 5 demonstrate that the ETICA-LSTM model consistently outperforms the other models (LSTM, CEEMDAN-LSTM, and EMD-LSTM) in predicting the S&P 500, NASDAQ, Dow Jones, SSE and SZSE indexes, based on both RMSE and MAE metrics. The ETICA-LSTM model achieves the lowest RMSE and MAE values across all indices, indicating its superior ability to minimize large errors. This suggests that the ETICA-LSTM model captures the underlying market trends and patterns more accurately than the other models. CEEMDAN-LSTM also performs reasonably well, but it still falls short of ETICA-LSTM in terms of accuracy. Both the standard LSTM

Table 3
The RMSE and MAE results for different epochs using 60% of the S&P500 training data.

Metrics	Epochs	Prices		Internal components		External trend	
		Train	Test	Train	Test	Train	Test
RMSE	200	22.36551	143.2097	13.06547	87.57098	21.70192	105.1294
	100	25.24784	474.9842	14.44803	82.41494	22.1677	103.9955
	75	25.05909	754.0312	15.97812	123.849	22.32238	85.18125
	50	31.05927	454.2541	14.30924	74.01685	22.10692	166.6111
	25	43.9962	712.2969	16.41208	114.1639	24.2606	129.6719
MAE	200	15.62731	103.5561	9.44937	63.65733	15.14469	76.62671
	100	17.86555	330.7606	10.16471	69.22759	15.40807	78.30137
	75	17.45219	542.768	11.16165	107.1391	15.61197	67.5348
	50	22.11518	308.4285	10.37416	61.44397	15.14839	139.3365
	25	32.58058	498.6752	11.93527	84.12145	16.99684	96.87104

Table 4
The RMSE and MAE results for different epochs using 80% of the S&P500 training data.

Metrics	Epochs	Prices		Internal components		External trend	
		Train	Test	Train	Test	Train	Test
RMSE	200	35.53548	190.7905	18.1127	34.43347	31.99144	53.94969
	100	35.81589	84.34714	19.03031	53.72668	30.63958	54.34137
	75	35.47269	198.8299	20.92899	46.47105	30.78812	55.50921
	50	36.87529	132.7711	25.86007	90.11835	31.2322	57.3451
	25	41.69296	143.7354	21.38984	50.48286	31.22885	58.26161
MAE	200	23.30286	152.4025	11.28833	27.14972	19.35455	40.54725
	100	24.74106	58.06614	11.98817	44.41505	19.28905	40.62808
	75	23.75857	146.0501	13.16748	34.98189	19.26047	41.22664
	50	23.78093	103.827	15.78541	76.89176	19.5384	43.57039
	25	26.67055	116.4197	13.24014	37.48881	19.28638	44.75618

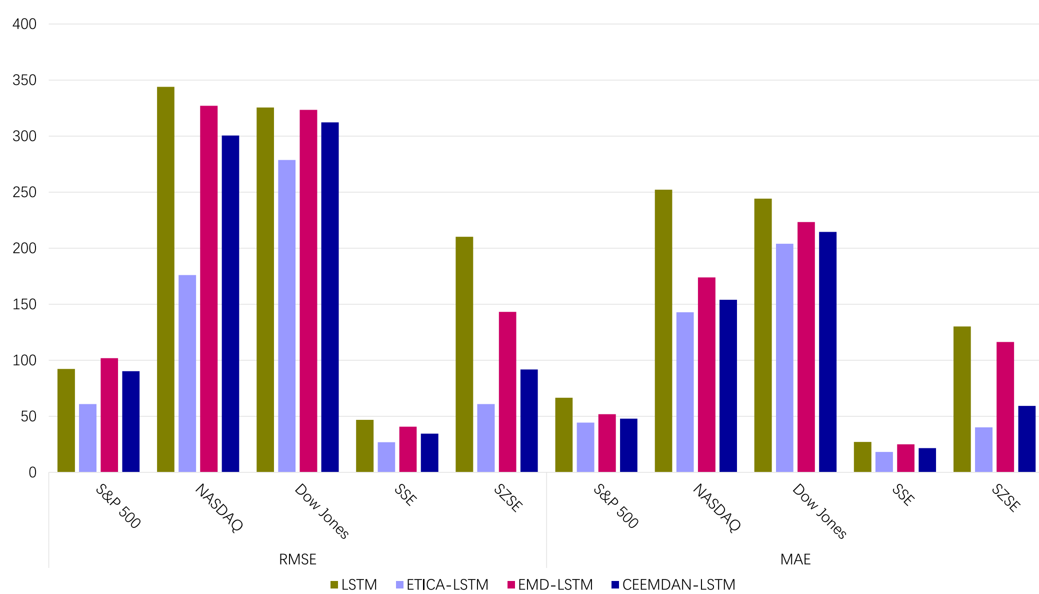

Fig. 9. RMSE and MAE comparison.

