ENGINEERING Management 2026, 13(1): 146–163
https://doi.org/10.1007/s42524-026-5240-1

RESEARCH ARTICLE

Wenya WANG, Hongyu WU, Jiali ZHAO, Jian SU

Incorporating security risks into natural gas forecasting: An interpretable, policy-oriented machine learning framework

Abstract Most existing predictive models remain demand-centric and fail to systematically incorporate supply-side risks such as import dependence, price volatility, and market concentration. Thus, this study proposes a SHAP-driven Weighted Rule Attention Mechanism (SWRAM), and explicitly embeds energy security risk factors including import dependence, market concentration, and price volatility into natural gas consumption forecasting. The model integrates explainable machine learning with a rule-constrained attention mechanism to enable both transparent feature attribution and robust predictive performance, and, when compared with the benchmark models, it demonstrates lower forecasting errors and more stable generalization, thereby validating the effectiveness of embedding SHAP-based rule-constrained weights within the attention mechanism. Using monthly data for China from 2012 to 2024, the results show that supply capacity and infrastructure remain the dominant drivers of natural gas consumption, while risk-related factors have gained importance since 2020, reflecting the impact of global supply-chain instability. Interaction analysis reveals a strong nonlinear coupling between domestic production and import dependence, indicating that insufficient domestic output amplifies exposure to external risks. Price volatility exerts an increasingly significant effect during global energy shocks, especially between 2021 and 2023. These findings suggest that natural gas consumption is shaped jointly by short-term demand cycles and long-term structural dependencies. The SWRAM framework provides an interpretable and policy-oriented tool for improving forecasting reliability and supporting data-driven energy security governance.

Keywords natural gas consumption, risk factors, attention mechanism, SHAP values, machine learning

Received Aug. 20, 2025; revised Dec. 11, 2025; accepted Dec. 25, 2025

Wenya WANG, Hongyu WU, Jiali ZHAO (✉)
School of Economics and Management & Institute for Energy Economics and Policy & Institute for Digital Transformation, China University of Petroleum, Qingdao 266580, China
E-mail: b23080018@s.upc.edu.cn

Jian SU
Research Institute of Petroleum Exploration and Development, Beijing 100083, China

This work was supported by the National Natural Science Foundation of China (Grant No. 72203235), and the Natural Science Foundation of Shandong, China (Grant No. ZR2022QG006).

1 Introduction

Among all energy sources, natural gas has gained growing significance as a clean and efficient fossil fuel (Agrell et al., 2025; Qiao et al., 2020; Xu and Lin, 2019). This is particularly evident in China, where natural gas consumption has risen rapidly in recent years. However, as a major natural gas importer, China faces mounting exposure to supply-side risks intensified by global instability, including cross-border disruptions, price volatility, and geopolitical fragility (Chai et al., 2018; Scharf and Möst, 2024; Wang and Lin, 2014). When consumers perceive future supply as unstable, they may adjust their consumption behavior, such as reducing demand, deferring procurement, or overstocking, which in turn undermines the reliability of traditional demand forecasts. For example, in 2021, global supply constraints and rising international demand caused liquefied natural gas (LNG) prices to surge. In China, LNG import prices increased by more than 50%, forcing price-sensitive industries and households to adopt cautious consumption strategies (Li et al., 2025). Similarly, after the outbreak of the Russia–Ukraine conflict in 2022, Western sanctions against Russia triggered gas supply shortages in Europe and produced spillover effects across global natural gas markets (Di Bella et al., 2024; Goodell et al., 2023; Sassi, 2025). Because some of China's imports arrive through Russian pipelines, domestic firms diversified supply sources and expanded strategic reserves, increasing short-term market uncertainty and price volatility. These events demonstrate the urgency of embedding structural supply risks into demand forecasting models to improve accuracy and inform policy under uncertainty.

Conventional natural gas forecasting models primarily focus on macroeconomic and demand-side factors, such as income growth, temperature, industrial activity, and energy prices. While these traditional models perform well under relatively stable conditions, they often neglect structural energy security risks, including import dependence, market concentration, infrastructure bottlenecks, and price volatility. These risks have become increasingly significant following the COVID-19 pandemic, as global supply-chain instability and geopolitical tensions have fundamentally altered the energy market (Dai et al., 2022; Demski et al., 2018; Kim et al., 2025; Li et al., 2023). Existing forecasting methods struggle to trade off predictive accuracy and interpretability. Econometric models are interpretable but limited in capturing the nonlinear effects of complex supply-side shocks, whereas machine learning models offer higher predictive accuracy but limited interpretability regarding how risk factors contribute and interact. Consequently, few tools can both quantify the impact of specific energy security risks on consumption and remain reliable in volatile market environments. A framework that embeds risk indicators into the forecasting process while providing interpretable measures of feature importance and interactions is therefore needed.

To address these gaps, this study develops a newly developed interpretable machine learning framework called the SHAP-driven Weight-based Rule Attention Mechanism (SWRAM), which systematically incorporates energy security-related risk factors into natural gas consumption forecasting. The framework integrates SHAP-based feature interpretability with a rule-based attention mechanism that assigns transparent, risk-aware weights to key variables. It allows for the simultaneous modeling and interpretation of supply capacity, external risks, and market volatility within a unified structure. First, SHAP-driven feature screening is employed to identify key predictors and assess the importance of risk-related variables. Then, SWRAM is compared with two benchmark models: SHAP-driven Feature Screening without Attention Mechanism (SHAP-noAtt) and SHAP-driven Feature Screening with Adaptive Attention Mechanism (SHAP-AdaptAtt). The results demonstrate that SWRAM achieves improved forecasting accuracy while maintaining high interpretability.

This study makes three key contributions.

1) Although previous studies have examined the factors influencing natural gas consumption, few have considered the impact of security-related risks, such as price volatility, import dependency, and transportation risks, on consumption forecasting. By embedding these risk dimensions, this paper identifies their key role in influencing demand uncertainty and enhances forecast accuracy under volatile market conditions.

2) We propose the SWRAM framework, which integrates SHAP value analysis with a rule-based attention structure. Unlike traditional attention or SHAP-augmented models, SWRAM introduces a rule-constrained weighting layer that incorporates domain-specific risk knowledge into the training process. By aligning model weights with quantifiable economic mechanisms, SWRAM achieves an effective trade-off between forecasting accuracy and structural interpretability.

3) Through SHAP-based feature importance ranking, analysis of feature interactions, and analysis of temporal variations, this study extends beyond conventional descriptive forecasting by providing interpretable insights into feature interactions in China's natural gas market. By clarifying these relationships, the study provides a novel approach to analyzing market dynamics and uncovers how supply-side vulnerabilities influence demand fluctuations, thereby improving the interpretability of the results.

This paper is structured as follows. Section 2 reviews the relevant literature. Section 3 outlines the proposed methodological framework. Section 4 presents the identification of key predictors and compares the proposed forecasting model against benchmark models. Section 5 provides a detailed analysis of variable interactions, with a focus on seasonal dynamics. Finally, Section 6 concludes the paper and discusses the main results.

2 Literature review

2.1 Drivers of natural gas consumption

Numerous studies have investigated the determinants of natural gas consumption from economic, demographic, and energy system perspectives. Early research mainly emphasized socioeconomic indicators such as economic growth, urbanization level, energy intensity, and industrial structure. For instance, Jiang et al. (2020) applied econometric models to provincial-level data and identified economic and structural indicators as major drivers. Chai et al. (2018) applied STIRPAT and LMDI models and found that urbanization, GDP per capita, and industrial energy intensity significantly affect consumption. At the national and cross-country levels, studies such as Azlina et al. (2014) and Aydin (2018) confirmed a unidirectional causal relationship between economic development and natural gas use.

Beyond economic factors, several studies have highlighted infrastructure and regulatory conditions, including population coverage, pipeline length, and environmental regulation, as essential drivers (Wang et al., 2023; Zhao et al., 2020). Comparable findings are observed in international contexts such as Europe and North America, where infrastructure connectivity, market integration, and diversification policies play a crucial role in stabilizing gas demand under external shocks. In Europe, diversified import routes and regional pipeline interconnections have strengthened consumption resilience to geopolitical disruptions, while in North America, the integration of the US, Canadian, and Mexican gas markets, facilitated by extensive pipeline networks and regulatory coordination, has enhanced cross-border trade flexibility and system efficiency amid price and policy fluctuations (Avraam et al., 2021; Avraam et al., 2020; Bilgili, 2025; Chen et al., 2025).

However, most existing studies have paid limited attention to supply-side security risks, such as import dependence, price volatility, and market concentration, which increasingly shape natural gas consumption dynamics under global energy uncertainty.

2.2 Feature selection and forecasting methods in energy consumption

To capture the complex, nonlinear nature of natural gas consumption, recent research has adopted feature selection and machine learning methods. Traditional linear models often fail to address high-dimensional interactions among variables (Liang et al., 2019). Feature selection methods, such as filter, wrapper, embedded, and dimensionality reduction techniques, have been widely applied to extract the most relevant predictors (Bhagat et al., 2021).

In recent years, hybrid machine learning frameworks integrating optimization algorithms with deep learning have achieved notable success. Examples include random forest combined with STAM-LSTM (Cao et al., 2024a; Wang et al., 2024b), MOPSO-SVR (Karasu et al., 2020), and GCMC-LSTM, all of which significantly improve forecasting accuracy. Nonetheless, most of these studies focus primarily on computational performance, with limited attention to how external risk variables, such as geopolitical events, import shocks, or price instability, can be systematically incorporated into forecasting models. Furthermore, few works explore how the inclusion of these risk dimensions affects model interpretability and feature weighting mechanisms, resulting in a persistent gap between technical modeling and policy-relevant understanding.

2.3 SHAP and attention mechanisms for interpretable energy risk assessment

Research in energy policy and risk management faces several challenges. Many studies focus on descriptive correlations or macro-level determinants (e.g., temperature, industrial activity) without explicitly linking supply-side risk factors, such as import dependence, market concentration, or price volatility, to actionable policy interventions (Benigno et al., 2025; Boţa-Avram et al., 2024; Eskandari et al., 2024). In addition, the static nature of feature importance in conventional SHAP-based models prevents us from understanding of how risks propagate over time, making it difficult to design dynamic policy responses or risk mitigation strategies. High-dimensional and temporally varying interactions between risks and demand are rarely captured, limiting the translation of model insights into practical decisions such as adjusting strategic reserves or diversifying supply sources.

