

Article

IQTN: An Interpretable Quantile Temporal Network for Systems-Oriented Tail-Risk Forecasting and Early Warning in Carbon Allowance Market

Tianli Huang  and Grace T. R. Lin 

Institute of Management of Technology, National Yang Ming Chiao Tung University, No. 1001, Daxue Rd., East Dist., Hsinchu City 300093, Taiwan; htlcjwd@gmail.com

* Correspondence: gtrl@nycu.edu.tw

Abstract

The carbon emission allowance (CEA) market is a complex socio-technical and environmental-management system in which regulatory design, trading activity, liquidity conditions, and price volatility interact dynamically. Accurate systems-level tail-risk forecasting and early warning remain challenging because carbon-market losses are affected by nonlinear dependence, episodic liquidity stress, and time-varying volatility. This study proposes an Interpretable Quantile Temporal Network (IQTN) as a systems-oriented risk-monitoring framework for China's national CEA market. By integrating a feature-gating mechanism, a causal temporal convolutional encoder, and a non-crossing quantile output layer, IQTN directly models the conditional tail distribution of future carbon-market losses. The framework produces multi-horizon Value-at-Risk (VaR) and Conditional Value-at-Risk (CVaR) forecasts for 1-day, 5-day, and 10-day horizons and converts predicted tail risk into operational early-warning signals. Compared with historical simulation, EWMA, GARCH-type models, machine-learning quantile models, and deep temporal benchmarks, IQTN achieved the lowest 95% VaR pinball loss across all horizons, with values of 0.1765, 0.3958, and 0.5732. VaR backtesting showed empirical exceedance rates of 5.23%, 6.04%, and 6.94%, closest to the nominal 5% level. Interpretability analysis identified rolling volatility, maximum loss, intraday range, trading value, and illiquidity as key risk drivers. The temporal importance results also show that recent observations dominated the risk forecasts, suggesting that the risk state of the CEA market is highly sensitive to short-term market information. This supports the use of a short-horizon temporal network as a systems-oriented tool for carbon-market tail-risk monitoring and early warning.

Keywords: carbon allowance market; carbon-market risk system; tail-risk forecasting; Value-at-Risk; Conditional Value-at-Risk; interpretable temporal network; early warning; systems-oriented risk monitoring



Academic Editor: Jun Wu

Received: 6 May 2026

Revised: 10 June 2026

Accepted: 18 June 2026

Published: 24 June 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Carbon allowance markets are increasingly important institutional systems for supporting low-carbon transition, emissions mitigation, and climate-risk governance. From a systems perspective, a carbon market is not merely a price-discovery venue but a complex socio-technical and environmental-management system in which regulatory design, compliance demand, trading behavior, liquidity conditions, and information flows interact dynamically. This view is consistent with the broader sustainability-transition literature,

which emphasizes that environmental transitions involve co-evolving technologies, institutions, user practices, regulations, and market structures [1,2]. China's carbon market has developed through regional pilot experiments and subsequent national-market construction, and existing studies show that carbon allowance prices in China are shaped by market design, institutional maturity, trading activity, and heterogeneous regional conditions [3,4]. As China's national carbon allowance market continues to expand, the need for a systems-oriented risk-monitoring framework becomes increasingly important.

To operationalize this systems perspective, this study treats the national CEA market as a set of interconnected market-state subsystems. The regulatory and compliance subsystem shapes institutional timing, allowance demand, and compliance-related pressure. The trading-behavior subsystem reflects how market participants adjust allowance positions through trading volume and trading value. The liquidity subsystem captures whether transaction demand can be absorbed without large price impact. The price-volatility and downside-pressure subsystem reflects intraday range, recent volatility, and realized losses. The information-adjustment subsystem reflects how policy expectations, compliance timing, and recent market signals are incorporated into trading behavior and prices. Based on these subsystems, we propose a risk-transmission hypothesis:

$$\begin{aligned} &\text{regulatory/compliance pressure} \rightarrow \text{trading concentration} \\ &\rightarrow \text{liquidity stress} \rightarrow \text{volatility and downside pressure} \\ &\rightarrow \text{tail risk.} \end{aligned} \quad (1)$$

This hypothesis guides the variable construction and model architecture. Feature gating is used to identify which subsystem-related variables are most predictive at a given time; causal temporal convolution is used to capture the short-horizon propagation of market-state signals; and non-crossing quantile estimation is used to translate the propagated system state into coherent VaR- and CVaR-relevant tail forecasts. This framework should be understood as a mechanism-oriented predictive framework rather than a strict causal-identification model.

The systemic nature of carbon-market risk becomes especially evident when the focus shifts from price prediction to downside-risk monitoring. For compliance entities, financial institutions, and regulators, the critical question is not only whether carbon prices will rise or fall, but also whether the market is approaching a high-risk state in which large losses may occur. Carbon prices can exhibit nonlinear dependence, volatility clustering, liquidity stress, and sensitivity to macroeconomic and policy-related information. Prior evidence indicates that carbon-market risk is associated with macroeconomic fundamentals and time-varying volatility conditions [5,6]. Therefore, effective carbon-market governance requires more than a static risk indicator. It requires a dynamic early-warning system that can forecast tail losses, update risk signals over time, and provide interpretable information about the market-state variables driving risk.

A growing body of research has examined carbon price forecasting, volatility modeling, and carbon-market risk measurement. Recent studies have introduced interval forecasting, decomposition-based learning, causal temporal convolutional networks, generative models, and dynamic interval-valued frameworks to improve carbon price prediction under nonlinear and non-stationary conditions [7–10]. These studies have substantially advanced the forecasting literature by moving beyond conventional linear models and by incorporating richer market information. However, price-level prediction, return prediction, volatility modeling, and interval prediction are not equivalent to tail-risk forecasting. Price and return models mainly address the expected direction or magnitude of market movement, volatility models describe conditional variability, and interval forecasts describe a range

of plausible prices. By contrast, VaR- and CVaR-oriented risk monitoring requires direct estimation of the conditional upper-tail distribution of future carbon-market losses. This distinction is important because a market can display moderate average price movement while still facing a high probability of severe downside loss under stressed liquidity or concentrated trading conditions.

Three research gaps, therefore, remain. First, current carbon-market forecasting models have not systematically modeled tail risk as a conditional upper-tail loss-distribution problem. Many existing models target the conditional mean, volatility, or price interval, but these targets are less directly aligned with Value-at-Risk, Conditional Value-at-Risk, and extreme-loss monitoring. Second, existing deep time-series models, including generic LSTM- and TCN-based forecasting architectures, can capture nonlinear temporal dependence but often provide limited interpretability. In carbon-market risk governance, this is problematic because regulators and market participants need to understand not only whether a warning signal is generated, but also whether the signal is driven by volatility, downside pressure, trading concentration, liquidity stress, intraday price instability, or compliance-related timing. Interpretable multi-horizon time-series models provide useful design principles for this purpose [11], and the broader interpretable machine-learning literature emphasizes the importance of transparent model structures in high-stakes applications [12]. Third, existing risk early-warning models have not sufficiently integrated the systemic characteristics of carbon markets. In particular, liquidity conditions, trading activity, intraday range, calendar pressure, and recent downside losses are often treated as separate predictors rather than as interacting components of a dynamic carbon-market risk system.

The theoretical advancement of IQTN is to align these three gaps within a single carbon-market tail-risk monitoring framework. Compared with conventional quantile models, IQTN does not only estimate isolated conditional quantiles; it embeds quantile estimation within a feature-gated temporal architecture that links market-state variables to interpretable risk-driver weights. Compared with generic TCN or LSTM forecasting models, IQTN does not use temporal representation merely for predictive accuracy; it uses causal temporal convolution to capture short-horizon propagation of volatility, liquidity, and trading-pressure signals and then maps these propagated states into non-crossing tail quantiles. Compared with generic early-warning approaches, IQTN integrates systemic carbon-market characteristics into the warning process by combining VaR, CVaR, feature-gating explanations, temporal importance, and hierarchical warning levels. In this sense, IQTN provides a carbon-market-specific extension of time-series quantile forecasting by introducing feature-gated interpretability into dynamic multi-horizon tail-risk prediction and early warning.

To address these gaps, this study proposes an Interpretable Quantile Temporal Network (IQTN) for dynamic tail-risk forecasting and early warning in China's national carbon allowance market. The proposed framework treats carbon-market risk as a systems-level prediction problem. It uses historical market-state sequences to estimate the conditional quantile function of future carbon-market losses over 1-day, 5-day, and 10-day horizons. The IQTN integrates three components: a feature-gating layer that identifies the relative contribution of market-state variables, a causal temporal convolutional encoder that captures short-term risk persistence, and a non-crossing quantile output layer that directly produces multi-level tail-risk forecasts. The predicted quantiles are further transformed into Value-at-Risk, Conditional Value-at-Risk, and four-level early-warning signals. In this way, the proposed model functions not only as a forecasting tool but also as an interpretable decision-support system for carbon-market risk monitoring.

The contributions of this study can be summarized in four aspects.

1. This study contributes to carbon-market systemic risk theory by reframing carbon-market forecasting as a systems-oriented tail-risk monitoring problem. Instead of treating carbon-market risk as a purely price-level or volatility-forecasting task, the proposed framework conceptualizes tail risk as the outcome of interactions among regulatory and compliance pressure, trading concentration, liquidity stress, volatility, and downside pressure. This provides a theoretical basis for interpreting carbon-market tail risk as a dynamic system-state outcome rather than as an isolated statistical feature of historical returns.
2. This study develops an interpretable quantile temporal network for multi-horizon carbon-market tail-risk forecasting. The model integrates feature gating, causal temporal sequence learning, and non-crossing quantile estimation to directly estimate the conditional upper-tail distribution of future CEA losses. This contributes a task-specific modeling framework for VaR and CVaR forecasting in an emerging carbon allowance market.
3. This study proposes an interpretable mechanism for monitoring systemic tail risk, thereby bridging the gap between black-box prediction and actionable regulation in carbon-market systemic risk modeling. By using global feature-gating weights, lag-specific temporal importance, and local severe-warning explanations, the framework identifies whether risk warnings are mainly driven by volatility, downside pressure, trading activity, liquidity stress, intraday price instability, or calendar-related market states.
4. This study provides a managerial and regulatory contribution by converting predicted VaR and CVaR into hierarchical early-warning signals. The normal, watch, warning, and severe-warning levels provide regulators, compliance entities, and trading institutions with an interpretable tool for dynamic risk monitoring, liquidity-risk review, exposure adjustment, compliance planning, and adverse-scenario stress testing. The framework therefore supports both predictive accuracy and actionable carbon-market risk governance.

The methodological contribution of IQTN should therefore be understood as task-specific architectural integration rather than the invention of a single isolated neural-network component. Feature gating, temporal convolution, and quantile estimation each have methodological antecedents, but in this study they are aligned with a specific systems-risk estimand: the conditional upper-tail distribution of future CEA losses across multiple forecast horizons. The feature-gating layer translates market-state variables into dynamic and interpretable risk-driver weights; the causal temporal convolutional encoder models short-horizon propagation of volatility, liquidity, and trading-pressure signals; the non-crossing quantile output layer produces coherent tail quantiles for VaR and CVaR estimation; and the early-warning conversion maps predicted tail risk into operational warning levels. Thus, IQTN is not intended as a generic stacking of existing techniques. Its novelty lies in integrating feature-level interpretability, temporal risk propagation, coherent tail-distribution estimation, and early-warning conversion into a unified framework for carbon-market systemic tail-risk monitoring.

The remainder of this paper is organized as follows. Section 2 presents the data, feature construction, tail-risk definitions, IQTN architecture, benchmark models, and evaluation strategy. Section 3 reports the empirical results, including quantile forecasting accuracy, VaR backtesting, CVaR evaluation, early-warning performance, interpretability analysis, robustness checks, and ablation results. Section 4 discusses the principal findings, implications for carbon-market risk monitoring, limitations, and conclusions.

2. Methodology

2.1. Overall Research Design

This study develops an interpretable quantile temporal network for dynamic tail-risk forecasting and early warning in China's national carbon allowance market. The empirical object is the China Emission Allowance (CEA) traded in the national carbon emissions trading market. The sample covers all official CEA trading days from 16 July 2021 to 30 January 2026. We acknowledge that this sample period is relatively short compared with mature financial markets. However, the short sample reflects the institutional age of China's national carbon market rather than discretionary sample truncation. The starting date corresponds to the opening of online trading in China's National Emissions Trading System, and the endpoint fixes the information set used in this study. We deliberately restrict the main analysis to the national CEA market rather than pooling earlier regional pilot-market observations because China's regional pilot markets differed in allocation rules, sectoral coverage, local trading arrangements, and market maturity. Combining pilot-market data with national-market CEA data would mix institutionally different carbon-market systems and could weaken the internal consistency of the systems-oriented risk-monitoring framework. Therefore, the sample is interpreted as the full available trading history of a single national carbon-market system. The CEA trading records are obtained from the official trading information released by the Shanghai Environment and Energy Exchange, including opening price, highest price, lowest price, closing price, trading volume, and trading value [13,14].

The methodological framework is designed for risk forecasting rather than carbon price point prediction. Let P_t denote the official CEA closing price on trading day t . The model uses the historical market state observed up to day t to forecast the conditional tail distribution of future carbon market losses over horizons $h \in \{1, 5, 10\}$. The one-day horizon captures immediate market risk, the five-day horizon captures weekly risk, and the ten-day horizon captures short-term compliance and trading-pressure risk. The proposed framework produces multi-quantile forecasts, estimates Value-at-Risk (VaR) and Conditional Value-at-Risk (CVaR), converts the predicted tail-risk level into early-warning signals, and explains the contribution of market-state variables through a dynamic feature-gating mechanism. The research design operationalizes the systems-risk hypothesis by linking observable variables to carbon-market subsystems. Institutional and compliance timing is represented by calendar variables and the year-end trading-pressure indicator. Trading behavior is represented by trading volume and trading value. Liquidity conditions are represented by the Amihud-type illiquidity proxy. Price-stress and downside-pressure states are represented by daily return, absolute return, squared return, intraday range, rolling volatility, and rolling maximum loss. The tail-risk outcome is represented by future loss, VaR, CVaR, and the four-level early warning signal. In the IQTN architecture, the feature-gating layer captures the time-varying relevance of these subsystem-related variables, the causal temporal convolutional encoder captures their short-term propagation through the input sequence, and the non-crossing quantile output layer maps the learned system state into coherent conditional tail-risk forecasts. Thus, the model structure is designed to reflect the hypothesized pathway from institutional and trading pressure to liquidity stress, volatility amplification, downside pressure, and future tail risk. The overall framework of the proposed Interpretable Quantile Temporal Network (IQTN) for carbon-market tail-risk forecasting and early warning is illustrated in Figure 1. As shown in the figure, the framework consists of four main components: market data collection, feature engineering, the core IQTN architecture, and the operational early-warning output.

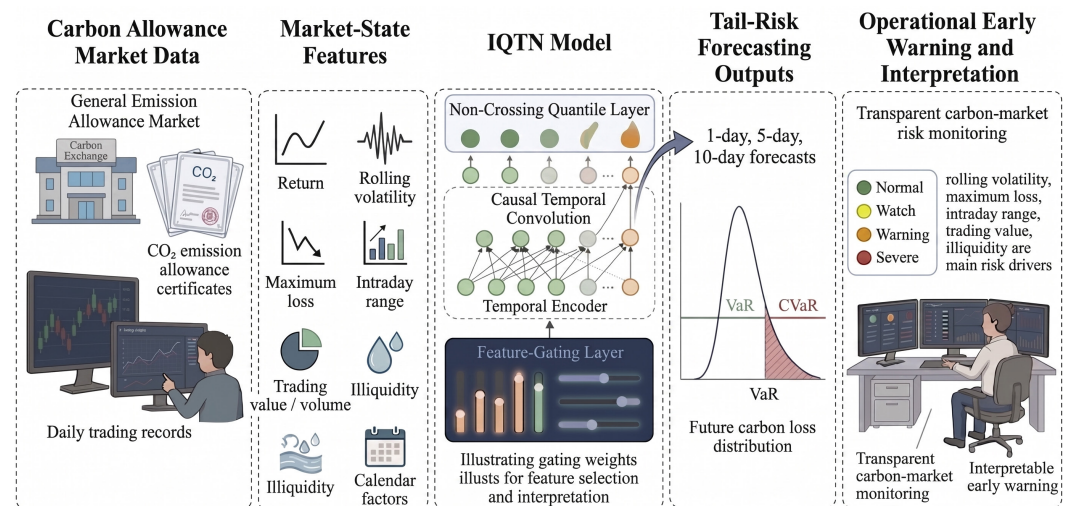


Figure 1. The overall research framework of the Interpretable Quantile Temporal Network (IQTN) for carbon-market tail-risk forecasting and early warning. The process encompasses (i) carbon allowance market data, (ii) multi-source market-state feature engineering, (iii) the IQTN model architecture consisting of feature-gating and temporal encoding, (iv) multi-horizon VaR and CVaR forecasting, and (v) operational risk monitoring and interpretation.

The full workflow consists of six steps. First, official CEA daily trading records are cleaned and aligned by trading day. Second, carbon market return, volatility, liquidity, range, and calendar variables are constructed. Third, future carbon market loss is defined for each forecast horizon. Fourth, the proposed Interpretable Quantile Temporal Network (IQTN) learns the conditional quantile function of future losses. Fifth, VaR, CVaR, and early-warning levels are generated from the estimated tail distribution. Sixth, the model is evaluated against traditional VaR models, machine-learning quantile models, and deep temporal quantile models using quantile loss, VaR backtesting, CVaR evaluation, and early-warning performance.

