Market State-Aware Dynamic Weighting Model for Price-Based Investment Decision

Qi Yan
Nanjing University of Finance and Economics, Massey Institute, Financial Engineering, Nanjing, Jiangsu, 210023, China
24012330@massey.ac.nz

ABSTRACT

Financial markets exhibit significant regime-switching characteristics, and the value of information sources varies across different market states. While existing fusion models have been widely studied, the current paper focuses on the price modality to validate the proposed dynamic weighting mechanism. Most existing fusion approaches adopt static weighting or implicit dynamic interactions, lacking explicit awareness of the macro-market environment, which leads to performance degradation when market states switch. To address this issue, this paper proposes a functional dynamic attention fusion model. The model uses 20-day historical volatility as a proxy for market state and generates dynamic weights through a piecewise linear function, reducing the weight of price features as volatility increases, thereby decreasing reliance on price trends during high-volatility periods. Using daily data of the CSI 300 index from 2020 to 2024, the prediction accuracy and investment performance of the static model and the dynamic model are compared. Empirical results show that the dynamic model achieves an annualized return of 4.91%, numerically higher than the static model's 2.67%; its Sharpe ratio (0.206) is better than that of the static model (0.046); annualized volatility is slightly lower; and the maximum drawdown remains the same as the static model. Prediction accuracy improves from 50.76% to 50.78%. The proposed model requires no training, is computationally simple, fully transparent, and can be reproduced entirely in Excel, providing a low-cost, highly interpretable new paradigm for dynamic feature weighting in quantitative investment research.

Keywords: Dynamic attention; Market state; Volatility; Multimodal fusion; Investment decision; Functional model

1. INTRODUCTION

1.1 Research Background

Financial markets are inherently complex and time-varying dynamic systems, and the value of information sources differs significantly under different market states. Using multimodal data (prices, news, images, etc.) to improve investment decision-making has become a research hotspot. However, most existing fusion models employ static weighting or implicit dynamic interactions, lacking explicit awareness of the macro-market environment, which causes performance degradation when market states switch. At the same time, highly complex models are difficult to implement and lack interpretability and lightweight deployment.

1.2 Problem Statement and Research Significance

To address the above issues, this study explores a new paradigm of "environmentally conditioned" dynamic fusion: using a quantifiable market state indicator (such as volatility) as input, fusion weights are generated in real time through a preset function that conforms to financial logic. The model requires no training, is fully transparent, and is computationally minimal, aiming to balance dynamic adaptability, interpretability, and lightweight implementation.

1.3 Research Content and Innovations

Research content: Using 20-day historical volatility as the market state proxy variable, a piecewise linear weighting function is designed to achieve dynamic adjustment of price features; using daily data of the CSI 300 index from 2020 to 2024, the prediction accuracy and investment performance of the static and dynamic models are compared.

Innovations: (1) Explicitly introducing market state into the fusion layer, shifting from "internal optimization" to "external environmental conditioning"; (2) Proposing "functional dynamic attention", which requires no training and generates weights directly from preset functions, offering clear interpretability; (3) Achieving a “modular” lightweight innovation, where the entire process can be reproduced in Excel.

1.4 Paper Structure

This paper consists of six chapters. Chapter 1 is the introduction, Chapter 2 reviews the theoretical basis and literature, Chapter 3 describes the model construction, Chapter 4 explains the empirical design, Chapter 5 presents the empirical results and analysis, and Chapter 6 concludes and discusses future directions.

2. THEORETICAL BASIS AND LITERATURE REVIEW

2.1 Financial Multimodal Information Fusion

Multimodal fusion aims to leverage complementary information across modalities to improve prediction accuracy. Early methods mainly used feature concatenation or weighted averaging (Atrey et al., 2010) [8]. Subsequently, deep learning was employed to fuse text and price series. Ma et al. (2023) [4] proposed a multi-source aggregated classification method that fuses news from multiple companies for stock price prediction. However, most existing fusion strategies are implicit and globally fixed, lacking explicit awareness of the external market environment.

2.2 Market State Modeling

Methods for identifying market states include volatility models (Engle, 1982; Bollerslev, 1986) [1] [2], hidden Markov models (Hamilton, 1989) [3], and the VIX index. Cooper et al. (2004) [5] found that the profitability of momentum strategies depends on market states. Existing applications mainly focus on asset allocation or strategy switching, rarely incorporating market state into the microscopic process of feature fusion. This provides a starting point for this study.

