Fund Performance Evaluation Based on Bayesian Model and Machine Learning Algorithm

Shuanbao Li and Shenming Qu

1School of Artificial Intelligence, Henan Finance University, Zhengzhou 450045, China
2School of Software, Henan University, Kaifeng 475001, China

Correspondence should be addressed to Shenming Qu; qushenming@whu.edu.cn

Received 21 January 2022; Revised 25 February 2022; Accepted 26 February 2022; Published 1 April 2022

Copyright © 2022 Shuanbao Li and Shenming Qu. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Based on Bayesian method, this paper constructs a model for estimating fund performance evaluation, and uses machine learning algorithm to construct a sampler that can sample on the basis of conditional distribution. Sampling is used for stress test, so as to give the closeness of all possible test results and data results. The results show that performance evaluation is affected by many factors, and the resistance to risk plays an important role in the whole performance evaluation. At the same time, the Bayesian model in machine learning can quickly and accurately approach the statistical results, which is of great significance for predicting performance evaluation.

1. Introduction

As a financial innovation tool endowed with strong vitality and institutional advantages, funds have gained rapid development in recent years. How to reasonably evaluate the fund’s business performance is of great importance to the general investors, fund managers and industry regulators. Fund performance evaluation is a complex issue. It involves not only objective and effective performance measurement standards, but also the sustainability of fund performance and performance attribution analysis [1–6]. From the current situation, the research on fund performance evaluation in China is still weak, not only in terms of theoretical research which is largely lacking market regulation is also relatively disorganized. And this is also true in empirical studies. In recent decades, theoretical research and practice at home and abroad have shown that quantitative analysis methods have been widely introduced into fund performance evaluation.

With the continuous development of modern mathematical finance theory, fund performance evaluation has been guided by theoretical studies with many empirical analyses. The results of the empirical analysis, in turn, validate the relevant financial theories and hypotheses. In fact, many controversial hypotheses in the field of financial performance evaluation, which can still be found in the field of fund performance evaluation. Foreign research on fund performance evaluation has a long history with a relatively sound theoretical system and a large number of empirical studies, while China is still basically in the development stage segment. This is because securities investment funds have been established in China for a relatively short period of time. Even the funds established around the world in the early 1990s have only a history of more than 30 years and are not standardized. Although they differ in their choice of metrics to measure risk (Sharpe Index uses all risks). Many evaluation models have a long history and are still widely used today. However, this theory has also met with strong opposition. Therefore, there is a need to select appropriate fund performance evaluation models to further develop the fund evaluation system [7–14].

For now, persistent attribution is another major aspect of research on performance evaluation. Domestic research on fund performance evaluation is still relatively scarce. From published articles and collected data, domestic research is still limited to the introduction of foreign theories and empirical studies, with more studies on a single theory or perspective, lacking holistic and systematic research, and lacking theoretical innovation.
With the synchronous growth of the number and scale of guidance funds, the management norms of government venture capital guidance funds by national and local government institutions at all levels have been gradually improved. On February 21, 2020, Government departments have issued government investment funds and improving the efficiency of financial investment, which puts forward higher requirements for strengthening government budget constraints, improving the use efficiency of financial funds, performance management and exit management. In order to comprehensively examine the policy guidance and investment operation effect of government venture capital guidance funds, Shandong, Henan and Zhejiang provinces have successively issued specific measures for the management of government investment funds and formulated the basic performance evaluation system of guidance funds. However, due to the unbalanced social and economic development in different regions and other factors, there are certain differences in the setting of performance evaluation indicators of guidance funds. Some government venture capital guidance funds even lack a normative description of the specific performance evaluation indicators and evaluation system of guidance funds, and have not established a sufficient and perfect fund performance evaluation system. As a result, there are great differences in fund use efficiency and performance of government venture capital guidance funds in different regions. Therefore, it is objectively necessary to establish a scientific, mature and operable performance evaluation system of government venture capital guidance fund that can fully reflect the allocation efficiency of government venture capital guidance fund, so as to fully activate the policy guidance effect of guidance fund. At present, performance evaluation can be systematically divided into three parts, as shown in Figure 1:

This paper focuses on the fund performance evaluation index system, which is based on the requirements of the Government Sector for the relevant performance evaluation policies of investment funds, and focuses on the realization of the fund policy objectives, the realization of the social benefit objectives and the standardized operation. In the process of designing the index system, on the basis of comprehensively investigating the basic characteristics of the fund performance evaluation index system launched in various pilot areas, the rows and columns of the matrix represent the total assets and benefits of the fund respectively. To test the performance of the network reconstruction method of row and column aggregation of the matrix composed of these fund data. Therefore, any stress test results will significantly depend on the method used to fill in the missing information. This paper proposes a new evaluation and testing method based on Bayesian model and machine learning algorithm.

2. Model Introduction

2.1. Bayesian Model

2.1.1. Basis of Bayesian Method. The concept of Bayesian method is relatively simple [15–20]. Generally, Bayesian reasoning uses a core tool called Bayesian theorem or Bayesian criterion. In fact, it is such a simple formula. We have basically learned this in probability. There are several basic objects in the formula, which we use $\theta$ to describe the parameters of the model. The model can be neural network, linear model or SVM. All parameters are used $\theta$. To describe; $D$ is our training set; $\pi(\theta)$ is a priori distribution, which is the description of the distribution of the model itself before we see the data; $p(D | \theta)$ is the likelihood function, given a model $\theta$. The case describes the likelihood of this data. Our goal is to obtain this a posteriori distribution, which is to look at the distribution of the model itself after seeing the data. It can be expressed as follows:

$$p(\theta | D) = \frac{p(D | \theta)}{p(D)}.$$  (1)

What can Bayesian rule do in machine learning?

Prediction: $m$ is a model class (for example, linear model in RD).

$$p(x | D, M) = \int p(x | \theta, D, M)p(\theta | D)d\theta.$$  (2)

Due to the IID assumption, $p(x | \theta, D, M)$ is usually simplified to $p(x | \theta)$.

Model comparison:

$M_1$ is a model class (e.g., liner model in Rd), and $M_2$ is another model class (e.g., polynomial model with order $p$)

$$p(D | M_1) \geq (? \text{ or } ?) p(D | M_2)p(D | M) = \int p(D | \theta)p(\theta | D)d\theta.$$  (3)

First, the prediction problem. We use big $M$ to describe model class, such as linear model and nonlinear model. There are many specific models in model class. We still use parameters $\theta$ express. Make a prediction for the new sample, such as calculating its likelihood, so that you can use the posterior distribution derived from the front to do an integral. This is the likelihood of the sample under a given model, which is the distribution of all possible models. In essence, an average is made. This idea is actually quite similar to many people’s idea of integrated learning. We also make a weighted average for many models, but there may be infinite models in this place. We use probability distribution to describe it.

This likelihood function usually assumes that the test data and training data are independent when the model parameters are given, which is what we usually call the assumption of independent and identically distributed.

In addition to prediction, we can also compare different models and select models. For example, if we want to solve the classification problem, whether to choose a linear model or a deep learning nonlinear model is the problem of model selection. This problem can be described as follows: we use $M_1$ to represent a model class, which may be a linear model, and $M_2$ to represent another model class, which is a nonlinear model. In the case of the same data set $D$, we can compare which of the two quantities is large. This quantity is described. Under $M_1$, we observe a likelihood of the training
set, the other is the set of statistics in the case of \( M_2 \). Through this comparison, we can see which model I should choose. This is a basic rule for model selection using Bayesian method.

