Hindawi
Mathematical Problems in Engineering
Volume 2022, Article ID 5239493, 10 pages
https://doi.org/10.1155/2022/5239493

Research Article

Deep Learning Model for Stock Excess Return Prediction Based on Nonlinear Random Matrix and Esg Factor

Tiantian Meng,[1,2] M. H. Yahya,[1] and Jingmin Chai[1,3]

[1]*School of Business & Economics, Universiti Putra Malaysia, Serdang 43400, Malaysia*
[2]*Shanghai University of Finance and Economics Zhejiang College, Jinhua 321000, Zhejiang, China*
[3]*Yuncheng University, Yuncheng 044000, Shanxi, China*

Correspondence should be addressed to Jingmin Chai; chaijingmin@ycu.edu.cn

Received 6 July 2022; Revised 23 August 2022; Accepted 25 August 2022; Published 14 September 2022

Academic Editor: Zaoli Yang

Aiming at the problem that the traditional model has low accuracy in describing stock excess return, in order to further analyze the change law of stock excess, based on the nonlinear random matrix and esg factor theory, the traditional learning model is analyzed, and the corresponding optimized deep learning model is obtained by introducing the single ring theorem and statistical data. Through the analysis and research of related indexes, the change rules of different indexes are obtained, and the optimization model is used to calculate and forecast the excess return of stocks. The results show that the statistics and spectral radius show typical local linear variation with the increase of eigenvalue. The corresponding statistics show a trend of gradual increase. The corresponding spectral radius has a decreasing variation law, and the two curves have obvious symmetry at some eigenvalues. It can be seen from the change curves under different factors that the change trend of the yield curve is mainly affected by the investment factor, while the change rule of the specific value of the yield curve is controlled by the profit factor. This shows that the two factors have the same influence on the stock excess return. The influence of optimized deep learning model on stock excess index has typical linear characteristics, which can be divided into linear increase and linear decline according to different change rules. The basic type has the greatest influence, while the corresponding pattern analysis type has the least influence. Finally, the method of experimental verification is used to verify the stock excess data, and the results show that the optimized deep learning model can better characterize the experimental results. Therefore, the optimized deep learning model based on nonlinear random matrix and esg factor can carry out targeted analysis of different types of stock returns, thus improving research ideas and calculation methods for the application of deep learning model in different fields.

1. Introduction

Nonlinear random matrix has obvious applications in different fields: optical phenomena [1], feature learning [2], structural evolution [3], finite element analysis [4], noise monitoring [5], critical value analysis [6], and so on. In view of the problems existing in graph analysis, model analysis method was adopted to explore the graph based on relevant theories of nonlinear random matrix [7]. First, the graph data with different features were integrated to obtain different kinds of graphs. Then, the data analysis of graphs with different features was carried out to obtain the optimized model indicators. Finally, the accuracy of the model was verified by experimental data. In order to further explore the change process of perception structure in maglev train, nonlinear random matrix was adopted to modify the deep learning model [8]. Thus, an optimization model that can reflect the changes of magnetic levitation perception structure can be obtained, through which the targeted analysis of the perception structure can be carried out. Relevant experiments were used to verify the superiority of the model and provide theoretical support for the application of the optimization model in other fields.

The above studies mainly explore the application of nonlinear random matrix in algorithms and other fields, and there were few studies on stock excess and other aspects. In order to carry out the application of deep learning model in stock excess returns, based on nonlinear random matrix and esg factor theory, the corresponding optimized deep learning model was obtained by modifying the traditional deep learning model. The accuracy of the model was verified through experiments. The research shows that the model can carry out accurate analysis and research on stock excess returns and other aspects. The model was used to predict stocks, and the research results can provide support for the application of deep learning in different fields.

2. Related Theories of Nonlinear Random Matrices

A nonlinear random matrix is a matrix whose elements are random variables in some probability space [9, 10]. To solve the problem of stock excess returns in nonlinear programming, an optimized random matrix can be established based on historical data. This matrix not only has statistical characteristics and obeys the relevant laws of random matrix theory but also can effectively reflect the operating state of the system. It is free from the limitation of simplification and hypothesis of traditional optimization method based on physical model and achieves the purpose of improving the accuracy and speed of solution. As a universal big data analysis method, the random matrix theory has many excellent characteristics of the empirical spectrum distribution function, such as the single ring theorem and Marchenko-Pastur law. Random matrix theory plays an increasingly important role in the field of big data analysis in recent years. The theory obtains some distribution characteristics of data by studying the eigenvectors and eigenvalues of random matrix. According to these theorems, the statistical characteristics of the data can be obtained [11, 12]. By analyzing and comparing these statistical characteristics, the results we need can be obtained. Random matrices should actually be divided into row random matrices and column random matrices. Row random matrix refers to the row sum of square matrix is equal to 1. And a column random matrix is a nonnegative matrix whose column sum is equal to 1. So a nonnegative matrix that has both a row and a column sum of 1 is a doubly random matrix, and the identity matrix is a doubly random matrix.

2.1. Monocyclic Theorem. The random matrix A is a complex matrix and $\overrightarrow{a}_i$ is an m-dimensional row vector. The matrix A is first transformed into the standard Hermitian matrix $\widetilde{A}$.

$$\widetilde{a}_{ij} = \overline{\widetilde{a}}_i + \left(a_{ij} - \overline{\widetilde{a}}_i\right) \times \frac{\sigma(\widetilde{a}_i)}{\sigma(\overrightarrow{a}_i)}. \tag{1}$$

where $\overline{a}_i$ and $\sigma(\overrightarrow{a}_i)$ are the mean and standard deviation of row vectors $\overrightarrow{a}_i$ respectively; the mean $\overline{\widetilde{a}}_i = 0$ and standard deviation $\sigma(\overrightarrow{a}_i) = 1$ of the transformed row vectors $\widetilde{A}$. In order to obtain the eigenvalues of the random matrix, the singular value equivalent square matrix $\widetilde{A}$ is transformed u$\widetilde{A}$.

$$\widetilde{A}u = \sqrt{\widetilde{A}\widetilde{A}^H}U. \tag{2}$$

In the formula, H is the conjugate transpose, U is the Haar unitary matrix. For L singular value equivalent random matrices u$\widetilde{A}$, the matrix product Z is computed.

