Research Article

Heston-GA Hybrid Option Pricing Model Based on ResNet50

Zheng Yang, Liqin Zhang, Xiangxing Tao, and Yanting Ji [ID]

Department of Mathematics, Zhejiang University of Science and Technology, Hangzhou 310023, China

Correspondence should be addressed to Yanting Ji; yanting.ji@zust.edu.cn

(1) Background. This study aims to improve the accuracy of the pricing model. *(2) Methods*. Heston model is combined with ResNet50 convolutional neural network model. Based on the optimization of Heston model parameters by genetic algorithm (GA), ResNet50 model is used to correct the deviation between market option price and Heston price, so a new hybrid option pricing model is established based on the empirical research on the European call options of Huatai-PB CSI 300ETF (code 510300), Harvest CSI 300ETF (code 159919), and SSE 50ETF (code 510050). *(3) Results*. The pricing result of the hybrid model is better than other single models and hybrid models. The model is applicable to the pricing of options with short and long remaining terms. *(4) Conclusions*. It is shown that the combination of Heston model and ResNet50 model with optimized parameters can ensure the interpretability of the model and enhance the nonlinear fitting ability of the model, which confirms the effectiveness of the hybrid model and provides a reference for investors and institutions to make scientific decisions.

1. Introduction

In 1973, Black, Scholes, and Merton made a major breakthrough in European option pricing and developed the classic Black-Scholes-Merton model (BSM model) [1, 2]. Aiming at the assumption that the volatility of BSM model is constant, Stein et al. [3] proposed Stein-Stein stochastic volatility model by introducing the stochastic process of the underlying asset price volatility, where the asset price and volatility follow their respective stochastic processes. In 1993, Heston [4] applied the mean-reversion square root process to address the problem of negative volatility values in Stein-Stein model. Heston model has been widely used and studied because of the advantage of portraying a reasonable implied-volatility shape and the existence of the closed-form solution [5–7]. Up to now, the optimization of Heston model from the mathematical perspective mainly involves five directions: considering the randomness of the constant terms [8–12], improving the geometric Brownian motion [13, 14], adding a jump diffusion process [15, 16], changing the power of the variance [17, 18], and considering the roughness of the volatility [19, 20]. The mathematical enhancement endowed Heston models with more rigorous fitting to real market dynamics, but two problems remain:

the lack of robustness due to insufficient consideration of the factors influencing the underlying asset prices and the lack of practical usability due to the complexity of the model structure.

The above BSM-based models are classified as parametric models, while the neural network (NN) option pricing models are called nonparametric models. NN is a nonlinear model driven by market data, whose structure and parameters are determined from the acquired financial sample data without any assumption. This property makes NN models more suitable for option pricing with unknown underlying asset price dynamics and the pricing equation without analytical solution. The early application of NNs in option pricing can be traced back to 1993, when the NN model used by Malliaris et al. [21] successfully outperformed BSM model. Mitra [22] improved NN's pricing performance from the perspective of structure optimization. According to BSM, the input layer of the NN model was designed as 5 variables in BSM solution, the hidden layer was the two normal distribution values of BSM, and the output layer was the option price. Yacin et al. [23] considered volatility as a random variable and input it into the NN model. Results showed that the NN model was more accurate than Heston model. In addition, many other researchers have extensively

studied the application of NNs in options pricing [24–28]. However, due to 'black box' problem, constructing the option price NN model still lacks rigorous mathematical proof, and the constructed model can hardly be explained by traditional financial theories.

Maciej et al. [29] found that NN models cannot adapt to various market conditions; i.e., they are not robust, and they proposed an idea to incorporate the NN model with BSM model or Heston model. Combining the financial model and the NN model to build a hybrid model can effectively overcome the drawbacks of both, improving the nonlinear fitting ability while enhancing the interpretability. Arin et al. [30] used a deep multilayer perceptron to fit BSM option price deviations where the ratio of strike price to stock price exceeds a certain threshold and thus built a hybrid deep neural network model. Hideharu [31] combined the advantages of AE (asymptotic expansion) and NN, used NN to train the residual between option price C and its asymptotic approximation $\tilde{C}$ to improve stability and approximation accuracy, and used the hybrid model for option pricing in both local and stochastic volatility models. Zhang et al. [32] replaced BSM model in the traditional hybrid NN models with Heston stochastic volatility model, and Yang et al. [33] modified the pricing bias of Heston model using a nested long short-term memory neural network (NLSTM). Their results are all better than single models. Inspired by the above hybrid BSM-NNs model, we optimize Heston model from outside rather than inside, i.e., using an NN model to correct the nonlinear deviation of Heston prices from market prices.

Both classical option pricing models and NN models have their own flaws: (1) the classical parametric option pricing model must follow strict assumptions, which will result in systematic deviation under large difference between the initial condition and the real market. (2) The NN models have some problems such as insufficient fitness for discontinuous trading. So, we explore a hybrid model which can avoid their weaknesses and make full use of their strengths. The empirical results also show that hybrid NNs models have quite adaptability to ETF options pricing. In fact, the option price, which is affected by many factors, is not unary time series. At present, the term of ETF option contracts in China is up to six months, and there are several option contracts which have different maturities and strike prices in the market at the same time. Therefore, it is necessary to comprehensively consider multiple option contracts for its pricing. Based on the above conclusion, we choose a convolutional neural network (CNN) model, i.e., construct a hybrid CNN model based on Heston model for option pricing. In our research, ResNet50 CNN model is used to correct the pricing bias between the Heston price and the market price. Our hybrid model gives full play to the precision pricing performance of classical option model and nonlinear fitting performance of deep NNs model. Furthermore, our hybrid model is used to carry out empirical research on Huatai-PB CSI 300ETF option, Harvest SCI 300ETF option, and SSE 50ETT option. Experimental results illustrate the effectiveness and high performance of our hybrid model.

2. Theoretical Model of Option Pricing

The hybrid CNN model in this paper includes both Heston option pricing model and ResNet50 CNN model. Both are described separately, followed by an explanation of the steps in genetic algorithm (GA) for estimating the parameters of Heston model, and finally there is an explanation of the overall model construction process.

2.1. Heston Model. Traditional BSM model assumes that the volatility of underlying asset price is constant, but, in real financial market, volatility can change and there is a "volatility smile" phenomenon. For this reason, Heston used the square root mean-reversion process [4] to describe the real-time volatility variance, with the underlying asset price and the volatility movement according to the following differential equation:

$$\begin{aligned} dS_t &= \mu S_t dt + \sqrt{V_t} S_t dW_t^1, \\ dV_t &= \kappa(\theta - V_t)dt + \sigma\sqrt{V_t} dW_t^2, \end{aligned} \tag{1}$$

$$dW_t^1 dW_t^2 = \rho dt, \tag{2}$$

where S_t and V_t represent, respectively, the underlying asset price and the variance of the asset price at time t and the parameter μ represents the mean value of the asset price. $\{V_t\}_{t\geq 0}$ follows a square root process. The parameter θ corresponds to the long-run average of V_t and κ is the mean-reversion speed of the volatility. The parameter ρ is the correlation between the underlying asset and the volatility while σ is the volatility of the variance of returns. Also, W_t^1 and W_t^2 are two standard Brownian motions.

Theorem 1. *Assume that the underlying asset price S_t and its volatility V_t obey (1) at time t $(0 \leq t \leq T)$; then the European call option price $C(S_t, V_t, t)$ satisfies the following partial differential equation (PDE), denoting S_t and V_t by S and V for simplified notation here:*

$$\frac{\partial C}{\partial t} + \frac{1}{2}VS^2\frac{\partial^2 C}{\partial S^2} + \frac{1}{2}\sigma^2 V\frac{\partial^2 C}{\partial V^2} + \sigma\rho SV\frac{\partial^2 C}{\partial S\partial V} + rS\frac{\partial C}{\partial S}$$
$$+ [\kappa(\theta - V) - \lambda(S, V, t)]\frac{\partial C}{\partial V} - rC = 0, \tag{3}$$

where the parameter r is the risk-free interest rate and $\lambda(S_t, V_t, t)$ is volatility risk premium. And, at time T, the PDE (2) satisfies the following boundary conditions: (1) $C(S, V, T) = Max(0, S - K)$, (2) $C(0, V, t) = 0$, (3) $\partial C/\partial S(\infty, V, t) = 1$, and (4) $(rS\ \partial C/\partial S(S, =, t)) + (\kappa\theta\ \partial C/\partial V(S, =, t)) - rC(S, =, t) + C_t(S, =, t) = 0$, (5) $C(S, \infty, t) = S$.

Theorem 2. *Based on the PDE of (3), the price $C(S_t, V_t, t)$ of a nonpaying dividend European call option with expiry date T and strike price K has the following closed-form solution:*

$$C(S_t, V_t, t) = S_t P_1(S_t, V_t, t) - Ke^{-r(T-t)}P_2(S_t, V_t, t), \tag{4}$$

where P_1 and P_2 represent the probability distribution function. The mathematical expression is shown as follows:

$$P_j = \frac{1}{2} + \frac{1}{\pi} \int_0^{+\infty} Re\left[\frac{\exp(-i\varnothing \ln(K) f_j(x, V_t, T-t, \varnothing))}{i\varnothing}\right] d\varnothing, \quad (5)$$

$$f_j(x, V_t, T-t, \varnothing) = \exp\{i\varnothing x + C_j(T-t, \varnothing) + D_j(T-t, \varnothing)V_t\}, \quad (6)$$

where $j = 1, 2$, f_1, f_2 represent the characteristic functions of P_1, P_2. $Re(y)$ represents the real part of y, i is an imaginary unit, and $\varnothing$ is the integral variable. The expressions for each part of (5) are shown as follows:

$$x = \ln(S_t),$$

$$C_j(T-t, \varnothing) = ri\varnothing(T-t) + \frac{a}{\sigma^2}\left\{(b_j - \rho\sigma\varnothing i - d_j) \cdot (T-t) - 2\ln\left[\frac{1 - g_j e^{d_j(T-t)}}{1 - g_j}\right]\right\},$$

$$D_j(T-t, \varnothing) = \frac{b_j - \rho\sigma\varnothing i + d_j}{\sigma^2} \cdot \left[\frac{1 - e^{d_j(T-t)}}{1 - g_j e^{d_j(T-t)}}\right], \quad (7)$$

$$g_j = \frac{b_j - \rho\sigma\varnothing i + d_j}{b_j + \rho\sigma\varnothing i + d_j},$$

$$d_j = \sqrt{(\rho\sigma\varnothing i - b_j)^2 - \sigma^2(2u_j\varnothing i - \varnothing^2)},$$

where $j = 1, 2$, $a = \kappa\theta$, $u_1 = 1/2$, $u_2 = -1/2$, $b_1 = \kappa + \lambda - \rho\sigma$, and $b_2 = \kappa + \lambda$.

