

RESEARCH

Open Access



Implementation of deep learning models in predicting ESG index volatility

Hum Nath Bhandari¹, Nawa Raj Pokhrel², Ramchandra Rimal^{3*} , Keshab R. Dahal⁴ and Binod Rimal⁵

*Correspondence:
ramchandra.rimal@mtsu.edu

¹ Department of Mathematics,
Roger Williams University, 1 Old
Ferry Rd., Bristol, RI 02809, USA

² Computer Science and Physics
Department, Xavier University
of Louisiana, 1 Drexel Dr, New
Orleans, LA 70125, USA

³ Department of Mathematical
Sciences, Middle Tennessee
State University, 1301 E Main St,
Murfreesboro, TN 37132, USA

⁴ Department of Mathematics,
State University of New York
Cortland, 22 Graham Ave,
Cortland, NY 13045, USA

⁵ Department of Mathematics,
The University of Tampa, 401 W
Kennedy Blvd, Tampa, FL 33606,
USA

Abstract

The consideration of environmental, social, and governance (ESG) aspects has become an integral part of investment decisions for individual and institutional investors. Most recently, corporate leaders recognized the core value of the ESG framework in fulfilling their environmental and social responsibility efforts. While stock market prediction is a complex and challenging task, several factors associated with developing an ESG framework further increase the complexity and volatility of ESG portfolios compared with broad market indices. To address this challenge, we propose an integrated computational framework to implement deep learning model architectures, specifically long short-term memory (LSTM), gated recurrent unit, and convolutional neural network, to predict the volatility of the ESG index in an identical environment. A comprehensive analysis was performed to identify a balanced combination of input features from fundamental data, technical indicators, and macroeconomic factors to delineate the cone of uncertainty in market volatility prediction. The performance of the constructed models was evaluated using standard assessment metrics. Rigorous hyperparameter tuning and model-selection strategies were implemented to identify the best model. Furthermore, a series of statistical analyses was conducted to validate the robustness and reliability of the model. Experimental results showed that a single-layer LSTM model with a relatively small number of neurons provides a superior fit with high prediction accuracy relative to more complex models.

Keywords: ESG investing, ESG index, Deep learning, Machine learning, Volatility prediction

Introduction

Traditional investors typically focus on investment returns in terms of profitability and meticulously scrutinize financial reports to determine the best-performing stocks in the market. A recent change in the mindsets of stakeholders and investors also considers the non-financial impacts of investment decisions. Companies are evaluated under a broad spectrum of environmental, social, and governmental (ESG) factors (Clementino and Perkins 2021). Environmental factors mainly focus on natural resources such as energy efficiency, biodiversity, pollution mitigation, water usage, and climate change. Similarly, social components primarily cover the welfare of the society as a whole, including labor standards, wages, benefits, affordable housing, education, workforce diversity, racial

justice, and health safety. Finally, the government strategically manages environmental and social issues such as corporate board composition and overall structure, strategic sustainability and oversight compliance, political contribution and lobbying, bribery, and corruption.

ESG investing has gained tremendous popularity recently, as society expects companies' corporate and social responsibility efforts (Tucker and Jones 2020). To flourish in the long run, financial institutions focus on the risk of investment and return on the portfolio and evaluate whether the companies have embraced the agendas raised by ESG. According to the Morningstar report, investors in the US poured a record \$69.2 billion into ESG funds in 2021, three times higher than in 2020 (CNBCNews, June 5, 2022). US investors had access to more than 550 ESG related mutual funds and exchange traded funds (ETFs) as of June 5, 2022, which is more than double in the past five years. Similarly, in the European market, a total of \$278 billion ESG related ETFs were under management by 2021 (IR Magazine, Jan 18, 2022). From the perspectives of consumers and investors, the global trend of sustainable investing is exponentially increasing. Prominent industry leaders are beginning to acknowledge the importance of ESG by providing the required information to ESG rating agencies, assuring ESG commitment, and issuing sustainability reports to the public.

There is no consensus on a framework for ESG. Several indices and frameworks are available in the market to better guide companies and inform investors. Some dominant international frameworks include the Global Reporting Initiative (GRI) standards, Sustainability Accounting Standards Board (SASB) standards, United Nations Principles for Responsible Investment (UNPRI), and United Nations Sustainable Development Goals (UNSDG). The scoring methodologies measure different parameters; thus, company names may appear in one framework but not in the other. Controversies exist regarding the agendas and their numerical quantification considered in every ESG framework. No universally accepted framework, model, algorithm, or rule of thumb is available for solving this problem. However, it always helps stakeholders to delineate the cone of uncertainty if human judgment and intuition are amalgamated over controversies based on the context of the problem.

Although the idea of ESG-focused investing is relatively new, several high-profile investment firms have begun to construct ESG indices by tracking companies committed to creating more environmentally friendly and sustainable business models. Some of these include the S &P 500 ESG index, the Dow Jones Sustainability World Index, MSCI World ESG Focus Index, and MSCI Emerging Markets ESG Focus Index. Several mutual funds and ETFs provide investment opportunities for ESG savvy investors. These include the Xtrackers S &P 500 ESG ETF (SNPE), SPDR S &P 500 ESG ETF (EFIV), Invesco MSCI Sustainable Future ETF (ERTH), iShares MSCI Global Sustainable Development Goals ETF (SDG), Fidelity International Sustainability Index Fund (FNIDX), and Vanguard FTSE Social Index Fund (VFTAX).

Stock market prediction is a complex and challenging task because of its nonparametric, nonlinear, and chaotic behavior (Ahangar et al. 2010). In addition, investment decisions are not always made simply by looking at structural data, such as balance sheets, financial report cards, company valuations, and volumes of shares traded in a specific range. Investors go beyond these factors and consider whether a company has

incorporated ESG agendas into its business models. These factors depend mainly on the nature of the company and its associated market structure from local and global perspectives. Consequently, ESG factors exhibit additional complexity in an already complex and volatile market. Thus, there is a pressing demand to develop a proper model that helps measure the performance and volatility of ESG indices to minimize related risks and better inform stakeholders before making responsible financial decisions.

Most classical time series models assume linear data relationships. However, this assumption raises significant concerns regarding the robustness of these classical models when applied to real-world time series data that frequently exhibit nonlinear behavior. Moreover, classical machine learning approaches struggle to capture long-term dependencies within time series data. This is where deep learning models have come to the forefront, as they effectively address these limitations. Deep learning excels at comprehending intricate patterns and connections within financial data, offering benefits such as automated feature extraction, nonlinear handling, temporal dependency capture, adaptability to changing conditions, and efficient management of extensive datasets. These attributes collectively position deep learning models as superior tools for precisely predicting ESG index volatility compared to conventional models.

Many studies have been conducted to build efficient predictive models using machine learning and deep learning techniques (Nabipour et al. 2020; Wang et al. 2020; Sen and Chaudhuri 2018). Some of these studies focus on predicting the price and/or volatility of ESG related indices (Guo et al. 2020; Lee et al. 2022; Raman et al. 2020). Varying degrees of success were observed, based on the accuracy and robustness of the models. The most widely used deep learning architectures are long short-term memory (LSTM), convolutional neural networks (CNN), gated recurrent units (GRU), and their respective hybridization techniques (Lin and Jin 2023)

We noticed several gaps in the literature. For instance, researchers often utilize the stated methods to speak of oneself with pride in terms of a model's accuracy. However, the model framework, underlying assumptions, and implementation differ. Thus, it is difficult to perform an unbiased comparison between published research articles, even if they use the same deep learning architecture to construct their predictive models. Furthermore, the authors could not find a transparent and data-driven approach for fine-tuning the model hyperparameters. In addition, several previous studies have focused on price prediction rather than volatility prediction, which is the focus of this study. This trend can be attributed to ESG savvy investors' concerns with the volatility and risks associated with their investment portfolios, rather than short-term returns. In addition, there is a significant lack of analysis on the robustness of the constructed models.

The current study aimed to fill these gaps by (a) providing an integrated computational framework to implement deep learning model architectures to predict the volatility of the ESG index in an identical environment; (b) gathering multifaceted information that directly and indirectly affects the ESG index, putting them together to construct a well-balanced set of input features; (c) implementing an extensive and data-driven approach for hyperparameter tuning and model selection; and (d) conducting statistical analyses to validate and verify the reliability and robustness of the model.

A complete roadmap for achieving this goal is presented in the schematic diagram in Fig. 1. Well-balanced input features were incorporated into the spheres of fundamental

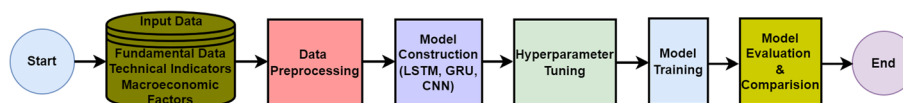


Fig. 1 Schematic diagram of the proposed research framework

data, macroeconomic data, and technical indicators. The collected data were normalized using the min-max technique, and input sequences for the models were created using a specific time step. Hyperparameters such as the number of neurons (or filters), epochs, learning rate, and batch size were tuned using regularization techniques to optimize the model performance. Once the hyperparameters were tuned, the models were trained to predict the volatility of the ESG index. Finally, the model quality was assessed using the RMSE, MAPE, and R-scores of a test set.

The remainder of this paper is organized as follows. Section explains the related work in this field. The data collection and feature selection procedure is explained in Sect. . Modeling approaches are discussed in Sect. . Section discusses the experiment and results, followed by discussion in Sect. . Section discusses the ethics and implications. Finally, Sect. presents the conclusions and future work, followed by acknowledgments and a list of references.