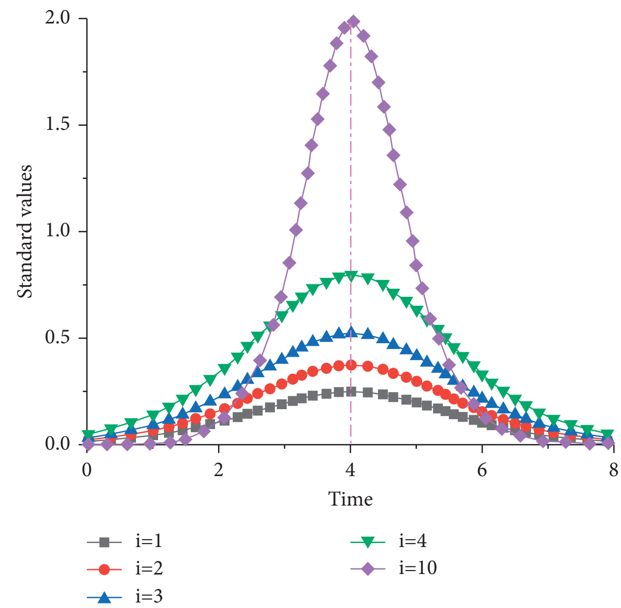

Figure 1: Standard matrix values of different row vectors.

$$Z = \prod_{i=1}^{L} \widetilde{A}u, i. \tag{3}$$

The matrix product Z is processed and converted to the standard matrix product $\widetilde{Z}$.

$$\widetilde{z}_i = \frac{\overrightarrow{z}_i}{\sqrt{N}\sigma\left(\overrightarrow{z}_i\right)}, \quad (i = 1, 2, \ldots, N), \tag{4}$$

where $\overrightarrow{z}_i$ is the row vector and $\sigma(\overrightarrow{z}_i)$ is the standard deviation of the i-th row vector of the matrix product Z.

The standard matrix values of different row vectors are different. In order to quantitatively analyze the influence of row vectors on the standard matrix values, the standard matrix of different row vectors is drawn by calculating relevant data, as shown in Figure 1. It can be seen from the figure that the curves between different row vectors show a symmetric change characteristic as a whole, and the axis of symmetry is 4. With the gradual increase of time, the standard value curves of corresponding different line vectors increase slowly at first, then increase rapidly to the maximum value, when it exceeds the maximum value, then decline rapidly, and finally slowly approach zero. It is worth explaining that the overall change trend of the curves of different line vectors is basically the same. Through analysis, it can be seen that when the variation range of row vector is 1–4, the corresponding curve standard value shows a gradually increasing trend with the gradual increase of row vector. And the lowest value was nearly three times the highest. When the curve corresponding to the row vector exceeds 4, the corresponding standard value data shows a trend of rapid increase. This shows that the increase of row vector will make the corresponding data show a significant improvement.

When N, M$\longrightarrow\infty$ and $c = N/M$ satisfies $c(0, 1)$, the empirical spectral distribution of $\widetilde{Z}$ is almost distributed on a single ring, and its probability density function is as follows:

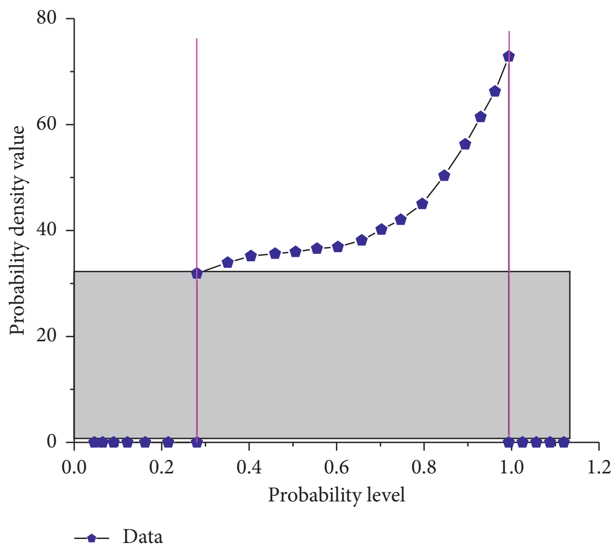

Figure 2: Probability exponential function change diagram.

$$f(\lambda_i)=\begin{cases}\dfrac{1}{\pi cL}\left|\lambda_i\right|^{2/L-2}, & (1-c)^{L/2}\le\left|\lambda_i\right|\le 1,\\ 0, & \text{other}.\end{cases}\tag{5}$$

Through the above analysis and calculation of the corresponding formula, the change relationship between the probability index and the corresponding probability curve density can be obtained, as shown in Figure 2. It can be seen from the correlation calculation curve that the corresponding curve has typical piecewise characteristics. When it is less than the critical value of probability index, the corresponding curve of the corresponding probability density value is parallel to the x axis, but when it exceeds 1, the corresponding curve is still zero. Only when it is between the critical value of the probability exponential curve and 1, the corresponding probability density value shows the characteristics of fluctuation: that is, the trend of slow increase at first and then rapid increase. The slope of the corresponding curve shows a trend of gradual increase, which indicates that different probability indexes will lead to certain deviation of the probability density value of the nonlinear random matrix, resulting in different results. And the corresponding curve variation range basically stays between 36 and 76.

Through the quantitative analysis of nonlinear random matrix, it can be seen that the change of nonlinear random matrix has typical nonlinear characteristics. In order to analyze the characteristic root distribution under the single ring theorem, the variation range of characteristic roots under different parameters was obtained through the above calculation, as shown in Figure 3. It can be seen from Figure 3 that both the real part and the imaginary part of features have influences on the solving parameters of the feature function. The distribution of the corresponding data is regular on the whole, and the change of the characteristic root tends to be circular on the whole. However, the feature roots on the right of the critical curve are mainly concentrated within the influence range of the real part, which

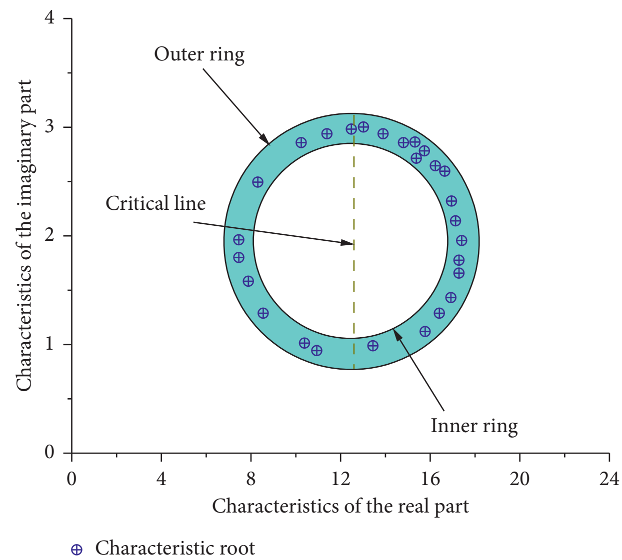

Figure 3: Distribution diagram of characteristic roots under single ring theorem.

indicates that the larger the real part of the feature is, the larger the variation range of the corresponding feature roots is, and the larger the corresponding feature imaginary part is, the more data corresponding to the critical value will be.