In addition, since volatility is not a tradable commodity in the market, the volatility risk premium is an unobservable measure. Heston assumes that $\lambda(S_t, V_t, t) = kV_t$, which is proportional to volatility, is independent of investors' risk attitudes. So, λ can be eliminated by equivalent martingale measure and volatility risk price $\lambda = 0$ in a risk-neutral world. Then, the above parameters can be expressed as

$$\kappa^* = \kappa + \lambda, \ \theta^* = \frac{\kappa\theta}{\kappa + \lambda}, \ a = \kappa^*\theta^*, \ b_1 = \kappa^* - \rho\sigma, \ b_2 = \kappa^*. \quad (8)$$

The parameters to be estimated of Heston model also change from $\{\kappa, \theta, \sigma, V_0, \rho, \lambda\}$ to $\{\kappa^*, \theta^*, \sigma, V_0, \rho\}$.

2.2. Convolutional Neural Network Model: ResNet50.

Convolutional neural network is a deep neural network model of multilayer supervised learning, with the convolutional layer and the sampling layer being the core modules that implement feature extraction. The convolutional layer performs local connectivity and weight sharing between neurons for feature extraction, and the sampling layer extracts the most representative features to achieve dimensionality reduction. Since Fukushima [34] first proposed convolutional neural network (CNN) in 1980, CNN has been widely used in machine learning tasks such as image recognition, face verification, speech recognition, and text classification. Classical models include LeNet, AlexNet, VGG, ResNet, and DenseNet. CNN, as the most representative network model in deep learning, was initially applied to image recognition, while option data has similar characteristics to image data, such as high dimensionality and local clustering, that is, multiple influencing factors, self-similarity, and memory of its own price movement process. Additionally, with the deepening of neural network layers, the accuracy of the model can be effectively improved, but the gradient disappearance and network degradation problems will also arise. Taking all factors into consideration, ResNet50 model is chosen in this paper to correct the deviation caused by Heston model.

ResNet model is a residual network with modular thinking [35]. The key idea is shortcut connection; that is, the input of the residual module is directly added to the output of the module through identity mapping. The basic structure of the residual module is shown in Figure 1, and its mathematical expression is

$$F(x) = W_2\sigma(W_1 x),$$
$$y = F(x, \{W_i\} + x), i = 1, 2, \quad (9)$$

where x is the input of the residual module, W_1 and W_2 are the weight matrix of the two-layer convolution layer, σ is the
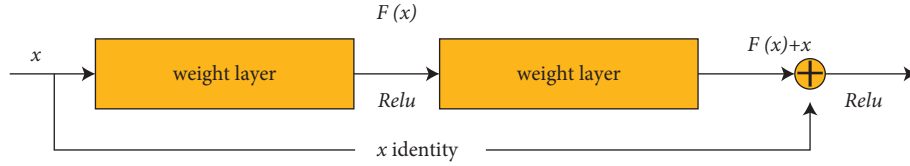
$$F(x)$$

FIGURE 1: Basic structure of residual module.

ReLu nonlinear activation function, and y is the output of the residual module.

ResNet50 consists of 16 residual modules, each of which has three convolutional layers. Since the dimensions of input and output are inconsistent in some ResNet50 residual modules, it is necessary to change the dimensions of input and then add it to the output, usually by adding a convolution process. To facilitate differentiation, the module whose input is mapped through the identity mapping is called Identity Block, and the module whose input is mapped through the convolution layer is called Conv Block. The structure diagrams of Identity Block and Conv Block of ResNet50 are shown in Figure 2, where batch norm represents the batch normalization layer and $1 * 1$ and $3 * 3$ represent the size of convolution kernel.

ResNet50 original model was made up of 49 convolutional layers and 1 fully connected layer (excluding pooling layer, batch normalization layer, and activation function layer). When He et al. [35] proposed this model, colour images with size of $224 * 224$ pixels were taken as input data. After feature extraction by one convolutional layer and one pooling layer, the image was flattened by entering a 16-residual module convolution operation and finally outputting a 1000-dimensional vector through a fully connected layer. Therefore, when ResNet50 model is applied to option pricing, it is necessary to construct data input in the form of images, that is, two-dimensional ($k = 1$) or three-dimensional ($k \geq 2$) data form of $m * n * k$. These data can be composed of several options with certain characteristics (such as the same expiration month). Specific data include factors affecting the option price, such as strike price, remaining term, and underlying asset price. The ZeroPadding layer, pooling layer, and activation function are also adapted to actual input data size and output data type.

2.3. Genetic Algorithm.
Option pricing with Heston model requires estimation of the model parameters firstly, which have five parameters to be estimated through an equivalent martingale measure transformation. In this paper, genetic algorithm is used to estimate parameters of Heston model [36]. Genetic algorithm is a random global search and optimization method based on biological genetic and evolutionary mechanism created by Professor Holland at the University of Michigan, inspired by biosimulation techniques, which overcomes the disadvantage of traditional algorithms that are prone to fall into local extremums. Genetic algorithm obeys the law of survival of the fittest in nature and simulates the phenomena of inheritance, crossover, and mutation that occur during the biological
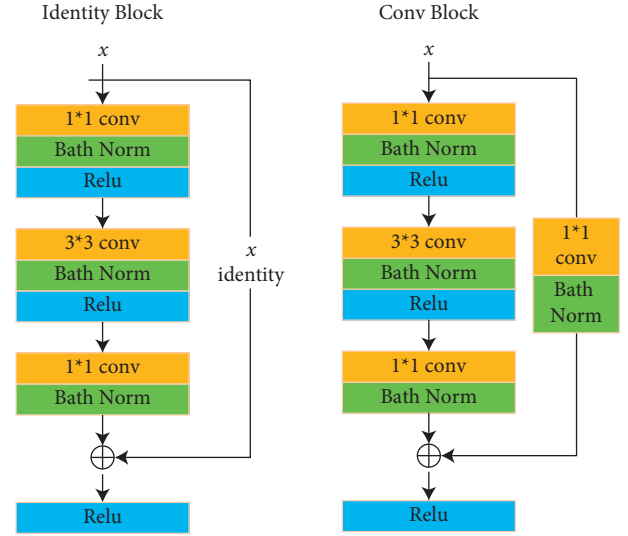

FIGURE 2: ResNet50 Identity Block and Conv Block.

evolution. Starting from a randomly generated generation of populations, the optimal solution to the problem is obtained through the selection of fitness functions and the generation after generation of individuals that are best adapted to their environment through selection, crossover, and mutation. The basic flow of genetic algorithm is shown in Figure 3.

The specific steps of the algorithm to estimate Heston model parameters are as follows.

2.3.1. Step 1. Population Initialization.
Encoding methods can be divided into three main categories: binary encoding, symbol encoding, and floating-point encoding, where we adopt binary encoding. The binary encoding method uses a fixed-length string of binary symbols to represent an individual. If there are multiple parameters, the binary strings corresponding to each parameter are concatenated as an individual. The binary string length of each parameter is determined according to the value interval and required precision. Suppose that the value interval of parameter x_j is $[a_j, b_j]$ and the required precision is k, then the calculation formula of binary string length m_j is

$$2^{m_j-1} < \left(b_j - a_j\right) \times 10^k \leq 2^{m_j} - 1. \tag{10}$$

Set the binary string as $U_1 U_2 U_3 \ldots U_{m_j}, U_i = 0$ or 1; then the binary string is converted to the actual value of the corresponding parameter; i.e., the decoding formula is
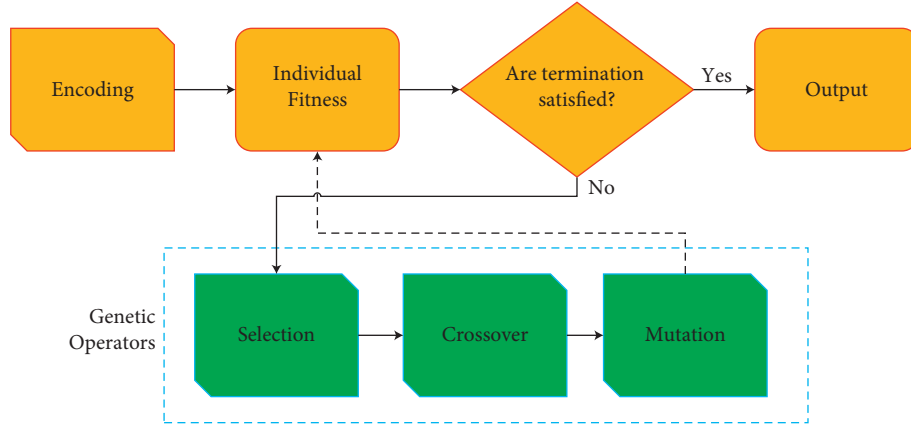
FIGURE 3: Basic flowchart of genetic algorithm.

$$x_j = a_j + \left(\sum_{i=1}^{m_j} U_i \cdot 2^{i-1} \right) \cdot \frac{b_j - a_j}{2^{m_j - 1}}. \tag{11}$$

2.3.2. Step 2. Calculate the Fitness Value.
Fitness is a function calculated to evaluate the fitness of each individual according to the optimization objective. In this paper, the optimization objective and the fitness function are given by the following equations, respectively:

$$\min_{\Omega} S(\Omega) = \frac{1}{n} \sum_{i=1}^{n} \left[C_i^H(\Omega) - C_i^M(\Omega) \right]^2, \tag{12}$$

where $\Omega = \{V_0, \kappa^*, \theta^*, \sigma, \rho\}$, $C_i^H(\Omega)$ and $C_i^M(\Omega)$ represent the Heston model price and the market price of the i th option, the market price of the option is the closing price, and n is the sample size.

$$F_j = e^{-f_j}, \tag{13}$$

where $f_j = 1/n \sum_{i=1}^{n} [C_i^H(\Omega_j) - C_i^M(\Omega_j)]^2$ and Ω_j represents the group j parameter in the population.

