Research on Pricing Methods of Convertible Bonds Based on Deep Learning GAN Models

Gui Ren * and Tao Meng

Department of Finance, School of Economics, Shanghai University, Shanghai 200444, China; 18916935072@163.com
* Correspondence: rengui@shu.edu.cn

Abstract: This paper proposes two data-driven models (including LSTM pricing model, WGAN pricing model) and an improved model of LSM based on GAN to analyze the pricing of convertible bonds. In addition, the LSM model with higher precision in traditional pricing model is selected for comparative study with other pricing models. It is found that the traditional LSM pricing model has a large error in the first-day pricing, and the pricing function needs to be further improved. Among the four pricing models, LSTM pricing model and WGAN pricing model have the best pricing effect. The WGAN pricing model is better than the LSTM pricing model (0.21%), and the LSM improved model (1.17%) is better than the traditional LSM model (2.26%). Applying the generative deep learning model GAN to the pricing of convertible bonds can circumvent the harsh preconditions of assumptions, and significantly improve the pricing effect of the traditional model. The scope of application of each model is different. Therefore, this paper proves the feasibility of the GAN model applied to the pricing of convertible bonds, and enriches the pricing function of derivatives in the financial field.

Keywords: convertible bond pricing; LSTM; GAN; least square Monte Carlo Simulation

1. Introduction

As a kind of securities with both bond and equity attributes, convertible bond has attracted the attention of the industry and academia since its birth in 1993. Comparing with A-shares, the convertible bond has developed more slowly and is still in an emerging market. The deep learning technology has the technology of mining the nonlinear laws in the data and the strict assumptions in the traditional pricing of convertible bonds. Some scholars began to carry out pricing of convertible bonds from another new perspective, that is, to introduce the relevant algorithm of deep learning to the valuation and pricing of convertible bonds, so as to reduce the pricing error. Though, the traditional convertible bond pricing theory can be roughly divided into four categories: B–S (Black–Scholes) option pricing model, tree graph pricing model, finite difference method and least square Monte Carlo pricing model (LSM).

However, due to the differences between the terms of domestic and foreign convertible bonds and the limitations of traditional models, it is found empirically that directly applying these methods to the pricing of domestic convertible bonds is not a good way to price convertible bonds. Considering some defects of the above models, this paper adopts a data-driven method (including the LSTM model and GAN model) to price convertible bonds, so as to solve the pricing imprecision problem of traditional methods. This method can discard some strict restrictions in traditional models and include more influential factors. According to the characteristics of different industries, reasonable influence factors can be selected for scientific pricing.

This paper also uses the method mentioned in Wiese et al. (2020) and Tan et al. (2022) for reference, and uses the GAN model instead of Monte Carlo to generate simulation
data. It is put into the traditional LSM model to form the LSM improvement model, so as to overcome the defects caused by the assumptions of LSM, so as to improve the pricing power of the traditional model and promote the traditional pricing model to keep pace with the times. Therefore, the focus of this paper is to put forward two kinds of pricing methods (including the data-driven pricing method and LSM improved model) to solve the problem of low precision and difficult pricing of convertible bonds.

Theoretically, the study of this paper enriches the existing pricing theory, and discusses the shortcomings, what needs to be improved, and applies the new method to the traditional pricing model, so that it can overcome the pre-defect, and then better complete its pricing task. In practice, a new financial data generation method is adopted, and it is applied to the traditional least square pricing model to price the convertible bond, in order to enrich the pricing function in the financial field and provide empirical evidence for strategic investment of financial assets. The research framework is shown in the following (Figure 1).
2. Literature Review

This paper summarizes the traditional pricing methods of convertible bonds according to the first division way. Among them, the analytical method is mainly the Black–Scholes option pricing method, and the numerical method includes the tree graph method, finite difference method and least square Monte Carlo simulation method. Secondly, because the machine learning algorithm can skip the overly strict assumptions in the traditional pricing methods, many scholars at home and abroad began to introduce machine learning methods to study the pricing of convertible bonds.

2.1. Pricing Method of Convertible Bonds Based on B–S Option Pricing Model

The application of the B–S option pricing method to the pricing of convertible bonds can be dated back to the studies of Brennan and Schwartz (1977) and Ingersoll (1977), who believed that convertible bonds were affected by three factors: interest payment, cash dividend and redemption terms. Since the volatility of the stock price of convertible bonds is an important factor in the B–S model, foreign scholars have made many improvements around it. Stochastic volatility (Hull and White (1988)), mean recovery Gaussian motion described volatility (Kalotay et al. (1993)) and GARCH option pricing theory (Duan 1995) were applied to the pricing of convertible bonds, and more accurate pricing results were obtained. In recent years, many new statistical methods have been applied to the valuation of convertible bonds. In 2018, the concept of index variance gamma model was applied to the valuation of convertible bonds (CB), which is a new attempt by Yang et al. (2018). This differs from the standard Black–Scholes approach to valuing derivatives by using the VG process to describe the dynamic underlying asset logarithm price.

Domestic scholars have done a lot of improvement research on the basis of the mature theoretical achievements abroad. Zheng and Lin (2004) proved that the B–S model could be applied to the pricing of Chinese convertible bonds by adding some special clauses (downward revision clauses of convertible bonds). Xie (2021) took the Opai convertible bond as an example and conducted a pricing study by using the B–S pricing model that considered sell-back clauses and redemption clauses, and found that the pricing effect of the model was more complete after considering the two clauses.

2.2. Pricing Method of Convertible Bonds Based on Tree Graph Method

It can be seen from the previous study that the pricing effect of the analytical method on the convertible bond is not very good, so scholars began to conduct in-depth research on the numerical method of pricing. Since Cox et al. (1979) proposed the binary tree pricing model in 1979, scholars at home and abroad began to use it for pricing convertible bonds. Hung and Wang (2002) used the binary tree method to derive the value of convertible bonds after considering the two factors of random interest rate and default risk. On this basis, Das and Sundaram (2007) added stochastic volatility as the third factor, in which CEV model was used to predict stochastic volatility. Ma et al. (2019) priced Chinese convertible bonds by using the willow model that included stock price and interest rate. Although the willow model is similar to the classical binary tree, the nodes in the willow model do not increase over time.

2.3. Pricing Method of Convertible Bonds Based on Finite Difference Method

Brennan and Schwartz (1977) first established the finite difference method to solve the structural model, and provided the convenient conditions of partial differential equations under some special conditions in 1980. Takahashi et al. (2001) applied the idea of investment fund recovery in the event of default of convertible bonds to the pricing model, adopted the credit risk model and obtained the value of convertible bonds through the finite difference method. Lau and Kwok (2004) verified the impact of two redemption clauses on the value of convertible bonds. Domestic scholars Xie et al. (2013), under the assumption that the discovery of the credit risk of underlying stocks will not decline to zero, found similar pricing effects when constructing a trinomial tree and a finite difference method.
for comparative analysis of pricing. In recent years, Chang and Wang (2020), based on the premise of Tsallis entropy distribution and instantaneous default risk, obtained the partial differential equation satisfied by the price of convertible bonds by applying the principle of no arbitrage, so as to derive the numerical solution of the price of convertible bonds (finite element method).

2.4. Pricing Method of Convertible Bonds Based on Least Square Monte Carlo Simulation Method

Since Longstaff and Schwartz (2001) first applied the least square Monte Carlo simulation method (LSM) to solve the American option problem, Lvov et al. (2004) used the CIR stochastic interest rate model to compare the accuracy of the LSM method and the finite difference method for convertible bonds. Yang et al. (2010) introduced the dilution effect of downward correction of convertibility and found through LSM model pricing that random interest rates could affect the pricing of convertible bonds, which was suggested to be included in the model, which was also verified by Batten et al. (2018). Feng et al. (2018) conducted an in-depth study on the impact of redemption clauses and sell-back clauses on the pricing of convertible bonds based on the LSM method.

2.5. Debentable Pricing Method Based on Machine Learning Model

It can be seen that traditional convertible bond pricing models are based on the displayed mathematical formula for pricing, but it is difficult for these models to reproduce the unique statistical characteristics of financial series, such as leverage effect, coarse-fine volatility correlation and the gain/loss asymmetry of financial time series (Takahashi et al. (2019); Dogariu et al. (2021)). Zhou et al. (2007) compared the pricing effect of the B–S model, the binary tree model and the artificial neural network model on convertible bonds and found that the estimation effect of the artificial neural network model is better. Niu and Ba (2021) used 31 factors of convertible bonds as input variables to predict the price of convertible bonds and found that the support vector regression model could well complete the prediction task.