2.2. Nonlinear Random Matrix Statistics. In nonlinear random matrices, linear eigenvalue statistics can effectively reflect the distribution characteristics of eigenvalues [13, 14].

Through the above single ring quantification, we can see that the statistical data of nonlinear random matrix have a certain correlation with specific calculation methods [15, 16]. In order to further analyze the calculation process of the nonlinear random matrix, we draw the specific flow of the corresponding nonlinear random matrix, as shown in Figure 4. The calculation process in the figure that first, the corresponding nonlinear original data need to be imported into the corresponding collection module, and quantitative data collection and analysis can be realized through the extraction of the collection module. Then, the nonlinear random matrix is used to calculate the feature weights of the relevant data. Then, the above feature weights are imported into the threshold value of the random function, and the calculation threshold value of the corresponding feature weights is obtained by setting the corresponding random function domain. On this basis, it is necessary to determine the corresponding characteristic function, which is mainly based on the selection of certain characteristic function according to different trend and range of change. The eigenvector of the nonlinear random matrix is calculated by the eigenfunction, and the corresponding value of the nonlinear random matrix is finally obtained, so as to obtain the correlation coefficient. The characteristic index of nonlinear random matrix is solved by further analysis of correlation coefficient. Finally, by setting the relevant optimization algorithm, we can eliminate the relevant error points, so as to make the results have higher accuracy, and

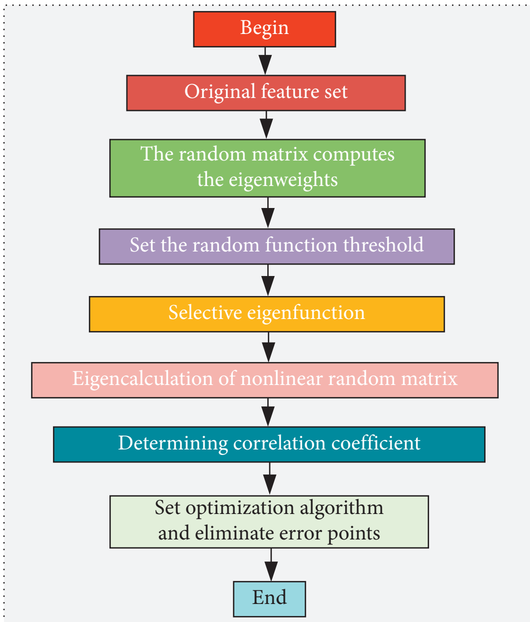

Figure 4: Flow chart of nonlinear random matrix calculation.

finally get the relevant data output. The statistical data function of the corresponding nonlinear random matrix is shown as follows:

$$\varsigma_N(\phi) = \sum_{i=1}^{N} \phi(\lambda_i), \tag{6}$$

where $\phi(\lambda_i)$ is the test function.

As a specific object of nonlinear random matrix, the mean spectral radius describes the average radius distribution of all eigenvalues of matrix A on the complex plane. It can effectively reflect the statistical characteristics of matrix eigenvalue distribution, and the corresponding spectral radius calculation formula is shown as follows:

$$r_{\text{MSR}} = \frac{1}{N} \sum_{i=1}^{N} |\lambda_i|. \tag{7}$$

From the above analysis, it can be seen that the index of eigenvalue will have a certain influence on the statistics and spectral radius [17, 18]. In order to analyze this influence law, relevant data in the above single ring theorem were imported into the calculation formula for solving, so as to obtain the change curve of statistics and spectral radius. The specific change curve is shown in Figure 5. It can be seen from the curves in the figure that the calculated curves between two different statistical radii and spectral radii have typical nonlinear fluctuation characteristics. To be specific, it can be seen from the data change curve of statistics that the curve first presents a linear change characteristic of gradual increase. When its characteristic value reaches 8, there will be a typical rapid rise stage. The reason for the above situation in this stage is that the jump of characteristic points leads to a certain increase of corresponding statistics. After the increase, the curve shows the same linear increase, and the corresponding slope of the curve is basically the same as

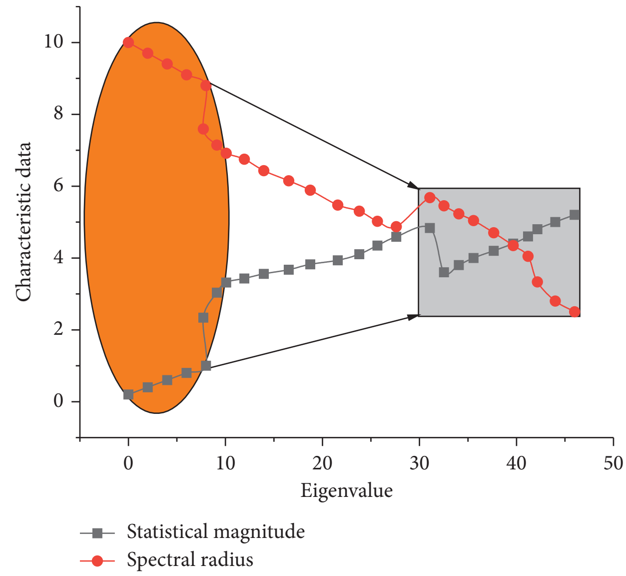

Figure 5: The variation curve of the statistic and spectral radius.

that of the first stage. And when the corresponding eigenvalue reaches 30, there is a sudden drop phenomenon, which is contrary to the above increase, and there is a certain decline. Subsequently, with the further increase of the eigenvalue, the corresponding statistic shows a trend of gradual increase, and the slope of the corresponding curve is also approximately constant, indicating that the curve at this stage is a nonlinear change. From the change of spectral radius, we can see that it shows a gradual downward trend as a whole, with a fluctuation trend of linear decline at first, then rapid decline, and then linear decline. The two curves are typically symmetrical in the first stage. Through the above analysis, it can be seen that the change of characteristic values will have an impact on the characteristic data corresponding to the statistics and spectral radius. And this effect has obvious regularity, indicating that these two analysis indicators can be used to analyze the model.