2.2. Systems-Theoretic Framing of Carbon-Market Tail Risk

To make the systems perspective explicit, we represent the national CEA market as a dynamic market-state system. Let

$$\mathbf{S}_t = (\mathbf{s}_t^{reg}, \mathbf{s}_t^{trd}, \mathbf{s}_t^{liq}, \mathbf{s}_t^{vol}, \mathbf{s}_t^{info}) \quad (2)$$

denote the carbon-market system state at trading day t . Here, $\mathbf{s}_t^{reg}$ represents regulatory and compliance-related timing pressure, $\mathbf{s}_t^{trd}$ represents trading-behavior intensity, $\mathbf{s}_t^{liq}$ represents liquidity conditions, $\mathbf{s}_t^{vol}$ represents volatility and downside-pressure states, and $\mathbf{s}_t^{info}$ represents the incorporation of recent market and policy-related information into trading behavior and prices. These states are observed indirectly through the constructed market-state variables, including calendar indicators, trading volume, trading value, illiquidity, intraday range, rolling volatility, rolling maximum loss, and recent return-based variables.

The system evolves through short-run state transitions. In abstract form, this can be written as

$$\mathbf{S}_t = \Phi(\mathbf{S}_{t-1}, \mathbf{x}_t, \boldsymbol{\varepsilon}_t), \quad (3)$$

where $\mathbf{x}_t$ is the observed market-state vector and $\boldsymbol{\varepsilon}_t$ represents unobserved market shocks. This transition equation is not estimated as a structural causal system. Instead, it provides a theoretical representation of how regulatory pressure, trading behavior, liquidity conditions,

volatility, and information adjustment can jointly evolve over time. The forecasting problem is then represented as

$$Q_{\alpha}(L_{t,h} | \mathcal{F}_t) = G_{\alpha,h}(\mathbf{S}_{t-w+1:t}), \quad (4)$$

where the conditional upper-tail quantile of future CEA loss depends on the recent sequence of carbon-market system states.

Figure 2 illustrates this theoretical system dynamics framework and its mapping to the IQTN architecture. The upper part of the figure summarizes the hypothesized risk-transmission pathway: regulatory and compliance pressure may induce trading concentration; trading concentration may generate liquidity stress when market depth is limited; liquidity stress may amplify intraday price fluctuations, rolling volatility, and downside pressure; and these stress signals may increase future tail risk. The figure also recognizes feedback loops. Higher volatility may reduce liquidity and further amplify price impact, while risk warnings and market stress may feed back into trading behavior, regulatory attention, and compliance responses.

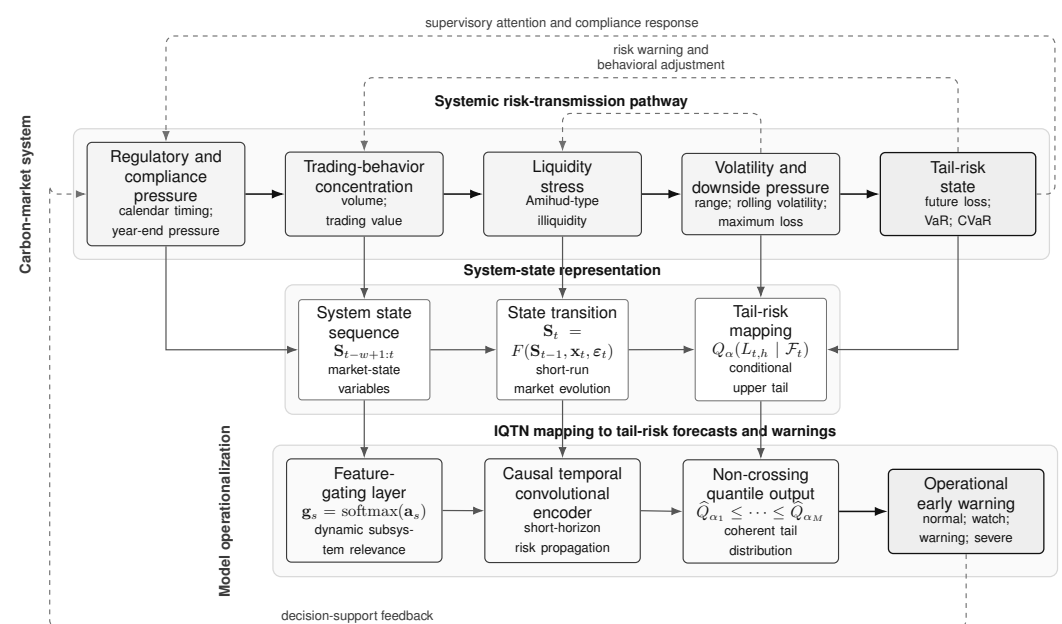


Figure 2. The theoretical system dynamics framework and its structural mapping to the proposed IQTN architecture. The upper panels illustrate the reinforcing and balancing feedback loops within the carbon-market risk system, while the lower panels show how the feature-gating mechanism and temporal convolutional encoder capture these dynamic state transitions.

The feature-gating mechanism provides the main link between this system framework and the IQTN model. Let $\mathcal{K}$ denote a subsystem-specific set of variables, such as liquidity variables or volatility-related variables. The subsystem-level gating relevance at time s can be summarized as

$$G_{\mathcal{K},s} = \sum_{j \in \mathcal{K}} g_{j,s}, \quad (5)$$

where $g_{j,s}$ is the feature-gating weight assigned to variable j at time s . A larger $G_{\mathcal{K},s}$ indicates that the corresponding subsystem is more important for the current tail-risk forecast. For example, high gating relevance for illiquidity and trading value indicates that liquidity and transaction concentration are important predictive drivers at that forecast origin, whereas high gating relevance for rolling volatility and maximum loss indicates that price stress and downside pressure dominate the forecast. Therefore, feature gating can be interpreted as a data-driven approximation of time-varying subsystem relevance

within the carbon market risk system. These weights should be interpreted as dynamic predictive importance rather than structural causal coefficients.

2.3. Research Hypotheses

Building on the systems-theoretic framing above, we formulate three mechanism-oriented predictive hypotheses. Because the purpose of this study is tail-risk forecasting and early warning rather than structural causal identification, these hypotheses are interpreted as testable predictive expectations about the relationship between market-state variables, model structure, and future carbon-market tail risk.

H1. *Recent volatility and downside-pressure variables are expected to have strong positive predictive relevance for short-horizon carbon-market tail risk. This hypothesis is derived from the idea that volatility clustering, intraday price instability, and recent realized losses contain information about the near-term upper tail of future losses. It corresponds to rolling volatility, rolling maximum loss, intraday range, absolute return, and squared return, and is evaluated through the feature-gating weights, temporal-importance results, and VaR/CVaR forecasting performance.*

H2. *Liquidity stress and trading concentration are expected to improve the prediction of multi-day carbon-market tail risk. This hypothesis follows the systemic risk-transmission mechanism in which concentrated allowance trading may increase liquidity pressure, price impact, and subsequent downside-risk accumulation. It corresponds to the Amihud-type illiquidity proxy, trading volume, trading value, and calendar-related trading-pressure variables, and is evaluated through the feature-gating results, local severe-warning explanations, and the ablation analysis that removes trading-activity and liquidity variables.*

H3. *A feature-gated temporal quantile architecture is expected to outperform ungated, non-temporal, and conventional benchmark models while providing interpretable information about risk-driving variables. This hypothesis is derived from interpretable machine-learning theory and the need for transparent risk monitoring in high-stakes regulatory settings. It corresponds to the feature-gating layer, causal temporal convolutional encoder, non-crossing quantile output layer, and early-warning conversion module and is evaluated through benchmark comparisons, Diebold–Mariano tests, ablation analysis, and global and local interpretability results.*

2.4. Data Processing and Feature Construction

The raw CEA trading data consist of daily open, high, low, close, trading volume, and trading value. Trading days with no official transaction but with a published closing price are retained because the closing price still represents the official market valuation for that day. Non-trading days, including weekends and holidays, are excluded and are not interpolated. Throughout the empirical analysis, the time index t refers to the ordered sequence of official CEA trading days rather than natural calendar days. Therefore, all rolling windows, input sequences, and forecast horizons are aligned by trading days. For example, a 10-day forecast horizon means 10 subsequent official trading days; when holidays occur, this horizon may span more than 10 natural calendar days, but the target remains aligned to the 10th subsequent trading day. All predictors are constructed using information available at or before trading day t , and all target variables are constructed using future trading-day prices only after the forecast origin has been fixed. This procedure prevents future information from entering the predictor matrix.

The forecast origin in this study is defined as the end of trading day t , after the official daily CEA trading information has become available. Therefore, the warning signal for horizon h is issued using the information set observed up to and including day t , and it forecasts the future loss from day t to day $t + h$. The daily open, high, low, close, trading

volume, and trading value are observable after market close on day t . Return, absolute return, squared return, intraday range, rolling volatility, rolling maximum loss, and the Amihud-type illiquidity proxy are all calculated only from day- t observations. Future losses $L_{t,h}$ are not used as predictors and are used only as supervised learning targets and out-of-sample evaluation outcomes. In addition, normalization parameters are estimated using the training set and then applied to validation and test observations, so the preprocessing step does not introduce future information. Thus, all predictor variables used by IQTN are observable at the forecast origin under an end-of-day risk-monitoring setting.

The daily log return is defined as

$$r_t = 100 \times \log\left(\frac{P_t}{P_{t-1}}\right), \quad (6)$$

where P_t is the official CEA closing price. The intraday log range is defined as

$$Range_t = 100 \times \log\left(\frac{High_t}{Low_t}\right), \quad (7)$$

where $High_t$ and Low_t denote the highest and lowest CEA trading prices on day t . The log trading volume and log trading value are defined as

$$LVOL_t = \log(1 + Volume_t), \quad (8)$$

and

$$LVAL_t = \log(1 + Value_t), \quad (9)$$

respectively. The constant one is added before logarithmic transformation to handle zero-volume days.

To describe short-term market stress, the squared return and absolute return are constructed as

$$RV_t = r_t^2, \quad (10)$$

and

$$AR_t = |r_t|. \quad (11)$$

Rolling volatility over a k -day window is calculated as

$$\sigma_{t,k} = \sqrt{\frac{1}{k-1} \sum_{i=t-k+1}^t (r_i - \bar{r}_{t,k})^2}, \quad (12)$$

where $k \in \{5, 10, 20\}$ and $\bar{r}_{t,k}$ is the mean return over the same rolling window. Rolling downside pressure is measured by the rolling maximum loss:

$$ML_{t,k} = \max_{i=t-k+1, \dots, t} (-r_i), \quad (13)$$

with $k \in \{5, 10, 20\}$. Market liquidity pressure is measured by an Amihud-type illiquidity proxy:

$$ILLIQ_t = 10^6 \times \frac{|r_t|}{Value_t + 1}. \quad (14)$$

The proxy is multiplied by 10^6 to improve numerical readability. Calendar variables are included because China's carbon allowance market has institutional trading cycles and compliance-related trading concentration. The model includes month-of-year indicators, day-of-week indicators, and a year-end trading-pressure indicator. The year-end indicator

equals one for trading days in November and December and zero otherwise. This setting is motivated by the compliance and allowance-surrender arrangements of China's national carbon market. Covered entities are required to surrender allowances corresponding to their verified emissions before the prescribed compliance deadline, and recent national ETS arrangements have placed surrender deadlines at or near the end of December [13,15]. Therefore, November and December are used to approximate a broad pre-deadline compliance window during which covered entities may review allowance balances, adjust trading positions, and prepare for fulfillment obligations. This window may be associated with more concentrated trading demand, greater transaction pressure, and stronger liquidity sensitivity. The year-end indicator should be interpreted as a parsimonious institutional-timing proxy rather than a precise structural measure of compliance pressure. A continuous variable such as the number of days remaining until the compliance deadline could provide a more refined measure when a longer sample and fully stable compliance calendar are available. However, given the short history of the national CEA market and the evolving compliance arrangements across cycles, imposing a precise countdown structure may introduce a false degree of timing precision. For this reason, the main specification uses the November–December indicator to capture the broad year-end compliance-pressure window. These calendar variables are treated as observed market-state variables rather than causal policy variables.

For each forecast origin t , the input matrix is

$$\mathbf{X}_{t-w+1:t} = [\mathbf{x}_{t-w+1}, \mathbf{x}_{t-w+2}, \dots, \mathbf{x}_t]^\top \in \mathbb{R}^{w \times p}, \quad (15)$$

where $\mathbf{x}_t \in \mathbb{R}^p$ is the feature vector on day t . The main specification uses a 20-trading-day lookback window, which represents approximately one trading month. All continuous predictors are standardized using the mean and standard deviation estimated from the training period only. The same training-period transformation is then applied to the validation and test periods.

2.5. Tail-Risk Target Definition

The forecasting target is future carbon market loss rather than future carbon price. Let d_t denote the t -th official CEA trading day in the ordered trading-day sequence. For each forecast horizon $h \in \{1, 5, 10\}$, the future log return is defined over trading days as

$$r_{t,h}^{TD} = 100 \times \log\left(\frac{P_{d_{t+h}}}{P_{d_t}}\right), \quad (16)$$

where P_{d_t} is the official CEA closing price on trading day d_t , and $P_{d_{t+h}}$ is the closing price on the h -th subsequent official trading day. Thus, $h = 1, 5, 10$ refer to 1, 5, and 10 trading days rather than natural calendar days. From the perspective of a market participant holding CEA positions, adverse price declines correspond to positive losses. The future carbon market loss is therefore defined as

$$L_{t,h} = -r_{t,h}^{TD}. \quad (17)$$

The conditional VaR at confidence level α is the conditional upper-tail quantile of $L_{t,h}$:

$$VaR_{\alpha,t,h} = Q_\alpha(L_{t,h} \mid \mathcal{F}_t), \quad (18)$$

where $\mathcal{F}_t$ denotes the information set available at the end of trading day t . The conditional CVaR is defined as the expected loss beyond the VaR threshold:

$$CVaR_{\alpha,t,h} = E[L_{t,h} \mid L_{t,h} \geq VaR_{\alpha,t,h}, \mathcal{F}_t]. \quad (19)$$

The main empirical analysis reports 95% VaR and 95% CVaR. The 90% and 99% VaR levels are used for robustness analysis. This target design follows the standard financial risk-management interpretation of VaR and CVaR [16–18].

2.6. Interpretable Quantile Temporal Network

2.6.1. Model Overview

The proposed model is named the Interpretable Quantile Temporal Network. IQTN combines three components: a dynamic feature-gating layer, a temporal convolutional encoder, and a non-crossing quantile output layer. The feature-gating layer assigns time-varying importance weights to carbon market predictors. The temporal convolutional encoder extracts short-term and medium-term risk patterns from the historical CEA trading sequence. The quantile output layer directly estimates the conditional quantile function of future losses.

The model estimates

$$\hat{Q}_{\alpha,t,h} = f_{\theta}(\mathbf{X}_{t-w+1:t}, \alpha, h), \quad (20)$$

where $\hat{Q}_{\alpha,t,h}$ is the predicted conditional quantile of $L_{t,h}$ at level α , and θ denotes all trainable parameters.

2.6.2. Methodological Distinction from Existing Components

Although IQTN draws on established ideas in feature gating, temporal convolution, and quantile estimation, its design differs from a simple additive combination of existing techniques in several ways. First, standard quantile regression models estimate conditional quantiles but usually do not learn an interpretable time-varying representation of market-state drivers. In IQTN, the feature-gating layer is linked to the carbon-market risk system by assigning dynamic weights to volatility, downside pressure, trading activity, liquidity, range, and calendar-related variables. The gating weights therefore provide model-attributed information about which market-state components are most relevant to tail-risk forecasts at a given time.

Second, standalone temporal convolutional networks are often used as generic sequence predictors. In IQTN, the causal temporal convolutional encoder is used specifically to capture short-horizon risk propagation from recent market stress to future tail losses. Because the convolution is causal, the representation at each forecast origin depends only on information available up to that point. This structure is consistent with the operational setting of real-time risk monitoring, in which future prices and future liquidity conditions cannot be observed when the warning signal is issued.

Third, the non-crossing quantile output layer is designed to produce a coherent conditional tail distribution rather than separate and potentially inconsistent point forecasts. This is important for VaR and CVaR estimation because higher confidence levels should correspond to weakly larger predicted losses. The complete IQTN framework therefore connects interpretable market-state weighting, causal temporal risk propagation, coherent multi-quantile tail estimation, and early-warning conversion. The subsequent ablation analysis evaluates these components separately, which further distinguishes the proposed framework from an arbitrary stacking of existing modules.

2.6.3. Dynamic Feature-Gating Layer

The feature-gating layer improves interpretability by assigning a normalized importance weight to each predictor at each time step. For each input vector $\mathbf{x}_s$, $s = t - w + 1, \dots, t$, the gating score is calculated as

$$\mathbf{a}_s = \mathbf{W}_a \phi(\mathbf{W}_z \mathbf{x}_s + \mathbf{b}_z) + \mathbf{b}_a, \quad (21)$$

where $\phi(\cdot)$ is the exponential linear unit activation function. The feature weight for variable j at time s is

$$g_{j,s} = \frac{\exp(a_{j,s})}{\sum_{m=1}^p \exp(a_{m,s})}. \quad (22)$$

The gated input is

$$\tilde{\mathbf{x}}_s = \mathbf{g}_s \odot \mathbf{x}_s, \quad (23)$$

where $\odot$ denotes element-wise multiplication. A larger $g_{j,s}$ indicates a stronger model-attributed contribution of variable j to the risk forecast at time s . The use of gating follows the principle of interpretable variable selection in neural forecasting systems [11].