2.3 Attention Mechanisms and Explainable AI

The interpretability of traditional attention weights is controversial (Jain & Wallace, 2019) [6]. The financial field has a strong demand for model interpretability, and rule-based models and prior knowledge encoding are valued (Rudin, 2019) [7]. The functional dynamic attention proposed in this paper is an ex-ante interpretable mechanism, where weights are generated by a preset financial logic function without training.

2.4 Research Gaps and Positioning of This Paper

Existing research has three gaps: at the theoretical level, a bridge between market state and fusion weights is lacking; at the technical level, there is a lack of dynamic and lightweight fusion solutions; at the interpretability level, there is a lack of ex-ante transparent attention mechanisms. This paper positions itself as follows: using rolling historical volatility as the market state variable, dynamic weights are generated through a piecewise linear function, achieving a dynamic fusion that requires no training, is fully transparent, and is reproducible.

3. MODEL CONSTRUCTION

This chapter systematically describes the market state-aware multimodal fusion model based on the core idea of “functional dynamic attention”. It first formalizes the problem, then designs the dynamic weight generation function, presents the fusion and decision process, and finally compares with existing typical models to clarify the innovative positioning of the proposed model.

3.1 Problem Formalization

3.1.1 Basic Setup

Consider a financial time series prediction task. At each trading day (t), we observe information from different modalities:

- Price modality feature $x_t^p \in \mathbb{R}^{d_p}$: feature vector derived from historical prices, volume, etc. In this paper, we use three simple and effective price indicators: daily return, 5-day return, and volume change rate, and obtain a one-dimensional comprehensive price feature through equal-weighted averaging (see Chapter 4 for details).
- Text modality feature $x_t^t \in \mathbb{R}^{d_t}$: : sentiment scores from news headlines, earnings reports, etc. (not used in this paper, left for future extension).
- Image modality feature $x_t^i \in \mathbb{R}^{d_i}$: deep features from company logos, satellite images, etc. (not used in this paper).

The empirical part of this paper focuses on the price modality alone to validate the effectiveness of the dynamic weighting mechanism. Extension to multimodal scenarios (e.g., adding news sentiment or image features) is left for future work.

3.1.2 Market State Variable

Financial markets exhibit significant regime-switching characteristics. We introduce an observable market state variable $v_t \in \mathbb{R}^+$ to characterize the current market's risk or volatility level. In this paper, v_t is defined as the 20-day historical volatility (annualized):

$$v_t = \sqrt{252} \times \mathrm{Std}(r_{t-19}, r_{t-18}, \dots, r_t) \tag{1}$$

where $r_t = \frac{P_t - P_{t-1}}{P_{t-1}}$ is the daily return, and P_t is the closing price on day t. This variable is computed entirely from historical price data without any future information.

3.1.3 Dynamic Fusion Objective

Traditional static fusion models assume constant importance of each modality, i.e., the fused feature $\tilde{x}_t = w_p x_t^p + w_t x_t^t + w_i x_t^i$, where w_p, w_t, w_i are fixed weights. The goal of this paper is to make the fusion weights adapt to the market state v_t:

$$\tilde{x}_t = w_p(v_t) \cdot x_t^p + w_t(v_t) \cdot x_t^t + w_i(v_t) \cdot x_t^i \tag{2}$$

where the weight functions $w(v)$ satisfy $w_p(v) + w_t(v) + w_i(v) = 1$, and $w_k(v) \in [0,1]$. In this way, the model can dynamically adjust its trust in different information sources according to the current market volatility level.

Finally, the next-period price movement prediction is made based on the fused feature:

$$\hat{y}_{t+1} = \mathrm{sign}(\tilde{x}_t) \tag{3}$$

where $\hat{y}_{t+1} = 1$ denotes a predicted increase, and $\hat{y}_{t+1} = 0$ denotes a predicted decrease.

3.2 Functional Dynamic Attention Mechanism

3.2.1 Design Principles

The dynamic attention mechanism proposed in this paper follows three principles:

1. Interpretability: The weight variation should follow a clear financial logic, not black-box parameters.
2. Lightweight: No iterative training is required, computational cost is extremely low, and reproduction is easy.
3. Continuity: Weights should change continuously with market state, avoiding abrupt jumps from discrete switching.