2.3. Esg Factors. Esg stands for Environmental, Social, and Corporate Governance. Different from traditional financial performance indicators, esg factors focus more on assessing the sustainability of stocks from environmental, social, and corporate governance perspectives [19, 20]. Ethical investment, responsible investment, green finance, and other concepts are commonly used to refer to esg investment concepts. Enterprise idiosyncratic risk, also known as nonsystematic risk, is different from systematic risk. Idiosyncratic risk is a risk closely related to internal factors such as the company's own financial status and governance structure and external factors such as industry development and social dynamics. For the measurement of idiosyncratic risk, the most direct and most extensive method is to estimate the residual by calculating the standard deviation of the three-factor model.

The three dimensions of esg factor focus on different responsibility themes, and each theme corresponds to several specific indicators to form an organic whole. From

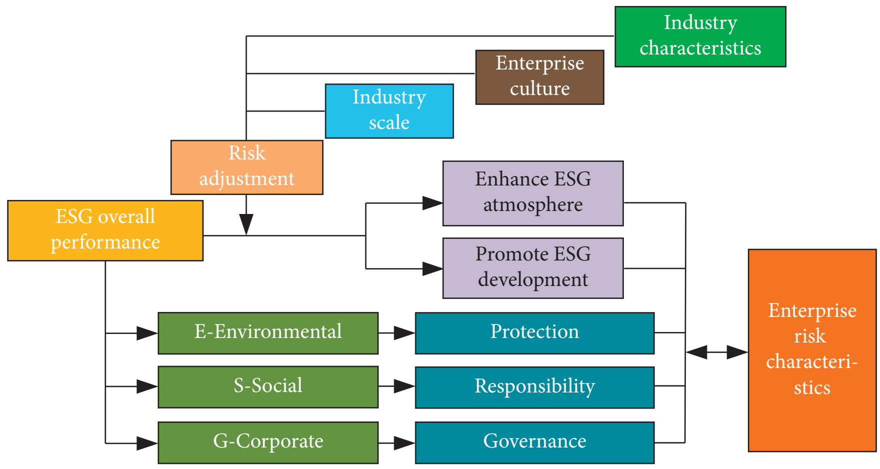

Figure 6: Esg factor analysis flowchart.

the perspective of environment, it mainly focuses on the green investment in the production and operation activities of enterprises, the recycling and sustainable utilization of natural resources and energy, and the treatment of harmful wastes [21, 22]. From the perspective of society, we mainly examine the expectations and demands of relevant stakeholders and pay attention to whether the balance and coordination between stakeholders of enterprises can be achieved. The angle of corporate governance mainly includes stock returns and ownership structure.

The change of esg factor will have a certain influence on the specific value of nonlinear random matrix. In order to further analyze the change rule of esg factor, the corresponding analysis process of esg factor is first drawn, as shown in Figure 6. The changes of esg have typical block characteristics. First, the comprehensive performance of esg can be divided into three parts: environmental performance, social performance, and corporate performance. According to the different research contents of the three parts, they correspond to three modules of environmental protection, social responsibility, and corporate governance, respectively. According to the different research contents of esg factor in the calculation process, it can be divided into two modules: promoting esg atmosphere and promoting esg development. Together with the above three modules, they constitute the changing characteristics of enterprise risk. In the process of dividing esg factors, it is necessary to consider the influence of different properties such as risk regulation, industry scale, enterprise culture, and industry characteristics on the comprehensive performance of esg. Therefore, through the above analysis flowchart, we can also see that esg has a strong integrity and generality, which can specifically reflect the characteristics of relevant changes with nonlinear random matrix. The original sample data are screened as follows: (1) due to the particularity of financial and insurance industries and their own business and financial statement standards, samples of these industries are firstly excluded; (2) exclude the samples with discontinuous ESG score data and missing annual report financial data; (3) samples with stock symbols ST and PT were excluded.

Risk measurement of esg factor: for the measurement of esg factor, at present, the most direct and most extensive method is to estimate the standard deviation of residual term by calculating the three-factor model. However, many studies copying the three factors have not achieved ideal results. In order to better analyze the calculation process of esg factor, an optimized three-factor model is proposed which is closer to the actual situation of stock market. This model has two obvious characteristics: (1) excluding the influence of stocks accounting for less than 30% when constructing esg factor; (2) esg factor is based on profit-to-market ratio.

The stock esg factor can be estimated by the standard deviation of the residual term of the optimized three-factor model, and the specific calculation method is shown as follows:

$$R_{it} = R_{Ft} + a_i + b_i\left(R_{Mt} - R_{Ft}\right) + s_i \mathrm{SMB}_t + h_i \mathrm{HML}_t + \varepsilon_{\mathrm{it}}, \tag{8}$$

where R_{it} is the stock monthly return rate; R_{Ft} is the risk-free interest rate; R_{Mt} is the market rate of return; $\mathrm{SMB}_{\mathrm{t}}$ is the company size factor; HML_t is the book-to-market ratio factor; b_{i} is the market regression coefficient; s_i is scale regression coefficient; h_i is the ratio factor regression coefficient; a_{i} is the model regression constant term; $\varepsilon_{\mathrm{it}}$ is the regression residual.

The idiosyncratic risk of the stock in year y can be described as:

$$\mathrm{IRISK}_{iy} = \mathrm{std}\left(\varepsilon_{iy}\right). \tag{9}$$

In order to complete the robustness test of the research results, the esg factor will be further estimated using a five-factor model, as shown below:

$$\begin{aligned} R_{it} = {} & R_{Ft} + a_i + b_i\left(R_{Mt} - R_{Ft}\right) + s_i \mathrm{SMB}_t + h_i \mathrm{HML}_t \\ & + \varepsilon_{it} + r_i \mathrm{RMW}_t + c_i \mathrm{CMA}_t, \end{aligned} \tag{10}$$

where RMW_t is the profit factor; is the investment factor; r_i is profit factor coefficient; c_{i} is the investment factor coefficient.