2.6.4. Temporal Convolutional Encoder

The gated input sequence

$$\tilde{\mathbf{X}}_{t-w+1:t} = [\tilde{\mathbf{x}}_{t-w+1}, \dots, \tilde{\mathbf{x}}_t]^\top \quad (24)$$

is passed into a temporal convolutional encoder. The encoder uses causal dilated convolutions so that the representation at day t depends only on information observed on or before day t . The l -th temporal convolutional block is

$$\mathbf{H}^{(l)} = \mathbf{H}^{(l-1)} + \text{Dropout} \left[\text{ReLU} \left(\text{Conv}_{k,d_l} \left(\mathbf{H}^{(l-1)} \right) \right) \right], \quad (25)$$

where $\text{Conv}_{k,d_l}(\cdot)$ is a one-dimensional causal convolution with kernel size k and dilation factor d_l . The main specification uses three residual temporal convolutional blocks with dilation factors 1, 2, and 4, kernel size 3, hidden dimension 64, and dropout rate 0.20. The final temporal representation is the last-step hidden vector:

$$\mathbf{h}_t = \text{Encoder}_\theta \left(\tilde{\mathbf{X}}_{t-w+1:t} \right). \quad (26)$$

Temporal convolutional networks provide stable sequence modeling through causal convolutions, dilations, and residual connections [19].

2.6.5. Non-Crossing Quantile Output Layer

The output layer produces quantile forecasts for all horizons and quantile levels. The quantile grid is

$$\mathcal{A} = \{0.50, 0.55, 0.60, 0.65, 0.70, 0.75, 0.80, 0.85, 0.90, 0.95, 0.975, 0.99\}. \quad (27)$$

For each horizon h , the network first produces an unrestricted base quantile $\eta_{1,t,h}$ and positive quantile increments $\Delta_{m,t,h}$:

$$\Delta_{m,t,h} = \text{softplus}(\eta_{m,t,h}), \quad m = 2, \dots, |\mathcal{A}|. \quad (28)$$

The predicted quantiles are then constructed as

$$\hat{Q}_{\alpha_1,t,h} = \eta_{1,t,h}, \quad (29)$$

and

$$\hat{Q}_{\alpha_m,t,h} = \hat{Q}_{\alpha_{m-1},t,h} + \Delta_{m,t,h}, \quad m = 2, \dots, |\mathcal{A}|. \quad (30)$$

This structure enforces

$$\hat{Q}_{\alpha_1,t,h} \leq \hat{Q}_{\alpha_2,t,h} \leq \dots \leq \hat{Q}_{\alpha_{|\mathcal{A}|},t,h}, \quad (31)$$

and therefore prevents quantile crossing.

2.7. Loss Function and Model Training

The model is trained by minimizing the multi-horizon multi-quantile pinball loss. For a quantile level α , the pinball loss is

$$\rho_\alpha(u) = u[\alpha - \mathbf{1}(u < 0)], \quad (32)$$

where

$$u = L_{t,h} - \hat{Q}_{\alpha,t,h}. \quad (33)$$

The full training objective is

$$\mathcal{L}(\theta) = \frac{1}{N} \sum_{t=1}^N \sum_{h \in \mathcal{H}} \sum_{\alpha \in \mathcal{A}} \rho_\alpha(L_{t,h} - \hat{Q}_{\alpha,t,h}) + \lambda \|\theta\|_2^2, \quad (34)$$

where $\mathcal{H} = \{1, 5, 10\}$, $\mathcal{A}$ is the quantile grid, and $\lambda = 10^{-5}$ is the L_2 regularization coefficient. Quantile regression and pinball loss follow the standard quantile-regression framework [20,21].

The sample is divided chronologically. The training period is 16 July 2021 to 31 December 2024. The validation period is 1 January 2025 to 30 June 2025. The test period is 1 July 2025 to 30 January 2026. These calendar dates define the outer boundaries of the three partitions, but all forecast origins and labels are aligned by the official trading-day sequence. Forecast origins are assigned to each partition only when the full h -trading-day future label $L_{t,h}$ remains inside the corresponding partition. For example, a 10-day horizon requires the closing price on the 10th subsequent official trading day to fall within the same partition. This rule keeps the validation and test labels separated from the training objective and prevents horizon-induced leakage across partitions.

The IQTN is trained with the Adam optimizer. The learning rate is 10^{-3} , the batch size is 64, the maximum number of epochs is 500, and early stopping is applied when the validation loss does not improve for 30 consecutive epochs. The model checkpoint with the lowest validation pinball loss is retained for final testing.

The main implementation settings of IQTN are summarized in Table 1. The model uses a 20-trading-day input window, three forecast horizons, and a fixed quantile grid from 0.50 to 0.99. The temporal encoder consists of three residual causal convolutional blocks with dilation factors 1, 2, and 4. The non-crossing quantile output layer uses positive increments to ensure monotonicity across quantile levels. All neural network models are trained using the same chronological training, validation, and test partitions. Model selection is based only on validation-period pinball loss, and the test period is used only once for final out-of-sample evaluation. Normalization parameters are estimated from the training set and then applied to validation and test observations.

Table 1. Implementation settings of the proposed IQTN model.

Setting Category	Implementation Choice
Input window	20 trading days
Forecast horizons	1, 5, and 10 trading days
Quantile grid	Fixed grid from 0.50 to 0.99
Feature-gating layer	Softmax-normalized dynamic feature weights
Temporal encoder	Three residual causal TCN blocks
Dilation factors	1, 2, and 4
Kernel size	3
Hidden dimension	64
Dropout rate	0.20
Quantile output	Positive-increment construction for non-crossing quantiles
Training loss	Multi-horizon multi-quantile pinball loss
Optimizer	Adam
Learning rate	10^{-3}
Batch size	64
Maximum epochs	500
Early stopping	Patience of 30 validation epochs
Regularization	$L_2 = 10^{-5}$
Model selection	Lowest validation-period pinball loss
Standardization	Training-period mean and standard deviation only
Test-period use	Final out-of-sample evaluation only

2.8. CVaR Estimation and Early-Warning Signal Construction

The estimated VaR is directly obtained from the predicted quantile:

$$\widehat{VaR}_{\alpha,t,h} = \widehat{Q}_{\alpha,t,h}. \quad (35)$$

The estimated CVaR is calculated by numerical integration over the predicted upper-tail quantiles:

$$\widehat{CVaR}_{\alpha,t,h} = \frac{1}{1-\alpha} \int_{\alpha}^1 \widehat{Q}_{u,t,h} du. \quad (36)$$

In implementation, the integral is approximated by the trapezoidal rule over the available quantile grid points at or above α :

$$\widehat{CVaR}_{\alpha,t,h} \approx \frac{\sum_{m=1}^{M-1} \frac{\widehat{Q}_{u_m,t,h} + \widehat{Q}_{u_{m+1},t,h}}{2} (u_{m+1} - u_m)}{u_M - u_1}, \quad (37)$$

where $u_1 = \alpha$ and $u_M = 0.99$ for the reported 95% CVaR.

The early-warning score combines standardized VaR and CVaR:

$$RiskScore_{t,h} = 0.5 \times \widetilde{VaR}_{0.95,t,h} + 0.5 \times \widetilde{CVaR}_{0.95,t,h}, \quad (38)$$

where $\widetilde{VaR}_{0.95,t,h}$ and $\widetilde{CVaR}_{0.95,t,h}$ are standardized using the mean and standard deviation of the training-period forecasts. Four early-warning levels are defined using training-period score thresholds:

$$EW_{t,h} = \begin{cases} 0, & RiskScore_{t,h} \leq c_{0.50}, \\ 1, & c_{0.50} < RiskScore_{t,h} \leq c_{0.75}, \\ 2, & c_{0.75} < RiskScore_{t,h} \leq c_{0.90}, \\ 3, & RiskScore_{t,h} > c_{0.90}, \end{cases} \quad (39)$$

where $c_{0.50}$, $c_{0.75}$, and $c_{0.90}$ are the 50th, 75th, and 90th percentiles of the training-period risk score. The four warning levels are labelled as normal, watch, warning, and severe warning. Because all thresholds are estimated from the training period, the early-warning rule is fully

out-of-sample in the test period. The four warning levels should be interpreted as decision-support signals rather than automatic intervention rules. Their purpose is to translate model-implied tail-risk intensity into differentiated monitoring and risk-management actions for regulators, compliance entities, and trading institutions. The normal level indicates that the predicted VaR–CVaR risk score remains within the routine range observed in the training period. The watch level indicates early accumulation of tail-risk pressure and calls for closer monitoring of volatility, trading concentration, and liquidity conditions. The warning level indicates an elevated tail-risk state and should trigger intensified risk review, stress testing, and adjustment of internal risk limits. The severe-warning level indicates that the predicted tail-risk score is in the most extreme training-period range and should trigger focused supervisory attention and conservative risk-management responses. Table 2 summarizes the operational interpretation of the four warning levels.

Table 2. Operational interpretation of the four-level early-warning signals.

Warning Level	Risk-Score Condition	Regulatory and Market-Surveillance Response	Compliance-Entity and Trading-Institution Response
Normal	$RiskScore_{t,h} \leq c_{0.50}$	Maintain routine market monitoring and standard surveillance of price, volume, and liquidity indicators.	Maintain regular exposure management, standard position limits, and routine VaR/CVaR reporting.
Watch	$c_{0.50} < RiskScore_{t,h} \leq c_{0.75}$	Increase monitoring frequency and review whether volatility, trading concentration, or liquidity pressure is beginning to accumulate.	Review allowance exposure, hedging needs, compliance demand, and short-term liquidity arrangements.
Warning	$c_{0.75} < RiskScore_{t,h} \leq c_{0.90}$	Conduct enhanced risk surveillance, with particular attention to abnormal volatility, concentrated trading, and liquidity deterioration.	Tighten internal risk limits, conduct VaR/CVaR stress testing, and consider earlier compliance or hedging adjustments.
Severe warning	$RiskScore_{t,h} > c_{0.90}$	Initiate focused risk review, consider issuing risk communication, and evaluate targeted market-stabilization measures within the applicable regulatory framework.	Reduce discretionary risk exposure, prepare liquidity buffers, reassess trading strategies, and conduct adverse-scenario stress testing.

2.9. Benchmark Models

The proposed IQTN is compared with nine benchmark models. The benchmark set covers traditional financial risk models, machine-learning quantile models, and deep temporal quantile models.

The traditional VaR benchmarks include Historical Simulation VaR, EWMA VaR, GARCH-t VaR, and GJR-GARCH-t VaR. Historical Simulation VaR estimates the empirical loss quantile using a 120-trading-day rolling window. EWMA VaR estimates time-varying volatility using the RiskMetrics decay parameter $\lambda = 0.94$ [22]. GARCH-t VaR models conditional volatility with Student- t innovations using the standard GARCH framework [23]. GJR-GARCH-t VaR adds an asymmetric response to negative return shocks and captures leverage-type volatility asymmetry [24]. GARCH-type models are re-estimated with expanding windows at each test forecast origin. The machine-learning benchmarks include Linear Quantile Regression, Quantile Random Forest, and LightGBM-Quantile. Linear Quantile Regression estimates the conditional quantile of future loss using the same predictor set as IQTN but without nonlinear temporal representation [20]. Quantile Random Forest estimates conditional quantiles from the distribution of responses in terminal leaves and uses 500 trees [25]. LightGBM-Quantile uses gradient-boosted decision trees with quantile loss, 500 boosting iterations, a learning rate of 0.03, and 31 leaves [26]. The deep temporal benchmarks include LSTM-Quantile and TCN-Quantile. LSTM-Quantile uses

two LSTM layers with a hidden dimension of 64 and the same quantile output grid as IQTN [27]. TCN-Quantile uses the same temporal convolutional encoder as IQTN but removes the feature-gating layer. This benchmark isolates the contribution of the interpretable gating component.

2.10. Evaluation Strategy

2.10.1. Quantile Forecasting Accuracy

Quantile forecasting accuracy is evaluated by the average pinball loss:

$$QPL_{\alpha,h} = \frac{1}{N_{test}} \sum_{t=1}^{N_{test}} \rho_{\alpha} \left(L_{t,h} - \widehat{Q}_{\alpha,t,h} \right). \quad (40)$$

Lower $QPL_{\alpha,h}$ indicates more accurate conditional quantile forecasting. The main comparison uses $\alpha = 0.95$ and horizons $h \in \{1, 5, 10\}$.

2.10.2. VaR Backtesting

VaR exceedance is defined as

$$I_{\alpha,t,h} = \mathbf{1} \left(L_{t,h} > \widehat{VaR}_{\alpha,t,h} \right). \quad (41)$$

The empirical exceedance rate is

$$ER_{\alpha,h} = \frac{1}{N_{test}} \sum_{t=1}^{N_{test}} I_{\alpha,t,h}. \quad (42)$$

For a correctly calibrated α -level VaR forecast, $ER_{\alpha,h}$ equals $1 - \alpha$. VaR calibration is assessed using the unconditional coverage test [28], the independence and conditional coverage tests [29], and the dynamic quantile test [30]. These tests evaluate whether VaR exceedances occur at the correct frequency and whether exceedances are serially independent.

2.10.3. CVaR Evaluation

CVaR forecasts are evaluated on VaR exceedance days. Let

$$\mathcal{E}_{\alpha,h} = \{t : L_{t,h} > \widehat{VaR}_{\alpha,t,h}\}. \quad (43)$$

The exceedance-tail absolute error is

$$ETA E_{\alpha,h} = \frac{1}{|\mathcal{E}_{\alpha,h}|} \sum_{t \in \mathcal{E}_{\alpha,h}} \left| L_{t,h} - \widehat{CVaR}_{\alpha,t,h} \right|. \quad (44)$$

Lower $ETA E_{\alpha,h}$ indicates more accurate conditional tail-severity forecasting. Joint VaR–CVaR forecast quality is also compared using consistent scoring principles for jointly elicitable VaR and expected shortfall [31,32].

2.10.4. Early-Warning Performance

Early-warning performance is evaluated by comparing the predicted warning signal with realized extreme-loss events. A realized extreme-loss event is defined by the training-period 95th percentile of future loss:

$$Y_{t,h} = \mathbf{1} \left(L_{t,h} > q_{0.95,h}^{train} \right), \quad (45)$$

where $q_{0.95,h}^{train}$ is calculated using only training-period losses for horizon h . The predicted binary warning indicator is

$$\hat{Y}_{t,h} = \mathbf{1}(EW_{t,h} \geq 2), \quad (46)$$

where levels 2 and 3 correspond to warning and severe warning. Performance is measured by precision, recall, F1-score, false alarm rate, missed alarm rate, area under the receiver operating characteristic curve, and area under the precision–recall curve. Recall, missed alarm rate, and precision–recall performance are treated as the primary early-warning metrics because extreme carbon market losses are rare events.

2.10.5. Statistical Significance and Uncertainty Analysis

To evaluate whether the predictive-performance gains of IQTN over benchmark models are statistically meaningful, we conducted pairwise Diebold–Mariano tests based on out-of-sample quantile loss differences. For each forecast horizon, the loss differential sequence was defined as

$$d_t = QPL_t^{benchmark} - QPL_t^{IQTN}, \quad (47)$$

where a positive value indicates that IQTN produced a lower pinball loss than the benchmark model at forecast origin t . The null hypothesis of the Diebold–Mariano test is equal predictive accuracy between IQTN and the benchmark model. Because multi-day loss forecasts may contain serial dependence due to overlapping forecast horizons, heteroskedasticity- and autocorrelation-consistent standard errors were used when computing the test statistic.

We also used a moving-block bootstrap to quantify uncertainty in the main out-of-sample performance metrics. The bootstrap resampled consecutive blocks of test-period forecast origins to preserve local temporal dependence. For each bootstrap replication, the main evaluation metrics were computed, including the average pinball loss, VaR exceedance rate, CVaR exceedance-tail absolute error, AUPRC, and F1-score. The reported 95% confidence intervals correspond to the 2.5th and 97.5th percentiles of the empirical bootstrap distribution based on 1000 replications. This procedure provides an uncertainty assessment for both predictive accuracy and early-warning performance.

2.11. Interpretability Analysis

The interpretability analysis has three components. First, the average feature-gating weight is computed for each predictor:

$$\bar{g}_j = \frac{1}{N_{test}w} \sum_{t=1}^{N_{test}} \sum_{s=t-w+1}^t g_{j,s}. \quad (48)$$

This statistic identifies the CEA market variables that contribute most strongly to tail-risk forecasts in the test period.

Second, temporal importance is summarized by lag-specific gating intensity:

$$TI_\ell = \frac{1}{N_{test}p} \sum_{t=1}^{N_{test}} \sum_{j=1}^p g_{j,t-\ell}, \quad \ell = 0, 1, \dots, w-1. \quad (49)$$

This statistic shows whether the model places more weight on immediate market movements or earlier observations within the 20-day input window.

Third, local event explanations are produced for all severe-warning days in the test period. For each severe-warning day, the model reports the top five predictors ranked by their gating weights at the forecast origin. These explanations are interpreted as model-attributed predictive contributions rather than causal effects.

