A Gold Futures Price Forecast Model Based on SGRU-AM

JINGYANG WANG [1], **YIFAN LI** [1], **TINGTING WANG** [1], **JIAZHENG LI** [1], **HAIYAO WANG** [2], **AND PENGFEI LIU** [3]

[1] School of Information Science and Engineering, Hebei University of Science and Technology, Shijiazhuang 050018, China
[2] School of Ocean Mechatronics, Xiamen Ocean Vocational College, Xiamen 361100, China
[3] Center for Information Technology, Hebei University of Chinese Medicine, Shijiazhuang 050200, China

Corresponding author: Pengfei Liu (liupengfei@hebcm.edu.cn)

ABSTRACT As the leading component of the financial market, the price formation mechanism of gold futures has been attracting extra attention of financiers and scholars. However, the data of gold futures price belongs to time series, and its forecast is very challenging owing to its chaotic, noisy, and non-stationary characteristics in data. A new forecast model named SGRU-AM, based on the special gated recurrent unit (SGRU) and attention mechanism (AM) is proposed in this paper to tackle these challenges. SGRU is the rectified model of gated recurrent unit (GRU) though executing the 1-tanh function on the reset gate output of the basic GRU to transform the value range of the reset gate value and adjusting the memory ratio between the current moment and the previous moment. Firstly, SGRU has the advantage of capturing long-distance information and can forecast the closing price of gold futures in the next trading day. Then, AM is introduced to adjust the SGRU time dimension's feature expression, so that the model can obtain more comprehensive feature information, learn the importance of current local sequence features and improve the forecast accuracy. Taking China's gold futures as an example, the gold futures data of the Shanghai Futures Exchange are selected from January 9, 2008, to May 31, 2021. Compared with the baseline methods, the experimental results show that SGRU-AM has the best performance among all baseline models in forecast efficiency and forecast accuracy.

INDEX TERMS GRU, attention mechanism, gold futures, time series forecast.

I. INTRODUCTION

In recent years, the world economic market has been turbulent and changeable. Gold has the dual attributes of currency and commodity and has a certain role in avoiding risks. Therefore, gold futures have become an effective investment method, and studying and forecasting the gold futures price has important practical significance.

The research on financial time series forecast has been the research direction of scholars at home and abroad for a long time. As for the gold futures price forecasting, it mainly includes exploring the influencing factors of gold futures price fluctuation [1], [2], [3] and forecasting gold futures price and its instability. The first classical

The associate editor coordinating the review of this manuscript and approving it for publication was Aasia Khanum [ID].

time series methods used in the field of financial time series analysis include autoregressive moving average model (ARMA) [4]–[6], generalized autoregressive conditional heteroskedasticity (GARCH) [7], etc. In addition to the classical econometric time series forecasting methods, decision tree [8]–[10], genetic algorithm [11], support vector machine (SVM) [12] and other machine learning methods have also been implemented to financial price forecasting [13]. It should be noted that the price of gold futures presents a very complex nonlinear trend. The traditional linear model forecast method usually has various difficulties, and the shallow neural network also has many limitations and shortcomings in time series forecast. In recent years, the research method of deep learning has developed rapidly, and has been successfully applied to many fields [14], [15]. As for the processing of financial time series forecast, deep

learning can better describe various kinds of highly nonlinear changes and has achieved good performance in quantitative investment, risk control, and other aspects.

In the prediction of financial time series, there are differences in the degree of influence or importance of different characteristics in the series in the financial time series prediction. Due to the differences in the influence or importance of different features on the series, important significant features often contain more information and have a greater impact on the future trend. Important salient features tend to contain more information and have a more significant greater impact on future trends. The features of attention mechanism enable GRU to pay more attention to the features of high importance, which can better achieve more effective extraction in short-term mode and avoid the problem of information loss caused by too-long time sequence.

Therefore, a new neural network of financial time series forecast model is proposed, it combines SGRU and AM and is denoted as the SGRU-AM. The model is fully aware of the factors among gold market factors and the measurable economic indicators of the relevant economic market. The feature set is constructed from these features, as the input of the model. The SGRU with a higher update efficiency is used for forecast, and the AM is introduced into SGRU to optimize the forecast. The AM is used for the hidden state training weight of all SGRU time steps to improve the model's ability to process information. In the experiment, the SGRU-AM model proposed in this paper forecasts the next trading day's closing price of gold futures. To verify the feasibility and effectiveness of the SGRU-AM, it is compared with the forecast results of the MLP, RNN, LSTM, GRU, SGRU, CNN-LSTM [16], and GRU-AM.

The following descriptions are the innovations and major contributions of this paper:

1) SGRU is a special structure of GRU. The 1-tanh function is executed on the reset gate output of the basic GRU to transform the value range of the reset gate value, and the memory ratio between the current moment and the previous moment is adjusted. The SGRU-AM network model is proposed in this paper, where SGRU is used with its simplicity of the structure to forecast the gold futures next trading day's closing price, and AM can optimize SGRU forecast results by obtaining the importance of different characteristics.

2) Applying the AM to SGRU in the field of financial timing sequence forecast can obtain the timing sequence information comprehensively and learn the importance of local time sequence characteristics. By enhancing the impact of vital time features and reducing the influence of unnecessary characteristics, the forecast efficiency of the SGRU can be effectively advanced.

3) Under the same data set and experimental conditions, the proposed model SGRU-AM is parallel with the benchmark models. Benchmark models include single models such as MLP, RNN, LTSM, GRU, and SGRU, and hybrid models, including CNN-LSTM

and GRU-AM. The higher forecast accuracy of SGRU-AM is verified by experiments.

II. RELATED WORK

In recent years, it has been prevalent to apply machine learning and deep learning to time series forecast in the scientific community. Using emotional news mining combining with deep learning networks, Peng and Jiang [17] predicted the trend of rising and falling stock price. They used a multilayer perceptron model (MLP). Eğrioğlu and Fildes [18] combined with an identically independent distributed residual bootstrap method and proposed a new hybrid ANN architecture to obtain probabilistic result.