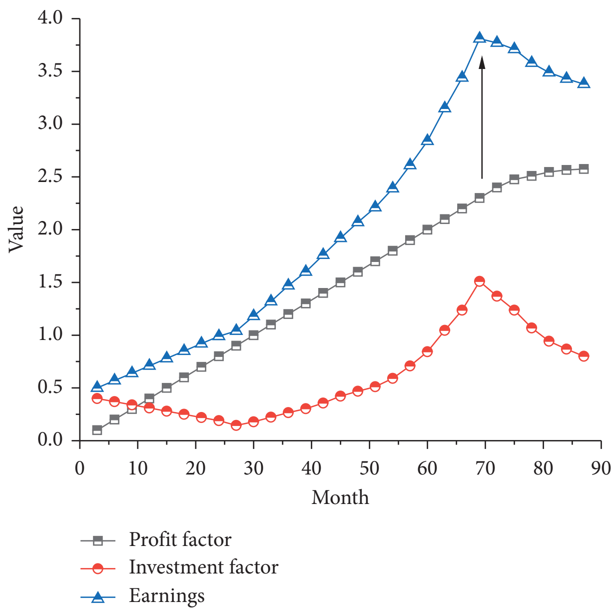

Figure 7: Yield curve.

Through the changes of esg factor, we can divide it into profit factor, investment factor, and corresponding related parameters. Through the changes of profit factor and investment factor under the action of different time, we draw the corresponding curve of return change, as shown in Figure 7. It can be seen from Figure 7 that the changes among different factors have certain volatility, and the changes of income factor and investment factor will lead to the increase and decline of income curve to different degrees. It can be seen from the profit factor that the corresponding curve shows an approximate linear increase trend with the gradual increase of corresponding time. When the corresponding time exceeds about 70, the increment of the corresponding curve gradually decreases, resulting in the slope of the curve gradually tending to equilibrium. This shows that with the gradual increase of time, the corresponding profit factor gradually tends to be stable, indicating that the longer the time, the corresponding profit factor increment is small. As can be seen from the investment factor: with the gradual increase of time, the curve of the investment factor first shows a linear downward trend, and then when it reaches the lowest point, with the further increase of time, the curve shows a gradual upward trend. When the time is around 70, the investment factor reaches its highest value and then continues to decline. This shows that the influence of time on investment factor shows typical fluctuation characteristics. The corresponding income change curve can be calculated by investment factor and profit factor. It can be seen from the income change curve that with the increase of time, the curve first presents an approximate linear increase trend, and when it reaches the highest point, it gradually decreases. It can be seen from the curve changes that its change pattern is basically consistent with that of the investment factor, which indicates that the investment factor mainly affects the change pattern of the income curve. The corresponding growth value is influenced by the profit factor. Therefore, we can see that both the investment factor and the profit factor have a great impact on the yield curve. Therefore, we need to comprehensively consider the influence of the two factors in the actual calculation process, so as to get the highest stock excess return.

3. Deep Learning Model Based on Nonlinear Random Matrix and Esg Factor

Driven by artificial intelligence, the deep learning model has entered a brand development platform, which requires more energy for further research and exploration in the future. Based on artificial neural network, the deep learning model captures useful key information in complex and changeable network data, constantly improves the learning mode, automatically learns key features from the data, and uses the features to carry out independent recognition and classification. Since the deep learning model considers various data calculation processes and calculation methods, it can effectively solve the problem of poor feature performance through its own active feature learning mode, thus significantly improving the recognition and prediction ability of the classification system. Deep learning trains each network layer, collects feature changes in each layer together and then maps these features to higher-dimensional space to fully extract features. The most fundamental features of data can be learned in high dimensional space, and the feature will not lose its obvious characteristics due to the large gradient, which solves the problem of multilayer gradient dispersion. The deep learning model is to learn the internal rules and representation levels of sample data, and the information obtained in the learning process is of great help to the interpretation of data such as text, image, and sound. The ultimate goal is for machines to be able to learn analytically, like humans, and to recognize data such as text, images, and sound. Deep learning is a complex machine learning algorithm that has achieved far more results in speech and image recognition than previous related technologies.

The deep learning model starts from the input layer, and every time features are extracted by convolution with the input vector through the convolution check, this feature will be taken as the input of the next layer. After many times of convolution and pooling, features will change from one-sided to comprehensive, and finally abstract and comprehensive information will be obtained. Deep learning models can improve their results with more data or better algorithms. For some applications, deep learning works better on large data sets than other machine learning methods. Deep learning is more suitable for unlabeled data, so it is not limited to the field of natural language processing, which is dominated by entity recognition. Deep learning models can be divided into convolution layer deep learning model, pooling layer deep learning model and full connection layer deep learning model according to different calculation characteristics and calculation methods. The specific basis and content of division are as follows:

3.1. Convolution Layer. The convolution layer in the deep learning model extracts the characteristic indicators of calculated data by using convolution operations. Convolutional neural network is a deep neural network with convolutional structure. It is a multilayer supervised learning neural network. Feature extraction is realized in two hidden layers, convolution layer, and pooling layer. The convolution operation in the convolution layer is particularly important. The important information of feature graph can be obtained by convolution operation, which provides great help for classification work. It is worth explaining that the deep learning model can enhance the data feature point index and reduce the data calculation process to reduce the noise.

$$y^{l(i,j)} = \sum_{j=0}^{n-1} W_i^{l(j)} x^{lj}, \tag{11}$$

where $W_i^{l(j)}$ is the weight function; x^{lj} is the local convolution region; n is the width of convolution kernel; $y^{l(i,j)}$ is the output result.