The reported feature-gating weights are calculated from the final validation-selected IQTN model checkpoint, which is the same trained model used to generate the out-of-sample VaR, CVaR, and early-warning forecasts. They are not averaged across multiple independently trained neural networks. This choice ensures that the interpretability results are directly aligned with the exact forecasting model used in the empirical evaluation. However, because neural-network attribution values may be affected by random initialization and optimization paths, the gating weights should be interpreted as model-attributed predictive relevance from the selected forecasting model rather than as invariant structural parameters or causal effects. To reduce over-interpretation of individual attribution values, the global feature-gating weights are averaged over all test-period forecast origins and input-window time steps, and local severe-warning explanations are reported as ranked predictive drivers rather than causal mechanisms. The use of a single validation-selected model therefore does not affect the out-of-sample predictive comparison, but it means that the interpretability results should be understood as model-specific explanatory evidence. Future applications may further assess attribution stability by averaging explanations across independently trained ensembles.

2.12. Robustness and Ablation Analysis

Three robustness and ablation analyses are conducted. First, alternative risk settings are evaluated by reporting 90%, 95%, and 99% VaR across the 1-day, 5-day, and 10-day horizons. This analysis verifies that the model performance is stable across moderate, high, and extreme tail-risk levels. Second, alternative input-window lengths of 10 and 30 trading days are compared with the main 20-day window. This analysis evaluates whether the results depend on a single historical lookback length. To clarify the marginal contribution of each IQTN component, we conduct a module-level ablation analysis. Four ablation variants are considered. The first removes the feature-gating layer and passes the standardized market-state sequence directly into the temporal encoder. This tests the value of dynamic market-state weighting and feature-level interpretability. The second removes the causal temporal convolutional encoder and replaces it with a non-temporal feedforward representation of the current market-state vector. This tests the value of short-run temporal risk-propagation learning. The third removes the non-crossing quantile output layer and allows each quantile level to be predicted independently. This tests whether the monotonic quantile construction improves the coherence of tail-distribution estimates and prevents quantile crossing. The fourth removes trading activity and liquidity variables from the input state vector. This tests whether transaction intensity and liquidity pressure provide incremental information for carbon-market tail-risk forecasting. All ablation variants use the same training, validation, and test partitions as the main model. To address the possibility that the results are driven by a particular chronological sample split, we further conduct alternative sample-split and rolling-origin robustness tests. Random sample splitting is not used because it would break the temporal ordering of the forecasting problem and could introduce look-ahead bias. In the alternative-split analysis, the training, validation, and testing cutoffs are shifted while preserving the chronological order of observations. The earlier test split advances the test partition boundary by six months, whereas the later test split evaluates a more recent high-stress window. In the rolling-origin analysis, the model is repeatedly re-estimated using an expanding training window and evaluated on the next non-overlapping forward test block. The rolling-origin mean and standard deviation summarize average predictive performance and temporal variability across evaluation windows. This design assesses whether the predictive advantage of IQTN remains stable across different evaluation periods rather than being specific to one fixed test sample.

2.13. Lead-Lag Diagnostics and Predictive Endogeneity Considerations

Trading behavior, liquidity conditions, volatility, and downside pressure may be jointly determined in the carbon-market system. For example, volatility may affect trading activity, while concentrated trading and liquidity stress may also amplify subsequent volatility. Because the objective of this study is out-of-sample tail-risk forecasting rather than structural causal estimation, the model does not interpret feature-gating weights or predictive contributions as causal effects. Instead, they are interpreted as dynamic predictive relevance conditional on the information available at the forecast origin.

To examine temporal relationships among key market-state variables, we conducted pairwise Granger causality diagnostics using trading-day aligned series. The tests focused on rolling volatility, trading value, and illiquidity because these variables correspond to the volatility, trading-behavior, and liquidity subsystems in the proposed risk-transmission framework. The Granger tests are interpreted as lead-lag predictive diagnostics rather than structural causal evidence. They indicate whether past values of one market-state variable contain incremental predictive information for another market-state variable.

We did not employ an instrumental-variable design as a main analysis because the target of the study is the conditional upper-tail distribution of future losses rather than a structural causal coefficient. In addition, potential instruments such as external energy prices or policy-event dummies may directly affect carbon prices, trading expectations, and tail risk, making the exclusion restriction difficult to justify. Therefore, IV-based causal identification is better suited to future research designed specifically for causal inference. In the present study, the main objective is to ensure temporal ordering, avoid look-ahead bias, and provide robust out-of-sample predictive evidence.

2.14. Mechanism-Oriented Predictive Pathway Analysis

To examine whether the important variables identified by the feature-gating analysis are connected through a risk-transmission pathway, we conducted a mechanism-oriented predictive pathway analysis. This analysis is designed as a predictive diagnostic rather than a causal mediation or structural equation model. The purpose is to evaluate whether the observed temporal relationships are consistent with the hypothesized pathway in which volatility or trading concentration is linked to future tail risk through liquidity stress.

We considered two pathways: volatility $\rightarrow$ illiquidity $\rightarrow$ future loss and trading value $\rightarrow$ illiquidity $\rightarrow$ future loss. For each pathway, the lagged predictor X_{t-1} was first related to current illiquidity $ILLIQ_t$, and current illiquidity was then related to future loss $L_{t,h}$ for horizons $h = 1, 5, 10$. The indirect predictive effect was calculated as the product of the path from X_{t-1} to $ILLIQ_t$ and the path from $ILLIQ_t$ to $L_{t,h}$. HAC-robust standard errors were used for the path coefficients, and 95% confidence intervals for the indirect predictive effects were obtained from 1000 moving-block bootstrap replications.

This analysis does not establish structural causality. Instead, it provides mechanism-oriented predictive evidence on whether the proposed risk-transmission logic is empirically consistent with the data. A positive indirect predictive effect indicates that the lagged predictor is associated with current liquidity stress and that this liquidity stress is in turn associated with higher future tail loss.

2.15. Policy-Window and Market-Phase Stability Diagnostics

To assess whether model performance changes during policy-sensitive periods, we conducted a policy-window stability diagnostic. We defined a policy-announcement-window indicator, I_t^{policy} , which equals one when the forecast origin falls within five trading days before or after major regulatory releases and zero otherwise. This window is designed to capture periods in which policy information may affect market expectations,

trading behavior, liquidity conditions, and carbon-price risk. We then computed conditional pinball loss separately for policy announcement windows and normal trading windows.

The diagnostic compares IQTN with GJR-GARCH-t VaR, the strongest traditional risk-model benchmark in the main quantile-loss comparison. Within each window type, we computed the 95% VaR pinball loss for the 1-day, 5-day, and 10-day horizons and used Diebold–Mariano statistics to assess whether the predictive accuracy of IQTN differed from that of GJR-GARCH-t VaR. This analysis is interpreted as a predictive stability diagnostic rather than a causal estimate of policy effects.

We did not include a pilot-phase versus nationwide-phase moderator in the main analysis because the study is deliberately restricted to China's national CEA market. Regional pilot markets and the national market differ in sectoral coverage, allocation rules, local trading arrangements, and liquidity conditions. Pooling pilot-market and national-market observations would mix institutionally different trading systems and weaken the internal consistency of the national-market risk-monitoring framework. Therefore, we assess stability within the national market by comparing policy announcement windows and normal trading windows.

2.16. Heterogeneity Analysis Across Market Conditions

To evaluate whether model performance is stable across heterogeneous market conditions, we conducted grouped performance comparisons. Two grouping schemes were considered. First, the evaluated out-of-sample observations were divided into high-volatility and low-volatility periods using the median of rolling volatility within the evaluated out-of-sample period. Second, near-compliance periods were compared with normal trading periods. Near-compliance periods were defined as trading days in November and December, consistent with the year-end compliance-pressure window used in the main feature construction.

For each group and forecast horizon, we computed the 95% VaR pinball loss, the empirical VaR exceedance rate, and AUPRC. The analysis compares IQTN with GJR-GARCH-t VaR because GJR-GARCH-t is the strongest traditional risk-model benchmark. We used 1000 moving-block bootstrap replications for statistical inference. The model-comparison bootstrap *p*-value evaluates whether IQTN significantly reduces QPL relative to GJR-GARCH-t within the same group. The group-difference bootstrap *p*-value tests whether IQTN's QPL differs across the two market environments in each panel.

3. Results

3.1. Descriptive Characteristics of the National Carbon Allowance Market

Table 3 reports the descriptive statistics of the main CEA market variables. The average CEA closing price was 74.38 RMB yuan per ton, with a standard deviation of 18.62. The closing price ranged from 41.26 to 118.40, indicating substantial price variation during the sample period. The mean closing price was slightly higher than the median value of 71.15, and the skewness was 0.42, showing a mildly right-skewed price distribution. The daily return was centered close to zero, with a mean of 0.08% and a median of 0.00%. However, the return distribution displayed clear downside risk, with a minimum daily return of −9.87% and a negative skewness of −0.37. The absolute return and squared return were strongly right-skewed, with skewness values of 3.12 and 7.45, respectively. These results indicate that carbon market volatility was concentrated in a limited number of high-fluctuation trading days. Trading activity also exhibited strong heterogeneity. The average trading volume was 982,541 tons, while the median was only 150,240 tons. Similarly, the mean trading value was 65.42 million RMB yuan, much higher than the median value of 9.25 million RMB yuan. The large gap between the mean and median, together with

the high skewness of trading volume and trading value, indicates that market activity was concentrated in a small number of highly active trading days.

The intraday range and Amihud illiquidity proxy were also right-skewed, with skewness values of 2.58 and 8.24, respectively. This suggests that both intraday price fluctuation and liquidity pressure were episodic rather than constant. Overall, the descriptive evidence shows that the national carbon allowance market was characterized by time-varying volatility, occasional large losses, uneven trading activity, and episodic liquidity stress. These features justify the use of dynamic VaR, CVaR, and early-warning models for carbon tail-risk forecasting. To address the relatively short history of the national CEA market, we further interpret the sample by calendar-year segment, as summarized in Table 4. This comparison should be read as an institutional-period comparison rather than a long historical regime analysis. The 2021 observations correspond to the launch and initial price-discovery stage of the national CEA market after online trading began in July. The year 2022 represents the first full calendar year of national-market operation and provides the first complete annual trading cycle for learning basic return, volatility, trading activity, and liquidity patterns. The years 2023–2024 represent subsequent consolidation years in which more national-market information became available for estimating recurring market-state relationships. The year 2025 is especially important for model validation and out-of-sample testing because it provides a more recent market environment for evaluating whether the risk model generalizes beyond the main training period. The January 2026 observations are retained only as the terminal part of the most recent test window. Across these calendar-year segments, the key empirical pattern is not a smooth annual trend but the repeated emergence of episodic volatility, concentrated trading activity, and liquidity-pressure episodes. This reinforces the need for a dynamic model that updates tail-risk forecasts using recent market-state information rather than relying on a static full-sample risk measure.

Table 3. Descriptive statistics of CEA market variables.

Variable	Mean	SD	Min	Median	Max	Skewness
Closing price	74.38	18.62	41.26	71.15	118.40	0.42
Daily return	0.08	1.45	−9.87	0.00	9.92	−0.37
Absolute return	0.89	1.15	0.00	0.41	9.92	3.12
Squared return	2.11	6.84	0.00	0.16	98.41	7.45
Trading volume	982,541	2,154,308	10	150,240	14,508,900	4.18
Trading value	65,423,800	145,182,700	512	9,250,400	1,154,200,000	5.03
Intraday range	1.42	1.38	0.00	1.05	10.21	2.58
Illiquidity proxy ($\times 10^6$)	28.65	135.42	0.00	0.04	1937.50	8.24

Note: Closing price is measured in RMB yuan per ton. Daily return, absolute return, squared return, and intraday range are reported in percentage units. Trading volume is measured in tons, and trading value is measured in RMB yuan. The Amihud illiquidity proxy is calculated as $10^6 \times |r_t| / (Value_t + 1)$.

Table 4. Year-by-year interpretation of the national CEA sample.

Year	Role in the Sample	Interpretation for Risk Monitoring
2021	Launch segment	Partial year after the start of national online trading; reflects initial price discovery.
2022	First, full trading year	Provides the first complete annual cycle for learning basic market-state patterns.
2023	Consolidation year	Adds subsequent national-market observations for volatility, trading, and liquidity dynamics.
2024	Final training year	Strengthens in-sample learning before validation and out-of-sample testing.
2025	Validation and test year	Provides the main recent out-of-sample environment for evaluating forecasting performance.
2026	Terminal test segment	January observations are retained as the most recent endpoint, not as a full annual regime.

3.2. Quantile Forecasting Accuracy

Table 5 reports the average pinball loss for 95% VaR forecasting across the 1-day, 5-day, and 10-day horizons. The proposed IQTN achieved the lowest forecasting loss at all

horizons, with QPL values of 0.1765, 0.3958, and 0.5732, respectively. The forecasting error increased with the prediction horizon for all models, confirming that multi-day carbon losses were more difficult to forecast than one-day losses. Among the traditional risk models, GJR-GARCH-t VaR performed best, with QPL values of 0.2037, 0.4604, and 0.6653 across the three horizons. This indicates that asymmetric volatility dynamics improved tail-risk estimation relative to Historical Simulation VaR, EWMA VaR, and standard GARCH-t VaR. Among the machine-learning models, LightGBM-Quantile achieved stronger performance than Linear Quantile Regression and Quantile Random Forest. Among the deep temporal models, TCN-Quantile outperformed LSTM-Quantile, showing the advantage of causal temporal convolution for modeling short-term carbon market risk. Compared with TCN-Quantile, IQTN reduced the pinball loss by 4.23%, 4.90%, and 5.21% for the 1-day, 5-day, and 10-day horizons, respectively. Compared with LightGBM-Quantile, the corresponding reductions were 6.91%, 7.54%, and 8.42%. These results show that the feature-gating and temporal quantile structure of IQTN provided stable gains beyond both tree-based quantile learning and standard temporal convolution. Figure 3 visualizes the realized future losses and IQTN-predicted 95% VaR during the out-of-sample test period. The predicted VaR increased markedly from November 2025 to early January 2026, when realized losses became more volatile. The VaR threshold also scaled upward from the 1-day to the 10-day horizon, consistent with the larger uncertainty of longer-horizon cumulative losses. The exceedance points were sparse and did not show persistent clustering, providing visual support for the formal VaR backtesting reported in the next subsection.

Table 5. Quantile forecasting accuracy measured by average pinball loss at the 95% VaR level.

Model	1-Day Horizon	5-Day Horizon	10-Day Horizon
Historical Simulation VaR	0.2418	0.5482	0.7715
EWMA VaR	0.2253	0.5019	0.7168
GARCH-t VaR	0.2114	0.4781	0.6892
GJR-GARCH-t VaR	0.2037	0.4604	0.6653
Linear Quantile Regression	0.2189	0.4827	0.6974
Quantile Random Forest	0.1982	0.4415	0.6481
LightGBM-Quantile	0.1896	0.4281	0.6259
LSTM-Quantile	0.1945	0.4356	0.6384
TCN-Quantile	0.1843	0.4162	0.6047
IQTN	0.1765	0.3958	0.5732

3.3. Economic Interpretation of Quantile-Loss Improvements

The pinball loss is a proper statistical scoring rule for quantile forecasts, but its numerical value may be difficult to interpret economically. To provide an intuitive scale for the improvement, we compare IQTN with GJR-GARCH-t VaR, which is the strongest traditional risk-model benchmark in Table 5. For each horizon, the improvement in pinball loss is defined as

$$\Delta QPL_h = QPL_h^{GJR} - QPL_h^{IQTN}.$$

Because future losses are measured in percentage points, this difference can be scaled by the monetary value of a representative allowance position. For an allowance price P and a position size A , the scaled VaR scoring-loss reduction is

$$\text{Scaled reduction}_h = P \times A \times \frac{\Delta QPL_h}{100}.$$

This calculation does not represent realized trading profit or realized loss reduction. Rather, it provides an economic scaling of the reduction in VaR-related quantile scoring loss.

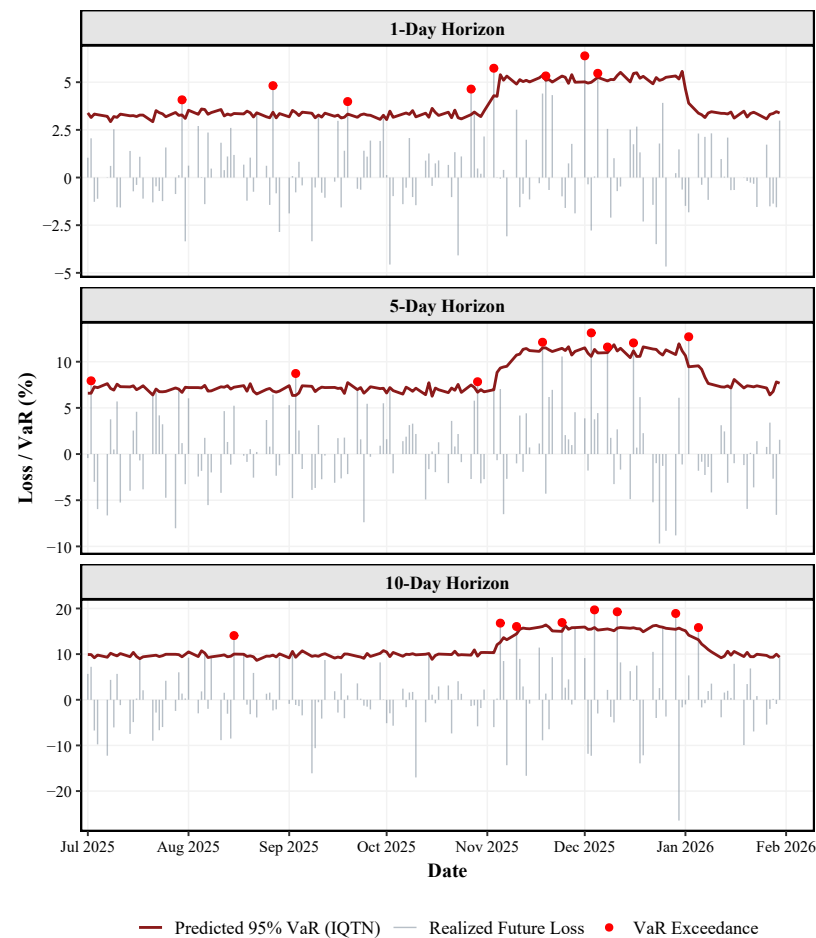


Figure 3. Realized future loss and IQTN-predicted 95% VaR across the 1-day, 5-day, and 10-day horizons during the out-of-sample test period. Red dots indicate VaR exceedance events.

