UNIVERSITI MALAYSIA SARAWAK

UNIMAS

Faculty of Economics and Business

Risk Contagion and Price Forecasting of China Carbon Market Based on High Order Moment Attribute

Ni Li

Doctor of Philosophy
2025

Risk Contagion and Price Forecasting of China Carbon Market Based on High Order Moment Attribute

Ni Li

A thesis submitted

In fulfillment of the requirements for the degree of Doctor of Philosophy

(Finance)

Faculty of Economics and Business
UNIVERSITI MALAYSIA SARAWAK
2025

DECLARATION

I declare that the work in this thesis was carried out in accordance with the regulations of Universiti Malaysia Sarawak. Except where due acknowledgements have been made, the work is that of the author alone. The thesis has not been accepted for any degree and is not concurrently submitted in candidature of any other degree.

NI LI

……………………………

Signature

Name:	Ni Li
Matric No.:	21010257

Faculty of Economics and Business

Universiti Malaysia Sarawak

Date :13/5/2025

ACKNOWLEDGEMENT

Firstly, I would like to express my sincere gratitude to the Universiti Malaysia Sarawak (UNIMAS) for providing me with this opportunity to pursue my doctoral studies, and now this agonizing but rewarding journey of learning is coming to an end.

I am very grateful to my main research supervisor Professor Dr Liew Khim Sen for his professional guidance, valuable suggestions and constructive comments throughout my research and thesis writing, Especially his rigorous academic attitude deeply influenced me,which are very beneficial for the research process and subsequently led to the completion of this PhD thesis. To my co-supervisor, Associate Professor Dr Shirly Wong Siew Ling, thank you for your moral support and motivation throughout my PhD journey.

I would also express my gratitude to all teachers of FEB Postgraduate Programme and all the staff of Centre for Graduate Studies (CGS) for their support and kindness in guiding me especially in students affairs.

Finally, I must express my profound gratitude to my parents and my husband for their material and spiritual support, especially my husband for providing me with unfailing support, unending encouragement, and motivation throughout my years of PhD study and the process of researching and writing this thesis. Without his endless guidance and support in every process, I don't have so much time to complete my doctoral thesis. I also want to thank my workplace and my classmates, because while working, I usually pursue a PhD and need to balance many things. Friends encourage and learn from each other together, which is a special experience that has given me more spiritual motivation encouragement.

ABSTRACT

As an effective mechanism for addressing climate issues, the carbon market plays an important role in reducing global greenhouse gas emissions. The core function of the market is achieving emissions reduction target by the manner of market price mechanism. So, the carbon price is the key point. This thesis focuses on studying the China carbon price forecasting and its price driving mechanism after considering the impact of high order risk contagion relationship, which supports a more convinced and innovative evidence for explaining the carbon price formation that is significantly different from previous researches.

In terms of the research variables, the carbon market has closely information linkages and spillovers with the capital market, homogeneous product market, and energy market products. In this regard, this thesis uses the daily transaction price of Hubei Carbon Emission Allowances (HBEA) as the representative indicators of China carbon price, takes the daily settlement price of the European Carbon Emission Allowance Futures (EUAF) as the representative variables of carbon homogeneous market, selects the daily trading price of Jiaotan futures (JTF), Jiaomei futures (JMF),and Brent crude oil futures (Oil) as the special variables of energy market, and selects the daily trading price of China security index 300 (CSI300) as the special variables of capital market.

In terms of model design and empirical discussion, firstly, this thesis measures the risk contagion relationship between the carbon market and its infected markets based on the idea of Markov theory, then designs a carbon price state transition model to classify the high, medium and low volatility state of the carbon market. Secondly, constructs and tests the risk contagion channels between carbon price and its infected markets, that is the low order moment risk contagion channel of Forbes Rigobon contagion (FR), and the high order moment risk contagion channels of Co-Skewness (CS), Co-Kurtosis (CK) and Co-Volatility (CV). And finally, designs a high order risk contagion carbon price forecasting model (HOC-LSTM) to forecasting the price of China carbon market.

The main conclusion of this thesis are as follows: firstly, the high order risk contagion carbon price forecasting framework support a convinced theoretical support for forecasting the China carbon price. Secondly, there is no risk contagion relationship in low order moment channels, but significant risk contagion relationship in high order moment channels no matter the carbon market in rapid and slow change. Thirdly, the HOC-LSTM model constructed in this thesis has a significant superiority in forecasting China carbon price then other comparative models, such as the Gated recurrent unit (GRU), Multi-Layer Perceptron(MLP), Gradient Boosting Decision Tree (GBDT), Extra Trees Regressor (ETR) and Back Propagation Neural Network (BPNN), the high order risk contagion channels are indispensable factors for explaining carbon price formation mechanism. Those results can not only provide reference for investors and emission reduction entities to make investment and financing decisions, analyze price trends, but also contribute to technical references for government departments to promote the construction of carbon market pricing mechanisms and market efficiency.

Keywords: China carbon market, risk contagion, price forecasting, high order moment, HOC-LSTM

PENULARAN RISIKO DAN RAMALAN HARGA DALAM PASARAN KARBON NEGARA CINA BERDASARKAN ATRIBUT MOMEN TERTIB TINGGI

ABSTRAK

Sebagai mekanisme yang efektif untuk mengatasi masalah iklim, pasaran karbon bermain peranan penting dalam pengurangan pelepasan gas rumah hijau secara global. Fungsi utama pasaran adalah mencapai sasaran pengurangan pelepasan gas melalui mekanisme harga pasaran. Jadi, harga karbon merupakan perkara utama. Tesis ini fokus pada mempelajari ramalan harga karbon negara Cina dan mekanisme penentuan harganya selepas mempertimbangkan kesan dari hubungan penularan risiko tertib tinggi, yang menyokong bukti yang lebih meyakinkan dan inovatif untuk menjelaskan penentuanan harga karbon yang ternyata berbeza berbanding dengan kajian sebelumnya.

Dari segi pembolehubah kajian, pasaran karbon mempunyai hubungan maklumat yang dekat dan limpahan dengan pasaran modal, pasaran produk homogen, dan pasaran produk tenaga. Dalam hal ini, tesis ini menggunakan harga transaksi harian Keizinan Pelepasan Karbon Hubei (HBEA) sebagai penunjuk yang mewakili harga karbon di negara Cina, sementara itu harga penyelesaian harian Keizinan Pelepasan Karbon di Eropah (EUAF) dianggap sebagai pembolehubah yang mewakili pasaran homogen karbon, manakala harga perdagangan harian bagi masa depan Jiaotan (JTF) dan masa depan Jiaomei (JMF) dianggap sebagai pembolehubah khas pasaran tenaga, dan harga perdagangan harian indeks China 300 (CSI300) digunna sebagai pembolehubah khas pasaran modal.

Dari segi rancangan model dan perbincangan empirik, pertama-tama, tesis ini mengukur

hubungan penyebaran risiko antara pasaran karbon dan pasaran yang terjejas olehnya berdasarkan idea Teori Markov, kemudian merancang model transaksi harga karbon negara untuk pengelasan keadaan kemeruapan tertib tinggi, tengah dan rendah pasaran karbon. Kedua, membina dan menguji saluran penularan risiko di antara harga karbon dan pasaran yang terpengaruh olehnya, iaitu saluran penularan risiko tertib rendah Forbes Rigobon (FR), dan saluran penularan risiko tertib tinggi Kecondongan Bersama (CS), Kurtosis Bersama (CK) dan Kemeruapan Bersama (CV). Dan akhirnya, kajian ini merancang model ramalan harga karbon (HOC-LSTM) untuk meramal harga pasaran karbon negara Cina.

Kesimpulan utama tesis ini adalah seperti berikut: pertama, rangkaian ramalan harga karbon penularan risiko tertib tinggi memberi sokongan teori yang menyakinkan untuk meramalkan harga karbon negara Cina. Kedua, tiada hubungan penularan risiko dalam salurantertib rendah, tetapi hubungan penularan risiko yang signifikan dalam saluran tertib tinggi telah dikesani dalam pasaran karbon pada frasa perubahan cepat dan lambat. Ketiga, model HOC-LSTM yang dibina dalam tesis ini mempunyai kelebihan yang signifikan dalam meramalkan harga karbon negara Cina berbanding dengan model pesaing lain, seperti Unit Berulang Berpagar, Perceptron Berbilang-Lapisan, Pohon Keputusan Membentuk Gradien, Regressor Ekstra Pohon dan Rangkaian Neural Penyedaran Belakang. Hasil kajian tersebut bukan sahaja memberikan rujukan bagi pelabur dan pihakpengurangan pelepasan gas rumah hijau untuk membuat keputusan pelaburan dan kewangan, serta menganalisa halatuju harga, tetapi juga menyumbangkan rujukan teknikal kepada jabatan kerajaan untuk mempromosikan pembangunan mekanisme penentuan harga pasaran karbon dan kecekapan pasaran.

Kata kunci: *Pasaran karbon Cina, penularan risiko, ramalan harga, momen terbtinggi, HOC-LSTM*

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

%	Percentage
°C	Degree Celsius
GDP	Gross Domestic Product
CO_2	Carbon Dioxide
NOAA	National Oceanic and Atmospheric Administration
EUCM	European Carbon Market
JI	Joint Implementation
IET	International Emission Trading
CDM	Clean Development Mechanism
EU ETS	European Union Emissions Trading System
EUA	European Union Allowance
ICAP	International Carbon Action Partnership
EEA	European Economic Area
MSR	Market Stability Reserve
NDRC	National Development and Reform Commission
CEA	Chinese Emission Allowances
CVER	Certified Voluntary Emission Reductions
NUCM	National Unified Carbon Market
HBEA	Hubei Carbon Market Emission Allowances
CMP	Carbon Market Pricing
UNGA	United Nations General Assembly
CEWC	Central Economic Work Conference
CAPM	Capital Asset Pricing Model
LSTM	Long and Short-Term Memory

HOC	High Order risk Contagion
EMH	Efficient Market Hypothesis
WFEM	Weak Form Efficient Market Hypothesis
SSEMH	Semi-Strong Efficient Market Hypothesis
SFEMH	Strong Form Efficient Market Hypothesis
CML	Capital Market Line
SML	Security Market Line
APT	Arbitrage Pricing Theory
CPT	Cumulative Prospect Theory
US	United States
FAP	Financial Asset Pricing
ANN	Artificial Neural Network
GRU	Gate Recurrent Unit
SVM	Support Vector Machine
ARCH	Autoregressive Conditional Heteroscedasticity
GARCH	Generalized Autoregressive Conditional Heteroscedasticity
SV	Stochastic Volatility
NYSE	New York Stock Exchange
EEMD	Ensemble Empirical Mode Decomposition
UN	United Nations
CER	Certified Emission Reduction
DMA	Dynamic Mean Analysis
TVP	Time-Varying Parameter
NARDL	Nonlinear Auto-Regressive Distributed Lag
LSSVM	Least Squares Support Vector Machine
PSO	Particle Swarm Optimization

GA	Genetic Algorithm
SNN	Spiking Neural Network
DNN	Deep Neural Network
IWOA	Improved Whale Optimization Algorithm
CEEMD	Complementary Ensemble Empirical Mode Decomposition
VMD	Variational Mode Decomposition
IMF	Intrinsic Mode Function
KDE	Kernel Density Estimation
IDWT	Interval Discrete Wavelet Transform
IEMD	Interval Empirical Mode Decomposition
IVMD	Interval Variational Mode Decomposition
FMRVR	Fast Multi Output Correlation Vector Regression
MOWOA	Multi Objective Whale Optimization Algorithm
MLP	Multi-Layer Perceptrons
BP	Back Propagation
GMDH	Group Method of Data Handling
CIT	Computational Intelligence Techniques
ANFI	Adaptive Neural Fuzzy Inference
FDL	Finite Distribution Lag
RRA	Ridge Regression Algorithm
MR	Mode Reconstruction
OCM	Optimal Combinatorial Model
GNN	Grey Neural Network
CS	Co-Skewness
CK	Co-Kurtosis
CV	Co-Volatility

EUAF	European Carbon Allowance Futures
CSI 300	CSI 300 Index
Oil	Brent Crude Oil Futures
JMF	Jiaomei Futures
JTF	Jiaotan Futures
FR	Forbes Rigobon Contagion
RNN	Recurrent Neural Network
CNN	Convolutional Neural Network
RMSE	Root-Mean-Square Error
MAE	Mean Absolute Error
ADF	Augmented Dickey-Fuller
JB	Jarque-Bera
BDS	Brlck Dechert Sheinkman
Std.Dev	Standard Deviation
QQ	Quantile Quantile
AIC	Akaike Information Criterion
BIC	Bayesian Information Criterion
HQ	Quinn Criterion
GBDT	Gradient Boosting Decision Tree
CART	Classification And Regression Tree
MART	Multiple Additive Regression Tree
RF	Random Forest
ETR	Extra Trees Regressor

CHAPTER 1

INTRODUCTION

1.1 Introduction

The carbon market, as an effective mechanism for addressing environmental issues, plays an important role in reducing global greenhouse gas emissions. The carbon market solves environmental problems through its function of market mechanisms. Generally, the government usually sets certain emission targets and grants a certain amount of carbon allowance to high emission enterprises. If enterprises use advanced emission reduction technologies to reduce emissions, they can sell the saved allowance in the carbon market to obtain economic benefits (Chevallier, 2009). While enterprises with insufficient emission allowance can purchase the required proportion in the market, for which they need to pay costs for high emissions. With the increase of companies participating in the carbon market, the incentive performance of the market becomes more significant, thereby encouraging the entire economy to achieve low-carbon goals.

The core function of the carbon market is achieving emissions reduction target by the manner of market price mechanism. As a result, the topic of carbon market price formation and driving are crucial. Furthermore, the price forecasting of carbon market in uncertain complex market condition has been regarded as the core of carbon market price research. Based on this, this thesis focus on studying the China carbon price forecasting mechanism and its price driving mechanism after consider the impact of high order risk contagion relationship on the carbon market, which support a more convinced and innovative evidence for explaining the carbon price formation that significant different from previous research. Based on the designed theoretical framework, this thesis uses long and short-time memory (LSTM) neural network to construct a model for forecasting carbon prices, solve the price forecasting issues of China carbon price that consider high order risk contagion channels.

According to research design, the framework of the first chapter is as follows. Firstly, this thesis analyzes the origin of the environmental problems, expounds the evolution of European carbon market (EUCM) and China carbon market, introduces the topic of risk contagion and price forecasting issues for aligning with the theme of this thesis. Secondly, this thesis suggests some specific research problems in the filed of carbon price risk contagion and price forecasting. Finally, proposes the general objective and special objective in different points. Furthermore, the contribution and some new innovations for solving the above research problems are suggested.

1.2 Study Background

1.2.1 Origin of the Environmental Problem

The environment externality theory was proposed by the famous economist Pigou in 1960, the key points of this theory were property rights and trading costs. As we known, the environment externalities have two types, that one is the positive externality, and another is the negative externality. Specifically, the negative externality can be produced when the behavior of a producer or consumer result in harm to others in society, while they do not provide any compensation for this damage. On the contrary, it creates positive externality. Actually, the environmental pollution is a typical manner of negative externality. The main reason is that the polluting enterprises have emitted pollutants into the atmosphere for a long time, and the world has paid a huge cost to digest this gas, including human health and environmental pollution, while those industrial enterprises do not take responsibility for this damage. Among them, the sharp increase in greenhouse gas emission is direct cause of environment negative externalities.

Since the industrial revolution in the 18th century, accompanied by advancements in production technology and a rise in fossil energy consumption, the world economy has achieved great development and created remarkable achievements. According Bradford DeLong, an economist at the University of Berkeley in the United States, the average annual growth of the global economy was only 0.1% before the industrial revolution. While, from the 18th century to end of the 20th century, the global per capita GDP increased by nearly 37 times (DeLong,2022). On the one hand, economic development has profoundly altered the production methods and lifestyles of human society, promoting the progress of social civilization. On the other hand, excessive resource extraction and consumption have also produced severe environmental problems. If this dilemma does not receive global attention and resolution, the global economic development will be difficult to maintain.

Although some environmentalists, such as Chevallier (2009) and Kim et al.(2010) commonly accepted that the moderate carbon dioxide proportion is an important component for stabilizing the global climate, it is evident that this share is undergoing significant changes due to human activities. According to the World Bank (2023), the global carbon emissions of CO_2 in 2022 are 36.8 billion tons, which had increased by 2.3 times compared to the level of 11.183 billion tons in 1965, with per capita carbon emissions increasing by nearly 68% (as shown in Figure 1.1).

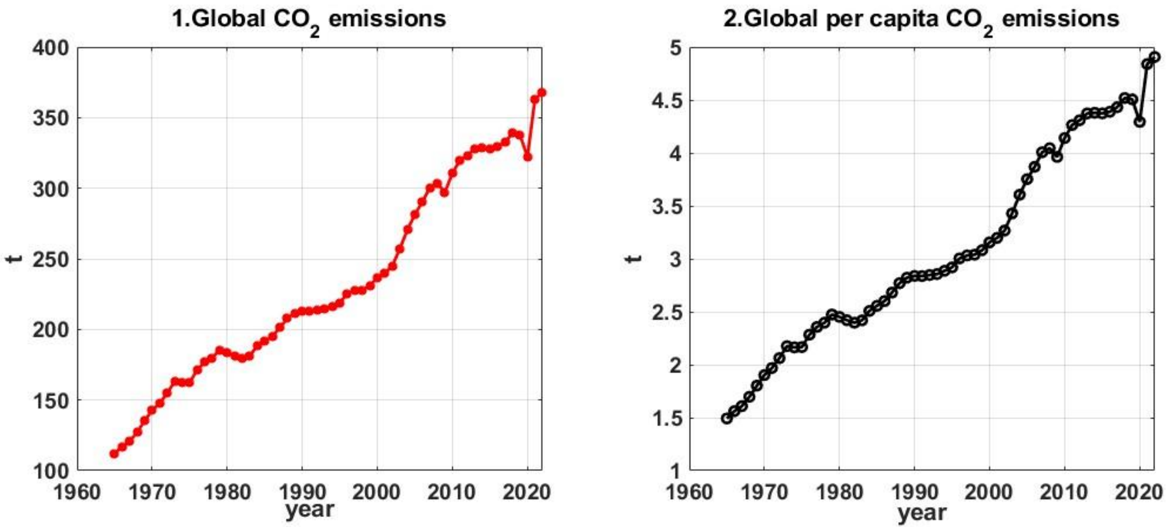


Figure1.1: The Curve of Main Greenhouse Gas Emission in 1965-2022

Source: Wind Database (2023)

In particular, with the rapid economic development over the past half century, the emission of various pollutant gases has produced a significant increase in the concentration of atmospheric pollutants. According to the latest data from the National Oceanic and Atmospheric Administration (NOAA) of the United States (Figure 1.2), global atmospheric CO_2 concentration rose from 316.91 ppm in 1960 to 417.1 ppm in 2022, with a significant increase of 31.6% and an average annual growth rate of 0.5%. With sharp increase in global carbon emissions, the average temperature of the earth has risen by 1.1°C in the past two hundred years. The increase of global temperature will not only trigger serious environmental problems, but also generate more serious social, economic and political problems. Therefore, for promoting the sustainable development, effectively curbing climate problems, suppressing global warming and reducing greenhouse gas emissions have become key issues that need to be urgently addressed by the global community.

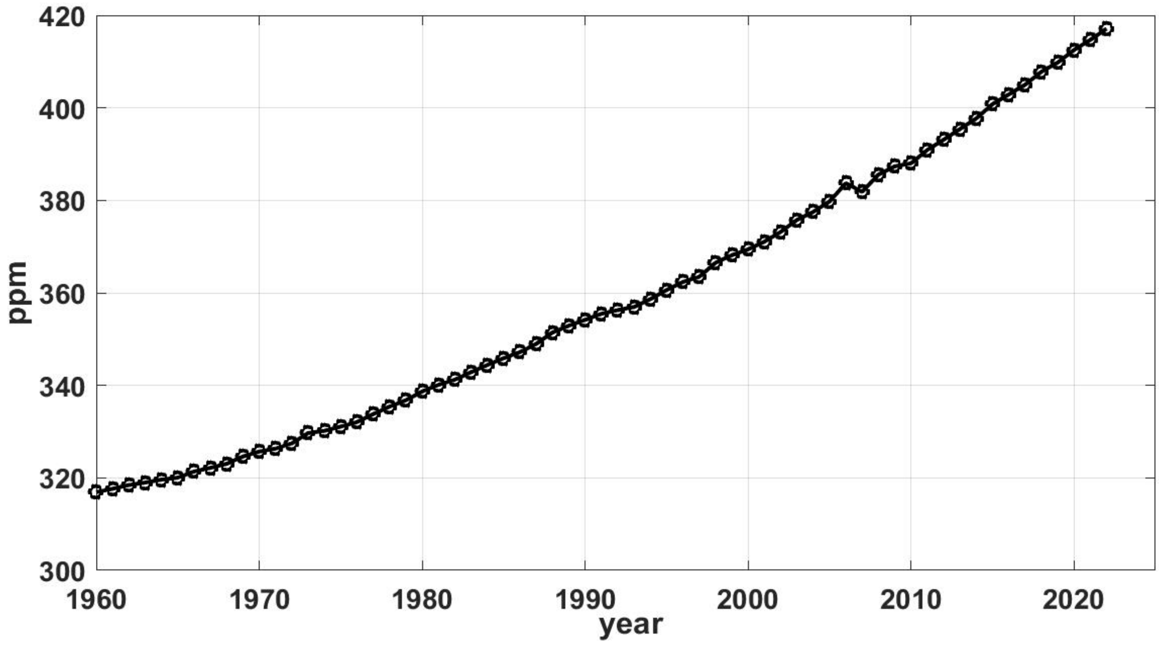


Figure1.2: The Curve of Global CO_2 Emission Concentration in 1960-2022

Source: Wind Database (2023)

1.2.2 European Carbon Market for Solving the Environmental Problems

1.2.2.1 The Evolution of European Carbon Market

The United Nations Framework Convention on Climate Change (UNFCCC), promoted in 1992 during the United Nations Conference on environment and development, established carbon market as an effective market mechanism to solve the climate problems and reduce global greenhouse gas emissions. Over 150 countries participated in formulating the convention. The main objective of this convention is to stabilize greenhouse gases concentration at a level that does not harm the climate system, and to achieve global sustainable development by suppressing greenhouse gas emissions. To construct common but differentiated responsibilities, the convention requires developed countries that with more greenhouse gas emissions to take positive measures to reduce pollution gas emissions, while developing countries do not bear the legal responsibility for emission reduction in the near future. The greatest contribution of this convention is providing the basic framework of international cooperation to solve climate and environmental problems.

The promotion of the Kyoto Protocol in 1997 is a supplementary clause to the UNFCCC. The Kyoto Protocol, adopted in 1997, further stipulates implementation of greenhouse gas emission reduction responsibilities by developed countries, especially in terms of the schedule and allocation of emission reduction quotas. To facilitate the implementation of national emission reduction actions, the Kyoto Protocol has established three market-based mechanisms for greenhouse gas emission reduction, namely Joint Implementation (JI), International Emission Trading (IET) and Clean Development Mechanism (CDM) (as shown in Table 1.1).

Table 1.1: Market Characteristics of the Phase I to Phase IV of the EUCM

	Object	Greenhouse gas	Scope covered	Industries covered	Total control	Quota method	Regulatory mechanism
Phase I	Establishing carbon market infrastructure	CO_2	European states	Power plants over 20MW,oil refining,coking,iron and steel, cement, glass,lime, brick making, ceramics,paper making industries	2.058 billion tons of CO_2	Free allocation of 95% quota	JI, CDM
Phase II	Reduce emissions by 8% on the basis of 1990	CO_2	European states, Norway, Iceland, Liechtenstein	Newly added aviation industry	1.859 billion tons of CO_2	Free allocation of 90% quota	JI, CDM
Phase III	Reduce emissions by 21% on the basis of 2005	CO_2, N_2O, PFC	European states, Norway, Iceland, Liechtenstein	Newly added industries such as aluminum production, petrochemical industry, ammonia production,nitric acid, oxalic acid, and aldehyde acid production,carbon capture, pipeline transportation, and underground storage of carbon dioxide	Linear decrease of 1.74% per year compared to 2013 levels	100% auction in the power industry, 80% free allocation in manufacturing industry in 2013, and reach 30% by 2020	JI, CDM
Phase IV	Reduce emissions by 43% on the basis of 2005	CO_2, N_2O, PFC	European states, Norway, Iceland, Liechtenstein	Same with Phase III	Linear decrease of 2.2% per year compared to 2020 levels	100% auction in the power industry, free allocation of 34% quota, and reach 0% by 2026	No offsetting allowed

Source: China Emissions Trading Network (2023)

Among them, the JI mechanism mainly refers to the transfer of emission reduction units achieved by developed countries to another developed country through project level cooperation, but at the same time, corresponding quotas must be deducted from the transferor's allocated quantity units. IET mechanism refers to the practice of developed countries purchasing emission allowances from another developed country with excess emission quotas when they may exceed their greenhouse gas emissions, in order to fulfill their emission reduction commitments, and at the same time deducting the corresponding transfer quota from the transferor's allowed emission quota. The CDM refers to the project level cooperation between developed countries and developing countries through the provision of funding and technology. The "Certified Emission Reductions" (CERs) achieved through projects are used by developed country Parties to fulfill their commitments under the Protocol. The implementation of the above three measures essentially gives the carbon emission right with commodity attribute, and promotes the formation of the global carbon market. The European carbon market (EUCM), the representative market is the European Union Emissions Trading System (EUETS), established in 2005, stands as the world's pioneering carbon market. With the development of nearly 20 years, the EUETS has become an famous carbon market with the largest trading volume, the strongest liquidity and mature market mechanism. The EUETS follows the "top-down" quota allocation principle, and each emission reduction entity determines the initial allocation of emission allowances, that is, the European Union Allowance (EUA). If the actual carbon emission quota is less than the allocated part, the excess quota can sell in the carbon market. On the contrary, if there is a shortfall, the entities can also buy the quota from the market.

The EUETS has established four stages to promote the development of carbon market, that is the first stage (from 2005.1.1 to 2007.12.31) belongs to the experimental stage, that is the informal trading stage with the aims to accumulate operational experience in the carbon market and does not mandate the achievement of emission reduction., and the emission reduction gas is limited to carbon dioxide. In the second stage (from 2008.1.1 to 2012.12.31), emissions reduction will be expanded to other greenhouse gases, including sulfur dioxide and chlorofluorocarbons, and the transportation industry will be included in the emission reduction range, the target is achieving a 19% reduction in the carbon emissions compared to 1980. In the third stage (from 2013.1.1 to 2020.12.31), the aviation industry had be included in the emission reduction scope, and the goal is reducing the total carbon emission by 20% compared with 2005. The important characteristic of the fourth stage (from 2020.1.1 to 2030.12.31) is the implementation of stricter emission reducing rules. At this stage, the European carbon market requires 2.2% reduction in the total annual quota and cannot use carbon credits to offset it.

The Paris Agreement that adopted at the 2015 Paris Climate Conference is the latest supporting consensus for the European carbon market. The Paris Agreement clearly puts forward two basic objectives, the first goal is to limit global temperature within 2 ℃ in this century and strive to control the temperature rise within 1.5 ℃. The second is achieving zero net global greenhouse gas emissions in the latter half of this century. The Paris Agreement engages all nations in a collective effort to protect the Earth's ecology. Although the European carbon market operation mechanism remain regards the trading rules such as CDM, JI and IET as the market core, the market governance pattern has from the "bottom-up" to "top-down". It is worth noting that, the bottom-up carbon allowance allocation method was implemented after the establishment of the European carbon market in 2005, which determines the emission allowance based on the actual carbon emissions of each enterprise. However, the entire region has not set a total emission limit. The practice of this method has led to excessive allocation and a decrease in carbon prices. Conversely, the top-down plan was implemented in 2008, and the allowance for this plan was uniformly set by the state, with enterprises developing their own allowance plans within a limited scope. This measure greatly improved the economic attributes of carbon allowance, and carbon prices steadily increased, providing the correct price signal for enterprises to reduce emissions.

1.2.2.2 The Price Trend of the European Carbon Market

The EUETS serves as the cornerstone of the European policy framework for combating climate change. According to data from the International Carbon Action Partnership (ICAP) in 2020-2021, EUETS covered approximately 36% of total emissions in the European Economic Area (EEA). The EUCM has played a positive role in reducing carbon emission in Europe. Since the market's formal operation, the carbon price trend has been upward, and has experienced several big fluctuations (as shown in Figure 1.3).

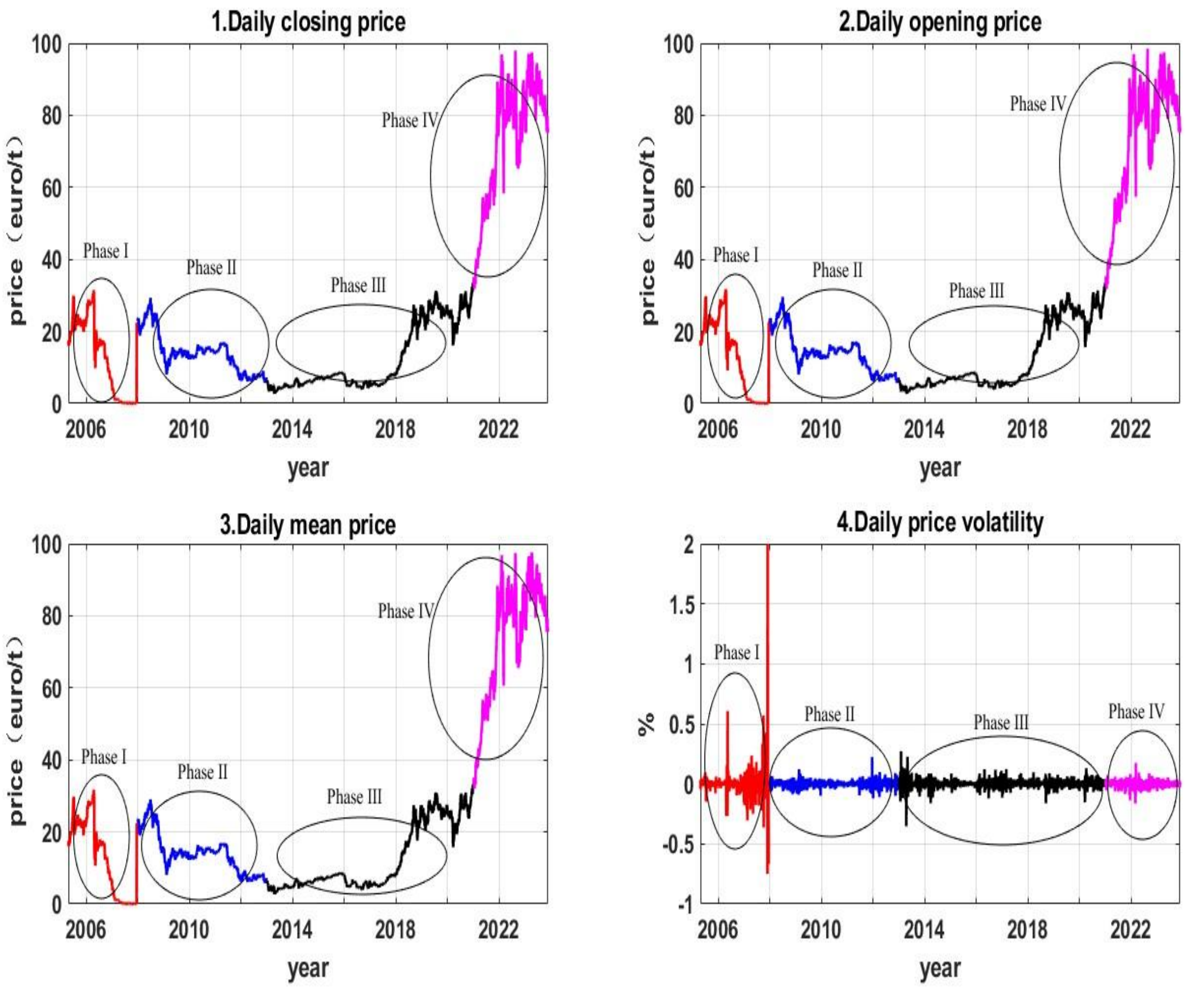


Figure 1.3: Prices and Volatility of EUA Continuous Futures Contracts in the EUCM
Source: European Energy Exchange(2023)

Specifically, in the first stage (from 2005 to 2007), the European carbon price has continued to fall. The EUA price continued to fall, dropping to almost zero in 2007. In the second stage (from 2008 to 2012), the carbon price remained consistently low. In particular, because of the 2008 financial crisis and immature carbon market mechanism, the emissions form the European corporate have significantly decreased, which resulting in a consistently low carbon price. In the third stage (from 2013 to 2020), the European carbon price has increased. Especially, the Backloading mechanism adopted in 2014, and then formally started to implement the Market Stability Reserve (MSR) in 2019, boosted market confidence and increased carbon price, exceeding 20 euros per ton. Actually, the Backloading mechanism mainly ensures that the carbon market's emission reduction targets match the macro emission reduction targets by adjusting the total allowance. For example, the emission reduction companies voluntarily purchase carbon credits to support emission reduction projects and offset their carbon emissions. In the fourth stage (from 2021 to 2030), the European carbon price has significantly increased, the possible reason is the ambitious emission reduction targets that Europe aims to achieve in 2023.

1.2.3 China Carbon Market for Solving the Environmental Problems

1.2.3.1 The Carbon Emission in China

Since the implementation of the reform and opening-up policy in 1980s, the Chinese economy has officially entered a stage of rapid development, especially from 2003 to 2010, where the growth rate of the Chinese economy remained at a level of around 10%. According the National Bureau of Statistics of China(2025), China's GDP grew from 362.4 billion yuan to 13.4 trillion yuan from 1978 to 2024, with an average annual economic growth rate of 8.9%, far higher than the world's average economic growth rate of 3% during the same period. China's GDP in 2024 was 13.4 trillion yuan, calculated at constant prices, an increase of 5.0% compared to the previous year. The total economic output ranks second in the world, and the economic growth rate ranks among the top in the world's major economies, making it an important driving force for global economic growth. Taking 2024 as an example, the added value of China's primary industry is 9.1414 trillion yuan, an increase of 3.5% over the previous year. The added value of the secondary industry is 4920.87 billion yuan, an increase of 5.3%. The added value of the tertiary industry was 765583 billion yuan, an increase of 5.0%. In general, the total economic output accounted for over 18% of the global economy, contributing 32% to global economic growth, and its per capita income was close to the high-income countries.

However, behind the rapid economic development, there are also concerns about carbon emissions. In fact, in the past two decades, with the economy development, China's carbon emissions have gradually increased, especially since 2007, the total carbon emissions have surpassed the United States for the first time, that make China becoming the world's largest emitter of CO_2. The excessive reliance on traditional fossil energy consumption is the primary cause for the increase in carbon emissions. That is to say, the China's economic growth miracle is largely at the cost of sacrificing the environment. Despite the Chinese government's increasing attention to environmental issues in recent years, and the dependence of economic growth on energy has also gradually decreased, the high carbon emissions in the short term still exists.

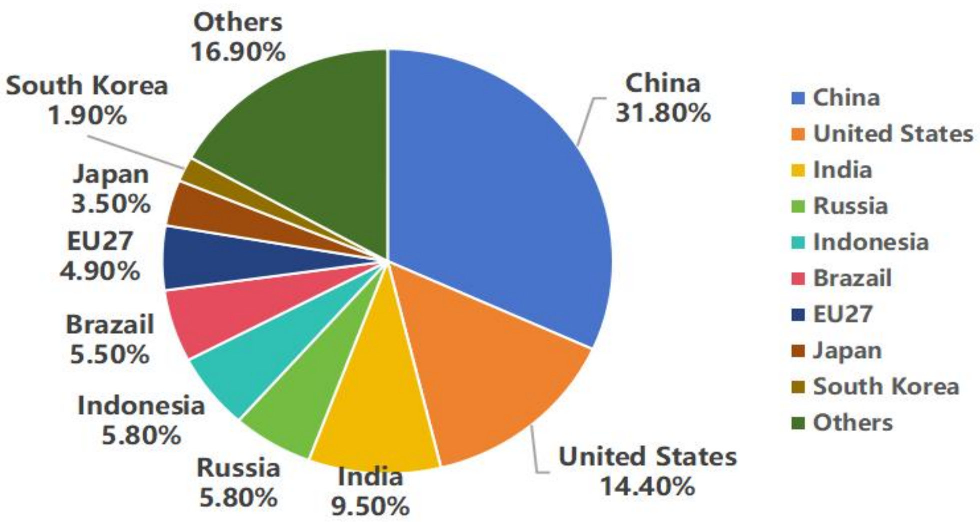


Figure 1.4: Country-by-Country breakdown of carbon emission for 2023

Source:National Bureau of Statistics of China (2024) and International Energy Agency (2024)

According the National Bureau of Statistics of China (2024) and The International Energy Agency (IEA,2024), the global energy related carbon dioxide emissions reached 37.4 billion tons in 2023, an increase of 1.1% compared to the previous year. There are four noteworthy observations. Firstly, the CO_2 is the largest emissions of the country, and China's carbon emissions reached 12.6 billion tons, accounting for 31.8% of the world's level. The second is the United States, with the carbon emissions of 4.85 billion tons in 2023, accounting for 14.4%. That is, China's carbon emissions still account for the highest proportion in the world (as shown in Table 1.2). India's global carbon emissions account for 9.5%, Russia and Indonesia both account for 5.8%, Brazil accounts for 5.5%, 27 European countries account for 4.9%, Japan accounts for 3.5%, South Korea accounts for 1.9%, and the remaining countries and regions account for 16.9%.

Secondly, the carbon emission growth rates of China from 2020 to 2022 were 0.6%, 0.6%, and 5.3%, respectively. From the perspective of carbon emissions sources, according the Shanghai Environment and Energy Exchange (2024), China's carbon emissions in 2023 mainly come from the energy sector, including energy supply and energy consumption industries, which account for 77% of the country's total emissions. Meanwhile, industrial process carbon emissions account for 14%, agricultural sector carbon emissions account for 7%, and waste carbon emissions account for 2%. Thirdly, China's carbon emission strength from 2020 to 2022 were 1.14, 1.11, and 1.03, respectively. Carbon emission strength is the ratio of total carbon emissions to GDP. The smaller the indicator, the more GDP output can be obtained under a given carbon emission level. Data from the Table 1.2 shows that since 2020, China's carbon emission strength has experienced a downward trend. Fourthly, from 2020 to 2022, China's per capita carbon emissions were 7.96, 8.0, and 8.42 tons per person, respectively, with a clear upward trend in per capita carbon emissions.

Table 1.2: Carbon Emission and its Growth Rate of China in 2020-2022

Items	2020	2021	2022
Carbon emission (Billion tons)	112.2	112.9	118.9
Growth rate of Carbon emission	0.6%	0.6%	5.3%
Carbon emission strength (Tons/10000 yuan)	1.14	1.11	1.03
Growth rate of Carbon emission strength	-6.2%	-2.1%	-7.2%
Per capita carbon emission (Tons per person)	7.96	8.0	8.42
Growth rate of per capita carbon emission	-0.4%	0.5%	5.3%

Source: National Bureau of Statistics of China Government (2023)

Therefore, to reduce the negative environmental impact of economic growth and maintain long-term high-quality economic development, taking effective measures to reduce carbon emissions has become a key issue that China urgently needs to solve. Among various means, relying on market-oriented carbon markets has become an significant breakthrough for achieving this goal, and has received strong support from Chinese government.

1.2.3.2 Mechanism Design of China Carbon Market

The institutional design of China's carbon market depends on the economic policies and energy policies currently implemented in China.The carbon peak and carbon neutrality strategy proposed by the Chinese government in September 2020 is the top-level design for formulating energy policies and promoting carbon market construction. The strategy of Carbon peak means the carbon dioxide emissions situation reach their peak and no longer increase. This means that China aims to achieve a peak in total carbon dioxide emissions before 2030, after which they will gradually decrease. The strategy of Carbon neutrality refers to China offsetting its own carbon dioxide emissions through afforestation, energy conservation and emission reduction, and industrial adjustment before 2060. Based on this, China's economic policies have also undergone adjustments, and economic growth is no longer based solely on the extensive model of pursuing economic growth, but on a high-quality growth model that places more emphasis on green and low-carbon development. To promote the reduction of carbon emission rights, the Ministry of Ecology and Environment began implementing the "Management Measures for Carbon Emission Trading (Trial)" in February 2021, and officially launched the construction of a national unified carbon emission trading market in July 2021.

Compared to the European carbon market, China carbon market development started relatively late. In 2011, the National Development and Reform Commission (NDRC) launched pilot carbon trading operations in seven regions: Beijing, Shanghai, Tianjin, Chongqing, Hubei, Guangdong and Shenzhen. In 2016, Fujian became the eighth pilot region. In the same year, the Sichuan carbon market also opened. Furthermore, the national unified carbon market was officially established in July 2021. It is estimated that annual industry emissions exceeds 4 billion tons of CO_2, that is China carbon market has become the world's largest in terms of emissions coverage.

The national unified carbon market is divided into two parts, one is the mandatory quota market, another is the voluntary emission reduction market. As for the mandatory quota market, the trading product is Chinese Emission Allowances(CEA), also known as carbon emission quotas. The mandatory quota market has developed since the regional pilot carbon market, that focusing initially on the power generation industry. The regional pilot and the national unified carbon market follow the same design principle, while the differences are mainly in the system design of the quota allocation. It is worth noting that the operation of national unified carbon market does not mean closure of pilot markets. Following the establishment of the national unified carbon market, the regional pilot market is still operating synchronously.

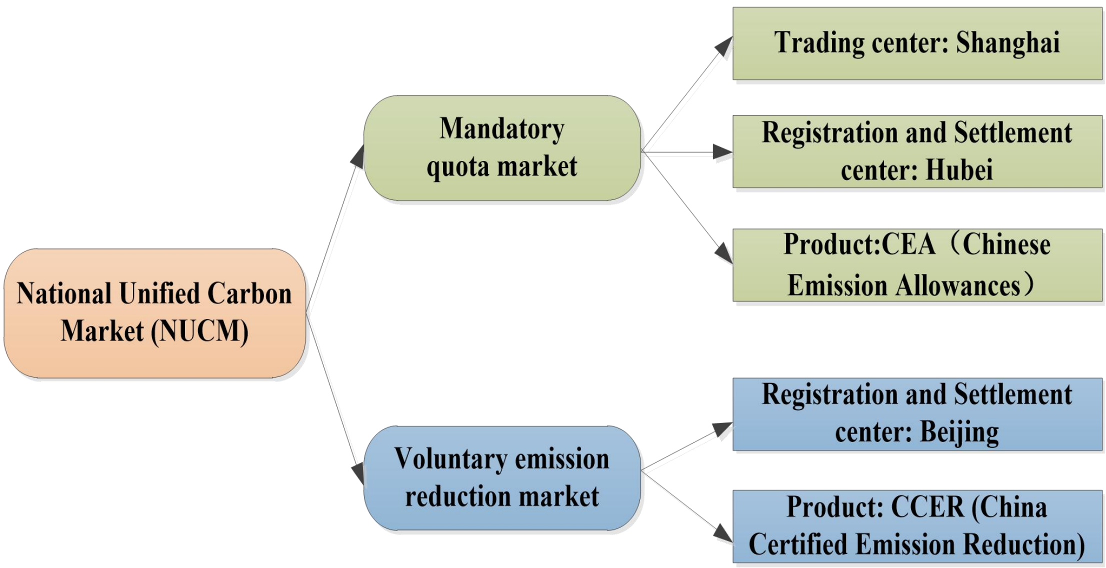


Figure 1.5: Policy Design Framework of the China NUCM

Source: China Carbon Emissions Trading Network (2024)

China carbon market adopts a "dual-centre" model, where Hubei is responsible for the construction of the national unified carbon market registration system, while Shanghai manages the operation of carbon trading system. The two regions jointly play a pillar role in the national unified carbon market (as shown in Figure 1.5). Among them, the registration system serves as the "warehouse" for storing carbon funds, and it undertakes the registration, settlement and cancellation of carbon price, and supervises registration system and its management institutions. Furthermore, Beijing Green Exchange is responsible for constructing the National Voluntary Emission Reduction Trading Centre (NVERTC), which serves as the national platform for future Certified Voluntary Emission Reductions (CVER) trading. This initiative encourages enterprises that do not bound by mandatory emission reduction obligations to develop greenhouse emission reduction projects. Emission-control enterprises taking part in the national carbon emissions trading market can also use CVER as a supplementary compliance method, up to 5% of actual emissions can be offset.