Based on the above analysis, the convolutional layer calculation formula under the deep learning model is obtained, and the corresponding calculation function under the nonlinear random matrix and esg factor is shown as follows:

$$\begin{aligned} a^{l(i,j)} &= f\left(y^{l(i,j)}\right) \\ &= \max\left\{0, y^{l(i,j)}, z, R_{\text{it}}\right\}. \end{aligned} \tag{12}$$

3.2. Pooling Layer. Pooling layer not only reduces network parameters in the full connection layer by reducing the size of the feature matrix but also connects neurons with different regions in the upper layer. Pooling layer is important for the computation of deep learning model. Pooling layer can reduce the structural complexity and improve the computation speed. At the same time, the optimal features of the data can be further extracted. Pooling operations include average pooling and maximum pooling, and the pooling layer needs to provide the filter size and step size. The accuracy and superiority of data can be improved by using pooling layer to analyze data.

$$p^{l(i,j)} = \max_{(j-1)n+1\leq t\leq jn} \left\{a^{l(i,j)}, z, R_{\text{it}}\right\}. \tag{13}$$

3.3. Full Connection Layer. After multilayer convolution and splicing, abstract features in network data have been extracted into specific and obvious features, completing an effective data feature extraction process. First, the full connection layer plays the role of "classifier" in the whole convolutional neural network. Second, due to the large number of parameters at the full connection layer, some recent network models with excellent performance use global average pooling to integrate the depth features learned. Loss function is used as network objective function to guide the learning process. Finally, the full connection layer can act as a "firewall" during the migration of model representation capabilities.

$$y^{l+(j)} = \sum_{i=1}^{n} W_{ij}^{l(j)} x^{lj} + b_j^l, \tag{14}$$

where bl/j is the bias quantity of neuron; $y^{l+(j)}$ is the corresponding output value.

$$a^{l+(j)} = \frac{e^{yl+(j)}}{\sum_j e^{yl+(j)}}. \tag{15}$$

Through the above analysis, we can see that nonlinear random matrix and esg factor have obvious characteristics of nonlinear change. In order to further apply it to the deep learning model, the exponential substitution method is adopted to modify and analyze the deep learning model. The nonlinear random matrix and esg factor are directly brought into the different computational layers of the deep learning model. The deep learning model considering nonlinear random matrix, the deep learning model considering esg factor and the comprehensive model considering two factors are obtained. In order to further analyze the influence of different models on stock forecast income, the variation curves of depth models with the number of iterations under different influencing factors are drawn, as shown in Figure 8.

It can be seen from the figure that the original deep learning model shows a slow increasing trend with the increase of iterations, and the corresponding nonlinear characteristics are not obvious and the variation range is small. Considering the nonlinear random matrix, the corresponding curve of the deep learning model shows an obvious trend of fluctuation as the number of iterations increases: That is, it increases slowly at first and then tends to be flat. Its maximum value is basically consistent with the traditional deep learning model. By considering the change curve corresponding to the deep learning model under the esg factor, it can be seen that the change factor of its initial value is large. However, as the number of iterations increases, the curve gradually tends to be gentle, and its maximum value is higher than that of traditional depth models, but the characteristics of deep learning indicators are not obvious. By comprehensively considering the curve of the deep learning model under the action of the two factors, it can be seen that the curve has typical segmentation characteristics: That is, the input prediction value at the initial value is high, which shows a linear increase trend with the increase of time. When the current iteration number is about 80, the slope of the curve tends to be gentle, which still shows a linear feature. It shows that the optimization model has obvious linear characteristics and can be used to simulate deep learning model. Therefore, it can be seen from the optimization deep learning model under different factors that the optimization depth model under the action of nonlinear random matrix and esg factor has good segmentation characteristics and can describe the typical change characteristics of stock return indicators.

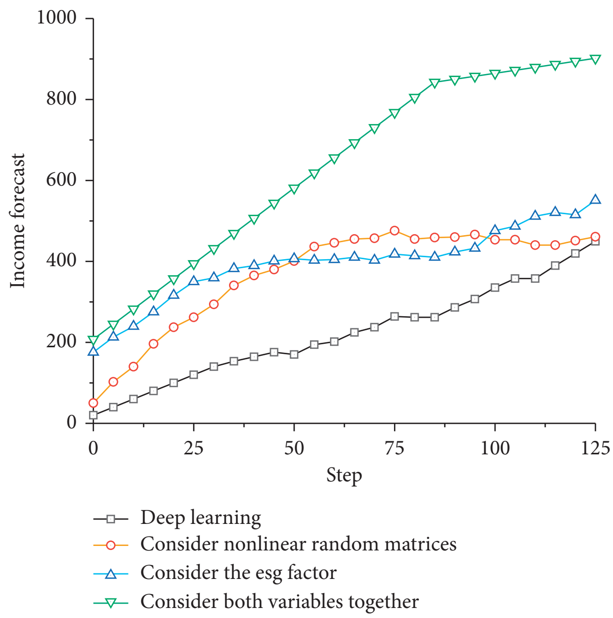

FIGURE 8: Deep learning model under different factors.

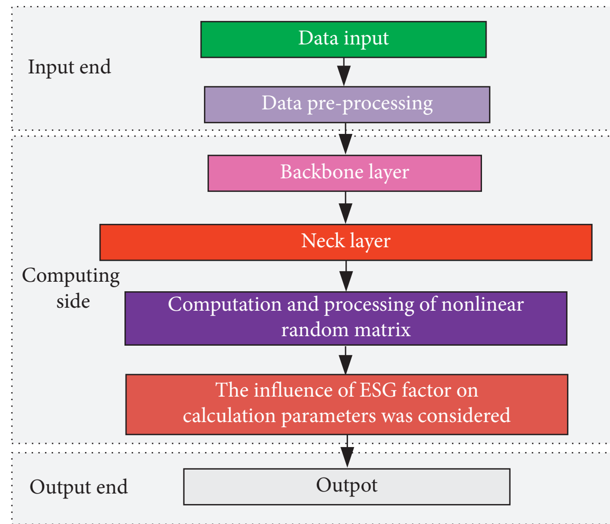

FIGURE 9: Calculation process of stock excess return based on optimized deep learning model.

4. Application of Optimized Deep Learning Model in Stock Excess Return Prediction

By introducing nonlinear random matrix and esg factor to modify the traditional deep learning model, the optimized deep learning model is obtained, which can describe and analyze different types of indicators [23, 24]. In order to analyze the calculation of stock excess return prediction by the optimized learning model, the corresponding calculation process is drawn, as shown in Figure 9. As can be seen from the figure, the calculation process of stock excess return based on optimized deep learning model can be divided into input module, calculation module, and output module according to different calculation contents. The specific calculation process is as follows: first, the corresponding data of stock excess returns are imported into the input terminal, and the input terminal preprocesses the data, so as to extract the corresponding data change characteristics. After preprocessing at the input end, import the stock excess data into the calculation module. At first, backbone computing layer is used for hierarchical processing of the data proposed, and then neck computing layer is used for further iterative calculation of the processed data. By analyzing the influence of the calculation parameters of the nonlinear random matrix and esg factor, the more accurate calculation data of stock excess return are obtained, and then the calculated data are imported into the output layer for analysis, and finally the corresponding results are derived.