The development of SHAP (Shapley Additive Explanations) has greatly enhanced interpretability in energy forecasting. SHAP enables both global and local feature importance analysis and is widely applied in studies involving gas consumption. However, SHAP-based models typically treat feature importance statically, without embedding it into the model structure for dynamic prediction. The Attention Mechanism has shown great potential for modeling temporal dependencies and cross-variable relationships in domains such as natural language processing, computer vision, and time series forecasting (Kong et al., 2023; Zang et al., 2021). Its adaptability enables models to focus on the most informative input features at each timestep. Despite its promise, few studies have attempted to link SHAP interpretability with attention structures, especially in the context of energy security risk modeling, limiting the ability to capture how structural and external risks dynamically influence consumption patterns.

To address these gaps, this study proposes the SWRAM framework, which embeds energy security risk factors directly into the forecasting process and dynamically adjusts attention weights according to SHAP-derived importance. By integrating interpretable attributions with a rule-based attention mechanism, SWRAM provides both accurate predictions and clear insights into how supply-side vulnerabilities influence demand fluctuations, thereby offering a more actionable basis for energy policy design and risk management.

3 Research framework and methodology

The literature on energy demand recurrently identifies supply stability, import exposure, market volatility, infrastructure development, and policy conditions as major dimensions shaping natural gas consumption(Peng et al., 2021; Wejnert-Depue et al., 2025; Zhang et al., 2021). Guided by this established perspective, we extend the predictor set by incorporating risk-related indicators that capture the growing influence of market disruptions and policy uncertainty. These dimensions were then translated into measurable indicators and used to construct SHAP-derived priors that inform a rule-based attention structure. Building on this expanded variable system, a multi-factor temporal modeling strategy is implemented, and four mainstream machine learning models (SVR, RF, LSTM, and XGBoost) are employed as baseline models to ensure robustness and comparability. The proposed SHAP-driven Weighted Rule Attention Mechanism (SWRAM) integrates feature attributions with rule-constrained attention weights, enabling dynamic factor weighting while preserving interpretability. The following sections detail the underlying rationale, the mapping of these dimensions to observable indicators, and the methodological design of the SWRAM model.

3.1 Categorization of predictors

This study classifies the predictor set for natural gas demand and market dynamics into five groups: supply-related factors, risk-related indicators, market and price conditions, demand-side fundamentals, and policy variables, following commonly adopted dimensions in the literature (Ata et al., 2025; Cao et al., 2024b; Li et al., 2024; Wang et al., 2024a; Wang et al., 2025; Yang et al., 2025).Within the risk-related group, and following previous research, the indicators are further refined into three categories to capture different sources of potential disturbance: supply-side and infrastructure vulnerabilities, external exposure through import dependence and geopolitical conditions, and market-price disturbances linked to international benchmark fluctuations. These indicators reflect the main channels through which disruptions propagate to domestic consumption. Policy variables operate across these groups by shaping system resilience, moderating exposure, and influencing the transmission of external shocks. A total of 18 variables are selected according to their functional roles, and these groups are incorporated into the SWRAM model to embed domain knowledge in a transparent and interpretable manner, as illustrated in Fig. 1.

This framework provides a coherent analytical logic that connects structural constraints with market-based transmission. When domestic supply capacity and infrastructure are limited, external risks and international price fluctuations exert stronger and more volatile impacts on domestic consumption. In contrast, higher production capacity, diversified supply portfolios, and effective fiscal or regulatory interventions weaken such transmission pathways. This conceptual structure offers a clear theoretical foundation for integrating domain-specific risk knowledge into the SWRAM model, allowing it to capture how structural vulnerabilities and price transmission jointly influence natural gas consumption dynamics. Within this logic, the use of SHAP-based feature contributions within the attention structure is theoretically justified because SHAP values quantify both the direction and the magnitude of factor influences in a manner consistent

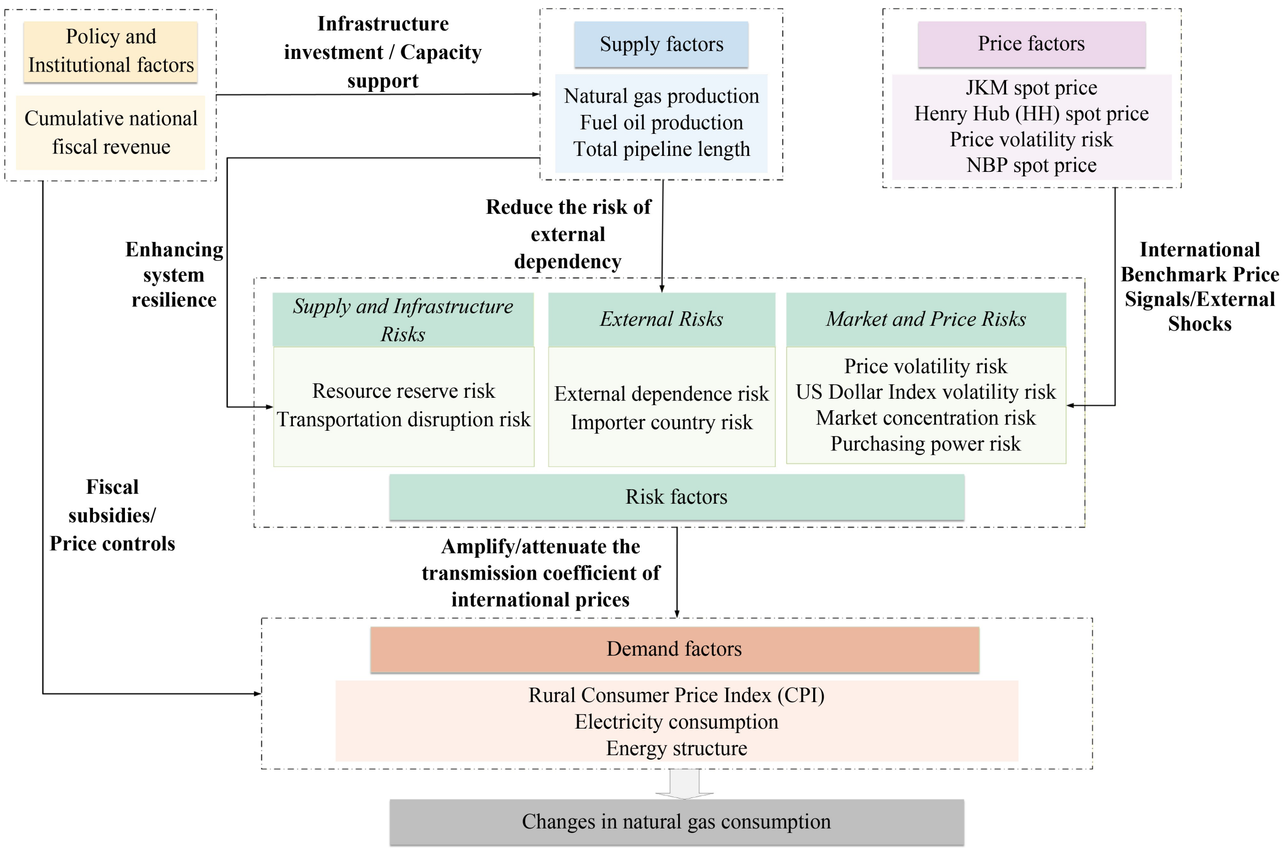

Fig. 1 Categorization of predictors.

with the causal relationships emphasized in the energy-security literature. Aligning these mechanisms with the model design enhances the theoretical grounding of the study and ensures continuity between conceptual reasoning and empirical implementation.

3.2 Methodological framework based on multi-dimensional energy risks

Because energy-security risks operate through nonlinear and interacting channels, a forecasting framework must account for both the marginal influence of individual factors and the interaction structures that shape their combined effects. SHAP offers an interpretable way to quantify these contributions, while attention mechanisms provide a flexible means of assigning dynamic importance under time-varying risk conditions. Building upon the energy-security framework introduced in Section 3.1, this section develops an interpretable forecasting system for natural gas consumption that reflects multidimensional risk transmission. The framework incorporates supply stability, import dependence, market volatility, infrastructure resilience, and policy support into the predictive structure to represent their joint impact on consumption dynamics. To operationalize this design, a baseline predictive model is augmented with a SHAP-based interpretability layer that measures the contribution of each risk dimension. This layer is further integrated into a weighted attention structure to align data-driven learning with factor-informed constraints. This methodological formulation enhances both predictive accuracy and theoretical transparency, enabling the model to reflect structural risk transmission while maintaining interpretability.

3.2.1 SHAP-based interpretability mechanism

SHAP (Shapley Additive Explanations) provides a rigorous approach to interpreting machine learning predictions. The method represents the model output as the sum of additive contributions from individual features, based on cooperative game theory. In this study, SHAP is used to quantify the marginal effect of every explanatory variable on predicted natural gas consumption. At the global level, SHAP ranks variables such as supply capacity, import dependence, and price volatility by their average contributions, thereby describing their system-wide influence. At the local level, SHAP attributes a specific share of each prediction to each variable, which allows case-by-case interpretation and makes the inference process transparent. The SHAP value for a feature is defined as the weighted average of its marginal contribution across all possible coalitions of features. This definition follows the Shapley value in cooperative game theory and satisfies key properties of local accuracy, consistency, and missingness, which ensures fair and model-agnostic attribution (Demski et al., 2018). The formal expression of the Shapley value used in this paper is given in Eq. (1).

$$\phi_i = \sum_{S\subseteq F\setminus\{i\}} \frac{|S|!(|F|-|S|-1)!}{|F|!}(f(S\cup\{i\})-f(S)). \quad (1)$$

In the above formula, ϕ_i represents the SHAP value of the i-th EMD-reconstructed component, F denotes the set of all components, S is a subset of components excluding the i-th component, and $f(S)$ is the model output given the subset S. The formula computes the SHAP value as a weighted average of the marginal contributions of each component across all possible combinations of components. This approach provides a theoretically grounded and consistent measure of each component's contribution to the model output.

