International Journal of Development Mathematics Vol 1 Issue Page No. 023 - 037

INTERNATIONAL JOURNAL OF DEVELOPMENT MATHEMATICS

ISSN: 3026-8656 (Print) | 3026-8699 (Online)

journal homepage: https://ijdm.org.ng/index.php/Journals

Integration of Modified Heston Model (Jump and Sentiment Factor) with Artificial Neural Networks (ANN) for Volatility Forecasting in Financial Markets

Itankan A. Wilfred[a*], Odekunle M. Remilekun[b], Alkali A. Muhammad[b], Alam Khursheed.[c] and Gupta Sangeeta [c]

[a]*Mathematics Unit, Dept. of General Studies, Federal polytechnic, Bali, Taraba State-Nigeria*

[b]*Department of Mathematics, Modibbo Adama University, Yola, Adamawa State-Nigeria.*

[c]*The A. H. Siddiqi Centre for Advanced Research in Applied Mathematics and Physics (CARAMP) Mathematics Dept. Sharda University, Knowledge Part III, Greater Noida, India.*

ARTICLE INFO

Article history:
Received 17 May, 2024
Received in revised form 05 July, 2024
Accepted 19 August, 2024

Keywords:
Artificial Neural Networks, Financial Markets, Modified Heston Model, Risk Management, Volatility Forecasting.

MSC 2020 Subject classification:
60H35, 60J65, 33F99, 49-04

ABSTRACT

This study examined a new methodology for predicting volatility in financial markets by combining the modified Heston model (JSF) with Artificial Neural Networks (ANN). Forecasting volatility is essential in various financial contexts, such as risk management, option pricing and portfolio optimization. Although the traditional Heston model is effective but it does not account for additional elements like jumps and investor sentiment factors, which impact asset price dynamics. To address this limitation, the extended Heston model with jump and sentiment variables helped in enhancing its ability to capture complex market trends. We utilized ANN's predictive capabilities on the modified Heston model to improve volatility predictions using historical random market data from NASDAQ Composite. The ANN was trained on simulated data from the modified Heston model and assessed on its effectiveness in forecasting volatility. This approach helped in developing a forecasting hybrid semi-parametric model known as the modified-Heston-ANN model. Our findings revealed that this integrated approach outperformed the individual model, offering more precise and robust volatility predictions.

1. Introduction

Forecasting volatility plays a pivotal role within financial markets, applying influence over a multitude of decision-making procedures including risk management, derivative pricing, and portfolio optimization. The Heston model, a conventional approach, has garnered extensive utilization in volatility forecasting owing to its proficiency in capturing the random characteristics of asset prices. Nevertheless, these models frequently fail to consider crucial components like jumps and investors sentiments, resulting in inaccuracies in forecasting volatility. Within this study, we suggest a unified methodology that merges the adjusted Heston model with Artificial Neural Networks (ANN) in order to enhance the precision of volatility predictions. For example, the allocation of funds into a particular economic industry could simultaneously offer both benefits and drawbacks. One method of investment available to investors involves the purchase and sale of shares of a publicly traded company on a securities exchange. Shares represent a form of security indicating partial ownership in a firm. The primary objective shared by all investors is the generation of profits. Nevertheless, engaging in transactions of shares leads to fluctuations in stock prices. These fluctuations introduce complexities in accurately anticipating the movement of stock prices, rendering the tasks of forecast and identification of optimal moments for transactions challenging (I-Ming, *et al*. 2014). Numerous research studies have been extensively carried out to formulate an effective strategy for predicting stock prices (Shunrong, *et al*. 2012). The fluctuations in stock prices are determined by a combination of economic, social, political and psychological variables that intertwine in a sophisticated manner. A considerable number of investors and financial experts display interest in determining the future trajectory of stock price movements. This foresight equips them with the essential guidance needed to devise profitable trading strategies in the market. Market volatility serves as a

* Corresponding author. Tel.: +2348033259471
E-mail address: wita4me@gmail.com (Itankan, W. A.)
https://doi.org/10.62054/ijdm/0103.03

metric to gauge the extent of fluctuations in a market's overall value over a specified period. Greater risks are typically associated with higher returns. Consequently, investors generally prefer markets characterized by price fluctuations. Predicting stock prices involves anticipating or estimating future price levels. Despite the stochastic nature of future stock prices, Chandan (2012) argued that prices adhere to particular criteria that are established based on historical data and understanding of share prices. The primary objective of stock price prediction is to achieve optimal outcomes while minimizing inaccuracies in price forecasts (Preethi and Santhi 2012). As a result, investors require a dependable approach to forecast volatility and, consequently, stock prices to facilitate well-informed investment choices. The objective of this article is to combine the Heston modified model, a widely recognized option pricing model, with Artificial Neural Networks (ANN) to establish a hybrid forecasting model known as the Heston (JSF)-ANN model. Through the integration of the stochastic volatility framework from the Heston model with the computational capabilities of ANN, the hybrid Heston-ANN model is positioned as a reliable method for predicting stock prices.