Chen *et al.* [19] applied LSTM to stock price forecasting in China's stock market, and its forecasting effect was greatly improved. Kim and Won proposed a new hybrid LSTM model which combined LSTM with various GARCH models to forecast stock price volatility. This model enhances forecast performance by combining the neural network model with multiple econometric models [20]. Jin used empirical mode decomposition (EMD) and LSTM to forecast stock price. At the same time, they analyzed the investor sentiment and added the investor sentiment factor to the forecast model, which further improved the forecast accuracy [21]. Su improved the LSTM algorithm named RFG-LSTM, and used the China's A stock market investment market to prove the effectiveness of the RFG-LSTM [22].

Qin proposed a two-stage cyclic neural network (DA-RNN) based on the attention mechanism. In the first stage (encoder), the input attention mechanism was introduced to extract the correlation adaptively for the external input at each moment. In the second stage (decoder), a temporal attention mechanism was introduced to capture the encoder's long-term temporal dependent information [23]. Moreover, Hu and Zheng proposed a new multistage attention network to capture the different influences. It consisted of the influential attention mechanism and temporal attention mechanism [24].

E and Ye used the ICA-GRU hybrid network to predict the gold price. The effectiveness of the proposed method for the gold price forecast was verified [25]. The results also show that GRU replacing LSTM has better performance in forecast accuracy and time. And compared with ICA-LSTM, ICA-GRU performs better. Andrés and Werner used the CNN-LSTM model to complete the prediction of gold volatility [26]. The CNN-LSTM model could capture static and dynamic data by converting the time series into images. Livieris proposed a new deep learning forecast model to forecast the U.S. gold price and its changes [16]. The CNN in the model had a difference in obtaining useful information and receiving time-series data's internal representation. The model used LSTM to forecast the gold price. The performance of CNN-LSTM in forecasting the US gold price and its volatility was impressive. Moreover, it verified that the hybrid model has better forecast accuracy. It should be noted that the optimal model parameters for forecasting gold price and

gold price fluctuation are different, and the optimal model parameters are also illustrated.

The SGRU-AM hybrid network model proposed in this paper firstly uses SGRU to forecast the gold futures the next trading day's closing price, and then AM is to optimize SGRU forecast results by obtaining the importance of different characteristics and improve the information processing capability. Thus, the model completes the forecast of the closing price of gold futures. And the first 10 consecutive trading days are used to forecast the next trading day's the closing price. The experimental results show that the SGRU-AM model can provide valuable information on future research in this field of financial time series forecast.

III. THE NOVEL HYBRID MODEL
A. SGRU
GRU, which is known as an improved model of LSTM networks, is a proposed model for improving the RNN gradient explosion and gradient messaging problems. Cho *et al.* [27] proposed GRU and applied it in machine translation. Compared with LSTM, GRU not only has a simple structure and higher forecast accuracy but also greatly reduces the forecast time. GRU can remember previous information and deal with long-term dependency problems effectively. GRU has no cell state, but it has gates that are used to control the influence of new information on the output value. GRU optimizes the gate function structure of LSTM, reduces the number of gates from three to two while mixing the neuron state and the hidden state, can effectively alleviate the problem of "gradient disappearance" in the RNN network. Moreover, compared with LSTM, GRU reduces replacement, and the model's the training time is shorter. GRU has two gates: update gate and reset gate. The reset gate determines how the new input information is combined with the previous memory. Update gate defines how much of the previous memory is saved to the current time step.

The improved model of GRU is named SRGU. Firstly, the gate mechanism is rectified based on the GRU. The 1-tanh function is used to transform the value range of the reset gate. This modification can effectively prevent oversaturation and improve the model learning efficiency. Secondly, the memory ratio between the current moment and the previous moment is adjusted, so that information could be better remembered. Thus, the SGRU has a higher update efficiency compared with GRU. The Fig. 1 is the SGRU memory cell's architecture.

The process of the SGRU memory module for a status update and information output is as follows:

$$z_t = \sigma \left(W_z \cdot h_{t-1} + U_z \cdot x_t + b_z \right) \quad (1)$$

$$r_t = \sigma \left(W_r \cdot h_{t-1} + U_r \cdot x_t + b_r \right) \quad (2)$$

$$tr_t = 1 - tanh \left(r_t \right) \quad (3)$$

$$\tilde{h}_t = tanh \left(W_h[tr_t * h_{t-1}] + U_h \cdot x_t + b_h \right) \quad (4)$$

$$h_t = z_t * h_{t-1} + (1 - z_t) * \tilde{h}_t \quad (5)$$
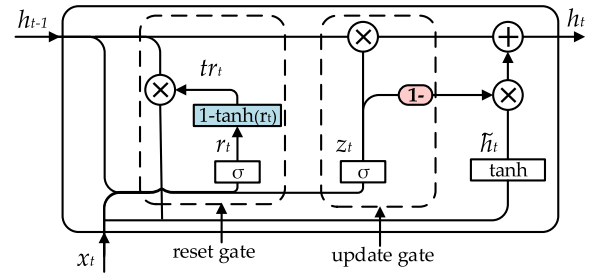

FIGURE 1. Architecture of GRU memory cell.

where, x_t is the input vector at time t, h_{t-1} is the state memory variable at time $t-1$, z_t is the value of the update gate, h_t is the state memory variable at time t, $\tilde{h}_t$ represents the state of the current state of the candidate set, W_t, W_z, W_h, U_t, U_z, U_h and b_t, b_z, b_h represent the weight matrix and bias vector respectively. σ is the logistic sigmoid function, $*$ is the dot product of two vectors.

The reset gate will determine how much of the last hidden state goes into the following process. In other words, it resets the cell's hidden state at time t-1, so this gate function is called the reset gate. The r_t in Eq. (2) is obtained by executing sigmoid function transformation through x_t and h_{t-1}. Then, Eq. (3) presents the process of obtaining tr_t. This rectification changes the degree of information preservation and can prevent the occurrence of oversaturation. r_t is the reset gate value of the basic GRU, but tr_t is the reset gate value of the SGRU. Eq. (1) represents the value of update gate processed at time t. Eq. (4) represents the process that obtained $\tilde{h}_t$. Update gate z_t will retain information about the current cell and pass it to the next cell. It should be noted that the ratio between the current moment and the previous moment is adjusted as shown in Eq. (5).