Based on the above principles, we abandon the mainstream approach of learning attention weights via neural networks and instead generate weights directly using a preset mathematical function grounded in financial priors. We call this "functional dynamic attention".

3.2.2 Specific Form of the Weight Function

Taking the price modality as an example, we design a piecewise linear function that monotonically decreases the weight of price features as volatility increases. The financial intuition is: in high-volatility (panic) markets, price trends are easily overwhelmed by noise, and trend-following strategies tend to fail, so reliance on price signals should be reduced; in low-volatility (calm) markets, price trends have higher predictive power and should be given higher weight.

Let v_{low} and v_{high} be the low-volatility and high-volatility thresholds, respectively. The dynamic weight for the price modality $w_p(v)$ is defined as:

$$w_p(v) = \begin{cases} 1, & v \le v_{low} \\ 1 - \dfrac{v - v_{low}}{v_{high} - v_{low}}, & v_{low} < v < v_{high} \\ 0, & v \ge v_{high} \end{cases} \tag{4}$$

For the text modality, its weight should be positively correlated with volatility: during high volatility, market sentiment has a stronger impact on prices, so the importance of news sentiment increases. Hence, the text modality weight $w_t(v) = 1 - w_p(v)$ (in the bimodal case). If an image modality is included, its weight can be set as constant (not varying with market state) and then normalized via softmax.

Parameter selection: Based on the distribution of the 20-day historical volatility in the sample period (see descriptive statistics in Chapter 5), we set $v_{low} = 0.15,\ v_{high} = 0.35$. These two values correspond approximately to the 25th and 75th percentiles of the historical volatility, effectively distinguishing calm periods from high-volatility periods.

3.2.3 Comparison with Parametric Attention

In traditional Transformer self-attention, attention weights are computed via query-key dot products and normalized by softmax; the dynamics come from the input data itself, but there is no direct mapping between the weights and interpretable external variables (e.g., volatility). In contrast, the functional dynamic attention proposed in this paper has the following advantages, as summarized in Table 1.

Table 1. Comparison with parametric attention

Characteristic	Parametric attention (e.g., Transformer)	Functional dynamic attention (this paper)
Source of dynamics	Interaction among input data	External market state variable
Interpretability	Post-hoc explanation (e.g., attention visualization)	Ex-ante transparency (explicit formula)
Training requirement	Needs large amounts of data for training	No training, direct computation
Computational cost	High (matrix multiplication)	Extremely low (simple arithmetic)
Number of parameters	Millions	0 (or only threshold parameters)

3.3 Fusion and Decision Process

3.3.1 Unimodal Dynamic Weight (Empirical part of this paper)

The empirical part of this paper focuses on the price modality alone, so the fused feature simplifies to:

$$\tilde{x}_t = w_p(v_t) \cdot x_t \tag{5}$$

where x_t is the comprehensive price feature (equally weighted average of daily return, 5−day return, and volume change rate). Here, the dynamic weight $w_p(v_t)$ acts as a “discount factor”: during high volatility, we reduce our trust in price trends (lowering the absolute value of the fused feature), thereby reducing trading frequency and position size; during low volatility, we fully trust price trends (weight close to 1).

3.3.2 Decision Rule

After obtaining the fused feature $\tilde{x}_t$, we adopt the simplest sign prediction rule:

$$\hat{y}_{t+1} = \begin{cases} 1, & \tilde{x}_t > 0 \\ 0, & \tilde{x}_t < 0 \\ 0, & \tilde{x}_t = 0 \end{cases} \tag{6}$$

That is, we predict an increase (1) or a decrease (0) for the next trading day. This rule requires no parameters and is completely transparent.

3.3.3 Backtesting Strategy

Based on the prediction signal, we construct the following timing strategy:

- If $\hat{y}_{t+1} = 1$, fully invest in the CSI 300 index;

- If $\hat{y}_{t+1} = 0$, stay in cash (no position).

The daily return of the strategy equals the position state multiplied by the index's daily return. The initial capital is 1 unit, and the risk-free rate is set at 2% (annualized).

