China’s 2060 Carbon Neutrality Pledge: The Effects of Changes in China’s Natural Gas Consumption on the Australian Economy

A Thesis by Compilation Submitted in Fulfilment of the Requirements of the Degree of Doctor of Philosophy (PhD)

Jinyong (Edward) Sung

Institute for Sustainable Futures
University of Technology Sydney, Sydney, Australia

December 2025

Certificate of Original Authorship

I, *Jinyong (Edward) Sung*, declare that this thesis is submitted in fulfilment of the requirements for the award of *Ph.D. in Sustainable Futures–Energy*, in the Institute for Sustainable Futures (ISF), at the University of Technology Sydney (UTS).

This thesis is wholly my own work unless otherwise referenced or acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

This document has not been submitted for qualifications at any other academic institution.

This research was supported by an Australian Government Research Training Program (RTP) Scholarship doi.org/10.82133/C42F-K220.

Signature: Production Note:
Signature removed prior to publication.

Date: 13 December 2025

Acknowledgements

The completion of this Doctoral research would not have been possible without the support and assistance of numerous individuals and institutions. I would like to express my deepest appreciation to my Principal Supervisor, Professor Xunpeng (Roc) Shi, for his invaluable guidance, encouragement, and expertise throughout the course of this research. I am equally grateful to my Co-Supervisor, Professor Sven Teske, whose insights and support were instrumental in shaping the direction of this work.

Special thanks are extended to Dr. Mengheng Li, who served as a research advisor and provided critical methodological advice, particularly in the area of time-series analysis. His thoughtful feedback and technical input have greatly enhanced the analytical rigour of this study.

I would also like to acknowledge Mr. Duncan Ford from the UTS Business School for granting me access to its Bloomberg Terminal and Refinitiv Workspace, which are essential for data collection and analysis.

My sincere thanks go to the staff of the UTS ISF, especially Professor Stuart White (Director), Dr. Tim Foster (Research Director), Ms. Wendy Wang (Graduate Research Program Officer), Mr. Brandon Nguyen (Systems Support Officer), for their support across various aspects of the research, including their generous assistance in facilitating my participation in conferences related to this study. In particular, Dr. Foster has provided extensive assistance in his role as a Responsible Academic Officer (RAO), which is greatly appreciated.

Finally, I would like to express my deepest gratitude to my wife, Juliana, and my children, Theresa and David. Their unwavering support, patience, and understanding have been a constant source of strength throughout this challenging yet genuinely enjoyable journey. Without their love and encouragement, this journey would not have been possible.

Statement on Thesis Format

This thesis is submitted in the format of a *Thesis by Compilation*, in accordance with the guidelines of the UTS. It comprises a series of published and unpublished manuscripts that collectively address the research objectives of the doctoral study. Each chapter is based on a distinct manuscript, and author contributions are detailed in the Declaration of Publications and Author Contributions section.

Declaration of Publications and Author Contributions

For each of the three co-authored publications—listed below as Papers 1, 2, and 3—included in this thesis, the contributions of each author have been clearly stated as follows:

Paper 1: Sung, J., Shi, X., Teske, S., and Li, M. (2024), *Chinese Natural Gas Phase-Out Pathways: A Novel Hybrid Scenario-Specific Projection Approach to Achieve Net Zero.*

- Status: Published in *Energy*, Vol. 328, 1 August 2025, Article 136387.
- Contributions
  - Sung, J (Candidate): Lead author; original draft writing & editing, conceptualisation, data collection and analysis, methodology design, software development, and visualisation.
  - Shi, X. (Principal supervisor) and Teske, S. (Co-supervisor): Supervision and resources, and manuscript review & editing.
  - Li, M (Research advisor): Methodology validation, resources, and manuscript review & editing.

Paper 2: Sung, J., Shi, X., Teske, S., and Li, M. (2025), *Identifying Key Natural Gas Drivers towards China's Carbon Neutrality through A Novel Temporal Shrinkage Framework*.

- Status: The paper has been revised and resubmitted to a journal for further consideration.
- Contributions: Same as above.

Paper 3: Sung, J., Shi, X., Teske, S., and Li, M. (2025), *Carbon-Neutral Planning Framework for China's Near-Term Natural Gas Demand: A Neural-enhanced Hybrid Statistical–Deep Learning Forecasting Approach*

- Status: The paper is currently under review at a peer-reviewed journal.
- Contributions: Same as above.

Paper 4: Sung, J. (2025), *Opportunities for Australia amid China's Carbon-Neutral Transition: Input–Output Approach with Analysis of Australian Renewable Potential*

- Status: The paper is currently under review at a peer-reviewed journal.
- Contributions
  - Sung, J (Candidate): Original draft writing, conceptualisation, data collection and analysis, methodology design, software development, and visualisation.
  - Shi, X. (Principal supervisor), Teske, S. (Co-supervisor), and Li, M (Research advisor): Provided feedback during the presentation stage of the research, with Shi and Li additionally involved in manuscript review and feedback.

The following co-authors' signatures confirm agreement with the stated contributions and the inclusion of these three publications.

[Principal supervisor] Dr. Xunpeng (Roc) Shi [a]

Production Note:
Signature removed prior to publication.

[Co-supervisor] Dr. Sven Teske [b]

Production Note:
Signature removed prior to publication.

[Research advisor] Dr. Mengheng Li [c]

Production Note:
Signature removed prior to publication.

[a] Professor of Energy and Environmental Economics and Sustainability, and Research Principal at the Australia-China Relations Institute, UTS, Sydney, Australia

[b] Professor and Research Director at the ISF, UTS, Sydney, Australia

[c] Senior Lecturer in the Economics Discipline Group, at the UTS, Sydney, Australia

Table of Contents

List of figures and tables

Chapter 1.0 Introduction

Chapter 2.0 Identifying Key Natural Gas Drivers towards China's Carbon Neutrality through a Novel Temporal Shrinkage Framework

Chapter 3.0 A Hybrid SARIMAX-LSTM Model Optimised by ANN for Near-Term Forecasting: An Application to China's Natural Gas Consumption

Chapter 5.0 Opportunities for Australia amid China's Carbon-Neutral Transition: Input-Output Approach with Analysis of Australian Renewable Potential

Abstract

China's Natural Gas (NG) market is undergoing a structural transformation driven by two key pillars: the 2060 carbon neutrality target and NG market liberalisation. These shifts are expected to have significant impacts on Australia, which is China's largest NG supplier. However, extensive research that integrates these two aspects and evaluates their combined effects on the Australian economy remains scarce. This study aims to address this gap by employing an interdisciplinary approach that integrates advanced big data analytics with established econometric methodologies to forecast both short- and long-term changes in China's NG market and analyse their implications for the Australian economy.

Chapters 2 to 4 examine the temporal changes in China's NG market driven by its carbon neutrality goals. Chapter 2 identifies the key short- and long-term drivers of NG consumption—such as temperature, coal and gas prices, import volumes, and infrastructure—using a novel temporal shrinkage framework. Chapter 3 develops a hybrid forecasting model that supports carbon-neutral near-term decision-making by optimally combining a statistical model with a deep learning model. This forecasting model is used to predict short-term (one-year-ahead) NG consumption, revealing that NG infrastructure is the most influential factor in shaping short-term demand. Chapter 4 forecasts NG consumption through 2060 using a scenario-based, long-term forecasting model that integrates traditional time-series and machine learning approaches, highlighting the need for a more radical transition to achieve genuine carbon neutrality in the NG market.

Finally, Chapter 5 assesses the impacts of changes in China's NG consumption on the Australian economy using both static and dynamic Input–Output (IO) models. The analysis reveals that the NG extraction and coal industries will experience the most significant setbacks, with negative ripple effects extending to related sectors such as iron ore mining. Notably, the dynamic IO analysis reveals that the Australian economy faces risks of prolonged downturns. Additionally, it finds that NG-oriented stimulus policies offer limited efficacy, whereas renewable energy-focused stimulus induces positive spillover effects across multiple industries, bolstering economic resilience.

Accordingly, Australia needs to proactively respond to China's carbon neutrality and energy transition by restructuring its economy to focus on renewable energy. Comprehensive and strategic support policies—such as tax incentives, investments in electric vehicles and related infrastructure, and the development of renewable energy industry clusters—can facilitate this restructuring. Such policies can foster Australia's new growth engines and serve as a foundation for a sustainable, long-term economic transformation.

CHAPTER 1.0

INTRODUCTION

1.1 Research background

Amid growing international pressure for concrete action to address climate change, China continues to rely heavily on carbon-intensive energy sources to sustain its economic growth. As a result, the country remains one of the leading contributors to environmental pollution both domestically and globally. In this context, President Xi Jinping's announcement at the 75th United Nations General Assembly in September 2020—committing China to peak carbon emissions by 2030 and achieve carbon neutrality by 2060—marks a significant turning point in the nation's energy trajectory. This pledge signals a formal shift from a carbon-intensive energy system towards a low-carbon development model, and has been widely regarded by the international community as a positive and hopeful step forward.

China's Natural Gas (NG) market is subject not only to the influence of its carbon neutrality targets, but also to structural changes driven by market liberalisation. NG is considered a transitional energy source with lower carbon emissions compared to coal, and is expected to play a key role in China's energy transition. However, this shift entails more than just increased demand—it also requires fundamental changes in the way the NG market operates. While often overshadowed by climate policy, market-oriented reforms under the Communist Party of China have introduced structural shifts with profound implications for international gas trade. In particular, the transition from rigid long-term contracts to more flexible spot market mechanisms is likely to affect both the volume and pricing of Australia's NG exports.

Australia's NG industry is expected to undergo substantial changes in light of China's commitment to carbon neutrality and market liberalisation. Currently, 37% of Australia's total exports—primarily composed of resource and mineral products—are directed to China [1]. Notably, Australia serves as China's largest supplier of NG, accounting for one-third of China's total NG imports [2]. As a key resource supplier, Australia is deeply tied to the Chinese economy, and its NG extraction industry is particularly responsive to fluctuations in Chinese energy demand and policy. As China transitions towards carbon neutrality with ongoing market liberalisation, structural changes in its NG consumption appear inevitable, which will have a direct impact on Australia, one of the world's leading Liquefied Natural Gas (LNG) exporter and a nation that heavily relies on the NG industry as a key pillar of its economy.

In summary, the potential implications of China's 2060 carbon neutrality pledge for Australia's NG industry and the broader economy warrant examination. Understanding how changes in China's NG market—driven by its decarbonisation goals—could impact the Australian economy is vital. In this context, it is also crucial to consider Australia's policy implications for its NG industry moving forward.

1.2 Problem statement

China's NG market is undergoing unprecedented changes, driven by not only its shift towards carbon neutrality but also ongoing market liberalisation. The country's 2060 carbon neutrality pledge provides a clear signal of the nation's ambitious commitment to achieving net-zero emissions. This pledge outlines a three-stage roadmap: (1) The first stage aims to peak carbon emissions between 2020 and 2030 and to foster the clean energy industry on a large scale by controlling coal consumption; (2) the second stage involves a transition to renewable energy systems from 2030 to 2050; and (3) the third stage targets full carbon neutrality by 2060, driven by widespread deployment of new energy-related science and technology from 2050 to 2060. In the early stages, the low-carbon transition is primarily supported by the phase-out of coal, followed by a gradual shift towards renewables and low-carbon technologies. Within this framework, NG plays a pivotal role as a transition fuel, facilitating the shift from coal and supporting renewables aligned with the country's low-carbon transition policies. However, NG, as a fossil fuel, is expected to be phased out in the final stage to achieve carbon neutrality. Simultaneously, China's NG market liberalisation is introducing structural changes that contribute to increased uncertainty and a potential decline in demand for Australian LNG, as competitively priced pipeline gas and flexible spot market mechanisms reduce Australia's market share and long-term contract stability.

In the context of these changes, a deeper understanding is needed of the potential structural shifts in China's NG market under the carbon neutrality goal, along with an analysis of current market conditions and future projections based on relevant and innovative methods. Furthermore, it is worth examining not only the impact of these changes on Australia's NG market but also the broader economic effects on the Australian economy.

1.2.1 Brief review: Changes in China's natural gas market and their effects on Australia

Chinese NG consumption's anticipated changes caused by the carbon neutrality goal. China has emphasised the strategic importance of NG in facilitating the *coal-to-gas* transition as outlined in the 14th Five-Year Plan (2021-2025)[1] [3]. This transition strategy, serving as a short-term and pragmatic approach to de-coalification while sustaining the country's economic growth, suggests that NG consumption is likely to increase through 2030. Alternatively, NG consumption will be gradually curtailed after 2030 in alignment with China's broader decarbonisation agenda, which involves limiting its fossil fuel consumption and expanding renewables [4]. In other words, while NG will play an increasingly critical role in the near future, NG consumption is expected to decline after 2030 due to policy shifts towards limiting the use of fossil fuels and expanding renewable energy. Alongside these changes, other potential shifts in China's NG market towards liberalisation also warrant close examination.

Potential shifts in China's NG market towards liberalisation. Under the leadership of the Communist Party of China, various changes have been implemented to promote liberalisation within the NG market. Shi and Variam [5] provide key insights into China's successful gas market liberalisation as well as its impact on Australia. Through policy scenario simulations— Scenario 1 (Liberalisation with China spot) and Scenario 2 (Liberalisation with China spot and

[1] Given that the 2060 carbon neutrality has no detailed roadmap, the 14th FYP has a straightforward background to understand China's energy policies regarding the country's carbon neutrality.

no Take-or-Pay (TOP))[2]—their study reveals contrasting outcomes. Under Scenario 1, although imports from Australia would decrease in China's future gas market due to an increase in domestic gas production and piped gas supply from Russia and Central Asia, Australia could maintain its exports up to the TOP level. Under Scenario 2, Australia's exports would decline at a relatively larger pace because Australia's LNG would be more expensive than pipeline gas from Russia and Central Asia. The study concludes that China's effective measures for gas market liberalisation are to establish the market with a combination of spot and contracted agreements, including TOP and destination clauses, based on single and national benchmarks. Additionally, under both scenarios, Australia's NG is likely to decrease in China's future liberalised NG market. Alternatively, Australia's economy is closely connected to China's NG market. Therefore, it is imperative to examine how changes in China's NG market impact the Australian economy.

Effects on the Australian economy. As China's NG market liberalisation advances, Australia is expected to lose its position as a dominant supplier to the Chinese NG market. Liberalisation is likely to render Australia's LNG exports less competitive relative to pipeline gas supplies from Russia and Central Asia, potentially reducing the total economic benefits to the Australian economy [5]. Furthermore, as China's NG market liberalisation is closely linked to global crude oil prices, the impact of these changes on the Australian economy must be analysed within the context of global NG price fluctuations [5].

As the low-carbon transition policies in Asian countries drive up global demand for NG, Australia could encounter expanded export opportunities; however, this increased demand could have a negative impact on the domestic industry in the form of elevated NG prices. According to Deloitte's analysis, Australia's NG prices are indexed to international prices, which could result in higher domestic NG prices and potential supply shortages [6]. These changes could place economic pressure on energy-intensive industries, particularly manufacturing. Ultimately, the low-carbon transition exerts the opposite effect on the Australian economy: it leads to more exports and higher costs.

As discussed above, China's NG market is closely associated with the country's carbon neutrality goal and ongoing market reforms aimed at liberalisation. These dynamics have a significant impact not only on China but also on Australia's NG market and the broader economy. To accurately analyse the impact on Australia, this study begins by establishing a detailed understanding of China's NG market. Accordingly, it is necessary to examine the key drivers of China's market by focusing on the country's carbon neutrality policies and the progression of market liberalisation, while also considering its interconnection with the global NG market. Based on this foundation, this study aims to forecast potential developments in China's NG market and assess their implications for Australia.

1.2.2 Bridging the gap

China's 2060 carbon neutrality target and the liberalisation of its NG markets constitute two central pillars of the structural transformation underway in the country's NG sector. However, no comprehensive study has yet jointly examined these two dimensions—

[2] Under Scenario 1: Chinese spot, China's Shanghai spot price could be the national benchmark for China's piped gas and LNG; Under Scenario 2: Chinese spot and no TOP, all the contracts in China exclude TOP and destination clauses.

decarbonisation and market liberalisation—in analysing the evolution of China's NG market. Additionally, although several relevant studies exist, most analyses of China's future NG market have been limited to near-term forecasts. For a more holistic understanding of China's NG market, it is essential to investigate China's structural market changes towards liberalisation and the transition towards carbon neutrality, using projections that span both near-term (by 2030) and the long-term (by 2060).

Conducting such an analysis necessitates a novel hybrid approach that integrates big data frameworks with time series analyses. Traditional theoretical models commonly employed in energy economics and policy are limited in their capacity to comprehensively and accurately analyse the diverse big data that drives structural changes in China's NG market. Consequently, there is a growing need to adopt data-driven methodologies that introduce innovative analytical techniques and yield more robust empirical and policy-oriented insights. In recent years, machine learning (ML) and deep learning (DL)-based data-driven methods have been increasingly applied in energy economics and policy research. For example, Shi et al. [5] employ an ML-based decomposition technique to analyse the primary factors influencing household carbon emissions, whereas Rezaei et al. [7] use a DL-based forecasting method to predict gasoline consumption trends in Iran. Nonetheless, few studies have employed such techniques to analyse the structural transformation of China's NG market through the dual lenses of decarbonisation and market liberalisation. Furthermore, there remains a notable gap in the literature concerning how these structural changes will affect the Chinese NG market and their temporal impacts—both short- and long-term—on the Australian economy.

To address this gap, this study seeks to identify the key drivers of structural change in China's NG market using a big data-based methodology integrated with time series analysis. Based on these drivers, this study will project short- and long-term NG demand and further examine the resulting impacts on the Australian economy.

1.3 Objective and key questions

This study's main objective is to analyse how changes in China's NG consumption—driven by the country's 2060 carbon neutrality goal and market liberalisation—impact the Australian economy and to develop effective policy responses for Australia. To achieve this, the study adopts an interdisciplinary methodology that integrates ML/DL techniques with econometric approaches. The analysis is structured in three main stages. In the first stage, a wide range of factors influencing China's NG market—identified through the literature—are collected, and key drivers are selected from among them. Based on these key drivers, the second stage involves forecasting short-term NG consumption in China and long-term demand using a scenario-based approach. In the third stage, the study assesses the implications of these demand projections for the Australian economy and derives corresponding policy responses. Through this process, this study addresses the following key questions:

Stage 1: Identifying short- and long-term key drivers in China's NG market.

- Based on existing literature that examines evolving trends in China's carbon neutrality goal and market liberalisation, what factors influence China's NG market?
- Among these diverse influencing factors, which key drivers have been identified from a temporal perspective—short-term (by 2030) and long-term (by 2060)?

Stage 2: Short-term forecast and scenario-specific long-term projection of China's NG demand

- What short-term forecast of China's NG consumption emerges from the selected key drivers?
- How do these short-term NG consumption forecasts compare with existing studies in terms of differences and implications?
- What does the scenario-based long-term NG consumption projection reveal about China's future NG market landscape?
- Based on the scenario-based long-term projections in NG consumption in this study, how can the feasibility of China achieving its carbon neutrality goal be evaluated?

Stage 3: Evaluating Australian economic impacts and designing policy responses

- What effects will the NG consumption projections from this study have on the Australian economy?
- What Australian industries are expected to be most impacted?
- How should Australia develop and execute its future policy responses?

1.4 Methodology and chapter linkage

This study proposes multiple innovative hybrid frameworks that integrate ML, DL, and time series analysis across three stages, as illustrated in Fig. 1.1. Each stage is discussed in detail in the corresponding chapters of this thesis. Below is a brief overview of each stage, along with the related chapters.

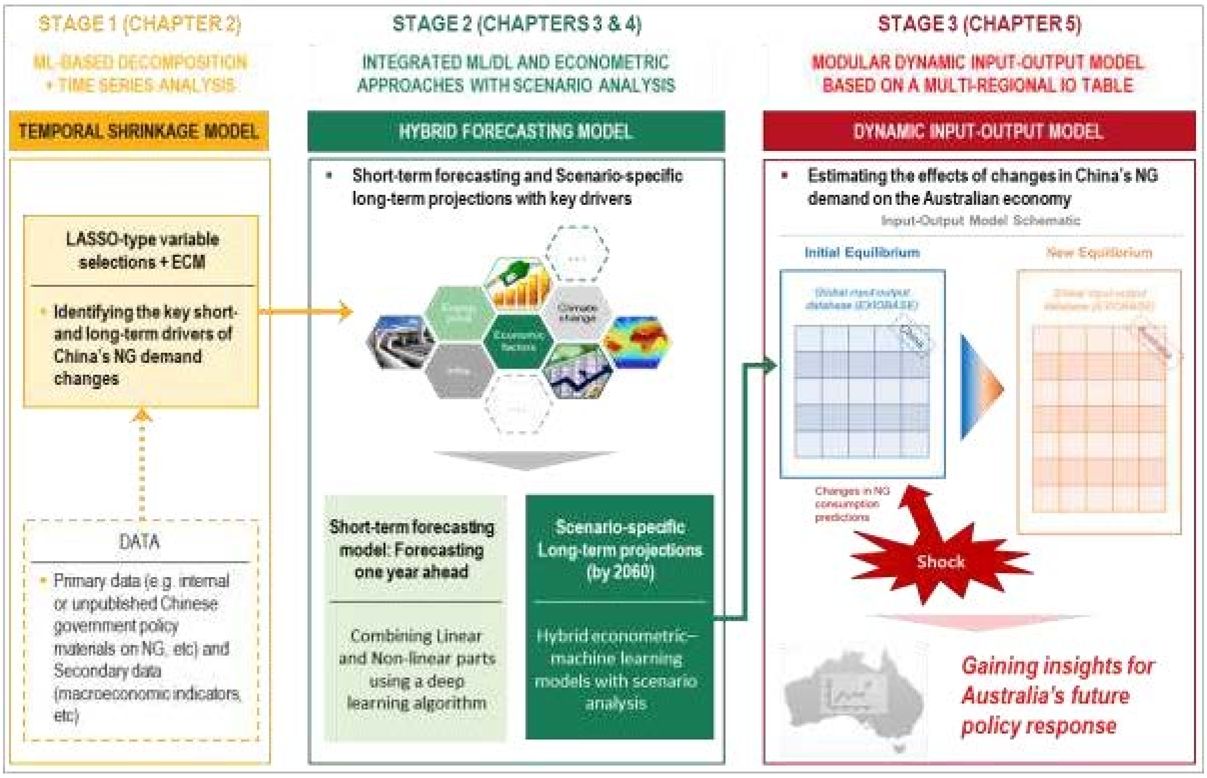


Figure 1.1 Methodology and its linkage to the relevant chapter

In the first stage, *Least Absolute Shrinkage and Selection Operator (LASSO)*-based variable selection techniques are integrated with the *Error Correction Model (ECM)*, a time series analysis method, to create a big data–driven temporal variable selection framework. This framework identifies temporal key drivers among various factors influencing changes in China's NG demand. These factors are derived from both primary data, including Chinese government policy documents, and secondary data, such as macroeconomic indicators. Chapter 2 provides an extensive discussion of the detailed methodology and specific contents of the first stage.

The second stage introduces hybrid methods for short-term forecasting of China's NG consumption and its scenario-based long-term projections. Initially, this study develops a short-term forecasting model by extracting linear and nonlinear features of the NG consumption based on the key drivers identified in the first stage, and by optimising these features using a DL algorithm. The model aims to achieve high predictive accuracy in one-year-ahead forecasts, offering practical implications for carbon-neutral-based near-term business strategies and policy formulation. Chapter 3 presents the detailed methodology and empirical results. Subsequently, this study integrates econometric tools for linear estimation with machine learning techniques for nonlinear estimation to construct a model suitable for scenario-specific long-term projections. Based on the key drivers, this scenario-based model projects China's NG consumption through 2060 and explores the feasibility of an NG phase-out under China's carbon-neutral scenarios. Chapter 4 provides the detailed methodology and results.

In the third stage, the study analyses the comprehensive economic impacts of changes in China's NG consumption on Australia. The analysis primarily employs a modular dynamic input–output (IO) model based on a multi-regional IO table. Based on the results, the study proposes policy recommendations on how Australia could respond to China's carbon-neutral transition. Chapter 5 elaborates on the methodology and corresponding results.

1.5 Concluding remarks

China's 2060 carbon neutrality pledge is likely to bring about fundamental shifts in its NG market, with significant implications for Australia's economy. If the Australian government fails to adequately prepare for the uncertainties arising from China's energy transition, it may be unable to respond effectively and promptly to potential disruptions. Additionally, analysing the implications of China's carbon-neutral transition for Australia's economic future can offer critical guidance for prioritising policies and implementing appropriate measures. In this context, this study conducts a comprehensive analysis of these issues with the aim of contributing to Australia's sustainable growth.

CHAPTER 2.0

IDENTIFYING KEY NATURAL GAS DRIVERS TOWARDS CHINA'S CARBON NEUTRALITY THROUGH A NOVEL TEMPORAL SHRINKAGE FRAMEWORK

This chapter is based on the manuscript: Sung, J., Shi, X., Teske, S., and Li, M. (2025). Identifying Key Natural Gas Drivers towards China's Carbon Neutrality through A Novel Temporal Shrinkage Framework. This manuscript has completed peer review and been resubmitted to Resources Policy in July 2025. Author contributions are detailed in the Declaration of Publications and Author Contributions section of this thesis.

Chapter outlook

Natural gas (NG) plays a critical short-term role but must be gradually phased out in the long term to achieve China's ambitious dual carbon goals (DCGs), making it essential to identify key drivers that reconcile these divergent temporal demands. This chapter employs a novel temporal shrinkage framework that integrates big data analytics (LASSO/ALASSO) with advanced time-series modelling (ECM), proposing the long-run model (sequentially combining LASSO/ALASSO and ECM) and the short-run model (simultaneously incorporating both ECM and LASSO/ALASSO). The long-run model first identifies the key drivers in the NG market—temperature, thermal coal and Henry Hub (HH) gas prices, piped NG (PNG) and LNG imports, NG market liberalisation, and NG infrastructure-related factors—and analyses their long-term effects. The short-run model then preliminarily assesses the short-term effects of these drivers while further exploring more comprehensive short-run dynamics by identifying additional drivers, such as ESG, global oil price, coking coal price, LNG and PNG import prices, and economic indicators. Key findings offer policymakers insights into these drivers, enabling the formulation of initiatives that balance short- and long-term effects to advance DCGs for China's sustainable development.

Keywords: Carbon Neutrality; LASSO; ALASSO; ECM; Natural Gas; China

2.1 Introduction

In the pursuit of China's *dual carbon goals (*DCGs)—carbon peak by 2030 and carbon neutrality by 2060—natural gas (NG) occupies a paradoxical position, presenting dilemmas on two distinct fronts. In the short term, NG serves as a bridge to a low-carbon future—most notably within the 'coal-to-gas' strategy—offering a pragmatic solution to reduce carbon emissions by 2030 [3, 8]. In contrast, looking ahead to 2060, long-term reliance on NG threatens carbon neutrality efforts. The technology is fraught with challenges, including methane leakage [9, 10], the crowd-out effect that may undermine cleaner energy adoption [11], the risk of carbon lock-in [11, 12], and the emergence of stranded assets [13]. To formulate effective initiatives for achieving DCGs, it is therefore crucial to understand the key drivers in the NG market from both short- and long-term perspectives. A core challenge for the industry lies in reconciling the burgeoning NG demand for an immediate energy transition with NG's gradual phasing out required for long-term carbon neutrality. This tension underscores the necessity for a rigorous, temporally informed analysis of the underlying NG drivers of these conflicting imperatives.

Adding to the inherent contradiction, China's NG market is shaped by dual layers of uncertainty. In the short term, market instability is exacerbated by rapid structural changes stemming from domestic market liberalisation [14, 15] and volatile global developments, such as the prolonged Russo–Ukrainian war [16]. In the long term, uncertainty increases further due to the planned fossil fuel phase-out [4] and fluctuations in electricity demand—driven by the accelerated electrification necessary for achieving China's carbon neutrality goals [17]. Under these uncertainties, current studies find it difficult to identify the key drivers, especially to reconcile the contradictory needs in the short and long term. These studies have highlighted various factors as primary NG drivers, including economic trends, energy policies, environmental considerations, technological advancements, and global trade dynamics (See Xu and Wang [18], Bu et al. [19], Jiang et al. [20], and Wang and Li [21] for recent examples). However, these studies lack consistency in the drivers' selection criteria and fail to account for the NG market's short- and long-term uncertainties. Accordingly, a more nuanced understanding of the underlying dynamics is crucial for formulating robust policies and industry strategies that balance short-term energy transition demands with long-term carbon neutrality goals, thereby supporting the effective realisation of China's DCGs and broader sustainable development agenda.

This study aims to fill these gaps by employing machine learning (ML) techniques, large datasets, and econometric approaches to analyse the key drivers of China's NG market. This study first employs ML-based shrinkage techniques—the *Least Absolute Shrinkage and Selection Operator (LASSO)* and *Adaptive LASSO (ALASSO)*—to identify these drivers from time-series data on 40 factors related to China's NG market. These methods streamline variable selection while improving prediction accuracy, model consistency, and interpretability in high-dimensional settings. Subsequently, the selected drivers' short- and long-run effects are examined using the *Error Correction Model (ECM)*, an econometric tool that simultaneously analyses both short-term fluctuations and long-term equilibrium dynamics of variables. This novel combined LASSO-ALASSO-ECM framework identifies key drivers—temperature, thermal coal and Henry Hub (HH) gas prices, piped NG and LNG imports, NG market liberalisation, and NG infrastructure-related factors—and analyses their temporal effects. These drivers link short-term energy transition needs with long-term carbon neutrality goals by capturing both structural shifts from market liberalisation and decarbonisation—areas overlooked in previous studies. The framework also examines additional short-run variables—ESG, global oil and coal prices, LNG/PNG import prices, and economic indicators—that remain underexplored in prior studies to provide a more comprehensive view of China's NG market dynamics. Key findings provide insights for policymakers to develop effective energy initiatives for carbon neutrality, considering both short- and long-run effects.

The primary contribution to the literature lies in the introduction of a novel temporal shrinkage framework that integrates ML and econometric methodologies into new energy policy analytics. Specifically, it explores the combined utilisation of LASSO/ALASSO-driven variable selection and ECM-based temporal analysis for identifying key short- and long-run NG drivers. To the best of our knowledge, this marks the first application of ALASSO's extended capability in the energy economics and policy field, contributing to a deeper analysis of the long-term effects. Compared to its previously recognised capability to overcome LASSO's limitations, such as bias and inconsistency [22], it has an extended capability to accurately select cointegrated (long-run) variables, thus identifying their long-run relationship [23]. This capability allows for analysis using the inherent long-term information within the data without altering it or encountering spurious regression [24]. Another key contribution is the identification of a robust set of NG drivers capable of reconciling the contradictory imperatives of short-term energy transition and long-term carbon neutrality. This integrated approach directly addresses the tension highlighted at the beginning of this chapter, providing policymakers with these drivers for developing evidence-based energy policies that reconcile short-term transition dynamics with the pursuit of long-term carbon neutrality. Furthermore, the drivers can be used as a foundational input for future forecasting research aimed at designing investment strategies for energy industry stakeholders navigating the short- and long-term complexities of the carbon-neutral transition.

The remainder of this chapter is organised as follows: Section 2.2 outlines the background of China's NG market and reviews related literature, including influencing factors of the NG market and LASSO- and ECM-based approaches and applications. Sections 2.3 and 2.4 describe data and methods, respectively. Section 2.5 presents the results and discussion, and the final section concludes the study. Supplementary materials include terms and abbreviations, tables, and figures used throughout this chapter.

2.2 Background and literature review

At the centre of China's NG market lie dual temporal uncertainties and the contrasting roles NG plays under the DCGs. In September 2020, China announced to peak carbon emissions by 2030 and achieve carbon neutrality by 2060. In the short term, NG is positioned as a key transition fuel under the 'coal-to-gas' initiative outlined in the 14th Five-Year Plan, intended to help meet carbon dioxide (CO_2) reduction targets and improve air quality [3, 8]. Accordingly, China's short-term NG consumption is expected to rise [3]. However, the uncertainties in NG consumption could also be expected to rise in the short term due to domestic structural changes resulting from market liberalisation [14, 25] and international NG price fluctuations caused by the prolonged Russo-Ukrainian war and Middle Eastern crisis [16, 26].

Conversely, China's carbon-neutral policies aim to decrease all fossil fuel consumptions, including NG, in the long term [4]. This long-term role of NG creates continuous uncertainties in the NG market, largely due to the absence of a clearly defined direction for phasing out NG entirely [27], which conflicts with more ambitious carbon neutrality goals suggested by NG's emerging critical environmental issues [28, 29]. For instance, methane emissions from the NG supply chain—approximately 80 times more potent than CO_2 as a greenhouse gas [9]—undermine its climate benefits in mitigating CO_2 emissions compared to coal [10]. Additionally, NG's role in potentially crowding out renewables [11] could result in carbon lock-in [11, 12] and stranded assets in the long term [13]. These long-term uncertainties could be exacerbated by China's primary carbon neutrality strategies—aggressive renewable expansion and electrification [30]—which intensify electricity demand fluctuations, ultimately raising the volatility in the NG consumption, including the use of NG to complement intermittent renewables. [17].

To analyse the shifting dynamics and temporal uncertainties of China's NG demand under DCGs, it is critical to understand both the intricate web of factors shaping NG demand and the methods used to recognise them. This literature review first examines the influencing factors behind China's NG demand and, subsequently, the methods used to identify key drivers among these factors and assess their short- and long-term effects.

2.2.1 Influencing factors of China's NG demand

Research has explored the factors influencing Chinese NG demand using diverse methodological frameworks. The literature generally identifies the most influencing factors as a mixture of economic, energy, urbanisation, and demographic characteristics. Xu and Wang [18] apply the logarithmic mean Divisia index (LMDI) to analyse China's NG consumption by decomposing economic growth, energy intensity, energy structure, and substitution effects, highlighting that as economic development increases, the energy substitution effect grows while the economic growth effect decreases. Bu et al. [19] employ social network analysis and LMDI to identify not only the economic effect reflecting per capita GDP, fossil energy structure, and energy intensity effect calculated as an inverse index of energy efficiency as main influencing factors of China's NG consumption at the national level, but also energy efficiency and population density as the main influencing factors at the provincial level. Jiang et al. [20] find that long-run NG demand is primarily driven by economic growth, urbanisation, and energy intensity. Wang and Li [21] propose a grey model to determine leading NG demand factors across regions of China, identifying economic development (GDP), industry structure, environmental policy's qualitative index, urbanisation rate, population density, energy consumption intensity and energy consumption structure as the influencing factors. Notably, they observe that the leading factors in eastern and central China are energy consumption structure, GDP, and urbanisation rate, whereas, in western China, the leading factors are urbanisation rate, industry structure, and population density.

Additionally, several studies consider energy-related infrastructure, technological progress, environmental policy, and climate factors as influencing factors, along with the combination of economic, energy, urbanisation, and demographic factors. Gao and Dong [31] use LMDI to identify economic growth, urban spatial expansion, pipeline network density and length, urbanisation rate, population density, and NG substitution as NG demand drivers in China. Liu et al. [32] apply generalised least squares to analyse the NG consumption of urban residents in China based on NG price, household income, pipeline length, household size, energy substitution, temperature, and NG consumption intensity. Mu and Li [33] develop a system dynamics model of China's NG supply–demand structure and conclude that economic development, urbanisation impact, and gas technological advancements are the main influencing factors on China's NG consumption rather than population growth. Overall, the main influencing factors of NG consumption in China can be categorised as follows: economic growth, including GDP; energy indicators, including NG/LNG prices and consumption and NG substitution (oil and coal); urbanisation rate; demographics, including population growth and density; NG infrastructure, such as pipeline length; energy-related technological progress; temperature; and environmental policy.

In addition to these factors, NG price distortions caused by the Chinese government's control have affected NG demand. Shi et al. [34] argue that government intervention in the domestic NG prices could affect NG demand by hindering the final consumers from benefiting from the low import prices. A case study on China's regulatory price distortion indicates that such interventions negatively affect economic growth due to the inefficient use of and misallocation of energy [35]. Paltsev and Zhang [36] suggest that China's efforts to transition towards a market-based NG pricing system could address the supply and demand imbalance caused by price distortion.

Furthermore, environmental, social, and governance (ESG) factors are increasingly recognised as influencing the NG demand. Research underscores ESG's rising influence in the oil and gas sector by positively impacting the value of oil and gas companies and their financial condition [37]. Wang et al. [38] highlight environmental regulation—a key constituent of the environmental pillar (E) within the ESG framework [39]—as a major driver of NG consumption in China.

By sorting out the literature thus far, the factors influencing the NG demand are classified by economic growth, energy indicators, urbanisation, demographics, environmental policy, infrastructure, climate, technological growth, price distortion, and ESG factors. However, identifying the key drivers of China's NG market from the literature is challenging due to the varying criteria for selecting the factors. Furthermore, the analysis of their short- and long-term effects is limited. Therefore, there is a growing need for a new data-driven decomposition approach involving time-series analysis.

2.2.2 Methods to reveal drivers and their short- and long-term effects

To address this need, this study explores three theoretical models: two LASSO-type variable selection models (LASSO and ALASSO) and one ECM. LASSO-type shrinkage and variable selection techniques have gained prominence in the ML and econometric literature. In the realm of time-series analysis, ECM techniques have been widely used to differentiate between short- and long-run relationships among variables. This study combines these techniques to disentangle the short- and long-run effects of key drivers on China's NG consumption.

Traditional econometric models, such as decomposition techniques including time-series analysis, have limitations in big data contexts. For instance, Adow et al. [40] examine Gulf Cooperation Council countries from 1990–2022 using a cross-sectional dependence (CSD)-robust panel auto-regressive distributed lag (ARDL) model within the environmental Kuznets curve framework, observing that financial development reduces CO_2 emissions in the long run, whereas energy use and urbanisation increase them. Similarly, Zhou et al. [41] apply the CSD-ARDL model to 41 high-emission countries from 1990 to 2021 and demonstrate that energy intensity and non-renewable energy consumption increase emissions, whereas renewable energy consumption decreases them. Unlike ML techniques, these traditional models lack automated variable selection capabilities from big data, which limits the big data-driven analyses of this study.

Recent advancements in energy economics and finance have increased ML analyses, particularly for big data-driven models. Regularisation techniques, such as LASSO, which shrinks the coefficients of specific variables to zero, have garnered substantial attention for their ability to control overfitting and identify key variable characteristics [42], thus propelling LASSO's adoption in the energy sector. Ghoddusi et al. [43] highlight the effectiveness of ML analyses, including LASSO, in energy-related research, covering classification, prediction, and policy analysis. Shi et al. [44] apply LASSO to identify key driving factors affecting China's household carbon emissions, thereby providing valuable insights for formulating low-carbon policies. Accordingly, LASSO has emerged as a valuable tool in energy economics and policy research for data-driven decision-making support and understanding of complex energy systems.

However, LASSO exhibits limitations, including bias and inconsistency, in high-dimensional settings where the number of variables exceeds the sample size [22]. ALASSO, which addresses these limitations, has been studied in various fields, including energy. Zou [22] introduces ALASSO, which achieves model selection consistency—meaning it correctly identifies the true variables—and efficient estimation; together, these features constitute the oracle property. This is accomplished by penalising different coefficients in LASSO's regularisation term. Mendes [23] extends ALASSO to cointegrated regressions, selecting the correct variable subset and demonstrating its oracle property. Additionally, Kock [45] demonstrates ALASSO's oracle property through consistent and conservative model selection in time series, including stationary and nonstationary autoregressions. Duras et al. [46] introduce ALASSO as a new variable selection method in energy economics, comparing its variable selection performance with other methods, including LASSO.

Building on these, this study aims to integrate LASSO/ALASSO's variable selection technologies with the ECM to identify NG demand drivers under DCGs and disentangle their short-run effects from long-run ones [24, 47]. Although prior research has applied LASSO-type techniques to ECMs, few have explicitly addressed temporal uncertainties. For example, Liao and Phillips [48] use ALASSO within a vector ECM (VECM) to simultaneously determine the cointegration relationships and autoregressive order. Similarly, Liang and Schienle [49] use the LASSO-VECM approach to explore the cointegration rank and autoregressive lag order determination in high-dimensional settings. However, these studies focus on the dynamic relationships between variables and the structural characteristics of the models rather than the short- and long-run effects of variables.

In summary, while previous literature identifies a multitude of diverse influencing factors within China's NG market, divergent selection criteria of these factors along with limited temporal analyses could obscure a clear understanding of the critical drivers and their dynamic effects. Such opacity and the lack of temporal effect analysis can impede the development of coherent and time-sensitive policies towards achieving DCGs. To address this, this study employs a novel temporal shrinkage framework integrating LASSO/ALASSO with ECM to identify key drivers in China's NG market and analyse their short- and long-run effects.

2.3 Data

This study examines 40 factors—identified through the literature review—as independent variables influencing NG consumption, which serves as the dependent variable. These factors are categorised into six groups: (1) Economics and Energy (15 factors); (2) Global Energy Index (5 factors); (3) Urbanisation and Demographics (4 factors); (4) Infrastructure (4 factors); (5) Environment, Climate, and Technological Growth (7 factors); and (6) Price Distortion (5 factors).

These factors consist of monthly time-series data collected from July 2013 to December 2019. The data series concludes in December 2019 due to the limited availability of price distortion data beyond this point and to avoid distortions caused by COVID-19-related market disruptions. Annual or quarterly data have been converted to monthly data through interpolation. Economic factors are expressed in constant or real values. Table 2.1 provides a detailed description of these factors and their respective sources (see Appendix B).

2.3.1 Data preparation

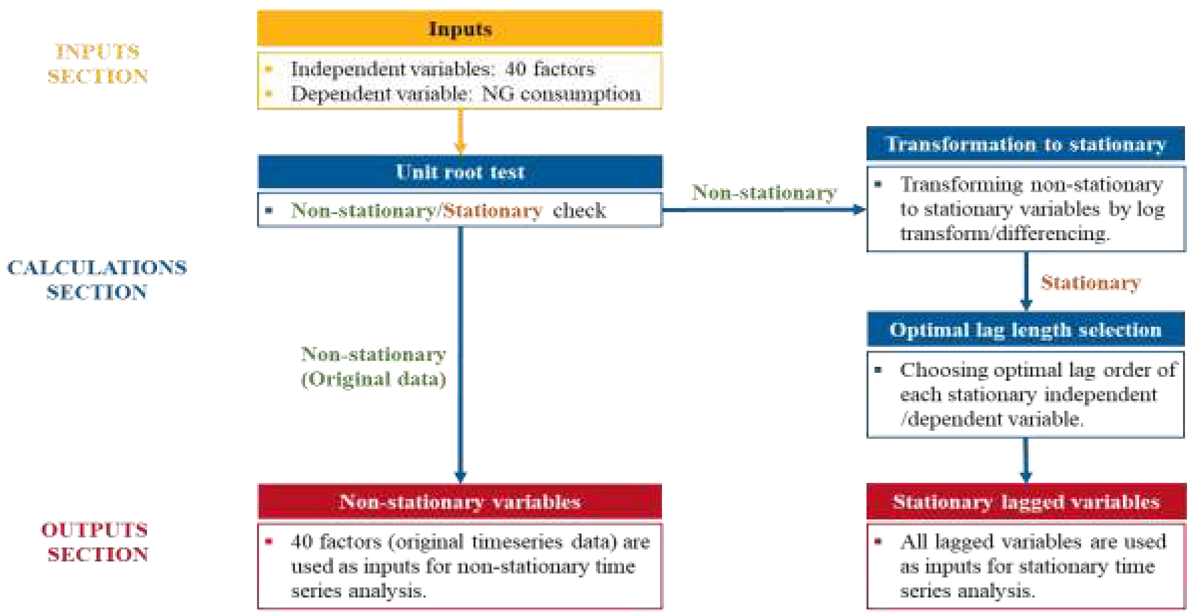


Figure 2.1 Data preparation

Stationary and non-stationary time-series analyses are conducted to prepare the input data for LASSO/ALASSO and ECM modelling.[3] First, to determine whether a time-series variable is non-stationary or stationary, two main unit root tests—augmented Dickey–Fuller (ADF) and Kwiatkowski–Phillips–Schmidt–Shin (KPSS) tests—are employed. For the U, AE, AF, and AG factors listed in Table 2.2 (see Appendix C), the Phillips–Perron test is used to further investigate the presence of a unit root, given the known finite-sample limitations in the KPSS test (see Zivot and Wang [50]). Second, the choice of whether the analysis focuses on non-stationary or stationary time series determines whether to use raw data or apply transformations. Non-stationary time-series analysis uses raw data, whereas stationary time-series analysis involves transforming non-stationary variables into stationary ones through log transformation or differencing [24]. Table 2.2 presents the transformation of variables into stationary ones—via log transformation or differencing—alongside their statistical significance assessment. Finally, as lagged stationary variables are used in the stationary analysis, their optimal lag order is selected using the vector autoregression (VAR) model with the Akaike information criterion (AIC) [51]. As presented in Table 2.2, the total number of the lagged stationary variables is 117. Figure 2.1 presents the schematic outlining the preparation of time-series data.

2.4 Methods

This section first outlines the hybrid regression techniques that sequentially integrate LASSO/ALASSO and ECM (hereafter, *long-run model*), and then presents the combined methodology that simultaneously incorporates both ECM and LASSO/ALASSO (hereafter, *short-run model*).

2.4.1 Long-run model

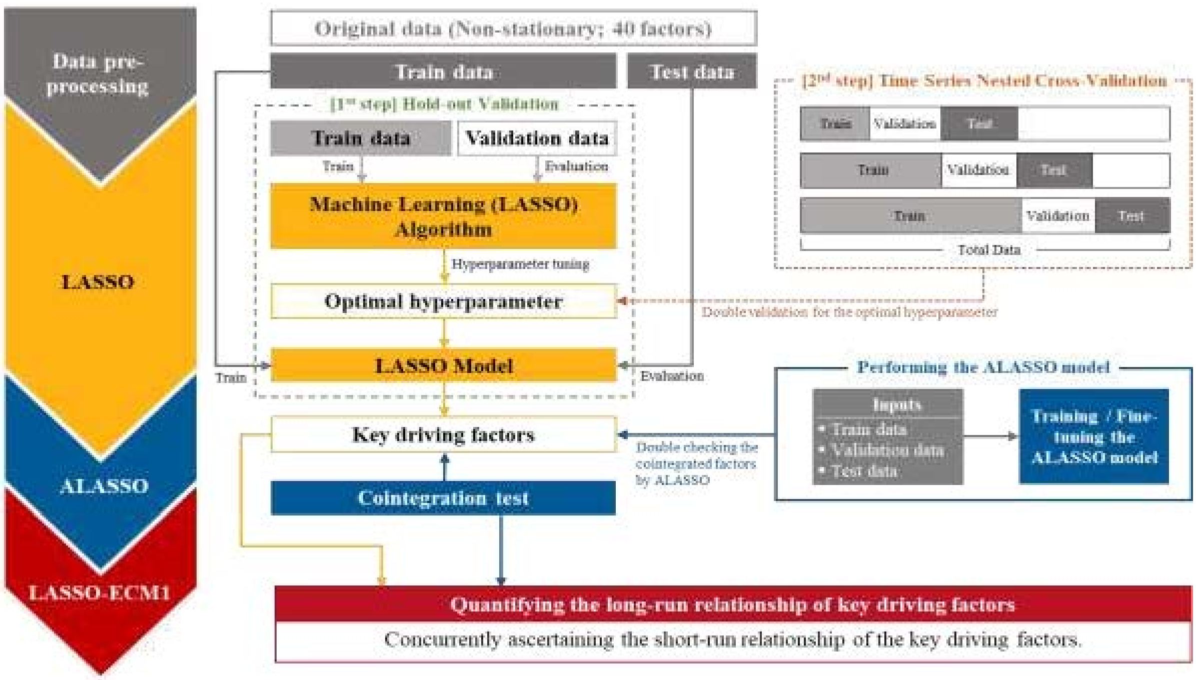


Figure 2.2 Long-run model

[3] A stationary time series has a constant mean and variance over time, while a non-stationary time series shows changing statistical properties.

This study utilises original (non-stationary) time-series datasets of 40 factors to analyse their key drivers and temporal effects. Here, a dilemma arises regarding non-stationary data. In general, the regression model using non-stationary time series can cause a spurious regression, where linear regression shows a misleading relationship between variables [24]. Therefore, for regression analysis, preference is given to time-series datasets that are either inherently stationary or transformed to stationarity. Moreover, transforming non-stationary time series to stationary can lead to the loss of their inherent long-run information [24].

The proposed long-run model identifies key drivers in China's NG market and explores their long-run effects without encountering spurious regression or losing long-run information. As shown in Figure 2.2, the long-run model progresses through the following stages. First, LASSO is applied directly to the original time-series datasets of 40 factors without converting them into stationary series to identify the key driving factors. Second, a residual-based cointegration test—the 1st step of the two-step procedure in Engle and Granger [47]—estimates the cointegrated regression (long-run relationship) among these factors. ALASSO then verifies the robustness of these cointegrated factors, indicating their stable relationship even if they exhibit short-term fluctuations [52]. Finally, ECM (hereafter, *LASSO-ECM1*) is used to quantify their long-run relationship. Further elaboration on each approach is provided in the subsequent sections.

2.4.1.1 LASSO and ALASSO

LASSO, an ML's technique, performs both feature selection and regularisation and is mathematically expressed as follows:

$$\boldsymbol{Minimize}_{\boldsymbol{\beta_0,\beta}} \left(\frac{1}{2n} \sum_{i=1}^{n} (\boldsymbol{y_i} - \boldsymbol{\beta_0} - \boldsymbol{x_i^T \beta})^2 + \boldsymbol{\alpha} \sum_{j=1}^{m} |\boldsymbol{\beta_j}| \right) \quad \textbf{(1)}$$

where n is the total number of observations; y_i is the dependent variable; m is the total number of independent variables $x_i = (x_{i1}, \ldots, x_{im})^T$; β_0 is the intercept; β_j represents the other parameters (coefficients); and α is the hyperparameter value.[4]

Equation (1) shows that LASSO penalises the cost function of the linear regression model (LRM). The penalty is the sum of the absolute values of the hyperparameter value α multiplied by coefficients, which is called *L1 Regularisation*. The cost function is the sum of the squared difference between y_i (actual result) and $\hat{y}_i = \beta_0 - x_i^T \beta$ (estimated result).

LASSO aims to find β_0 and β that minimise the cost function and penalty. Specifically, each coefficient in the penalty is reduced by increasing the hyperparameter value α. This process reduces some coefficients to zero. These unique properties of LASSO can lead to relatively choosing more essential variables, while excluding less important variables by setting their coefficients to zero [53].

However, LASSO can inadvertently eliminate features with small coefficients, causing overfitting [22]. To overcome this limitation, ALASSO penalises larger coefficients, thereby identifying an accurate set of true variables. This phenomenon—known as the *oracle property*—enables ALASSO to select cointegrated variables, indicating their long-run relationship [23].

[4] In ML algorithms, hyperparameters control the learning process and assist in determining the appropriate model parameters (coefficients).

ALASSO extends LASSO using adaptive weights [54]. The mathematical formulation for ALASSO is as follows:

$$\boldsymbol{Minimize}_{\boldsymbol{\beta_0,\beta}} \left(\frac{1}{2n} \sum_{i=1}^{n} (\boldsymbol{y_i} - \widehat{\boldsymbol{y}}_i)^2 + \boldsymbol{\alpha} \sum_{j=1}^{m} \widehat{\boldsymbol{w}}_j |\boldsymbol{\beta_j}| \right) \quad \textbf{(2)}$$

where $\widehat{w}_j$ represents adaptive weights, calculated reciprocally from $|\hat{\beta}_j|$ values derived from LASSO. The remaining parameters include n, y_i, $\hat{y}_i$ ($= \beta_0 - x_i^T \beta$), m, β_j, and α, mirroring LASSO's parameters. As demonstrated in Equation (2), ALASSO introduces adaptive weights to impose penalties on different coefficients within L1 Regularisation. In this study, both LASSO and ALASSO models are developed using Python, an open-source programming language.

2.4.1.2 LASSO-ECM1

The ECM constitutes the 2nd step of the Engle–Granger two-step procedure after a residual-based cointegration test [47], designed to estimate both short- and long-run effects of one time-series variable on another [24]. The fundamental version of ECM is mathematically expressed as follows:

$$\Delta \boldsymbol{Y_t} = \boldsymbol{\beta_0} + \boldsymbol{\beta_1} \Delta \boldsymbol{X_t} - \boldsymbol{\gamma} \widehat{\boldsymbol{\mu}}_{\boldsymbol{t-1}} + \boldsymbol{e_t} \quad \textbf{(3)}$$

where β_1 serves as the short-run coefficient, signifying the immediate impact of a change in X_t on a change in Y_t; $\hat{\mu}_{t-1}$ represents the lagged one-period residuals—referred to by the error correction term (ECT)—that are estimated through the LRM of two different series X_t and Y_t; γ stands as the coefficient of the ECT—also known as the adjustment effect or error correction coefficient—dictating the speed of adjustment towards the long-run equilibrium; e_t accounts for the white noise error term; β_0 denotes the constant term; and Δ represents the first difference operator.

In general, as presented in Equation (3), a statistically significant and negative ECT coefficient (γ) indicates that deviations from the long-run equilibrium are subsequently corrected towards the equilibrium. The adjustment effect, as governed by γ, is specifically observed in the two cases [55]: (1) When γ falls between -1 and 0, the adjustment monotonically converges to the long-run equilibrium; and (2) when γ ranges between -2 and -1, the adjustment oscillates around the long-run equilibrium with a damping effect. The long-run relationship of the LASSO-selected key drivers within these cases can be estimated, but not others. This analytical approach, carried out using the statistical software EViews, corresponds to LASSO-ECM1 as introduced earlier in this study. Simultaneously, LASSO-ECM1 also estimates the short-run relationship of these drivers.

2.4.2 Short-run model

The short-run model comprises LASSO-ECM1 and ECM-LASSO-ALASSO (hereafter, *LASSO-ECM2*), a simultaneously integrated approach for identifying a broader set of short-run variables. As depicted in Figure 2.3, the short-run model begins with LASSO-ECM1, a preliminary test, and then proceeds to LASSO-ECM2, a more advanced test. LASSO-ECM1 yields relatively straightforward short-run effects of the key drivers selected within the long-run model, as it relies on the fundamental ECM framework described in Equation (3), without incorporating lagged variables. Alternatively, LASSO-ECM2 aims to capture more comprehensive short-run dynamics by incorporating lagged stationary variables. In essence, the short-run model ultimately seeks to identify a more comprehensive set of short-run variables. LASSO-ECM2 is explained further in the subsequent sections.

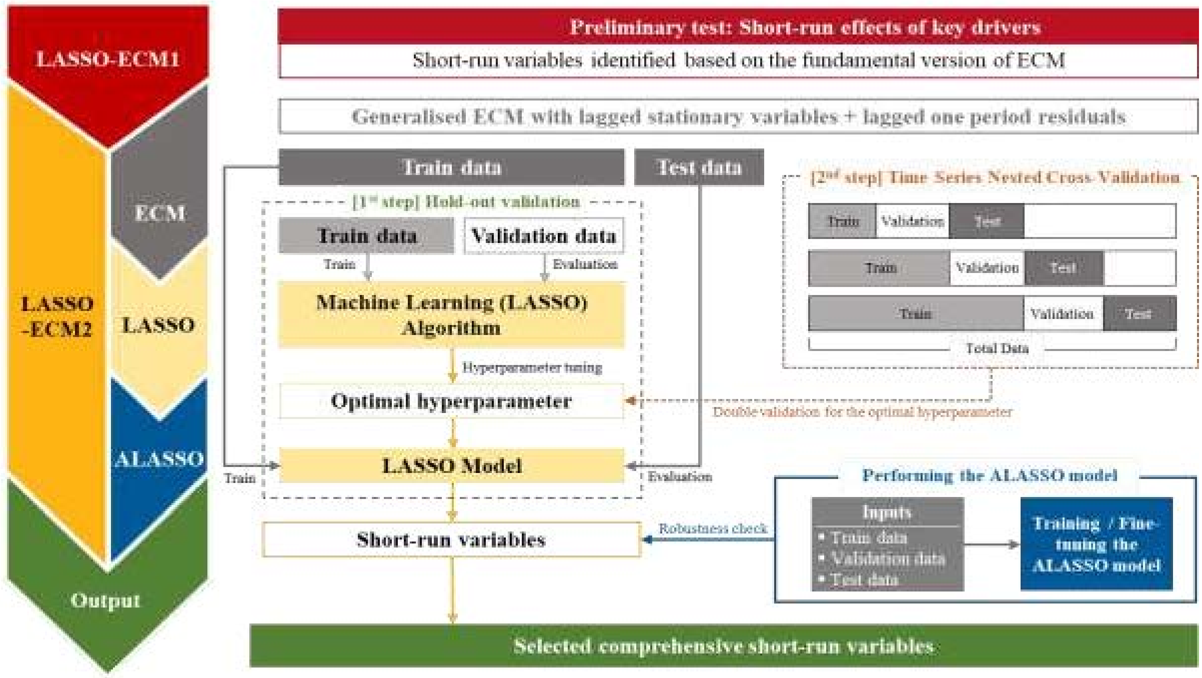


Figure 2.3 Short-run model

2.4.2.1 LASSO-ECM2

LASSO-ECM2 provides estimates for identifying more comprehensive short-run variables by applying LASSO/ALASSO techniques to lagged stationary variables, formulated in the structure of a generalised ECM. The generalised ECM is mathematically expressed as follows:

$$\boldsymbol{\Delta Y_t = \beta_0 + \sum_{i=0}^{n} \beta_i \Delta X_{t-i} + \sum_{j=1}^{m} \alpha_j \Delta Y_{t-j} - \gamma \hat{\mu}_{t-1} + e_t} \qquad \textbf{(4)}$$

where α_j and β_i are short-run coefficients; n and m are lags of variables; $\hat{\mu}_{t-1}$ is the lagged one-period residuals estimated from the LRM of two different series X_t and Y_t; γ is the coefficient of ECT; e_t is the white noise error term; β_0 is the constant term; and Δ represents the first difference operator.

In general, while raw time-series variables (at zero lag) might not impact a system in the short term, lagged variables can exert such effects. Incorporating lagged variables enables researchers to evaluate how past values impact the current system, improving their understanding of its dynamics [56, 57]. Hence, to identify more comprehensive short-run variables, a generalised ECM with lagged stationary variables is first formulated. Subsequently, following the general-to-specific approach,[5] LASSO is applied to eliminate less essential independent variables in a data-driven manner. Finally, ALASSO provides a robustness check

[5] The general-to-specific approach is an econometric methodology designed to simplify a general model that initially includes all potentially relevant variables. This simplification is achieved by systematically eliminating insignificant or irrelevant variables [58].

for the LASSO outputs using its oracle properties. In summary, this analytical approach—LASSO-ECM2—enables us to delve into more comprehensive short-run dynamics by introducing additional variables. This analysis is conducted using Python.

2.5 Results and discussion

2.5.1 Long-run model results

Long-run model results comprise three components: (1) identification of the key driving factors by LASSO; (2) verification of their long-run relationship through the residual-based cointegration test and validated by ALASSO; and (3) quantification of the relationship using LASSO-ECM1.

2.5.1.1 Key driving factors

Three phases are involved in LASSO modelling for selecting key driving factors: (1) data pre-processing; (2) training, tuning, and evaluation; and (3) final performance and evaluation.

In the first phase, data pre-processing, the original time-series data is split into a training set (60%), validation set (20%), and test set (20%) using a fixed random seed of 279. This seed, referred to as the *random_state* parameter in ML, represents the initial value used when shuffling data randomly, ensuring consistent output when the same ML function is run multiple times [59]. The same random_state parameter is constantly used thereafter. All input sets are standardised and scaled to unit variance.

In the second phase—training, tuning, and evaluation—the training set is used to fit a LASSO model, while the validation set is utilised to evaluate the model for hyperparameter tuning, that is, determining an optimal hyperparameter (α). The hyperparameter tuning is processed in two steps:

- **Step 1**: α is identified using hold-out validation [53], which entails holding out a portion of the training set as a validation set, which is then used to assess several candidate models. First, these models are trained with different hyperparameters on the reduced training set. Then, α is selected through a graphical analysis of metrics, such as mean absolute error (MAE), root mean square error (RMSE), and R-squared (R^2) on the validation set. Figure 2.4 indicates that α selected for the highest R^2 and lowest MAE and RMSE is 0.1.

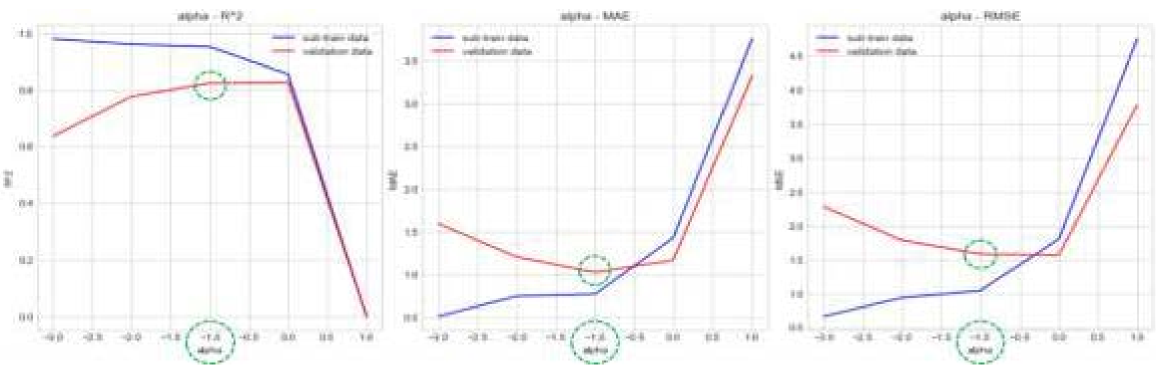


Figure 2.4 Optimal hyperparameter (α) by hold-out validation

Notes: A base-10 log scale is utilised for the Y-axis (‘α’ values) of all the above graphs. The α value is 0.1.

- **Step 2**: To validate the α chosen in the first step, time series nested cross-validation

(TSNCV) is employed [60]. TSNCV involves performing the train/validation/test splits of the original data in a time-ordered manner. A robust estimate of the model error (RMSE) is computed by training and tuning the model on the train/validation set of each fold and averaging the errors on the test sets. α is calculated based on RMSE. α calculated through TSNCV also yields the same value, 0.1. Figure 2.5 illustrates TSNCV process.

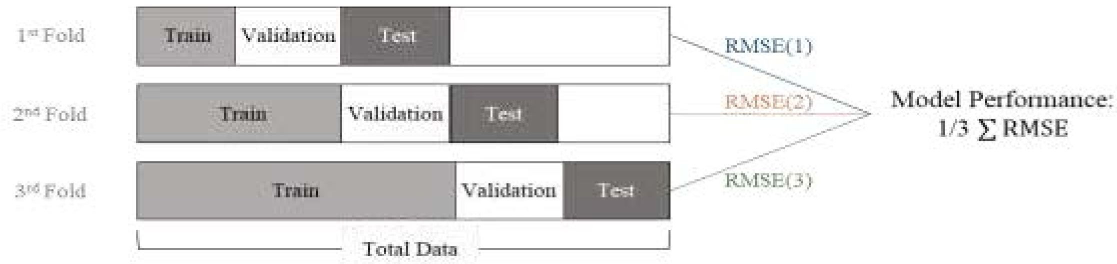
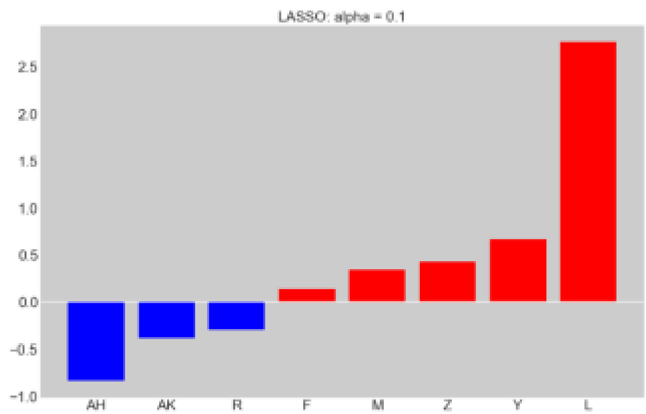


Figure 2.5 Process of time series nested cross-validation

In the third phase—final performance and evaluation—the best model with the optimal hyperparameter (α), selected through hold-out validation and TSNCV, is trained on the full training set (including the validation set), using the fixed random seed. This model is considered the finalised LASSO model, which is then used to evaluate the performance of the test set.

The finalised LASSO model, achieving the high accuracy (R^2) of 93.84% (on a train set) and 93.53% (on a test set), selects eight key driving factors among the 40 factors. Among these five factors—domestic thermal coal price, NG imports via pipelines, gas pipeline capacity, LNG terminal capacity, and LNG imports—are positively associated with NG consumption. The remaining three factors—temperature, price distortion, and HH gas price—exhibit a negative relationship with NG consumption. These results are illustrated in Figure 2.6, where red and blue bars represent positive and negative relationships, respectively.



#		Negative/Positive relationship
		Negative relationship
1	AH	Average national temperature (°C, Celsius)
2	AK	Price distortion (City gate gas price / Piped gas price, %)
3	R	HH gas price (USD/MMBTU)
		Positive relationship
4	F	Domestic thermal coal price: Datong (5500 kcal/kg; USD/MT)
5	M	NG imports via pipelines (Metric tonne)
6	Z	Gas pipeline capacity (BCM)
7	Y	Accumulated LNG terminal capa. per year (10KT)
8	L	LNG imports (Metric tonne)

Figure 2.6 LASSO results (α = 0.1): Key driving factors

Note: Figure 2.6 displays only variables with non-zero coefficients and its y-axis shows coefficient values.

2.5.1.2 Long-run relationship

The long-run relationship between the selected factors and NG demand is estimated using a cointegration test and ALASSO. The cointegration test preliminarily investigates whether a long-run relationship exists, whereas ALASSO validates the robustness of the cointegration test. This process involves three phases: (1) cointegration test, (2) ALASSO-specific data pre-processing, and (3) training and evaluating phases in ALASSO.

