Gold Price Forecast: A Summary of the Integration of Economic Factors and Calculation Methods

Bowen Xie[a]
East Los Angeles College, 1301 Avenida Cesar Chavez, Monterey Park, U.S.A.

Keywords: Gold Price, Machine Learning, Deep Learning, Economic Factors.

Abstract: As an essential safe-haven asset and investment tool in the global financial markets, the price movement of gold has been widely watched. The price of gold is affected by various economic factors, making its prediction challenging. In recent years, with the development of machine learning and deep learning technology, scholars have begun combining traditional econometric models with advanced computational models for gold price prediction to improve prediction accuracy and provide a reference for investment decisions. This paper synthesizes the research in gold price forecasting from the perspective of economics and computational methods. Based on the background of the gold market and the factors affecting the price, it compares the progress of the application of traditional time series models and machine learning models, discusses the performance, advantages, and disadvantages of the different models, and finally puts forward the challenges faced by the current research and the direction of future development. Several studies have shown that machine learning models incorporating economic factors have yielded promising gold price prediction results. This can better capture the nonlinear fluctuation characteristics of gold prices and provide valuable references for the investment market.

1 INTRODUCTION

Gold plays a pivotal role in the global financial system, as an essential jewelry and industrial material and as a key component of central banks' reserves and investors' asset allocation. Because of its stable value and high resistance to inflation, gold has always been seen as the backbone of the fight against financial market uncertainty. Studies have shown that the price of gold is closely related to the macroeconomic situation: for example, gold is often used as a hedge against stock market risks, against inflation, and as a "safe-haven" in times of currency depreciation or financial crisis, with its price showing a negative correlation with significant currencies and stock markets. However, the volatility of the gold price is also influenced by speculative trading and market sentiment, which can be unpredictable and drastic. This price volatility poses risks for investors and policymakers and highlights the importance of accurate gold price forecasting.

Gold price forecasts rely heavily on econometric methods and time series models. For example, the Autoregressive Integrated Moving Average (ARIMA) model is often combined with historical price trends to make forecasts, which, to a certain extent, captures the cyclical changes and trends in gold prices. At the same time, researchers also consider the impact of macroeconomic variables on gold prices by incorporating factors such as exchange rates, oil prices, and inflation rates into regression models or vector autoregression (VAR) models to improve the persuasiveness of their forecasts. However, such traditional models are often affected by their limitations when dealing with sudden situations in financial markets.

The good news is that in recent years, machine learning (ML) and deep learning (DL) methods have excelled in time-series forecasting in finance and have gradually begun to be applied in the gold price forecasting field. Rather than relying on purely statistical models, these data-driven methods are able to automatically understand complex patterns in large amounts of historical data while adapting to their nonlinear relationships. For example, machine

[a] https://orcid.org/0009-0006-3560-5386

Xie, B.
Gold Price Forecast: A Summary of the Integration of Economic Factors and Calculation Methods.
DOI: 10.5220/0013697400004670

In *Proceedings of the 2nd International Conference on Data Science and Engineering (ICDSE 2025)*, pages 360-365
ISBN: 978-989-758-765-8

learning models such as Support Vector Regression (SVR) and Random Forest have demonstrated superior performance in gold price prediction. At the same time, deep learning models such as Long Short-Term Memory (LSTM) can capture the time series' long- and short-term dependence, which also significantly improves the forecasting accuracy of financial series, including the price of gold. In this context, combining insights from economics with advanced computational models would be the best option to improve the accuracy of gold price forecasting. The remainder of this review will introduce the main forecasting methods and their applied research progress, analyze the effects of different models and suggest future research directions.

2 SUMMARY OF PREDICTION METHODS

Gold price forecasting methods can be broadly categorized into two types: traditional time series and econometric methods and machine learning and deep learning methods. The former represents an economic perspective, mainly modeling with historical patterns and considering macro factors; the latter reflects a computational perspective, mainly learning forecasting patterns from data. The following will summarize these two types of methods and related research separately.

2.1 Traditional Measurement and Time Series Model

The first is time series modeling; as one of the most widely used traditional methods, the ARIMA model is often used to capture the autocorrelation structure in time series data. Many studies have used ARIMA as a benchmark model to forecast gold prices. For example, Yang used the ARIMA model to model the forecasting trend of gold prices with some accuracy (Yang, 2019). This model works well for linear and smooth time series, but the actual gold price series tends to be nonlinear and unstable, which requires differencing and parameter tuning of the model to accommodate this complex pattern. In addition, many studies have combined ARIMA with other methods, such as using ARIMA to forecast long-term trends and then mixing it with Support Vector Machines (SVMs) to capture the nonlinear portion of the residuals, which can serve to improve forecasting performance.