2.6. Review of GAN Model Applications in Various Fields

Deep learning, especially GAN proposed by Goodfellow et al. (2014), has the potential to conduct dynamic modeling of complex data. Therefore, in this paper, we choose to use GAN as the generation model of stock price. In the financial field, GAN architecture also shows advantages in stock price prediction (Dogariu et al. (2021)), separation of market behavior from stock price movements (Hadad et al. (2017)), and systematic trading strategies (Koshiyama et al. (2021)). The application of the GAN model by domestic scholars mainly focuses on image generation and natural language processing, and a few of them are involved in the financial field. In the field of image generation, Yang and Jiao (2021) adopted a new structure, Retina-GAN, integrating the attention mechanism and RU-Net structure into the generator node, generating the adantagonistic network (GAN) for automatic vascular segmentation of fundus images. In the financial field, Yao et al. (2022) conducted a study on credit bond default data based on the WGAN model and the SMOTE sampling technique. The research results show that the generation technique of the GAN model can improve the predictive ability of related algorithms on bond default risk and provide a new way to study bond default risk forecast under non-balanced sample conditions.

2.7. Overview

Through the review of the relevant research literature, it is found that the research on pricing of convertible bonds in China started relatively late. Although compared with the development lag of foreign countries, domestic scholars pay more attention to exploring pricing methods of convertible bonds suitable for the Chinese market.

In the study of convertible bond pricing, domestic and foreign scholars have experienced the evolution from the B–S formula to the tree graph method and finite difference
method, and then to the least square Monte Carlo method, which gradually takes more terms and situations into account, thus improving the applicability and accuracy of the pricing model. With the development of machine learning, scholars at home and abroad have also begun to explore the application of relevant models to the pricing of convertible bonds, and found that the machine learning model can improve the pricing effect of convertible bonds.

It can be seen that existing research mainly revises and improves the mainstream pricing model, but lacks the innovative application of convertible bond pricing and the pricing accuracy of traditional pricing methods is poor. Therefore, this paper uses the traditional pricing model, LSTM model, GAN model and LSM improved model to price convertible bonds, and compares and analyzes the pricing effects and application scenarios of different models.

3. Research Design

Convertible bonds, referred to as convertible bonds, can be converted into a certain percentage of the underlying stocks in accordance with the agreed terms during the conversion period, or held until maturity to obtain all interest, or sold directly for earnings. Due to its complex structure and characteristics, this chapter mainly elaborates related background concepts of convertible bonds from the development history and basic concepts of convertible bonds. In addition, while discussing the traditional pricing theory of convertible bonds, this paper also expounds the deep learning models used in this paper. Finally, the evaluation index of the model pricing effect and the authenticity evaluation index theory, which needed to be used in the generation sequence of the LSM improved model, are analyzed.

3.1. Pricing Theory of Convertible Bonds

3.1.1. Traditional Convertible Bond Pricing Theory

(1) B–S option pricing theory

The Black–Scholes formula was put forward by Black and Scholes in April 1973, and later perfected by Morton, which became the classic algorithm in the option pricing model. Subsequently, Ingersoll (1977) and other scholars introduced it to analyze the pricing of domestic convertible bonds. As an analytical method, the B–S model mainly decomposed the intrinsic value of convertible bonds into two parts: pure bond value and option value, and then directly calculated the value of the two parts of convertible bonds through the B–S option pricing model. Finally, the total value of pure bond and option was added to obtain the analytical solution of convertible bonds, namely:

$$\text{Value of convertible bonds} = \text{Net debt value} + \text{Option value}$$  \hspace{1cm} (1)

The value of the pure debt part of the convertible bond can be obtained through the classical discounted cash model, as for the meaning of each letter in the model please refer to the following formula and Table 1, namely:

$$B_{t,T} = \sum_{t=0}^{T} \frac{I}{(1+R)^t} + \frac{FV}{(1+R)^T}$$  \hspace{1cm} (2)

**Table 1.** The meanings of each letter in the discounted cash flow model.

<table>
<thead>
<tr>
<th>Name</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>$B_{t,T}$</td>
<td>The theoretical price of the ordinary bond tranche</td>
</tr>
<tr>
<td>$I$</td>
<td>The coupon rate for each interest payment date</td>
</tr>
<tr>
<td>$FV$</td>
<td>Principal due for conversion</td>
</tr>
<tr>
<td>$R$</td>
<td>Convertible bond yield to maturity (discount rate)</td>
</tr>
<tr>
<td>$T$</td>
<td>The duration of a convertible bond</td>
</tr>
</tbody>
</table>
The value of the option part of the convertible bond is obtained through the B–S formula, which is derived as follows:

In the case of risk neutral, the intrinsic option of a convertible bond is simplified into a European call option. The exercise price of a convertible bond is the price of the equity transfer, denoted as X; 100 (par value)/X is the conversion ratio; the stock price at the current time t is \( S_t \), and the stock price at the expiration time \( T \) is \( S_T \), \( C_{t,T} \), indicating the partial value of the convertible bond option.

\[
E[\max\left(\frac{100}{X}(S_T - X), 0\right)] = \frac{100}{X} E[\max(S_T - X, 0)]
\]

\[
C_{t,T} = \frac{100}{X} E[\max(S_T - X, 0)] e^{-r_f(T-t)}
\]

According to the B–S formula, the logarithmic return rate of the stock price follows the normal distribution, and is as follows:

\[
E[\max(S_T - X, 0)] e^{-r_f(T-t)} = S_t N(d_1) - X e^{-r_f(T-t)} N(d_2)
\]

\[
C_{t,T} = \frac{100}{X} \left[ S_t N(d_1) - X e^{-r_f(T-t)} N(d_2) \right]
\]

where \( N(*) \) represents the cumulative probability distribution function of the standard normal distribution variable with a mean of 0 and a standard deviation of 1; \( d_1 = \frac{\ln(S_t/X) + (r_f + \frac{\sigma^2}{2})(T-t)}{\sigma \sqrt{T-t}} \), \( d_2 = d_1 - \delta \sqrt{T-t} \), and \( \delta \) represent stock price volatility.

Therefore, with the B–S option pricing method, the value of the convertible bond is

\[
CBV_{t,T} = B_{t,T} + C_{t,T}
\]

(2) Binary tree pricing method

Specifically, the option validity period of the convertible bond is divided into several time intervals \( \Delta t \). Suppose that the stock price of the company at the current moment rises or falls from moment \( t_0 \), and the stock price at the next moment \( t_0 + \Delta t \) is \( Su \) or \( Sd \), where \( u \) represents the rising range and \( d \) represents the falling range. Suppose that the probability of rising is \( P \), the probability of falling is \( P - 1 \) and the number of rising probabilities is \( m \), then:

\[
S_{t,T} = Su^m d^{n-m}
\]

According to the risk-neutral pricing theory, \( P = e^{r_f \Delta t} - d \), \( u = e^{\delta \sqrt{\Delta t}} \), \( d = \frac{1}{u} r f \) is the risk-free interest rate and \( \delta \) is the volatility of the stock price.

(3) Finite difference method

The application of the finite difference method to the pricing of convertible bonds was first proposed by scholars such as Brennan and Schwartz (1977). In the model of Tsiveriotis and Fernandes (1998), the value of convertible bonds (V) is divided into debt part value (B) and stock part value (E). The specific formula is as follows:

\[
\frac{\partial V}{\partial t} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS \frac{\partial V}{\partial S} - r(V - B) - (r_f + r_c)B = 0
\]

\[
\frac{\partial B}{\partial t} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 B}{\partial S^2} + rS \frac{\partial B}{\partial S} - (r_f + r_c)B = 0
\]

where \( S \) is the stock price of the underlying asset; \( r_f \) is the risk-free interest rate; \( r_c \) is credit risk spreads; \( \delta \) is the volatility of stock prices.
In this model, the impact of redemption clauses and sell-back clauses on convertible bonds can also be taken into account. Based on the above partial differential equation, Lai et al. (2005) took into account the downward revision clause and modified the terminal conditions and boundary conditions to make the model more suitable for Chinese market conditions.

3.1.2. Least Square Monte Carlo (LSM) Model

Basic principles

The idea of the LSM model is to discretize the duration of convertible bonds \([0,T]\) into \(\{t_0, t_1, t_2, \ldots, t_k\}\). The optimal exercise time in each simulated path is obtained by backward pushing. There are two main considerations in choosing the optimal exercise time: (1) the holding value of delayed exercise is less than the immediate conversion value; (2) the return of immediate stock conversion is greater than 0. According to the method proposed by Longstaff and Schwartz (2001) in 2001, the conditional expected income of the next exercise is

\[
E(e^{-rf\Delta t} h(S_k|S_{k-1})),
\]

taken as the holding value of the delayed exercise in a risk-neutral scenario. Specific conditions for exercising the right are as follows:

\[
\begin{align*}
E(e^{-rf\Delta t} h(S_k|S_{k-1})) &\leq h(t_n) \text{ and } h(t_n) > 0, \text{ Immediate exercise} \\
\text{Otherwise, No exercised}
\end{align*}
\]

where the conditional expected return \(E(e^{-rf\Delta t} h(S_k|S_{k-1})) = \sum_{i=0}^{\infty} a_i L_i(S_k)\), \(a_i\) refers to the parameter to be estimated of the Laguerre polynomial at time \(i\), \(L_i(S_k)\) is the basis function at time \(i\). The common linear combination of the basis function is

\[
E(e^{-rf\Delta t} h(S_k|S_{k-1})) = a_0 + a_1 S_k + a_2 S_k^2.
\]

In order to more intuitively show the pricing process of LSM and combine with the path classification method of stock prices proposed by Feng et al. (2016),

Figures 2 and 3 are drawn as the pricing situation in the simulation path, and path1 represents the path where early redemption occurs. path2 and path3 represent the path of in-price and out-of-price American options.