and EMD-LSTM models exhibit higher error rates, which implies that they are less effective in handling the complexities of stock market data, particularly in comparison to the decomposition-based models. Overall, the findings highlight that ETICA-LSTM is the most robust and reliable model for stock market prediction, benefiting from the added precision of the ETICA decomposition technique. In summary, based on these results, ETICA-LSTM stands out as the best performing model for predicting stock indices, followed by the standard LSTM model. The decomposition-based models (CEEMDAN-LSTM and EMD-LSTM) show higher error rates, indicating that while decomposition is helpful, ETICA might be better at capturing stock market patterns than the other decomposition techniques.

The RMSE and MAE comparison in Fig. 9 demonstrates that the ETICA-LSTM model consistently outperforms the other models across various stock indices, achieving the lowest errors in both metrics. For all indices, ETICA-LSTM exhibits lower RMSE and MAE values compared to the standalone LSTM model, as well as the EMD-LSTM and CEEMDAN-LSTM hybrid models, indicating its superior accuracy and reliability in time series forecasting. While CEEMDAN-LSTM and EMD-LSTM also reduce error compared to LSTM, they do not match the error reduction achieved by ETICA-LSTM. These results suggest that ETICA-LSTM is particularly well-suited for stock market prediction, as it minimizes both the root mean squared error and mean absolute error, providing a more accurate and consistent forecasting performance.

Table 5
Comparison of RMSE and MAE for LSTM, ETICA-LSTM, CEEMDAN, and EMD-LSTM.

Metrics	Index	LSTM	ETICA-LSTM	EMD-LSTM	CEEMDAN-LSTM
RMSE	S&P 500	92.2042	60.9296	101.8114	90.2642
	NASDAQ	343.9945	175.9742	327.0636	300.5837
	Dow Jones	325.5471	278.7129	323.4276	312.3018
	SSE	46.7960	26.8433	40.7021	34.4517
	SZSE	210.1326	60.8872	143.1594	91.7752
MAE	S&P 500	66.5955	44.3562	51.8878	47.8889
	NASDAQ	252.2192	142.7784	173.9248	153.9363
	Dow Jones	244.1752	203.9577	223.3155	214.4895
	SSE	27.1615	18.1522	24.9971	21.6138
	SZSE	130.1781	40.1122	116.2839	59.3111

5. Conclusion

To improve stock market predictions for the S&P 500, NASDAQ, Dow Jones, SSE and SZSE indices, this study presents a novel hybrid model that combines the LSTM model with the ETICA decomposition method. Extensive experiments demonstrate that the ETICA-LSTM model significantly outperforms the single LSTM model, achieving lower RMSE and MAE values, indicating improved predictive accuracy and reduced error margins. These results underscore the robustness and reliability of the ETICA-LSTM model, offering potential for transformative applications in investment strategies, risk management, and portfolio performance optimization. Future research could extend this hybrid model to other financial indices and integrate additional machine learning techniques for further accuracy and reliability enhancements.

CRediT authorship contribution statement

Fatene Dioubi: Writing – review & editing, Writing – original draft, Validation, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Negalign Wake Hundera:** Writing – review & editing, Visualization, Validation, Supervision. **Huiying Xu:** Visualization, Validation, Supervision, Project administration, Funding acquisition. **Xinzhong Zhu:** Visualization, Validation, Supervision, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (62376252); Key Project of Zhejiang Provincial Natural Science Foundation (LZ22F030003).

References

Abdulrahman, Umar Farouk Ibn, Ussiph, Najim, Hayfron-Acquah, Benjamin, 2021. A hybrid arima-LSTM model for stock price prediction. Int. J. Comput. Eng. Inf. Technol. 12 (8), 48–51.

Ahmadi, Elham, Jasemi, Milad, Monplaisir, Leslie, Nabavi, Mohammad Amin, Mahmoodi, Armin, Jam, Pegah Amini, 2018. New efficient hybrid candlestick technical analysis model for stock market timing on the basis of the Support Vector Machine and Heuristic Algorithms of Imperialist Competition and Genetic. Expert Syst. Appl. 94, 21–31.