Using an illustrative allowance position of 10,000 tons and a representative price of RMB 100 per ton, IQTN reduces the 95% VaR pinball loss relative to GJR-GARCH-t VaR by 13.35%, 14.03%, and 13.84% across the 1-day, 5-day, and 10-day horizons, respectively. Table 6 summarizes the corresponding monetary scaling. The scaled reductions are RMB 272, RMB 646, and RMB 921 per forecast origin for the three horizons. These values provide a practical interpretation of the magnitude of the forecasting improvement: even small reductions in quantile loss may become economically meaningful when applied to large carbon-allowance positions or repeated daily risk-management decisions.

Table 6. Economic scaling of QPL improvements relative to GJR-GARCH-t VaR.

Horizon	GJR-GARCH-t QPL	IQTN QPL	Relative Reduction	Scaled Reduction
1-day	0.2037	0.1765	13.35%	RMB 272
5-day	0.4604	0.3958	14.03%	RMB 646
10-day	0.6653	0.5732	13.84%	RMB 921

Note: The scaled reduction is calculated as $P \times A \times \Delta QPL_h / 100$, where P = RMB 100 per ton, A = 10,000 tons, and $\Delta QPL_h = QPL_h^{GJR} - QPL_h^{IQTN}$. The values represent monetary scaling of VaR scoring-loss improvement, not realized trading profit.

3.4. VaR Backtesting Results

Table 7 reports the 95% VaR backtesting results. The proposed IQTN produced empirical exceedance rates of 5.23%, 6.04%, and 6.94% for the 1-day, 5-day, and 10-day horizons, respectively. These values were the closest to the nominal exceedance probability of 5%

among all models. In contrast, Historical Simulation VaR generated much higher exceedance rates of 9.80%, 10.74%, and 12.50%, indicating systematic underestimation of downside tail risk. The diagnostic test results further confirm the calibration advantage of IQTN. For the 1-day horizon, IQTN obtained UC, IND, CC, and DQ p -values of 0.894, 0.765, 0.948, and 0.810, respectively. All tests failed to reject the correct VaR specification. The same pattern remained at the 5-day and 10-day horizons, where all IQTN p -values remained above conventional significance levels. This shows that IQTN produced VaR thresholds with accurate exceedance frequency and no evidence of excessive exceedance clustering. The benchmark models displayed weaker backtesting performance. Historical Simulation VaR was rejected by all tests at all horizons. EWMA VaR also failed most independence, conditional coverage, and dynamic quantile tests, especially at the 5-day and 10-day horizons. GARCH-t and GJR-GARCH-t improved coverage relative to historical and EWMA methods, but their exceedance rates remained higher than those of IQTN. Machine learning and deep learning quantile models further improved calibration, with TCN-Quantile performing best among the benchmarks. Nevertheless, IQTN consistently achieved the lowest exceedance rate deviation and the strongest overall backtesting performance. The backtesting results are consistent with the pinball-loss results in Table 5. IQTN not only minimized quantile forecasting error but also produced statistically reliable VaR forecasts in the out-of-sample test period. These findings confirm that the proposed model provides better-calibrated tail-risk thresholds for China's national carbon allowance market.

Table 7. VaR backtesting results at the 95% confidence level.

Model	Horizon	ER	UC p -Value	IND p -Value	CC p -Value	DQ p -Value
Historical Simulation VaR	1-day	9.80%	0.015	0.008	0.002	0.002
EWMA VaR	1-day	7.84%	0.136	0.035	0.036	0.015
GARCH-t VaR	1-day	7.19%	0.246	0.088	0.119	0.042
GJR-GARCH-t VaR	1-day	6.54%	0.416	0.152	0.258	0.095
Linear Quantile Regression	1-day	7.84%	0.136	0.095	0.082	0.065
Quantile Random Forest	1-day	7.19%	0.246	0.165	0.196	0.115
LightGBM-Quantile	1-day	6.54%	0.416	0.210	0.328	0.144
LSTM-Quantile	1-day	6.54%	0.416	0.320	0.440	0.255
TCN-Quantile	1-day	5.88%	0.644	0.450	0.677	0.388
IQTN	1-day	5.23%	0.894	0.765	0.948	0.810
Historical Simulation VaR	5-day	10.74%	0.005	0.002	<0.001	0.001
EWMA VaR	5-day	8.72%	0.056	0.012	0.007	0.005
GARCH-t VaR	5-day	8.05%	0.113	0.045	0.038	0.018
GJR-GARCH-t VaR	5-day	7.38%	0.213	0.110	0.129	0.062
Linear Quantile Regression	5-day	8.72%	0.056	0.055	0.026	0.028
Quantile Random Forest	5-day	8.05%	0.113	0.095	0.071	0.065
LightGBM-Quantile	5-day	7.38%	0.213	0.145	0.160	0.095
LSTM-Quantile	5-day	7.38%	0.213	0.210	0.210	0.145
TCN-Quantile	5-day	6.71%	0.358	0.280	0.368	0.210
IQTN	5-day	6.04%	0.576	0.550	0.719	0.495
Historical Simulation VaR	10-day	12.50%	<0.001	<0.001	<0.001	<0.001
EWMA VaR	10-day	10.42%	0.009	0.004	<0.001	0.001
GARCH-t VaR	10-day	9.03%	0.044	0.018	0.008	0.008
GJR-GARCH-t VaR	10-day	8.33%	0.092	0.055	0.039	0.035
Linear Quantile Regression	10-day	9.72%	0.020	0.025	0.005	0.012
Quantile Random Forest	10-day	9.03%	0.044	0.045	0.018	0.025
LightGBM-Quantile	10-day	8.33%	0.092	0.065	0.044	0.038
LSTM-Quantile	10-day	8.33%	0.092	0.095	0.060	0.075
TCN-Quantile	10-day	7.64%	0.173	0.125	0.122	0.082
IQTN	10-day	6.94%	0.308	0.315	0.362	0.260

Note: ER denotes the empirical exceedance rate. UC, IND, CC, and DQ denote the p -values of the unconditional coverage, independence, conditional coverage, and dynamic quantile tests, respectively. The number of valid forecast origins decreases across horizons ($N_1 = 153$, $N_5 = 149$, $N_{10} = 144$).

3.5. CVaR Forecasting Performance

Table 8 reports the CVaR forecasting performance measured by exceedance-tail absolute error. The proposed IQTN achieved the lowest CVaR forecasting error across all three horizons. The errors were 0.3784, 0.7452, and 1.1540 for the 1-day, 5-day, and 10-day horizons, respectively. The forecasting error increased with the horizon, reflecting the greater uncertainty of multi-day cumulative losses. The benchmark results show a pattern consistent with the VaR forecasting and backtesting results. Among traditional models, GJR-GARCH-t VaR produced the lowest CVaR error, outperforming Historical Simulation VaR, EWMA VaR, and GARCH-t VaR. Among machine-learning models, LightGBM-Quantile performed better than Linear Quantile Regression and Quantile Random Forest. Among deep temporal models, TCN-Quantile outperformed LSTM-Quantile, confirming the advantage of temporal convolution for carbon tail-risk modelling. Compared with TCN-Quantile, IQTN reduced the CVaR forecasting error by 8.95%, 9.56%, and 10.23% for the 1-day, 5-day, and 10-day horizons, respectively. Compared with LightGBM-Quantile, the corresponding reductions were 14.49%, 14.87%, and 15.43%. These results indicate that IQTN improved not only VaR threshold estimation but also the prediction of tail-loss severity beyond the VaR exceedance threshold.

Table 8. CVaR forecasting performance measured by exceedance-tail absolute error.

Model	1-Day	5-Day	10-Day
Historical Simulation VaR	0.6542	1.2540	1.8755
EWMA VaR	0.5821	1.1052	1.6548
GARCH-t VaR	0.5318	1.0215	1.5420
GJR-GARCH-t VaR	0.4955	0.9450	1.4585
Linear Quantile Regression	0.5512	1.0450	1.5720
Quantile Random Forest	0.4780	0.9125	1.4150
LightGBM-Quantile	0.4425	0.8754	1.3645
LSTM-Quantile	0.4580	0.8865	1.3920
TCN-Quantile	0.4156	0.8240	1.2855
IQTN	0.3784	0.7452	1.1540

Note: Values denote exceedance-tail absolute errors for 95% CVaR forecasts and are reported in percentage points of future loss. Lower values indicate more accurate tail-severity forecasts.

3.6. Early-Warning Performance

Table 9 reports the early-warning performance for realized extreme-loss events. The proposed IQTN achieved the best results across all forecasting horizons and all major evaluation metrics. For the 1-day horizon, IQTN obtained a precision of 76.82%, a recall of 88.54%, and an F1-score of 82.26%, together with the highest AUPRC of 0.8165. Its false alarm rate and missed alarm rate were only 5.12% and 11.46%, respectively. The same superiority remained at the 5-day and 10-day horizons, where IQTN consistently delivered the highest precision, recall, F1-score, and AUPRC, while maintaining the lowest FAR and MAR. A clear performance ranking is also observed across model classes. Traditional statistical models, especially Historical Simulation VaR and EWMA VaR, showed relatively weak warning performance. GARCH-based models improved classification accuracy but still lagged behind the machine-learning and deep-learning methods. Among the benchmark learning models, TCN-Quantile performed best, followed by LSTM-Quantile and LightGBM-Quantile. However, IQTN further improved upon TCN-Quantile. For example, the F1-score increased from 74.87% to 82.26% at the 1-day horizon, from 70.52% to 77.85% at the 5-day horizon, and from 64.83% to 72.74% at the 10-day horizon. This result indicates that the proposed interpretable quantile-temporal design translated superior risk forecasts into more effective warning signals.

Table 9. Early-warning performance for realized extreme-loss events.

Model	Horizon	Precision (%)	Recall (%)	F1-Score (%)	FAR (%)	MAR (%)	AUPRC
Historical Simulation VaR	1-day	42.15	51.35	46.30	13.52	48.65	0.4512
EWMA VaR	1-day	48.62	58.46	53.09	11.85	41.54	0.5125
GARCH-t VaR	1-day	52.18	62.15	56.73	10.95	37.85	0.5562
GJR-GARCH-t VaR	1-day	55.34	66.82	60.54	10.36	33.18	0.5840
Linear Quantile Regression	1-day	50.45	59.80	54.73	11.30	40.20	0.5315
Quantile Random Forest	1-day	58.60	69.45	63.57	9.42	30.55	0.6220
LightGBM-Quantile	1-day	62.18	73.55	67.39	8.58	26.45	0.6658
LSTM-Quantile	1-day	65.40	77.12	70.78	7.82	22.88	0.7012
TCN-Quantile	1-day	69.45	81.20	74.87	6.85	18.80	0.7420
Proposed IQTN	1-day	76.82	88.54	82.26	5.12	11.46	0.8165
Historical Simulation VaR	5-day	38.65	46.80	42.33	14.32	53.20	0.4055
EWMA VaR	5-day	44.50	53.25	48.48	12.75	46.75	0.4682
GARCH-t VaR	5-day	47.92	57.10	52.11	11.90	42.90	0.5015
GJR-GARCH-t VaR	5-day	51.15	61.20	55.73	11.24	38.80	0.5348
Linear Quantile Regression	5-day	46.12	54.85	50.11	12.30	45.15	0.4852
Quantile Random Forest	5-day	54.20	64.15	58.76	10.40	35.85	0.5740
LightGBM-Quantile	5-day	58.20	68.45	62.91	9.42	31.55	0.6150
LSTM-Quantile	5-day	61.35	72.18	66.33	8.75	27.82	0.6510
TCN-Quantile	5-day	65.38	76.54	70.52	7.78	23.46	0.6985
Proposed IQTN	5-day	72.45	84.12	77.85	6.15	15.88	0.7724
Historical Simulation VaR	10-day	34.20	41.50	37.50	15.65	58.50	0.3546
EWMA VaR	10-day	39.80	47.15	43.16	13.95	52.85	0.4150
GARCH-t VaR	10-day	42.85	51.05	46.59	13.35	48.95	0.4485
GJR-GARCH-t VaR	10-day	46.25	54.60	50.08	12.44	45.40	0.4820
Linear Quantile Regression	10-day	41.15	49.20	44.82	13.80	50.80	0.4310
Quantile Random Forest	10-day	49.30	58.12	53.35	11.72	41.88	0.5215
LightGBM-Quantile	10-day	53.40	62.15	57.45	10.62	37.85	0.5612
LSTM-Quantile	10-day	56.45	66.05	60.87	10.05	33.95	0.5982
TCN-Quantile	10-day	60.15	70.28	64.83	9.12	29.72	0.6425
Proposed IQTN	10-day	67.80	78.45	72.74	7.30	21.55	0.7218

Note: Realized extreme-loss events are defined by the training-period 95th percentile threshold. All performance metrics are calculated from the out-of-sample confusion matrix. FAR denotes the false alarm rate, calculated as $FP/(FP + TN)$, and MAR denotes the missed alarm rate, calculated as $FN/(TP + FN)$. Precision, recall, F1-score, FAR, and MAR are reported in percentages (%), whereas AUPRC is unitless.

As expected, early-warning performance declined as the forecasting horizon increased. For IQTN, the AUPRC decreased from 0.8165 at the 1-day horizon to 0.7724 at the 5-day horizon and 0.7218 at the 10-day horizon. Precision, recall, and F1-score followed the same downward pattern. This deterioration is consistent with the greater uncertainty involved in forecasting more distant extreme-loss events. Nevertheless, IQTN preserved a clear advantage over all benchmark models even at the 10-day horizon. Figure 4 illustrates the four-level warning signals generated by IQTN during the out-of-sample test period. The warning signals intensified visibly during periods of elevated realized losses, especially from November 2025 to January 2026. Severe and warning-level signals were concentrated around the largest realized positive loss spikes, whereas normal states dominated relatively calm periods. Watch-level signals appeared more frequently and often surrounded high-risk episodes, suggesting that the model provided gradual escalation rather than abrupt alarms. Overall, the figure visually confirms that IQTN generated timely and interpretable warning signals that were well aligned with subsequent extreme-loss realizations.

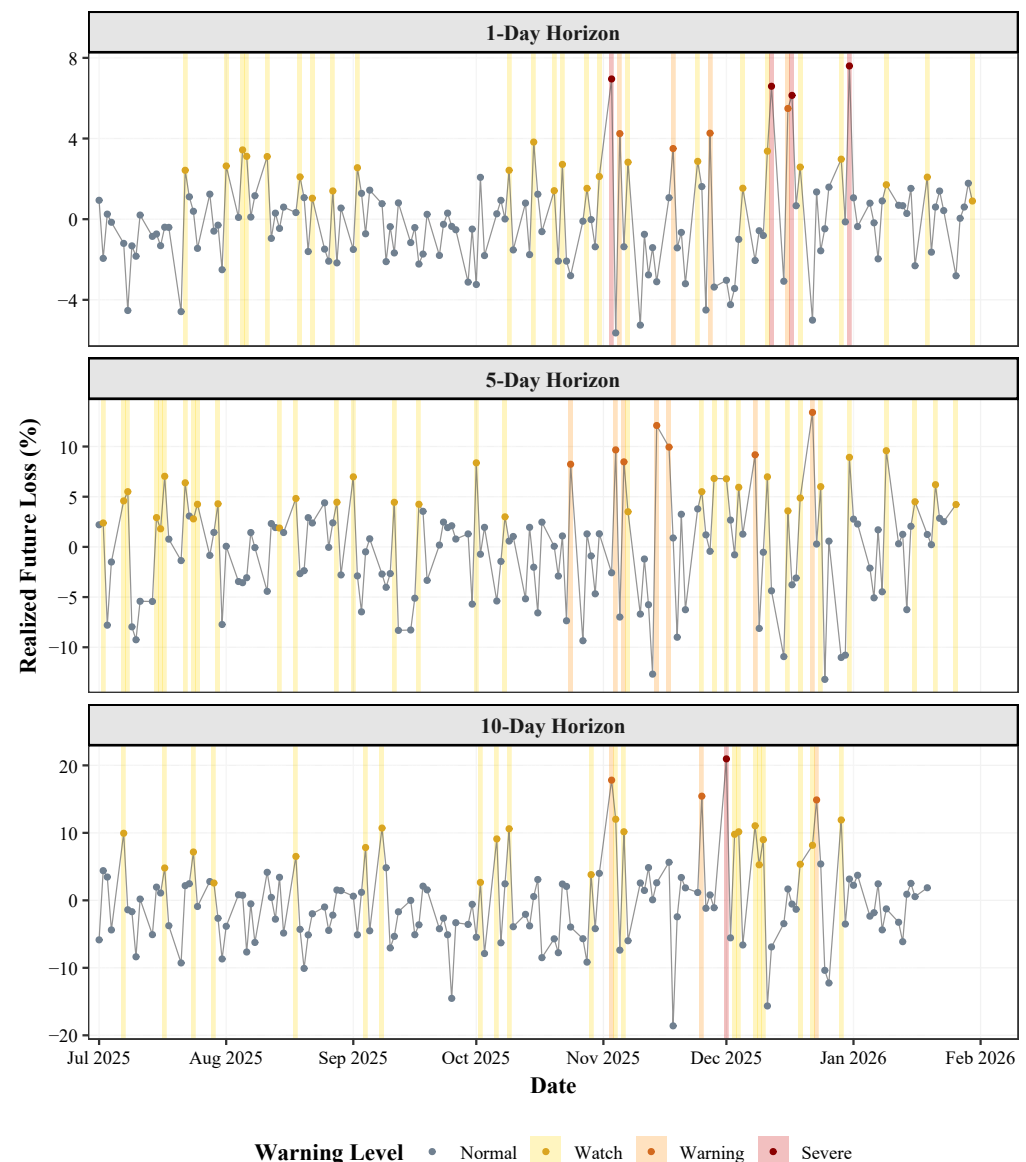


Figure 4. Four-level early-warning signals generated by IQTN across the 1-day, 5-day, and 10-day horizons during the out-of-sample test period. In each panel, the line shows the realized future loss for the corresponding horizon, and the shaded bands indicate the warning level assigned at the forecast origin.