2. LITERATURE REVIEW

The end of the 20th century was witness by the rise of a mathematical framework used to predict variations in stock prices, such as the Geometric Brownian Motion (GBM) and the Black-Scholes model. GBM utilizes a method that commences with calculating the return value in order to infer the level of confidence (Ladde and Ling, 2009). The Black-Scholes model holds significance in the area of stock pricing as well. It has proficiently evaluated the worth of derivatives while incorporating alternative investment tools, taking into account the impact of time and other essential risk factors. These pricing model play a crucial role for stakeholders in the financial sector as they ascertain and predict the values of various options and when engaging in stock transactions. According to Preethi and Santhi (2012), securities exhibit fluctuating prices over the course of a trading session. Consequently, the stock pricing models aid in scrutinizing and integrating the factors prompting variations to pinpoint the most suitable stock for trading. Nonetheless, the existing stock pricing models have introduced several assumptions and constraints that require reassessment to enhance the precision of stock price predictions. The Black-Scholes model, for example, results in inaccuracies and erroneous pricing of securities due to its reliance on constant volatility (Teneng, 2011). Volatility under actual market conditions exhibits non-constancy but rather stochastic in nature. As a result, these notable constraints have prompted an ongoing re-evaluation of foundational models and the advancement of improved models, as predicted in this study. These approaches guarantee an improved ability to forecast stock prices, thereby enabling advantageous profit margins in the stock market. A model was proposed by Steven Heston in 1993. This model, primarily utilized for valuing European Options over time, is a stochastic volatility model. The emphasis lies on stochastic volatility as opposed to the Black-Scholes model, which relies on the assumption of constant volatility. The ability of the Black-Scholes model can be attributed to its correlation with the distribution of spot returns concerning the cross-sectional attributes of option prices (Lauterbach *et al*. 1990). Heston (1993) expanded the Black-Scholes model while preserving that particular characteristic. The Heston model stands out due to its superior reliability in forecasting stock prices compared to other models, attributed to the arbitrary nature of volatility, as opposed to a fixed one. Implementation of stochastic volatility involves the generation of random variables using a Monte Carlo method, with the process being iterated multiple times through Monte Carlo simulation to produce a probability distribution of possible results. A key aspect of this model is its capability to introduce correlation between volatility and stock price, a crucial factor for stock price projections. Moreover, the model offers a closed-form solution and the results are based on complex mathematical computations, as highlighted by Gauthier and Possama (2010). Noteworthy is its focus on mean reversion and its reliance on a log-normal probability distribution. Teixeira (2017) noted that the Heston model is linked to the volatility smile model, which presents a visual representation featuring identical expiration dates. The model's importance lies in illustrating the increase in volatility as an option transitions from in-the-money to out-of-the-money. The reliability of the Heston model is better than that of other models due to these characteristics. Nevertheless, the model does possess certain limitations. It entails parameters requiring particular calibrations to establish a dependable estimation. Furthermore, a crucial drawback of the Heston model arises when predicting option prices for short-term options. This inability is attributed to the model's failure in capturing high implied volatility, representing the anticipated market volatility from the option's purchase to its expiration. Another noticeable constraint is the Heston model's complexity compared to the Black-Scholes model, discouraging investors from its utilization. Scholars are consistently addressing these deficiencies to enhance the efficiency of the model. Charlotte *et al*. (2022) focused on enhancing the Heston model by incorporating long-term jump diffusion to tackle the issue of predicting short-term maturities. The study conducted a comparison of the results obtained from the enhanced version (Heston model with a jump) with the original Heston model and the Black-Scholes model in relation to short maturities. By utilizing Mean Squared Error (MSE) as the metric for performance evaluation, it was observed that the Heston model with a jump surpassed both the original Heston model and the Black-Scholes model for short-term maturities. The enhancement demonstrated significant progress, revealing a 58.08% reduction in error when compared to the original Heston model and a 47.3% decrease in error in comparison to the Black-Scholes model. Despite the primary development of the Heston model for the financial market, recent studies have unveiled its potential applications in other sectors. Reichert, (2022) conducted a thorough literature review to evaluate the model's significance in the energy industry for energy generation prediction. The investigation covered 25 articles and demonstrated that the suitability of the Heston model for predicting energy generation is justified by its ability to provide a closed-form solution and accurately represent stochastic volatility, thus facilitating the prediction of mean energy generation costs. While the financial market remains unsettled with unexpected and unprecedented fluctuations, recent technological progress has ushered in a new era of forecasting. These advancements play a pivotal role in aiding investors to make informed decisions promptly. As suggested by Luca Di and Oleksandr (2016), investors aim to accurately forecast stock prices to sell before a decline or buy before an increase. The incorporation of machine learning into stock valuation signifies a notable and advantageous technological progress within the realm of financial market. The prediction of future stock prices and values with increased accuracy is facilitated by the utilization of machine learning techniques, addressing a key challenge in maximizing profits within the stock market (Ayala, *et al*. (2021). The emergence of Artificial Neural Networks (ANN) through machine learning has enabled the integration of technical analysis, a crucial element in stock market forecasting. ANN, unlike traditional machine-learning algorithms, is skilful at understanding complex structures. Like the human brain, ANN functions through a complex network of interconnections, established via algebraic equations that guide information flow within a specific model. Its application allows for the implementation of reliable modeling approaches for datasets, employing numerical equations to process variables and generate one or multiple outputs. The effectiveness of a model relies heavily on its predictive accuracy. The foundational back-propagation algorithm of ANN plays a pivotal role in ensuring the network excels in this aspect, solidifying its status as a highly valuable learning tool. Furthermore, the ability of Artificial Neural Networks (ANN) to accurately represent data with high volatility and varying variance further strengthens its effectiveness in delivering accurate and reliable forecasts. As a result, the application of ANN has demonstrated efficiency in predicting financial time series, especially in situations where the data shows considerable volatility. ANN is often preferred over traditional statistical methods due to their ability to approximate non-linear functions accurately, without imposing prior assumptions regarding their characteristics. Furthermore, neural networks exhibit flexibility towards the noisy and chaotic elements present in the market, while showing reduced sensitivity to error terms, assumptions, and constraints (Roumen, *et al*. (2017). In a research conducted by Maina, *et al*. (2023), the application of the Heston-ANN model was implemented to make predictions on datasets pertaining to three specific stocks: BA, IBM, and GOLD. Through the utilization of graphical examinations and the assessment of various model performance indicators such as Mean Absolute Percentage Error, Mean Absolute Error and Mean Squared Error, the investigation compares the integrated model with the basic Heston model. The results suggested that the Heston-ANN model provides more precise predictions in comparison to the original Heston model. The combined effect of the Heston model and Artificial Neural Network (ANN) results in the hybrid model being perceived as a more robust strategy for predicting stock prices. In their study, Shuaiqiang, *et al*. (2019) introduced a data-driven methodology employing an ANN to evaluate financial derivatives and calculate implied volatilities, with the aim of increasing computational techniques related to these procedures. Leveraging the universal function approximation capability of ANNs, the approach involves training an optimized ANN on a dataset derived from an advance financial model, subsequently using the trained ANN to speed up the original solver's operations in a swift and efficient manner. This methodology was tested across three different types of solvers, covering the analytic solution for the Black-Scholes equation, the COS method for the Heston stochastic volatility model, and Brent's iterative root-finding method for implied volatility calculations. The numerical results demonstrated a significant reduction in computing time achieved by the ANN solver. To fill this gap, a modified version of the Heston model (referred to as HJSF) was introduced in this research, incorporating Artificial Neural Networks (ANN). By means of a visual examination and the assessment of performance indicators such as Mean Squared Error (MSE), the research undertakes a comparative analysis between the hybrid model and the original Heston model. Utilizing a frequently applied feed-forward multilayer architecture, the research incorporates three distinct layers: the input layer, the hidden layer, and the output layer. Following this, the model implemented the Rectified Linear Unit (ReLU) activation function to effectively map historical observations (input data) to future results (output data).The choice of