2.3.3. Step 3. Select Operation.
The selection is made according to the fitness value. The higher the fitness value, the greater the probability of being inherited to the next generation. Selection methods include roulette selection, random competition selection, and optimal reservation selection. In this paper, we use the roulette selection method. Each individual divides a disk into M sectors according to the probability $P_j = F_j / \sum_{k=1}^{M} F_k$, where M is the population size. Rotating the disk randomly, we select the individual at which the disk stops where the pointer falls.

2.3.4. Step 4. Crossover Operation.
Crossover operation refers to two paired chromosomes (individuals) exchange part of genes with each other in a certain way; the main methods are single point crossover, uniform crossover, multipoint crossover, and so on. Due to the quantity of 5 parameters and high precision of Heston model, we choose

three-point crossover operation. Figure 4 is an operation diagram, where the cross point is randomly generated according to the crossover probability.

2.3.5. Step 5. Mutation Operation.
Mutation refers to a change in the value of a bit or bits on an individual coding string. The main methods include basic position mutation, uniform mutation, and boundary mutation. In this paper, we adopt basic position mutation. The specific operation method is to randomly determine the gene sites to be mutated in each individual according to the mutation probability and then change the "1" in binary code into "0" or "0" into "1."

2.3.6. Step 6. Determine If Evolution is Over.
In this paper, the judgment criterion is whether the initial set of evolution algebra has been reached. If not, go back to Step (2). If so, the algorithm will be terminated.

2.4. Hybrid Modelling Framework.
In summary, the model framework of hybrid CNN based on Heston model established in this paper is shown in Figure 5. The key steps of the modelling process are as follows.

2.4.1. Step 1. Parameter Estimation of Heston Model.
Taking the mean square error between Heston prices and market prices as the optimization objective, the genetic algorithm is used to estimate the five parameters $\{V_0, \kappa^*, \theta^*, \sigma, \rho\}$ of Heston model.

2.4.2. Step 2. Extract the Deviation Sequence.
Combining the parameters obtained by genetic algorithm, the option price C^H was obtained by Heston model, and the deviation sequence y was obtained by subtracting the model price C^H from the market price C^M.

2.4.3. Step 3. Build the ResNet50 Model.
Taking y as the expected output and the option sample information as the input data, the training sets and test sets are divided
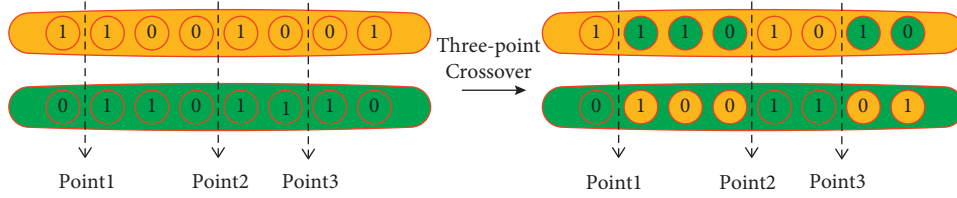
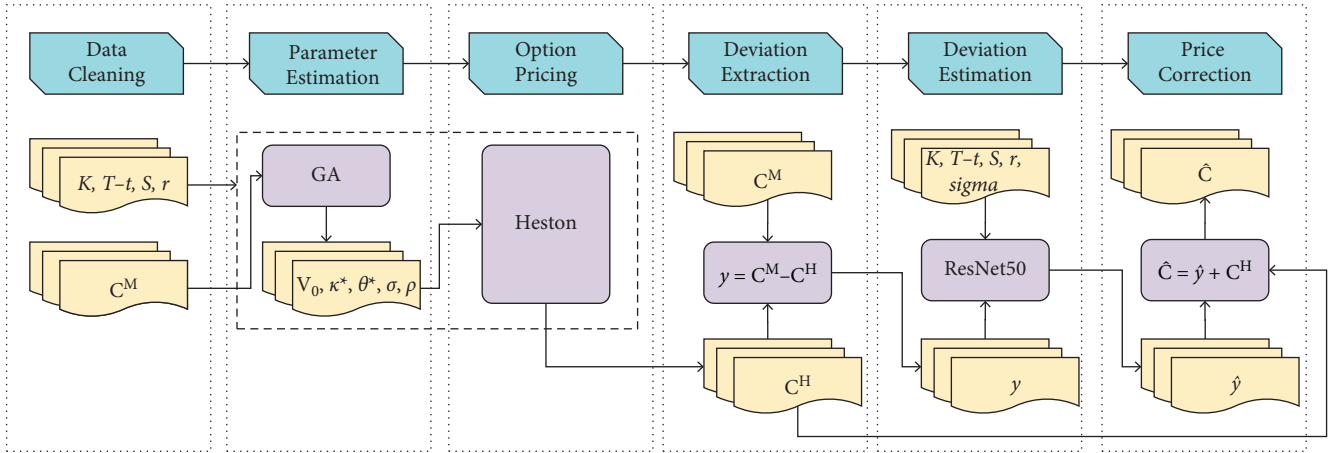FIGURE 4: Schematic diagram of three-point crossover operation.

FIGURE 5: Modelling framework diagram. Modelling in left-to-right order.

according to the obtained deviation sequence y and the corresponding option sample information, and then ResNet50 can be trained.

2.4.4. Step 4. Correct Pricing Deviations. Therefore, the trained hybrid model is used for data fitting or extrapolation analysis, and the option pricing result of Heston model after parameter optimization is integrated with the prediction result of ResNet50 deviation to complete the correction of pricing deviation, so the option pricing result of the hybrid model can be obtained as $\widehat{C}$.

3. Empirical Analysis

3.1. Data Selection and Preprocessing. This paper studies the call options of Huatai-PB CSI 300ETF, Harvest CSI 300ETF, and SSE 50ETF. The expiration months of the three options are the current month, the next month, and the following two quarterly months, and the expiration date is the fourth Wednesday of the expiration month. According to the above contract provisions and to reflect the integrity of data characteristics, the daily data from February 19 to December 31, 2021 are selected as model fitting samples and training samples in this paper. The daily data from January 4 to January 14, 2022 are selected as the prediction samples to conduct extrapolation analysis for each model. The daily data on February 19 and February 20, 2021 are selected as the parameter estimation data of Heston model. The above data are all from Wind database.

Data variables selected according to the model include strike price (K), remaining term ($T - t$, unit: year), closing price of the underlying asset (S), closing price of the option, risk-free interest rate (r), and historical volatility of the underlying asset price. In this paper, Shanghai Interbank Offered Rate (SHIBOR) is selected as the risk-free interest rate, with maturities of 3 months, 6 months, 9 months, and 1 year. The risk-free interest rate for the remaining maturities is calculated using linear interpolation of rates for adjacent maturities. The historical volatility of the underlying asset price is calculated using the volatility of the previous 30 trading days, a total of 250 trading days in a year. The specific calculation formula is $S = \sqrt{1/29 \sum_{i=1}^{29} (x_i - \overline{x})^2}$; then $\sigma = S\sqrt{250}$, where $x_i = \ln P_{i+1}/P_i$, $\overline{x} = 1/30 \sum_{i=1}^{30} x_i$, P_i is the closing price of the underlying asset in the previous trading day, and P_{i+1} is the closing price of the underlying asset in the current trading day.

Original data needs to be preprocessed, excluding nontrading day data, data of nonstandard contracts, data of option bid or offer price with null value and value less than 0.001, and data of underlying asset price with null value. The sample data table of the three ETF options is shown in Table 1.

3.2. Pricing Error Measure. In this paper, the mean squared error (MSE), mean absolute percentage error (MAPE), and mean absolute error (MAE) are selected to evaluate the model performance by referring to the comparative measures of Zhang et al. [32] and Yang et al. [33].

TABLE 1: Options sample data size.

Option	Parameter estimation data size	Fitting data size	Prediction data size	Total size
510300 ETF	62	8752	372	9186
159919 ETF	62	5887	373	6322
510050 ETF	82	7023	368	7473

$$\text{MSE} = \frac{1}{n} \sum \left(C^M - \widehat{C} \right)^2,$$

$$\text{MAPE} = \frac{1}{n} \sum \left| \frac{C^M - \widehat{C}}{C^M} \right|, \qquad (14)$$

$$\text{MAE} = \frac{1}{n} \sum \left| C^M - \widehat{C} \right|,$$

where C^M is the market price of the option and $\widehat{C}$ is the estimated price of the option model.

3.3. Parameter Setting and ResNet50 Model Adjustment

3.3.1. Parameter Estimation of Heston Model. Considering the precision and efficiency of parameter estimation, the parameters of genetic algorithm are set as population size 20, parameter precision 6 (retaining six decimal places), selection probability $pc = 0.6$, mutation probability $pm = 0.1$, evolutionary generation 500, and parameter value ranges $\{\kappa^*: [0, 2], \theta^*: [0, 20], \sigma: [0, 1], V_0: [0, 5], \rho: [-1, 1]\}$. An initial population is randomly generated and converges to an optimal individual after 500 iterations. The parameter results of Heston model are shown in Table 2, which will also be used in hybrid modelling.