Related work

Although ESG investing is a relatively new thematic investment idea yet to be fully adopted by the mainstream investment community, various studies have been conducted to understand the importance of ESG criteria in portfolio construction and optimization, the integration of ESG factors in machine learning models for price and volatility predictions, and the role of ESG factors during systemic crises.

Some researchers have explored the importance of ESG factors in portfolio construction and optimization. Vo et al. developed a deep responsible investment portfolio to predict quarterly and yearly stock returns, which they then combined with ESG ratings in their modified mean-variance ESG model to construct and rebalance socially responsible investment (SRI) portfolios (Vo et al. 2019).

Xidonas and Essner employed a minimax optimization approach to enhance portfolio optimization, which entailed the integration of key ESG risk performance factors. The minimax methodology facilitates the optimization of individual security weights within the portfolio, aiming to reduce deviations from ESG targets. This is achieved by simultaneously minimizing the maximum risks and maximizing the attainment of ESG investment objectives. They tested the models’ performance on multiple European and American stock indices and demonstrated better risk-adjusted returns than the benchmarks (Xidonas and Essner 2022). Berg et al. conducted an empirical analysis of ESG investments that quantified the returns associated with ESG investment strategies and assessed their financial performance. This study analyzed a diverse set of companies and industries to evaluate the impact of ESG metrics on investment outcomes. These findings highlight the correlation between ESG scores and stock returns and indicate a potential link between sustainable practices and financial success (Berg et al. 2023). Lucia et al. conducted a case study to explore whether ESG practices led to better financial performances in 1038 public enterprises in Europe. Their findings suggest a

relationship between ESG variables and improved financial performance (De Lucia et al. 2020). Hang and Chen proposed two SRI portfolio construction models, namely double-screening socially responsible investments I and II, which utilized a double-screening mechanism and an extreme learning machine model with genetic algorithm optimization to predict stocks and integrate ESG factors to determine the investment proportion of the screened stocks. The study claimed that the proposed models exhibited better performance (Zhang and Chen 2011). Umar et al. investigated the relationship between the cryptocurrency environmental attention index and the volatility and return on assets categorized as either green or dirty (Umar et al. 2022). Their findings suggest that dirty equities and bonds are the main drivers of return spillover, while dirty equities transmit volatility spillover, and that environmental attention has a greater effect on equities than on bonds. These findings provide insights into investment, hedging, and policymaking decisions as well as the potential usefulness of ESG investments in providing diversification. All of the above studies support the idea that ESG has a positive impact on portfolio construction and optimization.

Efforts have been made to integrate ESG factors into machine learning techniques to enhance the accuracy of stock price predictions by identifying the underlying ESG alpha. For instance, Chen et al. utilized ESG scholar data to establish an automatic trading strategy and proposed a practical machine learning approach to quantify a company's ESG premium and capture ESG alpha (Chen and Liu 2020). Their study involved creating an ESG investment universe, conducting feature engineering on the ESG scholar data of companies, and training the proposed models using financial indicators and ESG scholar data. They used an ensemble method to forecast stock prices and provided recommendations for portfolio construction, trading, and rebalancing. According to this study, the proposed ESG alpha strategy generated impressive cumulative returns from the proposed portfolio compared with several benchmarks. Similarly, Magrot et al. designed and implemented a machine learning algorithm capable of identifying patterns between ESG profiles and performance (Margot et al. 2021). Their algorithm generates a set of rules, each of which identifies a region in the high-dimensional space of ESG features in which excess stock returns can be predicted. This study empirically demonstrates the correlation between ESG profiles and financial performance.

ESG investors are typically savvy investors who prioritize the volatility and risk associated with their investment portfolios over short-term returns. A few researchers have focused on incorporating ESG criteria into the development of efficient volatility prediction models. For example, Sabbaghi conducted empirical investigations of asymmetric volatility in ESG investing using Morgan Stanley Capital International (MSCI) indices and found that the impact of news on the volatility of ESG firms is greater for bad news than for good news (Sabbaghi 2020). Additionally, the impact of bad news on the volatility of ESG firms is smaller for small-cap ESG firms than for large- and mid-cap ESG firms. By contrast, Guo et al. implemented a new deep learning framework called ESG2Risk to predict the future volatility of stock prices using ESG news (Guo et al. 2020). The study concluded that ESG news has a significant impact on the future returns and risks of companies and can therefore be considered a relevant factor when making investment decisions. The studies discussed above, including (Yu et al. 2022; Daniali

et al. 2021), provide evidence that machine learning models that incorporate ESG factors outperform other models in predicting volatility.

Market volatility increases during systemic crises, such as recessions, pandemics, and wars. The inclusion of specific factors in the model is required to capture these effects. Umar et al. investigated how social media coverage of the Covid-19 pandemic affected ESG leader indices in different regions, identifying periods of low, medium, and high coherence between the media coverage index and the price movements of the ESG leader indices Umar and Gubareva (2021). The periods of low coherence suggest that ESG investments could potentially provide diversification benefits during a systemic pandemic like Covid-19. Moreover, Akhtaruzzaman et al. found that media coverage contributed to the spread of the contagion in both advanced and emerging equity markets, with the US being the most severely impacted country (Akhtaruzzaman et al. 2022). Albuquerque et al. investigated the mechanism by which corporate social responsibility (CSR) and ESG policies affect firms' systematic risk by assuming CSR as a product differentiation strategy. They claim that strong ESG firms face relatively less price-elastic demand, which results in lower systematic risk due to a product differentiation strategy. They concluded that consumers play a vital role in influencing firm policies and risk profiles (Albuquerque et al. 2019).

In summary, limited research has been conducted on ESG-related stock market portfolios and volatility predictions compared to the volatility predictions of broad stock market indices (Cho and Lee 2022; Koo and Kim 2023; Mittnik et al. 2015; Lu et al. 2022). The reviewed studies made significant contributions to integrating ESG into portfolio construction, optimization, performance analysis, and risk assessment. However, some of these studies focused solely on building a complex model, whereas others implemented machine learning models without serious consideration of feature selection. An efficient model is required that utilizes a balanced combination of input features, while maintaining the simplicity of its architecture. Our study aimed to address these issues by developing an integrated framework for implementing state-of-the-art deep learning models trained with the best possible set of influencing factors. The main goal was to ensure a comprehensive understanding of the behavior of ESG investment portfolios from multiple dimensions and offer valuable insights for future research.

Data description and preparation

This study used the S &P 500 ESG index, a popular ESG focused index in the US. It is a broad-based market-cap-weighted index designed to measure the performance of securities meeting sustainability criteria while maintaining similar industry group weights to the S &P 500 (Winegarden 2019; Gary 2019). S &P Global maintains the index under the Dow Jones Indices (Indices 2016). The launch date of the index was January 28, 2019, and the backward data assumption date was May 3, 2010. Factors such as fundamental data, technical indicators, and macroeconomic variables may contribute directly or indirectly to index value fluctuations (Serfling and Miljkovic 2011; Tien et al. 2021). The core intrinsic fundamental data are extracted directly from the underlying index. Technical indicators are byproducts of fundamental data that utilize standard mathematical equations to produce final numerical values. Macroeconomic variables were selected based on their potential impact on the overall economy and broader markets.

Input features, such as fundamental data and technical indicators, provide crucial internal information about the overall quality of the underlying stocks as well as supply and demand situations in a given market environment. Other factors, namely macroeconomic variables, contribute by providing information about the potential external influence on the given index fluctuations, capturing the status of the overall economy and broader markets. The incorporation of these comprehensive data sources is pivotal for enhancing the predictive ability of the deep learning framework and ensuring a more robust and accurate analysis of the complex dynamics of stock markets. Consequently, insights gained from this holistic approach can significantly contribute to informed decision-making and more effective predictions.

The selected timeframe for the data was from 01–02-2013 to 12–30-2021, which incorporates a major bear market during the COVID-19 pandemic in 2020. Thus, the construction of the model, which includes both bear and bull markets, resembles the overall market scenario.

S &P 500 ESG index is constructed primarily from the popular US broad market index. Based on a thorough investigation of the related literature and also from the exploratory data analysis, we can identify the following evidence.

- *Finding 1:* The information presented in Table 1 and Fig. 2 vividly reveal the fact that the two indices are not identical in terms of their constituents and sector exposures (Indices 2016).
- *Finding 2:* S &P 500 and S &P 500 ESG have almost similar patterns in terms of daily returns and cumulative returns, as demonstrated in Fig. 3.
- *Finding 3:* Fig. 4 shows similar annualized rolling volatility and Sharpe ratio patterns of these two indices in the given time interval. S &P 500 ESG index’s annualized return is slightly higher than the S &P 500, but these higher returns come with higher risks, as illustrated in Fig. 5.