ReLU as the activation function over alternative options was attributed to its rapid computational speed and absence of overloading the threshold, thereby aiding gradient descent.

3. Methodology

3.1. The Heston Model

The origin of the Heston model can be traced back to the inherent stochastic characteristics of an asset's volatility. This particular model is characterized by a system of stochastic differential equations (SDEs) as outlined in the subsequent section. These SDEs serve to describe the advancement of both the stock price and the volatility diffusion process within a designated probability model denoted as $\mathbb{P}$.
We assume that the stock price adheres to the following diffusion process at time t.

$$dS_t = \mu S_t dt + \sqrt{V_t} S_t dZ_t^1 \quad (1)$$

$$dV_t = k\,(\theta - V_t)\, dt + \sigma\sqrt{V_t}dZ_t^2 \quad (2)$$

and where Z_t^1 and Z_t^2 are correlated Wiener processes with ρ, i.e

$$dZ_t^1 dZ_t^2 = \rho dt \quad (3)$$

where the symbol μ represents the drift factor pertaining to the stock price, while θ signifies the long-term mean of variance. Additionally, k denotes the rate of mean reversion, and σ stands for the volatility of volatility. St and Vt correspond to the price and volatility process, respectively; where $\{V_t\}_{t \geq 0}$ is characterized as a square root mean-reverting process. This process was originally introduced with a long-run mean of θ and a rate of reversion k. In order to incorporate the leverage effect, it is observed that stock returns and implied volatility display a negative correlation. Furthermore, Z_t^1 and Z_t^2 are associated with a correlated Wiener process, with the correlation coefficients represented by ρ.

It is essential to note that all parameters including (μ, k, θ, σ, and ρ) are considered as time and stated similarly. The volatility follows an Ornstein-Uhlenbeck process (Elias, *et al.* 1991 In Maina, *et al.* 2023). The model assumes a constant correlation, denoted by ρ, existing between the two Brownian movements. The relationship between the price of the asset and its volatility yields a solution expressed in a closed mathematical form, facilitating the integration of random interest rates within the framework. To simplify, let r be a drift under risk neutrality. Then, equation (1) may be formulated as follows:

$$dS_t = rS_t dt + \sqrt{V_t} S_t dZ_t^1 \quad (4)$$

We compute the numerical value of $\mathbb{P}\ddot{X}[f(S_t)]$ for a payoff function f: $\mathcal{Q}\mathcal{N} \rightarrow \mathcal{Q}$ where a stochastic function refers to the underlying stock S:

$$dS_t = \alpha(S_t, t)dt + \beta(S_t, t)dZ_t \quad (5)$$

The discretization of the time interval and the simulation of the dynamics of the state process on a discrete-time grid are conducted. In this scenario, a discrete-time approximation of equation (5) is employed to obtain a Monte Carlo estimation of the process of $\mathbb{P}\ddot{X}[f(S_t)]$. The simulation generates multiple sample paths for state variables like interest rates, volatility, and stock price (Brandon, 2017).
This research employs the natural log of stock prices of the underlying asset $Y_t = ln[S_t]$ to facilitate the process of discretization. The application of Ito's lemma to equation (4) results in:

$$dY_t = r - \frac{1}{2}V_t dt + \sqrt{V_t}dZ_t^1 \quad (6)$$

$$dV_t = k\,(\theta - V_t)\, dt + \sigma\sqrt{V_t}dZ_t^2 \quad (7)$$

The research employs the natural logarithm of variance to avoid the presence of negative values during the continuous-discrete time transition. By applying Ito's lemma once more, the subsequent outcome is obtained:

$$dln(V_t) = \frac{1}{V_t}(\,k\,(\theta - V_t - \frac{1}{2}\sigma^2)\, dt + \sigma\frac{1}{\sqrt{V_t}}dZ_t^2 \quad (8)$$

Implementing the Euler discretization method yields the discrete-time solutions as presented below.

$$ln(S_{t+\Delta t}) = lnS_t + (r - \frac{1}{2}V_t)\Delta t + \sqrt{V_t}\sqrt{\Delta t} \in_{S,t+1} \quad (9)$$

$$ln(V_{t+\Delta t}) = lnV_t + \frac{1}{V_t}\,(\,k\,(\theta - V_t) - \frac{1}{2}\sigma^2)\,\Delta t + \sigma\frac{1}{\sqrt{V_t}}\sqrt{\Delta t} \in_{V,t+1} \quad (10)$$

Taking exponentials gives the Heston model:

$$S_{t+\Delta t} = S_t \cdot exp\,(r - \frac{1}{2}V_t)\Delta t + \sqrt{V_t}\sqrt{\Delta t} \in_{S,t+1}) \quad (11)$$

$$V_{t+\Delta t} = V_t \cdot exp\left(\frac{1}{V_t}\left(\mathbf{k}\left(\theta - V_t\right) - \frac{1}{2}\sigma^2\right)\Delta t + \sigma\frac{1}{\sqrt{V_t}}\sqrt{\Delta t}\ \epsilon_{V,t+1}\right) \quad (12)$$

3.2 Modified Heston Model (HJSF):

We expanded the traditional Heston model by adding jump and sentiment variables, thereby improving its efficiency in capturing market performances. The incorporation of the jump element brings about sudden and substantial fluctuations in prices, whereas the sentiment factor considers the impact of investor sentiment on asset pricing. The modified Heston model is expressed through a set of stochastic differential equations, allowing the simulation of the asset price and volatility paths across different market situations.