3.3.2. ResNet50 Model Adjustment. When different types of options are applied to ResNet50, the input data in line with the actual situation should be constructed, and the ZeroPadding layer and pooling layer should be adjusted accordingly. This paper empirically selects the call ETF options in China and makes the following adjustments to ResNet50 model. The adjusted structure is shown in Figure 6.

(1) Adjustment of input size: the input sample point consists of 9 options with the same expiration month at time t: 1 at-the-money options, 4 in-the-money options, and 4 out-of-the-money options stipulated in the contract. Each option includes the data of five variables: strike price K, remaining maturity time $T - t$, underlying asset price S at time t, risk-free interest rate r, and historical volatility of the underlying asset. Therefore, the size of the input data is $9 * 5 * 1$.

(2) Add a ZeroPadding layer before the initial convolutional layer to make the data size $9 * 9 * 1$.

(3) Due to the small data size, the pooling layer of the original model is removed.

(4) The output data is the expected pricing deviation of the Heston model optimized by the genetic algorithm for 9 options; i.e., the output is a 9-dimensional vector.

(5) The pricing deviation is positive or negative, so the activation function of the whole connection layer is adjusted from softmax to tanh.

(6) ResNet50 model is used for regression problems, so the loss function is adjusted to *mse*.

3.3.3. Hyperparameter Setting of Neural Networks. In this paper, we use BP neural network model to make a comparison to the performance of ResNet50 model in correcting the option price deviation of Heston model. BP model consists of three layers, and the number of neurons in the hidden layer is calculated according to formula $\sqrt{m + n} + a$, where m and n are the number of neurons in the input layer and output layer, respectively, and a is an integer between $[0, 10]$. In the empirical trial by trial, the error and training time are compared, and finally the hidden layer of 6 layers is determined to be the optimal number of layers. In order to improve the accuracy of the model and ensure the speed of training convergence, the batch size is 64, the sample data training times (Epoch) is 500, and the ratio of training set to test set is $4 : 1$. ResNet50 model is different from BP model in the selection of learning rate. Learning rate is tested and selected from 0.01, 0.001, and 0.0001. The results show that the respective loss functions of ResNet50 model and BP model change at reasonable rates for learning rates of 0.0001 and 0.01, respectively. Figure 7 shows the change in loss for ResNet50 model training option deviation data. The horizontal coordinate is Epoch value and the vertical coordinate is loss/val_loss value. In addition, due to the large number of ResNet50 model parameters, in order to increase the amount of training data, when constructing input samples as described previously, the row data of each sample is rearranged into 9 different sample points, and the output results were averaged. The rearrangement is shown in Figure 8, where each number represents a row of option data.

3.4. Model Comparison

In order to test the accuracy of Heston-GA-ResNet50 pricing model, BSM model, ResNet50 model, Heston-CIR model, nonaffine Heston model, Heston-GA model, and Heston-GA-BP hybrid model are established at the same time. Among them, Heston-CIR model and nonaffine Heston model, whose parameters are estimated by genetic algorithm, are computed by Monte Carlo. The results of the parameter estimation are shown in Tables 5 and 6 in the Appendix. Therefore, the fitting and prediction effects of each model have been analysed and compared as follows.

TABLE 2: Heston parameters of ETF options.

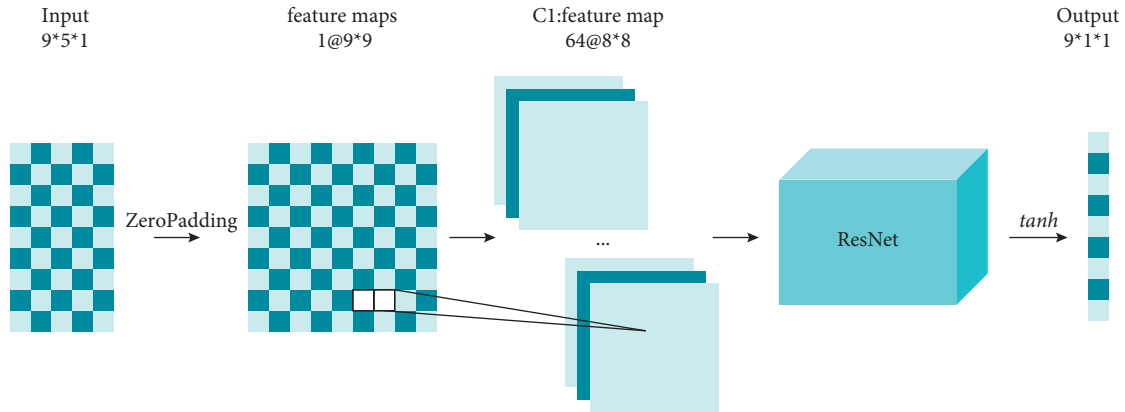| Option	Parameters				
	κ^*	θ^*	σ	V_0	ρ
510300 ETF	14.495668	0.016163	0.404808	0.020398	0.908735
159919 ETF	0.0960471	0.0853677	0.090974	0.015918	−0.7
510050 ETF	0.030086	4.962398	0.698335	0.03	0

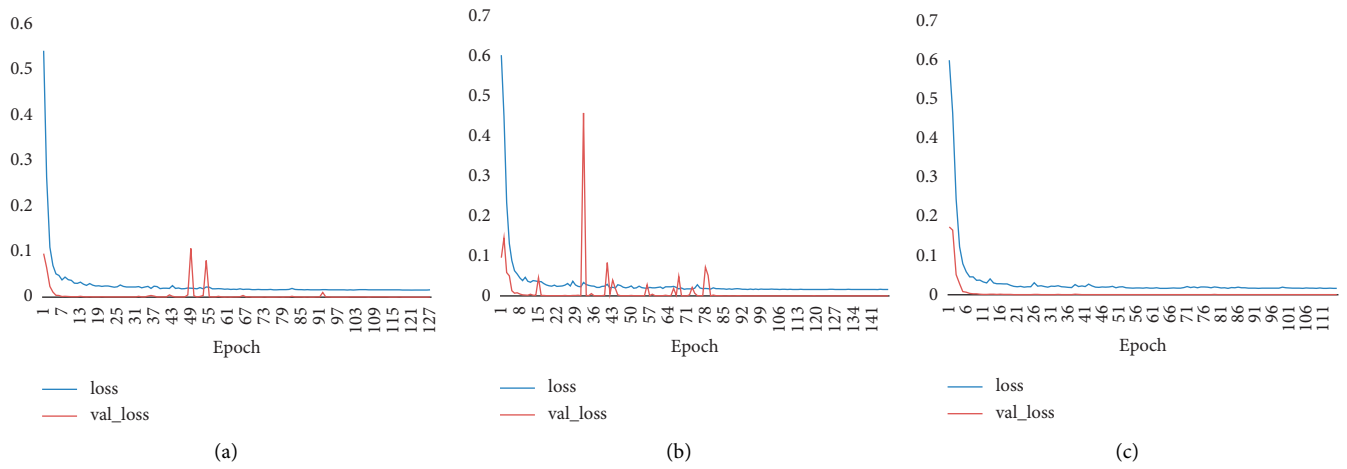FIGURE 6: Adjusted ResNet structure diagram.

FIGURE 7: Changes in loss of training set of Heston-GA-ResNet50 hybrid model. (a) 510300 ETF, (b) 159919 ETF, and (c) 510500 ETF.
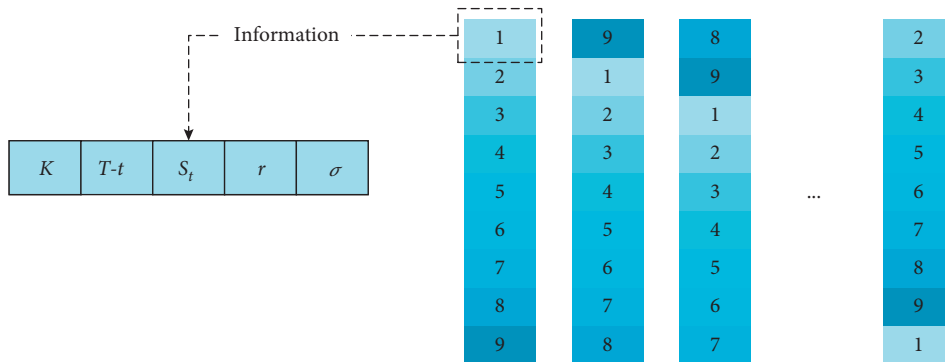

FIGURE 8: Rearrangement of sample points.

Table 3: Comparison of fitting errors of different models for different ETF options.

Model	Error measure		
	MSE (E)	MAPE	MAE
(a) 510300			
BSM	2.3158−3	0.3747	0.0308
ResNet50	4.8249−3	1.2926	0.0337
Heston-CIR	1.5582−3	0.3007	0.0278
Non-affine Heston	1.8931−3	0.3264	0.0309
Heston-GA	7.2447−4	0.2831	0.0193
Heston-GA-BP	4.2495−4	0.2788	0.0151
Heston-GA-ResNet50	4.3194−4	0.2724	0.0152
(b) 159919			
BSM	9.5349−4	0.3143	0.0210
ResNet50	9.6927−3	0.6950	0.0353
Heston-CIR	5.0487−4	0.2208	0.0155
Non-affine Heston	6.8583−4	0.2905	0.0194
Heston-GA	6.8167−4	0.3614	0.0191
Heston-GA-BP	4.7389−4	0.4480	0.0157
Heston-GA-ResNet50	3.2724−4	0.2303	0.0126
(c) 510050			
BSM	5.4316−4	0.3316	0.0151
ResNet50	4.4120−3	0.3065	0.0259
Heston-CIR	4.6814−4	0.3791	0.0158
Non-affine Heston	3.9161−4	0.2740	0.0135
Heston-GA	4.6465−4	0.2303	0.0132
Heston-GA-BP	4.0417−4	0.4231	0.0135
Heston-GA-ResNet50	3.1185−4	0.2387	0.0112

Table 4: Comparison of prediction errors of different models for different ETF options.