### 2.1.2. Why be Bayesian.

One of many answers. Infinite exchangeability:

\[
\forall n, \forall \sigma, \ p(x_1, \ldots, x_n) = p(x_{\sigma(1)}, \ldots, x_{\sigma(n)}). \tag{4}
\]

De Finetti’s theorem (1955): if \((x_1, x_2, \ldots)\) are infinitely exchangeable, then \( \forall n \)

\[
p(x_1, \ldots, x_n) = \prod_i p(x_i | \theta)p(\theta)d\theta. \tag{5}
\]

For some random variable \( \theta = \{x_1, x_2, \ldots\} \) This has a basic property called infinite exchangeability. If your data has \( n \) samples, if they are exchanged in any order, the joint distribution of these data does not change, that is, what we usually say is independent of the sequence, then its joint distribution can be written in the form of Bayesian—there is a model and corresponding distribution, you can describe the whole distribution in the form of integral. If you draw it graphically, given this model, the data is completely independent. We call it conditional independence, which is a very important concept in the probability graph model.

It is worth noting that the exchangeability is a little wider than the independent and identically distributed previously. The exchangeable data cannot be independent and identically distributed. In addition, this theorem only tells you that it exists, but the latter problem is that we don’t know what the specific model should be, such as linear model and nonlinear model, and what kind of model to describe, which is the problem to be solved by statistical modeling.

Where does a priori come from?

The main choice is through the following. Objective priors which are nonstatistically informative priors. Subjective a priori—a priori should reflect our thoughts and trends as much as possible.

Hierarchical a priori—multi-level a priori.

Empirical priors—some parameters for learning priors from the data.

\[
p(\theta | \alpha) = \text{argmax} \ p(D | \alpha). \tag{6}
\]

Pros: robust—overcomes some limitations of mis-specification

Cons: double counting of evidence/overfitting

There is a technology called “hierarchical priority.” There is a basic assumption that the parameters in \( \pi \) are called hyperparameters. The farther away it is from my data generation model, the weaker the impact. In fact, using this basic assumption, we can build multi-layer prior.

It can be seen from this that Bayesian itself is a multi-layer representation, which is essentially the same as the multi-layer representation of deep learning, but it is completely described in the way of probability. Of course, there are some approximate methods. Do some empirical priority and estimate it through data. The advantage is that the calculation is relatively simple. The disadvantage is that if you use the training data many times, you will get some over fitting problems.

How to calculate the integral?

Recall that:

\[
p(D | M) = \int p(D | \theta)p(\theta | M)d\theta. \tag{7}
\]

This can be a very high dimensional integral.

If we consider latent variable, it leads to additional dimensions to integrated out

\[
p(D | M) = \int \int p(D, H | \theta)p(\theta | M)d\theta dH. \tag{8}
\]

No matter in the most basic Bayesian operation or in multi-layer a priori, integral operation will be used repeatedly, which is the most annoying thing when using Bayesian method, because the integral here is not like the unary or binary simple integral operation when we learn calculus. The integral here may be a very high-dimensional integral. Suppose we use a linear model. When the feature is 100 dimensions, I’m the integral of 100 dimensions. If it is higher dimensions, such as thousands and thousands of dimensions, the integral is higher dimensions. There are very important calculation problems, The first point is the basic theoretical problem. The concept of Bayesian method I mentioned earlier is very simple, and its core is Bayesian
Bayesian theorem has existed for more than 250 years. It has some limitations. How can we reunderstand it from the perspective of basic information processing criteria and make more flexible reasoning.

The second point is about calculation. With the model and reasoning framework, how can I do efficient calculation and high-precision calculation.

Third, how to model different scenes. When applying Bayesian method to different scenarios, you have to understand your problem and find the appropriate model.

Bayesian theorem or Bayesian reasoning is a criterion of information processing. In information theory, we have a channel with input and output. In Bayesian reasoning, our input is a priori distribution and a likelihood function, and the output is a posteriori distribution. With this view, we can make many extensions. You can think that my input and output can remain unchanged, but I can replace the criteria of information processing. I can remove the criterion of information processing here and make it more universal than classical Bayes. The Bayesian method is organically combined with the machine learning framework based on risk minimization. Machine learning has deeply studied risk minimization and related optimization tools. Under our framework, Bayesian reasoning can be integrated to obtain a more flexible information processing criterion.