From the above information processing, SGRU can also remember the previous information well. SGRU uses the update gate and the reset gate, which control the final output information of the SGRU memory cell at the current time. Instead of discarding previous information over time, SGRU retains relevant information and passes it on to the next unit, so it takes advantage of all the information and avoids the gradient extinction problem. Moreover, 1-tanh function is used to transform the value range of the reset gate, so that the dependence between short-term information has been strengthened, while long-term information can also be better preserved. So, the possibility of SGRU oversaturation is less.

B. AM
In 1980, Treisman and Gelade proposed the AM based on the study of human vision. In order to reasonably utilize the limited resources of visual information processing [28], human beings need to select a specific part of the visual region and then focus on it. The AM can optimize the traditional model by calculating the probability distribution of attention, selecting more critical information from a large amount of information, and highlighting essential parts of the input

The calculation process of AM is as follows:

1) Attention mechanism calculates the weighted sum of the input vector expression, and the weight indicates the importance of the feature at each time point. Suppose the input is K feature vectors, h_i, $i = 1, 2, \ldots, k$. The attention mechanism can obtain the environment vector based on h_i. As shown in formula (6):

$$c_i = \sum_{i=1}^{k} a_i h_i \quad (6)$$

where the weight attached to the state is the weight of attention a_i.

2) The input vector was trained through a fully connected network to obtain a_i, and the score s_i of each hidden layer vector is calculated to evaluate its influence on the output. As shown in formula (7):

$$s_i = tanh\left(W^T h_i + b_i\right) \quad (7)$$

where W is the weight of AM, b_i is the bias of AM, h_i is the input vector, and W and b_i are shared weights in each layer.

3) The softmax function is used to normalize the score s_i to obtain the final weight coefficient a_i. As shown in formula (8):

$$a_i = softmax(s_i) \quad (8)$$

C. SGRU-AM

The model constructed in this paper consists of two parts: SGRU and AM. As a special structure of GRU, SGRU makes sequence forecast of input feature components. Then, the AM automatically fits the weight distribution of SGRU. It also multiplies and sums the output vector of the hidden layer of the SGRU at different time points with the corresponding weight, and SGRU can give more weight to the important feature components. Through these processes, the final characteristic expression of the model is the optimized forecast. Finally, the AM covers the output of each cell, which makes the model obtain more comprehensive and detailed feature information. The information processing capability of the model has been improved by combining with AM. The structure diagram of the SGRU-AM model is shown in Fig. 2.
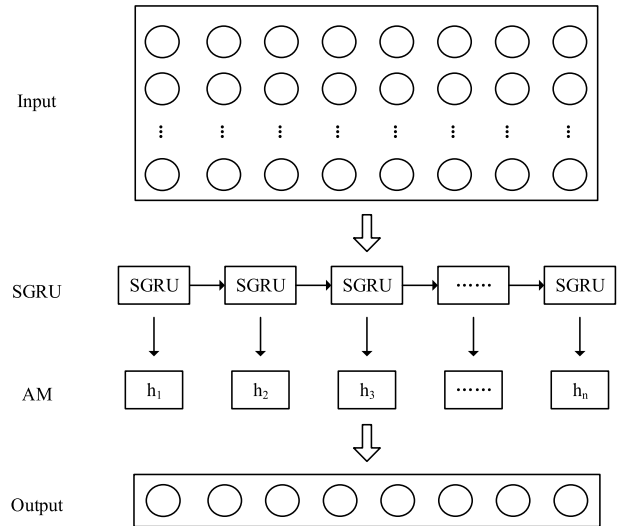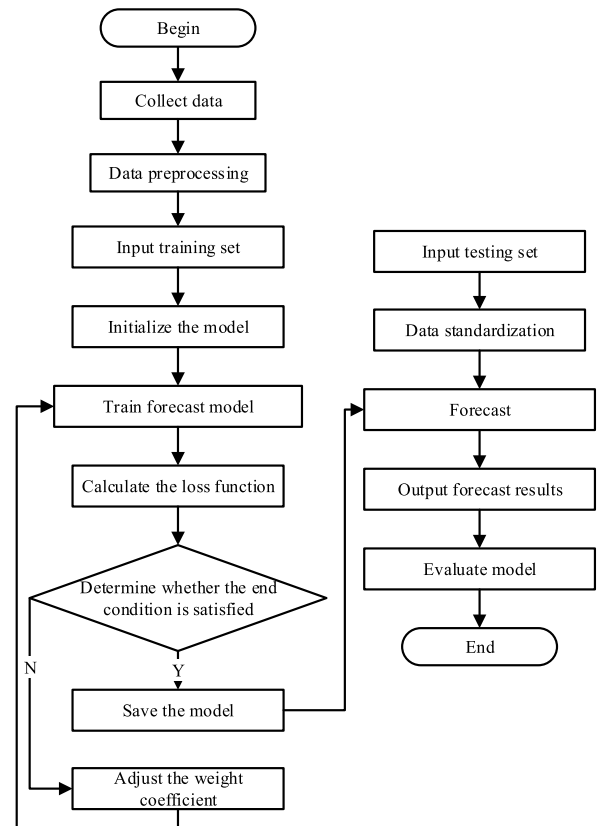
IV. EXPERIMENTS

A. EXPERIMENTAL PROCESS

The flow diagram of the experimental process is shown in Fig.3.

The main process is as follows:

1) Collect data: It is necessary to collect various data items used in the experiment.

2) Data preprocessing: Data preprocessing on the collected data is performed, including operations such as data cleaning, data integration, data standardization, and data set division. The z-score standardization method is adopted to standardize data, as shown

FIGURE 2. Structure diagram of SGRU-AM model.