It is worth noting that there are significant differences in the legal frameworks of carbon markets between China and the European Union, mainly reflected in the legal hierarchy and specific content. Specifically, firstly, the legal framework for the EU carbon market is relatively comprehensive. The EU has formulated numerous basic legal documents on carbon emissions trading, clarifying the common standards and procedures that member states must follow. The emission quotas and emission rights allocation plans formulated by various countries need to be reviewed and approved by the European Commission before they can take effect, which provides a solid legal guarantee for the operation of the carbon market. Secondly, the legal framework for China's carbon market is relatively weak. At present, China's legal framework for carbon emission trading mainly includes the "Interim Regulations on the Administration of Carbon Emission Trading" and the "Measures for the Administration of Carbon Emission Trading (Trial)". Although a "1+N" policy system for carbon peak and carbon neutrality has been established, policy requirements have not yet fully risen to the legal level, and there is a lack of clear regulations on many core issues in trading. Thirdly, in terms of market size and activity, the Chinese carbon market has shown a gradual growth trend. The EU carbon market is in a leading position globally in terms of market size and activity, especially with the extensive participation of numerous enterprises, financial institutions, and various investors, making the market highly liquid and the price discovery mechanism more effective.

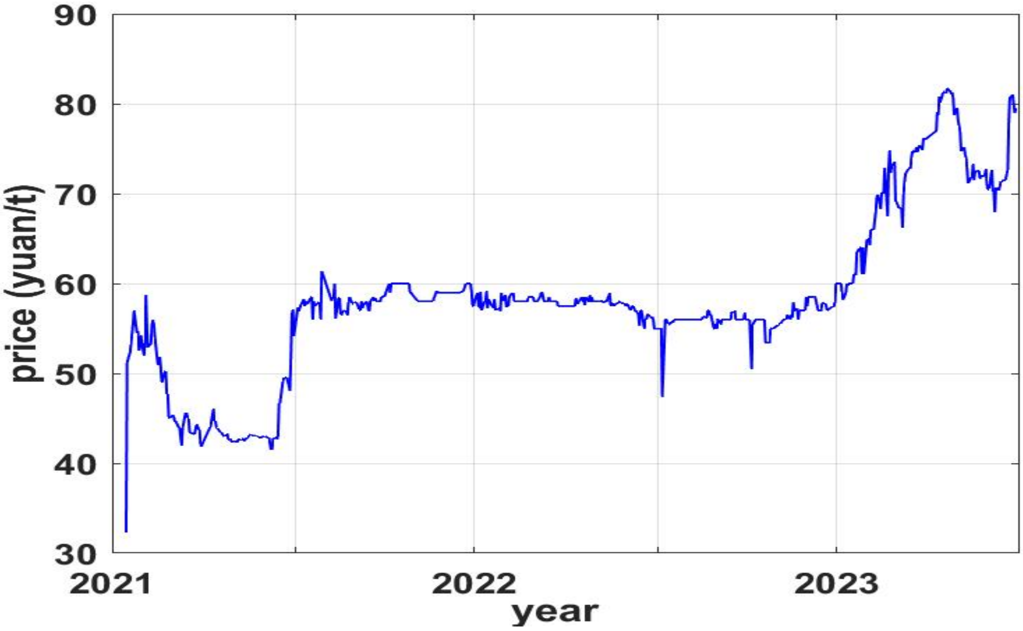


Figure 1.6: Transaction Prices of China NUCM from 2021-2023 (yuan/t)

Source: China Carbon Emissions Trading Network (2024)

1.2.3.3 The Price Trend in China Carbon Market

In terms of price and trading volume of the national unified carbon market, as of the end of June 2023, the accumulative trading volume reached 235 million tonnes, with a trading volume of 10.787 billion yuan, and an average carbon price of 45.83 yuan per tonne. The closing price was 60 yuan per tonne, reflecting 25% increase compared to the opening price on 16 July according to the Figure 1.5. After surveying the transaction prices of the national unified carbon market over the past two years, prices are generally smoothing. In particular, the trading price has risen steadily. For example, The national unified carbon market initially opened with a price of 48 yuan per tonne. However, by

November 2021, this price had dropped to an average of around 40 yuan per tonne. Starting from January 2022, the market saw a steady recovery, with transaction prices stabilizing between 50 and 60 yuan per tonne (as shown in Figure 1.6).

In terms of the price of Hubei carbon market, as the most active carbon market, Hubei carbon market ranked first in China in terms of transaction scale, continuity, amount of social capital introduced, and participation of enterprises. As of June 30, 2023, the secondary market of Hubei carbon market quota has accumulated a transaction volume of 365 million tonnes, turnover of 8.831 billion yuan, maintaining a leading level in the national pilot carbon market.

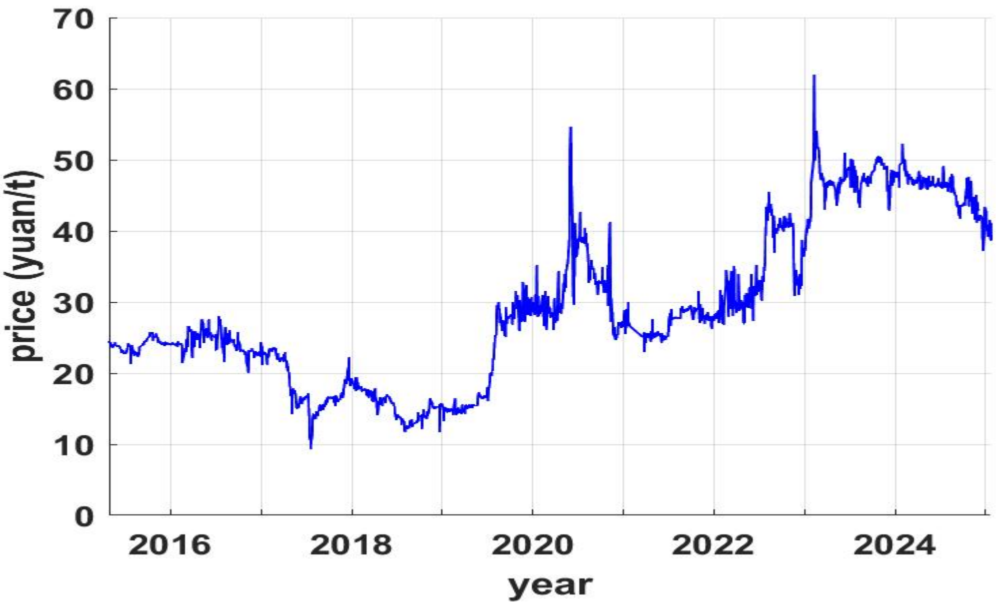


Figure 1.7: Transaction Prices of Hubei Carbon Market from 2014-2024 (yuan/t)
Source: China Carbon Emissions Trading Network (2024)

The latest data shows that, as of October 2023, total transaction volume of Hubei Emission Allowances (HBEA) is 184,500 tonnes, with a total turnover of 9,179,600 yuan. During the period, the highest transaction price was 51.40 yuan per tonne, while the lowest transaction price was 47.02 yuan per tonne. The average transaction price settled 49.76

yuan per tonne, on the last trading day of October, the closing price was 50.19 yuan per tonne, which was 2.70% higher than that of the closing price of 48.87 yuan per tonne on the last trading day of September. As of 31 October, the cumulative turnover of HBEA was 82.25 million tonnes, and cumulative turnover reached 1.99 billion yuan (as shown in Figure 1.7).

1.3 Problem Statement

1.3.1 Problem Statement from Macro Level

As the rapid economic development and increasing energy consumption, China has become the world's largest emitter of carbon emission. The environmental problems caused by the increasing carbon emissions are profoundly affecting the physical health and lifestyle of every Chinese person. If environmental problems are not effectively solved, sustaining long-term growth in the Chinese economy will also be difficult to maintain. Therefore, as a solution to environmental issues and carbon emission reduction, the establishment of the carbon market is a crucial measure for China. This approach aims to lead the global climate governance, overcome energy and environmental constraints, and get socioeconomic improvement and efficiency. It is generally believed that the faster the economic development speed, the more obvious the demand for fossil fuels, and the more pollution emissions come from the combustion of fossil fuels. Therefore, for China's rapidly developing economy, addressing environmental pollution and carbon emissions has become an important topic.

In September 2020, Chinese President Xi Jinping declared at the 75th United Nations General Assembly that, China will increase the strength of its national autonomous contribution, adopt stronger policies and measures, and strive to peak its carbon dioxide emissions before 2030 and work towards carbon neutrality before 2060. This important commitment plays a crucial role in promoting the development of China carbon market. In December 2020, the Chinese government further proposed to accelerate the construction of national energy rights and carbon emission trading market, and expect to achieve stable and moderate decrease in carbon emissions after reaching peak levels. The People's Bank of China in January 2021 clearly proposed to enhance the financial system's ability to manage climate related risks, and promote the construction of carbon market to reduce carbon emissions. In fact, the operation of the carbon market is an important means of reducing carbon emissions. An effective carbon market pricing mechanism can motivate the ability to address environmental issues, achieve effective reduction of carbon emissions, gradually alleviate climate conflicts, including preventing climate disasters, and responding to economic losses caused by abnormal and extreme weather conditions. That is to say, in the carbon market, companies with strong emission reduction capabilities can sell their excess quotas to companies with weak emission reduction capabilities, and through such active trading, achieve effective incentives for companies to reduce emissions. If the entire society's enterprises achieve emission reduction through this method, then the goal of carbon emissions will be eventually achieved.

According to the arrangement, in July 2021, the Ministry of Ecology and Environment of China officially launched the national unified carbon market. As of the end of 2023, the annual trading volume of China carbon market is 212 million tons. Among them, the trading volume of listed transactions was 35 million tons, and the trading volume of bulk agreement transactions was 177 million tons.

With the expansion of carbon market trading scale, the annual transaction volume of carbon emission quotas in 2023 is 14.44 billion yuan. Among them, the trading volume of listed agreement transactions was 2.57 billion yuan, and the trading volume of bulk agreement transactions was 11.88 billion yuan. The China carbon market has become the largest carbon market in the world. The price mechanism is the core of the carbon market (Yun et al., 2023), thus, in order to reduce emissions, enhance carbon market effectiveness, achieve the target of carbon peaking and carbon neutrality, it is essential to study price formation and forecasting mechanism of carbon market. Actually, an effective carbon price forecasting mechanism can improve the development of carbon market mechanism, but also serve the performance of carbon reduction.

1.3.2 Problem Statement from Micro Level

As a special financial innovative market, the price forecasting of carbon market need follows the basic methodology of general financial assets before considering the exclusive price characteristic. Compared with other financial markets such as the stock market, exchange rate market, interest rate market, etc, China carbon market has remarkable characteristics of sharp peaks and thick tails, market asymmetry, high sensitivity to policy shocks, and time-varying volatility (Zhang et al., 2017; Yun et al., 2023). Actually, the market asymmetry characteristics reflect the price fluctuations caused by irrational market shocks in the carbon market. Highly sensitivity to policy shocks means carbon price vulnerable to external events or policy changes, such as energy policies, emission reduction quota, global climate negotiations and financial crises in the capital market sector (Adekoya et al., 2021; Qiu et al., 2023). So, effective price forecasting research needs to capture those above characteristics, otherwise, such research needs to be strengthened. In statistics, the skewness of the third-order moments and kurtosis of the fourth-order moments can be used to reveal the impact of market asymmetry and extreme factors on the carbon market suggested by Kraus & Litzenberger (1976) and developed by Forbes et al.(2002). Effective price forecasting requires not only characterizing market characteristics, but also considering the cross market risk contagion relationship between carbon market and its infected markets. As the close connection of global financial networks, the price signal transmission and risk contagion relationship between carbon markets, energy markets, and capital markets cannot be ignored (Conrad et al., 2013; Zhu et al., 2022). Effectively identifying these relationships and incorporating them into carbon price forecasting models can enhance the explanatory of carbon price formation.

Actually, according to cross-market contagion and extended high order moment CAPM theory (Fry & Hsiao, 2018), the risk contagion from sourced carbon market to its infected markets also have an impact on carbon price. Relevant studies have found that, the volatility spillover originated from carbon market to the energy market, especially to the crude oil markets, is much larger than the degree of volatility spillover it receives (Tsai et al.,2024;Wang et al.,2024). The carbon prices contain complex volatility risk , which is easily transmitted to other closely markets through cross-market network, and forming a phenomenon of risk contagion (as shown in Figure 1.8).

After reviewing previous studies, mainly including Yu et al. (2020), Uddin et al. (2018) and Ji et al.(2018), as for the risk contagion measurement, this thesis found that previous studies mainly focused on measuring the information transmission and volatility spillover relationship between carbon market and its infected markets from the perspectives of market returns and volatility variance, using GARCH volatility techniques. Although these studies have certain advantages, they overlook the characterization of market asymmetry and highly sensitivity to policy shocks. That is to say, existing studies mainly focus on the low order moment perspective of mean and variance to study the risk contagion problem in carbon market, ignoring the risk contagion revealed from third-order and fourth-order moments such as skewness and kurtosis. As for the construction of forecasting model, the high order moment risk contagion factors were not included in the forecasting model, which has raised doubts about the forecasting accuracy.

In theory, the price transmission and volatility spillover relationship between carbon prices and their infected factors based on low order moments have received considerable attention. However, the impact mechanism of carbon prices has not been explained from the perspective of high order moments, especially from the perspective of reflecting the characteristics of carbon market asymmetry and extreme shocks. This makes it difficult for the existing theoretical framework to truly reflect the operating laws of carbon prices, and there are doubts about the accuracy of the relevant conclusions.

Therefore, to provide new and more convincing evidence, it is necessary to reveal the high order moment risk contagion relationship between China carbon market and its infected market before conducting the carbon price forecasting, and to construct machine learning models that can handle complex data, which is the core theme of this study.

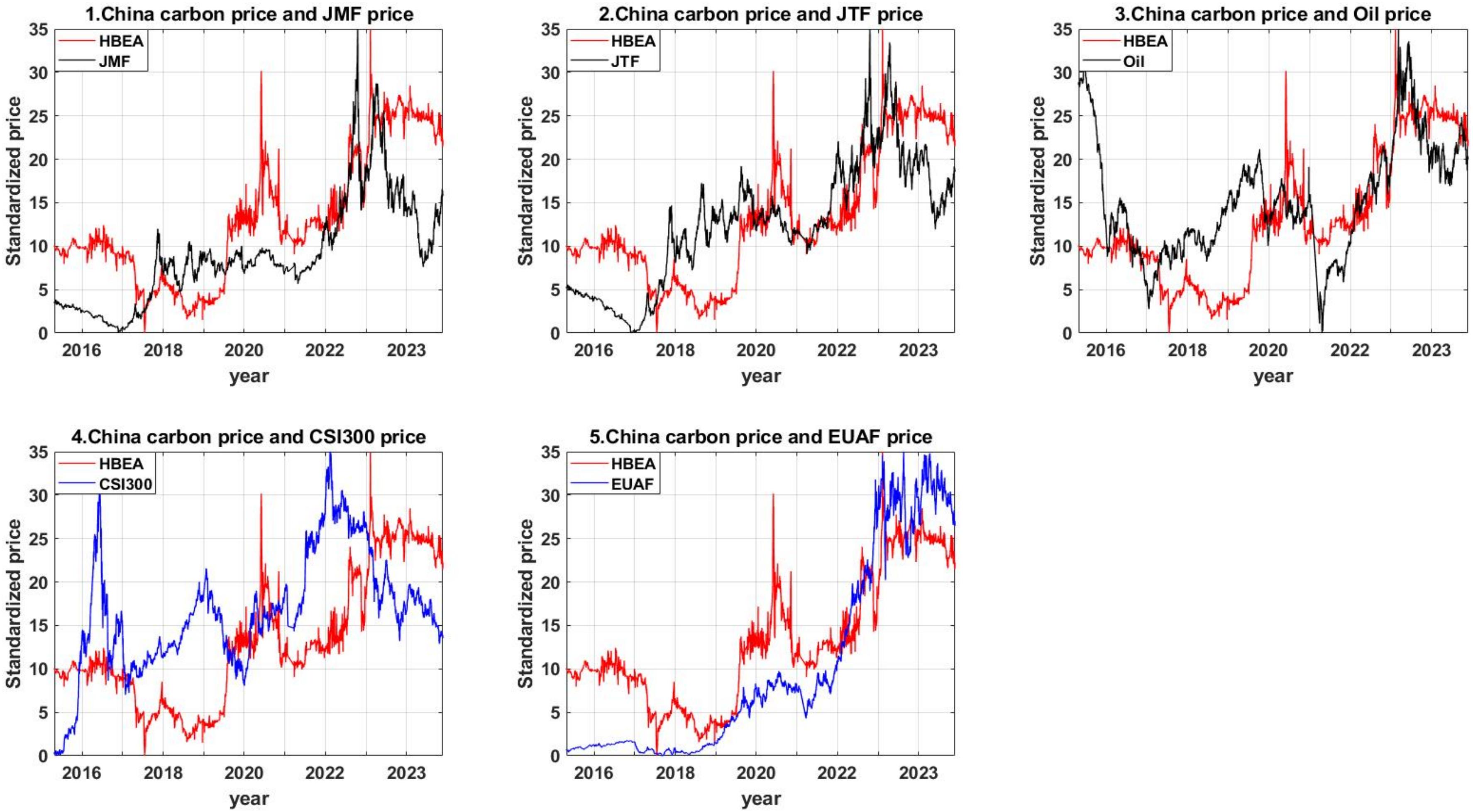


Figure 1.8: Price Fluctuation between Carbon Market and Infected Markets

Source: Investing App (2024)

1.4 Objective of the Study

1.4.1 General Objective

As environmental issues and carbon emissions have given more attention, relying on the carbon market to reduce carbon emissions has become the most effective means. Especially, the price mechanisms play a vital role in implementing emission reduction targets. Effective price forecasting can promote the achievement of emission reduction targets. Therefore, the general objective of this study is to forecast carbon price in China by high order risk contagion long and short-term memory model (HOC-LSTM).

1.4.2 Specific Objective

The specific objectives pertain to the following points:

i. To test the risk contagion channel from risk source carbon market to its infected markets in the manner of high order moment attribute.

ii. To construct a machine learning carbon price forecasting model that suitable for capturing the impact of high order moment risk contagion on carbon price.

iii. To forecast the China carbon price by the proposed high order risk contagion model (HOC-LSTM) to prove the risk contagion is useful to improve the forecasting performance.

1.5 Research Questions

In achieving the objectives of this study, the following questions research have to be addressed:

i. Under the high order moment attribute, how to test the risk contagion channel from risk source carbon market to its infected markets?

ii. How to construct a suitable carbon price forecasting model that consider the impact of the high order moment risk contagion?

iii. How to forecast the China carbon price by the proposed high order risk contagion machine learning model (HOC-LSTM) compared with others models?

1.6 Significance of the Research

Forecasting carbon market prices is central to its functioning. Building an effective mechanism between risk contagion and price forecasting through high order moment attributes, can not only conducive to improve the effectiveness of carbon market, but also better serve the capital allocation function of carbon market. This thesis has certain theoretical significance and practical significance.

1.6.1 Theoretical Significance

Firstly, the carbon price forecasting theoretical framework designed in this thesis can expands the theory and method of China carbon price forecasting research. Starting from the high order moment financial asset pricing theory, this thesis explores the basic connotation of high order moment risk contagion and its infected mechanism on China carbon price. According to extended high order moment pricing theory and cross market risk contagion theory, combined the special characteristics of carbon market, using LSTM neural network to forecast China carbon price.

Secondly, a new non parametric method has been developed, which helps to test the risk contagion relationship from sourced carbon market to infected market. On the one hand, the carbon market volatility is characterized into rising and declining trend, and furthermore, constructs different risk contagion relationship from sourced carbon market to its infected markets in the fast and slow market volatility. On the other hand, this study conducts a new non-parametric method to test high order risk contagion relationship in the channel of Co-Skewness (CS), Co-Kurtosis (CK) and Co-Volatility (CV).

Thirdly, this study promote the carbon price forecasting model development through designing a high order moment risk contagion model (HOC-LSTM). The LSTM has training advantage of conforming to long-term memory function of financial time series, especially through the special design of gate structure to control data information. The LSTM can effectively avoiding the problem of gradient explosion and gradient vanishing that other deep networks, so as to ensure the effective convergence of the model (Wang et al., 2021; Elsayed et al., 2022). An effective carbon price forecasting mechanism can significantly contribute to achieving carbon reduction targets. Therefore, the theoretical framework and research methods improve the forecasting accuracy of carbon price, as a result, the solution ideas and process of carbon price forecasting can also offer valuable insights and reference for solving forecasting problems in other capital markets or energy markets.

1.6.2 Practical Significance

Firstly, this study serves the emission reduction of carbon market and improve market trading mechanism. The price forecasting of carbon market is the core issue. An effective carbon price forecasting framework helps to improve the maturity and improvement of carbon market, promote the allocation of emission reduction funds, provide decision-making reference for government departments to formulate carbon market risk control policies, supervise the operation of carbon market, and deal with systemic financial risks. At the same time, these advancements will better support entity enterprises to reduce emissions of pollutants, enhance the role of the carbon market, and achieve carbon peak and carbon neutrality goals.

Secondly, this study help carbon market investors make investment decisions under uncertain environment. On the one hand, compared with the traditional price forecasting framework, constructing a carbon price forecasting mechanism that considers high order moment attribute risk contagion can offer valuable reference for market participants, such as investors and emission reduction entities to make investment decisions and risk management. On the other hand, the maturity of carbon price forecasting model can promote the development of market efficiency, thereby providing price operation rules and decision-making basis for carbon asset supply and demand enterprises to achieve optimal economic and ecological benefits.

1.7 Organization of the Study

The thesis is organized into five chapters. Chapter one analyzes the origin of greenhouse gas emissions and the environmental negative externalities caused by them, expounds the development of carbon market in European and China, analyzes the core issue of carbon market, that is, the relevant research background of carbon price forecasting, and then puts forward the research problems and research objectives. Then,this chapter emphasizes the significance and structural details of the research, in order to comprehensively understand the motivation and direction of the research.

Chapter two reviews the literature, which is related to the objective of the study. From perspective of financial asset moment attribute, this chapter summarizes the risk contagion theory of carbon market, carbon asset pricing theory, efficient market hypothesis theory and prospect theory. Collects and analyzes high order moment attribute financial asset pricing theory and high order moment attribute risk contagion theory, and combine them with the characteristics of carbon emission trading to classify, extract and summarize the risk fluctuation characteristics of carbon market. Through theoretical analysis, a carbon price forecasting framework suitable for China carbon market is constructed, which provides foundation for the model construction and experimental analysis in the following chapters.

Chapter three explains research design and methodology used in the study. The chapter is mainly the design of carbon price forecasting model based on high order moment attribute risk contagion. Firstly, identify and judge the high order moment risk contagion relationship caused by market asymmetric information and extreme event impact. Then, according to the identified risk contagion relationship, the model framework of carbon price forecasting is clarified and determined. Finally, machine learning method is used to predict the high order moment risk contagion framework.

Chapter four is the empirical analysis of high order moment risk contagion in China carbon market. Firstly, based on statistical modeling technology, this chapter makes basic statistical analysis, stationarity test, heterogeneity division of fluctuation trend of carbon price, calculation and analysis of high order statistics. Secondly, based on the non-parametric statistical hypothesis test method, identify the risk contagion relationship between carbon market and its infected markets, analyze variations risk contagion channels among different market volatility trends. Thirdly, this study utilizes machine learning methods to fit the pricing framework and forecast prices based on high order moment contagion.

Chapter five summarizes the empirical results and conclusions of this study, as well as explaining the limitations that must be noted. In addition, the following suggestions have been put forward in the thematic areas: policy recommendations for China carbon market price mechanism, analysis risk contagion prevention strategies in China carbon market, and some information that may be used for future research.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

From a global perspective, China, the United States, and Europe are the top three carbon emitters, accounting for approximately 51% of the total global carbon emissions by 2023, according to the International Energy Agency (2024). The total carbon emissions are closely related to the total economic output. China's carbon emissions have shown rapid growth since joining the WTO, mainly due to the continuous expansion of its economic scale. After the financial crisis, the overall carbon emissions in Europe and America showed a downward trend. The main driving force behind decarbonization in Europe and the United States is the removal of coal. The difference between the two is that Europe is a renewable energy source, mainly driven by the growth of wind and solar power offsetting the decline of coal-fired power, while the United States has benefited from the "shale revolution", with a large amount of cheap natural gas resources accelerating the process of replacing coal-fired power with gas power. Regardless of the economy, ultimately achieving the goal of reducing carbon emissions requires minimizing coal consumption, and filling the coal gap can only rely on renewable energy.

The topic of this thesis is price forecasting of China carbon market based on high order risk contagion. The relevant literature review mainly focuses on two aspects: theoretical literature and empirical literature. In terms of theoretical literature, this study first introduces the classical Efficient Market Hypothesis theory, and further analyzes the financial asset price forecasting research from the perspective of portfolio theory and capital asset pricing model theory. Secondly, this study reveals the research of prospect theory on financial asset price forecasting from the perspective of market inefficiency, and further explores the risk contagion theory and its specific performance in the carbon market. In terms of empirical literature, this study discusses the financial asset price forecasting research from both the perspective of low order moment and high order moment, and also analyzes the price characteristics of carbon market. Based on this, the latest development in carbon price forecasting is revealed through GARCH technology and artificial intelligence modeling technology.

2.2 Theoretical Literature

2.2.1 Efficient Market Hypothesis Theory

The Efficient Market Hypothesis (EMH) theory was proposed by the renowned economist Eugene Fama in 1970. It is a cornerstone of financial asset price forecasting research based on investor rational expectations in uncertain environments. Professor Fama was awarded the Nobel prize in economics in October 2013 for his new approach in the field of asset market trends. In his doctoral thesis "The behaviour of stock-market prices", Fama (1965) first studied the non-normal distribution characteristics of Dow Jones Industrial Average stocks, and revealed the relationship between time series characteristics and stock returns (Fama,1965a). The study is considered to be the earliest test of the efficient market hypothesis. Under the framework of random walk theory, based on commonly used strategies such as technical analysis and fundamental analysis in quantitative investment, Fama (1965b) formally proposed and elaborated the basic concepts of the efficient market hypothesis in his article "Random walks in stock market prices". This study has shown that the financial market under the random walk theory is actually equivalent to an efficient market, where stock prices reflect all information. In such a market, a large number of market investors makes the stock price to follow the

"Random walks ". If the deviation of stock returns and prices is triggered by systematic risk rather than random behaviour, then the rational investors trading behaviour driven by profit-seeking motives will bring the market price back to the real value, at the same time, the trading motivation will be offset by systematic changes in returns. Although the intrinsic price of the stock market is stable and unchanged, the market price will be wander randomly. In order to establish a complete framework for the efficient market hypothesis, Fama (1970) systematically summarized the previous research on the efficient market hypothesis in his article "Efficient market hypothesis: A review of theory and empirical work", and innovatively proposed a validity testing framework for the Joint Hypothesis, the study has suggested that it is necessary to establish a reasonable and effective asset price forecasting model to test market efficiency. Only on established pricing models can the relationship between expected returns and wandering behavior can be analyzed, and then the market efficiency problem represented by stock returns can be studied (Fama, 1970).

Based on the ideas of the above classical literature, the efficient market hypothesis suggests that market price of any financial asset already incorporates all available information accessible to investors, and investors can use those information to take asset valuation and price forecasting. Only the emergence of new information will trigger the price fluctuations, while the occurrence of new information is uncertain, so the financial assets price is also unpredictable. Therefore, any strategy attempts to obtain excess profits though technical tools and fundamental analysis is futile (Fama and MacBeth, 1973; Fama and Schwert, 1977). Depending on the extent to which market information reflects asset prices, the efficient market hypothesis can be divided into three forms. Among them, the Weak Form Efficient Market Hypothesis (WFEMH) states that asset prices reflect the historical information of all disclosed stocks assets, so the excessive profit by using technical analysis is no longer valid. The Semi-Strong Efficient Market Hypothesis (SSEMH) means that asset prices reflect all publicly available market information. Therefore, it is difficult to obtain excess returns when using publicly available information for asset valuation. The Strong Form Efficient Market Hypothesis (SFEMH) suggests that asset prices reflect all publicly available and undisclosed stocks information in the market, and any means of chasing excess profits will be ineffective (Fama, 1970; Fama, 1991).

The efficient market hypothesis essentially means that there is no free lunch in the world. As the starting point of traditional financial asset price forecasting theory, the efficient market hypothesis has two core assumptions: one is that stocks prices reflect all information and quickly adjust to it, the other is that all investors are completely rational. In fact, the efficient market hypothesis is only a hypothetical framework for analyzing asset price forecasting problems. As a matter of fact, not all stock prices can reflect market information, and investors are not entirely rational. To be honest, the hypothesis still occupies an important position in the basic framework of modern mainstream financial market theory (Barberis et al., 2016).

2.2.2 Financial Asset Price Forecasting Theory

2.2.2.1 Portfolio Theory

Portfolio theory is an asset management theory based on diversified investment portfolios, the advantage is diversifying investment risks and improving investment efficiency, the more portfolio assets there are, the greater the degree of non-systematic risk diversification. In 1952, Markowitz published the classical " Portfolio Selection ", marking the beginning of modern portfolio theory. Markowitz (1952) proposed a research conclusion that effective portfolio selection can reduce non systematic risk based on the "mean-variance" low order moment framework. According to the portfolio theory, rational investors prefer to seek portfolios that maximize the expected return under established risk, or minimize investment risk under expected return. The curve formed by connecting the points corresponding to expected return and standard deviation of each portfolio is called efficient frontier. The Markowitz portfolio model based on the mean-variance framework provides an analytical basis for quantifying the relationship between risk and return, while the model requires the calculation of covariance matrices for all portfolio assets, especially when the number of portfolio assets is enough large, which makes calculation process too complicated. On this basis, William Sharpe (1963) proposed Sharpe's one-way analysis of variance, which simplified the calculation of the covariance and significantly promoted the practical application of portfolio theory. The Sharpe's one-way model divides the market risk into systematic risk and non-systematic risk, the single factor only maps the systematic risk, and there is no contagion between the assets of non-systematic risk. Therefore,the factors that affect stocks returns are focused on the common systematic risk.

2.2.2.2 Capital Asset Pricing Model Theory

Based on Markowitz portfolio theory, William Sharpe (1964) further research the risk-return relationship of financial asset, and proposed the famous Capital Asset Pricing Model (CAPM). The Markowitz portfolio theory provides a framework for the analysis of asset portfolios based on the perspective of the individual investor, while the CAPM determines the equilibrium returns of asset portfolio at the efficient portfolio boundary. The CAPM assumes that investors make their investment decisions in accordance with Markowitz's asset selection theory, and focuses on the relationship between expected returns and risk-reward coefficients. It establishes a simple linear correlation between risk-taking and expected returns, and puts forward a series of assumptions for applying the model. Building on the findings of the CAPM, the relationship between the expected returns obtained by investors and the market risk they bear can be explained by the Capital Market Line (CML) and the Security Market Line (SML). By relaxing risk-free borrowing restriction in the classical CAPM, and assuming that there are no risk-free assets in the financial market, Black (1972) pointed that any asset portfolio composed of asset groups is still an efficient asset portfolio. Furthermore, relaxing the restriction of liquidity costs in the classical CAPM, assuming that there are trading frictions and liquidity costs in the financial market, Achary et al.(2005) proposed a liquidity-adjusted CAPM (LA-CAPM). In the consumption-based capital asset pricing model, Breeden et al.(1989) assume that the investor's utility function is maximizing the utility between immediate and future periods, with equilibrium utility satisfying the condition where the marginal cost of immediate consumption equals the marginal benefit of future consumption.

Compared to the optimal portfolio theory and CAPM theory, the multi factor price forecasting model based on Arbitrage Pricing Theory (APT) relaxes more assumptions, and incorporate numerous macro factors into the framework, thus presenting stronger explanatory power for financial price forecasting (Ross, 1976; Merton, 1987). APT theory uses factor models to explain the determining factors of asset prices and the formation mechanism of equilibrium prices. As an extension of the CAPM, Fama & French (1993) introduced the classical three-factor model, considering the market size factor and book-to-market ratio factor beside the risk premium measured by CAPM. Furthermore, based on the three-factor model, Carhart (1997) suggests that studying stock returns should add momentum trend factors into three-factor models, then construct a four-factor model. Similarly, based on three-factor model, Fama and French (2015) further incorporate profit level risk and investment level risk factors into the analytical framework and propose a five-factor pricing model.

2.2.3 Prospect Theory Based on the Market Inefficiency Hypothesis

Based on the limited rational decision-making practices and the development of investment psychology, Tversky & Kahneman (1979) proposed the classical Prospect Theory. This theory suggests that financial asset prices are not only determined by the intrinsic value, but also largely influenced by the subjective behaviour of rational investors. That is, the decision-making psychology and behavior of investors are also affect asset prices (Barberis & Thaler, 2003; Barberis et al., 2018). Therefore, prospect theory focuses on the hypothesis of limited investor rationality, and focus on studying how a limited rational investor evaluates and determines the optimal decision option when facing with multiple decision options (Tversky & Kahneman, 1981). Unlike traditional financial asset price forecasting theory under completely rational investment, prospect theory solves the decision-making problem that cannot be explained by traditional gain-loss utility. Under prospect theory, the value function evaluates the subjective value of uncertain decision outcomes to rational investor, while the weight function is the probability of an uncertain decision outcome.

Tversky & Kahneman (1992) proposed the Cumulative Prospect Theory (CPT), which further extends the analysis of prospect theory to multiple alternative decision options, and achieving quantitative expression of the value function and weighting function. Furthermore, prospect theory also confirms that, lower volatility risk indicates lower returns, which increases investors' expectations of high future returns and promotes buy financial assets with right-skewed distributions in the future. Therefore, there is a negative correlation between volatility and expected returns (Ang et al., 2006; Ang et al., 2009), there is a certain positive correlation between low volatility and future skewness, which is known as low heterogeneous volatility anomaly (Mitton & Vorkink, 2007; Boyer et al., 2010). Based on the analysis of the value function, it is found that rational investors exhibit a strong risk aversion preference when facing deterministic returns. When facing deterministic losses, investors tend to have a risk preference, which explains the fact that investors tend to arbitrage and sell profits as soon as possible when facing stock price increases. While when losses occur, they are delayed in selling (Barberis&Xiong, 2009; Ingersoll&Jin, 2013).

2.2.4 Risk Contagion Theory of the Carbon Market

Probability analysis and correlation analysis are commonly used in previous research to identify risk contagion of financial markets. Probability analysis is one of the early methods to test risk contagion among financial markets. If the probability of a financial risk in one country increases significantly in response to a risk in another country, it indicates that risk contagion has occurred (Dornbusch et al., 2000; Kumar et al., 2003). Correlation analysis suggests that the existence of risk contagion can be defined when the correlation between markets significantly increases following a financial crisis (King & Wadhwani, 1990). Based on this idea, the significant change in correlation between different financial markets before and after a risk shock becomes a proxy indicator for identifying risk contagion. In particular, with the development of econometric models, correlation analysis is able to satisfy the risk contagion measurement request within linear and nonlinear financial systems (Boyer et al., 1999; Loretan et al., 2000; Tjøstheim et al., 2013; Støve et al., 2014). Furthermore, correlation analysis can also capture the nonlinear interdependence and tail dependence between financial markets, this method can describes dynamic structural changes and analyzes the direction and intensity of risk contagion (Patton, 2006; Chollete et al., 2009; Arakelian and Dellaportas, 2012; Abbara et al., 2014).

The economic fundamentals are the direct cause of financial risk contagion, however, due to the investors limited rationality and the market incomplete efficiency, the impact of financial crisis on investors psychological expectations cannot be ignored in the formation of risk contagion. Karolyi (2003) proposed that the risk contagion theory based on market efficiency, which states that contagion occurs when the cross market linkage behavior has become irrational and cannot be explained by economic fundamentals. Risk contagion is essentially an irrational co-movement. Actually, it is the phenomenon of "net contagion" through high order moment channels, which excludes market linkage caused by economic fundamentals and rational investment decisions (Masson, 1999; Forbes and Rigobon, 2002; Boyer et al., 2006). For example, the financial markets globalization can lead to investor herd effects, affecting investor sentiment and trading psychological expectations (Calvo et al., 2000; Zhang et al., 2023; Ashfaq et al., 2024). The macro-risk exposure behaviour accompanying with the investors asset portfolios in other markets is commonly used to hedge their investment risks (Kodres & Pritsker, 2002). These behaviors can exacerbate the occurrence and transmission of financial crises. Therefore, "net contagion" theory of carbon market risk is essentially the study of investor expectations and behavioral changes caused by market asymmetry and extreme event shocks, leading to significant changes in market correlation.

2.3 Empirical Literature

2.3.1 Financial Pricing Research in Low Order Moment Risk Contagion Channel

2.3.1.1 Financial Pricing Research in the Perspective of Market Return

The idea of early research on financial asset pricing are commonly based on the point of returns, these studies mainly examining the influence mechanism of asset price from the perspective of first order moments of financial assets. In terms of research methodology, it is mainly based on linear or non-linear regression methods to explain premiums of financial asset prices.

On the basis of portfolio theory, the CAPM model based on rational expectations links excess returns with portfolio returns, and studies the relationship between financial asset risk premium and systemic risk. Sharpe (1964) used the stocks sample form the New York Stock Exchange (NYSE) discovered a nearly exact linear relationship between average returns and risk coefficients. Although using only the risk coefficient as an explanatory factor can clarify the compensation mechanism between risk and return, it is obvious that the strict assumptions is not adapted to the specific analyses of real problems. Subsequent studies have gradually relaxed the assumption, and more consideration have given attention to test the impact of non-market factors on the price formation of financial assets (Mayers, 1973). It is difficult to explain the risk premiums only based on the single or combined market factors. Multiple factors, including macroeconomic factors such as inflation rate and interest rate, should be included in the explanatory framework, as a result, the multi-factor Arbitrage Pricing Theory (APT) is constructed (Ross, 1976; Merton, 1984). Under the multi-factor pricing idea, consider the emerging financial anomalies and risk premiums, inflation factor (Moerman & Van, 2010), liquidity risk factors (Pastor & Stambaugh, 2003; Acharya & Pedersen, 2005; Asparouhova et al. 2010; Lam & Tam, 2011), exchange rate risk factors (Apergis et al., 2011), global risk factors (Bali & Cakici, 2010), homogeneity of financial standards (Griffin, 2002; Moerman, 2005; Biscarri & Espinosa, 2008), and tax heterogeneity risk factor (Eikseth & Lindset, 2009) have been successively incorporated into the APT framework and used to improve the explanatory power of pricing models for realistic risky asset returns.

Breaking through the limitation of considering the linear impact of a single risk factor on premium, Fama & French (1993) added company size and book-to-market ratio factors into regression model. Research found that the three-factor model was able to explain 70%-80% of price changes in US stock returns. Research results such as the long-term low performance caused by rights issues (Allen et al., 2024; Fritz et al., 2024), market efficiency and corporate performance under equity separation reform (Trakarnsirinont et al., 2023; Alodat et al., 2023), and the improvement of the three-factor model under state transition and unexpected stock returns (Kostin et al., 2022; Hong et al., 2022) all shown that the three-factor model can fit and analyse the market returns and abnormal changes in China's capital market (Jiang, 2014). As the three-factor model is unable to reveal the momentum phenomenon of financial asset returns, the four-factor model includes momentum factors have been proved to be significantly improves the explanatory power of pricing model for asset premiums (Carhart, 1997). Furthermore, Fama & French (2015) continued to incorporate profitability factor and investment capacity factor into three-factor model, and proposed a five-factor asset pricing model. It was found that the five-factor model could achieve an explanation level of 46%-58%. Empirical evidence from China also suggests that the five-factor model greatly improves the average portfolio returns (Guo et al., 2017).

Table 2.1: Summary of Financial Pricing in the Perspective of Return

Author & Year	Methodology	Pricing Factors	Changes
Sharpe (1964)	Linear regression model	Risk factor	+
Mayers (1973)	Linear regression model	Human capital, social insurance, pensions	
Ross (1976); Merton (1984)	Linear regression model	Inflation, interest rates	
Fama & French (1993) Jiang,(2014); Kostin et al.(2022) Trakarnsirinont et al.(2023) Allen et al.(2024)	Three factors CAPM	Risk factors, company size, book value ratio	+
Carhart (1997)	Four factors CAPM	Risk Factors, Firm Size, Book-to-Value Ratio, Asset Momentum Factors	+
Moerman (2005)	Linear regression model	Homogenisation of financial standards	
Eikseth & Lindset (2009)		Tax heterogeneity	
Acharya & Pedersen, (2005) Moerman & Van (2010) Asparouhova et al. (2010) Bali & Cakici (2010)		Inflation factor, liquidity risk factor, global and country-specific risk factors	+
Lam & Tam (2011) Apergis et al. (2011)		Liquidity risk factor, exchange rate risk factor	+
Fama & French (2015) Chen et al.(2022) Zerbib (2022) Mosoeu and Kodongo (2022)	Five factors CAPM	Risk Factors Firm Size, Book-to-Value Ratio, Asset Momentum Factor, Profitability Factor and Investability Factor	+
Roy & Shijin (2018) Hong et al.(2022) Alodat et al.(2023) Fritz et al.(2024)	Six factors CAPM	Risk Factors Company size, book to value ratio, asset momentum factor, profitability factor and investability factor, human capital factor	+

However, some studies have pointed out that the effectiveness of financial asset pricing models varies depending on market efficiency and investment philosophy, the efficiency of China capital market is weaker compared to developed countries, the fundamentals of the capital market are more sensitive to policy shocks, investors are not sufficiently concerned about company growth, etc. (Chen et al., 2022; Zerbib, 2022), which makes profitability factors and investment level factors ineffective in revealing premium volatility in asset returns. Contrary to the empirical experience in the US capital market, the five-factor model is not suitable for the China capital market (Mosoeu and Kodongo, 2022), especially for small-cap stocks with high investment ratio and low profitability, the five-factor model lacks explanatory power (Fama & French, 2017). Introducing non-marketable human capital factors into the five-factor model, Roy & Shijin (2018) proposed an equilibrium six-factor asset pricing model, this model revealed that human capital component possesses predictive power comparable other factors in explaining changes in portfolio returns. Table 2.1 summarizes the studies on premiums and changes in financial asset prices under the linear regression approach.