Figure 2. Unrevised path.
3.1.3. Comparison of Traditional Pricing Methods

The above two sections discuss the basic principles and specific steps of traditional pricing methods. In order to intuitively compare the three pricing methods, the advantages and disadvantages of the above four technologies are summarized (Table 2):

<table>
<thead>
<tr>
<th>Pricing Method</th>
<th>Advantage</th>
<th>Shortcoming</th>
</tr>
</thead>
</table>
| B–S model                | (1) The thinking is clear and the analytical solution can be obtained  
(2) Easy to calculate   | (1) Use only European options  
(2) Unable to reflect the value of the additional terms of the convertible bond |
| Binary tree              | (1) Ability to handle American options                         | (2) Additional clauses dealing with path dependence are relatively weak     |
| Finite difference method | (1) Ability to handle American options  
(2) By changing the terminal, more factors affecting the value of convertible bonds can be considered | (2) The method is limited by the partial differential equation of convertible bonds  
(1) The amount of calculation is large, and it needs to consume a large amount of memory, so it is difficult to achieve tens of thousands of simulations |
| LSM simulation           | (1) It can solve the problem of path dependence  
(2) Ability to handle American options                         |                                                                             |

Based on the above analysis, the least square Monte Carlo simulation is used as the pricing comparison group, compared with the deep learning pricing model alone, and the deep learning financial data generation technology is applied to the least square Monte Carlo pricing, so as to overcome the situation that the assumptions of traditional pricing methods are divorced from the real situation.

3.2. Theoretical Analysis of Neural Network Model

Considering the difficulties of existing pricing methods, this paper adopts two pricing ideas. One is to learn from the neural network model adopted by Yang et al. (2020) as the pricing model to price convertible bonds at any time within the duration. The other is to learn from the in-depth pricing method adopted by Tan et al. (2022) for the traditional pricing framework of convertible bonds. This method has the stylized fact of high real data reproducibility and can skip many overly stringent assumptions.

3.2.1. LSTM Model

The Long short-term memory network LSTM (Long short-term memory) is a variant of RNN, whose core concept lies in cell state and “gate” structure. Its essence is an improved version of RNN. On the basis of the original RNN, a cell state $c$ is added for long-term state
storage. The structure diagram is as follows Figure 4 (where \( C_t \) represents the memory state at time \( t \) and \( h_{t-1} \) represents the hidden layer state at time \( T - 1 \)):

\[
f_t = \text{Sigmoid}(W_f[h_{t-1}, X_t] + b_f) \tag{12}
\]

The function of the forgetting gate is to control the degree to which the cellular state of the previous moment is retained, that is, the degree of previous memory. The principle is to take the hidden state \( h_{t-1} \) and input value \( X_t \) at the moment \( t - 1 \) as input, and convert it into a probability value \( f_t \) with the interval between 0 and 1 through Sigmoid activation. The closer to 1, the more important it is, that is, the more important the past information is. \( W_f, b_f \) represent the weight and bias of the forgetting door, respectively.

\[
i_t = \text{Sigmoid}(W_i[h_{t-1}, X_t] + b_i) \tag{13}
\]

\[
\tilde{C}_t = \tanh(W_c[h_{t-1}, X_t] + b_c) \tag{14}
\]

The main function of the input gate, also known as the memory gate, is to update the cell state, and the forgetting gate input method; first of all, the hidden state \( h_{t-1} \) and the input value \( X_t \) at \( t - 1 \) moment are taken as the input, and converted into a probability value \( i_t \) with the interval between 0 and 1 through Sigmoid activation, which is used to control the updating degree of the new information. Both inputs are passed into the Tanh layer at the same time to activate, creating a new cell state \( \tilde{C}_t \).

\[
C_t = f_tC_{t-1} + i_t\tilde{C}_t \tag{15}
\]

The function of the cell state structure is mainly used to generate the cell state at the current moment. The work flow is to take the output of the forgetting gate and memory gate as the input to update the cell state, so as to obtain the cell state \( C_t \) at the moment \( t \).

\[
O_t = \text{Sigmoid}(W_o[h_{t-1}, X_t] + b_o) \tag{16}
\]
\[ h_t = O_t \cdot \tan h(C_t) \]  

(17)

3.2.2. Generative Adversarial Network (GAN)

(1) The basic principle of generating adversarial network model.

Since the GAN model was proposed by Goodfellow et al. (2014) in 2014, it has become one of the most popular generation task algorithms at present. The function of a large number of seeds is to ensure the diversity of the generated images. In the most original GAN papers, a multi-layer perceptron (i.e., MLP) was used to build generation models and discriminant models, and the specific structure is shown in Figure 5.

![Figure 5. Structure diagram of GAN model.](image)

The distribution of the generator and discriminator in the training process is shown in Figure 6 (the generated data distribution is the blue line, and the real data distribution is the green line).

![Figure 6. Distribution of real data and generated data during training.](image)

As shown in the figure above, the line trains the discriminator so that it can distinguish the real data from the generated data, and then trains the discriminator so that the distance between the generated data and the real data becomes smaller. After several rounds of training of the discriminator generator, we hope that the distribution of the generated sample and the real sample has been completely consistent. And the discriminator cannot tell them apart anymore.

As shown in Figure 4, random noise \( z \) (each element of \( z \) is generally set as independent and equally distributed and subject to standard normal distribution or uniform distribution \([0, 1]\)) is passed into generator \( G \) to obtain output distribution \( P_G \). The distribution of real data is represented by \( P_{\text{data}} \) and then the output distribution data \( P_G \) and real distribution data \( P_{\text{data}} \) are passed into discriminator \( 7D \), respectively, to obtain the probability value of sample generation and the probability value of the real sample. The specific optimization objective function is as follows

\[
\min_G \max_D V(D, G) = E_{x \sim P_{\text{data}}}[\log D(x)] + E_{z \sim P_z}[\log(1 - D(G(z)))]
\]  

(18)

where \( x \sim P_{\text{data}} \) represents the data distribution derived from real data, and \( z \sim P_z \) is equivalent to a sampling of noise data. \( E_{x \sim P_{\text{data}}}[\log D(x)] \) represents the discrimina-
tor’s expectation of the judgment result of the real sample. For the optimal discriminator $D$, the judgment result $D(x)$ for the real sample should be 1, that is, $\log(D(x))$ is 0; similarly, $E_{z \sim P_z}[\log(1 - D(G(z)))]$ represents the discriminator’s expectation of the spurious sample judgment result.

3.2.3. Wasserstein Generated Network Model (WGAN)

WGAN is the Wasserstein generation network model proposed by Arjovsky and Bottou (2017), whose core idea is to measure the distance between beam parts by Earth-Mover distance (also known as Wasserstein distance). Its advantage is that compared with KL divergence and JS divergence, even if the two distributions do not overlap, Wasserstein distance can still reflect their distance, which is exactly what will really happen under the condition of high dimension and insufficient sample sampling. Based on this, the loss function of the generator and discriminator in WGAN can be rewritten as:

$$G_{loss} = -E_{z \sim P_z}[f_w(G(z))]$$  \hspace{1cm} (19)

$$D_{loss} = E_{z \sim P_z}[f_w(G(z))] - E_{x \sim P_{data}}[f_w(x)]$$  \hspace{1cm} (20)

where $f_w(x)$ represents a series of functions $f$ that depend on the parameter $w$, which can be any function chosen, or a network model, where the parameter $w$ is the set of parameters in the network. $f_w(x)$ satisfies the Lipschitz condition (a condition of smoothness stronger than uniformly continuous), namely $|f_w(a) - f_w(b)| \leq K|a - b|$, where $K$ is the Lipschitz constant and generally takes a value of 1.

Later, Gulrajani et al. (2017) added a penalty term to the discriminator loss function on the basis of WGAN to replace weight clipping in WGAN, which was widely adopted to overcome the situation of the poor weight clipping effect, and then obtained the discriminator loss function in the WGAN-GP model, namely

$$D_{loss} = E_{z \sim P_z}[f(G(z))] - E_{x \sim P_{data}}[f_w(x)] + \lambda E_{x \sim P_{penalty}}[(||\nabla_x f(x)||_2 - 1)^2]$$  \hspace{1cm} (21)

where $\lambda E_{x \sim P_{penalty}}[(||\nabla_x f(x)||_2 - 1)^2]$ represents the gradient penalty, which is used to constrain function $f_w(x)$ to meet the Lipschitz condition, and $x \sim P_{penalty}$ represents the penalty distribution composed of the region of the generated data set, the region of the real data set and the region between them.

