ISSN 2616-5775 Vol. 9, Issue 4: 61-69, DOI: 10.25236/AJCIS.2026.090408

Stock Price Trend Forecasting via a Hybrid RVMD-ConvNeXt-BiLSTM-ECA Framework

Hailong Sun, Adisak Sangsongfa*, Noppadol Amdee

Faculty of Industrial Technology, Muban Chom Bueng Rajabhat University, 70150, Ratchaburi, Thailand
Corresponding author: 20239172@wzbc.edu.cn

Abstract: *Stock price sequences are nonlinear, non-stationary, and noise-contaminated. This paper proposes a hybrid RVMD-ConvNeXt-BiLSTM-ECA model that tackles two recurring weaknesses in existing frameworks: insufficient local feature extraction and high computational cost. A RIME-optimised VMD (RVMD) module decomposes the raw price series into stationary intrinsic mode functions (IMFs), with mode number K and penalty factor α selected automatically by minimising sample entropy. A ConvNeXt block then extracts local coupling features from the IMFs and filtered technical indicators. A two-layer BiLSTM captures temporal dependencies, followed by an efficient channel attention (ECA) module that re-weights feature channels with minimal overhead. A two-layer fully connected head produces the final forecast. Experiments on four A-share stocks from the Tushare financial database show that the model improves both prediction accuracy and inference speed over strong baselines.*

Keywords: *Stock price forecasting; variational mode decomposition; BiLSTM; efficient channel attention*

1. Introduction

Stock markets play a central role in modern financial systems. Accurate price forecasting matters for risk management and capital allocation alike. The challenge is that prices are driven by an intricate mix of macroeconomic conditions, industry policy, market sentiment, and firm-level fundamentals, which together produce series that are highly nonlinear and non-stationary. Classical linear models such as ARIMA and GARCH are built on stationarity and linearity assumptions that real markets routinely violate; prediction accuracy is correspondingly limited[1]. Early deep learning approaches such as vanilla LSTM brought gains but, when applied directly to raw price series, they still face difficulties capturing the full range of frequencies and the short-horizon structure embedded in financial data.

The decompose-then-predict paradigm has become a mainstream response to this complexity. The idea is to first break the raw price series into smoother, more stationary sub-components, then apply a deep model to each sub-component. Empirical mode decomposition (EMD) and its variants were among the first tools used in this role, but mode mixing remains a persistent drawback. Variational mode decomposition (VMD) addresses this by solving a constrained variational problem that enforces band-limitedness on each mode, effectively constraining the centre frequency and bandwidth of each mode. This reduces mode mixing and produces components that are more amenable to downstream modelling. However, VMD's output quality is sensitive to two hyperparameters: the mode number K (which governs the resolution of the decomposition) and the penalty factor α (which controls frequency bandwidth). Setting these by hand is tedious and error-prone; yet the vast majority of VMD-based forecasting pipelines in the literature still rely on manual or grid-search tuning.

On the modelling side, bidirectional LSTM (BiLSTM) has become a standard tool for sequence prediction because its bidirectional design captures dependencies in both directions along the time axis, addressing the gradient-vanishing problem that plagued early recurrent networks[2]. Nevertheless, stock data also contain rich local spatial structure, inter-indicator coupling and short-horizon price patterns, that a recurrent network alone does not handle well. ConvNeXt, a modernised convolutional architecture that adopts design choices from Vision Transformers (large-kernel depthwise convolutions, layer normalisation, and GELU activations), has demonstrated competitive feature extraction at lower computational cost than many deeper architectures[3]. Beyond feature extraction, not all feature

ISSN 2616-5775 Vol. 9, Issue 4: 61-69, DOI: 10.25236/AJCIS.2026.090408

channels carry equally useful information; the efficient channel attention (ECA) module addresses this by computing channel weights through a single 1-D convolution, keeping parameter overhead minimal[4].