The above research on financial asset pricing based on linear regression can reveal the influential strength and direction between expected asset returns and pricing factors. However, as the complexity of financial asset pricing factors, the nonlinear characteristics of financial assets make it difficult for traditional linear regression methods to accurately capture the nonlinear relationship. The modeling technology based on Artificial Neural Network (ANN) can capture the nonlinear structure through adaptive adjustment and optimization of model parameters, thus making the nonlinear regression between financial assets and pricing factors more accurate. For example, Szkuta et al. (1999), Yamin et al. (2004), Gonget et al. (2008), Singhal and Swarup (2011) constructed a comprehensive model for short-term electricity price adaptive forecasting based on artificial neural networks. It was found that the neural network-based electricity price forecasting model has general advantages, especially the Recurrent Neural Network (RNN) model represented by LSTM neural network and Gate Recurrent Unit (GRU) neural network have significant fitting advantages for time series data (Ugurlu et al., 2018). To solve the problem of under-learning and over-learning in the learning process of traditional neural networks, as well as the possibility of network training falling into local minima, regression based on Support Vector Machine (SVM) is used in the financial asset price forecasting research. For example, by constructing the nonlinear mapping relationship between daily return of the exchange rate and its lagged return, as well as pricing factor of lagged return, it is found that the exchange rate forecasting model based on support vector machine can effectively explain the price volatility mechanism (Yuan, 2013; Cao et al., 2005; Alamili, 2011; Plakandaras et al., 2011; Özorhan et al., 2017).

Machine learning models are essentially a type of regression models, the machine learning model is a deep network structure of multi-layer perceptron with self-organizing, self-adaptation and self-adjustment advantage, the significant advantage is promoting feature extraction and network learning, which can improve the fitting performance between input variables and output variables. For example, the improved deep belief network and the Copula LSTM model can both fit and predict exchange rate prices well (Shen et al., 2015; Ugurlu et al., 2018; Zheng et al., 2019; Cao et al., 2020; Cheng et al., 2023). It indicates that machine learning methods can use powerful feature learning and optimization methods, maximize the network regression and nonlinear mapping between financial asset prices and their pricing factors, providing technical support for grasping the determination mechanism of financial asset returns (Kim et al., 2008; Luo et al., 2024).

The Copula model can effectively reveal the nonlinear dependence relationship between carbon prices and their influencing factors, especially it can accurately reveal the characterization of this correlation by the multi-vine Copula model. As such, combining Copula models with various machine learning regressors can greatly improve the predictive performance of hybrid models.Table 2.2 summarizes the studies on premiums and changes in financial asset prices under the non-linear regression methods.

Table 2.2: Premiums and Changes in FAP under Non-Linear Regression Methods

Author & Year	Objectives	Methodology	Ability to Explain Asset Premiums
Szkuta et al. (1999) Yamin et al. (2002) Gonget al. (2008) Singhal and Swarup (2011)	Construction of an integrated model for adaptive, forecasting of short-term electricity prices in the electricity market	LSTM and GRU	Universal advantages
Cao et al. (2005) Alamili (2011) Plakandaras et al.(2011) Yuan (2013) Özorhan et al. (2017)	Constructing a non-linear mapping between daily returns on exchange rates and their lagged returns, lagged returns on pricing factors	Holding Vector Machines for Exchange Rate Prediction Models SVM	Effectively explain price fluctuations in exchange rate assets
Kim et al. (2008) Shen et al. (2015) Ugurlu et al. (2018) Zheng et al. (2019) Cao et al. (2020) Cheng et al. (2023) Luo et al. (2024)	Enabling the fitting and forecasting of exchange rate prices	Copula LSTM	Has a good fit and predict

2.3.1.2 Financial Pricing Research in the Perspective of Market Volatility

Risk contagion and information transfer triggered by market volatility are important foundation for the study of cross-market asset pricing mechanisms. Affected by sudden economic events and policy events, the financial assets prices show obvious characteristics such as sharp peaks and thick tails, volatility clustering, which makes the pricing model based on the first order moment attribute of return no longer applicable. Incorporating financial asset volatility into the pricing models, the Autoregressive Conditional Heteroscedasticity Model (ARCH), Generalized Autoregressive Conditional Heteroscedasticity Model (GARCH), and Stochastic Volatility Model (SV) modles have gradually been applied to forecast the second order moment attribute of volatility (Sajjad et al., 2013; Wang & Wong, 2017; Arashi & Rounaghi, 2022).

The classic Black Scholes (BS) volatility pricing model assumes that the financial assets price follow a Brownian motion with constant volatility, but it cannot reflect the time-varying characteristics of financial asset price in reality (Black & Scholes, 1973). In response to the drawbacks of constant volatility, Heston (1993) proposed the Heston stochastic volatility model, which not only captures the characteristics of financial assets volatility clustering, but also reflects the dynamics and time-varying nature of volatility. Considering the sudden price variability caused by market shocks, the jump process based on Poisson distribution can provide a better explanation for price volatility behaviour, and improve the pricing ability of financial assets derivatives (Chen et al., 2022). Based on the dynamic jump-diffusion stochastic process, Lian & Chen (2022) proposed a two-factor cross-feedback option pricing model with time-varying jump arrival rate and volatility. The study showed that the explanatory power of asset premium based on the jump process is better than that of risk premium based on volatility, and pricing power of the two-factor cross-feedback model is significantly better than that of the one-way feedback jump diffusion model (Büchner & Kelly, 2022). Based on the three factor term model (short-term, medium-term, long-term model) and random discount factor model, a futures pricing dynamic model with jumps is constructed to depict the impact of jump factors on energy market prices. The study found that the asset premium explanatory power of the futures price dynamic model is stronger (Sakariyahu et al., 2023; Fang et al., 2023; Cheng et al., 2023; Tronzano, 2024). Table 2.3 summarizes the studies on financial asset pricing for market volatility analysis.

Table 2.3: Summary Financial Pricing in the Perspective of Market Volatility

Author & Year	Methodology	Ability to Explain Premiums on Financial Assets
Sajjad et al. (2013) Wang & Wong (2017) Arashi & Rounaghi (2022)	ARCH, GARCH, SV	Stronger prediction and fitting of volatility of second-order moment properties
Black & Scholes (1973)	Black Scholes volatility pricing model	Failure to reflect the time-varying nature of financial asset price fluctuations in reality
Heston (1993)	Heston stochastic volatility model	Not only does it capture the characteristics of volatile aggregation of financial assets, but also the dynamic and time-varying nature of fluctuations in returns
Büchner & Kelly (2022)	Jump-diffusion modelling of unidirectional feedback	Weaker
Chen et al. (2022) Lian & Chen (2022)	A two-factor cross-feedback option pricing model	Relatively strong
Sakariyahu et al.(2023) Fang et al.(2023) Cheng et al.(2023) Tronzano (2024)	A jumpy model of futures pricing dynamics	More powerful

2.3.2 Financial Pricing Research in High Order Moment Risk Contagion Channel

2.3.2.1 Financial Pricing Research in the Perspective of Market Skewness

Kraus & Litzenberger (1976) added third-order moment skewness information to the CAPM model for revealing the financial asset price formation process, derived an asset pricing model based on third-order moment attributes, and used the two-stage regression method proposed by Fama & MacBeth (1973) for empirical analysis. The study found that co-skewness as a systematic risk can provide a good explanation for the NYSE stock returns from 1936 to 1970, and the market skewness of third-order moments is negatively correlated to expected returns. In order to obtain a portfolio with a right-skewed return distribution, rational investors are willing to sacrifice some risk premium, and only a portion, but not all, of the portfolio's return distribution is affected by skewness (Conrad et al., 2013). Harvey & Siddique (2000) found that co-skewness affects not only cross-sectional changes in expected returns, but also reveals the relationship between inertia effects and systematic co-skewness. Inertia portfolios with low expected returns have a larger skewness than those with high expected returns. Buckle et al. (2016) combined realized volatility measures with high order moments properties of the portfolio assets, and found that time varying high-order moment models can more accurately capture the high order moment changes of asset portfolios during financial crisis. The third order moment CAPM model that integrates first-order moment mean, second-order moment variance, and third-order moment skewness not only meets expectations of rational investors' risk preferences, but also has significantly explanatory power of asset premiums than the CAPM model and the three factor model (Smith, 2007).

Moreover, the third order moment CAPM model that integrates liquidity adjustments has better explanatory effect on financial asset premiums, especially in asset pricing research in the China A-share market, the liquidity factor and the third order moment skewness factor have strong explanatory power (Shafiullah et al., 2024). The financial asset pricing models considering high order moment attributes are more suitable for China capital market than the low order moments model (mean-variance based pricing model) (Yun et al., 2020).

Table 2.4: Summary Financial Pricing in the Perspective of Market Skewness

Author & Year	Methodology	Study Result
Kraus & Litzenberger (1976)	Three high-moment CAPM	Negative correlation
Conrad et al. (2013)		Partial positive correlation
Harvey & Siddique (2000)		Positive correlation
Buckle et al. (2016)		Positive correlation
Smith (2007)		Positive correlation
Boyer et al. (2010)	Fama-French three factors model	Positive correlation
Lin et al. (2019) Yun et al.(2020) Shafiullah et al.(2024)	Equilibrium assets and option pricing models	Negative correlation

Based on the Fama & French asset pricing framework, Boyer et al. (2010) found that stocks with high skewness have lower expected returns, and the expected skewness coefficient in the Fama-MacBeth cross-sectional regression is significantly negative, this conclusion indicates that a negative correlated between expected skewness and stock portfolio returns. The Fama & French Alpha value of stocks with lower expected skewness is one percentage point higher than that for stocks with high expected skewness on a monthly basis. Defining the skewness risk premium as the difference between statistical skewness and risk-neutral skewness, Lin et al. (2019) used equilibrium asset and option pricing models, the result shown that S&P500 index can be predicted, especially when the market skewness risk premium is high, and risk-averse individuals typically demand higher risk compensation. Table 2.4 summarizes the studies on asset pricing for market skewness based on high-order moment.

2.3.2.2 Financial Pricing Research in the Perspective of Market Kurtosis

Further incorporating the kurtosis moment attribute into high order moment CAPM, explanatory power of co-kurtosis on returns of emerging market financial assets is stronger than that of co-skewness attribute (Hwang and Satchell, 1999). For investment portfolios with different sample periods, the extension of the sample period can gradually reduce the explanatory power of the three factor model in terms of systematic co-skewness and co-kurtosis (Chung and Schill, 2006). By establishing a high order moment asset pricing framework under neutral volatility of security risk, Conrad et al. (2013) used option prices and stock market data to extract the density function of high order moment distribution and tested the relationship between cross-sectional differences of high order moment factors in security returns and later returns, and studied the negative correlation between expected kurtosis changes and later returns. The research results with those of Ang et al. (2006), Ang et al. (2009) and Amaya et al. (2015) on trait moment volatility studies are basically consistent.

To avoid the inherent bias of cross-sectional regression methods, Markowski (2024) pointed out that non-linear models have significantly better explanatory power with common skewness and kurtosis than traditional CAPM and Fama French models. In response to the shortcomings of traditional portfolio theory that does not consider high order moment risk, Guo et al. (2024) designed a Generalized Autoregressive Conditional Heteroscedasticity Skewness Kurtosis (GARCHSK) model that considers skewness and kurtosis information. The study found that the China stock market not only has high order moment attribute risk, but also risk has time-varying characteristics. Therefore, designing a model based on high order moment dynamic portfolio can improve the effectiveness of financial asset investment portfolios (Zhu et al., 2022; Luo et al., 2024).

Table 2.5: Summary Financial Pricing in the Perspective of Market Kurtosis

Author & Year	**Methodology**	**Study Result**
Hwang and Satchell (2001)	Four high moment CAPM	Emerging market financial asset returns better than the covariance property
Chung and Schill (2006)		Longer sample periods can reduce the explanatory power and significance of the three-factor model
Ang et al. (2006, 2009) Conrad et al. (2013) Amaya et al. (2015)		The change in kurtosis was negatively correlated with late yield, the Negative skewness is positively correlated with returns
Fry et al. (2014b)	Forex Option Models	Significant pricing errors in option prices
Fry et al. (2010a, 2014) Chan et al. (2018)	Monte Carlo non-parametric equation	Positive covariance risk and higher covariance risk contagion relationships can significantly explain the extent of the decline in risk premiums for financial assets during the financial crisis
Guo et al(2024)	GARCHSK model	Models based on higher-order moments dynamic portfolios can improve the effectiveness of financial asset portfolios.
Zhu et al.(2022) Markowski (2024) Luo et al.(2024)	Nonlinear option risk index estimation method	Non-linear models have strong fitting effects

Risk contagion, especially those occurring on high order moment attributes, are indispensable factors to conduct asset pricing research (Fry et al., 2010a). A high order moment asset pricing model that integrates the asymmetry and tail information shocks is used to study asset premiums under the assumptions of two-dimensional normal and skewed distributions. Monte Carlo non-parametric analysis is used to examine the contagion effects of co-skewness, co-kurtosis, and co-volatility between portfolio assets on risk asset premiums, research has found that the risk contagion relationship between positive co-skewness and high co-kurtosis can significantly explain the degree of decline in financial asset risk premiums during financial crises (Fry et al., 2010a; Fry et al., 2014, Chan et al., 2018). Sudden financial crisis are important triggers of risk contagion, and it is found that during financial crises, the high order moment attribute risk contagion relationship between capital markets is also more significant. Using the exchange option model, it is found that there is a greater risk exposure without considering the high order moment risk contagion (Fry et al. 2014b). Table 2.5 summarizes the studies on asset pricing for market kurtosis based on high-order moment.

2.3.3 Research on Price Characteristics of Carbon Market

2.3.3.1 General Price Characteristics

The heteroskedasticity of financial asset price originates from the significant non-stationary characteristics, while the non-stationary data can lead to strong volatility and time-varying variance changes, and then forming distribution characteristics such as volatility clustering and sharp peaks and thick tails (Engle, 1982). Among them, the ARCH model and the GARCH model created by Engle and Bollerslev have become effective tools for capturing and mapping the non-stationary heteroskedasticity time-series (Bollerslev, 1986; Bollerslev, 1996).

As a special financial markets, carbon market have basic financial attributes, and market returns have obvious characteristics of volatility clustering, sharp peaks and thick tails, as well as stronger time-varying variance fluctuations (Montagnoli et al., 2010; Chevallier, 2010a; Nazifi, 2013). It has been found that there is significant heteroskedasticity in spot price of European Union Allowance (EUA), with the tail distribution following a Poisson distribution rather than a normal distribution (Taschini & Paolella, 2008). Benz &Truck (2009) captured the heteroskedasticity characteristics of the carbon spot return using the Markov transformation model and AR(1)-GARCH(1,1) model. GARCH cluster models, which incorporate heteroskedasticity, time-varying variance, and time-varying jump-variance, demonstrate effective predictive capability for European carbon prices (Byun & Cho, 2013; Koop & Tole, 2013; Sanin et al., 2015). The return series of carbon futures assets in the European carbon market show sharp peaks and thick tails and volatility clustering characteristics, the volatility clustering phenomenon is more obvious as the increasing of transaction costs (Palao & Pardo, 2012). By constructing the Copula-ARMA-GARCH model to depict the nonlinear relationship among the multiple factors of carbon prices, the study shows that the carbon prices have volatility clustering and heteroskedasticity characteristics, the sharp peaks and thick tails characteristics of carbon price are significantly stronger than other markets (Zhang et al., 2020; Zhou et al., 2022). Based on Ensemble Empirical Mode Decomposition (EEMD) technique, it was found that there are spatio-temporal heterogeneity characteristics of carbon price volatility influenced by market institutional design and policy adjustments (Li et al., 2021; Kong et al., 2022). The Asymmetric Generalized Dynamic Conditional Correlation Model

(AGDCC-GARCH) was used to analyse the dynamic correlation between the European carbon market and the China stock market. The research results showed that, similar to return characteristics of traditional financial assets, the EUA carbon futures also existed more obvious characteristics of sharp peaks and thick tails, and volatility clustering (Lu & Wang, 2009).

2.3.3.2 Special Price Characteristics

Compared to general price characteristics, specific characteristics of carbon price primarily emphasize their commodity attributes. This is evident in their heightened sensitivity to policy shocks, asymmetric market volatility, and non-linear multi-fractal characteristics.

(1) High sensitivity to policy shocks

The carbon price has a significant policy driving impact, and price volatility is highly susceptible to emissions reduction technologies, quota policy adjustments, carbon tax policy, as well as energy and environmental policies in the energy market.

Although a well-designed carbon market will effectively reduce emission costs and promote innovation, achieving sustainable development of a low-carbon economy still requires constructive incentive and protection policies for the operation of carbon market (Peace et al., 2009). Research has shown that borrowing and storage system of carbon allowances forms a foundational element for effective operation of carbon market (Rubin, 1996). Excessive allocation of carbon allowances and the allowance storage policy are important forces that caused a serious decline of carbon price in the first stage of EUCM (2005-2007) (Ellerman and Montero, 2007). By studying the structural adjustments experienced by the European carbon price in the first stage, it was found that the structural adjustment of carbon allowances and the market response to carbon prices usually follow the European releases allowance information (Alberola et al., 2008). The restriction of EU ETS banning on cross period storage has a significant impact on the EUAs price (Alberola and Chevallier, 2009). Based on Hotelling CAPM analysis, the study suggests that policies of French and Poland banning banks from engaging in carbon business have also led to the price declining of EUAs in the first stage of EUCM.

The government's carbon tax is also an important external factors that leads to volatility of carbon price. Research has concluded that imposition of carbon tax can effectively reduce carbon emissions from manufacturing enterprises, especially those with greater responsibility to reduce emissions, to participate more actively in the carbon trading market and purchase excess stock of allowances, thus moderately increasing the market demand for carbon emission rights and driving up the price price (Xu et al., 2023; Hu et al., 2020; Zhou et al., 2022).

Due to rescue measures introduced by the European Union (EU) and the United Nations (UN), such as “volume auctions" and the implementation of storage management mechanisms aimed at controlling the supply and price of emission rights, as well as uncertainties surrounding carbon emission reduction policies in the post-Kyoto era, the trading price of Certified Emission Reduction (CER) showed a sudden increase in 2015 (Zhang, 2018). Through event analysis and wavelet analysis, it has been discovered that incidents such as data leakage in the European carbon market, the US subprime crisis, and the European debt crisis have all influenced the price volatility of the European carbon market. This underscores the carbon market's characterization as a typical "policy market." (Adekoya et al., 2021; Zeng et al., 2021; Qiu et al., 2023).

The above literature studies affect of relevant policies on carbon prices. Further research has found that policy events related to capital or energy markets can also have an impact on carbon prices through financial channel (Hammoudeh et al., 2015; Reboredo, 2014). The policy changes related to the carbon price can also be transmitted to carbon market itself through globalization of financial markets and cross market linkage mechanisms. The study has highlighted the significance of national energy and environmental policies, including energy-saving and emission-reduction measures, which exert a crucial influence on the price dynamics of the fossil energy market. Given that consumption of fossil energy are primary sources of carbon emissions. Therefore, when the fossil energy market is faced with the background of energy policy changes or industrial technology upgrading, energy consuming enterprises, especially high-energy-consuming power generation firms, can switch between various power generation fuels such as coal, natural gas and oil. This forms an inherent price transmission mechanism between the fossil energy market and the carbon market. That is, an increase in energy prices will drive up carbon prices, the decrease in energy prices will also lead to a decline in carbon prices (Convery, 2007; Chang et al., 2019; Gong et al., 2021; Jiang et al., 2023). Table 2.6 summarizes the studies on high policy sensitive characteristics of carbon price.

(2) Asymmetric market volatility

Using asymmetric GARCH models such as GJR-GARCH, EGARCH, TGARCH, it is found that the tail data of carbon price has obvious asymmetric characteristics (Chevallier, 2009). And there is a significant leverage impact on carbon price, the negative impact of carbon market excess on prices gradually weakens (Wang et al., 2018; Lin et al., 2019).

Table 2.6：Summary of High Policy Sensitive Characteristics of Carbon Price

Author & Year	Methodology	Influence Factor	Result	Country/Region
Peace et al.(2009)	CAPM	Carbon emission reduction technology innovation	Significant impact：+	European
Rubin(1996) Elleman and Montero(2007) Alberola et al. (2008) Alberola and Chevallier(2009)	Hotelling CAPM analysis	Adjustment of carbon emission quota policy	Significant impact limit：- Information disclosure：+	European
Hu et al.(2020); Adekoya et al.(2021) Zeng et al.(2021); Zhou et al.(2022) Xu et al.(2023); Qiu et al.(2023)	Event analysis and wavelet analysis	Carbon tax policy	Carbon tax collection: rising carbon price	European
Reboredo (2014) Hammoudeh et al.(2015)	Event analysis	Energy policy	+	European
Convery(2007); Chang et al.(2019) Gong et al.(2021); Jiang et al.(2023)		Environmental policy	+	European

Based on heterogeneity differences in market volatility, the asymmetric GARCH cluster models were used to find that when the volatility trend of carbon market rises, carbon prices show a positive Monday effect and a negative Tuesday effect, while when the volatility trend decreases, they show a negative Monday effect and a positive Tuesday effect (Zhang et al., 2019). Investor overconfidence and risk preference differences are the main causes of asymmetric volatility in carbon market, the higher degree of investor overconfidence, the higher expected price of carbon futures, and investors overconfident behaviour drives abnormal volatility of carbon futures returns (Zhang et al., 2018). According to data analysis of European power companies, it was observed that the spillover effects from carbon and electricity market have obvious asymmetric effect on the volatility spillover of stock returns, the spillover intensity of negative returns is higher than that of positive returns (Ji et al., 2018). The ARCH and GARCH cluster models were used to study the domestic carbon market, it was observed that price volatility of China carbon market is generally characterized by volatility asymmetry and market policies sensitivity (Yun et al., 2020; Yang et al., 2020). Further study suggests that the carbon market in Shenzhen shows persistent volatility and obvious time-varying jump behavior in terms of market returns, the jump direction exhibits significant asymmetry, meaning that after a market price increase, the next jump direction tends to be more positive (Han et al., 2019; Liu et al., 2021).

(3) Nonlinear multi-fractal characteristics

The price behaviour of EUETS is a nonlinear dynamical system with fractal characteristics. As the market efficiency is not completely efficient, so its price behaviour, trading mechanism cannot be studied with a linear paradigm (Fan et al., 2019; Zou et al., 2020). By studying price quantity multi-fractal characteristics between the European and Hubei carbon market, research shows that price relationship between European carbon futures market and Hubei carbon market has nonlinear multi-fractal and long-term memory characteristics, the fractal degree of Hubei carbon market is stronger than that of the European carbon futures prices (Zhu et al., 2023). Cross-correlation has observed in price of the EUCM that with multi-fractal characteristics. Moreover, there is a significant asymmetry in the multi-fractal characteristics between price and volume relationships of the carbon market, especially when returns and trading volume changes are on an upward trend, the price relationship becomes more complex, and the corresponding market risk increases (Wang et al., 2023).

Dividing carbon prices into different frequencies and using the Hilbert spectrum model to capture shock of extreme events, economic crises on carbon price at different time scales (Zhu et al., 2018a). The EUA carbon futures prices experienced multiple significant structural changes (Alberola et al., 2008), the regional transformation equation can reflect the nonlinear change characteristics of carbon price volatility. The Hurst index using the R/S method can determine long-term memory fractal characteristics of carbon price (Zhu et al., 2015). The multi-fractal model with long-term dependence and regional transformation found that the carbon spot price also has multi-fractal characteristics (Segnon et al., 2017). Table 2.7 summarizes the studies on market asymmetry and nonlinear multifractal characteristics of carbon price.

Table 2.7：Summary of Special Fluctuation Characteristics of Carbon Price

Special Characteristics	**Author & Year**	**Methodology**	**Impact Result**	**Country/Region**
Market asymmetry	Zhang et al.(2019)	GARCH	Monday：- Tuesday：+	European
	Ji et al(2018)		Significant impact	European, China
	Yang et al.(2020) Yun et al.(2020)	ARCH and GARCH	Asymmetry of price fluctuation	European
Nonlinear multifractal	Alberola et al.(2008) Zhu et al.(2015)	Hurst index of R/S method	Significant structural mutations	European
	Segnon et al.(2017) Fan et al.(2019) Zou et al.(2020)	Multi-fractal model	With multifractal characteristics	European, China
	Zhu et al.(2018a) Wang et al.(2023)	Hilbert spectrum	Nonlinear structural catastrophe characteristics	China
	Zhu et al.(2023)	Wavelet three-layer transform and neural network model	With local scale diversity	European

2.4 Carbon price Forecasting Research

2.4.1 Carbon Price Forecasting Under the GARCH Cluster Models

2.4.1.1 Linear GARCH Cluster Forecasting Models

Based on correlation analysis and multivariate linear modelling, it was found that energy market returns have the most significant impact on carbon prices. The GARCH cluster models are able to better fit the carbon futures price (Byun & Cho, 2013). Study has concluded that the GARCH model based on Markov regime switching is superior to other GARCH cluster models in forecasting short-term carbon prices (Oertel et al., 2022). Zeitlberger et al. (2016) suggest that AGARCH and GJR-GARCH models can accurately forecast European carbon futures price. Through the asymmetric threshold GARCH models, Chevallier (2009) studied empirical relationship between carbon futures prices and macroeconomic changes. Research has pointed out that stock and bond market, namely stock dividend and "junk bond" premiums, can effectively explain the asymmetric fluctuations in carbon futures prices. However, interest rate and commodity market returns are not robust in forecasting carbon futures prices, and research reveals that macroeconomic have relatively weak impact on the carbon prices volatility. Even if energy market factors are included in the study framework, the research conclusions remain broadly robust. Using a multivariate GARCH model, Oberndorfer (2009) pointed out that price changes in European emission allowances (EUA) are positively correlated with stock returns of power companies.

However, fluctuations in the stock market have not caused fluctuations in the EUA market. Using multiple factor model and panel quantile regression method, the study shows that stock market returns can effectively explain the European carbon market returns, and this correlation shows heterogeneous differences at different stages of EUCM. In first stage (2005-2007) and third stage (2013-2020) of the European carbon market, carbon returns have positive correlation with stock market returns, whereas in the second stage (2008-2012) show negative correlation (Zhu et al., 2018). Chevallier (2010b) constructed a dynamic AR(1)-GARCH(1,1) model based on 115 indicators to effectively forecast the carbon price volatility. Using daily data volatility, option price, and intraday data to measure three types of EUA volatility, the study found that policy is important evidence for explaining European carbon price (Chevallier, 2011). Conducting the DCC-GARCH and ARCH model to study volatility spillover between fossil energy and carbon market, it has found that returns in coal, crude oil and natural gas markets significantly impact short-term European carbon price (Zhang et al., 2016; Jie et al., 2021). Carbon market exhibits price fractal characteristics (Zhu et al., 2021;Liu et al., 2021).

2.4.1.2 Nonlinear GARCH Cluster Forecasting Models

The carbon price has nonlinear dynamic feature, traditional single and multiple linear regression models are difficult to reveal the driving path of carbon prices. Studies have shown that the European carbon spot price show random walk volatility, and carbon futures returns show unconditional tail behaviour and dynamic heteroskedasticity. The hybrid models considering jump diffusion, and GARCH models demonstrates a robust capability to accurately model and forecast carbon prices (Daskalakis et al., 2009; Taschini & Paolella, 2008; Seifert et al., 2008; Zhou et al., 2022).

Different from the findings of linear regression methods, studies based on nonlinear autoregressive models found that stock prices, especially those of clean energy companies, cannot explain the volatility of European carbon markets (Kumar et al., 2012). Koop et al.

(2013) used a Dynamic Mean Analysis (DMA) nonlinear model to forecast the European carbon futures prices by infected markets of energy market product returns, climate factors, capital market factors, corporate risk premiums, and carbon homogeneous products. They found that the price forecasting accuracy was significantly better than Bayesian models and Time-Varying Parameter (TVP) regression models. Due to the abnormal returns and fluctuations in the European carbon futures market, the ARMA-GARCH model based on the assumption of normal distribution is unable to forecast the carbon market returns, for improving the model predictability, the study incorporated the random jump process of carbon futures into the price formation framework. It was found that ARMAX-GARCH model based on the Gaussian mixture distribution with time-varying jump process can effectively reveal price changes of European carbon futures market (Sanin et al., 2009). Affected by the supply and demand of market carbon quotas, the price changes in European carbon market show heterogeneous volatility characteristics. The GARCH model based on Markov regime transfer has better forecasting performance for short-term carbon spot prices than other GARCH cluster models (Benz & Stefan, 2009). Using the improved linear quantile regression model and Nonlinear Auto-Regressive Distributed Lag (NARDL) model to test nonlinear and asymmetric relationship between energy market prices and carbon prices. The study has showed that crude oil prices exert a long-term negative asymmetric effect on European carbon prices. In the short term, declines in coal prices have a greater impact on carbon prices than increases do, while natural gas prices and electricity prices exert a symmetric effect on carbon prices (Hammoudeh et al., 2014). Based on the ARCH auto-regressive lag model, it was found that the volatility of coal market is the main cause for price changes in carbon markets. The dynamic fluctuations in crude oil, natural gas, and coal prices exert a profound and far-reaching influence on the dynamics of carbon prices, especially short-term carbon prices (Kim & Koo, 2010; Zhu et al., 2023). Especially, the model can reveal the non-linear effect of policy regulation events on price of carbon market (Ren et al., 2020). Table 2.8 summarizes the studies on carbon market pricing based on GARCH cluster models.

2.4.2 Carbon Price Forecasting Under the Artificial Intelligence Technology

2.4.2.1 Integrated Artificial Intelligence Models for Forecasting Carbon Price

Compared with single price forecasting model, integrated price forecasting model can effectively forecast nonlinear and non-stationary carbon price sequences. Empirical Mode Decomposition (EMD) technique can decompose nonlinear non-stationary carbon price signals into time-frequency-differentiated mode components, revealing formation mechanism of carbon price from different time-scale perspectives (Zhu et al., 2015). Zhu et al. (2018) proposed a carbon price mixed forecasting model that combines EMD and Least Squares Support Vector Machine (LSSVM). Research found that multi-scale nonlinear integrated method has high fitting accuracy for forecasting carbon price. The carbon price signal is divided into multiple time-scale modes, and GARCH and LSSVM models are used for component forecasting, and the forecasting models are optimized by Particle Swarm Optimization (PSO), Genetic Algorithm (GA), Spiking Neural Network (SNN) and Deep Neural Network (DNN). It is found that carbon price forecasting performance of the EMD-ARMA-LSSVM model as well as other combination models, is significantly better than the single model effect (Zhu et al., 2019; Elsayed et al., 2022). Based on the improved Empirical Mode Decomposition technique, Yang et al. (2020) used the Improved Whale Optimization Algorithm (IWOA) to optimize the LSTM prediction model, and the study showed that the hybrid model IWOA-LSTM was stable and robust in forecasting prices of the emerging carbon markets in Beijing, Fujian and Shanghai. To solve the issue of mode mixing that easily arises in EMD mode decomposition, Li et al. (2022) used Complementary Ensemble Empirical Mode Decomposition (CEEMD) and Variational Mode Decomposition (VMD) to decompose the original carbon prices and the obtained Intrinsic Mode Function (IMF) with maximum sample entropy. The results show that the mode decomposition technique has a obvious advantage in forecasting carbon price. Breaking through shortcomings of the traditional EMD technique in point forecasting, Ji et al. (2022) constructed a three-stage vertical carbon price interval forecasting model, and used Kernel Density Estimation (KDE) algorithm for interval estimation, the results suggested a stronger credibility. Liu et al. (2022) proposed interval multi-scale decomposition methods, including Interval Variational Mode Decomposition (IVMD), for regional trend decomposition and carbon price forecasting. Empirical evidence shows that the above models are effective means of interval price forecasting.

Carbon prices have nonlinear and non-stationary features, and a hybrid price forecasting framework that integrates multiple mode has more accuracy advantages (Zhou et al., 2019). Xiong et al. (2019) proposed a hybrid multi-step forecasting model that integrates Variational Mode Decomposition (VMD), and applied the proposed hybrid to forecast carbon price of China carbon market. The research found that compared with other multi-output models, the integrated VMD model has better performance in terms of carbon price forecasting accuracy and stability. Based on similar modelling ideas, a hybrid model for carbon price forecasting consisting of extreme value learning machine and grey wolf optimization algorithm is integrated (Zhang et al., 2019; Zhou et al., 2019). The hybrid model for carbon price forecasting integrating empirical mode decomposition, sample entropy, and particle swarm optimization improved extreme value learning machine (Sun and Duan, 2019) have good forecasting performance for price of China regional carbon market. Hao et al. (2020) proposed a carbon price forecasting model according to two-stage feature selection and multi-objective optimization algorithm. The study shows that after selecting pricing factor variables, the multi-objective grasshopper optimization algorithm is used to optimize weighted regularization extreme value learning machine, which can forecast European carbon price and China carbon price better. Used the CEEMD and VMD technologies to secondary decompose carbon price, conducting SVM, BP network for forecasting, previous researches shown that CEEMDVMD-BP, CEEMD-VMD and CEEMD-VMD-SVM model have obvious price forecasting advantages in China carbon market (Li et al., 2023;Yang et al., 2023).

2.4.2.2 Artificial Neural Networks for Carbon Price Forecasting

Carbon price forecasting models based on traditional statistical measures usually require the market returns to follow strict parametric assumptions and tail distribution assumptions (Ji et al., 2018; Zhang et al., 2019), which makes some parameter structures unable to show special characteristics of carbon price. Artificial intelligence technologies, such as ANN, SVM, LSSVM, and Multi-Layer Perceptrons (MLP), have obvious advantages compared to traditional statistical econometric models in solving the price fitting and forecasting of the carbon price (Atsalakis and Valavanis, 2009; Oliveiraet al., 2013; Zhang et al., 2017), because those models do not considering the tail distribution of return sequences, and can achieve the advantage of self-learning and adaptive adjustment.

According to the optimized Extreme Learning Machine (ELM) model, the Kidney Algorithm (KA) with scaling factor and cooperation factor (CKA) model are established to forecast China carbon price. The findings reveal that the MOEMD-CKA-ELM performs well in carbon price forecasting (Huang et al., 2020). Using Computational Intelligence

Techniques (CIT) such as new hybrid neural fuzzy controller (PATSOS), ANN to forecast carbon prices changes (Atsalakis, 2016). Considering correlation of various mode components, Sun et al.(2020) proposed a hybrid carbon price forecasting model incorporating factor analysis, empirical mode decomposition and least squares support vector machine, the superiority of the model was tested and proved on the price forecasting in China carbon markets. Based on the same mode decomposition modeling approach, Huang et al. (2021) constructed an integrated carbon price forecasting model based on Variational Mode Decomposition (VMD). The findings showed that the integrated models VMD-LSTM were able to effectively forecast European carbon price, especially during price increase stage, those models have the strongest stability, while the EMD-VMD-LSTM model and VMD-GRU models, on the other hand, present good forecasting accuracy only in China carbon market (Sun et al., 2020; Wang et al., 2021). According to maximum Lyapunov index and Kolmogorov entropy, the third stage of European carbon price was studied from a chaotic system. It was found that MLP neural network can forecast carbon price effectively (Fan X et al., 2015). The Finite Distribution Lag (FDL) model based on the Genetic Algorithm (GA) and Ridge Regression Algorithm (RRA) can independently select suitable regressors, and has better performance in forecasting and analyzing complex carbon prices than other GARCH cluster models (Han et al., 2015). According to integrated learning idea, GA-ANN is used to forecast various mode components (Zhang et al., 2020; Zhu et al., 2018b; Qin et al., 2020). A hybrid forecasting model is constructed, combination model is optimized by Spike Neural Network (SNN) and Grey Neural Network (GNN). The study pointed out that the combined model can effectively fit and forecast carbon prices (Sun et al., 2016; Zhu et al., 2019; Zhang et al., 2018). To overcome the neglect of high order moment terms such as market skewness and kurtosis, which represent market asymmetric information and external event shocks, Yun et al. (2020) proposed a multi factor carbon price forecasting framework based on the multi factor APT concept, which extends the high order moment. A multi-layer multivariate LSTM model was used to map the nonlinear relationships of carbon price. Study has found that compared to pricing frameworks that do not consider high order moment risk contagion, the LSTM models based on extended high order moment price forecasting frameworks have significant fitting advantages, with better performance and stability than other neural networks and nonlinear volatility models. Table 2.9 summarizes the studies on carbon market pricing based on artificial intelligence techniques.

2.5 Summary of the Chapter

According to the above analysis of the literature review, it shows more clearly that the price forecasting of carbon price is based on the traditional financial asset pricing theories. The research perspective have change from the low order moment attribute of asset returns and volatility to the high order moment attribute of skewness and kurtosis. The price forecasting models of carbon market also develops from linear or nonlinear GARCH cluster models to the artificial intelligence methods such as artificial intelligence and machine learning. These studies provide empirical references for further studying the topic of carbon price forecasting in China. While through the literature review, there are still some improvements need to be solved in the future.

(1) Existing carbon price forecasting research listed in the Section 2.4 above lacks a theoretical framework that considers high order moment risk contagion relationship.

Firstly, existing carbon price forecasting research mentioned in Section 2.4 mainly consider low order moment attribute of returns, adopts the influence factor analysis method to study the return transmission and volatility spillover between carbon price and its infected market. However, these research are limited to explaining the premium mechanism of low order moment attribute, and does not consider high order moment attribute. Because high order moment attribute of skewness and kurtosis reflect the asymmetry and sensitivity to extreme events, those characteristics are completely consistent with carbon market. That is to say, existing research listed in above literature neglects to consider influence mechanism of carbon price from the high order moment attribute, making it difficult to reflect the operation rules of carbon price, and the accuracy of relevant price forecasting conclusions is questionable.

Secondly, previous studies discussed in Section 2.4 neglect risk contagion relationship from sourced carbon market to infected markets. The theory of high order moment risk contagion can explain price changes purely due to irrational and synergistic fluctuations in the market beyond market fundamentals, and the explanatory perspective of this return fluctuation conforms to features of carbon prices. Therefore, existing research that neglects to consider risk contagion relationship of high order moment attribute, which will make it difficult to provide stronger evidence to explain the carbon premium.

(2) Existing research methods listed in the Section 2.4 for testing the high order risk contagion are difficult to satisfy the characterization the volatility trends heterogeneity in the carbon market.

Previous studies about measures high order moment risk contagion discussed in Section 2.4 mainly consider event affect, such as financial crisis. However, this event-based method, on the one hand, does not consider the volatility trend heterogeneity of the financial return. On the other hand, it also can not identify affect factors of risk contagion in carbon market. To thoroughly investigate high order risk contagion in the carbon market, it is crucial to uncover the risk contagion relationship from sourced carbon market to infected markets under varying market volatility trends.

(3) Existing modeling techniques listed in the Section 2.4 for carbon price forecasting are difficult to fit and map non-linear the special carbon price forecasting framework with high order moment risk contagion relationships.

Existing carbon price forecasting models do not capture carbon price characteristics of high order moment risk contagion well. Under the high order carbon price forecasting framework, the carbon price infected markets include not only low order moment information reflecting market returns and risks, but also high order moment skewness and kurtosis information. However, among the existing models, the statistical modeling technique represented by GARCH models require the return follow strict distributional assumptions, which is difficult to reflect the real volatility characteristics of carbon prices. The artificial intelligence modeling technique based on hybrid models focuses on the processing of single moment attribute dimension information. Artificial neural network modeling techniques are prone to forget or difficult to capture data characteristics of financial time series with longer time intervals. The algorithm based on reverse gradient descent may suffer from gradient explosion and gradient disappearance, resulting in forecasting models presenting under-fitting or over-fitting, and the model may not converge to the optimal solution.

Table 2.8: Summary of CMP Based on GARCH Cluster Models

Author & Year	Methodology		Research Target	Findings	Country/Region
Byun & Cho (2013) Oertel et al.(2022)	Linear GARCH cluster model	GARCH cluster model	Impact of energy market returns on carbon prices	+	European
Chevallier (2009) Zeitlberger et al.(2016)		Asymmetric threshold GARCH family models	Carbon Futures Returns and the Macroeconomy	The impact is relatively weak	European
Oberndorfer (2009)		Multivariate GARCH model	EUA price changes and electric utility stock returns	+	European
Chevallier (2010b);Chevallier (2011) Zhang et al.(2016); Jie et al.(2021)		AR(1)-GARCH(1,1) model	Measuring the three volatilities of the EUA	Carbon price volatility instability	European, China
Zhu et al. (2018); Zhu et al.(2021)		Multifactor modelling and panel quantile regression	Carbon financial asset returns versus equity market returns	-	European, China
Taschini & Paolella (2008) Seifert et al. (2008) Daskalakis et al. (2009)	Non-linear GARCH cluster model	Time-varying GARCH models with nonlinear parameters	EU carbon spot market price	Goodness of fit	European
Sanin et al. (2009) Benz & Stefan (2009)		ARMAX-GARCH model	European Carbon Futures Market Benefits	Effective explanation of changes in earnings	European
Kim & Koo (2010); Ren et al.(2020)		ARCH autoregressive lag model	Carbon financial asset return impact triggers	High impact of volatile coal market returns	European
Koop et al. (2013); Zhou et al.(2022)		Dynamic Mean Nonlinear (DMA) Modelling	EU Carbon Futures Yield Price Forecast	Significantly stronger pricing accuracy	European, China
Hammoudeh et al. (2014); Liu et al.(2021); Zhu et al. (2023)		NARDL model	Crude oil prices and CO2 quota prices	Negative asymmetric effects	European, China

Table 2.9: Summary of CMP Based on Artificial Intelligence Techniques

Author & Year	Methodology		Effectiveness on Carbon Price	Country/Region
Zhu et al. (2015,2018,2019) Elsayed et al.(2022);Zhou et al.,2019	Integrated modelling techniques for artificial intelligence	Multi-scale nonlinear integrated learning methods	Good fit	European, China
Xiong et al. (2019);Liu et al. (2022)		VMD-FMRVR-MOWOA integration	High accuracy and stability	China
Zhang et al. (2019);Zhou et al. (2019) Sun and Duan (2019);Hao et al. (2020) Sun et al. (2020);Li et al.(2023)		Hybrid models for carbon price forecasting	Better predictive performance	China
Zhu et al.(2018b); Zhang et al.(2020); Qin et al.(2020)	Artificial neural network	GA-ANN	Strong predictive ability	European, China
Atsalakis (2016)		PATSOS	Good prediction	European
Fan X et al. (2015);Han et al. (2015)		MLP model and FDL model	Effective forecasting	European
Zhu et al. (2018b);Qin et al. (2018)		GA-ANN model	High predictive accuracy and stability	China
Sun et al. (2016);Zhu et al. (2019);Zhang et al. (2018)		CFM Hybrid Predictive Modelling	Goodness of fit	China
Zhu et al. (2019);Zhang et al. (2021)		EMD-ARMA-LSSVM model	Significant predictive accuracy and stability	China
Sun et al. (2020);Wang et al. (2021);Huang et al. (2021)		VMD-GARCH and LSTM-LSTM	Stable and robust	European, China
Yun et al. (2020); Ji et al. (2022)		Multilayer multivariate LSTM models	Significant fit performance and stability	China
Li et al. (2021);Yang et al.(2023)		VMD and ICEEMD model	Improved credibility of forecasts	China
Liu et al. (2022); Guo et al. (2024)		IEMD and IVMD models	Good prediction	China

CHAPTER 3

METHODOLOGY

3.1 Introduction

This chapter discusses the research framework, research variables and data sources. This chapter also introduce contagion model for recognizing high order risk contagion, furthermore a high order contagion forecasting model (HOC-LSTM) has been constructed, which can provide theoretical support for research design and empirical analysis.

3.2 Research Framework

The objective of this chapter is constructing the risk contagion recognizing model and price forecasting model of the China carbon market. The first objective is to test the high order moment risk contagion from risk source carbon market to infected markets. During risk contagion test, using the Markov mechanism transformation model to classify price volatility carbon market into three states: high volatility, medium volatility and low volatility. For testing the specific manner of high order moment risk contagion relationship from the carbon market to its infected markets, this thesis constructs the Co-Skewness (CS) risk contagion model, Co-Kurtosis (CK) risk contagion model and Co-Volatility (CV) risk contagion model to identify the risk contagion relationship from risk source carbon market to its infected markets. The second objective is to construct a machine learning carbon price forecasting model that suitable for capturing the impact of high order moment risk contagion on carbon price. The last objective is to forecast the China carbon price by HOC-LSTM model to prove the high order risk contagion is useful to improve the forecasting performance. During the test, the high order risk contagion relationship recognized by the above models are as the input infected markets for forecasting, the machine learning LSTM model is used to take out-of-sample forecasting of the carbon prices. This chapter constructs a theoretical framework for forecasting carbon price from a theoretical perspective and proposes a theoretical solution to the scientific problem among other things (as shown in Figure3.1).