3.2.4. LSM Improved Model

Consistent with the discussion in the previous two chapters, the improved model of LSM mainly refers to the method mentioned in Wiese et al. (2020) and Tan et al. (2022), and uses the GAN model instead of the Monte Carlo method to generate simulation data and put it into the traditional LSM model, thus forming the improved model of LSM. The generator and discriminator of the GAN model both use the TCN model, so this part mainly analyzes the theory of time convolutional network.

4. Data and Empirical Analysis

4.1. LSM Pricing Model

4.1.1. Model Description

The research samples selected in this paper are bank convertible bonds. It can be seen from the issuance announcement of each sample that the trigger condition of call-back is the following: Only the holders of convertible bonds can trigger the call-back clause when they are identified by China Securities Regulatory Commission as changing the use of raised funds, so it cannot be quantified, so call-back clause is not considered in the model. Secondly, according to the inference in the literature of Zhou and Wu (2013), Zheng and Lin (2004) and the downward revision conditions of the special downward revision clause, it is found that the downward revision clause is equivalent to transferring
the old shareholders’ equity to the investors of convertible bonds, which directly affects the original shareholders’ equity.

4.1.2. Sample Data and Descriptive Statistics

As can be seen from the total issuance volume in chapter one, the largest convertible bonds issuing scale in our country are usually banks. In addition, convertible bonds with a higher rating are usually issued by banks. Therefore, the sample mainly selects the existing bank convertible bonds on 20 January 2023, whose pricing date is the first day of listing, and conducts multi-node pricing on the Shanghai Pudong Development convertible bonds (through analysis of pricing results with other models), with the pricing range from 20 January 2022 to 20 January 2023, every 30 days as a node. The convertible bond price of each node is estimated, respectively.

4.1.3. Parameter Estimation

The first step of the LSM pricing model is to estimate the volatility of the risk-free interest rate and stock price. The following is the estimation process and results of the two parameters. In addition to that, other input parameters are set as follows: Monte Carlo simulation compensation is 252 trading days in a year, the number of simulations is 5000 times and the initial stock price selected is recorded as $S_0$ on the day before the listing of convertible bonds.

(1) Risk-free interest rate

In the selection of risk-free interest rate, the maturity yield of ordinary bonds of China commercial banks with the same rating as the issuing subject is selected as the risk-free interest rate, which can consider the credit risk of enterprises on the basis of the risk-free interest rate, namely, credit spread.

(2) Stock price volatility

For the estimation of stock price volatility, the academic circle generally adopts the historical volatility estimation method and the GRACH model estimation method. The former uses the closing price data of the 252 trading days before the pricing date to estimate, while the latter uses the GARCH model to solve the unconditional stock price volatility, which is generally divided into four steps:

a. Descriptive statistics of the logarithmic rate of return of stock price;
b. Test the stationarity of logarithmic rate of return;
c. Autocorrelation and ARCH effect test;
d. GARCH model is used to estimate stock price volatility, and unconditional stock price volatility is obtained.

This paper refers to the research results of Zhao and Zhao (2009) and uses the GARCH model to solve stock price volatility. Next, take the Jiangbank Convertible Bond (128034.SZ) for example to illustrate the process of estimating stock price volatility as the following chart:

First of all, the logarithmic yields of the first 252 trading days of the CB were selected and visualized, as shown in Figure 7, and the descriptive statistics were carried out, as shown in Tables 3 and 4.

As can be seen from Table 5, the Jarque-Bera statistic is 38.85, $p$ value is $3.7 \times 10^{-9}$, and H0 hypothesis is rejected, indicating that the return rate of logarithm of stock does not obey the normal distribution.
### Table 3. Descriptive statistics of bank class samples (still existing on 20 January 2023).

<table>
<thead>
<tr>
<th>Security Code</th>
<th>Security Abbreviation</th>
<th>Term of Issue</th>
<th>Listing</th>
<th>Maturity</th>
<th>Transfer Price</th>
<th>Debt Rating</th>
<th>Subject Rating</th>
<th>Issue Size (100 Million)</th>
</tr>
</thead>
<tbody>
<tr>
<td>128034.SZ</td>
<td>Jiangyin convertible bonds</td>
<td>6</td>
<td>2018/2/14</td>
<td>2024/1/26</td>
<td>9.16</td>
<td>AA+</td>
<td>AA+</td>
<td>20</td>
</tr>
<tr>
<td>110059.SH</td>
<td>Development convertible Bond</td>
<td>6</td>
<td>2019/11/15</td>
<td>2025/10/27</td>
<td>15.05</td>
<td>AAA</td>
<td>AAA</td>
<td>500</td>
</tr>
<tr>
<td>110043.SH</td>
<td>Wuxi convertible bond</td>
<td>6</td>
<td>2018/3/14</td>
<td>2024/1/30</td>
<td>8.9</td>
<td>AA+</td>
<td>AA+</td>
<td>30</td>
</tr>
<tr>
<td>110053.SH</td>
<td>Su Yin convertible bond</td>
<td>6</td>
<td>2019/4/3</td>
<td>2025/3/13</td>
<td>7.9</td>
<td>AAA</td>
<td>AAA</td>
<td>200</td>
</tr>
<tr>
<td>110079.SH</td>
<td>Hangyi convertible bond</td>
<td>6</td>
<td>2021/4/23</td>
<td>2027/3/28</td>
<td>17.06</td>
<td>AAA</td>
<td>AAA</td>
<td>150</td>
</tr>
<tr>
<td>113011.SH</td>
<td>Everbright bonds</td>
<td>6</td>
<td>2017/4/5</td>
<td>2023/3/16</td>
<td>4.36</td>
<td>AAA</td>
<td>AAA</td>
<td>300</td>
</tr>
<tr>
<td>113021.SH</td>
<td>Citic bonds</td>
<td>6</td>
<td>2019/3/19</td>
<td>2025/3/3</td>
<td>7.45</td>
<td>AAA</td>
<td>AAA</td>
<td>400</td>
</tr>
<tr>
<td>113042.SH</td>
<td>On the silver bond</td>
<td>6</td>
<td>2021/2/10</td>
<td>2027/1/24</td>
<td>11.03</td>
<td>AAA</td>
<td>AAA</td>
<td>200</td>
</tr>
<tr>
<td>113050.SH</td>
<td>South Bank convertible bonds</td>
<td>6</td>
<td>2021/7/1</td>
<td>2027/6/14</td>
<td>10.1</td>
<td>AAA</td>
<td>AAA</td>
<td>200</td>
</tr>
<tr>
<td>113052.SH</td>
<td>CCB convertible bond</td>
<td>6</td>
<td>2022/1/14</td>
<td>2027/12/26</td>
<td>25.51</td>
<td>AAA</td>
<td>AAA</td>
<td>500</td>
</tr>
<tr>
<td>113055.SH</td>
<td>Convert bonds into silver</td>
<td>6</td>
<td>2022/4/06</td>
<td>2028/3/2</td>
<td>14.53</td>
<td>AAA</td>
<td>AAA</td>
<td>80</td>
</tr>
<tr>
<td>113056.SH</td>
<td>Heavy silver bond</td>
<td>6</td>
<td>2022/4/14</td>
<td>2028/3/22</td>
<td>11.28</td>
<td>AAA</td>
<td>AAA</td>
<td>130</td>
</tr>
<tr>
<td>113062.SH</td>
<td>Changyi convertible bond</td>
<td>6</td>
<td>2022/10/17</td>
<td>2028/9/14</td>
<td>8.08</td>
<td>AA+</td>
<td>AA+</td>
<td>60</td>
</tr>
<tr>
<td>113065.SH</td>
<td>Qilu convertible bond</td>
<td>6</td>
<td>2022/12/19</td>
<td>2028/11/28</td>
<td>5.87</td>
<td>AAA</td>
<td>AAA</td>
<td>80</td>
</tr>
<tr>
<td>113516.SH</td>
<td>Sunong convertible bonds</td>
<td>6</td>
<td>2018/8/20</td>
<td>2024/8/2</td>
<td>6.34</td>
<td>AA+</td>
<td>AA+</td>
<td>25</td>
</tr>
<tr>
<td>128129.SZ</td>
<td>Green farmers swap bonds</td>
<td>6</td>
<td>2020/9/18</td>
<td>2026/8/24</td>
<td>5.74</td>
<td>AAA</td>
<td>AAA</td>
<td>50</td>
</tr>
</tbody>
</table>

Data source: Choice database.