Table 1 Top 10 constituents of S &P 500 and S &P 500 ESG as of May 31, 2022

Rank	S &P 500		S &P 500 ESG	
	Company	Sector	Company	Sector
1	Apple Inc. (AAPL)	Information Technology	Apple Inc. (AAPL)	Information Technology
2	Microsoft Corp (MSFT)	Information Technology	Microsoft Corp (MSFT)	Information Technology
3	Amazon.com Inc. (AMZN)	Consumer Discretionary	Amazon.com Inc. (AMZN)	Consumer Discretionary
4	Alphabet Inc A (GOOGL)	Communication Services	Alphabet Inc A (GOOGL)	Communication Services
5	Tesla Inc (TSLA)	Consumer Discretionary	Alphabet Inc C (GOOG)	Communication Services
6	Alphabet Inc C (GOOG)	Communication Services	Unitedhealth Group Inc ((UNH)	Health Care
7	Berkshire Hathaway B (BRK.B)	Financials	Nvidia Corp (NVDA)	Information Technology
8	Johnson & Johnson (JNJ)	Health Care	Exxon Mobil Corp (XOM)	Energy
9	Unitedhealth Group Inc (UNH)	Health Care	JP Morgan Chase & Co (JPM)	Financials
10	Nvidia Corp (NVDA)	Information Technology	Procter and Gamble (PG)	Consumer Staples

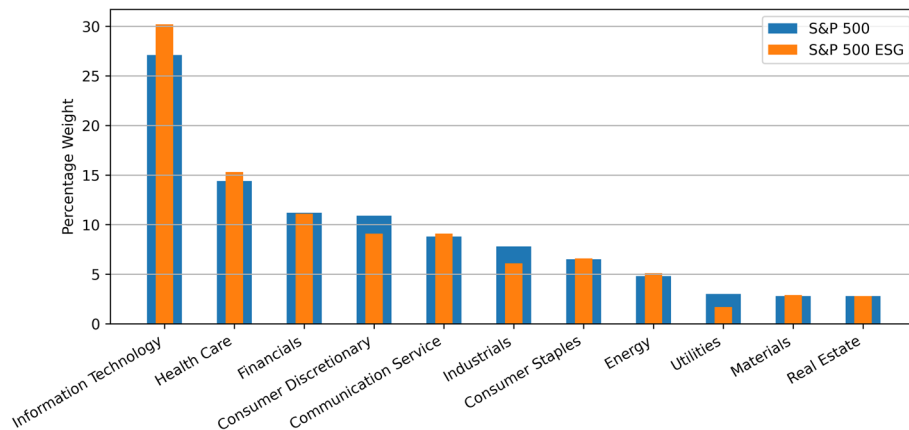


Fig. 2 Sector-wise composition of S &P 500 and S &P 500 ESG indices as of May 31, 2022

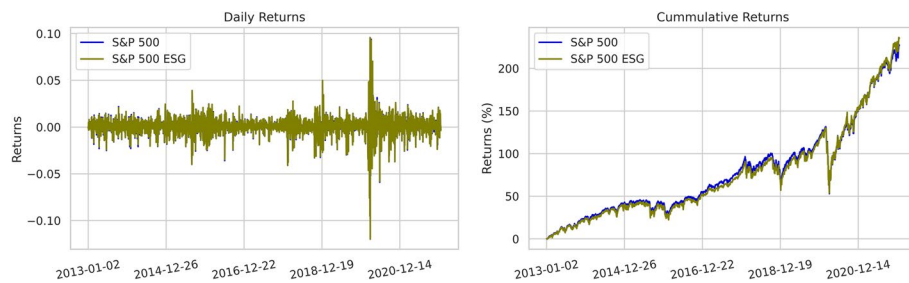


Fig. 3 Comparison of daily and cumulative returns of S &P 500 and S &P 500 ESG indices

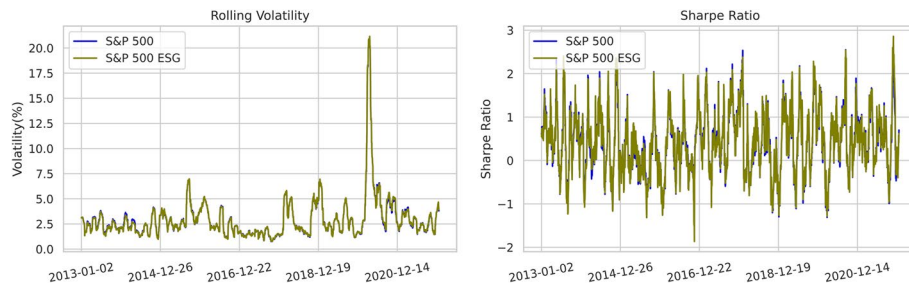


Fig. 4 Rolling volatility and Sharpe ratio of S &P 500 and S &P 500 ESG indices

- Finding 4:* The broad market macroeconomic features such as CBOE Volatility Index, Interest Rate, and US Dollar Index have a similar impact on both indices, as shown in Fig. 6. The data entries of the correlation matrix of broad market macroeconomic features to the closing price and volatility of both S &P 500 and S &P 500 ESG indices show similar correlation.



Fig. 5 Comparison of annualized returns and volatility(Left: Annualized returns and Right: Annualized volatility)

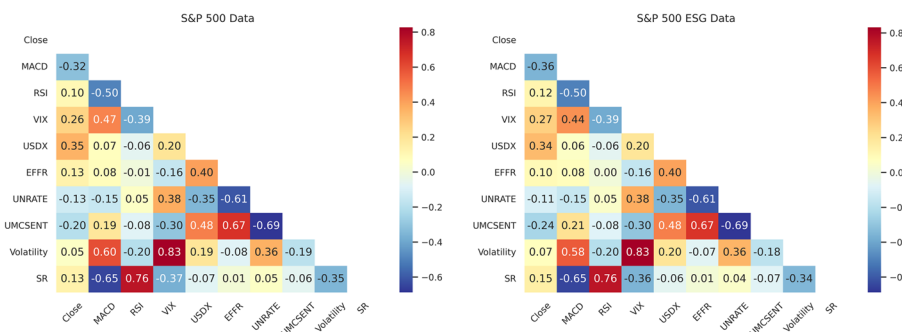


Fig. 6 Correlation heatmaps (Left: S &P 500 data, Right: S &P 500 ESG data)

From the aforementioned evidence, we conclude that the S &P 500 ESG index captures broad US financial market behavior and exhibits similar functionality to the S &P 500 index in terms of returns and volatility, irrespective of variations in their constituents and sector exposures. Therefore, the features, particularly the macroeconomic factors, used to predict the S &P 500 index (Bhandari et al. 2022a) can

Table 2 List of potential features for the model

Data	Source	Frequency	Abbreviation
<i>Fundamental</i>			
Close price	S &P Dow Jones indices	Daily	...
<i>Macroeconomic</i>			
Cboe volatility index	Yahoo	Daily	VIX
Interest rate	FRED	Daily	EFFR
Civilian unemployment rate	FRED	Monthly	UNRATE
Consumer sentiment index	FRED	Monthly	UMCSSENT
US dollar index	Yahoo	Daily	USDX
<i>Technical indicator</i>			
Volatility	...	Rolling	...
Moving average convergence divergence	...	Daily	MACD
Relative strength index	...	Daily	RSI
Sharpe ratio	...	Daily	SR

also be used in the S &P 500 ESG index. The complete input variables used in this study are listed in Table 2, and their short descriptions are presented in the following subsection.

Fundamental data

The first set of variables presented in Table 2 comprises fundamental or historical data that provide basic information regarding the performance of the index. The closing price is the final price of the index on a given trading day.

Macroeconomic data

The second set of variables shown in Table 2 comprises macroeconomic data that significantly influence stock market performance by reporting the overall health of the financial market (Bhandari et al. 2022a; Bhandari et al. 2022). We choose the CBOE volatility index (VIX), interest rate (EFFR), civilian unemployment rate (UNRATE), consumer sentiment index (UMCSENT), and US dollar index (USDIX) as macroeconomic factors (Chandra and Thenmozhi 2015; Ruan 2018; Bernanke and Kuttner 2005; Farsio and Fazel 2013; Bock 2018; Baker and Wurgler 2007). These variables are representative features that explain the overall status of the economy in the proposed model.

Technical indicators

The third set of variables, shown in Table 2, are technical indicators, including volatility, moving average convergence divergence (MACD), relative strength index (RSI), and the Sharpe ratio (SR). Volatility was used as both the input and response variables in this study. First, monthly volatility is calculated as the rolling standard deviation of monthly returns (21 trading days on average, based on the US market). Monthly volatility is then annualized by multiplying it by $\sqrt{12}$:

Active traders use them extensively in the market because they are primarily designed to analyze short-term price movements and are included in this study (Rodríguez-González et al. 2011; Wilder 1978; Anghel 2015; Chong et al. 2014; Chong and Ng 2008; Eric et al. 2009; Murphy 1999; Wang and Kim 2018; Schmidt 2022; Goyal and Aggarwal 2014).

Modelling approach

Deep learning models: LSTM, GRU, and CNN

Let (x_t, y_t) be an input–output pair of the model, where $x_t \in \mathbb{R}^{k \times 1}$ is the input feature, and $y_t \in \mathbb{R}$ is the output at times $t = 1, 2, \dots, n$. Here, k and n are the number of input features and total number of observations, respectively. Furthermore, to incorporate the time step into LSTM, GRU, and CNN architectures, the input sequence X_t was created by taking m continuous sequence $x_t : x_{t+m-1}$, which is a matrix of shape $k \times m$ for $t \in \{1, 2, \dots, n - m + 1\}$.