### 2.2. Machine Learning

In machine learning and pattern recognition, we are familiar with the loss function optimization problem. For example, if I want to do classification, I want to train the neural network. The first item is a loss function, which measures the error rate in the training set; The second term is the regularization term we want to add. The purpose is to protect the model from over-fitting and divergence. This is a basic framework. These things are basically out of touch with Bayes in machine learning and will not be discussed together with Bayes [21–27].

First, establish learning objectives in machine learning

\[
\min_{\theta} l(\theta; D) + r(\theta). \tag{9}
\]

Loss functions of our predictive task.

#### 2.2.1. Recap of SVMs

SVM learn a single decision boundary. As shown in the figure.

Here we give a simple example to show the application of Bayesian method in machine learning, such as linear SVM. Its purpose is to find a linear plane to distinguish different categories. As shown in Figure 2:

The first strategy is an averaging model and the second is a stochastic model. The two are theoretically related. With this different treatment, the understanding is achieved by dividing it up in space. With this loss function, optimisation can be carried out directly. We can understand this classifier as a division of the linear space, at its simplest, on a two-dimensional space, by straight lines. It is also understood as template matching, where each category can be seen as a template for one of the categories. The score for each category is, in effect, the element and template match for calculation. Through the loss function, also called the cost function, which is used to measure the degree of match, the aim is to minimise the loss function during the training process. From a Bayesian perspective, this is the SVM approach [28–31].

#### 2.2.2. Bayesian SVMs

The overall optimization problem

\[
\min_{q(\theta)} KL(q(\theta)\pi(\theta)) + C \cdot l(q(\theta); D). \tag{10}
\]

Strategy #1 (averaging model):

\[
l(q; D) = \sum_{i} \max_{\theta} \left(0, 1 - y_{i}x_{i}^{\top} \theta\right) \theta = E_{q(\theta)}[\theta]. \tag{11}
\]

Strategy #2 (Gibbs/stochastic model):

\[
l(q; D) = \sum_{i} E_{q(\theta; D)} \left[ \max_{\theta} \left(0, 1 - y_{i}x_{i}^{\top}\right) \right]. \tag{12}
\]

With the above problems, how to solve them? The first average model can be solved by convex optimization. When the model is linear, if the a priori is Gaussian, your a posteriori distribution is still Gaussian. This actually degenerates to traditional SVMs. You can also use kernel method to expand the nonlinear model, which is OK.

**SOLVE Bayesian SVMs.**

Strategy #1 (Averaging model):

If the prior is normal \( \pi(\theta) = N(0, 1) \), we have the solution:

\[
q(\theta) \propto \pi(\theta) \exp \left( \sum_{i} a_{i}y_{i}x_{i}^{\top} \right) = N(\sum_{i} a_{i}y_{i}x_{i}, 1), \tag{13}
\]

where \( a_{i} \) are the solution of the dual problem.

For the second, given a priori, the general solution can be obtained.

Bayesian view can bring some additional benefits. Consistent with the idea of representation learning in deep learning, in Bayesian SVM, we can consider implicit variables or add many layers of implicit variables to learn the distribution of these implicit variables. The above is the coupling of Bayesian model in machine learning. Relevant evaluation indicators will be introduced in the next section.

The application of Bayes’ theorem to machine learning to accomplish the task of model prediction and selection is machine learning from a Bayesian perspective. Since Bayesian theorems heavily involve the dependence of various explicit and hidden variables, they are usually described visually by probabilistic graphical models. Bayesianism treats the unknown parameters as random variables, and the uncertainty of the parameters before learning is described by the prior probability, while the uncertainty after learning is described by the posterior probability, and the elimination of uncertainty in between is where machine learning comes into play.

Unlike frequencyism, the output of Bayesian learning is not simply the optimal estimate \( \hat{\theta} \), but a probability distribution \( p(\theta) \) about the parameters, thus giving more
complete information. In the prediction problem, Bayesian learning also gives not just a most probable outcome, but a complete picture of all outcomes and their probabilities in the form of probability distributions.