### Table 4. Risk-free interest rates of 17 sample convertible bonds considering credit risk.

<table>
<thead>
<tr>
<th>Security Abbreviation</th>
<th>Debt Rating</th>
<th>Listing</th>
<th>Risk-Free Rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jiangyin convertible bonds</td>
<td>AA+</td>
<td>2018/2/14</td>
<td>5.6003</td>
</tr>
<tr>
<td>Shanghai Pudong Development convertible Bond</td>
<td>AAA</td>
<td>2019/11/15</td>
<td>3.8117</td>
</tr>
<tr>
<td>Wuxi convertible bond</td>
<td>AA+</td>
<td>2018/3/14</td>
<td>5.4230</td>
</tr>
<tr>
<td>Su Yin convertible bond</td>
<td>AAA</td>
<td>2019/4/3</td>
<td>4.0120</td>
</tr>
<tr>
<td>Hangyi convertible bond</td>
<td>AAA</td>
<td>2021/4/23</td>
<td>3.6957</td>
</tr>
<tr>
<td>Everbright bonds</td>
<td>AAA</td>
<td>2017/4/5</td>
<td>4.4287</td>
</tr>
<tr>
<td>Citic bonds</td>
<td>AAA</td>
<td>2019/3/19</td>
<td>4.0818</td>
</tr>
<tr>
<td>On the silver bond</td>
<td>AAA</td>
<td>2021/2/10</td>
<td>3.7587</td>
</tr>
<tr>
<td>South Bank convertible bonds</td>
<td>AAA</td>
<td>2021/7/1</td>
<td>3.6836</td>
</tr>
<tr>
<td>CCB convertible bond</td>
<td>AAA</td>
<td>2022/1/14</td>
<td>3.2402</td>
</tr>
<tr>
<td>Convert bonds into silver</td>
<td>AAA</td>
<td>2022/4/06</td>
<td>3.4224</td>
</tr>
<tr>
<td>Heavy silver bond</td>
<td>AAA</td>
<td>2022/4/14</td>
<td>3.3653</td>
</tr>
<tr>
<td>Changyi convertible bond</td>
<td>AA+</td>
<td>2022/10/17</td>
<td>3.0916</td>
</tr>
<tr>
<td>Qilu convertible bond</td>
<td>AAA</td>
<td>2022/12/19</td>
<td>3.3693</td>
</tr>
<tr>
<td>Sunong convertible bonds</td>
<td>AA+</td>
<td>2018/8/20</td>
<td>4.8510</td>
</tr>
<tr>
<td>Zhang Bank convertible bonds</td>
<td>AA+</td>
<td>2018/11/29</td>
<td>4.5052</td>
</tr>
<tr>
<td>Green farmers swap bonds</td>
<td>AAA</td>
<td>2020/9/18</td>
<td>3.6520</td>
</tr>
</tbody>
</table>

### Table 5. Descriptive statistics of the logarithmic yield rate of Jiangbank convertible bonds and regular shares.

<table>
<thead>
<tr>
<th>Name</th>
<th>Mean</th>
<th>Max</th>
<th>Min</th>
<th>STD</th>
<th>Jarque-Bera</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jiangyin convertible bonds</td>
<td>−0.001116</td>
<td>0.09561</td>
<td>−0.10552</td>
<td>0.03721</td>
<td>38.85 (3.7 × 10⁻⁹)</td>
</tr>
</tbody>
</table>
In the selection of risk-free interest rate, the maturity yield of ordinary bonds of China commercial banks with the same rating as the issuing subject is selected as the risk-free interest rate, which can consider the credit risk of enterprises on the basis of the risk-free interest rate, namely, credit spread.

(2) Stock price volatility

For the estimation of stock price volatility, the academic circle generally adopts the historical volatility estimation method and the GRACH model estimation method. The former uses the closing price data of the 252 trading days before the pricing date to estimate, while the latter uses the GARCH model to solve the unconditional stock price volatility, which is generally divided into four steps:

a. Descriptive statistics of the logarithmic rate of return of stock price;

b. Test the stationarity of logarithmic rate of return;

c. Autocorrelation and ARCH effect test;

d. GARCH model is used to estimate stock price volatility, and unconditional stock price volatility is obtained.

This paper refers to the research results of Zhao and Zhao (2009) and uses the GARCH model to solve stock price volatility. Next, take the Jiangbank Convertible Bond (128034.SZ) for example to illustrate the process of estimating stock price volatility as the following chart:

Figure 7. Yield distribution of convertible bonds and shares of Jiangbank.

Then the second step is the ADF unit root test. The specific results are shown in Table 6. It can be seen from the table that, at the significance level of 1%, the null hypothesis of the existence of unit root can be rejected, that is, the time series of logarithmic return rate of positive shares of unit root does not exist and is stable.

Table 6. Descriptive statistics of logarithmic yield rate of Jiangbank convertible bonds.

<table>
<thead>
<tr>
<th>Name</th>
<th>ADF</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jiangyin convertible bonds</td>
<td>−13.279</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Furthermore, the ARCH effect test was conducted on the square sequence of residual errors in the mean value equation (see Table 7). It can be seen that the Lagrange multiplier statistic is 27.649, the p value is zero and the ARCH effect exists.

Table 7. ARCH effect test.

<table>
<thead>
<tr>
<th>Name</th>
<th>Lagrange Multiplier Statistics</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jiangyin convertible bonds</td>
<td>27.649</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Finally, the GARCH (1,1) model is used to correct the volatility of stocks. According to the results of Stata, the GARCH (1,1) model of the stock volatility of the convertible bonds of Jiangbank can be obtained as follows:

\[
\sigma_t^2 = 0.0000566 + 0.1027115\alpha_{t-1}^2 + 0.8478677\sigma_{t-1}^2 
\]  

(22)

Thus, according to the estimation formula \( \sigma^2 = \frac{a_0}{1 - a_1 - b_1} \) of the unconditional variance formula, the estimated value \( \hat{\sigma}^2 \) of daily volatility can be obtained as 0.001145, which is converted into an annual volatility of 0.5372, shown in Table 8.
Table 8. Summary of estimated annualized volatility parameters of 17 convertible bonds.

<table>
<thead>
<tr>
<th>Security Code</th>
<th>Security Abbreviation</th>
<th>Historical Volatility</th>
<th>GARCH(1,1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>128034.SZ</td>
<td>Jiangyin convertible bonds</td>
<td>0.5896</td>
<td>0.5372</td>
</tr>
<tr>
<td>110059.SH</td>
<td>Shanghai Pudong Development convertible Bond</td>
<td>0.2140</td>
<td>0.2239</td>
</tr>
<tr>
<td>110043.SH</td>
<td>Wuxi convertible bond</td>
<td>0.5056</td>
<td>0.4960</td>
</tr>
<tr>
<td>110053.SH</td>
<td>Su Yin convertible bond</td>
<td>0.2098</td>
<td>0.3323</td>
</tr>
<tr>
<td>110079.SH</td>
<td>Hangyin convertible bond</td>
<td>0.3931</td>
<td>0.4133</td>
</tr>
<tr>
<td>113011.SH</td>
<td>Everbright bonds</td>
<td>0.1581</td>
<td>0.1602</td>
</tr>
<tr>
<td>113021.SH</td>
<td>Citic bonds</td>
<td>0.2493</td>
<td>0.2652</td>
</tr>
<tr>
<td>113042.SH</td>
<td>On the silver bond</td>
<td>0.1858</td>
<td>0.1871</td>
</tr>
<tr>
<td>113050.SH</td>
<td>South Bank convertible bonds</td>
<td>0.3129</td>
<td>0.3123</td>
</tr>
<tr>
<td>113052.SH</td>
<td>CCB convertible bond</td>
<td>0.3473</td>
<td>0.3467</td>
</tr>
<tr>
<td>113055.SH</td>
<td>Convert bonds into silver</td>
<td>0.3495</td>
<td>0.3532</td>
</tr>
<tr>
<td>113056.SH</td>
<td>Heavy silver bond</td>
<td>0.2510</td>
<td>0.2505</td>
</tr>
<tr>
<td>113062.SH</td>
<td>Changyin convertible bond</td>
<td>0.2987</td>
<td>0.2954</td>
</tr>
<tr>
<td>113065.SH</td>
<td>Qilu convertible bond</td>
<td>0.2859</td>
<td>0.2866</td>
</tr>
<tr>
<td>113516.SH</td>
<td>Sunong convertible bonds</td>
<td>0.3746</td>
<td>0.3764</td>
</tr>
<tr>
<td>128048.SZ</td>
<td>Zhang Bank convertible bonds</td>
<td>0.5175</td>
<td>0.5175</td>
</tr>
<tr>
<td>128129.SZ</td>
<td>Green farmers swap bonds</td>
<td>0.4297</td>
<td>0.4361</td>
</tr>
</tbody>
</table>

4.1.4. Empirical Results

(1) Pricing results on the first day of listing

We use the pricing error rate formula mentioned in Section 2 to evaluate the pricing results of the model, namely

\[
\text{Pricing error rate} = \frac{\text{Theoretical price} - \text{Actual price}}{\text{Actual price}}
\]

Based on the LSM model and combined with the above estimated parameters as input values, the estimated first-day price, actual price and error rate of 17 sample convertible bonds simulated in Python are shown in Table 9.