LSTM is a recurrent neural network consisting of an input, hidden state, cell state, and output. It is designed using a gate mechanism (Hochreiter and Schmidhuber 1997; Gers et al. 2000, 2003). LSTM has four gates: input, update, forget, and output, as shown

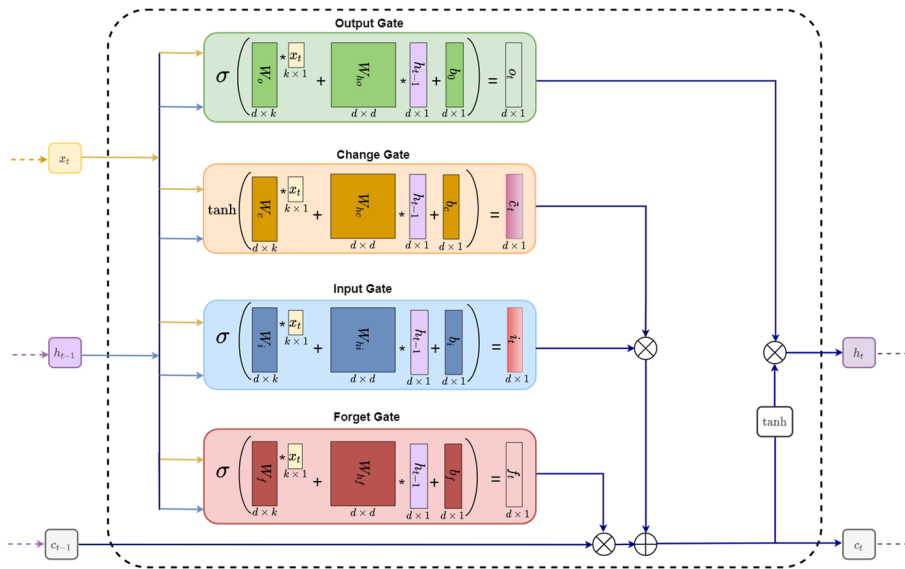


Fig. 7 Long short-term memory(LSTM) architecture (Bhandari et al. 2022a)

in Fig. 7 (Bhandari et al. 2022a). At time t , the gates and layers compute the following functions:

$$\begin{aligned}
 i_t &= \sigma(W_i x_t + W_{hi} h_{t-1} + b_i), \\
 f_t &= \sigma(W_f x_t + W_{hf} h_{t-1} + b_f), \\
 o_t &= \sigma(W_o x_t + W_{ho} h_{t-1} + b_o), \\
 \tilde{c}_t &= \tanh(W_c x_t + W_{hc} h_{t-1} + b_c), \\
 c_t &= f_t \otimes c_{t-1} + i_t \otimes \tilde{c}_t, \\
 h_t &= o_t \otimes \tanh(c_t)
 \end{aligned}$$

where σ and \tanh represent the sigmoid and hyperbolic tangent functions, respectively, the operator \otimes is the element-wise product, $W \in \mathbb{R}^{d \times k}$, $W_h \in \mathbb{R}^{d \times d}$ are the weight matrices, and $b \in \mathbb{R}^{d \times 1}$ is the bias vector. Moreover, d denotes the hidden size (Greff et al. 2017; Qiu et al. 2020; Lei et al. 2019).

The input gate identifies information that must be updated from the change gate. The output of the forget gate is between 0 and 1 through a sigmoid activation function. This identifies the information required to forget former cell state c_{t-1} . It stores all the information in the cell if the output is 1. However, it forgets all the information from the previous cell state if the output is 0. The output gate determines which information is to be taken as the output from the present cell state, and the output (h_t, c_t) of LSTM is a feature representation of the input sequence X_t at time t , which can be expressed as follows:

$$(h_t, c_t) = LSTM(X_t, h_{t-1}, c_{t-1}, w).$$

GRU is a simplified version of LSTM (Chollet 2017). The short-term (h_t) and long-term (c_t) information of LSTM are merged into a single vector h_t in GRU. In contrast to the four gates in LSTM, GRU has three gates: reset gate, change gate, and update gate, as shown in Fig. 8. The update gate of GRU is equivalent to the forget gate and input gate

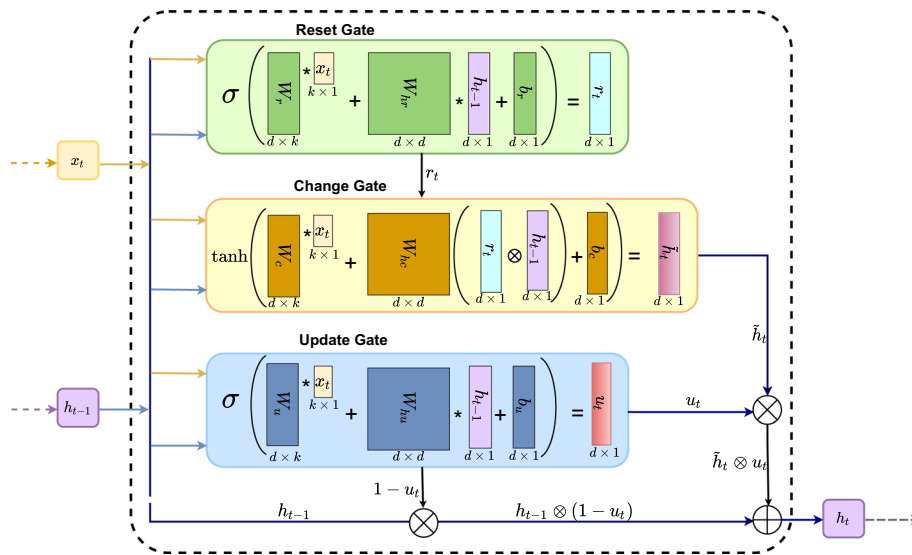


Fig. 8 Gated Recurrent Unit (GRU) architecture (Pokhrel et al. 2022)

of LSTM (Gáeron 2019). Thus, a single gate decides what to forget and update in GRU instead of the two gates in LSTM.

At time t , the gates and layers compute the following functions:

$$\begin{aligned}
 u_t &= \sigma(W_z x_t + W_{hz} h_{t-1} + b_u), \\
 r_t &= \sigma(W_r x_t + W_{hr} h_{t-1} + b_r), \\
 \tilde{h}_t &= \tanh(W_c x_t + W_{hc}(r_t \otimes h_{t-1}) + b_c), \\
 h_t &= (1 - u_t) \otimes h_{t-1} + u_t \otimes \tilde{h}
 \end{aligned}$$

The output h_t of GRU is a feature representation of the input sequence X_t at time t and is calculated as follows:

$$h_t = GRU(X_t, h_{t-1}, w).$$

The CNN architecture has the following components: input, convolutional layer with a nonlinear activation function, a pooling layer, a fully connected layer, and an output. All the layers in a CNN have training parameters, except for the pooling layer. A CNN views a time step as a sequence in which convolutional operations can be performed on a one-dimensional image. Because each series contains observations at the same time step, the input time series is parallel. We can reconfigure these three data arrays (no. of samples, time steps, and no. of features) as a single dataset, where each row is a time step, and each column is a separate time series (Brownlee 2018b, c). We have $n - T_s$ many matrices of size $T_s \times k$ as in LSTM and GRU, and each matrix is treated as an image of size $k \times T_s$ in the CNN. The output h_t of the CNN is a feature representation of the input sequence X_t at time t , which can be expressed as

$$h_t = CNN(X_t, w).$$

For each image, we use m filters and slide each filter on the time axis with a stride of one. Then, after the convolution operation, we obtain m feature maps from m filters.

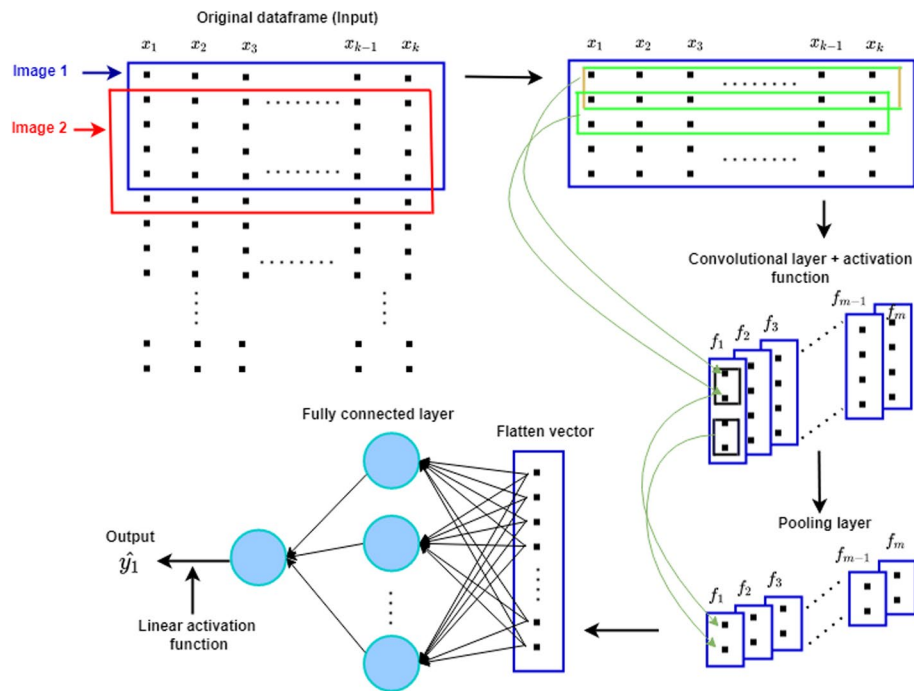


Fig. 9 CNN architecture with m filters for multivariate time series prediction (Pokhrel et al. 2022; Rimal 2022)

After the convolution operations, we use nonlinear activation functions such as ReLU or Leaky ReLU. A pooling operation is performed for downsampling. Subsequently, the feature maps from each filter are vectorized into a single sequence to form a fully connected layer. Finally, the output \hat{y}_1 is predicted using a linear activation function, as shown in Fig. 9.