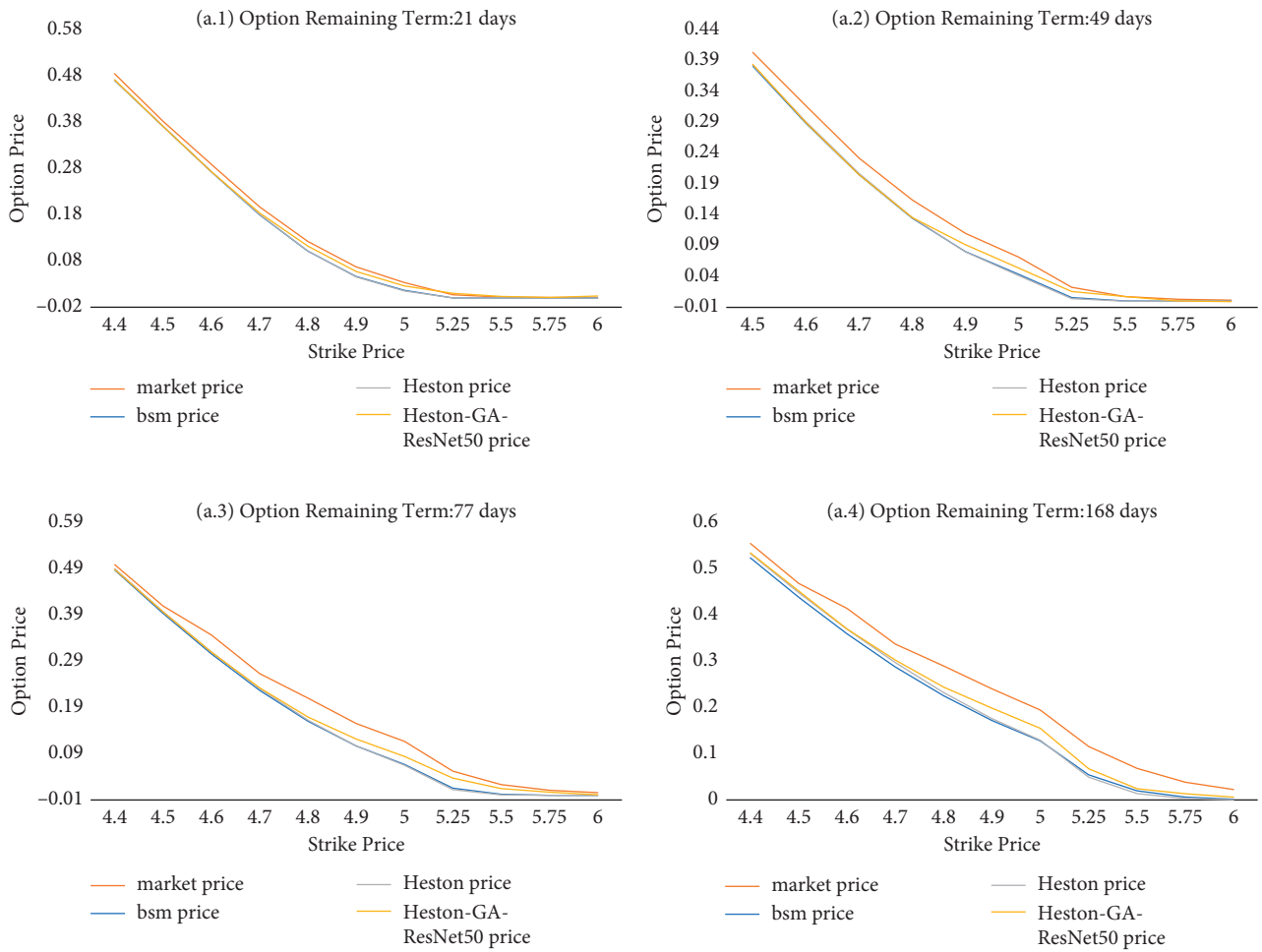| Model	Error measure		
	MSE (E)	MAPE	MAE
(a) 510300			
BSM	7.2966−4	0.4240	0.0219
ResNet50	5.1112−3	0.7137	0.0487
Heston-CIR	4.3061−4	0.2601	0.0155
Non-affine Heston	6.9010−4	0.3135	0.0200
Heston-GA	7.7971−4	0.3581	0.0222
Heston-GA-BP	4.8228−4	0.3478	0.0167
Heston-GA-ResNet50	6.2385−4	0.3110	0.0194
(b) 159919			
BSM	8.5637−4	0.4131	0.0240
ResNet50	5.5807−3	0.9124	0.0525
Heston-CIR	1.5082−4	0.2069	0.0089
Non-affine Heston	4.0113−4	0.2708	0.0160
Heston-GA	9.3303−4	0.4452	0.0251
Heston-GA-BP	5.8388−4	0.5907	0.0197
Heston-GA-ResNet50	5.5141−4	0.3050	0.0187
(c) 510050			
BSM	7.9299−4	0.4409	0.0231
ResNet50	1.8546−3	1.1366	0.0297
Heston-CIR	3.0460−4	0.3703	0.0143
Non-affine Heston	1.8652−4	0.2306	0.0106
Heston-GA	1.9828−4	0.1892	0.0095
Heston-GA-BP	2.0192−4	0.3287	0.0109
Heston-GA-ResNet50	1.5690−4	0.2407	0.0087
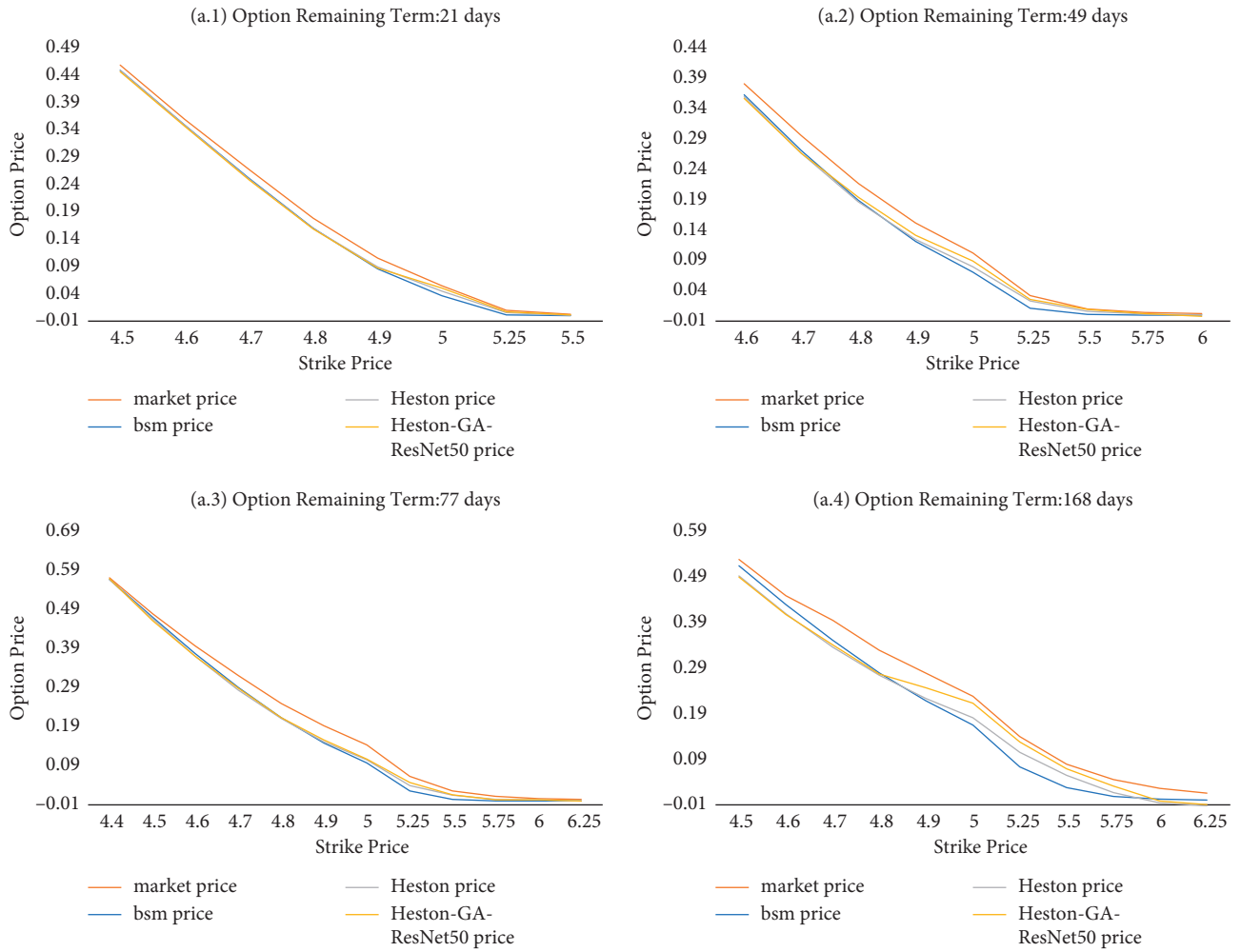
(a)

