A Study on New Energy Stock Price Prediction Based on Error Compensation Hybrid Model

Tian-ci Ma, Yong Zhang*

Abstract—**Stock price prediction is crucial in the financial field. To address the challenge of accurately predicting new energy stock prices, this paper takes the stock price of new energy leading enterprise Contemporary Amperex Technology Co., Limited (CATL) as the research object, and proposes a new energy stock price prediction method based on the error compensation hybrid model. Firstly, to tackle the issues of insufficient information and interference in new energy stock prices, a multi-source dataset is constructed by integrating various indicators, and 22 key features are screened out by dimensionality reduction of Pearson's correlation coefficient method, while the Variational Mode Decomposition (VMD) method is used to decompose the highly complex and non-stationary stock price series. Secondly, after combining the Convolutional Neural Network-Attention Mechanism (CNN-ATT) module to extract the feature information, the prediction is performed by the Long Short-Term Memory (LSTM) model, and Support Vector Machine (SVM) is introduced for error compensation to improve the prediction accuracy of the model. Furthermore, the proposed ELG-BKA—a Black-winged Kite Algorithm variant, tested on CEC2022 functions, can autommatically optimize parameters of the VMD and LSTM. Finally, the validity of each module in this model is verified by orthogonal experiments on the CATL stock price dataset and the accuracy and generalization ability of the model is verified by comparing with various representative models such as LSTM from different evaluation metrics (MAE, MAPE, RMSE, R^2) on 7 new energy stock datasets. Experimental results demonstrate that the proposed method achieves high prediction accuracy and stability in forecasting new energy stock prices.**

Index Terms—**stock price prediction, multi-source data, error compensation, BKA**

I. Introduction

In recent years, stock price forecasting has been a hot academic research topic. Stock markets are significant for economic growth, corporate finance and investor asset allocation[1]. In the context of global climate change, carbon neutrality and energy transition, China vigorously support the new energy automobile industry, policy and funding under the dual drive, the industry in the technology and market level are leading the new energy sector[2]. China Association of Automobile Manufacturers data show that in 2024, China's new energy vehicle production and sales reached 12.888 million vehicles, 12.866 million vehicles, up 34.4%, 35.5%, presenting the supply and demand situation.Ouyang Minggao, academician of the Chinese Academy of Sciences, predicted that in 2025 the sales of new energy vehicles are expected to exceed 16 million.

Manuscript received May 6, 2025; revised August 7, 2025.

Tian-ci Ma is a postgraduate student of School of Electronic and Information Engineering, University of Science and Technology Liaoning, Anshan, Liaoning 114051, PR China (e-mail: mtcwork163@163.com).

Yong Zhang is a professor of School of Electronic and Information Engineering, University of Science and Technology Liaoning, Anshan, Liaoning 114051, PR China (*Corresponding author, phone: 86-0412-889699; e-mail: zy9091@163.com).

Under the background of the rise of the new energy industry, CATL, as the leading power battery, is the wind vane of industry development. In 2024, the global installed capacity of power battery is 894.4 GWh, and CATL leads the market share with 37.9%; in January 2025, the domestic market share reached 47.1%, up 1.6 percentage points sequentially, and it has been the champion of the global installed capacity for 8 consecutive years, with a solid market position and a continuously increasing share. The study of CATL stock price is valuable for grasping new energy industry trends, guiding investment decisions and analyzing economic operations.

This paper proposes a new energy stock closing price prediction method based on error compensation:

First, the ELG-BKA, an improved Black-winged Kite Algorithm incorporating elite driven, lens reversal, and gold sine strategies, is proposed to automatically tune VMD and LSTM parameters, addressing accuracy limitations inherent in conventional manual parameter selection methods.

Second, fusing trading and technical indicators to construct a multidimensional stock dataset.

Third, correlation analysis is used to reduce the dimensionality and VMD is utilized to decompose the non-stationary stock price series to improve the data quality.

Fourth, integrating the advantages of CNN, ATT, and LSTM to predict the stock price, and introducing the SVM error compensation mechanism to correct the model bias, which significantly improves the prediction accuracy.

The subsequent chapters of this paper are organized as follows: section 2 reviews state-of-the-art research on stock price forecasting; section 3 introduces the ELG-BKA—improved Black-winged Kite Algorithm; section 4 describes the construction of the hybrid model; section 5 illustrates the experimental data and design methodology; section 6 demonstrates and analyzes the experimental results; and section 7 concludes the full paper and proposes future research directions.

II. Related Research

In stock market research, the common forecasting techniques include time series analysis and machine learning methods. Traditional approaches primarily employ time series models such as ARIMA[3] and GARCH[4]. While these methods demonstrate some predictive capability, they rely fundamentally on linearity assumptions. However, stock price data exhibit inherent nonlinear characteristics with multifactor interactions, leading to suboptimal performance of conventional techniques in price forecasting.

Advances in artificial intelligence and computing have facilitated widespread adoption of machine learning for stock price prediction. Support Vector Machines (SVM), capable of processing nonlinear data patterns, are particularly prevalent. Pavan et al. developed an integrated model combining Random Forest and SVM to create price predictor to formulate trading strategies[5]. Kang et al. further optimized SVM parameters using a Binary Gravitational Search Algorithm, achieving superior predictive performance compared to conventional methods[6].

The evolution of deep learning has established neural networks as a dominant approach for stock price forecasting, leveraging their superior capability to model complex nonlinear relationships. Among these architectures, Recurrent Neural Networks (RNNs) and their variants demonstrate exceptional performance. Long Short-Term Memory networks (LSTMs), particularly adept at capturing temporal dependencies in sequential data, have been extensively implemented in stock prediction models[7]. Yan et al. developed an LSTM framework for financial time-series forecasting that significantly outperformed both Backpropagation Neural Networks (BPNNs) and conventional RNNs in predictive accuracy[8]. Lu et al. pioneered a hybrid CNN-LSTM architecture where convolutional layers extract discriminative features subsequently processed by LSTM for price prediction, substantially enhancing forecasting precision[9]. Chen et al. integrated LSTM and ATT, transformed the feature space through multilayer perceptron, extracted temporal features through bidirectional long and short-term memory network (BiLSTM), and utilized the attention mechanism to highlight the key information, which significantly improved the prediction accuracy of the model, confirming the important value of the ATT in stock price prediction[10].

Although the above methods have significantly contributed to stock price prediction research, the complexity, nonlinearity, and high-noise characteristics of stock data continue to pose substantial challenges for accurate forecasting. Currently, many researchers are actively exploring innovative approaches across multiple dimensions to construct more efficient and reliable prediction models.

In the field of parameter optimization for stock price prediction, the integration of intelligent optimization algorithms into machine learning models has brought new ideas to the field. Das et al. used Particle Swarm Optimization and Crow Optimization Algorithm to optimize parameters of Extreme Learning Machines[11]. Gülmez et al. employed Artificial Rabbit Algorithm to optimize hyperparameters of LSTM networks[12]. Mustaffa et al. applied Barnacle Mating Algorithm to automatically optimize network parameters of Artificial Neural Networks (ANNs)[13], which significantly enhanced the accuracy of the stock price prediction model.

At the level of data selection for stock price forecasting, multivariate data fusion has become an important trend. Agustin demonstrates through comparative experiments that combining technical indicators with fundamental indicators significantly enhances stock price forecasting accuracy[14]. Muthukumar et al. further advanced this approach by integrating textual stock data with numerical data and incorporating sentiment analysis results into model inputs, significantly improving prediction accuracy[15]. These findings indicate that accurate stock price prediction remains challenging when relying solely on basic trading data, while fusing multiple data types (including trading data, technical indicators, sentiment indicators, etc.) provides distinct advantages. Therefore, this paper refines the model input data dimensions through a multi-data fusion strategy.

In the data processing stage of stock price forecasting, data decomposition techniques have emerged as critical methodologies for enhancing prediction accuracy. Researchers have demonstrated that preprocessing raw stock price series with decomposition algorithms significantly optimizes model performance. Ni et al. applied Variational Mode Decomposition (VMD) to process data[16], Liu et al. employed Empirical Wavelet Transform (EWT) for data preprocessing[17], and Lin et al. utilized Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN) to decompose original stock index series[18]. All subsequently implemented LSTM models for prediction, effectively enhancing forecast accuracy. Building upon the demonstrated efficacy of data decomposition for stock price prediction, this paper selects VMD to decompose stock price sequences and optimizes its key parameters using the proposed ELG-BKA algorithm to further extract discriminative data features.

In the domain of model construction for stock price forecasting, hybrid models have emerged as the predominant research direction. Researchers have significantly enhanced model performance through methodological integration. Burak et al. constructed the GA-Attention-Fuzzy-Stock-Net model, integrating genetic algorithms, attention mechanisms, and neuro-fuzzy systems—where the former optimizes hyperparameters while the latter enhances feature selection capabilities, validating its efficacy[19]. Li et al. proposed a GAN-LSTM-Attention framework that synergistically combines long short-term memory networks, attention mechanisms, convolutional neural networks, and generative adversarial networks, demonstrating robust predictive performance on U.S. stock samples[20]. Zhu designed an EEMD-SESHO-RF-BiLSTM architecture fusing data processing techniques, optimization algorithms, and multiple network structures, exhibiting high accuracy and strong generalization capabilities in global index data testing[21]. Collectively, these hybrid approaches substantiate the superior predictive advantage of integrated models through algorithmic optimization and structural fusion. Building on this foundation, this paper constructs a combined model via module fusion strategy to enhance stock price forecasting accuracy.

Above comprehensive review of existing stock price forecasting research reveals that model construction needs to address core challenges including data selection, sequence decomposition, feature extraction and hyper-parameter optimization. To tackle these challenges, this study integrates cutting-edge technologies and incorporates feedback mechanisms to innovatively propose an error compensation-based hybrid forecasting model, which aims to achieve multi-stage collaborative optimization, overcome inherent limitations of conventional models, and enhance prediction accuracy while ensuring reliability.

III. Improved algorithm for Black-winged Kite

A. *Black-winged Kite Algorithm*

Proposed in 2024, the Black-winged Kite Algorithm (BKA) simulates migration and predation behaviors of black-winged kites to achieve population optimization. Demonstrating competitive performance on benchmark functions and engineering design problems, BKA shows considerable potential for solving complex optimization challenges[22]. The algorithm comprises three core phases: initialization, predation, and migration. The corresponding specific mathematical models for each phase are as follows:

$$X_i = lb + rand(ub - lb) \tag{1}$$

Eq.(1) defines the position at initialization, i is an integer between 1 and N, N is the number of black-winged kite populations, $rand$ is a random value between 0 and 1, ub and lb are the upper and lower boundaries of the search range, respectively, and X_i is the position of the ith black-winged kite at the initialization.

$$X_{i,j}^{t+1} = \begin{cases} X_{i,j}^t + n\left(1+\sin\left(r\right)\right) \times X_{i,j}^t \ p < r \\ X_{i,j}^t + n \times \left(2r-1\right) \times X_{i,j}^t \text{ else} \end{cases} \tag{2}$$

$$n = 0.05 \times e^{-2\times\left(\frac{t}{T}\right)^2} \tag{3}$$

Eq.(2) describes the position update rule of the jth dimension of the ith black-winged kite in the $t+1$th and tth iteration of the predation stage, where r is a random number between 0 and 1, p is a constant with the value of 0.9, t is the current iteration number, T is the maximum iteration number, and the value of n is determined by Eq.(3). When $p < r$, the black-winged kite hovering and swooping hunting behaviors are simulated; conversely, its hovering and attacking postures are simulated.

$$X_{i,j}^{t+1} = \begin{cases} X_{i,j}^t + C\left(0,1\right) \times \left(X_{i,j}^t - L_j^t\right) \ F_i < F_{ri} \\ X_{i,j}^t + C\left(0,1\right) \times \left(L_j^t - m \times X_{i,j}^t\right) \ else \end{cases} \tag{4}$$

$$m = 2 \times sin\left(r + \pi/2\right) \tag{5}$$

Eq.(4) defines the position of the ith black-winged kite in the $t+1$th and tth iteration of the migration phase in the jth dimension. Where L_j^t is the leading position in the jth dimension of the tth iteration, F_i is the fitness of the current individual, F_{ri} is the fitness of the randomized position, and $C\left(0,1\right)$ denotes the standard Cauchy distribution whose probability density is given by Eq.(6).

$$f\left(z\right) = \frac{1}{\pi} \times \frac{1}{\left(z^2+1\right)}, -\infty < z < +\infty \tag{6}$$

B. Improvement strategies of Black-winged Kite Algorithm

1) Elite driven strategy: The elite driven strategy improves the probability of finding an optimal swarm intelligence optimization algorithm by including the inverse solution of the current solution in the search scope and retaining elite individuals, effectively enhancing the diversity and quality of the population and circumventing the local optimum[23]. The mathematical model is as follows: on the interval $[a, b]$, the inverse point of the number x is defined as x', extended to the D-dimensional space, the inverse point p' of the point p, where $x_i' = a_i + b_i - x_i$ and $i = 1, 2, 3, \cdots, D$.