First, the cointegration test is executed based on the 1st step of the Engle–Granger two-step procedure [47]. An LRM is constructed from the key driving factors using the ordinary least squares (OLS) method. Residuals from the LRM are then calculated and two widely-used unit root tests—ADF and KPSS—are applied to these residuals, the independent variables (key driving factors), and the dependent variables (NG consumption) to determine whether these variables exhibit stationarity. As shown in Tables 2.3 and 2.4, the residuals are stationary at level and all variables are stationary at the first difference, indicating cointegration and thus a long-run relationship among the key driving factors [47].

Table 2.3 LRM and Unit root test results

LRM results								
Coefficients of independent variables (key driving factors)								
F	L	M	R	Y	Z	AH	AK	C
0.0327	1.625e-06	4.809e-07	-0.8487	0.0011	0.0366	-0.1039	-2.7543	7.0185
Dependent Variables: NG consumption				Method: OLS			R^2 (Adjusted R^2) : 0.944 (0.938)	

Unit root test results (ADF/KPSS) of the residuals		
	Unit root test	Description
ADF	P-value: 0.00 (ADF statistic: -7.90)	Based on the significance level of 0.05 and the p-value of ADF, H_0 can be rejected. Hence, the residuals are stationary at level.
KPSS	P-value: 0.10 (KPSS statistic: 0.03)	Based on the significance level of 0.05 and the p-value of KPSS, H_0 cannot be rejected. Hence, the residuals are stationary at level.

Notes: In the ADF test, H_0 (null hypothesis) represents that the series has a unit root (non-stationarity), while H_1 (alternative hypothesis) represents that the series has no unit root (stationarity). In the KPSS test, H_0 denotes that the series is stationary, while H_1 denotes that the series is not stationary. Table 2.4 shows notations (F, L, M, R, Y, Z, AH, and AK) denoting relevant factors. 'e-0N' means ten to the minus 'N' power (i.e. 1.625e-06 = 1.625×10^{-6}). 'C' of OLS regression results represents a constant.

Table 2.4 ADF test result of independent and dependent variables

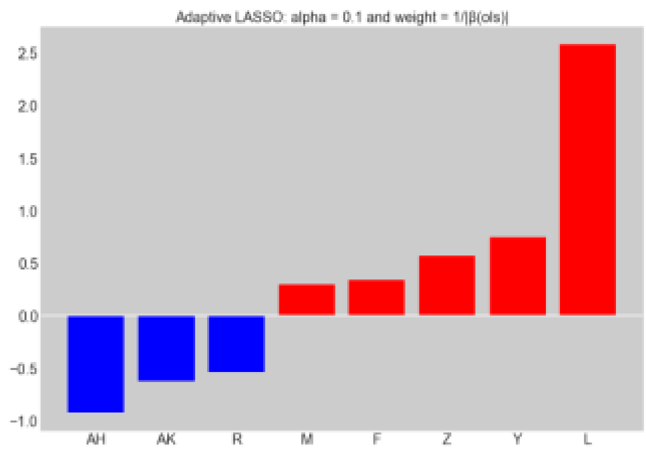
Factor		**t-Statistic**				**P-value**
		Test critical values			**ADF statistic**	
		1% level	5% level	10% level		
D(F)***	Domestic thermal coal price	- 4.085092	- 3.470851	- 3.162458	- 6.252749	0.0000
D(L)*	LNG imports	- 2.601024	- 1.945903	- 1.613543	- 1.932583	0.0515
D(M)***	NG imports via pipelines	- 4.083355	- 3.470032	- 3.161982	- 16.28665	0.0001
D(R)***	HH gas price	- 4.083355	- 3.470032	- 3.161982	- 10.05880	0.0000
D(Y)***	Accum. LNG terminal capa. per year	- 4.083355	- 3.470032	- 3.161982	- 7.152722	0.0000
D(Z)***	Gas pipeline capacity	- 4.083355	- 3.470032	- 3.161982	- 4.748410	0.0013
D(AH)***	Average national temperature	- 4.103198	- 3.479367	- 3.167404	- 9.971515	0.0000
D(AK)***	Price distortion	- 4.083355	- 3.470032	- 3.161982	- 9.983380	0.0000
D(NG)***	NG consumption	- 3.533204	- 2.906210	- 2.590628	- 7.644665	0.0000

Notes: ***, **, and * indicate ADF statistical significance at the 1%, 5%, and 10% levels, respectively. D ('factor') denotes the first difference of the factor. Based on the significance level of 0.01 and the p-value of ADF, independent (F, M, R, Y, Z, AH, and AK) and dependent (NG consumption) variables are stationary at the first difference. Based on the significance level of 0.1 and the p-value of ADF, an independent (L) variable is stationary at the first difference.

Second, in the ALASSO-specific data pre-processing phase, the original time-series data is divided into an 80% training set and a 20% test set, utilising the same fixed random seed employed in LASSO. All input sets are standardised and scaled to unit variance.

Finally, in the training and evaluating phases of ALASSO, adaptive weights ($\hat{w}_j$) are incorporated into a LASSO model to construct the ALASSO model, as per Equation (2). Following Zou [22], a LASSO model is initially solved, then OLS is used to compute the coefficients (selected by LASSO) denoted as $\hat{\beta}_j(ols)$. The adaptive weights are then calculated as the inverse of the absolute value of $\hat{\beta}_j(ols)$: $\hat{w}_j = 1/|\hat{\beta}_j(ols)|$.[6] The ALASSO model is fitted using the training sets, these weights, and the optimal hyperparameter from LASSO. The performance of the finalised ALASSO model is then evaluated on the test set.

As Figure 2.7 illustrates, the finalised ALASSO model attains a high accuracy (R^2) of 94.20% on a train set and 94.41% on a test set while selecting the same eight key driving factors as LASSO. Notably, as indicated in Table 2.5, the coefficients of the key driving factors in ALASSO differ from those in LASSO due to the adaptive weights. Nonetheless, the (positive/negative) relationships between the key driving factors and NG consumption in ALASSO align with LASSO. This implies that ALASSO results validate those of LASSO. Additionally, ALASSO confirms these selected factors' long-run relationship, as preliminarily indicated by the cointegration test [23].



#		Negative/Positive relationship
		Negative relationship
1	AH	Average national temperature (°C, Celsius)
2	AK	Price distortion (City gate gas price / Piped gas price, %)
3	R	HH gas price (USD/MMBTU)
		Positive relationship
4	M	NG imports via pipelines (Metric tonne)
5	F	Domestic thermal coal price: Datong (5500 kcal/kg; USD/MT)
6	Z	Gas pipeline capacity (BCM)
7	Y	Accumulated LNG terminal capa. per year (10KT)
8	L	LNG imports (Metric tonne)

Figure 2.7 ALASSO results ($\alpha = 0.1$; $\hat{w}_j = 1/|\hat{\beta}_j(ols)|$): Key driving factors

Table 2.5 Coefficients by LASSO vs ALASSO

Model	Coefficient values (= y-axis in Figure 2.7)							
	AH	AK	R	F	M	Z	Y	L
LASSO	- 0.829377	- 0.381393	- 0.291971	0.139179	0.345636	0.430412	0.672438	2.758738
ALASSO	- 0.916028	- 0.619827	- 0.538088	0.338705	0.293858	0.573565	0.74688	2.577915

Notes: Figure 2.7 shows notations (F, L, M, R, Y, Z, AH, and AK) denoting factors and displays only variables with non-zero coefficients.

2.5.1.3 Quantification of long-run relationship

LASSO-ECM1 represents a fundamental version of ECM, using the first-differenced key driving factors and lagged first-differenced NG consumption as independent variables, and the first-differenced NG consumption as the dependent variable. This model reaffirms the long-run relationship of the selected factors, as indicated by ALASSO.[7] As shown in Table 2.6, the lagged ECT in LASSO-ECM1 has a coefficient (γ) of -1.1646, which is negative and statistically significant at the 1 % level. According to Narayan and Smyth [55], a γ value between -2 and -1 suggests ECT convergence towards the long-run equilibrium in a dampening manner. This confirms that the selected factors have long-term effects on China's NG consumption. The γ value also reveals their short-term effects [55], further discussed in the

[6] Very small numbers (1.00e-20), close to zero, are used as substitutes for zero coefficients, which appear as denominators in the weight calculation.
[7] LASSO-ECM1 can be conducted once the cointegration test results are satisfied.

following section.

2.5.2 Short-run model results

LASSO-ECM1 serves as a preliminary test of the short-run model, indicating that the key driving factors have short-term effects on China's NG consumption. However, as Table 2.6 shows, temperature, domestic thermal coal price, and NG imports via pipelines among these selected factors are not statistically significant at the 10% level. These factors may be excluded from short-run variables in this model. This limitation may stem from the inherent nature of a fundamental ECM version, which excludes lagged variables. In case a more extensive set of lagged variables is utilised to analyse the short-run variables, it is possible that the variables excluded or not identified in the preliminary test (LASSO-ECM1) could exhibit short-term effects [56, 57]. Therefore, for a more comprehensive examination of potential short-run variables, this study simultaneously combines LASSO/ALASSO modelling with stationary lagged variables based on the generalised ECM (LASSO-ECM2).

Table 2.6 LASSO-ECM1 results: Short- and long-term effects of key driving factors

Variables		LASSO-ECM1 outputs			
		Coefficient	Std. Error	t-Statistic	P-value
D(AH)	Average national temperature	- 0.05959	0.033334	- 0.178754	0.8587
D(AK)*	Price distortion	- 2.955824	1.592273	- 1.856355	0.0679
D(F)	Domestic thermal coal price	0.0039266	0.035823	1.096104	0.2771
D(L)***	LNG imports	1.88e-06	2.20e-07	8.541529	0.0000
D(M)	NG imports via pipelines	3.38e-07	3.06e-07	1.104004	0.2737
D(R)*	HH gas price	- 0.533474	0.308705	- 1.728105	0.0887
D(Y)***	Accumulated LNG terminal capa. per year	0.002841	0.000994	2.857464	0.0057
D(Z)***	Gas pipeline capacity	0.079203	0.024306	3.258604	0.0018
D(NG)(-1)	Lagged one-period NG consumption	0.115895	0.087708	1.321375	0.1910
ECT***	Lagged error correction term	- 1.164600	0.159314	- 7.310104	0.0000
C	Constant	- 0.133114	0.130626	- 1.019050	0.3120

Notes: ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. D ('factor') denotes the first difference of the factor. 'e-0N' means ten to the minus 'N' power (i.e. 1.88e-06 = 1.88 × 10^{-6}). The series '(-1)' refers to the series lagged by one period. The lagged ECT represents lagged one-period residuals.

2.5.2.1 A more comprehensive set of short-run variables

LASSO-ECM2 involves four main steps: (1) creating all lagged stationary variables based on the generalised ECM; (2) LASSO-specific data pre-processing; (3) training, tuning, and evaluation phases within LASSO; and (4) final performance and evaluation of LASSO and robustness check of ALASSO.

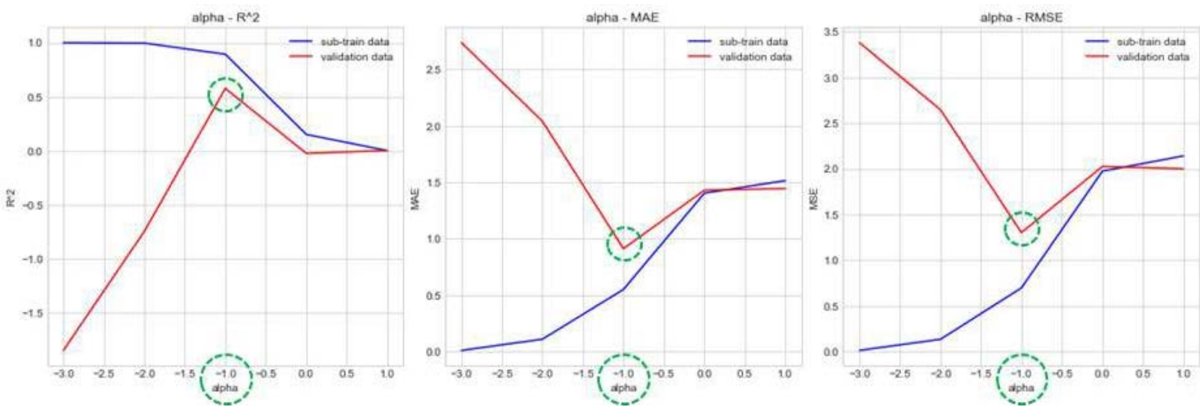


Figure 2.8 Optimal hyperparameter (α) by hold-out validation

Notes: A base-10 log scale is utilised for the Y-axis ('α' values) of all the above graphs. The optimal α value is 0.1.

In the initial step, 118 variables are prepared for data pre-processing. This comprises 117 lagged stationary variables derived from both independent and dependent variables (as detailed in Section 2.3.1 and Table 2.2; see Appendix C), along with one lagged one-period residual variable from the LRM based on the key driving factors (as specified in Table 2.3). Given their stationarity, as indicated in Table 2.2, these variables are organised in line with the generalised ECM shown in Equation (4).

In the second step, the 118 variables are split into training, validation, and test sets, mirroring the process used for LASSO in the long-run model.[8]

In the third step, the training set is used to fit a LASSO model, whereas the validation set is used to evaluate the LASSO model for hyperparameter tuning—the process of determining an optimal hyperparameter (α). Similar to the LASSO in the long-run model, hyperparameter tuning employs hold-out validation and TSNCV. As shown in Figure 2.8, α is selected based on the highest R^2 and lowest MAE and RMSE on the validation set.

In the last step, the model is trained on the full training set, including the validation set, with the fixed random seed used previously and α selected through hold-out validation and TSNCV. This results in the finalised LASSO model that selects a more comprehensive set of short-run variables on the test set. ALASSO then verifies their robustness, replicating the long-run model's approach.

Table 2.7 LASSO-ECM2 results: Selected comprehensive short-run variables

#		Negative/Positive relationship
		Negative relationship
1	Res(-1)	Lagged one-period residuals
2	AE(-1)	[ESG] Environmental score
3	AI	[ESG] Country score for climate change (%)
4	AH(-1)	Average national temperature (°C, Celsius)
5	E	Chinese coking coal import price (USD/MT)
6	Q(-1)	WTI oil price (USD/bbl)
7	AN(-1)	Price distortion (City gate gas price / HH, %)
8	AM	Price distortion (City gate gas price / JCC, %)
9	K(-3)	NG import price via pipelines (USD/MMBtu)
10	J(-1)	LNG average import price (USD/MMBtu)
11	C(-3)	Trend restored CLI
		Positive relationship
12	F	Domestic thermal coal price: Datong (5500 kcal/kg; USE/MT)
13	A(-5)	Real GDP by expenditure (CNY, Billions)
14	AE(-3)	[ESG] Environmental score
15	M	NG imports via pipelines (Metric tonne)
16	Y	Accumulated LNG terminal capa. per year (10KT)
17	A(-1)	Real GDP by expenditure (CNY, Billions)
18	L(-1)	LNG imports (Metric tonne)
19	Z	Gas pipeline capacity (BCM)
20	K	NG import price via pipelines (USD/MMBtu)
21	L	LNG imports (Metric tonne)

Notes: The series ('- Number') denotes the series lagged by '#' period. Table 2.1 describes the calculation of price distortion factors

As presented in Table 2.7 and Figure 2.9 (see Appendix D), the finalised LASSO model selects 21 comprehensive short-run variables from 118 lagged stationary variables, achieving a high accuracy (R^2) of 87.96% on a training set and 83.47% on a test set. As depicted in Figure 2.10 (see Appendix E), ALASSO selects the same short-run variables. Like the long-run model, while ALASSO's coefficients differ from those in LASSO due to adaptive weights, the

[8] To produce consistent output, the pseudorandom number generator for data shuffling is provided with a fixed random seed of 200.

(positive/negative) relationships between the variables and NG consumption in ALASSO align with those observed in LASSO.

The 21 selected variables follow the generalised ECM structure and are considered to have short-run informative content. As shown in Table 2.7, the lagged one-period residuals, RES(-1), are indicated with a negative value, confirming the alignment of the 21 variables with the generalised ECM structure. As time-series variables transition from non-stationarity to stationarity, they may lose their inherent long-run relationship or information [24]. Hence, it can be asserted that the 21 selected variables carry short-run information and represent a more comprehensive set of short-run variables.

This approach also identifies short-term effects that are not statistically significant in the preliminary test. For example, the preliminary test does not demonstrate short-term effects between domestic thermal coal price and NG demand because, as demonstrated in Table 2.6, the coefficient of the thermal coal price is not statistically significant. However, a more comprehensive estimation of short-term effects, as presented in Table 2.7, includes the thermal coal price as a short-run variable. The following section explores these short-run dynamics in greater detail.

2.5.3 Further discussion: Unpacking temporal dynamics of NG drivers

This study addresses the limitations of previous research on the determinants of NG demand in China, which has struggled to identify key drivers due to heterogeneous criteria for variable selection and limited analysis of short- and long-term effects. To overcome these challenges, this study introduces a novel, data-driven decomposition framework based on time-series analysis to identify the key drivers of China's NG market. These drivers effectively bridge the short-term energy transition needs and long-term carbon neutrality goals by reflecting both the structural changes resulting from market liberalisation and the pursuit of carbon neutrality—aspects insufficiently addressed in previous studies. Moreover, while prior studies have primarily focused on economic growth, urbanisation, and energy structure, this framework incorporates variables that have been relatively overlooked, including ESG factors and global energy price indicators (e.g., WTI oil price, coking coal import price, and LNG average import price), thereby capturing a more comprehensive set of short-run dynamics affecting China's NG demand.

Specifically, the short-run model additionally selects ESG, coking coal import price, WTI oil price, piped NG import price, LNG average import price, trend restored CLI, and real GDP. Among all the short-run variables identified by this model, domestic thermal coal price, real GDP, environmental score, NG imports via pipelines, LNG terminal capacity, gas pipeline capacity, NG import price via pipelines, and LNG imports have a positive relationship with NG consumption. In contrast, ESG, temperature, coking coal import price, WTI oil price, price distortion, NG import price via pipelines, LNG average import price, and trend restored CLI exhibit a negative relationship with NG consumption. Revealing these short-run dynamics—many of which have not been thoroughly explored in the existing literature—this study offers a deeper, temporally sensitive understanding of China's NG demand, thereby making a significant academic contribution. Further exploration provides these dynamics as follows:

2.5.3.1 ESG factors

The short-run model selects two ESG factors—Environmental score and Country score for climate change—as additions to the long-run variables. A lagged three-period Environmental score shows a positive relationship with NG consumption, whereas a lagged one-period Environmental score and the Country score for climate change at zero lag exhibit negative relationships. These relationships are explained as follows:

- *The negative relationship between Country score for Climate change and NG consumption*: NG is a fossil fuel that generates CO_2 emissions. Although its CO_2 emission intensity is lower than that of coal, it contributes to climate change. Therefore, increasing NG consumption is negatively related to reducing the Country score for climate change.
- *The positive and negative relationships between Environmental score and NG consumption*: The positive relationship between NG consumption and the Environmental score lagged by three periods could be attributed to underlying factors or dynamics. There could be a delayed effect of changes in environmental practices or policies on NG consumption. Alternatively, the negative relationship between NG consumption and the Environmental score lagged by one period could indicate a more immediate impact. Methane leakage from NG consumption currently raises environmental concerns [9, 10], explaining the negative relationship for the Environmental score lagged by one period. Both the Environmental score and the Country score for climate change are considered environmental factors within ESG factors, hence these two scores exhibit the same (negative) relationship with NG consumption.

2.5.3.2 Chinese coking coal and domestic thermal coal prices

The Chinese coking coal import price, as an additional short-run variable, exhibits a negative relationship with NG consumption. Coking coal is a specialised fuel in Chinese steelmaking, and NG is not typically used as a substitute [61].[9] Higher coking coal prices indicate higher steel prices, leading to a decrease in steel demand and, consequently, a reduction in steel production. In China, the steel industry is one of the high-electricity-consuming industries [64]. Therefore, reducing steel production decreases electricity demand from steel mills, which impacts the overall demand for NG. This decreases demand for gas-fired power plants.

Contrastingly, the short-term model's domestic thermal coal price selected has a positive relationship with NG consumption, similar to the long-run model. Power plants typically use thermal coal and NG for electricity generation. This positive relationship suggests that higher thermal coal prices could result in an increased demand for NG as an energy-generating substitute.

2.5.3.3 WTI oil price and HH gas price

WTI oil price, an additionally chosen short-run variable, and HH gas price, which demonstrates its short-term effects at the 10% significance level in the short-run model's preliminary test, both exhibit negative relationships with NG consumption. Notably, HH gas

[9] In recent research, the substitution of coking coal in steel production is most likely associated with the application of net-zero breakthrough technologies, such as CCS and green hydrogen [62, 63].

and WTI oil prices exhibit an interconnection in the short and long term [65]. Thus, a connection can be inferred between WTI oil price and HH gas price, suggesting that the latter has short-run characteristics, although it is not selected as a comprehensive short-run variable, as demonstrated in Table 2.7.

2.5.3.4 Other energy-related factors

The short-run model's preliminary test reveals the short-term effects of three energy-related factors—gas pipeline capacity, LNG terminal capacity, and LNG imports—at the 1% significance level. The short-run model identifies NG imports via pipelines, along with these three factors, as comprehensive short-run variables that are integral components of the key driving factors. Additionally, all four of these energy-related factors exhibit a negative relationship with NG consumption, thus indicating their short-term negative effects.

The short-run model additionally chooses two energy price factors: LNG average import price and NG import price via pipelines. LNG average import price lagged by one period exhibits a negative relationship with NG consumption, indicating that LNG price changes in the previous period can influence NG consumer and industrial behaviour, causing fluctuations in NG consumption in the current period [66]. NG import price via pipelines lagged by three periods shows a negative relationship with NG consumption, whereas NG import price via pipelines at zero lag shows a positive relationship. The negative relationship indicates that changes in piped NG import prices take some time to influence NG consumption patterns attributable to factors, including term contracts, price adjustment mechanisms, inventory levels, economic decision-making processes, and seasonal variations in demand [14]. Alternatively, the positive relationship might be attributed to the immediate impact of fluctuations in NG import price via pipelines on NG demand during the same period, such as increased use of NG as a peak-shaving tool in response to periods of high electricity demand when NG prices are elevated [67].

2.5.3.5 Economic factors

Two additional short-run variables are selected—trend restored composite leading indicator (CLI) and real GDP—which are not identified in the short-run model's preliminary test and long-run model. The trend restored CLI variable lagged by one period, representing an economic activity indicator, has a negative relationship with NG consumption, indicating that an economic downturn in the short-term business cycles precedes increased NG consumption. The trend restored CLI suggests that low-frequency movements in CLI are driven by the output of specific industries and financial and market indicators, which can affect the short-run dynamics of NG demand [68].[10]

Conversely, real GDP variables lagged by one period or five periods, representing a measure of China's economic output, exhibit a positive relationship, suggesting that economic growth leads to higher NG consumption [70]. These relationships suggest that changes in economic activity influence NG demand in the short term.

[10] China's CLI components include outputs from key industries—such as chemicals, steel, construction, and motor vehicles—the diffusion index of 5,000 industrial enterprises, and the turnover of the Shanghai Stock Exchange. [69].

2.5.3.6 Temperature and price distortion

Similar to the long-run model, the short-run model chooses temperature and price distortion factors. The short-run model's preliminary test reveals only the short-term effects of the price distortion factor at the 10% significance level. However, as a more comprehensive set of short-run variables includes all these factors, it can be concluded that these factors have short-term effects.

In brief, a more comprehensive short-run dynamic analysis helps us gain understand the dynamics of both the previously selected and additional short-run variables. Some variables that are statistically nonsignificant in the short-run model's preliminary test still exhibit short-run effects in its final estimation. Therefore, the key driving factors have both short- and long-run effects.

2.6. Conclusions and policy implications

China's DCGs place NG in a paradoxical position: it is indispensable as a short-term transitional fuel but poses a long-term risk to carbon neutrality. This tension is compounded by short-term market volatility and long-term uncertainties, underscoring the need to identify key drivers in China's NG market that reconcile these conflicting temporal demands. Leveraging advanced analytical methods, this study offers a systematic framework for isolating the most relevant market drivers. The insights derived from this framework lay the groundwork for policies and strategies that effectively balance immediate energy needs with long-term sustainability.

This study employs a novel temporal shrinkage framework integrating big data analytics (LASSO/ALASSO) with advanced time-series modelling (ECM), through two models: the long-run model (sequentially combining LASSO/ALASSO and ECM) and the short-run model (simultaneously incorporating both ECM and LASSO/ALASSO). The long-run model initially identifies key drivers in the NG market—average national temperature, price distortion, HH gas price, NG imports via pipelines, domestic thermal coal price, gas pipeline capacity, LNG terminal capacity per year, and LNG imports—and assesses their long-term effects. Subsequently, the short-run model shows that three of these drivers (temperature, domestic thermal coal price, and piped NG imports), although not confirmed as having short-term effects in the model's preliminary test, indeed exhibit such effects. Additionally, it identifies additional short-run variables, such as energy commodity prices, ESG, and economic factors. Together, these models clarify the key drivers' short- and long-term effects.

The study not only introduces a novel temporal shrinkage framework for energy policy analytics but also demonstrates ALASSO's extended capability in accurately selecting cointegrated variables. This framework leverages the inherent long-term information within the data without disrupting it or encountering spurious regression [24]. Furthermore, the study results offer policymakers insights into the key drivers, enabling the formulation of initiatives that balance short- and long-term effects to advance DCGs for China's sustainable development. Based on these findings, four policy implications for China's ambitious carbon-neutral initiatives are proposed.

First, environmental costs should be internalised separately for each fossil fuel in the short and long term (for coal until 2030 and LNG from 2030 to 2060) and utilised to promote net-

zero breakthrough technologies. This study indicates that, in both the short and long term, LNG imports and thermal coal prices positively correlate with NG consumption. In the short term, increasing domestic thermal coal prices can encourage a transition to NG, supporting a key policy in China's 2030 carbon peaking strategy—the 'coal-to-gas' switching in the 14th Five-Year Plan [3, 8]. Implementing environmental pricing mechanisms, such as a carbon tax, could make thermal coal more expensive and boost NG consumption [36]. In the long term (covering the period from its 2030 carbon peak to its 2060 carbon neutrality), these pricing mechanisms should be expanded to LNG and other sources to facilitate NG phase-out and the transition away from carbon-intensive energy [71]. Among the key long-term drivers, LNG imports have the largest LASSO coefficient in relation to NG consumption, confirming their critical role. Therefore, gradually imposing environmental costs on LNG import contracts as a mandatory measure could internalise externalities and reshape long-term contracting structures, thus reducing LNG import volumes. In the long term, this policy measure could enable the fundamental phase-out of NG.

Simultaneously, pricing mechanisms should support investments in net-zero breakthrough technologies—including renewables and hydrogen—across all periods, boosting their competitiveness over fossil fuels and complementing their phase-out [72]. This may enable an ambitious carbon-neutral pathway—the complete phase-out of fossil fuels through the technologies [62]—similar to Denmark's case, which now sources 89% of its electricity from renewables [73] and aims to phase out NG entirely by 2030 [74]. This pathway could serve as a novel feasible guideline for achieving China's carbon neutrality.

Second, NG import infrastructure should be optimised according to its short- and long-term usage purposes. The study finds that gas-related infrastructure has positively affects NG consumption in both the short and long term. Alongside LNG imports, it has a large LASSO coefficient, suggesting that it is a major determinant in NG consumption changes. In this context, infrastructure developed to support NG in alignment with China's 2030 carbon peak should shift to infrastructure that decreases NG consumption to align with carbon neutrality by 2060. In the short term, China could enhance the NG supply chain's capability and flexibility by diversifying supply sources and developing new gas infrastructure, such as pipelines and LNG terminals, to increase NG demand. In the long term, this infrastructure can be repurposed for hydrogen, which is emerging as an essential energy source towards carbon neutrality [75].

Third, price distortion in the NG market should be minimised. This study demonstrates that price distortion negatively impacts NG demand in both the short and long term, with a large LASSO coefficient, smaller than temperature but still impactful. In the short term, it enables consumers in China to choose cheaper energy sources like coal over NG, creating a supply–demand imbalance in the NG market. A market-based NG pricing system could promote NG use over carbon-intensive and inexpensive energy sources [36], aligning with China's short-term strategy for carbon neutrality through the coal-to-gas transition. In the long term, this market-based pricing system would align domestic gas markets with global pricing, allowing for environmental cost reflection [76]. This could reduce NG competitiveness against renewables [77, 78] and eventually facilitate the full phase-out of fossil fuels [79].

Finally, ESG criteria should be immediately integrated into energy policy. The study highlights the significance of ESG factors as a short-term driver of China's NG demand. While

ESG reporting is not yet mandatory in China, regulations requiring large public companies to report ESG metrics are expected by 2026 [80]. Additionally, certain carbon neutrality targets have seen limited progress [81, 82]. Therefore, policymakers should immediately incorporate ESG criteria into energy-related decisions and regulations to meet China's short-term low-carbon transition goals.

The study analyses key NG market drivers through the lens of China's 2030 carbon peak (short-term) and 2060 carbon neutrality (long-term) goals. However, some forecasts suggest China may not peak emissions until 2040 (medium term) [83]. It may be feasible to engineer medium-term policies based on identified long-term variables, but this introduces a methodological limitation. The proposed ECM framework primarily models short-term adjustments towards the long-term equilibrium of the variables and is thus not suitable for quantifying medium-term effects. On the transition path towards the medium-term goal, one cannot rule out the possibility that the long-term equilibrium will remain constant. Accordingly, a potential extension of our ECM is to allow for time variation in the cointegration relationship following Bierens and Martins [84] and Koop et al. [85]. Such extension warrants further investigation.

The study's temporal shrinkage framework (LASSO/ALASSO and ECM) reflects a reduced-form design aimed at identifying China's key natural gas drivers across short- and long-term perspectives rather than uncovering underlying mechanisms. Therefore, it does not seek to establish causal relationships. Causal inference is beyond the scope of this chapter but is acknowledged as another limitation. Future research will incorporate advanced econometric techniques such as VAR models and impulse response analysis, along with machine learning-based approaches suggested by Brand et al. [86]. Sectoral consumption analysis will also be explored to refine causal understanding.

An additional limitation is that LASSO-ECM2, while effective for short-term analysis, is restricted in examining long-term effects. Although the selected negative RES(-1) in Table 2.7 may suggest a long-run relationship, LASSO-ECM2 can distort the proper relationship between independent and dependent variables by shrinking less important coefficients towards zero, resulting in a biased coefficient for RES(-1) [87]. Additionally, LASSO-ECM2 cannot evaluate standard errors for the biased coefficient, indicating that the coefficient might not be statistically significant. Furthermore, LASSO-ECM2 does not use the ECM techniques for long-term effects. Future research should employ additional statistical tools to investigate these effects more comprehensively.

CHAPTER 3.0

CARBON-NEUTRAL PLANNING FRAMEWORK FOR CHINA'S NEAR-TERM NATURAL GAS DEMAND: A NEURAL-ENHANCED HYBRID STATISTICAL–DEEP LEARNING FORECASTING APPROACH

This chapter is based on the manuscript: Sung, J., Shi, X., Teske, S., and Li, M. (2025), Carbon-Neutral Planning Framework for China's Near-Term Natural Gas Demand: A Neural-enhanced Hybrid Statistical–Deep Learning Forecasting Approach. The manuscript is currently under review at a peer-reviewed journal. Author contributions are detailed in the Declaration of Publications and Author Contributions section of this thesis.

Chapter outlook

Chapter 2 identifies eight key drivers underlying the structural transformation of China's NG market, primarily driven by carbon neutrality targets and market liberalisation. These drivers encompass a range of variables related to climate, global energy, and infrastructure, capturing both short-term fluctuations and long-term structural trends. Considering these drivers, China's NG consumption is expected to increase through 2030, in line with the nation's carbon neutrality policy—its coal-to-gas transition strategy. However, China's NG market faces significant uncertainties due to extra volatility caused by geopolitical developments and the ongoing NG market liberalisation. In response to the growing need for accurate forecasting of NG demand within the context of China's carbon neutrality policy, this study develops a hybrid model that optimally integrates a statistical approach—Seasonal Autoregressive Integrated Moving-Average with eXogenous regressors (SARIMAX)—with a deep learning technique—Long Short-Term Memory (LSTM)—via an Artificial Neural Network (ANN) for one-year-ahead forecasting of Chinese NG consumption, utilising these eight drivers of the NG market. This study evaluates the forecasting performance of all models—single SARIMAX, LSTM, and hybrid models—based on widely accepted statistical measures of predictive accuracy. The results of the study indicate that ANN-optimised SARIMAX-LSTM combination, with the trend-capturing function, produces the best predictive accuracy. Furthermore, the study not only confirms the suitability of these drivers with varying characteristics as inputs for hybrid models through Principal Component Analysis, but also identifies NG-related infrastructure factors as the most influential contributors to near-term forecasting. These findings offer practical implications for policy, investment, and operational planning in the context of carbon-neutral energy transitions.

Keywords: Hybrid Forecasting Methods; SARIMAX; LSTM; ANN; Natural Gas; China

3.1 Introduction

Natural Gas (NG) has become a crucial energy component for sustainable development in China. The country has set a goal of achieving carbon neutrality by 2060 as part of its sustainable development agenda [88]. In this journey towards carbon neutrality, NG not only contributes to reducing carbon emissions by substituting coal, a major energy source of carbon emissions in China, but also serves as a bridge to future energy sources such as renewable energy. This leads to continued growth in China's NG market in the medium and short term [89]. However, the NG market has been experiencing significant uncertainties due to recent geopolitical developments, including Russia's invasion of Ukraine and the Israel-Hamas conflict. Besides, the liberalisation of the NG market, which causes its structural changes, has also exacerbated these uncertainties. In this context, accurate near-term forecasting of NG consumption becomes essential. It not only supports more resilient energy planning but also enables timely policy and investment decisions that align with China's carbon neutrality goals. To address this need, this study aims to develop a high-performing one-year-ahead forecasting model using Chinese NG consumption as a case to examine its merits. The resulting insights are expected to be highly valuable for energy industry stakeholders seeking to make carbon-neutral-based near-term investment and operational decisions.

Traditional Auto-Regressive Integrated Moving Average (ARIMA)-based and deep learning models have evolved in time series forecasting, with Long Short-Term Memory (LSTM) deep learning models demonstrating significant improvements in prediction performance. Time series forecasting techniques are traditionally used in various fields: econometrics, weather, statistics, earthquake, signal processing, astronomy, energy consumption, etc [90-92]. Influenced by advancements in computer science and artificial intelligence technology, energy forecasting techniques are experiencing rapid growth through the emergence of deep learning models and the revival of conventional models [92]. Representative traditional time series forecasting techniques include ARIMA-based models, while deep learning models encompass Artificial Neural Networks (ANNs), Recurrent Neural Networks (RNNs), and LSTM models. Elamin et al. employ a Seasonal ARIMA with eXogenous variables (SARIMAX) model, an ARIMA-based model that includes seasonality and exogenous variables, to forecast hourly load demand data. They compare a SARIMAX model containing only main effects with a SARIMAX model incorporating both main and cross-effect interactions [93]. They conclude that combining the interaction effects of the exogenous variables in the SARIMAX model can lead to more precise and reliable forecasts. Manigandan et al. utilise both SARIMAX and SARIMA models to forecast monthly NG production and consumption in the U.S. until 2025, comparing the performance of each model [94]. This study indicates an expected increase in the U.S. NG production (16%) and consumption (24%) by 2025. Vagropoulos et al. compare forecasting for grid-connected photovoltaic plants using SARIMA, SARIMAX, modified SARIMA (resulting from a posteriori modification of the SARIMA model), and ANN-based models [95]. In this research, SARIMAX, modified SARIMA, and ANN-based models demonstrate superior day-ahead forecasting performance. In recent years, LSTM, a cutting-edge RNN-based deep learning model, has emerged as a powerful tool for forecasting, outperforming traditional models like

ARIMA-based models in various fields, such as energy load, flood, housing price, stock price, and LNG price [96-101].

Beyond individual comparisons of traditional forecasting and deep learning models, existing studies have highlighted the superior performance of hybrid models that combine the two. Comparing ARIMA that is designed for capturing linear relationships with ANN that is tailored for capturing non-linear relationships, a recent study presents various hybrid ARIMA-ANN combinations, including additive and multiplicative hybrid models, and demonstrates that such hybrid models show improved forecasting abilities over the individual models [91]. Cerit et al. employ SARIMAX, ANN, and hybrid SARIMAX-ANN models for seaborne transport forecasting, showing that the hybrid model outperforms the others [102]. Dave et al. utilise a hybrid ARIMA-LSTM model for forecasting Indonesia's exports, indicating that the hybrid model produces the lowest error metrics compared to standalone models [103].

However, the hybrid models in the existing studies cannot be sufficient for predicting China's near-term NG consumption, influenced by various external factors such as economic conditions, weather-related events, and geopolitical developments [104] as well as market instability caused by endogenous factors such as gas price regulations [105]. These factors can engender irrational patterns or substantial fluctuations in the NG consumption, which existing hybrid models cannot effectively capture [105]. A recent study utilises ANN, LSTM, and support vector regression (SVR) models, along with hybrid models integrated with SARIMAX, specifically SARIMAX-ANN, SARIMAX-LSTM, and SARIMAX-SVR models, for highly volatile peak-load forecasting [106]. The study finds no significant performance difference between single and hybrid LSTM models. This indicates that the existing combination methods between traditional forecasting and deep learning models have limitations in forecasting the energy market with substantial fluctuations. This limitation underscores the need for a novel hybrid approach, incorporating carefully selected external drivers that reflect not only China's carbon neutrality goals and supporting NG policy directions but also structural changes in its NG market, to accurately predict its NG consumption under significant volatility—a task beyond the capabilities of existing hybrid prediction models.

This research seeks to create a hybrid model optimally combining statistical and deep learning approaches for one-year-ahead NG consumption prediction. This hybrid model utilises eight key drivers as inputs, identified from 40 factors related to China's NG market through a novel combination of machine learning and econometric techniques—Least Absolute Shrinkage and Selection Operator (LASSO)/Adaptive LASSO (ALASSO) and Error Correction Model (ECM) techniques [107]. The hybrid model employs a SARIMAX model with a trend-capturing function and LSTM to compute linear and non-linear predictions, and then optimally combines these using ANN to forecast one-year-ahead NG consumption. The performance of this optimally combined hybrid model enabling the trend-capturing function is compared against single SARIMAX and LSTM models, an additively combined hybrid SARIMAX-LSTM model, which is based on the existing combination method, and hybrid models without the trend-capturing function. The results suggest that the optimally combined hybrid model with the trend-capturing function is the most suitable approach for accurately capturing unreasonable patterns or significant variability in one-year ahead NG consumption prediction. Moreover, this study not only assesses the suitability of these key drivers as inputs for the optimally combined hybrid model through Principal Component Analysis (PCA), but also explores the factors most significantly affecting the forecasting among the drivers.

The contribution of this study is twofold. First, the study introduces a novel hybrid forecasting model that optimally integrates a statistical time series model—SARIMAX—with a deep learning model—LSTM—via an ANN-driven fusion technique. This approach advances forecasting methodology in the field of energy economics and policy by enhancing near-term predictions of NG consumption to support carbon-neutral decision-making. Specifically, this hybrid model innovates near-term forecasting by not only optimally combining linear and non-linear components but also utilising an advanced pattern-capturing technique—such as the trend-capturing function—using selected key drivers as model inputs, which reflect near-term policy and market dynamics in China's NG market [107]. This innovative amalgamation can ultimately demonstrate its highest one-year-ahead predictive performance. Second, the study not only demonstrates that these key drivers with varying characteristics [107] are valid inputs for the hybrid model, but also identify NG-related infrastructure factors of these drivers as the most significant contributors to the near-term forecasting. The key drivers with different characteristics, consisting of price distortion, temperature, NG-related infrastructure, and other energy factors, are assessed for their suitability as inputs for the hybrid model using PCA. Furthermore, the study identifies NG-related infrastructure factors as the most influential factor in forecasting Chinese NG consumption by analysing the output's sensitivity to changes in key drivers. Based on these findings, the study suggests carbon-neutral policy or investment implications regarding near-term NG demand for government agencies or industry experts.

The rest of this chapter is organised as follows: Sections 3.2 and 3.3 discuss data and methods for the forecasting models, respectively. Sections 3.4 and 3.5 presents the models' application as the results and the further discussion of those results, respectively. The final section concludes the study. The supplementary materials include terms and abbreviations, tables, and figures used throughout this chapter.

3.2 Data

This study is conducted employing eight key driving factors as independent variables and Chinese NG consumption as the dependent variable.[11] These eight factors are identified from 40 factors related to the NG consumption through a novel combination of machine learning—LASSO/ALASSO—and econometric—ECM—techniques [107]. These eight factors are price distortion, temperature, Henry Hub (HH) gas price, Piped Natural Gas (PNG) imports, Thermal coal price, Gas pipeline capa, Liquid Natural Gas (LNG) terminal capa, and LNG imports. The factors are organised into the five groups of different characteristics. They are monthly time series variables on a real basis, ranging from April 2009 to July 2023 due to the availability of the price distortion factor. Table 3.1 provides details of these factors along with their respective sources.

The eight key driving factors, selected from the 40 factors, are important inputs for the near-term forecasting of this study. The 40 factors, classified into six groups[12], include economic developments, energy and carbon-neutral policies, environmental impacts,

[11] Monthly NG consumption variables (BCM) are sourced from two global data providers, Bloomberg and Refinitive.

[12] 40 factors in the six groups are (1) Economics & Energy (15 factors); (2) Global Energy Index (5 factors); (3) Urbanisation & Demographics (4 factors); (4) Infrastructure (4 factors); (5) Environment, Climate, and Technological growth (7 factors); (6) Price distortion (5 factors).

technological progress, and global market dynamics [107]. Particularly, these factors reflect NG's paradoxical role in achieving carbon neutrality—a short-term transitional fuel [3] and long-term phase-out [4]. Besides, these dynamics are influenced by the uncertainties addressed by these factors—the short-term uncertainties from NG market liberalisation [14, 25] and geopolitical developments [16, 26] and long-term uncertainties from electricity demand variations tied to carbon-neutral initiatives [17]. These eight factors are selected based on these carbon-neutral-driven shifting dynamics and NG market uncertainties, considering their temporal effects. Hence, the eight factors are highly applicable to the near-term projections.

Like several case studies, most of the variables are non-stationary, as confirmed by an ADF test, a widely used unit root test, in Table 3.2 (see Appendix F). All variables except price distortion are non-stationary at level and stationary at the first difference. SARIMAX can be applied to non-stationary variables [95, 108]. Deep learning models can learn complex patterns and dynamic relationships in data, allowing them to handle non-stationary variables [109]. Thus, these non-stationary key drivers can be directly utilised for hybrid SARIMAX-LSTM modelling. Detailed properties of the variables are in Table 3.2 (see Appendix F).

Table 3.1 Key driving factors

Factors	Description	Data source[13]
	(1) Energy (3 factors)	
Domestic thermal coal price (5500 kcal/kg; USD/MT)	Monthly series	Bloomberg
LNG imports (Metric tonne)	Monthly series	
NG imports via pipelines (Metric tonne)	Monthly series	
	(2) Global Energy Index (1 factor)	
HH gas price (USD/MMBtu)	Monthly series	Bloomberg
	(3) Infrastructure (2 factors)	
Accumulative LNG terminal capa. per year (10KT)	Monthly series	KEEI
Gas pipeline capacity (BCM)	Monthly series	Chen et al.'s research [110]
	(4) Climate (1 factor)	
Average national temperature (°C, Celsius)[14]	Monthly series	Bloomberg
	(5) Price distortion (1 factor)	
Price distortion (City gate gas price[15] ÷ Piped gas price[16], %)	Monthly series	Internal source[17], Bloomberg, and own estimates[18]

3.3 Methods and Models

The hybrid model, integrating SARIMAX with LSTM via ANN, utilises a methodology that sequentially calculates linear and non-linear predictions and then optimally combines these. As depicted in Figure 3.1, SARIMAX initially forecasts one-year-ahead linear components of Chinese NG consumption using eight key drivers lagged by one year as exogenous variables (Mx). Its role is to capture the predictable, linear relationships in the data. Subsequently, LSTM forecasts one-year-ahead non-linear components using past residuals (Mres). It is designed to

[13] The monthly exchange rates, referenced from Bloomberg, are applied.
[14] This factor is calculated as the average temperature of all provinces per month.
[15] City gate gas prices from December 2005 to June 2013 are based on the work of Paltsev and Zhang [36]. Prices from July 2013 to December 2019 are obtained from internal sources. From January 2020 onwards, the price is assumed to remain fixed at the December 2019 level.
[16] The piped gas price is referenced from Bloomberg.
[17] UTS ACRI provides the internal source.
[18] Price distortion variables are calculated using the formula: (City gate gas price ÷ Piped gas price − 1) ×100.

capture more complex, non-linear patterns that SARIMAX may not be able to account for. Finally, ANN optimally combines the linear and non-linear predictions to determine a one-year-head forecast of the NG consumption. Its role is to optimise the combined outputs of both models to enhance overall prediction accuracy. This entire modelling framework is implemented in Python, leveraging its open-source nature to ensure transparency, reproducibility, and accessibility. More details are provided in the following sections.

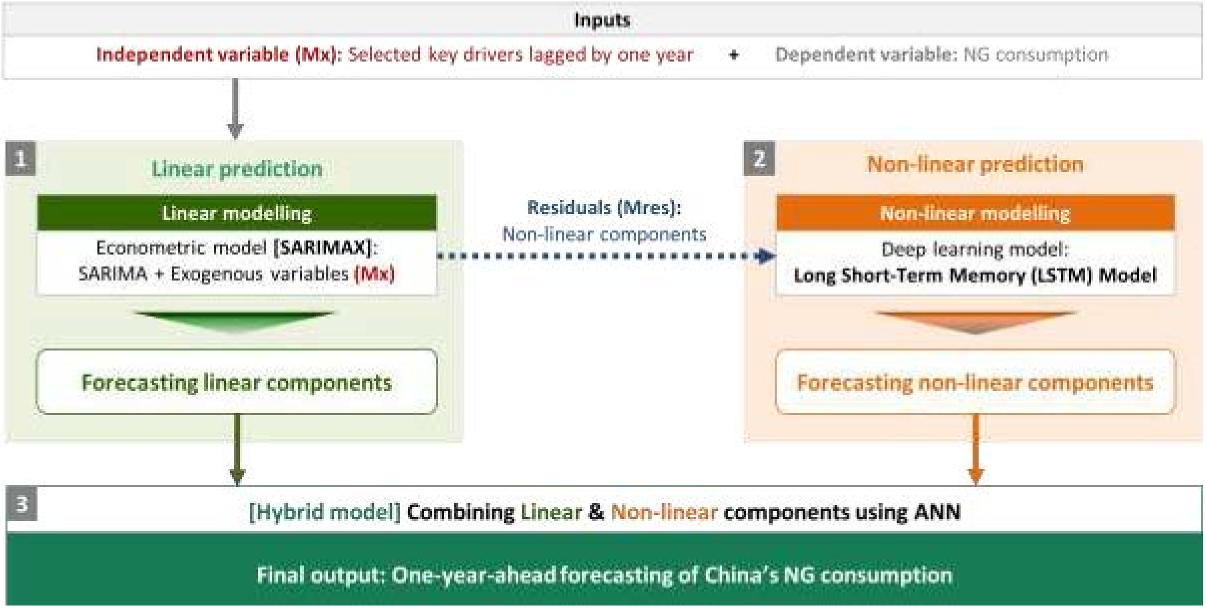


Figure 3.1 Flowchart of the hybrid model

3.3.1 Linear prediction

This section presents a concise overview of the SARIMAX model, which is employed for linear prediction and outlines its implementation in Python.

3.3.1.1 SARIMAX

SARIMAX is an extension of the Seasonal ARIMA (SARIMA) model that incorporates exogenous factors—denoted as X—and is formally denoted as SARIMAX (p, d, q) (P, D, Q)^S [94]. SARIMA represents an enhancement of the ARIMA framework, introduced to improve the performance of ARIMA models in capturing seasonal patterns in time series data [94]. Specifically, SARIMA can model repeating cycles or trends that occur at regular intervals using seasonal components. SARIMAX further extends SARIMA by incorporating exogenous regressors, allowing the model to account for external influences such as economic indicators, policy changes, and other relevant variables. Mathematically, SARIMAX can be expressed as:

$$\boldsymbol{SARIMAX}:\ \boldsymbol{\varphi_p(B)\phi_P(B^s)\nabla^d\nabla_s^D\, y_t = \beta_k x_{k,t}{}' + \theta_q(B)\Theta_Q(B^s)\varepsilon_t} \tag{1}$$

where:

- $\boldsymbol{\varphi_p(B)}$ $(= 1 - \sum_{i=1}^{p}\varphi_i B^i)$ is the characteristic AR polynomial of order p.
- $\boldsymbol{\theta_q(B)}$ $(= 1 + \sum_{i=1}^{q}\theta_i B^i)$ is the characteristic MA polynomial of order q.
- $\boldsymbol{\phi_P(B^s)}$ $(= 1 - \sum_{i=1}^{P}\phi_i B^{s,i})$ is the seasonal characteristic AR polynomial of order P.
- $\boldsymbol{\Theta_Q(B^s)}$ $(= 1 + \sum_{i=1}^{Q}\phi_i B^{s,i})$ is the seasonal characteristic MA polynomial of order Q.

- ∇^d $(= 1 - B)^d$ and ∇_s^D $(= 1 - B^s)^D$ denote the non-seasonal differencing operator and seasonal, respectively, where d and D denote the order of differencing and the seasonal differencing, respectively.
- y_t is the prediction variable at time t.
- $x_{k,t}{}'$ refers to the vector associated with the k^{th} exogenous input variables at time t, and β_k denotes the coefficient value of the k^{th} exogenous variable.
- B indicates the backshift operator (e.g., $B^s\ (y_t) = y_{t-k}$).
- s denotes the length of the seasonal pattern (e.g., s = 12 for monthly series).
- ε_t signifies the error terms.

3.3.1.2 Modelling process

The SARIMAX modelling process involves two sequential phases: data pre-processing and identification of the optimal order. After the two phases, SARIMAX forecasts one-year-ahead linear components of Chinese NG consumption. The two phases are as follows:

Phase 1: Data pre-processing

In the data pre-processing phase, all exogenous variables are first standardised to the range of -1 and 1 using MinMaxScaler, a widely adopted algorithm in both statistical and deep learning contexts [42].[19] The default range of MinMaxScaler is 0 to 1. Due to negative values in the data, this study adjusts the scaling range of MinMaxScaler to -1 and 1 [112, 113]. MinMaxScaler with the range of -1 and 1 is applied in the data pre-processing stage of all models in this study.

Subsequently, this study splits scaled exogenous and dependent variables into two sets – 80% as a training set and 20% as a test set – without shuffling, similar to other SARIMA and SARIMAX forecasting [94, 114]. Since this is a time series analysis where temporal order is crucial, the shuffling feature is turned off when dividing the data.

Phase 2: Identification of the optimal order

The optimal order of the SARIMAX model is typically chosen to model the characteristics of time series data and make the best predictions. To find the optimal order of SARIMAX, this study employs two Python libraries—auto ARIMA and statsmodels SARIMAX—with a training set. Initially, the optimal order is found by auto ARIMA, which automates the selection of the best SARIMAX model by identifying the model with the lowest AIC [115]. To demonstrate the robustness of the previously identified optimal order, the study subsequently use statsmodels SARIMAX, which manually selects the best SARIMAX model by specifying the model orders in advance. The optimal order is identified considering both the disabling and enabling of SARIMAX's trend-capturing parameter.[20] To find methods for enhancing forecasting accuracy, the study compares SARIMAX results without the trend-capturing parameter (referred to as SARIMAX 1) and SARIMAX with the parameter (referred to as SARIMAX 2). After finding the optimal order, Ljung-Box and Jarque-Bera tests are conducted on the fitted model with the optimal order. These tests are used as diagnostic tools for assessing the adequacy of SARIMAX models through their residuals, as shown in Table 3.3 (see

[19] MinMaxScaler is implemented in scikit-learn, a widely used Python library for machine learning [111].
[20] When the parameter is enabled, 'trend = ct' is set in SARIMAX, where 'ct' represents a constant multiplied by a linear trend with time.

Appendix G) [116]. The validated residuals are used for non-linear prediction through LSTM.

3.3.2 Non-linear prediction

This section introduces the non-linear modelling process, beginning with a brief explanation of how it is conceptualised and implemented. It then presents the LSTM model, which is used to capture non-linear components, followed by a detailed description of its implementation in Python.

3.3.2.1 Non-linear modelling

In the non-linear modelling process, it is assumed that the Chinese NG consumption variables $\boldsymbol{y_t}$ at time $\boldsymbol{t}$ is the sum of the linear $\boldsymbol{L_t}$ and non-linear components $\boldsymbol{N_t}$ as:

$$\boldsymbol{y_t = L_t + N_t} \quad \textbf{(2)}$$

SARIMAX is utilised to fit the linear components $\boldsymbol{L_t}$ and then obtain the linear predictions $\boldsymbol{\hat{L}_t}$. Non-linear components, residuals, are obtained by subtracting the linear predictions $\boldsymbol{\hat{L}_t}$ from the original variables $\boldsymbol{y_t}$ as:

$$\boldsymbol{N_t = y_t - \hat{L}_t} \quad \textbf{(3)}$$

The residuals are predicted using a deep learning approach—a LSTM model.

<u>Long Short-Term Memory</u>

The LSTM model is a specific type of RNN (see Appendix H) particularly designed to capture long-term dependencies and retain information over extended periods [97, 103]. RNN exhibits performance degradation in long-term sequences. LSTM has been introduced by S. Hochreiter and J. Schmidhuber in 1997 to overcome this degradation [117]. Figure 3.2 presents the schematic diagram of LSTM, in which there are three gates (the input gate, output gate, and forget gate) and the memory cell input and output, incorporating information from the current input and output and the hidden state from the previous time steps. The three gates enable LSTM to effectively capture long-term dependencies (see Appendix I).

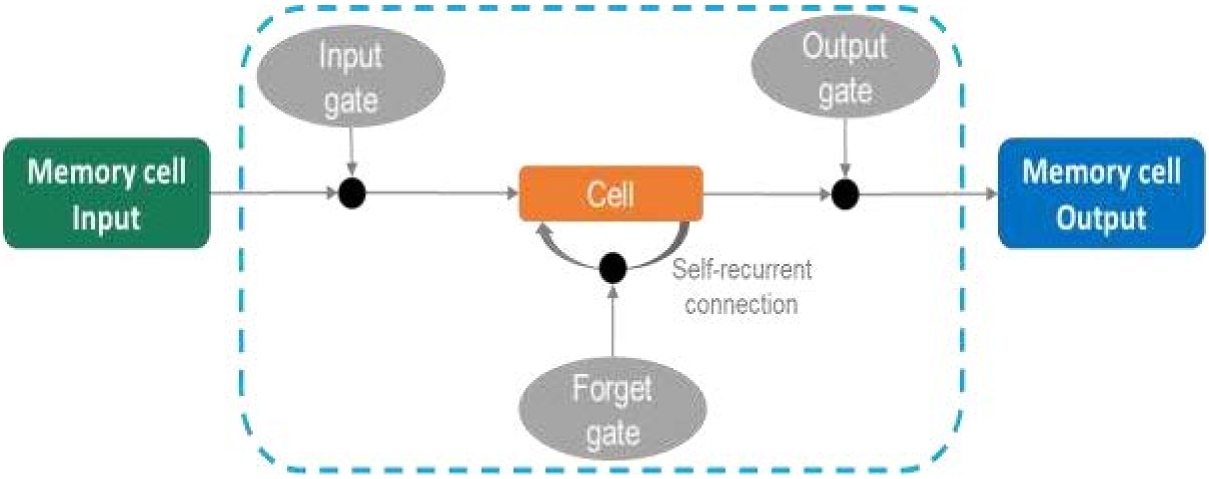


Figure 3.2 Schematic diagram of LSTM

The improvement in the ability to detect long-term dependencies enables LSTM to precisely model data with short- or long-term dependencies. As depicted in Figure 3.3, which presents historical data, the trend of China's NG consumption has steadily increased since 2010 with cyclical fluctuations. Although this historical trend appears largely linear, underlying non-linear patterns may exist that a linear model cannot capture. LSTM can effectively capture short-term fluctuations while retaining long-term upward trend information, which is highly suitable for predicting the non-linear trend of China's NG consumption.

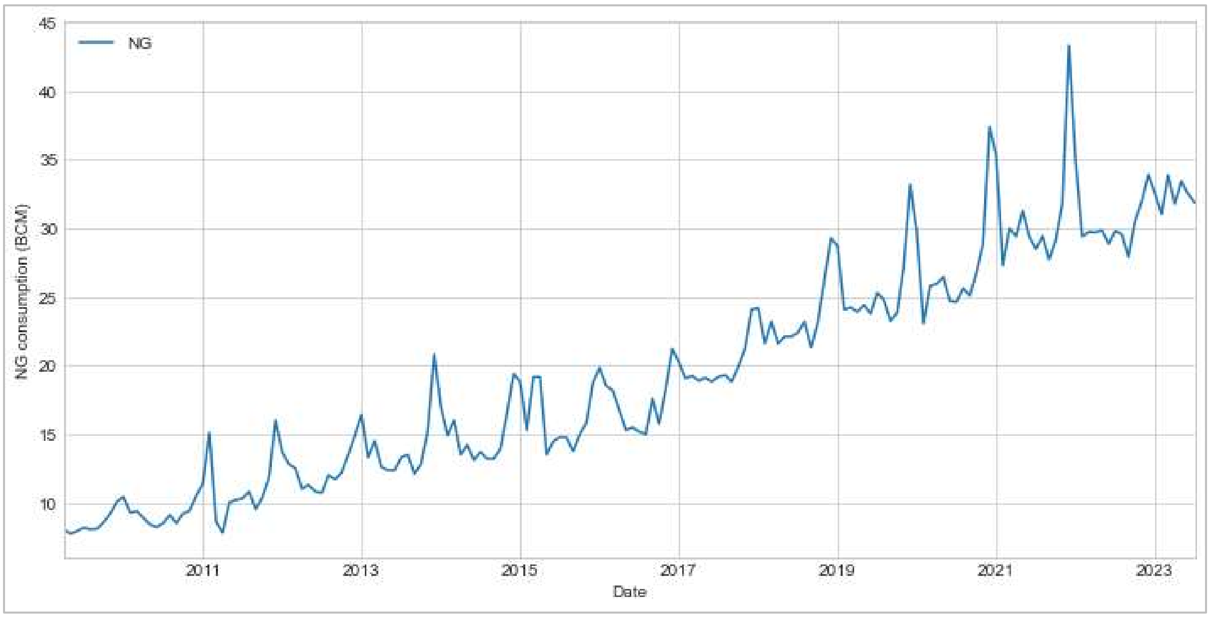


Figure 3.3 China's NG consumption

3.3.2.2 Modelling process

Initially, two types of single LSTM models – the LSTM model comprising three or four layers – are established with the key drivers as inputs to predict one-year-ahead Chinese NG consumption. Subsequently, for the evaluation of model performance, the LSTM models are compared among themselves, as well as with SARIMAX 1 and 2. Finally, the best-performing LSTM model is utilised in the hybrid SARIMAX-LSTM models to compute non-linear (residual) predictions.

Here, the two single LSTM modelling processes are elaborated upon, while LSTM modelling for residuals in the hybrid models will be provided in the section of hybrid models (see Section 3.3.3.1). The LSTM modelling processes involve three sequential phases: data pre-processing, model design, and compilation and training. LSTM modelling is conducted using TensorFlow, a Python library for deep learning. The three phases are as follows:

<u>Phase 1: Data pre-processing</u>

In the first step of data pre-processing, all variables are standardised within the range of -1 and 1, similar to the data pre-processing applied in SARIMAX modelling. In contrast to SARIMAX modelling, dependent variables are normalised in LSTM modelling. The variables are standardised to achieve better fitting and prevent divergence during model training [118]. At the final modelling stage, they are de-standardised using the mean and variance calculated earlier.

In the second step, the 'window size', representing the number of data points used as inputs at once for a single prediction in the LSTM, is set to the past six months. The window size of six months is assumed to provide valuable information for one-year-ahead forecasting by considering seasonal factors and the impact of economic activity. In the LSTM modelling, all variables are aggregated in six-month intervals as inputs during the model training due to the window size.

In the final step, scaled variables are split into two sets – 70% as a training set and 30% as a test set – with shuffling. This differs from the data pre-processing approach of SARIMAX in that the shuffling is used here. A LSTM model retains the information from previous steps and uses it along with the current step's input to generate the output. In the process, the data sequence impacts the model's learning of patterns. If the data is not shuffled and fed into the model in its original sequence, it tends to emphasise patterns related to that specific order [119]. Therefore, the generalisation of the LSTM network is improved by employing a training method that involves shuffling the data rather than training it sequentially [120]. In a recent study, the best results on test data for LSTM are achieved by shuffling the input data [119].

Phase 2: Model architecture

This study employs the LSTM model structure, drawing inspiration from the machine and deep learning literature [42, 121] as well as recent relevant studies [122, 123]. In detail, this model structure consists of LSTM as the initial LSTM layer, followed by multiple dense layers with 'tanh' activation functions, and concludes with the output layer featuring a single node applying a linear transformation to its input.

This study develops the LSTM model comprising three layers (referred to as LSTM 1 hereafter) and the LSTM model consisting of four layers (denoted as LSTM 2 hereafter) for comparison with SARIMAX models. The three layers of LSTM 1 are LSTM, 1st dense, and output layers. The four layers of LSTM 2 are LSTM, 1st dense, 2nd dense, and output layers. Given the focus on the hybrid model in this research, hyperparameters in LSTM modelling are not aimed at finding new ones but rather drawn from previous studies and literature. Table 3.4 provides the specific model architectures of LSTM 1 and 2.

Table 3.4 Mode architectures of LSTM 1 and 2

Layer	LSTM 1	LSTM 2
LSTM layer	128 nodes with a 'tanh' activation function[21] and a 'dropout rate' of 20%[22]	
1st dense layer	32 nodes with a 'tanh' activation function	64 nodes with a 'tanh' activation function
2nd dense layer	N/A	32 nodes with a 'tanh' activation function
Output layer	1 node with a 'linear' activation function	

Phase 3: Model compilation and training

The compilation step essentially prepares the model for training by specifying how it should learn and adjust its parameters. This involves setting up the optimiser, the loss function, and an additional metric to configure the learning process in the model training as follows:

- Loss function: Mean Squared Error (MSE)
- Optimiser: Adam
- Metric: Mean Absolute Error (MAE)

As this study is relevant to a regression problem, MSE and MAE are chosen as the loss function

[21] The hyperbolic tangent function (*tanh*) is a bounded activation function with outputs constrained to the interval [-1, 1] [53].
[22] Dropout is a regularisation technique for a neural network by randomly dropping out the nodes in the input and hidden layers [53].

and additional metric, respectively [124]. Furthermore, Adam, one of the most highly efficient optimisation algorithms among practitioners, is selected as the optimiser [53, 125].

The training step is where the model actually learns from the data based on the specifications defined during compilation. In detail, model training involves feeding the training set through the model, computing the loss (via the loss function), and adjusting the model's weights (using the optimiser) to minimises the loss, thereby improving the model's performance.

The hyperparameters required for the training are the batch size, number of epochs, and early stopping. The batch size is determined at a default value of 32 [53, 126].[23] In line with recent research findings, this study employs 100 epochs [127].[24] Early stopping was implemented with a threshold of 5.[25]

3.3.3 Prediction of hybrid models

Hybrid models forecast China's NG consumption by combining linear and non-linear predictions. Depending on the combination method, the hybrid models are classified into the hybrid model 1 and 2. The hybrid model 1 uses an additive combination, while the hybrid model 2 employs an optimal combination.

3.3.3.1 Hybrid model 1

The additive combination of hybrid model 1 (referred to as Hybrid 1 hereafter) indicates the sum of linear components previously determined by SARIMAX and residuals derived from LSTM. Further details are provided as follows:

<u>LSTM modelling with residuals</u>

The residuals are predicted using LSTM 2 through three phases: data pre-processing, model architecture, and model compilation and training.

In the data pre-processing phase, residuals are first calculated as the difference between original lagged NG consumption variables and linear predictions based on equation (3). The Secondly, the window size is set to one year. In other words, to predict an one-month-ahead residual, the past one-year residuals are utilised as inputs. In comparison to the LSTM models (LSTM 1 and 2) using scaled exogenous and dependent variables, the window size in the LSTM model using the residuals (the LSTM model in Hybrid 1) is larger than that of LSTM 1 or 2. During the training phase, the residuals serve as both independent and dependent variables for comparison with predictions. Therefore, to ensure better training on historical data, a window size of one year has been chosen. Finally, scaled residuals are split into two sets – 70% as a training set and 30% as a test set – without shuffling. This approach differs from the data pre-processing methodology employed in LSTM 1 and 2. The decision not to shuffle the data during data pre-processing aligns with the recent study's hybrid SARIMA-LSTM approach [128]. Besides, the LSTM model in Hybrid 1 is calculated based on the sequential residuals. Therefore, shuffling data with sequential significance in the LSTM model compromises the

[23] Batch size is the number of data samples used in a single forward and backward pass during each training iteration. [53].
[24] An epoch indicates one complete pass through the entire training dataset [53].
[25] Early stopping is the feature that halts training before overfitting occurs [53]. The threshold of 5 means that this model is set to have the patience for epochs as 5, even if the validation score does not improve.

integrity of its readings [129].

In a nutshell, the LSTM model in Hybrid 1 uses the past one-year scaled residuals as inputs to predict the one-month-ahead residual. A rolling one-month-ahead residual is continuously computed using the residuals from exactly one year prior to each predicted residual, ultimately predicting one-year-ahead residuals. The model design (including parameters), compilation, and training follow the same methodologies as introduced earlier for LSTM 1 and 2 (see Section 3.3.2.2).