In addition to individual feature contributions, SHAP also introduces SHAP interaction values to capture pairwise interaction effects between components (Liu et al., 2025). These interaction values extend the explanatory power of SHAP by quantifying not only the standalone importance of each component but also the combined effect of interactions. The SHAP interaction value is defined as follows:

$$\phi_{i_1,i_2} = \sum_{S\subseteq F\setminus\{i_1,i_2\}} \frac{|S|!(|F|-|S|-2)!}{2(|F|-1)!} [(f(S\cup\{i_1,i_2\})-f(S\cup\{i_2\})-(f(S\cup\{i_1\})-f(S))]. \quad (2)$$

3.2.2 SHAP-driven weighted rule attention mechanism (SWRAM)

The attention mechanism, originally inspired by cognitive neuroscience, has become a widely adopted approach in artificial intelligence, particularly in Natural Language Processing (NLP) and computer vision(Li et al., 2021; Zhong et al., 2024). Its core idea is to simulate how humans selectively focus on the most relevant information by assigning different weights to input features. This enables models to handle complex dependencies more effectively and to enhance both forecasting accuracy and interpretability.

However, in most energy forecasting applications, conventional attention mechanisms are implemented as black-box modules, limiting interpretability and transparency. To address this issue, this study introduces an interpretable dual-attention framework called the SHAP-driven Weighted Rule Attention Mechanism (SWRAM), which integrates adaptive attention and rule-based static attention to jointly capture data-driven feature relevance and domain-informed risk weighting.

The adaptive attention layer dynamically learns the importance of input features based on data variations. The attention weights are computed through the query–key–value mapping, as summarized in Eq. (3):

$$Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V. \tag{3}$$

The query matrix Q, the key matrix K, and the value matrix V were generated through linear transformations. The detailed mathematical formulation of the attention computation, including normalization and projection procedures, is provided in Appendix A2 for reference. This mechanism allows the model to identify which features should receive greater emphasis when forecasting natural gas consumption under varying conditions.

Four widely used machine learning algorithms, including Support Vector Regression (SVR), Random Forest (RF), Long Short-Term Memory (LSTM), and XGBoost, are employed as baseline models for comparative evaluation. These models represent different algorithmic paradigms: kernel-based (SVR), ensemble-based (RF), recurrent (LSTM), and boosting (XGBoost). Their standard mathematical forms are summarized in Appendix A1.

This study proposes a novel interpretable modeling framework called the SHAP-driven Weighted Rule Attention Mechanism (SWRAM) to enhance both predictive performance and interpretability in natural gas consumption forecasting under energy security risks. The SWRAM model integrates adaptive attention with a rule-constrained static attention component, forming a dynamic-static interpretability embedding structure. In this structure, the static component embeds domain knowledge through SHAP-derived feature importances, while the adaptive component fine-tunes weight interactions dynamically during training. This hybrid design explicitly links each attention weight to the marginal contribution of its corresponding feature, thereby improving model transparency, convergence stability, and interpretability.

Conventional attention networks rely solely on data-driven weight learning. In contrast, SWRAM incorporates explainable priors from SHAP analysis to constrain the learning process, thereby mitigating the instability of attention weights often observed in high-dimensional, non-stationary energy markets. The model embeds these priors into the attention structure, keeping its learning dynamics consistent with theoretically grounded risk pathways. As a result, feature weights reflect both empirical evidence and mechanisms informed by key risk factors.

The normalized SHAP-based static weight of each feature i is defined as:

$$\omega_i^{\text{SHAP}} = \frac{|\phi_i|}{\sum_{j=1}^{n} |\phi_i|}. \tag{4}$$

These weights are incorporated into the attention layer by forming a diagonal projection matrix ω_s, which interacts with the adaptive attention representation ω_a to yield a hybrid feature representation:

$$H = \alpha W_a X + (1 - \alpha) W_s X, \tag{5}$$

where X denotes the input feature matrix, and $\alpha \in [0, 1]$ controls the trade-off between adaptability and interpretability. This formulation ensures that the model's attention distribution is guided jointly by data-driven learning and domain-informed prior knowledge, effectively reducing overfitting and improving interpretive reliability in high-dimensional forecasting tasks.

The SWRAM architecture retains the nonlinear adaptability of traditional attention mechanisms while strengthening the model's focus on key determinants such as price volatility, import dependence, and supply-side risk (see Fig. 2 for the algorithmic structure). By embedding SHAP-based priors, the model mitigates instability in weight optimization and achieves a balanced trade-off between interpretability, computational efficiency, and forecasting accuracy.

To ensure methodological rigor and temporal consistency, we adopted a rolling-origin one-step-ahead forecasting framework using monthly data from January 2012 to August 2024. At each iteration, the model is trained on all available observations up to month t and then used to predict consumption for month $t+1$, generating a continuous sequence of 26 monthly forecasts from July 2022 to August 2024. To examine the impact of forecast horizon on model performance, we conducted an additional experiment evaluating RMSE, MAE, and R^2 across multiple horizons. The results, reported in Appendix C Fig. C1, indicate that the 26-month horizon provides the most favorable balance between prediction accuracy and stability, thereby supporting its selection in the main analysis. Placing this figure in the appendix allows the main text to remain concise while providing readers with detailed empirical evidence on horizon sensitivity.

During each rolling iteration, the latest 12 months of data are set aside for model validation. Hyperparameters are tuned through time-series cross-validation using the TimeSeriesSplit technique, maintaining temporal sequence and avoiding information leakage from forthcoming data. Given the limited number of monthly samples, all models underwent a comprehensive grid search to ensure convergence stability and robust generalization under repeated rolling training. The final optimal hyperparameter configurations for each prediction model are presented in Table 1, which provides a transparent basis for reproducibility and subsequent model comparisons.

4 Identifying policy-relevant drivers and evaluating predictive modeling strategies for natural gas consumption

4.1 Data processing

Building upon the five-dimensional energy security

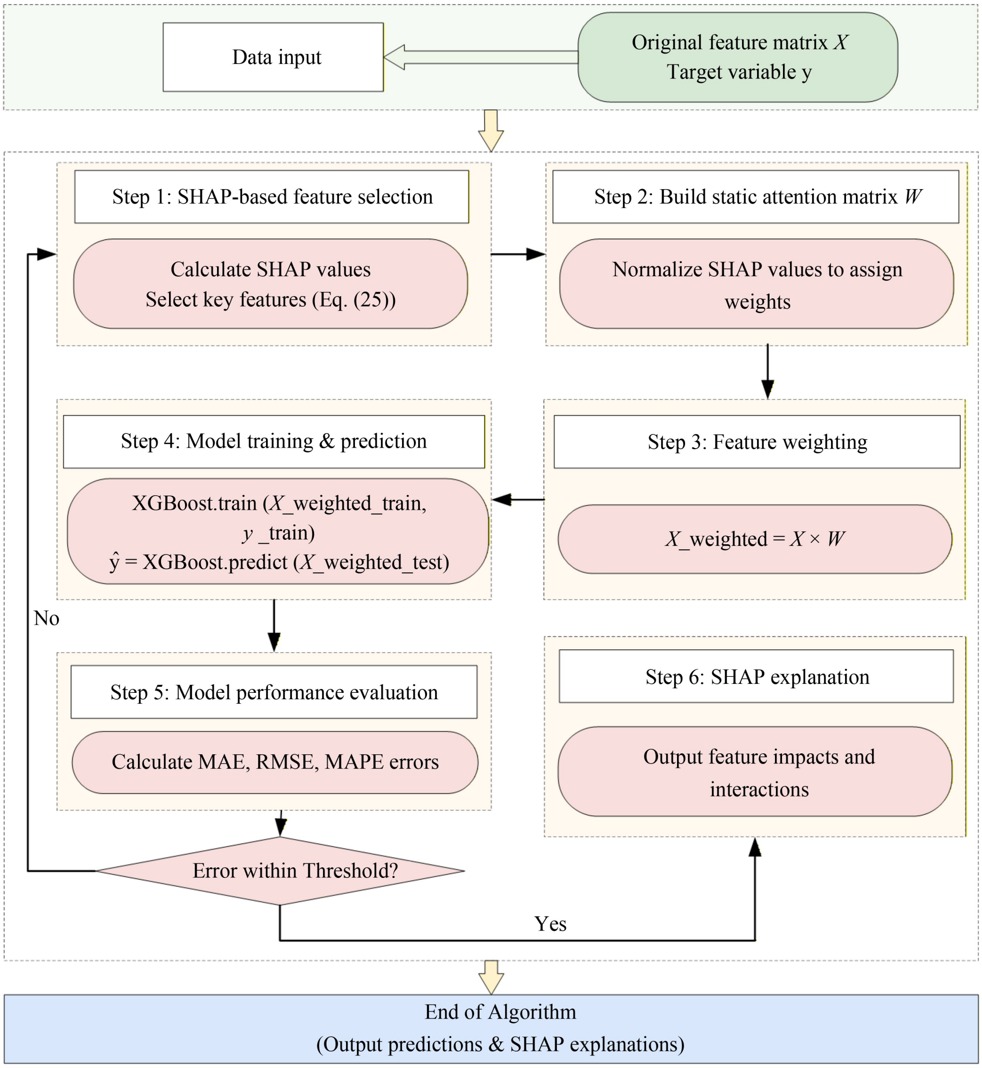

Fig. 2 Flowchart of the SHAP-driven Weight-based Rule Attention Mechanism (SWRAM) Algorithm.

framework established in Section 3.1, this study identifies 18 indicators (denoted as $x_i(i = 1, ..., 18)$) corresponding to the supply, risk, price, demand, and policy dimensions of natural gas consumption (see Fig. 1). This design ensures conceptual alignment between the theoretical framework and the empirical model, linking each indicator to a specific domain of energy security: availability, resilience, affordability, sustainability, and governance.

The data set spans January 2012 to August 2024 and includes variables related to production conditions, market and price dynamics, demand fundamentals, infrastructure development, and geopolitical and institutional risks. Specifically, the indicators comprise: natural gas production, fuel oil production, rural CPI (year-on-year), cumulative national fiscal revenue, electricity consumption, total length of natural gas pipelines, energy structure, JKM, Henry Hub (HH), and NBP spot prices, as well as eight risk indices capturing external dependence, resource reserve uncertainty, price volatility, US Dollar Index volatility, market concentration, importer country risk, transportation disruption risk, and purchasing power vulnerability.

Data were sourced from the National Bureau of Statistics of China, General Administration of Customs, UN Comtrade, BP Statistical Review of World Energy, Wind Database, US EIA, ICRG, and multiple industry and security risk reports (IMB, EGIG). Among these variables, some are reported annually (e.g., pipeline length, market concentration risk), others monthly (e.g., production and consumption indicators), while international benchmark prices are recorded daily. To unify frequencies for monthly forecasting, annual data were interpolated to monthly values, and daily prices were converted to monthly averages. Risk indices were constructed from the corresponding monthly data.