An elite driven strategy is introduced into the population initialization of the black-winged kite optimization algorithm in the following steps:

1) Generate N black-winged kite individual positions randomly in the search space to form the initial population X;

2) Generate the reverse population X' of each individual in X according to the definition of reverse point, as shown in Eq.(7), where ${X_i}'$ is the elite reverse position of the current individual X_i (generated by Eq.(1)).

$${X_i}' = lb + ub - X_i \tag{7}$$

3) Merge X and X', sort them by fitness value, and take the first N individuals to form the final initialized population X_{init}, $F_{min}(\cdot)$ represents the selection of the N individuals with the smallest fitness values.

$$X_{init} = F_{min}(X_i, {X_i}') \tag{8}$$

2) Lens reversal strategy: The lens reversal strategy is based on the principle of convex lens imaging to generate the reverse position centered on the current position, which can break through the local optimum, expand the search scope, and effectively improve the global search coverage capability and population diversity[24], and its mathematical model is as follows:

As shown in Fig.1, the lens reversal strategy takes the convex lens imaging as the principle: let y-axis be the convex lens, F is the focal length of the lens, $[m, n]$ be the search space, the projection height h of object i, the mapping point in x-axis is X_i, and it is imaged as i' by the convex lens, which corresponds to the mapping point ${X_i}'$ in x-axis and the projection height h'. Based on the imaging principle can be obtained:

$$\frac{\left(m+n\right)/2 - X_i}{{X_i}' - \left(m+n\right)/2} = \frac{h}{h'} \tag{9}$$

Let $h/h' = f$ be the scaling factor, and generalize the above principle to D-dimensional space transformations:

$${X_i}' = \frac{m_i + n_i}{2} + \frac{m_i + n_i}{2f} - \frac{X_i}{f} \tag{10}$$

By dynamically adjusting the f value, the diversity of the black-winged kite population can be enhanced to help search for better individuals. f is closely related to the generation of individual inverse solutions, and its updating formula is as shown below:

$$f = \left(1 + \left(\frac{2t}{T}\right)^{0.5}\right)^4 \tag{11}$$

The BKA incorporates a lens imaging reverse learning strategy in the predation phase, and the specific mathematical model is shown in Eq.(12), where $X_{i,j}^t$ is the lens reversal position of the current individual (generated by Eq.(2)) during the predation process, $(X_{i,j}^{t+1})'$ is the individual from its lens reversal solution.

$$(X_{i,j}^{t+1})' = \frac{lb + ub}{2} + \frac{lb + ub}{2f} - \frac{X_{i,j}^t}{f} \tag{12}$$

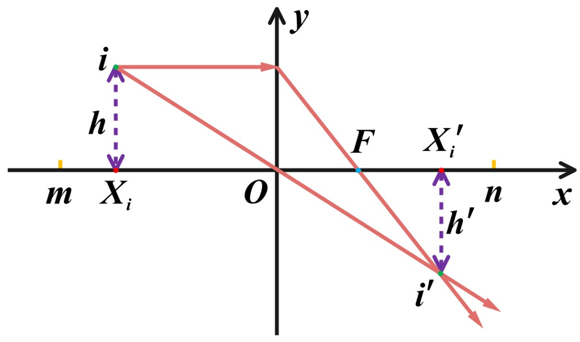

Fig. 1: Schematic of lens reversal strategy.

3) Gold sine strategy: The gold sine strategy replaces the original algorithm constant parameter with the sine function, searches in depth in the neighborhood of the local optimal solution, and improves the exploration capability of the algorithm; at the same time, relying on the association between the sine function and the unit circle, it expands the search range on the unit circle, and enhances the dynamic search performance and traversal[25], whose mathematical model is as follows:

$$\begin{cases} X_i^{t+1} = X_i^t \times |\sin R_1| - R_2 \times \sin R_1 \times \\ \qquad |\varepsilon_1 \times X_{best}^t - \varepsilon_2 \times X_i^t| \\ \varepsilon_1 = \alpha \times \tau + \beta \times (1-\tau) \\ \varepsilon_2 = \alpha \times (1-\tau) + \beta \times \tau \end{cases} \tag{13}$$

In Eq.(13), R_1, R_2 are random values, and $R_1 \in [0, 2\pi]$, $R_2 \in [0, \pi]$, X_{best}^t is the global optimal position in the tth iteration, ε_1, ε_1 are the golden section coefficients, α generally takes the value of $-\pi$, β generally takes the value of π, and τ is the number of golden sections, $\tau = \left(\sqrt{5}-1\right)/2$.

The gold sine strategy is introduced in the migration phase of the BKA, whose mathematical model is shown in Eq.(14), where $X_{i,j}^t$ is the current individual in the migration process (generated by Eq.(4)), $(X_{i,j}^{t+1})_G$ is the individual from its gold sine strategy solution.

$$\left(X_{i,j}^{t+1}\right)_G = X_{i,j}^t \times |\sin R_1| - R_2 \times \sin R_1 \times \\ |\varepsilon_1 \times X_{best}^t - \varepsilon_2 \; X_{i,j}^t| \tag{14}$$

C. Algorithm Workflow of ELG-BKA

The structure of the ELG-BKA algorithm, which incorporates the elite driven, lens reversal and gold sine strategies, is shown in Fig.2, and its process is detailed in Table I. First, the parameters of the algorithm are initialized. Based on the existing population, an initial elite population is obtained using the elite driven strategy. Then, during the predation phase and the migration phase, the lens reversal strategy and the gold sine strategy are introduced respectively to enhance the search range and accuracy in order to obtain the final optimal solution.

D. Performance Testing of ELG-BKA

The CEC2022 benchmark provides standardized test functions for evaluating single-objective constrained optimization algorithms. Comprising unimodal, multimodal, hybrid, and

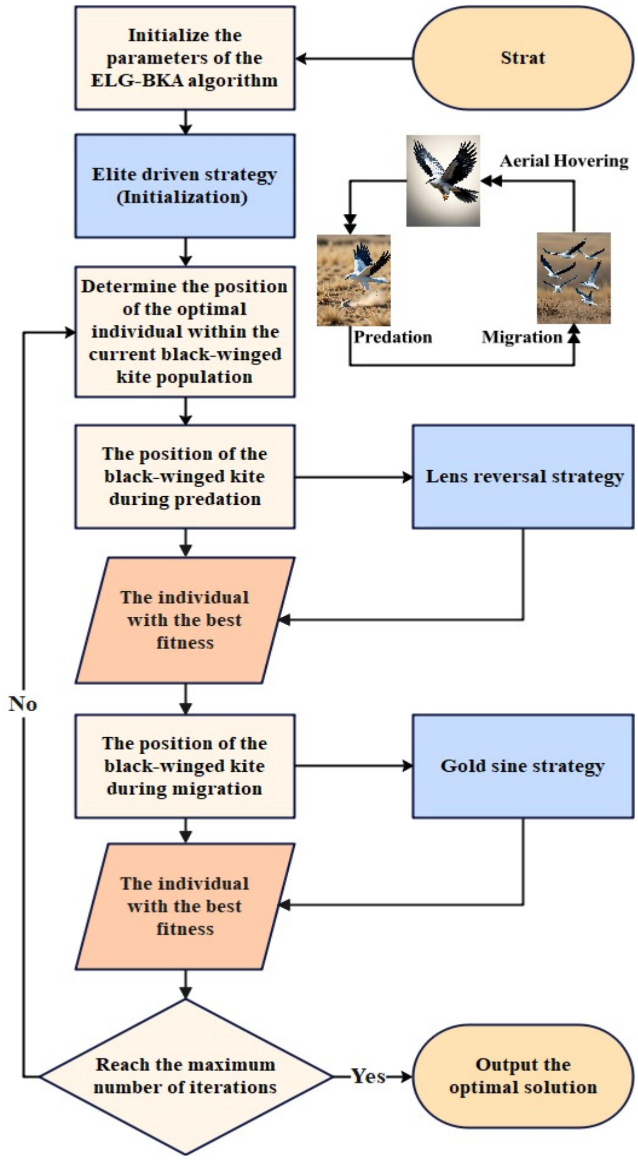

Fig. 2: Structure of ELG-BKA.

compositional functions, it establishes a unified performance assessment framework. Among these, F1 is a unimodal function, F2–F5 are multimodal functions, F6–F8 are hybrid functions, and F9–F12 are compositional functions. Foundational functions include classical optimization problems such as Zakharov, Rosenbrock, and Schaffer, which test fundamental algorithmic capabilities. Unimodal functions (e.g., Zakharov, Rosenbrock) primarily evaluate convergence behavior. Hybrid functions increase complexity by partitioning variables into subcomponents with distinct function properties. Compositional functions combine multiple base functions to simulate challenging environments featuring non-separability and asymmetry, testing algorithm adaptability. Multimodal functions (e.g., Rastrigin, Levy) assess global exploration capabilities. This diverse function set comprehensively evaluates algorithm performance under standardized conditions: search range [-100, 100] and dimension dim=10. Fig.3 visualizes representative CEC2022 function landscapes.

In complex optimization problems, convergence performance critically determines algorithmic efficacy. Figures 4-

TABLE I: ELG-BKA algorithm workflow

Input: N, T, lb, wb, Dim, $fobj$;
Output: X_{best}, F_{best}, $Curve$;

/*Initialization phase*/
1. Initialization of the population using the elite driven strategy and parameters;
/*Elite driven strategy*/
2. for $i \leftarrow 1$ to N do
3. Calculate the positions of the population $\{X_i, X_i' \mid i = 1, 2, \cdots, N\}$ according to Eqs.(1) and (7);
4. Calculate the fitness values $\{f(X_i), f(X_i') \mid i = 1, 2, \cdots, N\}$ and select the top N individuals with superior fitness values to form the elite population;
5. end
6. Calculate the initial fitness;

7. while $t < T$ do
8. Select the current optimal solution $X_{current}$ and its fitness value $F_{current}$;
/*Attacking behavior*/
9. for $i \leftarrow 1$ to N do
10. Calculate the positions of the population during the predation process according to Ep.(2);
11. Retain the better solution X_{best} and F_{best};
12. end
/*Lens reversal strategy*/
13. for $i \leftarrow 1$ to N do
14. Calculate the lens reversal positions during the predation process of the population according to Ep.(11);
15. Retain the better solution X_{best} and F_{best};
16. end
/*Migration behavior*/
17. for $i \leftarrow 1$ to N do
18. Calculate the positions of the population during the migration process according to Ep.(4);
19. Retain the better solution X_{best} and F_{best};
20. end
/*Gold sine strategy*/
21. for $i \leftarrow 1$ to N do
22. Calculate the golden sine positions of the population during the migration process according to Ep.(13);
23. Retain the better solution X_{best} and F_{best};
24. end
25. Compare the X_{best} and F_{best} with the $X_{current}$ and $F_{current}$;
26. Record the convergence curve;
27. end while
28. Return X_{best}, F_{best}, $Curve$.