The second is the economic factor regression model; in addition to the pure time series model, researchers also often use multiple regression models to incorporate macroeconomic variables into gold price forecasting; typical examples include Ismail and other scholars as early as 2009 on the application of multiple linear regression in gold price forecasting. In fact, many common economic indicators affect the price of gold, such as the price of crude oil, inflation rate, stock market index, and so on. By regressing these variables with gold price, the fluctuation of gold price can be explained from the perspective of economic fundamentals. However, such linear regression makes it challenging to characterize the more complex nonlinear relationship between variables. For this reason, using more advanced econometric models, such as GARCH-type models, to describe the volatility dynamics of the gold price or using structural equations and cointegration analyses to explore the long-run equilibrium relationship would be more effective. In general, however, traditional econometric models do frequently falter when dealing with highly volatile and nonlinear financial data.

2.2 Machine Learning Model

In addition to the typical economics methods described above, machine learning methods are also well worth using. This approach does not rely on strict modeling assumptions but instead extracts valuable features from historical data and applies them to forecast the price of gold with high accuracy. Regression trees and integrated models are two of the most commonly used methods. Decision tree regression fits nonlinear relationships by recursively splitting the data and is suitable for determining the key intervals and thresholds that affect the price of gold. Random Forest model, such as the decision tree-based integration method, has a higher level of robustness and accuracy because of the combination of the results of multiple trees. Studies have shown that when dealing with noisy data, Random Forest tends to achieve more minor mean square error (RMSE) than the simple linear model and has better prediction performance. Sharma et al. compared the effectiveness of various supervised learning models for predicting the price of diamonds and ultimately found that Random Forest has the smallest RMSE and the highest accuracy rate. Also, for the continuous variable prediction task of gold price, Random Forest proved to be one of the effective models.

Another machine learning method is SVR, which is a method that uses kernel function mapping to transform the input features into a high-dimensional space and find the optimal hyperplane in the space to minimize the prediction error value. SVR is widely used in financial time series prediction; Yuan et al. have combined market sentiment analysis and genetic algorithm optimization with the LS-SVR model to predict the gold price and achieved high accuracy. Yuan et al. combined market sentiment analysis and a genetic algorithm to optimize the LS-SVR model for gold price forecasting and achieved high accuracy. In addition to these methods, other algorithms, such as regularized variants of multiple linear regression (ridge regression, Lasso regression), k-nearest neighbor regression (KNN), etc., have also been used for modeling gold prices. Overall, these machine learning models can better capture the complex relationship between gold prices and multidimensional features than traditional statistical models and tend to provide better predictions with sufficient training data.

2.3 Deep Learning Model

Due to the nonlinear pattern of gold price changes, deep learning methods that automatically learn high-level features from data through a multilayer neural network structure have a significant advantage in predicting in the face of such nonlinear time series. Specifically, for example, networks like LSTM are a special kind of recurrent neural network, which is very well suited to meet the long-range dependence in processing time series data due to its gating structure that can store and update long-term states. In gold price prediction, the LSTM model is able to utilize the price information of the past ages to predict future trends, which has a data advantage over the traditional model that only refers to recent data. A study comparing the performance of LSTM and ARIMA models in gold price prediction shows that LSTM significantly outperforms ARIMA in terms of RMSE and MAPE, etc. Yurtsever also applies LSTM and its bidirectional LSTM and GRU variants to gold price prediction, demonstrating that deep learning models have higher prediction accuracy than traditional methods. Prediction accuracy compared to conventional methods (Yurtsever, 2021).

There is also a beneficial model called a hybrid neural network model, which is a hybrid model formed by combining different types of neural networks to improve prediction performance further. Among them, Convolutional Neural Network (CNN) is good at extracting local time-series features from data, while combining CNN with LSTM (i.e., CNN-LSTM model) can simultaneously capture the regional patterns of the gold price series and satisfy its global time-dependence, which is very effective. CNN-LSTM model proposed by Livieris et al. (Livieris et al., 2020). has also achieved good results in predicting the time series of the gold price. Good results. Similarly, some studies integrate signal processing methods such as variational modal decomposition (EMD) with LSTM, and some even incorporate attention mechanisms and optimization algorithms to enhance the model's adaptability to the dramatic fluctuations of gold prices. In addition, some studies combine integrated learning with deep learning, such as using the XGBoost algorithm to screen features or provide initial predictions and then fine-tuning the results by neural networks to improve the robustness of the model. Jabeur et al. utilized the Extreme Gradient Boosting (XGBoost) algorithm. They combined it with the SHAP interpretation method to scientifically predict and interpret the price of gold, demonstrating the model saturation and the economic meaning of feature influence (Jabeur et al., 2021). The financial implications of the model's characteristic influence are shown.