Experimental design and results

The primary goal of this study was to conduct a comparative analysis of the performance of LSTM, GRU, and CNN models in volatility prediction. Figure 10 shows the original time series of the annualized rolling volatility of the S &P500 ESG index for the

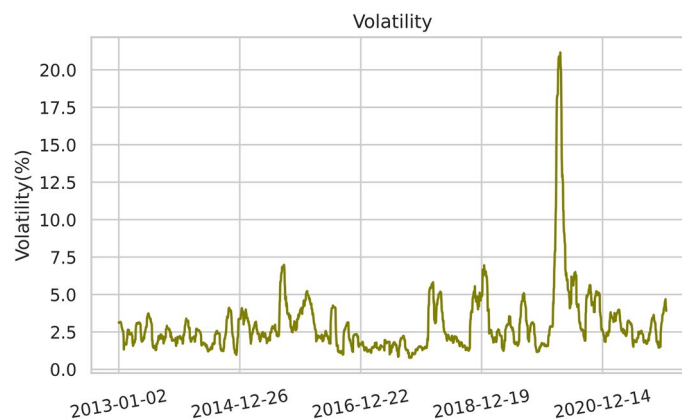


Fig. 10 S &P 500 ESG annualized rolling volatility

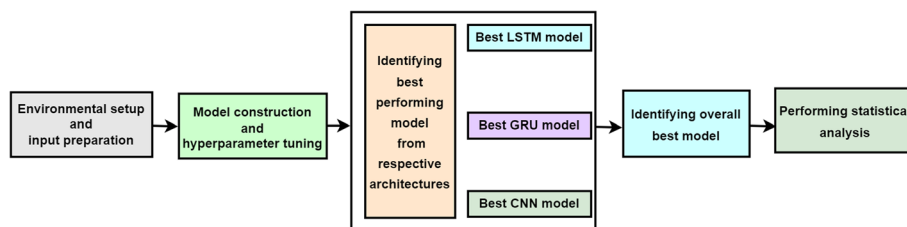


Fig. 11 Experimental design

01–02-2013 textemdash 12–30-2021 interval, which exhibits complex, noisy, and volatile behavior.

To achieve the stated goal, as shown in Fig. 11, the overall experiment was divided into five phases: (a) environmental setup and input preparation, (b) model construction and hyperparameter tuning, (c) identifying the best-performing models from the respective architectures, (d) identifying the overall best-performing model, and (e) performing statistical analysis.

Environmental setup and input preparation

Table 3 summarizes the computational framework of the experiments. The experiments used the Python programming environment and TensorFlow and Keras APIs. The machine configuration and architecture used in the experiments are also listed in the Table 3

As part of the input/output preparation, the original dataset was first divided into training and test sets at a ratio of 4:1. Among the training data, 25% was separated for validation, which accounted for 20% of the total data. A validation set was used for hyperparameter tuning. After obtaining the optimal hyperparameters, the validation data were added to the training set. The overall distribution of the data is presented in Table 4.

Table 3 Computing environmental setup

Machine configuration	Google Colab with NVIDIA-SMI 495.44 GPU
Environment	Python 3.6.0, TensorFlow, and Keras APIs
Architecture	LSTM, GRU, CNN

Table 4 Overall distribution of training, validation, and test data

Data	Dates	No. of samples
Complete data	2013–01–02 to 2021–12–30	2264
Training	2013–01–02 to 2020–03–16	1811
Validation	2018–05–29 to 2020–03–16	453
Test	2020–03–17 to 2021–12–30	453

Because the range of values for the input features varied widely, a min–max normalization technique was implemented. The normalized data were in the form of a 2D array (number of observations and features). However, the proposed model architecture requires 3D input data. Thus, it was converted into a 3D array (number of observations, time steps, and number of features) by incorporating the time step before being fed into the model. The prediction accuracy of the constructed model was assessed using three performance metrics: RMSE, MAPE, and R. The stated matrices help determine the best model in terms of accuracy and reliability.

Model construction and hyperparameter tuning

We constructed deep learning models, each of which consisted of an input layer, an LSTM/GRU/CNN layer, and a dense output layer with linear activation. Early stopping criteria were implemented to address the consequences of underfitting and overfitting that can occur when training neural networks. This approach allowed us to specify a large number of epochs and stop training when the model’s performance stopped improving on the validation data (Brownlee 2018a).

After constructing the model, we performed a hyperparameter tuning process in which each model identified its best set of hyperparameters from multiple avenues. This included three different optimizers (Adam, Adagrad, and Nadam), three different learning rates (0.1, 0.01, and 0.001), and three batch-size options (4, 8, and 16). Therefore, $3 \times 3 \times 3 = 27$ possible choices were available for each model for identifying the best combination. We performed ten independent replicates for each model before calculating the average scores to address the model’s stochastic behavior. The best model was selected based on the lowest possible average RMSE score calculated on the validation dataset. Thus, we executed three architectures—(LSTM, GRU, and CNN) \times six models for each architecture (number of different neurons) \times 27 (possible combinations for each model) = 486 instances—during the complete hyperparameter tuning process. The optimal set of hyperparameters for each model architecture is presented in Table 5.

Identifying the best performing models from respective architectures

Once the hyperparameter tuning process was completed, the models were set with their corresponding hyperparameters. Finally, all models ($6 * 3 = 18$) were trained in full scale with the best hyperparameters. Fully trained models were implemented on the test

Table 5 Optimal hyperparameters for LSTM, GRU, and CNN models

No. of neurons/ filters	LSTM			GRU			CNN		
	Optimizer	Learning rate	Batch size	Optimizer	Learning rate	Batch size	Optimizer	Learning rate	Batch size
10	Adam	0.001	8	Adagrad	0.1	16	Adagrad	0.01	8
30	Adagrad	0.1	4	Adagrad	0.01	8	Adagrad	0.01	8
50	Adagrad	0.01	16	Adagrad	0.01	16	Adagrad	0.01	4
100	Adagrad	0.001	4	Adagrad	0.001	16	Adagrad	0.01	4
150	Adagrad	0.001	4	Adagrad	0.001	16	Adagrad	0.001	4
200	Adagrad	0.001	16	Adagrad	0.001	16	Adagrad	0.001	16

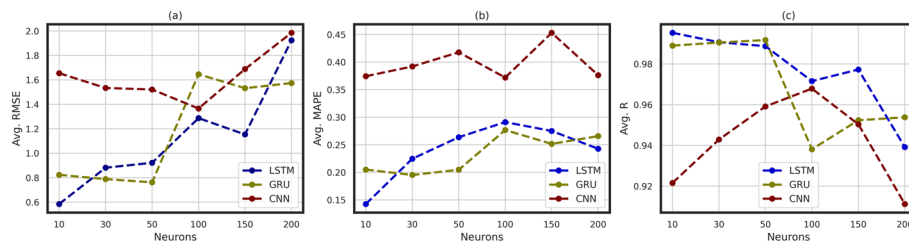
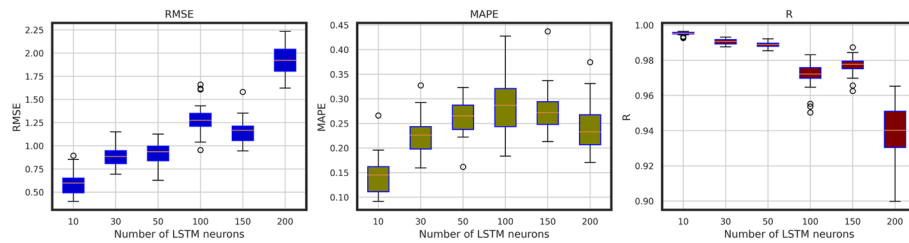
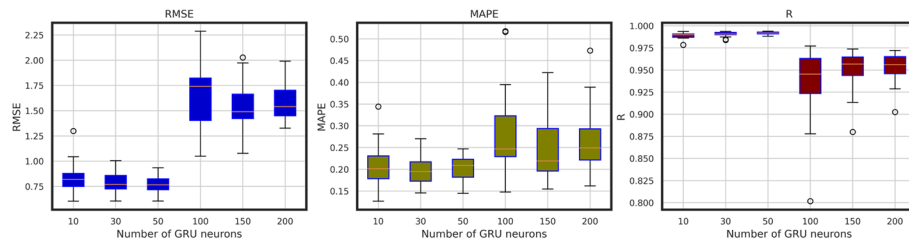


Fig. 12 Average scores obtained from LSTM, GRU, and CNN models: **a** RMSE, **b** MAPE, and **c** R on test dataset

(a) LSTM Models



(b) GRU Models



(c) CNN Models

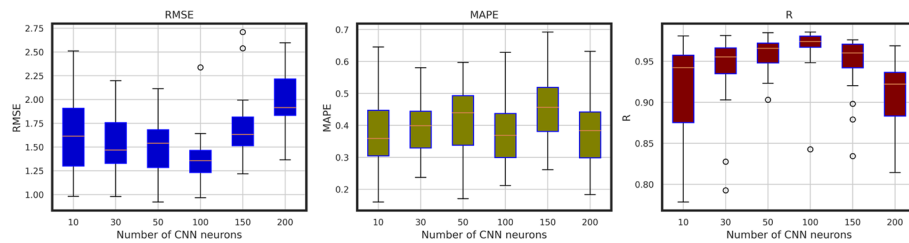


Fig. 13 Boxplots of evaluation metrics for **a** LSTM models, **b** GRU Models, and **c** CNN Models

data to verify their performance and reliability. We replicated each model 30 times to address the stochastic behavior of the deep learning models. Figure 12 shows a graphical representation of the average scores produced by the employed model architectures (LSTM, GRU, and CNN). The subplots (a), (b), and (c) show the overall patterns of the average RMSE, MAPE, and R-scores for each model architecture.