The assumptions of the new model

To introduce new assumptions for the modified Heston model with jump and sentiment factors, we consider the following extensions:

Correlated Jumps: We assume that the jumps in the asset price process are correlated with the jumps in volatility. This means that when there is a large positive (negative) jump in the asset price, there is also a tendency for volatility to increase (decrease) simultaneously. This correlation captures the idea that extreme movements in asset prices often coincide with changes in market volatility.

Time-Varying Sentiment: Instead of a constant sentiment coefficient (Ωμ), we introduce a time-varying sentiment factor that evolves over the simulation period. This allows us to model how investor sentiment changes over time in response to market conditions, news events, and other external factors. We specify a stochastic process for the sentiment factor like a jump-diffusion process or mean-reverting process.

Leverage Effect: We incorporated a leverage effect by allowing the correlation coefficient between asset price and volatility (ρ) to depend on the level of volatility. Empirical studies have shown that the correlation tends to be more negative when volatility is high, reflecting the tendency for asset prices to decline when volatility spikes. By incorporating these assumptions, we created a more realistic and flexible model that captures additional features of financial markets and enhances our ability to analyses and forecast asset price and volatility dynamics. We represented the modified Heston model with jump and sentiment factors using the following stochastic differential equations (SDEs):

$$\mathbf{dS_t} = (\mathbf{R} - \boldsymbol{\Omega\mu})\mathbf{S_t dt} + \sqrt{\mathbf{v_t}}S_t\mathbf{dZ_t^S} + J_t dt \quad (13)$$

$$\mathbf{dv_t} = k(\theta - \mathbf{v_t})dt + \sigma\sqrt{\mathbf{v_t}}d\mathbf{Z_t^v} + v dN_t \quad (14)$$

$$\mathbf{dJ_t} = (\alpha - \boldsymbol{\lambda}\mathbf{J_t})dt + \tilde{\eta}\, dQ_t \quad (15)$$

where:

S_t denotes asset price at time t.
V_t denotes the volatility at time t.
J_t i denotes the jump component representing the impact of jumps on the asset price.
N_t denotes, a Poisson process with intensity v, representing the occurrence of jumps.
μ denotes, the average jump size.
k denotes, volatility of the mean reversion rate.
θ denotes, volatility of the long-term average.
σ denotes, the volatility of volatility (variance).
ρ denotes, the correlation coefficient between Brownian motions.
ν denotes, the jump intensity.
R denotes, Risk free rate
α denotes, the mean jump size.
Z_t^v, Z_t^S and Q_t are standard Brownian motions.
$\boldsymbol{\lambda}$ denotes, the jump mean reversion rate.
η̃ denotes, the volatility of jumps.
Ω denotes, Sentiment factor

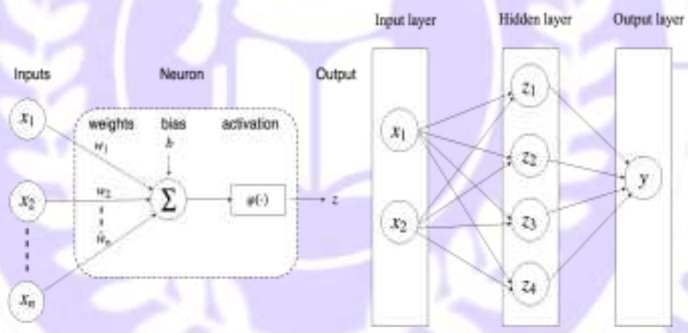
This system of SDEs captures the dynamics of the modified Heston model, incorporating the impact of jumps (modeled by the process Jt) and sentiment (reflected in the term Ωμ). The jump component J_t follows a mean-reverting process and its dynamics are influenced by both the mean jump size α and the jump mean reversion rate $\boldsymbol{\lambda}$. We used these equations to analyse the behaviour of the modified Heston model with jump and sentiment factors.

3.2. Artificial Neural Network

Artificial Neural Networks (ANNs) represent effective machine learning model with the ability of learning complex patterns from historical data. We utilized ANN to enhance volatility predictions produced by the modified Heston model. Through the training of ANN on simulated data derived from the modified Heston model, it becomes feasible to cover further non-linear correlations and enhance the accuracy of volatility forecasts. This study utilized a feed-forward neural network with a single hidden layer that underwent optimization through the adaptive moment estimation (Adam optimization) technique. The methodology was based on the foundational concept of gradient descent, modifying the learning rate for all parameters of the neural network by approximating the first and second-order statistical characteristics of the gradient. The Adam optimizer was chosen over other optimization methods because of its reduced need for hyper-parameter tuning and its exceptional computational efficiency.

3.3 Structural Architecture of *ANN*

Artificial Neural Networks (ANNs) typically comprised three hierarchical levels of constituents, namely neurons, layers, and the structure from the base to the summit. The arrangement of the structure was ascertained through a fusion of diverse layers, each composed of a multitude of synthetic neurons. A neuron, incorporating modifiable weights and biases, stands as the pivotal module of ANNs. By creating connections among neurons in successive layers, the output signals originating from a preceding layer goes into the next layer in the form of input signals. The transmission of signals from the input layer through hidden layers to the output layer was facilitated by the overlay of layers, potentially involving cyclic or recurrent links. This process allows the Artificial Neural Network (ANN) to establish a relationship between input and output pairs. As Illustrated in Figure 1a, an artificial neuron essentially comprises the subsequent three consecutive operations: The computation involves determining the summation of inputs that are weighted accordingly. In addition, a bias is included in the calculated summation. The output is then computed using a transfer function.