All temporal transformations were conducted in a forward-looking manner consistent with the forecasting task. Missing values were handled using forward-fill or

Table 1 Optimal hyperparameter settings for each prediction model

Model	Hyper parameterization	SWRAM	SHAP-AdaptAtt	SHAP-noAtt
SVR	penalty parameter (C)	10	100	100
	Kernel parameter (γ)	0.1	1.33	0.1
	ε-insensitive zone (ε)	0.001	0.01	0.01
RF	N_estimators	100	100	150
	Max_depth	10	5	10
	Minimum samples per leaf	1	1	1
XGBoost	Tree number	200	100	100
	Learning rate (eta)	0.1	0.1	0.1
	Maximum depth	5	3	3
	Subsample	0.8	0.8	0.8
	L2 regularization coefficient λ	0	1	1
LSTM	Number of hidden units	50	32	32
	Batch size	32	16	16
	Epochs	50	30	30
	Optimizer	adam	rmsprop	adam
	Learning rate	0.001	0.001	0.01

linear interpolation within the training sample only, avoiding information leakage. A detailed summary of the variables and data processing steps is provided in Appendix B.

4.2 Identification of key factors

To balance the nonlinear representation capability of the model with interpretability, this study employs the SHAP (SHapley Additive exPlanations) method. Within the XGBoost framework, SHAP values quantify the marginal contribution of each feature to the prediction output based on game-theoretic principles. These values serve as the primary basis for determining feature importance. Figure 3 presents the feature importance rankings obtained from the five methods.

The results indicate that natural gas consumption is primarily influenced by supply capacity and structural

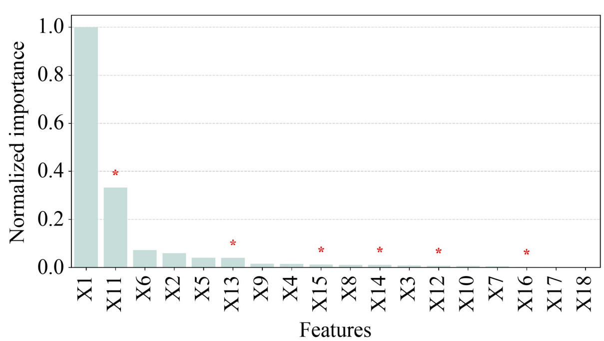

Fig. 3 Characteristic importance of features in the predictive model (* indicates risk-related factor).

risk factors. The external dependence risk index (X11) ranks immediately after the core supply–demand variables, suggesting that fluctuations in import dependence exert a persistent and systematic influence on China's natural gas consumption. This pattern reflects the country's reliance on imported LNG and pipeline gas from a limited number of source countries, which increases vulnerability to external disruptions. By contrast, domestic production and infrastructure expansion stabilize the system by mitigating the impact of import shocks. The high SHAP value of X11 thus underscores the strategic importance of energy self-sufficiency and supply diversification within China's energy security framework.

The price volatility risk index (X13) also shows substantial importance. Unlike X11, which captures long-term structural exposure, *X13* represents short-term market-driven shocks that affect consumer affordability and investment timing. Meanwhile, the US Dollar Index volatility (X14) occupies a mid-level position, implying that exchange rate movements influence natural gas imports primarily through indirect cost transmission mechanisms rather than direct physical constraints.

Other risk-related indicators, such as market concentration risk (X15) and transportation disruption risk (X17), have relatively smaller SHAP values but are consistently selected across model iterations, highlighting their relevance to overall market resilience. Collectively, these findings indicate that China's natural gas consumption dynamics are shaped not only by domestic demand or price levels but also by the interplay between supply-side vulnerabilities and international market volatility.

These findings confirm that incorporating multidimensional energy security risks, particularly import dependence and market concentration, is essential for accurate and policy-relevant demand forecasting. Integrating these risk dimensions enhances model robustness and provides empirical support for designing forward-looking, risk-informed energy policies. By quantifying the distinct channels through which external dependence, price shocks, and currency volatility affect consumption, the model establishes a more realistic analytical foundation for managing uncertainty in China's rapidly evolving natural gas market.

4.3 Performance comparison of SHAP-Screened models with and without attention mechanisms

To evaluate the predictive performance of the proposed SHAP-driven Weight-based Rule Attention Mechanism (SWRAM), this study compares it with two benchmark methods: (i) SHAP-driven Feature Screening without Attention Mechanism (SHAP-noAtt) and (ii) SHAP-driven Feature Screening with Adaptive Attention Mechanism (SHAP-AdaptAtt). The SHAP-noAtt framework integrates SHAP-based feature selection into four standard models, SVR, RF, LSTM, and XGBoost, without any attention mechanism, thereby isolating the contribution of SHAP to feature relevance and model accuracy. The results show that SHAP enhances predictive performance by improving variable selection and mitigating overfitting.

Building upon SHAP-noAtt, attention mechanisms were introduced to enhance interpretability. SHAP-AdaptAtt employs a dynamic, learnable attention structure, while SWRAM adopts a rule-constrained static attention strategy based on normalized SHAP values. This design allows SWRAM to integrate prior knowledge into the attention layer, ensuring that the relative importance of variables remains consistent with their economic and physical significance. Compared with adaptive attention, which may generate unstable weight fluctuations under non-stationary time-series conditions, the rule-based attention in SWRAM provides greater robustness and clearer interpretability.

Figure 4 compares the predictive performance of the three frameworks (SWRAM, SHAP-noAtt, and SHAP-AdaptAtt) using three key metrics: MAE (left panel), MAPE (right panel), and RMSE (represented by bubble size). The results indicate that SWRAM consistently achieves lower MAE and MAPE across most model configurations, particularly in the SVR and XGBoost settings, demonstrating improved accuracy and robustness. Notably, although SHAP-AdaptAtt is designed to capture adaptive attention, it performs significantly worse in the LSTM model, suggesting that the learned attention weights become unstable in high-dimensional temporal spaces. In contrast, SWRAM leverages rule-based SHAP-derived weights to produce more stable and interpretable forecasts. This stability is particularly important for risk-related variables such as import dependence and price volatility, whose influence on natural gas consumption is amplified under market uncertainty and policy adjustment cycles.

These findings confirm that incorporating a SHAP-based rule-constrained attention mechanism improves both performance and interpretability while maintaining manageable computational complexity. This balance between accuracy, stability, and explainability supports the practical applicability of SWRAM for policy-oriented energy forecasting and risk assessment tasks.

4.4 Error significance test

Figure 5 systematically compares the performance of three forecasting methods on both the validation and test sets: the best-performing model without attention (SHAP-noAtt), the best-performing model with adaptive attention (SHAP-AdaptAtt), and the proposed model with rule-based weighted attention (SWRAM). On the validation set, the SHAP-noAtt model achieved the highest performance, with an R^2 value of 0.995. However, its R^2 dropped sharply to 0.430 on the test set, indicating a clear overfitting issue.

SHAP-AdaptAtt exhibited consistently lower accuracy on both data sets. This suggests that the flexibility introduced by adaptive attention can result in excessive degrees of freedom under limited sample conditions, reducing model robustness. Instability arises particularly in non-stationary, high-dimensional time series, such as those encountered in energy markets subject to structural shocks.

In contrast, the SWRAM model outperformed both benchmarks on the test set, achieving the highest R^2 value of 0.553. Although its validation performance was slightly lower than that of SHAP-noAtt, SWRAM demonstrated substantially stronger out-of-sample generalization and error stability, confirming that embedding SHAP-based rule constraints improves robustness and interpretability. The improved generalization performance also reflects the model's ability to capture structural variations associated with import dependence and price volatility, two major sources of consumption uncertainty in China's natural gas market.

To formally assess the significance of forecast performance differences, the Diebold–Mariano (DM) test was

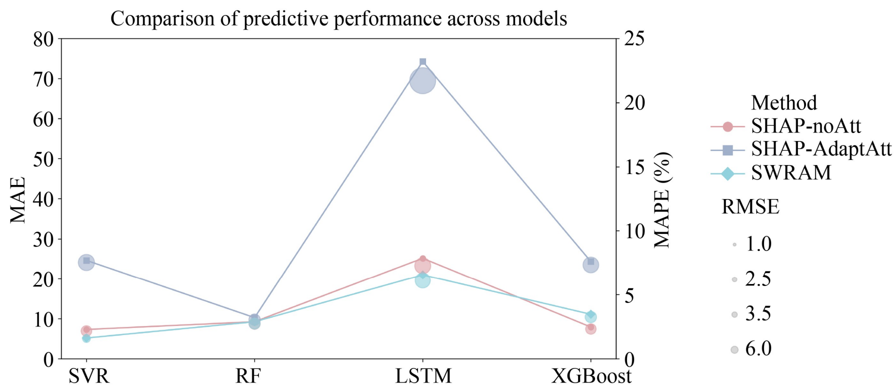

Fig. 4 Performance comparison of the SWRAM model and baseline methods (SHAP-noAtt and SHAP-AdaptAtt).