Observing the performance scores in a holistic approach, for LSTM, the average RMSE and MAPE scores were low with 10 neurons. Thereafter, no significant decreasing trend appeared. Similarly, the highest average R-score was observed for 10 neurons. In addition, GRU with 50 neurons provided the smallest average RMSE and MAPE, and the most significant average R score. The CNN model with 100 neurons had the smallest

average RMSE and MAPE and the largest R score. Furthermore, the distributions of the RMSE, MAPE, and R scores and their variabilities obtained from the 30 replicates are presented in Figs. 12 and 13.

Based on the comparative analysis, it can be concluded that the 10 neurons LSTM, 50 neurons GRU, and 100 neurons CNN were the best in their respective categories. The list of best-performing models from the respective architectures, along with their optimal hyperparameters, is highlighted in Table 5 using bold letters.

Identifying overall best model

After identifying the best models from the respective architectures, we compared the performance scores to identify the best model among the three. Table 6 presents the statistics of the performance scores obtained from the three best models. The LSTM with 10 neurons showed the smallest RMSE (0.5849), MAPE (0.1425), and R (0.9952) scores. The GRU with 50 neurons had the second-smallest average RMSE (0.7621) and MAPE (0.2046), and the second-largest R-score (0.9917). Similarly, the standard deviation of the R scores was the smallest and the standard deviations of RMSE and MAPE scores were slightly larger for the best-performing LSTM model compared with those of the best-performing GRU model. In addition, Fig. 13 illustrates that the overall distributions of the scores were approximately symmetric with relatively small variability, indicating the consistent performance of the three best-performing models. Thus, Table 6 and the distribution observed in Fig. 13 suggest that the LSTM model with 10 neurons is the winner, followed by GRU with 50 neurons and CNN with 100 neurons.

Figure 14 shows the true vs. predicted plots that gauge the goodness of fit to determine the quality of the prediction obtained from the training and test data. The blue dots represent the actual versus predicted values, and the olive dotted line shows the best fit of each plot ($y = x$). The overall fit of the training data is almost indistinguishable in all three subplots of Fig. 14a, despite the relatively better performance of LSTM. In the test data, the predicted values deviated to a greater extent from the actual values compared with the training data, as expected. Among the three subplots in Fig. 14b, LSTM shows a superior fit compared with GRU and CNN.

Figure 15 shows the actual time series together with the predicted volatility obtained from the three best models. The blue curves represent the actual values, whereas the

Table 6 Performance scores of the models on the test data

Models	Metrics →	RMSE	MAPE	R
LSTM	Mean ± Std	0.5849 ± 0.1136	0.1425 ± 0.0372	0.9952 ± 0.0008
	Minimum	0.40172	0.0917	0.9926
	Maximum	0.8939	0.2661	0.9964
GRU	Mean ± Std	0.7621 ± 0.0855	0.2046 ± 0.0247	0.9917 ± 0.001354
	Minimum	0.6039	0.1447	0.9881
	Maximum	0.9346	0.246957	0.9938
CNN	Mean ± Std	1.3661048 ± 0.2512	0.37198 ± 0.0935	0.9679 ± 0.0252
	Minimum	0.9666	0.2120	0.8427
	Maximum	2.3382	0.6286	0.9856

The bold values represent the scores associated with the best performing model on the test data

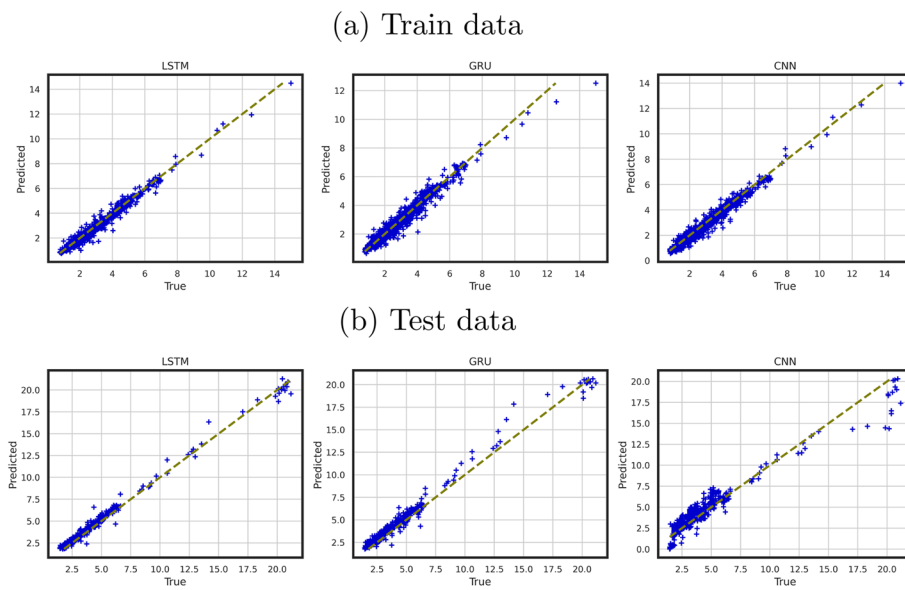


Fig. 14 True versus predicted value plots of the best performing LSTM, GRU, and CNN models

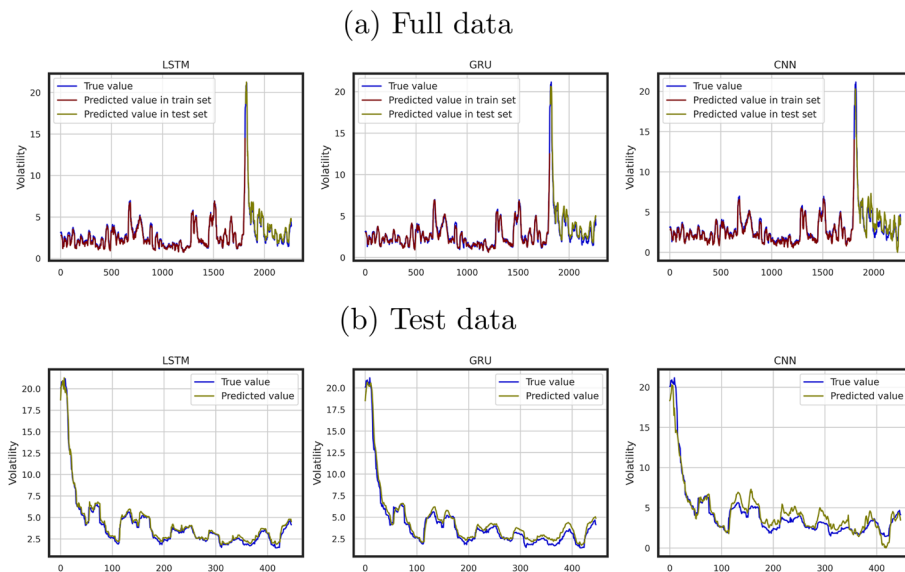


Fig. 15 Time series plots of the true and predicted values obtained from three best performing models

maroon and olive curves represent the values predicted from the training and test data, respectively. As shown in the subplots in Fig. 15a and b, the prediction curve obtained from the LSTM model captures the fluctuations more accurately in almost every situation. However, the GRU and CNN struggle to capture actual values, particularly in the test data. It is clear that the LSTM provided a superior fit compared with the others.

Table 7 Test statistics and p -values from normality test of RMSEs of the models

	LSTM	GRU	CNN
Test statistics	4.20	0.5329	0.7896
P -value	0.1223	0.766	0.673

Table 8 Test statistics and p -values from two samples t -test for pairwise comparison of model performance

	(LSTM, GRU)	(LSTM, CNN)	(GRU, CNN)
Test statistics	- 6.7063	- 18.8414	- 15.3499
P -value	1.2476×10^{-8}	1.2305×10^{-23}	2.3429×10^{-18}

Statistical analysis

To validate the reliability of the model outcome, we conducted a statistical analysis to identify whether the performances of the three best models differed significantly. We performed a pairwise comparison of the mean RMSEs of the three models using Welch's two-sample t -tests. The normality test of RMSEs based on D'Agostino and Pearson (D'agostino and Pearson 1973) ensures that the RMSEs of the three models follow normal distributions, as the p -values are significantly higher than the significance level $\alpha = 0.05$ as presented in Table 7.

The test statistics and p -values from the two-sample t -test are listed in Table 8. A significant difference exists between the mean RMSEs of the pairs (LSTM, GRU), (LSTM, CNN), and (GRU, CNN). The pairwise model comparison produced an outcome in favor of the LSTM model. Hence, we conclude that the LSTM model with 10 neurons best predicts the volatility of the S &P500 ESG index.

Discussion

This study developed an efficient model for predicting the volatility of the broader ESG index of the stock market using deep learning architectures, such as LSTM, GRU, and CNN. This study utilized a diverse set of features from multiple avenues that contribute to ESG index volatility and compared the model performance. The researchers collected data from various sources and prepared the data for modeling. The study rigorously followed standard guidelines for predictive modeling and identified the overall best model with the best fit and highest prediction accuracy. The models were trained using data from both bull and bear market conditions, including the great recession of 2007–2009 and the COVID-19 market downturn, and their performances were evaluated using several measures. Thus, the developed model can make reasonable predictions, even in highly volatile market situations.