(a)Aneuro (b)A exampl

Figure 1: MLP Configuration Illustration of ANN. (Zaheer, 2004)

A Multi-Layer Perceptron (MLP), which comprises a minimum of three layers, represents the most basic form of an Artificial Neural Network (ANN). Mathematically, Multilayer Perceptrons (MLPs) are characterized by a set of specific parameters.

$$\theta = (W_1, b_1, W_2, b_2 ... W_L, b_L) \quad \textbf{(16)}$$

The weight matrixes, denoted as Wj, and the bias vector, represented as bj, correspond to the parameters of the neural layer indexed by L-th. Subsequently, a function can be formally expressed in the subsequent manner.

$$y(x) = F(x/\boldsymbol{\theta}). \quad \textbf{(17)}$$

Let $Z_j^{(l)}$ denote the value of the j-th neuron in the l-th layer, then the corresponding transfer function would read,

$$Z_j^{(l)} = \phi^{(l)}\left(\sum_i w_{ij}^{(l)} Z_i^{(l-1)} + b_j^{(l)}\right), \quad \textbf{(18)}$$

where $Z_i^{(l-1)}$ is the output value of the i-th neuron in the (l-1) th layer and $\phi(\cdot)$ is an activation function, $w_{ij}^{(l)} \in W_l$, $b_j^{(l)} \in b_l$. When l = 0, $Z^{(0)} = x$ is the input layer; When l = L, $Z^{(L)} = y$ is the output layer; otherwise, $Z^{(l)}$

represents an intermediate variable. The activation function $\phi(\cdot)$ adds non-linearity to the system. For example, any of the following activation functions may be employed,

1. Relu, $\phi(x) = \max(x, 0)$,
2. Sigmoid, $\phi(x) = \frac{1}{1+e^{-x}}$
3. Leaky ReLu, $\phi(x) = \max(x, ax)$, $0 < a < 1$;

Equation (19) presents an example of the formula of an MLP with "one-hidden-layer",

$$\begin{cases} y = \phi^{(2)}\left(\sum_i w_j^{(2)} Z_j^{(1)} + b^2\right) \\ Z_j^{(1)} = \phi^{(1)}\left(\sum_i w_{ij}^{(1)} x_i^{(1)} + b_j^{(1)}\right) \end{cases} \quad \textbf{(19)}$$

The Universal Approximation Theorem states that a single-hidden-layer of Artificial Neural Network (ANN) possessing an adequate quantity of neurons which has the capability to approximate all continuous functions. The measurement of the difference between two functions is conducted through the utilization of the function norm $\|\cdot\|$.

$$D(f(x), F(x)) = \|f(x) - F(x)\|, \quad \textbf{(20)}$$

The objective function f(x) is represented as F(x), which serves as the approximated function of the neural network. An illustration of this concept can be seen in the representation of the p-norm.

$$\|f(x) - F(x|\theta)\|p = \sqrt[p]{\int x|f(x) - F(x|\theta)|p d\mu(x)},$$

In the context where $1 \leq p < \infty$ and $\mu(x)$ represents a measure defined on the space X, the selection of p = 2 is made for the purpose of evaluating the average accuracy, aligning with the concept of Mean Squared Error (MSE). Within the supervised learning, the loss function can be viewed as being identical to the distance mentioned above,

$$L(\theta) := D(f(x), F(x|\theta)). \quad \textbf{(21)}$$

The training procedure is designed to acquire the most suitable weights and biases for the application of Equation (17) in order to minimize the loss function. This procedure can be characterized as an optimization problem.

$$\arg\min L(\theta|(x,y)), \quad \textbf{(22)}$$

Given, are the predefined input-output pairs (x, y) accompanied by a loss function denoted as $L(\theta)$. Various methodologies rooted in back-propagation gradient descent have demonstrated to be efficient in tackling Equation (22), including techniques like Stochastic Gradient Descent (SGD) and its iterations Adam and RMSprop. The optimization procedures initiate with preset initial values and progress towards the minimization of the loss function. The formula for adjusting the parameters can be expressed as follows:

$$\begin{cases} W \leftarrow W - \eta^{(i)} \frac{\partial L}{\partial W,} \\ b \leftarrow b - \eta^{(i)} \frac{\partial L}{\partial b,} \\ i = 0,1,2, \ldots, \end{cases} \quad \textbf{(23)}$$

The parameter η represents a learning rate that has the potential to fluctuate throughout the iterative process. The significance of the learning rate becomes evident during the training phase, where a high value can lead to oscillations in the convergence of Artificial Neural Networks (ANNs), while a low value may hinder the learning process by causing ANNs to progress slowly and potentially become stuck in localized optimal areas. In such cases, an adaptive learning rate is often recommended.

3.4. Integration Approach

The combination of the modified Heston model with Artificial Neural Networks (ANN) comprises of two primary stages. Firstly, a large datasets of asset price and volatility paths is generated by implementing the modified Heston model across various scenarios. Subsequently, the ANN is trained using this simulated data to learn the mapping between input variables (such as historical asset prices, volatility levels) and the projected volatility. Following its training, the ANN is capable of creating volatility forecasts based on prevailing market conditions, enhancing the predictions derived from the modified Heston model.

3.5 Evaluation of Model Performance

We measured the performance of Artificial Neural Networks (ANNs) in addressing financial models through the examination of accuracy metrics, which served as the basis for the training process, as demonstrated in the equations provided below,

Mean Square Error (MSE) or Mean Square Prediction Error (MSPE)

$$\text{MSE} = \frac{1}{n}\sum(y_i - \hat{y}_i)^2 \quad \textbf{(24)}$$

Root Mean Square Error (RMSE

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum(y_i - \hat{y}_i)^2} \quad \textbf{(25)}$$

Mean Absolute Error (MAE)

$$\text{MAE} = \frac{1}{n}\sum|y_i - \hat{y}_i| \quad \textbf{(26)}$$

where y_i is the actual value and $\hat{y}_i$ is the ANN predicted value.