6 compare ELG-BKA's convergence behavior against seven benchmark algorithms (including BKA and SO) on CEC2022 test functions F3, F6, and F9. Key observations reveal: (1) While all algorithms exhibit high initial fitness values, ELG-BKA rapidly reduces fitness through its novel mechanisms, demonstrating superior convergence velocity and precision; (2) Comparative algorithms frequently stagnate in local optima during mid-late iterations, whereas ELG-BKA maintains solution refinement via balanced global-local search, consistently approaching near-optimal solutions; (3) Convergence curves confirm ELG-BKA's outperformance in both efficiency and accuracy across tested CEC2022 problems, establishing state-of-the-art optimization capability.

Figures 7-9 present comparative boxplot analyses of multiple algorithms on CEC2022 functions F3, F6, and F9. Key findings demonstrate ELG-BKA's superior performance: F3: ELG-BKA exhibits the lowest median fitness, minimal outlier dispersion, and efficient complex function handling; F6: Significantly lower quartiles than competitors with favorable outlier distribution, effectively evading local optima; F9: Outperforms GWO and WOA in solution quality while maintaining robust global exploration. Collectively, ELG-BKA achieves exceptional stability in fitness values and optimal solution discovery across test functions. Compared to benchmark algorithms, it consistently converges to superior solutions through balanced global-local search, demonstrating both evasion of local optima and precise exploitation capability.

Table II quantifies eight algorithms' performance on CEC2022 benchmarks using mean fitness (Mean), standard deviation (Std), best value (Best), and ranking metrics (Rank). Results demonstrate ELG-BKA's superiority: it achieves lower mean values and reduced Std across most functions (excluding F2, F9, F12), indicating exceptional solution quality and stability. ELG-BKA significantly outperforms comparative algorithms, securing the highest overall ranking. Compared to BKA, it excels in critical optimization metrics, validating the enhancement strategy's efficacy for complex optimization problems.

IV. LSTM HYBRID MODEL CONSTRUCTION BASED ON ERROR COMPENSATION

Variational Mode Decomposition (VMD) is a modal decomposition method with adaptive and non-recursive properties, which can decompose complex signals into multiple single-component signals, and is suitable for dealing with nonlinear and nonsmooth signals[26]. The method obtains the IMF components by solving the constrained variational problem, but the number of modal decomposition K and the quadratic penalty factor α need to be set manually, and their values have a direct impact on the decomposition accuracy-too large a value of K will easily generate redundant information, while too small will lead to omission of modes; too large a value of α will lead to loss of information in the frequency band, while too small will lead to redundancy of information. For this reason, the ELG-BKA algorithm is used in this paper to optimize these two parameters.

TABLE II: CEC2022 function test index results

Data	Index	ELG-BKA	BKA	SO	PSO	GWO	WOA	AVOA	DBO
F1	Mean	**3.00000E+02**	7.85189E+03	3.01255E+02	3.00001E+02	4.29176E+03	1.87301E+ 04	3.01808E+02	3.96026E+02
	Std	1.61945E-11	1.90742E+03	2.37189E+00	6.42722E-04	1.95251E+03	9.00684E+03	5.76114E+00	3.10550E+02
	Best	3.00000E+02	5.28210E+03	3.00001E+02	3.00000E+02	3.24881E+02	2.86664E+03	3.00000E+02	3.00000E+02
	Rank	**1**	7	3	2	6	8	4	5
F2	Mean	4.07850E+02	4.14007E+02	4.08891E+02	**4.06903E+02**	4.24339E+02	4.42395E+02	4.17089E+02	4.31726E+02
	Std	1.82322E+01	2.64018E+01	1.66951E+01	1.24511+-01	2.16795E+01	5.71972E+01	2.53548E+01	3.16844E+01
	Best	4.00023E+02	4.00001E+02	4.00002E+02	4.00009E+02	4.06727E+02	4.00699E+02	4.00000E+02	4.05289E+02
	Rank	2	4	3	**1**	6	8	5	7
F3	Mean	**6.00883E+02**	6.19967E+02	6.04611E+02	6.09846E+02	6.01179E+02	6.37655E+02	6.17691E+02	6.11161E+02
	Std	1.26402E+00	9.82784E+00	4.96155E+00	8.50195E+00	1.23745E+00	1.25365E+01	1.03598E+01	7.74521E+00
	Best	6.00010E+02	6.04572E+02	6.00052E+02	6.00619E+02	6.00055E+02	6.15089E+02	6.00541E+02	6.00902E+02
	Rank	**1**	7	3	4	2	8	6	5
F4	Mean	**8.14659E+02**	8.19468E+02	8.22353E+02	8.22320E+02	8.18267E+02	8.36118E+02	8.30210E+02	8.33059E+02
	Std	5.35433E+00	5.04417E+00	5.57796E+00	6.59297E+00	7.74000E+00	1.72823E+01	9.40828E+00	1.23284E+01
	Best	8.04975E+02	8.10945E+02	8.08955E+02	8.10945E+02	8.06095E+02	8.12052E+02	8.10945E+02	8.13929E+02
	Rank	**1**	3	5	4	2	8	6	7
F5	Mean	**9.00000E+02**	9.90365E+02	9.22402E+02	9.76567E+02	9.15004E+02	1.45877E+03	1.27160E+03	9.48855E+02
	Std	0.00000E+00	9.26659E+01	2.82247E+01	1.58551E+02	2.80383E+01	3.14844E+02	1.75556E+02	4.60901E+01
	Best	9.00000E+02	9.05908E+02	9.00000E+02	9.00000E+02	9.00100E+02	9.63942E+02	9.75818E+02	9.05518E+02
	Rank	**1**	6	3	5	2	8	7	4
F6	Mean	**1.87823E+03**	3.05733E+03	3.83924E+03	3.56283E+03	5.11840E+03	3.14209E+03	3.58913E+03	5.81409E+03
	Std	5.75221E+01	1.02132E+03	1.53155E+03	2.27539E+03	2.43754E+03	1.07620E+03	1.72806E+03	2.24633E+03
	Best	1.80538E+03	1.88034E+03	1.87184E+03	1.87176E+03	2.14801E+03	1.92406E+03	1.87095E+03	2.17033E+03
	Rank	**1**	2	6	4	7	3	5	8
F7	Mean	**2.02933E+03**	2.10651E+03	2.03187E+03	2.03137E+03	2.03180E+03	2.06871E+03	2.04261E+03	2.03876E+03
	Std	1.18274E+01	4.49292E+01	2.72115E+01	1.06845E+01	1.49580E+01	3.01740E+01	2.49638E+01	1.93981E+01
	Best	2.00796E+03	2.04747E+03	2.00242E+03	2.02000E+03	2.02010E+03	2.03140E+03	2.00562E+03	2.00542E+03
	Rank	**1**	8	4	2	3	7	6	5
F8	Mean	**2.21994E+03**	2.39075E+03	2.22410E+03	2.22073E+03	2.22520E+03	2.23446E+03	2.22633E+03	2.22794E+03
	Std	5.94857E+00	1.21717E+02	2.27665E+01	4.27146E-01	4.71913E+00	9.93222E+00	5.68010E+00	7.85960E+00
	Best	2.20037E+03	2.22319E+03	2.20071E+03	2.22004E+03	2.20400E+03	2.22334E+03	2.21764E+03	2.22114E+03
	Rank	**1**	8	3	2	4	7	5	6
F9	Mean	2.53011E+03	2.59322E+03	**2.50457E+03**	2.53429E+03	2.56452E+03	2.57609E+03	2.53908E+03	2.53961E+03
	Std	4.20671E+00	1.45229E+01	5.81925E+01	2.68059E+01	3.00239E+01	4.66762E+01	3.72778E+01	2.84367E+01
	Best	2.52928E+03	2.55894E+03	2.48550E+03	2.52928E+03	2.52929E+03	2.53022E+03	2.52928E+03	2.52928E+03
	Rank	2	8	**1**	3	6	7	4	5
F10	Mean	**2.52126E+03**	2.74666E+03	2.55562E+03	2.57962E+03	2.56915E+03	2.61923E+03	2.57449E+03	2.53538E+03
	Std	4.71265E+01	3.97895E+02	1.04236E+02	9.62933E+01	6.47013E+01	1.36895E+02	6.20865E+01	5.84529E+01
	Best	2.50037E+03	2.53789E+03	2.42853E+03	2.50020E+03	2.50029E+03	2.50052E+03	2.50047E+03	2.50054E+03
	Rank	**1**	8	3	6	4	7	5	2
F11	Mean	**2.62841E+03**	2.75000E+03	2.7297E+03	2.75339E+03	2.78676E+03	2.84535E+03	2.73649E+03	2.78602E+03
	Std	8.38598E+01	1.52564E+02	1.16408E+02	1.60243E+02	1.36656E+02	1.53207E+02	1.38391E+02	1.55712E+02
	Best	2.60000E+03	2.60000E+03	2.60000E+03	2.60000E+03	2.60193E+03	2.63904E+03	2.60000E+03	2.60000E+03
	Rank	**1**	4	2	5	7	8	3	6
F12	Mean	2.86566E+03	3.00385E+03	2.86984E+03	**2.86471E+03**	2.86663E+03	2.88742E+03	2.86964E+03	2.87481E+03
	Std	5.30522E+00	5.91039E+01	3.40325E+00	2.33349E+01	6.47263E+00	2.62556E+01	8.15949E+00	1.88649E+01
	Best	2.86144E+03	2.92092E+03	2.86319E+03	2.84613E+03	2.85955E+03	2.86304E+03	2.86257E+03	2.86327E+03
	Rank	2	8	5	**1**	3	7	4	6
Toal rank		**15**	73	41	39	52	86	60	66
Final rank		**1**	7	3	2	4	8	5	6

* The rank is based on Mean. When Mean values are the same, the algorithm with a smaller Std wins. The winner is shown in bold.

Convolutional Neural Network (CNN), as the core algorithm of deep learning, is able to mine features from stock market time-series data, and is a feed-forward network containing convolutional operations, which consists of a convolutional layer, a pooling layer, and a fully connected layer. The convolutional layer extracts stock price spatial distribution features through multiple convolutional kernels, but the output feature dimensions are high; the pooling layer reduces the dimensionality of the high-dimensional data to improve the network operation efficiency and extracts the secondary features; and the fully-connected layer integrates the local features into the global features to ultimately realize the stock price prediction[27].

Attention mechanism (ATT) simulates human brain perception and improves information processing efficiency and prediction accuracy by assigning feature weights to input data. Its core is a weighted summation based on the importance of the inputs to the prediction task, which is achieved in three steps: calculation of weights, normalization, and weighting. In this paper, ATT is incorporated into the prediction model, which can accurately extract the key information of the stock price sequence, filter the secondary content, and effectively improve the efficiency of feature extraction[28].