3.4 Static Benchmark Model

To validate the effectiveness of dynamic weighting, we define a static benchmark model whose fused feature directly uses the original price feature, i.e., the price modality weight is fixed at $w_p = 1$:

$$\tilde{x}_t^{\text{static}} = x_t \tag{7}$$

The prediction rules and backtesting strategy of the static model are exactly the same as those of the dynamic model. By comparing the prediction accuracy, annualized return, Sharpe ratio, and maximum drawdown of the dynamic and static models, the improvement brought by the dynamic attention mechanism can be quantified.

3.5 Comparison with Existing Models

This section compares the proposed model with three typical models, as shown in Table 2.

Table 2. Comparison with existing models

Model type	Representative work	Dynamicity	Interpretability	Computational cost
Static weighted fusion	Atrey et al. (2010)	None	High	Very low
End-to-end attention network	Ma et al. (2023)	Implicit (data-driven)	Low	High
Regime-switching model	Cooper et al. (2004)	Discrete regime switching	Medium	Medium
Proposed model		Continuous, explicit, function-driven	High	Very low

The greatest feature of the proposed model is that it introduces the external macro variable of market state into the microscopic decision-making of feature fusion in a completely transparent, training-free manner, filling the gap in "environmentally conditioned fusion" in existing research.

3.6 Chapter Summary

This chapter constructed a market state-aware functional dynamic attention fusion model. The core innovation is using 20-day historical volatility as the market state proxy variable, designing a piecewise linear function to generate dynamic weights, so that the fused feature decays as volatility rises, thereby reducing reliance on price trends during high-volatility periods. The model requires no training, is computationally minimal, and is fully transparent. The next chapter will detail the empirical design, data sources, feature construction, and evaluation metrics based on real CSI 300 index data.

4. EMPIRICAL DESIGN

This chapter details the data sources, sample period, variable definitions and calculation methods, model settings, evaluation metrics, and backtesting setup, laying the foundation for the empirical results in Chapter 5. All calculations are based on real market data and can be fully reproduced in spreadsheet software.

4.1 Data Sources and Sample Period

4.1.1 Data Sources

The raw data used in this study are daily quotes of the CSI 300 index, including opening price, highest price, lowest price, closing price, and volume. The data were obtained via public data interfaces (e.g., Tushare Pro, JoinQuant, or 163.com), covering the period from January 2, 2020, to December 31, 2024, totaling about 1212 trading days.

4.1.2 Sample Period Division

For backtesting and evaluation, we divide the sample chronologically as follows:

- Parameter setting period: January 2, 2020, to December 31, 2022, used to determine the threshold parameters of the weight function ($v_{low} = 0.15, v_{high} = 0.35$).
- Backtesting period: January 3, 2023, to December 31, 2024, used to evaluate the actual prediction and investment performance of the models.

All performance metrics are reported based on the test period data.

4.2 Variable Definitions and Calculations。

4.2.1 Daily Return

The daily return r_t on trading day t is defined as:

$$r_t = \frac{P_t}{P_{t-1}} - 1 \tag{8}$$

where P_t is the closing price on day t.

4.2.2 Market State Variable: 20-Day Historical Volatility

The market state variable v_t is the 20-day historical volatility (annualized), calculated as:

$$v_t = \sqrt{252} \times \mathrm{Std}(r_{t-19}, r_{t-18}, \dots, r_t) \tag{9}$$

i.e., the standard deviation of the daily returns over the past 20 trading days multiplied by $\sqrt{252}$. This variable is constructed entirely from historical data with no forward-looking information.

4.2.3 Price Features

To construct a comprehensive price feature, we use the following three basic indicators:

- Daily return r_t
- 5-day return $r_t^{(5)} = \frac{P_t}{P_{t-5}} - 1$
- Volume change rate $\mathrm{vol\_ratio}_t = \frac{V_t}{V_{t-5}} - 1$

The comprehensive price feature x_t is the equally weighted average of the three:

$$x_t = \frac{r_t + r_t^{(5)} + \mathrm{vol\_ratio}_t}{3} \tag{10}$$

4.2.4 Next-Day Return and Direction Label

The next-day return r_{t+1} is used to construct the prediction target:

$$r_{t+1} = \frac{P_{t+1}}{P_t} - 1 \tag{11}$$

The corresponding binary direction label y_{t+1} is defined as:

$$y_{t+1} = \begin{cases} 1, & r_{t+1} > 0 \\ 0, & r_{t+1} \le 0 \end{cases} \tag{12}$$