The MSE would be used as the training metric to update the weights, and to evaluate the selected ANN.

3.6 Experimental Results

Table 1 The Heston parameter ranges for training the ANN

HESTON-ANN	**Parameters**	**Value**
	Initial Asset stock price St	100
	Initial volatility Vt	0.05
	Time horizon T	1
	Number of time steps N	1000
	Time step size dt = T / N	
	Risk-free rate (**R**)	0.05
	Mean reversion rate (**λ**)	2
	Long-term volatility level (θ)	0.04
	Sentiment factor (Ω)	0.02
	Volatility of volatility (σ)	0.3
	Jump intensity (v)	0.1
	Average jump size (μ)	0.1
	Initial sentiment coefficient (Ωμ)	0.02
	Initial correlation Coefficient (**ρ**)	-0.5
Input	Threshold for leverage effect (ρ)	0.1
Output	European call price	

As shown in Figure 2, we employed random search for the hyper-parameters on a small-sized data set, and train the selected ANN on a large data set. The training and validation losses remained close, which indicates that there was no over-fitting.

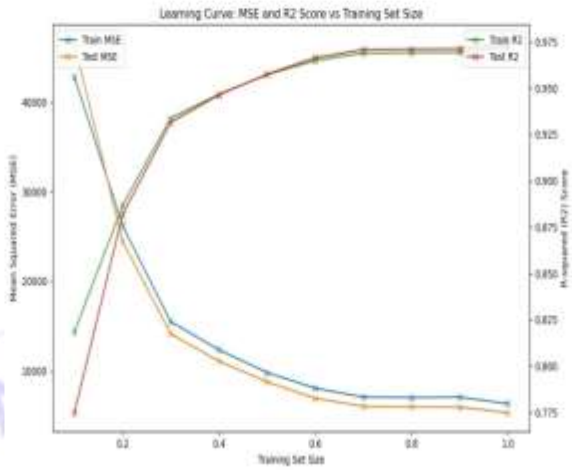


Figure 2 MSE vs. R^2 and size of the training set.

The figure shows improvement in the loss, which is an indicator of the model performance.

Table 2 the trained Heston (JSF)-ANN performance

Heston(JSF)-ANN	**MSE**	**MAE**	**MAPE**	R^2
Training	3.54×10^{-6}	6.92×10^{-5}	6.76×10^{-4}	0.8996
Testing	2.56×10^{-6}	7.51×10^{-5}	7.27×10^{-4}	0.8995

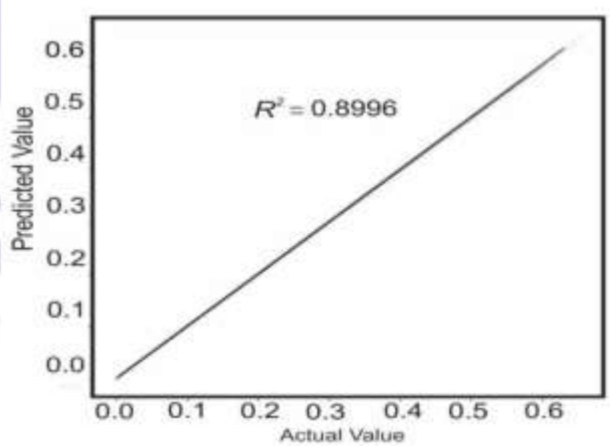


Figure 3 Out-of-sample Heston (JSF)-ANN performance.

Figure 3 shows a strong positive relationship between the predicted value and the actual value.

Table 3 Training Progress for performance

Training Progress

Unit	Initial Value	Stopped Value	Target Value
Epoch	0	9	1000
Elapsed Time	-	00:00:05	-
Performance	5.47	0.732	0
Gradient	15.6	0.0709	1e-07
Mu	0.001	0.001	1e+10
Validation Checks	0	6	6

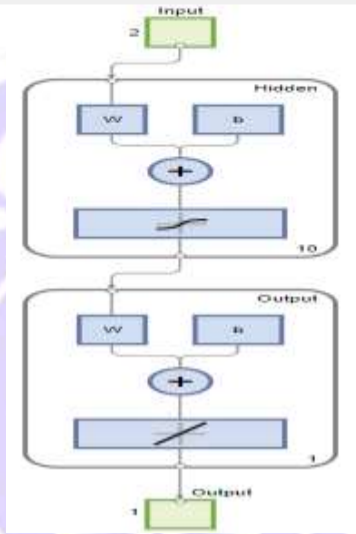


Figure 4 Training Plots

Figure 5 Validation Performance of the MH-ANN

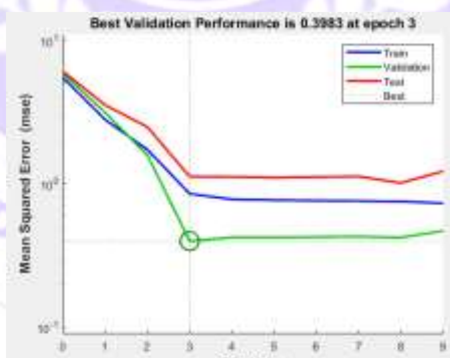
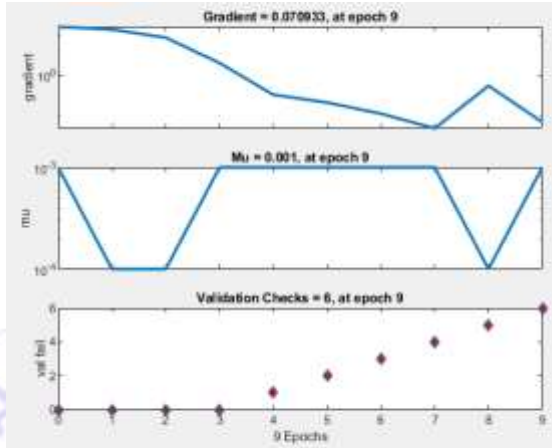


Figure 6 Training State of HM-ANN

Figure 6 shows the best validation performance of 0.3983 at epoch 3 using MSE performance indicator.