Long Short-Term Memory (LSTM) network is an improved version of RNN with long-term memory capability, which solves the problem of gradient vanishing and explosion by effectively controlling the flow of information, better captures time series dependencies, and utilizes memory units to store and update the historical state, which is suitable for time-series data prediction such as stock prices[29]. The

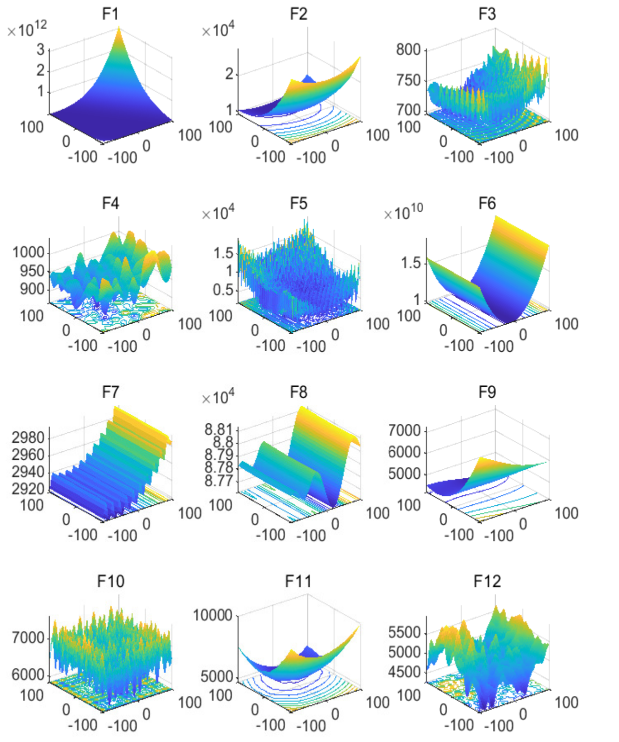

Fig. 3: 3D visualization of CEC2022 benchmark functions with dimension dim=10.

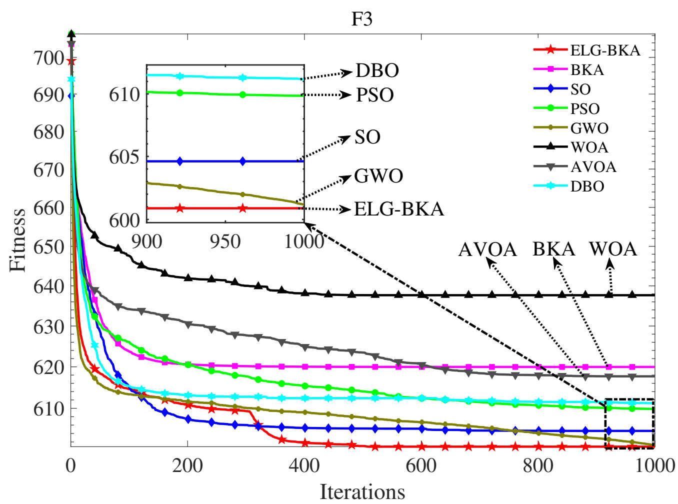

Fig. 4: Convergence curves of CEC2022 function F3.

LSTM unit consists of a forget gate, an input gate, an output gate and a memory cell: the forget gate filters redundant information, the input gate screens the memorized content, and the output gate regulates the output. The units share the parameters and are optimized in the loop learning. Given that LSTM hyperparameters affect the performance, this paper uses ELG-BKA to optimize the number of hidden layers, the initial learning rate and the L2 regularization parameters.

Support Vector Machine (SVM) is a classical supervised learning algorithm for classification and regression tasks, with significant advantages in classifying small and medium-

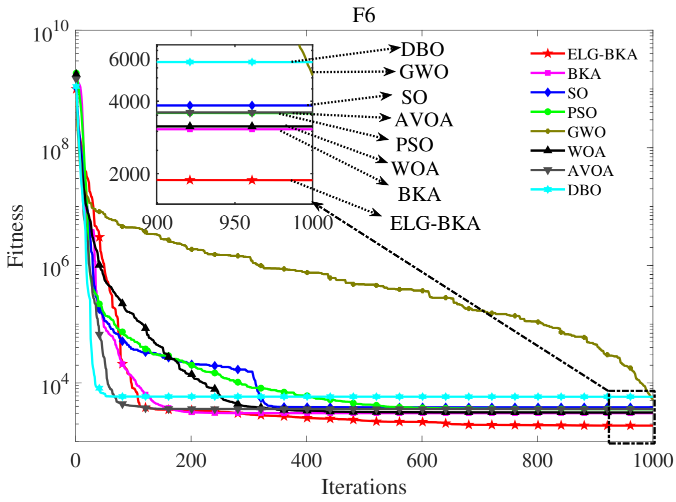

Fig. 5: Convergence curves of CEC2022 function F6.

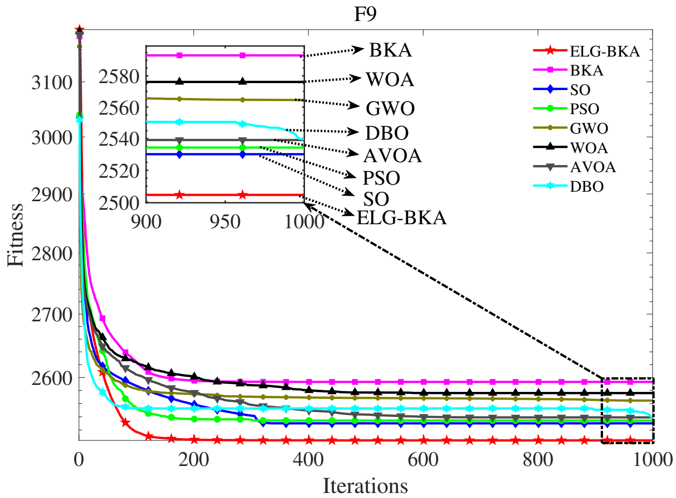

Fig. 6: Convergence curves of CEC2022 function F9.

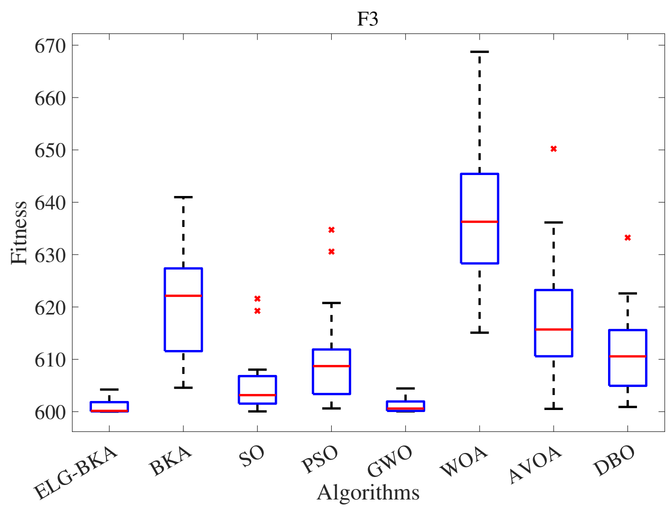

Fig. 7: Boxplot of different algorithms on CEC2022 function F3.

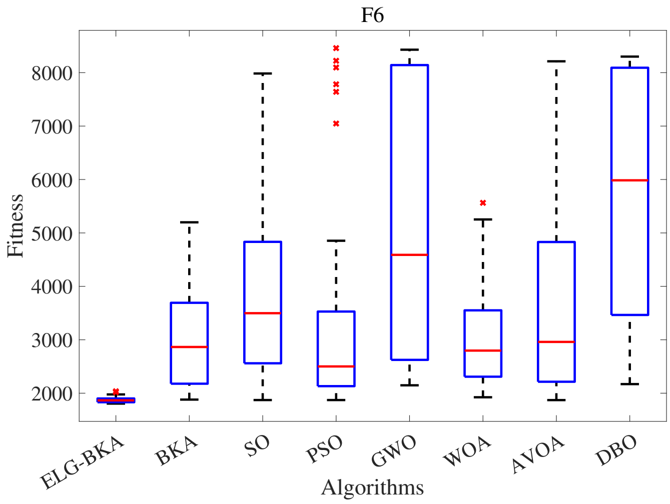

Fig. 8: Boxplot of different algorithms on CEC2022 function F6.

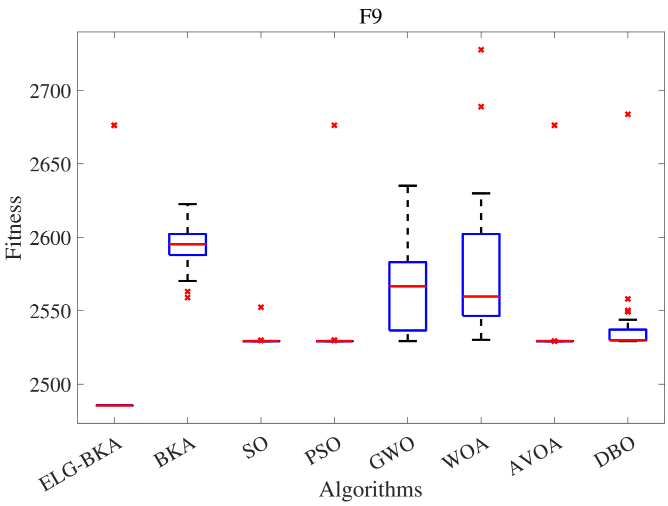

Fig. 9: Boxplot of different algorithms on CEC2022 function F9.

sized datasets. Its core is to maximize the category data interval by finding the optimal hyperplane to enhance the model generalization ability[30].

Consequently, this study constructs a hybrid LSTM model based on error compensation by integrating the advantages of ELG-BKA optimization, VMD, CNN-ATT feature extraction, LSTM prediction, SVM error correction. And the structure of this model is illustrated in Fig.10.

V. Experimental design

A. Data sources and description

The experimental data are taken from the Tushare database (https://tushare.pro/), covering the opening price and volume of seven stocks from July 9, 2018, to February 21, 2025, and other information. The seven new energy stocks are: BYD Company Limited (BYD), Gotion High-tech Co.,Ltd (Gotion), Shenzhen Senior Technology Material Co., Ltd (Senior), Shanghai Kelai Mechatronics Engineering Co.,Ltd (Kelai), Contemporary Amperex Technology Co., Ltd (CATL), Zhejiang Sanhua Intelligent Controls Co.,Ltd (Sanhua), Sunwoda Electronic Co., Ltd (Sunwoda). Influenced by the new crown epidemic, Russia-Ukraine conflict and the goals on carbon peak and carbon neutrality and other events, the stock price of the new energy industry fluctuates drastically, showing non-linearity and instability, which makes forecasting more difficult. Taking CATL stock as an example, Table III shows some of the data, and Fig.11 shows its closing price trend. It is evident that the closing prices fluctuate significantly during the sample period, presenting a certain level of forecasting difficulty.

B. Feature construction

The data characteristics of each stock cover basic trading data, fundamental data, technical indicators, and the basic trading data of stocks in the same industry is also an important supplement. In the case of CATL, Table IV presents its stock data characteristics. Among them, the basic stock trading indicators, directly obtained technical indicators and industry trading indicators can be obtained from the Tushare database, and some stock technical indicators are constructed according to their own definitions.

C. Data processing

1) VMD decomposition: ELG-BKA is used to optimize the number of decompositions K and the penalty factor α of VMD for decomposing the historical closing price series with minimum sample entropy as the objective function. The convergence curve is shown in Fig.12, and finally $K = 7$ and $\alpha = 4600$ are determined, and the results of the decomposition of the closing price of CATL stock are shown in Fig.13 and Fig.14.