From the analysis of the above models, they have their own advantages and disadvantages: traditional econometric models have good interpretability and rich economic implications, but it is challenging to capture the complexity of the nonlinear relationship; on the contrary, machine learning and deep learning models can provide higher prediction accuracy, but often they are regarded as a "black box" due to the lack of direct economic interpretation. ". Therefore, a significant trend in recent years has been to integrate the advantages of both, for example, by incorporating macroeconomic variables into machine learning models or by interpreting the decision-making rationale of complex models, which can be done in a way that ensures accuracy while improving the transparency and usefulness of the models.

3 ANALYSIS OF LITERATURE RESULTS

Since its realization, a large body of literature has examined and empirically compared the performance of different models in gold price prediction. The general trend is that machine learning and deep learning models have higher prediction accuracy compared to traditional methods. For example, Makala and Li compared the prediction of the gold price by ARIMA and the support vector machine. They found that the error of price prediction was reduced after introducing a nonlinear machine learning model (Makala and Li, 2021). Tripurana et al. reviewed the effectiveness of various machine learning algorithms for gold price prediction, including linear regression, decision trees, random forests, SVR, and neural networks, emphasizing the importance of integrating multiple economic factors and feature engineering (Tripurana, 2022). Their study shows that machine learning models are able to predict gold price movements more accurately when relevant macro variables are taken into account.

Let's look at the deep learning modeling side of the equation, where the results of many studies have demonstrated superior performance over both traditional and shallow models. A survey by Yurtsever showed that LSTM models already outperform classical methods and that combining bidirectional LSTMs or gated recurrent units (GRUs) can further improve the accuracy (Yurtsever, 2021). Another study utilized a combination of LSTM and fuzzy systems to predict the price of gold, again obtaining results below the error of the traditional approach. Livieris et al.'s hybrid CNN-LSTM model excelled in long-term prediction, with experiments on monthly gold price data achieving high goodness of fit (Livieris et al., 2020). Taking the metrics together, deep learning models typically excel in measures of prediction accuracy such as mean squared error (MSE or RMSE), while in directional prediction (e.g., up and down direction), some machine learning models such as classification trees also achieve good results like Sadorsky uses a random forest classifier to predict the up and down direction of the price of gold and silver, with an accuracy rate of 1.5 percent compared with that of randomized predictions (Sadorsky, 2021)

Nonetheless, some studies point out that the amount of data and parameter selection influences the advantage of deep learning models. If the training data is insufficient or the parameters are not tuned properly, the complex model may not always win. Therefore, scholars emphasize the importance of comparing benchmark models, which should be comprehensively judged by cross-validation and multi-indicator assessment. In addition, considering that gold prices are significantly affected by unexpected events, such as financial crises or public health events (e.g., the New Crown Epidemic) that can lead to inaccuracy of the model's prediction based on historical patterns, some studies have specifically improved the model for gold price prediction in times of epidemics. For example, Mohtasham Khani et al. explored a methodology that combines epidemic-related variables and deep learning in the presence of epidemic shocks to improve the reliability of gold price forecasts during special periods (Mohtasham Khani et al., 2021).

Overall, today's literature generally endorses the application of machine learning, intense learning, to gold price forecasting but also emphasizes the importance of incorporating economic theory and adequate data. Precisely because multi-model integration and hybrid approaches tend to achieve more robust results, some recent studies have attempted to introduce explanatory methods (e.g., SHAP values) to reveal the factors that the models consider most essential to influence the forecasts, thus making them more informative.

4 CHALLENGES AND PROSPECTS

Although researchers have made significant progress in the field of gold price prediction, there are still many challenges in further research. First, regarding data and feature selection, the gold price is still affected by multiple factors, and how to select and extract key features is still inconclusive. Future research can explore how to introduce more real-time data sources, such as news sentiment analysis, Internet search trends, etc., to enrich the model's input. Second, the generalization and robustness of the model remain a problem. Financial markets are often unstable, and the fact that a model performs well on historical data does not guarantee that it will remain valid in the future. Therefore, there is a need to develop adaptive forecasting models that can be updated promptly based on real-time data. Not only that, but the impact of abnormal events (black swan events) on the gold price is also often beyond the scope of experience accumulated by the model. Improving the model's ability to respond to extreme situations is a direction worthy of research. This can be done by combining scenario analysis or building a hybrid forecasting system that relies on data-driven models during smooth periods while combining rules or expert judgment during abnormal periods.