3.7. Mechanistic Interpretation of Model Performance

The superior performance of IQTN in the preceding tables can be interpreted through three complementary mechanisms: selective market-state weighting, short-run risk-transmission learning, and coherent tail-distribution estimation. First, the feature-gating layer can be understood from a limited-attention and information-friction perspective. In an emerging carbon market, participants and risk monitors may not process all market signals with equal intensity, especially when market conditions change quickly. Under such conditions, attention is likely to concentrate on salient variables such as volatility, recent losses, liquidity pressure, and trading concentration [33–35]. The feature-gating layer provides a model-based analogue of this selective signal weighting. By assigning larger weights to currently informative variables and smaller weights to less relevant or noisier variables, it helps the model focus on the market-state drivers that are most relevant for tail-risk forecasting.

Second, the causal temporal convolutional encoder helps capture short-run risk-transmission patterns. Carbon-market tail risk is not generated only by the contemporaneous level of volatility or liquidity. Trading concentration, liquidity stress, intraday price movement, and recent downside pressure may propagate over several recent trading days before appearing as future tail losses. The temporal convolutional structure therefore allows IQTN to learn local persistence and short-horizon propagation among trading behavior, liquidity conditions, volatility, and downside pressure. This explains why IQTN outperformed non-temporal quantile models and why the temporal-importance results showed that recent observations dominated the forecast.

Third, the non-crossing quantile output layer improves the coherence of the estimated tail distribution. Traditional volatility models estimate risk mainly through a volatility channel, and generic deep temporal models may improve sequence prediction without guaranteeing an ordered multi-quantile loss distribution. In contrast, IQTN directly estimates coherent conditional loss quantiles for multiple horizons, which supports more stable VaR, CVaR, and early-warning signals. These mechanisms explain why IQTN improved not only average pinball loss but also VaR calibration, CVaR tail-severity forecasting, and early-warning performance. These interpretations should be understood as mechanism-oriented predictive explanations rather than structural causal claims.

3.8. Statistical Significance and Uncertainty Analysis

To assess whether the predictive improvements of IQTN were statistically meaningful, we conducted pairwise Diebold–Mariano tests using the out-of-sample pinball-loss differential between IQTN and the strongest benchmark models from each model class. As reported in Table 10, IQTN significantly outperformed GJR-GARCH-t VaR, LightGBM-Quantile, and TCN-Quantile across all three forecast horizons. The mean loss differences were positive in every comparison, indicating lower pinball loss for IQTN. The p -values were below 0.05 for all comparisons, including the most demanding comparison against TCN-Quantile at the 10-day horizon. These results indicate that the superior quantile forecasting performance of IQTN was unlikely to be driven only by random sampling variation in the test period.

Table 10. Pairwise Diebold–Mariano (DM) test results across multiple horizons.

Comparison	Horizon	Mean Loss Difference	DM Statistic	p -Value
IQTN vs. GJR-GARCH-t VaR	1-day	0.0272	4.26	<0.001
IQTN vs. LightGBM-Quantile	1-day	0.0131	3.01	0.003
IQTN vs. TCN-Quantile	1-day	0.0078	2.34	0.019
IQTN vs. GJR-GARCH-t VaR	5-day	0.0646	3.89	<0.001
IQTN vs. LightGBM-Quantile	5-day	0.0323	2.74	0.006
IQTN vs. TCN-Quantile	5-day	0.0204	2.16	0.031
IQTN vs. GJR-GARCH-t VaR	10-day	0.0921	3.52	<0.001
IQTN vs. LightGBM-Quantile	10-day	0.0527	2.41	0.016
IQTN vs. TCN-Quantile	10-day	0.0315	1.97	0.049

Note: The loss differential sequence is defined as $d_t = QPL_t^{\text{benchmark}} - QPL_t^{\text{IQTN}}$. A positive mean loss difference and a positive DM statistic indicate that IQTN has higher predictive accuracy, corresponding to lower pinball loss, than the benchmark model.

We further quantified uncertainty in the main IQTN performance metrics using moving-block bootstrap confidence intervals. Table 11 reports the point estimates and 95% confidence intervals based on 1000 bootstrap replications. The intervals show that the QPL estimates were stable across horizons, although uncertainty increased as the forecast horizon became longer. The bootstrap intervals for VaR exceedance rates remained close to the nominal 5% exceedance probability, supporting the calibration results reported in the

VaR backtesting section. The confidence intervals for AUPRC and F1-score also confirm that the early warning performance remained strong, with the expected widening of uncertainty at the 10-day horizon. Overall, the significance tests and bootstrap intervals provide additional evidence that IQTN achieved robust and statistically credible improvements in tail-risk forecasting and early-warning performance.

Table 11. Point estimates and 95% moving-block bootstrap confidence intervals for IQTN.

Metric	1-Day Horizon	5-Day Horizon	10-Day Horizon
Quantile Pinball Loss (QPL)	0.1765 [0.1534, 0.2001]	0.3958 [0.3412, 0.4526]	0.5732 [0.4895, 0.6582]
VaR Exceedance Rate	5.23% [2.58%, 8.46%]	6.04% [3.31%, 9.35%]	6.94% [4.11%, 10.37%]
CVaR ETAE	0.3784 [0.2815, 0.4892]	0.7452 [0.5492, 0.9581]	1.1540 [0.8264, 1.5093]
AUPRC	0.8165 [0.7412, 0.8795]	0.7724 [0.6895, 0.8412]	0.7218 [0.6315, 0.7964]
F1-score	82.26% [74.31%, 88.82%]	77.85% [69.04%, 85.06%]	72.74% [63.11%, 80.43%]

Note: Point estimates correspond to the full out-of-sample test results reported in previous sections. Brackets contain the 2.5th and 97.5th percentiles of the empirical distribution generated via 1000 moving-block bootstrap replications.

3.9. Interpretability Results

Figure 5 presents the interpretability results of IQTN from both global and temporal perspectives. Panel A shows that the most influential predictors were rolling volatility (0.2512), rolling maximum loss (0.1800), and intraday range (0.1545). These three variables all belong to the market-stress category, indicating that recent turbulence and realized downside pressure were the dominant drivers of carbon tail-risk forecasts. Trading value (0.1147), illiquidity (0.0920), and trading volume (0.0742) also made non-negligible contributions, suggesting that market activity and liquidity conditions provided additional information beyond pure price dynamics. By contrast, lagged return and calendar factors received relatively small weights, implying that they played a secondary role in the model's risk assessment. Panel B shows the lag-specific temporal importance within the 20-day input window. The importance weights declined sharply as the lag became more distant. The most recent observation had the largest contribution, and the first five lags accounted for the majority of the cumulative temporal importance. After approximately the 10th lag, the marginal contribution became very small. This pattern indicates that IQTN mainly relied on recent market information when forming current tail-risk forecasts, which is consistent with the short-memory and fast-adjustment characteristics of the carbon allowance market.

Table 12 provides local explanations for selected severe warning days. The dominant predictors on these high-risk dates were again rolling volatility, rolling maximum loss, intraday range, trading value, and illiquidity. For the 1-day severe warning dates, rolling volatility was ranked first on both occasions, showing that sudden local increases in short-term market stress were the main source of near-term tail-risk alerts. For the 5-day and 10-day severe-warning cases, rolling maximum loss, trading value, trading volume, and illiquidity appeared repeatedly among the top-ranked predictors, indicating that sustained downside pressure and worsening market liquidity were important for medium-horizon risk accumulation. Overall, the global and local interpretability results are highly consistent. IQTN identified market-stress variables as the core determinants of carbon tail risk, while also capturing the complementary roles of liquidity and trading activity. The temporal importance analysis further shows that these signals exerted their strongest effects through the most recent observations. These findings support the economic interpretability of the proposed model and increase its usefulness for practical market surveillance.

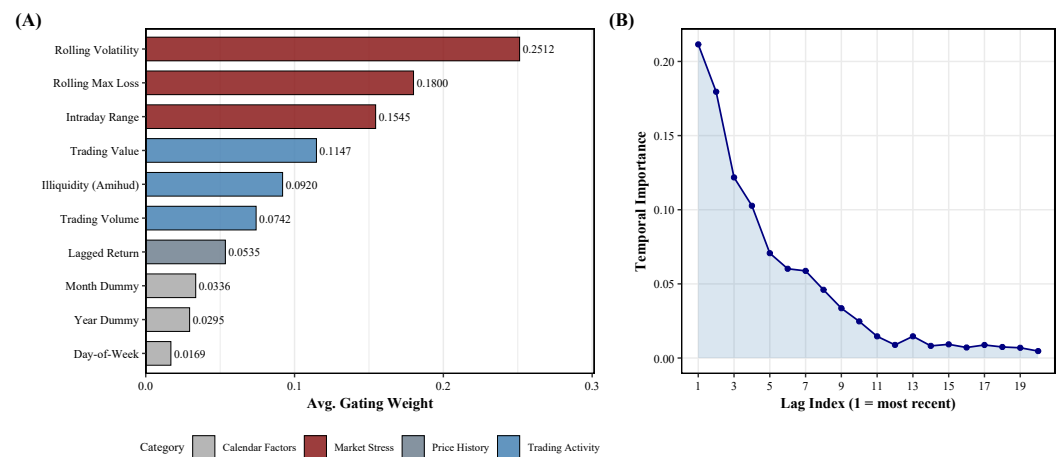


Figure 5. Interpretability analysis of the proposed IQTN model. **(A)** Average feature-gating weights during the test period, where predictors are grouped by their financial categories. **(B)** Lag-specific temporal importance within the 20-day input window, showing the relative contribution of historical observations to the current risk forecast.

Table 12. Local explanations for selected severe-warning days during the test period.

Forecast Date	Horizon	Top Predictor 1	Top Predictor 2	Top Predictor 3	Top Predictors 4–5
14 November 2025	1-day	Rolling Volatility	Rolling Maximum Loss	Intraday Range	Trading Value; Illiquidity (Amihud)
22 December 2025	1-day	Rolling Volatility	Intraday Range	Illiquidity (Amihud)	Rolling Maximum Loss; Trading Value
25 November 2025	5-day	Rolling Maximum Loss	Rolling Volatility	Trading Value	Intraday Range; Trading Volume
8 December 2025	5-day	Rolling Volatility	Illiquidity (Amihud)	Rolling Maximum Loss	Trading Value; Intraday Range
15 December 2025	10-day	Rolling Volatility	Rolling Maximum Loss	Intraday Range	Illiquidity (Amihud); Trading Volume

Note: Predictors are ranked according to their local feature-gating weights at the forecast origin. The reported predictors indicate model-attributed predictive contributions rather than causal effects.

3.10. Robustness and Ablation Results

3.10.1. Alternative Risk Levels

Table 13 reports the robustness results under alternative VaR confidence levels at the 1-day horizon. IQTN achieved the lowest pinball loss at the 90%, 95%, and 99% VaR levels, with values of 0.1428, 0.1765, and 0.2534, respectively. The advantage remained stable across moderate-tail, high-tail, and extreme-tail settings. This confirms that the superior performance of IQTN was not driven by a single VaR confidence level.

Table 13. Robustness analysis under alternative VaR confidence levels at the 1-day horizon.

Model	90% VaR	95% VaR	99% VaR
GJR-GARCH-t VaR	0.1892	0.2037	0.3421
LightGBM-Quantile	0.1683	0.1896	0.2942
LSTM-Quantile	0.1748	0.1945	0.3067
TCN-Quantile	0.1564	0.1843	0.2789
Proposed IQTN	0.1428	0.1765	0.2534

Note: Values denote average pinball loss at the 1-day horizon. Lower values indicate better quantile forecasting accuracy. Representative benchmark models from traditional risk models, machine-learning quantile models, and deep temporal quantile models are reported.

3.10.2. Alternative Lookback Windows

Table 14 compares alternative lookback windows for IQTN. The 20-trading-day window achieved the lowest pinball loss across all three horizons, with QPL values of 0.1765, 0.3958, and 0.5732. The 10-day window produced higher errors, indicating that a shorter history did not provide sufficient information for dynamic tail-risk forecasting. The 30-day window also performed slightly worse than the 20-day window, suggesting that more distant observations added limited additional value. These results support the use of the 20-day window as the main specification.

Table 14. Robustness analysis of IQTN under alternative lookback windows measured by 95% VaR pinball loss.

Lookback Window	1-Day QPL	5-Day QPL	10-Day QPL
10 trading days	0.1894	0.4215	0.6180
20 trading days	0.1765	0.3958	0.5732
30 trading days	0.1792	0.4024	0.5847

Note: Values denote average pinball loss for 95% VaR forecasts. Lower values indicate better quantile forecasting accuracy. The 20-day window corresponds to the main IQTN specification.

3.10.3. Alternative Sample-Split and Rolling-Origin Robustness

Table 15 reports the robustness results under alternative chronological sample splits and rolling-origin evaluation. Across all validation designs, IQTN consistently achieves lower QPL than GJR-GARCH- t across the 1-day, 5-day, and 10-day horizons. Under the earlier test split, IQTN achieves QPL values of 0.1712, 0.3785, and 0.5429, compared with 0.1984, 0.4412, and 0.6318 for GJR-GARCH- t . Under the later test split, both models show higher losses, reflecting the more stressed evaluation window, but IQTN still maintains lower QPL and higher AUPRC. The rolling-origin evaluation further confirms the stability of the main findings. The average rolling-origin QPL values of IQTN are 0.1772, 0.3934, and 0.5694, which are close to the baseline test-sample results. The rolling-origin standard deviations are also moderate and smaller for IQTN than for GJR-GARCH- t . These findings indicate that the predictive advantage of IQTN is not driven by a single sample partition, although the short history and evolving liquidity conditions of the national CEA market remain important constraints.

Table 15. Robustness results under alternative sample splits and rolling-origin evaluation.

Validation Design	Model	1-Day QPL	5-Day QPL	10-Day QPL	VaR ER	AUPRC
Baseline split	GJR-GARCH- t	0.2037	0.4604	0.6653	7.26%	0.5842
	IQTN	0.1765	0.3958	0.5732	5.12%	0.6895
Earlier test split	GJR-GARCH- t	0.1984	0.4412	0.6318	6.94%	0.5915
	IQTN	0.1712	0.3785	0.5429	4.91%	0.7024
Later test split	GJR-GARCH- t	0.2156	0.4812	0.6984	7.48%	0.5716
	IQTN	0.1854	0.4109	0.6015	5.28%	0.6752
Rolling-origin mean	GJR-GARCH- t	0.2045	0.4589	0.6612	7.11%	0.5821
	IQTN	0.1772	0.3934	0.5694	5.08%	0.6904
Rolling-origin sd	GJR-GARCH- t	0.0184	0.0342	0.0495	0.68%	0.0241
	IQTN	0.0142	0.0271	0.0388	0.43%	0.0215

Note: QPL denotes the 95% quantile pinball loss. VaR ER and AUPRC are evaluated at the representative 5-day forecast horizon ($h = 5$) against the 5.00% nominal target level. Baseline split employs the original partition. Earlier test split advances the test partition boundary by six months. Later test split evaluates a more recent high-stress window. Rolling-origin mean and sd report the average performance and standard deviation across a series of recursive rolling out-of-sample forward evaluation windows.

3.10.4. Ablation Analysis

Table 16 reports the module-level ablation results. The full IQTN achieved the lowest QPL across all three horizons and maintained a zero quantile crossing rate. Removing the feature-gating layer increased QPL from 0.1765 to 0.1843 at the 1-day horizon, from 0.3958 to 0.4162 at the 5-day horizon, and from 0.5732 to 0.6047 at the 10-day horizon. This indicates that dynamic market-state weighting improves predictive accuracy while also providing interpretable risk-driver information. The largest deterioration occurred when the temporal encoder was removed. The QPL increased to 0.2356, 0.5284, and 0.7652 across the 1-day, 5-day, and 10-day horizons, respectively. This confirms that short-run temporal propagation of volatility, liquidity, trading activity, and downside-pressure signals is central to the forecasting task. Removing the non-crossing quantile output layer produced a smaller increase in QPL but generated a quantile crossing rate of 4.28%. This shows that the non-crossing structure contributes mainly by enforcing coherent tail-distribution estimates for VaR and CVaR. Finally, removing trading activity and liquidity variables also substantially increased QPL, confirming that transaction intensity and liquidity pressure are important components of the carbon market risk state.

Table 16. Module-level ablation results for the IQTN framework.