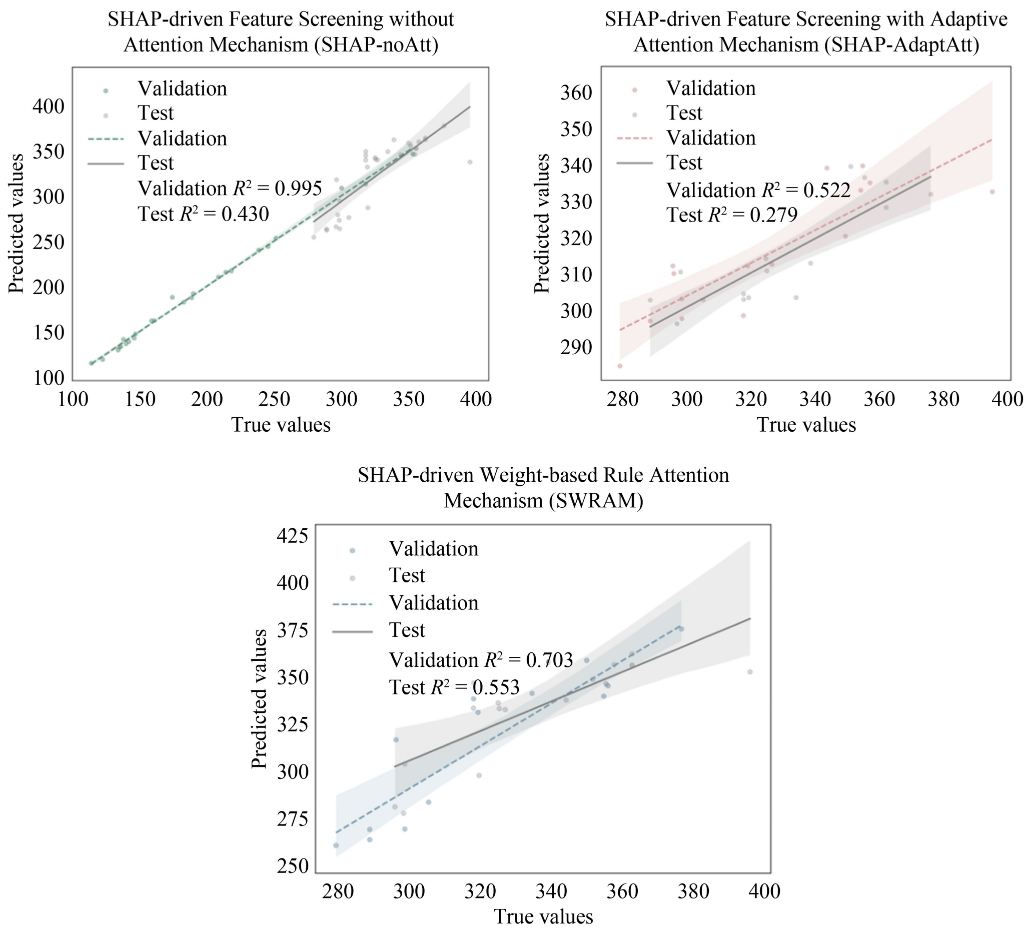

Fig. 5 Performance comparison of optimal natural gas consumption prediction models with no attention, adaptive attention, and SHAP-weighted rule-based attention.

applied to compare prediction errors among the three models. Table 2 reports the DM statistics and p-values for Mean Absolute Error (MAE) and Root Mean Square Error (RMSE). The results show that SWRAM significantly reduces MAE compared with SHAP-noAtt (DM = 2.9261, p = 0.0034) and SHAP-AdaptAtt (DM = 3.6576, p = 0.0003). Regarding RMSE, SWRAM significantly outperforms SHAP-AdaptAtt (DM = 2.4054, p = 0.0162), while the difference relative to SHAP-noAtt is not statistically significant (p = 0.3315). These findings indicate that SWRAM provides robust error control, particularly under conditions prone to extreme deviations.

To further validate the robustness, the models were re-estimated over two distinct time windows (2015–2019 and 2020–2024). Table 3 shows that the relative performance ranking remained consistent, with SWRAM consistently outperforming SHAP-noAtt and SHAP-AdaptAtt. This confirms that the superior performance of SWRAM is not driven by a particular sample structure or short-term market fluctuation, but reflects its inherent stability in capturing the nonlinear dynamics of natural gas consumption.

5 Analysis of key factor interactions and seasonal dynamics based on SHAP values

Within the proposed model, the interactions among risk factors, supply capacity, and market price dynamics are more effectively captured and interpreted, providing greater explanatory power for natural gas consumption forecasts. Through SHAP interaction analysis, this study further uncovers the combined influence of supply capability, market risks, and price volatility on consumption behavior. These insights offer more targeted and actionable recommendations for policymakers in the context of energy management and planning.

5.1 Contribution of influencing factors and interaction between influencing factors

5.1.1 Dominant factor contributions

According to the SHAP analysis presented in Fig. 6(a), natural gas consumption is primarily driven by supply capacity and risk-related factors. Supply-related variables (X1, X2, X6) account for 69.33% of the total contribu-

Table 2 DM Inspection results

Metric	Model comparison	DM statistic	*p*-value
MAE	SWRAM vs SHAP-noAtt	2.9261	0.0034
MAE	SWRAM vs SHAP-AdaptAtt	−2.4720	0.0134
MAE	SHAP-noAtt vs SHAP-AdaptAtt	−3.6576	0.0003
RMSE	SWRAM vs SHAP-noAtt	0.9711	0.3315
RMSE	SWRAM vs SHAP-AdaptAtt	−2.4054	0.0162
RMSE	SHAP-noAtt vs SHAP-AdaptAtt	−2.3901	0.0168

Table 3 Subsample stability results of three models across two time windows

Time window	Model	MAE	RMSE	R^2
2015–2019	SHAP-noAtt	0.068	0.092	0.463
	SHAP-AdaptAtt	0.071	0.098	0.452
	SWRAM	0.061	0.087	0.514
2020–2024	SHAP-noAtt	0.072	0.097	0.438
	SHAP-AdaptAtt	0.076	0.104	0.427
	SWRAM	0.064	0.089	0.498

tion, suggesting that policy efforts should prioritize enhancing domestic production, strengthening storage capacity, and optimizing transmission infrastructure to ensure stable market supply. In addition, risk-related factors (X11–X18) have a greater influence on consumption than price, demand, and macroeconomic policy variables, contributing 24.65% to the prediction outcome.

Among these, import dependence (X11) is the dominant driver, reflecting both material reliance on imports and exposure to external price shocks, exchange-rate volatility, and geopolitical risks. Market concentration risk (X12) and Dollar Index volatility (X14) represent complementary risk channels, influencing price adjustment efficiency and trade settlement costs, respectively. Together, these three types of risks, namely supply-side, market-structure, and financial-transmission, capture the main pathways through which external and internal shocks affect domestic consumption.

Demand-related factors (X3, X5, X7) and price-related variables (X8–X10) exhibit smaller contributions (3.22% and 1.94%, respectively), but their role will likely increase as the energy transition progresses. Policy and fiscal factors (X4) contribute only 0.86%, suggesting limited direct influence.

Overall, the SHAP analysis in this study reveals that natural gas consumption is primarily driven by supply capacity and risk-related factors, while demand, price, and policy variables contribute relatively less. This finding underscores the importance of ensuring stable supply and reducing market risks as key determinants of natural gas demand. Policymakers should therefore prioritize efforts to enhance domestic supply capacity, reduce import dependency, ensure the security of transportation infrastructure, optimize market structure, and stabilize prices. These measures are essential for strengthening the resilience of the energy market and improving the accuracy of natural gas demand forecasts. In addition to supporting more effective energy management, such strategies can also provide a robust policy foundation for ensuring long-term energy security and promoting sustainable development.

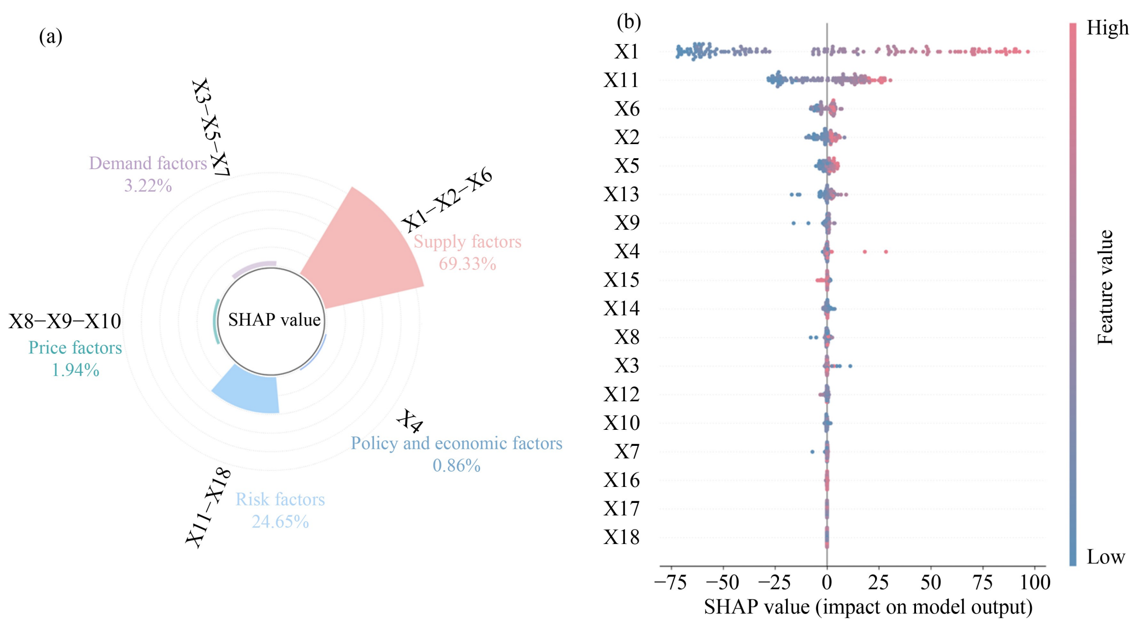

Fig. 6 SHAP-based analysis of key drivers of natural gas consumption: (a) SHAP value contributions of categorized factor groups: supply, risk, price, demand, and policy-related and (b) SHAP summary plot showing individual feature impacts on model output.

5.1.2 Interactions and seasonal dynamics

Based on the SHAP interaction visualization in Fig. 7, this study examines the interaction effects among different categories of variables in forecasting natural gas consumption. In the figure, bubble size represents the strength of pairwise nonlinear interactions, with larger bubbles indicating that the marginal impact of one variable on consumption changes substantially with the level of another variable.

Among all variable combinations, supply and risk-related factors exhibit the strongest interaction. Domestic natural gas production (X1) and external dependence risk (X11) reach an interaction strength of 3.94. This indicates that when domestic production is low and import dependence is high, consumption drops more sharply, demonstrating how excessive reliance on foreign supply amplifies supply–demand uncertainty. Conversely, when X1 is high, the marginal effect of X11 weakens significantly, suggesting that strengthening domestic production capacity can effectively mitigate external risk exposure. This mechanism reflects a dual-channel effect, in which domestic supply expansion both reduces vulnerability to import shocks and stabilizes demand expectations.

Infrastructure–supply interactions also play a critical role. The interaction strength between natural gas production (X1) and pipeline infrastructure length (X6) is 1.48, showing that infrastructure development magnifies the demand response to production increases. When *X6* is low, additional output cannot be efficiently delivered, making transportation capacity a binding constraint in the supply chain. As X6 increases, the supply elasticity of demand rises, highlighting infrastructure as both a capacity enhancer and a reliability stabilizer. Hence, expanding pipelines, LNG terminals, and interregional transmission networks remains essential for mitigating supply bottlenecks and regional imbalances.