2) Feature screening: Pearson's correlation coefficient method is used to measure the degree of linear correlation between two variables, based on the covariance and standard deviation calculation, taking the value of the range of $[-1, 1]$, the closer the absolute value of 1 correlation is the stronger, the sign characterizes the direction of correlation. This method is often applied to feature screening, retaining features that are significantly correlated ($|r| > Threshold$) with the target variable for dimensionality reduction. In this paper, this method is applied to eliminate irrelevant features to determine the model inputs, and the correlation coefficients are shown in Table V (the high correlation values are marked in bold), and the features with $|r| > 0.7$ are selected as the model inputs[31]. The 22 features selected are: Open, High, Low, Close, Pre_close, Total_mv, BOLL_LB, BOLL_MD, BOLL_UB, OBV, MA_5, MA_10, MA_20, Close_002050.SZ, Close_002460.SZ, Close_002594.SZ, Close_002886.SZ, Close_300207.SZ, Close_300274.SZ, Close_300568.SZ, Close_300648.SZ, Close_300712.SZ.

3) Data normalization: In order to eliminate the magnitude and scale differences between features and to improve the efficiency of model processing, the indicator dataset was normalized to $[0, 1]$ with the following calculation formula:

$$Y = \frac{(Y_{\max} - Y_{\min}) \times (X - X_{\min})}{X_{\max} - X_{\min}} + Y_{\min} \tag{15}$$

Where, Y is the normalized value; X is the original value; $X_{\min}$, $X_{\max}$ is the minimum and maximum value of the original data; $Y_{\min} = 0$, $Y_{\max} = 1$ is the minimum and maximum value of the target interval after normalization.

TABLE III: Selected data on CATL stock price

No.	Date	Open	High	Low	Close	Pre_close	Change	Pct_chg	Volume	Amount	Turnover_rate
1	20180709	37.76	38.17067	36.85867	38.16	70.88	0.67	0.95	310689.79	2183986.5	14.3014
2	20180710	40	40.47467	38.21867	38.61333	71.55	0.85	1.19	364844.01	2666043.24	16.7942
3	20180711	37.33333	42.10133	37.10933	41.09867	72.4	4.66	6.44	491208.05	3695050.77	22.6109
4	20180712	40.54933	45.21067	40.32	45.21067	77.06	7.71	10.01	430852.12	3430182.91	19.8327
5	20180713	43.73333	47.184	43.216	44.10667	84.77	-2.07	-2.44	482020.55	4049051.4	22.188

TABLE IV: CATL stock price data characteristics

Feature type	No.	Indicator	Indicator description
Basic trading indicators of individual stocks	1	Open	Opening price
	2	High	Highest price
	3	Low	Lowest price
	4	Close	Closing price
	5	Pre_close	Yesterday's closing price
	6	Change	Amount of change
	7	Pct_chg	Percentage change
	8	Volume	Amount of trading
	9	Amount	Trading value
Indicators of individual stocks that can be directly obtained	1	Turnover_rate	Turnover rate
	2	Volume_ratio	Volume ratio
	3	Pe	Price-to-earnings ratio
	4	Pe_ttm	Trailing-twelve-month price-to-earnings ratio
	5	Pb	Price-to-book ratio
	6	Ps	Price-to-sales ratio
	7	Ps_ttm	Trailing-twelve-month price-to-sales ratio
	8	Total share	Total share capital
	9	Total mv	Total market value
Constructed technical indicators of individual stocks	1	MACD	Moving Average Convergence Divergence
	2	KDJ_K	Stochastic oscillator K-line
	3	KDJ_D	Stochastic oscillator D-line
	4	KDJ_J	Stochastic oscillator J-line
	5	RSI_6	6-day Relative Strength Index
	6	RSI_12	12-day Relative Strength Index
	7	RSI_24	24-day Relative Strength Index
	8	W&R	Williams %R
	9	BOLL_LB	The lower band of the Bollinger Bands
	10	BOLL_MD	The middle band of the Bollinger Bands
	11	BOLL_UB	The upper band of the Bollinger Bands
	12	CCI	Commodity Channel Index
	13	PSY	Psychological Line
	14	OBV	On-Balance Volume
	15	MA_5	5-day moving average
	16	MA_10	10-day moving average
	17	MA_20	20-day moving average
	18	VMA_5	5-day average trading volume
	19	VMA_10	10-day average trading volume
	20	VMA_20	20-day average trading volume
Basic trading indicators of individual stocks within the same industry	1	Close_002050.SZ	Closing price of Zhejiang Sanhua Intelligent Controls Co.,Ltd
	2	Close_002460.SZ	Closing price of Ganfeng Lithium Group Co., Ltd
	3	Close_002594.SZ	Closing price of BYD Company Limited
	4	Close_002886.SZ	Closing price of Shenzhen WOTE Advanced Materials Co., Ltd
	5	Close_300207.SZ	Closing price of Sunwoda Electronic Co., Ltd
	6	Close_300274.SZ	Closing price of Sungrow Power Supply Co., Ltd
	7	Close_300568.SZ	Closing price of Shenzhen Senior Technology Material Co., Ltd
	8	Close_300648.SZ	Closing price of Fujian Nebula Electronics., Ltd
	9	Close_300712.SZ	Closing price of Fujian Yongfu Power Engineering Co.,Ltd
	10	Close_300713.SZ	Closing price of Shenzhen Increase Technology Co., Ltd

TABLE V: The correlation coefficients of the characteristics of CATL stock price data

No.	Feature	r	No.	Feature	r	No.	Feature	r
01	Open	**1**	17	Total_share	0.31	33	MA_5	**1**
02	High	**1**	18	Total_mv	**1**	34	MA_10	**1**
03	Low	**1**	19	MACD	-0.0048	35	MA_20	**0.99**
04	Close	**1**	20	KDJ_K	0.0033	36	VMA_5	0.14
05	Pre_close	**0.89**	21	KDJ_D	0.0006	37	VMA_10	0.14
06	Change	0.023	22	KDJ_J	0.0057	38	VMA_20	0.13
07	Pct_chg	0.0052	23	RSI 6	-0.0078	39	Close_002050.SZ	**0.78**
08	Volume	0.12	24	RSI_12	-0.024	40	Close_002460.SZ	**0.83**
09	Amount	0.69	25	RSI_24	-0.081	41	Close_002594.SZ	**0.95**
10	Turnover_rate	-0.48	26	WR	-0.0085	42	Close_002886.SZ	**0.71**
11	Volume_ratio	-0.0045	27	BOLL_LB	**0.98**	43	Close_300207.SZ	**0.76**
12	Pe	0.46	28	BOLL_MD	**0.99**	44	Close_300274.SZ	**0.95**
13	P_ttm	0.31	29	BOLL_UB	**0.99**	45	Close_300568.SZ	**0.83**
14	Pb	0.57	30	CCI	-0.005	46	Close_300648.SZ	**0.75**
15	Ps	0.4	31	PSY	-0.072	47	Close_300712.SZ	**0.84**
16	Ps_ttm	0.3	32	OBV	**0.91**	48	Close_300713.SZ	0.39

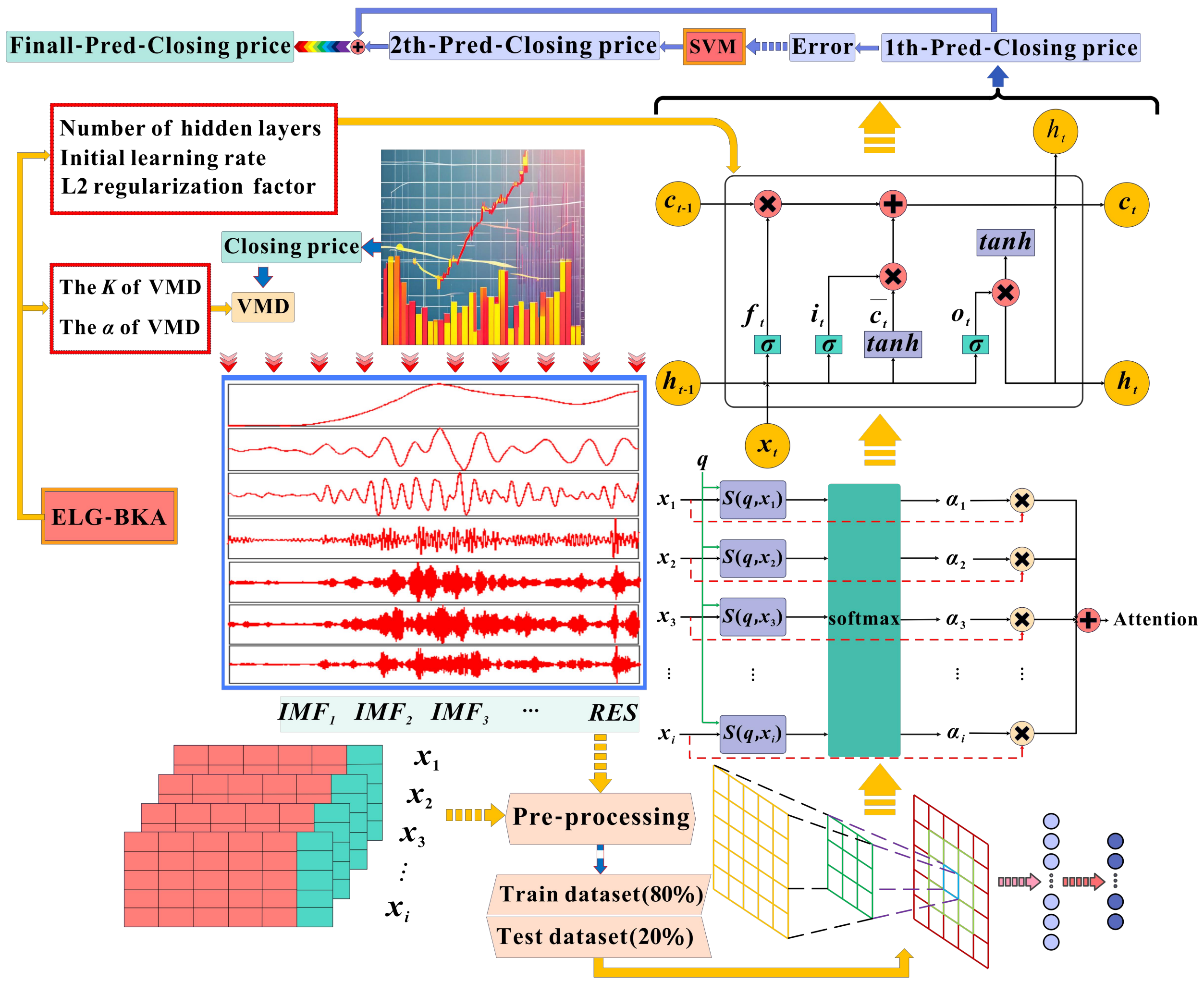

Fig. 10: Structure of LSTM hybrid model based on error compensation.

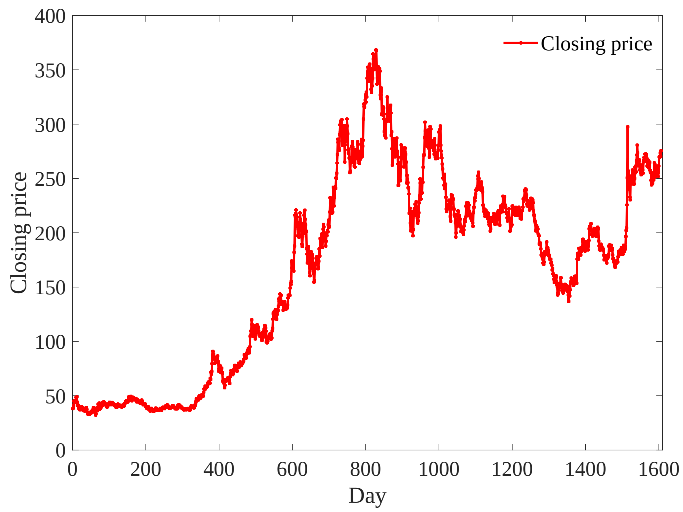

Fig. 11: CATL closing price.