Table 9. Pricing results on the first day of listing of the LSM model.

<table>
<thead>
<tr>
<th>Security Abbreviation</th>
<th>Actual Price</th>
<th>Estimated Price</th>
<th>Error Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jiangyin convertible bonds</td>
<td>97.7600</td>
<td>112.0035</td>
<td>14.56%</td>
</tr>
<tr>
<td>Shanghai Pudong Development convertible Bond</td>
<td>103.8901</td>
<td>107.4661</td>
<td>3.44%</td>
</tr>
<tr>
<td>Wuxi convertible bond</td>
<td>97.3947</td>
<td>115.2294</td>
<td>18.31%</td>
</tr>
<tr>
<td>Su Yin convertible bond</td>
<td>109.1290</td>
<td>117.3514</td>
<td>7.53%</td>
</tr>
<tr>
<td>Hangyin convertible bond</td>
<td>113.9763</td>
<td>120.0042</td>
<td>5.28%</td>
</tr>
<tr>
<td>Everbright bonds</td>
<td>102.9918</td>
<td>108.7224</td>
<td>5.56%</td>
</tr>
<tr>
<td>Citic bonds</td>
<td>107.9877</td>
<td>112.1466</td>
<td>3.85%</td>
</tr>
<tr>
<td>On the silver bond</td>
<td>100.1368</td>
<td>105.6658</td>
<td>5.52%</td>
</tr>
<tr>
<td>South Bank convertible bonds</td>
<td>120.2712</td>
<td>120.9079</td>
<td>0.52%</td>
</tr>
<tr>
<td>CCB convertible bond</td>
<td>111.0496</td>
<td>115.2653</td>
<td>3.79%</td>
</tr>
<tr>
<td>Convert bonds into silver</td>
<td>121.7108</td>
<td>124.3157</td>
<td>2.14%</td>
</tr>
<tr>
<td>Heavy silver bond</td>
<td>104.8973</td>
<td>109.5403</td>
<td>4.42%</td>
</tr>
<tr>
<td>Changyin convertible bond</td>
<td>118.7719</td>
<td>120.6737</td>
<td>1.60%</td>
</tr>
<tr>
<td>Qilu convertible bond</td>
<td>95.1164</td>
<td>111.7812</td>
<td>17.52%</td>
</tr>
<tr>
<td>Sunong convertible bonds</td>
<td>99.1781</td>
<td>118.5775</td>
<td>19.56%</td>
</tr>
<tr>
<td>Zhang Bank convertible bonds</td>
<td>102.3814</td>
<td>122.2723</td>
<td>19.42%</td>
</tr>
<tr>
<td>Green farmers swap bonds</td>
<td>109.0568</td>
<td>119.8994</td>
<td>9.94%</td>
</tr>
</tbody>
</table>

From Table 9, we can see that the traditional pricing method has a certain deviation without considering the revised terms, and the maximum value of the error is within 20%. Among them, the deviation rate of the first day estimated price of five convertible bonds is more than 10%, with a maximum value of 19.42% and a minimum value of 0.52% of
Nanyin convertible bonds. The overall mean absolute percentage error (MAPE) was 8.41%. In summary, these results show that the estimated results of the model are higher than the first-day prices of convertible bonds, indicating that the pricing function of the model needs to be improved.

(2) Multi-node pricing results within the duration

Consistent with the above, this paper compared with the pricing model in the following paper and selected the Pudong Development convertible bond with a relatively low error rate as an example to conduct multi-node pricing analysis. The pricing range was from 20 January 2022 to 20 January 2023, and the price of each node was estimated every 30 days, namely 12 nodes. Among them, stock price volatility and stock transfer price are updated every 30 days, shown Table 10 below:

Table 10. Multi-node pricing of Shanghai Pudong Development convertible bonds.

<table>
<thead>
<tr>
<th>Date</th>
<th>GARCH(1,1)</th>
<th>Actual Price</th>
<th>Estimated Price</th>
<th>Error Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>2022/2/21</td>
<td>0.1756</td>
<td>105.9092</td>
<td>102.7697</td>
<td>−2.96%</td>
</tr>
<tr>
<td>2022/3/21</td>
<td>0.1684</td>
<td>104.7241</td>
<td>102.1407</td>
<td>−2.46%</td>
</tr>
<tr>
<td>2022/4/20</td>
<td>0.1563</td>
<td>104.8308</td>
<td>102.3051</td>
<td>−2.40%</td>
</tr>
<tr>
<td>2022/5/20</td>
<td>0.1513</td>
<td>105.3475</td>
<td>102.4839</td>
<td>−2.71%</td>
</tr>
<tr>
<td>2022/6/20</td>
<td>0.1559</td>
<td>105.0701</td>
<td>102.3051</td>
<td>−2.40%</td>
</tr>
<tr>
<td>2022/7/20</td>
<td>0.1539</td>
<td>105.4568</td>
<td>103.1392</td>
<td>−2.19%</td>
</tr>
<tr>
<td>2022/8/22</td>
<td>0.1568</td>
<td>106.3042</td>
<td>103.3065</td>
<td>−2.81%</td>
</tr>
<tr>
<td>2022/9/20</td>
<td>0.1588</td>
<td>106.3561</td>
<td>103.5676</td>
<td>−2.62%</td>
</tr>
<tr>
<td>2022/10/20</td>
<td>0.1578</td>
<td>106.3608</td>
<td>103.8714</td>
<td>−2.34%</td>
</tr>
<tr>
<td>2022/11/21</td>
<td>0.1618</td>
<td>105.3742</td>
<td>102.8606</td>
<td>−2.78%</td>
</tr>
<tr>
<td>2022/12/20</td>
<td>0.1667</td>
<td>104.0123</td>
<td>103.2316</td>
<td>−0.75%</td>
</tr>
<tr>
<td>2023/1/20</td>
<td>0.1590</td>
<td>104.916</td>
<td>103.5403</td>
<td>−1.31%</td>
</tr>
</tbody>
</table>

Figure 8 describes the LSM pricing results of the multi-node duration of Shanghai Pudong Development CB. It can be seen that the trend of estimated price and actual price is basically the same, and the trend of inflection point is earlier than the real situation. The deviation of pricing results is also mostly maintained within plus or minus 3%, meeting the pricing accuracy of the pricing error rate within 95%, and MAPE is 2.26%. Duration pricing results are excellent.

Figure 8. Multi-node pricing chart of Shanghai Pudong Development Co., LTD.
4.2. Long and Short-Term Memory Network (LSTM) Pricing Model

4.2.1. Model Description

The LSTM model is one of the most frequently used neural network models among the current series prediction models, and has excellent effects in natural language processing, series prediction and other aspects. This paper will use the modified model to predict the price of convertible bonds. The LSTM’s model super parameters are set as follows:

(1) Step Length

Step size in LSTM refers to the length of input sequence in the model, that is, the sequence length to be considered by the model during prediction. However, since LSTM is a supervised neural network model, a reasonable time step cannot be known before training, so it is necessary to set this parameter. The main task of the predictive pricing model of convertible bonds in this paper is to predict the price of the fifth trading day in the future, so as to avoid large errors in the future. In this paper, the experimental time step is set as 5 consecutive trading days.

(2) Hidden layer

In the process of multiple hyperparameter adjustment, it is found that the number of hidden layers and neurons of LSTM is set to 1 and 128, respectively, and the number of neurons in the two fully connected layers between the input layer and gate structure and the LSTM gate structure and output layer is set to 16 and 128, respectively.

(3) Activation function

In the neural network model, in order to improve the effect of the model, learn some nonlinear factors and prevent the neural network from falling back into a perceptron model after training for many times, an activation functions need to be added between layers. The activation function adopted in this paper is the ReLU function, whose expression is \( f(x) = \max(0, x) \), where \( x \) representing the input value, and \( f(x) \) representing the output result.

4.2.2. Sample Data and Descriptive Statistics

As the data-driven modeling approach is inconsistent with traditional thinking, another modeling approach is adopted in this paper, that is, daily data of all maturing and unmaturing bank convertible bonds from 1 January 2010 to 20 January 2023 are selected for pricing analysis of convertible bonds. There are 25 convertible bonds in total, with a total of 18,134 sample values. Twenty-three of them were selected as training sets and the remaining two as test sets to test the pricing ability of the model (all data came from Choice financial data terminal).

In the test set, samples Zhongpu convertible bond and Jiangyin convertible bond were selected. The former had a total issuance of CNY 50 billion, was rated AAA at the time of issuance, and the issuance period was 6 years. The latter issued a total of CNY 2 billion, with a rating of AA+ at the time of issuance and a maturity of six years.