Seasonal dynamics further shape these nonlinear interactions. China's natural gas demand exhibits a clear winter–summer dual-peak pattern. Winter heating demand in northern provinces produces sharp consumption surges, while summer electricity demand for air conditioning drives a secondary peak, particularly in eastern and southern regions. During winter shortages, the coupling between supply capacity (X1) and import risk (X11) intensifies, as heating loads coincide with limited pipeline throughput. In contrast, in summer, the interaction between electricity consumption (X5) and pipeline length (X6) becomes more pronounced, reflecting the increasing role of gas-fired power generation in balancing seasonal load fluctuations. These findings suggest that infrastructure capacity acts as a seasonal buffer, reducing volatility caused by extreme weather and uneven regional demand growth.

Price-related interactions, although relatively weaker, remain significant. The interaction between pipeline infrastructure (X6) and international price variables (X8, X9) ranges from 0.09 to 0.10. When infrastructure is underdeveloped, international price fluctuations are transmitted more directly to domestic users, amplifying consumption volatility. Conversely, a well-developed infrastructure network cushions this pass-through effect, thereby enhancing market resilience. Similarly, the interaction between electricity consumption (X5) and pipeline length (X6) demonstrates that as the power sector becomes increasingly gas-dependent, infrastructure

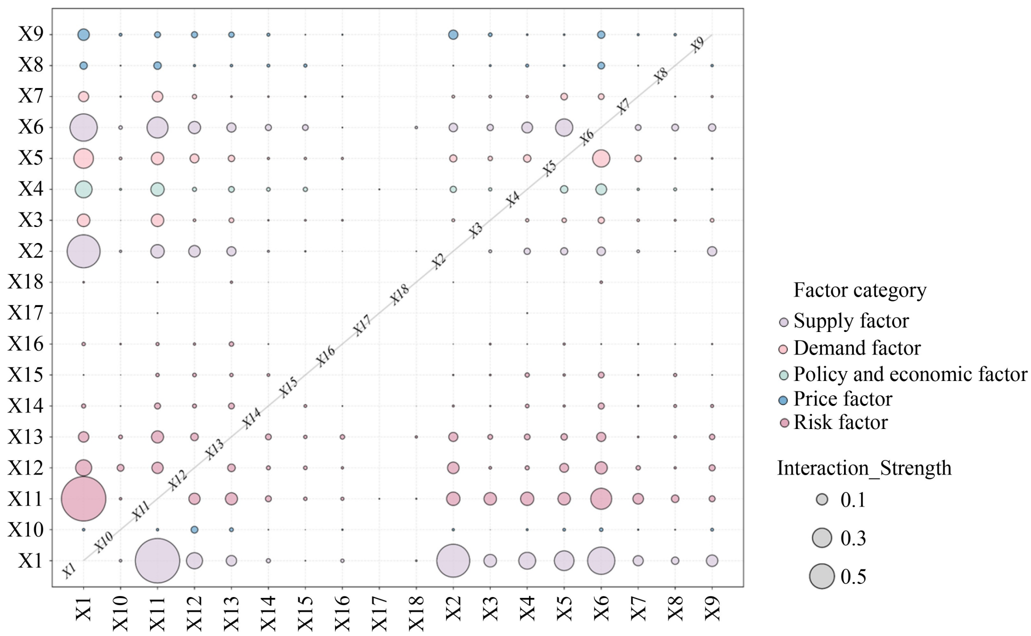

Fig. 7 Visualization of feature interaction strengths using SHAP values.

improvements facilitate greater gas penetration in the final energy mix, improving load matching and reducing bottleneck risks.

Finally, international prices and geopolitical risks, represented by Henry Hub spot prices (X9) and import-country risk (X16), interact with a strength of 0.17. High import risk amplifies the suppressive effect of rising international prices on domestic consumption, illustrating how geopolitical instability propagates through the price channel. Diversifying import origins and strengthening long-term bilateral agreements can help stabilize consumption by reducing domestic demand sensitivity to external shocks.

Overall, the SHAP interaction analysis reveals that nonlinear synergies rather than single-factor effects drive natural gas consumption dynamics. The combined influences of supply, infrastructure, and external risks determine consumption stability and responsiveness. These findings emphasize that effective energy governance requires a systemic approach. Enhancing domestic production, reinforcing infrastructure, and diversifying import sources can jointly mitigate volatility, balance seasonal load, and strengthen national energy security.

5.2 Analysis of seasonal variation in key factors based on SHAP values

Building on the interaction analysis in Subsection 5.1.2, Fig. 8 illustrates the temporal evolution and seasonal differentiation of the main determinants of natural gas consumption, based on the time-series dynamics of SHAP values. The results reveal that the relative importance of supply, infrastructure, price, and risk-related factors varies systematically across seasons, reflecting the interaction between domestic demand cycles and external market fluctuations.

In winter, when heating demand peaks and supply margins tighten, the SHAP value of domestic production capacity (X1) rises markedly, indicating that constraints in upstream output directly shape seasonal consumption variability. By contrast, infrastructure capacity (X6) maintains a steady and persistent influence throughout the year, confirming its structural role in mitigating inter seasonal imbalances and ensuring delivery continuity. The stability of X6 underscores that long-term investments in storage and transmission systems constitute a fundamental condition for moderating short-term supply shocks.

From 2021 to 2023, price-related variables (X9, X13) show a marked increase in their SHAP contributions, coinciding with heightened global energy market volatility. This pattern reflects the growing transmission of international price disturbances to the domestic market and suggests that consumption has become more sensitive to external price fluctuations. At the same time, import dependence (X11) and market concentration (X15) have risen in importance since 2020, indicating that global supply chain instability has become an additional structural driver of domestic consumption uncertainty. The co-movement of these two indicators suggests that systemic vulnerability stems not only from quantitative import reliance but also from asymmetries in supplier concentration.

Taken together, the seasonal and temporal patterns identified through SHAP analysis indicate that variations in China's natural gas consumption arise from the interaction between short-term demand cycles and long-term structural constraints. Seasonal elasticity is primarily determined by domestic supply capacity and infrastructure adequacy, which shape the system's ability to absorb winter–summer fluctuations. In contrast, exposure to global volatility is governed by external price shocks and import-related risks, reflecting the economy's increasing integration with international energy markets. These differentiated temporal effects reveal that consumption dynamics are driven not only by immediate demand

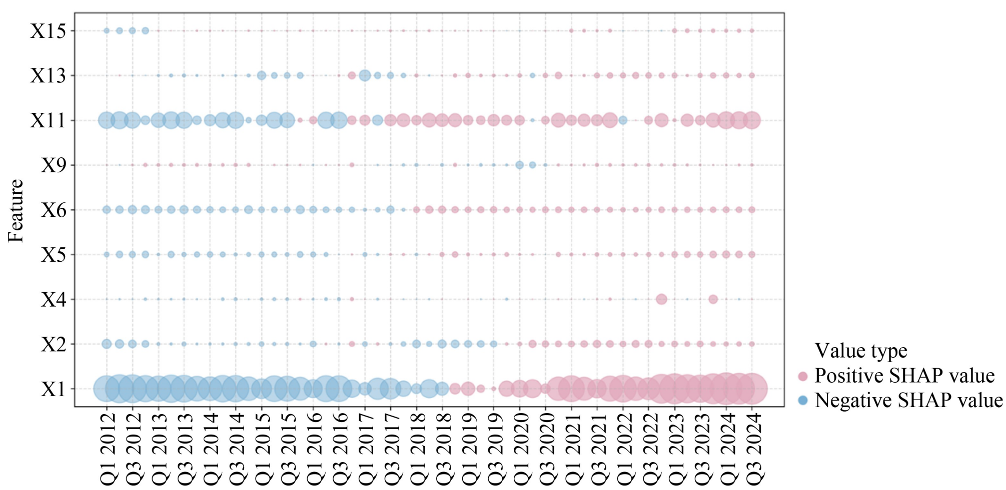

Fig. 8 SHAP value bubble timeline for key influencing factors of natural gas consumption.

pressures but also by the persistence of structural dependencies in supply and trade. Understanding this dual mechanism provides a theoretical basis for developing adaptive policy responses that link short-term stabilization with long-term resilience in the natural gas sector.

6 Conclusions and policy implications

This study develops an interpretable weighted attention mechanism model that integrates energy security risk factors (SWRAM) into natural gas consumption forecasting. The model advances the balance between predictive accuracy and interpretability by embedding structured risk dimensions, namely, import dependence, market concentration, and price volatility, within a data-driven analytical framework. By incorporating SHAP-based interpretability analysis, the model quantifies both the direct importance and nonlinear interactions of influencing factors, thus providing a transparent empirical foundation for policy-relevant energy forecasting.

The empirical findings reveal that supply capacity and infrastructure development remain the dominant determinants of natural gas consumption, with their combined SHAP contributions exceeding 69% across all forecast periods. In contrast, the importance of risk-related variables, including import dependency and market concentration, has risen significantly since 2020, accounting for approximately 25% of total model attribution. This shift reflects the growing influence of global supply-chain instability and domestic structural adjustments on consumption dynamics. Furthermore, the interaction analysis identifies a strong nonlinear coupling between domestic production and import dependence (interaction strength 3.94), confirming that insufficient domestic output amplifies the impact of external risks. The time-varying SHAP results also reveal that price factors exert a stronger marginal effect during global energy crises (2021–2023), while seasonal elasticity peaks in winter and summer, consistent with demand surges under extreme weather conditions. These findings collectively indicate that China's natural gas consumption is simultaneously shaped by cyclical demand pressures and structural dependencies, underscoring the need for adaptive policy coordination between short-term stability and long-term resilience.

Based on the above insights, this study offers the following differentiated policy implications:

1) Enhancing seasonal supply resilience. Establishing a flexible reserve mechanism tailored to seasonal cycles, particularly for winter heating demand, can mitigate short-term supply–demand imbalances. Expanding domestic exploration and storage capacity will improve self-sufficiency and reduce the transmission of external risks.

2) Designing an adaptive price-stabilization framework. The rise in SHAP-based price contributions since 2019 signals increasing market sensitivity. Introducing adaptive stabilization tools, such as fiscal buffers, risk-based subsidies, and derivative hedging instruments, can reduce volatility while preserving market efficiency.

3) Expanding infrastructure as a systemic buffer. The persistent importance of infrastructure variables (X6) highlights their moderating effect on price and supply shocks. Prioritizing investment in interregional pipeline networks and LNG terminals will strengthen cross-seasonal balancing capacity and support integrated electricity–gas dispatch.