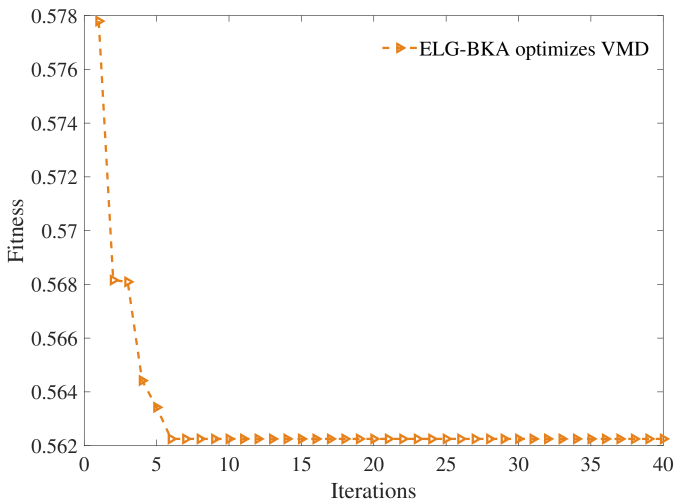

Fig. 12: Convergence curve of VMD for ELG-BKA.

4) Evaluation metrics: In this paper, four metrics, mean absolute error (MAE), mean absolute percentage error (MAPE), root mean square error (RMSE), and fit coefficient (R^2), are used to assess the model prediction effect, and the calculation is shown in Table VI. In the equation: y_i is the actual value, y_i' is the predicted value, $\bar{y}$ is the average value, and n is the number of data.

5) Experimental methods: In order to validate the prediction accuracy and stability of the error complement-based LSTM hybrid model proposed in this paper, experiments are conducted on 7 stock datasets (BYD, Gotion, Kelai, CATL, Sanhua, Sunwoda, and Senior). First, a fully orthogonal experiment of the hybrid model is designed based on the stock price of CATL to verify the validity of the sub-

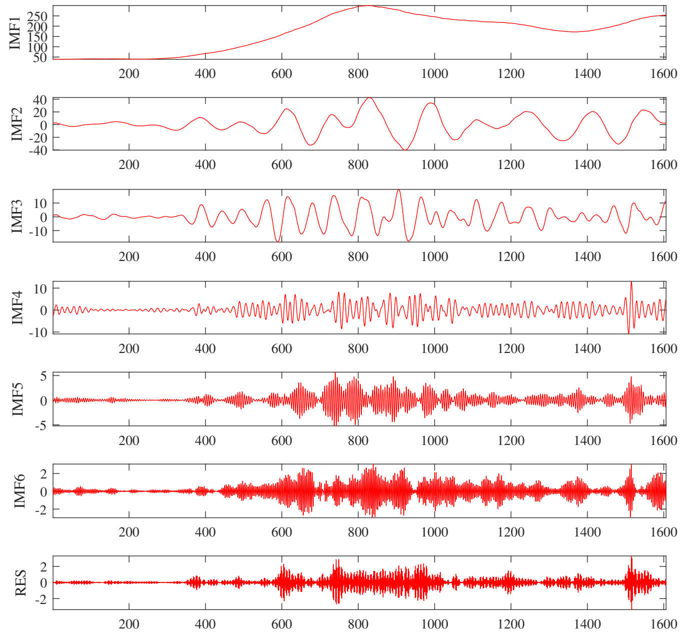

Fig. 13: Breakdown curve of the closing price of CATL stock.

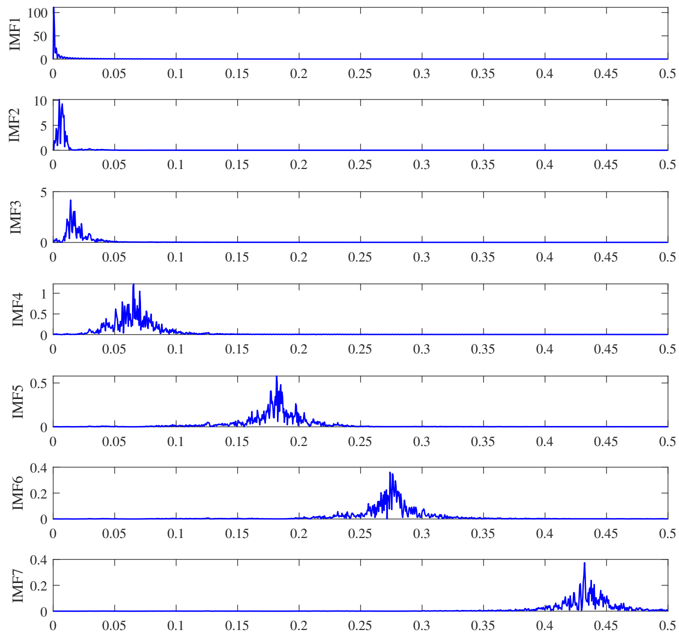

Fig. 14: Spectrogram of the decomposition of the closing price of CATL stock.

TABLE VI: Evaluation metrics

Metrics	Equation
MAE	$MAE = \frac{1}{n}\sum_{i=1}^{n}\lvert y_i - y_i'\rvert$
MAPE	$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left\lvert \frac{y_i - y_i'}{y_i}\right\rvert$
RMSE	$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - y_i')^2}$
R^2	$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - y_i')^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$

module for the hybrid model. Then, a model comparison experiment is designed based on seven representative new energy stocks, and BiLSTM, BiGRU, Transformer-LSTM, Transformer-GRU, Transformer-BiLSTM, and Transformer-BiGRU models with high competitiveness are selected to comprehensively compare the four prediction evaluation metrics to verify the superiority of the hybrid model proposed in this paper.

VI. Experimental results and analysis

A. Results of orthogonal experiment

The results of the orthogonal experimental model evaluation metrics are shown in Table VII, and the prediction results are shown in Fig.15. From the data in the table, it can be seen that the LSTM hybrid model based on error complement proposed in this paper outperforms the other comparative models in all the metrics, reflecting a good prediction ability. The comparison reveals that the model performance decreases significantly when key components in the model are gradually removed. For example, after removing the VMD decomposition step, the model shows a significant increase in MAE, MAPE and RMSE metrics and a decrease in R^2 value, which indicates that the preprocessing of the original stock price series by the VMD decomposition is able to effectively extract different features and improve the model prediction accuracy. After adding the SVM error compensation mechanism, the R^2 value of the model is further improved, and the MAE, MAPE and RMSE metrics are all decreased, which indicates that the error compensation mechanism can significantly improve the accuracy of the model. Similarly, the absence of components such as ELG-BKA optimization parameters and CNN-ATT feature extraction prominence can lead to deterioration of the performance of the LSTM hybrid model, which fully proves the importance of the various parts of the model constructed in this paper working in concert, and the absence of one is indispensable. Specifically, the LSTM hybrid model based on error compensation improves 18.72%-89.44% on MAE, 19.55%-88.67% on MAPE, 17.47%-89.28% on RMSE, and 31.25%-98.83% on R^2.

B. Results of the individual stock multi-model comparison experiment

The evaluation metrics for the stock prices of the 7 new energy representatives across the 7 comparative models are presented in Table VIII. As shown in the table, the error compensation-based hybrid model proposed in this paper demonstrates superior performance to the other comparative models in predicting different stocks. This result validates its outstanding performance in stock price forecasting and underscores its strong generalization ability. Specifically, the metric improvements are detailed in Table IX. For individual stock prediction, the error compensation-based hybrid model achieves lower MAE, MAPE, and RMSE values compared to the other models, while attaining a higher R^2 metric. This indicates that the model more effectively captures the relationship between features and the target variable, exhibiting a high degree of goodness-of-fit.

As shown in Fig.16 and Fig.17, taking the prediction results of BYD and Senior stock as examples, all models

TABLE VII: Orthogonal experimental model evaluation metrics results

No.	Factors					Models	Metrics			
	VMD	ELG-BKA	CNN-ATT	SVM	LSTM		MAE	MAPE	RMSE	R^2
1	0	0	0	0	1	LSTM	9.9326	4.7501	12.782	0.9024
2	0	0	0	1	1	LSTM-SVM	3.4551	1.6723	5.5946	0.9811
3	0	0	1	0	1	CNN-ATT-LSTM	13.037	6.1575	16.106	0.8450
4	0	0	1	1	1	CNN-ATT-LSTM-SVM	3.5515	1.7169	5.7437	0.9801
5	0	1	0	0	1	ELG-BKA-LSTM	2.8449	1.4923	3.5705	0.9924
6	0	1	0	1	1	ELG-BKA-LSTM-SVM	3.2621	1.4599	4.6290	0.9871
7	0	1	1	0	1	ELG-BKA-CNN-ATT-LSTM	1.9297	1.0004	2.3023	0.9968
8	0	1	1	1	1	ELG-BKA-CNN-ATT-LSTM-SVM	2.7382	1.3168	4.1893	0.9894
9	1	0	0	0	1	VMD-LSTM	14.853	7.1002	17.721	0.8123
10	1	0	0	1	1	VMD-LSTM-SVM	3.5922	1.7367	5.7895	0.9798
11	1	0	1	0	1	VMD-CNN-ATT-LSTM	7.4626	3.9454	9.0123	0.9515
12	1	0	1	1	1	VMD-CNN-ATT-LSTM-SVM	2.7190	1.3522	4.0808	0.9900
13	1	1	0	0	1	VMD-ELG-BKA-LSTM	4.1973	1.9338	6.2074	0.9770
14	1	1	0	1	1	VMD-ELG-BKA-LSTM-SVM	2.0894	0.9822	3.5805	0.9923
15	1	1	1	0	1	VMD-ELG-BKA-CNN-ATT-LSTM	5.6389	2.7006	7.8565	0.9631
16	1	1	1	1	1	VMD-ELG-BKA-CNN-ATT-LSTM-SVM	1.5684	0.8048	1.9002	0.9978

Note: 0 represents the absence, and 1 represents the presence.