The research can be extended to model unusual volatility during a systemic crisis, which may require a close attention to the specific crisis, and identify additional features that can influence investors sentiment during that crisis. Some studies have discussed the importance of studying this scenario and suggested that studying the

performance of equities during a systemic crisis requires special treatment because several unusual factors contribute to volatility (Jabeur et al. 2021; Kou et al. 2019; Chatzis et al. 2018; Lee et al. 2019; Engelhardt et al. 2021). Another potential extension could be to utilize the predictive power of the proposed model for investment portfolio construction, optimization, and analyzing risk-adjusted returns. Recent studies in this demanding research area include the development of automatic clustering and fuzzy system-based approaches to optimize investment portfolios by analyzing large-scale financial data (Li et al. 2021; Kou et al. 2021).

Ethics and implications

The model development process is not driven by profit maximization. All major ethical attributes, such as transparency, integrity, and candor, are internalized to maintain the trustworthiness of the stakeholders. This study used a publicly available dataset without manipulation. Machine learning scripts are completely inspected, interpretability of the final outcome concerning domain knowledge is not sacrificed. The reported performance of the model is the average performance of the out-of-sample data based on several replications. Thus, the results can be used as additional information to make an investment decision that upholds investors' confidence. However, investment decisions should not rely entirely on the research outcomes. Investors are expected to perform due diligence and consider their risk tolerance under various market conditions. A reasonable forecast depends not only on the outcome of the specific model but also on the volatile nature of the stock market, especially during geopolitical tension, global supply chain disturbances, war, pandemics, and various other market risks. Thus, stakeholders can benefit if the market's current behavior is appropriately analyzed and amalgamated with the model's outcome.

Equity traders, individual investors, and portfolio managers intrinsically want to predict volatility using projected risks. This study demonstrates the potential of a neural network architecture to delineate the cone of uncertainty in market volatility prediction. Moreover, academic researchers can build the proposed model framework to expand horizons in the field of sequential data modeling.

Conclusion

Predicting the volatility of the stock market is of great interest to finance practitioners to best allocate their assets and academics to build an optimal model for consistent predictions with a high level of accuracy. Predicting a volatile market is challenging because of its noisy and nonlinear behavior. Multifaceted factors, both local and global, may directly or indirectly affect predictions. This study built predictive models using 10 predictors that fall under fundamental, macroeconomic, and technical data.

A comparative analysis of S & P500 ESG index volatility prediction was performed using deep learning architectures, namely LSTM, GRU, and CNN. An extensive data-driven approach was implemented to optimize the model hyperparameters. The performance of the model was evaluated using RMSE, MAPE, and R. The experimental results showed that the LSTM model with 10 neurons provided a superior fit and high prediction accuracy, followed by GRU with 50 neurons and CNN with 100 neurons.

The outcome was further validated by a statistical analysis of the performance metrics. The proposed model can be tailored to other broad-market ESG indices for which the data show similar characteristics.

In the near future, we plan to develop hybrid predictive models by combining the implemented models with other neural network architectures such as transformers. Another potential direction is the amalgamation of classical and deep learning model architectures to build a new predictive model. We also plan to implement a hybrid optimization algorithm that trains model parameters by combining local and global optimizers. Finally, the implementation of evolutionary algorithms to achieve state-of-the-art performance is a topic for future research.

Abbreviations

ESG	Environment, Social, and Governmental
GRI	Global Reporting Initiative
SASB	Sustainability Accounting Standards Board
UNPRI	United Nations Principles for Responsible Investment
UNSDG	United Nations Sustainable Development Goals
ETF	Exchange Traded Fund
ML	Machine Learning
LSTM	Long-Short Term Memory
GRU	Gated Recurrent Unit
CNN	Convolution Neural Networks
RMSE	Root Mean Square Error
MAPE	Mean Absolute Percentage Error
R	Correlation Coefficient
MVP-ESG	Modified Mean-Variance ESG
SRI	Socially Responsible Investment
CSR	Corporate Social Responsibility
ROE	Return of Equity
ROA	Return of Assets
VADER	Valence Aware Dictionary and Sentiment Reasoner
DSSRI-I	Double-Screening Socially Responsible Investment-I and
DSSRI-II	Double-Screening Socially Responsible Investment-II
ELM	Extreme Learning Machine
MSPR	Maximum Sharpe Ratio
GA	Genetic Algorithm
MSCI	Morgan Stanley Capital International
PLS	Partial Least Square
VIX	CBOE Volatility Index
EFFR	Effective Federal Funds Rate
UNRATE	Civilian Unemployment Rate
UMCSENT	Consumer Sentiment Index
USDIX	US Dollar Index
MACD	Moving Average Convergence Divergence
RSI	Relative Strength Index
SR	Sharpe Ratio
SNPE	Ticker of Xtrackers S &P 500 ESG ETF
EFIV	Ticker of SPDR S &P 500 ESG ETF
ERTH	Ticker of Invesco MSCI Sustainable Future ETF
SDG	Ticker of iShares MSCI Global Sustainable Development Goals ETF
FNIDX	Ticker of Fidelity International Sustainability Index Fund
VFTAX	Ticker of Vanguard FTSE Social Index Fund
GPU	Graphic Processing Unit

Acknowledgements

The authors would like to thank the Association of Nepalese Mathematicians in America for creating a collaborative research opportunity that resulted in this work. We are also thankful to Google LLC for providing GPU supported open-source cloud computing platform, Google Colab.

Author Contributions

All authors contributed to the study conception and design. Data collection and software development were performed by HNB. Methodology and analysis were done by all authors together. The outline was created by HNB and all authors contributed to the sections of the manuscript. All authors read and approved the final manuscript.

Funding

No funding was received to assist with the preparation of this manuscript.

Availability of data and materials

The data used in this work are open source, the readers can access the data following the manuscript. Codes will be made available to public once the manuscript is accepted for publication.

Declarations**Competing interests**

The authors have no competing interests to declare that are relevant to the content of this article.

Received: 14 October 2022 Accepted: 29 December 2023

Published online: 08 July 2024

References

- Ahangar RG, Yahyazadehfar M, Pournaghshband H (2010) The comparison of methods artificial neural network with linear regression using specific variables for prediction stock price in Tehran stock exchange. [arXiv:1003.1457](https://arxiv.org/abs/1003.1457)
- Akhtaruzzaman M, Boubaker S, Umar Z (2022) COVID-19 media coverage and ESG leader indices. *Finance Res Lett* 45:102170
- Albuquerque R, Koskinen Y, Zhang C (2019) Corporate social responsibility and firm risk: theory and empirical evidence. *Manage Sci* 65(10):4451–4469
- Anghel GDI (2015) Stock market efficiency and the MACD. Evidence from countries around the world. *Procedia Econ Finance* 32:1414–1431
- Baker M, Wurgler J (2007) Investor sentiment in the stock market. *J Econ Perspect* 21(2):129–152
- Berg F, Lo AW, Rigobon R, Singh M, Zhang R (2023) Quantifying the returns of ESG investing: an empirical analysis with six ESG metrics. Available at SSRN
- Bernanke BS, Kuttner KN (2005) What explains the stock market's reaction to Federal Reserve policy? *J Finance* 60(3):1221–1257
- Bhandari HN, Rimal B, Pokhrel NR, Rimal R, Dahal KR (2022) LSTM-SDM: an integrated framework of LSTM implementation for sequential data modeling. *Softw Impacts* 14:100396. <https://doi.org/10.1016/j.simpa.2022.100396>
- Bhandari HN, Rimal B, Pokhrel NR, Rimal R, Dahal KR, Khatri RK (2022) Predicting stock market index using LSTM. *Mach Learn Appl* 100320
- Bock J (2018) Quantifying macroeconomic expectations in stock markets using google trends. [arXiv:1805.00268](https://arxiv.org/abs/1805.00268)
- Brownlee J (2018) Better deep learning: train faster, reduce overfitting, and make better predictions. Ebook: Machine Learning Mastery
- Brownlee J (2018) Deep learning for time series forecasting: predict the future with MLPs, CNNs and LSTMs in python. Ebook: Machine Learning Mastery
- Brownlee J (2018c) How to develop convolutional neural network models for time series forecasting. Machine Learning Mastery
- Chandra A, Thenmozhi M (2015) On asymmetric relationship of India volatility index (India VIX) with stock market return and risk management. *Decision* 42(1):33–55
- Chatzis SP, Siakoulis V, Petropoulos A, Stavroulakis E, Vlachogiannakis N (2018) Forecasting stock market crisis events using deep and statistical machine learning techniques. *Expert Syst. Appl.* 112:353–371. <https://doi.org/10.1016/j.eswa.2018.06.032>
- Chen Q, Liu XY (2020). Quantifying ESG alpha using scholar big data: an automated machine learning approach. In: Proceedings of the first ACM international conference on AI in finance, pp 1–8
- Cho P, Lee M (2022) Forecasting the volatility of the stock index with deep learning using asymmetric Hurst exponents. *Fractal Fract* 6(7):394
- Chollet F (2017) Deep learning with python. Simon and Schuster, New York
- Chong TT-L, Ng W-K (2008) Technical analysis and the London stock exchange: testing the MACD and RSI rules using the FT30. *Appl Econ Lett* 15(14):1111–1114
- Chong TT-L, Ng W-K, Liew VK-S (2014) Revisiting the performance of MACD and RSI oscillators. *J Risk Financ Manag* 7(1):1–12
- Clementino E, Perkins R (2021) How do companies respond to environmental, social and governance (ESG) ratings? Evidence from Italy. *J Bus Ethics* 171(2):379–397
- D'agostino R, Pearson ES (1973). Tests for departure from normality. Empirical results for the distributions of b^2 and \sqrt{b} . *Biometrika* 60(3):613–622
- Daniali SM, Barykin SE, Kapustina IV, Mohammadbeigi Khortabi F, Sergeev SM, Kalinina OV, Senjyu T (2021) Predicting volatility index according to technical index and economic indicators on the basis of deep learning algorithm. *Sustainability* 13(24):14011
- De Lucia C, Paziienza P, Bartlett M (2020) Does good ESG lead to better financial performances by firms? machine learning and logistic regression models of public enterprises in Europe. *Sustainability* 12(13):5317
- Engelhardt N, Ekkenga J, Posch P (2021) ESG ratings and stock performance during the COVID-19 crisis. *Sustainability* 13(13):7133
- Eric D, Andjelic G, Redzepagic S (2009) Application of MACD and RVI indicators as functions of investment strategy optimization on the financial market. *Zbornik radova Ekonomskog fakulteta u Rijeci:casopis za ekonomsku teoriju i praksu* 27(1):171–196