4) Promoting demand-side flexibility and clean substitution. Encouraging natural gas utilization in the power and industrial sectors enhances the flexibility of the energy mix and contributes to the dual-carbon goals. Linking demand forecasting with infrastructure planning can prevent localized congestion and improve system coordination.

Overall, this study contributes to the literature by integrating model interpretability and energy security analysis within a unified framework. The SWRAM model not only improves forecasting accuracy but also enhances transparency in identifying key drivers and interactions, thereby increasing its practical relevance for policy design and risk assessment. Future research may extend this approach by incorporating real-time market indicators and geopolitical risk measures to further refine the temporal adaptability of natural gas forecasting models in an increasingly volatile global energy environment.

Appendix A

Appendix A1 Machine learning algorithms

SVR (Support Vector Regression) models nonlinear relationships by mapping input features from the original space to a high-dimensional feature space through a kernel function, and constructing a linear regression model in that space. Its prediction function is of the form:

$$y = w^T \varphi(x) + b. \tag{A1}$$

To solve the above model, the optimization objective of SVR is to minimize the combination of model complexity and prediction error. Its standard optimization problem is defined as follows:

$$\begin{aligned} &\min \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}(\xi_i + \xi_i^*) \\ &s.t.\ \ y_i - w^T\varphi(x_i) - b \leqslant \varepsilon + \xi_i \\ &w^T\varphi(x_i) + b - y_i \leqslant \varepsilon + \xi_i^* \\ &\xi_i, \xi_i^* \geqslant 0 \end{aligned} \tag{A2}$$

In Eq. (A2), $\varphi(x)$ denotes the nonlinear mapping of the input features into a high-dimensional feature space. $C \geqslant 0$ is the penalty parameter that balances model complexity and prediction error. ε represents the width of the ε-insensitive zone, while ξ_i and ξ_i^* are slack variables used to account for deviations beyond the ε-insensitive margin. In practical applications, kernel techniques are employed to avoid explicitly constructing $\varphi(x)$, enabling inner product computation in high-dimensional space via kernel functions. This study adopts the Radial Basis Function (RBF) kernel, defined as:

$$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2). \tag{A3}$$

Thus, indirectly realizing the nonlinear mapping:

$$K(x_i, x_j) = \varphi(x_i)^T \varphi(x_i). \tag{A4}$$

RF is an ensemble learning method that allows to improve the accuracy and stability of the proposed model by combining the predictions of multiple decision trees. Each decision tree was constructed using a subset of features to predict the dependent variable y. The splitting rule of the decision tree was based on the optimal split point of the multiple input factors, which was calculated as follows:

$$G = Impurity(parent) - \sum_{k=1}^{K} \frac{n_k}{N} \cdot Impurity(k), \tag{A5}$$

where G represents the gain; k is the child node after splitting; n_k is the number of samples in the child node; and N is the number of samples in the parent node. RF allows to aggregate the predictions of T decision trees and to average them to obtain the final output, as follows:

$$\hat{y} = \frac{1}{T} \sum_{t=1}^{T} f_t(X). \tag{A6}$$

LSTM is a type of special recurrent neural network (RNN) that can effectively address the long-term dependency problem. By introducing gating mechanisms, LSTM can selectively remember and forget information. The multi-factor input sequence $\{x_t, x_{t-1}, \cdots, x_{t-n}\}$ was processed through time steps t to capture both long-term and short-term dependencies. Its input consisted of the current time step's multi-factor features X_t and historical information. The core calculations were as follows:

(1) Forget gate:

$$f_t = \sigma(W_f \cdot [h_{t-1}, X_t] + b_f). \tag{A7}$$

(2) Input gate:

$$i_t = \sigma(W_i \cdot [h_{t-1}, X_t] + b_i), \tilde{C}_t = \tan h(W_C \cdot [h_{t-1}, X_t] + b_C). \tag{A8}$$

(3) Cell state update:

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t. \tag{A9}$$

where $\odot$ represents element-wise multiplication.

(4) Output gate: to determine how much information is transferred to the output.

$$o_t = \sigma(W_o \cdot [h_{t-1}, X_t] + b_o), h_t = o_t \odot \tan h(C_t). \tag{A10}$$

In the basic computation of LSTM, the output hidden state h_t represents the information extraction result at the current time step.

(5) Predicting the dependent variable: For forecasting tasks, the hidden state h_t is mapped to the dependent variable y_t through a fully connected layer f_{out}, as follows:

$$y_t = f_{\text{out}}(h_t) = W_{\text{out}} \cdot h_t + b_{\text{out}}, \tag{A11}$$

where σ is the sigmoid activation function; tan$h$ is the hyperbolic tangent function; W and U are weight matrices; and b is the bias vector.

XGBoost is an efficient implementation tool based on Gradient Boosting Decision Trees (GBDT), which allows to improve model performance by iteratively adding weak learners (decision trees). XGBoost builds a predictive model through incremental optimization by splitting multi-factor features. Each iteration minimizes the objective function, as follows:

$$\Gamma = \sum_{i=1}^{n} l\left(y_i, \hat{y}_i^{(t)}\right) + \sum_{k=1}^{T} \Omega(f_k), \tag{A12}$$

where l is the loss function; and $\Omega(f_k) = \gamma T + \frac{1}{2}\lambda\|w\|^2$ is the regularization term. The optimal solution for the leaf node weights is as follows:

$$w_j = -\frac{\sum_{i \in I_j} g_i}{\sum_{i \in I_j} h_i + \lambda}. \tag{A13}$$

The predicted value was updated as follows:

$$\hat{y}_i^{(t)} = \hat{y}_i^{(t-1)} + f_t(X_i). \tag{A14}$$

Appendix A2 Attention mechanism

(1) Adaptive attention mechanism

The adaptive Attention Mechanism dynamically learns weights based on input data, emphasizing attention to important features. The core idea is to generate attention weights through the interaction of query (Query), key (Key), and value (Value), with the weights automatically adjusted according to the input features. For the input feature matrix $X \in R^{n \times d}$ (where n is the number of features and d is the feature dimension), the calculation steps are as follows:

● Generation of the query, key, and value matrices: The query matrix Q, the key matrix K, and the value matrix V were generated through linear transformations, as follows:

$$Q = XW_Q,\ K = XW_K,\ V = XW_V, \tag{A15}$$

where W_Q and $W_V \in R^{d \times d_a}$ are learnable parameter matrices; and d_a represents the attention dimension.

● Attention score computation:

$$A = softmax\left(\frac{QK^T}{\sqrt{d_a}}\right), \tag{A16}$$

where $A \in R^{n \times n}$ is the attention weight between features; and each element A_{ij} represents the attention score between the i-th query and the j-th key.

● Weighted feature fusion: The value matrix was weighted and summed using the attention weights to generate a new feature representation X', as follows:

$$X' = AV. \tag{A17}$$

Finally, the output X' from the attention mechanism was further inputted into the downstream model for the purposes to predict natural gas consumption.

(2) Rule-based static attention mechanism

To improve interpretability and ensure model robustness, we construct a complementary static mechanism using pre-defined feature importance weights from SHAP. The procedure includes:

● Feature importance evaluation

$$s_i = f_i^{\text{SHAP}},\ I = \{I_1, I_2, \ldots, I_n\}, \tag{A18}$$

where s_i represents the SHAP value-based contribution of feature i.

● Normalization

$$w_i = \frac{s_i}{\sum_{j=1}^{n} s_j},\ i = 1, 2, \ldots, n. \tag{A19}$$

● Weight matrix and projection

We converted the weight vector w into a diagonal matrix $W_h \in R^{n \times n}$, as follows:

$$W = \text{diag}(w_1, \ldots w_n), \tag{A20}$$

$$X_{\text{static}} = WX. \tag{A21}$$

The resulting matrix X_{static} is merged with the adaptive representation X to yield a hybrid feature matrix for prediction.

This dual-attention architecture ensures that both fixed domain knowledge and data-driven feature dynamics are embedded in the model, improving both prediction accuracy and interpretability.

Prediction accuracy is an evaluation metric that measures the degree of agreement between a model's predicted and actual values. In this study, we used the Mean Absolute Error (MAE), the Root Mean Square Error (RMSE), and the Mean Absolute Percentage Error (MAPE) to assess accuracy. The calculation formulas are as follows:

$$MAE = \frac{1}{m} \sum_{i=1}^{m} |y_i - \hat{y}_i|, \tag{A22}$$

$$RMSE = \sqrt{\frac{1}{m} \sum_{i=1}^{m} (y_i - \hat{y}_i)^2}, \tag{A23}$$

$$MAPE = \frac{1}{m} \sum_{i=1}^{m} \left|\frac{y_i - \hat{y}_i}{y_i}\right| \times 100\%, \tag{A24}$$

where y_i and $\hat{y}_i$ are the actual and predicted values, respectively; and m is the number of data sets.

Appendix B

Table B1 Data frequency harmonization and variable handling methods

Type	Name	Processing method	Notes
Annual data	Natural gas price volatility risk index	The index is constructed based on daily data from three major spot prices: JKM, Henry Hub (HH), and NBP. First, the daily returns for each price series within a given month are calculated, using either log returns or relative changes. Then, the volatility level is quantified to capture the extent of price fluctuations during the month. $Price\ Volatility\ Index_t = \sigma\left(r_t^{(d)}\right)$ $r_t^{(d)} = \frac{P_t^{(d)} - P_{t-1}^{(d)}}{P_{t-1}^{(d)}}$	$P_t^{(d)}$ denotes the spot price on the d trading day, and σ represents the standard deviation. This method captures the intensity of intra-month price fluctuations and serves as a quantitative indicator of market uncertainty.
	US Dollar Index volatility risk index	This index is constructed based on the daily values of the US Dollar Index (DXY). The daily returns of the DXY for all trading days within each month are used to compute the coefficient of variation (CV), which serves as the monthly exchange rate risk indicator and reflects the relative volatility risk. $Dollar\ Volatility\ Index_t = \frac{\sigma\left(r_t^{(d)}\right)}{\bar{r}_t^{(d)}}$	$\sigma\left(r_t^{(d)}\right)$ represents the standard deviation of daily returns of the US Dollar Index in month t, and $\bar{r}_t^{(d)}$ denotes the mean of daily returns during the same month.
	Energy structure	Natural gas to primary energy consumption ratio	

(*Continued*)

Type	Name	Processing method	Notes
	Length of natural gas pipeline infrastructure Natural gas external dependence risk index Natural gas resource reserve risk index Natural gas market concentration risk index Importer country risk index Natural gas transportation disruption risk index Natural gas purchasing power risk index	Monthly replication. That is, the current year's data are evenly replicated to each month of the year.	
Daily data	JKM, Henry Hub (HH), and NBP spot prices	The average is calculated using historical daily data within the corresponding month, that is: $P_m = \frac{1}{n_m} \sum_{i=1}^{n_m} p_i$	P_m represents the monthly average price in month m, n_m is the number of trading days in that month, and p_i denotes the actual price on the i-th day. The entire process relies solely on data available within the given month, with no use of future information.