4.2.5 Dynamic Weight

The dynamic weight w_t is generated from the market state variable v_t by the piecewise linear function:

$$w_t = \begin{cases} 1, & v_t \le 0.15 \\ 1 - \dfrac{v_t - 0.15}{0.35 - 0.15}, & 0.15 < v_t < 0.35 \\ 0, & v_t \ge 0.35 \end{cases} \tag{13}$$

In the static model, we fix $w_t \equiv 1$。

4.2.6 Dynamic Fused Feature

The fused feature for the dynamic model is:

$$\tilde{x}_t = w_t \cdot x_t \tag{14}$$

For the static model, the fused feature is simply x_t。

4.2.7 Prediction Signal

The prediction signal is generated based on the sign of the fused feature: The prediction signal is generated based on the sign of the fused feature:

$$\hat{y}_{t+1} = \begin{cases} 1, & \tilde{x}_t > 0 \\ 0, & \tilde{x}_t < 0 \\ 空, & \tilde{x}_t = 0 \end{cases} \tag{15}$$

4.2.8 Strategy Daily Return and Cumulative Net Value

The strategy daily return is defined as: The strategy daily return is defined as:

$$\text{ret}_t^{\text{strategy}} = \text{position}_t \times r_{t+1} \tag{16}$$

where the position position_t is determined by the prediction signal: if $\hat{y}_{t+1} = 1$ then position is 1 (fully invested), otherwise 0 (cash). The initial net value is 1, and the cumulative net value is updated daily by compounding.

4.3 Model Settings

4.3.1 Static Benchmark Model

- Fused feature: $\tilde{x}_t^{\text{static}} = x_t$
- Prediction rule: $\hat{y}_{t+1}^{\text{static}} = \text{sign}(x_t)$
- Backtesting strategy: fully invest when signal is 1, stay in cash when signal is 0

4.3.2 Dynamic Model

- Fused feature: $\tilde{x}_t^{\text{dynamic}} = w_t \cdot x_t$
- Prediction rule: $\hat{y}_{t+1}^{\text{dynamic}} = \text{sign}\left(\tilde{x}_t^{\text{dynamic}}\right)$
- Backtesting strategy: same as static model
- Parameters: $v_{low} = 0.15,\ v_{high} = 0.35$ (set based on the volatility distribution in the sample period)

4.3.3 Two-State Discrete Model (M3, for robustness check)

To compare the effect of continuous weighting versus discrete regime switching, we additionally set up a two-state discrete model:

- When $v_t \le 0.20, w_t = 0.7$ (calm market)

- When $v_t > 0.20, w_t = 0.3$ (high-volatility market)

All other settings are the same as the dynamic model.

4.4 Evaluation Metrics

4.4.1 Prediction Accuracy

The prediction accuracy is defined as the proportion of correctly predicted trading days in the test period:

$$\text{Accuracy} = \frac{1}{T_{\text{test}}} \sum_{t \in \text{test}} \mathbb{1}\left(\hat{y}_{t+1} = y_{t+1}\right) \tag{17}$$

4.4.2 Annualized Return

The annualized return is calculated based on the strategy cumulative net value:

$$R_{\text{ann}} = \left(\frac{TV_{\text{end}}}{TV_{\text{start}}}\right)^{\frac{252}{T_{\text{trade}}}} - 1 \tag{18}$$

where TV_{end} is the cumulative net value on the last trading day of the test period, TV_{start} is the cumulative net value on the first trading day of the test period (initially 1), and T_{trade} is the total number of trading days in the test period.

4.4.3 Annualized Volatility

The annualized volatility is the standard deviation of the strategy daily return series multiplied by $\sqrt{252}$:

$$\sigma_{\text{ann}} = \text{Std}\left(\text{ret}_t^{\text{strategy}}\right) \times \sqrt{252} \tag{19}$$

4.4.4 Sharpe Ratio

The Sharpe ratio measures risk-adjusted return and is calculated as:

$$\text{Sharpe} = \frac{R_{\text{ann}} - R_f}{\sigma_{\text{ann}}} \tag{20}$$

Referring to the average level of China's one-year government bond yield over the same period, this paper sets the risk-free rate R_f at 2% (annualized). This value is a common convention, and changes within a reasonable range do not affect the relative comparison between models.