FIGURE 3. Flow diagram of the experimental process.

in Eq. (9).

$$x_{norm} = \frac{x_i - \bar{x}}{s} \quad (9)$$

where, x_{norm} is the standardized value, x_i is the input data, $\bar{x}$ is the average of the input data, and s is the standard deviation of the input data.

Influence factor	Feature symbol	Feature description
Gold market factors	AUL_SP	the settlement price of the main gold futures market
	AUL_VOL	the trading volume of the main gold futures market
	AUL_POS	the open position of the main gold futures market
	AU99.99	the gold spot price in China
Relevant economic measurable factors	DJIA	Dow Jones Industrial Average
	SP500	Standard & Poor's 500 Index
	IXIC	Nasdaq Composite Index
	USN-CNY	USD/RMB exchange rate
	USDX	US Dollar Index

3) Input training set: The training data set required for model training is inputted.

4) Initialize the model: The parameters in the model, number of iterations, learning rate, and other parameters are set.

5) Train forecast model: The forecast model with set parameters is trained on the training set data.

6) Calculate the loss function: The loss function of the model is calculated.

7) Determine whether the end condition is satisfied: It is judged whether the model meets the end condition. If the condition is met, it means that the current model is the trained forecast model, and then continue to the step 8; if the condition is not satisfied, jump back to the step 9.

8) Save the model: The trained forecast model is saved.

9) Adjust the weight coefficient: The model weight coefficient is adjusted, jump back to the step 5.

10) Input testing set: The testing data set needed for model training is inputted.

11) Data standardization: The data of the testing set are standardization.

12) Forecast: The standardized testing data set is inputted into the trained model to make forecast.

13) Output forecast results: The forecast results are obtained and output.

14) Evaluate model: Three indicators of root mean square error (RMSE), the mean absolute error (MAE) and R-square (R^2) are used to evaluate different models' forecast effects. The Equations of the three indicators are as follows:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2} \tag{10}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|\hat{y}_i - y_i| \tag{11}$$

$$R^2 = 1 - \frac{\left(\sum_{i=1}^{n}(y_i - \hat{y}_i)^2\right)/n}{\left(\sum_{i=1}^{n}(\bar{y}_i - \hat{y}_i)^2\right)/n} \tag{12}$$

where $\hat{y}_i$ is the predicted value and y_i is the real value. The value of R^2 is between 0 and 1. The smaller the MSE and MAE values, the closer the R^2 value is

to 1, which indicates that the prediction accuracy of the model is higher.

B. DATASETS

Data series of gold futures price is the time-series data, which can be analyzed and forecasted based on its historical price. A reasonable construction of the forecast feature set is vital to the model's convergence and forecast accuracy. Moreover, many factors are affecting the price fluctuation of gold futures. Through existing studies and analysis of the current market, the economic factors affecting the price of gold futures mainly come from two aspects. On the one hand, the fluctuation of the gold market would affect the gold futures price fluctuation. The factors selected from the gold market include the daily settlement price, the open position [29], the trading volume [30], and the gold spot price in China [31], [32]. On the other hand, the supply and demand of gold would fluctuate with the change of the economy, so the price of gold futures will be affected by the macro-economy. As for the relevant economic impact of other markets, the measurable factors selected in this paper include the Dow Jones Industrial Average (DJIA), Standard & Poor's 500 Index, Nasdaq Composite Index, USD/RMB exchange rate, and US Dollar Index (USDX).

To sum up, a total of 9-dimensional feature vectors, including gold market factors and relevant economic measurable factors are constructed as the input of the forecast model, as shown in Table 1. Data are collected once a day, and the input characteristic dimension is 9, as shown in Table 2.

The experimental data are obtained from the Wind database and the financial website (https://cn.investing.com). The main contract of gold futures on the Shanghai Futures Exchange is selected and the date range is from January 9, 2008, to May 31, 2021. Excluding holidays and weekends, a total of 3245 days of data as a data set. According to the ratio of 8:2, the data is divided into training set and testing set, China's gold futures closing prices are shown in Fig. 4.

C. EXPERIMENTAL SETUP

Firstly, all experiments are carried out in the same test environment. The experiment is taken on Jupyter notebook based on Python 3.7 interpreter under Windows 10, 64-bit operating

Data	AUL_SP	AUL_CP	AUL_VOL	AUL_POS	AU99.99	DJI	SP500	IXIC	USN-CNY	USDX
2008/1/9	224.74	223.30	103364	16342	208.80	7.2636	76.42	12590.21	1409.13	2474.55
2008/1/10	217.68	215.12	87984	23494	205.60	7.2714	75.93	12733.11	1420.33	2488.52
2008/1/11	217.79	218.10	85418	25424	208.39	7.2625	75.98	12850.74	1401.02	2439.94
2008/1/14	219.28	220.10	60776	27714	211.19	7.2515	75.61	12613.78	1416.25	2478.3
2008/1/15	223.11	223.80	88242	29866	212.30	7.2421	75.57	12777.50	1380.95	2417.59
2008/1/16	217.51	215.89	94362	29582	206.8	7.2324	76.26	12476.81	1373.2	2394.59
2008/1/17	215.5	216.47	61962	30916	207	7.2486	76.17	12467.05	1333.25	2346.9
2008/1/18	213.72	214.7	58006	29538	205.88	7.2423	76.38	12159.94	1325.19	2340.02

FIGURE 4. Daily gold price trend from January 9, 2008 to May 31, 2021.

TABLE 3. Parameter setting of SGRU-AM.

Parameters	value
Number of hidden units in SGR-GRU layer	8.60777
SGRU layer activation function	7.96892
time_step	10
batch_size	64
learning rate	0.001
optimizer	Adam
loss function	mean_absolute_error
Epochs	100

FIGURE 5. Convergence of GRU, SGRU, GRU-AM and SGRU-AM convergence.

system. The model is built and implemented based on the TensorFlow framework. The TensorFlow version is 1.14.0. The hardware configuration is Intel (R) Core (TM) i5-4300u CPU@1.90GHz 2.49 GHz, memory 8GB.