4.2.3. Data Preprocessing

In the following section, the input characteristic variables are preprocessed and converted into daily data that can be input.

(1) Quarterly financial data/monthly macro indicators

Generally speaking, the financial data referred by investors is delayed, and the financial data of the last quarter is generally released in the next quarter, so the data of this quarter adopts the data of the last quarter and adjusts the value at the beginning of each quarter. The same is true for the monthly macro indicators.

(2) Debt rating and subject rating

For debt rating and subject rating, AAA and AA+ in the sample are mapped as 1 for AAA and 2 for AA+ in this paper.
(3) Normalization

In order to increase the stability of model training and solve the dimensional problem between indexes, this paper carries out dimensionless processing on the original data. Generally speaking, dimensionless processing includes regularization, zero-mean standardization and min-max normalization. Referring to the method of previous scholars (Tang et al. (2020)), min-max normalization is adopted to de-dimensionalize data, namely

\[ x = \frac{x - \min(x)}{\max(x) - \min(x)} \]  

(23)

where \( x \) represents the feature data after normalization processing, \( \min(x) \) represents the minimum value of the feature data set, and \( \max(x) \) represents the maximum value of the feature set.

4.2.4. Empirical Results

There were 25 convertible bonds in the sample interval. In the experiment, Shanghai Development convertible bonds and Jiangbank convertible bonds were selected as test sets to verify the model pricing results. All the data in the training set were divided into several single samples with characteristics and label values according to the input sample requirements of LSTM, that is, the characteristic values of a single sample were 33 characteristic values of the past five trading days as the input of the model. The closing price of the convertible bond on the sixth trading day is used as the output of the model. Then, the samples from all training sets were divided into a training set and verification set in a ratio of about 3:1 (the number of training set samples was 10,000, and the verification set was 3670).

Figures 9 and 10, respectively, describe the loss change process of each round of LSTM in the training set and verification set, where the abscissa axis represents the epoch of training rounds, with a total of 500 rounds of training; the ordinate represents the total training error of each round, namely the summary value of the mean square error of each round. It can be seen that when the training reaches about 20 rounds, the model can simulate the samples in the training set very well, that is, better parameters are obtained in the training set.

![Figure 9. Loss diagram of training set.](image-url)
4.3. WGAN Pricing Model

4.3.1. Model Description

As explained in the theory part of Section 2, WGAN, as a derivative model of basic GAN, inherits the advantages of the former, strengthens the training stability of the model and is widely used in the industry. Due to the special structure of the dual training model of WGAN, this paper refers to the modeling method of Kang (2019). It intends to use its discriminator to enhance the pricing effect of the generator model. The structure of the model is shown as follows in Figure 11:

Figure 10. Loss diagram of verification set.

Figure 11. Structure of WGAN pricing model.

(1) Generator structure

As LSTM has many parameters, this section adopts the gated cycle unit structure (GRU) with faster computing performance as WGAN’s generator model, and uses Leaky ReLU as the activation function. The specific parameters are listed in the following table:

(2) Discriminator structure

Compared with traditional GAN, the discriminator of WGAN cancels the Sigmoid activation function of the output layer, resulting in the final output not being a value
between 0 and 1. Instead, the W distance between the generated distribution and the real distribution is fitted, thus making the training of the model more stable. This paper chooses a multi-layer perceptron as the discriminator of the WGAN model and Leaky ReLU as the activation function. The specific parameters are shown in the following table:

Where \( f_w(x) \) represents the function satisfying the Lipschitz condition, which is replaced by the discriminator network model in this paper, \( x' \sim P_{fake-data} \) represents the price sequence formed after adding the price generation result and \( x \sim P_{data} \) represents the real price sequence. In particular, according to the practice of Arjovsky and Bottou (2017), the training frequency ratio of generator and discriminator is 1:5, that is, the discriminator model trains the generator model once every five times.

4.3.2. Empirical Results

In this empirical study, it can be seen that the pricing effect of test sample 1 (Jiangbank CB) is better in the first 1000 trading days, but higher in the last 150 trading days. On the whole, the MAPE of sample 1 is 0.806%, which is slightly higher than the 0.763% of the LSTM pricing model. However, for test sample 2 (SPD CB), the results of Figure 12, Figure 13 and MAPE (0.231%) are both better than the LSTM pricing model (0.3% MAPE) and the traditional LSM model (2.25% MAPE for multi-node duration), indicating that the application of WGAN helps to improve the pricing effect of the model.

As can be seen from the above empirical results, the LSTM model and WGAN model have a better pricing effect, and the average absolute error of the two training samples in the test set is less than 1%, which fully meets the pricing requirements. In addition, both of the two models are data-driven pricing models, mainly based on the sample data to be learned, rather than relying on the preconditions in other models. They are able to mine the nonlinear relationship in the data and adapt to the real market environment.

Next, this paper takes the method mentioned by Takahashi et al. (2019) as reference and Wiese et al. (2020) and Tan et al. (2022) as reference, and uses Quant GANs proposed by Wiese et al. (2020) to replace Monte Carlo as the generation model of stock price series. In order to improve the pricing effect of the model, it avoids the premise assumptions in the traditional model and reuses the unique statistical characteristics of financial data, such as leverage effect, coarse-fine volatility correlation and the gain/loss asymmetry of financial time series.

![Figure 12. LSTM model pricing effect of test sample 1.](image-url)
4.4. Improved Model of LSM

4.4.1. Model Description

The improved model of LSM in this paper adopts the Quant GANs model proposed by Wiese et al. (2020) as the generation model of stock price series, thus avoiding the assumptions in the traditional model, and then prices the convertible bonds according to the traditional LSM pricing method. Quant GANs is a financial time series generation model proposed by Wiese et al. (2020) in 2019.

4.4.2. Empirical Results

The selection of samples is consistent with the traditional LSM pricing model, and the first-day pricing of 17 samples of continued bonds is carried out. Among them, the training sample of the Quant GANs model adopts the stock price data of 17 samples of bonds, and the sample interval is from 1 January 2010 to 20 January 2023, which is representative after several bull and bear markets.

In order to make a horizontal comparison with the pricing results of other models, this paper also adopts the Shanghai Pudong Development convertible bond for multi-node pricing. The pricing method is consistent with the above, and the node selection method is consistent with that described in Section 4.1.4.

(1) Generate sequence authenticity evaluation index

From the comparison of the following groups of figures, it can be intuitively seen that, except for the leverage effect, the generated sequence reproduces the style characteristics of the real sequence. In addition, it can be seen from Figures 14–17 that the Quant GANs model can generate a return sequence basically consistent with the real sequence.
In addition, it can be seen from Figures 14–17 that the Quant GANs model can generate a return sequence basically consistent with the real sequence.

Figure 14. Real sequence feature diagram.

Figure 15. Generated sequence feature diagram.

Figure 16. Real sequence rate of return.

Consistent with the above, the multi-node pricing range of Shanghai Pudong Development convertible bonds is from 20 January 2022 to 20 January 2023, with 12 nodes every 30 days. The price of convertible bonds at each node is estimated, respectively, and the stock price volatility and stock price are updated every 30 days. Pricing results are shown in Figure 18 below:

Figure 18. Multi-section pricing chart of Shanghai Pudong Development convertible bond.
Figure 16. Real sequence rate of return.

Figure 17. Generate a sequence yield graph.

(2) Pricing results of multi-node pricing within the duration

Consistent with the above, the multi-node pricing range of Shanghai Pudong Development convertible bonds is from 20 January 2022 to 20 January 2023, with 12 nodes every 30 days. The price of convertible bonds at each node is estimated, respectively, and the stock price volatility and stock price are updated every 30 days. Pricing results are shown in Figure 18 below:

Table 11 and Figure 19 show the pricing error rate of the four models, and the best pricing effect is the two deep learning models. From the perspective of the mean error rate, the pricing effect of WGAN (0.10%) is better than that of LSTM (0.13%), and the pricing effect of LSM’s improved model (−0.98%) is better than that of the traditional LSM model (−2.26%).
Table 11. Pricing comparison of each model in details.