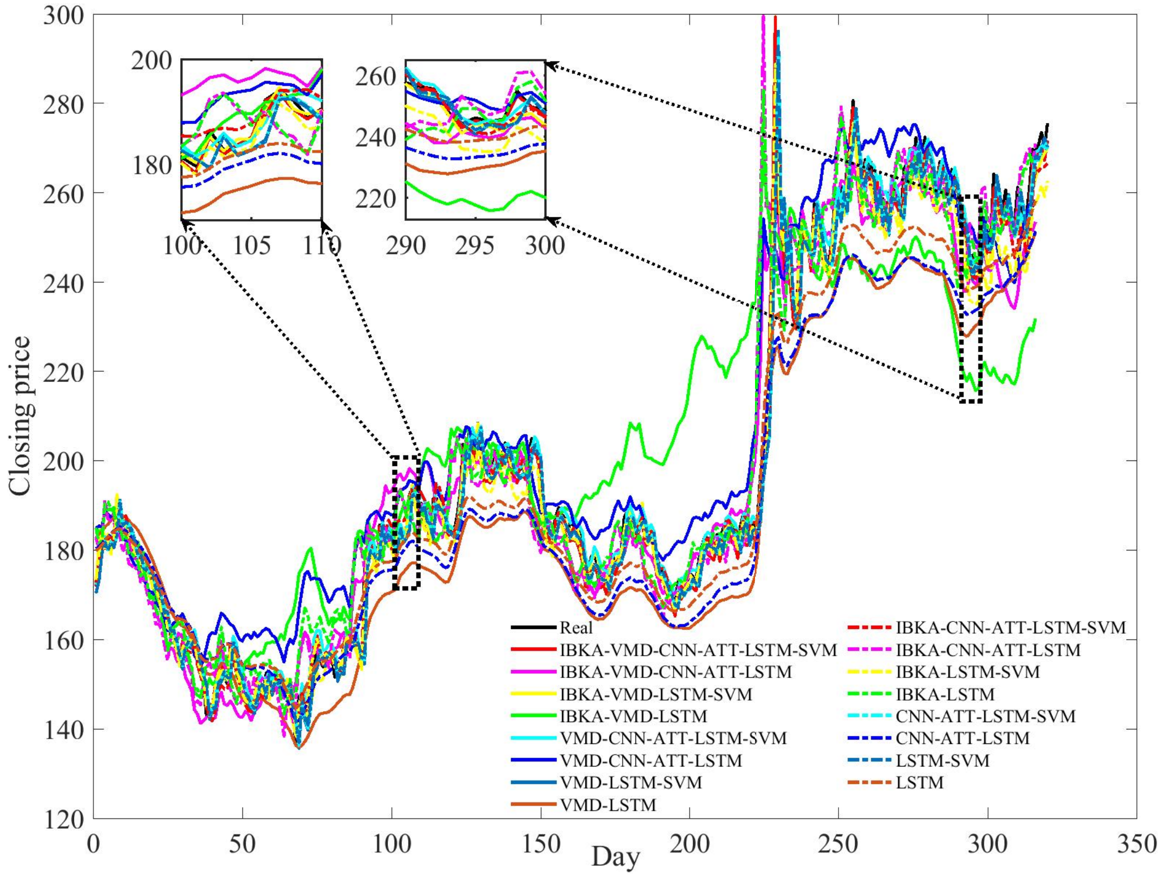

Fig. 15: Orthogonal experimental model prediction results.

can track the stock price trend, but with a lag. In contrast, the hybrid model based on error compensation is closer to the real stock price in the multi-stage prediction and shows higher prediction accuracy. Specifically, BYD's stock price exhibited an initial slight decline followed by a significant increase, while Senior' stock price showed an initial sharp drop followed by fluctuating movements. Yet, the proposed model accurately predicted both trends, demonstrating its robust capability in stock price forecasting. In summary, the error compensation-based hybrid model delivers superior performance in predicting new energy stock prices, demonstrating strong potential and value for stock prediction applications.

VII. Summary

A. *Summary of research results*

This study constructs an error compensation-based hybrid LSTM model for new energy stock price prediction. The non-stationary price series are processed using VMD, combined with Pearson correlation coefficients for dimensionality reduction. CNN-ATT extracts salient features to optimize LSTM inputs. The SVM error compensation module enhances prediction accuracy, while the proposed ELG-BKA algorithm optimizes VMD and LSTM parameters. Experimental results verify that the proposed model achieves superior prediction accuracy and stability compared to benchmark models, demonstrating adaptability to market dynamics while

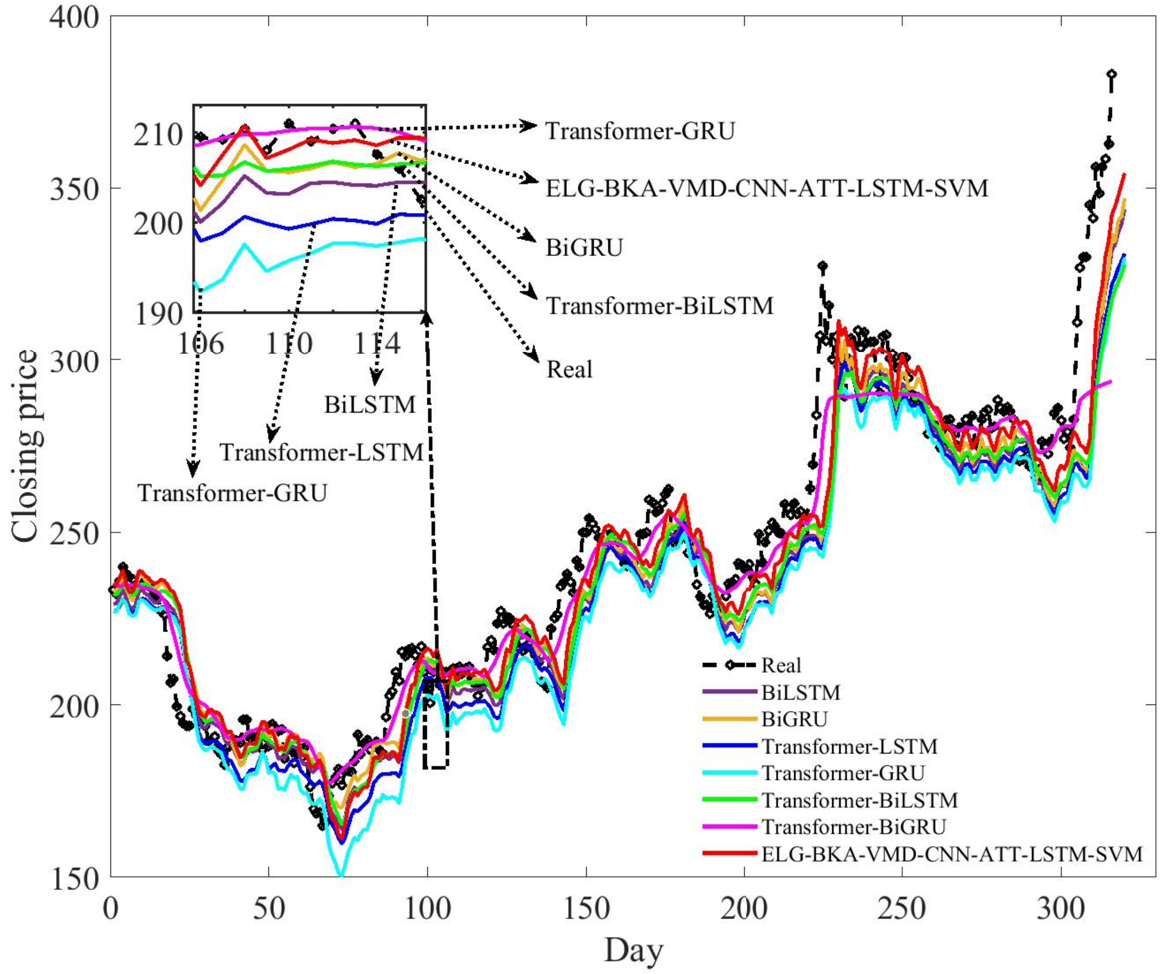

Fig. 16: BYD stock price prediction results.

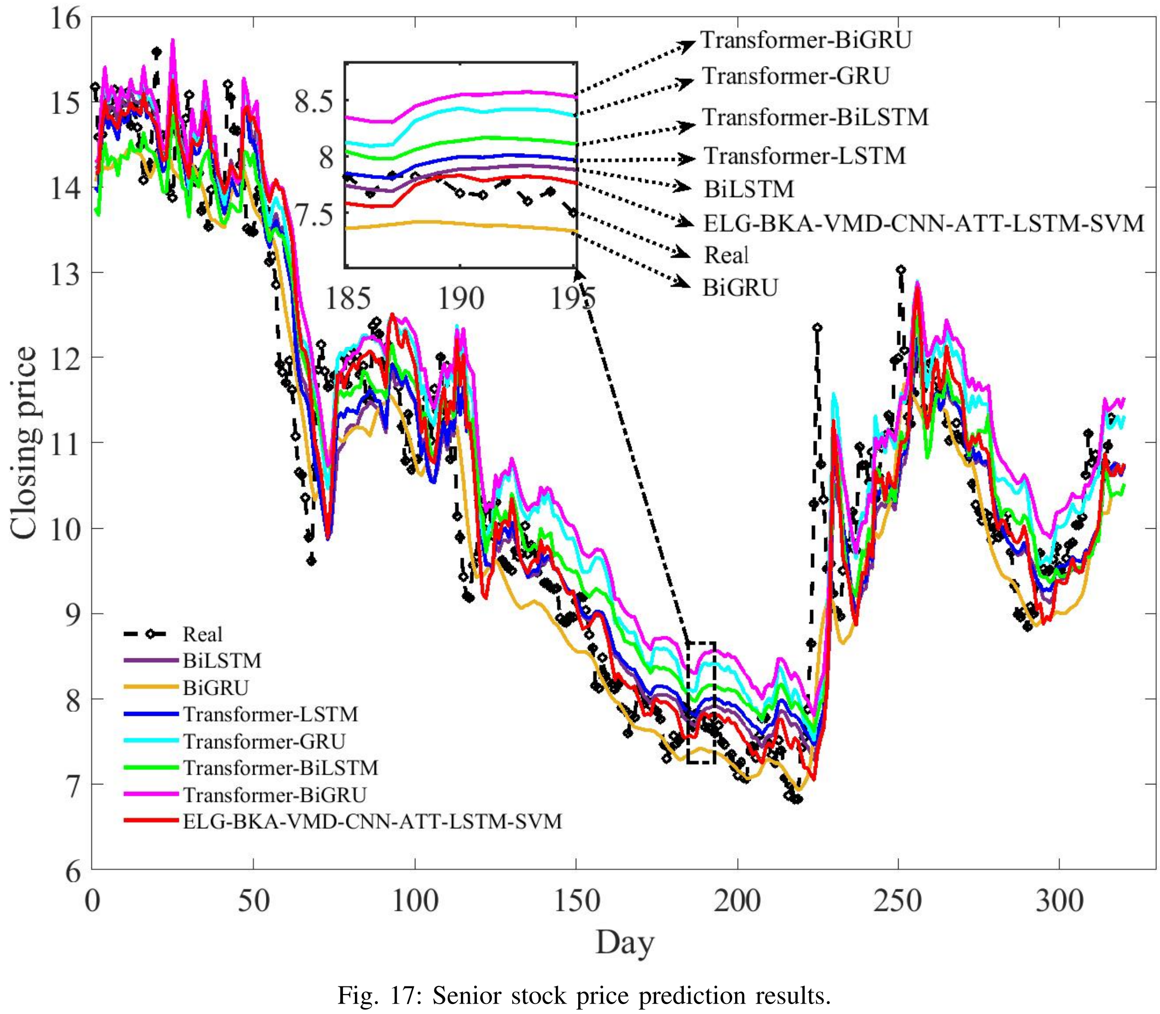

Fig. 17: Senior stock price prediction results.