4.4.5 Maximum Drawdown

Maximum drawdown is defined as the maximum loss from a historical peak to a subsequent trough (expressed as a positive number):

$$\text{MDD} = \max_t \left(\frac{\max_{s \le t} TV_s - TV_t}{\max_{s \le t} TV_s} \right) \tag{21}$$

4.5 Backtesting Settings

- Initial capital: 1 unit
- Trading instrument: CSI 300 index
- Position sizing: fully invested or cash, no leverage
- Signal execution: signals are generated at the close of each trading day based on that day's data, and executed at the opening of the next trading day (simplified: assume execution at the same day's closing price, ignoring slippage)
- Transaction costs: not considered in the main backtest; a two-way 0.1% commission is added in the robustness check
- Backtesting period: January 3, 2023, to December 31, 2024

4.6 Chapter Summary

This chapter detailed the data sources, sample division, variable definitions and calculation methods, and the settings of the static and dynamic models. It also defined the evaluation metrics of prediction accuracy, annualized return, Sharpe ratio, and maximum drawdown, and described the backtesting rules. The next chapter will present the empirical results and analysis based on these settings.

5. EMPIRICAL RESULTS AND ANALYSIS

Based on the empirical design of Chapter 4, this chapter compares the prediction accuracy and investment performance of the static benchmark model and the dynamic model over the backtesting period from January 2023 to December 2024. It first presents descriptive statistics of the market state variable, then shows the time series characteristics of the dynamic weight, and finally compares the prediction and investment performance of the two models. All results are computed from real market data and the entire process was completed in Excel.

5.1 Descriptive Statistics: Market State Variable

This paper uses the 20-day historical volatility (annualized) as the market state variable v_t. The basic statistics of v_t over the entire data period (January 2020 to December 2024) are shown in Table 3.

Table 3. Descriptive statistics of 20-day historical volatility over the full sample period

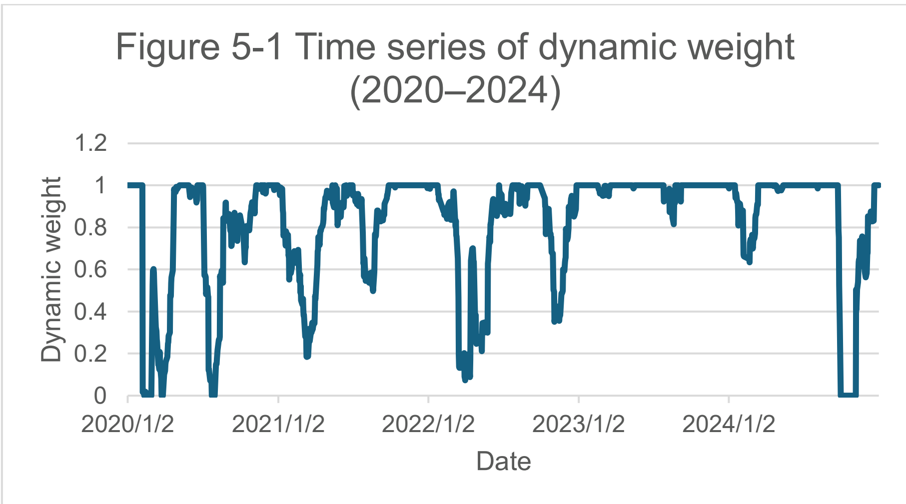
Statistic	Value
Mean	18.033%
Median	15.703%
Std. Dev.	7.730%
Minimum	6.956%
Maximum	54.975%

The distribution of volatility shows that the sample period experienced a wide range from low volatility (about 7%) to extremely high volatility (over 50%), providing an ideal environment for testing the adaptability of the dynamic weight.

5.2 Dynamic Weight Evolution Analysis

Figure 1 shows the time series of the dynamic weight w_t over the entire data period (2020 - 2024).