- Farsio F, Fazel S (2013) The stock market/unemployment relationship in USA, China and Japan. *Int J Econ Financ* 5(3):24–29
- Gary SN (2019) Best interests in the long term: fiduciary duties and ESG integration. *U. Colo. L. Rev.* 90:731
- Gáeron A (2019) Hands-on machine learning with Scikit-Learn, Keras, and Tensorflow: concepts, tools, and techniques to build intelligent systems. O'Reilly Media, California
- Gers FA, Schmidhuber J, Cummins F (2000) Learning to forget: continual prediction with LSTM. *Neural Comput* 12(10):2451–2471
- Gers FA, Schraudolph NN, Schmidhuber J (2003) Learning precise timing with LSTM recurrent networks. *J Mach Learn Res* 3:115–143. <https://doi.org/10.1162/153244303768966139>
- Goyal M, Aggarwal K (2014) ESG index is good for socially responsible investor in India. *Asian J Multidiscip Stud* 2(11):92–96
- Greff K, Srivastava RK, Koutník J, Steunebrink BR, Schmidhuber J (2017) LSTM: a search space odyssey. *IEEE Trans Neural Netw Learn Syst* 28(10):2222–2232. <https://doi.org/10.1109/TNNLS.2016.2582924>
- Guo T, Jamet N, Betrix V, Piquet L-A, Hauptmann E (2020) Esg2risk: A deep learning framework from ESG news to stock volatility prediction. [arXiv:2005.02527](https://arxiv.org/abs/2005.02527)
- Hochreiter S, Schmidhuber J (1997) Long short-term memory. *Neural Comput* 9(8):1735–1780
- Indices SDJ (2016) S & P Dow Jones indices. Retrieved 12 April 2016
- Jabeur SB, Khalfaoui R, Arfi WB (2021) The effect of green energy, global environmental indexes, and stock markets in predicting oil price crashes: evidence from explainable machine learning. *J Environ Manag* 298:113511
- Koo E, Kim G (2023) A new neural network approach for predicting the volatility of stock market. *Comput Econ* 61(4):1665–1679
- Kou G, Chao X, Peng Y, Alsaadi FE, Herrera-Viedma E (2019) Machine learning methods for systemic risk analysis in financial sectors. *Technol Econ Dev Econ* 25(5):716–742
- Kou G, Olgu Akdeniz Ö, Dinççer H, Yüksel S (2021) Fintech investments in European banks: a hybrid IT2 fuzzy multidimensional decision-making approach. *Financ Innov* 7(1):39
- Lee O, Joo H, Choi H, Cheon M (2022) Proposing an integrated approach to analyzing ESG data via machine learning and deep learning algorithms. *Sustainability* 14(14):8745
- Lee TK, Cho JH, Kwon DS, Sohn SY (2019) Global stock market investment strategies based on financial network indicators using machine learning techniques. *Expert Syst Appl* 117:228–242
- Lei J, Liu C, Jiang D (2019) Fault diagnosis of wind turbine based on long short-term memory networks. *Renew Energy* 133:422–432. <https://doi.org/10.1016/j.renene.2018.10.031>
- Li T, Kou G, Peng Y, Philip SY (2021) An integrated cluster detection, optimization, and interpretation approach for financial data. *IEEE Trans Cybern* 52(12):13848–13861
- Lin S-L, Jin X (2023) Does ESG predict systemic banking crises? A computational economics model of early warning systems with interpretable multi-variable LSTM based on mixture attention. *Mathematics* 11(2):410
- Lu X, Ma F, Wang J, Dong D (2022) Singlehanded or joint race? Stock market volatility prediction. *Int Rev Econ Finance* 80:734–754
- Margot V, Geissler C, de Franco C, Monnier B, Advestis F, Ossiam F (2021) ESG investments: filtering versus machine learning approaches. *Appl Econ Finance* 8(2):1–16
- Mittnik S, Robinzonov N, Spindler M (2015) Stock market volatility: identifying major drivers and the nature of their impact. *J Bank Finance* 58:1–14
- Murphy JJ (1999) Technical analysis of the financial markets: a comprehensive guide to trading methods and applications. Penguin, New York
- Nabipour M, Nayyeri P, Jabani H, Shahab S, Mosavi A (2020) Predicting stock market trends using machine learning and deep learning algorithms via continuous and binary data; a comparative analysis. *IEEE Access* 8:150199–150212
- Pokhrel NR, Dahal KR, Rimal R, Bhandari HN, Khatri RK, Rimal B, Hahn WE (2022) Predicting nepse index price using deep learning models. *Mach Learn Appl* 100385
- Qiu J, Wang B, Zhou C (2020) Forecasting stock prices with long-short term memory neural network based on attention mechanism. *PLOS ONE* 15(1):1–15. <https://doi.org/10.1371/journal.pone.0227222>
- Raman N, Bang G, Nourbakhsh A (2020) Mapping ESG trends by distant supervision of neural language models. *Mach Learn Knowl Extract* 2(4):453–468
- Rimal B (2022) Financial time-series analysis with deep neural networks (Unpublished doctoral dissertation). Florida Atlantic University
- Rodríguez-González A, García-Crespo Á, Colomo-Palacios R, Iglesias FG, Gómez-Berbis JM (2011) Cast: Using neural networks to improve trading systems based on technical analysis by means of the RSI financial indicator. *Expert Syst Appl* 38(9):11489–11500
- Ruan L (2018) Research on sustainable development of the stock market based on VIX index. *Sustainability* 10(11):4113
- Sabbaghi O (2020) The impact of news on the volatility of ESG firms. *Glob Finance J* 100570
- Schmidt AB (2022) Optimal ESG portfolios: an example for the Dow Jones index. *J Sustain Finance Invest* 12(2):529–535
- Sen J, Chaudhuri TD (2018) Stock price prediction using machine learning and deep learning frameworks. In: Proceedings of the 6th international conference on business analytics and intelligence, Bangalore, India, pp 20–22
- Serfling MA, Miljkovic D (2011) Time series analysis of the relationships among (macro) economic variables, the dividend yield and the price level of the S & P 500 index. *Appl Financ Econ* 21(15):1117–1134
- Tien NH, Jose RJS, Ullah SE, Thang HV (2021) The impact of world market on Ho Chi Minh city stock exchange in context of COVID-19 pandemic. *Turk J Comput Math Educ (TURCOMAT)* 12(14):4252–4264
- Tucker JJ III, Jones S (2020) Environmental, social, and governance investing: investor demand, the great wealth transfer, and strategies for ESG investing. *J Financ Serv Prof* 74(3):56

- Umar Z, Abrar A, Zaremba A, Teplova T, Vo XV (2022) Network connectedness of environmental attention green and dirty assets. *Finance Res Lett* 50:103209
- Umar Z, Gubareva M (2021) The relationship between the COVID-19 media coverage and the environmental, social and governance leaders equity volatility: a time–frequency wavelet analysis. *Appl Econ* 53(27):3193–3206
- Vo NN, He X, Liu S, Xu G (2019) Deep learning for decision making and the optimization of socially responsible investments and portfolio. *Decis Support Syst* 124:113097
- Wang J, Kim J (2018) Predicting stock price trend using MACD optimized by historical volatility. *Math Probl Eng* 2018:1–12
- Wang Z, Hong T, Piette MA (2020) Building thermal load prediction through shallow machine learning and deep learning. *Appl Energy* 263:114683
- Wilder JW (1978) *New concepts in technical trading systems*. Trend Research, New York
- Winegarden W (2019) *Environmental, social, and governance (ESG) investing: an evaluation of the evidence*. Pacific Research Institute
- Xidonas P, Essner E (2022) On ESG portfolio construction: a multi-objective optimization approach. *Comput Econ* 1:25
- Yu, G., Liu, Y., Cheng, W., Lee, C.-T. (2022). Data analysis of ESG stocks in the Chinese stock market based on machine learning. In: 2022 2nd international conference on consumer electronics and computer engineering (ICCECE), pp 486–493
- Zhang J, Chen X (2021) Socially responsible investment portfolio construction with a double-screening mechanism considering machine learning prediction. *Discrete Dyn Nat Soc* 2021

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

Submit your next manuscript at ▶ [springeropen.com](https://www.springeropen.com)