Figure 9: Continued.
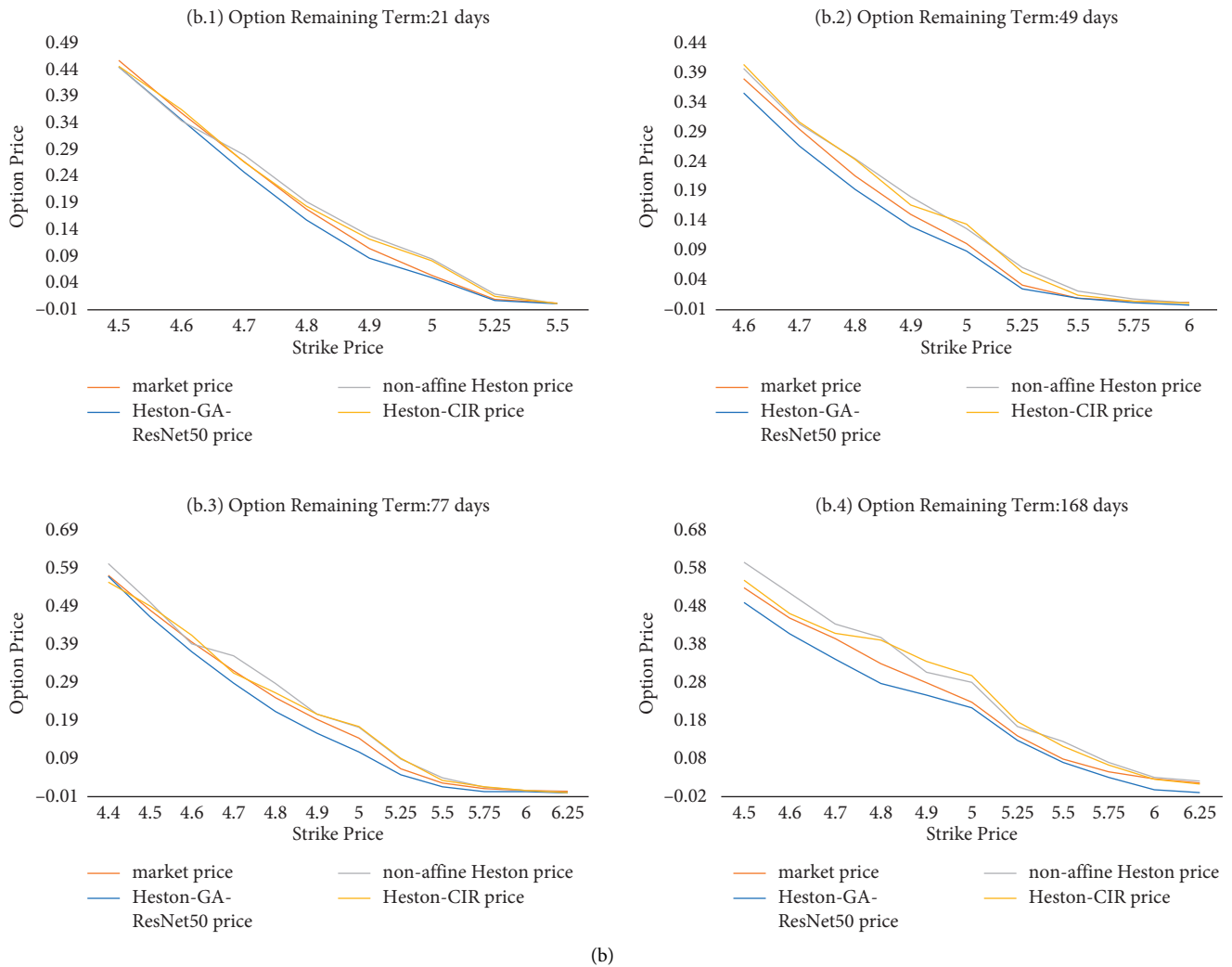
FIGURE 9: Option price prediction of different models with different remaining terms (510300). (a) Pricing comparison of Heston-GA-ResNet50 model with BSM model and Heston model for 510300 ETF options. (b) Pricing comparison of Heston-GA-ResNet50 model with nonaffine model and Heston-CIR model for 510300 ETF options.

Table 3 lists the fitting results of six different models applied to data sets of three listed ETF options in China. MSE, MAPE, and MAE of Heston-GA model are basically lower than those of BSM model, indicating that Heston model will be closer to the actual market situation after relaxing the BSM model's assumption that volatility is constant. In the comparison of three option fits, it can be concluded that the ResNet50 model does not outperform the other four parametric models, i.e., BSM, Heston, Heston-CIR, and nonaffine Heston. This supports that the ResNet50 model indeed lacks an explanation for the factors affecting option prices, leading to a large deviation in option pricing. The above conclusions confirm the necessity of hybrid modelling. Table 3 demonstrates the basic reduction in three measures for two hybrid models over five single models, with the ResNet50 hybrid model offering a modelling advantage over the BP hybrid model.

To validate the marketability of the model developed in this paper and to show that the trained ResNet50 model does not suffer from overfitting problems, an extrapolation analysis is performed on the model. We perform model pricing error analysis with data not involved in NNs training, i.e., out-of-sample predictive analysis. Table 4 illustrates the prediction errors of each model. The relative prediction error of each model is basically similar to the fitting error in Table 3, where three measures of Heston-GA-ResNet50 model of three ETF options all decrease. In particular, it can be argued from the empirical evidence on 510300 ETF options that if the hybrid model performs poorly in the fitting experiment, it will perform similarly in the prediction. Therefore, if the NNs models are to be applied to the real market, the training of the model determines the quality of the model.

(a.1) Option Remaining Term:21 days

(a.2) Option Remaining Term:49 days

(a.3) Option Remaining Term:77 days

(a.4) Option Remaining Term:168 days
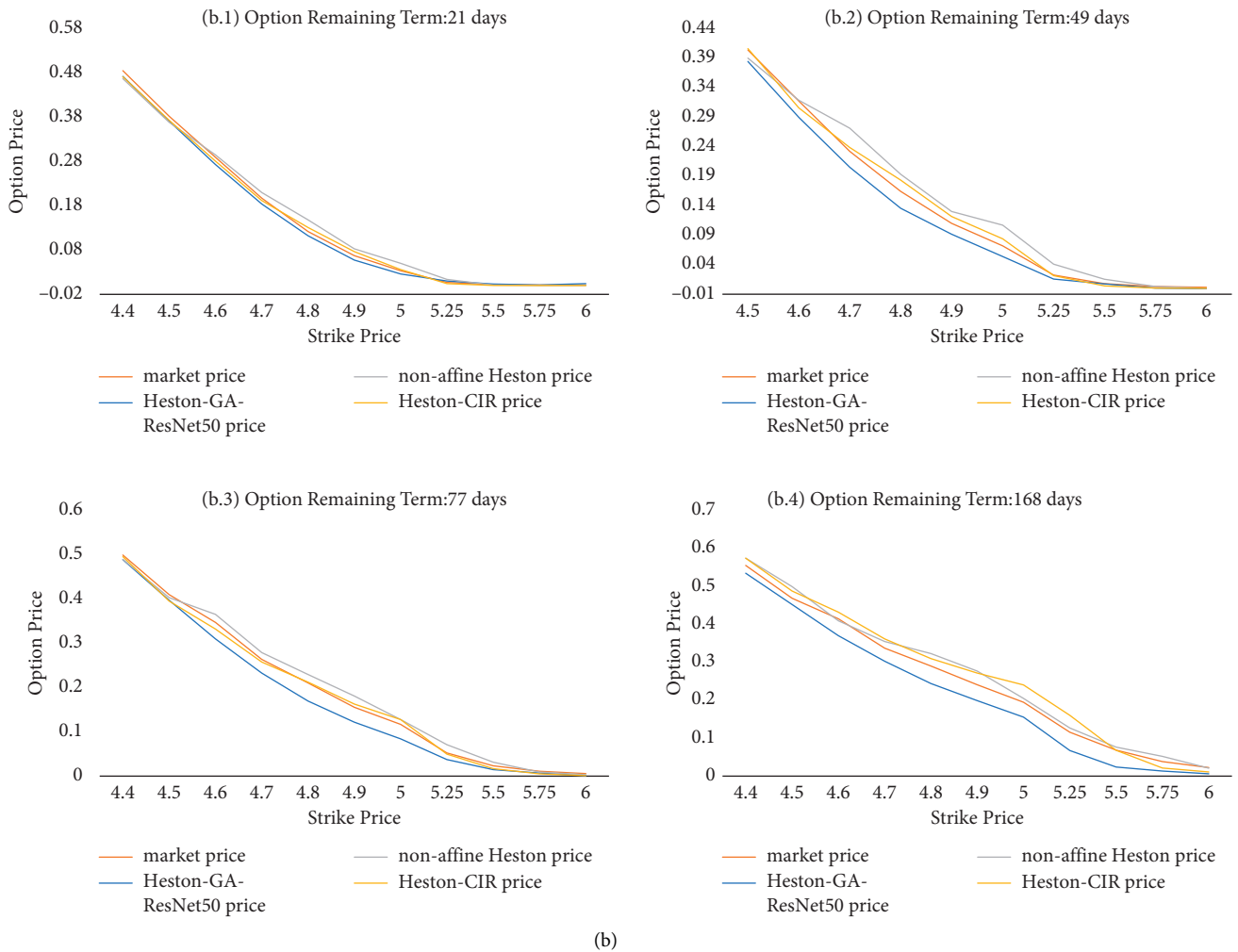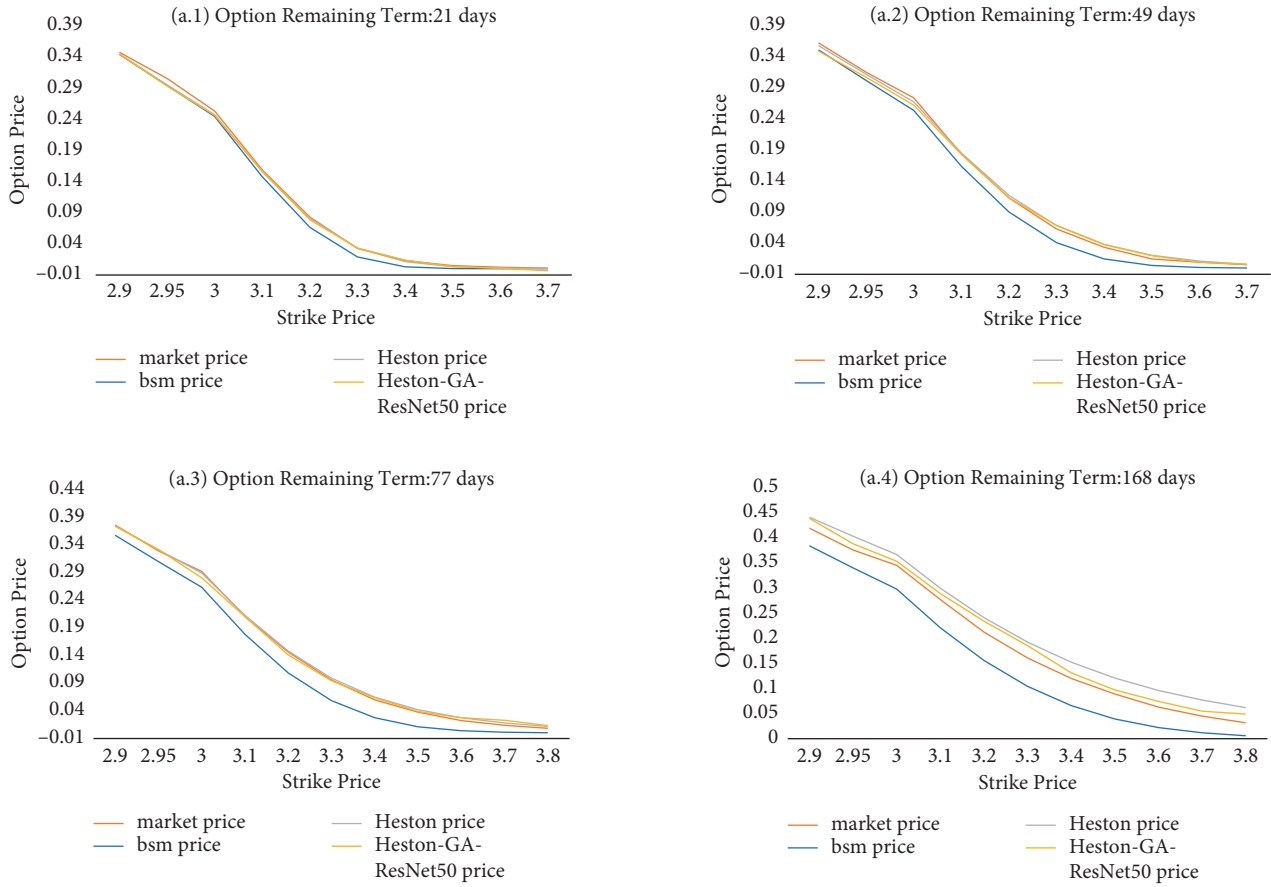
(a)