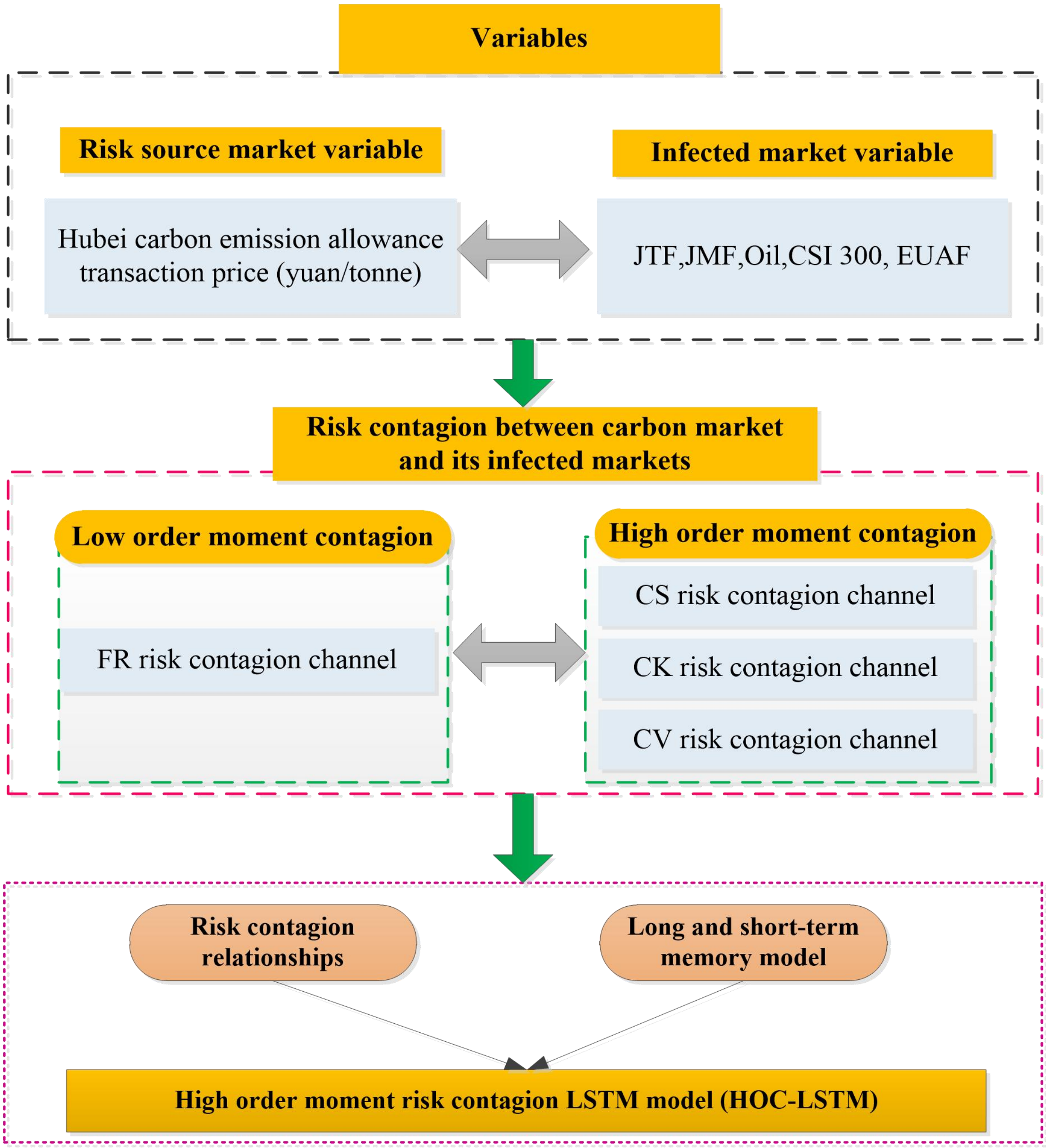


Figure 3.1: The Theoretical Framework of Methodology

It is worth noting that the meanings of some variable abbreviations in Figure 3.1 are as follows: the full name of JTF is Jiaotan Futures, the full name of JMF is Jiaomei Futures, the full name of Oil is Brent Crude Oil Futures, the full name of CSI300 is China Securities Index 300Futures, the full name of EUAF is European Carbon Allowance Futures.And the FR means the low order moment correlation coefficient risk contagion channels, means the Forbes Rigobon Contagion.

3.3 Research Variables and Data

3.3.1 Risk Source Market Variable

In this study, the average transaction price of Hubei carbon emission allowance (HBEA) is chosen as return series of carbon assets, so the Hubei carbon price is the risk source market variable of the forecasting model. The reason for choosing the Hubei carbon emission allowance is that Hubei carbon market has already led the China regional market in terms of market transaction scale, introduction of social capital and participation of emission reduction enterprises, and has undertaken the task of operating role of national carbon emission right registration system. So, the market system construction and operation regulations of Hubei carbon market are relatively mature. Another important reason is that, with the official establishment of China national unified carbon market in 2021, the Hubei carbon market, with its mature operational experience in the introduction of social capital and enterprise participation, has play a key role in the operation of the national carbon market in market registration, enroll, and settlement. Therefore, Hubei carbon market is representative.

3.3.2 Infected Market Variable

As a product innovation with financial attributes, the carbon market not only has closely related information linkages and spillovers with the capital market and homogeneous product market, but also has correlation with the energy market (Nazifi et al., 2010; Aatola et al., 2013). Based on this, this study selects the product instruments of carbon homogeneous market, capital market and energy market as the carbon pricing factors, which are the infected market variables of the model. Among them, the homogeneous products are selected as EUAF, the European carbon allowance futures product traded in the European carbon market. The capital market is selected as CSI 300 index, which represents China's macroeconomy. The energy market products are selected as Jiaotan futures (JTF), Jiaomei futures (JMF), and Brent crude oil futures. The specific variable definitions are shown in Table 3.1. Using R_t to represent the return of carbon assets, one can define: $R_t = 100\times(lnP_t - lnP_{t-1})$, where P_t represents the daily price or index series of carbon price and its infected markets.

Table 3.1: Research Variable Design

Financial Market	**Representative Trading Products**	**Abbreviation**	**Product Meaning**
Panel A: Carbon Market (Risk Source Market)			
Carbon market	Hubei Carbon Emission Allowances	HBEA	Carbon emission allowance transaction price (yuan/tonne)
Panel B: Infected Markets			
Carbon homogeneous products market	EUA Futures	EUAF	EUA Continuous Futures Settlement Price (EUR/tonne)
Capital market	CSI 300 Index	CSI 300	CSI 300 Daily Closing Price
Energy market	Jiaomei Futures	JMF	Coking Coal Futures Daily Settlement Price
	Jiaotan Futures	JTF	Coke Futures Daily Settlement Price
	Crude Oil Futures	Oil	Brent crude oil futures daily settlement price

3.3.2.1 JT Futures (JTF)

JTF is a kind of carbon product. JTF is a commodity futures contract with coke as the underlying material, taking advantage of the standardized characteristics of futures contracts. The product of JTF is listed on the Dalian Commodity Exchange. Participants can hedge risks or make investments by buying and selling coke futures contracts. The main role of JTF is to provide price risk management tools for producers and consumers. Producers can lock in future selling prices by selling futures contracts, while consumers can lock in future purchasing prices by buying futures contracts. In addition, investors can also invest and speculate by buying and selling coke futures.

3.3.2.2 JM Futures (JMF)

JMF, also known as metallurgical coal, is a kind of bituminous coal with medium and low volatile matter, medium bonding and strong bonding. In the national standard for coal classification in China, JMF is a type of bituminous coal with high coaling degree and good coking property. JMF is a commodity futures contract with coking coal as the underlying material, using the standardized characteristics of futures contracts.

3.3.2.3 Brent Crude Oil Futures (Oil)

Brent crude oil, sourced from the Brent region of the North Sea in the North Atlantic Ocean, is extensively traded across futures, over-the-counter swaps, forwards, and spot markets. Futures trading on the Intercontinental Exchange in London and the New York Mercantile Exchange serves as a key benchmark for global oil prices. Brent crude oil futures contracts are contracts that can be physically delivered, while the contracts can be settled with the option to convert from futures to cash.

3.3.2.4 CSI 300 Index (CSI 300)

The China Securities Index 300 (CSI300) comprises the largest and most representative 300 securities with high liquidity from the Shanghai and Shenzhen markets. Officially released on April 8, 2005, the index reflects the overall performance of securities listed in these markets, acting as a "barometer" for the overall trend. The CSI300 index samples are selected from both the Shanghai and Shenzhen stock markets, encompassing the majority of the circulating market value.

3.3.2.5 European Carbon Allowance Futures (EUAF)

EUA futures are financial derivatives traded on the European Emissions Trading System (EUETS), which is a European Union policy designed to combat climate change by limiting greenhouse gas emissions. EUA futures serve as mature instruments for representing the European carbon price.

3.3.3 Data Sources

The time span of above risk source carbon market and infected market variables is from 28 April 2014 to 24 January 2024. Excluding the inconsistent data, a total of 2337 time series data sample were obtained. For machine learning modeling process, the first 80% of the time series data were used as the training set and the last 20% of the data were used as the test set.

For the data source, the daily transaction price of Hubei carbon market is selected for China carbon emission right price data, and the data is sourced from China Carbon Emission Trading Network (https://www.hbets.cn). The price of European carbon allowance futures (EUAF) product is the daily trading settlement price, and the CSI 300 index uses the daily closing price, the data is sourced from the Wind database. The prices of JTF, JMF, and Brent crude oil futures are also selected as daily settlement price data, and the data are sourced from China Dalian Commodity Exchange.

3.4 Risk Contagion Model of the Carbon Market

3.4.1 The Model for Dividing the Market Volatility Trend

The price of carbon market has obvious volatility clustering characteristics, that is, one fluctuation trend usually hides another bigger fluctuation, and this fluctuation has strong time-varying, stochastic and unobservability characteristics. Therefore, it is necessary to establish a fluctuation state recognition mechanism to divide the carbon returns into different state characteristics. Considering the persistence of carbon price fluctuations, the uncertainty of the volatility states number and the unobservability of the mutual transformation between different states, this study establishes a carbon price volatility state transition model to classify asymmetrically volatility of the carbon price based on the idea of Markov state transition model.

Considering carbon price complexity, it is assumed that state of carbon return fluctuation following the Markov stochastic process only depends on the n states before the state. That is, the current return fluctuation is only related to the probability of the current state and previous state, which is the n-order Markov stochastic process. However, due to the independence assumption of state transition, the first-order Markov stochastic processes are sufficient for problem analysis in applications. Therefore, the discussion in this study is also based on the first-order Markov state transition model. That is, it is assumed that the probability of carbon price fluctuation in state M depends only on the probability of state M-1 is $P(M_t \mid M_{t-1}, M_{t-2}, \cdots, M_1) = P(M_t \mid M_{t-1})$, where the formula denotes that in the first-order stochastic Markov process, the historical fluctuation state of the carbon price is irrelevant for forecasting future fluctuations.

Based on this, assuming that the carbon returns with heteroscedasticity characteristics follow a first-order autoregressive AR process, and the variance sequence has an M-volatility states, according to the mechanism transition model proposed by Hamilton (1989), the fluctuation distribution of the carbon returns sequence satisfies the following model:

$$R_t = v(M_t) + \sum_{i=1}^{p} \phi_a(M_t) R_{t-a} + \varepsilon_t \tag{3.1}$$

where R_t is the carbon returns series, $\varepsilon_t \sim N(0, \sigma(M_t)^2)$, represents a stochastic process in which the variance sequence follows the state of M zone system; $t \in \{1, 2, \cdots, k\}$ is an unobservable discrete variable describing the number of fluctuating states, and M_t obeys a first-order Markov chain, so the probability of a transition denoted $P_{ab} = pr(M_t = b \mid M_{t-1} = a, M_{t-2} = \alpha, M_{t-3} = \beta, \cdots) = pr(M_t = b \mid M_{t-1} = a)$; $\{a, b\} \in t$ represents the state variable; $v(M_t)$, $\phi_a(M_t)$ and $\sigma(M_t)$ represent the intercept term, autoregressive coefficient and standard deviation,respectively.

Under the assumption of a normal distribution in the residual series, the conditional probability density of the carbon return R_t in the state M_t is:

$$f(R_t \mid M_t = b, I_{t-1}; \theta) = \frac{1}{\sqrt{2\pi}\sigma(b)} \exp\left[\frac{-(R_t - v(b))^2}{2\sigma^2(b)}\right] \tag{3.2}$$

When the probability $f(M_t = b \mid I_{t-1}; \theta)$ is known, under the condition that I_{t-1} is determined, the probability density of R_t is denoted as:

$$
\begin{aligned}
f(R_t \mid I_{t-1};\theta) = & p(M_t = 1 \mid I_{t-1};\theta) f(R_t \mid M_t = 1, I_{t-1};\theta) \\
& + p(M_t = 2 \mid I_{t-1};\theta) f(R_t \mid M_t = 2, I_{t-1};\theta) + \cdots \\
& + p(M_t = k \mid I_{t-1};\theta) f(R_t \mid M_t = k, I_{t-1};\theta)
\end{aligned} \qquad (3.3)
$$

where I_{t-1} represents the observed values of all variables R_t in state M_t up to moment t-1, that is all the information available up to moment t-1, $\theta = \{p_{ab},\ v_i(M_t), \phi_a(M_t), \sigma_a(M_t)\}$ denotes parameters to be estimated for the volatility transition model, which can be estimated using the $\ln f(\theta) = \frac{1}{n}\sum \ln f(R_t \mid I_{t-1};\theta)$ function applied to the data collected over the observation period.

Using smooth probability to describe the likelihood of carbon returns in various fluctuation states, denoted as:

$$
\begin{aligned}
p(M_t = b \mid I_T;\theta) &= \sum_{i=1}^{k} p(M_t = b, M_{t+1} = a \mid I_T;\theta) \\
&= p(M_t = b \mid I_t;\theta) \times \sum_{i=1}^{k} \frac{p_{ab} \times p(M_{t+1} = a \mid I_T;\theta)}{p(M_{t+1} = a \mid I_t;\theta)}
\end{aligned} \qquad (3.4)
$$

Due to the differences in the smoothing probability among different states, a larger probability indicates a greater likelihood of a specific fluctuation state occurring, while a lower probability indicates a lower probability. Therefore, after calculating the smoothing probability, considering the state value corresponding to maximum smoothing probability, this thesis can refer to the processing method of Jiang et al. (2013), and take the value of 0.5 as the screening threshold for the smoothing probability of each state system. Therefore, to identify the carbon returns related to maximum smoothing probability, and provide an analytical basis for analyzing the high order moment attributes of carbon asset and its infected factors, this thesis uses screening criteria of $p(M_t = b \mid I_T;\theta) > 0.5 \Rightarrow R(M_t)$ to decide the maximum smoothing probability.

3.4.2 Risk Contagion Measure of the Carbon Market

The carbon market is marked by high sensitivity to policy changes and low market efficiency, which requires the design of risk contagion models be able to capture these two features. Based on the non-parametric high order moment risk contagion test method introduced by Fry et al. (2014, 2018), this thesis focuses on the Co-Skewness (CS) risk contagion relationship of market asymmetric information shocks, the Co-Kurtosis (CK) risk contagion relationship of extreme event shocks, and the Co-Volatility (CV) risk contagion relationship of market volatility shocks. For comparability reasons, this thesis also comparatively analyses the risk contagion indicator of Forbes Rigobon Contagion (FR) based on low order moment attributes proposed by Forbes et al. (2002).

3.4.2.1 CS Indicator for Co-Skewness Risk Contagion

The co-skewness risk contagion is an indicator that measures whether there is a significant change in asymmetry of the return distribution between carbon assets and its infected markets. It measures whether the risk contagion coefficient is significantly changes before and after the market volatility trend under the impact of market asymmetry. In fact, calculating the risk contagion strength of the co-skewness is also a measure of the degree to which the portfolio deviates from the normal distribution. According to the different that returns and squared returns when calculating the co-skewness statistics, the co-skewness risk contagion indicators is categorized into two types: CS_{12} and CS_{21} , where CS_{12} represents the contagion channel from carbon returns to infected markets variance, and CS_{21} represents the contagion relationship from carbon asset variance to infected markets returns. A statistically significant contagion coefficient indicates that there is a risk contagion relationship between carbon asset and infected markets with co-skewness attribute. Smaller contagion coefficients, indicating that distribution of the portfolio composition is basically close to standard normal distribution, with less exposure to asymmetric risk, and can achieve portfolio objective of risk sharing and return sharing. In constrast, higher risk contagion coefficients indicate that the asymmetric risk of the portfolio is high, making it difficult to achieve the portfolio investment objective.

The co-skewness channels risk contagion are given as the following equation:

$$CS_{12}(i \to j; r_i^1, r_j^2) = \left(\frac{\hat{\psi}_y(r_i^1, r_j^2) - \hat{\psi}_x(r_i^1, r_j^2)}{\sqrt{(4\hat{\upsilon}_{y/x_i}^2 + 2)/T_y + (4\hat{\rho}_x^2 + 2)/T_x}} \right)^2 \qquad (3.5)$$

$$CS_{21}(i \to j; r_i^2, r_j^1) = \left(\frac{\hat{\psi}_y(r_i^2, r_j^1) - \hat{\psi}_x(r_i^2, r_j^1)}{\sqrt{(4\hat{\upsilon}_{y/x_i}^2 + 2)/T_y + (4\hat{\rho}_x^2 + 2)/T_x}} \right)^2 \qquad (3.6)$$

And,

$$\hat{\psi}_y(r_i^1, r_j^2) = \frac{1}{T_y} \sum_{t=1}^{T_y} \left(\frac{y_{i,t} - \hat{\mu}_{yi}}{\hat{\sigma}_{yi}}\right)^1 \left(\frac{y_{j,t} - \hat{\mu}_{yj}}{\hat{\sigma}_{yi}}\right)^2 \qquad (3.7)$$

$$\hat{\psi}_y(r_i^2, r_j^1) = \frac{1}{T_y} \sum_{t=1}^{T_y} \left(\frac{y_{i,t} - \hat{\mu}_{yi}}{\hat{\sigma}_{yi}}\right)^2 \left(\frac{y_{j,t} - \hat{\mu}_{yj}}{\hat{\sigma}_{yi}}\right)^1 \qquad (3.8)$$

$$\hat{\psi}_x(r_i^1, r_j^2) = \frac{1}{T_x} \sum_{t=1}^{T_x} \left(\frac{x_{i,t} - \hat{\mu}_{xi}}{\hat{\sigma}_{xi}}\right)^1 \left(\frac{x_{j,t} - \hat{\mu}_{xj}}{\hat{\sigma}_{xi}}\right)^2 \qquad (3.9)$$

$$\hat{\psi}_x(r_i^2, r_j^1) = \frac{1}{T_x} \sum_{t=1}^{T_x} \left(\frac{x_{i,t} - \hat{\mu}_{xi}}{\hat{\sigma}_{xi}}\right)^2 \left(\frac{x_{j,t} - \hat{\mu}_{xj}}{\hat{\sigma}_{xi}}\right)^1 \qquad (3.10)$$

$$\hat{\upsilon}_{y/x_i} = \frac{\hat{\rho}_y}{\sqrt{1+(\frac{s_{y,i}^2 - s_{x,i}^2}{s_{x,i}^2})(1-\hat{\rho}_y^2)}} \quad (3.11)$$

In the above equation, *i* and *j* represent the risk sourced market and infected market, respectively; x and y represent volatility states of the carbon markets, respectively; T_x and T_y represent the market capacity under different market volatility states, respectively; $x_{i,t}$, $x_{j,t}$, $y_{i,t}$, and $y_{j,t}$ represent returns of the risk sourced market and infected market under the market states of x and y; $\hat{\mu}_{xi}$, $\hat{\mu}_{xj}$, $\hat{\mu}_{yi}$, and $\hat{\mu}_{yj}$ represent corresponding mean values of returns; $\hat{\sigma}_{xi}$, $\hat{\sigma}_{xj}$, $\hat{\sigma}_{yi}$, and $\hat{\sigma}_{yj}$ represent the standard deviation of returns; $\hat{\upsilon}_{y/x_i}$ represents the market correlation coefficient of the adjusted volatility state transformation; $\hat{\rho}_x$ and $\hat{\rho}_y$ represent the unconditional correlation between the two markets; and $s_{x,i}^2$ and $s_{y,i}^2$ represent the variance of the risk sourced market under different market volatility states.

To test whether risk contagion of co-skewness attribute occurs from carbon market to its infected markets, the original hypothesis that there is no high order moment risk contagion is assumed to be:

$$H(CS_{12})_0 : \hat{\psi}_y(r_i^1, r_j^2) = \hat{\psi}_x(r_i^1, r_j^2) \quad (3.12)$$

$$H(CS_{12})_1 : \hat{\psi}_y(r_i^1, r_j^2) \neq \hat{\psi}_x(r_i^1, r_j^2) \quad (3.13)$$

$$H(CS_{21})_0 : \hat{\psi}_y(r_i^2, r_j^1) = \hat{\psi}_x(r_i^2, r_j^1) \quad (3.14)$$

$$H(CS_{21})_1 : \widehat{\psi}_y(r_i^2, r_j^1) \neq \widehat{\psi}_x(r_i^2, r_j^1) \tag{3.15}$$

In the absence of the risk contagion channels of co-skewness, the contagion coefficients are verified for obeying the chi-square distribution, and the coefficients significance is used to determine whether co-skewness contagion relationship occurs under different market volatility trend:

$$CS_{12}, CS_{21}(i \to j) \xrightarrow{df} \chi_1^2 \tag{3.16}$$

where the Lagrange polynomials that validate CS_{12}, CS_{21} obey the chi-square distribution validation are denoted as:

$$LM(CS_{12}) = \frac{T}{4\hat{\rho}^2 + 2}\left[\frac{1}{T}\sum_{t=1}^{T}\left(\frac{r_{i,t} - \hat{\mu}_i}{\hat{\sigma}_i}\right)^1\left(\frac{r_{j,t} - \hat{\mu}_j}{\hat{\sigma}_j}\right)^2\right]^2 \tag{3.17}$$

$$LM(CS_{21}) = \frac{T}{4\hat{\rho}^2 + 2}\left[\frac{1}{T}\sum_{t=1}^{T}\left(\frac{r_{i,t} - \hat{\mu}_i}{\hat{\sigma}_i}\right)^2\left(\frac{r_{j,t} - \hat{\mu}_j}{\hat{\sigma}_j}\right)\right]^2 \tag{3.18}$$

3.4.2.2 CK Indicator for Co-Kurtosis Risk Contagion

The co-kurtosis attribute risk contagion is a measure of whether and to what extent the portfolio consisting of carbon market and its infected markets is affected by extreme event during the transformation process of carbon market volatility. It is to say, whether there is a obvious change in the high order moment co-kurtosis contagion coefficient before and after a change in the market volatility trend under extreme risk factor shocks. Like the co-skewness indicator, the co-kurtosis risk contagion tests is divided into two categories: CK_{13} and CK_{31} , where CK_{13} denotes the risk contagion of carbon returns on infected market skewness, and CK_{31} denotes the risk contagion of carbon market skewness on returns of the infected markets. If the contagion coefficient is statistically significant, it indicates that there is a risk contagion of co-kurtosis between the carbon market and the infected markets. The larger contagion coefficient, indicating that the portfolio returns are exposed to a larger external systematic risk shock, conversely, it indicates that the systematic risk is low.

The co-kurtosis channels risk contagion are given as the following equation:

$$CK_{13}(i \to j; r_i^1, r_j^3) = \left(\frac{\hat{\xi}_y(r_i^1, r_j^3) - \hat{\xi}_x(r_i^1, r_j^3)}{\sqrt{(18\hat{\upsilon}_{y/x_i}^2 + 6)/T_y + (18\hat{\rho}_x^2 + 2)/T_x}} \right)^2 \qquad (3.19)$$

$$CK_{31}(i \to j; r_i^3, r_j^1) = \left(\frac{\hat{\xi}_y(r_i^3, r_j^1) - \hat{\xi}_x(r_i^3, r_j^1)}{\sqrt{(18\hat{\upsilon}_{y/x_i}^2 + 6)/T_y + (18\hat{\rho}_x^2 + 2)/T_x}} \right)^2 \qquad (3.20)$$

Among them,

$$\hat{\xi}_y(r_i^1, r_j^3) = \frac{1}{T_y} \sum_{t=1}^{T_y} \left(\frac{y_{i,t} - \hat{\mu}_{yi}}{\hat{\sigma}_{yi}}\right)^1 \left(\frac{y_{j,t} - \hat{\mu}_{yj}}{\hat{\sigma}_{yi}}\right)^3 - (3\hat{\upsilon}_{y/x_i}) \qquad (3.21)$$

$$\hat{\xi}_y(r_i^3, r_j^1) = \frac{1}{T_y} \sum_{t=1}^{T_y} \left(\frac{y_{i,t} - \hat{\mu}_{yi}}{\hat{\sigma}_{yi}}\right)^3 \left(\frac{y_{j,t} - \hat{\mu}_{yj}}{\hat{\sigma}_{yi}}\right)^1 - (3\hat{\upsilon}_{y/x_i}) \qquad (3.22)$$

$$\hat{\xi}_x(r_i^1, r_j^3) = \frac{1}{T_x} \sum_{t=1}^{T_x} \left(\frac{x_{i,t} - \hat{\mu}_{xi}}{\hat{\sigma}_{xi}}\right)^1 \left(\frac{x_{j,t} - \hat{\mu}_{xj}}{\hat{\sigma}_{xi}}\right)^3 - (3\hat{\rho}_x) \qquad (3.23)$$

$$\hat{\xi}_x(r_i^3, r_j^1) = \frac{1}{T_x} \sum_{t=1}^{T_x} \left(\frac{x_{i,t} - \hat{\mu}_{xi}}{\hat{\sigma}_{xi}}\right)^3 \left(\frac{x_{j,t} - \hat{\mu}_{xj}}{\hat{\sigma}_{xi}}\right)^1 - (3\hat{\rho}_x) \qquad (3.24)$$

The relevant parameter definitions are consistent with the above. In order to test whether a risk contagion from sourced carbon market to its infected markets in the co-kurtosis channels, the original hypothesis that there is no high order moment risk contagion is assumed to be:

$$H(CK_{13})_0 : \hat{\xi}_y(r_i^1, r_j^3) = \hat{\xi}_x(r_i^1, r_j^3) \quad (3.25)$$

$$H(CK_{13})_1 : \hat{\xi}_y(r_i^1, r_j^3) \neq \hat{\xi}_x(r_i^1, r_j^3) \quad (3.26)$$

$$H(CK_{31})_0 : \hat{\xi}_y(r_i^3, r_j^1) = \hat{\xi}_x(r_i^3, r_j^1) \quad (3.27)$$

$$H(CK_{31})_1 : \hat{\xi}_y(r_i^3, r_j^1) \neq \hat{\xi}_x(r_i^3, r_j^1) \quad (3.28)$$

In absence of the risk contagion channels of co-kurtosis, the contagion coefficients are verified for obeying the chi-square distribution, and the coefficients significance is used to determine whether co-kurtosis contagion relationship occurs under different market volatility trend:

$$CK_{13}, CK_{31}(i \to j) \xrightarrow{df} \chi_1^2 \quad (3.29)$$

where the Lagrange polynomials that validate CK_{13}, CK_{31} obey the chi-square distribution validation are denoted as:

$$LM(CK_{13}) = \frac{1}{T(18\hat{\rho}^2 + 6)}\left[\sum_{t=1}^{T}\left(\frac{r_{i,t} - \hat{\mu}_i}{\hat{\sigma}_i}\right)^1\left(\frac{r_{j,t} - \hat{\mu}_j}{\hat{\sigma}_j}\right)^3 - T(3\hat{\rho})\right]^2 \quad (3.30)$$

$$LM(CK_{31}) = \frac{1}{T(18\hat{\rho}^2 + 6)}[\sum_{t=1}^{T}(\frac{r_{i,t} - \hat{\mu}_i}{\hat{\sigma}_i})^3(\frac{r_{j,t} - \hat{\mu}_j}{\hat{\sigma}_j}) - T(3\hat{\rho})]^2 \qquad (3.31)$$

3.4.2.3 CV Indicator for Co-Volatility Risk Contagion

The co-volatility risk contagion refers to the degree of contagion between the risk of carbon market and its infected market during transition of carbon market volatility. The co-volatility risk contagion occurs between the second-order moment variance risk of carbon market and infected markets. If contagion coefficient is statistically significant, it indicates the risk contagion phenomenon of co-volatility between carbon market and the infected markets. Among them, a smaller contagion coefficient indicates that carbon market risk has a weak impact on the market risk of its infected markets, and the portfolio can satisfy investors expectation. While a larger contagion coefficient indicates that the risk of the carbon market can easily trigger risk in the infected markets.

The co-volatility channels risk contagion are given as the following equation:

$$CV_{22}(i \to j; r_i^2, r_j^2) = \left(\frac{\hat{\varphi}_y(r_i^2, r_j^2) - \hat{\varphi}_x(r_i^2, r_j^2)}{\sqrt{(4\hat{\upsilon}_{y/x_i}^4 + 16\hat{\upsilon}_{y/x_i}^2 + 4)/T_y + (4\hat{\rho}_x^4 + 16\hat{\rho}_x^2 + 4)/T_x}} \right)^2 \qquad (3.32)$$

Among them,

$$\hat{\varphi}_y(r_i^2, r_j^2) = \frac{1}{T_y}\sum_{t=1}^{T_y}(\frac{y_{i,t} - \hat{\mu}_{yi}}{\hat{\sigma}_{yi}})^2(\frac{y_{j,t} - \hat{\mu}_{yj}}{\hat{\sigma}_{yi}})^2 - (1 + 2\hat{\upsilon}_{y/x_i}^2) \qquad (3.33)$$

$$\hat{\varphi}_x(r_i^2, r_j^2) = \frac{1}{T_x}\sum_{t=1}^{T_x}(\frac{x_{i,t} - \hat{\mu}_{xi}}{\hat{\sigma}_{xi}})^2(\frac{x_{j,t} - \hat{\mu}_{xj}}{\hat{\sigma}_{xi}})^2 - (1 + 2\hat{\rho}_x^2) \qquad (3.34)$$

In the risk contagion channels of co-volatility, the contagion coefficients are verified for obeying the chi-square distribution, and the coefficients significance is used to determine whether co-volatility contagion relationship occurs under different market volatility trend:

$$H(CV_{22})_0 : \hat{\varphi}_y(r_i^2, r_j^2) = \hat{\varphi}_x(r_i^2, r_j^2) \tag{3.35}$$

$$H(CV_{22})_1 : \hat{\varphi}_y(r_i^2, r_j^2) \neq \hat{\varphi}_x(r_i^2, r_j^2) \tag{3.36}$$

In the absence of risk contagion of co-volatility attributes, the contagion coefficients are verified for obeying the chi-square distribution, and the significance of the coefficients is used to determine whether co-volatility risk contagion relationship occurs under the different market volatility trend transitions:

$$CV_{22}(i \to j) \xrightarrow{df} \chi_1^2 \tag{3.37}$$

where the Lagrange polynomials that validate CV_{22} obey the chi-square distribution validation are denoted as:

$$LM(CV_{22}) = \frac{1}{T(4\hat{\rho}^4 + 16\hat{\rho}^2 + 6)}\left[\sum_{t=1}^{T}\left(\frac{r_{i,t} - \hat{\mu}_i}{\hat{\sigma}_i}\right)^2\left(\frac{r_{j,t} - \hat{\mu}_j}{\hat{\sigma}_j}\right)^2 - T(1 + 2\hat{\rho}^2)\right]^2 \tag{3.38}$$

3.4.2.4 FR Indicator for Low Order Moment Risk Contagion

The low order moment correlation coefficient risk contagion (Forbes Rigobon Contagion, FR) is a measure that whether there is a significant change in the cross-market correlation between carbon market and its infected markets base on the low order moment. If this indicator has statistical significance, it indicates a risk contagion relationship from carbon market to infected markets. The risk contagion indicator based on low order moment attributes is:

$$FR(i \to j) = \left(\frac{\hat{\upsilon}_{y/x_i} - \hat{\rho}_x}{\sqrt{Var(\hat{\upsilon}_{y/x_i} - \hat{\rho}_x)}} \right)^2 \tag{3.39}$$

$$\hat{\upsilon}_{y/x_i} = \frac{\hat{\rho}_y}{\sqrt{1 + (\frac{s_{y,i}^2 - s_{x,i}^2}{s_{x,i}^2})(1 - \hat{\rho}_y^2)}} \tag{3.40}$$

The relevant parameter definitions are consistent with the above. In order to test whether low order moment risk contagion occurs between carbon market and infected markets, the original hypothesis that there is no risk contagion of low order moments is assumed:

$$H(FR)_0 : \hat{\upsilon}_{y/x_i} = \hat{\rho}_x \tag{3.41}$$

$$H(FR)_1 : \hat{\upsilon}_{y/x_i} \neq \hat{\rho}_x \tag{3.42}$$

Without low order moment risk contagion, the contagion coefficients are verified for obeying the chi-square distribution, and the significance of the coefficients is used to determine whether risk contagion relationships occur under different market volatility trend transitions:

$$FR(i \to j) \xrightarrow{df} \chi_1^2 \tag{3.43}$$

3.5 HOC-LSTM Carbon Price Forecasting Model in China

According to results of identified risk contagion channels, a High Order risk Contagion LSTM model (HOC-LSTM) is constructed to forecast the carbon price.

3.5.1 Carbon Price Forecasting Framework Based on HOC

The CAPM pricing model with high order moment attributes points out that financial asset returns are not only affected by systematic risk, but also by investors limited rational behaviour and extreme event shocks. Therefore, compared to traditional low order moment price forecasting models, the CAPM pricing framework with high order moment can better capture the volatility characteristics of financial assets (Hwang et al., 1999). The CAPM pricing model for carbon market with high order moment attributes is defined as:

$$E(R_{carbon}) = \alpha\sigma^2_{carbon,m} + \beta S^3_{carbon,m} + \gamma K^4_{carbon,m} \tag{3.44}$$

Where,

$$\sigma_{carbon,m} = \frac{E[(r_{carbon,t} - E(r_{carbon,t})) \times (r_{m,t} - E(r_{m,t}))]}{E(r_{carbon}) \times E(r_m)} \tag{3.45}$$

$$S_{carbon,m} = \frac{E[(r_{carbon,t} - E(r_{carbon,t})) \times (r_{m,t} - E(r_{m,t}))^2]}{E(r_{carbon}) \times E(r_m)^2} \tag{3.46}$$

$$K_{carbon,m} = \frac{E[(r_{carbon,t} - E(r_{carbon,t})) \times (r_{m,t} - E(r_{m,t}))^3]}{E(r_{carbon}) \times E(r_m)^3} \tag{3.47}$$

$E(R_{carbon})$ represents the excess returns of carbon assets; $\sigma^2_{carbon,m}$ represents the covariance coefficient of carbon assets and market portfolio returns, that is, the impact of the first order central moments (returns) of the portfolio returns on the first order central moments (returns) of the carbon returns; $S_{carbon,m}^{3}$ represents the co-skewness coefficient of carbon assets and market portfolio, that is, the impact of the second order central moments (variances) of the market portfolio returns on the first order central moments (returns) of the carbon asset returns; $K_{carbon,m}^{4}$ represents the co-kurtosis coefficient of carbon asset and market portfolio, that is, the impact of third order central moment (skewness) of the market portfolio on the first order central moment (return) of the carbon return. α,β,γ represents the risk premium coefficient. Co-skewness reflects the asymmetric behaviour of carbon assets relative to market portfolio, negative co-skewness indicates the probability of carbon returns declining is higher than the probability of returns rising. Co-kurtosis reflects the the kurtosis of carbon assets relative to the market portfolio, a higher co-kurtosis indicate carbon returns are subject to extreme event shocks than the portfolio, implying that carbon assets have a higher order moment risk. In order to compensate for the holding losses, investors tend to demand higher premium returns for compensation.

3.5.2 Carbon Price Forecasting Model of HOC-LSTM

The multi-factor CAPM pricing framework of carbon market has significant nonlinear characteristics, namely the difference in the order moment attribute dimension of data information, resulting in a complex nonlinear structure between carbon price and its infected markets. Furthermore, the proposed pricing framework has many parameters need to be estimated. Based on these two features, this thesis adopts a multi-layer multivariate long and short-term memory (LSTM) model based on machine learning algorithms to forecast carbon prices with risk contagion relationships.

3.5.2.1 Advantages of the LSTM Model

LSTM model is a machine learning method to address issues of gradient explosion and vanishing, as well as insufficient long memory ability in traditional recurrent neural networks according to the previous studies that suggested in above literature review section, such as Hochreiter et al.(1997), Sun et al.(2020), Huang et al.(2021),Kong et al.(2022),Yun et al.(2023). The neural network structure of LSTM model with special gate structure stemmed from the optimization and updating of traditional recurrent neural networks.

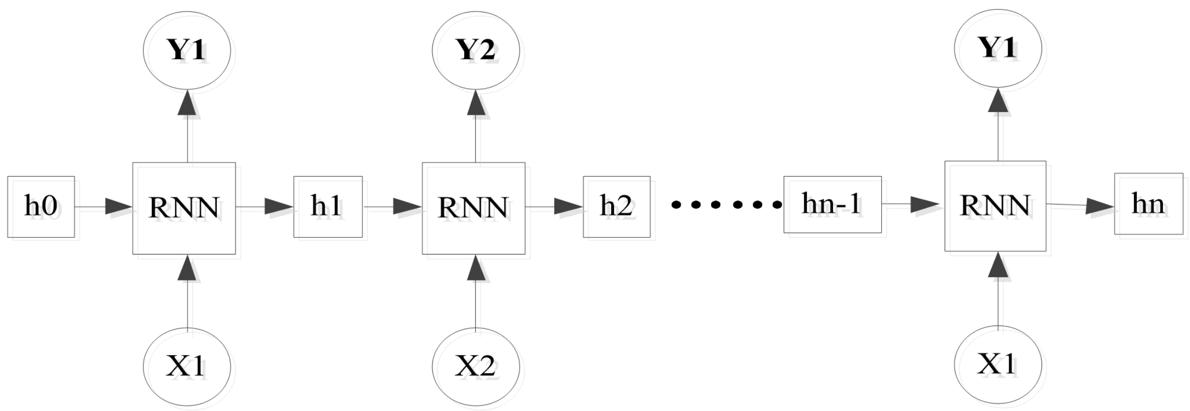


Figure 3.2: The Structure of RNN

Recurrent Neural Network (RNN) is a kind of chained network with special memory ability, the current output of the model structure is related to the previous output. That is, the network structure remembers front-end information of the sequence and applies it to the current output calculation. In terms of model structure, the nodes between hidden layers are also no longer unconnected, but are linked. The input to a hidden layer now includes output from input layer, also includes the output from the previous moment's hidden layer. The model structure is shown in Figure 3.2. Figure 3.2 shows a Recurrent Neural Network (RNN) diagram. It represents the flow of information in an RNN, showing how hidden states (h_t) are passed through time steps, processing sequential inputs (X_t)

and generating corresponding outputs (X_t). The function of each RNN cell is to process the input at that time step and updates its hidden state. Compared to classical BP algorithm and CNN, the biggest advantage of RNN is to achieve memory function of the input information. However, when the interval between the information of the previous moment of the financial time series and the current forecasting position is long, the algorithm may cause the gradient disappearance, thereby making it difficult for the RNN to learn the features of long-term information. Based on this, during the training process of LSTM, in addition to continue to pass the hidden layer information backward, but also through the special design of the Cell structure to transmit past longer period information, so as to effective control the gradient.

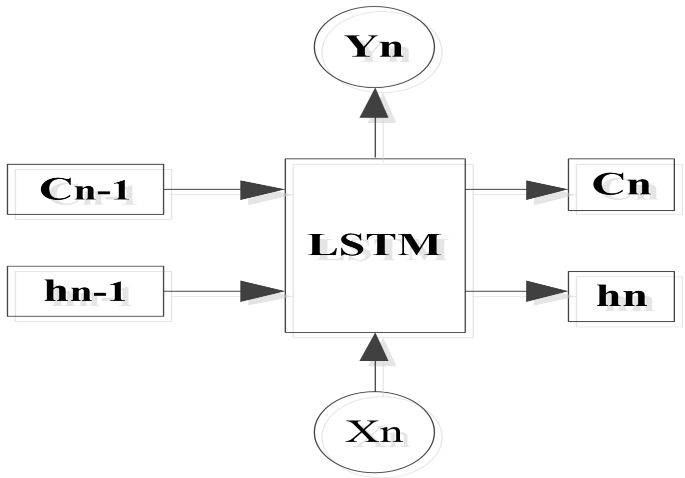


Figure 3.3: The Structure of LSTM

The LSTM model structure is shown in Figure 3.3,where.

Xn=The input to the LSTM cell at the current time step(n);

hn-1=The hidden state from the previous time step(n-1);

LSTM Cell=The core computational unit that updates both the cell and hidden state;

Cn=The updated cell state, which carries long-term information;

hn=The updated hidden state, which is used for generating outputs passed to the next time step;

Yn=The output generated at this time step(n).

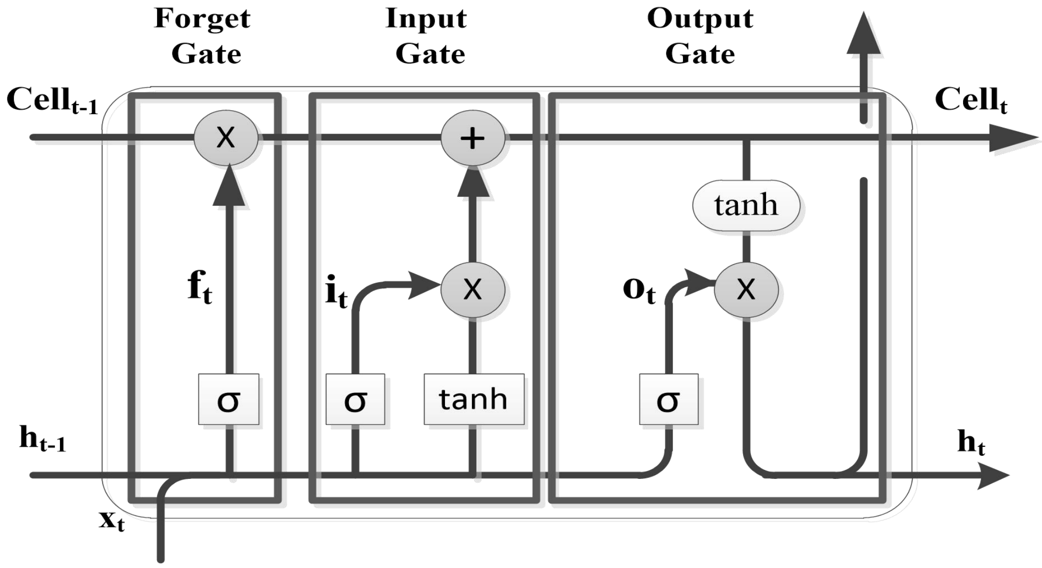
LSTMs improve upon standard RNNs by introducing a cell state (cn) and three gating mechanisms that regulate the flow of information: First, the Forget Gate, which decides what information should be discarded from the previous cell state. Second, the Input Gate which determines what new information should be added to the cell state. And lastly, the Output Gate, which controls what information from the cell state should be output at the current time step. This gating mechanism helps prevent the vanishing gradient problem, allowing LSTMs to retain information overlong sequences.