Albeladi, Khulood, Zafar, Bassam, Mueen, Ahmed, 2023. Time series forecasting using LSTM and ARIMA. Int. J. Adv. Comput. Sci. Appl. (IJACSA) 14 (1).

Ali, Muhammad, Khan, Dost Muhammad, Alshanbari, Huda M., El-Bagoury, Abd Al-Aziz Hosni, 2023. Prediction of complex stock market data using an improved hybrid emd-lstm model. Appl. Sci. 13 (3), 1429.

Argollo de Menezes, M., Barabási, A.-L., 2004. Separating internal and external dynamics of complex systems. Phys. Rev. Lett. 93 (6), 068701.

Barthélemy, Marc, Nadal, Jean-Pierre, Berestycki, Henri, 2010. Disentangling collective trends from local dynamics. Proc. Natl. Acad. Sci. 107 (17), 7629–7634.

Bengio, Yoshua, Simard, Patrice, Frasconi, Paolo, 1994. Learning long-term dependencies with gradient descent is difficult. IEEE Trans. Neural Netw. 5 (2), 157–166.

Bhandari, Hum Nath, Rimal, Binod, Pokhrel, Nawa Raj, Rimal, Ramchandra, Dahal, Keshab R., Khatri, Rajendra K.C., 2022. Predicting stock market index using LSTM. Mach. Learn. Appl. 9, 100320.

Bousoño-Calzón, Carlos, Bustarviejo-Muñoz, Josué, Aceituno-Aceituno, Pedro, Escudero-Garzás, José Joaquín, 2019. On the economic significance of stock market prediction and the no free lunch theorem. IEEE Access 7, 75177–75188.

Brandi, Giuseppe, Gramatica, Ruggero, Di Matteo, Tiziana, 2020. Unveil stock correlation via a new tensor-based decomposition method. J. Comput. Sci. 46, 101116.

Bustos, Oscar, Pomares-Quimbaya, Alexandra, 2020. Stock market movement forecast: A systematic review. Expert Syst. Appl. 156, 113464.

Cao, Jian, Li, Zhi, Li, Jian, 2019. Financial time series forecasting model based on CEEMDAN and LSTM. Phys. A 519, 127–139.

Chen, Lu, Chi, Yonggang, Guan, Yingying, Fan, Jialin, 2019. A hybrid attention-based EMD-LSTM model for financial time series prediction. In: 2019 2nd International Conference on Artificial Intelligence and Big Data. ICAIBD, IEEE, pp. 113–118.

Chung, Junyoung, Gulcehre, Caglar, Cho, KyungHyun, Bengio, Yoshua, 2014. Empirical evaluation of gated recurrent neural networks on sequence modeling. arXiv preprint arXiv:1412.3555.

Dave, Emmanuel, Leonardo, Albert, Jeanice, Marethia, Hanafiah, Novita, 2021. Forecasting Indonesia exports using a hybrid model ARIMA-LSTM. Procedia Comput. Sci. 179, 480–487.

Deng, Yamin, Fan, Huifang, Wu, Shiman, 2023. A hybrid ARIMA-LSTM model optimized by BP in the forecast of outpatient visits. J. Ambient Intell. Humaniz. Comput. 1–11.

Eapen, Jithin, Bein, Doina, Verma, Abhishek, 2019. Novel deep learning model with CNN and bi-directional LSTM for improved stock market index prediction. In: 2019 IEEE 9th Annual Computing and Communication Workshop and Conference. CCWC, IEEE, pp. 0264–0270.

Fama, Eugene F., 1995. Random walks in stock market prices. Financ. Anal. J. 51 (1), 75–80.

Hajirahimi, Zahra, Khashei, Mehdi, 2019. Hybrid structures in time series modeling and forecasting: A review. Eng. Appl. Artif. Intell. 86, 83–106.

Hiransha, M.E.A.G., Gopalakrishnan, E. Ab, Menon, Vijay Krishna, Soman, K.P., 2018. NSE stock market prediction using deep-learning models. Procedia Comput. Sci. 132, 1351–1362.

Hochreiter, Sepp, 1998. The vanishing gradient problem during learning recurrent neural nets and problem solutions. Internat. J. Uncertain. Fuzziness Knowledge-Based Systems 6 (02), 107–116.

Hoseinzade, Ehsan, Haratizadeh, Saman, Khoeini, Arash, 2019. U-cnnpred: A universal cnn-based predictor for stock markets. arXiv preprint arXiv:1911.12540.

Hu, Y., 2018. Stock market timing model based on convolutional neural network–a case study of Shanghai composite index. Finance Econ. 4, 71–74.