Relevant studies show that the index of stock excess return will have a great impact on the calculation of deep learning model, and the different selection of model parameters will lead to the difference between the calculation process and calculation formula, which ultimately makes the calculation results cannot fully reflect the real situation of data. By comprehensively considering the influence of different factors, this paper chooses five different indexes to

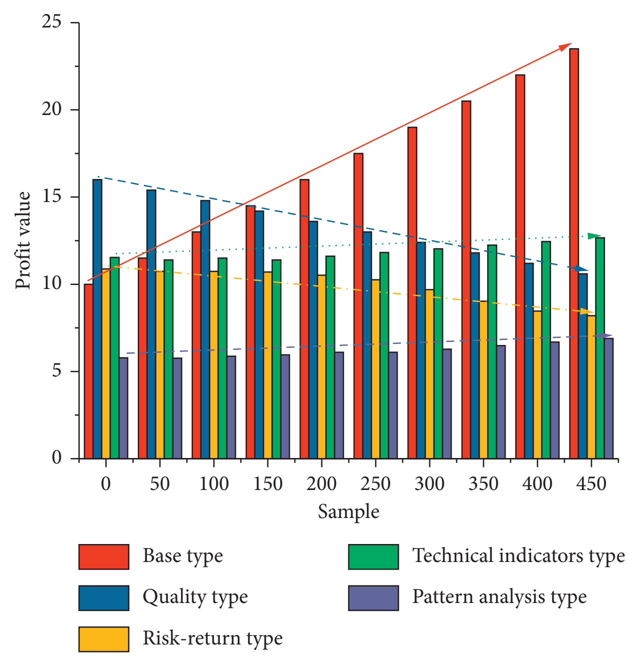

FIGURE 10: Optimized deep learning model for stock excess return calculation bar chart.

characterize stock returns, including basic type, quality type, risk-return type, technical index type, and pattern analysis type.

By using the deep learning optimization algorithm of the above analysis to analyze the calculation process of stock excess return, the histogram changes corresponding to calculated return indicators under different types are obtained, as shown in Figure 10. According to the change trends of different indicators in the figure, we can see that the optimized deep learning model has different influences on the calculation of different indicators. The details can be shown as follows: in the basic type, with the gradual increase of the number of samples, the corresponding profit value shows an obvious linear increase trend, and its increase range is large, with the maximum value 2.5 times of the corresponding minimum value. As can be seen from the variation trend of the quality type index, with the increase of stock excess return samples, the corresponding quality return value shows a trend of gradual decline. This indicates that the increase of sample number will inhibit the increase of quality type index. The corresponding risk-return types also show a linear downward trend. However, its decreasing range is small compared with the change of quality type, while the change of technical index type can be seen that the data show an increasing trend, but the increasing range is small. The corresponding pattern analysis types also show a gradually increasing trend with the increase of sample size. But its increase and total amount are small. Therefore, we can see that the basic type, as the highest index of returns, has the greatest impact on the excess returns of stocks, while the corresponding pattern analysis type has the least impact.

5. Discussion

Aiming at the problem that the traditional deep learning model fails to better analyze and describe the stock excess return, in order to further improve the accuracy of the analysis of stock excess return, so as to improve the relevant model and calculation method for the research of stock excess return. Therefore, the single ring theorem and esg factor in nonlinear random matrix are respectively introduced into the corresponding deep learning model to obtain the corresponding analysis optimization model. Based on the above analysis, it can be seen that the model can carry out targeted analysis of deep learning indicators under different sample numbers.

In order to further analyze and verify the influence of modular deep learning model on stock excess returns, we draw the prediction curve of the deep learning model, as shown in Figure 11. From the change curve in the figure, we can see that the original data show a fluctuating change trend, with a typical three-stage change rule: in the first stage, the curve shows a smooth increase, and with the improvement of time, the corresponding forecast data and curve show a gradually increasing trend. When the corresponding time exceeds 4.5, the curve drops rapidly, and the corresponding time at this stage is small and the variation range is large. When the time exceeds 5, the curve enters the third stage, at which the curve becomes an approximately constant change rule with the gradual increase of time, indicating that the time in this stage has almost no influence on the predicted value. However, the original deep learning model can better describe the corresponding change patterns of stock excess return data, but the traditional model cannot carry out more accurate analysis and description on specific data. It shows that the traditional deep learning model cannot well analyze and characterize the fluctuations of stock excess returns. The optimized deep learning model can not only accurately describe the changing trend of stock returns but also better describe the predicted value at the key nodes. Therefore, it can be seen from the curve changes that

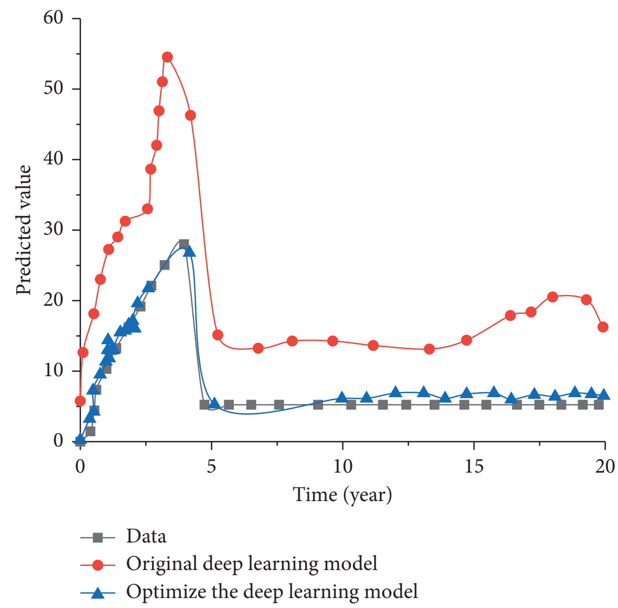

Figure 11: Prediction graph of deep learning model.

the optimized deep learning model can better describe the excess returns of stocks. At the same time, the accuracy of the optimized deep learning prediction model based on nonlinear random matrix and esg factor is also demonstrated.

6. Conclusion

(1) It can be seen from the standard matrix values of different row vectors that, with the gradual increase of row vectors, the corresponding nonlinear matrix data shows an obvious increase trend. This shows that the improvement of row vector can promote the improvement of the corresponding data of nonlinear random matrix to a certain extent. It is worth explaining that the data under different parameters have typical symmetry characteristics.