Additive combination

Hybrid 1 predicts one-year-ahead Chinese NG consumption based on the combination of linear and non-linear (residual) predictions through addition. This method is introduced in a recent study [91]. The linear predictions are calculated using a SARIMAX model – SARIMAX 1 or SARIMAX 2, while the non-linear predictions are computed through LSTM 2. In Hybrid 1, the combination with SARIMAX 1 is referred to as Hybrid 1-S1, while the combination with SARIMAX 2 is labelled as Hybrid 1-S2.

3.3.3.2 Hybrid model 2

The hybrid model 2 (denoted as Hybrid 2 hereafter) is derived by optimally combining linear and non-linear predictions through ANN. The following sections provide more detailed explanations regarding ANN and the optimal combination used in Hybrid 2:

Artificial Neural Network

An Artificial Neural Network (ANN) is a data-driven mathematical tool that demonstrates the capacity to address problems through machine-learning neurons [91, 95, 97]. ANN, as a general, flexible, and non-linear tool, can approximate any type of relationship between inputs and outputs [91]. This facilitates the optimal combination between the linear and non-linear predictions of China's NG consumption. Figure 3.4 provides a schematic diagram of ANN with six layers utilised in this study.

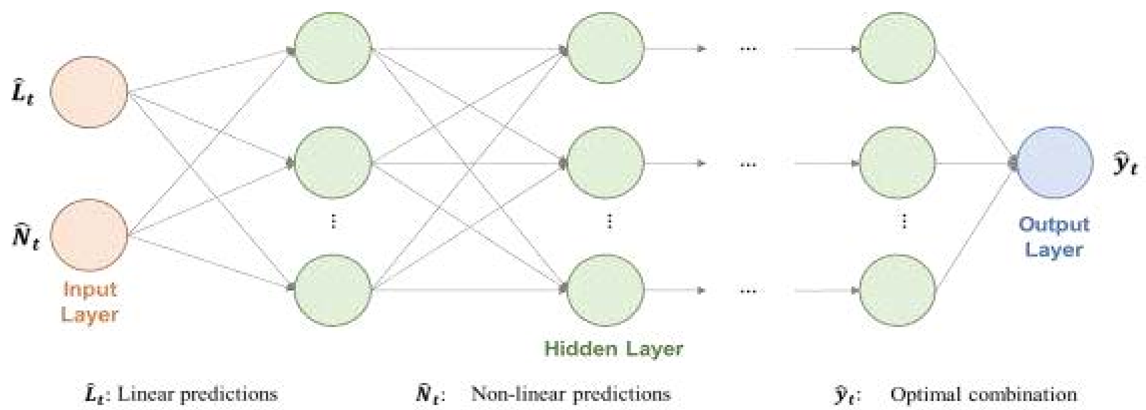


Figure 3.4 The structure of ANN

Optimal combination

The ANN-based optimal combination in Hybrid 2 is performed using both the linear and non-linear predictions as inputs. This process involves three key phases: data pre-processing, model architecture, and model compilation and training. The data pre-processing phase for the ANN mirrors that of the LSTM model in Hybrid 1, with the exception of the train-test split ratio—80% for training and 20% for testing. In the model architecture phase, the ANN is structured as a six-layer network. It comprises an input layer, followed by two dense layers with *tanh* activation functions, then two additional dense layers with *elu* activation functions. The model concludes with an output layer consisting of a single node that applies a linear transformation to its input. These activation function choices are informed by prior research, which highlights the effectiveness of tanh and elu as continuous functions capable of achieving relatively high accuracy in ANN-based continuous time series analysis [130]. The compilation and training phases of the ANN follow the same methods as introduced earlier for LSTM 1 and 2. Table 3.5 presents the detailed model architecture of the ANN.

Table 3.5 Model architecture of ANN

Layer	ANN
Input layer	2 nodes for inputs - the linear and non-linear predictions
1st dense layer	128 nodes with a *tanh* activation function[26]
2nd dense layer	64 nodes with a *tanh* activation function
3rd dense layer	32 nodes with a *elu* activation function[27]
4th dense layer	16 nodes with a *elu* activation function
Output layer	1 node with a *linear* activation function

The linear and non-linear predictions used in Hybrid 2 are identical to those employed in Hybrid 1. Depending on which SARIMAX model is used to generate the linear predictions—SARIMAX 1 or SARIMAX 2—Hybrid 2 is further categorised into two variants: one combined with SARIMAX 1, referred to as Hybrid 2-S1, and another combined with SARIMAX 2, referred to as Hybrid 2-S2. This naming convention mirrors that of Hybrid 1.

3.3.4 Evaluation of forecasting performance

The forecasting performance of each model is measured by comparing the actual value $\boldsymbol{y_t}$ of Chinese NG consumption and its predicted value $\boldsymbol{\hat{y}_t}$ at time $\boldsymbol{t}$. To evaluate the model's generalisation ability and estimate its actual performance, the test data, referred to as 'out of sample', is used. The metrics employed for this evaluation are out-of-sample R-squared (R^2) defined by Equation (4), Root Mean Square Error (RMSE) defined by Equation (5), and MAE defined by Equation (6) [131].

In addition to the three metrics, the performance is also compared using the cumulative squared forecast error (CSFE), a common out-of-sample forecasting performance measure widely used in applied economics, particularly in energy and computational contexts [132-134]. The CSFE adds up the squared errors (the differences between the actual and predicted values) for each month of the test data, as defined by equation (7).

$$\boldsymbol{R^2} = \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2} \quad \textbf{(4)}$$

$$\boldsymbol{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad \textbf{(5)}$$

[26] The hyperbolic tangent function ('tanh') is the same as that introduced for an LSTM model (See Section 3.3.2.2).
[27] As 'elu' include an exponential function in the negative range, it create a smooth connection between the linear segment in the positive range and the exponential curve in the negative range [130].

$$\boldsymbol{MAE} = \frac{1}{n}\sum_{i=1}^{n} |\boldsymbol{y_i} - \widehat{\boldsymbol{y}_i}| \tag{6}$$

$$\boldsymbol{CSFE} = \sum_{i=1}^{n} (\boldsymbol{y_i} - \widehat{\boldsymbol{y}_i})^2 \tag{7}$$

where $\boldsymbol{y_i}$ is the observed value; $\widehat{\boldsymbol{y}}_i$ the predicted value; $\overline{\boldsymbol{y}}_i$ the mean of the observed value; and $\boldsymbol{n}$ number of predictions.

3.3.5 Principal component analysis

As an extension of the hybrid forecasting model, this study investigates the suitability of incorporating identified key drivers with varying characteristics [107] as inputs for the forecasting model, using PCA.

PCA is a widely used machine learning technique for dimensionality reduction [53]. In this context, the dimensions of data are typically referred to as features. To reduce these dimensions, PCA first identifies directions in the data that exhibit the greatest variance. A vector aligned with the direction of maximum variance is termed a principal component (PC). The original dataset is then projected onto this PC, thereby reducing its dimensionality. In essence, the data is transformed and compressed along the direction that best preserves its inherent variability. Since the PC captures the highest variance, the transformed data retains the most salient characteristics of the original dataset.

Scikit-learn, a standard machine learning library, is utilised to perform PCA in this study. The PCA procedure consists of two main steps: determining the optimal number of dimensions and evaluating model performance using the transformed data. In the first step, a plot of the explained variance ratio against the number of dimensions is generated to guide the selection of an appropriate dimensionality reduction level. The explained variance ratio indicates the extent to which each PC captures the variability of the original data. By examining this plot, one can identify a point beyond which additional dimensions contribute minimally to the explained variance, thereby informing the choice of dimensions to retain. In the second step, the data transformed via PCA is used to train the hybrid forecasting model. The performance of the model using PCA-transformed data is then compared with that of the model trained on the raw data, enabling an assessment of the impact of dimensionality reduction on forecasting accuracy.

3.4 Application and analysis for forecasting near-term natural gas consumption in China

This section applies multiple forecasting approaches—including statistical models, deep learning techniques, and hybrid methods—to near-term NG consumption in China and subsequently provides a comprehensive analysis to evaluate model performance, input suitability, and the influence of contributing factors.

3.4.1 Model comparison

Model comparisons are initially conducted using three metrics—R^2, RMSE, and MAE—in the following sequence: first, SARIMAX is compared with LSTM; next, LSTM is compared with the additively combined hybrid SARIMAX-LSTM model (Hybrid 1); and finally, Hybrid 1 is compared with the ANN-based optimally combined hybrid SARIMAX-LSTM model (Hybrid 2). Additionally, comparisons are also conducted using the CSFE metric. Accordingly, to develop a high-performing model for one-year-ahead Chinese NG consumption forecasting based on key drivers, the optimal choice is the hybrid model (Hybrid 2-S2), which integrates SARIMAX with a trend-capturing function and LSTM through an ANN-based combination. The detailed comparison results are presented below.

3.4.1.1 Comparison 1: SARIMAX vs. LSTM

SARIMAX 1 and SARIMAX 2 begin generating predictions from December 2020, with their optimal orders identified as (2, 1, 1)(0, 1, 1, 12) and (0, 0, 2)(1, 1, 0, 12), respectively. Diagnostic tests, including the Ljung-Box and Jarque-Bera, confirm a good fit for both SARIMAX models, as detailed in Table 3.3 (see Appendix G). LSTM 1 and LSTM 2 commence predictions from September 2019. The forecasting performance of all models is evaluated using three metrics: R^2, RMSE, and MAE. Figure 3.5 and Table 3.6 present a comparative analysis of the models based on these metrics.

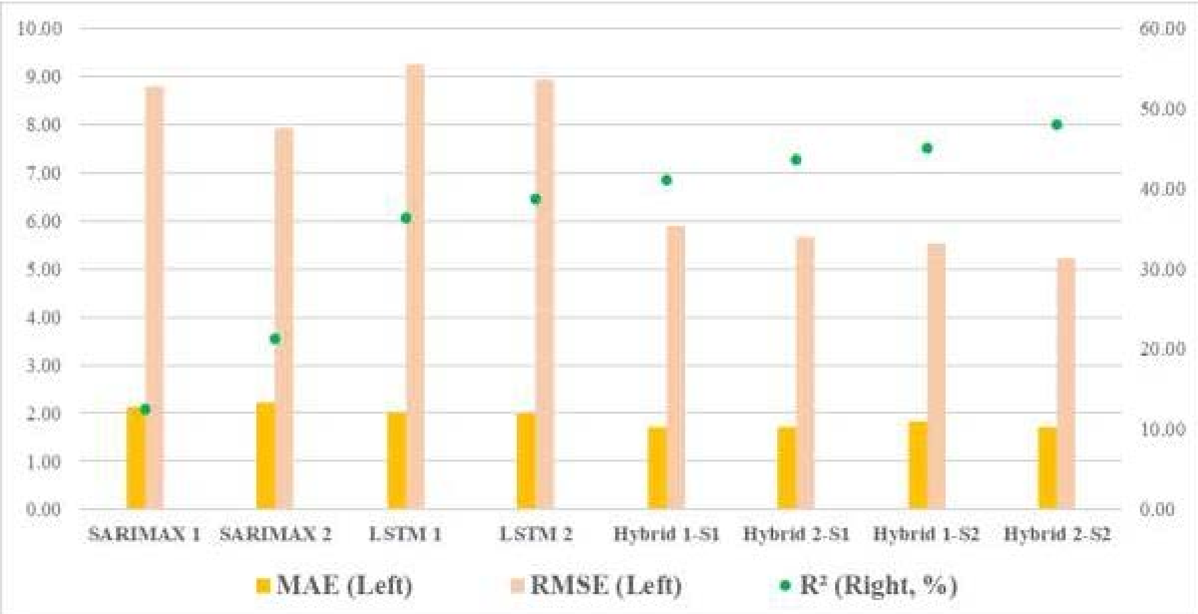


Figure 3.5 Comparison of SARIMAX, LSTM, and Hybrid models

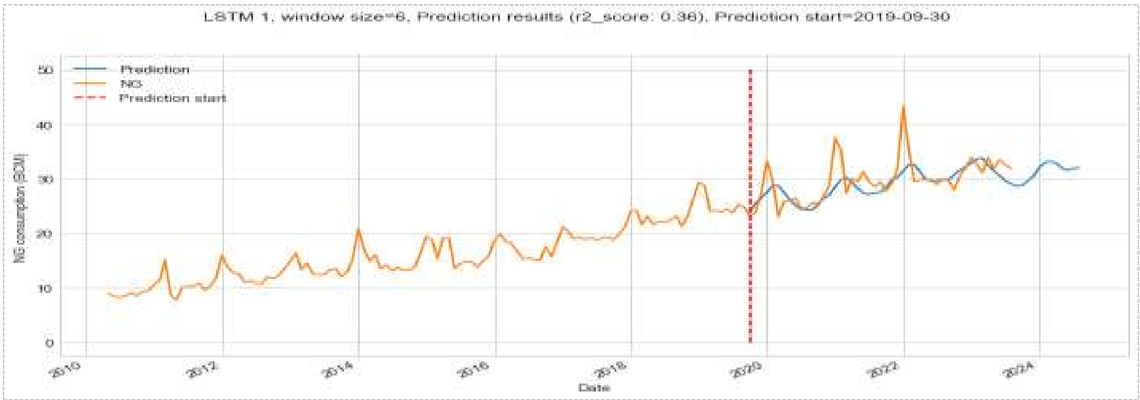
Table 3.6 Metrics of SARIMAX, LSTM, and Hybrid models

Model	R^2 (%)	RMSE	MAE
SARIMAX 1	12.55	8.79	2.12
SARIMAX 2	21.25	7.92	2.21
LSTM 1	36.46	9.27	2.01
LSTM 2	38.74	8.94	1.99
Hybrid 1-S1	41.22	5.91	1.70
Hybrid 2-S1	43.63	5.67	1.71
Hybrid 1-S2	45.08	5.52	1.82
Hybrid 2-S2	47.94	5.23	1.70

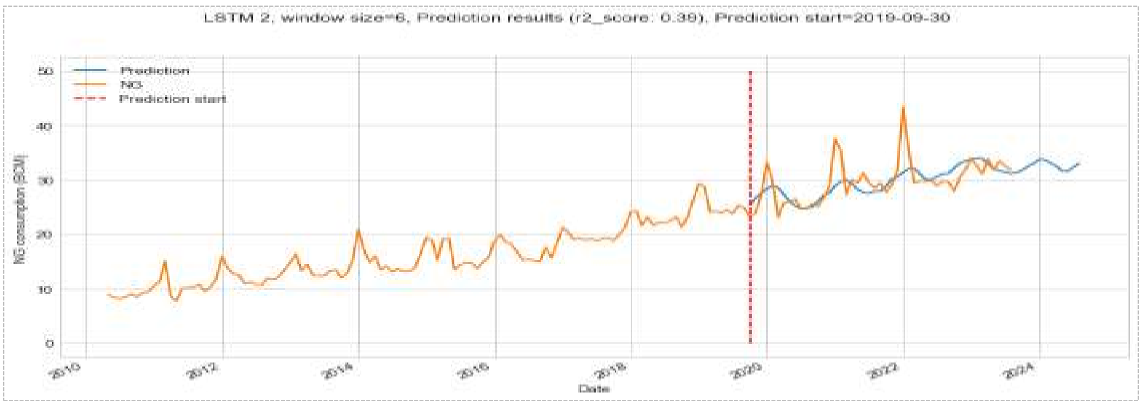
Notes: All results are rounded to two decimal places. Detailed descriptions of SARIMAX 1, SARIMAX 2, LSTM 1, and LSTM 2 are provided in Section 3.3.1 and 3.3.2. Details of Hybrid 1-S1, Hybrid 1-S2, Hybrid 2-S1, and Hybrid 2-S2 are presented in Section 3.3.3. S1 and S2 refer to SARIMAX 1 and SARIMAX 2, respectively.

SARIMAX 1 (2,1,1),(0,1,1,12) - Prediction start: Dec-31-2020

SARIMAX 2 (0,0,2),(1,1,0,12) - Prediction start: Dec-31-2020



LSTM 1 (Window size=6) - Prediction start: Sep-30-2019



LSTM 2 (Window size=6) - Prediction start: Sep-30-2019

Figure 3.6 SARIMAX vs. LSTM

Notes: The orange line represents the actual values of Chinese NG consumption, and the blue line indicates the predicted values.

As shown in Figure 3.5 and Table 3.6, SARIMAX 2 (R^2 = 21.25%, RMSE = 7.92) demonstrates superior performance compared to SARIMAX 1 (R^2 = 12.55%, RMSE = 8.79), with a higher R^2 and lower RMSE. However, SARIMAX 1 achieves a slightly lower MAE (2.12) than SARIMAX 2 (2.21). Overall, SARIMAX 2 outperforms SARIMAX 1 in two of the three metrics—R^2 and RMSE—highlighting its enhanced forecasting capability, likely attributable to the inclusion of a trend-capturing parameter.

In the comparison of the LSTM models, LSTM 2 (R^2 = 38.74%, RMSE = 8.94, MAE = 1.99) outperforms LSTM 1 (R^2 = 36.46%, RMSE = 9.27, MAE = 2.01) across all three metrics—achieving a higher R^2, and lower RMSE and MAE—thereby demonstrating superior forecasting performance. Furthermore, when compared with the SARIMAX models, both LSTM models exhibit higher R^2 values and lower MAE, underscoring their strong predictive capabilities. These results are consistent with findings in the literature, which frequently highlight the superior forecasting performance of LSTM models [96-101].

Among the SARIMAX and LSTM models, LSTM 2 emerges as the most effective, exhibiting the highest R^2 and the lowest MAE. Consequently, LSTM 2 is selected for modelling the non-linear (residual) components in Hybrid 1 and Hybrid 2. The graphical outputs for the SARIMAX and LSTM models are presented in Figure 3.6.

3.4.1.2 Comparison 2: LSTM vs. Hybrid 1

Hybrid 1-S1 and S2 outperform LSTM 1 and 2 in terms of the three metrics: higher R^2, lower RMSE, and MAE. As shown in Table 3.6, the lowest R^2 value of 41.22% between Hybrid 1 models (Hybrid 1-S1 and S2) is higher than the highest R^2 value of 38.74% between LSTM models (LSTM 1 and 2). Likewise, the highest RMSE value of 5.91 and MAE of 1.82 between the Hybrid 1 models are lower than the lowest RMSE value of 8.94 and MAE of 1.99 between the LSTM models, respectively. These results suggest that the hybrid models outperform not only the LSTM models but also the single SARIMAX models. Moreover, this trend is consistent with prior literature, where hybrid forecasting models demonstrate higher accuracy compared to traditional time series or deep learning models [91, 102, 103].

3.4.1.3 Comparison 3: Hybrid 1 vs. Hybrid 2

As indicated in Table 3.6, among the Hybrid 1 models—Hybrid 1-S1 and S2—Hybrid 1-S2 outperforms Hybrid 1-S1, achieving a higher R^2 (45.08%) and a lower RMSE (5.52) compared to Hybrid 1-S1 (R^2 = 41.22%, RMSE = 5.91). Similarly, within the Hybrid 2 models, Hybrid 2-S2 demonstrates superior performance, with an R^2 of 47.94%, RMSE of 5.23, and MAE of 1.70—outperforming Hybrid 2-S1 (R^2 = 43.63%, RMSE = 5.67, MAE = 1.71). These results highlight the enhanced performance of hybrid models that incorporate SARIMAX 2 (S2), which includes the trend-capturing parameter. This underscores the important of capturing linear trends effectively to enhance the forecasting accuracy of the hybrid SARIMAX-LSTM framework.

As shown in Table 3.7, the hybrid models demonstrate a synergistic effect between SARIMAX and LSTM. The R^2 values for SARIMAX 1 and SARIMAX 2 are 12.55% and 21.25%, respectively. The LSTM models, trained on the residuals from SARIMAX 1 and 2, achieve R^2 values of 18.41% and 25.73%, respectively. In contrast, the hybrid models achieve

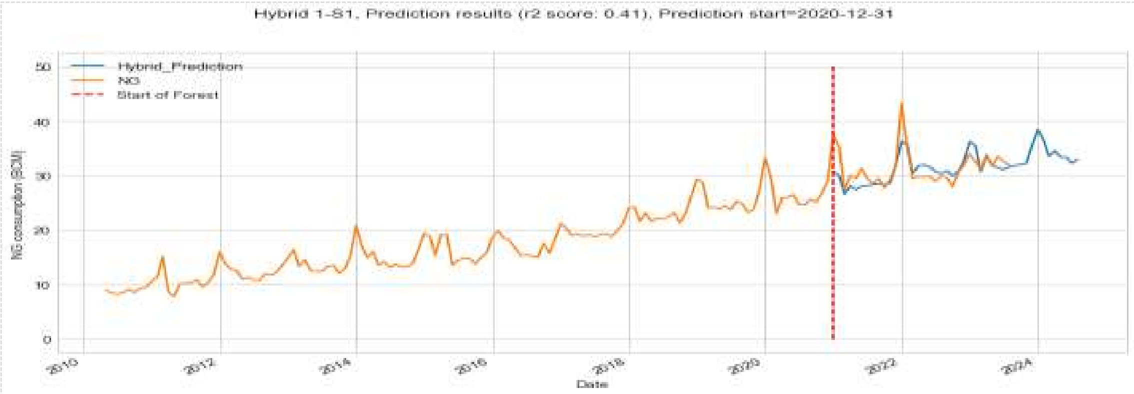
R^2 values of at least 41.22% (Hybrid 1-S1). This indicates that creating a hybrid model by combining the two—SARIMAX and LSTM—has a synergistic effect.

Table 3.7 The synergitic effect between SARIMAX and LSTM

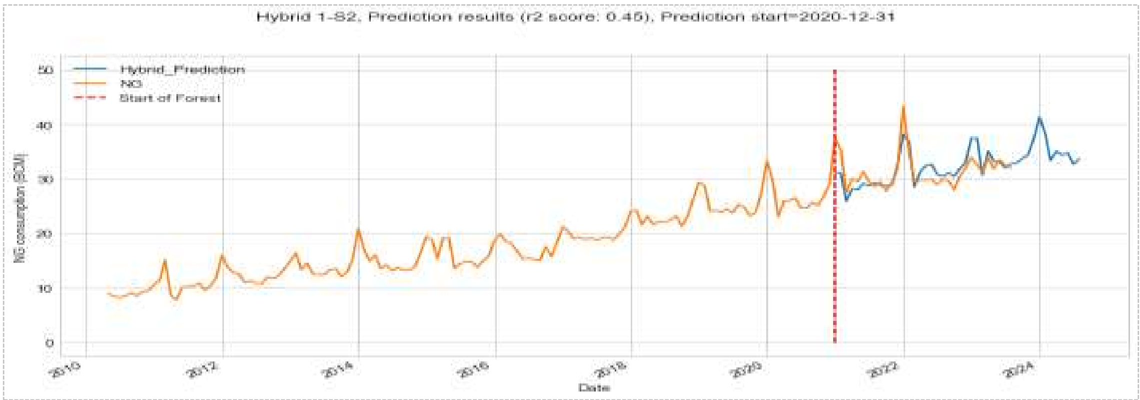
	Model	R^2 (%)
Hybrid 1 and 2 models based on S1	SARIMAX 1 (S1)	12.55
	LSTM for residuals (derived from S1)	18.41
	Hybrid 1-S1	41.22
	Hybrid 2-S1	43.63
Hybrid 1 and 2 models based on S2	SARIMAX 2 (S2)	21.25
	LSTM for residuals (derived from S2)	25.73
	Hybrid 1-S2	45.08
	Hybrid 2-S2	47.94

As depicted in Table 3.6, when comparing Hybrid 1 and Hybrid 2 models that share the same SARIMAX component (S1 or S2), Hybrid 2 consistently outperforms Hybrid 1. Specifically, Hybrid 2-S1 (R^2 = 43.63%, RMSE = 5.67) achieves a higher R^2 and lower RMSE than Hybrid 1-S1 (R^2 = 41.22%, RMSE = 5.91). Likewise, Hybrid 2-S2 (R^2 = 47.94%, RMSE = 5.23, MAE = 1.70) surpasses Hybrid 1-S2 (R^2 = 45.08%, RMSE = 5.52, MAE = 1.82) across all three metrics. These findings suggest that the Hybrid 2 models offer superior performance. This implies that optimally integrating SARIMAX and LSTM through ANN yields significantly better predictions than simply adding their outputs together.

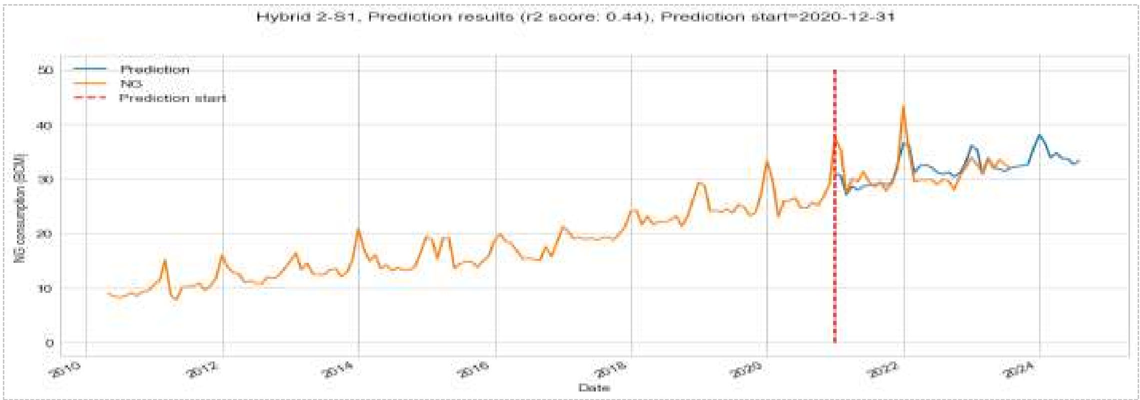
In conclusion, to develop a high-performing model for forecasting one-year-ahead Chinese NG consumption using key drivers, Hybrid 2-S2 is the most suitable choice. It effectively combines SARIMAX with trend-capturing functionality and LSTM through ANN. The performance graphs for all hybrid models are presented in Figure 3.7.



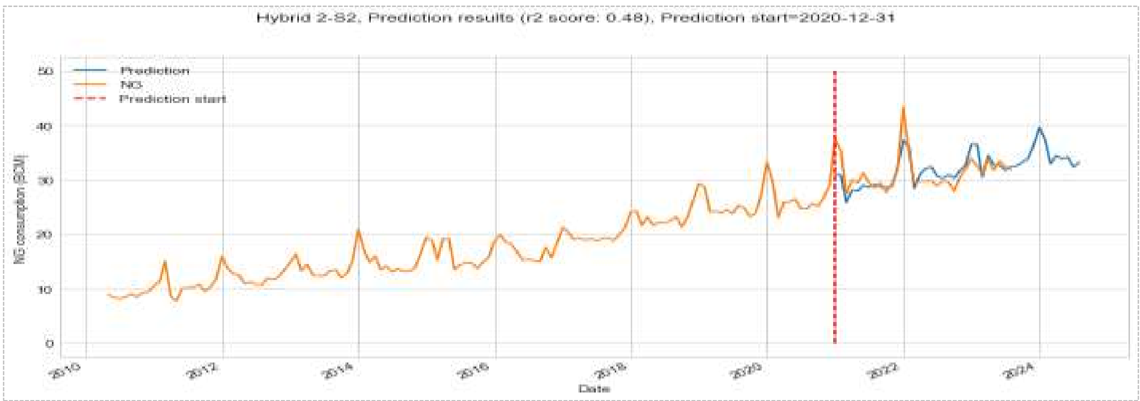
Hybrid 1-S1 (Prediction start: Dec-31-2020)



Hybrid 1-S2 (Prediction start: Dec-31-2020)



Hybrid 2-S1 (Prediction start: Dec-31-2020)



Hybrid 2-S2 (Prediction start: Dec-31-2020)

Figure 3.7 Hybrid 1 vs. Hybrid 2

Notes: The orange line represents the actual values of Chinese NG consumption, and the blue line indicates the predicted values.

3.4.1.4 Cumulative squared forecast error

In addition to the three primary evaluation metrics, model performance is also assessed using the CSFE. The CSFE for each model is calculated over the period from December 2020 to July 2023, which corresponds to the test data range for most models. However, due to differences in window size and train-test split ratios, the test data for LSTM 1 and LSTM 2 begin earlier, in September 2019. To ensure consistency in comparison, the CSFE evaluation uses December 2020 as the common starting point across all models. Figure 8 presents the CSFE results for each model.

The model that exhibits the smallest CSFE is considered the most accurate, and evaluating whether CSFE decreases over time serves as an indicator of predictive performance. A declining CSFE suggests that the model's forecasts become increasingly accurate as time progresses. As shown in Figure 8, the models with relatively lower CSFE values are the Hybrid 1 and Hybrid 2 models. Among these, the variants combined with SARIMAX 2 (S2) consistently demonstrate lower CSFE values compared to those combined with SARIMAX 1 (S1). This finding highlights the contribution of S2's trend-capturing functionality in enhancing predictive accuracy.

As with the comparison of the three primary metrics, Hybrid 2-S2 (represented by the red line) also demonstrates the strongest predictive performance in terms of CSFE. When comparing Hybrid 1-S2 and Hybrid 2-S2—both of which exhibit relatively low CSFE values—Hybrid 1-S2 initially records lower CSFEs. However, over time, Hybrid 2-S2 maintains a lower CSFE, indicating that its predictive accuracy improves as time progresses.

Furthermore, Hybrid 2-S2 shows strong capability in forecasting sudden changes. Given that CSFE is based on squared errors, even small deviations are heavily penalised. As shown in Figure 3.7, there is a sharp increase in Chinese NG consumption at the end of 2022. Correspondingly, Figure 3.8 illustrates that Hybrid 2-S2 experiences the smallest increase in CSFE during this period, suggesting that it effectively captures abrupt shifts in consumption.

However, the LSTM models—LSTM 1 and 2—do not accurately predict such sudden increases. In terms of the three primary metrics, the LSTM models perform better than the SARIMAX models (S1 and S2). Similarly, in the CSFE comparison, the LSTM models initially exhibit lower CSFE values than the SARIMAX models. However, over time, the CSFEs of the LSTM models increase, eventually surpassing those of the SARIMAX models. As shown in Figure 8, this divergence becomes more pronounced from the end of 2021, coinciding with a noticeable rise in NG consumption. This suggests that the LSTM models are relatively less effective than the SARIMAX models in forecasting sudden increases in NG consumption.

In the CSFE comparison with the SARIMAX models, they ultimately record lower CSFE values than the LSTM models, but higher than those of the hybrid models. Among the SARIMAX models, although the final CSFE of S1 is higher than that of S2, the CSFE values for S2 increase more steeply over time. Notably, the CSFE for S2 rises most sharply during the sudden increase in NG consumption at the end of 2022. This suggests that when NG consumption is forecast solely using S2—the SARIMAX model with the trend-capturing function—it fails to accurately predict abrupt changes in NG consumption. However, as demonstrated by Hybrid 2-S2, SARIMAX models with the trend-capturing function, when optimally integrated with LSTM via ANN, significantly improve predictive accuracy.

In summary, as with the comparison of the three primary metrics, the CSFE analysis also identifies Hybrid 2-S2 as the optimal model for one-year-ahead forecasting of NG consumption using key drivers. This model effectively integrates S2 and LSTM via ANN, resulting in superior predictive performance. Furthermore, the Ljung-Box and Jarque-Bera test results for Hybrid 2-S2 validate the robustness of its output—an ANN-optimised combination of linear SARIMAX and nonlinear LSTM predictions—as presented in Table 3.3 (see Appendix G).

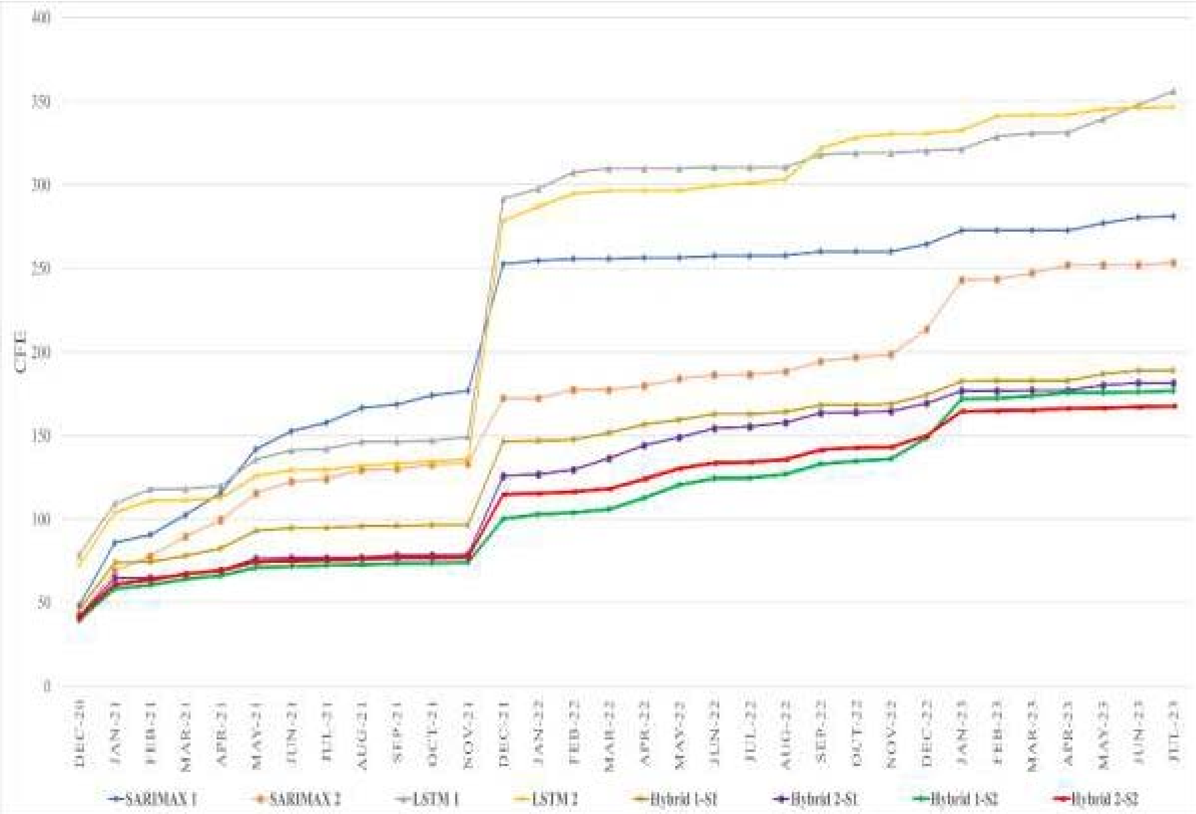


Figure 3.8 Cumulative Squared Forecast Errors of Models

3.4.2 Validity analysis of inputs

In addition to establishing an effective hybrid forecasting model, this study undertakes additional research to investigate whether the identified key drivers with different characteristics are valid inputs for the forecasting model [107]. For this investigation, the PCA-transformed data is applied to hybrid models, and then their forecasting performances are compared.

First, the principal components (PCs) that capture the key features of the dataset are selected through a graphical analysis of the explained variance ratio against the number of components [42, 53]. As indicated in Figure 3.9, the first four PCs account for the majority of the explained variance, while subsequent components contribute relatively little. Accordingly, four PCs are selected. The transformed data, derived from these PCs, is then applied across all hybrid models, following the modelling approach previously introduced for Hybrid 1 and Hybrid 2. The hybrid models incorporating PCA—Hybrid 1-S1, Hybrid 1-S2, Hybrid 2-S1, and Hybrid 2-S2—are denoted as Hybrid 1P-S1, Hybrid 1P-S2, Hybrid 2P-S1, and Hybrid 2P-S2, respectively.

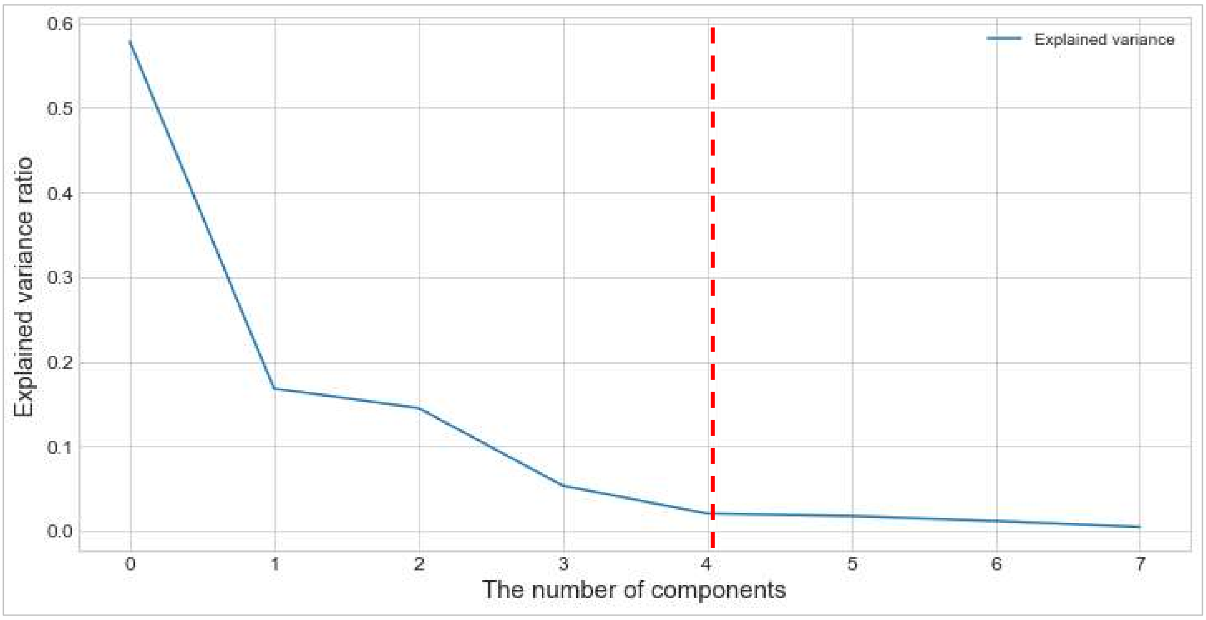


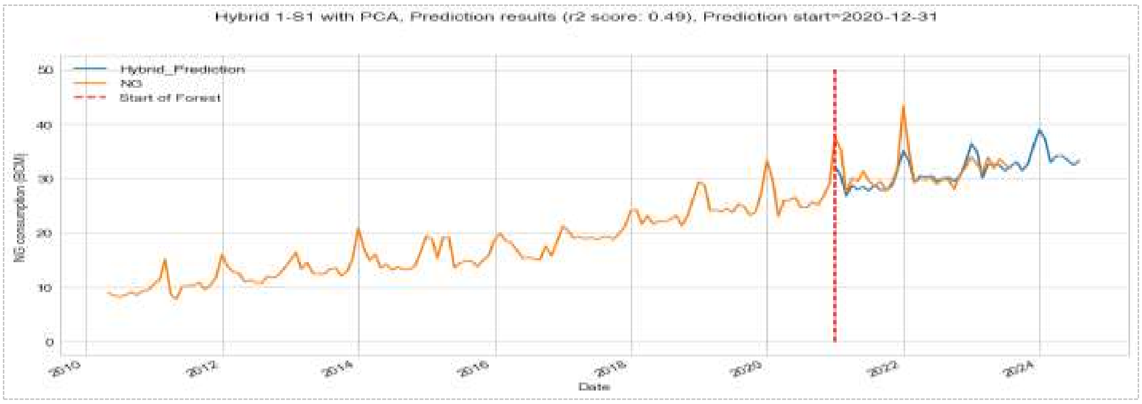
Figure 3.9 Explained variance ratio

Table 3.8 Metrics of Hybrid models with PCA

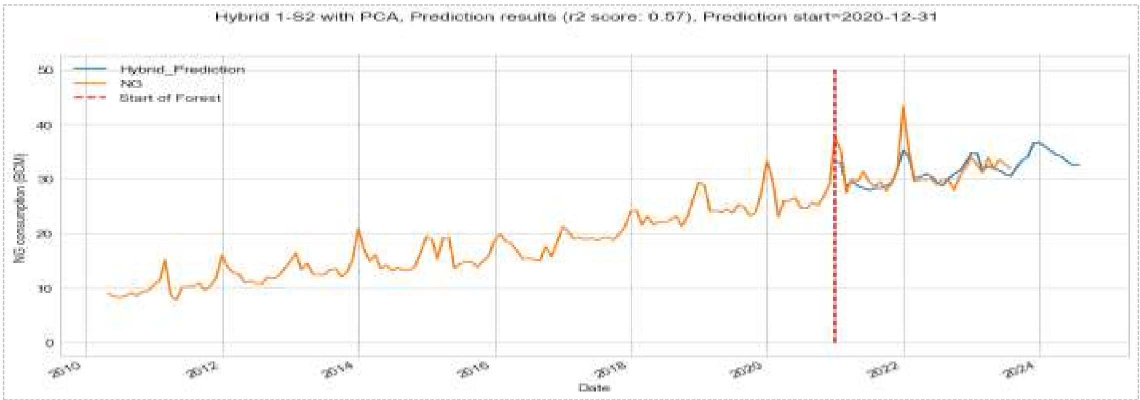
Model	R^2 (%)	RMSE	MAE
Hybrid 1P-S1	48.54	5.17	1.47
Hybrid 2P-S1	54.46	4.58	1.47
Hybrid 1P-S2	56.96	4.33	1.41
Hybrid 2P-S2	58.50	4.17	1.49

As shown in Tables 3.6 and 3.8, hybrid models incorporating PCA achieve higher R^2 values and lower MAE and RMSE compared to those without PCA. This suggests that the PCs, which capture the key features of the input variables while filtering out noise, contribute to improved model performance. This finding also indicates that the key drivers contain meaningful features that are valuable for hybrid forecasting models. Thus, the key drivers with diverse characteristics serve as valid inputs for prediction.

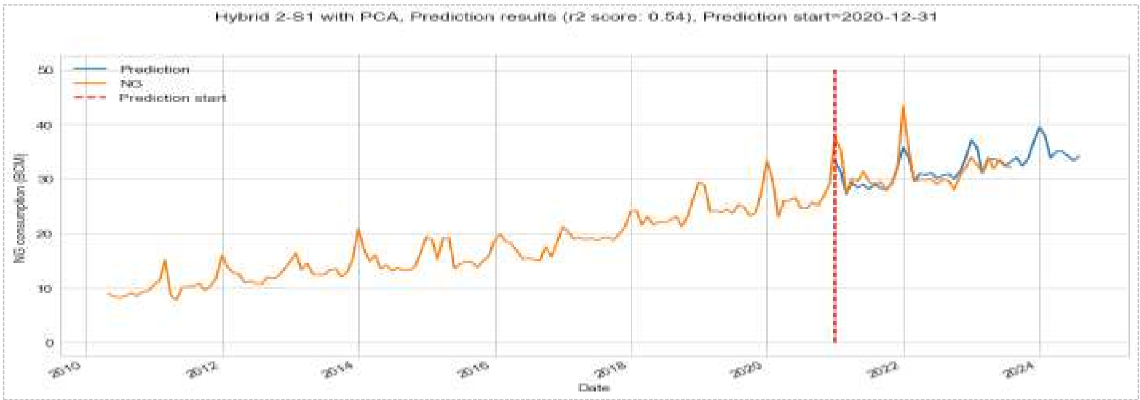
However, using PCA-transformed data alone makes it difficult to identify which specific factors most influence the forecasts. Therefore, an additional analysis is conducted to determine which key drivers contribute most significantly to predictive performance. The graphs for the hybrid models incorporating PCA are presented in Figure 3.10.



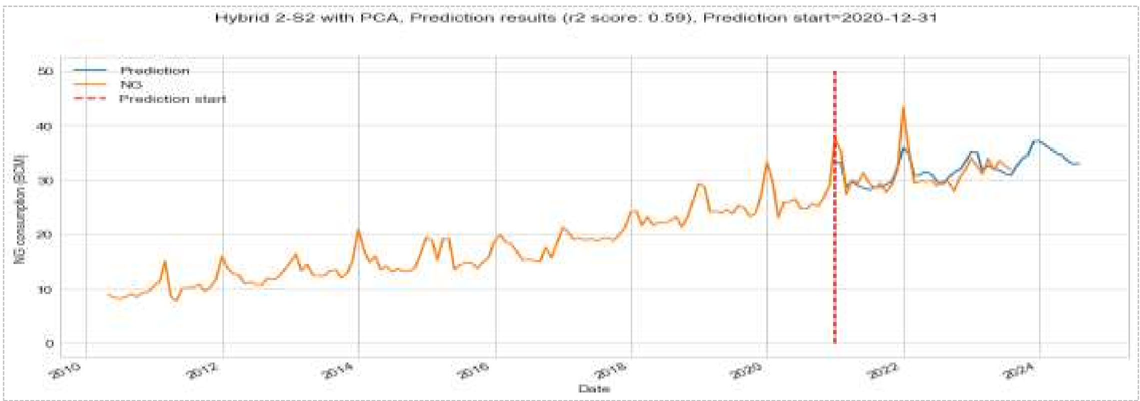
Hybrid 1P-S1 (Prediction start: Dec-31-2020)



Hybrid 1P-S2 (Prediction start: Dec-31-2020)



Hybrid 2P-S1 (Prediction start: Dec-31-2020)



Hybrid 2P-S2 (Prediction start: Dec-31-2020)

Figure 3.10 Hybrid models with PCA: Hybrid 1P-S1, Hybrid 1P-S2, Hybrid 2P-S1, and Hybrid 2P-S2

Notes: The orange line represents the actual values of Chinese NG consumption, and the blue line indicates the predicted values.

3.4.3 Contributing factors to the forecasting

To identify contributing factors, this study examines the changes in the output—predicted values of Chinese NG consumption—as an input—a key driver—increases from twofold to tenfold in the high-performing hybrid forecasting model, Hybrid 2-S2. This model incorporates data lagged by one year and is designed to account for underlying trends, thereby allowing input changes to influence the direction of predicted outcomes. However, for the purpose of identifying contributing factors, the analysis considers only the absolute magnitude of changes in output, irrespective of whether those changes are positive or negative.

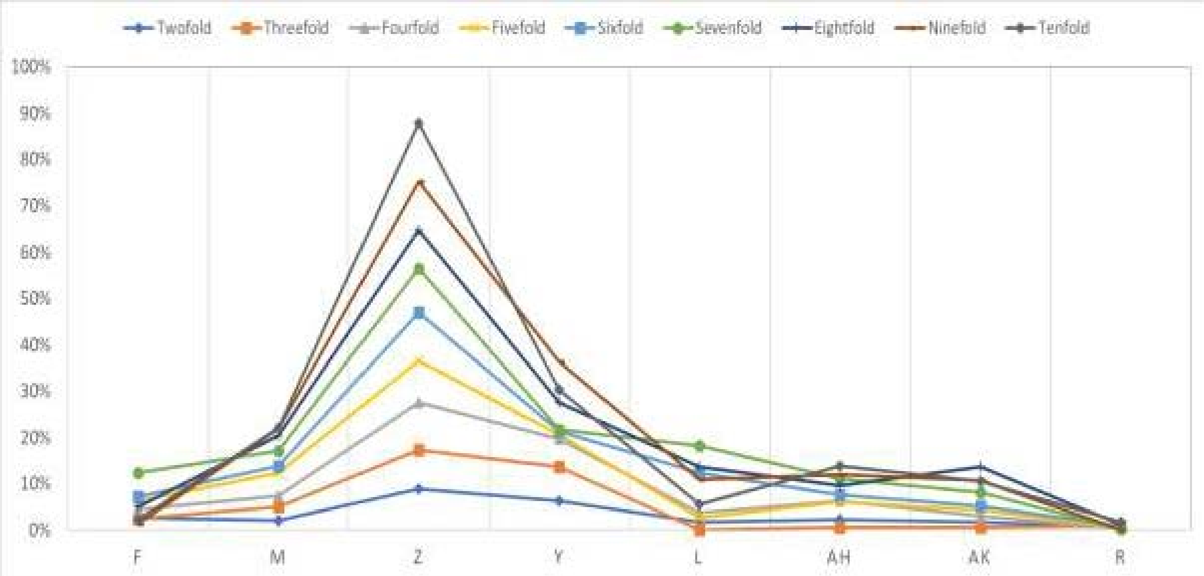


Figure 3.11 The output changes for each multiple of input

Table 3.9 The ranks of the output changes

Factors		Multiples of Input								
		2X	3X	4X	5X	6X	7X	8X	9X	10X
F	Domestic Thermal coal price	3%	2%	4%	6%	7%	12%	5%	1%	2%
M	PNG imports	2%	5%	7%	12%	14%	17%	20%	22%	22%
Z	Gas pipeline capacity	9%	17%	27%	37%	47%	56%	65%	75%	88%
Y	Accum. LNG terminal capa. per year	6%	14%	20%	21%	21%	22%	28%	36%	30%
L	LNG imports	2%	6%	4%	3%	12%	18%	14%	11%	6%
AH	Avg. national temperature	2%	1%	6%	6%	8%	11%	10%	12%	14%
AK	Price distortion	2%	1%	3%	4%	5%	8%	14%	11%	10%
R	HH gas price	1%	1%	0%	1%	1%	0%	1%	0%	2%

Notes: The 1st place is coloured in red, the 2nd in orange, the 3rd in green, and the 4th in blue.

As shown in Figure 3.11, which displays the output changes for each multiple of input, it becomes evident that, among the key drivers, NG-related infrastructure factors, Gas pipeline capacity and Accumulated LNG terminal capacity, have the most significant impact on output changes compared to other factors.

As shown in Table 3.9, output changes are ranked with the first place in red, the second in orange, the third in green, and the fourth in blue. Compared to other factors, NG-related infrastructure factors consistently maintain the first and second positions across all input multiples. Beyond these infrastructure factors, an examination of those ranked third and fourth—indicated in green and blue, respectively—reveals that PNG imports, LNG imports, and temperature are also influential in driving output variation within the forecasting model. This is aligned with the findings of Liu et al.'s research, indicating that temperature influences the near-term forecasting of NG consumption [92].

The remaining three factors—thermal coal price, price distortion, and HH gas price—exert a comparatively smaller influence on output variation than the other variables and display similar patterns of change. As a result, it is difficult to classify them as contributing factors. This analysis highlights that, among the key drivers, NG-related infrastructure factors have the most pronounced impact on the forecasting outcomes.

Practical application: Enhancing forecasting and gas networks via telemetry

The contributing factors influencing the short-term forecasting in this study—NG-related infrastructure (gas pipeline capacity and accumulated LNG terminal capacity), PNG and LNG imports, and temperature—can be monitored in real time via the telemetry control system of the gas transportation network [135]. Gas pipeline capacity and accumulated LNG terminal capacity can be estimated in real time through telemetry, based on pressure and flow rate for pipeline [136], and storage volume and real-time send-out capacity for LNG terminals [137], respectively. Additionally, PNG imports can be measured using pipeline flow sensors, LNG imports within transport tanks, and temperature via monitoring systems installed in pipelines or storage facilities [135, 138].

These real-time factors not only enhance the accuracy of forecasting but can also be utilised for safety management and operational efficiency improvements of gas distribution networks. Real-time data is more valuable for short-term forecasting than recent historical data, as it enables more effective detection of anomalies, thereby improving forecasting performance [139, 140]. Accordingly, the integration of real-time factors can further strengthen the high-performing hybrid forecasting model—Hybrid 2-S2—used in this study. This enhanced model, when combined with the telemetry control system, can support the design of an efficient system capable of implementing proactive safety measures. For instance, if a sharp increase in predicted NG consumption is detected, the system can automatically adjust pipeline pressure, flow rate, terminal storage volume, and send-out rate via the telemetry control system, thereby improving system stability and overall efficiency.

The telemetry control system and Hybrid 2-S2 can be integrated via an Application Programming Interface (API) and a data pipeline, ultimately connecting with gas distribution networks. An API enables data exchange between software applications and systems [141], while a data pipeline collects, transforms, and transfers data to storage, thereby enabling connectivity between the model and the telemetry control system [142]. Developed in Python, an open-source programming language, Hybrid 2-S2 can be integrated with a web-based API, Representational State Transfer (REST) API [141], and an Extract, Transform, Load (ETL) data pipeline [142]. This integration allows the model to collect real-time telemetry data and enables the operational management system of the gas networks to utilise the model's forecasts. As a result, the system can not only automate safety and operational processes but also optimise the overall efficiency of the gas distribution network. Further details regarding this practical application are reserved for future research.

3.5 Further discussion: Carbon-neutral and broader implications of the hybrid method

The hybrid forecasting model developed in this study presents key findings that highlight its practical and potential implications for future applications and advancement. The model's enhanced accuracy can aid governmental agencies in making evidence-based short-term NG policy decisions and support energy companies in undertaking informed near-term investment and operational planning. In particular, by enabling more precise forecasting of NG demand and utilisation, the model could contribute to carbon neutrality initiatives through improved resource efficiency and the reduction of avoidable emissions. Its strong performance also indicates broader applicability across other energy sectors as well as in non-energy domains. These implications are delineated across four key dimensions.

First, the hybrid forecasting approach proposed in this study should be actively adopted to advance carbon-neutral and efficient operations related to NG demand. Its enhanced predictive capability enables policymakers to improve the precision of energy demand projections aligned with carbon neutrality goals, facilitating more effective energy planning and resource allocation. This can reduce reliance on carbon-intensive sources while supporting the integration of renewable energy. Specifically, for NG producers and distributors, the hybrid forecasting model provides a reliable tool for short-term demand prediction. Accurate forecasting helps reduce overproduction and excessive imports, thereby avoiding unnecessary carbon emissions [10]. It also enables better coordination with variable renewable energy sources, supporting a more balanced and flexible energy system. When integrated with telemetry control systems, the model enhances supply chain efficiency and operational safety, contributing to stable and sustainable emissions-reduction efforts.

Second, this hybrid forecasting model can effectively support Environmental, Social, and Governance (ESG)-oriented investment strategies, which financial institutions increasingly prioritise. Near-term NG demand forecasting on a monthly and seasonal basis—using selected NG drivers such as carbon-neutral transition dynamics and policy shifts in the NG market—supports strategic ESG-aligned investment decision-making in energy planning and infrastructure management. As a transition fuel, NG plays a vital role in complementing the intermittency of renewables, helping to ensure a reliable supply during periods of reduced solar or wind generation (e.g., lower solar output in winter or weaker wind conditions in summer). Accurate forecasting enables the timely procurement of NG, enhancing both cost efficiency and energy system reliability. This allows institutions to optimise risk–return trade-offs while meeting ESG compliance and contributing to carbon neutrality objectives. Notably, in China, ESG regulations are undergoing formalisation and refinement. The government has introduced a national ESG disclosure framework, with mandatory reporting for large listed companies set to begin in 2026 and full implementation expected by 2030 [80]. These standards are becoming increasingly investor-focused, emphasising financial materiality and alignment with global ESG frameworks. In this context, this hybrid forecasting approach helps Chinese financial institutions meet evolving ESG requirements while supporting more informed and sustainable investment decisions.

Third, NG-related infrastructure can be regarded as a strategic lever for short-term operational efficiency, contributing to achieving long-term decarbonisation objectives. This hybrid forecasting model identifies NG-related infrastructure as a critical determinant in near-term NG consumption forecasting, highlighting the need to prioritise its development in response to anticipated demand growth. In particular, expanding pipelines, storage facilities, and import terminals can support increased NG use, which emits less carbon than coal, thereby reinforcing its role as a transitional fuel in the shift towards cleaner energy. Furthermore, targeted private investment and technological innovation in NG-related infrastructure can enhance supply resilience and support stable NG demand growth in high-consumption markets such as China, contributing to the long-term development of a carbon-neutral energy system.

Lastly, the high predictive accuracy of this hybrid forecasting model demonstrates its scalability to various fields where short-term forecasting is critical, including other energy sources. Other energy sectors, such as oil consumption and solar and wind energy production, share similarities with NG consumption regarding extreme volatility. Beyond these energy sectors, other prediction areas using complex and volatile time series data include tourism demand and import/export forecasting. The model, capable of handling such complex and volatile temporal data, could be a universal forecasting tool for efficiently supporting near-term supply and demand management in various other fields, including the energy sector.

3.6 Conclusions

This study develops a hybrid forecasting model that integrates SARIMAX with LSTM through ANN, demonstrating superior performance compared to individual SARIMAX, LSTM, and a simple additive SARIMAX-LSTM hybrid model. The effectiveness of the proposed model is validated by comparing the forecasting performance of both single and hybrid models for one-year-ahead NG consumption in China, using key identified drivers of NG consumption as inputs. Furthermore, the study not only conducts a validity analysis of these inputs—drivers with varying characteristics—using PCA, but also investigates which factors most significantly influence forecasting accuracy. The key findings of the study are summarised in the following three points.

The first significant finding of the study is that the optimally combined hybrid forecasting models using ANN (Hybrid 2) outperform all other approaches, including the individual statistical (SARIMAX) and deep learning (LSTM) models, as well as the additively combined hybrid models (Hybrid 1). The forecasting performance of these models is evaluated using four key metrics: R^2, RMSE, MAE, and CSFEs. In the initial comparison, both SARIMAX and LSTM models are assessed individually against the two hybrid models. The hybrid models consistently demonstrate superior predictive performance, achieving higher R^2 values and lower MAE and RMSE scores. They also record the lowest CSFEs, further confirming their enhanced accuracy. These results suggest that hybrid models offer greater forecasting precision than either statistical or deep learning models alone. In the subsequent comparison between the two hybrid approaches, Hybrid 2 models—based on ANN-driven optimal combination—outperform Hybrid 1 models across all evaluation metrics. This highlights the effectiveness of using ANN to optimally combine statistical and deep learning methods, significantly improving the predictive capability of hybrid forecasting models.

The second substantial finding is that incorporating a trend-capturing function into the SARIMAX model (SARIMAX 2) significantly enhances forecasting accuracy. This underscores the importance of capturing structural temporal patterns in complex short-term time series predictions, such as one-year-ahead NG consumption in China. Compared to SARIMAX 1, which does not include the trend-capturing function, SARIMAX 2 (S2) achieves a higher R^2 of 21.25%, a lower RMSE of 7.92, and reduced CSFEs, demonstrating superior performance. Furthermore, hybrid models that incorporate S2—Hybrid 1-S2 and Hybrid 2-S2—deliver even greater accuracy. Notably, the ANN-optimised hybrid model with S2 (Hybrid 2-S2) outperforms all other models, achieving the highest R^2 of 47.94%, the lowest RMSE of 5.23, and the lowest MAE of 1.70. This superior performance is further substantiated by its lowest final CSFE value, confirming the robustness of the ANN-based optimisation when combined with trend-aware statistical modelling.

The third major finding is that the study not only validates the selected key drivers as appropriate inputs for forecasting but also identifies NG-related infrastructure factors as the most significant contributors to near-term predictions. To assess the validity of these drivers, the study employs PCA to transform the input data—capturing essential features while reducing noise—and incorporates these into the hybrid models. The models using PCA-transformed data achieve higher R^2 values and lower MAE and RMSE compared to those using raw input data, indicating that the original drivers contain meaningful features relevant to short-term forecasting. Furthermore, to evaluate the relative influence of each driver, the study analyses changes in predicted NG consumption in response to increased values of individual inputs. This analysis reveals that NG-related infrastructure factors exert the greatest influence on near-term forecasting outcomes, highlighting their critical role in shaping consumption patterns.

The findings highlight the model's potential to support data-driven planning for carbon-neutral energy policies and corporate strategies aligned with decarbonisation goals. In particular, by enabling more accurate and timely forecasting of energy consumption, the model contributes meaningfully to carbon-neutral strategies through optimised resource allocation, reduced greenhouse gas emissions, and the facilitation of sustainable energy transitions. Moreover, the model shows strong potential for extension into other sectors where near-term forecasting is critical, such as oil, renewable energy, tourism, and trade. Leveraging cutting-edge deep learning algorithms and advanced trend-capturing techniques, the model's performance can be further enhanced, broadening its applicability in carbon-neutral energy operations and investment planning.

As a final note, a limitation of this study is that the hybrid model is unsuitable for medium- and long-term forecasting due to its sole reliance on historical data. The model with the enabled trend-capturing function can produce unreliable medium- and long-term predictions. For example, if China's NG consumption shows a consistently increasing trend in historical data, the model may incorrectly predict that the NG consumption will continue over the long term, despite expectations of its decrease due to China's carbon neutrality goal—the long-term mitigation of fossil fuels [4]. This discrepancy arises as the trend-capturing feature aligns predictions with the historical increasing trend, potentially distorting medium- and long-term outputs. To address this, additional research is needed to develop medium- and long-term forecasting models that can be applied to scenarios alongside the key drivers used in this study.

Incorporating both various scenarios that account for medium- and long-term uncertainties and these drivers having long-term effects [107] can enhance the realism of these models [143, 144]. This aspect is left for future research.

Another limitation of this study is the potential loss of temporal relevance of the forecasting results at the time of application. Since this study focuses on a short-term one-year-ahead forecast, the results may become outdated as time progresses. In consideration of this, the forecasting exercise aims to demonstrate the methodological contribution of the proposed hybrid forecasting model for supporting carbon-neutral and efficient operations related to NG demand. As a practical solution to this limitation, the study proposes incorporating telemetry-based real-time data collection to continuously update the model's forecasting outputs. This approach remains a subject for future research.

CHAPTER 4.0

CHINESE NATURAL GAS PHASE-OUT PATHWAYS: A NOVEL HYBRID SCENARIO-SPECIFIC PROJECTION APPROACH TO ACHIEVE NET ZERO

This chapter is based on the publication: Sung, J., Shi, X., Teske, S., and Li, M. (2025). "Chinese Natural Gas Phase-Out Pathways: A Novel Hybrid Scenario-Specific Projection Approach to Achieve Net Zero", published in Energy, Vol. 328. The author contributions are detailed in the Declaration of Publications and Author Contributions section of this thesis.

Chapter outlook

Building on the robust carbon-neutral near-term planning framework established in Chapter 3, this chapter shifts to a longer-term perspective, exploring how China's natural gas trajectory could evolve in light of its 2060 carbon neutrality target. Achieving this goal necessitates a fundamental shift away from natural gas reliance, driven by increased adoption of renewable energy, electrification, and energy efficiency measures. However, the natural gas consumption projections through 2060 suggest a trajectory that is unlikely to align with this goal. In addition, uncertainties stemming from domestic structural changes associated with market liberalisation and natural gas price volatility due to ongoing international developments render these projections uncertain. In light of these uncertainties, this chapter develops a novel hybrid Prophet–Seasonal AutoRegressive Integrated Moving Average with eXogenous regressors model for long-term projections of China's natural gas consumption under three scenarios. The study subsequently reevaluates the likelihood of achieving genuine carbon neutrality based on these projections. The chapter demonstrates that this hybrid model is well-suited for scenario-specific projections, employing two conservative estimation methods: comparisons with other hybrid models and error bands. Key findings suggest that a complete phase-out of natural gas in China by 2060 is unlikely. In response, this chapter provides key driver pathways towards achieving zero NG consumption by utilising the One Earth Climate Model scenario. Additionally, this chapter highlights the advantages of the hybrid model for formulating long-term policies and business plans and illustrates how the proposed driver pathways can offer valuable policy implications for advancing China's carbon neutrality agenda.

Keywords: Carbon Neutrality; Natural Gas Phase-Out Pathways; Hybrid Scenario-Based Projection; Prophet; SARIMAX; China

4.1 Introduction

China's Natural Gas (NG) market finds itself at a crossroads in the context of both the global and domestic markets. On the one hand, China's carbon neutrality goal by 2060 mitigates carbon and methane emissions by reducing the use of NG and other fossil fuels [145]. On the other hand, China considers NG a crucial transitional energy source to achieve carbon neutrality by 2060 [3]. China's carbon neutrality policy aims to reduce the consumption of all fossil fuels, including NG, but does not explicitly target achieving zero NG consumption by 2060 [27]. Specifically, the primary strategies for achieving carbon neutrality involve expanding renewable energy and other low-carbon technologies, including negative-emission technologies, to reduce overall carbon emissions [30]. During this transition period, NG is expected to play a crucial role temporarily. As a result, China's short- and mid-term NG consumption growth is anticipated to be driven by its decarbonisation targets, such as carbon emission reduction and coal-to-gas substitution [89]. However, such coal-to-gas switching may not reduce greenhouse gas emissions due to differences in methane and carbon dioxide emissions stemming from the supply chains of NG and coal, as well as generation technology [146]. Moreover, methane is approximately 80 times more powerful as a greenhouse gas than carbon dioxide in the long term [9]. It is being leaked from China's NG supply chain [10]. Such methane leakage undermines the advantages of NG in reducing carbon dioxide emissions compared to coal as long as NG is used, causing significant climate impacts [9]. Given the significant environmental challenges associated with methane leakage from NG, China's ultimate carbon neutrality objective is explicitly towards achieving zero NG consumption [29].

However, uncertainties in China's NG market indicate that the projected NG consumption is unlikely to align with its carbon neutrality goal. Short- to medium-term uncertainties in the NG market are rising due to structural changes from domestic market liberalisation and international NG price volatility [25]. Moreover, these uncertainties are exacerbated by supply-demand instability from the prolonged Russo-Ukrainian war and geopolitical tensions in the Middle East. In the long term, China's electrification and high renewable energy penetration, essential for its 2060 carbon neutrality, increase variability due to greater fluctuations in electricity demand caused by the acceleration of electrification [17]. This, in turn, extends the uncertainties in China's NG market through 2060, ultimately making the projection of China's NG consumption by 2060 uncertain. Therefore, given the uncertain domestic and international NG market conditions in China, it is necessary to reassess whether China's NG consumption projection aligns with its carbon neutrality goal. For this reassessment, this study aims to project China's long-term NG consumption under its climate policy scenarios and then reevaluate the likelihood of achieving its carbon neutrality target with the projected NG consumption.

Current studies on scenario-specific long-term projections mainly utilise energy-economy models, climate and renewable models, regression analysis, and system dynamics models. However, these models limit this study's long-term NG consumption projections, which rely on big datasets and various climate policy scenarios. Energy-economy models focus on evaluating policy impacts [147]. In the same vein, climate and renewable models are also designed to assess these effects [148]. Regression analysis depends heavily on past trends [149]. System dynamics models can be inefficient when dealing with complex scenarios [150].