According to advantages of LSTM network structure in parameter learning, this thesis uses it as an empirical fitting and forecasting method to study the carbon price forecasting framework. Firstly, during processing financial time series, especially time series data with long memory characteristics, the LSTM model can maximize the capture of features and information from longer time dimensions into the current structure, addressing the issue of long-term dependencies in data. Secondly, as a type of recurrent neural network with a specialized memory function, the LSTM model optimizes and adjusts the parameter structure by performing forward unsupervised learning and backward supervised learning process, namely the self-learning and self-adaptation process of the parameter structure, so as to train the optimal model structure and solve the non-linear problem of the carbon price forecasting framework. Thirdly, due to the large number of parameter structure in the forecasting model, especially during the training process of neural networks, model gradient may increase as the increasing of the network layers, the neural network structure has the possibility of gradient explosion or gradient disappearance, resulting in interruption of the self-learning process. While the LSTM model is designed through a special threshold structure and activation function, provides a guarantee for the training of neurons and parameter weights, and can solves the interruption of model training and weight update , and ensures the effective learning, training and convergence of the carbon price forecasting framework.

Although LSTM models have the obvious advantages mentioned above, they also have some drawbacks: Firstly, the model training time is relatively long, which is related to its relatively complex gate structure and more participation. Secondly, there is a risk of overfitting, where the LSTM model, due to its strong memory ability, may remember the noise in the training data during the training process, leading to overfitting. Thirdly, the problem of gradient vanishing has not been fully resolved. Although LSTM alleviates the gradient vanishing problem through gating mechanisms, in some cases, especially when dealing with very long sequences, gradients may still disappear. This requires more complex optimization algorithms and techniques to solve.

3.5.2.2 Training Process of the LSTM Model

As a special neural network structure, LSTM model consists of an input layer, output layer and hidden layer. As for forecasting carbon price with the high order moment risk contagion, the input layer data mainly refers to the high order moments infected markets that recognized by the above models. The output layer mainly refers to the carbon returns, which is also a label item for supervised learning in the process of parameter training and model optimization. On the one hand, the hidden layer includes the features and weights of the input layer learned by the network structure, that is, the short-term memory capability of the model. On the other hand, it also includes the special cell structure that can remember the long-term carbon price infected markets. Among them, the long memory function of the cell structure is primarily achieved through specially designed gate mechanisms, specifically forget gates, input gates, and output gates.



Note: ⊗ represents the Kronecker product, which is an operation between two matrices, and ⊕ represents the logical operation of yes or no.

Figure 3.4 : The Training Structure of LSTM Model

In the LSTM training process (as shown in Figure 3.4), the forget gate first determines which parts of the previous network output should be forgotten. It maps the current input x_t and the hidden layer h_{t-1} to a value between 0 (forget everything) and 1(keep everything) by sigmoid activation function(σ), thus obtaining forget gate output f_t. Secondly, the input gate determines which information should be added to the current memory unit $Cell_t$. It uses sigmoid activation function to obtain the preservation of original input x_t and previous moment of hidden layer h_{t-1} in structure of this layer i_t . And then, the output of this layer(o_t) is obtained through the tanh function $\tilde{C}_t$, and $i_t \times \tilde{C}_t$ represent the preservation characteristics of the information in this layer. Combined with the return information of the forgetting gate, the characteristics of the input information are summarized. Thirdly, the output gate is used to decide how many neuron Cells are filtered, that is, a sigmoid activation function is used to obtain a value in the interval of 0 and 1, and then the current memory cell, $cell_t$ is processed through the tanh function to obtain the output of this layer h_t. The multi-layer LSTM model is a network stack of LSTM model by adding hidden layers and memory cell neurons on top of a single layer.

The output of forget gate (f_t) is determined by screening the original infected markets including high order moment term and the hidden layer features. The training structure of forget gate is as follows:

$$f_t = \sigma(W_f \times [h_{t-1}, x_t] + b_f) \tag{3.48}$$

The output of input gate is determined by saving and updating the input data. The training structure of input gate is as follows:

$$i_t = \sigma(W_i \times [h_{t-1}, x_t] + b_i) \tag{3.49}$$

$$\widetilde{C}_t = tanh(W_C \times [h_{t-1}, x_t] + b_C) \tag{3.50}$$

$$C_t = f_t \times C_{t-1} + i_t \times \widetilde{C}_t \tag{3.51}$$

The data filtering output of the output gate (o_t) under the current memory neural unit is as follows:

$$o_t = \sigma(W_o \times [h_{t-1}, x_t] + b_o) \tag{3.52}$$

$$h_t = o_t \times tanh(C_t) \tag{3.53}$$

Where the above formula weight functions need to be calculated separately during the learning process. $W_f = W_{fx} + W_{fh}$, $W_i = W_{ix} + W_{ih}$, $W_c = W_{cx} + W_{ch}$, $W_o = W_{ox} + W_{oh}$. The input of the output layer is $Y_t^i = W_{yi} h_t$ and the output is $Y_t^o = \sigma(Y_t^i)$. i_t , $\tilde{C}_t$, and $\tilde{C}_t$ represent the information update vector of the input gates, the candidate vector, and the output gate update vector of current state respectively. h_t represents the final hidden layer output of LSTM model; W_f , W_i , W_C and W_o denote weight vectors; b_f , b_i , b_C and b_o represent the bias of the training process; σ represents the sigmoid activation function.

This study utilized LSTM model to estimate the cross market risk contagion relationship between carbon market and its infected markets, such as energy markets, and capital. In this respect, the Higher-Order Coupled Long Short-Term Memory (HOC-LSTM) is an advanced variation of LSTM designed to improve sequence modeling by incorporating higher-order interactions between different elements in the LSTM cell. It enhances the learning capability of standard LSTMs by capturing complex dependencies in sequential data. The characteristics of HOC-LSTM model and its comparative models are summarized in Table 3.2. The comparative models are described in Section 3.6.

3.6 Select of the Comparative Models

To compare the superiority of the proposed HOC-LSTM model in fitting and forecasting China carbon price, the following commonly used machine learning classifiers were selected as the comparative benchmarks. During the experiment, the neural network structure and some parameters were adopted above design. As shown in Table 3.2.

Table 3.2: The characteristics of HOC-LSTM model and its comparative models

Models	Characteristics	Advantages	Disadvantages
GRU (Gated Recurrent Unit)	1. Combines input and forget gates into a single update gate 2. Uses a reset gate for short-term dependencies. 3. Eliminates cell units	1.Simpler structure 2.Fewer parameters, faster training 3. Suitable for long time series data	May not capture some complex long-term dependencies as effectively as LSTM
MLP (Multi-Layer Perceptron)	1. Each layer performs a nonlinear transformation through an activation function. 2. Trained using the backpropagation algorithm	1.Strong nonlinear modeling ability 2. Excellent regression and prediction capability 3.Flexible for various types of data	1.Non-convex loss function can lead to different results 2. Slow training, prone to local minima 3.Requires setting many parameters.
GBDT (Gradient Boosting Decision Tree)	1.Trains weak classifiers based on the negative gradient of the current loss function 2. Combines weak classifiers into a stronger model	1. Strong predictive performance 2. Handles nonlinear problems well	1.Training can be time-consuming 2.Sensitive to noise, requiring careful tuning
ETR (Extra Trees Regressor)	1.Does not use random sampling for training data, instead uses all available data 2.Randomly selects features for building decision trees	1.Faster training time compared to RF 2.Excellent regression results	1.Sensitive to outliers in some cases 2.Requires careful tuning of parameters
BPNN (Back Propagation Neural Network)	1.Optimizes network structure and parameters through supervised and unsupervised learning 2.Strong ability to model nonlinear relationships	1. Strong nonlinear mapping ability 2.Flexible network structure 3.Excellent regression ability	1.Slow learning speed 2.Prone to getting stuck in local minima 3. Requires manual selection of network layers and neurons
HOC-LSTM	1.Uses gates to optimize memory retention 2.Effective for handling long-term dependencies in sequential data	1.Can capture long-term dependencies effectively 2. Ideal for time series.	1.Complex model structure 2.Large number of parameters

3.6.1 Gated Recurrent Unit Neural Network

Gated recurrent unit (GRU) neural network is another improved structure of RNN model, which can better capture longer memory characteristic of the time series data. The construction of this model is used to solve issues of insufficient long memory and gradient explosion in the process of back propagation training (Cho et al.,2014; Zhu et al.,2023). Compared to the special input gate, forget gate, and output gate structures of LSTM model as suggested in previous research of Hochreiter et al.(1997) and Yun et al.(2023), the GRU model further integrates gate structure of LSTM to make the model structure more concise. That is, the input gate and forget gate are combined into a new update gate, which determines which information should be discarded and added. The reset gate, on the other hand, decides the amount of information to forget from past time series, aiding in capturing short-term dependencies in the input time series.

Different from LSTM models that rely solely on cell units to obtain long term data memory characteristic, GRU model discards the cell units and uses the hidden layers to transmit information. Generally, the GRU model is a effective variant of LSTM networks, with a simpler structure and fewer parameter and sample requirements, and also has the advantage of faster training and fitting performance. So, it can solve the problem of long dependencies in traditional RNN networks.

3.6.2 Multi-Layer Perceptron Neural Network Model

Essentially, the Multi-layer perceptron (MLP) neural network model is a typical of artificial neural network (ANN), which consists of a feed-forward neural network with one or more hidden layers (Fan et al.,2015; Zhu et al.,2024). As for the MLP network, each layer undergoes nonlinear transformation through an activation function apart from the input layer.

As a neural network structure based on back-propagation algorithm, the MLP structure includes: input layer, hidden layer and output layer. Among them, input layer neurons are used to receive the high order moment risk contagion price information , the hidden layer and output layer contain functional neurons that can compare the received price with a certain threshold, and then process it through an activation function to generate the output of the neurons. According to Fan et al.(2015), the training method of MLP is the back-propagation algorithm, which has the greatest advantage of improving the network data learning ability and non-linear price forecasting ability. However, the hidden layer of the MLP model has a non convex loss function, and different random initial weights may lead to different errors. Additionally, the MLP model requires preset a series of initial parameters, such as the number of hidden neurons, hidden layers, and the number of iterations.

3.6.3 Gradient Boosting Decision Tree Model

Gradient boosting decision tree model (GBDT) is a type of ensemble algorithm, whose basic classifier is the classification and regression tree (CART), and the ensemble method is gradient boosting based on the previous studies of Sun et al.(2020). Its idea is inspired by the gradient descent method, which trains a newly added weak classifier based on the negative gradient information of the current loss function, and then combines the trained weak classifier with the existing model to obtain a new ensemble model.

The gradient boosting algorithm that uses decision trees as the weak classifiers is called GBDT, sometimes also known as multiple additive regression tree (MART).

Actually, the gradient boosting and random forest (RF) are both belong to ensemble algorithms. The difference is that the random forest algorithm improves the forecasting performance by using a large number of trees in parallel each time, while when the size of the trees reaches a certain level, the performance cannot be further improved. However, the gradient boosting algorithm assigns a weight value to each classification result in a serial manner, and finally obtains the final result through accumulation to achieve satisfactory results.

3.6.4 Extra Trees Regressor Model

Extra trees regressor (ETR) model is a regressor that integrates multiple decision trees. Different from the idea of the traditional random forest (RF) model, the ETR model has two main advantages, firstly, as for each decision tree's training set, the RF model uses random sampling bootstrap to select the sampling set as the training data, while ETR model generally does not use the random sampling method. Secondary, as for model ensemble idea, the RF model applies the Bagging model to obtain the sampling data, while the ETR model uses all samples with randomly selected characteristic (Yahşi et al.,2019; Yun et al.,2020). Because the data sampling are random, the ETR model has better regression results and better generalization ability than the RF model to some extent. Therefore, ETR model is better than RF model about forecasting effect and training cost.

3.6.5 Back Propagation Neural Network

The back propagation neural network (BPNN) model is a classic algorithm for learning and training multi-layer neural networks, achieving parameter optimization and structural adjustment of network models, and improving the regression mapping ability. Its essence is to establish the input-output relationship between price data and its influencing factors data, and then extract the structural price data characteristic by the positive supervised learning and negative unsupervised learning processes according to the previous studies that listed in Literature review section,such as the research of Shen et al.,2015; Zhang et al.,2020; Yun et al.,2023. During the training process, gradient boosting algorithm is used to optimize model parameter, map and fit input and output data, and enhance effect of price forecasting. The core of model is using error function to calculate gradient of each training parameter, and conducting gradient to optimize the parameters and reduce the losses.

The significant advantages of BPNN model are the nonlinear mapping ability and flexible network structure. As for the network structure, the quantity of hidden layers and neurons can be set according to characteristics of input data, as a result, the forecasting ability of the BPNN model is excellent. However, in some applications of the price forecasting, the BPNN model also has theoretical shortcomings such as slow learning speed, easy falling into local minima, and subjective selection of network layers and neurons nodes.

3.7 Evaluation Criteria of the Proposed Forecasting Model

In order to evaluate the performance of the carbon price forecasting framework, highlight the rationality of considering the risk contagion relationship of the high order moment attribute in theory framework, this study adopts the following evaluation indicators to measure and analyse the performance of regression classifier.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{T}(y_i - \hat{y}_i)^2}{T}} \qquad (3.54)$$

$$MAE = \frac{1}{T}\sum_{i=1}^{T}\left|y_i - \hat{y}_i\right| \quad (3.55)$$

$$MAPE = \frac{1}{T}\sum_{i=1}^{T}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \quad (3.56)$$

The Kendall correlation coefficient is define as:

$$Kendall = \frac{2P}{0.5n(n-1)} - 1 \quad (3.57)$$

where n is the number of items, and P denotes the sum of the number of items ranked after the given item by both rankings.The correlation coefficient value is from 0 to 1. The closer the coefficient is to 1, the stronger the correlation is.

In the above evaluation criteria, the Root-Mean-Square Error (RMSE) measures the deviation between true returns and predicted returns. Mean Absolute Error (MAE) measures square of difference between true and predicted returns, and then sums and averages it, it is often used to evaluate the loss function in linear regression analysis. Mean Absolute Percentage Error (MAPE) evaluates the extent to which the predicted returns deviate from the true returns, and is often used to judge the stability and accuracy of regression classifiers. The above three indicators are more commonly used to evaluate the strength of the model. The larger the value indicates that the forecast and actual returns deviation is larger, suggesting worse model performance. As for correlation coefficients, the Kendall correlation is suitable for measuring the relationship between two sequential variables, when the data does not follow a normal distribution, the Kendall correlation coefficient is more accurate than the Pearson correlation.

3.8 Summary of the Chapter

This chapter describes the research framework, data collection, methods constructed and used of the study. The target of this study is to construct a model for risk contagion and price forecasting in China carbon market based on high order moment attribute. Time series data is used in the study, the daily transaction price of Hubei carbon market is selected for China carbon price data, this study selects the product instruments of carbon finance homogeneous market, capital market and energy market as carbon pricing factors. The study period covers 10 years, which is from the year 2014 to 2024. Mainly studying the risk contagion relationship from sourced carbon market to its infected markets, as well as forecasting carbon price under the impact of high order moment risk contagion. This thesis adopts a machine learning LSTM to forecast carbon asset prices with risk contagion relationships.

CHAPTER 4

RESULTS AND DISCUSSIONS

4.1 Introduction

The chapter is an empirical analysis based on high order moment risk contagion model and price forecasting model designed in Chapter 3. Firstly, this thesis conducts the basic descriptive statistical analysis on all the sample and variables to grasp the basic price characteristics of carbon market and its infected markets. Secondly, measures risk contagion relationship between carbon market and infected markets, especially high order risk contagion channels that represent the effects of asymmetric information and extreme factors on carbon prices. And also compare differences with the type of low order moment risk contagion. Thirdly, incorporates the identified high order moment risk contagion factors into the HOC-LSTM model to improve the carbon price forecasting accuracy and robustness.

4.2 Descriptive Statistical Analysis and Data Preprocessing

Descriptive statistics can help us understand the basic characteristics of research data and provide a foundation for subsequent data analysis and model construction. The collection range of all sample data in this thesis is from April 28, 2014 to January 24, 2024. After excluding the inconsistent data, a total of 2337 price data were obtained, the price trend of China carbon market and its infected markets can be shown in Figure 4.1. Furthermore, 2336 returns data can be obtained by using the operation of first-order difference, the return trend of China carbon market and its infected markets can be shown in Figure 4.2.

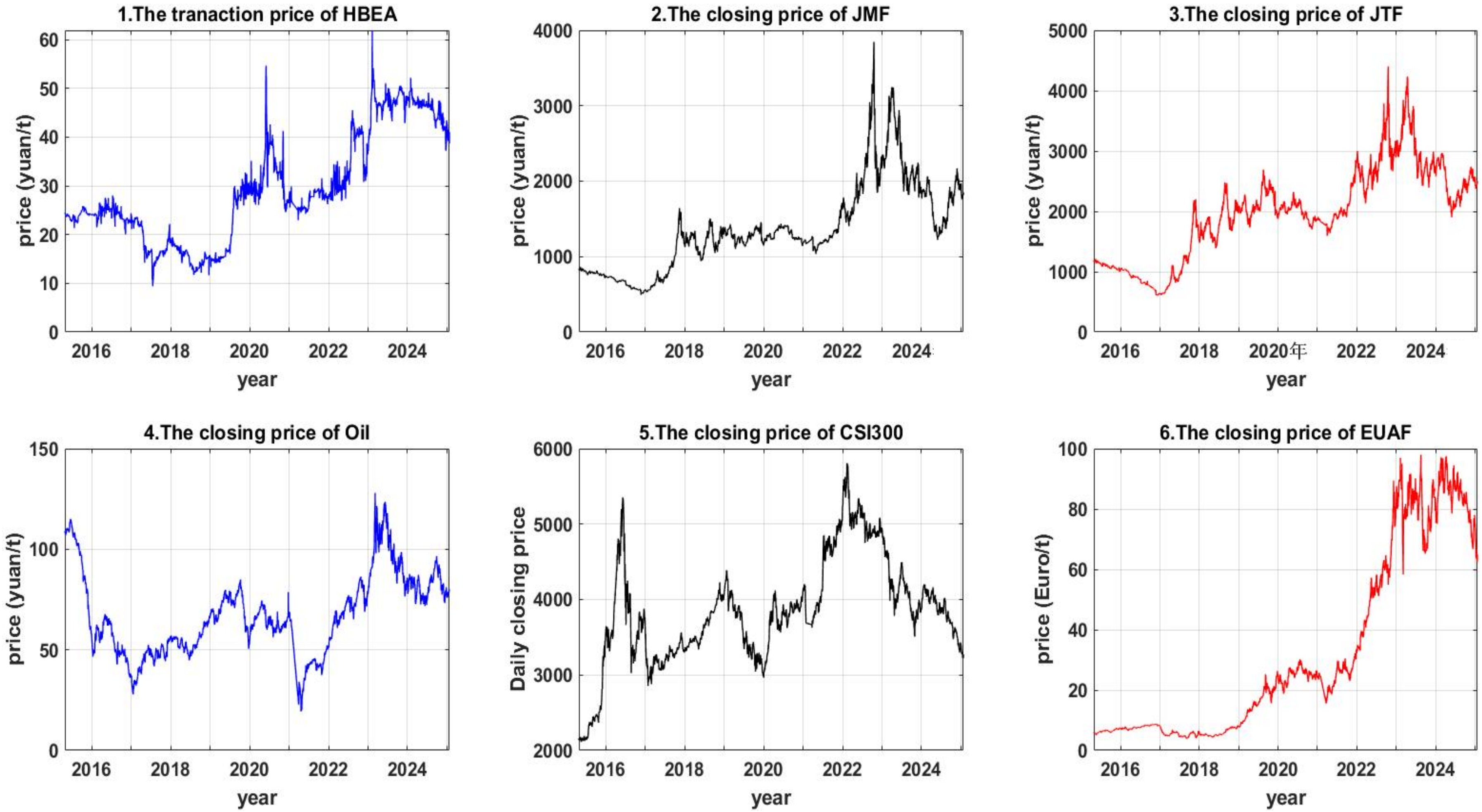


Figure 4.1: The Price Trend of Carbon Market and its Infected Markets

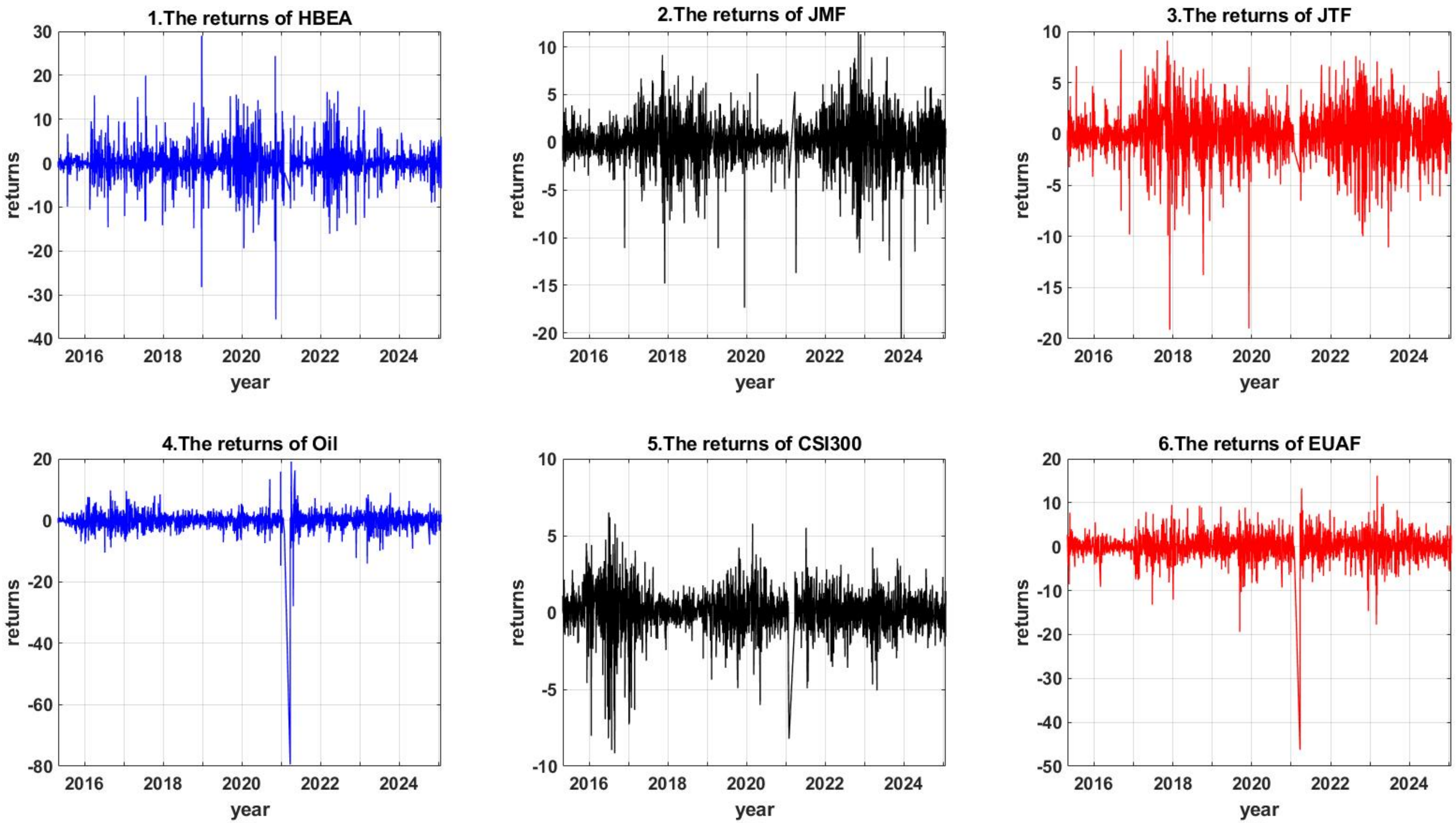


Figure 4.2: The Return Volatility of Carbon Price and its Infected Markets

According to the basic descriptive statistics results in Table 4.1, this thesis can summarize the following findings: Firstly, in terms of average returns, the European carbon price EUAF has the highest return of 0.104, while the China carbon price HBEA has a relatively low return of 0.021, which is only one-fifth of the European carbon price. The market return of crude oil is the lowest of all the sample, that the value is -0.013.

Table 4.1: Basic Statistical of China Carbon Price and its Infected Markets

	Mean	Std.Dev.	Skewness	Kurtosis	ADF	JB-Stat	ZBDS$_{(10)}$	Obs.
HBEA	0.021	3.777	-0.208	13.110	40.615	9965.856***	40.563***	2336
JMF	0.034	2.438	-0.808	9.691	49.371	4611.777***	29.782***	2336
JTF	0.032	2.336	-0.833	8.975	50.609	3745.413***	23.217***	2336
Oil	-0.013	3.016	-8.035	213.798	47.695	4350.220***	40.407***	2336
CSI300	0.018	1.389	-0.842	9.4333	46.564	4304.193***	22.823***	2336
EUAF	0.104	2.857	-2.109	35.486	49.242	1044.515***	21.867***	2336

Note: * * * indicates the significance under the 1% level.

Secondly, in terms of standard deviation, the standard deviation of China carbon price is the highest, that the value is 3.777, while the standard deviation of CSI300, which reflects the macroeconomics, is relatively the lowest. So, this thesis believe that the high market variance and low market returns of China carbon price indicate a significant return risk and market uncertainty. Generally, market risk is a compensation for returns, higher risks usually correspond to higher returns. After comparing the return and standard deviation of all the sample, a surprising finding is that the carbon prices do not seem to conform to this conclusion, which may indicate irrational trading behavior in carbon market, thesis findings are completely consistent with the research of Yun et al (2020).

Thirdly, regarding skewness and kurtosis, the market skewness of all samples is negative, and the kurtosis is significantly greater than the critical value of 3. In particular, the skewness and kurtosis of China carbon price are -0.208 and 13.110, respectively. This finding indicates that the price of China market and infected markets have obvious characteristics of sharp peaks and thick tails and fluctuating clusters, which means the extreme external factors with low occurrence probability can easily trigger the price changes in carbon market. According to Yun et al.(2020), a negative skewness indicates a leftward bias in the distribution of market return, that means a significant outlier on left side of distribution. A higher kurtosis is usually related to the low-frequency outliers in the return series. If the variance fluctuation is largely caused by the outliers, then the probability of return series having a higher kurtosis is also high.

Fourthly, in terms of the stationarity test, the unit root Augmented Dickey-Fuller (ADF) test results of China carbon market and its infected markets are not significant, meaning that all the sample return sequences accept the null hypothesis of the existence of unit roots. So price of China carbon market and its infected market are non-stationary series. This conclusion is completely consistent with the majority of research that listed in above literature part, that is carbon price series are basically non-stationary.

Fifthly, in terms of the normality distribution test, the results show that the Jarque-Bera (JB) statistic of China market and its infected market are significant at 1% level, so this thesis should reject the null hypothesis of normal return distribution, as a result, all sample series are non normality. In addition, according to the Quantile Quantile plots of each series (as shown in Figure 4.3), it can also be found that there is a significant difference between the sample percentile and the theoretical percentile, indicating that there is a systematic bias in the return distribution of each series, and show the characteristic of the non-normality.

Finally, in terms of the non-linear test, when the embedding correlation dimension is 10, the Brock Dechert Sheinkman (BDS) statistical results of all series show significance at 1% level, indicating that the return series of China market and its infected markets are non-linear.

In summary, this thesis concludes that price of China carbon market and its infected markets have features of sharp peaks and thick tails, fluctuation clustering, non normal, nonlinear and non-stationary.

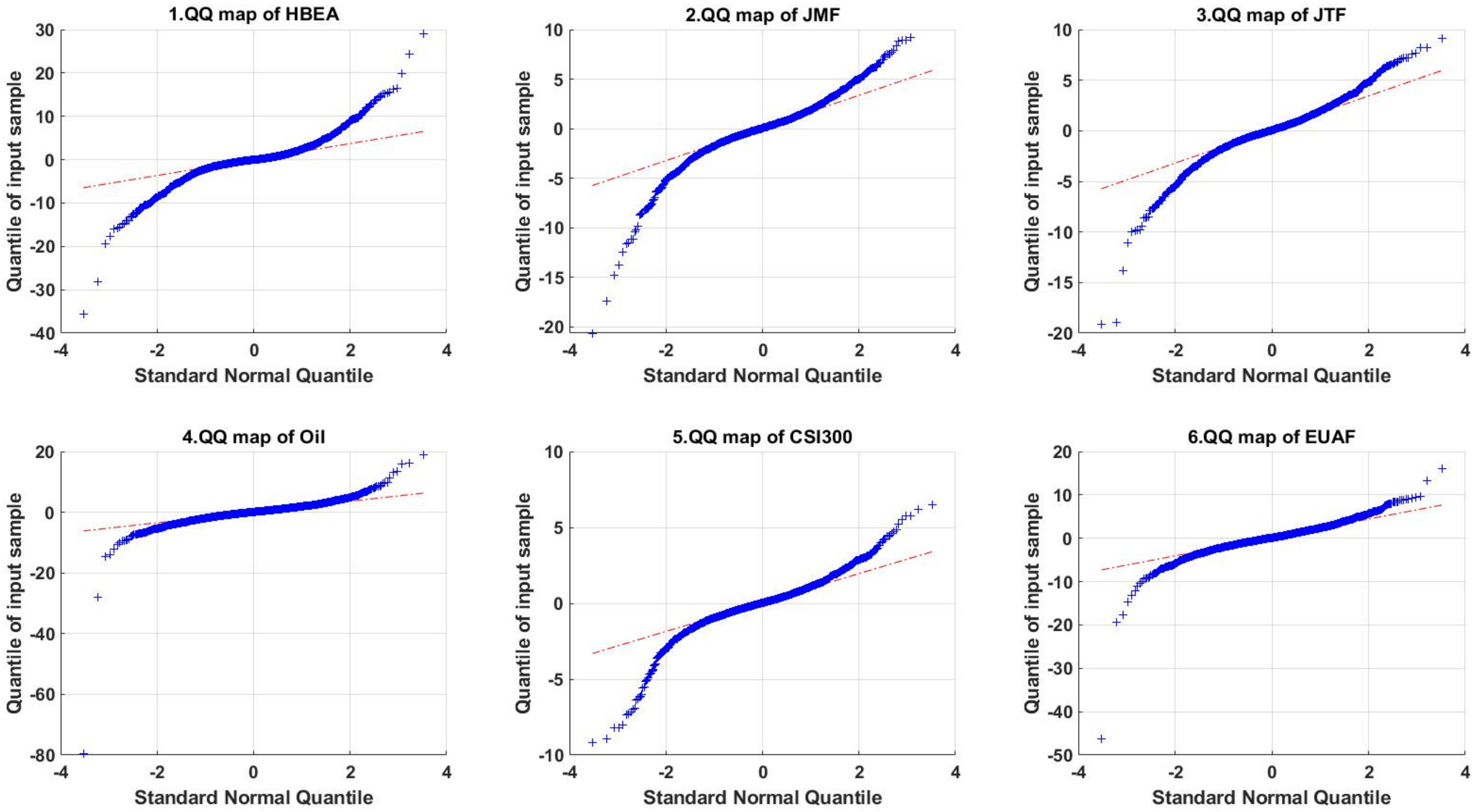


Figure 4.3: Q-Q Distribution of Carbon Price and its Infected Market Prices

4.3 Test High Order Moment Risk Contagion of China Carbon Market

4.3.1 Price Trend of China Carbon Market

To explore high order moment risk contagion relationship between China carbon market and its infected markets under different volatility states, this thesis uses the Markov state transition model to divide the market volatility states and design different market trends.

4.3.1.1 Select the State Transition Model

This thesis uses the experimental comparison method to determine the suitable volatility state transformation model. That is, referring to the experience of Hamilton (1989), this thesis test the data fitting ability of each Markov state model when the state number is 3 and 4, and the autoregressive lag order is 1, 2, 3 and 4. The reason does not select the state of 1 and 2 is that, the Markov mechanism transition model requires the states number to be greater than 1, while it is obvious that 2 states cannot fully characterize the state volatility division of the carbon market and cannot obtain more accurate risk contagion characterization. Therefore, this thesis only measures cases where the volatility states are 3 and 4.

As can be seen from Table 4.2, when the number of Markov state is 3, the autoregressive lag order is 1, the tail residual follows a normal distribution assumption, and the number of model parameters is 18, the value of akaike information criterion (AIC), bayesian information criterion (BIC) and Hannan-Quinn Criterion (HQ) are 18.7836, 122.4028, and 19.6575, respectively, which are the smallest values among all the data. This indicates that the MS (3) - AR (1) - N model can fit China carbon price data well, and is suitable for identifying the fluctuation status and features of China carbon price. This indicates that Markov state transition model of MS (3) - AR (1) - N can relative accurately fit China carbon price data, and is suitable for identifying the volatility status and market trend of China carbon market.

Table 4.2: Parameter Estimation Comparison of Different Markov State Models

Transition Models	**Residual Distribution**	**Number of Parameters**	**Likelihood**	**AIC**	**BIC**	**HQ**
MS(3)-AR(1)	T	21	-5488.0616	24.7793	145.6684	25.7988
	N	**18**	**-5476.3631**	**18.7836**	**122.4028**	**19.6575**
MS(3)-AR(2)	T	24	-5473.0071	30.7848	168.9438	31.9500
	N	21	-5469.8959	24.7860	145.6751	25.8055
MS(3)-AR(3)	T	27	-5439.9652	36.7969	192.2258	38.1077
	N	24	-5463.1481	30.7884	168.9474	31.9536
MS(3)-AR(4)	T	30	-5435.591	42.7986	215.4973	44.2550
	N	27	-5460.1448	36.7895	192.2184	38.1003
MS(4)-AR(1)	N	28	-5435.1261	38.7987	199.9842	40.1580
	T	32	-5448.7731	46.7937	231.0057	48.3472
MS(4)-AR(2)	T	36	-5451.6946	54.7926	262.0311	56.5403
	N	32	-5450.1248	46.7932	231.0052	48.3467
MS(4)-AR(3)	T	40	-5401.5329	62.8111	293.0761	64.7530
	N	36	-5438.9415	54.7973	262.0358	56.5450
MS(4)-AR(4)	T	44	-5399.5888	70.8118	324.1033	72.9479
	N	40	-5424.9103	62.8025	293.0674	64.7444

4.3.1.2 Recognize Market State and Price Trend

According to the selected model above, this thesis continues to use the MS (3) - AR (1) - N model to recognize the market volatility state of China carbon market. As a result, three market states with average standard deviations of 2.9346%, 8.1299% and 0.8410% can be obtained (as shown in Table 4.3). By comparing the frequency differences of the three states, this study defines them as low volatility state (State 1), high volatility state (State 2) and stable state (State 3). The transition probability changes of the three states throughout the entire sample period are shown in Figure 4.4.

Table 4.3: State Classification of Carbon Price Based on the MS (3)-AR (1)

State	Average Volatility (%)	State Classification	Transition Probabilities	State Duration	Obs.
State1	2.9346***	Low volatility	0.85	6.89	1101
State2	8.1299***	High volatility	0.82	5.56	264
State3	0.8410***	Stability	0.88	8.13	971

Note: *** means the significant in the level of 1%.

Specifically, the average volatility of the low volatility state (State 1) is 2.9346%, with a transition probability of 0.85 and a state duration period of 6.89, which is total of 1101 sample points. The average volatility of high volatility state (State 2) is 8.1299%, with a transition probability of 0.82 and a duration period of 5.56, and a total of 264 sample points were obtained. The average volatility of stable state (State 3) is 0.8410%, with a transition probability of 0.88 and a duration period of 8.13, and a total of 971 sample points were obtained.

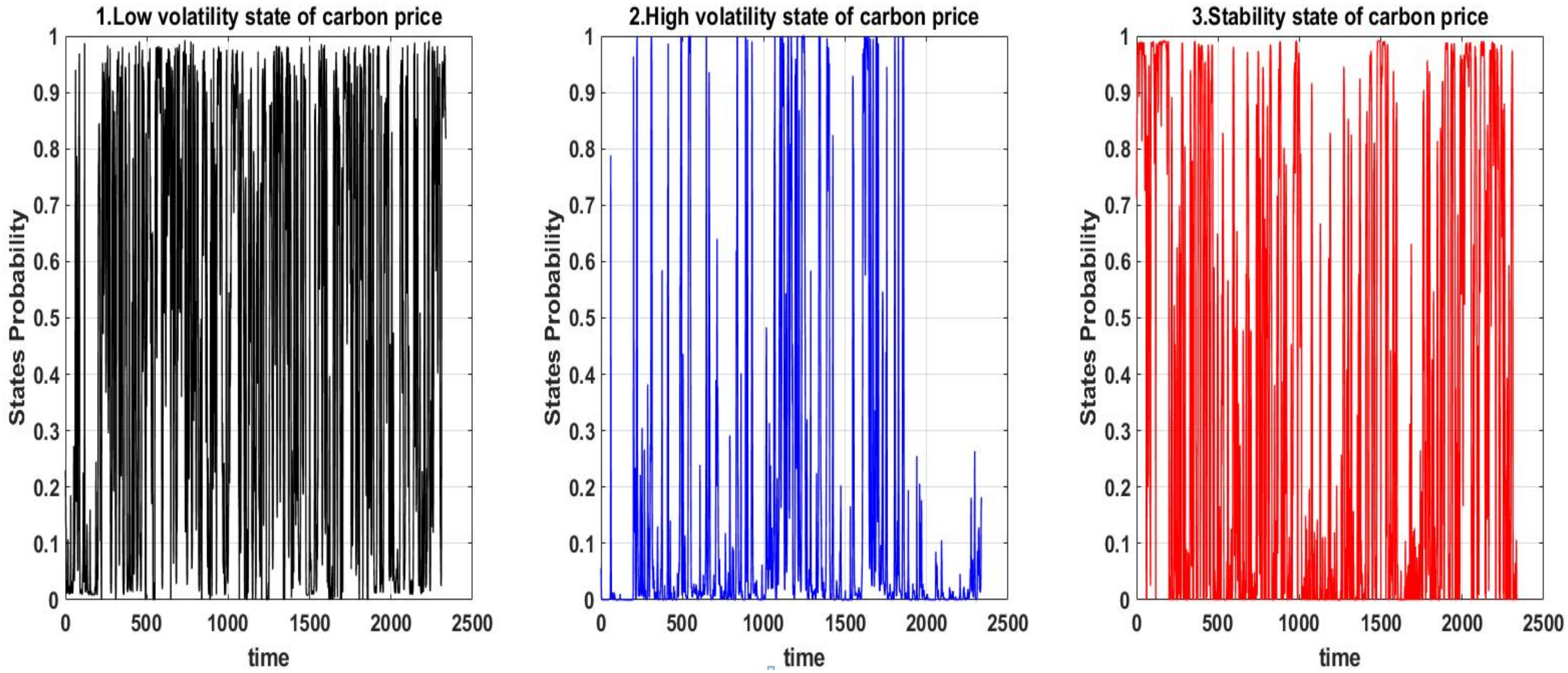


Figure 4.4: Smooth Probability Curve of the Three Volatility State Recognized by the MS(3)-AR(1) Model

Furthermore, this study find that the average volatility of the high volatility state (State 2) is 10 times greater than average volatility of the stable state (State 3) and 2.8 times the low volatility state (State 1). Therefore, this study defines transition between high volatility state (State 2) and stable volatility state (State 3) as rapid market change, that is transition from high volatility state (State 2) to stable volatility state (State 3) is considered as a rapid declining market trend, and the transition from stable state (State 3) to high volatility state (State 2) is considered as a rapid rising market trend. Similarly, the transition between low volatility state (State 1) and stable volatility state (State 3) is defined as slow market change, that is transition from low volatility state (State 1) to stable volatility state (State 3) is considered as a slow declining market trend, and the transition from stable state (State 3) to low volatility state (State 1) is a slow rising market trend. The carbon market trend classification can be shown in Table 4.4.

Table 4.4: Market Trend Classification Results of China Carbon Market

Market State	Market Trend	State Transition Direction	Specific Trends
State2-State3	Rapid change	From State3 to State2	Rapid rising trend
		From State2 to State3	Rapid declining trend
State1-State3	Slow change	From State3 to State1	Slow rising trend
		From State1 to State3	Slow declining trend

4.3.2 The High Order Risk Contagion under the Market Rapid Change

The carbon market volatility state transition implies a switching of carbon price risk, during this process, new risk exposures and profit opportunities can be released. Especially the extreme shocks of the carbon price volatility can encourage investors to engage in cross market investment portfolios and fund allocation through the irrational trading, consequentially, the high order risk contagion can be produced(Zhang et al.,2020). The carbon market rapid change indicates a drastic change in market price risk, which means that investors may face a rapidly changing market situation and serious challenges for avoiding market risks. It can be said that quick decision-making is more important than how to make decisions in this changing situation (Chevallier,2012). Due to the fact that rapid market volatility do not leave enough time for investors to conduct rational analysis and decision-making research, as a result, the irrational characteristics of investor trading motivation are more evident in the process of rapid change situation.

Research has found that in a rapid change market situation, there is not only a significant risk contagion relationship between China carbon market and infected markets in low order moment attributes, but also a significant risk contagion in the majority of high order moment channels (as shown in Tables 4.5 and 4.6). This evidence indicates that high order moment risk factors originated from market asymmetric and extreme shocks can spread across markets. Therefore, only depend on the traditional low order moment channels to determine the existence of risk contagion and conduct price forecasting may be difficult to obtain accurate conclusions.

4.3.2.1 Risk Contagion Analysis on the Rapid Rising Market Trend

When the carbon market volatility is in a rapid rising market trend, it means that the market uncertainty is rapid increasing, and price signals usually hidden more systemic risk factors, that posing an urgent need for risk aversion according to the classical Prospect Theory proposed by Tversky & Kahneman (1979). As an emerging policy market, the China carbon market was established later than the European carbon market, the market-oriented mechanism need to be further improved. Therefore, it is easy to trigger some systemic risks and become a risk contagion sourced market. To avoid those systemic risks, carbon market investors can diversify their risks through cross market transactions, resulting in a high order moment risk contagion phenomenon to its infected markets.

Table 4.5: Risk Contagion Test Results of Carbon Market Based on State3-State2

	HBEA-JMF	HBEA-JTF	HBEA-Oil	HBEA-CSI300	HBEA-EUAF
FR	0.0005	0.0058	0.0117	0.0007	0.0435
CS12	5.7055**	3.4529***	51.0723***	0.0408	12.8467***
CS21	3.3755***	5.6102***	0.0323	0.0248**	0.2981
CK13	0.0102	22.3302*	2527.4***	21.3872	599.7096***
CK31	2.8141***	2.5781***	1.3978	0.1248	0.0455**
CV22	3.5584***	6.7437***	15.7527***	0.0027	14.7678***

Note: *,**,*** means the significant in the level of 10%,5% and 1%, respectively.

(1) Basic analysis of research results

It can be concluded from Table 4.5 that when the carbon market under the rapid rising trend, there are obvious contagion phenomenon from sourced carbon market to JMF in the channel of CS12, CS21, CK31 and CV22, the contagion coefficients are 5.7055, 3.3755, 2.8141 and 3.5584, respectively. In terms of the risk contagion from sourced carbon market to JTF market, there are significant risk contagion phenomenon in the channel of CS12, CS21, CK13, CK31 and CV22, the contagion coefficients are 3.4529, 5.6102, 22.3302, 2.5781 and 6.7437, respectively. In terms of the risk contagion from sourced carbon market to Oil market, there are significant risk contagion phenomenon in channel of CS12, CK13 and CV22, the contagion coefficients are 51.0723, 2527.4 and 15.7527, respectively. In terms of the risk contagion from sourced carbon market to CSI300 market, there are significant risk contagion phenomenon in the channel of CS21, the contagion coefficients are 0.0248. In terms of the risk contagion from sourced carbon market to EUAF market, there are significant risk contagion phenomenon in channel of CS12,CK13,CK31 and CV22, the contagion coefficients are 12.8467, 599.7096, 0.0455 and 14.7678, respectively.