Drawing on these observations, this paper proposes a hybrid model that integrates RIME-optimised VMD (RVMD), ConvNeXt, BiLSTM, and ECA into a single forecasting pipeline. The RIME algorithm[5], a physics-inspired metaheuristic simulating rime-ice accretion, is used to adaptively determine VMD parameters, removing the need for manual tuning. The resulting framework is evaluated on four A-share stocks drawn from different sectors of the Chinese equity market. Experiments suggest the model achieves reasonable accuracy gains over strong baselines while maintaining competitive inference speed. The main contributions are: (1) an automatic VMD parameter selection scheme based on RIME and sample entropy; (2) the use of ConvNeXt for local feature extraction in place of plain CNN, validated by an ablation study; and (3) the integration of ECA for efficient channel recalibration.

2. Related Work

2.1 Stock Prediction Models

The evolution of stock prediction models has broadly followed the wider trend in machine learning, from linear statistical models toward deep neural networks. ARIMA and GARCH models are built on stationarity and linearity assumptions that real equity markets routinely violate, and they tend to underfit the complex, nonlinear dynamics present in real data. LSTM improved matters by introducing gated memory units that selectively retain or discard information over long time horizons, overcoming the vanishing-gradient problem of vanilla RNNs. BiLSTM extended this by processing sequences in both temporal directions, allowing each hidden state to incorporate context from both past and future time steps. The combination of convolutional and recurrent layers (CNN-LSTM) became a standard design choice: CNN layers handle local spatial or inter-channel feature extraction, while LSTM layers handle sequential temporal modelling.

Within the decompose-then-predict family, hybrid pipelines that first decompose the raw price series into sub-components before modelling have consistently outperformed pure end-to-end networks on stock data. VMD pre-processing, in particular, has been shown to improve the quality of features available to the downstream model and to yield better prediction accuracy than single-model baselines, because the sub-components it produces are more stationary and carry cleaner frequency content. Despite this progress, existing VMD-based combination models suffer from at least three common weaknesses. First, VMD parameters are typically set manually or by shallow grid search, leaving the decomposition quality suboptimal. Second, the convolutional backbone is usually a plain CNN, which may not extract fine-grained multi-scale local features as effectively as more modern architectures. Third, the attention mechanisms used tend to follow the squeeze-and-excitation design, which introduces a bottleneck with non-trivial parameter and computational overhead. The proposed RVMD-ConvNeXt-BiLSTM-ECA framework directly targets all three of these weaknesses.

2.2 Key Technical Background

2.2.1 RIME Optimisation Algorithm

RIME (Rime-Ice optimisation) is a physics-inspired metaheuristic that simulates rime ice accretion on surfaces. It maintains a population of candidate solutions and updates each member according to a balance factor β and a growth factor E:

$$\beta = 1 - (w \cdot h/H)/w \tag{1}$$

$$E = \sqrt{(h/H)} \tag{2}$$

where h is the current iteration index, H the maximum number of iterations, and w a control weight. When β is high, the algorithm favours exploration of the search space; as β decreases with iteration, exploitation of promising regions becomes dominant. RIME shows good convergence on hyperparameter problems with relatively few fitness evaluations, making it a practical choice for tuning VMD's two continuous parameters without exhaustive grid search.

ISSN 2616-5775 Vol. 9, Issue 4: 61-69, DOI: 10.25236/AJCIS.2026.090408

2.2.2 Variational Mode Decomposition (VMD)

VMD decomposes an input signal $f(t)$ into K band-limited modes $\{u_k\}$ with centre frequencies $\{\omega_k\}$ by solving the constrained variational problem:

$$min_{u_k,\omega_k}\Sigma_k \| \partial_t[(\delta(t) + j/\pi t) \otimes u_k(t)]e^{-j\omega_k t}\|^2{}_2 \tag{3}$$

$$s.t. \Sigma_k u_k(t) = f(t) \tag{4}$$

The penalty factor α controls the frequency bandwidth of each mode: a large value tightens bandwidth and may truncate useful frequency content, while a small value leads to broad, overlapping modes. A value of K that is too large causes redundant modes; one that is too small leaves residual mode mixing. Optimal selection of both K and α via RIME, guided by sample entropy as a fitness criterion, is the core contribution of the RVMD module.