Additionally, the parameter setting of SGRU-AM is shown in Table 3. The SGRU-AM model is composed of a layer of SGRU and combines the AM layer to forecast the gold futures' closing price. The forecasted value is the next trading day's closing price.

D. MODEL CONVERGENCE

It is should be attached attention to verify the convergence of the model after the model parameters have been set. The convergence comparison is shown in Fig. 5. When all parameters are the same, the convergence of GRU, SGRU, GRU-AM, and SGRU-AM is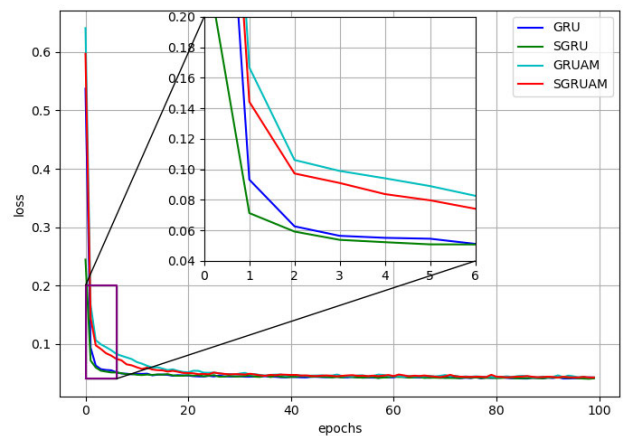 compared. In the first 6 training times, SGRU has a faster convergence speed than GRU. Moreover, the convergence speed of the single model is faster than that of the hybrid model, and the convergence speed of SGRU-AM is faster than GRU-AM. When the training reaches 6 times, the convergence speeds of GRU and SGRU are basically the same and tend to be stable; when the number of training times exceeds 20, the convergence speeds of the four models tend to be stable and the speeds are equal. From the overall analysis of the models, the convergence effect of SGRU is better than the other three. The modification of GRU has a certain improvement in convergence performance.
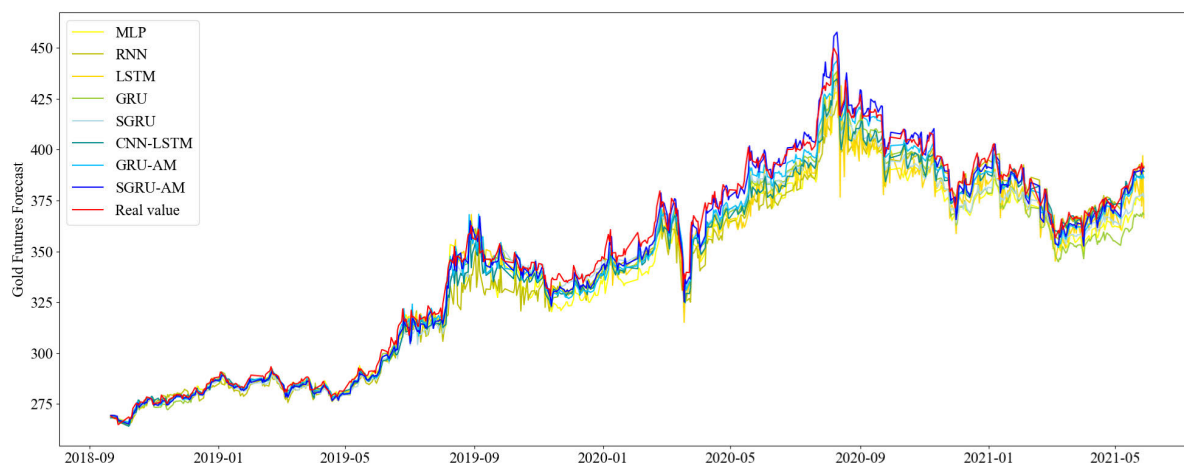
FIGURE 6. Comparison of forecasted value and real value.

E. FORECAST COMPARISON

In order to verify the effectiveness and superiority of the SGRU-AM model, the performance of the SGRU-AM model is compared with benchmark models. Furthermore, the benchmark models have consisted of single models such as MLP, RNN, LSTM, GRU, and SGRU, and hybrid models, including CNN-LSTM, and GRU-AM. Among the above models, the CNN-LSTM model is the gold price forecast model proposed in the paper [16]. It is worth noticing that the training and testing of each model are carried out in the same environments.

The forecast effect of SGRU-AM is analyzed and compared with that of the single model and mixed model. The forecast effect of the model is evaluated by the three evaluation indicators MAE, RMSE and R^2, and the results are shown in Table 4. MLP has the worst performance among the deep learning models, it has the largest value of MAE and RMSE, and the smallest value of R^2. SGRU-AM has the best forecast effect and can provide a reliable reference for investors. SGRU performed best in the single model. Compared with RNN, the MAE value of SGRU decreases from 8.12228 to 6.87255, the RMSE value of SGRU decreases from 10.86029 to 8.52930, and its R^2 value increases from 0.94631 to 0.96688. Compared with GRU, the MAE value of SGRU decreases from 7.08941 to 6.87255; the RMSE value of SGRU decreases from 9.16237 to 8.52930 and the R^2 value of SGRU increases from 0.96178 to 0.96688. Furthermore, the performance of the hybrid models is better than that of the single models. Moreover, among the hybrid models, the SGRU-AM model performs the best. Compared with GRU-AM, the MAE value of SGRU-AM decreases from 4.58940 to 3.59882; the RMSE value of SGRU-AM decreases from 5.88732 to 4.79393 and the R^2 value of SGRU-AM increases from 0.98422 to 0.98953.

Moreover, Table 5 presents a comparison of the time spent on model training and testing. It can be concluded from Table 5 that the single model takes less time than the

TABLE 4. Comparison of model forecast results.

Model	MAE	RMSE	R^2
MLP	9.27862	11.46375	0.94018
RNN	8.12228	10.86029	0.94631
LSTM	7.21362	9.70735	0.95710
GRU	7.08941	9.16237	0.96178
SGRU	6.87255	8.52930	0.96688
CNN+LSTM	5.48835	7.18771	0.97648
GRU-AM	4.58940	5.88732	0.98422
SGRU-AM	**3.59882**	**4.79393**	**0.98953**

TABLE 5. Comparison of training and testing time of different models.