Moreover, they are not suited for handling large datasets [151]. This highlights the need for new advanced long-term projection modelling techniques based on big data analysis and various scenarios.

This research endeavours to utilise an innovative hybrid combination of machine learning and econometric approaches for reassessing China's NG consumption under scenarios. This hybrid model initially employs a machine learning model, Prophet, to project the selected key drivers in China's NG market [107] by 2060 under three scenarios, including its carbon neutrality target. Subsequently, the hybrid model uses an AutoRegressive Integrated Moving Average (ARIMA)-based model, Seasonal ARIMA with eXogenous regressors (SARIMAX), with the projected key drivers as inputs to forecast the NG consumption by 2060. The three scenarios are derived from International Energy Agency (IEA) scenarios [152] - Stated Policies Scenario (STEPS), Announced Pledges Scenario (APS), and Net Zero Emissions by 2050 (NZE) Scenario - as a reference point along with British Petroleum (BP) Energy [153] and Representative Concentration Pathways (RCP) scenarios [154]. The final findings of this study indicate that China's NG consumption projected by the hybrid model does not align with its genuine carbon neutrality target due to the climate impacts of methane emissions from NG. In response, this study, utilising the One Earth Climate Model (OECM) scenario of University of Technology Sydney (UTS) [155], suggests the country's NG phase-out pathways, which align with its carbon neutrality target.

The contribution of this study is threefold. First, the study introduces a novel hybrid model that combines the machine learning-based nonlinear model Prophet and the traditional linear forecasting model SARIMAX, applied to a scenario-based long-term projection framework, in the energy modelling and prediction field. Specifically, the study delves into the combined utilisation of Prophet and SARIMAX with the selected eight drivers in China's NG market [107] under the three scenarios. It demonstrates the superior performance of the Prophet–SARIMAX combination in scenario-based long-term projections, employing conservative estimations involving comparisons with other hybrid projection models or error bands. Second, the study not only demonstrates that China's NG consumption projected by 2060 under the current scenarios is incompatible with the country's 2060 carbon neutrality target, but also proposes alternative approaches that align with its carbon neutrality target—its NG phase-out pathways. In addition, the study also suggests policy implications drawn from the NG phase-out pathways. Finally, the study quantifies monthly projections of Chinese NG consumption until 2060 through the hybrid Prophet–SARIMAX model. While prestigious energy institutions like the IEA provide the NG consumption projections for specific years (e.g., 2030 and 2060), the hybrid model offers continuous monthly projections, useful for economic or policy analyses.

The rest of this chapter is organised as follows: Section 4.2 provides related literature pertaining to SARIMAX and Prophet models. Sections 4.3, 4.4, and 4.5 discuss data, scenarios, and methods and models for long-term projections, respectively. Section 4.6 presents the results, and the final section concludes the study. The supplementary materials include terms and abbreviations, tables, and figures used throughout this chapter.

4.2 Literature review

This chapter relates to two research strands: (1) energy modelling methods for scenario-based long-term projections and (2) the hybrid econometric-machine learning approach combining Prophet and SARIMAX.

Few studies have applied various modelling methods to scenario-based long-term projections in the energy field. A recent study, using climate and different renewable power models with emission scenarios, demonstrates that the uncertainties associated with climate change, such as seasonal differences, can have a significant impact on national power systems and play a crucial role in Europe's future power system planning and design by 2050 [148]. It indicates that climate models show significant variability in the spatio-temporal patterns of climate change's impacts, and even in the direction of wind and solar energy, and suggests that deeper climate uncertainty analysis is needed. Edelenbosch et al. introduce various long-term energy models used in the industry, including Integrated Assessment Models (IAMs), Computable General Equilibrium (CGE), and energy system models, and project the final energy demand up to 2100 based on two climate policy scenarios: baseline and mitigation scenarios [147]. The long-term energy models, a combination of simulation and intertemporal optimisation models, are referred to as Energy-Economy and IAM models, which are frequently used to analyse emission mitigation strategies and relevant investment costs. Aydin employs regression analysis to model the consumption of energy sources (CES) using trends and presents its projections by 2035 [149]. It demonstrates that the calculated results for CES, when compared with CES published by BP Energy and the IEA, appear to be within the acceptable range. Mu et al. use a system dynamics model with the VENSIM system dynamics software to forecast China's NG consumption, its CO_2 emissions from NG, etc., during 2016–2035 [156]. The system dynamics model explores the factors influencing NG's total demand and consumption structure across seven scenarios related to population growth, economic development, and technological progress.

By sorting out the literature so far, recent energy models for scenario-based long-term projections are primarily designed to simulate the broader energy economy or system under climate change or policy scenarios. However, these models can be limited in their applicability and accuracy when used solely for long-term NG projections, especially considering the key drivers in the NG market under various climate policy scenarios. First, IAMs, energy system, and CGE models based on climate scenarios have been used to assess long-term policy impacts [147]. Additionally, the climate and renewable power models based on emission scenarios have also been instrumental in this analysis [148]. However, these models are inappropriate for forecasting NG consumption, as they focus on evaluating policy scenarios. Second, regression analysis predicts future outcomes based on past trends, performs well when data patterns are consistent, and is relatively simple and fast in the calculation [149]. Yet, it can fail to capture potential future trends affecting long-term NG projections due to its reliance on historical trends. Lastly, system dynamic models using the VENSIM software are effective for complex scenarios and structural changes [156]. However, the system dynamics models are unsuitable for long-term monthly NG consumption forecasts using scenario-driven big data, as they become complicated with complex scenarios [150]. Additionally, they are time-consuming when dealing with large datasets, making them inefficient for such forecasts [151].

As for the methodologies, this study explores ARIMA-based models, including SARIMAX, and Prophet individually, as well as their integration. ARIMA-based models have long been used in predictive studies, while the newly developed Prophet model has gained popularity and demonstrated strong performance in recent predictive research.

Single ARIMA-based models cannot be suitable for scenario-based long-term projections. In general, ARIMA models and their variants have been proven useful with satisfactory long-term predictive performance in various fields of long-term prediction. Valipour compares the ARIMA and Seasonal ARIMA (SARIMA) models using the data from 1901 to 2010 for long-term runoff forecasting in the United States, finding that SARIMA outperforms ARIMA [157]. Chen et al. analyse long-term predictions of emergency department revenue and visitor volume using ARIMA, exploring correlations with meteorological, clinical, and economic factors [158]. McRae presents a long-term forecasting method for regional hospital demand, combining population and per-capita demand forecasts through ARIMA [159]. Andrews et al. demonstrate that both ARIMA and ARIMA with eXogenous regressors (ARIMAX) models provide accurate long-term disability claims forecasts, showing that including exogenous variables improves predictive performance [160]. Biswas et al. forecast wheat production in the Punjab region from 2010–11 to 2020–21 using the ARIMA method with 60 years of data [161]. In energy predictive studies, ARIMA models and their extensions are also utilised as useful tools. Sutthichaimethee and Ariyasajjakorn apply ARIMA to forecast 30-year energy use in Thailand's construction and materials sectors [162]. S. Ozturk and F. Ozturk forecast Turkey's energy consumption using ARIMA, projecting a continued increase through 2040 [163]. Citroen et al. use ARIMAX with Gross Domestic Product (GDP) and demographic factors to model Morocco's long-term electricity load [164]. However, ARIMA-based models, which rely on linear forms, can exhibit reduced accuracy in long-term projections with various scenarios. This is because these scenarios, which simulate real-world conditions, often reflect nonlinear characteristics that such models cannot capture effectively [165].

Prophet, which can handle nonlinear trends and complex seasonal patterns, overcomes the limitations of ARIMA and is widely used due to its practicality and flexibility. Facebook introduces the Prophet model in a 2017 article [166]. According to this article, the Prophet model provides four practical advantages: (1) the model has the flexibility to incorporate seasonal elements and trends easily; (2) unlike the ARIMA model, it is not necessary for the data measurements to be regularly spaced, and the model is unaffected by missing values; (3) the model's fast fitting speed allows for various trials; and (4) the model can be easily extended by intuitively controlling parameters. This practical and flexible Prophet model has been utilised in diverse time-series fields, consistently demonstrating robust results in the prediction. Basak et al. demonstrate that the Prophet model outperforms Support Vector Regression (SVR) and Multiple Linear Regression (MLR) models in meteorological drought forecasting [167]. Sivaramakrishnan et al. indicate that the Prophet model enables easier data analysis than the ARIMA model for forecasting airline passengers [168]. Khayyat et al. propose the Prophet model as a prediction and analysis model for COVID-19 in Saudi Arabia, showcasing high accuracy in predicting deaths [169]. Jha and Pande conclude that the Prophet model shows lower errors and better predictive performance than ARIMA in supermarket sales forecasting [170]. Likewise, Yenidoğan et al. indicate that the Prophet model performs better than ARIMA in Bitcoin forecasting [171]. Moreover, compared with ARIMA-based models, the Prophet model also demonstrates its superiority in long-term prediction [172].

Hybrid frameworks combining Prophet and ARIMA-based models can perform better than individual models in scenario-specific long-term projections. Recent studies have compared single Prophet and ARIMA-based models with other machine learning models, demonstrating each model's predictive superiority. Haider et al. argue that SARIMAX and Prophet models are efficient for long-term forecasting by comparing individual SARIMAX and Prophet models with machine learning approaches, such as Long Short-Term Memory (LSTM), Convolutional Neural Networks (CNN), and Artificial Neural Networks (ANN) [173]. Moreover, compared to the individual Prophet and SARIMAX models showcasing excellence in long-term forecasting, hybrid forms of the models enhance predictive performance. Luo et al. employ the combined SARIMAX and Prophet model to forecast the incidence of AIDS in Henan, suggesting using the combined model in AIDS surveillance [174]. Saqware and Beary develop hybrid SARIMA-Prophet models to predict staple food prices in Tanzania, and their results show that the hybrid models have a more significant predictive performance than individual models [175]. Zhao and Zhang show that their combined Prophet–SARIMAX model has higher prediction accuracy in sales volume than individual models [176].

Combining models in the order of Prophet–SARIMAX can enhance the performance of a hybrid model. Like Prophet, the model that estimates and analyses various data components to explain data variability is the Unobserved Components Model (UCM), which comprises trend, cycle, seasonal, and irregular components [177]. UCM can flexibly model changes in seasonality and trends similar to Prophet [177]. However, unlike UCM, Prophet is designed to handle nonlinear elements [166]. By using Prophet's nonlinear projections as inputs for SARIMAX, which is based on a linear form, Prophet–SARIMAX can incorporate the nonlinear elements captured by Prophet into SARIMAX, thus improving prediction performance [176].

From the literature, recent energy models that simulate climate change or policy scenarios are limited in applicability and accuracy for long-term NG projections, and existing regression analysis and system dynamics models are inadequate for scenario-based big data analysis. On the other hand, hybrid models that combine Prophet and ARIMA-based models demonstrate excellent capabilities in long-term forecasting. Furthermore, hybrid models incorporating machine learning approaches can benefit big data analysis. Therefore, the hybrid models based on these superior predictive modelling techniques can be leveraged for scenario-based long-term projections in the energy field [178]. This study introduces a hybrid Prophet–SARIMAX model that has yet to be utilised in scenario-based long-term projections in the energy field, and demonstrates its suitability for scenario-specific projections. Additionally, it reassesses the feasibility of China's carbon-neutral pathways through NG consumption projections generated by the hybrid model.

4.3 Data

In this study, Prophet projects each key driver using one of the eight key driving factors as the independent variable and Chinese NG consumption as the dependent variable. Then, SARIMAX uses the projected key drivers to project Chinese NG consumption until 2060. These eight drivers are selected from 40 factors related to the NG consumption through an innovative amalgamation of machine learning and econometric techniques - LASSO/ALASSO and ECM techniques [107]. In detail, LASSO effectively identifies key drivers influencing NG consumption, while ALASSO serves as a validation method to ensure the reliability of LASSO results. ECM distinguishes the short- and long-term impacts of these drivers, demonstrating that they have both immediate and lasting effects. The selected drivers are Price distortion, Temperature, HH gas price, piped NG (PNG) imports, Thermal coal price, Gas pipeline capacity, LNG terminal capacity, and LNG imports. The drivers belong to the five groups, implying that they incorporate different characteristics. All the variables consist of time series data every month, spanning from April 2009 to July 2023 due to the availability of the price distortion factor. Price factors among the variables are represented on a constant or real basis. Table 4.1 (see Appendix J) provides a detailed description of these factors, their sources, and the LASSO results.

The selected drivers are crucial inputs for the long-term projections of this study. The 40 factors, categorised into six groups from existing literature[28], encompass economic trends, energy policies, environmental considerations, technological advancements, and global trade dynamics [107]. Specifically, these factors capture the paradoxical role of NG in achieving carbon neutrality—serving as a transitional fuel in the short term [3] while being phased out in the long term [4]. They also reflect short-term uncertainties in the NG market due to market structure changes resulting from market liberalisation [14] and global market instability caused by geopolitical conflicts, including the Russian-Ukraine war [16], as well as long-term uncertainties from electricity demand variations related to electrification goals and renewables expansion for carbon neutrality [17]. The eight key drivers are identified from these shifting dynamics and uncertainties by considering their short- and long-term effects [107]. Hence, these drivers are highly suitable for the long-term projections.

Like many real-world cases, most of the key drivers are non-stationary variables. As shown in Table 4. 2 (see Appendix J), an ADF test, a widely used unit root test, is carried out to check stationarity in the variables. The variables, except Price distortion, are non-stationary at level and become stationary at the first difference. The successful construction of SARIMAX involves differencing non-stationary data [108]. Besides, it is an appropriate method for automating this process [95]. Prophet extracts temporal features, such as trend and seasonality, from actual historical (non-stationary) variables and can be readily extended by intuitively adjusting the parameters related to the variables [179]. Therefore, the key drivers, which are non-stationary variables, can be directly used for hybrid Prophet–SARIMAX modelling. Table 4.2 (see Appendix J) provides detailed properties of the variables.

4.4 Scenarios

This section introduces three scenarios from IEA, BP Energy, and RCP, and then present the OECM scenario used for the NG phase-out pathways. Details of the scenarios are provided below.

4.4.1 Three scenarios

The hybrid Prophet–SARIMAX model generates projections of each key driver for

[28] The six groups for 40 factors are (1) Economics & Energy (15 factors); (2) Global Energy Index (5 factors); (3) Urbanisation & Demographics (4 factors); (4) Infrastructure (4 factors); (5) Environment, Climate, and Technological growth (7 factors); and (6) Price distortion (5 factors). These factors are sourced from global data providers (Bloomberg, Refinitiv, CEIC, CCTD, and Sxcoal), government institutions (EIA, NBS, and KEEI), international organizations (World Bank and OECD), and other sources (UTS ACRI and Chen et al.'s research [110]).

Chinese NG consumption up to 2060 under three scenarios and then utilises the projected drivers to estimate the NG consumption projection until 2060. These scenarios are derived from IEA scenarios: STEPS, APS, and NZE [152]. In this study, the same scenario names are used. To be specified, the STEPS offers a projection based on the most recent policy configurations, encompassing energy, climate, and related industrial policies. The APS assumes that all national energy and climate targets governments declare are achieved fully and punctually. In the NZE, significant additional progress is required to restore global warming to 1.5 °C. These IEA scenarios serve as reference points for the three scenarios of the projected key drivers. In contrast, the scenarios for these drivers not covered by the IEA scenarios are drawn from BP Energy [153] and RCP scenarios [154].

In the Prophet model, IEA scenarios are utilised as inputs for projecting each key driver until 2060. In cases where IEA scenarios cannot be applied to certain key driving factors, analogous scenarios from BP Energy and RCP, along with relevant academic journal articles, are employed as alternatives. Specifically, the three scenarios from BP Energy (New Momentum, Accelerated, and Net Zero) and the three scenarios from RCP (RCP 8.5, RCP 6.0, and RCP 4.5) are considered equivalent to IEA scenarios (STEPS, APS, and NZE), respectively. The IEA scenarios reflect current policies, with emissions peaking around 2030 and then declining [152]. The timing and extent of this peak and reduction vary by scenario (STEPS, APS, and NZE). Similarly, BP [153] and RCP scenarios [154] are built around an emissions peak and decline, as detailed in the tables of drivers for each scenario (see Tables 4.4 to 4.9). The scenarios presented by each institution are summarised in Table 4.3 (see Appendix K), grouped as analogous scenarios. The detailed scenarios applied to key driver projections are as follows.

4.4.1.1 International Energy Agency scenarios

The IEA has published fossil fuel prices for 2030 and 2050 for each scenario—STEPS, APS, and NZE [152]. Rather than using the published prices directly, this study applies the Compound Annual Growth Rate (CAGR) [29], calculated from those prices, to Prophet, generating monthly key driver projections. The CAGRs for 2030 are calculated based on IEA’s prices for 2022 and 2030, and the CAGRs for 2050 are calculated using IEA’s prices for 2030 and 2050, with the CAGRs for 2050 assumed for the period from 2050 to 2060. The CAGRs across the three scenarios are shown in Table 4.4.

The key drivers influenced by the IEA scenarios are Thermal coal price, HH gas price, and Price distortion. On the one hand, the steam coal price scenarios for China and the NG price scenarios for the U.S., published by the IEA, are directly used in thermal coal price and HH gas price projections, respectively. On the other hand, the fossil fuel price scenarios from the IEA are not applied to price distortion itself but are indirectly used as input variables for price distortion calculation. Table 4.1 (see Appendix J) shows that the price distortion calculation is computed using City Gate Gas Price (CGGP) and PNG price. First, this study assumes that CGGP is calculated based on HH gas and Brent oil prices [36]. Thus, the study projects the CGGP using the HH gas prices predicted from the IEA's U.S. NG price scenarios and the Brent oil prices forecasted from the IEA's crude oil price scenarios. Second, it projects the PNG prices

[29] CAGR is calculated as $(V_{final}/V_{begin})^{1/t}$ - 1, where V_{begin} is the beginning value; V_{final} the final value; and t the time in years.

based on the IEA's Chinese NG price scenarios. Finally, the predicted CGGP and PNG prices are used to project the price distortion. It is worth noting that the IEA assumes a reduction of fossil fuel prices in all scenarios between 2030 and 2050, with the most significant price drop in the NZE scenario due to a substantial decrease in oil, gas and coal demand.

Table 4.4 Fossil fuel prices by IEA scenarios

Item	STEPS		APS		NZE	
	2030	2050	2030	2050	2030	2050
Relevant key driver: Thermal coal price						
Steam coal in China (USD/MT)	96	80	79	62	64	49
CAGR (%)	-9.05	-0.91	-11.24	-1.20	-13.54	-1.33
Relevant key driver: HH gas price & Price distortion – CGGP (HH gas)						
NG in the U.S. (USD/MMBtu)	4.0	4.3	3.2	2.2	2.4	2.0
CAGR (%)	-2.99	0.36	-5.66	-1.86	- 8.99	0.91
Relevant key driver: Price distortion – CGGP (Brent oil)						
Crude oil (USD/barrel)	85	83	74	60	42	25
CAGR (%)	-1.76	-0.12	-3.45	-1.04	-10.05	-2.56
Relevant key driver: Price distortion – PNG price						
NG in China (USD/MMBtu)	8.4	7.7	7.8	6.3	5.9	5.3
CAGR (%)	-5.93	-0.43	-6.80	-1.06	-10.00	-0.53

Notes: The fossil fuel prices are presented in real terms. The prices of China's steam coal, the U.S. NG, crude oil, and China's NG price in 2022, as reported by the IEA, are 205 USD/MT, 5.1 USD/MMBtu, 98 USD/barrel, and 13.7 USD/MMBtu, respectively.

4.4.1.2 British Petroleum Energy scenarios

The key driver influenced solely by BP Energy scenarios is LNG imports. Key drivers influenced by both BP Energy scenarios and relevant studies include PNG imports, Gas pipeline capacity, and Accumulated LNG terminal capacity. The LNG and PNG imports, sourced from BP Energy scenarios and the papers, are measured in billion Cubic metres (BCM), while the units for LNG and PNG imports in the key drivers are in million metric tons (MMT). To ensure unit compatibility, the values converted from BCM to MMT are provided in Tables 4.5 and 4.6.[30] Similar to the IEA scenarios, CAGRs calculated from these key driver scenarios are used for monthly projections of the key drivers. BP Energy estimates a significant demand for LNG in 2050, even under the 'Net-Zero' scenario.

Table 4.5 LNG imports by BP Energy scenarios

Item	**New Momentum**		**Accelerated**		**Net Zero**	
	2030	**2050**	**2030**	**2050**	**2030**	**2050**
Relevant key driver: LNG imports						
LNG imports (BCM)	149	133	130	46	130	33
LNG imports (MMT)	109	98	95	33	95	24
CAGR (%)	5.56	-0.56	4.21	-5.08	4.25	-6.59
Ratio of imports compared to New Momentum	-	-	0.87	0.34	0.87	0.25

Note: The LNG imports in 2019, as reported by BP Energy, are 82 BCM (60 MMT).

Table 4.5 shows BP Energy's projections for LNG imports in 2030 and 2050 across three scenarios: New Momentum, Accelerated, and Net Zero [153]. Like the IEA scenarios, LNG import projections are generated by applying the CAGRs calculated based on the announced imports to the Prophet model. Based on BP Energy's announced 2019 LNG imports, CAGRs

[30] According to the conversion factors provided by BP Energy [180] and the Institut français des relations internationales (IFRI) [181], the value of 1 BCM is equivalent to 0.73 million tonnes of LNG for LNG imports. Additionally, for other types of NG, the conversion factor of 1 BCM = 0.86 million tonnes of oil equivalent is applied, based on reports from the World Bank [182] and the IEA [152].

for three scenarios are calculated up to 2030. The CAGRs for 2050 are derived from announced imports for 2030 and 2050, and these 2050 CAGRs are assumed for the period 2050 - 2060.

In both the IEA and BP Energy scenarios, projections for PNG imports are not included. Consequently, the projections for PNG imports derived from reports that reflect current plans or policies are considered to align with the New Momentum scenario. According to Wang et al., in 2020, the PNG imports mainly from Turkmenistan, Uzbekistan, Kazakhstan, Myanmar, and Russia was about 48 BCM [183]. By 2025, the China-Russia East-Route Natural Gas Pipeline is expected to increase imports by 30 BCM [183]. As a result, the total PNG imports are expected to reach 78 BCM by 2025. This is validated by “Figure 7. Chinese gas production, demand and imports (bcm)” in the Gas Exporting Countries Forum (GECF) report [184]. Beyond 2025, PNG imports reflect the projections from the GECF report. Imports are expected to reach 99 BCM in 2030, 148 BCM in 2040, and 132 BCM in 2050 [184]. Based on these two complementary reports, PNG import projections from 2025 to 2050 are assumed to follow the New Momentum scenario. The projections for PNG imports in the Accelerated and Net Zero scenarios are calculated by using the 'Ratio of imports compared to New Momentum' outlined in Table 4.5. The CAGRs for the years 2025, 2030, 2040, and 2050 are estimated based on projected PNG imports for each scenario, with the CAGRs for 2050 assumed for 2050 to 2060. Table 4.6 illustrates the PNG import projections and CAGRs across the three scenarios.

Table 4.6 PNG imports by BP scenarios

Item	New Momentum				Accelerated				Net Zero			
	2025	2030	2040	2050	2025	2030	2040	2050	2025	2030	2040	2050
Relevant key driver: PNG imports												
PNG imports (BCM)	78	99	148	132	67	86	89	45	68	86	83	33
PNG imports (MMT)	67	85	127	113	58	74	77	39	58	74	71	28
CAGR (%)	10.27	4.97	4.11	-1.17	7.20	4.97	0.36	-6.65	7.29	4.97	-0.45	-8.86

Note: The 2020 PNG imports of 48 BCM (41 MMT) are used for the CAGR calculation for 2025.

The projections for Accumulated LNG terminal capacity and Gas pipeline capacity are made incrementally based on current policies and operational plans, without using the Prophet model (see Section 4.5.1.3 for details). The stepwise projections, based on relevant studies on accumulated LNG terminal capacity and gas pipeline capacity, are assumed to represent the New Momentum scenario.

The designed capacity of Chinese LNG terminals is anticipated to achieve 190 million metric tons per annum by 2025, according to Wang et al.'s report [183]. As per the GECF report, the Chinese government aims to attain an accumulated annual LNG import capacity of 247 million metric tons by 2035 [184]. The phased projections of annual LNG terminal capacity up to 2035 are already determined by current plans and policies. Thus, the phased projections for each year in 2025 and 2035 are uniformly applied across all scenarios. By using the CAGRs of LNG imports for each scenario from Table 4.5, the accumulated LNG terminal capacity is projected to decrease from 2035 to 2060. Table 4.7 shows the accumulated LNG terminal capacity and CAGRs for each scenario.

With the assumption that Central Asia-D and Power of Siberia-3 are expected to be operational by 2026, this research projects a substantial increase in the annual gas pipeline capacity, reaching 40 BCM [185]. This will bring the total gas pipeline capacity to 145 BCM by that year, including the existing capacity of 105 BCM currently in operation [110]. Similarly, with the projected operational commencement of Power of Siberia-2 assumed for 2030, the research anticipates another reliable increase in the annual gas pipeline capacity, reaching 50 BCM. This will result in a cumulative gas pipeline capacity of 195 BCM by that time [110, 185]. The annual cumulative gas pipeline capacity (195 BCM) projected by 2030 (as shown in Table 4.8) exceeds the annual PNG imports (147 BCM) in the New Momentum scenario for 2040 (as seen in Table 4.6). Therefore, it is assumed that the gas pipeline capacity in the New Momentum and Accelerated scenarios is anticipated to accumulate to 195 BCM by 2040. By 2050, it is assumed that the gas pipeline capacity in these scenarios is projected to decrease based on the CAGRs of PNG imports (as indicated in Table 4.6). In the Net Zero scenario, the gas pipeline capacity is assumed to be consistent with other scenarios until 2030 and, starting from 2040, is supposed to decrease based on the CAGRs of PNG imports. The CAGRs for 2050 across all scenarios are applied to the period from 2050 to 2060. Table 4.8 illustrates the gas pipeline capacity and CAGRs for each scenario.

Table 4.7 Accumulated LNG terminal capacity by BP scenarios

Item	New Momentum				Accelerated				Net Zero			
	2025	2035	2050	2060	2025	2035	2050	2060	2025	2035	2050	2060
Relevant key driver: Accumulative LNG terminal capa												
LNG terminal capa (10kt)	19,000	24,700	22,709	11,357	19,000	24,700	11,294	6,704	19,000	24,700	8,877	4,487
CAGR (%)	29.34	2.66	-0.56	-0.56	29.34	2.66	-5.08	-5.08	29.34	2.66	-6.59	-6.59

Note: The accumulated LNG terminal capa in 2023 is 113,570 kt, which is used for the CAGR calculation in 2025.

Table 4.8 Gas pipeline capacity by BP scenarios

Item	New Momentum				Accelerated				Net Zero			
	2026	2030	2040	2050	2026	2030	2040	2050	2026	2030	2040	2050
Relevant key driver: Gas pipeline capa												
Accumulative capa (BCM)	145	195	195	174	145	195	195	98	145	195	186	74
CAGR (%)	-	-	-	-1.17	-	-	-	-6.65	-	-	-0.45	-8.86

4.4.1.3 Representative Concentration Pathways scenarios

Temperature projections are generated based on the RCP scenarios in Wang et al.'s journal article [154]. Specifically, the temperature projections for 2040 and 2070 are generated by applying scenario-specific CAGRs to the Prophet model; these CAGRs are derived from the temperature data in the journal article. The CAGRs for 2040 are calculated based on the 2018 temperature and 2040 in the journal article, and the CAGRs for 2070 are computed utilising the temperatures published for 2040 and 2070, with the CAGRs for 2070 assumed for 2050 to 2060. The projected temperatures and CAGRs across the three scenarios are shown in Table 4.9.

Table 4.9 Temperature by RCP scenarios

Temperature by RCP scenarios	RCP 8.5		RCP 6.0		RCP 4.5	
	2040	2070	2040	2070	2040	2070
Relevant key driver: Temperature						
Average national temperature (°C, Celsius)	17.4	18.8	16.9	17.6	17.0	17.6
CAGR (%)	0.33	0.26	0.22	0.14	0.25	0.12

Notes: The temperature in 2018 for RCP 8.5 is 16.2 ℃, while for RCP 6.0 and 4.5, it is 16.1 °C.

4.4.2 One Earth Climate Model

As an alternative approach to carbon neutrality in China, this study delves into the trajectories of key drivers under the OECM scenario, specifically the Net-zero 1.5 °C scenario

[155]. According to the OECM scenario, by around 2050, Chinese primary energy demand, including non-energy use, accounts for 88% of the total energy demand in China. Within this context, 98% of Chinese primary energy demand is sourced from renewable energy, with the remaining 2% coming from nuclear power. This OECM scenario assumes that China's demand for fossil fuels, such as NG, coal, and oil, would disappear by 2060. Accordingly, this study assumes that by 2060, the prices of fossil fuels will decline significantly to the point where their economic value is effectively considered to be zero. Based on the OECM scenario and its assumptions, the study utilises the hybrid Prophet–SARIMAX model to illustrate the key driver pathways necessary to achieve zero NG demand in China by 2060. The OECM scenario is depicted in Figure 4.1.

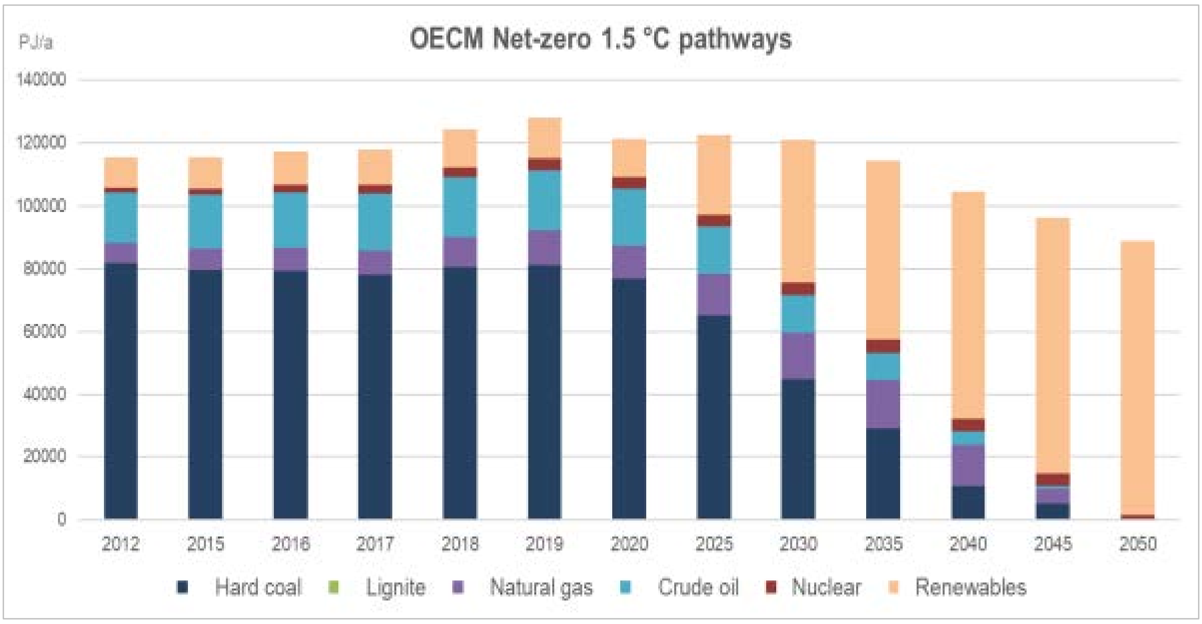


Figure 4.1 Chinese primary energy demand in the OECM scenario

4.5. Methods and Models

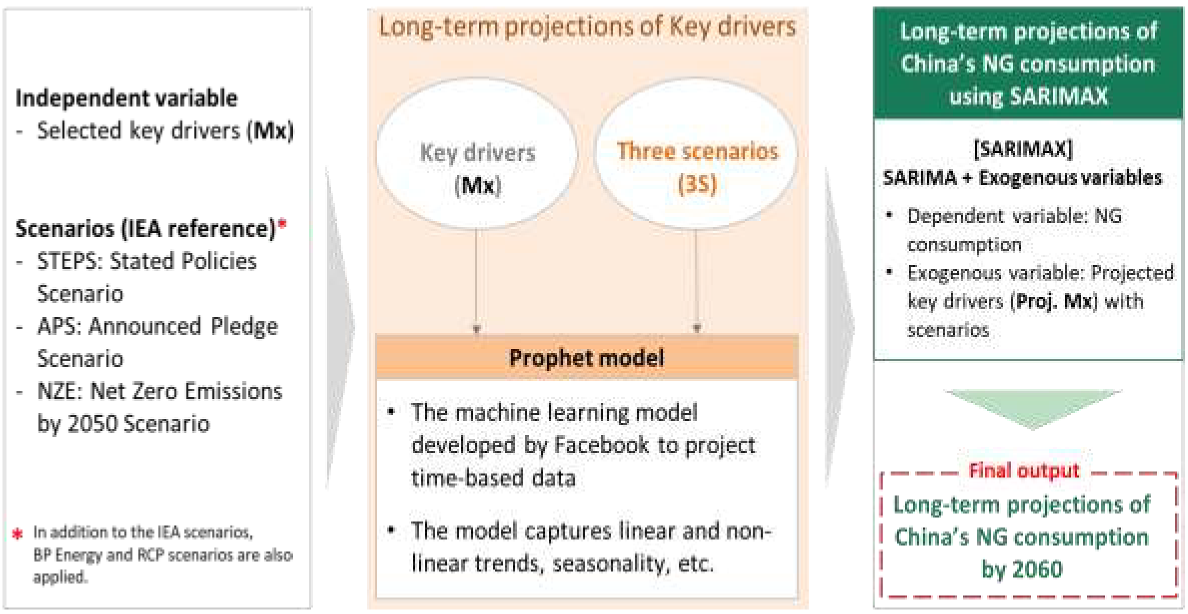


Figure 4.2 Scenario-based long-term projection model

The scenario-based long-term projection model employs a hybrid approach, sequentially integrating Prophet and SARIMAX, to estimate Chinese NG consumption under three specific scenarios (3S) until 2060. This sequential integration leverages the strengths of both models: Prophet's capacity to capture nonlinear trends and SARIMAX's linear framework. By combining these approaches, this hybrid model enhances prediction accuracy, retaining Prophet's nonlinear insights while refining them through SARIMAX's optimal linear predictive function. In contrast, using SARIMAX's linear input for Prophet underutilizes its nonlinear capabilities and contradicts its intended design [186]. Therefore, the hybrid method sequentially combining Prophet and SARIMAX is highly effective for scenario-based forecasting.

As depicted in Figure 4.2, the long-term projection model involves a two-stage process. In the initial stage of the hybrid modelling process, each historical key driver (Mx) is utilised to fit a Prophet model. Subsequently, the key driver (Mx) is projected up to 2060 using the fitted Prophet model within the three scenarios. In the second stage, the projected key drivers (Proj. Mx) serve as inputs for SARIMAX to project Chinese NG consumption for each scenario.

4.5.1 Long-term projection

This section first explains Prophet and SARIMAX models, and then describes the hybrid method that combines them sequentially.

4.5.1.1 Prophet

Prophet is designed as a model that considers the characteristics of time series data while remaining independent of time dependencies, enabling relatively straightforward modelling of large time series datasets [166]. In general, ARIMA, like other traditional time series models, exhibits a structure that is dependent on time. However, Prophet is not time-dependent. When Prophet considers time-dependent factors such as seasonality and holiday effects, it can adjust them using additional regressors. This flexibility allows Prophet to handle various data patterns more effectively than conventional time series models.

Moreover, Prophet is widely used in time series data forecasting due to its user-friendliness. Prophet addresses problems through curve fitting, where ‘curve fitting’ refers to the process of identifying an appropriate curve for the data [166]. This feature enables Prophet to provide high-level prediction performance without requiring users to configure the model explicitly.

The parameters constituting Prophet are Trend function, Seasonality, and Holidays [187]. The trend function models non-periodic changes in the time series, with change points controlling the trend's smoothness. Seasonality controls the frequency of seasonal patterns, with the seasonality mode specifying whether the seasonal cycle is additive or multiplicative. Holidays control the strength of holiday effects. The mathematical representation is as follows:

$$\boldsymbol{Prophet}:\; \boldsymbol{y_t} = \boldsymbol{g}(\boldsymbol{t};\, \boldsymbol{\alpha}) + \boldsymbol{s}(\boldsymbol{t};\, \boldsymbol{\beta}) + \boldsymbol{h}(\boldsymbol{t};\, \boldsymbol{\gamma}) + \varepsilon_t \qquad \textbf{(1)}$$

where $\boldsymbol{g}(\boldsymbol{t};\, \boldsymbol{\alpha})$ represents the piecewise-linear trend (or “growth term”), where $\boldsymbol{\alpha}$ is a model parameter that controls the smoothness or signal-to-noise ratio of the trend component; $\boldsymbol{s}(\boldsymbol{t};\, \boldsymbol{\beta})$ refers to various seasonal patterns (periodic changes), where $\boldsymbol{\beta}$ is a vector of model parameter that specifies the mode of and frequency of each seasonal pattern; $\boldsymbol{h}(\boldsymbol{t};\, \boldsymbol{\gamma})$ captures holiday effects, where $\boldsymbol{\gamma}$ is a model parameter that reflects holidays or special events; and $\boldsymbol{\varepsilon_t}$ denotes the white noise error term (idiosyncratic changes) [166]. All these parameters are optimised based on maximum a posteriori estimation with cross-validation regularization or an analyst-in-the-loop modelling approach that enables direct expert manipulation [166].

4.5.1.2 Seasonal AutoRegressive Integrated Moving Average with eXogenous regressors

SARIMAX is a development of seasonal ARIMA (SARIMA) with exogenous factors (X), known as SARIMAX (p, d, q) * $(P, D, Q)^s$, where SARIMA is an enhanced model designed to improve the effectiveness of ARIMA models, especially for modelling time series data with seasonal patterns [94]. SARIMAX (p, d, q) and (P, D, Q) represent non-negative integers indicating the polynomial order of the autoregressive (AR), integrated (I), and moving average (MA) parts of the non-seasonal and seasonal components of the model, respectively [94]. SARIMAX is mathematically expressed as follows:

$$\boldsymbol{SARIMAX}\text{:}\ \ \boldsymbol{\varphi_p(B)\phi_P(B^s)\nabla^d\nabla_s^D\, y_t = \beta_k x_{k,t}{}' + \theta_q(B)\Theta_Q(B^s)\varepsilon_t} \quad \textbf{(2)}$$

where:

- $\boldsymbol{\varphi_p(B)}$ $(= 1 - \sum_{i=1}^{p} \varphi_i B^i)$ is the characteristic AR polynomial of order p.
- $\boldsymbol{\theta_q(B)}$ $(= 1 + \sum_{i=1}^{q} \theta_i B^i)$ is the characteristic MA polynomial of order q.
- $\boldsymbol{\phi_P(B^s)}$ $(= 1 - \sum_{i=1}^{P} \phi_i B^{s,i})$ is the seasonal characteristic AR polynomial of order P.
- $\boldsymbol{\Theta_Q(B^s)}$ $(= 1 + \sum_{i=1}^{Q} \phi_i B^{s,i})$ is the seasonal characteristic MA polynomial of order Q.
- $\boldsymbol{\nabla^d}$ $(= 1 - B)^d$ and $\boldsymbol{\nabla_s^D}$ $(= 1 - B^s)^D$ denote the non-seasonal differencing operator and seasonal, respectively, where $\boldsymbol{d}$ and $\boldsymbol{D}$ denote the order of differencing and the seasonal differencing, respectively.
- $\boldsymbol{y_t}$ is the prediction variable at time $\boldsymbol{t}$.
- $\boldsymbol{x_{k,t}{}'}$ refers to the vector associated with the $\boldsymbol{k^{th}}$ exogenous input variables at time $\boldsymbol{t}$, and $\boldsymbol{\beta_k}$ denotes the coefficient value of the $\boldsymbol{k^{th}}$ exogenous variable.
- $\boldsymbol{B}$ indicates the backshift operator (e.g., $B^s\ (y_t) = y_{t-k}$).
- $\boldsymbol{s}$ denotes the length of the seasonal pattern (e.g., s = 12 for monthly series).
- $\boldsymbol{\varepsilon_t}$ signifies the error terms.

4.5.1.3 Projection of key driving factors

As shown in Figure 4.3, using the Prophet model to project scenario-specific projections involves six of eight key drivers: Coal price, PNG imports, LNG imports, Temperature, HH gas price, and Price distortion. The initial step for the projection of each key driver involves training the Prophet model by fitting it with full sample data of the driver (①), historical escalation (growth) rate (②), and historical caps (③). The full sample data is available from April 2009 to July 2023; the historical escalation rate is assumed to be set at a constant rate of '1'[31], and the historical caps are determined based on the historical data's average annual or monthly values. In the training phase, the historical data and the historical caps are indispensable independent variables, and the historical escalation rate serves as an additional

[31] Since the historical (full sample) data already includes the escalation rate, it is assumed to be '1' without applying an additional rate.

regressor. Subsequently, the fitted model is utilised to generate the projection of each driver within three scenarios. In the projection modelling phase, the future caps or escalation rates are used as scenario-specific inputs (④) for the projection. Only the future caps are essential independent variables, while the future escalation rate is an additional regressor. The future caps or escalation rates are derived from the CAGRs, which are calculated based on the presented prices and quantities for each scenario (see Section 4.4.1 Three scenarios for the CAGR calculation). Eventually, projected key drivers are used as inputs in SARIMAX to project Chinese NG consumption. The detailed modelling elaboration of each driver's projection—scenario reference, inputs for model fitting and projection, and parameters for Prophet modelling—is provided in Table 4.10 (see Appendix L).

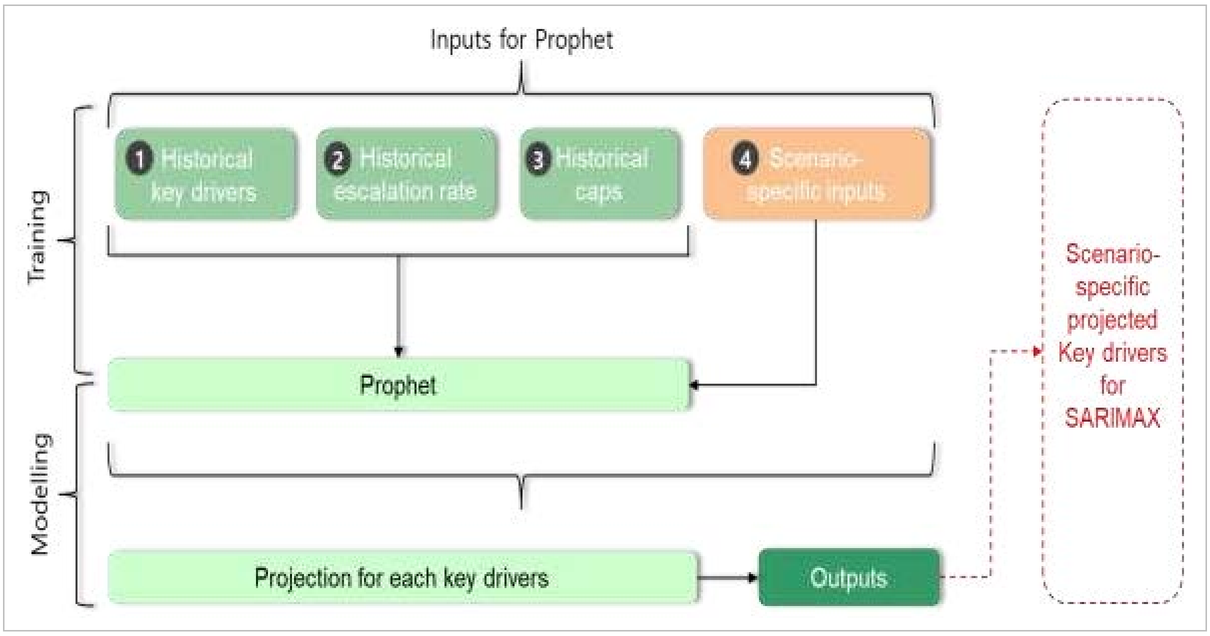


Figure 4.3 Prophet-based projection of six key drivers

In Prophet modelling, the CAGR has been chosen due to its reasonable estimation. The CAGR does not measure the uncertainty, which is hard to quantify without strong assumptions about future policy and market responses [188]. However, it offers an analytically trackable growth pattern [189]. Additionally, assuming a consistent pattern in trend changes, Prophet can generate smoother and more coherent projections for key drivers using CAGRs [166]. Moreover, since energy consumption has an exponential relationship with societal growth [190], using the compound growth rate is an effective approach for these projections.

On the other hand, two other NG infrastructure-related factors, Gas pipeline capacity and Accumulated LNG terminal capacity, are projected step by step without utilising the Prophet model. This is because considering the operational aspects of these two factors regarding both construction periods and relevant policies makes continuous increases and decreases challenging for projection. Until 2035/2040, the two factors are projected to increase step by step according to government plans, and the step-by-step increased projections are assumed to be the same across the three scenarios. From 2035/2040 onwards, it is assumed that they would decrease according to the three scenarios of the LNG and PNG imports presented by BP Energy. This choice of using the BP Energy scenarios is determined because two other NG infrastructure-related factors, Gas pipeline capacity and Accumulated LNG terminal capacity, are not provided in existing scenarios presented by the IEA. Therefore, the stepwise projection for the factors is established primarily based on the government's plans until 2035, forming the base case. From 2035 onwards, it is assumed that the factors decrease in line with the declining trend in the three LNG/PNG import scenarios of BP Energy. In the end, two projected NG infrastructure-related factors are utilised as inputs in SARIMAX for the projection of Chinese NG consumption. The details in the stepwise projection are presented in Table 4.11 (see Appendix M).

4.5.1.4 Projection of Chinese NG consumption

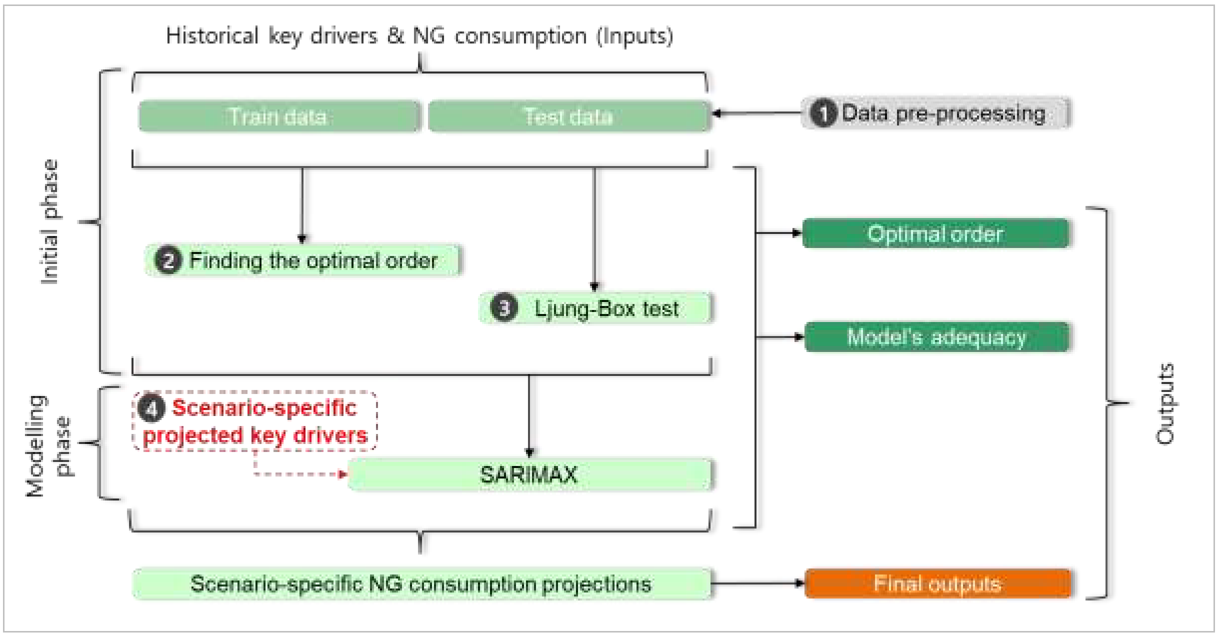


Figure 4.4 SARIMAX-based projection of Chinese NG consumption

As indicated in Figure 4.4, Prophet-based or stepwise projected key drivers are ultimately used as inputs to project Chinese NG consumption through SARIMAX. The projection of Chinese NG consumption is processed in two phases: the initial phase and the modelling phase. In the initial phase, firstly, data pre-processing (①) is conducted. Specifically, the data pre-processing standardises eight historical key drivers, exogenous variables, through MinMaxScaler, a popular scaling algorithm in machine learning [42]. Then, this procedure splits scaled exogenous variables and dependent variables (Chinese NG consumption) into two sets [94]—80% training and 20% test sets—without shuffling because the temporal order of the variables is crucial [114]. Secondly, the optimal order (②) for SARIMAX is determined through two Python libraries—auto ARIMA and statsmodels SARIMAX—with the training and test sets. The optimal order is initially found by auto ARIMA, which automates the selection of the best SARIMAX model by identifying the model with the lowest AIC. Subsequently, to demonstrate the robustness of the previously identified optimal order, this research uses statsmodels SARIMAX, which manually selects the best SARIMAX model by specifying the model orders in advance. The optimal order is identified with the enabled trend-capturing feature of SARIMAX to reflect the trend in historical data. Thirdly, after finding the optimal order, the Ljung-Box test (③) is conducted on the fitted model with the optimal order. This serves as a diagnostic test for assessing the adequacy of the model.

In the modelling phase, scenario-specific projected key drivers (④) are utilised as inputs in SARIMAX for the projection of Chinese NG consumption. The projection is generated under the three scenarios—STEPS, APS, and NZE. When the SARIMAX model performs the projection, the trend-capturing feature is turned off, in contrast to the process of finding the optimal order.[32] In the case that the trend-capturing feature is enabled, the model can produce unreliable projections. Specifically, due to the continuously increasing trend in the historical data, the model can incorrectly project the output as an increasing trend even if the input indicates a decreasing trend. This discrepancy occurs because the trend-capturing feature forces the projection to align with the historical increasing trend, thereby skewing the output [191]. Furthermore, based on the assumption that the trends of key drivers can explain the trend in China's NG consumption, the trend-capturing feature is disabled in the scenario-specific projection of the NG consumption. The long-run behaviour of gas consumption can be explained by other trending variables [192]. In particular, China's NG consumption is cointegrated with the key drivers, demonstrating that these drivers can explain the long-term dynamics of the NG consumption [107]. Moreover, scenario-specific trends are already embedded in the key drivers. Hence, given that the scenario-specific trends of the key drivers account for the trend in the projected NG consumption, explicit trend specification in SARIMAX is unnecessary.

4.5.2 Comparison of hybrid models

As illustrated in Figure 4.5, the hybrid Prophet–SARIMAX model employed in this study demonstrates robust projections of Chinese NG consumption through two comparison methods from a conservative estimation. The first comparison method ("Comparison 1") involves comparing the projections derived from two hybrid models—Prophet–Prophet and Prophet–SARIMAX—with those presented by IEA or BP Energy. The second comparison method ("Comparison 2") is based on the error bands of two hybrid models: Linear Interpolation Method (LIM)–SARIMAX and Prophet–SARIMAX.

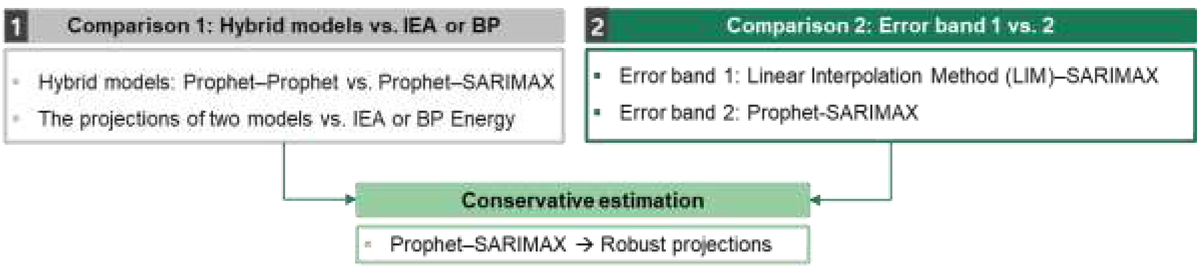


Figure 4.5 Two conservative comparison methods

4.5.2.1 Comparison 1

Like the long-term projection model—the hybrid Prophet–SARIMAX model—in this study, the hybrid Prophet–Prophet model generates projections for Chinese NG consumption with key drivers projected by Prophet as inputs, utilising Prophet instead of SARIMAX. Similar to SARIMAX in the Prophet–SARIMAX model, the final Prophet in the Prophet–

[32] During the process of finding the optimal order, a linear trend is added to the model, which refers to the enabled trend-capturing feature. In contrast, during the projection, the model is set to no trend, which indicates the disabled trend-capturing feature.

Prophet model generates projections with its trend-capturing feature disabled. The projections by the Prophet–Prophet model are first compared to those by the Prophet–SARIMAX model. Subsequently, to conservatively validate the Prophet–SARIMAX model producing robust results, the projections of the two models are compared with those announced by prestigious institutions such as IEA or BP Energy [149].

4.5.2.2 Comparison 2

The hybrid LIM–SARIMAX model employs key drivers projected by LIM instead of Prophet as inputs to ultimately project Chinese NG consumption via SARIMAX. The reason for comparing these models is to validate the uncertainty in projections by the Prophet–SARIMAX model. For this purpose, a conservative estimation is adopted to compare the error bands of the two hybrid models, along with the magnitude of their difference: the error band of LIM–SARIMAX ("Error band 1") and the error band of Prophet–SARIMAX ("Error band 2"). Details in the development of error bands for each model are as follows:

<u>Error band 1: LIM–SARIMAX</u>

Error band 1 establishes lower and upper bounds for each interpolated key driver. These bounds are determined using the 95% confidence intervals derived from the linear regression slope under the three scenarios—STEPS, APS, and NZE. These scenarios align with those utilised in the Prophet-based projections of key drivers. It is important to note that since this method relies on linear regression, the resulting intervals predominantly capture the trend and may not fully account for seasonality or other fluctuations. Finally, the bounds obtained from the confidence intervals are utilised as inputs in SARIMAX to generate the error band. The detailed LIM elaboration—modelling and inputs for projection—is provided in Table 4.12 (see Appendix N).

<u>Error band 2: Prophet–SARIMAX</u>

Error band 2 is created based on the uncertainty intervals (lower and upper bounds) for each key driver generated by Prophet at a 95% confidence level. The uncertainty intervals in the Prophet are estimated within a Bayesian framework through Markov Chain Monte Carlo (MCMC) sampling, preventing overconfidence in prediction intervals [193]. In other words, Prophet appropriately considers uncertainty and does not excessively expand prediction intervals.

Typically, when a model makes predictions, it sets prediction intervals by considering uncertainty. If these intervals are too narrow or too wide, they may fail to reflect the genuine uncertainty of accurate predictions. The Prophet model uses a Bayesian framework and MCMC sampling to prevent such situations, ensuring that the model does not set prediction intervals excessively narrowly or widely [193]. Specifically, in the Bayesian framework, the prior information combined with data yields the posterior distribution of the parameters. The Monte Carlo aspect of MCMC involves drawing samples from the Bayesian posterior distribution. The Markov Chain process of MCMC converges the posterior distribution to the target distribution on the parameters through the sampling. This mechanism enables Prophet to determine the mean and confidence interval of the parameters used in the prediction.

In a nutshell, Prophet in the Prophet–SARIMAX model initially projects intervals for key drivers that reflect the underlying uncertainty. Subsequently, SARIMAX uses these intervals as inputs to generate Error band 2, the confidence interval for Chinese NG consumption.

4.6 Results

For the three scenarios—STEPS, APS, and NZE, the hybrid Prophet–SARIMAX model initially projects six key drivers—Coal price, PNG imports, LNG imports, Temperature, HH gas price, and Price distortion—using Prophet. Two other NG infrastructure-related factors, Gas pipeline capacity and Accumulated LNG terminal capacity, are projected incrementally without using Prophet. The graphs illustrating the projected key drivers are presented in Figure 4.6 (see Appendix P). Subsequently, SARIMAX in the Prophet–SARIMAX model identifies its optimal order (0,0,0)(0,1,1,12), and the Ljung-Box test result for the SARIMAX indicates its good fit in Table 4.13 (see Appendix O). The SARIMAX, with the optimal order, generates scenario-specific monthly projections of Chinese NG consumption using the projected key drivers as inputs. The monthly projections are depicted in Figure 4.7 (see Appendix Q).

Figures 4.6 and 4.7 illustrate that the long-term projections of drivers and NG consumption, as predicted by Prophet–SARIMAX, differ by scenario. However, all these projections are based on the assumptions of an emissions peak around 2030 and net-zero emissions by 2060. To ensure the robustness of the final result of Prophet–SARIMAX, the NG consumption projection, this research employs two conservative comparison methods—Comparison 1 and Comparison 2—in a thorough validation process.

4.6.1 Comparison 1

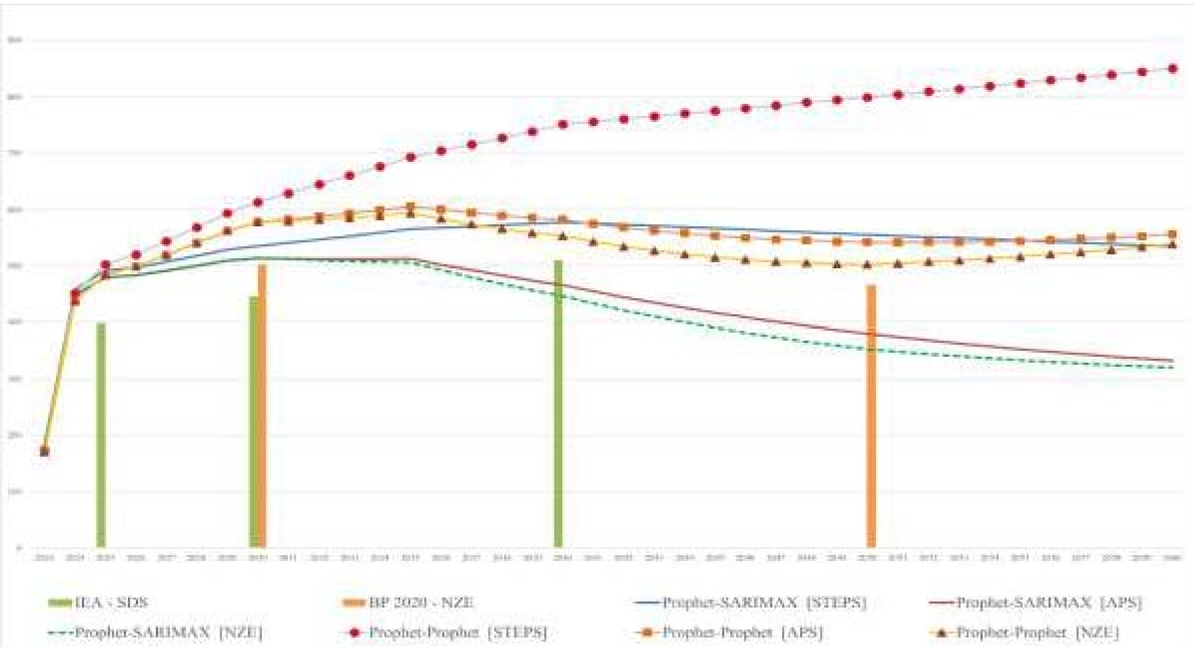


Figure 4.8 Chinese NG consumption projections (BCM)

Comparison 1 involves comparing Chinese NG consumption projections from two hybrid models—the hybrid Prophet–Prophet and Prophet–SARIMAX models—with those from IEA and BP Energy. The two hybrid models each generate monthly projections for Chinese NG consumption under the three scenarios: STEPS, APS, and NZE. The monthly projections are converted into yearly projections to facilitate a comparison between the two models. Then, these yearly projections are compared with those announced by prestigious institutions such as IEA or BP Energy, as shown in Figure 4.8. This comparative analysis aligns with one of the methodologies employed in the previous study, where its results for consumption of energy sources (CES) are compared with the CES projections presented by IEA and BP Energy [149].

The Chinese annual NG consumption projections of the two models for APS or NZE are compared with those recently estimated in the studies for IEA–Sustainable Development Scenario (SDS) and BP 2020–Net-Zero Emission (NZE) [183]. The IEA–SDS represents a 'well below 2 °C pathway', positioning it between APS and NZE [194]. The BP 2020–NZE aims to limit global warming to 1.5 °C, similar to the NZE scenario in the three scenarios. In the comparison, the annual projections of the Prophet–Prophet model do not fall within the range of IEA–SDS and BP 2020–NZE. However, the annual projections of the Prophet–SARIMAX model for APS or NZE are close to those of IEA–SDS and BP 2020–NZE in 2030 and lower than those in 2040 and 2050, thereby falling within the range. Therefore, it is reasonable to contend that the Prophet–SARIMAX model produces robust results.

4.6.2 Comparison 2

Comparison 2 involves comparing uncertainty estimates generated by two hybrid models—the hybrid LIM–SARIMAX and Prophet–SARIMAX models—for Chinese NG consumption projections under the three scenarios. These uncertainty estimates are represented by both the error band of the model and the magnitude of the difference. The uncertainty estimates, initially generated from the two hybrid models, are on a monthly basis. To facilitate the comparison of the two models, monthly uncertainty estimates are converted to annual uncertainty estimates. From a conservative estimation, the uncertainty estimates for each model are compared based on the error band and the magnitude of the difference. The two models' monthly and annual error bands are provided in Figures 4.7 and 4.9, respectively (see Appendices Q and R for further details).

The comparison of the error bands demonstrates that Prophet–SARIMAX can mitigate the uncertainty in projections. As depicted in Figure 4.9 (see Appendix R), annual error bands (highlighted in light pink) represent the error bands of the Prophet–SARIMAX model, while annual error bands (highlighted in green) signify the error band of the LIM–SARIMAX model. The error bands of LIM–SARIMAX appear significantly wider across all scenarios compared to those of Prophet-SARMAX. In fact, LIM–SARIMAX excludes seasonality and other fluctuations of inputs due to its reliance on the linear interpolation method. Nevertheless, the error bands of LIM–SARIMAX exhibit greater width than those of Prophet–SARIMAX, which reflects more seasonality or fluctuations. This indicates that Prophet–SARIMAX, by showing a narrower range of error bands, can minimize uncertainty in projections.

Comparing the magnitude of the difference in the outputs of the two models reveals a substantial difference between them. As illustrated in Figure 4.9 (see Appendix R), the difference between annual Prophet–SARIMAX outputs (highlighted in orange) and annual LIM–SARIMAX outputs (highlighted in blue) is clearly evident. This difference is further highlighted in Figure 4.10 (see Appendix S) through the magnitude of the difference between the two models. To be specific, the magnitude of the difference between the two models is substantial at certain points, reaching at least 20 BCM in all scenarios. This disparity clearly illustrates that the Prophet–SARIMAX model differs significantly from the LIM–SARIMAX model, which reflects only linear trends of the data.

In brief, Prophet–SARIMAX, the district from the model that simply captures linear trends in the data, can reduce uncertainty in projected outputs. Besides, unlike projections of other prestigious institutions (IEA, BP Energy, etc.) that only present particular periods—2030 and 2050—in China, the projections by Prophet–SARIMAX are quantified every month until 2060. These findings imply that the predictions using Prophet–SARIMAX can be beneficial for formulating relevant future policies or business plans.

4.6.3 Accelerated Net Zero Emission (ANZE) pathways

This study explores the pathways of key drivers to achieve zero NG consumption in China using the hybrid Prophet–SARIMAX model. As indicated in Table 4.14, projections of various reports for Chinese NG consumption do not anticipate zero NG consumption around 2050 or 2060. Additionally, the projections by Prophet–SARIMAX using a recent NZE scenario in this study also do not reach a zero by 2060 (highlighted in orange in Table 4.14), despite Chinese NG consumption projections by Prophet–SARIMAX lower than those in the reports.

Table 4.14 Chinese NG consumption projections by 2060

Items	Scenarios	Projection of NG consumption (BCM)			
		By 2030	By 2040	By 2050	By 2060
BP [195]	Net-zero emission	502	-	467	-
IEA [194]	Sustainable development	446	511	-	-
Tsinghua University [196]	1.5 ℃	580	338	150	-
World Resources Institute [197]	Intensified actions	-	-	460	-
State Grid Energy Research Institute [198]	Accelerated electrification	-	600	440	410
Global Energy Interconnection Development and Cooperation Organization [199]	Carbon neutrality	500 (by 2035)	-	-	-
Zhou et al. [200]	Carbon neutrality	650 (by 2035)	-	550	430
Projections by Prophet–SARIMAX	NZE	573	437	252	187
Projections by Prophet–SARIMAX	ANZE	430	105	7	0

Note: These results are derived from a recent study of Wang et al. [183].