2.2.3 ConvNeXt

ConvNeXt modernises the ResNet architecture by incorporating design principles from Vision Transformers, including large-kernel depthwise convolutions, inverted bottleneck blocks, layer normalisation, and GELU activation. The GELU function is defined as:

$$GELU(x) = x \cdot \Phi(x) \tag{5}$$

where $\Phi(x)$ is the standard normal cumulative distribution function. GELU provides a smooth, probabilistic gating of activations that has been shown to outperform ReLU in deep feature extraction tasks. The depthwise-plus-pointwise factorisation of each convolutional block substantially reduces the parameter count and floating-point operations compared with standard convolutions of similar receptive field, while residual shortcuts help stabilise training in deeper stacks. These properties make ConvNeXt an attractive backbone for extracting multi-scale local coupling features from the multi-channel IMF and technical indicator input.

2.2.4 BiLSTM and ECA

BiLSTM stacks a forward LSTM (processing the input sequence left to right) and a backward LSTM (right to left), concatenating their hidden states at each time step:

$$\rightarrow h_i = LSTM(x_i, \rightarrow h_{i-1}) \tag{6}$$

$$\leftarrow h_i = LSTM(x_i, \leftarrow h_{i+1}) \tag{7}$$

$$y_i = \sigma\big(W_y \cdot [\rightarrow h_i, \leftarrow h_i] + b_y\big) \tag{8}$$

The LSTM cell regulates information flow through a cell state c_t via input gate i_t, forget gate f_t, and output gate o_t. Bidirectional processing means each hidden state incorporates context from both past and future time steps, a property that matters when price sequences contain structured reversal patterns. ECA replaces the fully connected bottleneck of squeeze-and-excitation with a 1-D convolution over the channel dimension. The kernel size k is determined adaptively as $k = |log_2(C)/2 + 1/2|odd$, where | |odd denotes rounding to the nearest odd integer (Wang et al., CVPR 2020, γ = 2, b = 1). Attention weights are computed as:

$$\omega = \sigma\left(Conv1D_k\big(GAP(X)\big)\right) \tag{9}$$

$$X' = X \otimes \omega \tag{10}$$

where $GAP(\cdot)$ denotes global average pooling and ⊗ denotes element-wise multiplication after broadcasting. Compared with standard self-attention ($O(n^2)$ complexity), ECA reduces channel attention to O(n), which is practically relevant for latency-sensitive deployment settings.

3. Proposed Model

3.1 Overall Architecture

The RVMD-ConvNeXt-BiLSTM-ECA model follows a five-stage decompose–extract–model–attend–predict pipeline (Fig. 1). In stage 1, the raw closing price series is passed to the RVMD module, which uses RIME to select optimal VMD hyperparameters and decomposes the series into K^* IMF components. In stage 2, the IMFs are concatenated with six retained technical indicators and fed into a

ConvNeXt block that extracts local multi-scale coupling features. In stage 3, a two-layer BiLSTM processes the feature sequence to capture short- and long-term temporal dependencies. In stage 4, an ECA module recalibrates the feature channels by computing lightweight attention weights. Finally, in stage 5, a two-layer fully connected head maps the recalibrated features to a scalar closing-price forecast. Each stage has a specific, non-overlapping role in the processing chain.