Model	Training Time / s	Testing Time / s
MLP	24.08354	0.70301
RNN	28.93133	0.78345
LSTM	63.62535	1.30838
GRU	46.40687	0.61164
SGRU	48.58915	0.78999
CNN+LSTM	217.40810	2.15040
GRU-AM	95.56980	1.78236
SGRU-AM	**97.00768**	**1.85847**

mixed model. In the mixed model, GRU-AM consumes the shortest time, with the training time of 95.56980 and the testing time of 0.79065, while the training time of SGRU-AM is 97.00768 and the its testing time is 1.85847.

Additionally, for a more intuitive comparison, the forecasted values of all models and real value are plotted as shown in Fig. 6. Furthermore, in order to display the results more clearly, the data after September 1, 2020 are selected to draw a comparison diagram of the forecast results of different models, as shown in Fig. 7. The legends of the Fig. 6 and Fig. 7 are the abbreviations of corresponding models.

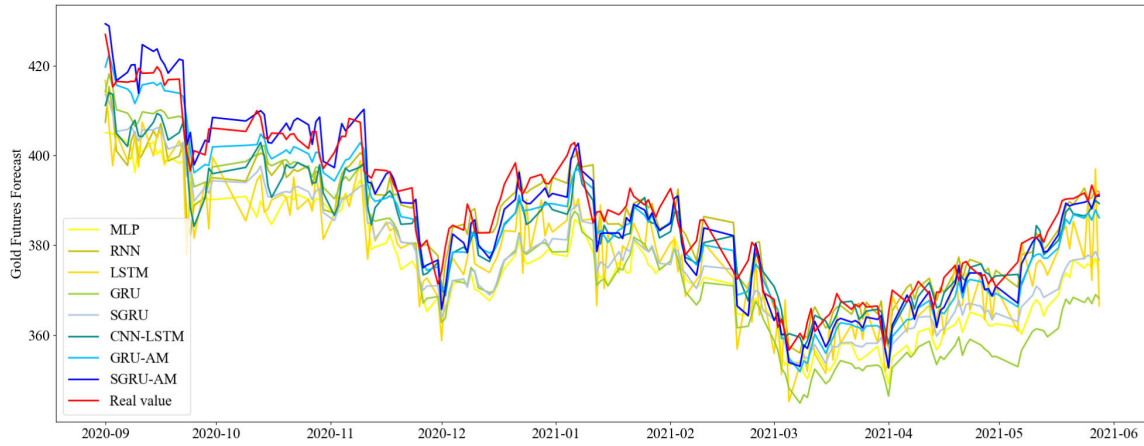Moreover, the SGRU-AM model is applied to predict the closing price of U.S. gold futures. The data from

FIGURE 7. Comparison of forecasted value and real value after September 1, 2020.
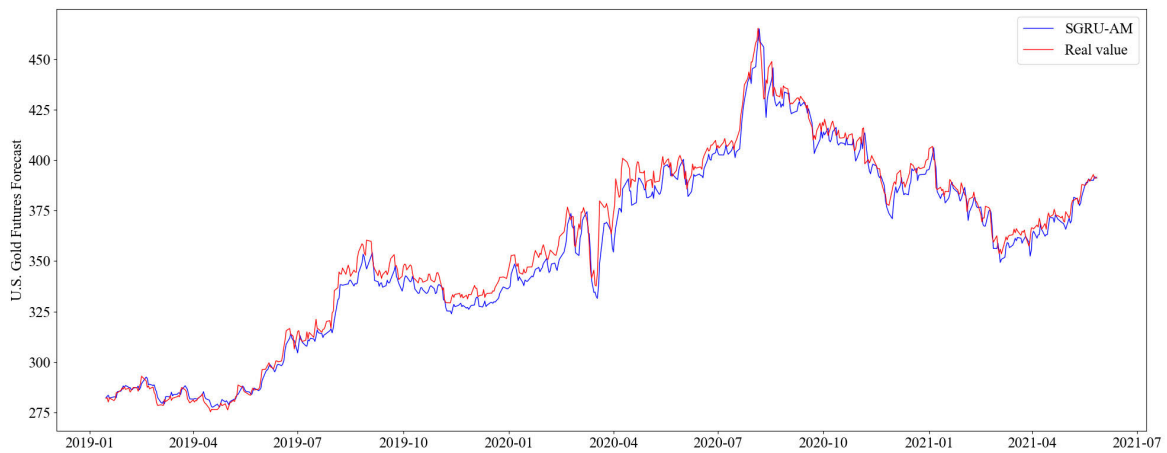

FIGURE 8. Comparison of U.S. Gold Futures real value and SGRU-AM forecasted value.

January 1, 2019 to May 31, 2020 of the U.S. COMEX exchange is selected as the testing data set to verify the generalization ability of the SGRU-AM model. The MAE value of SGRU-AM is 4.78802, the RMSE value of SGRU-AM is 6.16506, the value of R^2 is 0.98196. The comparison of the predicted value and the real value is shown in Fig. 8. In Fig. 6-8, the abscissa represents time, and the ordinate represents the value of the gold futures settlement price (the unit of ordinate is yuan).

F. EXPERIMENTAL ANALYSIS

It's obvious that the worst performing model among the deep learning models is MLP with the MAE value of 9.27862, RMSE value of 11.46375 and R^2 value of 0.94018, while the SGRU-AM has the best predictive effect among all models, its MAE value is 3.59882, the value of RMSE is 4.79393, and the value of R^2 is 0.98953. In addition, as a special structure of GRU, SGRU have better forecast accuracy than GRU. Moreover, although hybrid model costs more time, it is worth noticing that the hybrid model performs better than the

single deep learning model, which reduces the fitting error and improves the forecast accuracy. Compared with GRU-AM, the MAE value of SGRU-AM is reduced by 27.51%, the RMSE value is reduced by 22.8%, and the R^2 value is increased by 0.53%.