(2) Comparative analysis of research conclusions

Specifically, as shown in Table 4.5, firstly, because carbon market is a risk contagion source market, there is obviously high order moment risk contagion in the majority of contagion channels. In particular, the significant risk contagion between carbon and JMF and JTF market occur in major of high order moment channels, but not in the low order moment channel. This explanation indicates that carbon market is easily transmitted systemic risk to the JMF and JTF market by investors' cross market operations when facing extreme market risks. Similarly, investors of JMF and JTF market are also tend to regard the carbon market as the portfolio tools, resulting in contagion relationship. In fact, the nonlinear dynamic relationship between China's carbon market and energy market has been verified by many literature studies,such as Chang et al.(2019),Han et al.(2019),Ji et al.(2019),Jie et al.(2021), which to some extent supports the findings of this study.

Secondly, there are three significant risk contagion channels from the carbon market to the Oil market. The possible reason is that the carbon market investors tend to choose Oil product as the risk management tool, while the Oil market can diversify market risks through more channels, not limited to the carbon market.

Thirdly, the risk contagion from the carbon market to CSI300 market is significant only in the CS21 channel. Specifically, there is a notable risk contagion relationship with a contagion coefficient of 0.0248 in this channel. The possible reason is that, as an indicator reflecting China's macroeconomic situation, the CSI300 reflects the macroeconomic trend, which is guiding the price changes and risk volatility of the carbon market, as a result, the price trend of China carbon price is relatively consistent with the CSI300. As a matter of fact, the CSI300 market usually hidden macroeconomic events such as extreme events and policy black swan events can transmit risks to the carbon market through the high order moment attributes.This conclusion is completely consistent with the majority of research that listed in above literature part, such as Liu et al.(2021), Luo et al.(2024), Sun et al.(2019), that is Macroeconomic indicators are important external references for leading the carbon market.

Finally, the China carbon price and EUAF are homogeneous products, while in terms of market influence, market capacity, transaction scale and market-oriented construction, the China carbon market is still relatively insufficient. The asymmetric and extreme shocks in the European carbon market will largely be transmitted to the China carbon market through the high-order moment channels, while the risk contagion that from China market to European carbon market is relatively weak. For example, the China carbon market only has an risk contagion impact on European carbon market through channel of CS12,CK13,CK31 and CV22, with the coefficient are 12.8467, 599.7096, 0.0455 and 14.7678, respectively.

4.3.2.2 Risk Contagion Analysis on the Rapid Declining Market Trend

The rapid declining market trend means the risk is gradually decreasing, and the market uncertainty is also gradually decreasing. During this process, investors have more time to assess the market risk, which make the investment portfolios and risk management operations more rational (Zhang et al.,2019;Zhang et al.,2020). Although the carbon market and its infected markets are facing a decreasing asymmetric risk and extreme risks caused by irrational shocks during the rapid market fluctuations, these contagion phenomena cannot be ignored in this process.

Table 4.6: Risk Contagion Test Results of Carbon Market Based on State2-State3

	HBEA-JMF	HBEA-JTF	HBEA-Oil	HBEA-CSI300	HBEA-EUAF
FR	0.0177	0.0002	0.0001	0.0076	0.0508
CS12	5.6457**	3.4223***	50.1257***	0.0406	12.146***
CS21	3.3401***	5.5604***	0.0317	0.0247**	0.2818
CK13	4.0624	51.2922*	2726.7***	21.5646	311.5333***
CK31	0.0629***	0.7619***	2.1802	0.1002	31.7965**
CV22	3.7324***	5.2181***	11.1737***	0.0030	3.2836***

Note: *,**,*** means the significant in the level of 10%,5% and 1%, respectively.

(1) Basic analysis of research results

It can be concluded from Table 4.6 that when the carbon market under the rapid declining trend, there are significant risk contagion phenomenon from the sourced carbon market to JMF market in the channel of CS12, CS21, CK31 and CV22, the contagion coefficients are 5.6457, 3.3401, 0.0629 and 3.7324, respectively. In terms of the risk contagion from the sourced carbon market to JTF market, there are significant risk contagion phenomenon in the channel of CS12, CS21, CK13, CK31 and CV22, the contagion coefficients are 3.4223, 5.5604, 51.2922, 0.7619 and 5.2181, respectively. In terms of the risk contagion from the sourced carbon market to Oil market, there are significant risk contagion phenomenon in the channel of CS12, CK13 and CV22, the contagion coefficients are 50.1257, 2726.7 and 11.1737, respectively. In terms of the risk contagion from the sourced carbon market to CSI300 market, there are significant risk contagion phenomenon in the channel of CS21, the contagion coefficients are 0.0247. In terms of the risk contagion from the sourced carbon market to EUAF market, there are significant risk contagion phenomenon in the channel of CS12, CK13, CK31 and CV22, the contagion coefficients are 12.146, 311.5333, 31.7965 and 3.2836, respectively.

(2) Comparative analysis of research conclusions

Specifically, as shown in Table 4.6, firstly, the carbon market has a risk contagion relationship with the JMF market in major high order moment channels. Secondly, the risk contagion channels between the carbon market and the JTF market in all the high order moment channels. Thirdly, the risk contagion between China carbon market and Oil market occurs in major of high order moment channels. The possible reason is that the rapid declining trend means a reduction in market risk, the normal or small market risks are not sufficient to trigger cross market risk contagion, the common investor trading can enough to reduce the market risk. While for some market asymmetric and extreme risks, investors may adopt cross market risk management, that leading to risk contagion in high-order moment channels. Fourthly, during the period of rapid declining market trend, there is a significant risk contagion relationship between the carbon market and the CSI300 market in the high order channel of CS21. When market volatility is declining, the market risk factors are basic related to the fundamentals. Therefore, the carbon market and CSI300 market maintain a high order moment channel risk contagion relationship. In this situation, unless extreme events occur, the whole cross market risk contagion of the carbon market will remain low. Finally, during the rapid declining trend, the carbon market has an risk contagion relationship with European carbon prices in major of the high order moment channels. Due to the lower market risk, it can be confirmed that this contagion is more related to fundamental factors. Macroeconomic fundamentals are fundamental factors that affect the volatility of carbon market prices, and the flow of information and spillover effects between different markets are related to fundamental factors (Zhang et al.,2018; Zhang et al.,2023; Zeng et al.,2021).

4.3.3 The High Order Risk Contagion under the Market Slow Change

The carbon market slow change indicates that the market price risk changes relatively smoothly, and the market situation is conducive to investors making more effective investment decisions and risk avoidance strategies through cautious analysis. The research has found that under the market slow change, there is only a risk contagion relationship between the carbon market and its infected market in some high order moment channels, while there is no contagion in low order moment channels (as shown in Table 4.7 and Table 4.8). It is worth noting that the number of significant high order moment contagion channels between the carbon market and its infected market in market slow change is significantly less than the situation of market rapid change analyzed earlier.

4.3.3.1 Risk Contagion Analysis on the Slow Rising Market Trend

The slow rising market trend means the carbon market risk is gradually increasing, but the growth is not high and belongs to limited growth. Although the increasing in risk implies an increasing in market uncertainty, it is may difficult to trigger a significant risk contagion phenomenon (Zhang et al.,2019; Chevallier,2011). Especially when the whole macro-market volatility is weak, it is difficult for market fundamentals to cause widespread risk contagion. As a comparison, the asymmetric risk and extreme risks that sourced from the carbon market and its infected market may trigger cross market risk transmission, and final form a risk contagion phenomenon.

Table 4.7: Risk Contagion Test Results of Carbon Market Based on State3-State1

	HBEA-JMF	HBEA-JTF	HBEA-Oil	HBEA-CSI300	HBEA-EUAF
FR	0.0005	0.0015	0.0115	0.0007	0.0428
CS12	1.4786***	0.2551***	0.0504***	0.7077	0.5951
CS21	0.2662	0.3621	0.6070	0.2215***	0.0037
CK13	20.2555**	16.4573*	88.6551	1.4498	35.4678
CK31	0.1417	1.2033	1.9406	1.0464	0.4633
CV22	1.4553	2.1052	1.5088**	23.3554***	7.1255**

Note: *,**,*** means the significant in the level of 10%,5% and 1%, respectively.

(1) Basic analysis of research results

It can be concluded from Table 4.7 that when the carbon market under the slow rising trend, there are significant risk contagion phenomenon from the sourced carbon market to JMF market in the channel of CS12 and CK13, the contagion coefficients are 1.4786 and 20.2555, respectively. In terms of the risk contagion from the sourced carbon market to JTF market, there are significant risk contagion channels of CS12 and CK13, with the coefficients are 0.2551 and 16.4573, respectively. In terms of the risk contagion from the sourced carbon market to Oil market, there are significant risk contagion phenomenon in the channel of CS12 and CV22, the contagion coefficients are 0.0504 and 1.5088, respectively. In terms of the risk contagion from the sourced carbon market to CSI300 market, there are significant risk contagion phenomenon in the channel of CS21 and CV22, the contagion coefficients are 0.2215 and 23.3554, respectively. In terms of the risk contagion from the sourced carbon market to EUAF market, there is significant risk contagion phenomenon in the channel of CV22, the contagion coefficients is 7.1255.

(2) Comparative analysis of research conclusions

Specifically, as shown in Table 4.7, firstly, there are two significant risk contagion channels between carbon market and JMF, JTF, Oil and CSI300 markets. Secondly, the contagion risk between China carbon market and Europe carbon market only occurs in the channel of CV22, and there is no significance in other channels.

4.3.3.2 Risk Contagion Analysis on the Slow Declining Market Trend

The slow declining market trend represents the carbon market risk is gradually decreasing and ultimately reaching a stable state that consistent with the market fundamentals. During this trend, various risk factors gradually decrease, investor risk is controllable, market risk exposure gradually closes, and the investors demand for cross market speculative arbitrage and risk management decreases. Most investors can only obtain the fundamentals returns based on the findings of Zhang et al.(2019). That is, the carbon market and its infected market have risk contagion relationships in some high order moment channels rather than the low order moment channel.

(1) Basic analysis of research results

It can be concluded from Table 4.8 that when the carbon market under the slow declining trend, there are obvious contagion phenomenon from sourced carbon market to JMF market in the channel of CS12,CK13 and CK31, the contagion coefficients are 1.4776, 13.8393 and 0.1615, respectively. In terms of the risk contagion from the sourced carbon market to JTF market, there is no significant risk contagion phenomenon. In terms of the risk contagion from the sourced carbon market to Oil market, there are significant risk contagion phenomenon in the channel of CS12 and CV22, the contagion coefficients are 0.0501 and 1.2907, respectively. In terms of the contagion from the sourced carbon market to CSI300 market, there are significant risk contagion phenomenon in the channel of CS21 and CV22, the contagion coefficients are 0.2215 and 23.4001, respectively. In terms of the risk contagion from the sourced carbon market to the EUAF market, there is significant contagion phenomenon in the channel of CV22, and the contagion coefficients is 5.086.

Table 4.8: Risk Contagion Test Results of Carbon Market Based on State1-State3

	HBEA-JMF	HBEA-JTF	HBEA-Oil	HBEA-CSI300	HBEA-EUAF
FR	0.0050	0.0001	0.0050	0.0001	0.0100
CS12	1.4776*	0.1907	0.0501***	0.7075	0.5847
CS21	0.2660	0.9084	0.6036	0.2215***	0.0036
CK13	13.8393**	13.2721	73.2137	0.8508	12.4902
CK31	0.1615***	1.9209	0.3219	0.5495	2.7933
CV22	1.4735	8.7063	1.2907**	23.4001***	5.086**

Note: *,**,*** means the significant in the level of 10%,5% and 1%, respectively.

(2) Comparative analysis of research conclusions

Specifically, as shown in Table 4.8, firstly, there is no significant low order moment risk contagion relationship from carbon market to infected markets. This conclusion is completely consistent with finding of Fry et al.(2018) that low order moment channels represented by mean and variance are no longer able to trigger more risk contagion.

Secondly, carbon market and JMF market have three significant risk contagion channels. Carbon market and Oil and CSI300 market have two significant risk contagion channels. Especially there are two high order moment risk contagion channels of CS12 and CV22 that from carbon market to Oil market. Similarly, there are two high order moment risk contagion channels of CS21 and CV22 that from carbon market to CSI300 market.

This is basically consistent with the previous findings in the slow rising market trend, that there is a obvious contagion risk relationship between Oil market, CSI300 market and China carbon market in high order moment channels. Finally, there is only a risk contagion relationship from carbon market and European carbon market in the CV22 channel.

4.3.4 Summary of China Carbon Market Risk Contagion Channels

Based on the above analysis, this thesis finds that there are various style of risk contagion relationships among China carbon market and infected markets, including the low order moment contagion channels and the high order moment contagion channels.

Especially when market is in rapid and slow market trend, there are significant differences in the forms and channels of risk contagion. Actually, as this study mentioned earlier, the risk contagion is essentially a phenomenon of cross market risk transmission, while risk contagion in high order moment attributes refers to cross-market risk transmission arise through market irrationality and extreme risk shocks beyond market fundamentals (Fry et al.,2014;Chen et al.,2022). Measuring the high order moment channels risk contagion between the carbon market and infected markets aligns with special characteristics of the carbon market. Clarifying the specific risk contagion forms is a necessary condition for understanding carbon price formation path and revealing nonlinear carbon price driving mechanism.

Based on the previous measure results of contagion among carbon market and infected markets, this thesis summarizes and identifies specific risk contagion channels that have commonly significance in different market change state and market trend (Yun et al.,2020). Furthermore, these risk contagion channels will be incorporated into the theoretical model of the high order moment risk contagion carbon price forecasting framework, and thus support a foundation for the carbon price forecasting research in the following contents.

According to Tables 4.5, Tables 4.6, Tables 4.7 and Tables 4.8, this thesis obtains that the following risk contagion channels have obvious stability, the significant high order moment risk contagion channels of CS12 between carbon market and JMF market, the significant high order moment contagion channels of CS12 and CV22 between carbon market and Oil market, the significant high order moment risk contagion channels of CS21 between carbon market and CSI300 market, and the significant CV22 channel between carbon market and European carbon market. The summary results are shown in Table 4.9.

Table 4.9: Risk Contagion Channels between Carbon and Infected Markets

Contagion Market	**Risk Contagion Channels**
HBEA-JMF	CS12
HBEA-JTF	-
HBEA-Oil	CS12,CV22
HBEA-CSI300	CS21
HBEA-EUAF	CV22

4.4 Carbon Price Forecasting Based on the High Order Moment Risk Contagion

To make effective nonlinear price forecasting of China carbon market, reveal the impact of high order moment risk contagion factors on carbon price formation, this thesis incorporates the risk contagion factors identified in previous study into the China carbon price forecasting framework constructed in Chapter 3, and conducts the HOC-LSTM model for the out of sample forecasting. The experimental operation is a program design based on Matlab 2018b and Python 3.7. In terms of data processing, the first 80% of all the time series data is selected for model training, and the last 20% are used to test the model's forecasting performance.

4.4.1 Basic Parameter Design of the HOC-LSTM Model

4.4.1.1 Decide the Iteration Number

In the field of computer science, iteration numbers refers to repeatedly performing the same operation under certain conditions to achieve a optimal goal. For a machine learning model, iteration is a common used method, which can repeatedly train the model, continuously adjusts parameters, and ultimately makes the model forecasting results more accurate (Plakandaras et al.,2011; Zhou et al.,2019; Zhu et al.,2019). When training a machine learning model, an appropriate iterations number can effectively improve the forecasting performance and training efficiency of the model.

When determine the iterations number, it is first necessary to design a convergence condition. When the model forecasting results reaches a certain level, the training stops. The determination of this convergence condition needs to be based on specific problems. The loss function is an error criterion in machine learning that determines the iteration loss, that indicates the error between the model forecasting and actual results. During training a machine learning model, the gradient descent and other optimal algorithms are commonly used to minimize the loss function and achieve better forecasting results (Shen et al.,2015;Zhang et al.,2020; Ren et al.,2020). If the loss function trend starts to slow down, it can be considered that the model has started to converge, and training can be stopped at this time.

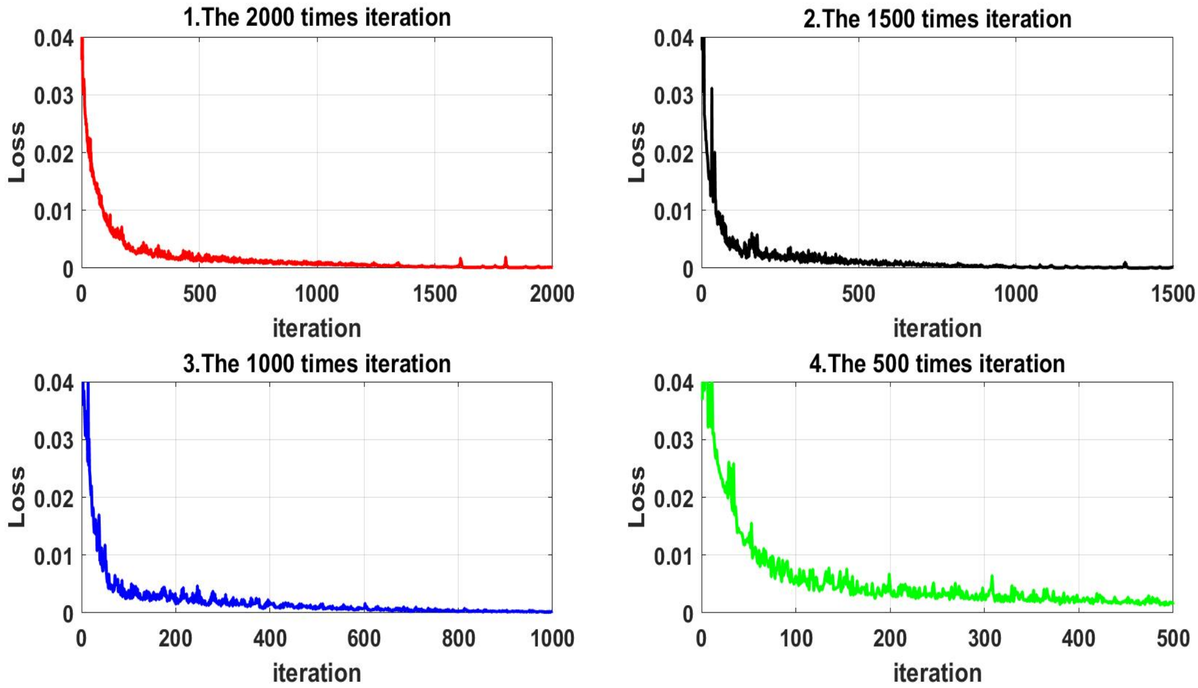


Figure 4.5: The Model Convergence Errors in Different Iteration Times

In conducting the experimental methods, this study refers to the experience of Shen et al.(2015) and Yun et al.(2023), and initializes and trains the HOC-LSTM forecasting model at the iterations numbers of 2000,1500,1000 and 500 respectively, the convergence results of the model as shown in Figure 4.5. According to the Figure 4.5, when the number of iterations is 2000, the model can basically converge quickly, and gradually approaching the minimum loss value. That is the model loss is basically stable after 2000 times parameter updating and gradient adjusting. Therefore, this thesis sets the iteration number of the HOC-LSTM forecasting model to 2000 times.

4.4.1.2 Decide the Learning Rate

Learning rate is a key concept in machine learning model, which affects the speed and stability of model training, it is one of the key parameters for model optimization. Briefly, the learning rate controls the update amplitude of a machine learning model parameters during the training process. A suitable learning rate can improve the training efficiency while ensuring the model convergence (Li et al., 2023;Yang et al., 2023). However, choosing the appropriate learning rate is not an easy task, because a bigger learning rates may cause the model divergence and difficult to achieve optimal solution, while a smaller learning rate can ensure the final model convergence, but it may reduce the model training speed and improve the training costs. Experiments have shown that different model structures and datasets typically require different learning rate settings.

The fixed step learning rate used in traditional gradient descent algorithms may cause model training not being able to obtain effectively result, which have weak model generalization ability. Based on this, this thesis uses the adaptive moment estimation Adam algorithm to update the model parameters. That is, make larger updates for the low-frequency parameters and smaller updates for high-frequency data. Furthermore, this thesis determines reasonable range of parameter updates through momentum adjustment of moment attributes. By optimizing model based on dynamic learning rate, the model achieves better robustness. According to the basic formula of adaptive moment estimation Adam algorithm mentioned earlier, this thesis sets the initial learning rate based on experimental methods. After setting the initial learning rate, the model algorithm can determine the dynamically adjusted learning rate based on the difference in the moment attribute dimensions of the input data, and then obtain the dynamic adjustment parameters.

The specific design is as follows: take the risk contagion factors identified in previous research into the HOC-LSTM model for parameter training. During the training process, this thesis refers to the experience of Shen et al.(2015) and Yun et al.(2023), and conducts the learning rates of 0.0001, 0.0003, 0.0006, 0.0009, 0.01, 0.03, 0.06, 0.09, 0.1, 0.3, 0.6, 0.9 as the alternative parameter, the average loss function during the supervised training process is used as the evaluation criterion, and the corresponding loss error performance of each learning rate is shown in Table 4.10 and Figure 4.6.

Table 4.10: Training Error Comparison of Different Parameter Learning Rates

Learning Rate	RMSE	MAE	Learning Rate	RMSE	MAE
0.0001	1.7188	1.1778	0.01	2.1220	1.3698
0.0003	1.7926	1.4148	**0.03**	**1.3716**	**0.9663**
0.0006	2.1589	1.5607	0.06	1.5057	1.0751
0.0009	2.7027	2.0412	0.09	1.6907	1.1928
0.001	2.8407	2.1117	0.1	1.6654	1.3071
0.003	2.9907	2.0890	0.3	1.8189	1.1773
0.006	2.1400	1.5148	0.6	2.0307	1.4355
0.009	2.0321	1.4724	0.9	2.1421	1.4402

Note: Bold represents the error loss corresponding to the optimal learning rate.

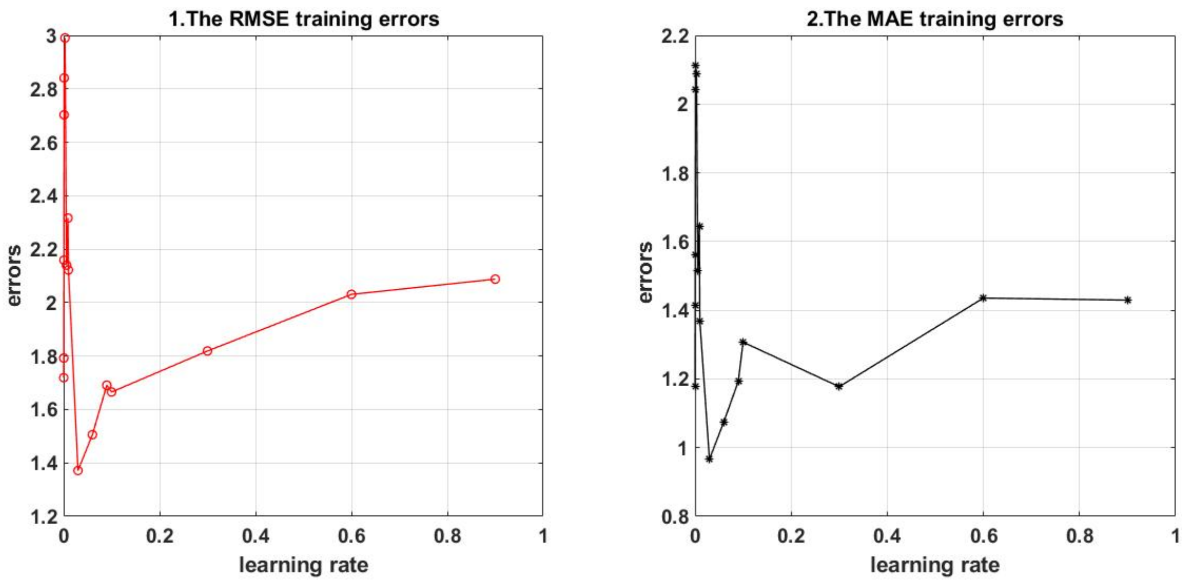


Figure 4.6: The Model Training Errors Based on Different Initial Parameter Learning Rate

This study find that when learning rate of adaptive moment estimation (Adam) algorithm is 0.03, the training errors RMSE and MAE are 1.3716 and 0.9663, respectively, which are the lowest values of all experimental results. That is, as the learning rate increases, the average training error gradually decreases and reaches its lowest point at a learning rate of 0.03. As learning rate reaches after point of 0.03, average error gradually increases (as shown in Figure 4.6). This evidence indicates that the HOC-LSTM model can achieve fast fitting performance at a learning rate of 0.03, which can effectively reduce training losses. Therefore,this study designs the initial learning rate of the HOC-LSTM forecasting model to 0.03.

4.4.1.3 Decide the Hidden Layers Number

As for the machine learning models, the hidden layer is a hierarchical structure composed of one or more layers of neurons located between input and output layers. The function of hidden layer is to enhance learning and generalization abilities of neural networks by introducing nonlinear transformations. The hidden layer can map input data to a higher characteristic space, and better understand input data features that need to be learned and extracted (Le and Bengio, 2008; Li et al.,2024).

Increase the number of hidden layers can help machine learning models capture complex features more accurately, reduce training errors, but it may also causes complex network structures and many parameters to be estimated, trigger gradient vanishing or gradient exploding problems, which may lead to training interruption. However, reduce the number of hidden layer may also decrease learning capacity of model, making it is not easy to approach training good result. Some famous studies have shown that neural networks with two hidden layers can already be sufficient to solve most problems (Le and

Bengio, 2008). This thesis uses the experimental methods to calculate the model training errors when hidden layers is 1, 2, 3, 4, 5, 6, 7 and 8. The results suggested that when hidden layers is 2, the model training errors RMSE and MAE are both the minimum values of the entire experimental results, that obtain the optimal training performance (as shown in Figure 4.7). Therefore, this thesis sets the initial hidden layers number of the HOC-LSTM forecasting model to 2.

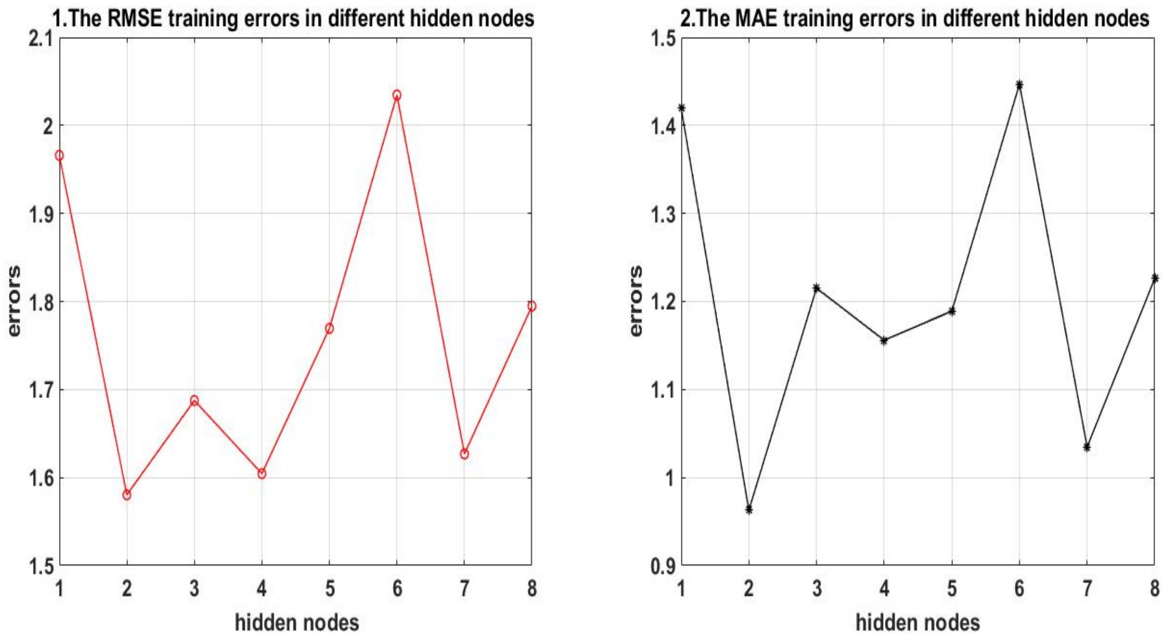


Figure 4.7: The Training Errors on Different Hidden Layers Nodes of HOC-LSTM

4.4.1.4 Decide the Neuron Nodes

Fewer neuron nodes in hidden layer may cause under-fitting. In contrast, too many neuron nodes can also cause over-fitting. When a neural network has a large number of nodes, the limited amount of information contained in training set is insufficient to train all the neurons in hidden layer, that resulting in over-fitting (Le and Bengio, 2008; Shen et al.,2015; Yun et al.,2023; Li et al.,2024). Even training data includes enough information, design too many neurons in hidden layer can also enhance the training time, it is not easy to achieve the expected results. Obviously, choose a suitable hidden layer neuron nodes is key for a machine learning model. Generally, use the same neuron nodes for all hidden layers is sufficient for most training cases. For some datasets, design larger first layer neuron nodes followed by smaller neuron nodes can also obtain expected result, the first layer can obtain many low-level features, which are then fed into the subsequent layers to extract high order features.

Table 4.11: Training Errors of HOC-LSTM on Different Hidden Layers and Neuron Nodes

Hidden Layers	Neuron Nodes	RMSE	MAE	Hidden Layers	Neuron Nodes	RMSE	MAE
1	2	2.1814	1.7869	4	2	2.5023	2.0493
	4	1.5347	0.9957		4	1.4528	1.0366
	8	1.5377	1.1744		8	1.9842	1.5448
	16	1.6423	1.2069		16	1.2251	0.8930
	32	2.0523	1.4718		32	1.3202	0.9366
	64	2.0762	1.4904		64	1.5500	1.1241
	128	1.9940	1.3735		128	1.2333	0.8170
2	2	1.9613	1.3250	5	2	1.4794	1.2196
	4	1.2820	0.9139		4	1.2905	0.9318
	8	2.3935	1.5474		8	1.5552	1.0912
	16	2.1615	1.6828		16	1.4062	1.0611
	32	1.3705	0.9620		32	1.5311	1.0973
	64	1.4501	1.1504		64	1.5699	1.1047
	128	**1.0322**	**0.7154**		128	1.4367	0.9974
3	2	1.4934	1.1821	6	2	2.0940	1.4341
	4	1.3201	0.8221		4	1.7711	1.2336
	8	1.7715	1.0613		8	2.0890	1.5216
	16	1.4448	1.0236		16	1.7011	1.1282
	32	1.4272	0.9637		32	1.9104	1.2942
	64	1.2083	0.8838		64	1.5126	1.0967
	128	1.2464	0.9307		128	1.5120	1.0574

Note: Bold represents the optimal neural network training errors.

Compared to other neural network networks, the LSTM network is a cyclic chain structure, that each hidden layer has a similar network structure and neuron nodes. The function of hidden layer neurons is to learn and map input data, which is similar to the function of hidden layers number. This study also employs the experimental methods and refers to the experience of Shen et al. (2015) and Yun et al.(2023) to calculate the training errors of the forecasting models with neuron nodes of 2, 4, 8, 16, 32, 64,128 and hidden layers of 1, 2, 3, 4, 5 and 6, respectively. Research indicate that when LSTM model has two hidden layers and neuron node structure are 128-128, the training error is the lowest. As shown in Table 4.11, when the neuron structure of HOC-LSTM model is 128-128, training errors RMSE and MAE are 1.0322 and 0.7154, respectively. Therefore, this thesis designs the initial hidden layer structure of the HOC-LSTM forecasting model to 128-128.

4.4.2 Carbon Price Forecasting Results Based on the Proposed HOC-LSTM Model

4.4.2.1 Forecast Performance Based on the Errors and Correlation

The HOC-LSTM model and its comparative models constructed in this thesis were used to conduct the out of sample forecast of the China carbon price. The results are shown in Table 4.13.

According to the Panel A and Panel B of the Table 4.12, as for the error indicators of RMSE, MAE and MAPE, the forecasting performance of the suggested HOC-LSTM model show the relatively lower errors. As for the Kendall correlation indicators, the proposed HOC-LSTM model is relatively high in the entire experiment, the correlation between predicted one and the real one is strong.

Table 4.12: Out-of-Sample Forecasting Error and Correlation of HOC-LSTM Model Based on the 20% Testing Sample

Model	RMSE	MAE	MAPE	Kendall
Panel A:high order risk contagion forecasting results				
HOC-LSTM	1.4667	0.9378	0.9746	0.9012
GRU	3.1695	2.4005	8.8951	0.3758
MLP	1.6837	1.1502	1.0832	0.8934
GBDT	1.7162	1.1500	1.3023	0.8944
ETR	1.7525	1.1692	1.1223	0.8759
BPNN	1.8212	1.0833	7.1593	0.6851
Panel B: non high order risk contagion forecasting results				
HOC-LSTM	1.6010	1.0765	1.0089	0.9784
GRU	2.6104	1.9204	8.0534	-0.0310
MLP	1.6197	1.0919	1.0410	0.9746
GBDT	2.4931	1.7303	5.9037	0.0670
ETR	1.6620	1.1173	1.4492	0.9602
BPNN	3.1077	1.7105	4.7064	0.0407

Specifically, firstly, as for the forecasting errors based on the high order risk contagion framework in Panel A of Table 4.12, the forecasting error indicators RMSE, MAE and MAPE of HOC-LSTM model are relatively minimum values, with the values of 1.4667, 0.9378 and 0.9746, respectively. While the forecasting performance of the MLP model, GBDT model, ETR model and BPNN model are basically the same, the error values of RMSE, MAE and MAPE are significantly greater than the performance of the HOC-LSTM model. For example, the forecasting error indicators RMSE, MAE and MAPE of MLP model are 1.6837, 1.1502 and 1.0832, respectively. The forecasting error indicators RMSE, MAE and MAPE of GBDT model are 1.7162, 1.1500 and 1.3023, respectively. The forecasting error indicators RMSE, MAE and MAPE of ETR model are 1.7525, 1.1692 and 1.1223, respectively. The forecasting error indicators RMSE, MAE and MAPE of BPNN model are 1.8212, 1.0833 and 7.1593, respectively. Furthermore, the GRU model performs poorly, the forecasting error of GRU model is the largest, the results suggest that carbon price prediction effect of GRU model is poor, and it is difficult to obtain more accurate conclusions based on this model. This conclusion is completely same with the studies of Yun et al. (2023) that the GRU-type single and hybrid models are commonly weaker than other machine learning models, especially the LSTM-type models.

In terms of the Kendall correlation indicators, based on the Panel A of Table 4.12, in the high order risk contagion forecasting framework, the Kendall correlation between the predicted carbon price of the HOC-LSTM and the real price is 0.9012, this value is bigger than the correlation of other comparative models, it shows that the correlation between predicted one and the real one is stronger. Furthermore, the Kendall correlation between the predicted carbon price of the GRU model and the real price is 0.3758, which is the lowest of the entire sample. While other comparative models shown smaller correlation, that is, the Kendall correlation between the predicted carbon price of the MLP model and the real price is 0.8934, the Kendall correlation between the predicted carbon price of the GBDT model and the real price is 0.8944, the Kendall correlation between the predicted carbon price of the ETR model and the real price is 0.8759. The Kendall correlation between the predicted carbon price of the BPNN model and the real price is 0.6851. These findings suggest that these four models have a relatively consistent forecasting performance of the carbon prices. The conclusion can provide effective explanation and mechanism analysis for revealing the effect of high order moment risk contagion on carbon price. In fact, the network training and forecasting ability of LSTM model is theoretically superior to MLP, BP and other models, and this experimental result is completely consistent with the theoretical conclusion and some findings in previous studies of Yun et al.(2023), Yang et al.(2023) and Zhu et al.(2024).

Therefore, as for the forecasting results in capturing the effect of high order moment risk contagion on China carbon price, the HOC-LSTM model appears the relatively lower errors, and a smaller dynamic deviation between predicted and actual price (as shown in Figure 4.8). From the correlation plots of the forecasting and actual value of HOC-LSTM model and its comparative models, the HOC-LSTM model has a better correlation, while the forecasting and actual value of GRU and BPNN model are more dispersed, indicating a bad correlation (as shown in Figure 4.9). This indicates that compared to other models, HOC-LSTM model can not only capture nonlinear characteristics of China carbon price, but also reveal the impact mechanism of the high order moment risk contagion on the carbon price, especially the complex price driving mechanism. Therefore, the forecasting effect of proposed HOC-LSTM model can provide an effective technical tools for investors, emission enterprises to take the carbon trading, quantify investments, and risk management. This conclusion is completely consistent with the findings of Yun et al. (2020), that is, the forecasting model considering high order moment risk contagion is significantly better than the forecasting models without considering risk contagion, and the forecasting errors is also significantly smaller.

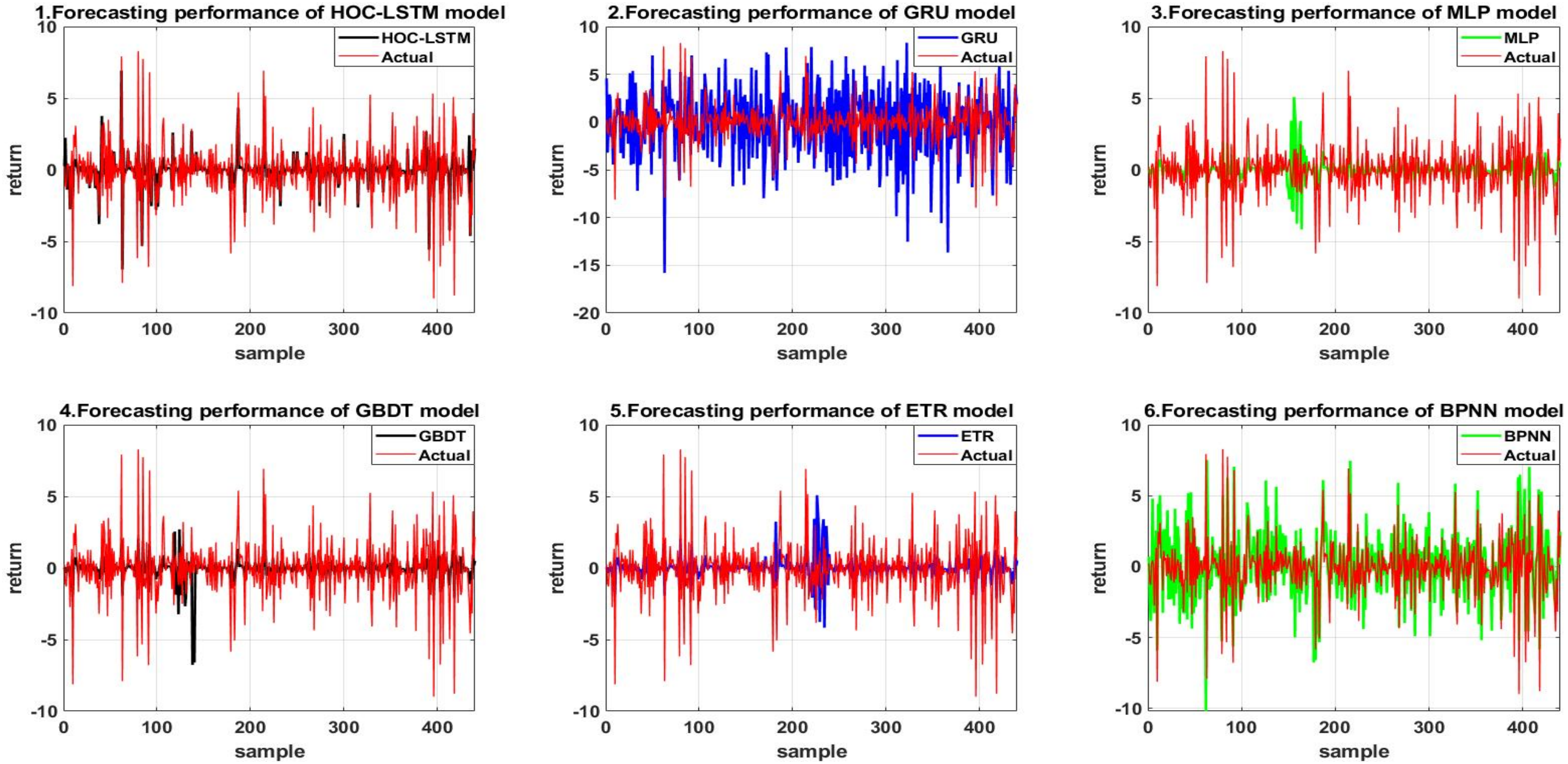


Figure 4.8: The China Carbon Price Out-of-Sample Forecasting Curve Based on the Impact of High Order Risk Contagion (The 20% Testing Sample of HOC-LSTM Model and its Comparative Models)

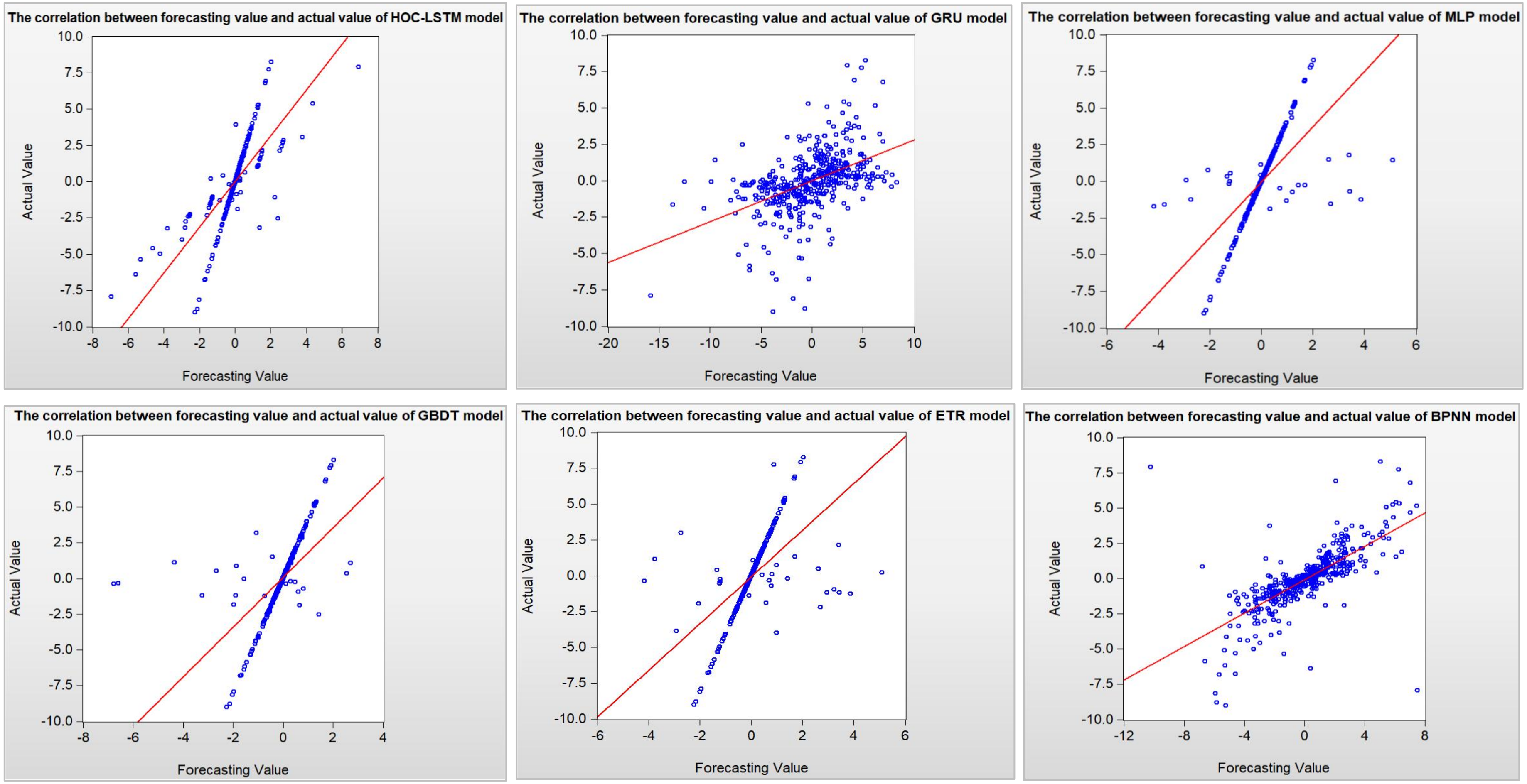


Figure 4.9: The Correlation Between Forecasting Value and Actual Value Based on the Impact of High Order Risk Contagion (The 20% Testing Sample of HOC-LSTM model and its comparative models)

Secondly, according to the Panel B of the Table 4.12, the forecasting performance of HOC-LSTM model remains stable and excellent in RMSE, MAE and MAPE based on the non high order moment risk contagion forecasting framework, the finding is completely same with result in Figure 4.10, and further proving that the HOC-LSTM model constructed in this thesis has a robust performance. For example, as shown in Panel B of the Table 4.12, the forecasting error indicators RMSE, MAE and MAPE of HOR-LSTM model are 1.6010, 1.0765 and 1.0089, respectively. The forecasting error indicators RMSE, MAE and MAPE of GRU model are 2.6104, 1.9204 and 8.0534, respectively. The forecasting error indicators RMSE, MAE and MAPE of MLP model are 1.6197, 1.0919 and 1.0410, respectively. The forecasting error indicators RMSE, MAE and MAPE of GBDT model are 2.4931, 1.7303 and 5.9037, respectively. The forecasting error indicators RMSE, MAE and MAPE of ETR model are 1.6620, 1.1173 and 1.4492, respectively. The forecasting error indicators RMSE, MAE and MAPE of BPNN model are 3.1077, 1.7105 and 4.7064, respectively.