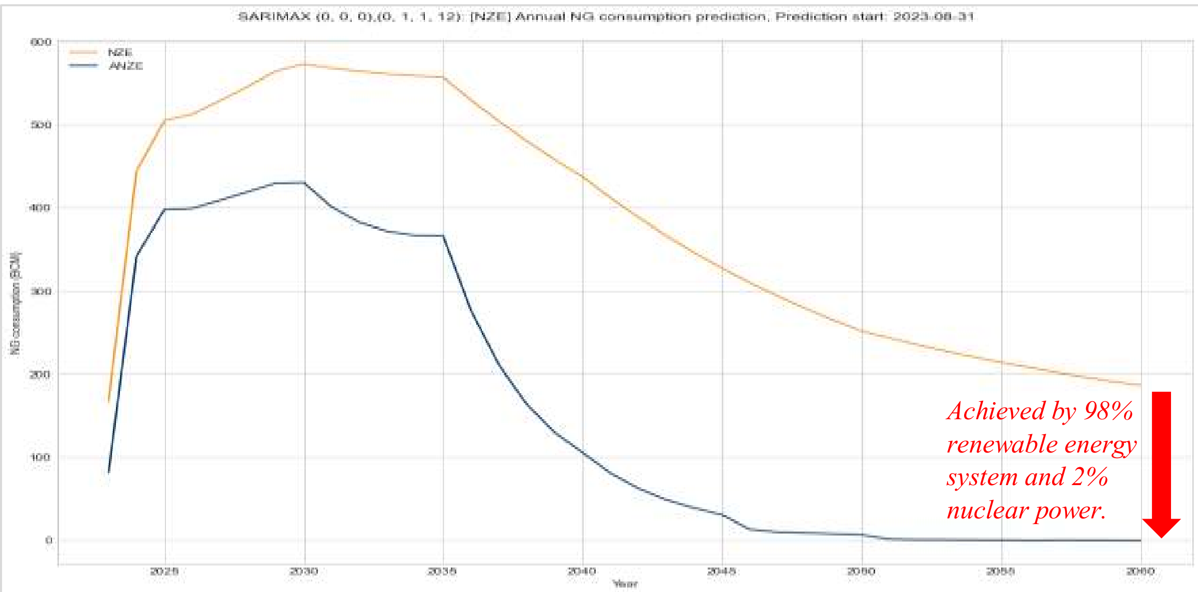


Figure 4.11 Annual NG consumption projection by Prophet–SARIMAX: NZE vs. ANZE

To illustrate the pathways of key drivers for the zero NG consumption projection, this study applies the OECM scenario (see Section 4.4.2) to the Prophet–SARIMAX model; this scenario suggests that China is anticipated to achieve a 98% renewable energy system, with a remaining 2% from nuclear power, by 2050. Applying the OECM scenario in this study is referred to as Accelerated Net Zero Emission (ANZE) pathways. As seen in Figure 4.11 and Table 4.14, Prophet–SARIMAX can generate China's zero NG consumption projection under the OECM scenario.[33] Moreover, the study can derive the pathways of the key drivers in China's NG consumption—Price distortion, Temperature, HH gas price, PNG imports, Thermal coal price, Gas pipeline capacity, Accumulated LNG terminal capacity, and LNG imports—based on the zero NG consumption projection.

As illustrated in Table 4.15, the ANZE pathways for key drivers are determined by incorporating fossil fuel demand projections or relevant CAGRs of the OECM scenario into these driver projections up to 2060. To be specific, the ANZE pathways are as follows:

Table 4.15 ANZE pathways of key drivers

Key Drivers			**ANZE pathways**			
			2025	**2030**	**2040**	**2050/60**
Coal price		Price (USD/MT)	-	31	-	0
		CAGR (%)	-	-20.00	-	-24.99
HH gas price		Price (USD/MMBtu)	-	1.0	-	0
		CAGR (%)	-	-18.43	-	-10.87
Price distortion	PNG price	Price (USD/MMBtu)	-	3.27	2.94	0
		CAGR (%)	-	-10.00	-1.07	-43.55
	CGGP	Brent oil (USD/barrel)	-	13	-	0
		CAGR (%)	-	-22.69	-	-21.45
LNG imports		Imports (BCM)	-	102	-	0
		CAGR (%)	-	2.00	-	-28.18
PNG imports		Imports (BCM)	60	66 (70 by 2035)	58 (22 by 2045)	0
		CAGR (%)	4.77	2.00 (1.10 by 2035)	-3.66 (-17.90 by 2045)	-86.00
Temperature			Adhering to Temperature's NZE scenario (in Table 4.9)			
Accumulated LNG terminal capacity			Reflecting LNG imports' ANZE scenario			
Gas pipeline capacity			Reflecting PNG imports' ANZE scenario			

Note: For LNG and PNG imports, the MMT to BCM conversion rates from Section 4.4.1—BP Energy Scenarios—are applied.

Coal and HH gas prices

According to the OECM scenario, the coal and HH gas prices are projected to drop to zero by 2050 and remain at that level through 2060. To align with the projected zero prices in 2050, it is assumed that the prices in 2030 will be set at half of the prices predicted under the NZE scenario for these drivers. This results in the coal price at 31 USD/MT and the HH gas price at 1 USD/MMBtu. The CAGRs are calculated based on the projections for 2030 and 2050, and these CAGRs are used as inputs for the Prophet-based projections of these driver prices.

Price distortion

As mentioned in Section 4.4.1.1, IEA scenarios are utilised to calculate price distortion instead of directly affecting it. Therefore, the OECM scenario is also applied to PNG price and

[33] The projected outputs of "Temperature" and "NG-related infrastructure" factors do not reach complete zero. These outputs generate the constant small values in the Chinese NG projections from 2050 onwards. Consequently, manually adjusting the final constant small values from the NG projections enables the model to achieve zero NG consumption.

CGGP, both of which are essential for calculating price distortion. The PNG price in the ANZE scenario remains the same as its NZE scenario until 2030, reaches the NZE scenario's 2050 price level by 2040, and is projected to reach zero by 2050, continuing at that level through 2060. For CGGP, which is calculated based on Brent oil and HH gas prices [36], it is assumed that the HH gas price in CGGP follows the ANZE pathways mentioned earlier. Similarly, the Brent oil price is projected to reach zero by 2050, akin to other fossil fuel prices in the ANZE pathways. To align with this projection of reaching zero by 2050, it is assumed that by 2030, the Brent oil price is expected to achieve half of the IEA crude oil's 2050 price level under its NZE scenario. The CAGRs are calculated for projections of PNG price and CGGP's Brent oil price for each period and then these CAGRs are used for price distortion projections generated by Prophet.

<u>LNG/PNG imports and Temperature</u>

According to the OECM scenario, which sets a 2% increase in NG demand by 2030 and a subsequent decrease by -28.18% by 2060, LNG imports are projected to reach 102 BCM by 2030 and decrease to zero by 2060. PNG imports are projected to decline to zero by 2050, following the CAGR calculations based on the OECM scenario's NG demand projections from 2025 to 2050 in five-year increments. The CAGRs for LNG and PNG import projections are utilised as inputs in the Prophet-based projection.

The final driver in the Prophet-based projection, Temperature, is applied according to its NZE scenario in Table 4.9, not the OECM scenario. Since the OECM scenario aims to adhere to the Net-zero 1.5 °C target, the temperature projection in the ANZE scenario follows the NZE scenario.

<u>Accumulated LNG terminal and Gas pipeline capacity</u>

Unlike other driver projections, Accumulated LNG terminal capacity and Gas pipeline capacity are projected step by step without utilising the Prophet model. The accumulated LNG terminal capacity is anticipated to increase according to its scenario in Table 4.7 until 2035. However, it is expected to decrease to 1,720 kt by 2050 and further to 60 kt by 2060, following the -28.18% CAGR in LNG import projections in the ANZE scenario from 2050 to 2060. The gas pipeline capacity is assumed to increase until 2030, similar to its BP Energy scenarios in Table 4.8, and then decrease in 2040, 2045, and 2050/60, following the CAGRs of the PNG import projections under the ANZE scenario. As a result, it is anticipated to decrease to 134, 50, and 0 BCM in 2040, 2045, and 2050/60, respectively.

In the end, even in the ANZE scenario, we use both Prophet-based projections of the six key drivers—Coal price, HH gas price, Price distortion, LNG imports, PNG imports, and Temperature—and the stepwise projections of two NG-related infrastructure drivers—Accumulated LNG terminal and Gas pipeline capacity—as inputs for the SARIMAX model in the Prophet–SARIMAX model. This demonstrates that Chinese NG consumption achieves zero in the ANZE scenario, which is based on the OECM scenario with a target of 98% renewable and 2% nuclear energy technologies. Here, an ambitious NG phase-out pathway can be developed through the technologies. As seen in the case of Denmark, which supplies 49% of its total energy consumption and 89% of its electricity from renewable sources in 2023 [73] and aims to phase out NG entirely by 2050 by significantly increasing these sources [201], this pathway could be a new feasible guideline towards China's carbon neutrality.

4.7 Conclusions

This chapter explores pathways for China's NG consumption through a novel hybrid scenario-based projection model sequentially combining a machine learning-based nonlinear model (Prophet) and a traditional linear forecasting model (SARIMAX). For scenario-specific projections of Chinese NG consumption, the hybrid Prophet–SARIMAX model initially uses Prophet to project the selected key drivers for China's NG consumption by 2060 under the three scenarios and subsequently employs SARIMAX to project China's NG consumption using the projected key drivers as inputs. These drivers are identified by the hybrid model combining machine learning (LASSO/ALASSO) and econometric model (ECM) [107]. The three scenarios are drawn from the IEA scenarios [152]—STEPS, APS, and NZE—serving as a reference point, alongside the BP Energy [153] and RCP scenarios [154]. This study adopts the same scenario names as those defined by the IEA.

Furthermore, this study indicates that the phasing out of NG in China by 2060 under the country's carbon neutrality target is unlikely. Specifically, the Prophet–SARIMAX projections based on the existing Net Zero scenario do not reach zero NG consumption by 2060, indicating that methane leakage from NG can undermine China's effort of reaching its genuine carbon neutrality without additional interventions. In response, the study derives the ANZE pathways of key drivers leading to zero Chinese NG consumption by applying the OECM scenario. According to the ANZE pathways derived from the OECM scenario, which anticipates China achieving a 98% renewable energy system with the remaining 2% from nuclear power by 2050, the key drivers—coal/gas prices, price distortions, and LNG/PNG imports—must be driven to decline even more significantly than in the NZE scenario to achieve zero NG consumption. The ANZE pathways demonstrate that China must shift away from the framework of existing scenarios in the country's NG market to achieve genuine carbon neutrality and adopt a new scenario perspective. The key driver pathways for the NG market proposed within the OECM scenario can serve as a new horizon for carbon-neutral policies.

This study also proves that the hybrid Prophet–SARIMAX model is highly suitable for scenario-specific projections, as it can reproduce the results of the three published scenarios from IEA and BP Energy with high accuracy. Integrating a nonlinear model (Prophet) and a linear model (SARIMAX) has demonstrated superior performance in long-term forecasting, employing two conservative estimation methods: Comparison 1 and 2. Specifically, in Comparison 1, where the projections from two hybrid models (Prophet-Prophet and Prophet–SARIMAX) are compared with those from IEA or BP Energy, Prophet–SARIMAX is validated to generate robust projections. In Comparison 2, where the error bands of two hybrid models (LIM–SARIMAX and Prophet–SARIMAX) are compared, Prophet–SARIMAX can mitigate uncertainty in the scenario-specific projections. Proven effective for scenario-specific long-term projections, Prophet–SARIMAX can be a valuable tool for government agencies or corporate executives making decisions regarding long-term policies or business plans.

In this vein, three policy implications for achieving carbon neutrality by 2060 could be derived from the ANZE pathways identified through Prophet–SARIMAX. First, a more robust carbon neutrality policy juxtaposing a carbon tax and renewables stimulus is imperative. The decrease in fossil fuel prices and imports in the ANZE pathways is much higher than the NZE scenario. To realise this, a more forceful carbon neutrality policy should be implemented, one that concurrently imposes both a carbon tax and investment promotion for renewables. A carbon tax could increase the cost of using fossil fuels, thereby inducing both businesses and consumers to switch to lower-carbon energy sources [202]. At the same time, investments in renewables could be stimulated by utilising the carbon tax. This enables renewables to be competitive and complement the diminished use of fossil fuels led by the carbon tax.

Second, China's NG market liberalisation reform should be accelerated to minimise price distortions. Compared to the NZE scenario, the ANZE pathways show a more significant decline in PNG, Brent oil, and HH gas prices, which are incorporated into price distortion calculations. Hence, to achieve carbon neutrality, China needs to reduce price distortions caused by government subsidies that artificially lower natural gas prices [203]. In other words, the NG market should rapidly adopt a market-based pricing system to enhance price transparency. This could strengthen China's collaboration with the global market [76], which reflects environmental costs, such as the European emissions tax on LNG imports [204]. These costs could penalise fossil fuels like NG [71], making them less competitive than renewables [78] and promoting investment in them [77]. Eventually, this could help to achieve the country's long-term carbon neutrality target, aiming to reduce the use of fossil fuels [79].

Third, China's gas transmission infrastructure and relevant regulatory framework should be upgraded in line with the country's hydrogen energy development. The ANZE pathways indicate a significant decrease in LNG/PNG imports from 2030 to 2060, rendering the existing gas transmission infrastructure stranded in the future. Meanwhile, hydrogen has garnered significant attention from industries and policymakers in China as a critical energy source towards the country's carbon neutrality [205]. Utilising existing gas pipelines for hydrogen transportation is considered a practical, effective solution [206]. However, this poses significant challenges. While hydrogen transportation may be tolerated in existing low-pressure networks [206], high-pressure facilities are necessary for its efficient distribution [207]. Moreover, more comprehensive national policies and regulations for hydrogen storage and transportation are required [208].

One limitation of this study is that its analysis relies on the OECM scenario to investigate potential pathways towards zero NG demand for achieving carbon neutrality in China, which may not directly align with current policies. The country's existing carbon-neutral policies do not explicitly specify zero NG consumption by 2060 [27]. In other words, China could have various specific pathways of NG consumption for its 2060 carbon neutrality. Through the OECM scenario, the study has proposed conceptual NG phase-out pathways for China to achieve its genuine carbon neutrality. Future research could evaluate the practical feasibility of achieving zero NG consumption, considering present or future economic and technological constraints.

The other limitations of the study include data constraints and scenario scope. Not all driver projections for 2060 are reflected in the IEA, BP, and RCP scenarios, requiring additional sources. While the drivers applied in these scenarios follow similar assumptions for carbon peak and carbon neutrality based on realistic policies, applying these assumptions to data from other sources may reduce reliability. Additionally, the scenarios are based on the premise of stable markets and policies, which may not account for unexpected events like energy crises or policy changes. Therefore, the scope of the scenarios may not adequately capture sudden fluctuations. Future research could integrate and cross-validate various data sources or adopt modelling techniques that incorporate uncertainties for more robust analysis.

CHAPTER 5.0

OPPORTUNITIES FOR AUSTRALIA AMID CHINA'S CARBON-NEUTRAL TRANSITION: INPUT–OUTPUT APPROACH WITH ANALYSIS OF AUSTRALIAN RENEWABLE POTENTIAL

This chapter is based on the manuscript: Sung, J. (2025). Opportunities for Australia amid China's Carbon-Neutral Transition: Input–Output Approach with Analysis of Australian Renewable Potential. The manuscript is currently under review at a peer-reviewed journal. Author contributions are detailed in the Declaration of Publications and Author Contributions section of this thesis.

Chapter outlook

Chapter 4 forecasts that the complete phase-out of natural gas will be challenging under China's 2060 carbon neutrality scenario and proposes alternative pathways to achieve this goal. Despite this challenge, a tsunami of changes in China's energy mix is already underway as the country strives to achieve carbon neutrality by 2060. Amid these changes, Australia—as a key supplier of fossil fuel and iron ore to China—faces emerging challenges that could bring significant structural changes to its economy. To address these challenges, this chapter conducts a comprehensive economic impact analysis of energy flows between Australia and China, utilising static and dynamic input–output analyses based on the Environmentally eXended multi-regional Input–Output dataBASE (EXIOBSE). Additionally, this chapter integrates two competing economic stimulus scenarios—one focused on natural gas and another on renewables—into the dynamic input–output model to explore potential policy responses to these challenges. Key findings reveal that China's carbon-neutral transition negatively impacts Australia's major resource and mineral export industries, triggering spillover effects across other sectors. Furthermore, the findings indicate that Australia's transition to a renewable-based energy structure could foster economic and environmental sustainability by reorienting these challenges into new growth drivers. Finally, the chapter highlights the advantages of the EXIOBASE-based modular dynamic input–output model, particularly its capacity for granular sectoral and temporal analysis, and demonstrates the model's potential for customisation across various policy planning domains.

Keywords: Carbon-Neutral Transition; Renewable Energy; Static and Dynamic Input–Output Analysis; EXIOBASE; Australia; China

5.1 Introduction

A tsunami of changes in China's energy mix is approaching with its goal of achieving carbon neutrality by 2060, and Australia must prepare a breakwater in response to this transformation. In China's carbon-neutral transition, Natural Gas (NG) is emerging as a crucial energy source to achieve its Dual Carbon Goals (DCGs): peaking carbon emissions by 2030 and reaching carbon neutrality by 2060. While NG plays a key role in replacing coal as a major source of carbon emissions in the energy mix to help China peak its carbon emissions by 2030 [149], NG, as a fossil fuel, is destined to be phased out by 2060 [4]. Amid this shift in the role of NG, Australia's position as a leading LNG supplier to China is projected to remain relatively stable through 2030 [209]. Meanwhile, the ultimate shift of phasing out fossil fuels, including NG, by 2060 underscores the need for Australia to diversify and future-proof its energy exports in anticipation of the global transition towards low-carbon alternatives. Against this backdrop, this study aims to analyse how China's decarbonisation transition—specifically focusing on shifts in China's energy mix driven by its NG policy—will impact the Australian economy. Furthermore, this analysis seeks to propose policy initiatives for Australia to mitigate these impacts and transform the challenges of China's DCGs into new growth opportunities.

In this study, a dynamic and flexible economic analysis tool is developed to facilitate the analysis effectively. Input–Output (IO) analysis and Computable General Equilibrium (CGE) models have been widely used for economic impact analysis. IO analysis uses an inter-industry transaction table—representing relationships between various economic agents such as industries, households, and the government at a given point in time [210]—to evaluate the impact of various hypothetical scenarios [211]. CGE models, which build on the IO framework, allow for comprehensive economy-wide analysis and can incorporate assumptions about technological change, making them particularly useful for the broader social impacts of policy proposals [212]. However, the Leontief IO analysis is inherently static, as it does not account for dynamic changes over time [213]. CGE models, while offering comprehensive analysis, can not only be highly complex [210] but also reduce flexibility due to their reliance on specialised software like the Global Trade Analysis Project (GTAP) [214]. To address these limitations, this study proposes an economic analysis tool that expands the IO framework to enable dynamic analysis at a more granular sectoral level, with the flexibility to adjust the level of detail according to the research objectives.

This study utilises a customised modular dynamic tool developed in Python with the multi-regional and granular sectoral IO table to analyse the economic impact of China's decarbonisation goals on the Australian economy. The study also proposes Australia's potential policy initiatives to alleviate the economic impact through the modular dynamic tool. For the economic impact analysis, the study uses China's NG projection scenarios developed by the Prophet-Seasonal AutoRegressive Integrated Moving Average with eXogenous regressors (SARIMAX) model [215] and its coal and oil demand scenarios based on the International Energy Agency (IEA) [152]. The Leontief IO analysis, a static IO model, initially examines snapshot analyses of economic shocks caused by these scenarios. Subsequently, a dynamic IO model, incorporating a policy framework developed in a modular approach, analyses these economic shocks dynamically and proposes policy initiatives to address the shocks. Additionally, the results of this static model are used to verify the robustness of the dynamic model.

This study presents two main contributions. First, it proposes a modular model that integrates the economic stimulus framework with the dynamic IO model, allowing for more comprehensive analyses. Specifically, this modular approach shows the potential to customise and extend the model for various economic and policy analyses or research objectives in the future. Second, this study contributes to a more comprehensive assessment of energy flows between Australia and China by employing time-based IO analyses of more granular sectors - beyond the level of detail provided by CGE models - through the Environmentally eXtended multi-regional Input–Output dataBASE (EXIOBASE) [216]. These approaches explore new opportunities for the potential development of Australia's renewable energy sector, offering a more nuanced understanding of the economic dynamics between the two countries.

The rest of this chapter is organised as follows: Section 5.2 provides related literature pertaining to Australia's challenges and opportunities amid China's carbon-neutral transition, as well as static and dynamic IO analyses. Sections 5.3 and 5.4 discuss data and the methods and models for IO modelling, respectively. Section 5.5 presents the models' results and relevant discussion, and the final section concludes the study. The supplementary materials include tables, and figures used throughout this chapter.

5.2. Literature review

This literature review not only outlines the economic implications of China's carbon-neutral transition on Australia, but also establish the validity of the proposed methodological framework. Specifically, it first investigates the economic challenges and opportunities that China's transition towards carbon neutrality presents for Australia. Second, it critically reviews existing analytical approaches to assessing such challenges and opportunities. In this way, the review defines the scope and direction of the study while highlighting the strengths and contributions of the proposed research framework.

5.2.1 Australia's challenges and opportunities amid China's carbon-neutral transition

China's transition towards carbon neutrality by 2060 necessitates a fundamental transformation of its industrial and energy sectors, which is expected to have significant implications for the Australian economy. As China is the largest importer of Australia's iron ore, coal, and LNG, changes in China's energy policy directly affect Australia's export structure [217]. Especially, because of the substantial contribution of these resource exports to Australia's GDP and employment, its regions heavily dependent on the mining and energy sectors are vulnerable to economic disruption. Consequently, growing concerns exist that its regional economic disparities could be exacerbated [217].

China's carbon-neutral transition, which is expected to reduce the demand for Australia's carbon-intensive commodities, drives Australia to redefine its economic model and export strategies. Huang et al. suggest that China's carbon-neutral transition urges Australia to shift from carbon-intensive to low-carbon industries, implying that Australia should explore new growth opportunities through this transition [218]. In particular, Song suggests that the decarbonisation of China's steel industry is crucial for achieving global carbon goals, and that its decarbonisation can be realised through cooperation with Australian renewable electricity and hydrogen [219].

In brief, China's push for carbon neutrality is anticipated to significantly impact the Australian economy. This poses challenges for Australia, necessitating the shift to transition from a fossil fuel-intensive economic model to one focused on green technologies and sustainability. Furthermore, this shift offers Australia's opportunities to turn structural vulnerabilities into new economic growth drivers. This study conducts scenario-based quantitative economic impact analyses with a focus on these challenges and opportunities.

5.2.2 Economic impact analysis for the Sino-Australian carbon relationship

Recent studies have widely utilised IO and CGE models in the economic impact analysis of the Sino-Australian carbon relationship. In particular, these models provide valuable insights into the external effects triggered by China's transition to carbon neutrality. Thus, it is necessary to systematically review them to gain a comprehensive understanding of their applicability, especially in economic analyses related to policy evaluation.

IO analysis, created by Wassily Leontief, examines how outputs from one economic sector serve as inputs for another [211]. This framework has been widely employed to evaluate inter-country trade transactions and carbon emissions in the context of China's carbon emissions reduction. Zhao et al. utilise a non-competitive IO table to analyse how different carbon tax rates in China affect the economy and emission reductions [220]. Their analysis demonstrates how carbon taxes influence production and income flows, especially highlighting potential changes in export relations involving carbon-intensive industries, such as coal and metals. Wang et al. show that a significant share of China's CO_2 emissions stems from export production with considerable emissions embodied in China-Australia trade through the simulation based on IO analysis and panel regression models [221]. Their findings highlight the impact of China's DCGs on both domestic emissions and Australia's resource-export industries, particularly coal and iron ore. Wei et al. use multi-regional IO models to underscore the contribution of international trade to carbon emissions, suggesting that technological improvements are the most viable way of mitigating China's carbon emissions [222].

However, traditional IO models are limited by their static nature, specifically the assumption that inter-sectoral relationships remain constant. This static characteristic fails to account for dynamic factors, such as policy developments and changes in consumer demand [223]. Moreover, the static nature—fixed ratios for intermediate inputs and production—can result in inadequate responses to new policies (e.g., carbon taxes or emissions regulations), as they do not dynamically adjust to a changing economic landscape [224]. To address these limitations, studies need to apply advanced IO modelling techniques integrating dynamic factors [225].

In a few recent studies, the CGE model serves as an effective tool for analysing the ripple effects of China's carbon-neutral transition, both within China and on Australia. Pan et al. suggest that China's DCGs promote green technology and consumption to drive its high-quality economic growth, using the CGE framework [226]. Using a single-country, multi-sectoral CGE model, Wei et al. indicate that China's DCGs have a negative impact on agriculture and highlight that advancing agricultural technology—particularly through a 1% annual productivity increase in key crops—is essential to offset the losses [227]. Peng et al.

point out that China's carbon reduction policies are expected to cause significant structural impacts on Australia's mining sector and fossil fuel-dependent regions, while its overall economic impact remains limited [217]. This aligns with the Reserve Bank of Australia (RBA)'s analysis on the impact of China, Japan, and South Korea's net-zero carbon emission goals [228].

However, CGE models have limitations in quickly and flexibly responding to carbon-neutral policies under rapidly changing economic conditions due to their complex structure, simulation inefficiencies, and low industry disaggregation. First, CGE models, which require advanced software and computational resources [214], involve substantial effort to modify their underlying equations due to their complex structures, limiting their ability to quickly adapt to new policy measures in dynamic market environments [210]. Moreover, CGE models, which demand multiple time-consuming stages such as modelling strategy formulation, code modification, sensitivity analysis, and other tasks [229], often result in lengthy simulations [230]. While these simulations can be valuable for detailed analysis, they may be less practical in time-sensitive policy contexts where quick scenario evaluations are essential [230]. Lastly, the GTAP, the latest version of the representative CGE model, includes only 65 industries [231], which is relatively low compared to the 163 industries in EXIOBASE [232]. This can prevent the GTAP from fully capturing policy effects at a detailed industry level, especially in countries with complex industrial structures like Australia.

As outlined thus far, China's push for carbon neutrality could significantly impact Australia, presenting challenges for economic structural transformation and opportunities for new economic growth. While providing valuable frameworks for assessing such economic implications, IO and CGE models are limited by their static characteristics and analytical complexity, respectively, which include lengthy simulations and low industry disaggregation.

To overcome these limitations, this study develops a dynamic IO model using EXIOBASE, which is disaggregated into 163 industries. This model not only reflects temporal dynamics but also enables a more granular sectoral analysis of changes in the Australian economic structure. Additionally, by applying a Python-based modular modelling approach, the model allows for easy incorporation of additional impact analyses, policy measures, or scenario changes. Consequently, this study enhances both the sectoral analytical capabilities and the realism of simulations for analysing the impact of China's carbon neutrality on Australia.

5.3 Data

The data employed in this research are derived from three sources: EXIOBASE version 3 (the third version of EXIOBASE, hereafter EXIOBASE 3), the projected scenario of China's NG consumption developed by the Prophet–SARIMAX model (hereafter Prophet–SARIMAX) [215], and the IEA scenarios for China's oil and coal demand. The IEA scenarios consist of three key pathways: Stated Policies Scenario (STEPS), Announced Pledges Scenario (APS), and Net Zero Emissions by 2050 Scenario (NZE) [152]. This study assumes that the IEA's NZE scenario—designed to limit global warming to 1.5°C [152]—most closely aligns with China's carbon neutrality pathway and is conducted based on this scenario. Below are the details of EXIOBASE 3 and the scenario.

5.3.1 Environmentally eXtended multi-regional Input–Output dataBASE

This study utilises EXIOBASE 3, a comprehensive multi-regional environmentally extended Supply and Use (SU) / Input–Output (IO) database, based on the 2022 reference year [216]. EXIOBASE 3 integrates detailed sectoral data of 44 countries[34] from the national SU and IO tables, enabling the estimation of international trade by commodity type or industry. Additionally, it allows for assessing emissions and resource extractions by industry. This database covers 28 European Union (EU) countries and 16 main economies—which together account for approximately 90% of the world's GDP—plus five Rest-of-World (RoW) regions (see Figure 5.1 and Table 5.1 in Appendix T), along with 163 industrial sectors by region (see Figure 5.2 and Table 5.2 in Appendix T). The tables for the list of countries and sectors are provided in Appendix T—List of regions and sectors.

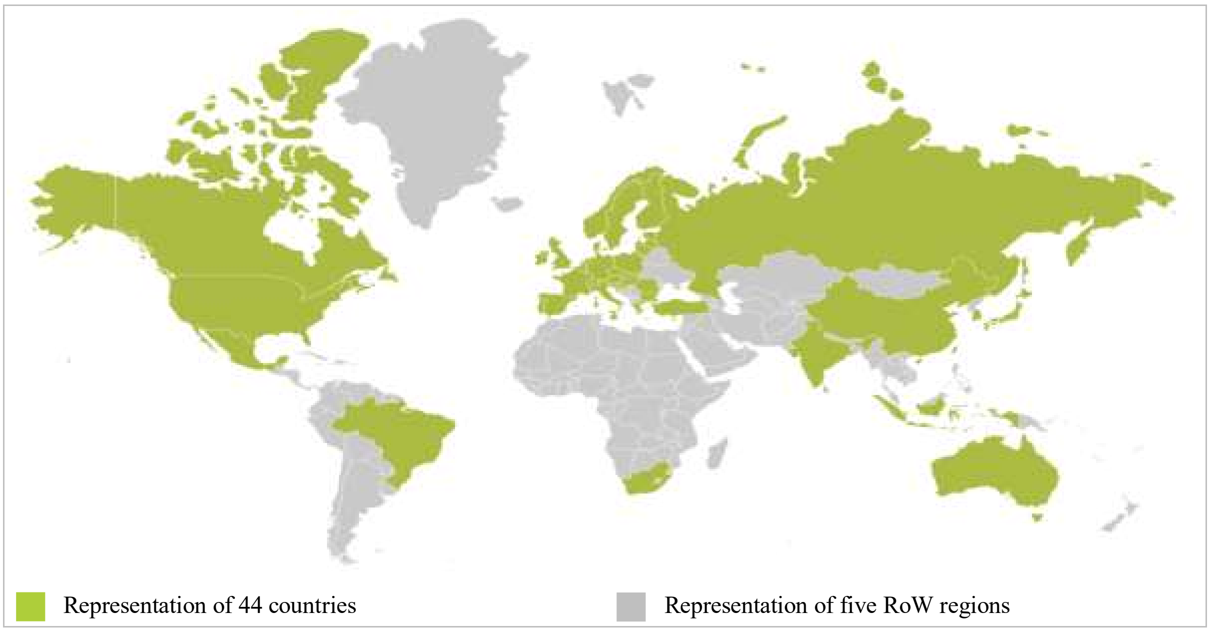


Figure 5.1 Map of 44 countries and five RoW regions in EXIOBASE 3

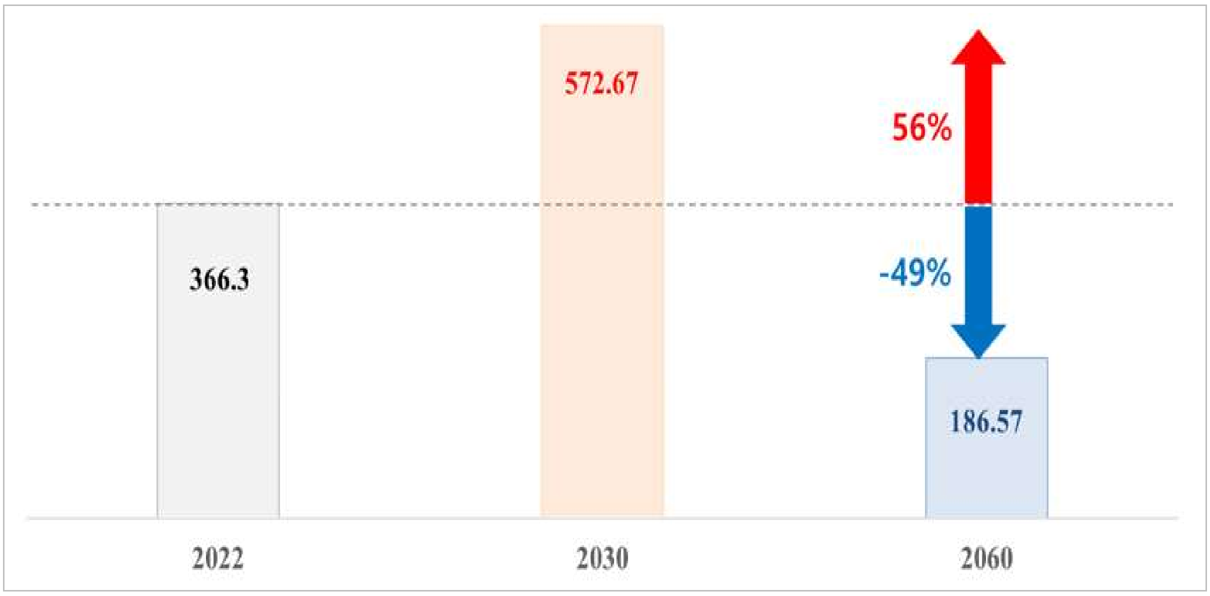


Figure 5.2 2030 and 2060 NG demand projections (BCM)

[34] The list of 44 countries does not include Qatar, which imports LNG second to Australia from China [2], and Qatar is categorised under the ROW regions. Since this study focuses on the relationship between Australia and China, it is assumed that the presence of Qatar data, even if it exists separately, does not significantly impact the IO analysis in the study.

5.3.2 Projected scenario of China's natural gas consumption

The projected scenario of China's NG consumption, used to analyse the economic impact of China's carbon neutrality target on Australia, are derived from Prophet–SARIMAX [35] [215]. This scenario presents the projected changes in NG consumption for the years 2030 and 2060, compared to the 2022 baseline. The NG consumption is expected to increase by approximately 56% by 2030 and then decrease by around 49% by 2060. This reflects China's carbon neutrality policy, which aims to reduce coal consumption by 2030—leading to an increase in NG use as a transitional energy source—and subsequently reduce overall fossil fuel consumption by 2060, resulting in a decline in NG demand [3]. The scenario is visually presented in Figure 5.2.

5.3.3 International Energy Agency scenarios for oil and coal demand

Apart from the NG consumption scenario from Prophet–SARIMAX, the study also considers China's coal and oil demand projections based on the assumptions of the IEA's NZE scenario for the economic impact analysis [152]. The IEA provides projections for China's coal and oil demand under the STEPS and APS, but not under the NZE scenario. Hence, China's coal and oil demand under the NZE scenario is estimated based on the world coal and oil demand projections in the IEA's NZE scenario. As shown in Table 5.3, the ratio of China's coal and oil demand to the world demand remains relatively consistent under both the STEPS and APS. For coal demand, China accounts for approximately 58% of the world demand in 2030 and about 44% in 2050. For oil demand, a similar approach was applied: under the NZE scenario, China's share of world oil demand was assumed to be about 16% in 2030 and around 12% in 2050. Based on these assumptions, coal demand in the NZE scenario is projected to decrease by 43% in 2030 and by 93% in 2050. As for oil demand, it is expected to decline by 14% in 2030 and by 80% in 2050 under the NZE scenario. Despite the sharp decline in coal and oil demand by 2050 in line with carbon neutrality targets, a complete phase-out is not assumed. However, it is assumed that the remaining use of fossil fuels produces no emissions because of the extensive implementation of carbon capture, utilization, and storage (CCUS) technologies at industrial facilities and power plants [152].

Table 5.3 China's coal and oil demand under IEA's NZE scenario

Item	**STEPS**		**APS**		**NZE**	
	2030	**2050**	**2030**	**2050**	**2030**	**2050/60**
Coal demand (Mtce)						
World	5,007	3,465	4,337	1,530	3,257	499
China	2,878	1,563	2,530	672	1,889	220
China/World (%)	58%	45%	58%	44%	58%	44%
% change compared to 2022	-13%	-53%	-23%	-80%	**-43%**	**-93%**
Oil demand (mb/d)						
World	101.5	97.4	92.5	54.8	77.5	24.3
China	16.4	12.0	15.1	6.9	12.4	2.9
China/World (%)	16%	12%	16%	13%	16%	12%
% change compared to 2022	14%	-17%	5%	-52%	**-14%**	**-80%**

Notes: Mtce = million tonnes of coal equivalent; mb/d = million barrels per day; China's coal and oil demand in 2022, as reported by the IEA, is 3,300 Mtce and 14.4 mb/d, respectively.

[35] China's NG demand is projected by Prophet–SARIMAX using its key drivers based on the IEA's NZE scenario [215]. Since the IEA's NZE scenario and other published sources do not provide detailed information on NG demand by 2060, this study employs the NG demand projections generated by Prophet–SARIMAX [215].

The coal and oil demand projected under the NZE scenario by 2050 shows a greater reduction compared to the 2060 coal and oil demand projections under the STEPS and APS presented in the IEA's Chinese energy sector roadmap report for carbon neutrality [233]. Therefore, this study assumes that the 2050 coal and oil demand scenarios apply to 2060.

5.4. Methods and models

This study employs an IO model based on the IO table (EXIOBASE 3) to analyse the impact of China's NG consumption scenarios on the Australian economy. There are two main approaches to modelling with IO tables. The first is the static IO model, which analyses changes in the economic structure at a given point in time in response to external shocks. The second is the dynamic IO model, which incorporates temporal elements to capture structural changes in the economy over time.

Prior to explaining the modelling methods, it is essential to understand the components of the IO table. Accordingly, this study first presents the structure and content of the IO table, followed by a description of the modelling methods based on it.

5.4.1 Input–Output table

An IO table captures the flow of goods and services in an economy by illustrating inter-industry transactions in terms of inputs and outputs. These transactions are expressed in a matrix format showing how industries produce goods (rows) and how other industries or consumers use them (columns). Modelling the transactions within EXIOBASE 3, this study estimates the impact of changes in one country's industry on the demand or supply of various industries in other countries. Figure 5.3 shows a simplified IO table with four quadrants:

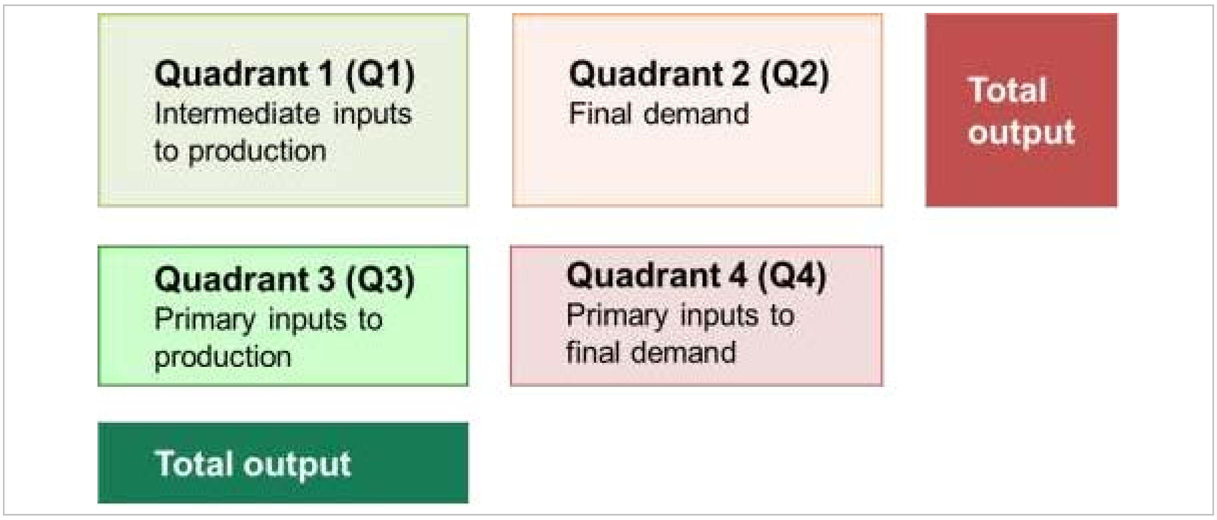


Figure 5.3 The simplest IO table represented by four quadrants

where:

- The core of the IO system is the symmetric industry-by-industry intermediate input matrix (Q1), which underpins IO modelling.
- The total output of an industry consists of intermediate goods (Q1) and final demand goods and services (Q2).
- The total output of an industry combines the intermediate goods and services (Q1) with primary inputs like labour and capital (Q3).

5.4.2 Static Input–Output model

The static IO model is designed to perform what-if scenario analysis. Specifically, it uses the Leontief inverse matrix to analyse the impact of external shocks—such as changes in demand or supply—within a specific industrial sector of one country on both the domestic economy and those of other countries. The Leontief inverse matrix is computed from the IO table [211]. In this study, it is derived from EXIOBASE 3 using Pymrio, a Python library, and the static IO modelling is implemented through tailored Python coding. The process of deriving the Leontief inverse matrix from the IO table is structured in three stages.

First, the inter-industry section of the IO table (Q1 in Figure 5.3) is transformed into a symmetric matrix:

$$\begin{array}{cc} & \textbf{Buying sector} \\ \begin{array}{c}\textbf{Selling}\\ \textbf{sector}\end{array} & \left[\begin{array}{c|ccccc} & 1 & \cdots & j & \cdots & n \\ \hline 1 & z_{11} & \cdots & z_{11} & \cdots & z_{11} \\ \vdots & \vdots & & \vdots & & \vdots \\ i & z_{i1} & \cdots & z_{i1} & \cdots & z_{i1} \\ \vdots & \vdots & & \vdots & & \vdots \\ n & z_{n1} & \cdots & z_{n1} & \cdots & z_{n1} \end{array}\right] \end{array}$$

This matrix illustrates the transactions between different industries, indicating how much each industry purchases from and sells to every other industry. The matrix is defined as the inter-industry transaction matrix *Z*. Each column (buying sector) reflects the inputs that industry *j* acquires for its intermediate consumption, while each row (selling sector) represents the output that industry i provides to the other industries and final demand.

Second, the inter-industry transaction matrix *Z* is normalised by the total output of each sector, yielding the input coefficient matrix *A*. The final demand *f* and total output *X* are derived from the IO table. Together with the inter-industry transaction matrix *Z*, defined in the first step, they are expressed as follows:

$$\underbrace{\boldsymbol{z_{ij}} = \begin{bmatrix} z_{11} & \cdots & z_{1n} \\ z_{21} & \cdots & z_{2n} \\ \vdots & & \vdots \\ z_{n1} & \cdots & z_{nn} \end{bmatrix}}_{\text{Inter-industry transaction}} \qquad \underbrace{\boldsymbol{f_i} = \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{bmatrix}}_{\text{Final demand}} \qquad \underbrace{\boldsymbol{x_i} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}}_{\text{Total output}}$$

where: $\boldsymbol{z_{ij}}$ is the output of sector *i* sold to sector *j* (the inter-industry transaction matrix), $\boldsymbol{f_i}$ denotes the output of sector *i* sold to final demand (the final demand vector), and $\boldsymbol{x_i}$ is the total output of sector *i* (the total output vector).

Equation (1) is constructed based on this matrix and the associated vectors. By normalizing Equation (1) by the total output of each industry, Equation (2)—the input coefficient matrix *A*—is obtained.

$$\boldsymbol{x_i} = \boldsymbol{z_{i1}} + \boldsymbol{z_{i2}} + \cdots + \boldsymbol{z_{in}} + \boldsymbol{f_i} \qquad \textbf{(1)}$$

$\boldsymbol{a_{ij}} = \frac{\boldsymbol{z_{ij}}}{\boldsymbol{x_j}}$, where a_{ij} is the proportion of input from sector *i* in the total output of sector *j*. **(2)**

Third, the Leontief inverse matrix is computed from the input coefficient matrix A. Equation (3) is derived by combining Equations (1) and (2).

$$\boldsymbol{x_i = a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n + f_i} \quad \textbf{(3)}$$

Equation (3) can be represented in matrix form as follows:

$$\underbrace{\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}}_{\mathbf{A}} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}}_{\mathbf{X}} + \underbrace{\begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{bmatrix}}_{\boldsymbol{f}} = \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}}_{\mathbf{X}}$$

This can be further expressed as Equation (4), where X is the total output vector, A is the input coefficient matrix, and f is the final demand vector.

$$\boldsymbol{X = AX + f} \quad \textbf{(4)}$$

Rearranging Equation (4) gives:

$$\boldsymbol{(I - A)X = f} \quad \textbf{(5)}$$

where I is the identity matrix. This formulation isolates the net output required to meet final demand. Solving for X yields:

$$\boldsymbol{X = (I - A)^{-1}f} \quad \boldsymbol{or} \quad \boldsymbol{\Delta X = (I - A)^{-1}\Delta f} \quad \textbf{(6)}$$

Here, $\boldsymbol{(I - A)^{-1}}$ is the Leontief inverse matrix, which captures the total (both direct and indirect) output required from each sector to meet a unit change in final demand, as shown in Equation (6).

In this study, Equation (6) has been developed using Python to apply to either a single sector or multiple sectors, and it has been implemented as a modular approach, separated from the entire model. Specifically, this implementation leverages Python's powerful libraries, including *NumPy* for efficient matrix operations and *Pandas* for handling multi-indexed economic data. This enables flexible and scalable analysis of shock propagation. The modular approach stores such static impact analysis separately, allowing additional economic shocks or stimuli to be easily applied to this analysis in the future.[36]

5.4.3 Dynamic Input–Output model

The static IO model has primarily been used for snapshot analyses of economic shocks, wherein a negative shock to the economy is assumed to cause a contraction in output under steady-state conditions. However, the economy is assumed to respond to shocks with both contraction and recovery in an immediate and dynamic manner, which requires an analytical framework that considers its inherently dynamic characteristics. Therefore, this study proposes a dynamic IO model that can not only simulate the immediate response of the economy to external shocks but also comprehensively analyse the recovery process resulting from policy interventions.

[36] The static impact analysis and the economic stimulus framework are stored separately in the *LeonTradeModel.py* file.

To derive more detailed dynamical responses, the dynamic IO model follows the approach proposed by Klimek et al. [234]. This method models how the economy reacts over time to a demand shock occurring in a steady-state setting, using a first-order ordinary differential equation (ODE). To simplify the analysis, it is assumed that the IO matrix remains constant throughout the specific studied period. This period is defined as the time from the onset of the economic shock until its effects have largely diminished. Based on this assumption and the approach proposed by Klimek et al. [234], the study employs the first-order ODEs to analyse the dynamic response of the economy, which are mathematically expressed as follows:

$$\frac{\boldsymbol{dy}}{\boldsymbol{dt}} = (\boldsymbol{A} - \boldsymbol{I})\boldsymbol{y} + \boldsymbol{S} + \frac{\boldsymbol{dk}}{\boldsymbol{dt}} \tag{7}$$

$$\frac{\boldsymbol{dk}}{\boldsymbol{dt}} = \boldsymbol{a} \times \sum \boldsymbol{y} \tag{8}$$

where:

- $\boldsymbol{dy/dt}$ is the rate of each sector's output change in continuous time.
- $(\boldsymbol{A} - \boldsymbol{I})$ is the internal dynamics of the system: how different sectors of the system are interconnected.
- $\boldsymbol{y}$ is the vector representing the current state of the system.
- $\boldsymbol{S}$ is external shocks.
- $\boldsymbol{dk/dt}$ is the dynamic economic stimulus—change in stimulus over continuous time.
- $\boldsymbol{a}$ represents the percentage of the economic stimulus relative to total output.

As shown in Equation (7), this first-order differential equation describes how the output of each industrial sector changes over time in response to an external shock. Here, $\boldsymbol{y}$ is a vector representing the output of each sector at a given point in time, and $\boldsymbol{dy/dt}$ indicates the rate at which this output changes over time. More specifically, the first term, $(\boldsymbol{A} - \boldsymbol{I})\boldsymbol{y}$, represents the internal interactions among sectors within the economic system. Within this framework, $\boldsymbol{A}$ is the IO matrix that captures the interconnections between sectors, and $\boldsymbol{I}$ is the identity matrix. This term reflects how a change in one sector propagates through the supply chain and influences other sectors, reflecting the cascading impact mechanism within the economic structure. For example, a decline in China's NG demand initially reduces production in Australia's NG extraction industries. These industries then lower their demand for intermediate goods, which in turn affects other mining sectors and transport services. This dynamic propagation captures second-round effects and feedback loops that unfold over time, reflecting gradual adjustment processes rather than an instantaneous shift. The second term, $\boldsymbol{S}$, denotes the external shock. This represents an exogenous impact to the system, such as a sudden drop in demand or a disruption in supply chains, and is active on the economy over a specific period (e.g., from t=0 to t=1).

This study does not merely apply the first-order ODE model proposed by Klimek et al. [234], but extends it by incorporating an economic stimulus framework, thereby enabling the analysis of potential policy measures to offset external economic shocks. The policy measure represents the magnitude of the stimulus injected into the economic system over time, which is expressed as the third term, $\boldsymbol{dk/dt}$, in Equation (7). As shown in Equation (8), this term specifies that the size of the economic stimulus changes over time as a constant proportion of the total output of all industries. Here, the constant proportion, a, is input as a fixed coefficient, and $\sum y$ refers to the total output of all industrial sectors at a given point in time. The product of these represents the process through which a set number of economic stimuli are injected into the total output of the entire industry at a constant ratio over a predefined period, following the dissipation of the external shock. For example, if an external shock occurs in the renewable energy sector and the government adjusts $\boldsymbol{a}$ to allocate more stimulus to this sector, the stimulus amount will vary according to the overall economic scale ($\sum \boldsymbol{y}$), thereby influencing the recovery trajectory of individual industries.

The dynamic IO modelling is implemented using a numerical solution to accurately solve the complex data of EXIOBASE, along with a modularised Python approach that considers future extensibility. Firstly, the numerical solution uses an explicit Runge-Kutta method. Runge-Kutta methods, such as RK23, RK45, and DOP853, are commonly used for solving ODEs in non-stiff systems, which are characterised by their stable numerical behaviour and the ability to achieve sufficient accuracy with relatively small time steps [235]. Configured as a first-order ODE with small time intervals, the dynamic IO model demonstrates stable behaviour and sufficient numerical accuracy [234]. Accordingly, this study considers the EXIOBASE 3-based dynamic IO model, implemented on a quarterly basis, to be a non-stiff system and applies an explicit Runge-Kutta method for its solution. Among the available algorithms suitable for non-stiff problems, such as RK23, RK45, and DOP853, the high-order method DOP853 is selected due to its superior performance in achieving high-precision results [235]. Further technical details on the use of DOP853 in Python, as well as an additional explanation on the non-stiff system, are provided in Appendix AG.

Secondly, the dynamic IO model is modularised to facilitate the integration of additional impact analyses or policy measures in the future. Specifically, the analysis framework and policy implementation, developed in Python using the *NumPy* library for numerical computations, are separated from the model. The code utilises programming features, such as conditional statements, to apply external shocks or stimuli based on the period. Relevant processes (e.g., dynamic shock and stimulus calculations) are stored in separate files[37], so new analyses or policies can be added by modifying only those files—without changing the main model structure.

5.5 Results and discussion

This study first presents the results of an analysis based on the static IO model, followed by additional analysis using the dynamic IO model. The results derived from the dynamic model are compared with those of the static model to assess the robustness of the analysis. Furthermore, incorporating the economic stimulus framework, the dynamic model applies two competing economic stimulus scenarios—one focused on NG and the other on renewable energy. This approach can provide insights into Australia's policy initiatives in response to China's carbon neutrality goals. This comparative analysis offers insights into the broader implications of each scenario.

[37] Dynamic impact and stimulus models are stored in the *Dynamic_Iot_Basic.py* and *Dynamic_Iot.py* files, respectively.

5.5.1 Static Input–Output analysis

The Static IO analysis provides a comprehensive picture of the impact of China's carbon neutrality policy on the Australian economy. In particular, it initially evaluates the effect of changes in China's NG demand on the Australian economy based on the NG demand projections for 2030 and 2060 presented in Figure 5.2, using the 2022 baseline version of EXIOBASE 3. Subsequently, the analysis incorporates the changes in China's coal and oil demand under the IEA's NZE scenario, as shown in Table 3, to assess the overall impact of changes in China's fossil fuel (NG, coal, and oil) consumption on the Australian economy. These analyses provide how changes in China's fossil fuel demand, driven by its carbon-neutral transition, affect the Australian economy. Further details of the analysis are as follows.

China’s single-sector impact: natural gas demand

The EXIOBASE 3-based static single-sector IO analysis shows that the overall impact of changes in China’s NG consumption on the Australian economy under China’s carbon-neutral policy is negligible. In this analysis, the related NG consumption sectors in EXIOBASE 3 include *Production of electricity by gas*, *Manufacture of gas*, and *Distribution of gaseous fuels through mains*. As shown in Table 5.4, the NG consumption is projected to increase by 56% by 2030 compared to 2022, and decrease by 49% by 2060. In this scenario, the Australian economy experiences a 0.64% rise in total output by 2030 and a 0.56% decline by 2060, indicating that the overall effect on Australia’s economy is insignificant.

This outcome becomes more plausible—particularly when considering Australia’s gas export strategy, which outlines a strategic shift toward emerging LNG markets in Southeast Asia [209]—as Australia could divert some of China’s declining natural gas (NG) demand after 2030 to these markets. Australia has a significant portion of its LNG export contracts set to expire between 2030 and 2035 [209]. Meanwhile, Indonesia and Vietnam have scheduled their coal phase-out for the 2040s and 2050s [236]. Hence, the Australian NG can serve as a transitional fuel to replace coal in these two countries during this interim period. As a result, the impact of declining NG demand from China on the Australian economy could be limited.

Table 5.4 The impact of China’s NG sector on Australia

Impact (% Change compared to 2022)		**2030**	**2060**
One sectoral impact in China	NG consumption	+ 56 %	- 49 %
Economic impact	Total output change in Australia	0.64%	- 0.56%

Additionally, this analysis shows that changes in China’s NG consumption affect Australia’s key resource export industries, such as coal, NG, and iron ore [237]. In terms of the normalised responses of Australia’s sectors for the NG consumption scenarios (see Figure 5.8 in Appendix U for 2030 and Figure 5.10 in Appendix W for 2060), which are defined as the ratio of each sector's output to the maximum output across all sectors, the increase in China's NG consumption in 2030 has the most significant impact on Australia's NG extraction sector, followed by its coal and iron ore mining sectors. Similarly, in 2060, China's reduced NG consumption most significantly impacts Australia's NG extraction, followed by its coal and iron ore mining sectors. These results are also reflected in the relative responses for Australia’s sectors (see Figure 5.9 in Appendix V for 2030 and Figure 5.11 in Appendix X for 2060).

However, this analysis fails to reflect the expected consumption of China's NG and Coal by 2030, indicating that NG and coal consumption are anticipated to move in the same direction by that year. According to the DCGs, by 2030, coal demand is expected to decrease, and NG consumption is projected to increase, replacing the reduction in coal demand. To align with these anticipated carbon-neutral scenarios, a static IO analysis for multiple sector impacts, including coal and oil consumption, is conducted.

China's multiple-sector impact: natural gas, coal, and oil demand

The EXIOBASE 3-based static multiple-sector IO analysis, which incorporates not only changes in China's NG consumption but also coal and oil demand scenarios under the country's carbon-neutral policy, assesses the overall impact of these energy transitions on the Australian economy, suggesting a progressively intensifying negative effect by 2060. The NG consumption is defined by integrating the same sectors used in static single-sector IO analysis. For the coal and oil demand, the relevant sectors from EXIOBASE 3 are *Production of electricity by coal* and *Production of electricity by petroleum and other oil derivatives*, respectively. According to Table 5.5, the NG consumption is projected to increase by 56% in 2030 compared to 2022, while the coal demand sector, *Production of electricity by coal*, is expected to decrease by 43%, and the oil demand sector, *Production of electricity by petroleum and other oil derivatives*, by 14%. Under this scenario, Australia's total output is projected to decrease by 0.08% in 2030. In the 2060 scenario, the NG consumption is expected to decrease by 49% compared to 2022, the coal demand by 93%, and the oil demand by 80%. As a result, Australia's total output is projected to decrease by 2.43% in 2060. The larger total output decrease in 2060 compared to 2030 suggests that China's decarbonisation achievement could give rise to structural changes in the Australian economy. These changes are expected to occur primarily in Australia's key resources export industries.

Table 5.5 The impact of China's fossil fuel sector on Australia

	Impact (% Change compared to 2022)	**2030**	**2060**
Multiple sectoral impacts in China: NG, Coal, and Oil	NG consumption	+ 56 %	- 49 %
	Production of electricity by coal	- 43%	- 93%
	Production of electricity by petroleum and other oil derivatives	- 14%	- 80%
Economic impact	Total output change in Australia	- 0.08%	- 2.43%

Consistent with the static single-sector IO analysis of NG demand, this analysis also reveals that Australia's key resource export industries [237]—coal, NG, and iron ore—are the most affected, along with other sectors that are impacted. In particular, as shown in Figure 5.12 (see Appendix Y), the normalised responses of Australia's sectors for China's fossil fuel demand scenarios in 2030 indicate that Australia's NG extraction sector experiences the most significant increase, while its coal mining sector shows the most significant decrease. Meanwhile, as shown in Figure 5.14 (see Appendix AA), the normalised responses in 2060 indicate that the coal mining, NG extraction, and iron ore mining sectors are the most affected in that order. These results align with the relative responses of Australia's sectors for 2030 and 2060 (see Figure 5.13 in Appendix Z for 2030 and Figure 5.15 in Appendix AB for 2060). Furthermore, the relative responses for 2060 (see Figure 5.15 in Appendix AB) show that, after the NG extraction, coal, and iron ore mining sectors, other sectors, including construction, agriculture, and fishing, also experience significant impacts.

These results imply that the achievement of China's carbon neutrality goals by 2060 could have broad impacts on China's electricity demand and industrial structure as well as Australia's major economic sectors. According to the EXIOBASE 3-based static IO analysis, China's NG and coal consumption is closely linked with its electricity demand, which is also associated with China's steel industry, which requires large amounts of electricity [64]. China's carbon-neutral energy transition, focused on the phase-out of fossil fuels, the increase in renewable energy, and improvements in electricity efficiency, naturally lead to a reduction in fossil fuel-based power consumption, resulting in decreased demand for thermal coal and NG. This, in turn, directly impacts Australia's coal and NG export industries. Furthermore, China's industrial decarbonisation transition could result in structural shifts in the carbon-intensive steel industry, particularly through increased use of scrap steel and wider adoption of energy-efficient Electric Arc Furnace (EAF) methods [62].[38] The shift towards scrap-based EAF could reduce the need for iron ores and reducing agents, such as coking coal, leading to a reduction in imports of these commodities from Australia. The anticipated decreases in coal, NG, and iron ore demand in China signal a structural risk of long-term demand reduction for Australia's resource export industries—coal, NG, and iron ore—which are key sectors of the national economy [237]. The contraction of these resource export industries could also have a cascading negative effect on its other industries, such as construction and agriculture.

5.5.2 Dynamic Input–Output analysis

The dynamic IO model analyses sectoral economic outputs both in aggregate and over time by applying the 2030 and 2060 coal, NG, and oil demand scenarios from Table 5 used in the static multiple-sector IO analysis. As in the static estimation, the 2030 scenario reflects a 56% increase in NG consumption, along with 43% and 14% decreases in coal- and oil-based power generation, respectively, in China. This scenario represents a shift in China's energy consumption, replacing reduced coal usage with increased NG consumption. Under this scenario, this analysis for 2030 (see Figure 5.16 in Appendix AC) shows that, over time, Australia's NG extraction sector experiences the most considerable growth, while the coal mining sector undergoes the most significant decline. These changes are consistent with the 2030 static multiple-sector IO analysis results. However, the analysis indicates that the 14% reduction scenario in oil-derived power generation—reflecting the decline in China's oil consumption—does not significantly impact the Australian sectors.

As shown in the quarterly total output changes for 2030 in Figure 5.4, total output declines significantly in the early stages, increases temporarily, and drops again. This is because the reduction in coal and oil consumption has a greater negative impact than the increase in NG consumption, and this impact then spreads across various sectors. Over time, as endogenous economic adjustment processes are activated, output fluctuates between contraction and recovery phases before gradually converging to a new equilibrium. The negative output impact lasts for about five years before becoming positive from the ninth year onward.

[38] The EXIOBASE 3-based IO analysis does not reflect the increase in the use of scrap-based EAF. However, due to China's carbon neutrality goals, its industrial decarbonisation transition is inevitable. Hence, this study assumes an increase in the use of scrap-based EAF.

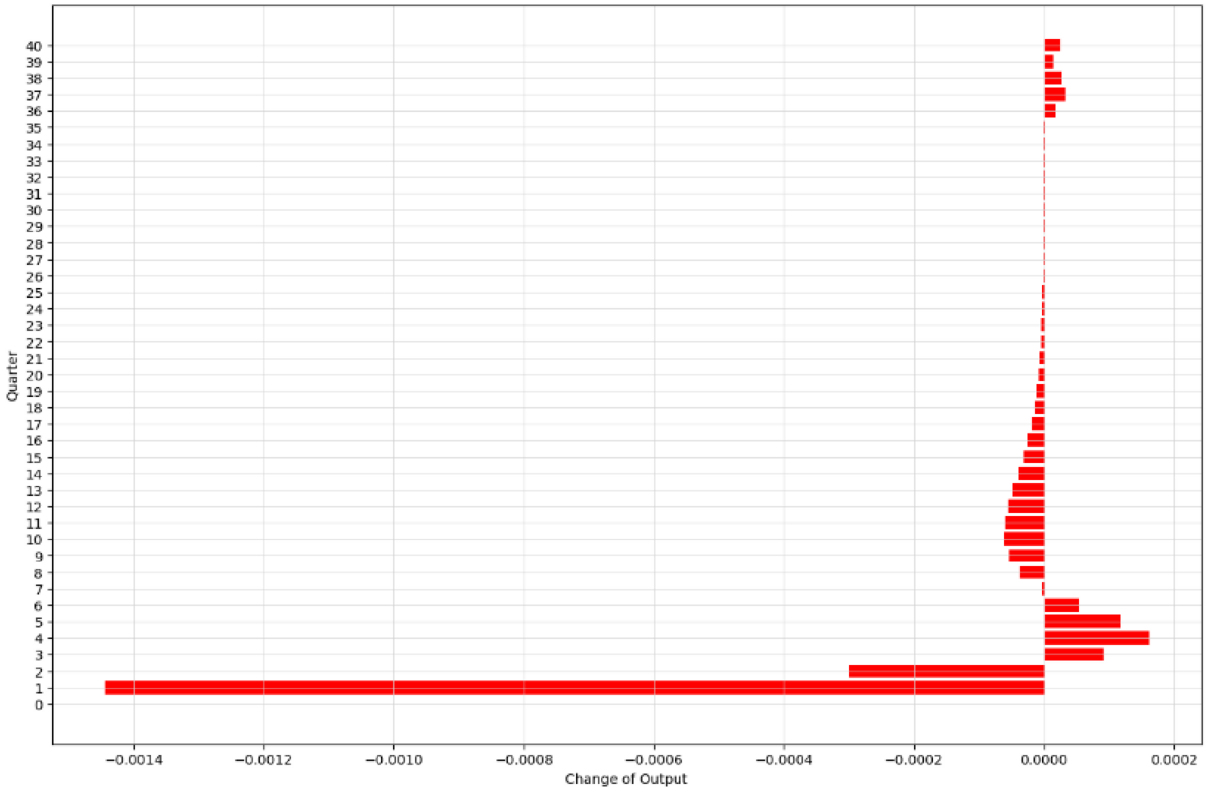


Figure 5.4 Total output changes per quarter for 2030

Notes: Quarter 0 is the baseline period in 2030 before the shock. Quarter 1 is the first quarter after the shock.

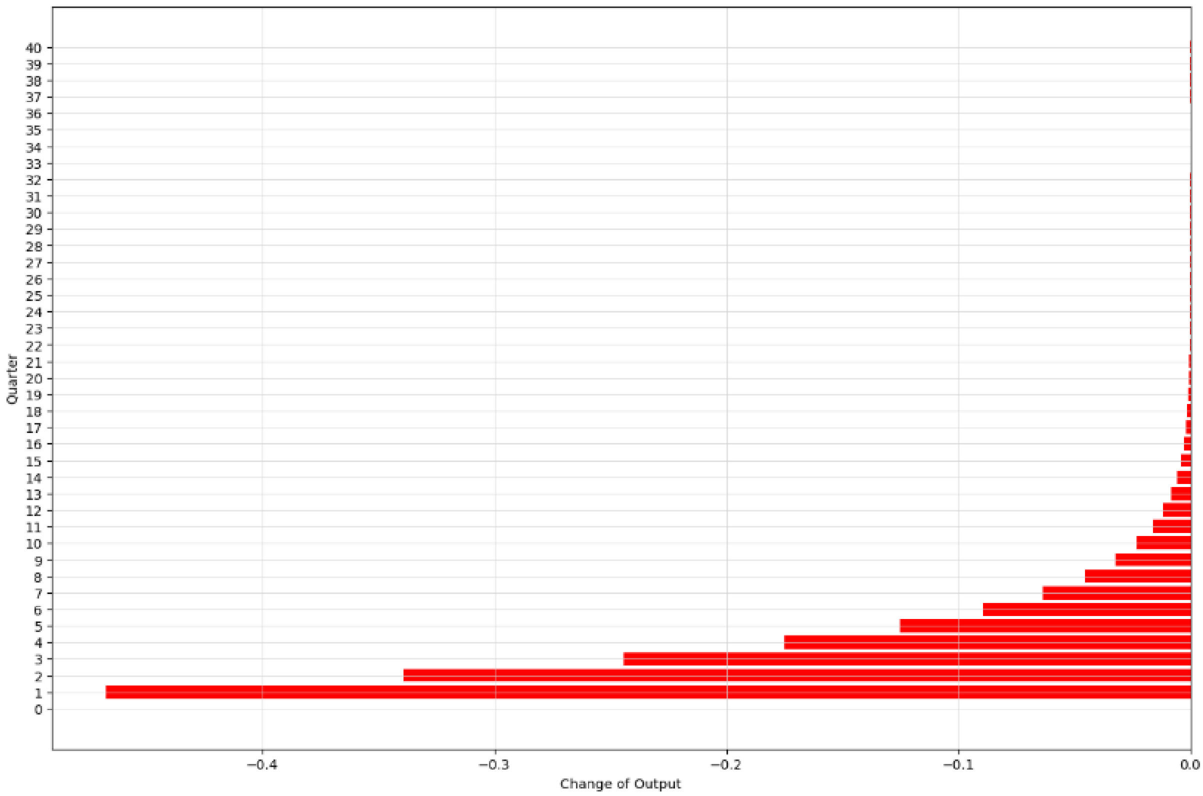


Figure 5.5 Total output changes per quarter for 2060

Notes: Quarter 0 is the baseline period in 2060 before the shock. Quarter 1 is the first quarter after the shock.