It can be concluded from Fig. 6 that the forecast accuracy of SGRU-AM is the highest among all the models, and the distance between the forecast value of the SGRU-AM and the real value is the smallest. As shown in Fig. 7, when there is a large fluctuation, it is the time to test the forecast performance of the model. It should be noted that with the increase of information amount, the forecast accuracy of all forecast models is declining, but in this case, SGRU-AM also performs best. It is obvious from Fig. 8 that SGEU-AM performed very well in forecasting the closing price of U.S. gold futures. This shows that SGRU-AM has a good generalization ability.

Based on the empirical analysis of gold futures price forecast and the double aspects of forecast accuracy and forecast time consumption, the SGRU-AM model proposed in this paper has a good effect on the gold futures price forecast.

In other words, the application of SGRU-AM has a certain reference value for grasping the trading opportunity and avoiding risks of gold futures.

V. CONCLUSION

In this paper, a novel SGRU-AM model is proposed and applied to forecast the gold futures closing price. SGRU has a simple structure and can solve the problem that traditional neural networks cannot remember and use historical information. AM is introduced to regulate the feature expression of GRU time dimension, so that the model can obtain more comprehensive feature information and learn the importance of current local sequence features. Selecting the gold market's factors and other markets' relevant economic measurable factors which affect the price volatility of gold futures construct the feature set. MLP, RNN, LSTM, GRU, SGRU, CNN-LSTM, and GRU-AM are used as benchmark models to forecast the gold futures closing price in China respectively. The main findings of the paper are as follows:

1) SGRU is the rectified model of GRU though executing the 1-tanh function on the reset gate output of the basic GRU to transform the value range of the reset gate value and adjusting the memory ratio between the current moment and the previous moment. The use of SGRU can fully consider the timing characteristics that affect the variation characteristics of gold futures prices, and it has good time series data fitting regression ability and high forecast efficiency.

2) SGRU-AM combines the advantages of SGRU and AM. Firstly, the SGRU model has a higher update efficiency compared with GRU. And the SGRU-AM improves information processing capabilities by combined with AM to give heavier weight to more important feature information.

3) The experimental results show that compared with the benchmark models, the SGRU-AM model proposed in this paper has a higher forecast accuracy while maintaining a faster model training speed.

To sum up, using SGRU-AM to forecast the gold futures closing price of the next trading day can help investors reduce investment risks and seek investment opportunities. However, there are many factors affecting the price change of gold futures, because the SGRU-AM model does not fully consider the impact of other external factors, so the model needs to be further improved. For example, it's worth trying to pay more attention to various factors, such as the futures market's own system characteristic, the national macroeconomic policy, and the development direction of the national economy. In further research, it can be considered that adding investor sentiment analysis to optimize the feature set so as to further optimize the model. In addition, the validity of SGRU-AM has been verified in forecasting the timing data of gold futures. Furthermore, it is worth exploring its forecast effect on other futures prices and financial time series forecast yield.

REFERENCES

[1] S. Liu and Z. Hu, "An analysis the stability of long-run determinants of the gold price," *World Econ. Study*, vol. 168, no. 2, pp. 35–41 and 87, Feb. 2008.

[2] H. Cao and S. Zhang, "Study on price volatility of China's gold futures market," *Prices Monthly*, vol. 2, pp. 1–5, Feb. 2012.

[3] L. A. Smales, "News sentiment in the gold futures market," *J. Banking Finance*, vol. 49, pp. 275–286, Dec. 2014.

[4] M. Askari and H. Askari, "Time series grey system prediction-based models: Gold price forecasting," *Trends Appl. Sci. Rese.*, vol. 6, no. 11, p. 1287, Nov. 2011.

[5] A. A. Adebiyi, A. O. Adewumi, and C. K. Ayo, "Comparison of ARIMA and artificial neural networks models for stock price prediction," *J. Appl. Math.*, vol. 2014, Mar. 2014, Art. no. 614342.

[6] J. He, Y. Yu, N. Liu, S. Zhao, B. Chen, and L. Yu, "Modeling and analyzing liquidity in stock market using macroeconomic factors based on state space model," *Chin. J. Manage. Sci.*, vol. 5, pp. 42–49, May 2019.

[7] S. A. Basher and P. Sadorsky, "Hedging emerging market stock prices with oil, gold, VIX, and bonds: A comparison between DCC, ADCC and GO-GARCH," *Energy Econ.*, vol. 54, pp. 235–247, Feb. 2016.

[8] P.-C. Chang, C.-Y. Fan, and J.-L. Lin, "Trend discovery in financial time series data using a case based fuzzy decision tree," *Expert Syst. Appl.*, vol. 38, no. 5, pp. 6070–6080, May 2011.

[9] R. C. Cavalcante, R. C. Brasileiro, V. L. F. Souza, J. P. Nobrega, and A. L. I. Oliveira, "Computational intelligence and financial markets: A survey and future directions," *Expert. Syst. Appl.*, vol. 55, pp. 194–211, Aug. 2016.

[10] B. Soybilgen and E. Yazgan, "Nowcasting US GDP using tree-based ensemble models and dynamic factors," *Comput. Econ.*, vol. 57, no. 1, pp. 387–417, Jan. 2021.

[11] N. Srinivasan and C. Lakshmi, "Stock price prediction using fuzzy time-series population based gravity search algorithm," *Int. J. Softw. Innov.*, vol. 7, no. 2, pp. 50–64, Apr. 2019.

[12] Y. Guo, S. Han, C. Shen, Y. Li, X. Yin, and Y. Bai, "An adaptive SVR for high-frequency stock price forecasting," *IEEE Access*, vol. 6, pp. 11397–11404, 2018, doi: 10.1109/ACCESS.2018.2806180.

[13] S. Urolagin, "Text mining of tweet for sentiment classification and association with stock prices," in *Proc. Int. Conf. Comput. Appl. (ICCA)*, Doha, Qatar, Sep. 2017, pp. 384–388, doi: 10.1109/COMAPP.2017.8079788.

[14] W. Lu, J. Li, Y. Li, A. Sun, and J. Wang, "A CNN-LSTM-based model to forecast stock prices," *Complexity*, vol. 2020, Nov. 2020, Art. no. 6622927.