<table>
<thead>
<tr>
<th>Date</th>
<th>Actual Price</th>
<th>LSM</th>
<th>Error Rate</th>
<th>LSTM</th>
<th>Error Rate</th>
<th>WGAN</th>
<th>Error Rate</th>
<th>LSM Improved</th>
<th>Error Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>2022/2/21</td>
<td>105.90</td>
<td>102.76</td>
<td>−2.96%</td>
<td>106.36</td>
<td>0.43%</td>
<td>106.21</td>
<td>0.29%</td>
<td>106.89</td>
<td>0.93%</td>
</tr>
<tr>
<td>2022/3/21</td>
<td>104.72</td>
<td>102.14</td>
<td>−2.46%</td>
<td>104.97</td>
<td>0.24%</td>
<td>105.01</td>
<td>0.28%</td>
<td>104.51</td>
<td>−0.19%</td>
</tr>
<tr>
<td>2022/4/20</td>
<td>104.83</td>
<td>102.30</td>
<td>−2.40%</td>
<td>105.10</td>
<td>0.26%</td>
<td>105.02</td>
<td>0.18%</td>
<td>105.07</td>
<td>0.22%</td>
</tr>
<tr>
<td>2022/5/20</td>
<td>105.34</td>
<td>102.48</td>
<td>−2.71%</td>
<td>105.42</td>
<td>0.08%</td>
<td>105.35</td>
<td>0.01%</td>
<td>104.30</td>
<td>−0.99%</td>
</tr>
<tr>
<td>2022/6/20</td>
<td>105.07</td>
<td>102.84</td>
<td>−2.11%</td>
<td>105.20</td>
<td>0.13%</td>
<td>105.26</td>
<td>0.19%</td>
<td>104.38</td>
<td>−0.65%</td>
</tr>
<tr>
<td>2022/7/20</td>
<td>105.45</td>
<td>103.13</td>
<td>−2.19%</td>
<td>105.52</td>
<td>0.07%</td>
<td>105.51</td>
<td>0.05%</td>
<td>104.81</td>
<td>−0.61%</td>
</tr>
<tr>
<td>2022/8/22</td>
<td>106.30</td>
<td>103.30</td>
<td>−2.81%</td>
<td>106.61</td>
<td>0.30%</td>
<td>106.64</td>
<td>0.32%</td>
<td>103.80</td>
<td>−2.35%</td>
</tr>
<tr>
<td>2022/9/20</td>
<td>106.35</td>
<td>103.56</td>
<td>−2.62%</td>
<td>106.19</td>
<td>−0.15%</td>
<td>106.46</td>
<td>0.10%</td>
<td>104.06</td>
<td>−2.15%</td>
</tr>
<tr>
<td>2022/10/20</td>
<td>106.36</td>
<td>103.87</td>
<td>−2.34%</td>
<td>106.13</td>
<td>−0.21%</td>
<td>106.29</td>
<td>−0.06%</td>
<td>104.12</td>
<td>−2.10%</td>
</tr>
<tr>
<td>2022/11/21</td>
<td>105.37</td>
<td>102.86</td>
<td>−2.38%</td>
<td>105.23</td>
<td>−0.13%</td>
<td>105.38</td>
<td>0.01%</td>
<td>103.03</td>
<td>−2.22%</td>
</tr>
<tr>
<td>2022/12/20</td>
<td>104.01</td>
<td>103.23</td>
<td>−0.75%</td>
<td>104.27</td>
<td>0.25%</td>
<td>104.00</td>
<td>−0.01%</td>
<td>103.55</td>
<td>−0.44%</td>
</tr>
<tr>
<td>2023/1/20</td>
<td>104.91</td>
<td>103.54</td>
<td>−1.31%</td>
<td>105.21</td>
<td>0.28%</td>
<td>104.70</td>
<td>−0.21%</td>
<td>103.75</td>
<td>−1.10%</td>
</tr>
<tr>
<td>Mean</td>
<td>105.38</td>
<td>104.35</td>
<td>−2.26%</td>
<td>105.52</td>
<td>0.13%</td>
<td>105.49</td>
<td>0.10%</td>
<td>104.35</td>
<td>−0.98%</td>
</tr>
</tbody>
</table>

Figure 19. Multi-node pricing error rate of each model.

4.5. Comparative Analysis of Pricing Effect of Each Model

Due to the different pricing methods of each model, the scope of application of the model is also different. The traditional LSM can price convertible bonds at any time point, but the efficiency is low and the presupposition is not in line with the reality. The pure data-driven pricing model can achieve the purpose of improving the pricing effect by adding multiple characteristic variables describing the real situation. However, due to the model input problem, the price of the first day of listing of convertible bonds cannot be predicted. The improved model of LSM uses a data-driven approach to replace the generation of stock price trend sequence, so as to avoid the assumptions in the traditional model and meet the requirements that the generation sequence should meet the reality, but there are still requirements on the sample data of the underlying stock price of the convertible bond.

To sum up, this paper takes Shanghai Pudong Development Co., LTD as an example and uses error rate, MAPE and RMSE to evaluate the pricing effect of each model, as shown in Table 12.

Table 12 summarizes the MAPE value of each pricing model and describes the overall pricing effect of the model. It can be seen that the two deep learning models still have the best overall pricing effect, and the pricing effect of WGAN (MAPE 0.14%, RMSE 0.1902) is better than that of LSTM (MAPE 0.21%, RMSE 0.2463), and the pricing effect of LSM’s
improved model (MAPE 1.17%, RMSE 1.4895) is better than that of the traditional LSM model (MAPE 2.26%, RMSE 2.4699).

<table>
<thead>
<tr>
<th>Model Name</th>
<th>MAPE</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>LSM</td>
<td>2.26%</td>
<td>2.4699</td>
</tr>
<tr>
<td>LSTM</td>
<td>0.21%</td>
<td>0.2463</td>
</tr>
<tr>
<td>WGAN</td>
<td>0.14%</td>
<td>0.1902</td>
</tr>
<tr>
<td>LSM improved</td>
<td>1.17%</td>
<td>1.4895</td>
</tr>
</tbody>
</table>

5. Conclusions

This paper adopts the traditional LSM pricing model, and selects 17 bank convertible bonds that still exist on 20 January 2023 as pricing samples. The pricing method is the first day of listing pricing. In addition, in order to compare with other models, the Shanghai Pudong Development convertible bonds with higher pricing accuracy on the first day of listing are selected as the pricing samples of multi-node pricing within the duration. The pricing range is from 20 January 2022 to 20 January 2023, and there are 12 pricing nodes every 30 days. In the two data-driven models, the daily data of all maturing and unmaturing bank convertible bonds from 1 January 2012 to 20 January 2023 were selected as the sample data, with a total of 25 samples, totaling 18,134 sample values. Among them, 23 were selected as training sets, and the remaining two were selected as test sets (Jiangbank convertible bonds and Shanghai Pudong Development convertible Bonds) to test the pricing power of the model. It is worth noting that the characteristic variables of each sample point are obtained from the macro environment, the fundamentals of listed companies, bond base, options, convertible bond trading data and clause design according to the existing literature.

The improved LSM model mainly uses the Quant GANs model proposed by previous researcher and scholars as the generation model of stock price series, thus avoiding the assumptions in the traditional model, and then prices the convertible bond according to the traditional LSM pricing model. The pricing sample is consistent with the traditional LSM pricing model. The first day of listing pricing of 17 sample bonds and the multi-node pricing of Shanghai Development convertible bonds within the duration were all conducted. Among them, the training sample of the Quant GANs model adopted the stock price data of 17 sample bonds. The sample interval was from 1 January 2010 to 20 January 2023, after several bull and bear markets; the samples were representative.

By sorting out the empirical results of each model, we reach the following conclusions:

1. The traditional LSM pricing model has a large error in the first-day pricing, indicating that the pricing function of this model needs to be further improved;
2. Among the four pricing models, the LSTM pricing model and WGAN pricing model have the best pricing effect. From the perspective of the MAPE index, the pricing effect of the WGAN pricing model (0.14%) is better than that of the LSTM pricing model (0.21%), and the pricing effect of the LSM improved model (1.17%) is better than that of the traditional LSM model (2.26%);
3. Applying the generative deep learning model GAN to the pricing of convertible bonds can avoid strict assumptions and significantly improve the pricing effect of the traditional model.

Though there are some limitations of the existing research in this paper, the prospects for the future research direction are put forward also:

1. The factors affecting the pricing of convertible bonds can be selected more accurately and the prediction accuracy of the model can be improved;
2. Improve the Quant GANs model in the future, so that the generated samples can better reproduce the statistical characteristics of real sequences;
(3) In order to avoid model failure on new data, we need to update the model in time and adjust and improve it according to new market conditions.

Author Contributions: Writing—original draft preparation, G.R. and T.M.; writing—review and editing, G.R., T.M. All authors have read and agreed to the published version of the manuscript.

Funding: National Social Science Fund Project (22BJY117).

Informed Consent Statement: Not applicable.

Data Availability Statement: Date are available on request.

Conflicts of Interest: The authors declare no conflict of interest.

References


Duan, Jin-Chuan. 1995. The GARCH Option Pricing Model. Mathematical Finance 1: 13–32. [CrossRef]


Feng, Yun, Bing-Hua Huang, and Yu Huang. 2016. Valuing resettable convertible bonds: Based on path decomposing. Finance Research Letters 19: 279–90. [CrossRef]


Hung, Mao-Wei, and Yan Wang, Jr. 2002. Pricing Convertible Bonds Subject to Default Risk. Derivations 10: 75–87. [CrossRef]


Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.