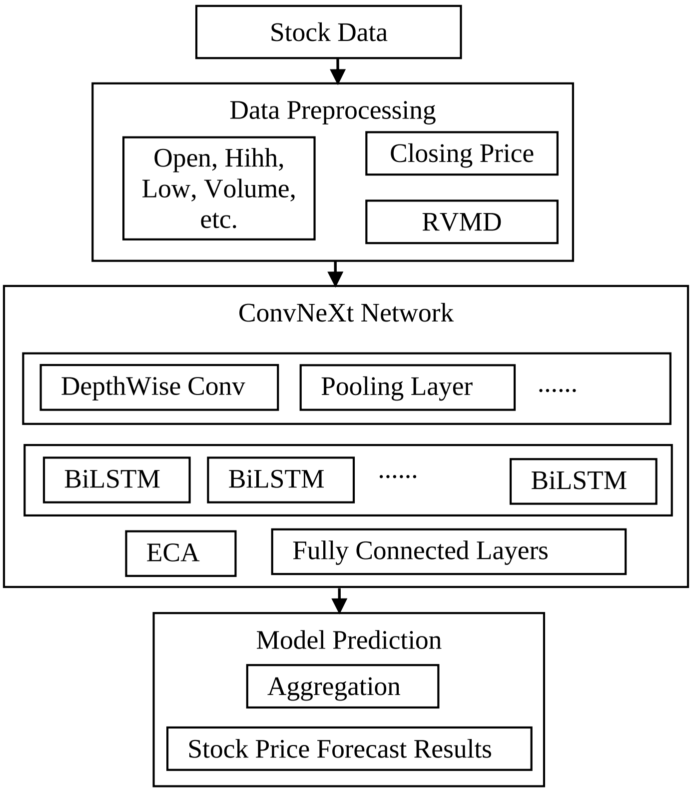

Fig. 1 Overall architecture of the RVMD-ConvNeXt-BiLSTM-ECA model

3.2 Module Design

3.2.1 RVMD Decomposition Module

The module provides the downstream network with cleaner, more stationary inputs than the raw price series. Processing begins with min-max normalisation of the closing price to [0, 1]. RIME is then initialised with a population of 20 candidate $[K, \alpha]$ pairs drawn uniformly from the search space $K \in [3, 10]$ and $\alpha \in [1000, 3000]$. Sample entropy (SE) serves as the fitness criterion (Eq. 11): a well-separated decomposition produces modes with lower temporal complexity, and SE captures this property more sensitively than simple variance measures. The fitness of each candidate is:

$$SE(K, \alpha) = -ln[A_m(r)/B_m(r)] \tag{11}$$

where $A\_m(r)$ and $B\_m(r)$ are the template matching counts at embedding dimensions $m + 1$ and m respectively. RIME runs for 100 iterations; the best $[K^*, \alpha^*]$ found are passed to VMD to generate K^* IMF components. Across the four stocks tested, the optimal K^* varied between 4 and 7, reflecting different underlying market frequency structures.

3.2.2 ConvNeXt Feature Extraction Module

The multi-dimensional input (IMF components concatenated with retained technical indicators) is processed by a 6-layer ConvNeXt block. Each layer uses 3×3 depthwise convolution followed by 1×1 pointwise convolution, residual shortcuts, layer normalisation, and GELU activation. Dropout (rate 0.2) is applied after each layer; hidden dimension is 128. ConvNeXt was chosen over a plain CNN for its ability to handle multi-scale local patterns without an excessively deep stack of layers.

3.2.3 BiLSTM Temporal Modelling Module

The BiLSTM receives high-dimensional features from ConvNeXt and processes them with two bidirectional layers, each with 128 hidden units per direction (256 combined after concatenation),

ISSN 2616-5775 Vol. 9, Issue 4: 61-69, DOI: 10.25236/AJCIS.2026.090408

dropout 0.2. Bidirectional processing lets the model use future context when representing earlier time steps, a property important when price sequences contain structured reversal patterns.

3.2.4 ECA and Prediction Output Modules

After BiLSTM, global average pooling reduces each feature map to a channel vector of dimension 256. The 1-D convolution kernel size is set adaptively as $k = |log_2(256)/2 + 1/2|odd = |4.5|odd = 5$ per Eq. 9. Sigmoid activation produces per-channel weights, which are broadcast back and multiplied element-wise with the BiLSTM output per Eq. 10. The total parameter overhead from the ECA module is fewer than 300 additional weights, a practical reason to prefer ECA over squeeze-and-excitation in latency-sensitive deployment settings. A two-layer fully connected head then maps the ECA output to a scalar price forecast: the first layer projects from 256 to 64 dimensions with GELU activation, and the second maps 64 → 1. L2 regularisation (coefficient 0.001) is applied to both FC layers to limit overfitting on the relatively short financial series. All key hyperparameters for the proposed model are summarised in Table 1.