In 2060, according to the IEA NZE scenario, which assumes that China will significantly reduce fossil fuel consumption to achieve its carbon neutrality goals, its NG consumption is projected to decrease by 49% compared to 2022, with coal- and oil-derived power generation also decreasing by 93% and 80%, respectively. Under this scenario, the dynamic IO multiple analysis for 2060 (see Figure 5.17 in Appendix AD) shows that, over time, Australia's coal mining sector experiences the largest decline, followed by the NG extraction sector, which is also significantly impacted. Meanwhile, other sectors are relatively less affected; however, it is particularly noteworthy that the iron ore mining sector is gradually impacted following the coal mining and natural gas extraction sectors. This finding is consistent with the results of the static multiple-sector IO analysis, which identified a similar pattern of sectoral impact. As discussed in the static IO analysis, this reflects the impact of China's industrial decarbonisation transition on its carbon-intensive steel industry, which in turn negatively affects Australia's iron ore export sector. Moreover, these changes imply that the contraction of Australia's key sectors—coal and NG exports—negatively impact the overall industry, as shown in Figure 5.5.

The quarterly total output changes for 2060 in Figure 5.5 indicate that initially, there is a sharp decline in economic output, decreasing by 0.48 in the first quarter. It then takes about 24 quarters (around 6 years) for this impact to be fully adjusted. Since this dynamic analysis assumes no additional stimulus for economic recovery, the adjustment process is considerably prolonged. To address this, the following section presents a comparative analysis of two economic stimulus scenarios that could help mitigate the impact.

5.5.3 Comparative analysis of two competing economic stimulus cases

This comparative analysis derives insights into Australia's policy response by incorporating its potential economic stimulus measures into the dynamic IO model to mitigate the impact of China's carbon neutrality. Two competing economic stimulus cases are applied to generate practical policy insights for addressing the economic impact on Australia identified in the dynamic IO analysis for the year 2030. In contrast, the 2060 analysis is excluded from applying the two economic stimulus cases, as it represents a time frame considered too distant for actionable policy planning.

Table 5.6 Two competitve economic stimulus cases

Cases		Applied sectors
Case 1	Economic stimulus for the Australian NG sectors	Production of electricity by gas; Manufacture of gas; and Distribution of gaseous fuels through mains
Case 2	Economic stimulus for the Australian renewable sector	Production of electricity by solar photovoltaic; Production of electricity by solar thermal; and Production of electricity by wind

As depicted in Table 5.6, the first economic stimulus case focuses on the NG industry, one of Australia's core economic sectors, while the second targets the renewable energy industry. The first economic stimulus case is based on the projected growth potential of NG through 2030 and Australia's current energy policy, the *Future Gas Strategy* (see Appendix AH for more details). The Australian government projects shifts in the external environment of the gas industry—specifically, the growth potential of NG as a transition fuel through 2030, followed by a decline expected thereafter—driven by the decarbonisation policies of Australia's major Asian trading partners [209]. Additionally, this strategy aims to position the NG sector as playing a key role in Australia to achieve its decarbonisation target by 2050 [209]. Hence, although this policy may change in the future, this study assumes that the Australian government could continue to support the use of NG through economic stimulus measures, given its importance as a driver of the national economic growth. Based on this assumption, the first economic stimulus case is examined using the dynamic IO analysis.

The second economic stimulus case for the renewable energy industry explores the potential for Australia to transition from a fossil fuel-based industrial structure to one centred on renewable energy. Drawing upon a representative global case of transitions to renewable energy, Denmark—which, like Australia, has historically relied on fossil fuel exports, including NG [238]—has successfully transitioned to a renewable energy-focused system and aims to phase out NG entirely by 2050 [74] (see Appendix AI for more details on the Denmark transition to renewables). As demonstrated by the Danish case, Australia, with its abundant renewable energy resources, could also have the potential to transition its economy towards a renewable energy-driven growth model. Based on this assumption, the second economic stimulus case is explored using the dynamic IO analysis.

In the two economic stimulus cases applied to the dynamic model, the model is structured such that the multi-sector impact occurs between time t = 0 and t = 1, and the stimulus is assumed to be implemented immediately after the impact, continuing over a specified period, two quarters (six months). The stimulus is assumed to be equivalent to approximately 0.1 of total output. In particular, it is allocated as 0.115 for the NG industry and 0.1 for the renewable energy industry, so that the maximum economic stimulus effect for each sector is approximately the same—around 0.585—in the fourth quarter, as shown in Appendices AE and AF. As illustrated in Table 5.6, the economic stimulus is applied to the three sectors related to the NG consumption—*Production of electricity by gas, Manufacture of gas*, *and Distribution of gaseous fuels through mains*—and the three sectors related to renewable energy consumption—*Production of electricity by solar photovoltaic*, *Production of electricity by solar thermal*, *and Production of electricity by wind*.

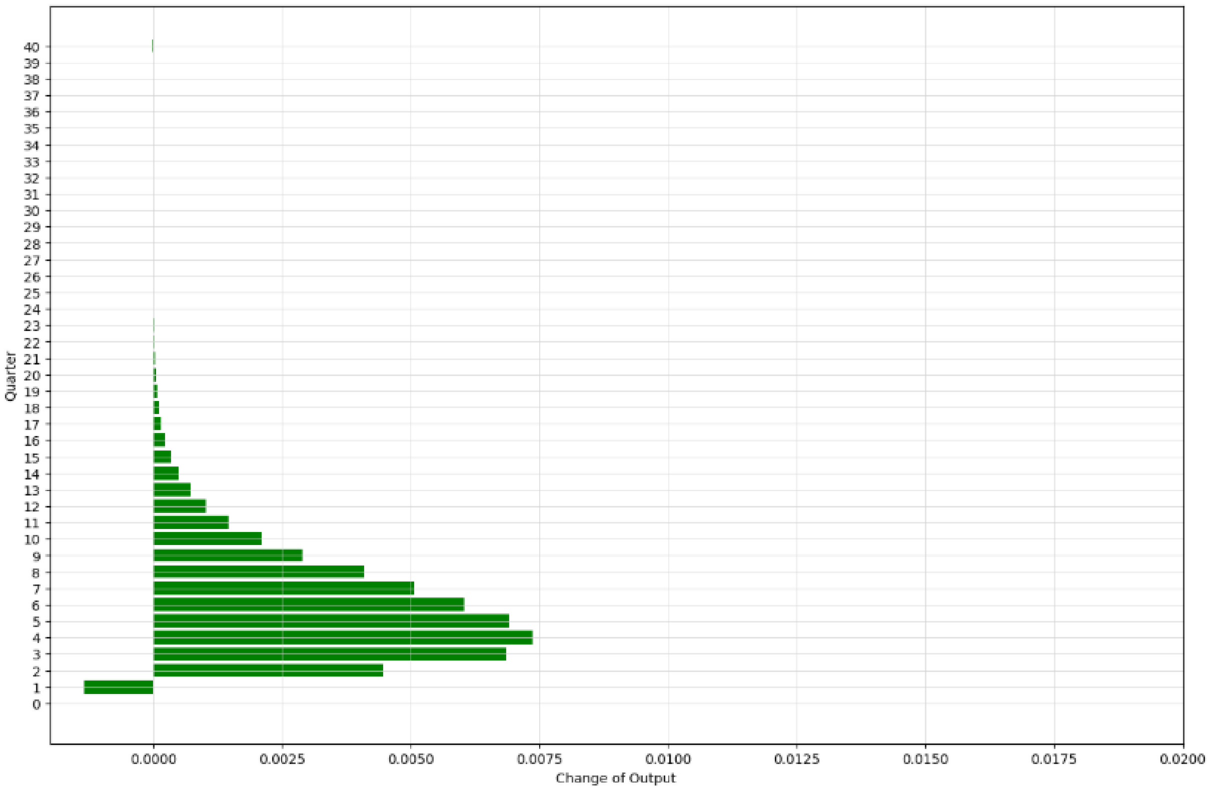


Figure 5.6 Total output change per quarter with the economic stimulus in Australia's NG industry

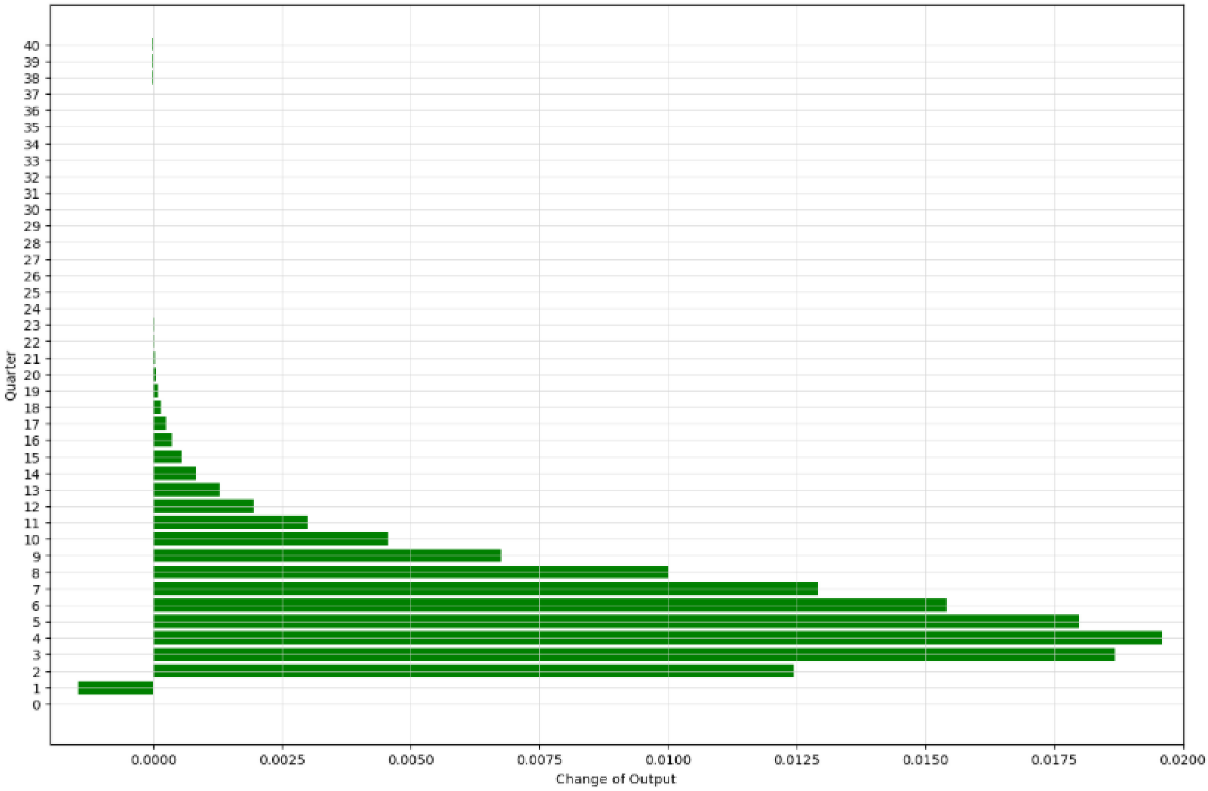


Figure 5.7 Total output changes per quarter with the economic stimulus in Australia's RE industry

According to the first economic stimulus case results presented in Figure 5.18 (see Appendix AE), fossil fuel sectors have the most significant impact. As expected, the most affected industries are those directly related to NG consumption. Among the ripple effects, the most notable impact appears in the coal mining industry, while no apparent effects are observed in other industrial sectors. These results suggest that the economic stimulus targeting the NG industry could increase the utilisation and operational efficiency of fossil fuel-related logistics infrastructure, such as railroads and ports, thereby indirectly benefiting the coal mining industry, which shares the use of this infrastructure. Furthermore, as shown in Figure 5.6, The first economic stimulus case triggers a recovery from the second quarter.

Compared to the first case, the results of the second economic stimulus case for the renewable energy industry presented in Figure 5.19 (see Appendix AF) show that a broader range of sectors is affected. The industries most impacted are those directly related to renewable energy. Additionally, spillover effects are observed across various other industries. In descending order, these spillover-affected industries are the sale, maintenance, and repair of motor vehicles and motorcycles (including parts and accessories), other land transport, financial intermediation excluding insurance and pension funding, and hotels and restaurants. Of particular interest is the fact that the automotive and motorcycle-related industry appears to be the second most impacted sector after the renewable energy industry itself. This could be interpreted as a result of accelerated electrification driven by the promotion of renewable energy, which in turn could lead to the proliferation of electric vehicles (EVs). Moreover, the second economic stimulus appears to expand to various industries, including logistics and financial industries. This implies that the expansion of renewables in Australia could create wider spillover effects across related industries. As indicated in Figure 5.7, the second economic stimulus case also initiates a recovery starting the second quarter. Lastly, when comparing the quarterly total output changes under the economic stimulus cases for the NG industry and the renewable energy industry, as seen in Figures 5.6 and 5.7, it is evident that the economic stimulus targeting the renewable energy sector demonstrates a significantly higher level of resilience.

As observed thus far, the economic stimulus case targeting the renewable energy sector not only contributes more comprehensively to mitigating the economic impact of China's carbon neutrality goals on the Australian economy, but also exerts a broader positive influence across various sectors, while exhibiting a more rapid overall economic recovery.

5.6. Conclusions and policy implications

This study analyses the economic impact on Australia from changes in China's fossil fuel consumption—focusing on shifts in NG demand—driven by its carbon neutrality targets. The study also proposes policy initiatives for Australia to adapt to the estimated economic changes. The research methods involve China's NG consumption scenarios built by Prophet-SARMIAX [215] as well as coal and oil demand scenarios based on the IEA NZE pathways [152]. These scenarios are then applied to static and dynamic IO models to evaluate the impact of China's 2030 and 2060 decarbonisation goals on the Australian economy. Furthermore, two economic stimulus cases are implemented in the dynamic IO analysis to assess Australia's economic resilience and responses dynamically to this impact. Lastly, the study introduces a modelling approach that separates the impact analysis and economic stimulus calculations from the models, applying a modularised Python approach to simplify the integration of future impact analysis or policy implementations.

The study indicates that Australia's NG extraction and coal mining sectors are the industries most affected by China's carbon neutrality targets. The static IO analysis shows the largest increase in output for the NG extraction sector under China's carbon neutrality scenario by 2030, whereas it suggests the most significant decline in the coal mining sector. By 2060, however, both the coal mining and NG extraction sectors in Australia show significant reductions. The contraction of these two sectors due to China's achievement of carbon neutrality is expected to have spillover effects on Australia's other sectors, particularly the iron ore mining sector, gradually leading to a decrease in Australia's total output by 2060. Subsequently, the same carbon neutrality scenario—the phase-out of fossil fuels—is applied in the dynamic IO analysis. The results of the dynamic IO analysis are consistent with those of the static IO analysis, thereby supporting the robustness of the findings. Through temporal analysis, the dynamic IO model shows that by 2030, the Australian economy experiences alternating periods of negative downturns and partial recoveries due to adjustment mechanisms in the economic system. Over time, Australia's total output stabilises to a more consistent level. However, by 2060, the Australian economy could be in a prolonged and persistent recession period. In other words, the decline in Australia's key economic sectors—NG extraction and coal mining—stemming from China's carbon neutrality goals could have significant negative impacts on Australia's overall economy.

To mitigate the negative economic impacts of China's carbon neutrality targets on Australia, the study integrates economic stimulus cases into the dynamic IO model to explore potential policy responses for Australia. Specifically, two competing types of economic stimulus scenarios are examined: one targeting the NG sector and the other focusing on the renewable energy sector. The results suggest that the NG-focused stimulus primarily benefits industries directly related to NG consumption, with some spillover effects observed in the coal mining sector. However, the impact on other industries appears to be limited. In contrast, the stimulus aimed at renewable energy has a more significant and widespread effect, positively impacting industries directly involved in renewable energy and producing spillover effects across various sectors. Moreover, the renewable energy-focused stimulus demonstrates a higher degree of economic resilience, suggesting that targeted investments in the renewable energy sector could be more effective in offsetting the adverse economic impacts of China's carbon-neutral transition on the Australian economy.

Building on this analysis, three key policy implications emerge for mitigating the economic impacts of China's carbon neutrality targets on Australia. First, providing direct financial support—such as tax credits and fiscal subsidies—to the renewable energy sector is a vital initiative for driving sustainable growth in Australia. As demonstrated in this study, industries directly affected by the renewable energy-focused economic stimulus show immediate increases in total output. This implies that the economic stimulus package, such as tax credits and fiscal support, helps reduce costs or provide direct funding to renewable firms, thereby supporting the development of diverse renewables, particularly in Australia, which has abundant renewable resources [239]. For example, Australia could strategically support China's hydrogen-based steel industry by exporting green hydrogen produced using its competitively priced, high-quality wind and solar energy [219]. This could, in turn, contribute to the growth of the hydrogen sector in Australia [219]. Such green energy exports could also help offset the potential decline in Australia's coal and NG exports, triggered by China's transition towards carbon neutrality [228]. Consequently, such financial support could foster economic and environmental sustainability by reorienting coming challenges into new drivers of growth [72].

Second, policies supporting EVs and related infrastructure should be implemented in parallel with financial support for renewable energy. As indicated in the study, the industry most indirectly impacted by the economic stimulus case for renewables is the sale, maintenance, and repair of motor vehicles and motorcycles. This suggests that financial incentives for renewables also imply an expansion of EVs. To facilitate the widespread adoption of EVs, government support for electrification acceleration, charging infrastructure, and battery technologies is essential [240]. In this context, promoting the adoption of EVs through subsidies and tax incentives can also stimulate the renewable energy industry associated with EVs [241]. This economic stimulus package for renewables and EVs can mutually reinforce each other, ultimately contributing to overall economic revitalisation.

Third, developing industrial clusters based on renewable energy should be supported. The study suggests that the economic stimulus case for renewables could create spillover effects across a diverse array of industries, including logistics and finance. To maximise these spillover effects, it is important not only to provide direct support for connecting to renewable energy, but also to establish industrial clusters where related companies can gather, and to offer policy support that promotes mutual cooperation and innovation within these clusters [242]. This approach can boost collaboration and innovation, enhancing the competitiveness of the renewable energy sector and its associated industries in Australia.

The limitation of this study lies in the difficulty of reflecting current technological advancements. In the static IO analysis, China's carbon neutrality policy reduces fossil fuel use, negatively impacting Australia's iron ore mining industry and thereby reducing iron ore exports. This is because, in the IO table, a reduction in fossil fuel use in China could lower electricity consumption, which is linked to a contraction of the steel industry. Since the steel industry accounts for a large portion of electricity consumption [16], this decline in electricity use results in a reduction in steel industrial activity, ultimately lowering the demand for iron ore imports from Australia. In practice, however, China is replacing fossil fuel-based electricity with renewable energy-based electricity and developing environmentally friendly steel production methods using renewable energy and hydrogen [26]. The IO table is based on existing economic transactions, so such technological advancements are not reflected. To overcome this limitation, future research could utilise machine or deep learning-based prediction and simulation techniques to incorporate advancements such as green steel technologies into the IO table, enabling more practical analysis.

CHAPTER 6.0

COMPREHENSIVE CONCLUSION

As China advances towards its 2060 carbon neutrality target, structural shifts in its energy mix—particularly the reduction in fossil fuel dependency—pose significant challenges for Australia, which has served as a main energy supplier to China. Simultaneously, China's efforts to liberalise its NG market—through increased domestic production, diversification of import sources, and a shift towards spot pricing mechanisms—further complicate Australia's export prospects. Together, these dynamics in China represent significant shifts in the external demand environment facing Australia's NG industry. This concluding chapter synthesises the key findings and implications, and offers policy recommendations for navigating this evolving bilateral energy relationship.

This study addresses a critical gap in the literature by jointly examining the impact of China's carbon neutrality goal and NG market liberalisation— two key structural drivers of change in China's NG sector—on the Australian economy. Previous research has tended to focus on either decarbonisation or market reform in isolation, often adopting a short-term perspective. In contrast, this study integrates both dimensions and extends the analytical horizon to 2060. To generate a more robust and forward-looking understanding, it employs a set of novel hybrid methodologies—combining big data analytics, machine learning, deep learning, and time-series analysis—each tailored to distinct research aims across the chapters. These interdisciplinary approaches provide both empirical insights and actionable recommendations for navigating a rapidly evolving China-Australia energy relationship.

This study is structured in three main stages. First, it identifies the key drivers of China's NG demand from a temporal perspective. Second, based on these drivers, it develops a short-term forecasting framework and derives long-term net-zero projections under various scenarios, thereby supporting both near- and long-term decarbonisation planning. Third, drawing on these demand projections, the study evaluates the economic implications for Australian industries and outlines policy directions for strategic adaptation.

Following this analytical structure, Chapters 2 through 5 unfold the study's content in a stepwise manner, outlining the key methodological approaches applied in each chapter. Chapter 2 introduces a novel temporal shrinkage framework that integrates big data decomposition techniques (LASSO/ALASSO) with time-series modelling (ECM) to analyse the key drivers of change in China's NG demand. This framework systematically identifies the complex dynamics influencing NG consumption, including short- and long-term factors, including temperature, thermal coal and HH gas prices, PNG and LNG imports, market liberalisation, and NG infrastructure-related variables. Chapter 3 builds on the drivers identified in Chapter 2 to develop a hybrid near-term forecasting model that optimally combines a trend-sensitive statistical model (SARIMAX2) and a deep learning model (LSTM) via an ANN. This model offers an empirical basis for guiding near-term investment and operational decisions amid the volatility of the carbon neutrality transition. Chapter 4 extends the analysis by projecting NG consumption to 2060 under China's carbon neutrality scenarios, using a hybrid model that integrates Prophet and SARIMAX to outline net-zero NG demand pathways. Chapter 4 extends this analysis by projecting NG consumption to 2060 under China's carbon neutrality scenarios and then suggesting net-zero NG demand pathways using a hybrid Prophet–SARIMAX model. Finally, Chapter 5 evaluates the economic impacts of these projections on Australia using both static and dynamic IO models, and derives policy implications to guide strategic responses.

6.1 Key findings

The findings presented below are organised by six key thematic issues and draw on evidence from Chapters 2 through 5.

First, the temporal dual role of NG: From a transitional fuel to a structural constraint. In China's carbon neutrality strategy, NG plays a critical role in the short term as a transitional fuel replacing coal. However, over the long term, it presents a dual character that could pose risks to achieving net-zero targets. Chapter 2 identifies key drivers of NG demand—including market liberalisation, import, and infrastructure-related factors—demonstrating how these elements contribute to structural changes in NG consumption. This suggests that NG could serve both as a tool for carbon emission mitigation and as a potential obstacle to long-term decarbonisation efforts. Chapter 4 substantiates these concerns through scenario-based forecasting, showing that under existing NZE scenarios, the complete phase-out of NG remains challenging. Together, these chapters highlight the need to redefine the strategic approach to NG's role along a temporal axis, calling for a carefully calibrated balance between short-term utilisation and long-term phase-out.

Second, Decarbonisation-aligned NG market liberalisation. China's NG market liberalisation is increasingly viewed not merely as a mechanism for improving market efficiency, but as a reform strategically aligned with national decarbonisation objectives. Chapter 2 identifies this market liberalisation as a key driver of NG demand, proposing that a market-based pricing system enhances NG's short-term competitiveness, thereby accelerating coal-to-gas substitution. Over the longer term, such a system can internalise environmental costs and provide a structural basis for transitioning to renewable energy. Chapter 4 builds on this premise by embedding the market liberalisation within a broader decarbonisation framework. Under the net-zero NG demand scenario, market-oriented reforms reduce artificial price distortions and improve price transparency. This recalibrates NG's long-term competitiveness, facilitates the shift to renewables, and ultimately contributes to a decline in NG consumption—demonstrating the potential of a decarbonisation-aligned NG market liberalisation pathway.

Third, strategic reconfiguration of China's NG Infrastructure. NG infrastructure is regarded as a critical asset within China's energy transition strategy. Chapter 2 highlights its strategic significance in shaping both short- and long-term demand trajectories, emphasising the need for a redefinition of its role in light of China's dual carbon targets—peaking emissions by 2030 and achieving carbon neutrality by 2060. Chapter 3 empirically demonstrates, through the study's hybrid forecasting model, that infrastructure factors play a pivotal role in short-term demand prediction. Chapter 4 shifts the focus to long-term infrastructure adaptation, proposing that, in scenarios where NG demand undergoes substantial reduction, existing infrastructure should be repurposed to support the production and transport of alternative energy carriers such as hydrogen. Accordingly, NG infrastructure must be approached through a flexible strategic lens—one that accommodates both immediate responsiveness and long-term transformation in alignment with decarbonisation imperatives.

Fourth, advancement of forecast accuracy via hybrid modelling approaches. This study introduces and empirically validates a suite of hybrid forecasting models aimed at enhancing the accuracy of NG demand projections across both short- and long-term horizons. Chapter 3 presents a hybrid model that optimally integrates a trend-sensitive statistical model (SARIMAX2) and a deep learning model (LSTM) via an ANN. This approach significantly enhances short-term forecasting performance, even under conditions of high market volatility and frequent fluctuations in the NG sector. This ANN-optimised hybrid model outperforms conventional time-series models (SARIMAX), standalone deep learning models (LSTM), and simple integrated hybrid models—additively combined linear–nonlinear models—demonstrating superior predictive capability. Chapter 4 demonstrates the effectiveness of a hybrid Prophet-SARIMAX model for scenario-specific long-term projections. By accurately replicating outcomes from the IEA and BP Energy Outlooks, this model exhibits strong predictive performance. Its capacity to mitigate uncertainty further underscores its value as a strategic tool for long-term energy planning and policy formulation.

Fifth, NG demand trajectories towards net zero in China. Under China's existing carbon neutrality scenarios, this study identifies the limitations of the current NG consumption pathway in achieving net-zero emissions and proposes an alternative pathway that necessitates a more radical transition. Chapter 4 projects NG consumption towards China's 2060 carbon neutrality target using a scenario-based hybrid Prophet-SARIMAX model. The IEA NZE scenario suggests that NG consumption continues at a certain level, which demonstrates the limitations in achieving carbon neutrality. In response, an alternative pathway—referred to as the ANZE pathway—based on the renewable energy-centred OECM scenario is proposed to underscore the need for a more aggressive shift. This pathway necessitates a more accelerated reduction in the key drivers of NG consumption than the standard pathway, demanding far-reaching structural reforms and a comprehensive reconfiguration of the energy system beyond what current policies envisage. Rather than offering a simple forecast, this analysis aims to redefine a practical and implementable trajectory for achieving a genuine net-zero future within the NG sector.

Sixth, realignment of Australia's economic strategy in response to China's carbon neutrality goals: Transitioning towards renewables. As China's carbon neutrality goals pose significant challenges to Australia's fossil fuel–dependent industries, a transition towards a renewable energy–driven economy is becoming increasingly necessary. Chapter 5 analyses the economic impacts on Australia stemming from changes in China's fossil fuel consumption—particularly NG demand—driven by its carbon neutrality targets. Specifically, both static and dynamic IO models are employed to assess the effects of China's decarbonisation goals for 2030 and 2060 on the Australian economy. The analysis reveals that the NG extraction and coal mining industries suffer the most significant impacts, with spillovers into other sectors, such as iron ore mining, potentially leading to an overall decline in Australia's total output by 2060. Notably, the dynamic IO analysis indicates that the Australian economy could face prolonged long-term downturns over time. Two economic stimulus scenarios—NG-centred and renewable energy-centred—are applied to the dynamic IO model to evaluate Australia's resilience and response strategies. The results show that the NG-centred stimulus has limited effectiveness, whereas the renewable energy-centred stimulus generates positive ripple effects across multiple industries, enhancing economic resilience.

6.2 Major policy implications

Based on the key findings above, four major policy implications can be drawn for China's NG market in achieving its carbon neutrality goals. These findings are also applicable to Australia, given its close energy trade relationship with China, and offer meaningful insights for shaping Australia's future policy directions.

First, to achieve carbon neutrality in the NG sector, China must implement strong decarbonisation policies that simultaneously introduce carbon pricing and expand investments in renewable energy. This dual approach can gradually internalise the environmental costs of fossil fuels, regulate NG demand, and support a phase-out of NG in the long term. Chapter 2 demonstrates that both LNG imports and coal prices are positively correlated with NG consumption in the short and long term. Accordingly, the introduction of carbon pricing mechanisms—such as a carbon tax or an emissions trading scheme—could lead to increased NG consumption through rising coal prices. In the long run, these mechanisms could be extended to LNG and other fossil fuels to facilitate the phase-out of NG. Chapter 4 highlights that, under current carbon neutrality scenarios, the complete phase-out of NG remains unachievable. To overcome this limitation, more radical policy interventions are required, premised on the large-scale expansion of renewable energy and technological advancement. Synthesising the findings across chapters, it becomes evident that strong policy intervention—combining carbon pricing and technological investment—can provide a practical foundation for internalising the environmental costs of fossil fuels and regulating NG consumption. This, in turn, offers a more viable pathway towards achieving China's carbon neutrality targets.

Second, NG-related infrastructure—identified as a key driver of NG consumption in China—should be designed not only to meet short-term demand but also to facilitate a transition towards alternative energy sources, such as hydrogen. When combined with these decarbonisation policies, such infrastructure planning could effectively reduce NG consumption. As discussed in the key findings, Chapter 2 identifies NG infrastructure as a key driver influencing both short- and long-term NG demand. Chapter 3 demonstrates its role as a critical factor in short-term demand forecasting. In contrast, Chapter 4 suggests that, in the event of a substantial reduction in NG demand, existing infrastructure should be repurposed to support the production and transport of alternative energy sources, such as hydrogen. Based on these analyses, NG infrastructure should be designed not only to meet immediate energy needs but also to support energy transition pathways aligned with decarbonisation objectives. This approach can serve as a key policy instrument for effectively curbing NG consumption and advancing China's carbon neutrality goals.

Third, promptly adopting a market-based NG pricing system is essential to address price distortions in the NG market and ensure that the environmental costs of fossil fuels are adequately reflected. This forms a critical institutional foundation for the substantive achievement of carbon neutrality. Chapter 2 suggests that such market liberalisation can enhance NG's short-term competitiveness, thereby facilitating coal substitution, while also creating institutional conditions conducive to a long-term transition towards renewable energy. Chapter 4 builds on this by demonstrating how a market-based pricing system can reduce price distortions and enhance transparency through scenario-specific long-term NG projections, thereby recalibrating NG's long-term competitiveness and accelerating the shift towards renewables. This ultimately leads to a structural decline in NG demand. Drawing on the findings from both chapters, it becomes evident that a market-based pricing framework serves as a key institutional mechanism for leveraging NG as a transitional fuel in the short term, while enabling a shift towards a renewable energy–centred system in the long term—thus supporting the realisation of China's carbon neutrality objectives.

Fourth, these structural changes must be integrated with ESG policies to create a market-based framework that can drive investment in renewable energy and other sustainability initiatives. As previously discussed, strong decarbonisation policies—combining the introduction of carbon pricing mechanisms with expanded investment in renewables—are essential. In addition, the timely implementation of a market-based pricing system for NG is critical for correcting price distortions and internalising the environmental costs of fossil fuels. Together, these combined measures provide the institutional foundation necessary for the substantive achievement of carbon neutrality. Beyond these measures, however, the structural transformation of the NG market also requires an institutional framework—specifically, an ESG-focused investment architecture—that actively mobilises capital into the renewable energy sector. As demonstrated in Chapter 2, ESG factors serve as short-term drivers of NG demand. This suggests that, as Chinese firms increasingly align their investment decisions with ESG standards—reducing exposure to high-emission energy sources, such as coal, while gradually shifting capital towards renewables—NG demand is initially boosted by the coal-to-gas transition but is expected to decline over time as cleaner alternatives gain traction. To ensure the prompt and tangible effectiveness of these measures in advancing China's carbon neutrality goals, it is essential to promote the widespread adoption of ESG frameworks—particularly those focused on renewable energy investment—and to strengthen institutional support for ESG implementation within China.

Finally, Australia should respond to China's carbon neutrality targets and broader energy transition by strategically restructuring its economy towards renewable energy. As discussed in Chapter 5, renewable energy-driven economic stimulus can generate spillover effects across multiple industries and enhance economic resilience. Accordingly, Australia requires comprehensive support measures for the renewable energy sector. These include tax incentives and subsidies for renewable energy industries, investments in electric vehicles (EVs) and related infrastructure, as well as the development of renewable energy industry clusters. Such policies could catalyse new growth drivers—such as a hydrogen export industry leveraging Australia's abundant renewable resources—and contribute to a sustainable, long-term economic transformation.

6.3 Future research directions

This study provides a foundational analysis of how China's carbon neutrality goal and NG market liberalisation jointly affect the Australian economy. To ensure effective implementation of the proposed policies in both countries, it is crucial to develop concrete research and implementation plans for each policy domain, including the accelerated adoption of EVs. As demonstrated by this study's findings, achieving this requires an innovative hybrid approach that combines advanced data-driven models with traditional, validated analytical methods. This methodological synthesis enhances the precision of policy formulation and evaluation while serving as a practical tool for strategic planning in the private sector over both the short and long term. Ultimately, this integrated approach can facilitate a more effective energy transition and make a substantive contribution to the global sustainability agenda, particularly in the pursuit of carbon neutrality.

Building upon this methodological foundation, several areas warrant further investigation to deepen and extend the research. First, future studies could incorporate supply-side dynamics in Australia, such as investment behaviour in LNG infrastructure, regulatory changes, and the role of emerging hydrogen markets. A detailed assessment of domestic transition risks and sectoral adjustment costs would complement the demand-side focus of this study. Second, while this study uses scenario-based long-term projections, future work could explore integrated modelling frameworks that link energy system models with CGE or agent-based economic models. This would enable richer simulations of bilateral trade impacts and more nuanced inter-industry feedback effects. Third, with the rapid advancement of machine learning and deep learning technologies, further research could refine the methodological toolkit by applying explainable artificial intelligence techniques and large language models. These approaches would enhance interpretability and policy relevance by clarifying how key variables influence forecast outcomes. Lastly, extending the analysis to include other major LNG-importing countries—such as Japan, South Korea, and India—would help situate Australia's China-facing exposure within a broader regional and global context. This would support more comprehensive trade diversification strategies and comparative policy analysis.

Appendix

A. Terms and abbreviations

TERM	DESCRIPTION
Accum.	Accumulative
ACRI	Australia-China Relations Institute at UTS
ADF	Augmented Dickey-Fuller
AIC	Akaike Information Criterion
ALASSO	Adaptive LASSO
ANN	Artificial Neural Networks
ARMA	Auto-Regressive Moving Average
ARIMA	Auto-Regressive Integrated Moving Average
ARMAX	Auto-Regressive Moving Average with Exogenous Variable
Avg.	Average
Bbl	Barrel
BCM	Billion cubic metres
BP	British Petroleum
Brent	A global oil benchmark
Capa	Capacity
CCTD	China Coal Transport & Distribution Association
CGGP	City Gate Gas Price
CGE	Computable General Equilibrium
CLI	Composite Leading Indicator; Early signals of turning points in business cycles
CNN	Convolutional Neural Networks
CNY	Chinese yuan renminbi; The official currency of the People's Republic of China
CPI	Consumer Price Index
DCGs	Dual Carbon Goals
ECM	Error Correction Model
ECT	Error Correction Term
ELU	Exponential Linear Unit
Envir.	Environment
EIA	U.S. Energy Information Administration.
ESG	Environmental, Social, and Governance
ETL	Extract, Transform, Load
GDP	Gross Domestic Product
HH	Henry Hub; A distribution hub on the NG pipeline system in Louisiana in the U.S.

TERM	DESCRIPTION
IEA	International Energy Agency
JCC	Japan Crude Cocktail; In general, JCC, the average price of customs-cleared crude oil imports into Japan, is used by LNG traders in APAC to price LNG contracts.
KEEI	Korea Energy Economics Institute
KPSS	Kwiatkowski–Phillips–Schmidt–Shin
KT	Kiloton (1,000 tons)
LASSO	Least Absolute Shrinkage and Selection Operator
LIM	Linear Interpolation Method
LNG	Liquefied Natural Gas
LSTM	Long Short-Term Memory
LRM	Linear Regression Model
MAE	Mean Absolute Error
MCMC	Markov Chain Monte Carlo
MGMT	Management
ML	Machine Learning
MMBtu	Metric Million British Thermal Unit
MMT	Million Metric Tons
MSE	Mean Square Error
MT	Metric Ton
N/A	Not Available
NBS	National Bureau of Statistics of China
OECD Stat	Organization for Economic Cooperation and Development Statistics
OECM	One Earth Climate Model of UTS
PCA	Principal Component Analysis
PNG	Piped Natural Gas
RCP	Representative Concentration Pathways
RMSE	Root Mean Squared Error
RNN	Recurrent Neural Networks
REST	Representational State Transfer
SARIMA	Seasonal Autoregressive Integrated Moving-Average
SARIMAX	Seasonal Autoregressive Integrated Moving-Average with Exogenous Regressors
SVR	Support Vector Regression
USD	US Dollar; the currency of the United States
UTS	University of Technology Sydney
WTI	West Texas Intermediate; A global oil benchmark

B. Description of data: 40 factors

Table 2.1 Description and relevant sources of the 40 factors

Factors	Description	Data source[39]
(1) Economics and Energy (15 factors)		
Real GDP by expenditure (CNY, Billions)	Quarterly → Monthly series	Bloomberg
CPI: Overall average	Monthly series	Refinitiv; NBS
Trend restored CLI[40]	Monthly series	OECD Stat
Average city gate gas price (USD/MMBtu)	Monthly series	Bloomberg; Internal source
Chinese coking coal import price (USD/MT)	Monthly series	Bloomberg
Domestic thermal coal price: Datong (5500 kcal/kg; USD/MT)	Monthly series	Sxcoal; Internal source[41]
Domestic thermal coal price: Datong (5800 kcal/kg; USD/MT)	Monthly series	
Newcastle coal index: (6000 kcal/kg; USD/MT)	Monthly series	
Crude petroleum oil import price (USD/bbl)	Average monthly price[42]	Refinitiv
LNG average import price (USD/MMBtu)	Monthly series	Bloomberg
NG import price via pipelines (USD/MMBtu)	Monthly series	
LNG imports (Metric tonne)	Monthly series	
NG imports via pipelines (Metric tonne)	Monthly series	
Energy intensity level of primary energy (MJ / GDP measured in constant 2017 USD at Purchasing Power Party)	Monthly series	World Bank Indicator
Industry structure[43]	Yearly → Monthly series	Bloomberg
(2) Global Energy Index (5 factors)		
Brent oil price (USD/bbl)	Monthly series	Bloomberg
LNG JCC import price (USD/MMBtu)	Monthly series	
WTI oil price (USD/bbl)	Monthly series	EIA
HH gas price (USD/MMBtu)	Monthly series	
Int'l spot thermal coal price benchmark (USD/MT)	Monthly series	CCTD[44]
(3) Urbanisation and Demographics (4 factors)		
Urbanisation rate (%, Urban population of total population)	Monthly series	World Bank Indicator
Per capita income (constant 2015 USD)	Monthly series	
Population density (people per sq. km of land area)	Monthly series	
Population (Millions)	Monthly series	Bloomberg
(4) Infrastructure (4 factors)		
Accumulated LNG terminal capa. per year (10KT)	Monthly series	KEEI
Gas pipeline capacity (BCM)	Monthly series	External source[45]
City public utilities: Length of gas pipelines (10KT)	Monthly series	NBS
Pipeline transportation: Length of gas pipelines (10KT)	Monthly series	CEIC[46]
(5) Environment, Climate, and Technological Growth (7 factors)		
[ESG] Environmental score	Yearly → Monthly series	Refinitiv
[ESG] Social score	Yearly → Monthly series	
[ESG] Governance score	Yearly → Monthly series	
[ESG] Country score for climate change (%)	Yearly → Monthly series	
Investment % of Chg Y/Y: Water conservancy, Envir. & Public utilities mgmt.	Monthly series	
Average national temperature (°C, Celsius)[47]	Monthly series	NBS
Technology and Innovation: Envir.-related patents of total patents (%)	Yearly → Monthly series	OECD Stats
(6) Price Distortion (5 factors)		
Price distortion (City gate gas price / LNG import price, %)	Monthly series	Bloomberg, Internal source, and estimation[48]
Price distortion (City gate gas price / Piped gas price, %)	Monthly series	
Price distortion (City gate gas price / Average of LNG import price and Piped gas price, %)	Monthly series	
Price distortion (City gate gas price / JCC, %)	Monthly series	
Price distortion (City gate gas price / HH, %)	Monthly series	

[39] Bloomberg and Refinitiv, global economic and financial data providers, are accessed through UTS Business School; The internal source is provided by UTS ACRI; Sxcoal, a Chinese coal data provider, is referenced from http://www.sxcoal.com/site/about-us/en.
[40] The trend restored CLI provides early signals of turning points in business cycles, reflecting economic trends such as GDP. This index is calculated by multiplying the raw CLI, which varies around 100, by the trend of the reference series such as GDP [243].
[41] Currently, China imports more coking coal than it domestically produces because coking coal, in contrast to thermal coal, is relatively scarce [244]. Thus, the domestic coking coal price is not factored into data collection, whereas both domestic and import prices of thermal coal are considered.
[42] Crude Petroleum Oil Import Price is the average monthly price calculated as Total monthly price / Total monthly volume.
[43] Industry structure is calculated as Secondary Industry's real GDP / Total real GDP. Secondary industry includes construction and manufacturing, both of which consume a significant amount of energy. This calculation is suggested by Wang and Li [21].
[44] CCTD is a Chinese coal data provider, referenced from http://www.cctdcoal.com/.
[45] This reference comes from Chen et al. [110].
[46] CEIC is a global economic, industry, and financial data provider.
[47] Average national temperature is assumed to represent overall NG consumption patterns across China.
[48] Price distortion factors are calculated using the formula: '(City gate gas price / each energy price - 1) ×100'.

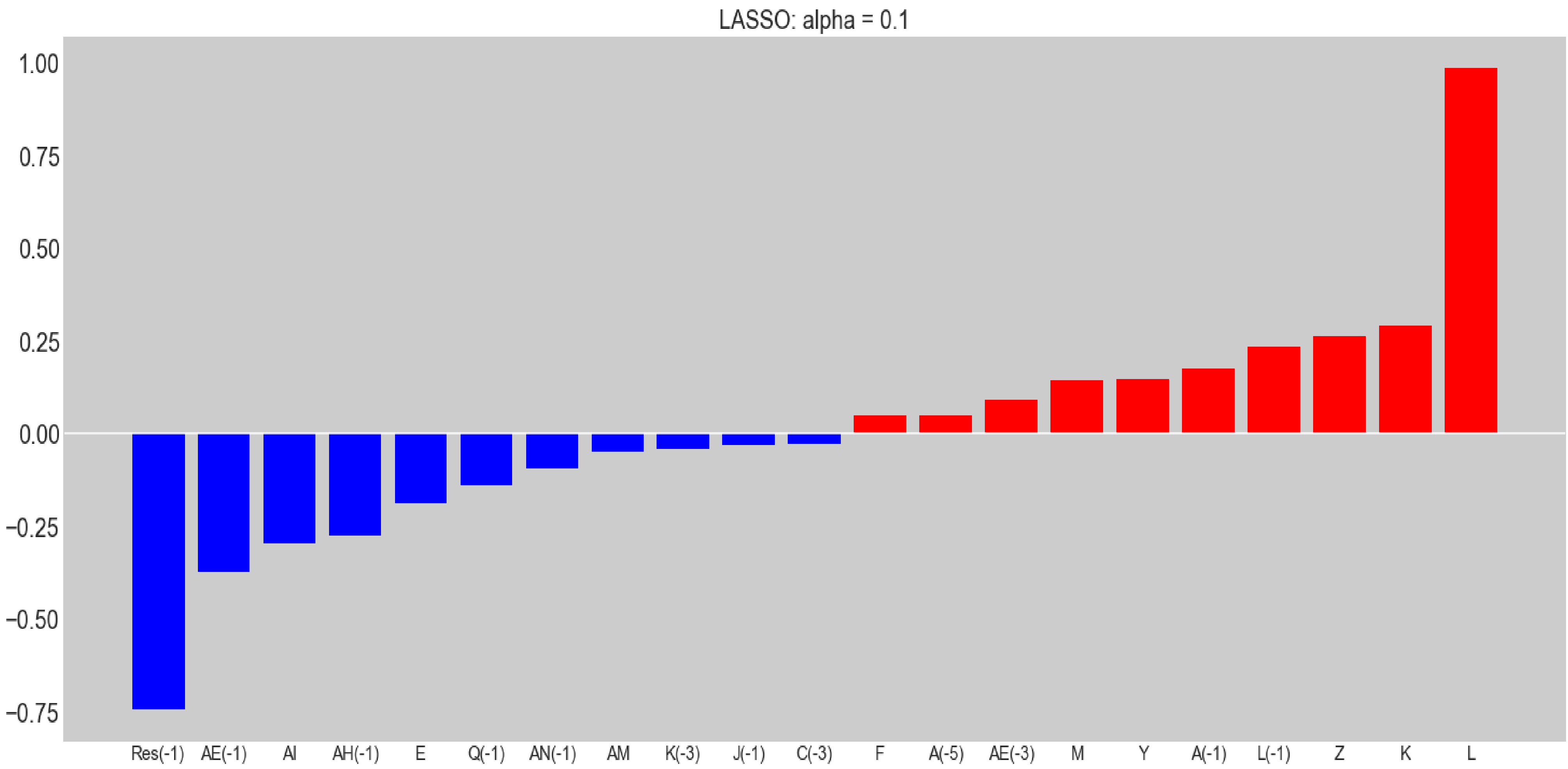
C. Lagged stationary variables

Table 2.2 Lagged stationary variables of 40 factors

		Factors	Stationary status	Lag order by AIC	Num. of lagged variables
1	A	Real GDP	Log transform & 2nd difference***	6	7
2	B	CPI	1st difference*	1	2
3	C	Trend restored CLI	Log transform & 2nd difference***	3	4
4	D	Average city gate gas price	1st difference***	0	1
5	E	Chinese coking coal import price	1st difference***	1	2
6	F	Domestic thermal coal price: Datong - 5500 kcal/kg	1st difference***	6	7
7	G	Domestic thermal coal price: Datong - 5800 kcal/kg	1st difference***	3	4
8	H	Newcastle coal index: 6600 kcal/kg	1st difference***	1	2
9	I	Crude petroleum oil import	1st difference***	1	2
10	J	LNG average import price	1st difference***	0	1
11	K	NG import price via pipelines	1st difference***	3	4
12	L	LNG imports	1st difference***	4	5
13	M	NG imports via pipelines	1st difference***	2	3
14	N	Industry structure	2nd difference***	5	6
15	O	Energy intensity level of primary energy	2nd difference***	0	1
16	P	Brent oil price	1st difference***	1	2
17	Q	WTI oil price	1st difference***	1	2
18	R	HH gas price	1st difference***	1	2
19	S	Int'l spot thermal coal price benchmark	1st difference***	1	2
20	T	LNG JCC import price	1st difference***	1	2
21	U	Urbanisation rate	Non-difference**	2	3
22	V	Per capita income	Log transform & 2nd difference***	0	1
23	W	Population	Log transform & 2nd difference***	0	1
24	X	Population density	Log transform & 2nd difference***	0	1
25	Y	Accumulated LNG terminal capa. per year	1st difference***	0	1
26	Z	Gas pipeline capacity	1st difference***	0	1
27	AA	City public utilities: Length of gas pipelines	Log transform & 2nd difference***	0	1
28	AB	Pipeline transportation: Length of gas pipeline	Log transform & 2nd difference***	0	1
29	AC	Water conservancy, Envir. & Public utilities mgmt.	2nd difference***	6	7
30	AD	Envir.-related patents of total patents	1st difference***	0	1
31	AE	[ESG] Environmental score	2nd difference***	5	6
32	AF	[ESG] Social score	2nd difference***	5	6
33	AG	[ESG] Governance score	2nd difference***	5	6
34	AH	Average national temperature	1st difference***	6	7
35	AI	[ESG] Country score for climate change	1st difference*	0	1
36	AJ	Price distortion (City gate gas price / LNG import price)	1st difference***	0	1
37	AK	Price distortion (City gate gas price / Piped gas price)	1st difference***	1	2
38	AL	Price distortion (City gate gas price / Average of LNG import price and Piped gas price)	1st difference***	5	2
39	AM	Price distortion (City gate gas price / JCC)	Non-difference*	2	3
40	AN	Price distortion (City gate gas price / HH)	1st difference***	1	2
41		[Dependent variable] NG consumption	1st difference***	2	2
		Total number of lagged stationary variables			**117**

Notes: ***, **, and * indicate statistical significance of unit root tests at the 1%, 5%, and 10% levels, respectively. The actual optimal lag order of the factor (AL) is 5, but a lag order of 1 is practically use due to a limit of available data. In this study, an alphabetical assignment is attributed to each facto

D. LASSO results in LASSO-ECM2

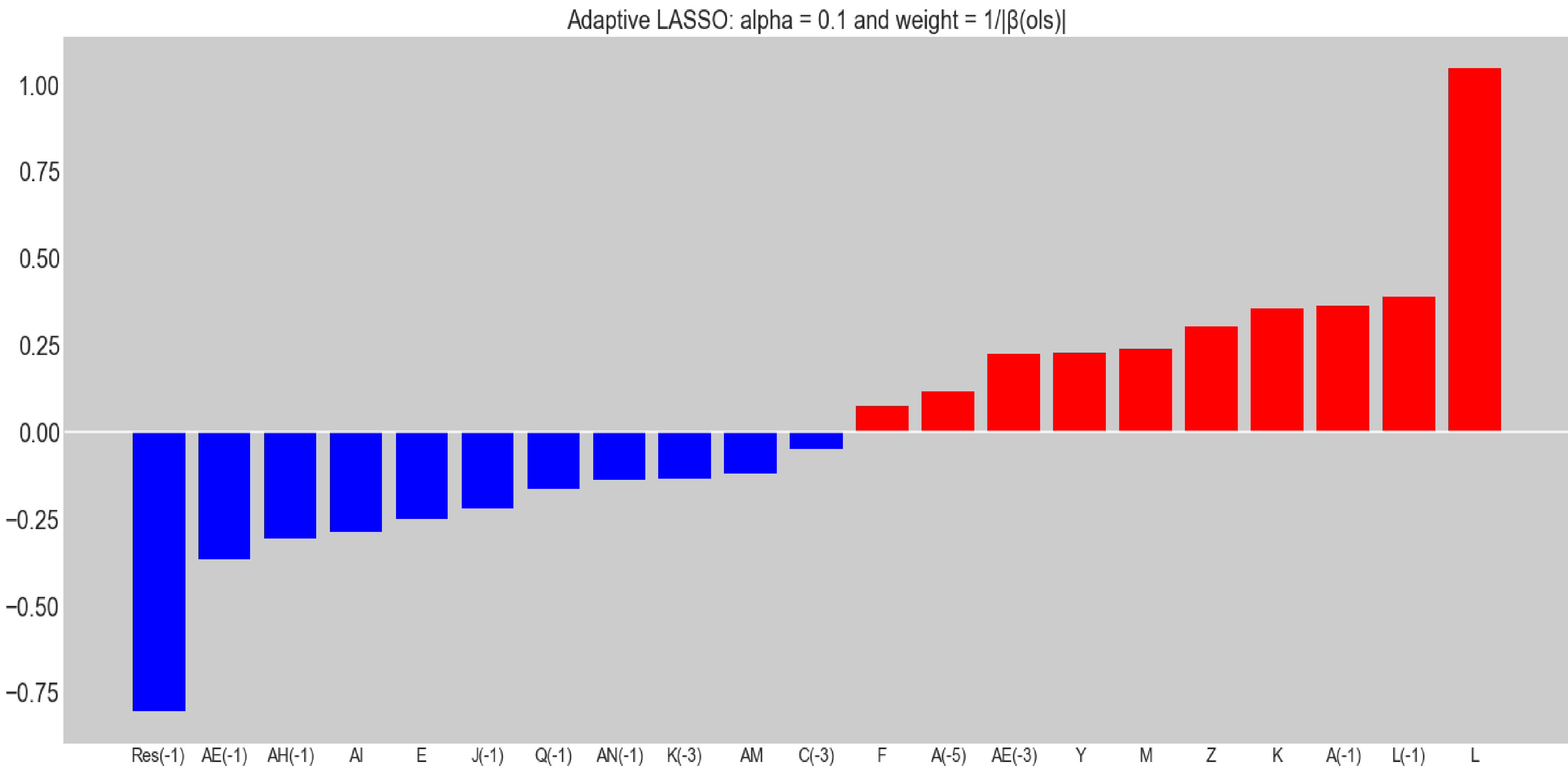


21 Coefficients selected by LASSO																				
Res(-1)	AE(-1)	AI	AH(-1)	E	Q(-1)	AN(-1)	AM	K(-3)	J(-1)	C(-3)	F	A(-5)	AE(-3)	M	Y	A(-1)	L(-1)	Z	K	L
- 0.7424	- 0.3733	- 0.2958	-0.2738	- 0.1889	- 0.1395	- 0.0925	- 0.0467	- 0.0414	- 0.0314	- 0.0286	0.0492	0.0498	0.0906	0.143	0.1487	0.1742	0.2363	0.2614	0.2905	0.9852

Figure 2.9 LASSO results (α = 0.1) in LASSO-ECM2: Selected comprehensive short-run variables

Notes: Table 2.2 shows notations denoting relevant factors. These coefficients are rounded to four decimal places. Figure 2.9 displays only variables with non-zero coefficients, and its y-axis shows coefficient values.

E. ALASSO results in LASSO-ECM2



21 Coefficients selected by ALASSO																				
Res(-1)	AE(-1)	AH(-1)	AI	E	J(-1)	Q(-1)	AN(-1)	K(-3)	AM	C(-3)	F	A(-5)	AE(-3)	Y	M	Z	K	A(-1)	L(-1)	L
-0.8028	-0.3673	-0.3060	-0.2874	-0.2492	-0.2209	-0.1664	-0.1396	-0.1339	-0.1217	-0.0489	0.0731	0.1146	0.2230	0.2288	0.2383	0.3039	0.3565	0.3627	0.3891	1.0455

Figure 2.10 ALASSO results (α = 0.1 and $\hat{w}_j = 1/|\hat{\beta}_j(ols)|$) in LASSO-ECM2: Selected comprehensive short-run variables

Notes: Table 2.2 shows notations denoting relevant factors. These coefficients are rounded to four decimal places. Figure 2.10 displays only variables with non-zero coefficients, and its y-axis shows coefficient values.

.

F. ADF test result

Table 3.2 ADF test result of independent and dependent variables

Factor	At level		At the first difference	
	ADF statistic	P-value	ADF statistic	P-value
Domestic thermal coal price	-2.1282	0.2333	-11.1934	0.0000***
LNG imports	-1.1527	0.6936	-1.9802	0.0456**
NG imports via pipelines	-1.4232	0.5711	-13.3244	0.0000***
HH gas price	0.1180	0.9673	-6.5136	0.0000***
Accum. LNG terminal capa. per year	0.6752	0.8621	-13.1604	0.0000***
Gas pipeline capacity	-2.4997	0.1156	-13.2892	0.0000***
Average national temperature	-1.6946	0.4339	-11.9628	0.0000***
Price distortion	-3.8151	0.0027***	-14.8039	0.0000***
NG consumption	0.9819	0.9941	-13.4945	0.0000***

Notes: ***, **, and * indicate ADF statistical significance at the 1%, 5%, and 10% levels, respectively. The ADF test is conducted using the Bayesian Information Criterion (BIC). All numerical values are rounded to four decimal places.

G. Ljung-Box and Jarque-Bera test results

Table 3.3 Ljung-Box Q and Jarque-Bera test results

Tests	Models	Results	Description
Ljung-Box	SARIMAX 1	P-value: 0.990 (Statistic: 0.000)	Based on the significance level of 0.01 and the p-value of Ljung-Box, H_0 cannot be rejected. This implies that the model effectively captures the temporal dependencies in the data (residuals), indicating that the model is well-specified in terms of autocorrection.
	SARIMAX 2	P-value: 0.970 (Statistic: 0.000)	
	Hybrid 2-S2	P-value: 0.115 (Statistic: 15.496)	
Jarque-Bera	SARIMAX 1	P-value: 0.360 (Statistic: 2.060)	Based on the significance level of 0.01 and the p-value of Jarque-Bera, H_0 cannot be rejected. This indicates that the data (residuals) can be considered normally distributed, suggesting that the model satisfies the normality assumption for residuals.
	SARIMAX 2	P-value: 0.340 (Statistic: 2.160)	
	Hybrid 2-S2	P-value: 0.011 (Statistic: 9.026)	

Notes: The number of lags used in the Ljung-Box test is the default value of 10; In the Ljung-Box test, H_0 (null hypothesis) represents that there is no autocorrelation in the residuals of the time series, which means that the data behaves like white noise, while H_1 (alternative hypothesis) represents that there is significant autocorrelation in the residuals, suggesting that the time series is not white noise; In the Jarque-Bera test, H_0 (null hypothesis) indicates that the data, the residuals of the time series, follow a normal distribution, while H_1 (alternative hypothesis) suggests that they do not follow a normal distribution.

H. Recurrent Neural Network (RNN)

Recurrent Neural Networks (RNNs) exhibit structures resembling chains of repeating modules, designed to leverage these modules as a memory for storing critical information from preceding processing steps [97, 99]. In essence, RNNs rely on the preceding elements within the sequence. They inherently possess a 'memory', incorporating information from prior inputs to influence both the current input and output. This can be conceptualised as a hidden layer retaining information over time. The structure of the RNN is shown in Figure 3.12.

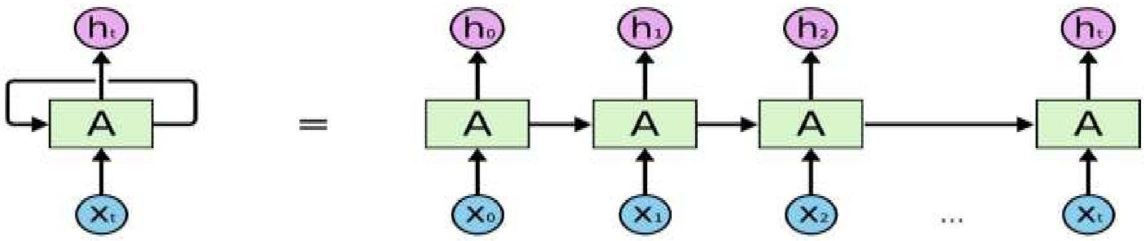


Figure 3.12 The structure of RNN

I. Long Short-Term Memory (LSTM)

The Long Short-Term Memory (LSTM) model is a distinct type of RNN specifically designed to capture long-term dependencies [97, 103]. The LSTM is structured in the form of a chain; however, its repeating module differs from that of a standard RNN. Unlike a typical RNN, which has a single neural network, the LSTM features four interacting layers with a distinctive flow method, as illustrated in Figure 3.13. Without delving into the details inside the box, one can observe that the state of the LSTM cell is divided into two vectors: h_t and c_t. Here, h_t represents the short-term output from the previous block, while c_t signifies the long-term memory from the previous block.

When examining the contents inside the box, the core idea of the operational principle becomes evident: the network learns what to store, discard, and read from the long-term state. The long-term memory, c_{t-1}, traverses the network from left to right. During this journey, it passes through (a) the Forget gate, losing some memories. The summation operation incorporates newly selected memories from (b) the Input gate. The resulting c_t is sent directly to the output without further transformations. Due to this operational principle, some memories are deleted, and new ones are added at each time step. Additionally, after the summation operation, the long-term state is copied and passed through the tanh function. Subsequently, this result is filtered by (c) the Output gate. This process creates the short-term state h_t, which is equivalent to the cell's output y_t at this time step.

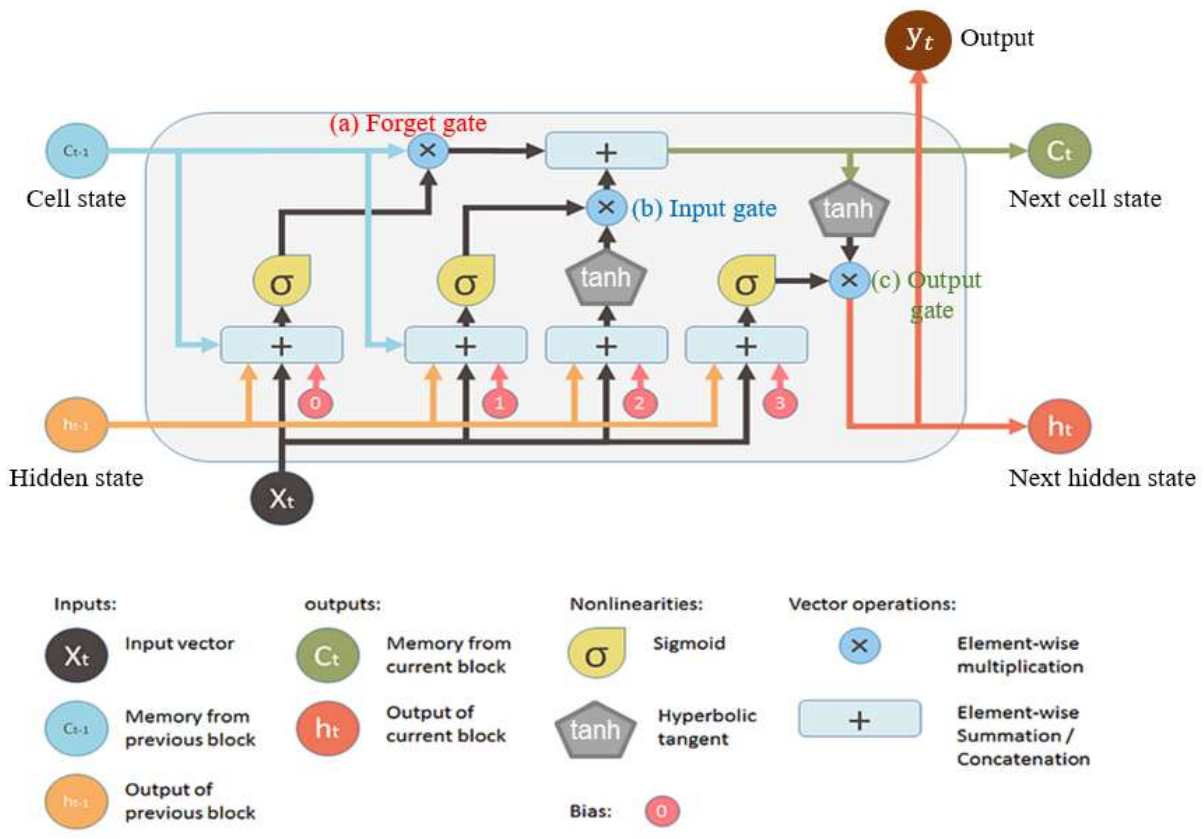


Figure 3.13 The structure of LSTM

Note: This figure is derived from Le et al [97]

J. Data source and property: Tables 4.1 and 4.2

Table 4.1 Key driving factors and LASSO results

	Key driving factors	Description	Data source[49]
	(1) Energy (3 factors)		
DT	Domestic thermal coal price (5500 kcal/kg; USD/MT)	Monthly series	Bloomberg
LI	LNG imports (Metric tonne)	Monthly series	
NI	NG imports via pipelines (Metric tonne)	Monthly series	
	(2) Global Energy Index (1 factor)		
HG	HH gas price (USD/MMBtu)	Monthly series	Bloomberg
	(3) Infrastructure (2 factors)		
AL	Accumulative LNG terminal capa. per year (10KT)	Monthly series	KEEI
GP	Gas pipeline capacity (BCM)	Monthly series	External source[50]
	(4) Climate (1 factor)		
AN	Average national temperature (°C, Celsius)[51]	Monthly series	Bloomberg
	(5) Price distortion (1 factor)		
PD	Price distortion (CGGP[52] / PNG price[53], %)	Monthly series	Internal source[54], Bloomberg, and, Estimation[55]

Notes: The monthly dependent variable, Chinese NG consumption (BCM), is sourced from Bloomberg and Refinitiv.

Model	LASSO results: Coefficients							
	AN	PD	HG	DT	NI	GP	AL	LI
LASSO	- 0.829377	- 0.381393	- 0.291971	0.139179	0.345636	0.430412	0.672438	2.758738

Notes: AN, PD, HG, DT, NI, GP, AL, and LI are acronyms for the key driving factors.

Table 4.2 ADF test result of independent and dependent variables

Factor	At level		At the first difference	
	ADF statistic	P-value	ADF statistic	P-value
Domestic thermal coal price	-2.1282	0.2333	-11.1934	0.0000***
LNG imports	-1.1527	0.6936	-1.9802	0.0456**
NG imports via pipelines	-1.4232	0.5711	-13.3244	0.0000***
HH gas price	0.1180	0.9673	-6.5136	0.0000***
Accum LNG terminal capa. per year	0.6752	0.8621	-13.1604	0.0000***
Gas pipeline capacity	-2.4997	0.1156	-13.2892	0.0000***
Average national temperature	-1.6946	0.4339	-11.9628	0.0000***
Price distortion	-3.8151	0.0027***	-14.8039	0.0000***
NG consumption	0.9819	0.9941	-13.4945	0.0000***

Notes: ***, **, and * indicate ADF statistical significance at the 1%, 5%, and 10% levels, respectively. The ADF test is conducted using the Bayesian Information Criterion (BIC). All numerical values are rounded to four decimal places.

K. Climate and energy scenarios

Table 4.3 Key global climate and energy scenario comparison

IEA	BP Energy	RCP
STEPS: Stated policies scenario	New Momentum	RCP 8.5: Business as usual
APS: Announced pledges scenario	Accelerated	RCP 6.0: Moderate emissions
NZE: Net zero emissions by 2050 scenario	Net Zero	RCP 4.5: Low emissions

49 The monthly exchange rates, referenced from Bloomberg, are applied.
50 This reference is taken from Chen et al.'s research [110].
51 This factor is calculated as the average temperature of all provinces per month.
52 The CGGP from December 2005 to June 2013 is based on Paltsev and Zhang's research [36]. From July 2013 to December 2019, it comes from an internal source. From January 2020, the price is assumed to remain at the December 2019 level.
53 The piped gas price is referenced from Bloomberg.
54 The internal source is provided by UTS ACRI.
55 Price distortion variables are calculated using the formula: '(CGGP / PNG price - 1) ×100'.