In terms of the Kendall correlation indicators, according to the Panel B of Table 4.12, the Kendall correlation between the predicted carbon price of the HOC-LSTM model and the real price is 0.9784, which is bigger than other comparative models. The Kendall correlation of GRU model is -0.0310. The Kendall correlation between the predicted carbon price of GBDT model and BPNN and the real price is 0.0670 and 0.0407, respectively. These models have almost no correlation. The Kendall correlation between the predicted carbon price of MLP and ETR models and real price is 0.9746 and 0.9602, respectively. As shown in Figure 4.11, high correlation means that the trend consistency between the predicted carbon price and the actual value is high, which can provide direction reference for market investment and financing.

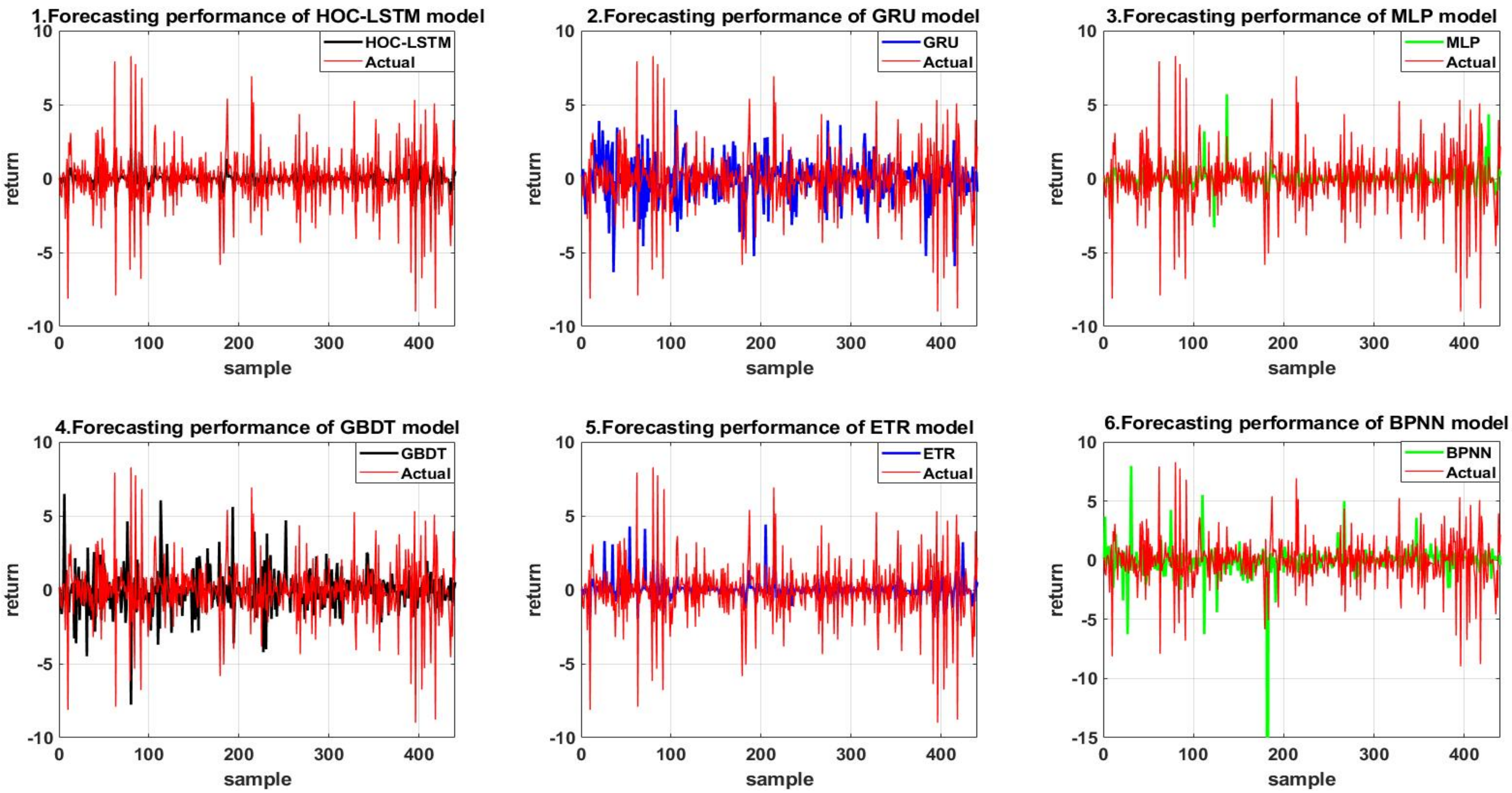


Figure 4.10: The China Carbon Price Out-of-Sample Forecasting Curve that Without Impact of High Order Risk Contagion (The 20% Testing Sample of HOC-LSTM Model and its Comparative Models)

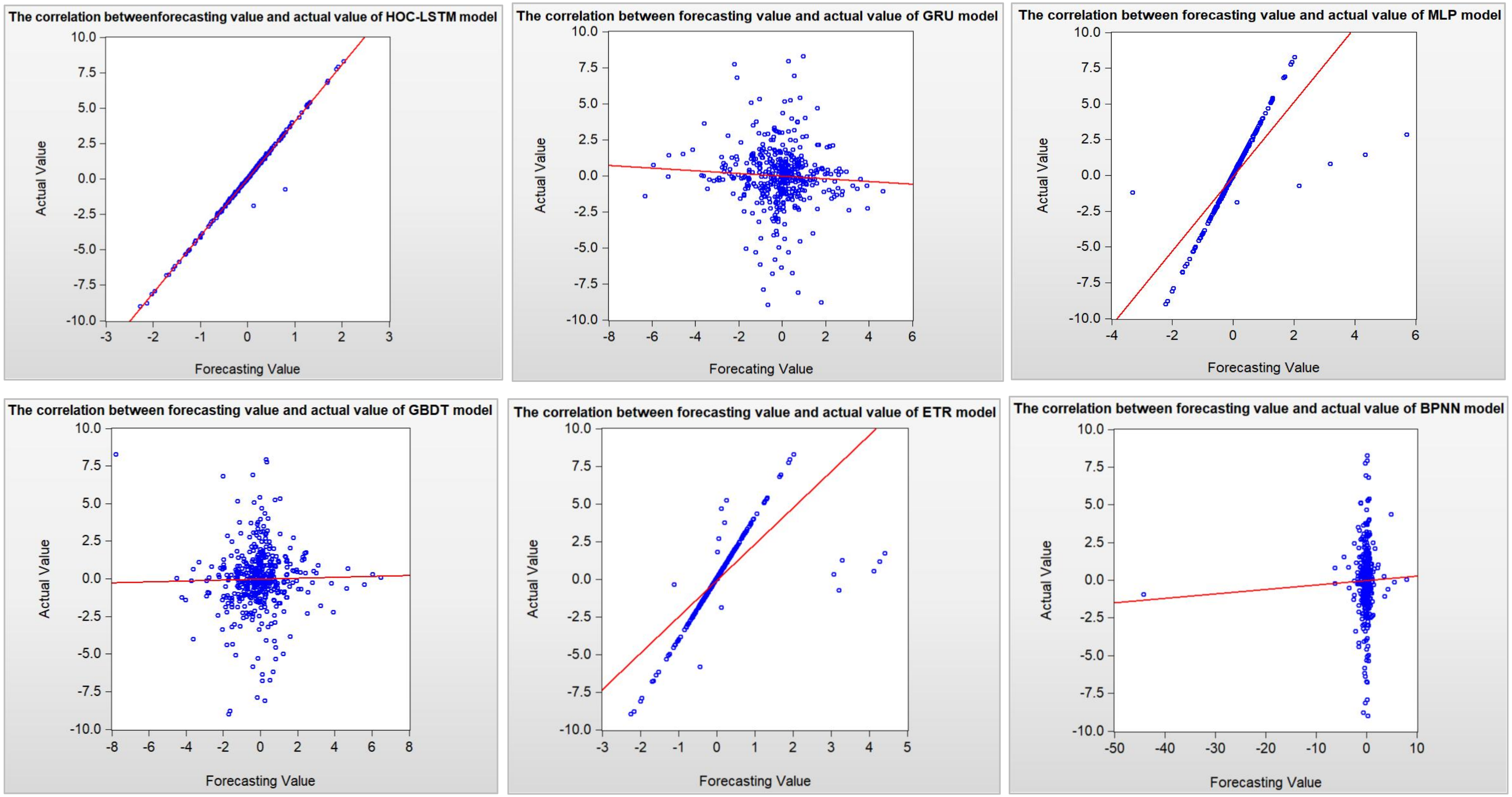


Figure 4.11: The Correlation Between Forecasting Value and Actual Value Without Impact of High Order Risk Contagion
(The 20% Testing Sample of HOC-LSTM Model and its Comparative Models)

Thirdly, after comparing the RMSE, MAE and MAPE errors results on Panel A and Panel B of the Table 4.12, this thesis finds that it can provide more accurate explanation for the carbon premium formation by incorporating the high order moment risk contagion factors into price forecasting model. That is, the high order moment risk contagion factors are indispensable component to explain the carbon price, and ignore those risk contagion factors will significantly reduce the carbon price forecasting ability. For example, the forecasting errors RMSE, MAE and MAPE of HOC-LSTM model in Panel A are 1.4667, 0.9378 and 0.9746, respectively, which are obviously lower than the forecasting errors in Panel B. Therefore, considering high order moment risk contagion relationship from risk source carbon market to infected markets can significantly enhance forecasting performance. This conclusion also indicates that carbon price is not only related to low order moment risk factors, but high order moment risk contagion factors that commonly ignored by previous scholars are also important price driving factors, these conclusion are completely consistently with the findings of Fry et al.(2014), Yun et al.(2020) and Chen et al.(2022). The superior performance of the HOC-LSTM model give support for investors and emission reduction firms to analyze market, forecast profit trends and make risk management decisions.

Therefore, based on the analysis above, this thesis concludes that the HOC-LSTM model has obvious advantages in forecasting China carbon price both in the high order and non high order risk contagion forecasting frameworks. This evidence indicates that HOC-LSTM model can effectively fit and map complex nonlinear, non normal, and non-stationary China carbon price data, out of sample forecasting effect is satisfactory. As for the purpose of this study, it focuses more on the model forecasting error, that is, the model that can bring smaller errors is the superior model. Actually, in order to further demonstrate the robustness of the conclusions in this thesis, sample adjustment testing and *W* testing will continue in next step.

4.4.2.2 Forecast Performance Based on the *W* Test

To estimate the contemporaneously correlation forecast errors problems between the proposed HOC-LSTM and its comparative models, this thesis uses the *w* test that suggested by Granger and Newbold (2004), and adopted in Liew (2006). By conducting the *w* test, it can be given the relative errors and the statistical significance of one model over another. Based on this, this thesis constructs three new error indicators $Z\text{-}_{RMSE}$, $Z\text{-}_{MAE}$ and $Z\text{-}_{MAPE}$. Among them, $Z\text{-}_{RMSE}$ is calculated by dividing RMSE of HOC-LSTM model by RMSE of other comparative models. $Z\text{-}_{MAE}$ is calculated by dividing MAE of HOC-LSTM model by MAE of other comparative models. $Z\text{-}_{MAPE}$ is calculated by dividing MAPE of HOC-LSTM model by MAPE of other comparative models. If this value is greater than 1, it indicates that the error of the suggested model proposed in this thesis is greater than that of other comparative models, as a result, the prediction effect of HOC-LSTM model is poor. Otherwise, the model in this thesis has good superiority.

Furthermore, following Granger and Newbold's (2004) , this study constructs two error sequences, $x_i = e_{1i} + e_{2i}$ and $z_i = e_{1i} - e_{2i}$, $i = 1, 2, \cdots, H$. where, e_{1i} and e_{2i} indicate the forecasting errors between the proposed HOC-LSTM model and a comparative model. Given that the first two assumptions above are valid, under the null hypothesis of equal forecast accuracy, x_i and z_i should be uncorrelated. Consider: $\rho_{xz} = Ex_i z_i = E(e_{1i}^2 - e_{2i}^2)$, If ρ_{xz} is positive, it indicates that the $Z\text{-}_{RMSE}$ error of HOC-LSTM is larger than comparative models, while if ρ_{xz} is negative, it indicates that the error of HOC-LSTM is smaller. Assuming that r_{xz} represents the Kendall correlation coefficient between e_{1i} and e_{2i} , and according to Granger and Newbold (2004), the test statistic $r_{xz} / \sqrt{(1-r_{xz}^2)/(H-1)}$ follows the *t*-distribution under *H*-1 degrees of freedom.

Table 4.13: *W* Test Results of HOC-LSTM Model Based on the 20% Testing Sample

Model	*Z-RMSE*	*Z-MAE*	*Z-MAPE*	*W* test	*P* value
Panel A: High order risk contagion *w* test results					
GRU	0.4628	0.3907	0.1096	-7.7486***	0.0000
MLP	0.8711	0.8153	0.8997	-0.2674***	0.0065
GBDT	0.8546	0.8155	0.7484	-0.2729***	0.0060
ETR	0.8369	0.8021	0.8684	-0.2729***	0.0060
BPNN	0.8053	0.8657	0.1361	-0.9516***	0.0093
Panel B: Non high order risk contagion *w* test results					
GRU	0.6133	0.5606	0.1253	-4.2510***	0.0000
MLP	0.9885	0.9859	0.9692	-0.0602	0.3444
GBDT	0.6422	0.6221	0.1709	-0.1991	0.3208
ETR	0.9633	0.9635	0.6962	-0.1991	0.3208
BPNN	0.5152	0.6293	0.2144	-7.0943***	0.0000

Note: *** means the significant in the level of 1%.
$Z_{\text{-RMSE}}$ is calculated by dividing RMSE of HOC-LSTM model by RMSE of other comparative models.
$Z_{\text{-MAE}}$ is calculated by dividing MAE of HOC-LSTM model by MAE of other comparative models.
$Z_{\text{-MAPE}}$ is calculated by dividing MAPE of HOC-LSTM model by MAPE of other comparative models.

The results shown in Panel A of Table 4.13 that, as for the *w* test under high order moment risk contagion, the relative errors $Z_{\text{-RMSE}}$ of the GRU, MLP, GBDT, ETR and BPNN model are all less than 1, with values of 0.4628, 0.8711, 0.8546, 0.8369 and 0.8053, respectively. The relative errors $Z_{\text{-MAE}}$ of the GRU, MLP, GBDT, ETR and BPNN model are also all less than 1, with values of 0.3907, 0.8153, 0.8155, 0.8021 and 0.8657, respectively. The relative errors $Z_{\text{-MAPE}}$ of the GRU, MLP, GBDT, ETR and BPNN model are also less than 1, with values of 0.1096, 0.8997, 0.7484, 0.8684 and 0.1361, respectively. This indicates that the carbon price forecasting performance of the HOC-LSTM model proposed in this study is better than other comparative models, and forecasting error is smaller based on the *w* test results. This result is completely same with discussion results in 4.4.3.1. Furthermore, the *w* test values for the GRU, MLP, GBDT, ETR, and BPNN model are -7.7486, -0.2674, -0.2729, -0.2729 and -0.9516, respectively, and there are statistical significance in all the comparative models according to the *p* value in Table 4.14.

Similarly, in Panel B of Table 4.13, as for the *w* test under non high order moment risk contagion, the relative errors $Z_{\text{-RMSE}}$ of the GRU, MLP, GBDT, ETR, and BPNN model are also less than 1, with values of 0.6133, 0.9885, 0.6422, 0.9633 and 0.5152, respectively. The relative errors $Z_{\text{-MAE}}$ of the GRU, MLP, GBDT, ETR, and BPNN model are also less than 1, with values of 0.5606, 0.9859, 0.6221, 0.9635 and 0.6293, respectively. The relative errors $Z_{\text{-MAPE}}$ of the GRU, MLP, GBDT, ETR, and BPNN model are also less than 1, with values of 0.1253, 0.9692, 0.1709, 0.6962 and 0.2144, respectively. Those evidence indicate that carbon price prediction performance of HOC-LSTM model is relatively better than other comparative models. Furthermore, *w* test values for the comparative models of GRU, MLP, GBDT, ETR and BPNN model are -4.2510, -0.0602, -0.1991, -0.1991 and -7.0943, respectively, which indicate statistical significance only in the models of GRU and BPNN.

Therefore, according to the results of the *w* statistical test analyzed above, the HOC-LSTM model constructed in this study not only has relatively good forecasting accuracy, but also the forecasting results still have significant statistical significance. Those explanation mean that proposed HOC-LSTM model has obvious effect and reliability in solving price forecasting problem in China carbon market compared to other models.

4.4.3 Re-Test of Forecasting Performance Based on Readjustment of Training Data

To further demonstrate out of sample forecasting advantage of HOC-LSTM model and provide more credible conclusions, following the idea of Zhang et al.(2020) and Yun et al.(2023), this study readjusts the ratio of training and testing data. Based on the initial 80% training data, this study continues reduce the ratio of training set, and test forecasting performance of HOC-LSTM model on 70% of the training data respectively, and 30% data for testing respectively. Based on the experiment design, if the results are consistent with previous analysis, it indicates that forecasting superiority of the HOC-LSTM model has strong robustness, and the model conclusion is reliable. It is worth noting that this study only readjust the ratio of the training set and test data, while design of other neural network structures and initial parameters remained consistent with the previous research.

Table 4.14: Out-of-Sample Forecasting Errors and Correlation of HOC-LSTM Model Based on the 30% Testing Sample

Sample Classification	**Model**	**RMSE**	**MAE**	**MAPE**	**Kendall**
70% training sample and 30% testing sample	HOC-LSTM	2.1356	1.3208	0.8875	0.9580
	GRU	2.5316	1.5589	5.0271	0.5784
	MLP	2.2545	1.4124	0.9222	0.9513
	GBDT	2.2065	1.3863	0.9500	0.9560
	ETR	2.4694	1.4557	0.8995	0.9420
	BPNN	2.3141	1.4744	2.8997	0.6828

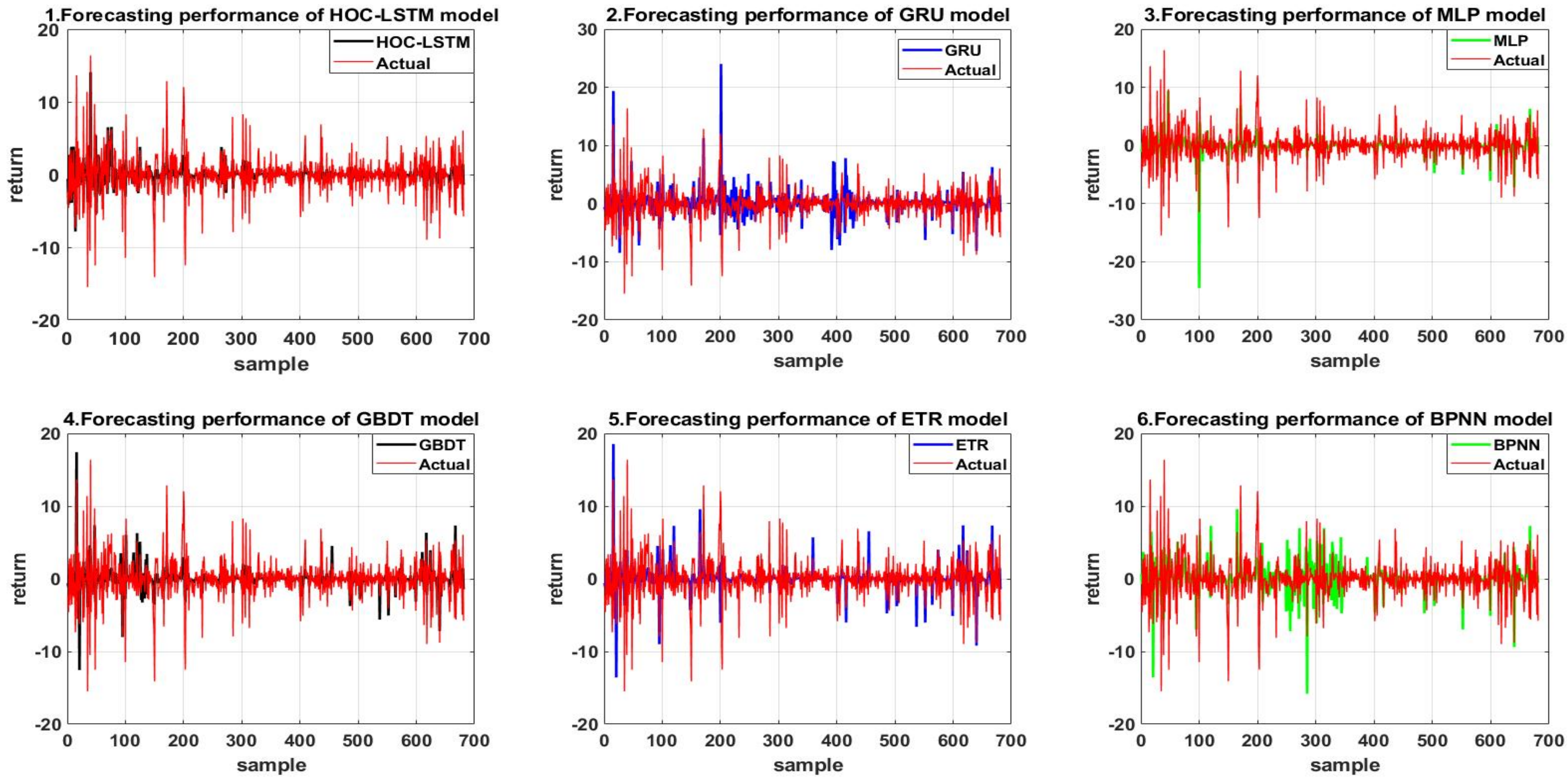


Figure 4.12: The China Carbon Price Out-of-Sample Forecasting Curve Based on the High Order Risk Contagion (The 30% Testing Sample of HOC-LSTM Model and its Comparative Models)

Firstly, the results have put that, as shown in Table 4.14, when the ratio of training and testing data is designed at 70% and 30%, the HOC-LSTM model has the smallest RMSE, MAE and MAPE forecasting errors compared with other comparative models. As shown in Figure 4.12, the curve fitting of predicted price and the real price is completely the same with minor errors. The forecasting errors of RMAE, MAE and MAPE of HOC-LSTM model are 2.1356, 1.3208 and 0.8875, respectively. While other comparative models have relatively higher error, that is the forecasting error indicators RMSE, MAE and MAPE of GRU model are 2.5316, 1.5589 and 5.0271, respectively. The forecasting error indicators RMSE, MAE and MAPE of MLP model are 2.2545, 1.4124 and 0.9222, respectively. The forecasting error indicators RMSE, MAE and MAPE of GBDT model are 2.2065, 1.3863 and 0.9500, respectively. The forecasting error indicators RMSE, MAE and MAPE of ETR model are 2.4694, 1.4557 and 0.8995, respectively. The forecasting error indicators RMSE, MAE and MAPE of BPNN model are 2.3141, 1.4744 and 2.8997, respectively. Those evidence indicate that no matter the different ratio of training data, it does not affect HOC-LSTM model for prediction China carbon price. The proposed model can effectively depict the high order moment risk infection factors triggered by market asymmetric and extreme shock, that this finding has also been proved in previous studies that listed in literature review such as Fry et al.(2018) and Yun et al.(2020).

Secondly, as for the Kendall correlation indicators, the proposed HOC-LSTM model is relatively higher than other comparative models. Specifically, the Kendall correlation between the predicted carbon price of the HOC-LSTM and the real price is 0.9580, the Kendall correlation between the predicted carbon price of the GRU model and the real price is 0.5784, which is the lowest of the entire sample, this conclusion is completely sample with previous analysis. The Kendall correlation between the predicted carbon price of the

MLP model and the real price is 0.9513, the Kendall correlation between the predicted carbon price of the GBDT model and the real price is 0.9560, the Kendall correlation between the predicted carbon price of the ETR model and the real price is 0.9420, which is the biggest value of the sample. The Kendall correlation between the predicted carbon price of the BPNN model and the real price is 0.6828.

Table 4.15: *W* Test Results of HOC-LSTM model Based on the 30% Testing Sample

Model	$Z_{\text{-RMSE}}$	$Z_{\text{-MAE}}$	$Z_{\text{-MAPE}}$	*W* test	*P* value
GRU	0.8436	0.8473	0.1765	-1.8484***	0.0003
MLP	0.9473	0.9351	0.9624	-0.5222***	0.0000
GBDT	0.9679	0.9528	0.9342	-1.5374***	0.0000
ETR	0.8648	0.9073	0.9867	-1.5374***	0.0000
BPNN	0.9229	0.8958	0.3061	-0.7941***	0.0076

Note:*** means the significant in the level of 1%.
$Z_{\text{-RMSE}}$ is calculated by dividing RMSE of HOC-LSTM model by RMSE of other comparative models.
$Z_{\text{-MAE}}$ is calculated by dividing MAE of HOC-LSTM model by MAE of other comparative models.
$Z_{\text{-MAPE}}$ is calculated by dividing MAPE of HOC-LSTM model by MAPE of other comparative models.

Thirdly, as for the *w* test based on the last 30% of the training sample in Table 4.15, the relative errors $Z_{\text{-RMSE}}$ of the GRU, MLP, GBDT, ETR, and BPNN model are also less than 1, with values of 0.8436, 0.9473, 0.9679, 0.8648, and 0.9229, respectively. The relative errors— $Z_{\text{-MAE}}$ of the GRU, MLP, GBDT, ETR, and BPNN model are also less than 1, with values of 0.8473, 0.9351, 0.9528, 0.9073, and 0.8958, respectively. The relative errors $Z_{\text{-MAPE}}$ of the GRU, MLP, GBDT, ETR, and BPNN model are also less than 1, with values of 0.1765, 0.9624, 0.9342, 0.9867 and 0.3061, respectively. Those evidence indicate that carbon price prediction performance of HOC-LSTM model is relatively better than other comparative models. Furthermore, *w* test values for the comparative models of GRU, MLP, GBDT, ETR and BPNN model are -1.8484, -0.5222, -1.5374, -1.5374 and -0.7941, respectively, which indicates statistical significance in all the comparative models. Those explanation mean that proposed HOC-LSTM model has obvious reliability in solving price forecasting problem in China carbon market.

4.4.4 The Economic Implications of the Empirical Results

Firstly, carbon pricing is an important policy tools for addressing climate issues, and has significant practical implications for sustainable development of society. The carbon market is a market-oriented mechanism to promote carbon emission reduction. Carbon market participants are set limits on greenhouse gas emissions, so the participants who exceed the emission quota are required to purchase emission rights, while those with emissions below the quota can sell their excess quotas in the carbon market (Chevallier,2012; Aatola et al.,2013; Zhu et al.,2024). This cap and trade mechanism can creates supply and demand of emission allowance, discovers market prices for greenhouse gas emissions, and guides carbon market participants to reduce carbon emissions that below government set emission limits at reasonable costs.

Secondly, carbon pricing can achieve a "carbon emitter pays" approach, effectively reducing global carbon emissions according to the research of Adekoya et al.(2021) and Qiu et al.(2023). The forecasting price of carbon emission rights inherits the principle of "polluter pays", by setting economic costs for carbon emission behavior, and assuming the environmental damage caused by carbon emissions. This internalization process is known as carbon emission pricing, which imposes clear economic responsibilities on carbon emission behavior, thereby addressing challenges such as free riding and moral hazard, and gradually realizing the principle of "carbon emitter pays". The gradual promotion of this mechanism on a global scale has laid a solid foundation for sustainable human development and provided solutions for global cooperation.

Thirdly, the determining factor of carbon price is balance between market allowance supply and demand, as well as the emission reduction costs of various industries. The carbon market can be divided into two segments: the primary and secondary markets. Much like the stock and bond markets, the primary market is where carbon emission allowances are created and issued. These allowances can be obtained through either free distribution or auction, with the government typically holding dominant control. The secondary market, on the other hand, is where carbon assets and carbon derivatives are traded and circulated. The primary market serves as the foundation and prerequisite for the secondary market (Conrad et al.,2013; Zhu et al.,2018). Carbon emission allowances and credits must first pass through the primary market before they can be traded in the secondary market.

Fourthly, the focus of this study is studying the formation and determination mechanism of carbon prices in the secondary market. It considers the effect path of high order moment risk contagion on carbon prices, and constructs a machine learning model to effectively forecast carbon prices in the secondary market. Research has found that there is a risk contagion relationship between carbon market and its infected markets through high order moment channels, which has an undeniable impact on the carbon prices trends. The risk contagion of this high order moment attribute is closely related to the market asymmetry and high sensitivity to policy shocks of carbon market (Fry & Hsiao, 2018; Elsayed et al., 2022; Tsai et al.,2024;Wang et al.,2024). Further research has used these identified high order moment risk contagion relationships as influencing factors for out of sample forecasting of carbon prices. The results show that using machine learning model HOC-LSTM can effectively and accurately forecast China's carbon prices with less error and higher accuracy. Therefore, this is a reliable price forecasting model. For investors, especially those in secondary market, carbon price changes are more complex. Obtaining effective and superior forecasting models to quantify and analyze carbon prices can improve market decision-making efficiency, increase investor market profits, and enhance the foresight of market decisions. So, forecasting carbon price in secondary market conforms to economic laws and practical management needs.

4.5 Summary of the Chapter

Based on the identification results of the high order moment risk contagion and the forecasting effect of proposed HOC-LSTM model, the main findings of this chapter are as follows:

Firstly, no matter the market volatility is under the rapid or slow trend, there is only high order moment risk contagion channels between carbon market and its infected markets. Therefore, revealing carbon price driving mechanism only through low order moment risk contagion channels may difficult to provide complete evidence for explaining risk premiums.

Secondly, the HOC-LSTM model constructed in this study has a significant superiority in fitting and forecasting the high order moments risk contagion factors, and indirectly convinced the rationality of taking the risk contagion relationship into the price forecasting models. It is also an indispensable factors in explaining carbon price formation mechanism and exploring driving mechanism of carbon premium.

Thirdly, carbon price forecasting performance of HOC-LSTM model has strong stability and robustness, that is, no matter how ratio of training data changes, it does not affect the carbon price forecasting results of HOC-LSTM model. The model can effectively catching high order moment risk contagion relationship that triggered by information asymmetric and extreme shock.

CHAPTER 5

CONCLUSION AND RECOMMENDATIONS

5.1 Introduction

This chapter provides an overall of the study on risk contagion and price forecasting of China carbon market based on high order moment attribute. The study addresses two different objectives. The first objective is to recognize risk contagion between China carbon market and its infected markets in manner of high order moment attribute. The second objective is to forecast the China carbon price by the proposed high order risk contagion machine learning model (HOC-LSTM) to prove the risk contagion is useful to improve the forecasting performance. This chapter begins with main findings, followed by discussions on economic and policy implications especially analyzing the origin of carbon price forecasting issues, the measurement results. Furthermore, limitations of the study are discussed, and some future suggestions are presented. Finally, concluding with a summary of this study.

5.2 Main Findings

As an innovative mechanism to address global climate change, the creation of carbon markets is based on various agreements of the international community fulfills its emission reduction responsibilities, and develops into policy tools to promote the operation of finance markets. Compared to other financial markets, the carbon market is highly susceptible to major black swan events such as energy policies, emission reduction quota, global climate negotiations and financial crises in the capital market sector. However, because of low efficiency of carbon market that concluded in previous studies of Zhang et al.(2019) and Yun et al.(2023) , trading tools of the carbon market do not reflect all value information. Especially, under the promote of the irrational investors, there is a clear market asymmetry effect on carbon market. Therefore, carbon price forecasting research should incorporate the effect of the market asymmetric and event change into the forecasting framework.

Based on the identification results of the high order moment risk contagion and carbon price forecasting effect of proposed HOC-LSTM model, main findings of this study are as follows:

(1) Test risk contagion relationship from risk source carbon market to its infected markets in the manner of high order moment attribute.

Because of heterogeneity and asymmetry of carbon price fluctuations, this study mainly measures high order moment attribute risk contagion relationship between carbon market and infected markets under market rapid and slow trend. The results suggest that there is no risk contagion relationship in low order moment channels, but significant risk contagion relationship through high order moment attributes have detected, including the significant high order moment risk contagion channels of CS12 from carbon market to JMF market, the significant high order moment contagion channels of CS21 and CV22 from carbon market to Oil market, the significant high order moment risk contagion channels of CS21 from carbon market to CSI300 market, and the significant CV22 channel from carbon market to European carbon market. Those finding indicates that previous carbon price forecasting studies through low order moments attribute may be insufficient to accurately reveal the carbon premium formation. Incorporating risk contagion relationship between carbon markets and infected markets into price forecasting framework can help improve the forecasting accuracy (Kumar et al., 2003; Tjøstheim et al., 2013; Støve et al., 2014; Zhang et al., 2023).

(2) Construct a machine learning carbon price forecasting model that suitable for capturing the impact of high order moment risk contagion on carbon price.

Based on the tested high order moment channel risk contagion relationship between carbon and its infected markets. The high order moment CAPM financial asset price forecasting framework is extended from two factors to multiple factors, and then the high order risk contagion carbon price forecasting model is formed. This model not only reveals the high order moment impact relationship of infected markets on carbon price, but also reflects the nonlinear impact path of high order moment attribute risk contagion on carbon price (Fama & French, 2017; Mosoeu and Kodongo, 2022).

Constructing high order moment risk contagion carbon price forecasting model has characteristic of multiple parameters and complex nonlinear network structures, which require machine learning model to extract deep features and improve the fitting performance (Yun et al., 2020). Furthermore, LSTM model has advantage of processing financial time series and preventing gradient vanishing and exploding(Sun et al., 2020; Adekoya et al., 2021). This evidence indicates that HOC-LSTM model is suitable for capturing the impact of high order moment risk contagion on carbon price.

(3) The HOC-LSTM model performance better in forecasting China carbon price than other comparative models.

Based on identified high order moment attribute risk contagion relationship between carbon market and its infected markets, this study reconstructed carbon price forecasting model, and the machine learning HOC-LSTM model is carried out to forecast carbon price.

According to the empirical analysis in chapter 4, the results show that the forecasting error indicators RMSE, MAE, MAPE of the HOC-LSTM model are smaller than other comparative models, the Kendall correlation indicator is also high value in the entire experiment. Furthermore, sample adjustment testing and *W* testing are conducted to further demonstrate the robustness of the HOC-LSTM model. This evidence indicates that HOC-LSTM model can effectively fit and map complex nonlinear, non normal, and non-stationary China carbon price data that with impact of high order moment risk contagion, out of sample forecasting effect is satisfactory. This conclusion indirectly proves risk contagion relationships between carbon market and infected markets are also a key element for explaining carbon price formation mechanism. Furthermore, those finding not only convinced the forecasting ability of HOC-LSTM model in fitting the complex carbon price, but also give stronger evidence that the price forecasting framework integrates the high order moment risk contagion can provide a more accurate explanation of carbon premium.

5.3 Policy Implications

Under the background of the global capital flows and cross market allocation of carbon assets, this study considers the risk contagion among carbon market and its infected markets, studies carbon price forecasting issue under impact of high order moment risk contagion. Based on the research findings of this study, the following are several policy implications aimed at improving the operational efficiency and stability of the carbon market, and providing reference for future carbon price forecasting.

(1) Establish and improve a high order moment risk contagion monitoring mechanism between carbon market and infected markets

Policy makers should develop and implement effective risk management policies, monitor and respond to high order risk contagion phenomena in carbon market, and establish risk contagion monitoring mechanisms. By establishing a carbon market risk warning system, real-time monitoring of market dynamics, timely detection and response to high order moment risk contagion, corresponding policy measures can be taken in a timely manner to prevent market risks and stabilize carbon prices. In addition, the Chinese government should cooperate with the international community, share and exchange risk management experience, and jointly enhance the stability of the global carbon market.

Firstly, the government agencies should strengthen carbon market risk dynamic analysis of the fundamental prices and risk information, introduce market risk trigger warning mechanisms and crisis response mechanisms. Especially, effective real-time monitoring mechanisms for carbon market risks should be established to ensure the tracking and response to the risk contagion process caused by unexpected or policy events.

Secondly, government agencies should strengthen the market-oriented nature of carbon asset pricing and make institutional arrangements for serving carbon pricing. The establishment of the carbon market is policy dependent and dominant, while its operation process, especially the pricing mechanism, needs to be carried out in accordance with the market-oriented mechanisms. Therefore, focusing on government service measures for carbon pricing, some policies should continue to improve the institutional norms for the operation of the carbon market. Some feasible measures, such as strengthening the innovation of quota allocation system, information disclosure system, report review system, national carbon trading registration management system, reporting and verification system, and other systems. The effective legal regulations will ensure promotion of trading market, enhance ability to monitor and control risk and better serve the development of carbon trading.

(2) Build and promote a multi factor carbon price forecasting framework

Given the impact of high order moment risk contagion on carbon price, it is recommended that policy makers construct and promote a multi factor carbon price forecasting framework. This framework should incorporate the high order moment characteristics of carbon market and various factors, such as energy policies, emission reduction quotas, global climate negotiations, and financial crises. This will help to have a more comprehensive understanding of the price fluctuations in the carbon market, improve market predictability, and reduce market volatility caused by unexpected events.

(3) Promote the application of HOC-LSTM model in carbon price forecasting

This study indicates that the HOC-LSTM model has significant advantages in carbon price forecasting. The government can promote the promotion and application of models through policy guidance. Encourage and support the application of HOC-LSTM model in carbon price forecasting to improve the accuracy carbon price prediction. The government can organize training and promotion activities to introduce the advantages and application methods of the HOC-LSTM model to carbon market participants. At the same time, the actual effectiveness of the model can be verified through pilot projects, and comprehensive promotion can be carried out after success.

Incorporating high order moment risk contagion correlation among carbon market and infected markets into forecasting framework to enhance forecasting accuracy.This conclusion suggests that it is reasonable to explore carbon price formation mechanism from channel of high order moment risk contagion. When making the carbon market decisions, apart from focus on the systemic risks caused by fundamental factors, it is also necessary to analyze impact of investor irrationality and external policy events. Government agencies should pay more attention to the price driving relationship between carbon market and its infected markets, and provides a technical analysis basis for market traders to catch price trend and forecast price changes.

(4) Enhance the financial knowledge and risk management capabilities of market participants

The irrational investor behavior in the carbon market may lead to asymmetric effects in the market. Policy makers can reduce the negative impact of irrational behavior on the market by strengthening regulation and guidance. For example, through training and education, enhance market participants' understanding and risk management capabilities, strengthen their ability to cope with market fluctuations and risk contagion. The government can also cooperate with universities and financial institutions to regularly hold training courses and seminars to popularize carbon market knowledge and risk management techniques. Meanwhile, convenient learning channels can be provided to the public through online courses and open resources.

On the one hand, the entire process from carbon trading registration, trading condition review to transaction completion to real-time monitor the market information, to reveal the potential risk points that may cause significant anomalies. On the other hand, introducing the specialized investment management institutions into the carbon market to provide more professional services. With the improvement of carbon asset management efficiency, the activity and market liquidity of carbon trading products will also improve. These management measures have important significance for the emerging China carbon market.

5.4 Conclusion

The topic of this study is solving the price forecasting issues in China. The contribution of this study are designing a high order risk contagion carbon price forecasting theoretical framework, constructing a new non-parametric method for testing risk contagion relationship among carbon market and infected markets, constructing a high order moment risk contagion model (HOC-LSTM) to forecast the carbon price. After constructing the machine learning model, the conclusion found that the carbon price forecasting theoretical framework with high order moment risk contagion can effective improve the carbon forecasting accuracy. This indicates that theoretical framework of this study contributes to significantly improve carbon price forecasting accuracy.

5.5 Future Research

Firstly, using price data from national unified carbon market to forecast the China carbon price as the maturity and improvement of China carbon market.

Secondly, using data mining methods to obtain unstructured pricing factors such as policy factors, extreme market factors, and international negotiation factors, and incorporating these factors into price forecasting models to improve accuracy.

Thirdly, genetic algorithm and whale optimization algorithm can be used to optimize and improve the training process of HOC-LSTM model, so as to provide the model's out of sample generalization ability and enhance the accuracy.

Fourthly, the risk measurement involved in price forecasting and the interdependence between risks contagion need to be further optimized and expanded in future research. In addition, considering the significant volatility of the carbon market at different stages and its susceptibility to policy events, in the future, structural fracture event analysis methods can be considered to more profoundly reveal the event dynamics behind carbon price fluctuations.

5.6 Limitation of the Study

The study has several limitations that should be noted. Firstly, in terms of sample selection, this study conducts the price of Hubei carbon market as the dependent variable. In theory, for studying the issue of China carbon prices forecasting, it is better to use the national unified carbon market price that officially started operating in July 2021. This may limit the selection of representative indicators. This problem is related to the current lack of unified carbon market price data.

Secondly, in terms of the influencing factors of carbon price, although the study considers the high order moment contagion relationship between the carbon market and its infected market, there are still many unpredictable factors in the market, such as policy changes, changes in international relations, etc., which are difficult to fully incorporate into the model.

Thirdly, in the construction of the HOC-LSTM model, although our model has a relatively small forecasting errors and has advantages in forecasting China carbon price, the superiority is not significant. This may be related to the fact that the proposed model has more parameters and lacks improvement in optimization algorithms.

REFERENCES

Aatola, P., Ollikainen, M., & Toppinen, A. (2013). Price determination in the EU ETS market: Theory and econometric analysis with market fundamentals. *Energy Economics*, 36, 380-395.