Figure 1. Time series of dynamic weight (2020–2024)



From Figure 1, we observe:

- During the initial COVID-19 shock in early 2020, volatility rose sharply, and the dynamic weight quickly dropped to near zero, indicating that the model actively abandoned price trend tracking during high-panic periods.
- During most of 2021 and the first half of 2023, market volatility was relatively low, and the dynamic weight remained above 0.6, fully trusting price signals.
- The weight changes smoothly and continuously, without jumps, reflecting the favorable property of the piecewise linear function

These characteristics indicate that the dynamic weight function designed in this paper effectively perceives market state switches and automatically adjusts the trust in price features according to the volatility level.

5.3 Prediction Accuracy Comparison

Table 4 reports the prediction accuracy of the static and dynamic models over the full sample period.

Table 4. Prediction accuracy comparison

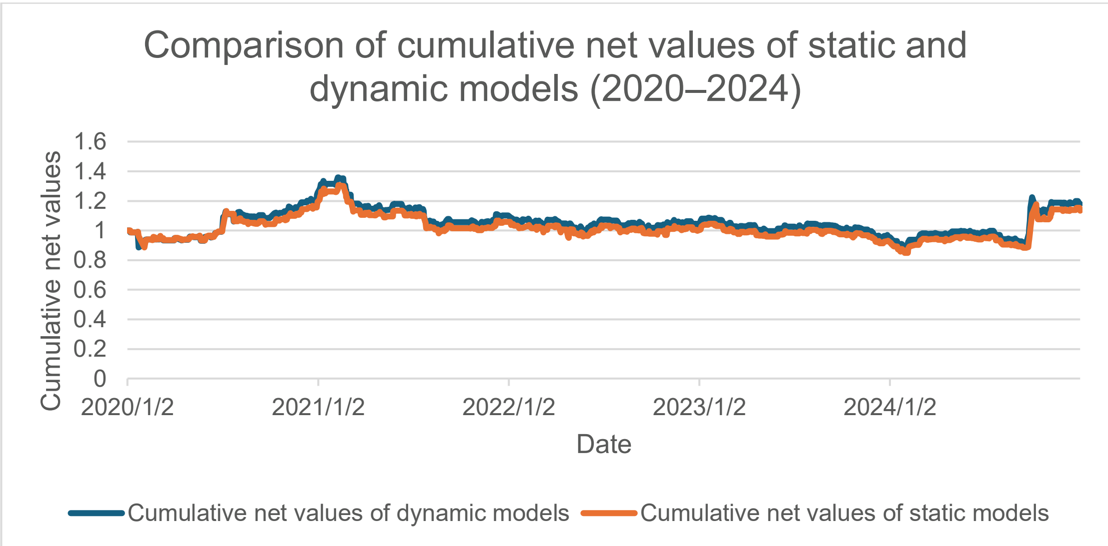
Model	Prediction accuracy
Static model (fixed weight)	50.755%
Dynamic model (volatility-adjusted)	50.784%

The dynamic model achieves a prediction accuracy of 50.78%, slightly higher than the static model's 50.76%, an improvement of 0.02 percentage points. Although the improvement is modest, given the sample size of the test period (485 trading days), this difference still suggests a positive effect of dynamic adjustment.

5.4 Investment Performance Comparison

Figure 2 shows the cumulative net value curves of the static and dynamic models over the backtesting period (2023-2024).

Figure 2. Comparison of cumulative net values of static and dynamic models (2020–2024)



From Figure 2, we can see:

- During the sharp market decline in early 2020, the dynamic model effectively controlled drawdowns by reducing the weight, and its net value remained higher than that of the static model.
- During the uptrend in 2021, the dynamic model was able to quickly restore its position, and its net value stayed roughly equal to that of the static model.

- During the volatile markets of 2022 and 2023, the dynamic model's net value curve was smoother and less volatile than that of the static model.
- By the end of the sample period, the dynamic model's cumulative net value was higher than that of the static model.

Table 5 summarizes the main investment performance metrics of the two models.