In addition to providing more complete information in prediction, Bayesian learning also has its advantages in model selection. In Bayesianism, the so-called different models are actually parametric representations of different probability distributions, using parameters with their own prior distributions, but what all models have in common is that they all generate training datasets, and the task of model selection is to pick the best one from these probability distributions.

2.3. Indicators for Fund Performance Evaluation. The most traditional and direct method of fund performance evaluation is to understand the net value and return on investment of the fund.

(1) NAV of net assets of fund units
Deduct various expenses and liabilities from the total assets of the fund, and the balance is the net assets of the fund:

\[
\text{NAV} = \frac{\text{net assets of the fund}}{\text{number of fund units issued}}
\]  

(14)

(2) Return on investment of the fund

\[
R = \frac{(\text{NAV}_1 - \text{NAV}_0 + C)}{\text{NAV}_0}
\]  

(15)

where \(C\) is the fund interest distribution or dividend.

The above two indicators measure the return of the fund. As a rational investor, we should also consider the risk of return. Therefore, we introduce three risk adjusted return indicators:

2.3.1. Introduction to Sharp Index. William sharp, an American economist, published the performance of funds and proposed to measure the fund performance by the excess return brought by the fund’s total risk (including systematic risk and nonsystematic risk), which is the sharp index. Sharp index measures the performance of the fund by the ratio of the average return of the fund portfolio exceeding the standard deviation of the risk-free return to the fund return in a certain evaluation period. The calculation formula is:

\[
S_p = \frac{(r_p - r_f)}{\sigma_p},
\]

(16)

where \(S_p\) is the index, \(r_p\) is the actual return of the fund portfolio, and \(r_f\) is the risk-free return, \(\sigma_p\) is the standard deviation corresponding to the fund portfolio. The theoretical basis of sharp index is the capital asset model (CAPM model), which takes the capital market line (CML) as the evaluation basis.

2.3.2. Introduction to Treynor Index. The Treynor index uses the method of comparing the average risk return of the portfolio with its systematic risk over a period of time to evaluate the performance of the investment fund.

\[
T_p = \frac{(r_p - r_f)}{\beta_p},
\]

(17)

where \(T_p\) is the \(T\) index; \(\beta_p\) represents the risk factor associated with the fund system. The capital asset pricing model is also the theoretical foundation of the Treynor performance index, but it is evaluated on the basis of the security market line (SML), where all asset portfolios fall on the SML when the market is in equilibrium, i.e., the slope of the SML represents the Treynor index of the market portfolio. When the Treynor index of a fund portfolio exceeds the slope of the SML, the portfolio is above the SML line, suggesting that it outperforms the market; conversely, when the Treynor index of a fund portfolio falls below the SML, the investment fund portfolio underperforms the market. As a result, the higher the Treynor performance index, the better.

2.3.3. Jensen Index. The additional return on fund inputs expresses the value of fund information. It is also a measure of the fund’s performance. Its calculation formula is:

\[
J_p = r_p - [r_f + \beta_p(r_m - r_f)].
\]

(18)

A positive value indicates that the evaluated fund is higher than the market average level and has good investment performance; a negative value indicates that the evaluated fund is higher than the market average level and
has good investment performance; The examined fund outperforms the market in terms of investment performance and is higher than the market average. A negative rating implies that the fund manager’s stock picking expertise is poor, that he or she cannot go beyond the index, and that the evaluated fund’s overall performance is bad in comparison to the market. If it is 0, it suggests that fund managers’ stock selection ability is broad and can only be in accordance with the index.

3. Data Analysis Based on Coupling Model

In order to better understand the linear regression of the model, the relevant data are analyzed and studied. Here, Changxin healthcare industry (163001. O) is selected as the research object of the fund. The market portfolio is approximated by CSI 300 index and the risk-free interest rate is approximated by 3-month Shibor. The research period is 2016 in order to assess the fund’s performance throughout that time. First, as an example, consider the net present value of the shunt fund in the third quarter, as indicated in Figures 3 and 4:

Figures 3 and 4 show the significant growth rate of open-end stocks and the high overall return. This is crucial for performance evaluation. Under normal circumstances it is worth choosing open-end stocks. A detailed analysis of the fund companies is shown in Figure 5.