Huang, Yusheng, Gao, Yelin, Gan, Yan, Ye, Mao, 2021. A new financial data forecasting model using genetic algorithm and long short-term memory network. Neurocomputing 425, 207–218.

Jin, Zhigang, Yang, Yang, Liu, Yuhong, 2020. Stock closing price prediction based on sentiment analysis and LSTM. Neural Comput. Appl. 32, 9713–9729.

Jing, Nan, Wu, Zhao, Wang, Hefei, 2021. A hybrid model integrating deep learning with investor sentiment analysis for stock price prediction. Expert Syst. Appl. 178, 115019.

Jothimani, Dhanya, Yadav, Surendra S., 2019. Stock trading decisions using ensemble-based forecasting models: a study of the Indian stock market. J. Bank. Financ. Technol. 3 (2), 113–129.

Kulshreshtha, Sakshi, et al., 2020. An ARIMA-LSTM hybrid model for stock market prediction using live data. J. Eng. Sci. Technol. Rev. 13 (4).

Kumar, Deepak, Meghwani, Suraj S., Thakur, Manoj, 2016. Proximal support vector machine based hybrid prediction models for trend forecasting in financial markets. J. Comput. Sci. 17, 1–13.

Kumbure, Mahinda Mailagaha, Lohrmann, Christoph, Luukka, Pasi, Porras, Jari, 2022. Machine learning techniques and data for stock market forecasting: A literature review. Expert Syst. Appl. 197, 116659.

Kurani, Akshit, Doshi, Pavan, Vakharia, Aarya, Shah, Manan, 2023. A comprehensive comparative study of artificial neural network (ANN) and support vector machines (SVM) on stock forecasting. Ann. Data Sci. 10 (1), 183–208.

Lahmiri, Salim, 2016. Intraday stock price forecasting based on variational mode decomposition. J. Comput. Sci. 12, 23–27.

Le, Xuan-Hien, Ho, Hung Viet, Lee, Giha, Jung, Sungho, 2019. Application of long short-term memory (LSTM) neural network for flood forecasting. Water 11 (7), 1387.

Lee, Si Woon, Kim, Ha Young, 2020. Stock market forecasting with super-high dimensional time-series data using ConvLSTM, trend sampling, and specialized data augmentation. Expert Syst. Appl. 161, 113704.

Lu, Wenjie, Li, Jiazheng, Li, Yifan, Sun, Aijun, Wang, Jingyang, 2020. A CNN-LSTM-based model to forecast stock prices. Complexity 2020, 1–10.

Murphy, John J., 1999. Technical Analysis of the Financial Markets: A Comprehensive Guide to Trading Methods and Applications. Penguin.

Nabipour, Mojtaba, Nayyeri, Pooyan, Jabani, Hamed, Mosavi, Amir, Salwana, Ely, S, Shahab, 2020. Deep learning for stock market prediction. Entropy 22 (8), 840.

Niu, Hongli, Xu, Kunliang, Wang, Weiqing, 2020. A hybrid stock price index forecasting model based on variational mode decomposition and LSTM network. Appl. Intell. 50, 4296–4309.

Rehman, Ziad Ur, Boulaaras, Salah, Jan, Rashid, Ahmad, Imtiaz, Bahramand, Salma, 2024. Computational analysis of financial system through non-integer derivative. J. Comput. Sci. 75, 102204.

Rezaei, Hadi, Faaljou, Hamidreza, Mansourfar, Gholamreza, 2021. Stock price prediction using deep learning and frequency decomposition. Expert Syst. Appl. 169, 114332.

Temur, Ayse Soy, Akgün, Melek, Temur, Gunay, 2019. Predicting housing sales in Turkey using ARIMA, LSTM and hybrid models. J. Bus. Econ. Manag. 20 (5), 920–938.

Xuan, Yuze, Yu, Yue, Wu, Kaisu, 2020. Prediction of short-term stock prices based on EMD-LSTM-CSI neural network method. In: 2020 5th IEEE International Conference on Big Data Analytics. ICBDA, IEEE, pp. 135–139.

Yan, Binbin, Aasma, Memon, et al., 2020. A novel deep learning framework: Prediction and analysis of financial time series using CEEMD and LSTM. Expert Syst. Appl. 159, 113609.

Zolfaghari, Mehdi, Gholami, Samad, 2021. A hybrid approach of adaptive wavelet transform, long short-term memory and ARIMA-GARCH family models for the stock index prediction. Expert Syst. Appl. 182, 115149.