TABLE VIII: Individual stock prediction evaluation metrics results

Data	Metrics	Models						
		BiLSTM	BiGRU	Transformer-LSTM	Transformer-GRU	Transformer-BiLSTM	Transformer-BiGRU	ELG-BKA-VMD-CNN-ATT-LSTM-SVM
BYD	MAE	7.8672	6.3581	10.6660	13.4076	7.6147	7.7690	4.9651
	MAPE	3.0890	2.4858	4.2872	5.5174	2.9370	2.9779	2.0128
	RMSE	10.5454	9.0793	13.3617	15.8324	11.4870	13.8102	7.3237
	R^2	0.9394	0.9551	0.9028	0.8635	0.9281	0.8974	0.9708
	Rank	3	2	5	7	4	6	1
Gotion	MAE	0.9634	0.2935	0.2002	0.5028	0.4494	0.2730	0.1675
	MAPE	4.8421	1.4732	0.9865	2.5124	0.0217	1.2940	0.8119
	RMSE	1.0367	0.3458	0.2603	0.5954	0.5251	0.3399	0.2271
	R^2	0.6454	0.9606	0.9776	0.8828	0.9090	0.9619	0.9830
	Rank	7	4	2	6	4	3	1
CATL	MAE	6.8818	7.2436	7.2577	8.7236	8.3439	7.8002	6.7230
	MAPE	3.0796	3.3536	3.3680	3.9463	4.0097	3.6430	3.1618
	RMSE	10.1838	10.4471	10.1328	12.4056	10.9907	10.7477	9.6753
	R^2	0.9375	0.9342	0.9381	0.9072	0.9272	0.9303	0.9436
	Rank	3	4	2	8	6	5	1
Senior	MAE	0.3694	0.5108	0.3545	0.5550	0.4366	0.6809	0.3197
	MAPE	3.5161	4.7740	3.4562	5.6979	4.3812	7.0460	2.9811
	RMSE	0.5094	0.6674	0.4826	0.6476	0.5624	0.7856	0.4608
	R^2	0.9511	0.9134	0.9561	0.9210	0.9404	0.8837	0.9600
	Rank	3	6	2	5	4	7	1
Kelai	MAE	1.0723	1.0293	0.9511	0.8732	0.8588	0.9416	0.7907
	MAPE	4.6692	4.4929	4.5884	4.3349	3.7515	4.7757	3.5015
	RMSE	1.6140	1.5564	1.3165	1.2402	1.3919	1.4555	1.1661
	R^2	0.9238	0.9292	0.9493	0.9550	0.9433	0.9381	0.9602
	Rank	7	6	3	2	4	5	1
Sanhua	MAE	1.2713	1.0533	1.2985	1.4215	1.4216	1.3253	0.9475
	MAPE	5.9715	4.8197	5.4322	6.4471	6.4931	6.0732	4.3487
	RMSE	1.4935	1.2475	1.7963	1.6491	1.6394	1.5728	1.1513
	R^2	0.8675	0.9075	0.8090	0.8384	0.8403	0.8530	0.9213
	Rank	3	2	7	6	5	4	1
Sunwoda	MAE	0.6646	0.6380	0.5652	0.5768	0.5696	0.6748	0.5651
	MAPE	3.5589	3.8266	3.1972	3.4892	3.4119	3.7187	3.3539
	RMSE	0.9921	0.8477	0.9050	0.8060	0.8195	0.9953	0.7919
	R^2	0.9318	0.9503	0.9439	0.9550	0.9535	0.9314	0.9566
	Rank	6	4	5	2	3	7	1

TABLE IX: Enhancements of individual stock prediction based on error-compensated hybrid models

Data	Metrics and enhancements			
	MAE	MAPE	RMSE	R^2
BYD	21.91%-62.97%	19.03%-63.52%	19.34%-53.74%	34.93%-78.60%
Gotion	16.34%-82.62%	17.70%-83.23%	12.74%-78.09%	23.88%-95.20%
CATL	7.37%-22.93%	6.12%-19.88%	4.52%-22.01%	8.93%-39.18%
Senior	9.81%-53.04%	13.75%-57.69%	4.53%-41.35%	8.86%-65.60%
Kelai	9.44%-26.26%	19.23%-25.01%	5.97%-27.75%	11.61%-47.81%
Sanhua	10.04%-27.03%	9.77%-19.95%	7.71%-35.91%	14.83%-58.77%
Sunwoda	2.04%-16.26%	3.88%-10.88%	1.79%-20.44%	3.47%-36.70%

providing actionable insights for investment decisions. This approach holds substantial theoretical and practical implications for the research of new energy stock price prediction.

B. Research limitations and future directions

This study exhibits several avenues for enhancement: First, the experimental data covers only a limited set of new energy stocks (e.g., CATL, BYD). Subsequent research should expand the sample coverage to include diverse market environments and cross-sector equities, thereby further validating the model's generalization capability. Second, the model remains inadequate in handling extreme markets and contingencies. Methods capturing the impact of emergencies (e.g., sentiment analytics) should be introduced to enhance robustness. Third, novel intelligent optimization algorithms or improved algorithmic combinations could be explored, with error compensation mechanisms integrated into new frameworks to elevate prediction performance. Future work will focus on the above directions to refine the proposed stock price prediction method, providing stronger support for financial market research.

REFERENCES

[1] X. Chen, F. Yang, Q. Sun, and W. Yi, "Research on stock prediction based on ced-pso-stocknet time series model," *Scientific Reports*, vol. 14, no. 1, p. 27462, 2024.

[2] J. Cai, M. Zhu, J. Wu, X. Chen, J. Xu, and Z.-Y. Shen, "Evaluating the impact of new energy vehicle subsidies on urban carbon emissions: Evidence from 261 chinese cities," *Renewable Energy*, vol. 240, p. 122261, 2025.

[3] A. A. Adebiyi, A. O. Adewumi, and C. K. Ayo, "Comparison of arima and artificial neural networks models for stock price prediction," *Journal of Applied Mathematics*, vol. 2014, no. 1, p. 614342, 2014.

[4] A. Ansari Saleh, S. Pawan Kumar, T. Nguyen Van, T. Nguyen Viet, and H. Vo Minh, "Prediction of bric stock price using arima, suttearima, and holt-winters," *Computers, Materials & Continua*, vol. 70, no. 1, pp. 523–534, 2022.

[5] P. K. Illa, B. Parvathala, and A. K. Sharma, "Stock price prediction methodology using random forest algorithm and support vector machine," *Materials Today: Proceedings*, vol. 56, pp. 1776–1782, 2022.

[6] H. Kang, X. Zong, J. Wang, and H. Chen, "Binary gravity search algorithm and support vector machine for forecasting and trading stock indices," *International Review of Economics & Finance*, vol. 84, pp. 507–526, 2023.

[7] S. Usmani and J. A. Shamsi, "Lstm based stock prediction using weighted and categorized financial news," *PloS one*, vol. 18, no. 3, p. e0282234, 2023.

[8] X. Yan, W. Weihan, and M. Chang, "Research on financial assets transaction prediction model based on lstm neural network," *Neural Computing and Applications*, vol. 33, no. 1, pp. 257–270, 2021.

[9] W. Lu, J. Li, Y. Li, A. Sun, and J. Wang, "A cnn-lstm-based model to forecast stock prices," *Complexity*, vol. 2020, no. 1, p. 6622927, 2020.

[10] Q. Chen, W. Zhang, and Y. Lou, "Forecasting stock prices using a hybrid deep learning model integrating attention mechanism, multi-layer perceptron, and bidirectional long-short term memory neural network," *Ieee Access*, vol. 8, pp. 117 365–117 376, 2020.

[11] S. Das, T. P. Sahu, R. R. Janghel, and B. K. Sahu, "Effective forecasting of stock market price by using extreme learning machine optimized by pso-based group oriented crow search algorithm," *Neural Computing and Applications*, vol. 34, no. 1, pp. 555–591, 2022.

[12] B. Gülmez, "Stock price prediction with optimized deep lstm network with artificial rabbits optimization algorithm," *Expert Systems with Applications*, vol. 227, p. 120346, 2023.

[13] Z. Mustaffa and M. H. Sulaiman, "Stock price predictive analysis: An application of hybrid barnacles mating optimizer with artificial neural network," *International Journal of Cognitive Computing in Engineering*, vol. 4, pp. 109–117, 2023.

[14] I. N. Agustin, "The integration of fundamental and technical analysis in predicting the stock price," *Jurnal Manajemen Maranatha*, vol. 18, no. 2, pp. 93–102, 2019.

[15] P. Muthukumar and J. Zhong, "A stochastic time series model for predicting financial trends using nlp," *arXiv preprint arXiv:2102.01290*, 2021.

[16] H. Niu, K. Xu, and W. Wang, "A hybrid stock price index forecasting model based on variational mode decomposition and lstm network," *Applied Intelligence*, vol. 50, pp. 4296–4309, 2020.

[17] H. Liu and Z. Long, "An improved deep learning model for predicting stock market price time series," *Digital Signal Processing*, vol. 102, p. 102741, 2020.

[18] Y. Lin, Z. Lin, Y. Liao, Y. Li, J. Xu, and Y. Yan, "Forecasting the realized volatility of stock price index: A hybrid model integrating ceemdan and lstm," *Expert Systems with Applications*, vol. 206, p. 117736, 2022.

[19] B. Gülmez, "Ga-attention-fuzzy-stock-net: An optimized neuro-fuzzy system for stock market price prediction with genetic algorithm and attention mechanism," *Heliyon*, p. e42393, 2025.

[20] P. Li, Y. Wei, and L. Yin, "Research on stock price prediction method based on the gan-lstm-attention model." *Computers, Materials & Continua*, vol. 82, no. 1, p. 609, 2025.

[21] X. Zhu, "The role of hybrid models in financial decision-making: Forecasting stock prices with advanced algorithms," *Egyptian Informatics Journal*, vol. 29, p. 100610, 2025.

[22] J. Wang, W.-c. Wang, X.-x. Hu, L. Qiu, and H.-f. Zang, "Black-winged kite algorithm: a nature-inspired meta-heuristic for solving benchmark functions and engineering problems," *Artificial Intelligence Review*, vol. 57, no. 4, p. 98, 2024.

[23] T. Liu, H. Yang, J. Yu, K. Zhou, and F. Jiang, "A global harmony search algorithm based on tent chaos map and elite reverse learning," in *2022 14th International Conference on Advanced Computational Intelligence (ICACI)*. IEEE, 2022, pp. 320–325.

[24] X. Ruize and Y. Lizhu, "A study of gray wolf optimization algorithm based on lens imaging learning strategy for solving vehicle routing problems with capacity and time window constraints," in *IET Conference Proceedings CP872*, vol. 2023, no. 41. IET, 2023, pp. 285–291.

[25] O. R. Adegboye, A. K. Feda, O. R. Ojekemi, E. B. Agyekum, B. Khan, and S. Kamel, "Dgs-scso: Enhancing sand cat swarm optimization with dynamic pinhole imaging and golden sine algorithm for improved numerical optimization performance," *Scientific Reports*, vol. 14, no. 1, p. 1491, 2024.

[26] X. Li, G. Wu, and P. Wu, "Vmd-se-dcnn-cbam: An intelligent cavitation recognition method for canned motor pump," *IEEE Transactions on Instrumentation and Measurement*, pp. 1–13, 2024.

[27] A. Arora, Y. Upadhyay, S. Shukla *et al.*, "Forecasting stock price by lstm-cnn hybrid model and compares deep learning models," in *2024 IEEE International Conference on Smart Power Control and Renewable Energy (ICSPCRE)*. IEEE, 2024, pp. 1–6.

[28] Z. Shi, Y. Hu, G. Mo, and J. Wu, "Attention-based cnn-lstm and xgboost hybrid model for stock prediction," *arXiv preprint arXiv:2204.02623*, 2022.

[29] Z. Xing, R. Chai, K. Chen, Y. Xia, and S. Chai, "Online trajectory planning method for autonomous ground vehicles confronting sudden and moving obstacles based on lstm-attention network," *IEEE Transactions on Cybernetics*, vol. 55, no. 1, pp. 421–435, 2024.

[30] I. T. Akinola, Y. Sun, I. G. Adebayo, and Z. Wang, "Daily peak demand forecasting using pelican algorithm optimised support vector machine (poa-svm)," *Energy Reports*, vol. 12, pp. 4438–4448, 2024.

[31] Z. Ning, B. Wang, S. Li, X. Jia, S. Xie, and J. Zheng, "Pipeline risk factors analysis using the pearson correlation coefficient method and the random forest importance factor method," in *2023 8th International Conference on Smart and Sustainable Technologies (SpliTech)*. IEEE, 2023, pp. 1–5.