In the context of the modified Heston model, the gradient indicates that the speed at which the volatility of the asset is expected to change over time or under different market conditions. A positive gradient suggests an increasing trend in volatility over time, while a negative slope suggests a decreasing trend. From Figure 6 gradient value (0.070933) suggests an increasing trend in volatility over time, indicating positive gradient.

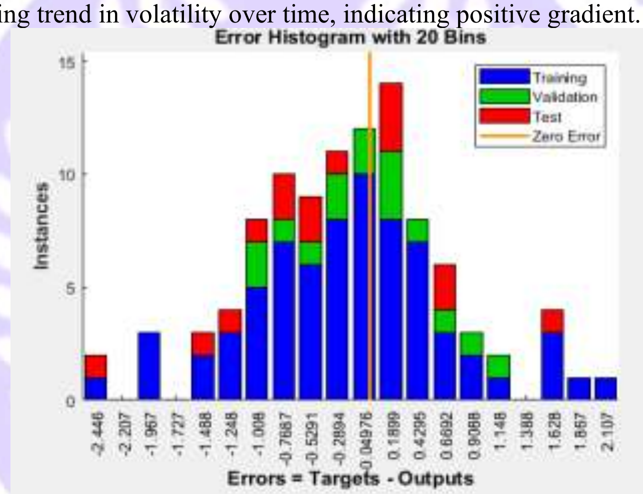


Figure 7 Error Histogram of MH-ANN

Figure 7 shows that the Error Histogram of MH-ANN is 0.04976 indicating good performance of the model.

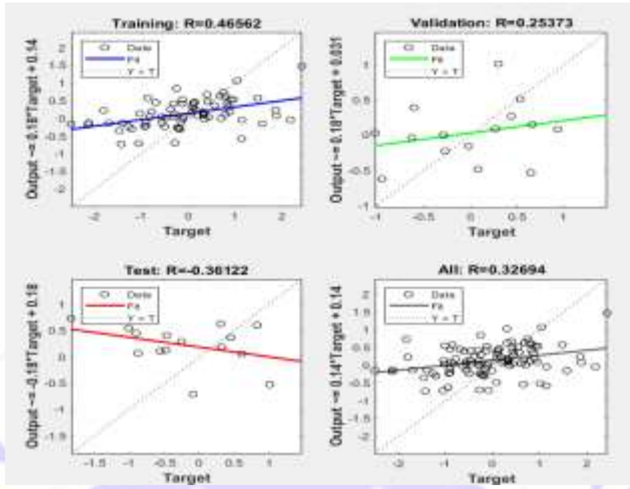


Figure 8 Regression Analysis of HM-ANN

Figure 8 shows Regression Analysis (R^2) of HM-ANN Training, Testing, Validation and All.

Interpretation of the regression line analysis of a modified Heston model involves understanding how the model parameters (slope, intercept) relate to volatility dynamics over time or under different market conditions and assessing the goodness of fit of the model to the observed data. The regression analysis provides measures of goodness of fit such as R^2 (coefficient of determination), which indicates the proportion of the variance in the dependent variable that is explained by the independent variable. A high R^2 value suggests that the regression line fits the data well, while a low R^2 value suggests poor fit.

Training-R= 0.46562, Testing-R = -0.36122, Validation-R = 0.25373 and All-R = 0.32694

From Figure 8 above, the goodness of fit of the model to the observed data R^2 value suggests that the regression line fits the data well at 0.32694.

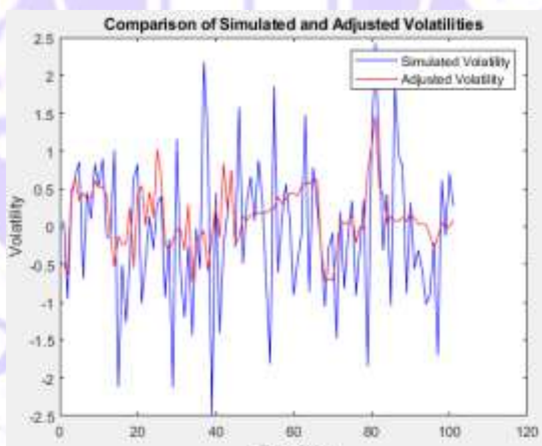


Figure 9 comparison of simulated and adjusted volatility

Adjusted volatilities typically refer to historical volatilities that have been adjusted or smoothed in some manner to better reflect the underlying risk of a financial instrument. They are often calculated using statistical techniques such as moving averages, exponential smoothing, or other methods to filter out noise or short-term fluctuations. Adjusted volatilities are commonly used in options pricing models, risk management, and portfolio optimization to estimate future volatility based on historical data while accounting for trends or anomalies in the data. While simulated volatilities, on the other hand, are volatilities that are generated through simulations rather than being directly observed from historical data. They are often used in Monte Carlo simulations or other stochastic modeling techniques to estimate the range of possible future outcomes for a financial instrument or portfolio; simulated volatilities can be based on assumed distributions of future returns or can be generated using historical data in combination with assumptions about future market conditions. These volatilities are useful for scenario analysis, stress testing, and risk assessment by providing a range of potential outcomes under different market conditions. In summary, adjusted volatilities are derived from historical data and are smoothed or adjusted to better represent underlying risk patterns, while simulated volatilities are generated through simulations and provide a range of possible future outcomes based on assumed or observed market dynamics.