As can be seen in Figure 5, the average absolute value of the fund managers is in an upward phase, with some funds having large gains. In terms of net selection returns in terms of combined returns and value-at-risk, it is possible to select the top tier of funds among the above funds. The high rate of return in the case of taking the same risk, these are more important in the performance assessment. The performance of the funds in 2016 is measured based on the above indicators and their performance is shown in Table 1.

Here, through three linear regression in machine learning, in the H-M model, if $\beta_1$ is most greater than 0, which means that the fund charger will take the initiative to raise when the market rises $\beta$. Value to follow the bull market and lower in the falling bear market $\beta$. In the C-L model, when the market rises, use $\beta_1$. Measure the systematic risk of fund portfolio. When the market falls, use $\beta_2$ measure the systematic risk of the fund portfolio, and finally pass $\beta_1 - \beta_2$ to judge the timing ability of fund managers. As shown in Table 2:

From the regression results, we can see that under the three models, alpha is significantly different from 0, but the sign is negative, which can be considered that the fund manager’s stock selection ability is under the market. In the T-M model, beta2 was significantly greater than 0; And H-M model, beta2 was not significant; In the C-L model, beta1 and beta2 are significant, and beta1 is significantly greater than beta2. In conclusion, it can be considered that the timing ability of the fund manager is significantly above the market.

In the analysis of beta value, only one factor of market risk is considered, but it is often necessary to analyze the source of fund performance more carefully, so there is a multifactor attribution analysis model. In this part, we will consider the attribution analysis of four factors: reversal (turnover), beta, volatility (standard deviation) and size (fund size). Using the past time series for rolling regression, we can get the exposure degree of the fund on these factors every day, which is summarized in the exposition. In factor. From the observation results, the performance of the fund is mainly attributed to the volatility factor and partly to the reversal factor as shown in Table 3:

Using the past time series for rolling regression, we can get the exposure degree of the fund on these factors every day, which is summarized in the exposition. In factor. From the observation results, the performance of the fund is mainly attributed to the volatility factor and partly to the reversal factor. As shown in Figure 6:

Profitability, risk resistance and risk adjusted return indicators. Among them, stage return and Jason alpha are commonly used for profitability, volatility (standard deviation), loss ratio (frequency of net loss in the stage), average loss (mean value of loss when loss occurs) and leverage are commonly used for anti risk ability, and Sharpe ratio, sotino ratio, information ratio, uplink capture rate, downlink capture rate are commonly used for risk adjusted indicators. After analyzing the above indicators, we can get the following results as shown in Figures 7 and 8:

With the above Figures 7 and 8, we can clearly understand that the impact of risk tolerance on fund performance evaluation is very important. Therefore, we understand that it is necessary to reasonably improve the evaluation through machine learning and related training algorithms. More accurate prediction estimation of fund trend is trained by
more data to ensure the accuracy of performance evaluation. Here, we use Bayesian models to address data accuracy issues in machine learning to help achieve accurate performance evaluations quickly. The Bayesian models in Figures 9 and 10 provide further insight into the machine training process and the degree of performance statistics during training, as shown in Figures 9 and 10.

As can be seen from Figures 9 and 10, the degree of approximation of the machine learning results gradually increases with the number of training sessions. In Figure 9...
presents probability distributions getting closer to the true values as the training of the machine is prolonged. And the trend of change of the fund performance value the more statistical value is completed consistent through Figure 10. This proves that the learning method of Bayesian model is important in fund performance evaluation.

| Date       | Explosion_turn | Explosion_scale | Explosion_beta | Explosion_volatility |
|------------|----------------|-----------------|----------------|-----------------------|
| 2016-02-22 | -0.094073      | -0.02039        | 0.087151       | -0.03889              |
| 2016-02-23 | -0.122745      | -0.011229       | 0.061206       | -0.237779             |
| 2016-02-24 | -0.150539      | -0.002904       | 0.027562       | -0.209910             |
| 2016-02-25 | -0.142255      | -0.003158       | 0.027469       | -0.244269             |
| 2016-02-26 | -0.160840      | 0.004011        | 0.000641       | -0.346814             |

Figure 6: Multi factor fluctuations in different periods.