[15] J. Wang, J. Li, X. Wang, J. Wang, and M. Huang, "Air quality prediction using CT-LSTM," *Neural Comput. Appl.*, vol. 2020, pp. 1–14, Nov. 2020, doi: 10.1007/s00521-020-05535-w.

[16] I. E. Livieris, E. Pintelas, and P. Pintelas, "A CNN-LSTM model for gold price time-series forecasting," *Neural Comput. Appl.*, vol. 32, no. 23, pp. 17351–17360, 2020.

[17] Y. Peng and H. Jiang, "Leverage financial news to predict stock price movements using word embeddings and deep neural networks," *Comput. Sci.*, vol. 2015, pp. 1–5, Jun. 2015, doi: 10.18653/v1/n16-1041.

[18] E. Eğrioğlu and R. Fildes, "A new bootstrapped hybrid artificial neural network approach for time series forecasting," *Comput. Econ.*, vol. 4, pp. 1–29, Nov. 2020.

[19] K. Chen, Y. Zhou, and F. Dai, "A LSTM-based method for stock returns prediction: A case study of China stock market," in *Proc. IEEE Int. Conf. Big Data*, Oct. 2015, pp. 2823–2824.

[20] H. Y. Kim and C. H. Won, "Forecasting the volatility of stock price index: A hybrid model integrating LSTM with multiple GARCH-type models," *Expert Syst. Appl.*, vol. 103, pp. 25–37, Aug. 2018.

[21] Z. Jin, Y. Yang, and Y. Liu, "Stock closing price prediction based on sentiment analysis and LSTM," *Neural Comput. Appl.*, vol. 32, no. 13, pp. 9713–9729, Jul. 2020.

[22] Z. Su, H. Xie, and L. Han, "Multi-factor RFG-LSTM algorithm for stock sequence predicting," *Comput. Econ.*, vol. 57, no. 4, pp. 1041–1058, Apr. 2021.

[23] Y. Qin, D. Song, H. Chen, W. Cheng, G. Jiang, and G. W. Cottrell, "A dual-stage attention-based recurrent neural network for time series prediction," in *Proc. 26th Int. Joint Conf. Artif. Intell.*, Aug. 2017, pp. 2627–2633, doi: 10.24963/ijcai.2017/366.

[24] J. Hu and W. Zheng, "Multistage attention network for multivariate time series prediction," *Neurocomputing*, vol. 383, pp. 122–137, Mar. 2020.

[25] J. E. , J. Ye, and H. Jin, "A novel hybrid model on the prediction of time series and its application for the gold price analysis and forecasting," *Phys. A, Stat. Mech. Appl.*, vol. 527, no. 1, pp. 1–15, Aug. 2019.

[26] V. Andrs and K. Werner, "Gold volatility prediction using a CNN-LSTM approach," *Expert Syst. Appl.*, vol. 157, no. 1, pp. 1–9, Nov. 2020.

[27] K. Cho, B. Merrienboer, C. Gulcehre, and F. Bougares, "Learning phrase representations using RNN encoder-decoder for statistical machine translation," *Atmos. Environ.*, vol. 124, pp. 22–27, Nov. 2014.

[28] A. M. Treisman and G. Gelade, "A feature-integration theory of attention," *Cognit. Psychol.*, vol. 12, no. 1, pp. 97–136, Jan. 1980.

[29] H. F. Baklaci, Ö. Säer, and T. Yelkenci, "Volatility linkages among gold futures in emerging markets," *Emerg. Markets Finance Trade*, vol. 52, no. 1, pp. 1–9, Jan. 2016.

[30] Y. Jiang, N. Kellard, and X. Liu, "Night trading and market quality: Evidence from Chinese and US precious metal futures markets," *Environ. Pollut.*, vol. 40, no. 10, pp. 1324–1333, Jul. 2020.

[31] Q. Sha and X. Zhang, "Research on the influencing factors of gold futures market price in China," *Times Finance*, vol. 8, p. 261, Aug. 2013.

[32] G. Tang, H. Qu, and Q. Lao, "Research on the dynamic relationship between China's gold futures and spot prices–based on set empirical mode decomposition and sate space model," *Price, Theory Pract.*, vol. 6, pp. 112–115, Jun. 2019.

TINGTING WANG is currently pursuing the bachelor's degree in computer science and technology with the Hebei University of Science and Technology. Her research interests include machine learning and deep learning.
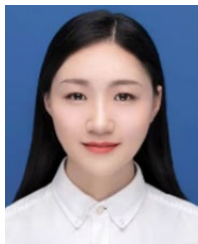
JIAZHENG LI is currently pursuing the master's degree with the Hebei University of Science and Technology. His research interests include machine learning and deep learning.

JINGYANG WANG received the B.Eng. degree in computer software from Lanzhou University, China, in 1995, and the M.Sc. degree in software engineering from the Beijing University of Technology, China, in 2007. He is currently a Professor with the School of Information Science and Engineering, Hebei University of Science and Technology, Shijiazhuang, Hebei, China. His research interests include machine learning, natural language processing, big data processing, and distributed computing.

HAIYAO WANG received the B.Eng. degree in mechanical design and manufacturing and the M.Sc. degree in industry engineering from the Hefei University of Technology, China, in 1998 and 2009, respectively. She is currently an Associate Professor with the School of Ocean Mechatronics, Xiamen Ocean Vocational College, Xiamen, Fujian, China. Her research interests include machine learning, process optimization, and efficiency improvement.

YIFAN LI is currently pursuing the master's degree with the Hebei University of Science and Technology. Her research interests include machine learning and deep learning.

PENGFEI LIU received the B.Eng. degree in computer science and technology and the M.Sc. degree in computer science from the Hebei University of Science and Technology, Shijiazhuang, Hebei, China, in 2012 and 2015, respectively. He is currently an Engineer with the Center for Information Technology, Hebei University of Chinese Medicine, China. His research interests include machine learning, natural language processing, big data processing, and distributed computing.

• • •