Figure 9 shows the comparison of simulated and adjusted volatility using the MH-ANN model. It enables the investor to know the patterns of a particular investment. Volatilities are adjusted or smoothed in some manner to better reflect the underlying risk of a financial instrument. Adjusted volatilities are commonly used in options pricing models, risk management, and portfolio optimization to estimate future volatility based on historical data while accounting for trends or anomalies in the investment. These volatilities are useful for scenario analysis, stress testing, and risk assessment by providing a range of potential outcomes under different market conditions.

We conducted an experiments to evaluate the performance of the integrated approach in volatility forecasting compared to standalone models. Using historical market data, we compare the accuracy of volatility forecasts generated by the modified Heston model, ANN, and the integrated approach. Our results revealed that the integrated approach outperforms standalone models, providing more accurate and robust volatility predictions across different market conditions.

4. Conclusion

In this paper, we presented an innovative approach for volatility forecasting in financial markets by integrating the modified Heston model with Artificial Neural Networks. By combining the strengths of both models, we offer a robust framework for improved volatility predictions, enhance the accuracy and reliability of volatility forecasts and enabling better risk management and decision-making in financial markets. Future research directions should include exploring more sophisticated neural network architectures and incorporating additional market factors for further improvement. This research paper provides a comprehensive overview of the integration of the modified Heston model with Artificial Neural Networks for volatility forecasting in financial markets.

5. Recommendations

Based on the research findings, we provide the following recommendations:

i. Practitioners in financial markets should consider adopting integrated approaches that combine traditional models with machine learning techniques for volatility forecasting.

ii. Further research is warranted to explore alternative neural network architectures and data pre-processing techniques to enhance the performance of integrated models.

iii. Market participants should continuously monitor and update their forecasting models to adapt to changing market conditions and incorporate new information.

Reference

Ayala, J., Garc´ıa-Torres, M., Jos´e, L., V´azquez, N., G´omezVela, F. and Divina, F. (2021). *Technical analysis strategy optimization using a machine learning approach in stock market indices*. Knowledge-Based Systems, 225: 107119.

Beni, L. and Schultz, P. (1990). Pricing warrants an empirical study of the Black-Scholes model and its alternatives. *The Journal of Finance,* 45 (4): 1181–1209.

Brandon, H. (2017). *Implementing the Heston option pricing model in object-oriented cython*.

Chandan, S. (2004). *Financial modeling using excel and VBA.* John Wiley and Sons.

Elias, M. S. and Jeremy, C. S. (1991). *Stock price distributions with stochastic volatility*: an analytic approach. The review of financial studies, 4 (4): 727–752,

I-Ming, J., Jui-Cheng, H., and Chuan-San, W. (2014). *Volatility forecasts*: Do volatility estimators and evaluation methods matter? *Journal of Futures Markets*, 34 (11): 1077– 1094.

Kazuhisa, M. (2004). *Introduction to Merton jump diffusion model*. Department of Economics. The Graduate Center, the City University of New York.

Ladde, G., S. and Ling, W. (2009). Development of modified geometric Brownian motion models by using stock price data and basic statistics. *Nonlinear Analysis: Theory, Methods and Applications*, 71 (12): 1203–1208.

Looney, C., G. (1997). *Pattern recognition using neural networks*: theory and algorithms for engineers and scientists: Oxford University press. New York,

Luca, D. P., and Oleksandr, H. (2016). Artificial neural networks approach to the forecast of stock market price movements. *International Journal of Economics and Management Systems* *http://iaras.org/iaras/journals/ijems*

Maina, A., Mwalili, S., Malenje, B. (2023). Forecasting Stock Prices Using Heston-Artificial Neural Network Model. *International Journal of Data Science and Analysis*. Vol. 9, No. 2, pp. 22-33. doi: 10.11648/j.ijdsa.20230902.11

Naiga, B. C., Mung'atu, J., Nafiu, L. A., and Adjei, M. (2022): On Modified Heston Model for Forecasting Stock Market Prices," *International Journal of Mathematics Trends and Technology (IJMTT)*, 68(1), 115-129. https://doi.org/10.14445/22315373/IJMTT-V68I1P513

OpenAI. (2024). *ChatGPT* (Mar 20 version) [Large language model]. https://chat.openai.com/chat

Ormonde, F, T. (2017). *On the numerical methods for the Heston model*. PhD thesis. *All Graduate Plan B and other Reports, Spring 1920 to Spring 2023*. 1146. https://digitalcommons.usu.edu/gradreports/1146

Pierre, G. and Possama, D. (2010). Efficient Simulation of the Double Heston Model *(January 11, 2010). Available at SSRN: https://ssrn.com/abstract=1434853 or http://dx.doi.org/10.2139/ssrn.1434853*.

Preethi, G. and Santhi, B., (2012). Stock market forecasting techniques: A survey. *Journal of Theoretical and Applied Information Technology,* 46 (1).

Reichert, B. and Souza, A. M. (2022.). Can the Heston Model Forecast Energy Generation? A Systematic Literature Review. *International Journal of Energy Economics and Policy*, *12*(1), 289–295. https://doi.org/10.32479/ijeep.11975

Roumen, T., Radoslav, Y., Galyac, P., and Georgi, T. (2017). Artificial neural network intelligent method for prediction. In AIP Conference Proceedings, volume 1872, page 020021. AIP Publishing LLC.

Shuaiqiang, L., Cornelis, W., Oosterlee, S. and Bohte, M. (2018) Pricing options and computing implied volatilities using neural networks: *Applied Mathematics (DIAM),* Delft University of Technology, Amsterdam, Netherlands. *7*(1), 16: doi: 10.3390/risks7010016

Shunrong, S., Haomiao, J., and Tongda, Z. (2012). Stock market forecasting using machine learning algorithms. Department of Electrical Engineering, Stanford University, Stanford, CA, pages 1–5.

Steven L Heston (1993). A closed-form solution for options with stochastic volatility, with applications to bond and currency options," review of financial studies.

Teneng, D. (2011). Limitations of the Black-Scholes model. Collection of Papers, 1: 143.