Figure 7: Evaluation of fund performance value at different times.
Figure 8: Size of each risk value at different times.

Figure 9: Training probability density of Bayesian model in machine learning.
4. Conclusions and Recommendations

(1) This paper describes the poor return of a fund in 2016, with little fluctuation, but a high proportion of losses, and the information ratio is almost 0. The stock selection ability of the fund is below the market, but the timing ability of the fund is significantly above the market. From the multi factor point of view, the performance of the fund is mainly attributed to the volatility factor, some to the reversal factor, and other factors are not significant. The indexes that fit well with the performance of the fund are pharmaceutical, biological, mechanical equipment, computing and other industries, and the index fitting of each industry is poor.

(2) As a basic algorithm, Bayesian statistics plays an important role in machine learning. Especially in data processing, it has a good classification effect on the analysis of event probability and event reliability.

In addition to not mechanically copying the quantitative analysis indicators of fund evaluation, the following recommendations should be noted when evaluating investment funds, taking into account the actual situation of the fund market.

(1) Quantitative consideration of evaluation indicators of nonquantitative factors. In the actual evaluation process of fund performance, it should also be combined with other aspects for comprehensive evaluation. Such as, the fund company’s management and management style. In addition, the economic environment and the trend of the securities industry involved in the fund can hardly be judged fairly by numerical analysis alone. It is better to consider the fund based on various factors to prevent generalization.

(2) The evaluation of investment funds should focus on performance and the sustainability of performance. Also pay attention to consider the stability of the fund, as long as to ensure a stable return is worth investing.

(3) The choice of benchmark needs a reasonable and comprehensive reference standard to compare the fund’s performance with the market situation. However, there is no unified and authoritative market index in China. The method of comparing the fund with Shanghai index or Shenzhen index to reflect the fund performance is not objective enough. Because, on the one hand, neither the Shanghai index nor the Shenzhen index is a unified index of the two markets, and the fund invests in the two markets; On the other hand, the fund can only invest in the circulating shares of the two cities, and the Shanghai index and Shenzhen index take the total share capital rather than the circulating share capital as the weight. Therefore, the above index or Shenzhen index is not effective in evaluating the fund performance. Some people think that we should take the weighted average of the two indexes as the benchmark, or make appropriate adjustments to the two indexes, but there are also some difficulties in practical operation.

(4) In the evaluation of fund performance, the principle of comparative performance of similar funds must be adhered to. If it is necessary to compare the performance of different types of funds, certain transformation of indicators must be performed.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research was supported by the National Natural Science Foundation of China (Grant nos. U1636107 and 61972297), the Science and Technology Project of Henan Province (China) (Grant nos. 182102210215, 192102210288, and 212102210538), the Soft Science Project of Henan Province (China) (Grant no. 182400410482), and Postgraduate education innovation and quality improvement project of Henan University (Grant no. SYL20040121).

References

[1] N. Hütter, D. T. Robinson, S. Sievers, and T. Hartmann-Wendels, “Paying for performance in private equity: evidence from venture capital partnerships,” Management Science, vol. 66, no. 4, pp. 1756–1782, 2020.
[2] H. Sharifi, A. Roozbahani, and S. M. Hashemy Shahdany, “Evaluating the performance of agricultural water distribution systems using FIS, ANN and ANFIS intelligent models,” Water Resources Management, vol. 35, no. 6, pp. 1797–1816, 2021.

[3] W. Jung, “The needs of incentives on long-term listed stock’s capital gains tax in planned financial investment gains taxation,” Seoul Tax Law Review, vol. 27, no. 1, pp. 325–362, 2021.

[4] M. Schmitt and A. D. Vries, “Ranking dairy sires using genetic selection indices based on financial investment methods versus lifetime net merit: description and spreadsheet,” Environmental Data and Information Service, vol. 102, no. 3, 2021.