(2) The probability data in the probability exponential function change graph has obvious segmentation characteristics. With the increase of the probability index of the curve, the corresponding probability density has obvious nonlinear characteristics. The main reason for this is that the corresponding calculation formula in the single ring theorem has different piecewise functions according to different analysis contents, resulting in different corresponding curves.

(3) The deep learning model under the influence of different factors has different influence rules on stock excess data. The deep learning model corresponding to the nonlinear random matrix and esg factor can only analyze and describe part of the variation characteristics of stock excess return data. The overall change trend of data cannot be well analyzed,

while the optimized deep learning model considering the two factors can better analyze and describe the excess returns of stocks.

Data Availability

The data used to support the findings of this study can be obtained from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] A. C. Mitus, M. Saphiannikova, W. Radosz, V. Toshchevikov, and G. Pawlik, "Modeling of nonlinear optical phenomena in host-guest systems using bond fluctuation Monte Carlo model: a review," *Materials*, vol. 14, no. 6, pp. 1454–1468, 2021.

[2] T. Ouyang, "Feature learning for stacked ELM via low-rank matrix factorization," *Neurocomputing*, vol. 448, no. 75, pp. 725–736, 2021.

[3] G. D. Pasparakis, V. C. Fragkoulis, and M. Beer, "Harmonic wavelets based response evolutionary power spectrum determination of linear and nonlinear structural systems with singular matrices," *Mechanical Systems and Signal Processing*, vol. 149, no. 10, 108121 pages, 2021.

[4] H. Qiu, L. Weng, B. Yu, and Y Zhang, "Finite-time anti-synchronization and parameters identification of nonlinear coupled multiple-weight markovian switching complex networks with stochastic perturbations," *Advances in Mathematical Physics*, vol. 2020, no. 12, pp. 1–15, 2020.

[5] M. Wang and N. Xia, "Estimation of high-dimensional integrated covariance matrix based on noisy high-frequency data with multiple observations," *Statistics & Probability Letters*, vol. 170, no. 12, 109116 pages, 2021.

[6] Z. Wang and E. Fan, "Critical edge behavior in the singularly perturbed Pollaczek–Jacobi type unitary ensemble [J]," *Random Matrices: Theory and Applications*, vol. 132, no. 15, pp. 225–246, 2021.

[7] M. Yang and S. Xu, "A novel patch-based nonlinear matrix completion algorithm for image analysis through convolutional neural network," *Neurocomputing*, vol. 389, no. 10, pp. 56–82, 2020.

[8] Y. Zhao, F. Peng, L. Ren, G. Lin, and J Xu, "A levitation condition awareness architecture for low-speed maglev train based on data-driven random matrix analysis," *IEEE Access*, vol. 8, no. 12, pp. 176575–176587, 2020.

[9] A. Arumugam, Y. Liu, and S. Rathinasamy, "Distributed event-triggered nonfragile H∞ control for networked nonlinear systems with energy constraints and redundant channels: observer-based case," *International Journal of Robust and Nonlinear Control*, vol. 30, no. 17, pp. 369–378, 2020.

[10] A. Babkin, "Incorporating side information into robust matrix factorization with bayesian quantile regression," *Statistics & Probability Letters*, vol. 165, no. 15, pp. 725–739, 2020.

[11] J. Cocola, P. Hand, and V. Voroninski, "No statistical-computational gap in spiked matrix models with generative network priors," *Entropy*, vol. 23, no. 1, p. 115, 2021.

[12] Y. Hosoe, D. Peaucelle, and T. Hagiwara, "Linearization of expectation-based inequality conditions in control for discrete-time linear systems represented with random polytopes," *Automatica*, vol. 122, no. 45, 2020.

[13] B. Lacroixacheztoine and Y. V. Fyodorov, "Counting equilibria in a random non-gradient dynamics with heterogeneous relaxation rates," *Journal of Physics A: Mathematical and Theoretical*, vol. 55, no. 14, 151452 pages, 2022.

[14] O. Ledoit and M. Wolf, "Analytical nonlinear shrinkage of large-dimensional covariance matrices[J]," *Annals of Statistics*, vol. 23, no. 5, pp. 142–156, 2020.

[15] Y. Li, W. Chen, and L. Yang, "Multistage linear gauss pseudospectral method for piecewise continuous nonlinear optimal control problems," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 120, no. 10, pp. 1–15, 2021.

[16] G. Liu, S. Shen, and X. Jia, "Dissipative filter design for nonlinear time-varying-delay singular systems against deception attacks," *Complexity*, vol. 42, no. 17, pp. 710–724, 2021.

[17] J. Liu, L. Wang, and Z. Tian, "Adaptive identification methods of nonlinear random electrical signal," *International Journal of Applied Electromagnetics and Mechanics*, vol. 64, no. 1-4, pp. 431–438, 2020.

[18] A. Malekan, S. Saber, and A. A. Saberi, "Exact finite-size scaling for the random-matrix representation of bond percolation on square lattice," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 32, no. 2, pp. 023112–023126, 2022.

[19] M. B. Fakoya and S. E. Malatji, "Integrating ESG factors in investment decisions by mutual fund managers: a case of selected Johannesburg Stock Exchange-listed companies[J]," *Investment Management and Financial Innovations*, vol. 17, no. 4, pp. 258–270, 2020.

[20] Q. J. Jin, "ESG Factors as a determinant on credit ratings[J]," *Asian Review of Financial Research*, vol. 34, no. 2, pp. 31–65, 2021.

[21] A. Madhavan, A. Sobczyk, and A. Ang, "Towards green alpha: analyzing ESG exposures through a factor lens[J]," *SSRN Electronic Journal*, vol. 42, no. 6, pp. 4669–4678, 2020.

[22] T. Y. Shemyakina, "Application of esg-factors in investment construction projects[J]," *Vestnik Universiteta*, vol. 12, no. 3, pp. 132–139, 2021.

[23] L. Qiu, X. Yang, and U. Ahsan, "Model predictive position tracking control for motion system with random communication delay[J]," *IET Control Theory & Applications*, vol. 14, no. 20, pp. 7785–7796, 2020.

[24] A. A. Semenov and Y. M. Beltukov, "Nonlinear elastic moduli of composite materials with nonlinear spherical inclusions dispersed in a nonlinear matrix[J]," *International Journal of Solids and Structures*, vol. 191, no. 10, pp. 726–735, 2020.