L. Prophet-based projection of six key drivers

Table 4.10 Prophet-based projection of six key drivers

<table>
<tr><th colspan="2">Key driver (Scenario)</th><th>Prophet modelling: Inputs for model fitting and projections & Parameters for Prophet modelling</th></tr>
<tr><td colspan="2">Coal price (IEA)</td><td>Inputs for model fitting
∘ Historical monthly data; Escalation rate of 1; and Average yearly prices (Caps)
Inputs for projection
∘ Future average yearly prices (Caps) calculated based on CAGRs (from China's steam coal price projections in Table 4.4) for 2030 to 2060 across different scenarios
∘ Escalation rates calculated based on the CAGRs for 2030 to 2060 across different scenarios
Parameters for Prophet modelling
∘ Growth is set as 'logistic'; Yearly_seasonality to 'True'; Seasonality_mode as 'additive'; MCMC_samples to '500'; Interval_width at '0.95'; and Escalation rate included as an additional regressor with the multiplicative mode[56]</td></tr>
<tr><td colspan="2">HH gas price (IEA)</td><td>Inputs for model fitting
∘ Historical monthly data; Escalation rate of 1; and Average yearly price (Caps)
Inputs for projection
∘ Future average monthly prices (Caps) calculated based on CAGRs (from the U.S. NG price projections in Table 4.4) for 2030 to 2060 across different scenarios
∘ Escalation rates calculated based on the CAGRs for 2030 to 2060 across different scenarios
Parameters for Prophet modelling
∘ The settings are the same as those for coal price</td></tr>
<tr><td rowspan="2">Price distortion</td><td>PNG price (IEA)</td><td>Inputs for model fitting
∘ Historical monthly data; Escalation rate of 1; and Average yearly prices (Caps)
Inputs for projection
∘ Future average monthly prices (Caps) calculated based on CAGRs (from China's NG price projections in Table 4.4) for 2030 to 2060 across different scenarios
∘ Escalation rates calculated based on the CAGRs for 2030 to 2060 across different scenarios
Parameters for Prophet modelling
∘ The settings are the same as those for coal price</td></tr>
<tr><td>CGGP (IEA)</td><td>Inputs for model fitting
∘ Historical monthly data; Average yearly Brent oil and HH gas prices (Regressors)
Inputs for projection
∘ Future average yearly HH gas and Brent oil prices calculated based on CAGRs (from the U.S. NG and crude oil price projections in Table 4.4, respectively) for 2030 to 2060 across different scenarios
Parameters for Prophet modelling
∘ Yearly_seasonality is set to 'False'; and Brent oil and HH gas prices are included as additional regressors with the additive mode[57]</td></tr>
<tr><td colspan="2">PNG imports (BP Energy)</td><td>Inputs for model fitting
∘ Historical monthly data; Escalation rate of 1; and Average monthly prices (Caps)
Inputs for projection
∘ Future average monthly imports (Caps) estimated based on scenario-specific PNG import projections (based on CAGRs in Table 4.6) for 2025 to 2060
∘ Escalation rates calculated based on the CAGRs, derived from the future PNG import projections.
Parameters for Prophet modelling
∘ The settings are the same as those for coal price</td></tr>
<tr><td colspan="2">LNG imports (BP Energy)</td><td>Inputs for model fitting
∘ Historical monthly data; Escalation rate of 1; and Average monthly imports (Caps)
Inputs for projection
∘ Future average monthly imports (Caps) calculated based on CAGRs (in Table 4.5) for 2030 to 2060 across different scenarios
∘ Escalation rates calculated based on the CAGRs for 2030 to 2060 across different scenarios
Parameters for Prophet modelling
∘ The settings are the same as those for coal price</td></tr>
</table>

[56] The model's output scales proportionally with the escalation rate. Please refer to Section 4.5.2.2 for details on MCMC samples.
[57] Due to the strong policy influence of CGGP, the seasonality feature has been turned off. Additionally, following the formula introduced by Paltsev et al. [36] for CGGP, it is assumed that Brent oil and HH gas prices consistently impact the CGGP.

Temperature (RCP)	*Inputs for model fitting* ∘ Historical monthly data; Escalation rate of 1; and Average monthly temperature data (Caps) *Inputs for projection* ∘ Future average monthly temperature data (Caps) calculated based on CAGRs (in Table 4.9) for 2040 and 2070 across different scenarios ∘ Escalation rates calculated based on the CAGRs for 2040 to 2070 across different scenarios *Parameters for Prophet modelling* ∘ The settings are the same as those for coal price	

Notes: CGGP and PNG prices, necessary for the calculation of price distortion, have been projected by the Prophet model for each.[58] For the projection of CGGP, it is assumed that CGGP are calculated based on Brent oil and HH gas prices [36]. All the variables of Brent oil and HH gas prices consist of time series data on a monthly basis, spanning from April 2009 to July. Brent oil prices are sourced from the EIA, and HH gas prices are sourced from Bloomberg.

M. Stepwise projection of two natural gas infrastructure-related factors

Table 4.11 Stepwise projection of two natural gas infrastructure-related factors

Key driver (Scenario)	Stepwise projection details
Accumulated LNG terminal capacity (BP Energy)	∘ As shown in Table 4.7, the phased increases in the accumulated LNG terminal capacity for each year in 2025 and 2035 are applied uniformly across all scenarios. ∘ It is assumed that the accumulated LNG terminal capacity is projected to decrease based on the CAGRs (in Table 4.7) from 2035 to 2060.
Gas pipeline capacity (BP Energy)	∘ As illustrated in Table 4.8, the gas pipeline capacity is anticipated to increase incrementally by 50 BCM by 2030, leading to an accumulative gas pipeline capacity of 195 BCM by that year. ∘ As seen in Table 4.8, in the New Momentum and Accelerated scenarios, it is assumed that the gas pipeline capacity is projected to increase step by step until 2040 and decrease based on the CAGRs of PNG imports (in Table 4.6) by 2060. ∘ As indicated in Table 4.8, in the Net zero scenario, it is assumed that the gas pipeline capacity projections increase by 2030 and decrease based on the CAGRs of PNG imports (in Table 4.6) from 2040 to 2060.

[58] Please see Appendix J for the price distortion formula.

N. Scenario-specific Linear Interpolation Method modelling and projection

Table 4.12 Scenario-specific Linear Interpolation Method modelling and projection

Scenario-specific LIM modelling

- The slope of LIM is calculated using the values of each factor for two periods presented in scenarios by IEA, BP, etc. (e.g., 2022 vs. 2030): the slope is calculated as "Change in value / Change in period".
- The slope is used to estimate monthly values of each scenario-specific factor using each month as an input.
- The lower and upper bounds of each factor are projected using the 95% confidence intervals for the coefficients (slope and intercept), calculated through the Excel tools - data analysis (regression).

Scenario-specific LIM projection

Key Drivers	Details in LIM
Coal Price	◦ Inputs for LIM: Scenario-specific coal prices for 2022, 2030, and 2050 (in Table 4.4–China's steam coal price)
PNG imports	◦ Inputs for LIM: Scenario-specific PNG imports for 2020, 2025, 2030, 2040, and 2050 (in Table 4.6)
LNG Imports	◦ Inputs for LIM: Scenario-specific LNG imports for 2019, 2030, and 2050 (in Table 4.5)
Temperature	◦ Inputs for LIM: Scenario-specific temperature data for 2018, 2040, and 2070 (in Table 4.9) ◦ Seasonality is not applied in the LIM modelling
HH gas price	◦ Inputs for LIM: Scenario-specific US NG prices for 2022, 2030, and 2050 (in Table 4.4–the U.S. NG price)
Price distortion	*Price distortion: CGGP* (②) */ PNG price* (①) *-1* ◦ Price distortion is calculated by subtracting one from the ratio of CGGP to PNG price. *PNG price* (①) ◦ Inputs for LIM: Scenario-specific PNG prices for 2022, 2030, 2025, and 2050 (in Table 4.4–China's NG price) *CGGP* (②) Inputs ◦ HH gas price: Scenario-specific US NG prices for 2022, 2030, and 2050 (in Table 4.4–the U.S. NG price) ◦ Brent oil price: Scenario-specific crude oil prices for 2022, 2030, and 2050 (in Table 4.4–Crude oil price) Modelling process: Linear interpolation and Multi regression method ◦ Calculating future monthly Brent oil and HH gas prices based on the linear interpolation method. ◦ Calculating future CGGP using the multi regression method with future monthly Brent oil and HH gas prices.

Notes: Scenario-specific values for coal price, PNG imports, LNG imports, Temperature, CGGP and PNG prices are provided in Tables 4.4 to 4.9.

O. Ljung-Box test

Table 4.13 Ljung-Box test results for SARIMAX

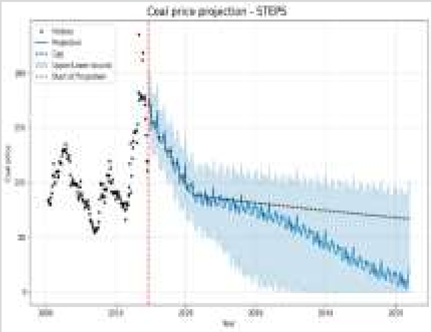
Model		Description
SARIMAX	P-value: 0.72 (Statistic: 0.13)	Based on the significance level of 0.01 and the p-value of Ljung-Box, H_0 cannot be rejected. This implies that the model effectively captures the temporal dependencies in the data, indicating a good fit.

Notes: In the Ljung-Box test, H_0 (null hypothesis) represents that there is no autocorrelation in the residuals of the time series, which means that the data behaves like white noise, while H_1 (alternative hypothesis) represents that there is significant autocorrelation in the residuals, suggesting that the time series is not white noise.

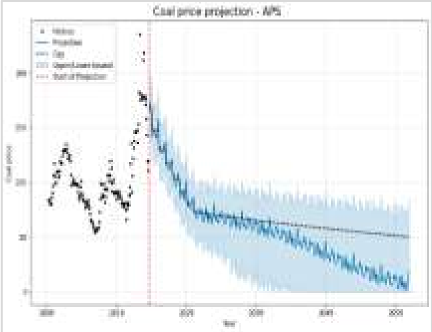
P. Projections of key drivers

<u>Prophet-based projections: Coal price, PNG imports, LNG imports, Temperature, HH gas price, and Price distortion (with CGGP and PNG price)</u>

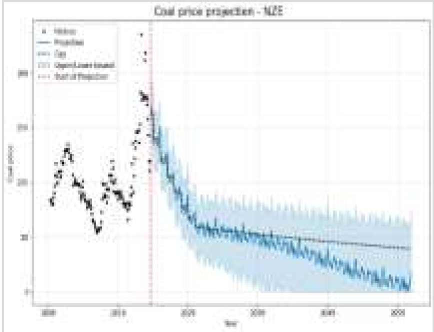
Coal price (USD/MT): STEPS



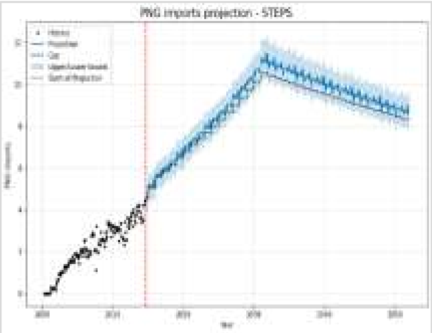
Coal price (USD/MT): APS



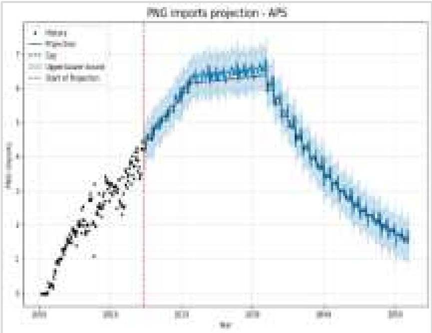
Coal price (USD/MT): NZE



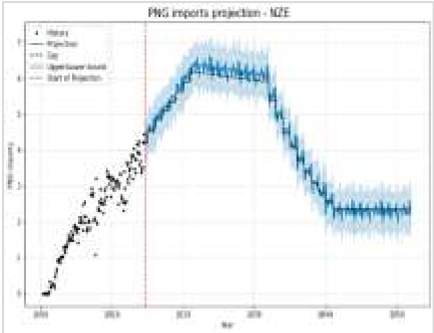
PNG imports (MMT): STEPS



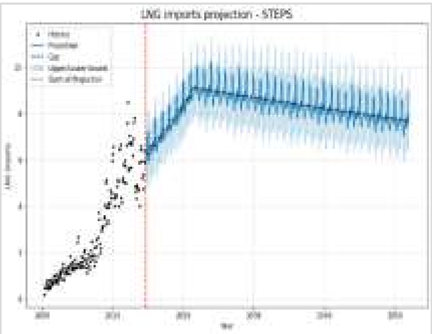
PNG imports (MMT): APS



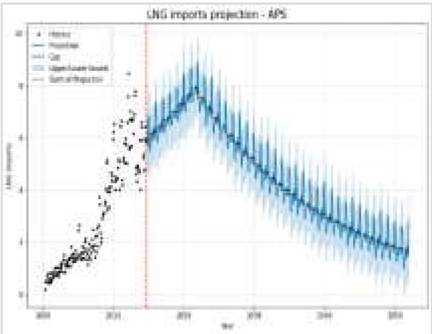
PNG imports (MMT): NZE



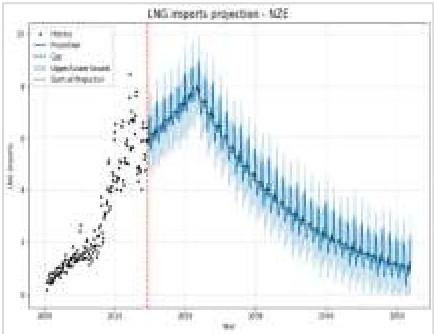
LNG imports (MMT): STEPS



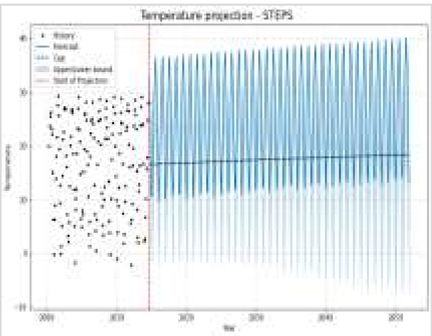
LNG imports (MMT): APS



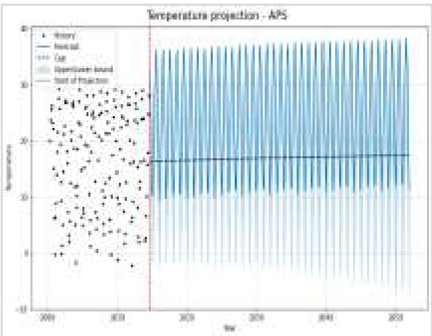
LNG imports (MMT): NZE



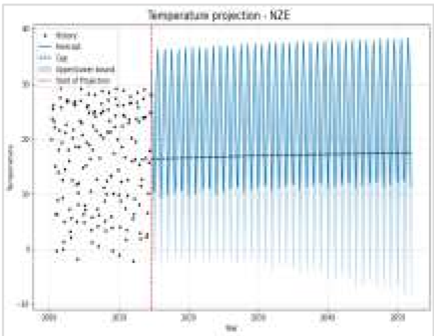
Avg. Temperature (°C): STEPS



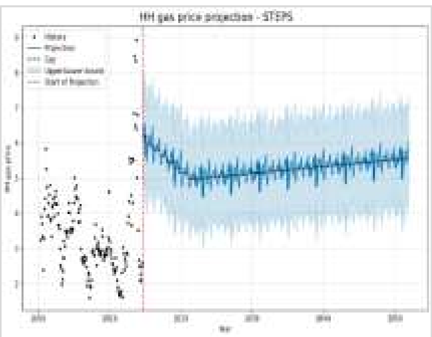
Avg. Temperature (°C): APS



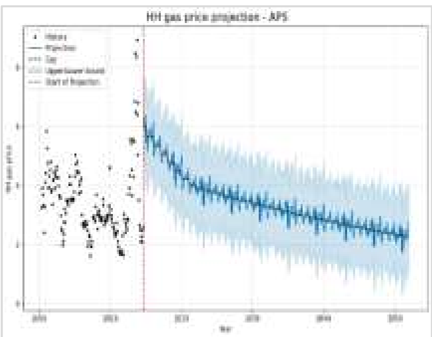
Avg. Temperature (°C): NZE



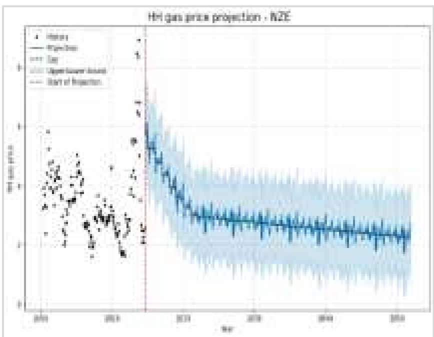
HH gas price (USD/MMBTU): STEPS



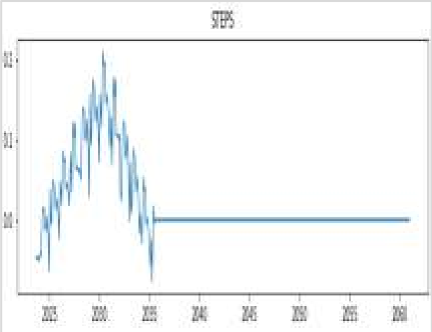
HH gas price (USD/MMBTU): APS



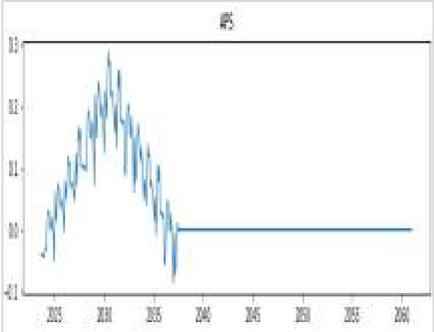
HH gas price (USD/MMBTU): NZE



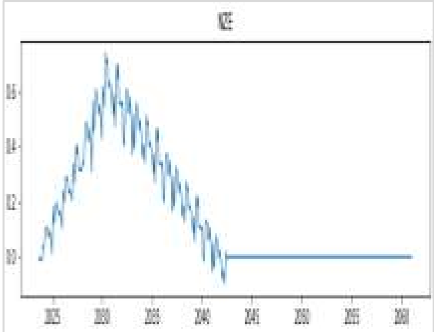
Price distortion: STEPS



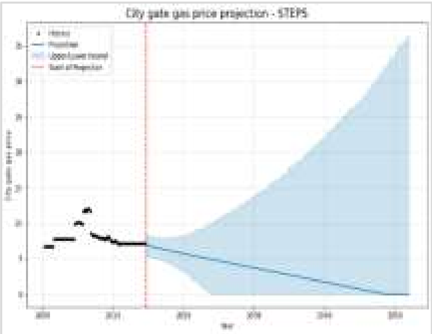
Price distortion: APS



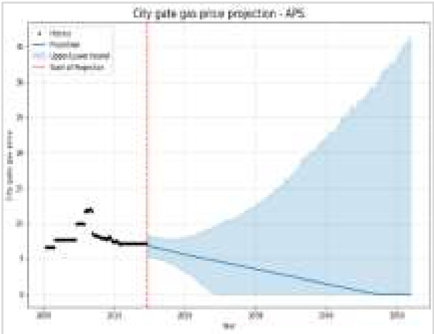
Price distortion: NZE



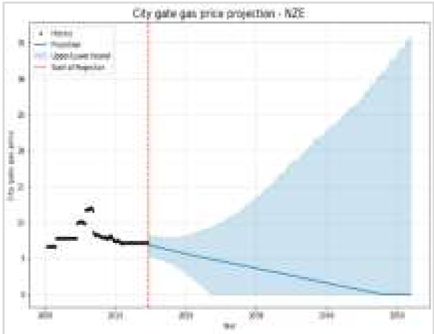
CGGP (USD/MMBTU): STEPS



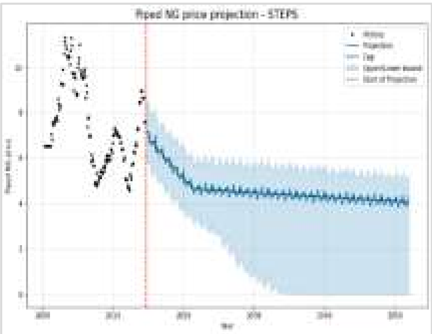
CGGP (USD/MMBTU): APS



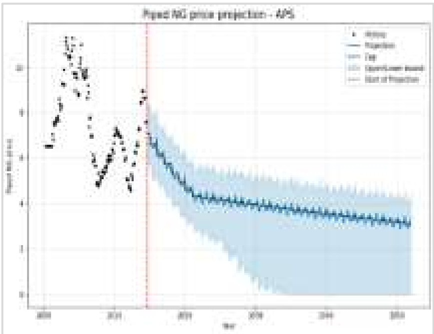
CGGP (USD/MMBTU): NZE



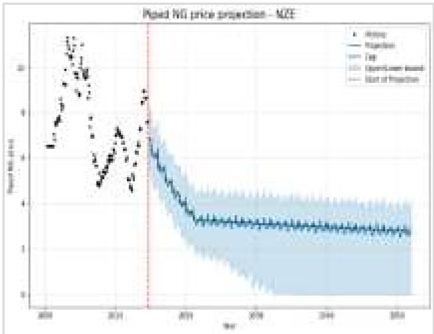
PNG price (USD/MMBTU): STEPS



PNG price (USD/MMBTU): APS



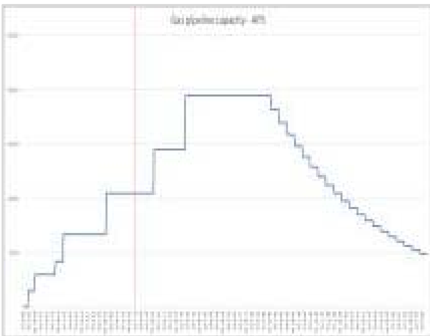
PNG price (USD/MMBTU): NZE



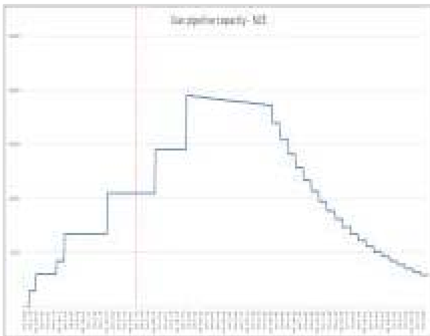
Stepwise projections of two NG infrastructure-related factors, Gas pipeline capacity and Accumulated LNG terminal capacity

Gas pipeline capa (BCM): STEPS

Gas pipeline capa (BCM): APS

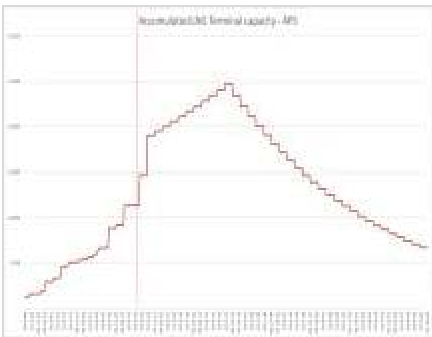


Gas pipeline capa (BCM): NZE



Accum LNG terminal capa (10kt): STEPS

Accum LNG terminal capa (10kt): APS



Accum LNG terminal capa (10kt): NZE

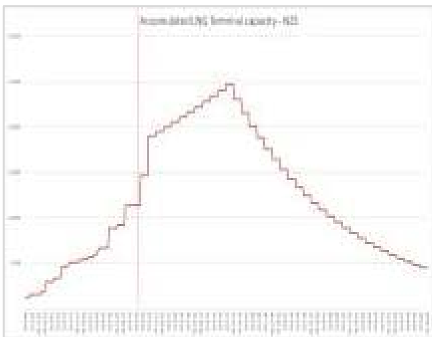
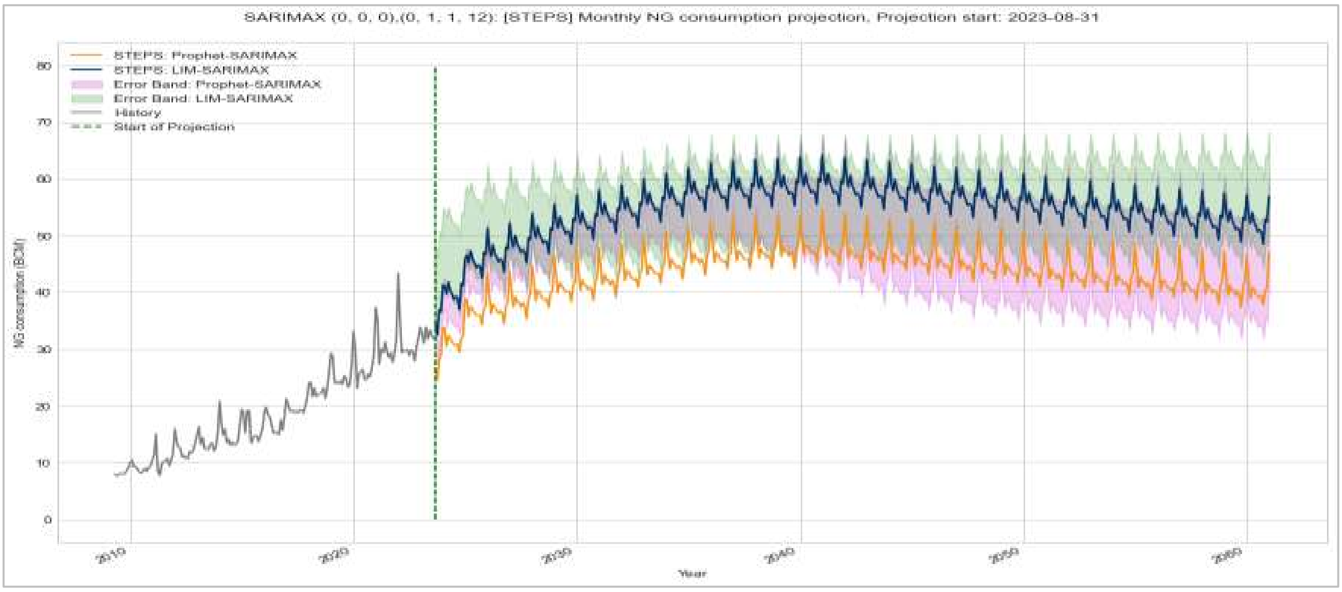


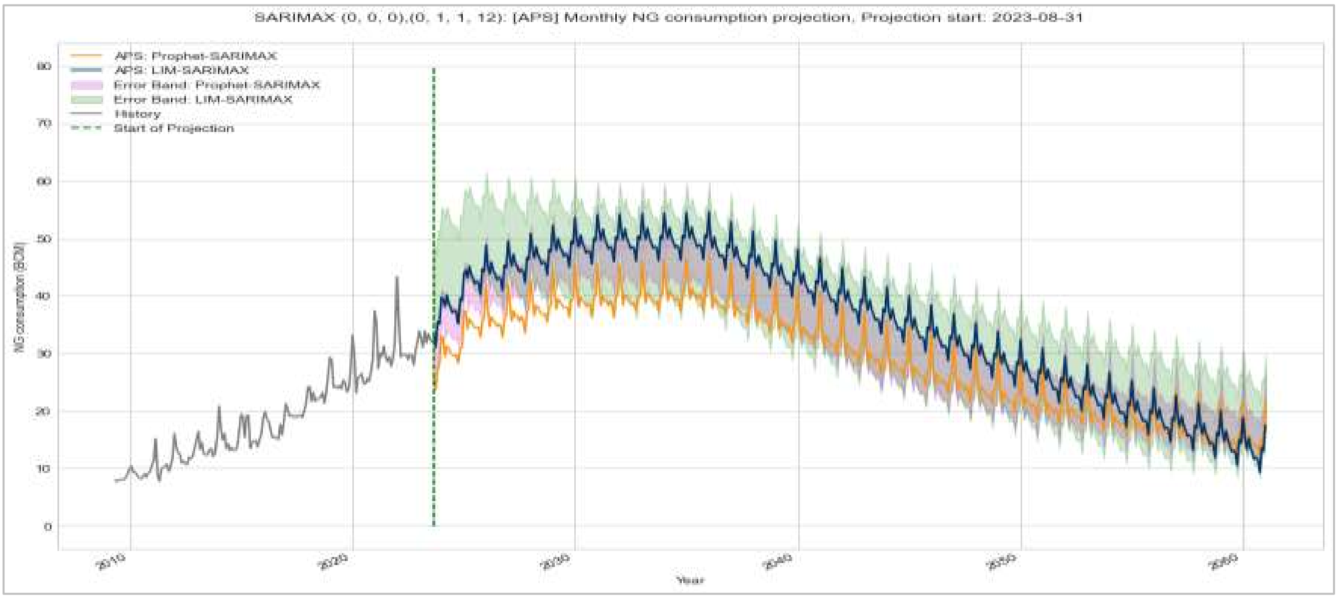
Figure 4.6 Projections of key drivers

Q. Monthly projection and error band

STEPS



APS



NZE

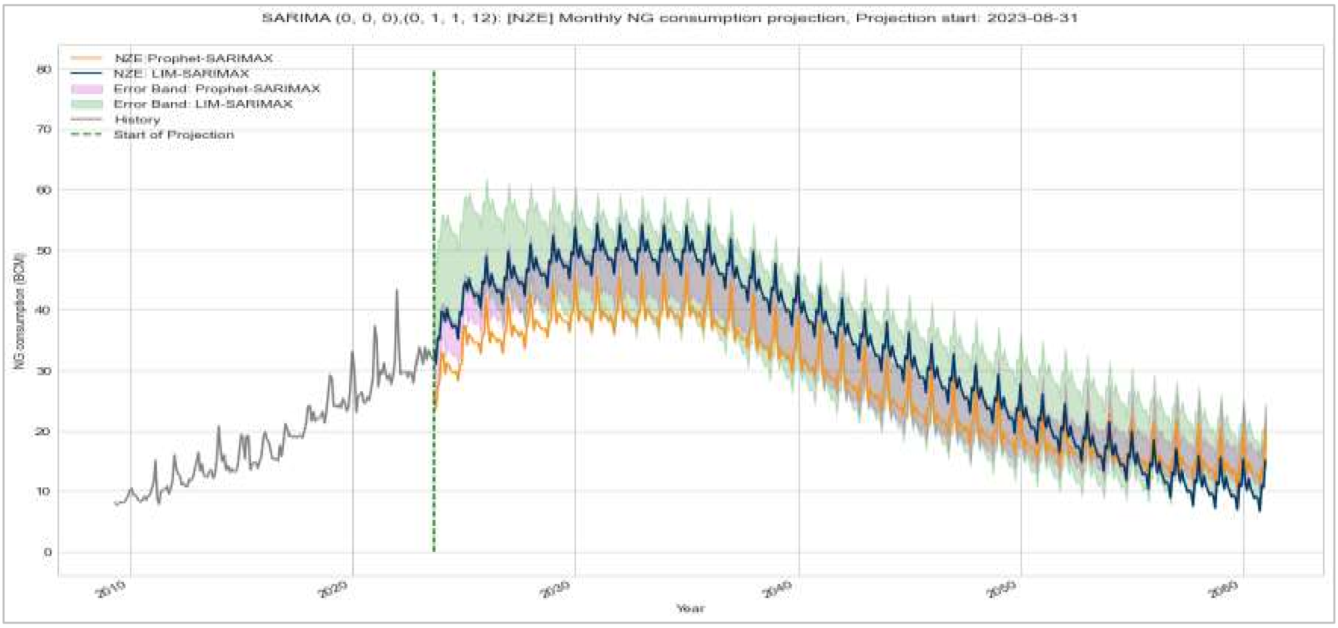
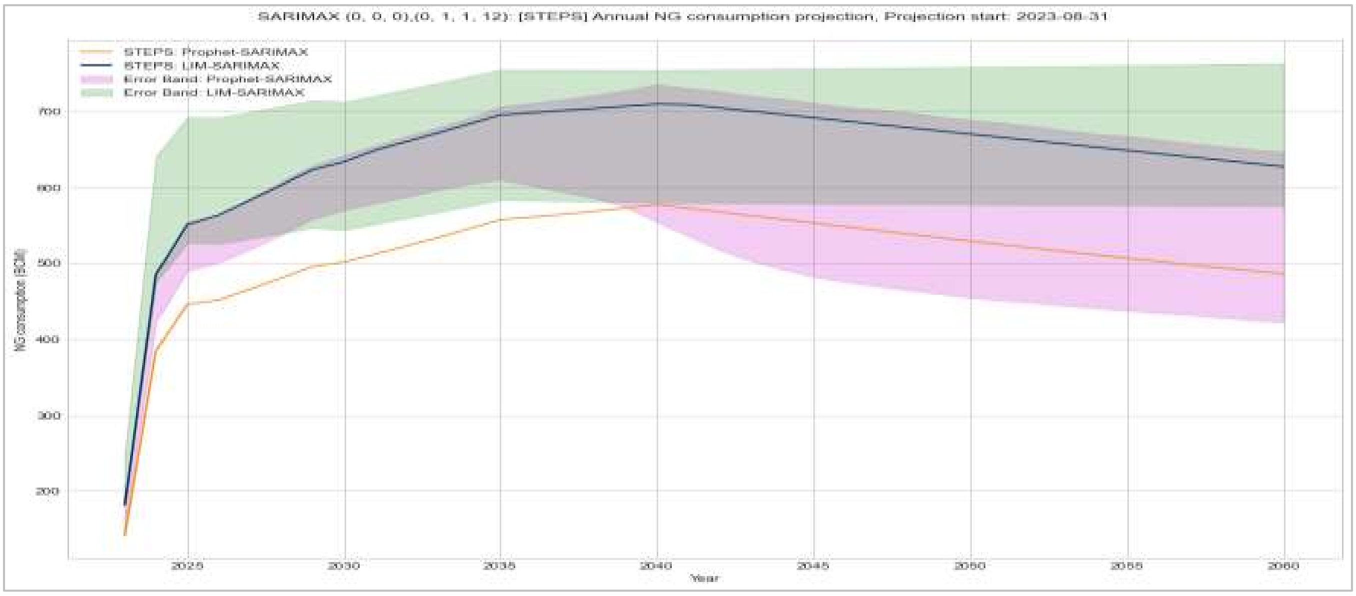


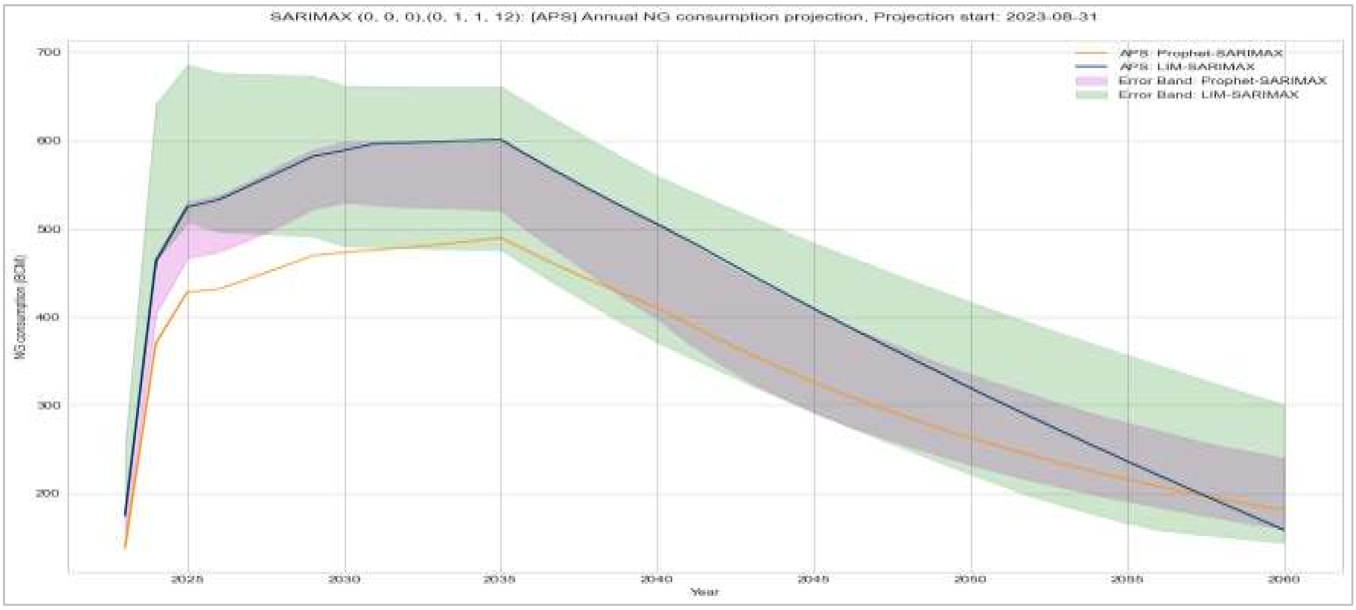
Figure 4.7 Monthly projection and error band

R. Annual projection and error band

STEPS



APS



NZE

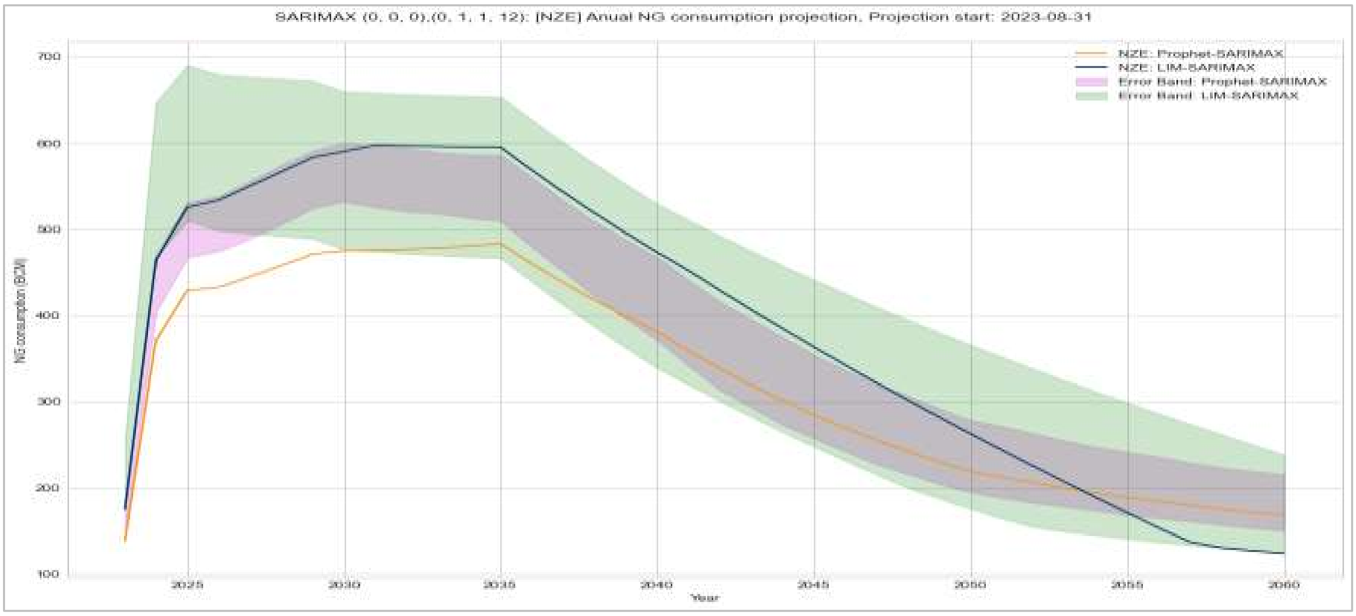
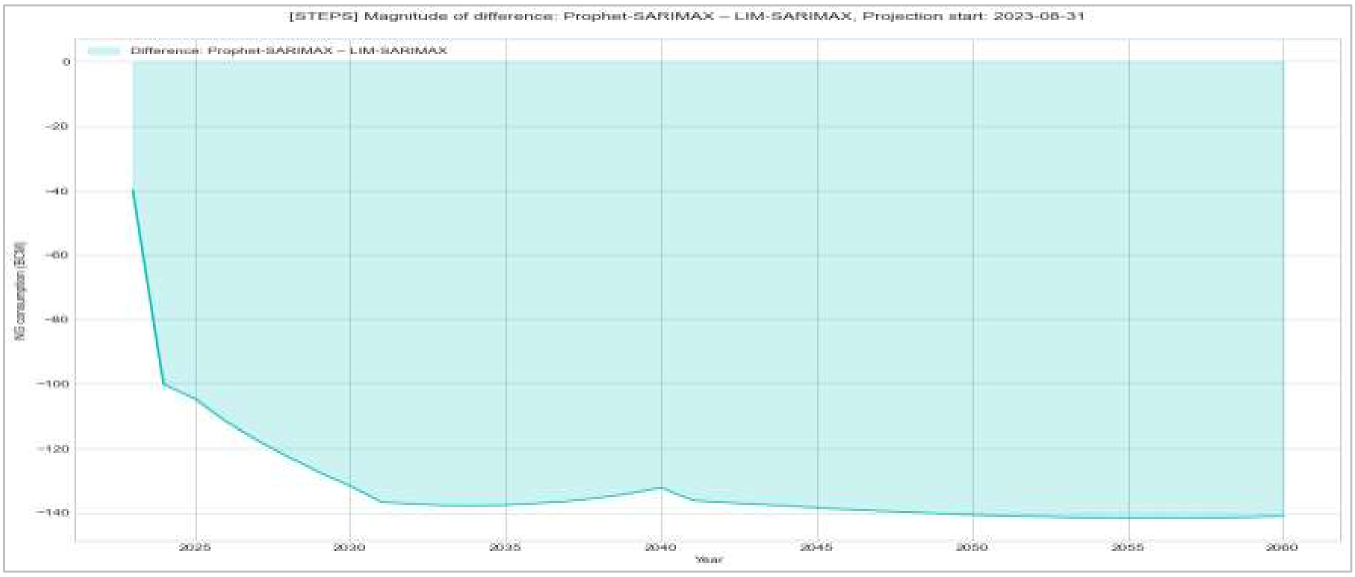


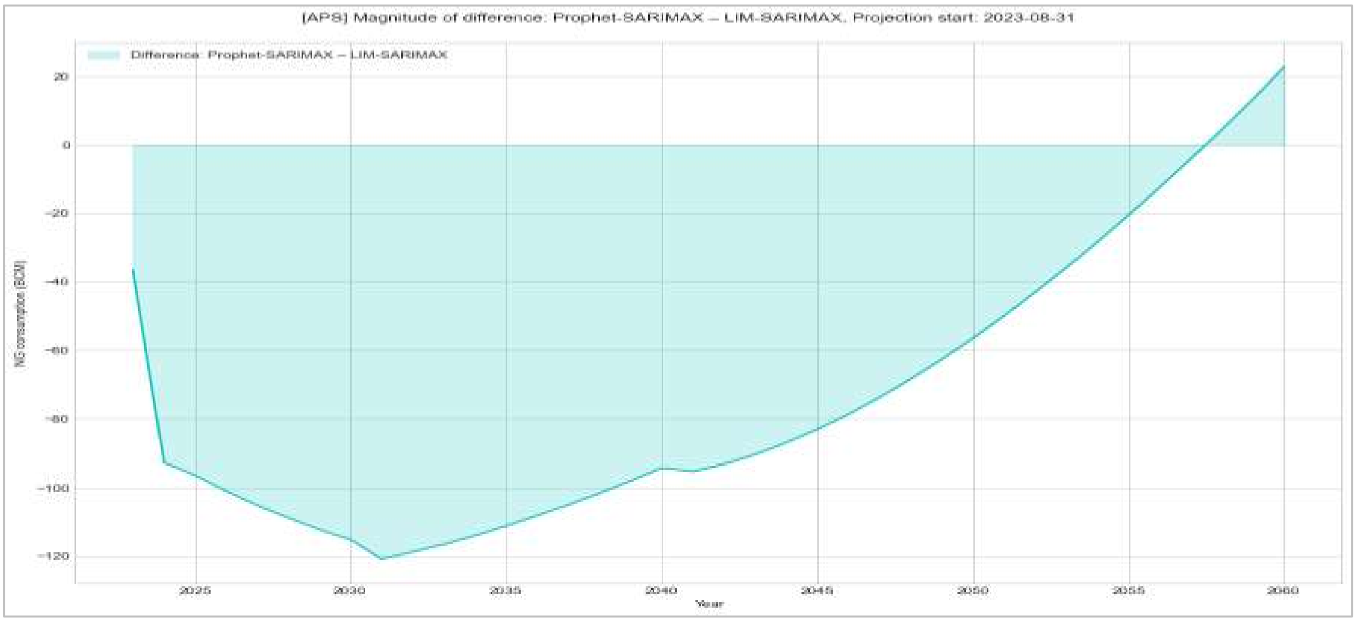
Figure 4.9 Annual projection and error band

S. Magnitude of difference

STEPS



APS



NZE

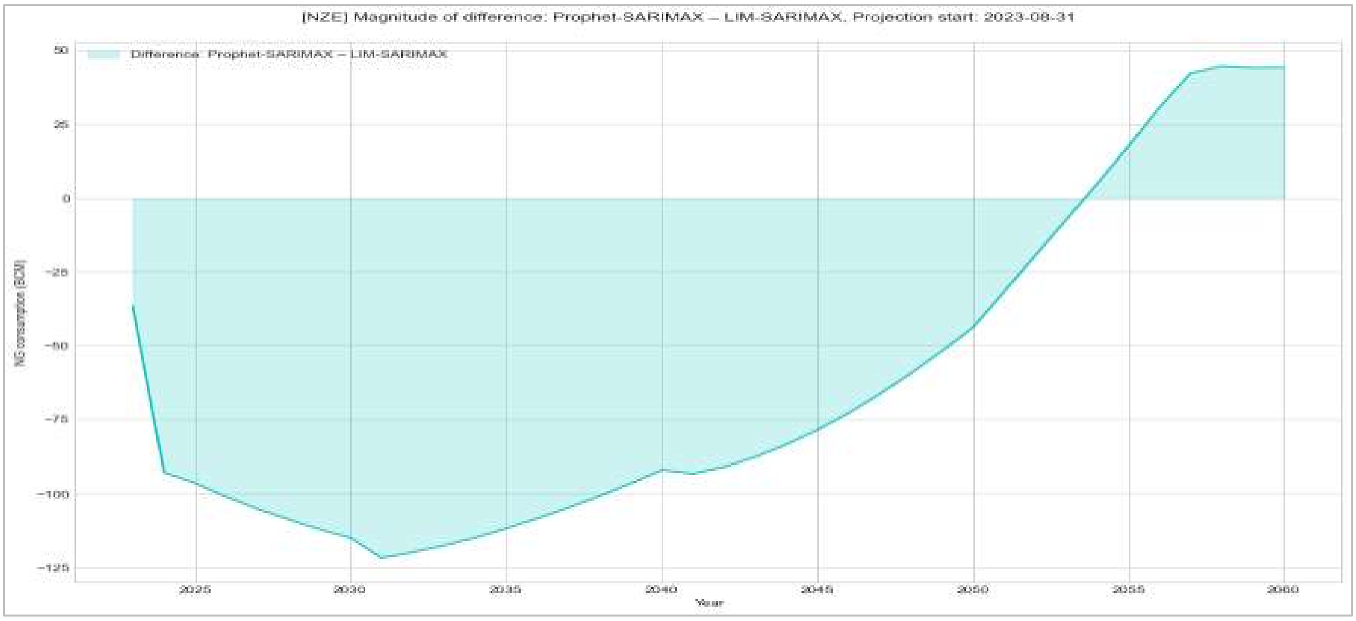


Figure 4.10 Magnitude of difference

T. List of regions and sectors: Tables 5.1 and 5.2

Table 5.1 List of countries and regions in EXIOBASE 3

28 EU countries	16 main economies	Five RoW regions
Austria	Australia	RoW Africa
Belgium	Brazil	RoW America
Bulgaria	China	RoW Asia and Pacific
Croatia	Canada	RoW Europe
Cyprus	India	RoW Middle East
Czech Republic	Indonesia	
Denmark	Japan	
Estonia	Mexico	
Finland	Norway	
France	Russia	
Germany	South Africa	
Greece	South Korea	
Hungary	Switzerland	
Ireland	Taiwan	
Italy	Turkey	
Latvia	United States	
Lithuania		
Luxembourg		
Malta		
Netherlands		
Poland		
Portugal		
Romania		
Slovakia		
Slovenia		
Spain		
Sweden		
United Kingdom		

Table 5.2 List of sectors in EXIOBASE 3

Sector name
1. Cultivation of paddy rice
2. Cultivation of wheat
3. Cultivation of cereal grains nec
4. Cultivation of vegetables, fruit, nuts
5. Cultivation of oil seeds
6. Cultivation of sugar cane, sugar beet
7. Cultivation of plant-based fibers
8. Cultivation of Crops nec
9. Cattle farming
10. Pigs farming
11. Poultry farming
12. Meat animals nec
13. Animal products nec
14. Raw milk
15. Wool, silk-worm cocoons
16. Manure (conventional treatment)
17. Manure (biogas treatment)
18. Forestry, logging and related service activities
19. Fishing, operating of fish hatcheries and fish farms; service activities incidental to fishing
20. Mining of coal and lignite; extraction of peat
21. Crude petroleum and services related to crude oil extraction, excluding surveying
22. Extraction of natural gas and services related to natural gas extraction, excluding surveying
23. Extraction, liquefaction, and regasification of other petroleum and gaseous materials
24. Mining of uranium and thorium ores
25. Mining of iron ores
26. Mining of copper ores and concentrates
27. Mining of nickel ores and concentrates
28. Mining of aluminium ores and concentrates
29. Mining of precious metal ores and concentrates
30. Mining of lead, zinc and tin ores and concentrates
31. Mining of other non-ferrous metal ores and concentrates
32. Quarrying of stone
33. Quarrying of sand and clay
34. Mining of chemical and fertilizer minerals, production of salt, other mining and quarrying n.e.c.
35. Processing of meat cattle
36. Processing of meat pigs
37. Processing of meat poultry
38. Production of meat products nec
39. Processing vegetable oils and fats
40. Processing of dairy products
41. Processed rice
42. Sugar refining
43. Processing of Food products nec
44. Manufacture of beverages
45. Manufacture of fish products
46. Manufacture of tobacco products
47. Manufacture of textiles
48. Manufacture of wearing apparel; dressing and dyeing of fur
49. Tanning and dressing of leather; manufacture of luggage, handbags, saddlery, harness and footwear
50. Manufacture of wood and of products of wood and cork, except furniture; manufacture of articles of straw and plaiting materials
51. Re-processing of secondary wood material into new wood material
52. Pulp
53. Re-processing of secondary paper into new pulp
54. Paper
55. Publishing, printing and reproduction of recorded media
56. Manufacture of coke oven products
57. Petroleum Refinery
58. Processing of nuclear fuel
59. Plastics, basic
60. Re-processing of secondary plastic into new plastic
61. N-fertiliser
62. P- and other fertiliser
63. Chemicals nec
64. Manufacture of rubber and plastic products (25)
65. Manufacture of glass and glass products
66. Re-processing of secondary glass into new glass
67. Manufacture of ceramic goods
68. Manufacture of bricks, tiles and construction products, in baked clay
69. Manufacture of cement, lime and plaster
70. Re-processing of ash into clinker
71. Manufacture of other non-metallic mineral products n.e.c.
72. Manufacture of basic iron and steel and of ferro-alloys and first products thereof
73. Re-processing of secondary steel into new steel
74. Precious metals production
75. Re-processing of secondary preciuos metals into new preciuos metals
76. Aluminium production
77. Re-processing of secondary aluminium into new aluminium
78. Lead, zinc and tin production
79. Re-processing of secondary lead into new lead, zinc and tin
80. Copper production
81. Re-processing of secondary copper into new copper
82. Other non-ferrous metal production
83. Re-processing of secondary other non-ferrous metals into new other non-ferrous metals
84. Casting of metals
85. Manufacture of fabricated metal products, except machinery and equipment (28)
86. Manufacture of machinery and equipment n.e.c. (29)
87. Manufacture of office machinery and computers (30)
88. Manufacture of electrical machinery and apparatus n.e.c. (31)
89. Manufacture of radio, television and communication equipment and apparatus (32)
90. Manufacture of medical, precision and optical instruments, watches and clocks (33)
91. Manufacture of motor vehicles, trailers and semi-trailers (34)
92. Manufacture of other transport equipment (35)
93. Manufacture of furniture; manufacturing n.e.c. (36)
94. Recycling of waste and scrap
95. Recycling of bottles by direct reuse
96. Production of electricity by coal
97. Production of electricity by gas
98. Production of electricity by nuclear
99. Production of electricity by hydro
100. Production of electricity by wind
101. Production of electricity by petroleum and other oil derivatives
102. Production of electricity by biomass and waste
103. Production of electricity by solar photovoltaic
104. Production of electricity by solar thermal
105. Production of electricity by tide, wave, ocean
106. Production of electricity by Geothermal
107. Production of electricity nec
108. Transmission of electricity
109. Distribution and trade of electricity
110. Manufacture of gas; distribution of gaseous fuels through mains
111. Steam and hot water supply
112. Re-processing of secondary construction material into aggregates
113. Sale, maintenance, repair of motor vehicles, motor vehicles parts, motorcycles, motor cycles parts and accessoiries
114. Retail sale of automotive fuel
115. Wholesale trade and commission trade, except of motor vehicles and motorcycles (51)
116. Retail trade, except of motor vehicles and motorcycles; repair of personal and household goods (52)
117. Hotels and restaurants (55)
118. Transport via railways
119. Other land transport
120. Transport via pipelines
121. Sea and coastal water transport
122. Inland water transport
123. Air transport (62)
124. Supporting and auxiliary transport activities; activities of travel agencies (63)
125. Post and telecommunications (64)
126. Financial intermediation, except insurance and pension funding (65)
127. Insurance and pension funding, except compulsory social security (66)
128. Activities auxiliary to financial intermediation (67)
129. Real estate activities (70)
130. Renting of machinery and equipment without operator and of personal and household goods (71)
131. Computer and related activities (72)
132. Research and development (73)
133. Other business activities (74)
134. Public administration and defence; compulsory social security (75)
135. Education (80)
136. Health and social work (85)
137. Incineration of waste: Food
138. Incineration of waste: Paper
139. Incineration of waste: Plastic
140. Incineration of waste: Metals and Inert materials
141. Incineration of waste: Textiles
142. Incineration of waste: Wood
143. Incineration of waste: Oil/Hazardous waste
144. Biogasification of food waste, incl. land application
145. Biogasification of paper, incl. land application
146. Biogasification of sewage slugde, incl. land application
147. Composting of food waste, incl. land application
148. Composting of paper and wood, incl. land application
149. Waste water treatment, food
150. Waste water treatment, other
151. Landfill of waste: Food
152. Landfill of waste: Paper
153. Landfill of waste: Plastic
154. Landfill of waste: Inert/metal/hazardous
155. Landfill of waste: Textiles
156. Landfill of waste: Wood
157. Activities of membership organisation n.e.c. (91)
158. Recreational, cultural and sporting activities (92)
159. Other service activities (93)
160. Private households with employed persons (95)
161. Extra-territorial organizations and bodies
162. Re-processing of secondary construction material into aggregates
163. Sale, maintenance, repair of motor vehicles, motor vehicles parts, motorcycles, motor cycles parts and accessoiries

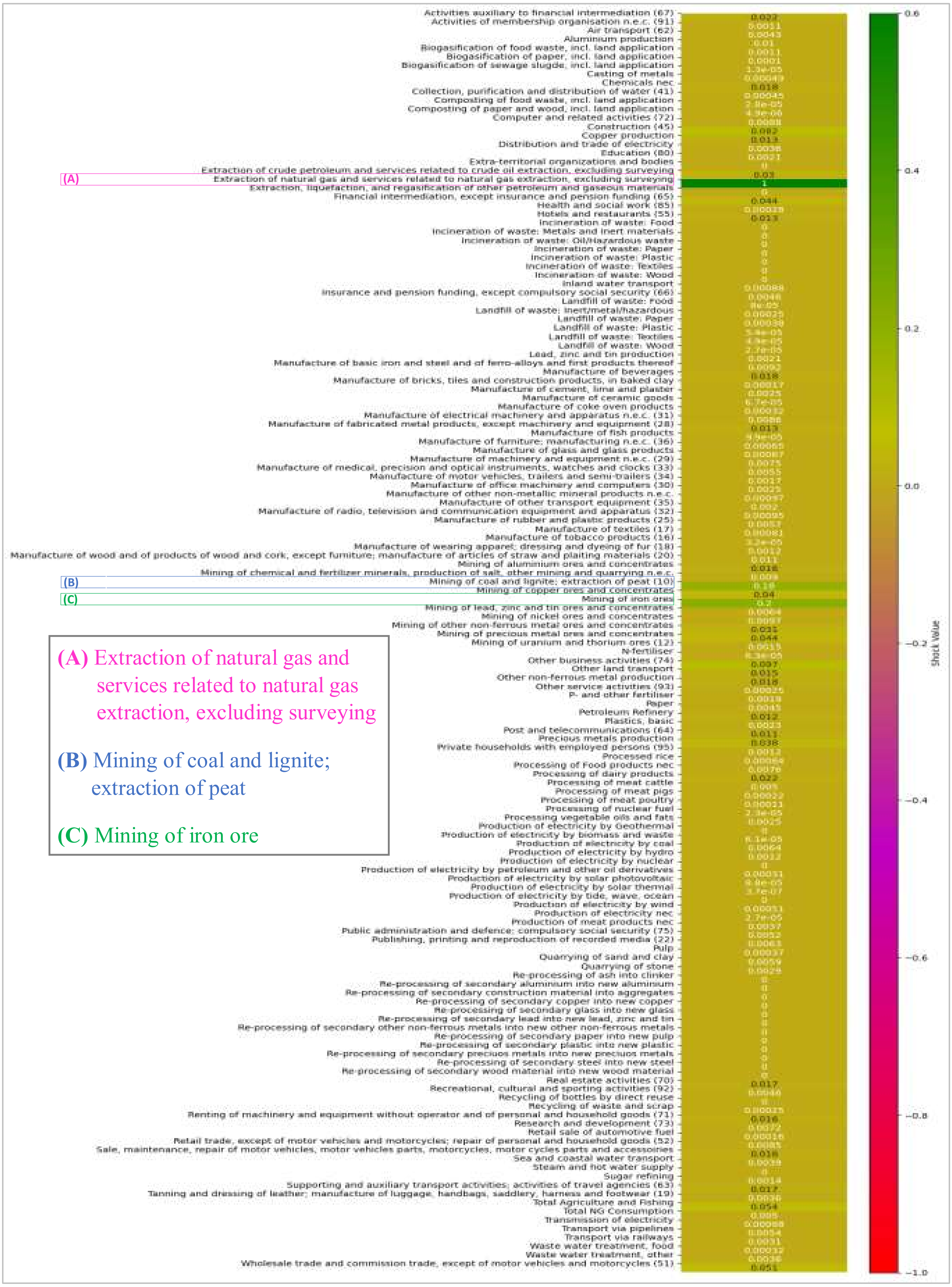


Figure 5.8 Static single-sector impact analysis for 2030: Normalised responses in Australia

Notes: The sectors most affected are marked as (A), (B), and (C).

Figure 5.9 Static single-sector impact analysis for 2030: Relative responses in Australia

Notes: The sectors most affected are marked as (A), (B), and (C).

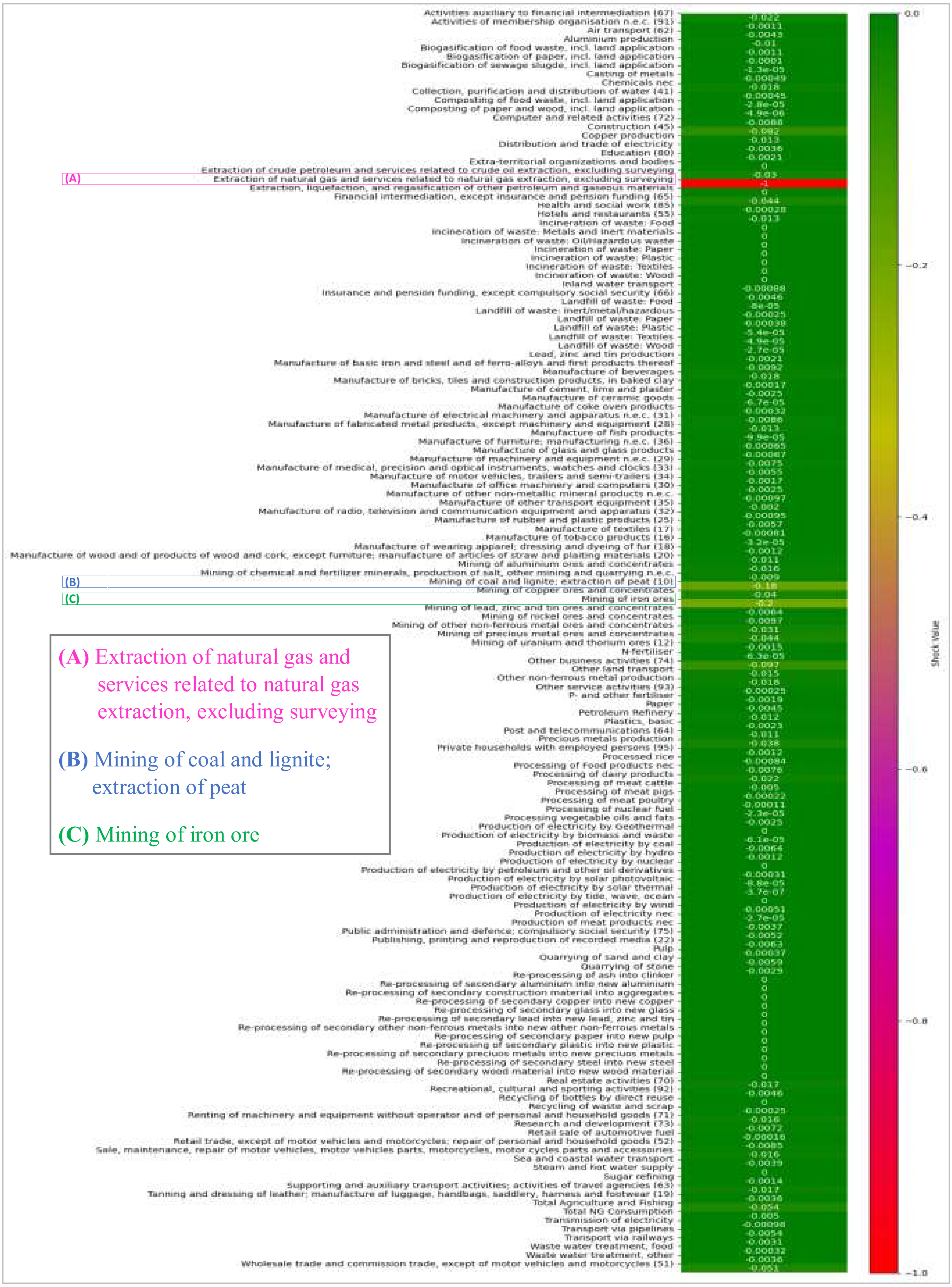


Figure 5.10 Static single-sector impact analysis for 2060: Normalised responses in Australia

Notes: The sectors most affected are marked as (A), (B), and (C).

Figure 5.11 Static single-sector impact analysis for 2060: Relative responses in Australia

Notes: The sectors most affected are marked as (A), (B), and (C).

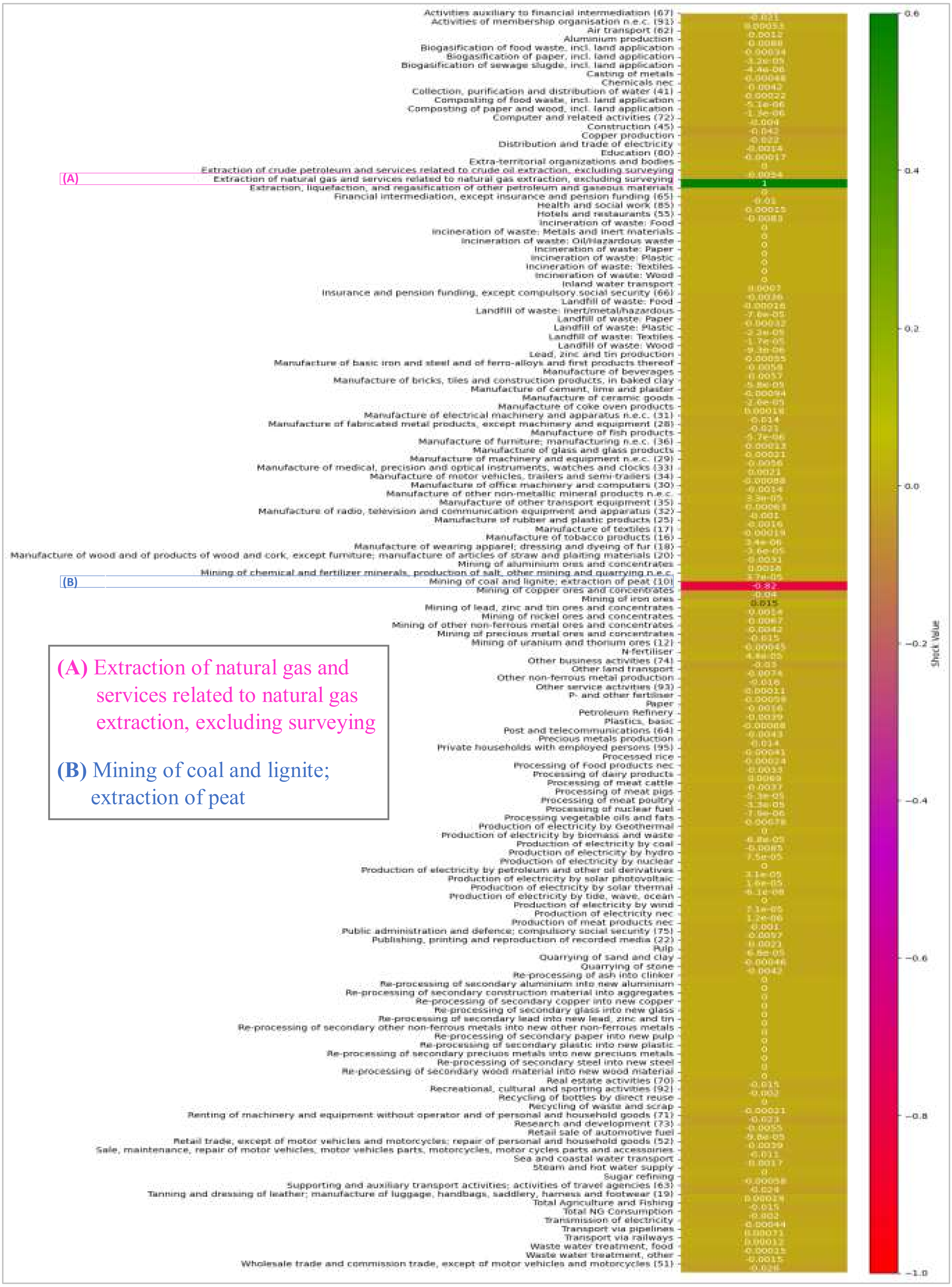


Figure 5.12 Static multiple-sector impact analysis for 2030: Normalised responses in Australia

Notes: The sectors most affected are marked as (A) and (B).