Model Variant	1-Day QPL	5-Day QPL	10-Day QPL	Crossing Rate	Main Interpretation
Full IQTN	0.1765	0.3958	0.5732	0.00%	Complete framework
w/o feature gating	0.1843	0.4162	0.6047	0.00%	Lower interpretability and selective weighting
w/o temporal encoder	0.2356	0.5284	0.7652	0.00%	Loses short-run risk propagation
w/o non-crossing output	0.1798	0.4034	0.5861	4.28%	Independent quantiles suffer from crossing
w/o trading-liquidity variables	0.2082	0.4658	0.6744	0.00%	Loss of transaction and liquidity states

Note: The table reports the 95% quantile pinball loss and empirical quantile crossing rate across ablation variants. The feature-gating ablation removes the dynamic state-weighting layer. The temporal-encoder ablation deactivates the causal TCN module. The non-crossing-output ablation removes the monotonic quantile construction. The trading-liquidity ablation excludes transaction intensity and Amihud-type illiquidity proxies from the input state vector.

3.11. Lead-Lag Diagnostics for Market-State Variables

To address potential predictive endogeneity and mutual temporal dependence among market-state variables, we conducted pairwise Granger causality diagnostics for rolling volatility, trading value, and illiquidity. Table 17 reports the results using five trading lags. The null hypothesis that trading value does not Granger-cause rolling volatility was rejected, indicating that past transaction intensity contains predictive information for subsequent volatility. The null hypothesis that illiquidity does not Granger-cause rolling volatility was also rejected, suggesting that liquidity stress precedes part of the variation in market turbulence.

The results also show feedback between volatility and liquidity. Rolling volatility Granger-causes illiquidity, while trading value Granger-causes illiquidity. These findings are consistent with the systems-risk interpretation that trading concentration, liquidity stress, and volatility are dynamically linked. They also support the use of a temporal sequence model that can learn short-run propagation among these market-state variables. They do not eliminate all endogeneity concerns, but they provide additional evidence that the main variables used by IQTN contain meaningful temporal information for carbon-market risk monitoring.

Table 17. Pairwise Granger causality test results for market-state variables.

Null Hypothesis (H_0)	Selected Lag	F-Statistic	p-Value	Interpretation
Trading value does not Granger-cause rolling volatility	5	4.82	0.002	Rejected
Illiquidity does not Granger-cause rolling volatility	5	2.94	0.034	Rejected
Rolling volatility does not Granger-cause illiquidity	5	3.61	0.013	Feedback evidence
Trading value does not Granger-cause illiquidity	5	4.15	0.005	Trading-liquidity linkage

Note: The table reports formal F-statistics and p -values for pairwise Granger causality tests using five trading lags. Rolling volatility is measured over a 5-day window ($\sigma_{t,5}$), trading value by $LVAL_t$, and illiquidity by $ILLIQ_t$. The tests are interpreted as lead-lag predictive diagnostics rather than structural causal evidence.

3.12. Mechanism-Oriented Predictive Pathway Analysis

Table 18 reports the results of the mechanism-oriented predictive pathway analysis. For the pathway from volatility to illiquidity to future loss, the path from lagged volatility to illiquidity was positive, and the path from illiquidity to future loss was also positive across the 1-day, 5-day, and 10-day horizons. The corresponding indirect predictive effects were 0.0444, 0.0972, and 0.1372, and all moving-block bootstrap confidence intervals excluded zero. This suggests that recent volatility is linked to future tail risk partly through liquidity stress.

For the pathway from trading value to illiquidity to future loss, the indirect predictive effects were also positive across all horizons. The estimated indirect effects increased from 0.0582 at the 1-day horizon to 0.1239 at the 5-day horizon and 0.1771 at the 10-day horizon. The bootstrap confidence intervals again excluded zero, indicating that trading concentration is linked to future tail loss through illiquidity. These results support the proposed systems-risk mechanism in which volatility pressure and transaction concentration interact with liquidity conditions before appearing as future tail-risk outcomes.

Table 18. Empirical results for mechanism-oriented predictive pathway analysis.

Predictive Pathway	Horizon	Path a (SE)	Path b (SE)	Indirect Effect	95% CI
Pathway 1: Volatility $\rightarrow$ Illiquidity $\rightarrow$ Loss	1-day	0.2843 (0.098)	0.1562 (0.066)	0.0444	[0.0078, 0.0921]
Volatility $\rightarrow$ Illiquidity $\rightarrow$ Loss	5-day	0.2843 (0.098)	0.3418 (0.103)	0.0972	[0.0294, 0.1816]
Volatility $\rightarrow$ Illiquidity $\rightarrow$ Loss	10-day	0.2843 (0.098)	0.4825 (0.134)	0.1372	[0.0513, 0.2485]
Pathway 2: Trading Value $\rightarrow$ Illiquidity $\rightarrow$ Loss	1-day	0.3926 (0.112)	0.1483 (0.064)	0.0582	[0.0094, 0.1183]
Trading Value $\rightarrow$ Illiquidity $\rightarrow$ Loss	5-day	0.3926 (0.112)	0.3156 (0.096)	0.1239	[0.0381, 0.2312]
Trading Value $\rightarrow$ Illiquidity $\rightarrow$ Loss	10-day	0.3926 (0.112)	0.4512 (0.121)	0.1771	[0.0624, 0.3149]

Note: Path a reports the coefficient of the lagged predictor X_{t-1} on current illiquidity $ILLIQ_t$. Path b reports the coefficient of the mediator on future tail loss $L_{t,h}$. Volatility is measured by $\sigma_{t-1,5}$ and trading value by $LVAL_{t-1}$. Standard errors are HAC-robust. The 95% confidence intervals for the indirect effect are constructed via 1000 moving-block bootstrap replications.

3.13. Policy-Window Stability Results

Table 19 reports the conditional 95% VaR pinball loss during policy announcement windows and normal trading windows. The results show that policy announcement windows were associated with higher absolute prediction errors for both GJR-GARCH-t VaR and IQTN. For example, the 1-day IQTN QPL increased from 0.1732 during normal trading windows to 0.1894 during policy announcement windows, and the 10-day IQTN QPL increased from 0.5634 to 0.6128. This pattern suggests that policy-sensitive periods are more difficult to forecast, which is consistent with the possibility that regulatory information affects market expectations, trading activity, and liquidity conditions.

Importantly, IQTN maintained its advantage over GJR-GARCH-t VaR in both policy-sensitive and normal windows. During policy announcement windows, IQTN reduced QPL by 12.87%, 13.25%, and 13.06% across the 1-day, 5-day, and 10-day horizons. During normal trading windows, the corresponding reductions were 13.05%, 13.22%, and 13.07%. The DM

statistics were significant at the 1% level in all comparisons. These findings indicate that policy announcement periods increase the absolute level of forecasting difficulty, but the relative advantage of IQTN remains stable across policy-sensitive and normal trading windows.

Table 19. Conditional pinball loss during policy announcement and normal trading windows.

Market Window	Horizon	GJR-GARCH-t QPL	IQTN QPL	Relative Reduction	DM Statistic
Policy announcement windows ($I_t^{policy} = 1$)	1-day	0.2174	0.1894	12.87%	3.82 ***
	5-day	0.4859	0.4215	13.25%	4.15 ***
	10-day	0.7049	0.6128	13.06%	3.96 ***
Normal trading windows ($I_t^{policy} = 0$)	1-day	0.1992	0.1732	13.05%	4.26 ***
	5-day	0.4485	0.3892	13.22%	4.85 ***
	10-day	0.6481	0.5634	13.07%	4.62 ***

Note: The table reports conditional 95% quantile pinball loss during policy announcement windows, defined as five trading days before and after major regulatory releases, and during normal trading windows. The DM statistic evaluates the null hypothesis of equal predictive accuracy between GJR-GARCH-t VaR and IQTN within each market window. *** indicates statistical significance at the 1% level.

3.14. Heterogeneity Results Across Market Conditions

Table 20 reports the heterogeneity results across high- versus low-volatility periods and near-compliance versus normal trading periods. In Panel A, both GJR-GARCH-t VaR and IQTN had higher QPL during high-volatility periods than during low-volatility periods, indicating that elevated volatility increases forecasting difficulty. However, IQTN consistently achieved lower QPL than GJR-GARCH-t across all horizons. IQTN also produced VaR exceedance rates closer to the nominal 5% exceedance probability and higher AUPRC values in both high- and low-volatility periods. The model-comparison bootstrap p -values were below 0.05 in all cases, indicating statistically significant QPL reductions by IQTN.

Table 20. Heterogeneity and conditional regime stability results at the 95% VaR level.

Horizon	Group/Model	QPL	VaR ER	AUPRC	Model p -Value	Group p -Value
Panel A: High-volatility vs. low-volatility periods						
1-day	High-volatility periods					
	GJR-GARCH- t	0.2341	7.18%	0.6724	—	—
	IQTN	0.2015	5.26%	0.7815	0.003	—
	Low-volatility periods					
	GJR-GARCH- t	0.1558	4.15%	0.5412	—	—
	IQTN	0.1362	4.91%	0.6328	0.012	0.001
5-day	High-volatility periods					
	GJR-GARCH- t	0.5126	7.34%	0.6915	—	—
	IQTN	0.4382	5.41%	0.8034	0.002	—
	Low-volatility periods					
	GJR-GARCH- t	0.3684	4.08%	0.5638	—	—
	IQTN	0.3195	4.88%	0.6512	0.009	0.002
10-day	High-volatility periods					
	GJR-GARCH- t	0.7419	7.62%	0.7108	—	—
	IQTN	0.6354	5.56%	0.8247	0.004	—
	Low-volatility periods					
	GJR-GARCH- t	0.5215	3.92%	0.5842	—	—
	IQTN	0.4618	4.81%	0.6735	0.015	0.004

Table 20. Cont.

Horizon	Group/Model	QPL	VaR ER	AUPRC	Model <i>p</i> -Value	Group <i>p</i> -Value
Panel B: Near-compliance vs. normal periods						
1-day	Near-compliance periods					
	GJR-GARCH- <i>t</i>	0.2482	7.42%	0.6945	—	—
	IQTN	0.2124	5.31%	0.8112	0.002	—
	Normal trading periods					
	GJR-GARCH- <i>t</i>	0.1695	4.26%	0.5592	—	—
	IQTN	0.1481	4.94%	0.6473	0.011	0.002
5-day	Near-compliance periods					
	GJR-GARCH- <i>t</i>	0.5348	7.61%	0.7184	—	—
	IQTN	0.4526	5.39%	0.8351	0.001	—
	Normal trading periods					
	GJR-GARCH- <i>t</i>	0.3912	4.19%	0.5746	—	—
	IQTN	0.3394	4.86%	0.6628	0.007	0.005
10-day	Near-compliance periods					
	GJR-GARCH- <i>t</i>	0.7716	7.84%	0.7329	—	—
	IQTN	0.6582	5.62%	0.8514	0.002	—
	Normal trading periods					
	GJR-GARCH- <i>t</i>	0.5493	3.85%	0.5961	—	—
	IQTN	0.4815	4.78%	0.6843	0.013	0.003

Note: QPL denotes the 95% quantile pinball loss. VaR ER denotes the actual exceedance rate relative to the 5% nominal exceedance probability. AUPRC measures early-warning performance for severe risk states. High- and low-volatility periods are defined using the median of rolling volatility within the evaluated out-of-sample period. Near-compliance periods refer to trading days in November and December, and normal periods refer to the remaining trading days. The model *p*-value evaluates the significance of the QPL reduction achieved by IQTN over GJR-GARCH-*t* within the same group. The group *p*-value tests whether IQTN's QPL differs between the two market environments in each panel. All inference is based on 1000 moving-block bootstrap replications.

Panel B shows a similar pattern for near-compliance periods. Near-compliance periods had higher QPL values than normal trading periods, suggesting that year-end compliance-related trading pressure increases tail-risk forecasting difficulty. Nevertheless, IQTN again outperformed GJR-GARCH-*t* VaR across all horizons, with lower QPL, better VaR calibration, and higher AUPRC. The group-difference bootstrap *p*-values indicate that IQTN's QPL differs significantly between the two market environments in both panels. The higher AUPRC values in high-volatility and near-compliance periods should be interpreted together with the greater concentration of extreme-loss events in these strata. Overall, the heterogeneity results show that market conditions affect the absolute level of forecasting difficulty, but IQTN retains a stable advantage across stressed and normal market environments.

4. Discussion and Conclusions

4.1. Principal Findings and Interpretation

This study developed an interpretable quantile temporal network for dynamic tail-risk forecasting and early warning in China's national carbon allowance market. From a systems perspective, the national CEA market can be viewed as a complex market system in which price volatility, liquidity conditions, trading activity, compliance-related pressure, and regulatory expectations interact over time. The empirical results show that IQTN consistently outperformed traditional VaR models, machine-learning quantile models, and deep temporal quantile benchmarks in average pinball loss, VaR backtesting, CVaR tail-severity evaluation, and early-warning performance. Rather than treating these results only as evidence of higher predictive accuracy, they should be interpreted in relation to the institutional and market structure of China's national CEA market.

The findings are consistent with the broader carbon-market literature, which shows that carbon prices are affected by energy fundamentals, macroeconomic conditions, regulatory adjustments, structural breaks, and market-design features [36–40]. These studies imply that carbon-market risk cannot be understood as a purely statistical property of historical returns. Instead, risk is shaped by interactions between market fundamentals, policy design, institutional constraints, and trading behavior. This point is especially important for China's national carbon market because it is a young and institutionally evolving market system. China's carbon markets developed after regional pilot-market experimentation, and national CEA online trading began in July 2021. During the sample period, the national market was still developing its trading conventions, compliance arrangements, data-disclosure practices, and regulatory infrastructure. The institutional framework was further strengthened when the Interim Regulations on the Administration of Carbon Emission Trading came into effect in May 2024, providing a clearer legal basis for the operation and supervision of the national carbon market [13,41]. Recent studies on integrated energy systems and coupled electricity–carbon markets suggest that energy and carbon-market processes are dynamic, interdependent, and shaped by system coordination, market interactions, and participant decisions [42,43]. This supports the systems-oriented view adopted in this study, while IQTN differs by focusing on the conditional upper-tail distribution of future CEA losses and translating short-horizon market-state information into multi-horizon VaR, CVaR, and early-warning outputs.

China's carbon markets also display heterogeneous trading behavior, regime variation, and market immaturity during pilot and national-market development [3,4,44]. This institutional background helps explain why carbon-market tail risk may arise not only from price volatility but also from compliance-related trading needs, changing regulatory expectations, uneven market participation, and liquidity conditions. For an emerging national market, trading activity may be episodic, liquidity provision may be uneven, and market participants may adjust positions in response to compliance cycles or policy expectations. Therefore, a systems-oriented interpretation is needed to understand why the same market can exhibit calm periods, sudden volatility, and concentrated tail-risk episodes within a relatively short sample window.

From a mechanism perspective, the empirical results are consistent with a risk-transmission process linking institutional pressure, trading behavior, liquidity stress, price volatility, and tail losses. Compliance-related expectations and regulatory signals may influence when market participants adjust allowance positions. When many participants adjust positions within similar periods, trading activity and trading value may become concentrated. If market depth is limited or liquidity provision is uneven, such trading concentration can be accompanied by higher illiquidity and larger price impact. Liquidity pressure and concentrated trading may then amplify intraday price ranges, rolling volatility, and recent maximum losses. These short-term stress signals can ultimately increase the upper tail of the future loss distribution, which is the direct target of VaR and CVaR forecasting.

The interpretability results further support this mechanism-based interpretation. IQTN assigned the largest weights to rolling volatility, rolling maximum loss, intraday range, trading value, and illiquidity. These variables are not isolated predictors; they correspond to different components of the carbon market risk system. Rolling volatility captures recent market turbulence, rolling maximum loss captures realized downside pressure, intraday range captures within-day price instability, trading value reflects transaction concentration, and illiquidity captures the degree to which trading pressure is associated with price impact. The dominance of these variables suggests that carbon-market tail risk is formed through the interaction of price stress, transaction intensity, and liquidity conditions rather

than through a single volatility channel. This is consistent with previous evidence that carbon-market risk is shaped by macroeconomic states, volatility clustering, asymmetric tails, and market-specific liquidity conditions [5,6,45,46].

This mechanism also explains why static historical simulation and conventional volatility models produced weaker risk forecasts in this study. Static historical simulation relies on past empirical loss distributions and cannot update flexibly when the current market state changes. Conventional GARCH-type models capture conditional volatility but are less able to represent nonlinear interactions among volatility, liquidity, trading concentration, and calendar-related market states. Tree-based and recurrent neural benchmarks can improve nonlinear prediction, but they do not simultaneously provide feature-level attribution, short-horizon risk propagation, and non-crossing multi-quantile tail estimates. In contrast, IQTN combines these elements in a single risk-monitoring architecture. The feature-gating layer dynamically identifies which market-state variables receive greater predictive weight under current conditions; the causal temporal convolutional encoder captures the short-term propagation of volatility, liquidity, and trading-pressure signals; and the non-crossing quantile output layer directly represents the conditional loss distribution needed for VaR and CVaR estimation.

The temporal importance results further support this interpretation. The model placed the greatest weight on the most recent observations, and the importance of earlier lags declined quickly within the 20-day input window. This pattern suggests that the national CEA market's tail-risk state is highly sensitive to recent information. For a young carbon market with evolving rules, uneven liquidity, and compliance-related trading concentration, recent market stress may be more informative than distant historical averages. Therefore, the superior performance of IQTN is not merely a technical result. It reflects the suitability of a short-horizon, interpretable, quantile-based temporal architecture for monitoring a market system in which institutional pressure, trading concentration, liquidity conditions, and price volatility interact dynamically.