Appendix C

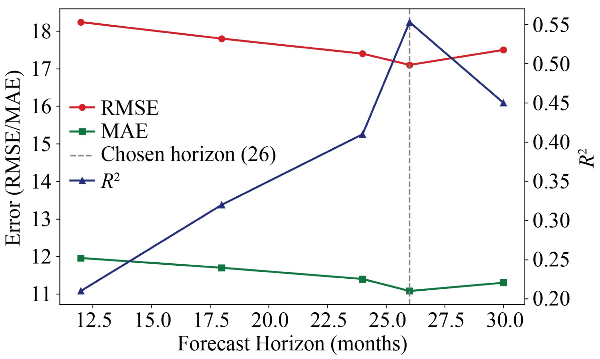

Fig. C1 Sensitivity of forecasting performance to horizon length: RMSE, MAE, and R^2 for rolling-origin predictions (SWRAM).

Competing Interests The authors declare that they have no competing interests.

References

Agrell P J, Dehaybe H, Rodriguez M H (2025). Balancing supply security and decarbonization: Optimizing Germany's LNG port infrastructure under the European Green Deal. Energy Policy, 198: 114484

Ata B, Pénzes J, Pakrooh P, Lotfata A, Mohammed S (2025). Comprehensive analysis of factors influencing EU-countries household energy consumption (2000–2022): Trends and vulnerability to COVID-19 and the Russia-Ukraine war. Energy Strategy Reviews, 59: 101728

Avraam C, Bistline J E T, Brown M, Vaillancourt K, Siddiqui S (2021). North American natural gas market and infrastructure developments under different mechanisms of renewable policy coordination. Energy Policy, 148: 111855

Avraam C, Chu D, Siddiqui S (2020). Natural gas infrastructure development in North America under integrated markets. Energy Policy, 147: 111757

Aydin M (2018). Natural gas consumption and economic growth nexus for top 10 natural gas-consuming countries: A granger causality analysis in the frequency domain. Energy, 165: 179–186

Azlina A A, Law S H, Nik Mustapha N H (2014). Dynamic linkages among transport energy consumption, income and CO2 emission in Malaysia. Energy Policy, 73: 598–606

Benigno P, Morana C, Tirelli P (2025). Introduction to the special issue on macroeconomic regime changes: Theory, evidence, and policy challenges ahead. European Economic Review, 180: 105164

Bhagat S K, Tiyasha T, Awadh S M, Tung T M, Jawad A H, Yaseen Z M (2021). Prediction of sediment heavy metal at the Australian Bays using newly developed hybrid artificial intelligence models. Environmental Pollution, 268: 115663

Bilgili L (2025). Environmental life cycle assessment of the Russian, Canadian, and Qatari natural gas supply to Europe. Journal of Environmental Management, 387: 125922

Boţa-Avram C, Apostu S A, Ivan R, Achim M V (2024). Exploring the impact of macro-determinant factors on energy resource depletion: Evidence from a worldwide cross-country panel data analysis. Energy Economics, 130: 107341

Cao W, Wang G, Liang X, Hu Z (2024a). A STAM-LSTM model for wind power prediction with feature selection. Energy, 296: 131030

Cao X, Zhang Z, Qian Y, Wen Z (2024b). The spatial pattern, driving factors and evolutionary trend of energy cooperation and consumption in the "Belt and Road Initiative" countries. Energy, 306: 132416

Chai J, Liang T, Lai K K, Zhang Z G, Wang S (2018). The future natural gas consumption in China: Based on the LMDI-STIRPAT-PLSR framework and scenario analysis. Energy Policy, 119: 215–225

Chen Y, Wang C, Miao J (2025). Marginal impact of European gas shocks on EU inflation variations: Evidence from a TVP-SVAR-SV approach. Economic Analysis and Policy

Dai X, Xiao L, Li M C, Wang Q (2022). Toward energy finance market transition: Does China's oil futures shake up global spots market? Frontiers of Engineering Management, 9: (3)409–424

Demski C, Poortinga W, Whitmarsh L, Böhm G, Fisher S, Steg L, Umit R, Jokinen P, Pohjolainen P (2018). National context is a key determinant of energy security concerns across Europe. Nature Energy, 3(10): 882–888

Di Bella G, Flanagan M, Foda K, Maslova S, Pienkowski A, Stuermer M, Toscani F (2024). Natural gas in Europe: The potential impact of disruptions to supply. Energy Economics, 138: 107777

Eskandari H, Saadatmand H, Ramzan M, Mousapour M (2024). Innovative framework for accurate and transparent forecasting of energy consumption: A fusion of feature selection and interpretable machine learning. Applied Energy, 366: 123314

Goodell J W, Gurdgiev C, Paltrinieri A, Piserà S (2023). Global energy supply risk: Evidence from the reactions of European natural gas futures to Nord Stream announcements. Energy Economics, 125: 106838

Jiang H, Dong X, Jiang Q, Dong K (2020). What drives China's natural gas consumption? Analysis of national and regional estimates. Energy Economics, 87: 104744

Karasu S, Altan A, Bekiros S, Ahmad W (2020). A new forecasting model with wrapper-based feature selection approach using multi-objective optimization technique for chaotic crude oil time series. Energy, 212: 118750

Kim J, Jaumotte F, Panton A J, Schwerhoff G (2025). Energy security and the green transition. Energy Policy, 198: 114409

Kong F, Li J, Jiang B, Wang H, Song H (2023). Integrated generative model for industrial anomaly detection via bidirectional LSTM and attention mechanism. IEEE Transactions on Industrial Informatics, 19(1): 541–550

Li C, Yang X, Wang L (2025). Public energy conservation behavior and the impact of tiered pricing: Based on water, electricity and gas consumption measurements in typical urban households in China. Energy Policy, 202: 114593

Li F, Zhang J, Li X (2023). Energy security dilemma and energy transition policy in the context of climate change: A perspective from China. Energy Policy, 181: 113624

Li M, Liu X, Yang M (2024). Analyzing the regional inequality of renewable energy consumption and its driving factors: Evidence from China. Renewable Energy, 223: 120043

Li Y, Yang Y, Zhu K, Zhang J (2021). Clothing sale forecasting by a composite GRU–prophet model with an attention Mechanism. IEEE Transactions on Industrial Informatics, 17(12): 8335–8344

Liang T, Chai J, Zhang Y J, Zhang Z G (2019). Refined analysis and prediction of natural gas consumption in China. Journal of Management Science and Engineering, 4(2): 91–104

Liu J, Peng C, Zhang J (2025). Understanding the relationship between rural morphology and photovoltaic (PV) potential in traditional and non-traditional building clusters using shapley additive exPlanations (SHAP) values. Applied Energy, 380: 125091

Peng C, Chen H, Lin C, Guo S, Yang Z, Chen K (2021). A framework for evaluating energy security in China: Empirical analysis of forecasting and assessment based on energy consumption. Energy, 234: 121314

Qiao W, Yang Z, Kang Z, Pan Z (2020). Short-term natural gas consumption prediction based on Volterra adaptive filter and improved whale optimization algorithm. Engineering Applications of Artificial Intelligence, 87: 103323

Sassi F (2025). The (Un)Intended consequences of power: The global implications of EU LNG strategy to reach independence from Russian gas. Energy Policy, 198: 114494

Scharf H, Möst D (2024). Gas power — How much is needed on the road to carbon neutrality? Energy Policy, 187: 114026

Wang C, Song J, Shi D, Reyna J L, Horsey H, Feron S, Zhou Y, Ouyang Z, Li Y, Jackson R B (2023). Impacts of climate change, population growth, and power sector decarbonization on urban building energy use. Nature Communications, 14(1): 6434

Wang C N, Nguyen T T V, Chiang C C, Le H D (2024a). Evaluating renewable energy consumption efficiency and impact factors in Asia-pacific economic cooperation countries: A new approach of DEA with undesirable output model. Renewable Energy, 227: 120586

Wang H, Li Y F, Ren J (2024b). Machine learning for fault diagnosis of high-speed train traction systems: A review. Frontiers of Engineering Management, 11(1): 62–78

Wang T, Lin B (2014). China's natural gas consumption and subsidies—From a sector perspective. Energy Policy, 65: 541–551

Wang Y, Song M, Han Y, Wang Y (2025). Multi-scale coupling measurement and driving factor analysis between carbon emission from energy consumption and ecological environment—The case study of Pearl River Delta urban agglomeration. Sustainable Cities and Society, 127: 106425

Wejnert-Depue C, Zhang Y, Casper K, O'Neill B C, Waldhoff S T (2025). A conceptual framework for residential energy security in the context of clean energy transitions. Energy Research & Social Science, 126: 104096

Xu B, Lin B (2019). Can expanding natural gas consumption reduce China's CO_2 emissions? Energy Economics, 81: 393–407

Yang K, Han Q, Yang D, de Vries B (2025). Historical changes and driving factors of food-water-energy footprint consumption: A Case study of the Beijing-Tianjin-Hebei city agglomeration. Sustainable Cities and Society, 122: 106222

Zang H, Xu R, Cheng L, Ding T, Liu L, Wei Z, Sun G (2021). Residential load forecasting based on LSTM fusing self-attention mechanism with pooling. Energy, 229: 120682

Zhang L, Bai W, Xiao H, Ren J (2021). Measuring and improving regional energy security: A methodological framework based on both quantitative and qualitative analysis. Energy, 227: 120534

Zhao J, Jiang Q, Dong X, Dong K (2020). Would environmental regulation improve the greenhouse gas benefits of natural gas use? A Chinese case study. Energy Economics, 87: 104712

Zhong Y, Li G, Yang J, Zheng H, Yu Y, Zhang J, Luo H, Wang B, Weng Z (2024). Learning motif-based graphs for drug–drug interaction prediction via local–global self-attention. Nature Machine Intelligence, 6(9): 1094–1105