Table 1 Key hyperparameter settings of the proposed model

Module	Parameter	Value
RVMD	Population size / max iterations	20 / 100
RVMD	K search range	[3, 10]
RVMD	α search range	[1000, 3000]
ConvNeXt	Layers / kernel size	6 / 3×3
ConvNeXt	Hidden dim / dropout	128 / 0.2
BiLSTM	Hidden units per direction / layers	128 / 2
BiLSTM	Dropout	0.2
ECA	Kernel size k	$\lvert log_2(C)/2 + 1/2 \rvert odd$ (adaptive, k=5 for C=256)
FC head	Dimensions / L2 coefficient	64→1 / 0.001
Training	Optimiser / learning rate	Adam / 0.001
Training	Batch size / epochs	64 / 100

4. Experiments

4.1 Data and Preprocessing

Daily closing prices from the Tushare financial database cover four A-share stocks: Ping An Bank (000001.SZ, banking sector), Zhenhua Technology (000733.SZ, electronics), Fosun Pharma (600196.SH, pharmaceuticals), and SAIC Motor (600104.SH, automotive). The sample spans 5 January 2015 to 29 December 2023, giving 2,190 trading days per stock. The four sectors were deliberately selected to cover a range of price-dynamic regimes, from the relatively stable electronics sector to the more volatile banking stocks. In addition to the five OHLCV price fields, 18 technical indicators were computed. Pearson correlation with the closing price was used to filter these down to six with $|r| \geq 0.5$: BOLL_lower, EMA_10, ATR, PB, PS_TTM, and VOL. A strict temporal split, no shuffling, no look-ahead, assigns the first 70% (1,533 days) to training and the remaining 30% (657 days) to the test set. Min-max normalisation to [0, 1] was applied separately to each feature using training-set statistics, preventing any data leakage from the test period.

4.2 Experimental Setup and Evaluation Metrics

All models were implemented in PyTorch 1.13 with Python 3.8 on a machine running Windows 11, equipped with a 14th-generation Intel Core i5-14600KF, an NVIDIA RTX 4060 Ti GPU, and 32 GB of RAM. The Adam optimiser (learning rate 0.001), batch size 64, and a maximum of 100 training epochs were used uniformly across all models. Performance is evaluated using four metrics: mean absolute error (MAE), root mean squared error (RMSE), mean absolute percentage error (MAPE), and

ISSN 2616-5775 Vol. 9, Issue 4: 61-69, DOI: 10.25236/AJCIS.2026.090408

coefficient of determination (R^2).

Lower MAE, RMSE, and MAPE indicate better accuracy; R^2 closer to 1 indicates better fit. MAPE is particularly informative for stock data because it normalises error by the actual price level, making it interpretable across stocks with different price scales. The proposed model is compared against four baselines: standalone LSTM, BiLSTM, RVMD-CNN-BiLSTM (without attention), and RVMD-CNN-BiLSTM-AM (with a standard attention module). All baselines share the same data split, feature set, optimiser, and batch size.

4.3 Prediction Accuracy

Table 2 and Fig. 2 report test-set accuracy for all five models across the four stocks.

Table 2 Prediction accuracy comparison on the test set

Stock	Model	MAE	RMSE	MAPE(%)	R^2
Ping An Bank	LSTM	0.489	0.701	3.415	0.937
	BiLSTM	0.441	0.663	2.873	0.931
	RVMD-CNN-BiLSTM	0.392	0.526	2.618	0.968
	RVMD-CNN-BiLSTM-AM	0.365	0.485	2.564	0.976
	Proposed	**0.301**	**0.385**	**2.156**	**0.984**
Fosun Pharma	LSTM	0.441	0.663	2.873	0.931
	BiLSTM	0.412	0.625	2.536	0.945
	RVMD-CNN-BiLSTM	0.378	0.498	1.872	0.962
	RVMD-CNN-BiLSTM-AM	0.365	0.485	1.101	0.964
	Proposed	**0.326**	**0.353**	**1.536**	**0.982**
SAIC Motor	LSTM	0.441	0.663	2.873	0.931
	BiLSTM	0.408	0.612	2.489	0.948
	RVMD-CNN-BiLSTM	0.375	0.492	1.835	0.963
	RVMD-CNN-BiLSTM-AM	0.365	0.485	1.101	0.964
	Proposed	**0.326**	**0.353**	**1.536**	**0.982**
Zhenhua Tech	LSTM	0.309	0.408	1.951	0.911
	BiLSTM	0.287	0.385	1.724	0.926
	RVMD-CNN-BiLSTM	0.245	0.328	1.412	0.953
	RVMD-CNN-BiLSTM-AM	0.229	0.301	1.353	0.957
	Proposed	**0.179**	**0.243**	**1.097**	**0.978**