[5] Q. Zhang, X. Guo, and H. Li, “The impact of financial risks on financial investment in infrastructure: based on a two-factor stochastic differential equation,” Discrete Dynamics in Nature and Society, vol. 2021, Article ID 9112739, 14 pages, 2021.

[6] A. Takahashi and S. Takahashi, “A new interval type-2 fuzzy logic system under dynamic environment: application to financial investment,” Engineering Applications of Artificial Intelligence, vol. 100, no. 1, p. 104154, 2021.

[7] P. Ranalli, D. Hobbs, and L. Lindegren, “Astrometry and exoplanets in theGaiaera: a Bayesian approach to detection and parameter recovery,” Astronomy & Astrophysics, vol. 614, p. A30, 2018.

[8] S. Wu, M. Yamada, K. Tamariibuchi, and J. L. Beck, “Multi-events earthquake early warning algorithm using a Bayesian approach,” Geophysical Journal International, vol. 200, no. 2, pp. 789–806, 2015.

[9] J. V. Candy, “Environmentally adaptive processing for shallow ocean applications: a sequential Bayesian approach,” Journal of the Acoustical Society of America, vol. 138, no. 3, pp. 1268–1281, 2015.

[10] P. M. Schnell, Q. Tang, W. W. Offen, and B. P. Carlin, “A Bayesian credible subgroups approach to identifying patient subgroups with positive treatment effects,” Biometrics, vol. 72, no. 4, pp. 1026–1036, 2016.

[11] S. J. Eder, D. Steyrl, M. M. Stefanczyk et al., “Predicting fear and perceived health during the COVID-19 pandemic using machine learning: a cross-national longitudinal study,” PLoS One, vol. 16, no. 3, Article ID e0247997, 2021.

[12] G. Pfeiffer, “The diagnostic power of motor unit potential analysis: an objective bayesian approach,” Muscle & Nerve, vol. 22, no. 5, pp. 584–591, 2015.

[13] E. V. Moran and J. S. Clark, “Estimating seed and pollen movement in a monocoeous plant: a hierarchical Bayesian approach integrating genetic and ecological data,” Molecular Ecology, vol. 20, no. 6, pp. 1248–1262, 2015.

[14] X. Wu, H. Liu, L. Zhang, M. J. Skibniewski, Q. Deng, and J. Teng, “A dynamic Bayesian network based approach to safety decision support in tunnel construction,” Reliability Engineering & System Safety, vol. 134, pp. 157–168, 2015.

[15] C. H. Wilson, T. T. Caughlin, D. J. Civitello, and S. L. Flory, “Combining mesocosm and field experiments to predict invasive plant performance: a hierarchical Bayesian approach,” Ecology, vol. 96, no. 4, pp. 1084–1092, 2015.

[16] S. J. Moe, S. Haande, and R.-M. Couture, “Climate change, cyanobacteria blooms and ecological status of lakes: a Bayesian network approach,” Ecological Modelling, vol. 337, pp. 330–347, 2016.

[17] D. H. Huang, S.-H. Wu, W.-R. Wu, and P.-H. Wang, “Cooperative radio source positioning and power map reconstruction: a sparse bayesian learning approach,” IEEE Transactions on Vehicular Technology, vol. 64, no. 6, pp. 2318–2332, 2015.

[18] T. H. Edwards and S. Stoll, “A Bayesian approach to quantifying uncertainty from experimental noise in DEER spectroscopy,” Journal of Magnetic Resonance, vol. 270, pp. 87–97, 2016.

[19] F. Ozel and D. Psaltis, “Statistics of measuring neutron star radii: assessing A frequentist and A bayesian approach[1],” The Astrophysical Journal, vol. 810, no. 2, pp. 268–308, 2015.

[20] Z. C. Quiroz, M. O. Prates, and H. Rue, “A Bayesian approach to estimate the biomass of anchovies off the coast of Perú,” Biometrics, vol. 71, no. 1, pp. 208–217, 2015.