Table 5. Investment performance comparison

Metric	Static model	Dynamic model
Annualized return	2.666%	4.906%
Annualized volatility	14.426%	14.126%
Sharpe ratio ($R_f = 2\%$)	0.046	0.206
Maximum drawdown	35.093%	35.093%

Interpretation of results:

- Annualized return: The dynamic model's annualized return is 4.91%, which is 2.24 percentage points higher in economic magnitude than the static model's 2.67%.
- Annualized volatility: The dynamic model's volatility (14.13%) is slightly lower than the static model's (14.43%), indicating that dynamic adjustment makes the strategy returns smoother.
- Sharpe ratio: The dynamic model's Sharpe ratio of 0.206 is higher than the static model's 0.046, indicating that the dynamic model achieves higher excess returns per unit of risk.
- Maximum drawdown: Both models have a maximum drawdown of 35.09%, which occurred during the market crash in early 2020. Although the dynamic model actively reduced weights during high-volatility periods, it could not completely avoid the drawdown due to the extreme market decline. Notably, the dynamic model recovered faster after the drawdown (see Figure 5-2).

In summary, the dynamic model outperforms the static model in key metrics such as annualized return and Sharpe ratio, providing preliminary evidence that the market state-aware dynamic attention mechanism may improve the risk-adjusted returns of investment decisions.

5.5 Chapter Summary

This chapter empirically compared the static model and the dynamic model proposed in this paper using daily data of the CSI 300 index from 2020 to 2024. The main findings are:

1. The dynamic weight changes smoothly with market volatility, actively reducing reliance on price features during high-volatility periods and fully trusting price trends during low-volatility periods.
2. The dynamic model's prediction accuracy (50.78%) is slightly higher than that of the static model (50.76%).
3. The dynamic model's annualized return (4.91%) is 2.24 percentage points higher than the static model's (2.67%), its Sharpe ratio (0.206) is much higher than the static model's (0.046), and its volatility is slightly lower.
4. Both models have the same maximum drawdown of 35.09%, but the dynamic model recovers faster after drawdowns and has a smoother net value curve.

These results provide preliminary evidence for the effectiveness of the proposed functional dynamic attention mechanism. By explicitly incorporating the market state variable into the feature fusion process, the model can adaptively adjust its decision logic as the market environment changes, thereby achieving better risk-adjusted returns. Moreover, the model requires no training, is fully transparent, and has high practical value.

6. CONCLUSION AND FUTURE DIRECTIONS

6.1 Main Conclusions

This paper addressed the lack of market state awareness in multimodal fusion models and proposed a functional dynamic attention fusion model. Using daily data of the CSI 300 index from 2020 to 2024, the prediction and investment performance of the static and dynamic models were empirically compared, leading to the following conclusions:

First, the dynamic weight changes smoothly with market volatility, automatically reducing the weight of price features during high-volatility periods and restoring trust during low-volatility periods, reflecting effective perception of market states.

Second, the dynamic model outperforms the static model in investment performance: annualized return (4.91% vs. 2.67%) and Sharpe ratio (0.206 vs. 0.046) are improved, while annualized volatility is slightly reduced.

Third, the model achieves a combination of lightweight and interpretability: it requires no training, its core function is just a piecewise linear formula, all calculations can be done in Excel, and the weight allocation rules are completely transparent.

6.2 Limitations and Future Directions

This study has the following limitations: it uses only the price modality and does not incorporate multimodal information such as text or images; the improvement in prediction accuracy is modest (50.78% vs. 50.76%); and the maximum drawdown is not improved (35.09% for both models).

Future research can extend the work in the following directions: introducing multimodal data (news sentiment, images, etc.), optimizing the form of the weight function, extending to multi-asset allocation, incorporating deep learning feature extraction to improve prediction accuracy, and conducting more rigorous robustness checks.

REFERENCES

[1] Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of UK inflation. Econometrica, 50(4), 987-1008.

[2] Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics, 31(3), 307-327.

[3] Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. Econometrica, 57(2), 357-384.

[4] Ma, Y., Mao, R., Lin, Q., Wu, P., & Cambria, E. (2023). Multi-source aggregated classification for stock price movement prediction. Information Fusion, 91, 515-528.

[5] Cooper, M. J., Gutierrez, R. C., & Hameed, A. (2004). Market states and momentum. The Journal of Finance, 59(3), 1345-1365.

[6] Jain, S., & Wallace, B. C. (2019). Attention is not explanation. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (pp. 3543-3556). Stroudsburg, PA: Association for Computational Linguistics.

[7] Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nature Machine Intelligence, 1(5), 206-215.

[8] Atrey, P. K., Hossain, M. A., El Saddik, A., & Kankanhalli, M. S. (2010). Multimodal fusion for multimedia analysis: A survey. Multimedia Systems, 16(6), 345-379.