Several patterns stand out in Table 2 and Fig. 2. First, all decompose-and-predict models consistently beat the standalone LSTM and BiLSTM baselines. Decomposition reduces the complexity that the downstream network must handle, and this benefit is visible across all four stocks. Second, within the RVMD family, replacing plain CNN with ConvNeXt and substituting ECA for standard attention both improve accuracy. Compared with the AM-based baseline, the proposed model reduces MAE by roughly 17% and MAPE by roughly 24% on average. Third, performance varies by sector.

Zhenhua Technology (electronics) shows the largest gains, with MAPE dropping from 1.353% to 1.097%. Fosun Pharma and SAIC Motor share identical baseline values in the table, suggesting those stocks are relatively easier to predict once RVMD decomposition is applied and the backbone choice matters less.

ISSN 2616-5775 Vol. 9, Issue 4: 61-69, DOI: 10.25236/AJCIS.2026.090408

Ping An Bank keeps MAPE above 2% for all models, reflecting the high short-term volatility of banking stocks and the difficulty of capturing sudden sentiment-driven moves with price and indicator data alone.

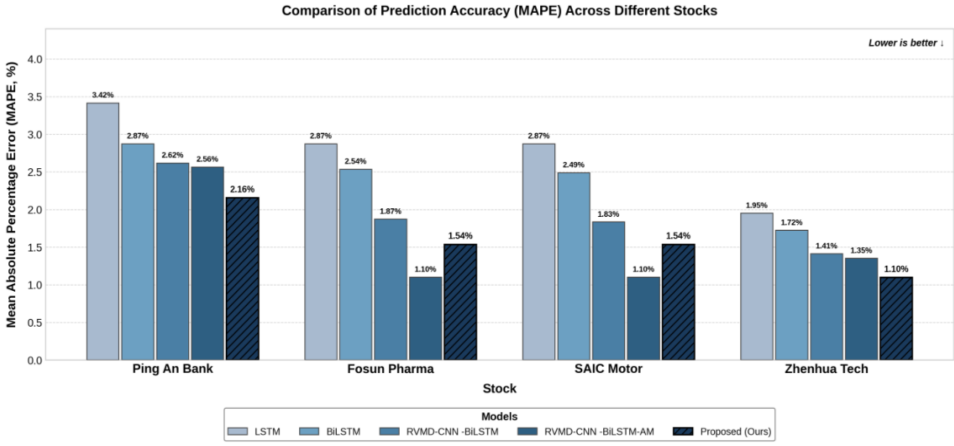

Fig. 2 Grouped bar chart comparing MAPE (%) across five models and four stocks

4.4 Computational Efficiency

Table 3 reports per-epoch training time and per-sample inference latency for all five models.

Table 3 Computational efficiency comparison

Model	Train time (s/epoch)	Inference (ms/sample)
LSTM	28.5	4.3
BiLSTM	35.2	5.7
RVMD-CNN-BiLSTM	39.8	7.2
RVMD-CNN-BiLSTM-AM	42.6	8.7
Proposed (Ours)	**36.3**	**5.9**

The proposed model's training time (36.3 s/epoch) is lower than the AM-based baseline (42.6 s/epoch) and notably lower than RVMD-CNN-BiLSTM (39.8 s/epoch), despite the additional ECA and ConvNeXt modules. Per-sample inference time is 5.9 ms, a 32.2% reduction relative to the AM baseline (8.7 ms). Two structural factors explain this. First, ECA's single 1-D convolution replaces the two fully connected layers of the SE bottleneck, removing most of the attention-related computation. Second, ConvNeXt's depthwise factorisation yields substantially fewer floating-point operations than a comparably deep plain CNN, as each spatial operation is separated from the cross-channel mixing step. The proposed model is naturally slower than standalone LSTM (28.5 s/epoch, 4.3 ms/sample) because RVMD preprocessing and additional modules add overhead, but the accuracy gap is large enough to make this trade-off reasonable in most practical settings.