This interpretation should not be read as a causal identification claim. The feature-gating weights indicate model-attributed predictive relevance rather than structural causal effects. Nevertheless, the consistency among the forecasting results, VaR backtesting, early-warning performance, ablation analysis, and interpretability results provides mechanism-oriented predictive evidence that carbon-market tail risk is jointly associated with recent volatility, downside pressure, transaction concentration, and liquidity stress. In this sense, IQTN contributes not only a forecasting tool but also a systems-oriented framework for interpreting and monitoring tail-risk formation in China's national carbon allowance market.

4.2. Implications for Carbon-Market Risk Monitoring

The main contribution of this study is to shift the modeling target from carbon price prediction to carbon tail-risk forecasting. Recent carbon price forecasting studies have increasingly used interval prediction, decomposition methods, deep learning, causal temporal convolution, and multi-source information to handle the nonlinear and non-stationary nature of carbon prices [7–10]. This study extends that direction by directly forecasting VaR and CVaR and by converting predicted tail risk into operational early-warning signals. The proposed framework has practical value for carbon-market risk monitoring because it links each risk level to a differentiated monitoring or risk-management response. Regulators can use the normal level as a routine monitoring state, the watch level as an indication that market stress should be observed more closely, the warning level as a trigger for enhanced surveillance of volatility and liquidity conditions, and the severe-warning level as a trigger for focused risk review. Compliance entities can use the same signals to review allowance exposure, fulfillment-related demand, and liquidity buffers before compliance deadlines.

Trading institutions can use the warning levels to adjust allowance positions, tighten internal VaR and CVaR limits, and conduct adverse-scenario stress testing. The operational value of the framework comes not only from the warning level itself but also from the explanation attached to the warning. Because IQTN provides feature-gating and lag-specific explanations, the warning signal is not a black-box alarm; it indicates whether the risk is mainly driven by rolling volatility, recent maximum loss, intraday range, trading activity, or illiquidity. For example, a severe-warning signal accompanied by high illiquidity and trading-value weights would suggest that liquidity pressure and transaction concentration should receive immediate attention. A warning signal mainly driven by rolling volatility and intraday range would instead suggest that price turbulence and short-term market instability are the dominant concerns. The early-warning design is especially useful for an emerging carbon market. A single VaR number is informative but not always operational. By transforming VaR and CVaR into normal, watch, warning, and severe-warning levels, the proposed framework provides a direct risk-monitoring interface. The empirical results show that warning and severe-warning signals were concentrated around high-loss periods, while watch signals captured intermediate risk states. This structure supports gradual escalation rather than abrupt binary alarms. Importantly, the proposed warning levels are not intended to impose mechanical trading restrictions or automatic regulatory intervention. They are designed as structured decision-support triggers that help regulators and market participants decide when to move from routine monitoring to enhanced observation, intensified risk review, or conservative risk-management action.

4.3. Specific Policy and Managerial Implications

The proposed framework also provides several specific policy and managerial implications. First, regulators and carbon exchanges can use IQTN as part of a market-surveillance dashboard. Instead of monitoring price changes alone, the dashboard can jointly track RiskScore, VaR, CVaR, and the feature-gating weights of volatility, maximum loss, trading value, and illiquidity. This would allow supervisors to identify not only whether tail risk is increasing but also which market-state variables are driving the increase.

Second, the model can support targeted liquidity and transaction-concentration review. When a warning or severe warning signal is accompanied by high weights on illiquidity and trading value, the relevant authority can examine whether risk is being amplified by concentrated trading, weak market depth, or deteriorating liquidity conditions. In this case, appropriate responses may include enhanced market monitoring, risk communication, or further review of liquidity conditions within the applicable regulatory framework.

Third, compliance entities and trading institutions can use the warning signals for internal risk management. A watch signal may trigger exposure review and allowance-procurement planning. A warning signal may justify tighter internal VaR and CVaR limits, additional stress testing, and earlier preparation for compliance demand. A severe warning signal may call for more conservative position management, stronger liquidity buffers, and adverse-scenario analysis. These applications do not imply automatic trading restrictions or mechanical regulatory intervention. Rather, IQTN should be used as a structured decision-support tool that combines quantitative risk forecasts with institutional judgment and periodic model recalibration.

4.4. Actionable Early-Warning Protocols

To make the policy implications more operational, the four-level warning system can be translated into several decision-support protocols. These protocols are not automatic intervention rules; rather, they provide structured triggers for further review by regulators, exchanges, compliance entities, and trading institutions. To avoid arbitrary fixed

cutoffs, driver-specific thresholds are defined using training-period quantiles and can be recalibrated as the market matures.

First, a severe systemic tail-risk review can be initiated when $RiskScore_{t,h} > c_{0.90}$, corresponding to a severe-warning state. In this case, regulators or exchanges may review whether the warning is driven mainly by volatility, recent maximum loss, trading concentration, or liquidity stress. Second, a liquidity-risk review can be triggered when $RiskScore_{t,h} > c_{0.90}$ and the liquidity-related gating relevance or the observed illiquidity proxy exceeds its training-period upper-quartile or 90th-percentile threshold. This protocol is designed to identify situations in which tail risk is accompanied by weak market depth, concentrated trading, or abnormal price impact.

Third, a downside-pressure and position risk review can be triggered when the warning level is at least warning, rolling maximum loss increases for three consecutive trading days, and the predicted 95% VaR exceeds its training-period 90th percentile. Under this condition, compliance entities and trading institutions may tighten internal VaR and CVaR limits, conduct adverse-scenario stress testing, and review discretionary allowance positions. Fourth, a compliance-window monitoring protocol can be activated when warning or severe warning signals occur during November–December and trading value or trading volume exceeds its training-period upper-quartile threshold. This protocol supports earlier allowance-procurement planning, liquidity-buffer preparation, and monitoring of concentrated year-end trading pressure.

These protocols connect the statistical warning levels to concrete operational actions while preserving human judgment. The model identifies when and why the risk state has intensified, but the final response should depend on market context, regulatory rules, liquidity conditions, and the cost of false alarms.

4.5. Limitations and Conclusions

This study has three limitations. First, China's national carbon market began trading in 2021, so the available national-market time series is still relatively short compared with mature financial markets. Second, this study used daily trading data and did not include intraday order-book information, block-trade microstructure, or firm-level compliance behavior. Third, the empirical analysis focused on the national CEA market. Future research can extend the framework to regional carbon markets, CCER-related trading, and cross-market spillovers between carbon, energy, and financial markets. Several additional limitations should also be acknowledged. First, policy announcements, regulatory changes, allowance-allocation adjustments, and compliance-rule changes may introduce structural breaks in the carbon-market data-generating process. The added policy window and heterogeneity analyses improve the assessment of conditional predictive stability, but they do not fully replace formal structural-break modelling or long-horizon regime-stability assessment. Second, false-positive warnings may generate regulatory and managerial costs, including unnecessary supervisory review, excessive position adjustment, avoidable liquidity-buffer preparation, and possible market overreaction. Future research could incorporate cost-sensitive threshold calibration, asymmetric utility functions, or decision-curve analysis to balance missed alarms against false positives. Third, the short history of the national CEA trading limits cross-period stability assessment. Because national CEA trading began in July 2021, the available sample does not permit a long pre-pandemic, pandemic, and post-pandemic comparison within the same national market. Future studies with longer samples should evaluate model stability across structural-break episodes, pandemic-related market regimes, and future phases of carbon-market expansion. Despite these limitations, the results provide clear evidence that interpretable quantile temporal modeling is effective for carbon-market tail-risk management. IQTN achieved the best performance in VaR

forecasting, VaR backtesting, CVaR tail-severity prediction, and early warning classification. The robustness and ablation analyses show that the model benefits jointly from temporal sequence learning, trading-liquidity information, and interpretable feature gating. Overall, this study provides a practical and transparent framework for dynamic tail-risk forecasting and early-warning in China's national carbon allowance market.

Author Contributions: Conceptualization, T.H. and G.T.R.L.; methodology, T.H.; software, T.H.; validation, T.H. and G.T.R.L.; formal analysis, T.H.; investigation, T.H.; resources, G.T.R.L.; data curation, T.H.; writing—original draft preparation, T.H.; writing—review and editing, G.T.R.L.; visualization, T.H.; supervision, G.T.R.L.; project administration, G.T.R.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data analyzed in this study were obtained from publicly available official trading information released by the Shanghai Environment and Energy Exchange through the National Carbon Market Trading Information Publishing Platform. The open-source *iqtn* package is available on GitHub at <https://github.com/HTL1230/iqtn> (accessed on 10 April 2026).

Conflicts of Interest: The authors declare no conflicts of interest.

References

- Geels, F.W. Technological Transitions as Evolutionary Reconfiguration Processes: A Multi-Level Perspective and a Case-Study. *Res. Policy* **2002**, *31*, 1257–1274.
- Markard, J.; Raven, R.; Truffer, B. Sustainability Transitions: An Emerging Field of Research and Its Prospects. *Res. Policy* **2012**, *41*, 955–967.
- Munnings, C.; Morgenstern, R.D.; Wang, Z.; Liu, X. Assessing the Design of Three Carbon Trading Pilot Programs in China. *Energy Policy* **2016**, *96*, 688–699.
- Chang, K.; Pei, P.; Zhang, C.; Wu, X. Exploring the Price Dynamics of CO₂ Emissions Allowances in China's Emissions Trading Scheme Pilots. *Energy Econ.* **2017**, *67*, 213–223.
- Jiao, L.; Liao, Y.; Zhou, Q. Predicting Carbon Market Risk Using Information from Macroeconomic Fundamentals. *Energy Econ.* **2018**, *73*, 212–227.
- Liu, T.; Guan, X.; Wei, Y.; Xue, S.; Xu, L. Impact of Economic Policy Uncertainty on the Volatility of China's Emission Trading Scheme Pilots. *Energy Econ.* **2023**, *121*, 106626.
- Zhu, B.; Wan, C.; Wang, P. Interval Forecasting of Carbon Price: A Novel Multiscale Ensemble Forecasting Approach. *Energy Econ.* **2022**, *115*, 106361.
- Li, D.; Li, Y.; Wang, C.; Chen, M.; Wu, Q. Forecasting Carbon Prices Based on Real-Time Decomposition and Causal Temporal Convolutional Networks. *Appl. Energy* **2023**, *331*, 120452.
- Huang, Z.; Zhang, W. Forecasting Carbon Prices in China's Pilot Carbon Market: A Multi-Source Information Approach with Conditional Generative Adversarial Networks. *J. Environ. Manag.* **2024**, *359*, 120967.
- Wang, Z.; Wei, Y.; Wang, S. Forecasting the Carbon Price of China's National Carbon Market: A Novel Dynamic Interval-Valued Framework. *Energy Econ.* **2025**, *141*, 108107.
- Lim, B.; Arik, S.Ö.; Loeff, N.; Pfister, T. Temporal Fusion Transformers for Interpretable Multi-horizon Time Series Forecasting. *Int. J. Forecast.* **2021**, *37*, 1748–1764.
- Rudin, C. Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead. *Nat. Mach. Intell.* **2019**, *1*, 206–215.
- Ministry of Ecology and Environment of the People's Republic of China. Progress Report of China's National Carbon Market (2024). Technical Report, Ministry of Ecology and Environment of the People's Republic of China, 2024. Available online: <https://www.mee.gov.cn/ywdt/xwfb/202407/W020240722528850763859.pdf> (accessed on 26 April 2026).
- Shanghai Environment and Energy Exchange. National Carbon Market Trading Information Publishing Platform, 2026. Available online: <https://www.cneex.com/qgtpfqjy/mrgk/> (accessed on 26 April 2026).

15. International Carbon Action Partnership. China Releases 2024–2025 Allowance Allocation Plan for Industrial Sectors in National ETS. 2025. Available online: <https://icapcarbonaction.com/en/news/china-releases-2024-2025-allowance-allocation-plan-industrial-sectors-national-ets> (accessed on 10 June 2026).
16. Jorion, P. *Value at Risk: The New Benchmark for Managing Financial Risk*, 3rd ed.; McGraw-Hill: New York, NY, USA, 2007.
17. Rockafellar, R.T.; Uryasev, S. Optimization of Conditional Value-at-Risk. *J. Risk* **2000**, *2*, 21–41. <https://doi.org/10.21314/JOR.2000.038>.
18. Rockafellar, R.T.; Uryasev, S. Conditional Value-at-Risk for General Loss Distributions. *J. Bank. Financ.* **2002**, *26*, 1443–1471.
19. Bai, S.; Kolter, J.Z.; Koltun, V. An Empirical Evaluation of Generic Convolutional and Recurrent Networks for Sequence Modeling. *arXiv* **2018**, arXiv:1803.01271.
20. Koenker, R.; Bassett, G. Regression Quantiles. *Econometrica* **1978**, *46*, 33–50.
21. Koenker, R. *Quantile Regression*; Cambridge University Press: Cambridge, UK, 2005.
22. Morgan, J.P.; Reuters. *RiskMetrics Technical Document*; Technical Report; Morgan Guaranty Trust Company; New York, NY, USA, 1996.
23. Bollerslev, T. Generalized Autoregressive Conditional Heteroskedasticity. *J. Econom.* **1986**, *31*, 307–327.
24. Glosten, L.R.; Jagannathan, R.; Runkle, D.E. On the Relation Between the Expected Value and the Volatility of the Nominal Excess Return on Stocks. *J. Financ.* **1993**, *48*, 1779–1801.
25. Meinshausen, N. Quantile Regression Forests. *J. Mach. Learn. Res.* **2006**, *7*, 983–999.
26. Ke, G.; Meng, Q.; Finley, T.; Wang, T.; Chen, W.; Ma, W.; Ye, Q.; Liu, T.Y. LightGBM: A Highly Efficient Gradient Boosting Decision Tree. In *Advances in Neural Information Processing Systems*; Curran Associates, Inc.: Red Hook, NY, USA, 2017; Volume 30, pp. 3146–3154.
27. Hochreiter, S.; Schmidhuber, J. Long Short-Term Memory. *Neural Comput.* **1997**, *9*, 1735–1780.
28. Kupiec, P.H. Techniques for Verifying the Accuracy of Risk Measurement Models. *J. Deriv.* **1995**, *3*, 73–84.
29. Christoffersen, P.F. Evaluating Interval Forecasts. *Int. Econ. Rev.* **1998**, *39*, 841–862.
30. Engle, R.F.; Manganelli, S. CAViaR: Conditional Autoregressive Value at Risk by Regression Quantiles. *J. Bus. Econ. Stat.* **2004**, *22*, 367–381.
31. Fissler, T.; Ziegel, J.F. Higher Order Elicitability and Osband’s Principle. *Ann. Stat.* **2016**, *44*, 1680–1707.
32. Patton, A.J.; Ziegel, J.F.; Chen, R. Dynamic Semiparametric Models for Expected Shortfall and Value-at-Risk. *J. Econom.* **2019**, *211*, 388–413.
33. Grossman, S.J.; Stiglitz, J.E. On the Impossibility of Informationally Efficient Markets. *Am. Econ. Rev.* **1980**, *70*, 393–408.
34. Sims, C.A. Implications of Rational Inattention. *J. Monet. Econ.* **2003**, *50*, 665–690.
35. Peng, L.; Xiong, W. Investor Attention, Overconfidence and Category Learning. *J. Financ. Econ.* **2006**, *80*, 563–602.
36. Alberola, E.; Chevallier, J.; Chèze, B. Price Drivers and Structural Breaks in European Carbon Prices 2005–2007. *Energy Policy* **2008**, *36*, 787–797.
37. Hintermann, B. Allowance Price Drivers in the First Phase of the EU ETS. *J. Environ. Econ. Manag.* **2010**, *59*, 43–56.
38. Chevallier, J. A Model of Carbon Price Interactions with Macroeconomic and Energy Dynamics. *Energy Econ.* **2011**, *33*, 1295–1312.
39. Creti, A.; Jouvet, P.A.; Mignon, V. Carbon Price Drivers: Phase I versus Phase II Equilibrium? *Energy Econ.* **2012**, *34*, 327–334.
40. Koch, N.; Fuss, S.; Grosjean, G.; Edenhofer, O. Causes of the EU ETS Price Drop: Recession, CDM, Renewable Policies or a Bit of Everything?—New Evidence. *Energy Policy* **2014**, *73*, 676–685.
41. The State Council of the People’s Republic of China. China Releases Carbon Emissions Trading Regulations. 2024. Available online: https://english.www.gov.cn/policies/latestreleases/202402/04/content_WS65bf7f70c6d0868f4e8e3c94.html (accessed on 10 June 2026).
42. Zhang, N.; Sun, Q.; Yang, L.; Li, Y. Event-Triggered Distributed Hybrid Control Scheme for the Integrated Energy System. *IEEE Trans. Ind. Inform.* **2022**, *18*, 835–846.
43. Han, Z.; Fang, D.; Yang, P. Managing Coupled Electricity and Carbon Markets: A Multiagent Equilibrium Perspective. *IEEE Trans. Eng. Manag.* **2026**, *73*, 2597–2610.
44. Fan, J.H.; Todorova, N. Dynamics of China’s Carbon Prices in the Pilot Trading Phase. *Appl. Energy* **2017**, *208*, 1452–1467.
45. Yang, X.; Zhang, C.; Yang, Y.; Wang, W.; Wagan, Z.A. A New Risk Measurement Method for China’s Carbon Market. *Int. J. Financ. Econ.* **2022**, *27*, 1280–1290.
46. Duan, Y.; He, C.; Yao, L.; Wang, Y.; Tang, N.; Wang, Z. Research on Risk Measurement of China’s Carbon Trading Market. *Energies* **2023**, *16*, 7879.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.