Figure 10: Continued.

FIGURE 10: Option price prediction of different models with different remaining terms (159919). (a) Pricing comparison of Heston-GA-ResNet50 model with BSM model and Heston model for 159919 ETF options. (b) Pricing comparison of Heston-GA-ResNet50 model with nonaffine model and Heston-CIR model for 159919 ETF options.

In general, an important observation can be derived from Tables 3 and 4, i.e., on three options data, MSE, MAPE, and MAE of Heston-GA-ResNet50, which are commonly smaller than those of the other six comparison models. This implies that correcting Heston price bias with ResNet50 gives more accurate results. As for the Heston optimization models, i.e., Heston-CIR and nonaffine Heston in this paper, they do perform relatively not badly for 510300 and 159919 ETF options in Table 4 empirically. But the combined information in Tables 3 and 4 demonstrates that their pricing effects are uncontrollable, which are limited by the market factors and the accuracy of the parameter estimates. This validates the previous statement that parametric option pricing models lack the ability of capturing nonlinear factors, i.e., robustness, to a certain extent compared to nonparametric models. To sum up, Heston-GA-ResNet50 model

has the highest accuracy and robustness and better pricing performance, which has certain application significance in the option market. This point will also be demonstrated in the following empirical analysis.

To better illustrate the applicability of the Heston-GA-ResNet50 model, the option contracts for January 5, 2022, including the four maturity months of January, February, March, and June, are chosen as examples. Figures 9–11 depict the pricing for different model for different remaining terms and the comparison with real market prices. For clarity of comparison, we divide models into two groups a and b. Group a includes BSM model, Heston model, and Heston-GA-ResNet50 model, and group b includes Heston-CIR model, nonaffine Heston model, and Heston-GA-ResNet50 model. In Figures 9–11, the pricing curves of the model in this paper
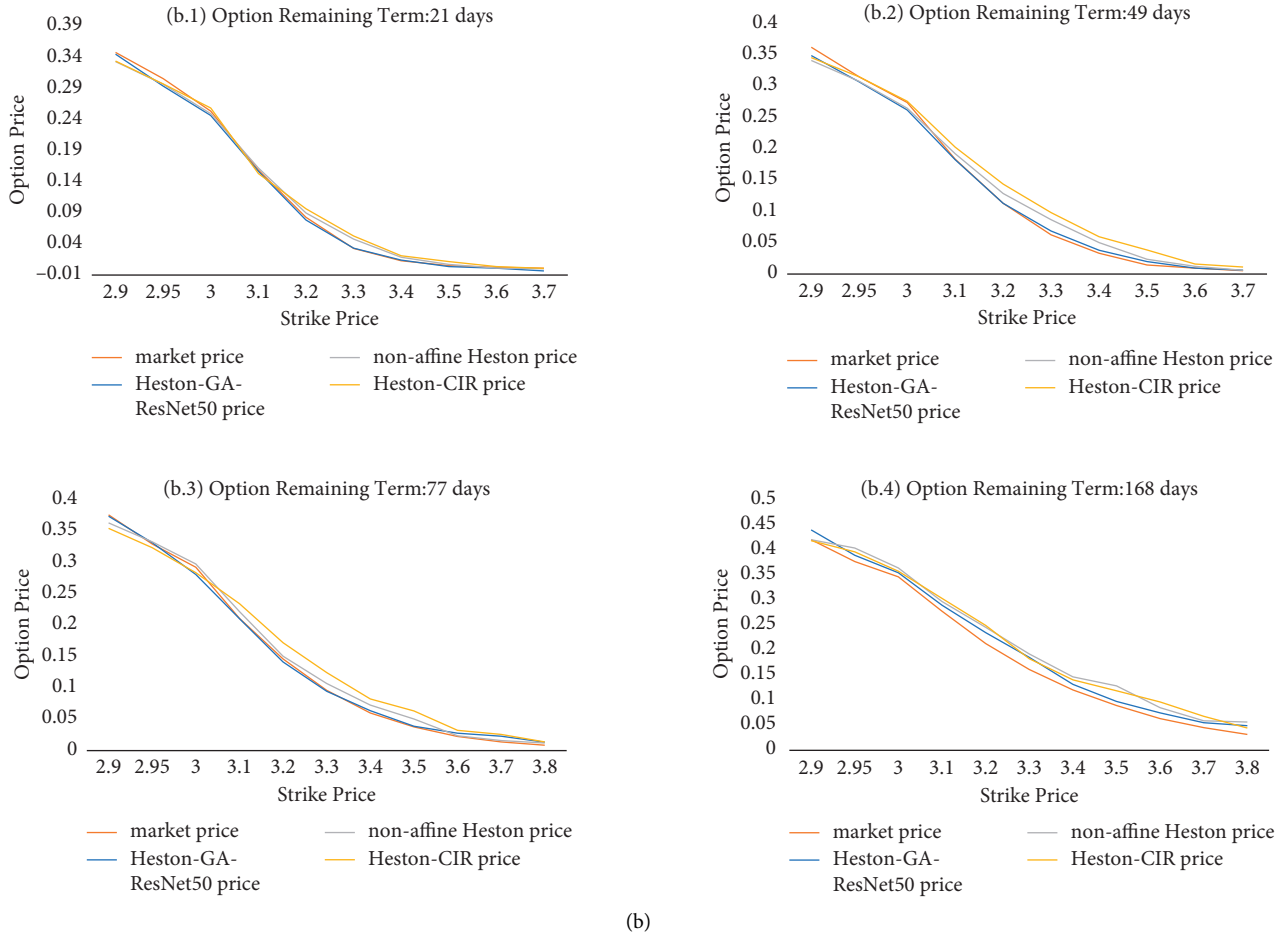
(a)

Figure 11: Continued.

(b)

Figure 11: Option price prediction of different models with different remaining terms (510050). (a) Pricing comparison of Heston-GA-ResNet50 model with BSM model and Heston model for 510050 ETF options. (b) Pricing comparison of Heston-GA-ResNet50 model with non–affine model and Heston-CIR model for 510050 ETF options.

Table 5: The parameters of Heston-CIR model.

Option	Parameters							
	κ^*	θ^*	σ	V_0	ρ	α	β	η
510300 ETF	0.29599	0.03604	0.04210	0.04700	−0.4	0.14506	0.02487	0.04506
159919 ETF	0.06208	0.03646	0.06224	0.03644	−0.6	0.05343	0.04537	0.06016
510050 ETF	0.05441	0.04355	0.0436	0.06456	−0.5	0.05432	0.04361	0.03472

Table 6: The parameters of nonaffine Heston model.

Option	Parameters					
	κ^*	θ^*	σ	V_0	σ	γ
510300 ETF	0.050145	0.051107	0.039328	0.051053	−0.6354	0.6675
159919 ETF	0.058008	0.058008	0.061538	0.057898	−0.7	0.478984
510050 ETF	0.049317	0.772893	0.045445	0.049504	−0.4937	0.755048

are essentially closer to the actual market price curves for both short and long remaining terms of the options, for which the conclusion further confirms the superiority of Heston-GA-ResNet50 model consistently with the previous empirical findings.