Abbara, O., & Zevallos, M. (2014). Assessing stock market dependence and contagion. *Quantitative Finance*, 14(9), 1627-1641.

Acharya, V. V., & Pedersen, L. H. (2005). Asset pricing with liquidity risk. *Journal of Financial Economics*, 77(2), 375-410.

Alamili, M. (2011). Exchange rate prediction using support vector machines: a comparison with artificial neural networks. *Lap Lambert Academic Publishing*.

Alberola, E., Chevallier, J., & Chèze, B. (2008). Price drivers and structural breaks in European carbon prices 2005–2007. *Energy policy*, 36(2), 787-797.

Alberola, E., & Chevallier, J. (2009). European carbon prices and banking restrictions: Evidence from Phase I (2005-2007). *The Energy Journal*, 30(3).

Amaya, D., Christoffersen, P., Jacobs, K., & Vasquez, A. (2015). Does realized skewness predict the cross-section of equity returns?. *Journal of Financial Economics*, 118(1), 135-167.

Ang, A., Hodrick, R. J., Xing, Y., & Zhang, X. (2009). High idiosyncratic volatility and low returns: International and further US evidence. *Journal of Financial Economics*, 91(1), 1-23.

Ang, A., Hodrick, R. J., Xing, Y., & Zhang, X. (2006). The cross-section of volatility and expected returns. *The Journal of Finance*, 61(1), 259-299.

Apergis, N., Artikis, P., & Sorros, J. (2011). Asset pricing and foreign exchange risk. *Research in International Business and Finance*, 25(3), 308-328.

Arakelian, V., & Dellaportas, P. (2012). Contagion determination via copula and volatility threshold models. *Quantitative Finance*, 12(2), 295-310.

Asparouhova, E., Bessembinder, H., & Kalcheva, I. (2010). Liquidity biases in asset pricing tests. *Journal of Financial Economics*, 96(2), 215-237.

Atsalakis, G. S., & Valavanis, K. P. (2009). Surveying stock market forecasting techniques–Part II: Soft computing methods. *Expert Systems with Applications*, 36(3), 5932-5941.

Atsalakis, G. S. (2016). Using computational intelligence to forecast carbon prices. *Applied Soft Computing*, 43, 107-116.

Adekoya, O. B., Oliyide, J. A., & Noman, A. (2021). The volatility connectedness of the EU carbon market with commodity and financial markets in time-and frequency-domain: The role of the US economic policy uncertainty. *Resources Policy*, 74, 102252.

Arashi, M., & Rounaghi, M. M. (2022). Analysis of market efficiency and fractal feature of NASDAQ stock exchange: Time series modeling and forecasting of stock index using ARMA-GARCH model. *Future Business Journal*, 8(1), 14.

Alodat, A. Y., Al Amosh, H., Khatib, S. F., & Mansour, M. (2023). Audit committee chair effectiveness and firm performance: The mediating role of sustainability disclosure. *Cogent Business & Management*, 10(1), 2181156.

Ashfaq, S., Liangrong, S., Waqas, F., Gulzar, S., Mujtaba, G., & Nasir, R. M. (2024). Renewable energy and green economic growth nexus: insights from simulated dynamic ARDL. *Gondwana Research*, 127, 288-300.

Allen, F., Qian, J., Shan, C., & Zhu, J. L. (2024). Dissecting the long-term performance of the Chinese stock market. *The Journal of Finance*, 79(2), 993-1054.

Bali, T. G., & Cakici, N. (2010). World market risk, country-specific risk and expected returns in international stock markets. *Journal of Banking & Finance*, 34(6), 1152-1165.

Barberis, N.,& Xiong, W. (2009). What Drives the Disposition Effect? An Analysis of a Long-Standing Preference-Based Explanation. *The Journal of Finance*, 64(2),751-784.

Barberis, N., Greenwood, R., Jin, L., & Shleifer, A. (2018). Extrapolation and bubbles. *Journal of Financial Economics*, 129(2), 203-227.

Barberis, N., Mukherjee, A., & Wang, B. (2016). Prospect theory and stock returns: An empirical test. *The review of financial studies*, 29(11), 3068-3107.

Barberis, N., & Thaler, R. (2003). A survey of behavioral finance. *Handbook of the Economics of Finance*, 1, 1053-1128.

Barillas, F., & Shanken, J. (2018). Comparing asset pricing models. *The Journal of Finance*, 73(2), 715-754.

Bastianin, A. (2020). Robust measures of skewness and kurtosis for macroeconomic and financial time series. *Applied Economics*, 52(7), 637-670.

Benz, E., & Trück, S. (2009). Modeling the price dynamics of CO2 emission allowances. *Energy Economics*, 31(1), 4-15.

Biscarri, J. G., & Espinosa, G. L. (2008). The influence of differences in accounting standards on empirical pricing models: An application to the Fama–French model. *Journal of Multinational Financial Management*, 18(4), 369-388.

Black, F. (1972). Capital market equilibrium with restricted borrowing. *The Journal of Business*, 45(3), 444-455.

Black, F., & Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3), 637-654.

Baillie, R. T., Bollerslev, T., & Mikkelsen, H. O. (1996). Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 74(1), 3-30.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. Journal of econometrics, 31(3), 307-327.

Boyer, B. H., Gibson, M. S., & Loretan, M. (1999). Pitfalls in tests for changes in correlations. *International Finance Discussion Papers, Washington, DC: Board of Governors of the Federal Reserve System,(597)*.

Boyer, B. H., Kumagai, T., & Yuan, K. (2006). How do crises spread? Evidence from accessible and inaccessible stock indices. *The Journal of Finance*, 61(2), 957-1003.

Boyer, B., Mitton, T., & Vorkink, K. (2010). Expected idiosyncratic skewness. *The Review of Financial Studies*, 23(1), 169-202.

Breeden, D. T., Gibbons, M. R., & Litzenberger, R. H. (1989). Empirical tests of the consumption-oriented CAPM. *The Journal of Finance*, 44(2), 231-262.

Buckle, M., Chen, J., & Williams, J. M. (2016). Realised higher moments: theory and practice. *The European Journal of Finance*, 22(13), 1272-1291.

Byun, S. J., & Cho, H. (2013). Forecasting carbon futures volatility using GARCH models with energy volatilities. *Energy Economics*, 40, 207-221.

Büchner, M., & Kelly, B. (2022). A factor model for option returns. *Journal of*

Financial Economics, 143(3), 1140-1161.

Calvo, G. A., & Mendoza, E. G. (2000). Rational contagion and the globalization of securities markets. *Journal of International Economics*, 51(1), 79-113.

Cao, D. Z., Pang, S. L., & Bai, Y. H. (2005, August). Forecasting exchange rate using support vector machines. *In 2005 international conference on machine learning and cybernetics (Vol. 6, pp. 3448-3452). IEEE*.

Cao, W., Zhu, W., Wang, W., Demazeau, Y., & Zhang, C. (2020). A deep coupled LSTM approach for USD/CNY exchange rate forecasting. *IEEE Intelligent Systems*, 35(2), 43-53.

Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of Finance*, 52(1), 57-82.

Chanˆ, J. C., Fry-McKibbin, R. A., & Hsiao, C. Y. L. (2018). A regime switching skew-normal model of contagion. *Studies in Nonlinear Dynamics & Econometrics*, 23(1), 1-24.

Chevallier, J. (2009). Carbon futures and macroeconomic risk factors: A view from the EU ETS. *Energy Economics*, 31(4), 614-625.

Chevallier, J. (2011). Detecting instability in the volatility of carbon prices. *Energy Economics*, 33(1), 99-110.

Chevallier, J. (2010). EUAs and CERs: Vector autoregression, impulse response function and cointegration analysis. *Economics Bulletin*, 30(1), 558-576.

Chevallier, J. (2010). Volatility forecasting of carbon prices using factor models. *Economics Bulletin*, 30(2), 1642-1660.

Chevallier, J. (2012). Global imbalances, cross-market linkages, and the financial crisis: A multivariate Markov-switching analysis. *Economic Modelling*, 29(3), 943-973.

Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning Phrase Representations using RNN Encoder–Decoder for Statistical Machine Translation. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP).*

Chollete, L., Heinen, A., & Valdesogo, A. (2009). Modeling international financial returns with a multivariate regime-switching copula. *Journal of Financial Econometrics*, 7(4), 437-480.

Chang, K., Ye, Z., & Wang, W. (2019). Volatility spillover effect and dynamic correlation between regional emissions allowances and fossil energy markets: New evidence from China's emissions trading scheme pilots. *Energy*, 185, 1314-1324.

Christoffersen, P., Fournier, M., Jacobs, K., & Karoui, M. (2021). Option-based estimation of the price of coskewness and cokurtosis risk. *Journal of Financial and Quantitative Analysis*, 56(1), 65-91.

Chung, Y. P., Johnson, H., & Schill, M. J. (2006). Asset pricing when returns are nonnormal: Fama-french factors versus higher-order systematic comoments. *The Journal of Business*, 79(2), 923-940.

Conrad, J., Dittmar, R. F., & Ghysels, E. (2013). Ex ante skewness and expected stock returns. *The Journal of Finance*, 68(1), 85-124.

Convery, F. J., & Redmond, L. (2007). Market and price developments in the European Union emissions trading scheme. *Review of Environmental Economics & Policy*, 1(1):88-111.

Chen, J. L., Glabadanidis, P., & Sun, M. (2022). The five-factor asset pricing model, short-term reversal, and ownership structure–the case of China. *International Review of Financial Analysis*, 82, 102147.

Chen, J., Taylor, N., Yang, S., & Han, Q. (2022). Hawkes processes in finance: market structure and impact. *The European Journal of Finance*, 28(7), 621-626.

Cheng, D., Liao, Y., & Pan, Z. (2023). The geopolitical risk premium in the commodity futures market. *Journal of Futures Markets*, 43(8), 1069-1090.

China Emissions Trading Network (2023).http://m.tanpaifang.com/article/93601.html

China Emissions Trading Network (2024). http://www.tanpaifang.com/2024/0402/105476_2.html

Daskalakis, G., Psychoyios, D., & Markellos, R. N. (2009). Modeling CO2 emission allowance prices and derivatives: Evidence from the European trading scheme. Journal of Banking & Finance, 33(7), 1230-1241.

DeLong, J. B. (2022). Slouching Towards Utopia: An Economic History of the Twentieth Century. New York: Basic Books.

De Oliveira, F. A., Nobre, C. N., & Zárate, L. E. (2013). Applying Artificial Neural Networks to prediction of stock price and improvement of the directional prediction index–Case study of PETR4, Petrobras, Brazil. *Expert Systems With Applications*, 40(18), 7596-7606.

Dornbusch, R., Park, Y. C., & Claessens, S. (2000). Contagion: Understanding how it spreads. *The World Bank Research Observer*, 15(2), 177-197.

European Energy Exchange (2023). https://www.eex.com/en/market-data/environmentals/futures

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50(4),987-1007.

Eichengreen, B., Rose, A. K., & Wyplosz, C. (1996). Contagious currency crises.*Scandinavian Journal of Economics*, 98(4): 463-484.

Eikseth, H. M., & Lindset, S. (2009). A note on capital asset pricing and heterogeneous taxes. *Journal of Banking & Finance*, 33(3), 573-577.

Ellerman, A. D., & Buchner, B. K. (2007). The European Union emissions trading scheme: origins, allocation, and early results. *Review of Environmental Economics And Policy*, 1(1): 66-87.

Elsayed, A. H., Naifar, N., Nasreen, S., & Tiwari, A. K. (2022). Dependence structure and dynamic connectedness between green bonds and financial markets: Fresh insights from time-frequency analysis before and during COVID-19 pandemic. *Energy Economics*, 107, 105842.

Fama, E. F. (1965). Random walks in stock market prices. *Financial Analysts Journal*, 51(1): 75-80.

Fama, E. F. (1965).The behavior of stock-market prices. *The Journal of Business*, 38(1): 34-105.

Fama, E. F. (1970). Efficient market hypothesis: A review of theory and empirical work. *Journal of Finance*, 25(2), 28-30.

Fama, E. F.(1991). Efficient capital markets: II. T*he Journal of Finance*, 46(5): 1575-1617.

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33(1), 3-56.

Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, 116(1), 1-22.

Fama, E. F., & French, K. R. (2017). International tests of a five-factor asset pricing model. J*ournal of Financial Economics*, 123(3), 441-463.

Fama, E. F., & MacBeth, J. D. (1973). Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy*, 81(3), 607-636.

Fama, E. F., & Schwert, G. W. (1977). Asset returns and inflation. *Journal of Financial Economics*, 5(2), 115-146.

Fan, X., Li, S., & Tian, L. (2015). Chaotic characteristic identification for carbon price and an multi-layer perceptron network prediction model. *Expert Systems with Applications*, 42(8), 3945-3952.

Fang, H., & Lai, T. Y. (1997). Co-kurtosis and capital asset pricing. *Financial Review*, 32(2), 293-307.

Forbes, K. J., & Rigobon, R. (2002). No contagion, only interdependence: measuring stock market comovements. *The Journal of Finance*, 57(5), 2223-2261.

Fry-McKibbin, R., & Hsiao, C. Y. L. (2018). Extremal dependence tests for contagion. *Econometric Reviews*, 37(6), 626-649.

Fry-McKibbin, R., Hsiao, C. Y. L., & Tang, C. (2014). Contagion and global financial crises: Lessons from nine crisis episodes. *Open Economies Review*, 25, 521-570.

Fry, R., Martin, V. L., & Tang, C. (2010). A new class of tests of contagion with applications. *Journal of Business & Economic Statistics*, 28(3), 423-437.

Fry-McKibbin, R., Martin, V. L., & Tang, C. (2014). Financial contagion and asset pricing. *Journal of Banking & Finance*, 47, 296-308.

Fan, X., Lv, X., Yin, J., Tian, L., & Liang, J. (2019). Multifractality and market efficiency of carbon emission trading market: Analysis using the multifractal detrended fluctuation technique. *Applied Energy*, 251, 113333.

Fang, T., Miao, D., Su, Z., & Yin, L. (2023). Uncertainty-driven oil volatility risk premium and international stock market volatility forecasting. *Journal of Forecasting*, 42(4), 872-904.

Fritz, M., Gries, T., & Wiechers, L. (2024). An early indicator for anomalous stock market performance. *Quantitative Finance*, 24(1), 105-118.

Gong, D. S., Che, J. X., Wang, J. Z., & Liang, J. Z. (2008, December). Short-term electricity price forecasting based on novel SVM using artificial fish swarm algorithm under deregulated power. *In 2008 Second International Symposium on Intelligent Information Technology Application. IEEE*.

Griffin, J. M. (2002). Are the Fama and French factors global or country specific?. *The Review of Financial Studies*, 15(3), 783-803.

Granger, C. W. J., & Newbold, P. (2004). Forecasting economic time series. Academic press.

Guo, B., Zhang, W., Zhang, Y., & Zhang, H. (2017). The five-factor asset pricing model tests for the Chinese stock market. *Pacific-Basin Finance Journal*, 43, 84-106.

Gong, X., Shi, R., Xu, J., & Lin, B. (2021). Analyzing spillover effects between carbon and fossil energy markets from a time-varying perspective. *Applied Energy*, 285, 116384.

Guo, H., Xi, Y., Yu, F., & Sui, C. (2024). Time–frequency domain based optimization of hedging strategy: Evidence from CSI 500 spot and futures. *Expert Systems with Applications*, 238, 121785.

Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica: Journal of the Econometric*

Society, 357-384.

Hammoudeh, S., Lahiani, A., Nguyen, D. K., & Sousa, R. M. (2015). An empirical analysis of energy cost pass-through to CO2 emission prices. *Energy Economics*, 49, 149-156.

Hammoudeh, S., Nguyen, D. K., & Sousa, R. M. (2014). Energy prices and CO2 emission allowance prices: A quantile regression approach. *Energy policy*, 70, 201-206.

Han, M., Ding, L., Zhao, X., & Kang, W. (2019). Forecasting carbon prices in the Shenzhen market, China: The role of mixed-frequency factors. *Energy*, 171, 69-76.

Han, S. K., Ahn, J. J., Oh, K. J., & Kim, T. Y. (2015). A new methodology for carbon price forecasting in EU ETS. *Expert Systems*, 32(2), 228-243.

Hao, Y., Tian, C., & Wu, C. (2020). Modelling of carbon price in two real carbon trading markets. *Journal of Cleaner Production*, 244, 118556.

Harvey, C. R., & Siddique, A. (2000). Time-varying conditional skewness and the market risk premium. *Research in Banking and Finance*, 1(1), 27-60.

Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The Review of Financial Studies*, 6(2), 327-343.

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, 9(8), 1735-1780.

Huang, Y., Dai, X., Wang, Q., & Zhou, D. (2021). A hybrid model for carbon price forecasting using GARCH and long short-term memory network. *Applied Energy*, 285, 116485.

Huang, Y., & He, Z. (2020). Carbon price forecasting with optimization prediction method based on unstructured combination. Science of the Total Environment, 725, 138350.

Hu, X., Yang, Z., Sun, J., & Zhang, Y. (2020). Carbon tax or cap-and-trade: Which is more viable for Chinese remanufacturing industry?. *Journal of Cleaner Production*, 243, 118606.

Hong, Y., & Jin, X. (2022). Pricing of variance swap rates and investment decisions of variance swaps: Evidence from a three-factor model. *European Journal of Operational Research*, 303(2), 975-985.

Hwang, S., & Satchell, S. E. (1999). Modelling emerging market risk premia using higher moments. *International Journal of Finance & Economics*, 4(4), 271-296.

Ingersoll, J. E., & Jin, L. J. (2013). Realization utility with reference-dependent preferences. *The Review of Financial Studies*, 26(3), 723-767.

Ingersoll, J. (1975). Multidimensional security pricing. *Journal of Financial And Quantitative Analysis*, 10(5), 785-798.

Investing App (2024). https://cn.investing.com/commodities/energy

Ji, Q., Xia, T., Liu, F., & Xu, J. H. (2019). The information spillover between carbon price and power sector returns: Evidence from the major European electricity companies. *Journal of Cleaner Production*, 208, 1178-1187.

Ji, Q., Zhang, D., & Geng, J. B. (2018). Information linkage, dynamic spillovers in prices and volatility between the carbon and energy markets. *Journal of Cleaner Production*, 198, 972-978.

Ji, Z., Niu, D., Li, M., Li, W., Sun, L., & Zhu, Y. (2022). A three-stage framework for vertical carbon price interval forecast based on decomposition–integration method. *Applied Soft Computing*, 116, 108204.

Jiang, T., Zhou, X. H.,& Dong, Y,W. (2013). Stock market cycle fluctuation in China: Markov regime switching model. *Systems Engineering-Theory & Practice*, 33(8): 1934-1939.

Jiang, W., Liu, Y., & Xue, W. (2023). How do fossil energy markets and carbon markets affect each other? Based on the grainger causal relationship on quantiles. *Journal of Cleaner Production*, 423, 138799.

Jiang, Y. (2014). Testing asset pricing models under non-linear assumptions: evidence from UK firm level panel data *(Doctoral dissertation, University of Birmingham)*.

Jie, D., Xu, X., & Guo, F. (2021). The future of coal supply in China based on non-fossil energy development and carbon price strategies. *Energy*, 220, 119-144.

Kahneman, D., & Tversky, A.(1979) . Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, 47(2):263-292.

Karolyi, G. A. (2003). Does international financial contagion really exist?. *International Finance*, 6(2), 179-199.

Kim, H. S., & Koo, W. W. (2010). Factors affecting the carbon allowance market in the US. *Energy Policy*, 38(4), 1879-1884.

Kim, S. J., & Boyd, S. (2008). A minimax theorem with applications to machine learning, signal processing, and finance. *SIAM Journal on Optimization*, 19(3), 1344-1367.

Kim, S., Lee, G., & Park, Y. J. (2017). Skewness versus kurtosis: Implications for pricing and hedging options. *Asia-Pacific Journal of Financial Studies*, 46(6), 903-933.

Kimball, M. S.(1990). Precautionary Saving in the Small and in the Large. *Econometrica*, 58(1):53-73.

King, M. A., & Wadhwani, S. (1990). Transmission of volatility between stock markets. *The Review of Financial Studies*, 3(1), 5-33.

Kodres, L. E., & Pritsker, M. (2002). A rational expectations model of financial contagion. *The Journal of Finance*, 57(2), 769-799.

Koop, G., & Tole, L. (2013). Forecasting the European carbon market. *Journal of the Royal Statistical Society Series A: Statistics in Society*, 176(3), 723-741.

Kraus, A., & Litzenberger, R. H. (1976). Skewness preference and the valuation of risk assets. *The Journal of Finance*, 31(4), 1085-1100.

Kumar, M., Moorthy, U., & Perraudin, W. (2003). Predicting emerging market currency crashes. *Journal of Empirical Finance*, 10(4), 427-454.

Kim, H. S., & Koo, W. W. (2010). Factors affecting the carbon allowance market in the US. *Energy Policy*, 38(4), 1879-1884.

Kumar, S., Managi, S., & Matsuda, A. (2012). Stock prices of clean energy firms, oil and carbon markets: A vector autoregressive analysis. *Energy Economics*, 34(1), 215-226.

Kostin, K. B., Runge, P., & Charifzadeh, M. (2022). An analysis and comparison of multi-factor asset pricing model performance during pandemic situations in developed and emerging markets. *Mathematics*, 10(1), 142.

Kong, F., Song, J., & Yang, Z. (2022). A novel short-term carbon emission prediction model based on secondary decomposition method and long short-term memory network. *Environmental Science and Pollution Research*, 29(43), 64983-64998.

Liew, K. S. (2006). Forecasting Performance of Exponential Smooth Transition

Autoregressive Exchange Rate Models. *Open Economies Review*,17(7),261-277.

Lam, K. S., & Tam, L. H. (2011). Liquidity and asset pricing: Evidence from the Hong Kong stock market. *Journal of Banking & Finance*, 35(9), 2217-2230.

Le Roux, N., & Bengio, Y. (2008). Representational power of restricted Boltzmann machines and deep belief networks. *Neural computation*, 20(6), 1631-1649.

Lin, Y., Lehnert, T., & Wolff, C. (2019). Skewness risk premium: Theory and empirical evidence. *International Review of Financial Analysis*, 63, 174-185.

Loretan, M., & English, W. B. (2000). Evaluating correlation breakdowns during periods of market volatility. *Board of Governors of the Federal Reserve System International Finance Working Paper*, 2000 (658).

Lu, W., & Wang, W. J. (2009). Dynamic correlation between carbon market and Chinese stock market based on AGDCC-GARCH. *2009 International Conference on Management Science and Engineering. IEEE*, 779-784.

Li, N., &Venus, K.S.L.(2024). Carbon emission price forecasting in China using a novel secondary decomposition hybrid model of CEEMD-SE-VMD-LSTM. *Systems Science and Control Engineering*, 12(1), 2291409.

Liu, Z., & Huang, S. (2021). Carbon option price forecasting based on modified fractional Brownian motion optimized by GARCH model in carbon emission trading. *The North American Journal of Economics and Finance*, 55, 101307.

Lin, B., & Jia, Z. (2019). Impacts of carbon price level in carbon emission trading market. *Applied energy*, 239, 157-170.

Liu, J., Jiang, T., & Ye, Z. (2021). Information efficiency research of China's carbon markets. *Finance Research Letters*, 38, 101444.

Li, H., Jin, F., Sun, S., & Li, Y. (2021). A new secondary decomposition ensemble learning approach for carbon price forecasting. Knowledge-Based Systems, 214, 106686.

Li, G., Zheng, C., & Yang, H. (2022). Carbon price combination prediction model based on improved variational mode decomposition. *Energy Reports*, 8, 1644-1664.

Liu, J., Wang, P., Chen, H., & Zhu, J. (2022). A combination forecasting model based on hybrid interval multi-scale decomposition: Application to interval-valued carbon price forecasting. *Expert Systems with Applications*, 191, 116267.

Li, J., & Liu, D. (2023). Carbon price forecasting based on secondary decomposition and feature screening. Energy, 278, 127783.

Lian, Y. M., & Chen, J. H. (2022). Foreign exchange option pricing under regime switching with asymmetrical jumps. *Finance Research Letters*, 46, 102294.

Luo, C., Qu, Y., Su, Y., & Dong, L. (2024). Risk spillover from international crude oil markets to China's financial markets: Evidence from extreme events and US monetary policy. *The North American Journal of Economics and Finance*, 70, 102041.

Markowitz, H.(1952).Portfolio Selection. Journal of Finance, 7(1):77-91.

Masson, P.(1999). Contagion: macroeconomic models with multiple equilateral. *Journal of International Money and Finance*, 18(4): 587-602.

Mayers, D. (1973). Nonmarketable assets and the determination of capital asset prices in the absence of a riskless asset. *The Journal of Business*, 46(2), 258-267.

Merton, R. C.(1987). A Simple Model of Capital Market Equilibrium with Incomplete

Information. *The Journal of Finance*, 42(3):483-510.

Mitton, T., & Vorkink, K. (2007). Equilibrium underdiversification and the preference for skewness. *The Review of Financial Studies*, 20(4), 1255-1288.

Moerman, G. A., & Van Dijk, M. A. (2010). Inflation risk and international asset returns. *Journal of Banking & Finance,* 34(4), 840-855.

Moerman, G. A. (2005). How domestic is the Fama and French three-factor model? An application to the Euro area. *ERIM Report Series Reference No. ERS-2005-035-F&A*.

Montagnoli, A., & De Vries, F. P. (2010). Carbon trading thickness and market efficiency. *Energy Economics*, 32(6), 1331-1336.

Mosoeu, S., & Kodongo, O. (2022). The Fama-French five-factor model and emerging market equity returns. *The Quarterly Review of Economics and Finance*, 85, 55-76.

Markowski, L. (2024). Conventional and downside CAPM with higher-order moments: Evidence from emerging markets. Equilibrium. *Quarterly Journal of Economics and Economic Policy*, 19(1), 93-138.

National Bureau of Statistics of China Government. (2023). https://www.stats.gov.cn/sj/ndsj/2023/indexch.htm

Nazifi, F., & Milunovich, G. (2010). Measuring the impact of carbon allowance trading on energy prices. *Energy & Environment*, 21(5), 367-383.

Nazifi, F. (2013). Modelling the price spread between EUA and CER carbon prices. *Energy Policy*, 56, 434-445.

Oberndorfer, U. (2009). EU emission allowances and the stock market: evidence from the electricity industry. *Ecological Economics*, 68(4), 1116-1126.

Özorhan, M. O., Toroslu, İ. H., & Şehitoğlu, O. T. (2017). A strength-biased prediction model for forecasting exchange rates using support vector machines and genetic algorithms. *Soft Computing*, 21, 6653-6671.

Oertel, C., Kovaleva, E., Gleißner, W., & Bienert, S. (2022). Stochastic framework for carbon price risk estimation of real estate: a Markov switching GARCH simulation approach. *Journal of Property Investment & Finance*, 40(4), 381-397.

Palao, F., & Pardo, A. (2012). Assessing price clustering in European carbon markets. *Applied Energy*, 92, 51-56.

Pástor, Ľ., & Stambaugh, R. F. (2003). Liquidity risk and expected stock returns. *Journal of Political Economy*, 111(3), 642-685.

Patton, A. J. (2006). Estimation of multivariate models for time series of possibly different lengths. *Journal of Applied Econometrics*, 21(2), 147-173.

Peace, J., & Juliani, T. (2009). The coming carbon market and its impact on the American economy. *Policy and Society*, 27(4), 305-316.

Plakandaras, V., Gogas, P., & Papadimitriou, T. (2011). Directional Forecasting in financial time series using Support Vector machines: the USD/Euro exchange rate. *Conference of the Financial Engineering & Banking Society*,2011.

Qin, Q., He, H., Li, L., & He, L. Y. (2020). A novel decomposition-ensemble based carbon price forecasting model integrated with local polynomial prediction. *Computational Economics*, 55, 1249-1273.

Qiu, L., Chu, L., Zhou, R., Xu, H., & Yuan, S. (2023). How do carbon, stock, and renewable energy markets interact: Evidence from Europe. *Journal of Cleaner Production*, 407, 137106.

Reboredo, J. C. (2014). Volatility spillovers between the oil market and the European Union carbon emission market. *Economic Modelling*, 36, 229-234.

Ripley, D. M. (1973). Systematic elements in the linkage of national stock market indices. *The Review of Economics and Statistics*, 55(3),356-361.

Ross, S. A.(1976). The arbitrage theory of capital asset pricing. *Journal of Economic Theory*, 13(3):341-360.

Roy, R., & Shijin, S. (2018). A six-factor asset pricing model. *Borsa Istanbul Review*, 18(3), 205-217.

Rubin, J. D. (1996). A model of intertemporal emission trading, banking, and borrowing. *Journal of Environmental Economics And Management*, 31(3), 269-286.

Ren, X., & Zhu, L. (2020). Influence of allowance allocation events on prices in China's carbon market pilots–an AR-GARCH-based analysis. *Energy Sources, Part B: Economics, Planning, and Policy*, 15(3), 157-171.

Sajjad, R., Hedayati, S. H., Hedayati, S. H.(2013). Comparative between SV (stochastic volatility) & GARCH models by VaR. *Financial Engineering And Securities Management*, 4(15):79-97.

Sanin, M. E., Violante, F., & Mansanet-Bataller, M. (2015). Understanding volatility dynamics in the EU-ETS market. *Energy Policy*, 82, 321-331.

Sanin, M. E., & Violante, F. (2009). Understanding volatility dynamics in the EU-ETS market: lessons from the future. *Université catholique de Louvain, Center for Operations Research and Econometrics (CORE)*, 2009.

Segnon, M., Lux, T., & Gupta, R. (2017). Modeling and forecasting the volatility of carbon dioxide emission allowance prices: A review and comparison of modern volatility models. *Renewable and Sustainable Energy Reviews*, 69, 692-704.

Seifert, J., Uhrig-Homburg, M., & Wagner, M. (2008). Dynamic behavior of CO2 spot prices. *Journal of Environmental Economics and Management*, 56(2), 180-194.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance*, 19(3), 425-442.

Shen, F., Chao, J., & Zhao, J. (2015). Forecasting exchange rate using deep belief networks and conjugate gradient method. *Neurocomputing*, 167, 243-253.

Siddique, H. A.(2000). Conditional Skewness in Asset Pricing Tests. *The Journal of Finance*, 55(3):1263-1295.

Singhal, D., & Swarup, K. S. (2011). Electricity price forecasting using artificial neural networks. *International Journal of Electrical Power & Energy Systems*, 33(3), 550-555.

Smith, D. R. (2007). Conditional coskewness and asset pricing. *Journal of Empirical Finance*, 14(1), 91-119.

Støve, B., Tjøstheim, D., & Hufthammer, K. O. (2014). Using local Gaussian correlation in a nonlinear re-examination of financial contagion. *Journal of Empirical Finance*, 25, 62-82.

Sun, G., Chen, T., Wei, Z., Sun, Y., Zang, H., & Chen, S. (2016). A carbon price forecasting model based on variational mode decomposition and spiking neural networks. *Energies*, 9(1), 54.

Sun, W., & Duan, M. (2019). Analysis and forecasting of the carbon price in China's regional carbon markets based on fast ensemble empirical mode decomposition, phase space reconstruction, and an improved extreme learning machine. *Energies*, 12(2), 277.

Szkuta, B. R., Sanabria, L. A., & Dillon, T. S. (1999). Electricity price short-term forecasting using artificial neural networks. *IEEE transactions on power systems*, 14(3), 851-857.

Sun, R., Wang, G., Zhang, W., Hsu, L. T., & Ochieng, W. Y. (2020). A gradient boosting decision tree based GPS signal reception classification algorithm. *Applied Soft Computing*, 86, 105942.

Sun, W., & Huang, C. (2020). A novel carbon price prediction model combines the secondary decomposition algorithm and the long short-term memory network. *Energy*, 207, 118294.

Sun, W., & Wang, Y. (2020). Factor analysis and carbon price prediction based on empirical mode decomposition and least squares support vector machine optimized by improved particle swarm optimization. Carbon Management, 11(3), 315-329.

Sakariyahu, R., Johan, S., Lawal, R., Paterson, A., & Chatzivgeri, E. (2023). Dynamic connectedness between investors' sentiment and asset prices: A comparison between major markets in Europe and USA. *Journal of International Financial Markets, Institutions and Money*, 89, 101866.

Shafiullah, M., Senthilkumar, A., Lucey, B. M., & Naeem, M. A. (2024). Deciphering asymmetric spillovers in US industries: Insights from higher-order moments. *Research in International Business and Finance*, 70, 102313.

Taschini, L., & Paolella, M. S. (2008). An econometric analysis of emission trading allowances. *Journal of Banking and Finance*, 32(10), 6-26.

Tjøstheim, D., & Hufthammer, K. O. (2013). Local Gaussian correlation: A new measure of dependence. *Journal of Econometrics*, 172(1), 33-48.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science,* 211(4481), 453-458.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and uncertainty*, 5, 297-323.

Trakarnsirinont, W., Jitaree, W., & Buachoom, W. W. (2023). Political Uncertainty and Financial Firm Performance: Evidence from the Thai Economy as an Emerging Market in Asia. *Economies*, 11(1), 18.

Tsai, I. C. (2024). Fossil energy risk exposure of the UK electricity system: The moderating role of electricity generation mix and energy source. *Energy Policy*, 188, 114065.

Tronzano, M. (2024). What Drives Asset Returns Comovements? Some Empirical Evidence from US Dollar and Global Stock Returns (2000–2023). *Journal of Risk and Financial Management*, 17(4), 167.

Uddin, G. S., Hernandez, J. A., Shahzad, S. J. H., & Hedström, A. (2018). Multivariate dependence and spillover effects across energy commodities and diversification potentials of carbon assets. *Energy Economics*, 71, 35-46.

Ugurlu, U., Oksuz, I., & Tas, O. (2018). Electricity price forecasting using recurrent neural networks. *Energies*, 11(5), 1255.

Wang, Y., Wong, H. Y. (2017). Forecast Under Different Volatility Specifications. *Asia-Pacific Financial Markets*, 24(2):131-148.

Wang, C., Wang, Z., Ke, R. Y., & Wang, J. (2018). Integrated impact of the carbon quota constraints on enterprises within supply chain: direct cost and indirect cost. *Renewable and Sustainable Energy Reviews*, 92, 774-783.

Wang, J., Sun, X., Cheng, Q., & Cui, Q. (2021). An innovative random forest-based nonlinear ensemble paradigm of improved feature extraction and deep learning for carbon price forecasting. *Science of the Total Environment*, 762, 143099.

Wang, J., Jiang, W., Yan, Y., Shao, W., Wu, X., & Hua, Z. (2023). Exploring the Asymmetric Multifractal Characteristics of Price–Volume Cross-Correlation in the Chinese Rebar Futures Market Based on MF-ADCCA. *Fluctuation and Noise Letters*, 22(03), 2350029.

Wang, J., Dai, P. F., Chen, X. H., & Nguyen, D. K. (2024). Examining the linkage between economic policy uncertainty, coal price, and carbon pricing in China: Evidence from pilot carbon markets. *Journal of Environmental Management*, 352, 120003.

Wind.(2023). Wind Database.https://www.wind.com.cn/

Xiong, S., Wang, C., Fang, Z., & Ma, D. (2019). Multi-step-ahead carbon price forecasting based on variational mode decomposition and fast multi-output relevance vector regression optimized by the multi-objective whale optimization algorithm. *Energies*, 12(1), 147.

Xu, H., Pan, X., Li, J., Feng, S., & Guo, S. (2023). Comparing the impacts of carbon tax and carbon emission trading, which regulation is more effective?. *Journal of Environmental Management*, 330, 117156.

Yahşi, M., Çanakoğlu, E., & Ağralı, S. (2019). Carbon price forecasting models based on big data analytics. *Carbon Management*, 10(2), 175-187.

Yamin, H. Y., Shahidehpour, S. M., & Li, Z. (2004). Adaptive short-term electricity price forecasting using artificial neural networks in the restructured power markets. *International Journal of Electrical Power & Energy Systems*, 26(8), 571-581.

Yang, H., Yang, X., & Li, G. (2023). Forecasting carbon price in China using a novel hybrid model based on secondary decomposition, multi-complexity and error correction. *Journal of Cleaner Production*, 401, 136701.

Yang, X., Zhang, C., Yang, Y., Wu, Y., Yun, P., & Wagan, Z. A. (2020). China's Carbon Pricing Based on Heterogeneous Tail Distribution. *Sustainability*, 12(7), 2754.

Yu, L., Li, J., & Tang, L. (2015). Dynamic volatility spillover effect analysis between carbon market and crude oil market: a DCC-ICSS approach. International *Journal of Global Energy Issues*, 38(4-6), 242-256.

Yuan, Y. (2013). Forecasting the movement direction of exchange rate with polynomial smooth support vector machine. *Mathematical and Computer Modelling*, 57(3-4), 932-944.

Yun, P., Zhang, C., Wu, Y., Yang, X., & Wagan, Z. A. (2020). A novel extended higher-order moment multi-factor framework for forecasting the carbon price: Testing on the multilayer long short-term memory network. *Sustainability*, 12(5), 1869.

Yun, P., Huang, X., Wu, Y., & Yang, X. (2023). Forecasting carbon dioxide emission price using a novel mode decomposition machine learning hybrid model of CEEMDAN-LSTM. *Energy Science & Engineering*, 11(1), 79-96.

Zeitlberger, A. C., & Brauneis, A. (2016). Modeling carbon spot and futures price returns with GARCH and Markov switching GARCH models: Evidence from the first commitment period (2008–2012). *Central European Journal of*

Operations Research, 24, 149-176.

Zeng, S., Jia, J., Su, B., Jiang, C., & Zeng, G. (2021). The volatility spillover effect of the European Union (EU) carbon financial market. *Journal of Cleaner Production*, 282, 124394.

Zerbib, O. D. (2022). A sustainable capital asset pricing model (S-CAPM): Evidence from environmental integration and sin stock exclusion. *Review of Finance*, 26(6), 1345-1388.

Zhang, C., Yang, Y., & Yun, P. (2020). Risk measurement of international carbon market based on multiple risk factors heterogeneous dependence. *Finance Research Letters*, 32, 101083.

Zhang, C., Yun, P., & Wagan, Z. A. (2019). Study on the wandering weekday effect of the international carbon market based on trend moderation effect. *Finance Research Letters*, 28, 319-327.

Zhang, J., Li, D., Hao, Y., & Tan, Z. (2018). A hybrid model using signal processing technology, econometric models and neural network for carbon spot price forecasting. *Journal of Cleaner Production*, 204, 958-964.

Zhang, X., Zhang, C., & Wei, Z. (2019). Carbon price forecasting based on multi-resolution singular value decomposition and extreme learning machine optimized by the moth–flame optimization algorithm considering energy and economic factors. *Energies*, 12(22), 4283.

Zhang, Y. J., & Sun, Y. F. (2016). The dynamic volatility spillover between European carbon trading market and fossil energy market. *Journal of Cleaner Production*, 112, 2654-2663.

Zhang, Y. J., & Wei, Y. M. (2010). An overview of current research on EU ETS:

Evidence from its operating mechanism and economic effect. *Applied Energy*, 87(6), 1804-1814.

Zhang, Y. J., & Zhang, J. L. (2018). Volatility forecasting of crude oil market: A new hybrid method. *Journal of Forecasting*, 37(8), 781-789.

Zhang, Y. J., & Sun, Y. F. (2016). The dynamic volatility spillover between European carbon trading market and fossil energy market. *Journal of Cleaner Production*, 112, 2654-2663.

Zhang, J., Pal, S., Mahalik, M. K., & Gozgor, G. (2023). Does financial development promote economic globalization? Evidence from the top and the bottom globalized emerging economies. *Emerging Markets Finance and Trade*, 59(13), 3915-3927.

Zhang, Y., Pan, G., Chen, B., Han, J., Zhao, Y., & Zhang, C. (2020). Short-term wind speed prediction model based on GA-ANN improved by VMD. *Renewable Energy*, 156, 1373-1388.

Zheng, J., Fu, X., & Zhang, G. (2019). Research on exchange rate forecasting based on deep belief network. *Neural Computing and Applications*, 31, 573-582.

Zhou, F., Huang, Z., & Zhang, C. (2022). Carbon price forecasting based on CEEMDAN and LSTM. *Applied energy*, 311, 118601.

Zhou, J. F., Wu, D., & Chen, W. (2022). Cap and trade versus carbon tax: an analysis based on a CGE model. *Computational Economics*, 59(2), 853-885.

Zhou, J., Huo, X., Xu, X., & Li, Y. (2019). Forecasting the carbon price using extreme-point symmetric mode decomposition and extreme learning machine optimized by the grey wolf optimizer algorithm. *Energies*, 12(5), 950.

Zhu, B., Ma, S., Xie, R., Chevallier, J., & Wei, Y. M. (2018). Hilbert spectra and empirical mode decomposition: a multiscale event analysis method to detect the impact of economic crises on the European carbon market. *Computational Economics*, 52, 105-121.

Zhu, B., Ye, S., Wang, P., He, K., Zhang, T., & Wei, Y. M. (2018). A novel multiscale nonlinear ensemble leaning paradigm for carbon price forecasting. *Energy Economics*, 70, 143-157.

Zhu, B., Wang, P., Chevallier, J., & Wei, Y. (2015). Carbon price analysis using empirical mode decomposition. Computational *Economics*, 45, 195-206.

Zhu, D., Zhang, C., Pan, D., & Hu, S. (2021). Multifractal cross-correlation analysis between carbon spot and futures markets considering asymmetric conduction effect. *Fractals*, 29(07), 176-215.

Zhu, M., Xu, H., Wang, M., & Tian, L. (2024). Carbon price interval prediction method based on probability density recurrence network and interval multi-layer perceptron. *Physica A: Statistical Mechanics and its Applications*, 636, 129543.

Zhu, J., Wu, P., Chen, H., Liu, J., & Zhou, L. (2019). Carbon price forecasting with variational mode decomposition and optimal combined model. *Physica A: Statistical Mechanics and Its Applications*, 519, 140-158.

Zhu, P., Lu, T., & Chen, S. (2022). How do crude oil futures hedge crude oil spot risk after the COVID-19 outbreak? A wavelet denoising-GARCHSK-SJC Copula hedge ratio estimation method. *Physica A: Statistical Mechanics and its Applications*, 607, 128217.

Zhu, P., Wei, Y., Lu, T., Tang, Y., & Zhang, C. (2023). The Price–Volume Dependences in the European and Chinese Carbon Markets: New Evidence from the Fractal Analysis. *Fluctuation and Noise Letters*, 22(05), 2350031.

Zhu, T., Wang, W., & Yu, M. (2023). A novel hybrid scheme for remaining useful life prognostic based on secondary decomposition, BiGRU and error correction. *Energy*, 276, 127565.

Zou, S., & Zhang, T. (2020). Multifractal detrended cross-correlation analysis of the relation between price and volume in European carbon futures markets. *Physica A: Statistical Mechanics and Its Applications*, 537, 122310.

APPENDICES

1. Ni, L., & Khim-Sen Liew, V. (2024). Carbon emission price forecasting in China using a novel secondary decomposition hybrid model of CEEMD-SE-VMD-LSTM. *Systems Science & Control Engineering*, 12(1), 2291409.
2. Ni, L., & Khim-Sen Liew, V. (2024). THE DRIVING RELATIONSHIP OF CHINA CARBON PRICE BASED ON DIFFERENT MARKET VOLATILITY STATES. *Asian Academy of Management Journal of Accounting & Finance*, 20(2),211-227.
3. Ni, L., & Liew, V. K. S. (2023). A Novel Hybrid Model for Forecasting China Carbon Price Using CEEMDAN and Extreme Learning Machine Optimized by Whale Algorithm. *Applied Mathematics, Modeling and Computer Simulation, 657-665.*