At a time when the interpretability of models is increasingly emphasized, it is more important for investors to understand the reasons why models give predictions rather than being satisfied with highly accurate predictions. Future research could make greater use of explainable artificial intelligence (XAI) methods to explain the decision-making process of complex models. Use feature importance analysis to clarify which macro indicators have the highest weight in model predictions or use locally interpretable models to analyze the drivers of predictions at specific points in time. This will facilitate the translation of forecasts into the market insight process.

Interdisciplinary integration will be an essential development trend in gold price forecasting. Gold price movement is not only a financial phenomenon but is also affected by multiple factors, such as politics and society. If the theories of financial economics are combined with machine learning techniques, new forecasting frameworks can be generated. Just like the idea of integrating game theory or multi-agent simulation, a complementary effect can be achieved with data-driven models. In addition, with the increasing computational power, the realization of real-time prediction and automated trading is no longer impossible, and future research can further embed the prediction models into trading decision-making systems so as to assess their actual returns and risks.

5 CONCLUSIONS

In general, gold price forecasting is a valuable and challenging research topic in the field of finance. This paper summarizes the significant research advances in this area from both economic and computational methodological perspectives. Both traditional econometric models, which provide the underlying framework and economic explanations for gold price forecasting, and machine learning and deep learning models, which are capable of handling large-scale historical data with high accuracy, play a crucial role. The approach of incorporating macroeconomic factors into the forecasting models, in turn, takes the performance of the models to the next level and also enhances their real-world applicability, providing more reliable theoretical and practical support for gold price forecasting.

However, even with modeling, there are still many sources of uncertainty in accurately predicting the price of gold, not to mention that the applicability of models needs to be tested in a changing market. In order to improve the reliability of forecasts, future research directions need to focus on the generalization ability of models and the adaptation to abnormal situations, which may require the integration of multi-model approaches and the accumulation of experts' experience. At the same time, enhancing the interpretability of model results is crucial for the application of forecasts in practical decision-making. All in all, gold price forecasting should be a cross-disciplinary collaborative process that combines economic insights with data science techniques, utilizing machine learning models to improve forecasting accuracy on one hand and financial theories to explain the results and guide model improvement on the other. Only in this way can forecasting research provide investors and policymakers with truly valuable guidance on how to respond to the rapidly changing gold market.

REFERENCES

Arouri, M. E. H., Lahiani, A., & Nguyen, D. K. 2015. World gold prices and stock returns in China: Insights for hedging and diversification strategies. Economic Modelling, 44, 273-282.

Bampinas, G., & Panagiotidis, T. 2015. Are gold and silver a hedge against inflation? A two-century perspective. International Review of Financial Analysis, 41, 267-276.

Jabeur, S. B., Mefteh-Wali, S., & Viviani, J.-L. 2021. Forecasting gold price with the XGBoost algorithm and SHAP interaction values. Annals of Operations Research. https://doi.org/10.1007/s10479-021-04233-1

Livieris, I. E., Pintelas, E., & Pintelas, P. 2020. A CNN–LSTM model for gold price time-series forecasting. Neural Computing and Applications, 32(23), 17351-17360.

Madziwa, L., Pillalamarry, M., & Chatterjee, S. 2022. Gold price forecasting using a multivariate stochastic model. Resources Policy, 76, 102544.

Makala, D., & Li, Z. 2021. Prediction of gold price with ARIMA and SVM. Journal of Physics: Conference Series, 1767(1), 012022. https://doi.org/10.1088/1742-6596/1767/1/012022

Nguyen, Q. N., Bedoui, R., Majdoub, N., Guesmi, K., & Chevallier, J. 2020. Hedging and safe-haven characteristics of gold against currencies: A multivariate dynamic copula approach. Resources Policy, 68, 101766.

Sadorsky, P. 2021. Predicting gold and silver price direction using tree-based classifiers. Journal of Risk and Financial Management, 14(5), 198.

Tripurana, N., Kar, B., Chakravarty, S., Paikaray, B. K., & Satpathy, S. 2022. Gold price prediction using machine learning techniques. In Proceedings of ACI@ISIC 2022: Advances in Computational Intelligence Workshop (pp. 274-281). CEUR-WS.

Yuan, F.-C., Lee, C.-H., & Chiu, C. 2020. Using market sentiment analysis and genetic algorithm-based least squares support vector regression to predict gold prices. International Journal of Computational Intelligence Systems, 13(1), 234-246.

Yurtsever, M. 2021. Gold price forecasting using LSTM, Bi-LSTM, and GRU. European Journal of Science and Technology, 21, 1-8.