4. Conclusions

In this paper, we propose a new hybrid neural network option pricing model Heston-GA-ResNet50 combining traditional Heston model with ResNet50 model, where the

parameters of Heston model are optimized through genetic algorithm. In particular, we propose a new way of constructing data facets by combining data from several option contracts according to some criteria. This image-like data is suitable as input data for CNN, as it is in this paper for the adjusted ResNet50 model. Our approach provides a reference for the application of CNN models or existing classical CNN models to options pricing, which has profound theoretical implications. Then, the empirical study on three kinds of listed ETF call options shows that our hybrid model is superior to other models in both terms of fitting and prediction. For the comparison, we not only use BSM model and Heston model, but also add BP hybrid model and two Heston optimization models, namely, Heston-CIR model and nonaffine Heston model. The more outstanding performance of our model compared to these excellent models confirms the validity of the modelling ideas in our research. Further analysis of the applicable scope of model pricing shows that Heston-GA-ResNet50 model has a high accuracy for both short and long maturities. And, compared with other models, Heston-GA-ResNet50 model has a stronger applicability for long maturities. Therefore, the modelling process of our hybrid model is effective, in which the parametric pricing model ensures the interpretability of the hybrid model and the nonparametric model enhances the nonlinear fitting ability of the hybrid model, making our Heston-GA-ResNet50 model more accurate and robust. In summary, the model proposed in this paper is of both theoretical and practical significance, which will provide a reference for academics and investors.

In the future research, we will extend our model to the pricing application of other kinds of options. Besides, we will also concern how to establish a hybrid model with low complexity, fast training speed, high accuracy, and strong applicability in future improvement.

Appendix

In the risk-neutral setting, Heston-CIR and nonaffine Heston have the following mathematical expressions, respectively.

Heston-CIR model [9] is

$$dS_t = r_t S_t dt + S_t \sqrt{v_t} dW_{1,t},$$

$$dv_t = \kappa(\theta - v_t)dt - \sigma\sqrt{v_t}\left(\rho dW_{1,t} + \sqrt{1-\rho^2}\, dW_{2,t}\right), \quad (A.1)$$

$$dr_t = \left\{\alpha\beta - \left[\alpha + B\eta^2\right]r\right\}dt + \eta\sqrt{r_t}dW_{3,t},$$

where $B = 2(e^{m(T-t)} - 1)/2m + (\alpha + m)\ (e^{m(T-t)} - 1)$ $m = \sqrt{\alpha^2 + 2\eta^2}$。

Nonaffine Heston model [37] is

$$dS_t = r S_t dt + \sqrt{v_t} S_t dW_{1,t},$$

$$dv_t = \kappa(\theta - v_t)dt + \sigma v_t^{\gamma/2}\left(\rho dW_{1,t} + \sqrt{1-\rho^2}\, dW_{2,t}\right), \quad (A.2)$$

where γ is a nonaffine parameter.

In this paper, the parameters of above two models are estimated by GA. The results are shown in Tables 5 and 6.

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] F. Black and M. Scholes, "The pricing of options and corporate liabilities," *Journal of Political Economy*, vol. 81, no. 3, pp. 637–654, 1973.

[2] R. C. Merton, "Theory of rational option pricing," *Bell Journal of Economics and Management Science*, vol. 4, no. 1, pp. 141–183, 1973.

[3] E. M. Stein and J. C. Stein, "Stock price distributions with stochastic volatility: an analytic approach," *Review of Financial Studies*, vol. 4, no. 4, pp. 727–752, 1991.

[4] S. L. Heston, "A closed-form solution for options with stochastic volatility with applications to bond and currency options," *Review of Financial Studies*, vol. 6, no. 2, pp. 327–343, 1993.

[5] X. He and S. Lin, "An analytical approximation formula for barrier option prices under the Heston model," *Computational Economics*, pp. 1–13, 2021.

[6] M. K. Lee, S. J. Yang, and J. H. Kim, "A closed form solution for vulnerable options with Heston's stochastic volatility," *Chaos, Solitons & Fractals*, vol. 86, pp. 23–27, 2016.

[7] X. Lu, S.-P. Zhu, and D. Yan, "Nonlinear PDE model for European options with transaction costs under Heston stochastic volatility," *Communications in Nonlinear Science and Numerical Simulation*, vol. 103, Article ID 105986, 2021.

[8] L. A. Grzelak and C. W. Oosterlee, "On the Heston model with stochastic interest rates," *SIAM Journal on Financial Mathematics*, vol. 2, no. 1, pp. 255–286, 2011.

[9] X. He and S. P. Zhu, "A closed-form pricing formula for European options under the Heston model with stochastic interest rate," *Journal of Computational and Applied Mathematics*, vol. 335, pp. 323–333, 2018.

[10] X. He and W. Chen, "A closed-form pricing formula for European options under a new stochastic volatility model with a stochastic long-term mean," *Mathematics and Financial Economics*, vol. 15, no. 2, pp. 381–396, 2021.

[11] X. He and W. Chen, "Pricing foreign exchange options under a hybrid Heston-Cox- Ingersoll-Ross model with regime switching," *IMA Journal of Management Mathematics*, vol. 33, no. 2, pp. 255–272, 2022.

[12] Z. Guo, "Option pricing under the Heston model where the interest rate follows the vasicek model," *Communications in Statistics - Theory and Methods*, vol. 50, no. 12, pp. 2930–2937, 2019.

[13] X. He and S. Lin, "A Fractional Black-Scholes Model with Stochastic Volatility and European Option Pricing," *Expert Systems with Applications*, vol. 178, Article ID 114983, 2021.

[14] F. Mehrdoust, A. R. Najafi, S. Fallah, and O. Samimi, "Mixed fractional Heston model and the pricing of American options," *Journal of Computational and Applied Mathematics*, vol. 330, pp. 141–154, 2018.

[15] Y. Jiao, C. Ma, S. Scotti, and C. Zhou, "The Alpha-Heston Stochastic Volatility Model," *Mathematical Finance*, vol. 31, no. 3, pp. 1–36, 2021.

[16] Y. Yang, S. Liu, Y. Wu, and B. Wiwatanapataphee, "Pricing of volatility derivatives in a heston–CIR model with markov-modulated jump diffusion," *Journal of Computational and Applied Mathematics*, vol. 393, pp. 1–25, 2021.

[17] M. Grasselli, "THE 4/2 stochastic volatility model: a unified approach for the Heston and the 3/2 model," *Mathematical Finance*, vol. 27, no. 4, pp. 1013–1034, 2017.

[18] S. Zhu and B. Wang, "Unified approach for the affine and non-affine models: an empirical analysis on the S&P 500 volatility dynamics," *Computational Economics*, vol. 53, no. 4, pp. 1421–1442, 2019.

[19] O. El Euch and M. Rosenbaum, "The characteristic function of rough Heston models," *Mathematical Finance*, vol. 29, no. 1, pp. 3–38, 2019.

[20] X. He and S. Lin, "A semi-analytical pricing formula for European options under the rough heston-CIR model," *ANZIAM Journal*, vol. 61, no. 4, pp. 431–445, 2020.

[21] M. Malliaris and L. Salchenberger, "A neural network model for estimating option prices," *Applied Intelligence*, vol. 3, no. 3, pp. 193–206, 1993.

[22] S. K. Mitra, "An option pricing model that combines neural network approach and black Scholes formula," *Global Journal of Computer Science and Technology*, vol. 12, no. 4, pp. 6–16, 2012.

[23] Y. Jerbi and S. Chaabene, "European call price modelling using neural networks in considering volatility as stochastic with comparison to the Heston model," *Journal of Statistical Computation and Simulation*, vol. 90, no. 10, pp. 1793–1810, 2020.

[24] C. F. Ivaşcu, "Option Pricing Using Machine Learning," *Expert Systems with Applications*, vol. 163, Article ID 113799, 2021.

[25] Z. Iltuzer, "Option pricing with neural networks vs. Black-Scholes under different volatility forecasting approaches for BIST 30 index options," *Borsa Istanbul Review*, vol. 22, no. 4, pp. 1–18, 2021.

[26] J. Yao, Y. Li, and C. L. Tan, "Option price forecasting using neural networks," *Omega*, vol. 28, no. 4, pp. 455–466, 2000.

[27] L. Xu, M. Dixon, B. Eales, F. Cai, B. Read, and J. Healy, "Barrier option pricing: modelling with neural nets," *Physica A: Statistical Mechanics and Its Applications*, vol. 344, no. 1-2, pp. 289–293, 2004.

[28] H. Park, N. Kim, and J. Lee, "Parametric Models and Non-parametric Machine Learning Models Forpredicting Option Prices: Empirical Comparison Study over KOSPI 200 index Options," *Expert Systems with Applications*, vol. 44, no. 11, 2014.

[29] M. Wysocki and R. Ślepaczuk, "Artificial neural networks performance in WIG20 index options pricing," *Entropy*, vol. 24, no. 1, pp. 35–53, 2021.

[30] E. Arin and A. M. Ozbayoglu, "Deep learning based hybrid computational intelligence models for options pricing," *Computational Economics*, vol. 59, pp. 39–58, 2020.

[31] H. Funahashi, "Artificial neural network for option pricing with and without asymptotic correction," *Quantitative Finance*, vol. 21, no. 4, pp. 575–592, 2021.

[32] L. Zhang and W. Zhang, "Option pricing model by applying hybrid neural network based on Heston model and genetic algorithm (in Chinese)," *Journal of Industrial Engineering/Engineering Management*, vol. 32, no. 3, pp. 142–149, 2018.

[33] C. Yang, Z. Shao, C. Liu, and F. Chao, "A hybrid modeling framework and its application for exchange traded fund options pricing (in Chinese)," *Chinese Journal of Management Science*, vol. 28, pp. 44–53, 2020.

[34] K. Fukushima, "Neocognitron: a self-organizing neural network model for A mechanism of pattern recognition unaffected by shift in position," *Biological Cybernetics*, vol. 36, no. 4, pp. 193–202, 1980.

[35] K. M. He, X. Zhang, and S. Ren, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, NV, USA, 2016.

[36] B. Li and W. L. He, "A new method of finding the parameters of heston's option pricing model (in Chinese)," *The Journal of Quantitative and Technical Economics*, vol. 32, pp. 129–146, 2015.

[37] Y. F. Sun, B. Y. Wu, T. Guo, and L. Caiyan, "Fourier-Cosine Based Option Pricing for SSE 50 ETF under non-affine stochastic volatility model (in Chinese)," *Systems Engineering —— Theory & Practice*, vol. 40, no. 4, pp. pp888–904, 2020.