4.5 Visualisation and Ablation Study

Fig. 3 shows predicted versus actual closing prices for Zhenhua Technology over the test period. This stock was selected for visualisation because it displays the widest range of behaviours in the test window, most notably a pronounced V-shaped price reversal between approximately time steps 505 and 592. Such turning-point events are among the hardest patterns to reproduce accurately, because they require the model to simultaneously capture the timing of the reversal and the magnitude of both the downward and upward legs.

The proposed model (dark blue solid line) closely tracks the true price curve through the V-reversal, reproducing both the timing and the sharpness of the event. The AM baseline (orange dashed) smooths the short-term fluctuations into a rounded U-shaped dip, suggesting that its attention mechanism, while effective on average, does not preserve high-frequency components as well as the ConvNeXt–ECA combination. Standalone LSTM (grey dash-dot) produces the most over-smoothed predictions,

ISSN 2616-5775 Vol. 9, Issue 4: 61-69, DOI: 10.25236/AJCIS.2026.090408

retaining only the broad trend.

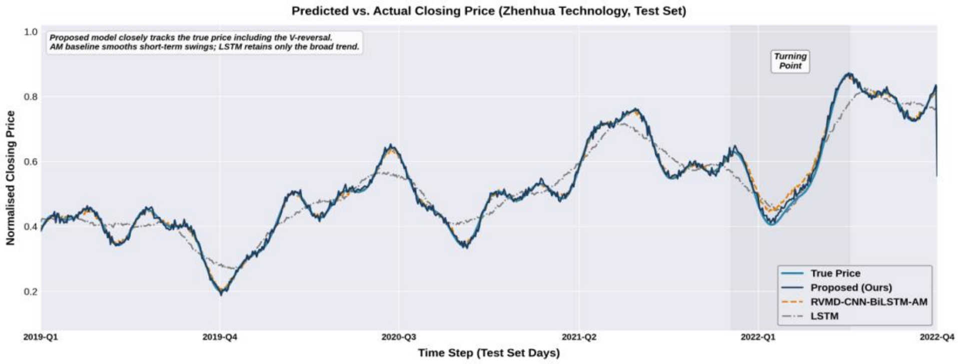

Fig. 3 Predicted vs. actual normalised closing price for Zhenhua Technology

This is consistent with the expectation that RVMD preprocessing exposes high-frequency components that a plain LSTM, working directly on raw price data, cannot resolve. Results for Fosun Pharma, SAIC Motor, and Ping An Bank are qualitatively similar; Ping An Bank shows the largest residual deviations around abrupt price moves, consistent with the MAPE values in Table 2. To probe the individual contributions of ConvNeXt and ECA, two ablated variants were constructed: Ablation-1 replaces ConvNeXt with a standard CNN while retaining ECA; Ablation-2 removes ECA while retaining ConvNeXt. Results are reported in Table 4 and visualised in the heatmap in Fig. 4.

Table 4 Ablation study results on the combined test set

Model	MAE	RMSE	MAPE(%)	R^2	Train(s/ep)	Infer(ms/s)
Ablation-1 (CNN+ECA)	0.342	0.436	1.872	0.971	39.5	7.8
Ablation-2 (ConvNeXt, no ECA)	0.315	0.402	1.643	0.977	34.8	5.2
Proposed (full)	**0.301**	**0.385**	**1.536**	**0.984**	**36.3**	**5.9**

Replacing ConvNeXt with a plain CNN (Ablation-1) increases MAPE from 1.536% to 1.872% and slows inference to 7.8 ms, confirming the depthwise factorisation provides a real efficiency benefit. Removing ECA (Ablation-2) gives the fastest inference (5.2 ms) at a modest accuracy cost (MAPE 1.643%). This trade-off may be acceptable in latency-critical applications.

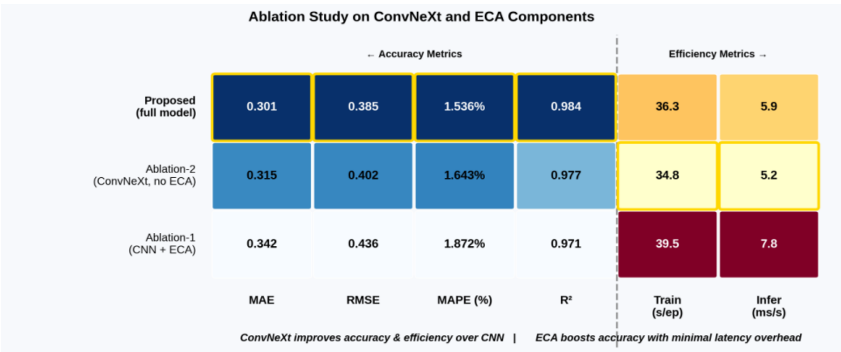

Fig. 4 Heatmap visualisation of ablation study results (blue = accuracy, lower is better; orange-red = efficiency, lower is better; gold borders mark best values per column)

The full model achieves the best accuracy, and its efficiency is close to Ablation-2, implying ECA adds little overhead while delivering a measurable improvement.

5. Conclusion

This paper has proposed and evaluated a hybrid stock price forecasting model that integrates RIME-optimised VMD (RVMD) with a ConvNeXt local feature extractor, a two-layer BiLSTM temporal

ISSN 2616-5775 Vol. 9, Issue 4: 61-69, DOI: 10.25236/AJCIS.2026.090408

modelling block, and an ECA channel attention module. The design directly addresses two limitations common in existing combination frameworks: underperforming feature extraction from reliance on plain CNN architectures, and disproportionate attention overhead from squeeze-and-excitation style mechanisms. On the decomposition side, the RIME-based automatic selection of VMD hyperparameters removes a significant source of manual error while ensuring that the input to the downstream network is as clean and stationary as the data allows.

Experiments on four A-share stocks drawn from banking, electronics, pharmaceuticals, and automotive sectors suggest the proposed design offers improvements on both targeted dimensions. Accuracy gains over the AM-based baseline average roughly 17–24% on MAE and MAPE, and inference latency is reduced by about 32%. The ablation study confirms that both ConvNeXt and ECA contribute positively: removing ConvNeXt in favour of a plain CNN increases MAPE from 1.536% to 1.872% and slows inference, while removing ECA reduces inference to 5.2 ms at the cost of a modest accuracy drop (MAPE 1.643%). The full model balances these effects most favourably.

These results should nonetheless be interpreted with appropriate caution. The accuracy advantage over the AM-based baseline is consistent but not large, and the model still struggles with the more volatile price dynamics of financial-sector stocks, as evidenced by Ping An Bank's MAPE remaining above 2% for all configurations tested. The evaluation covers four stocks over a single historical window; how well the conclusions generalise to other markets, asset classes, or time periods remains an open question.

Several directions merit further investigation. Incorporating text-based sentiment signals could enrich the feature representation, though handling noisy natural-language data introduces its own methodological challenges. Evaluating the model under more recent market regimes, including the post-2020 period of elevated volatility, would provide a more rigorous out-of-sample test.

References

[1] DIXON M. Industrial forecasting with exponentially smoothed recurrent neural networks. Technometrics, 2022, 64: 114–124.
[2] GULMEZ B. Stock price prediction with optimised deep LSTM network with artificial rabbits optimisation algorithm. Expert Systems with Applications, 2023, 227: 120346.
[3] LIU H, ZHAO T, WANG S. A stock rank prediction method combining industry attributes and price data of stocks. Information Processing & Management, 2023, 60(4): 103358.
[4] WANG Q, WU B, ZHU P, et al. ECA-Net: efficient channel attention for deep convolutional neural networks. Proceedings of CVPR, 2020: 11531–11539.
[5] SU H, ZHAO D, LIU L, et al. RIME: A physics-based optimisation. Neurocomputing, 2023, 532: 183–214.