Z. Static multiple-sector impact analysis for 2030: Relative responses in Australia

Figure 5.13 Static multiple-sector impact analysis for 2030: Relative responses in Australia

Notes: The sectors most affected are marked as (A) and (B).

AA. Static multiple-sector impact analysis for 2060: Normalised responses in Australia

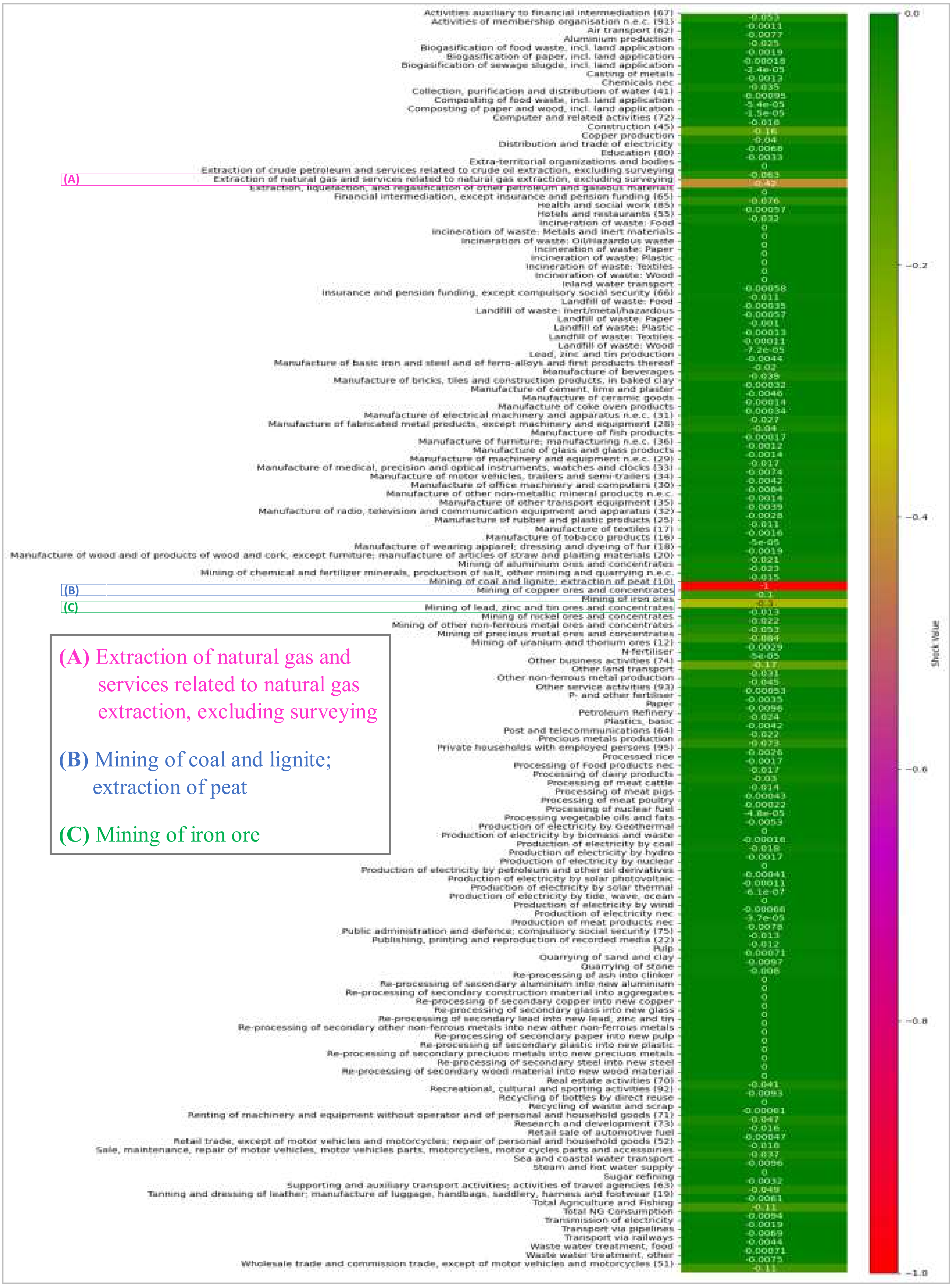


Figure 5.14 Static multiple-sector impact analysis for 2060: Normalised responses in Australia

Notes: The sectors most affected are marked as (A), (B), and (C).

AB. Static multiple-sector impact analysis for 2060: Relative responses in Australia

Figure 5.15 Static multiple-sector impact analysis for 2060: Relative responses in Australia

Notes: The sectors most affected are marked as (A), (B), and (C).

AC. Dynamic multiple-sector impact analysis in Australia for 2030

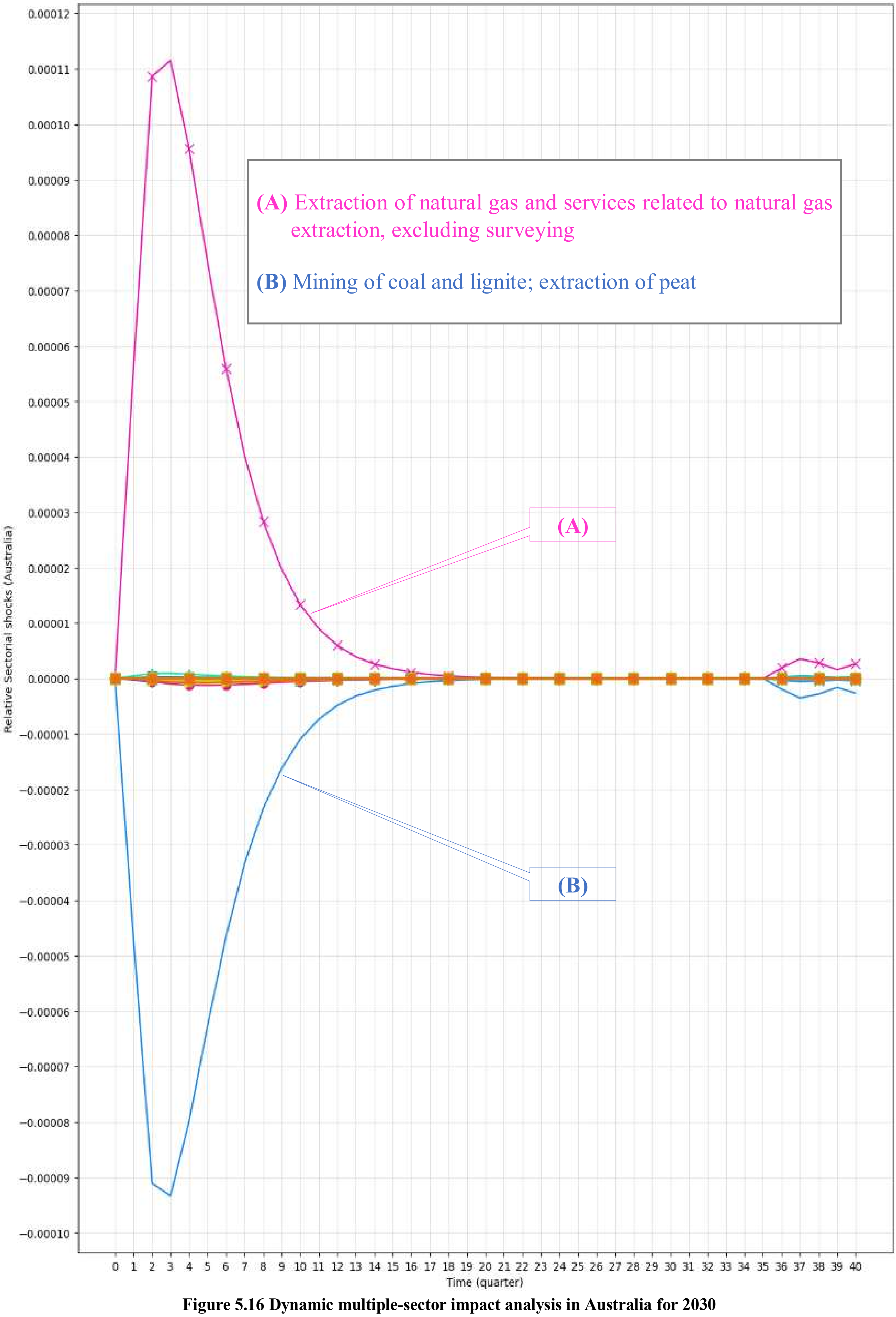


Figure 5.16 Dynamic multiple-sector impact analysis in Australia for 2030

Notes: The sectors most affected are marked as (A) and (B).

AD. Dynamic multiple-sector impact analysis in Australia for 2060

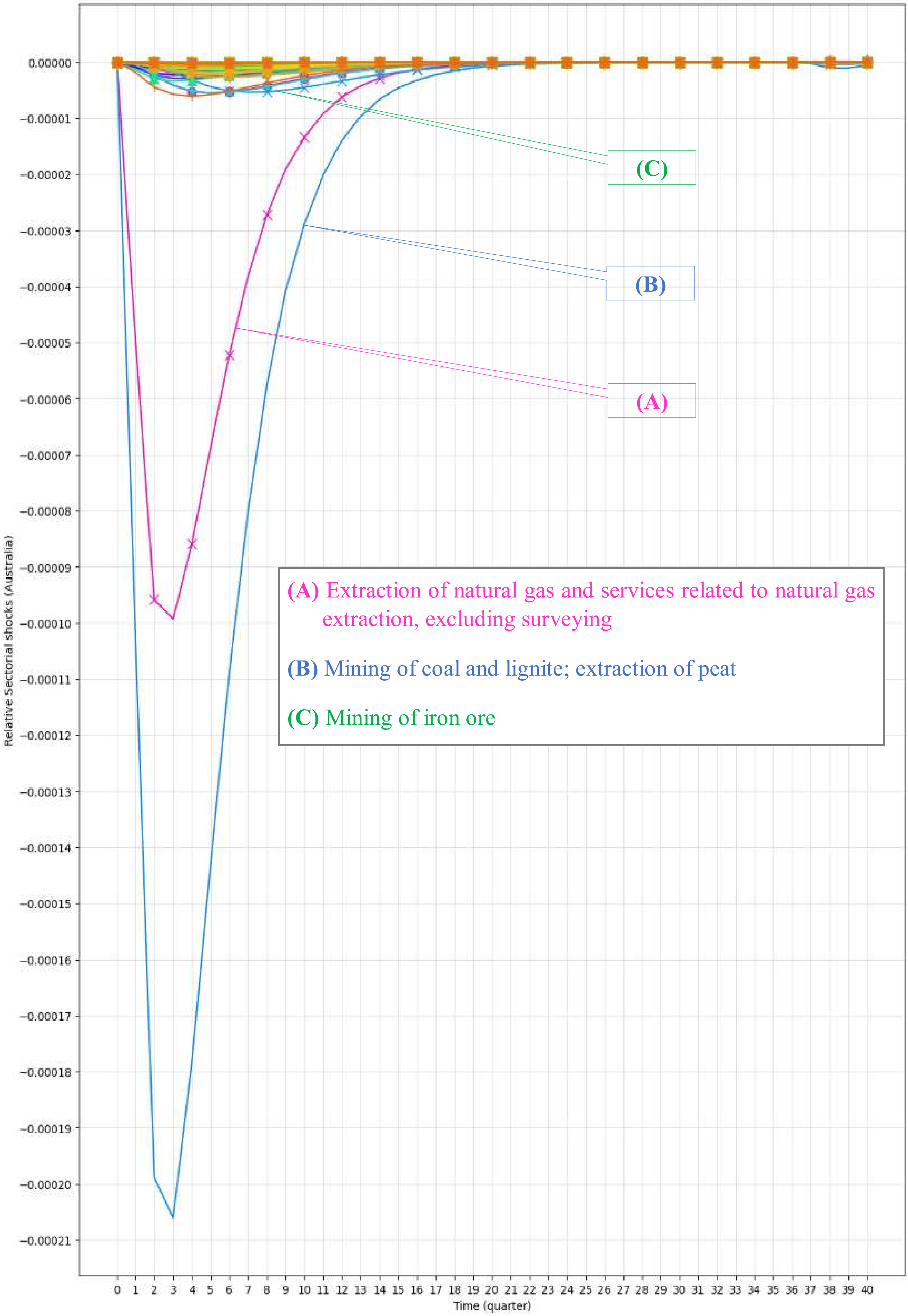


Figure 5.17 Dynamic multiple-sector impact analysis in Australia for 2060

Notes: The sectors most affected are marked as (A) and (B).

[Legend for Appendices AC and AD: Industries]

AE. The economic stimulus in Australia's NG sector

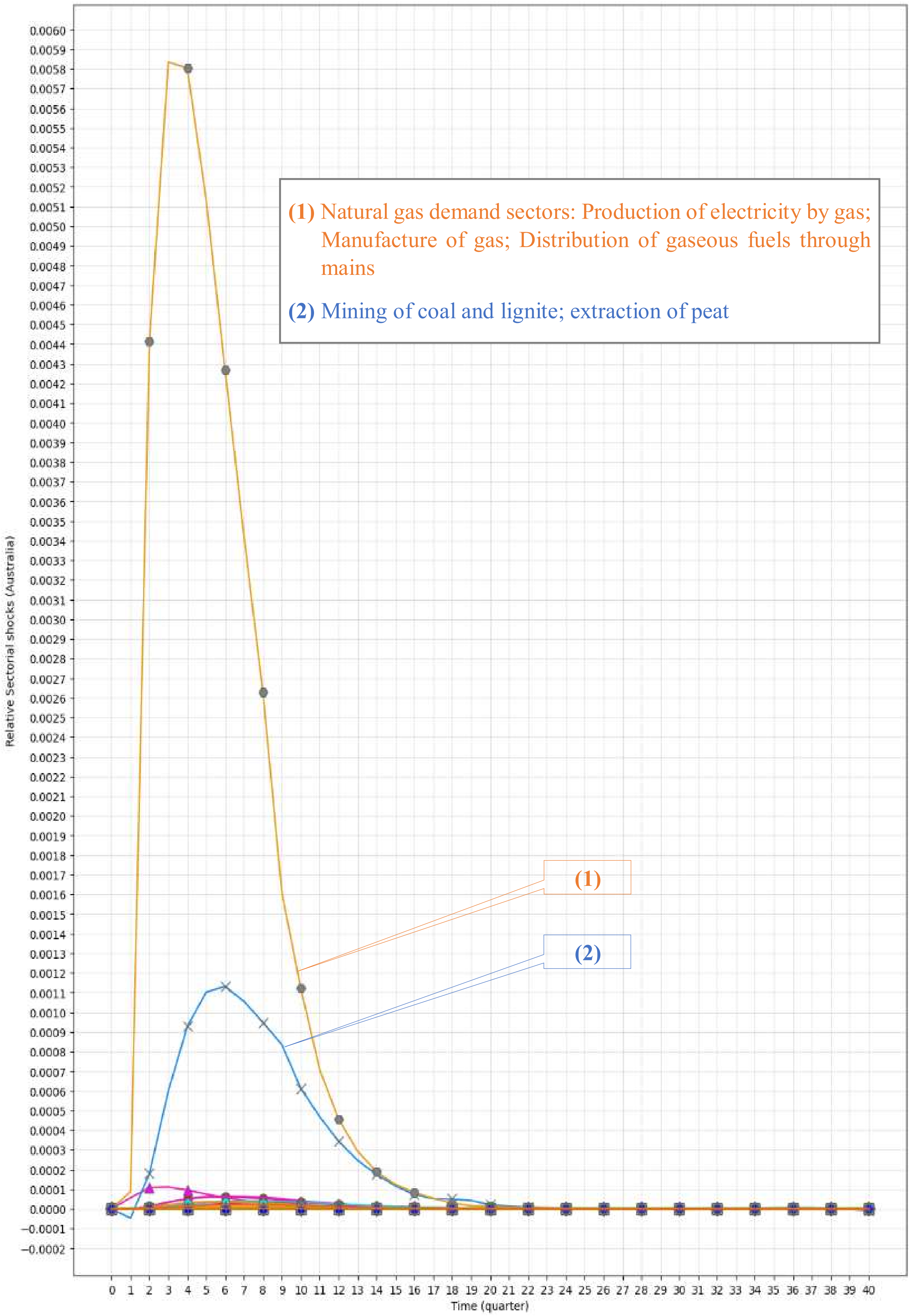


Figure 5.18 The economic stimulus in Australia's NG sector

Notes: The sectors most affected are marked as (1) and (2).

AF. The economic stimulus in Australia's renewable sector

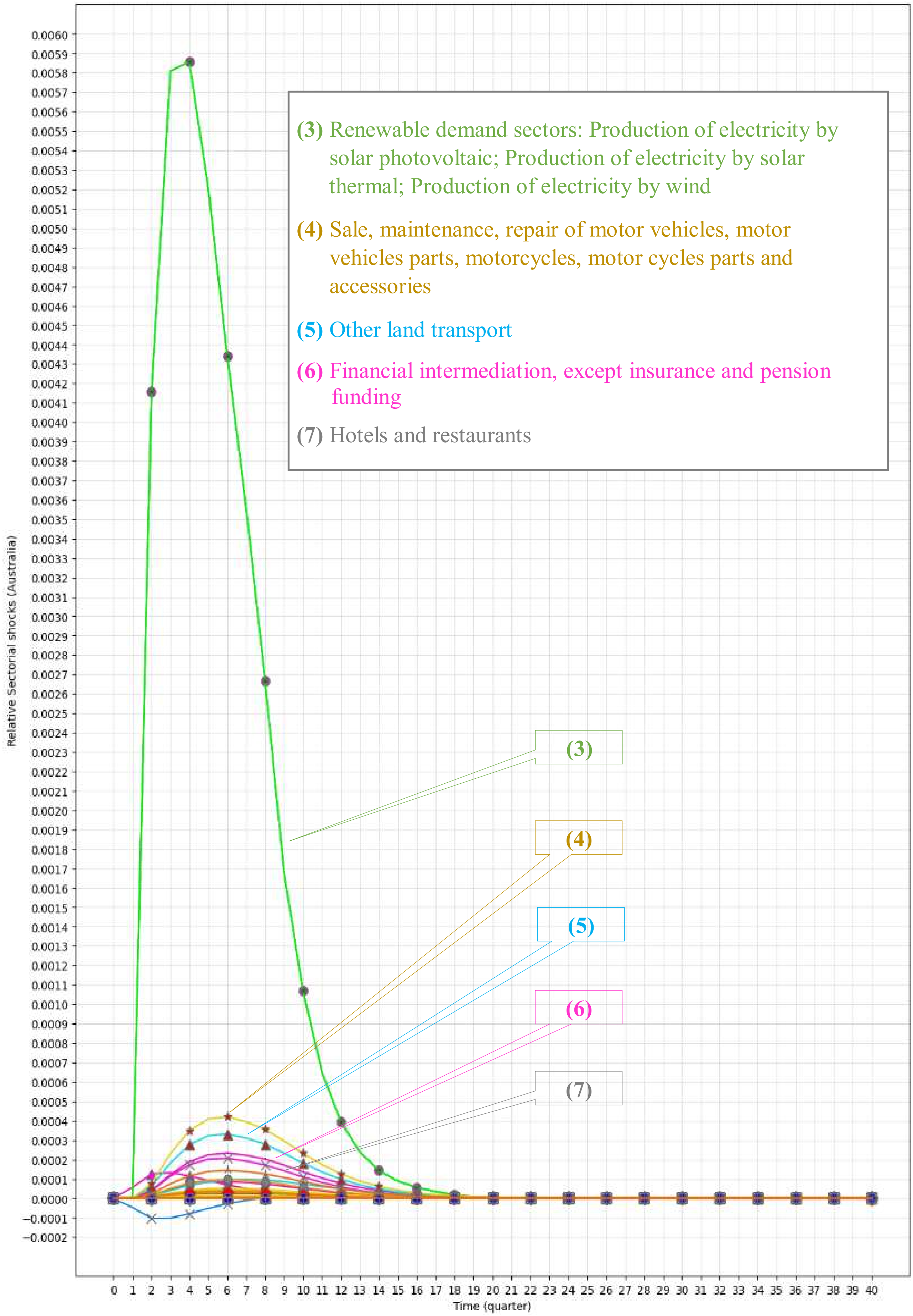


Figure 5.19 The economic stimulus in Australia's renewable sector

Notes: The sectors most affected are marked as (3), (4), (5), (6), and (7).

[Legend for Appendices AE and AF: Industries]

AG. Numerical implementation details of dynamic input–output modelling

The dynamic IO simulation framework in this study is developed in Python using SciPy's solve_ivp function with various Runge-Kutta solvers [235]. After comparative testing, the DOP853 algorithm was chosen for its ability to handle large systems with high numerical precision and computational efficiency.

In economic modelling, a system is typically considered non-stiff if its numerical solution does not display rapid changes in scale and remains stable over time. This stability enables accurate integration, even with relatively small time intervals. For such systems, high-order explicit Runge-Kutta methods are sufficient and advantageous due to their improved precision and stability.

AH. Australia's future gas strategy

	Future Gas Strategy
Background	• NG consumption In May 2024, the Australian government announced the Future Gas Strategy, which emphasises gas's role in achieving carbon neutrality (net zero) by 2050. • The government also supports using gas in the industrial sector if alternative energy sources are unavailable or too costly beyond 2050
Policy guideline	• Australia aims for net zero emissions by 2050. • Gas must remain affordable during the net zero transition • New gas supplies are needed to meet demand during the economy-wide transition. • A reliable gas supply will support higher-value gas uses, offering households energy choices. • Gas and electricity markets must adapt for the energy transformation. • Australia will remain a reliable energy trading partner, particularly in LNG and low-emission gases.

AI. Denmark's green energy transition

Denmark's successful energy transition, driven by reducing NG consumption and expanding renewable energy development, provides important insights into the feasibility of Australia's renewable energy-driven growth model. In the 1960s and 70s, Denmark heavily relied on fossil fuel imports [238], but after the 1973 oil crisis, the country pursued energy diversification and, as of 2023, supplies 89% of its electricity from renewable energy sources [73]. In particular, wind power plays a significant role, and the share of gas in electricity generation has decreased significantly from 2000 to 2023, now accounting for only 2.9% of electricity generation and 9% of total energy supply [73]. Furthermore, Denmark plans to phase out NG entirely by 2030 by significantly increasing onshore wind, solar, and offshore wind energy and switching heating gas to biomethane [74]. This is a Danish crucial step towards eliminating fossil fuels dependence to achieve carbon neutrality by 2050 [74].

References

[1] *Australia Exports By Country*. 2025; Available from: https://tradingeconomics.com/australia/exports-by-country.

[2] Observatory of Economic Complexity, *Where does China import Liquefied Natural Gas from? (2023)*. 2023; Available from: https://oec.world/en/visualize/tree_map/hs92/import/chn/show/5271111/2023.

[3] Stern, N. and C. Xie, *China's new growth story: linking the 14th Five-Year Plan with the 2060 carbon neutrality pledge.* Journal of Chinese economic and business studies, 2022. **ahead-of-print**(ahead-of-print): p. 1-21.

[4] Meidan, M., *Unpacking China's 2060 carbon neutrality pledge.* Oxford Energy Comment, 2020(The Oxford Institute For Energy Studies.).

[5] Shi, X. and H. Variam, *China's Gas Market Liberalisation-The impact on China–Australia gas trade*. 2015. p. 137-174.

[6] Deloitte Access Economics, *Gas Market Transformations: Economic Consequences for the Manufacturing Sector*. 2014. p. The Australian Industry Group, the Australian Aluminium Council, the Australian Food and Grocery Council, the Australian Steel Institute, the Energy Users Association of Australia and the Plastics and Chemicals Industries Association; Available from: https://1earthmedia.com/wp-content/uploads/2024/11/Deloitte_Gas_Market_Transformations_Manufacturing_Impacts_Report_web_final_July_14_2014.pdf.

[7] Rezaei, M., et al., *A new approach based on scenario planning and prediction methods for the estimation of gasoil consumption.* International journal of environmental science and technology, 2020. **17**: p. 3241-3250.

[8] Hepburn, C., et al., *Towards carbon neutrality and China's 14th Five-Year Plan: Clean energy transition, sustainable urban development, and investment priorities.* Environmental Science and Ecotechnology, 2021. **8**: p. 100130.

[9] Howarth, R.W., *The greenhouse gas footprint of liquefied natural gas (LNG) exported from the United States.* Energy Science & Engineering, 2024. **12**(11): p. 4843-4859.

[10] Qin, Y., et al., *Challenges of using natural gas as a carbon mitigation option in China.* Energy Policy, 2018. **117**: p. 457-462.

[11] Gürsan, C. and V. de Gooyert, *The systemic impact of a transition fuel: Does natural gas help or hinder the energy transition?* Renewable and Sustainable Energy Reviews, 2021. **138**: p. 110552.

[12] Unruh, G.C., *Understanding carbon lock-in.* Energy Policy, 2000. **28**(12): p. 817-830.

[13] Santillán Vera, M., et al., *Drivers of electricity GHG emissions and the role of natural gas in mexican energy transition.* Energy Policy, 2023. **173**: p. 113316.

[14] Wang, T., et al., *Market reforms and determinants of import natural gas prices in China.* Energy, 2020. **196**: p. 117105.

[15] Wei, Q., P. Zhou, and X. Shi, *The congestion cost of pipeline networks under third-party access in China's natural gas market.* Energy, 2023. **284**: p. 128521.

[16] Mbah, R.E. and F. Wasum, *Russian-Ukraine 2022 War: A Review of the Economic Impact of Russian-Ukraine Crisis on the USA, UK, Canada, and Europe.* Advances in Social Sciences Research Journal, 2022. **9**: p. 144-153.

[17] Li, J., et al., *China's flexibility challenge in achieving carbon neutrality by 2060.* Renewable and Sustainable Energy Reviews, 2022. **158**: p. 112112.

[18] Xu, H. and Z. Wang, *Analysis of driving factors of natural gas consumption in china: Based on regional integration and LMDI method.* Energy Engineering: Journal of the Association of Energy Engineering, 2021. **118**(4): p. 1027-1043.

[19] Bu, Y., et al., *Spatial pattern and driving factors for interprovincial natural gas consumption in China: Based on SNA and LMDI.* Journal of Cleaner Production, 2020. **263**: p. 121392.

[20] Jiang, H., et al., *What drives China's natural gas consumption? Analysis of national and regional estimates.* Energy Economics, 2020. **87**: p. 104744.

[21] Wang, J. and N. Li, *Influencing factors and future trends of natural gas demand in the eastern, central and western areas of China based on the grey model.* Natural Gas Industry B, 2020. **7**(5): p. 473-483.

[22] Zou, H., *The Adaptive Lasso and Its Oracle Properties.* Journal of the American Statistical Association, 2006. **101**(476): p. 1418-1429.
[23] Mendes, E.F., *Model Selection Consistency for Cointegrating Regressions.* arXiv: Methodology, 2011.
[24] Shrestha, M.B. and G.R. Bhatta, *Selecting appropriate methodological framework for time series data analysis.* The Journal of finance and data science, 2018. **4**(2): p. 71-89.
[25] Jia, W., et al., *Potential changes of regional natural gas market in China amidst liberalization: A mixed complementarity equilibrium simulation in 2030.* Energy, 2023. **284**: p. 129254.
[26] Lee, S. and M. Lee, *The Israel-Hamas War, Oil Price Volatility, and Anticipated Impacts: Implications for Korean Industries.* Korea Institute for Industrial Economics and Trade Research Paper No, 2023. **23**.
[27] Tuyu Zhou, H.X., Jorrit Gosens, and Frank Jotzo, *China's net-zero plans*, in *Policy brief on near-term policy challenges & Australia-China links in decarbonization*. 2022, Crawford School of Public Policy; Available from: https://iceds.anu.edu.au/files/ANU%20policy%20brief%20-%20Chinese%20climate%20and%20energy%20policy%20trends_0.pdf.
[28] Jia, X., et al., *Multi-objective energy planning for China's dual carbon goals.* Sustainable Production and Consumption, 2022. **34**: p. 552-564.
[29] Patel, A. *Experts: What to expect in China's climate pledge for 2035*. 2024.
[30] Liu, Z., et al., *Challenges and opportunities for carbon neutrality in China.* Nature Reviews Earth & Environment, 2022. **3**(2): p. 141-155.
[31] Gao, J. and X. Dong, *Stimulating factors of urban gas consumption in China.* Natural Gas Industry, 2018. **38**(3): p. 130-137.
[32] Liu, G., et al., *Natural gas consumption of urban households in China and corresponding influencing factors.* Energy Policy, 2018. **122**: p. 17-26.
[33] Mu, X. and G. Li, *Study on natural gas consumption forecast and influence factors in China based on system dynamics model.* Journal of Engineering Studies, 2018. **10**(1): p. 56-67.
[34] Shi, X., V. Hari Malamakkavu Padinjare, and T. Jacqueline, *Global impact of uncertainties in China's gas market.* Energy Policy, 2017. **104**: p. 382-394.
[35] Shi, X. and S. Sun, *Energy price, regulatory price distortion and economic growth: A case study of China.* Energy Economics, 2017. **63**: p. 261-271.
[36] Paltsev, S. and D. Zhang, *Natural gas pricing reform in China: Getting closer to a market system?* Energy Policy, 2015. **86**: p. 43-56.
[37] Ramírez-Orellana, A., et al., *Is the corporate financial strategy in the oil and gas sector affected by ESG dimensions?* Resources Policy, 2023. **81**: p. 103303.
[38] Wang, X., et al., *Does Environmental Regulation Affect Natural Gas Consumption? Evidence from China with Spatial Insights.* Sustainability, 2020. **12**(8): p. 3354.
[39] Puttachai, W., R. Phadkantha, and W. Yamaka, *The threshold effects of ESG performance on the energy transitions: A country-level data.* Energy Reports, 2022. **8**: p. 234-241.
[40] Adow, A.H., H. Mahmood, and M. Furqan, *Exploring the Effects of Energy Consumption, Financial Market Development, and Urbanization on CO2 Emissions in GCC Countries: Cross-Sectional Dependence Analysis.* International Journal of Energy Economics and Policy, 2025. **15**(3): p. 146.
[41] Zhou, Z.Z., et al., *From pollution to sustainable environment: unlocking the role of energy intensity and renewable energy in highly pollutant nations.* Environmental Research Communications, 2025.
[42] Albon, C., *Machine Learning with Python Cookbook: Practical Solutions from Preprocessing to Deep Learning*. 2018, Sebastopol: O'Reilly Media, Incorporated.
[43] Ghoddusi, H., G.G. Creamer, and N. Rafizadeh, *Machine learning in energy economics and finance: A review.* Energy Economics, 2019. **81**: p. 709-727.
[44] Shi, X., et al., *Prioritizing driving factors of household carbon emissions: An application of the LASSO model with survey data.* Energy Economics, 2020. **92**: p. 104942.

[45] Kock, A.B., *CONSISTENT AND CONSERVATIVE MODEL SELECTION WITH THE ADAPTIVE LASSO IN STATIONARY AND NONSTATIONARY AUTOREGRESSIONS.* Econometric Theory, 2016. **32**(1): p. 243-259.
[46] Duras, T., et al., *Using machine learning to select variables in data envelopment analysis: Simulations and application using electricity distribution data.* Energy Economics, 2023. **120**: p. 106621.
[47] Engle, R.F. and C.W.J. Granger, *Co-Integration and Error Correction: Representation, Estimation, and Testing.* Econometrica, 1987. **55**(2): p. 251-276.
[48] Liao, Z. and P.C.B. Phillips, *AUTOMATED ESTIMATION OF VECTOR ERROR CORRECTION MODELS.* Econometric theory, 2015. **31**(3): p. 581-646.
[49] Liang, C. and M. Schienle, *Determination of vector error correction models in high dimensions.* Journal of Econometrics, 2019. **208**(2): p. 418-441.
[50] Zivot, E. and J. Wang, *Modeling Financial Time Series with S-PLUS®*. 2nd ed. 2006. ed. 2006, New York, NY: Springer New York.
[51] Stock, J.H. and M.W. Watson, *Introduction to econometrics*. The Pearson series in economics. 2019: Pearson Education.
[52] Yu, E.S.H. and J.C. Jin, *Cointegration tests of energy consumption, income, and employment.* Resources and Energy, 1992. **14**(3): p. 259-266.
[53] Géron, A.l., *Hands-on machine learning with Scikit-Learn, Keras and TensorFlow : concepts, tools, and techniques to build intelligent systems*. Third edition. ed. 2022, Sebastopol: O'Reilly.
[54] Shortreed, S.M. and A. Ertefaie, *Outcome-adaptive lasso: variable selection for causal inference.* Biometrics, 2017. **73**(4): p. 1111-1122.
[55] Narayan, P.K. and R. Smyth, *WHAT DETERMINES MIGRATION FLOWS FROM LOW-INCOME TO HIGH-INCOME COUNTRIES? AN EMPIRICAL INVESTIGATION OF FIJI-U.S. MIGRATION 1972-2001.* Contemporary economic policy, 2006. **24**(2): p. 332-342.
[56] Kondratieff, N.D., *The Static and the Dynamic View of Economics.* The Quarterly Journal of Economics, 1925. **39**(4): p. 575-583.
[57] Reed, W.R., *On the Practice of Lagging Variables to Avoid Simultaneity.* Oxford bulletin of economics and statistics, 2015. **77**(6): p. 897-905.
[58] Campos, J., N.R. Ericsson, and D.F. Hendry, *General-to-specific Modeling: An Overview and Selected Bibliography.* International finance discussion papers, 2005. **2005**(838): p. 1-94.
[59] Müller, A.C. and S. Guido, *Introduction to Machine Learning with Python: A Guide for Data Scientists*. 2016, Sebastopol: O'Reilly Media, Incorporated.
[60] Kumar, P.K. and B. Sarojamma, *Time Series Cross-Validation Techniques for determining the order of the Autoregressive models.* International journal of advanced research in computer science, 2017. **8**(3).
[61] Wang, Y., et al., *Decarbonization pathways of China's iron and steel industry toward carbon neutrality.* Resources, Conservation and Recycling, 2023. **194**: p. 106994.
[62] Yu, X. and C. Tan, *China's pathway to carbon neutrality for the iron and steel industry.* Global Environmental Change, 2022. **76**: p. 102574.
[63] Shen, J., et al., *Future CO2 emission trends and radical decarbonization path of iron and steel industry in China.* Journal of Cleaner Production, 2021. **326**: p. 129354.
[64] Zhang, C., et al., *Analysis of electricity consumption in China (1990–2016) using index decomposition and decoupling approach.* Journal of Cleaner Production, 2019. **209**: p. 224-235.
[65] Villar, J.A. and F.L. Joutz, *The relationship between crude oil and natural gas prices.* Energy Information Administration, Office of Oil and Gas, 2006. **1**: p. 1-43.
[66] Maxwell, D. and Z. Zhu, *Natural gas prices, LNG transport costs, and the dynamics of LNG imports.* Energy Economics, 2011. **33**(2): p. 217-226.
[67] Gu, Y., et al., *Overall review of peak shaving for coal-fired power units in China.* Renewable & sustainable energy reviews, 2016. **54**: p. 723-731.
[68] Gyomai, G. and E. Guidetti, *International Seminar on Early Warning and Business Cycle Indicators 14 to 16 December 2009 Scheveningen, The Netherlands.*
[69] OECD, *OECD Composite Leading Indicators dataset*. Available from: https://data-explorer.oecd.org/vis?fs[0]=Topic%2C1%7CEconomy%23ECO%23%7CLeading%20indicators

%23ECO_LEA%23&pg=0&fc=Topic&bp=true&snb=1&vw=tb&df[ds]=dsDisseminateFinalDMZ&df[id]=DSD_STES%40DF_CLI&df[ag]=OECD.SDD.STES&df[vs]=&dq=.M.LI...AA...H&lom=LASTNPERIODS&lo=10&to[TIME_PERIOD]=false.
[70] Li, Z.-G., H. Cheng, and T.-Y. Gu, *Research on dynamic relationship between natural gas consumption and economic growth in China.* Structural Change and Economic Dynamics, 2019. **49**: p. 334-339.
[71] Zhang, X., et al., *Carbon emissions in China: How far can new efforts bend the curve?* Energy Economics, 2016. **54**: p. 388-395.
[72] Stern, N. and A. Valero, *Innovation, growth and the transition to net-zero emissions.* Research Policy, 2021. **50**(9): p. 104293.
[73] IEA, *Denmark (Energy Mix)*. 2023; Available from: https://www.iea.org/countries/denmark/energy-mix.
[74] IEA, *Denmark 2023*. 2023, IEA, Paris; Available from: https://www.iea.org/reports/denmark-2023.
[75] Ogden, J., et al., *Natural gas as a bridge to hydrogen transportation fuel: Insights from the literature.* Energy Policy, 2018. **115**: p. 317-329.
[76] Lu, X., et al., *Government regulation and China's natural gas price distortion: A sectoral perspective.* Natural resources forum, 2024.
[77] Yang, X., et al., *How does China's green institutional environment affect renewable energy investments? The nonlinear perspective.* Science of The Total Environment, 2020. **727**: p. 138689.
[78] Wesseh, P.K. and B. Lin, *Environmental policy and 'double dividend' in a transitional economy.* Energy Policy, 2019. **134**: p. 110947.
[79] Green, F., et al., *No new fossil fuel projects: The norm we need.* Science (American Association for the Advancement of Science), 2024. **384**(6699): p. 954-957.
[80] Interesse, G. *China Unveils Its First Set of Basic Standards for Corporate Sustainability (ESG) Disclosure*. 2024.
[81] Lauri Myllyvirta; Qi Qin; Xinyi Shen; Chengcheng Qiu; Belinda Schäpe; Xunpeng Shi, M.Y., Ying Zhang, *China's Climate Transition: Outlook 2024*. 2024: Centre for Research on Energy and Clean Air (CREA); Available from: https://energyandcleanair.org/wp/wp-content/uploads/2024/11/Report_-CREA_Chinas-Climate-Transition-Outlook-2024_EN-1.pdf.
[82] Cushing, D.S.C. *China off track on emissions goals as energy demand offsets renewables push, researchers say*. 2024.
[83] Fang, K., et al., *Will China peak its energy-related carbon emissions by 2030? Lessons from 30 Chinese provinces.* Applied Energy, 2019. **255**: p. 113852.
[84] Bierens, H.J. and L.F. Martins, *Time-varying cointegration.* Econometric Theory, 2010. **26**(5): p. 1453-1490.
[85] Koop, G., R. Leon-Gonzalez, and R.W. Strachan, *Bayesian inference in a time varying cointegration model.* Journal of Econometrics, 2011. **165**(2): p. 210-220.
[86] Brand, J.E., X. Zhou, and Y. Xie, *Recent developments in causal inference and machine learning.* Annual Review of Sociology, 2023. **49**(1): p. 81-110.
[87] Ranstam, J. and J.A. Cook, *LASSO regression.* British Journal of Surgery, 2018. **105**(10): p. 1348-1348.
[88] Wang, C., et al., *The roles of hydro, nuclear and biomass energy towards carbon neutrality target in China: A policy-based analysis.* Energy, 2023. **262**: p. 125303.
[89] *Shell LNG Outlook*. 2024, Shell.
[90] Wei, N., et al., *Conventional models and artificial intelligence-based models for energy consumption forecasting: A review.* Journal of Petroleum Science and Engineering, 2019. **181**: p. 106187.
[91] Alsuwaylimi, A.A., *Comparison of arima, ann and hybrid arima-ann models for time series forecasting.* Information Science Letters An International Journal, 2023. **12**(2): p. 1003-1016.
[92] Liu, J., et al., *Natural gas consumption forecasting: A discussion on forecasting history and future challenges.* Journal of Natural Gas Science and Engineering, 2021. **90**: p. 103930.
[93] Elamin, N. and M. Fukushige, *Modeling and forecasting hourly electricity demand by SARIMAX with interactions.* Energy, 2018. **165**: p. 257-268.

[94] Manigandan, P., et al., *Forecasting Natural Gas Production and Consumption in United States-Evidence from SARIMA and SARIMAX Models, Energies, 14 (19)*. 2021.
[95] Vagropoulos, S.I., et al., *Comparison of SARIMAX, SARIMA, modified SARIMA and ANN-based models for short-term PV generation forecasting.* 2016 IEEE International Energy Conference (ENERGYCON), 2016: p. 1-6.
[96] Muzaffar, S. and A. Afshari, *Short-Term Load Forecasts Using LSTM Networks.* Energy Procedia, 2019. **158**: p. 2922-2927.
[97] Le, X.-H., et al., *Application of long short-term memory (LSTM) neural network for flood forecasting.* Water, 2019. **11**(7): p. 1387.
[98] Chen, X., L. Wei, and J. Xu, *House price prediction using LSTM.* arXiv preprint arXiv:1709.08432, 2017.
[99] Bao, W., J. Yue, and Y. Rao, *A deep learning framework for financial time series using stacked autoencoders and long-short term memory.* PloS one, 2017. **12**(7): p. e0180944.
[100] SEO, S.H., *Forecasting Korean LNG import price using ARIMAX, VECM, LSTM and hybrid models.* 2021.
[101] Siami-Namini, S., N. Tavakoli, and A.S. Namin. *A Comparison of ARIMA and LSTM in Forecasting Time Series*. in *2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)*. 2018.
[102] Cerit, A.G., et al., *Time series forecasting of domestic shipping market: comparison of SARIMAX, ANN-based models and SARIMAX-ANN hybrid model.* International journal of shipping and transport logistics, 2022. **14**(3): p. 193.
[103] Dave, E., et al., *Forecasting Indonesia Exports using a Hybrid Model ARIMA-LSTM.* Procedia Computer Science, 2021. **179**: p. 480-487.
[104] Szoplik, J., *Forecasting of natural gas consumption with artificial neural networks.* Energy, 2015. **85**: p. 208-220.
[105] He, Y. and B. Lin, *The impact of natural gas price control in China: A computable general equilibrium approach.* Energy Policy, 2017. **107**: p. 524-531.
[106] Lee, J. and Y. Cho, *National-scale electricity peak load forecasting: Traditional, machine learning, or hybrid model?* Energy, 2022. **239**: p. 122366.
[107] Sung, J., et al., *Short- and Long-term effects of key drivers in China's Natural Gas market (Working Paper).* 2023.
[108] Shadkam, A., *Using SARIMAX to forecast electricity demand and consumption in university buildings*. 2020, University of British Columbia.
[109] Hammami, Z., et al., *Neural networks for online learning of non-stationary data streams: a review and application for smart grids flexibility improvement.* Artificial Intelligence Review, 2020. **53**(8): p. 6111-6154.
[110] Chen, Z., et al., *Demand for Storage and Import of Natural Gas in China until 2060: Simulation with a Dynamic Model.* Sustainability, 2021. **13**(15): p. 8674.
[111] Jayalakshmi, T. and A. Santhakumaran, *Statistical normalization and back propagation for classification.* International Journal of Computer Theory and Engineering, 2011. **3**(1): p. 1793-8201.
[112] Kishan, S.A., et al., *Handwritten character recognition using cnn.* Published in, 2018. **5**(3).
[113] Nugroho, R.A., et al. *Covid-19 classification using K-nearest neighbor with various distances*. in *AIP Conference Proceedings*. 2024. AIP Publishing.
[114] Duangchaemkarn, K., et al. *SARIMA model forecasting performance of the COVID-19 daily statistics in Thailand during the Omicron variant epidemic*. in *Healthcare*. 2022. MDPI.
[115] Şahinli, M.A., *Potato Price Forecasting with Holt-Winters and ARIMA Methods: A Case Study.* American journal of potato research, 2020. **97**(4): p. 336-346.
[116] Danielsson, J., *Financial Risk Forecasting: The Theory and Practice of Forecasting Market Risk with Implementation in R and Matlab*. 1st ed. ed. Wiley finance series. Vol. 588. 2011, Newark: Wiley. xxi-xxi.
[117] Hochreiter, S. and J. Schmidhuber, *Long short-term memory.* Neural computation, 1997. **9**(8): p. 1735-1780.

[118] Kumar, D., et al., *Forecasting of solar and wind power using LSTM RNN for load frequency control in isolated microgrid.* International journal of modelling & simulation, 2021. **41**(4): p. 311-323.
[119] Kaushik, S., et al., *AI in healthcare: time-series forecasting using statistical, neural, and ensemble architectures.* Frontiers in big data, 2020. **3**: p. 4.
[120] Shah, D., W. Campbell, and F.H. Zulkernine. *A comparative study of LSTM and DNN for stock market forecasting*. in *2018 IEEE international conference on big data (big data)*. 2018. IEEE.
[121] Chollet, F.o., *Deep learning with Python*. Second edition. ed. 2021, Shelter Island, New York: Manning Publications Co. LLC.
[122] Wang, X., et al. *LSTM-based short-term load forecasting for building electricity consumption*. in *2019 IEEE 28th International Symposium on Industrial Electronics (ISIE)*. 2019. IEEE.
[123] Chamatidis, I., et al., *Short-Term Forecasting of Rainfall Using Sequentially Deep LSTM Networks: A Case Study on a Semi-Arid Region.* Environmental Sciences Proceedings, 2023. **26**(1): p. 157.
[124] Malhathkar, S. and S. Thenmozhi. *Deep Learning for Time Series Forecasting–With a focus on Loss Functions and Error Measures*. in *2022 IEEE World Conference on Applied Intelligence and Computing (AIC)*. 2022. IEEE.
[125] Chang, Z., Y. Zhang, and W. Chen. *Effective adam-optimized LSTM neural network for electricity price forecasting*. in *2018 IEEE 9th international conference on software engineering and service science (ICSESS)*. 2018. IEEE.
[126] Kandel, I. and M. Castelli, *The effect of batch size on the generalizability of the convolutional neural networks on a histopathology dataset.* ICT express, 2020. **6**(4): p. 312-315.
[127] Budiharto, W., *Data science approach to stock prices forecasting in Indonesia during Covid-19 using Long Short-Term Memory (LSTM).* Journal of big data, 2021. **8**: p. 1-9.
[128] Wu, D.C.W., et al., *Forecasting tourist daily arrivals with a hybrid Sarima–Lstm approach.* Journal of hospitality & tourism research, 2021. **45**(1): p. 52-67.
[129] Elhalwagy, A. and T. Kalganova, *Hybridization of Capsule and LSTM Networks for unsupervised anomaly detection on multivariate data.* arXiv preprint arXiv:2202.05538, 2022.
[130] Kim, J., et al., *Impact of activation functions on flood forecasting model based on artificial neural networks.* Journal of The Korean Society of Agricultural Engineers, 2021. **63**(1): p. 11-25.
[131] Tatachar, A.V., *Comparative assessment of regression models based on model evaluation metrics.* International Research Journal of Engineering and Technology (IRJET), 2021. **8**(09): p. 2395-0056.
[132] Zhang, Y., et al., *Forecasting the prices of crude oil: An iterated combination approach.* Energy Economics, 2018. **70**: p. 472-483.
[133] Reschenhofer, E., *Forecasting volatility: double averaging and weighted medians.* International Journal of Computational Economics and Econometrics, 2010. **1**(3-4): p. 317-326.
[134] Pettenuzzo, D. and F. Ravazzolo, *Optimal portfolio choice under decision-based model combinations.* Journal of Applied Econometrics, 2016. **31**(7): p. 1312-1332.
[135] Rosa, A., *Innovative handling of telemetry enhances gas system automation.* Pipeline & gas journal, 2005. **232**(11): p. 54-57.
[136] Yu, W., et al., *A methodology to quantify the gas supply capacity of natural gas transmission pipeline system using reliability theory.* Reliability Engineering & System Safety, 2018. **175**: p. 128-141.
[137] Özelkan, E.C., A. D'Ambrosio, and S.G. Teng, *Optimizing liquefied natural gas terminal design for effective supply-chain operations.* International journal of production economics, 2008. **111**(2): p. 529-542.
[138] Kim, D., et al., *Development of an FBG-based low temperature measurement system for cargo containment of LNG tankers*. Optics East. Vol. 6770. 2007: SPIE.
[139] Raiyn, J. and T. Toledo, *Real-time short-term forecasting based on information management.* Journal of Transportation Technologies, 2014. **4**(1): p. 11-21.
[140] Stark, T. and D. Croushore, *Forecasting with a real-time data set for macroeconomists.* Journal of Macroeconomics, 2002. **24**(4): p. 507-531.

[141] Massé, M., *REST API Design Rulebook*. 1st edition ed. Representational State Transfer API design rulebook. 2011, Place of publication not identified: O'Reilly Media Incorporated.
[142] Pandey, B.K. and E.R. Schoof, *Building ETL Pipelines with Python: Create and Deploy Enterprise-Ready ETL Pipelines by Employing Modern Methods*. 1 ed. 2023, Birmingham: Packt Publishing, Limited.
[143] Schoemaker, P.J., *How historical analysis can enrich scenario planning.* Futures & Foresight Science, 2020. **2**(3-4): p. e35.
[144] Khuntia, S.R., J.L. Rueda, and M.A. van Der Meijden, *Forecasting the load of electrical power systems in mid-and long-term horizons: a review.* IET Generation, Transmission & Distribution, 2016. **10**(16): p. 3971-3977.
[145] Zhu, R., J. Gordon, and F. Dai, *China's Climate Action Brief: Methane*. 2023, California-China Climate Institute; Available from: https://ccci.berkeley.edu/sites/default/files/China%E2%80%99s%20Climate%20Action%20Brief.pdf.
[146] Rosselot, K.S., D.T. Allen, and A.Y. Ku, *Comparing Greenhouse Gas Impacts from Domestic Coal and Imported Natural Gas Electricity Generation in China.* ACS Sustainable Chemistry & Engineering, 2021. **9**(26): p. 8759-8769.
[147] Edelenbosch, O., et al., *Comparing projections of industrial energy demand and greenhouse gas emissions in long-term energy models.* Energy, 2017. **122**: p. 701-710.
[148] Bloomfield, H., et al., *Quantifying the sensitivity of european power systems to energy scenarios and climate change projections.* Renewable Energy, 2021. **164**: p. 1062-1075.
[149] Aydin, G., *The Modeling and Projection of Primary Energy Consumption by the Sources.* Energy sources. Part B, Economics, planning and policy, 2015. **10**(1): p. 67-74.
[150] Kasperska, E., E. Mateja-Losa, and R. Marjasz, *Sensitivity Analysis and Optimization for Selected Supply Chain Management Issues in the Company – Using System Dynamics and Vensim.* Journal of entrepreneurship, management and innovation, 2013. **9**(2): p. 29-44.
[151] Sumari, S., et al., *Comparing three simulation model using taxonomy: System dynamic simulation, discrete event simulation and agent based simulation.* International Journal of Management Excellence, 2013. **1**(3): p. 54-59.
[152] IEA, *World energy outlook*. 2023: OECD/IEA Paris.
[153] *BP energy outlook.* Report–BP Energy Economics: London, UK, 2023.
[154] Wang, X., et al., *Climate change projections of temperature over the coastal area of China using SimCLIM.* Frontiers in Environmental Science, 2021. **9**: p. 548.
[155] Teske, S., et al., *Net-zero 1.5° C sectorial pathways for G20 countries: energy and emissions data to inform science-based decarbonization targets.* SN Applied Sciences, 2023. **5**(9): p. 252.
[156] Mu, X.-Z., G.-H. Li, and G.-W. Hu, *Modeling and scenario prediction of a natural gas demand system based on a system dynamics method.* Petroleum science, 2018. **15**(4): p. 912-924.
[157] Valipour, M., *Long-term runoff study using SARIMA and ARIMA models in the United States.* Meteorological Applications, 2015. **22**(3): p. 592-598.
[158] Chen, C.-F., et al., *Long-term prediction of emergency department revenue and visitor volume using autoregressive integrated moving average model.* Computational and mathematical methods in medicine, 2011. **2011**.
[159] McRae, S., *Long-term forecasting of regional demand for hospital services.* Operations Research for Health Care, 2021. **28**: p. 100289.
[160] Andrews, B.H., et al., *Building ARIMA and ARIMAX models for predicting long-term disability benefit application rates in the public/private sectors.* Society of Actuaries, 2013: p. 1-54.
[161] Biswas, B., et al., *Forecasting wheat production using ARIMA model in Punjab.* International Journal of Agricultural Sciences, 2014. **10**(1): p. 158-161.
[162] Sutthichaimethee, P. and D. Ariyasajjakorn, *Forecasting energy consumption in short-term and long-term period by using arimax model in the construction and materials sector in thailand.* Journal of Ecological Engineering, 2017. **18**(4): p. 52-59.
[163] Ozturk, S. and F. Ozturk, *Forecasting energy consumption of Turkey by Arima model.* Journal of Asian Scientific Research, 2018. **8**(2): p. 52.

[164] Citroen, N., M. Ouassaid, and M. Maaroufi. *Long term electricity demand forecasting using autoregressive integrated moving average model: Case study of Morocco*. IEEE.
[165] Zhang, G.P., *Time series forecasting using a hybrid ARIMA and neural network model.* Neurocomputing, 2003. **50**: p. 159-175.
[166] Taylor, S.J. and B. Letham, *Forecasting at scale.* The American Statistician, 2018. **72**(1): p. 37-45.
[167] Basak, A., et al., *Drought forecasting using the Prophet model in a semi-arid climate region of western India.* Hydrological Sciences Journal, 2022. **67**(9): p. 1397-1417.
[168] Sivaramakrishnan, S., et al., *Forecasting time series data using arima and facebook prophet models*, in *Big data management in Sensing*. 2022, River Publishers. p. 47-59.
[169] Khayyat, M., et al., *Time Series Facebook Prophet Model and Python for COVID-19 Outbreak Prediction.* Computers, Materials & Continua, 2021. **67**(3).
[170] Jha, B.K. and S. Pande. *Time series forecasting model for supermarket sales using FB-prophet*. in *2021 5th International Conference on Computing Methodologies and Communication (ICCMC)*. 2021. IEEE.
[171] Yenidoğan, I., et al. *Bitcoin forecasting using ARIMA and PROPHET*. in *2018 3rd international conference on computer science and engineering (UBMK)*. 2018. IEEE.
[172] Chan, W.N., *Time series data mining: comparative study of ARIMA and Prophet methods for forecasting closing prices of Myanmar Stock Exchange.* J. Comput. Appl. Res, 2020. **1**(1): p. 75-80.
[173] Haider, S.A., et al., *Deep learning and statistical methods for short- and long-term solar irradiance forecasting for Islamabad.* Renewable Energy, 2022. **198**: p. 51-60.
[174] Luo, Z., et al., *A combined model of SARIMA and prophet models in forecasting AIDS incidence in henan province, China.* International Journal of Environmental Research and Public Health, 2022. **19**(10): p. 5910.
[175] Saqware, G. and I. Beary, *Hybrid SARIMA-Facebook Prophet Model for Prediction and Forecasting of the Staple Food Prices in Tanzania.* International Journal of Operations Research, 2023. **20**: p. 1-11.
[176] Zhao, J. and C. Zhang. *Research on sales forecast based on prophet-SARIMA combination model*. in *Journal of Physics: Conference Series*. 2020. IOP Publishing.
[177] Harvey, A.C. and N. Shephard, *10 structural time series models.* 1993.
[178] Suganthi, L. and A.A. Samuel, *Energy models for demand forecasting—A review.* Renewable & sustainable energy reviews, 2012. **16**(2): p. 1223-1240.
[179] Atasever, S., B. Öztürk, and G. Bilgiç, *A new approach to short-term wind speed prediction: the prophet model.* Energy sources. Part A, Recovery, utilization, and environmental effects, 2022. **44**(4): p. 8826-8841.
[180] Petroleum, B., *Approximate conversion factors*. 2022, Statistical Review of World Energy. p. 2; Available from: https://www.bp.com/content/dam/bp/business-sites/en/global/corporate/pdfs/energy-economics/statistical-review/bp-stats-review-2022-approximate-conversion-factors.pdf.
[181] Cornot-Gandolphe, S., *China's Quest For Gas Supply Security.* IFRI Report, 2019: p. 34.
[182] Kubota, S., *Natural gas trade in Asia and the Middle East.* 1996.
[183] Wang, Z., Y. Kong, and W. Li, *Review on the development of China's natural gas industry in the background of "carbon neutrality".* Natural Gas Industry B, 2022. **9**(2): p. 132-140.
[184] Moghaddam, H., *Role of Natural Gas in China 2050*. 2019, Energy Economics and Forecasting Department (EEFD), GECF; Available from: https://www.gecf.org/_resources/files/events/gecf-expert-commentary---the-role-of-natural-gas-in-china-2050/gecf_expertcommentary_role_naturalgaschina2050.pdf.
[185] Kim, N., *The outlook for China's natural gas demand and LNG imports in 2023*, in *World Energy Market Insight*. 2023, KEEI; Available from: https://www.keei.re.kr/web_energy_new/insight.nsf/0/27F6B42036E9D95A4925897F001E6D72/$File/%5B23-6-%ED%8F%AC%EC%BB%A4%EC%8A%A4-2%5D%20%EC%A4%91%EA%B5%AD%EC%9D%98%202023%EB%85%84%20%EC%B2%

9C%EC%97%B0%EA%B0%80%EC%8A%A4%20%EC%88%98%EC%9A%94%20%EB%B0%8F%20LNG%20%EC%88%98%EC%9E%85%20%EC%A0%84%EB%A7%9D.pdf.
[186] Chavan, H.K. and R.K. Sinharay, *Prophet modeling for oil production forecasting in an enhanced oil recovery field.* Physics of fluids (1994), 2024. **36**(9).
[187] Yusof, U.K., et al., *Financial Time Series Forecasting Using Prophet*. 2021, Springer International Publishing AG: Switzerland. p. 485-495.
[188] Oleksandr, L., *SALES FORECASTS USING CAGR: ADVANTAGES AND DISADVANTAGES.* Editorial board, 2022: p. 272.
[189] Center, P.E.S., *The corridor marked Scrub Corridor on the Project Drawings is a sterile corridor.*
[190] Boyd, R., *Energy and the Financial System : What Every Economist, Financial Analyst, and Investor Needs to Know*. 1st ed. 2013. ed. Energy Analysis. 2013, Cham: Springer International Publishing.
[191] Dheri, G., et al., *Hands-on training on "statistical tools and database management in agriculture".* Icarnahep-caast-snrm department of soil science, 2019. **109**.
[192] Apergis, N. and J.E. Payne, *Natural gas consumption and economic growth: A panel investigation of 67 countries.* Applied Energy, 2010. **87**(8): p. 2759-2763.
[193] Aguilera, H., et al., *Towards flexible groundwater-level prediction for adaptive water management: using Facebook's Prophet forecasting approach.* Hydrological sciences journal, 2019. **64**(12): p. 1504-1518.
[194] IEA, *World Energy Outlook*. 2020: OECD/IEA Paris.
[195] BP, *BP energy outlook* 2020, Accessed.
[196] He, J., et al., *Research on China's long-term low-carbon development strategy and transformation path.* China Popul. Resour. Environ, 2020. **30**: p. 1-25.
[197] Wenyi XI, X.J., Mofan Zhang, et al., *Accelerating the net-zero transition: strategic action for China's 14th Five-Year Plan*. 2020, World Resources Institute.
[198] *China energy and power development prospects*, in *China Electric Power Press*. 2020, State Grid Energy Research Institute Co., Ltd: Beijing.
[199] Organization, G.E.I.D.a.C., *China carbon neutral research report 2060*. 2021.
[200] Zhou, S., J. Wang, and Y. Liang, *Development of China's natural gas industry during the 14th Five-Year Plan in the background of carbon neutrality.* Natural gas industry, 2021. **41**(2): p. 171-182.
[201] Madsen, P.T., et al., *Abandoning fossil fuel production: What can be learned from the Danish phase-out of oil and gas?* Energy Research & Social Science, 2023. **103**: p. 103211.
[202] Dong, H., et al., *Exploring impact of carbon tax on China's CO2 reductions and provincial disparities.* Renewable and Sustainable Energy Reviews, 2017. **77**: p. 596-603.
[203] Lin, B. and Z. Li, *Does natural gas pricing reform establish an effective mechanism in China: A policy evaluation perspective.* Applied Energy, 2021. **282**: p. 116205.
[204] Odoardo, M.D., et al., *Call of duties: how emission taxes on imports could transform the global LNG market*. 2024, Wood Mackenzie; Available from: file:///C:/Users/13773041/Downloads/horizons_march-2024-whitepaper-updated.pdf.
[205] Li, Y., X. Shi, and H. Phoumin, *A strategic roadmap for large-scale green hydrogen demonstration and commercialisation in China: A review and survey analysis.* International Journal of Hydrogen Energy, 2022. **47**(58): p. 24592-24609.
[206] Lipiäinen, S., et al., *Use of existing gas infrastructure in European hydrogen economy.* International Journal of Hydrogen Energy, 2023. **48**(80): p. 31317-31329.
[207] Hui, Y., et al., *Comprehensive review of development and applications of hydrogen energy technologies in China for carbon neutrality: Technology advances and challenges.* Energy Conversion and Management, 2024. **315**: p. 118776.
[208] Zhang, Q., W. Chen, and W. Ling, *Policy optimization of hydrogen energy industry considering government policy preference in China.* Sustainable Production and Consumption, 2022. **33**: p. 890-902.
[209] Office of the Chief Economist, *Future Gas Strategy Analytical Report*. 2024, Department of Industry, Science and Resources, Australian Government Available from: https://www.industry.gov.au/sites/default/files/2024-05/future-gas-strategy-analytical-report.pdf.

[210] Bulavskaya, T., et al., *EXIOMOD 2.0: EXtended Input-Output MODel: A full description and applications*. 2016: Den Haag: TNO.
[211] Leontief, W., *Input-output economics*. 1986: Oxford University Press.
[212] Andre, F.J., M.A. Cardenete, and C. Romero, *Designing public policies : an approach based on multi-criteria analysis and computable general equilibrium modeling*. 1st ed. 2010. ed. Lecture notes in economics and mathematical systems, 642. 2010, Heidelberg: Springer.
[213] Kerimkhulle, S., et al. *The Use Leontief Input-Output Model to Estimate the Resource and Value Added*. 2022. IEEE.
[214] van der Mensbrugghe, D., *Flexible nesting using standard GAMS.* 2024.
[215] Sung, J.E., et al., *Chinese Natural Gas Phase-Out Pathways: A Novel Hybrid Scenario-Specific Projection Approach to Achieve Net Zero.* ENERGY, 2025. **328**.
[216] Leclerc, A., et al., *Toward global and national toxicity footprints.* Journal of Industrial Ecology, 2023. **27**(6): p. 1521-1537.
[217] Peng, X., et al., *The transition to carbon neutrality in China and its impacts on Australia.* China's Transition to a New Phase of Development, 2022: p. 257.
[218] Huang, R., et al., *The effect of technology spillover on CO2 emissions embodied in China-Australia trade.* Energy Policy, 2020. **144**: p. 111544.
[219] Song, L., *Decarbonising China's Steel Industry*, in *The Superpower Transformation: Making Australia's Zero-Carbon Future*, R. Garnaut, Editor. 2022, La Trobe University Press: Melbourne. p. 157-170.
[220] Zhao, A., et al., *Effects of carbon tax on urban carbon emission reduction: Evidence in China environmental governance.* International Journal of Environmental Research and Public Health, 2023. **20**(3): p. 2289.
[221] Wang, S., Y. Zhao, and T. Wiedmann, *Carbon emissions embodied in China–Australia trade: A scenario analysis based on input–output analysis and panel regression models.* Journal of Cleaner Production, 2019. **220**: p. 721-731.
[222] Wei, B., X. Fang, and Y. Wang, *The effects of international trade on Chinese carbon emissions.* Journal of Geographical Sciences, 2011. **21**(2): p. 301-316.
[223] van Leeuwen, E.S., P. Nijkamp, and P. Rietveld, *Regional Input–Output Analysis*, in *Encyclopedia of Social Measurement*, K. Kempf-Leonard, Editor. 2005, Elsevier: New York. p. 317-323.
[224] Gretton, P., *On input-output tables: uses and abuses.* Productivity Commission Staff Research Note, 2013.
[225] Tognoli, M., *A Model for Optimal Economic Lockdown to Contain Epidemic.* Journal of Economic Research & Reviews, 2022. **2**(2).
[226] Pan, J., L. Zhang, and S. Cheng. *Study on the Economic Effects of Dual-Carbon Economy from the Perspect ive of General Equilibrium Theory*. 2023. EAI.
[227] Wei, W., Q. Cui, and Y. Sheng, *Dual carbon goals and the impact on future agricultural development in China: a general equilibrium analysis.* China Agricultural Economic Review, 2022. **14**(4): p. 664-685.
[228] Kemp, J., M. McCowage, and F. Wang, *Towards net zero: Implications for Australia of energy policies in East Asia| Bulletin–September 2021.* 2021.
[229] Government, S. *Computable General Equilibrium modelling: introduction*. 2016; Available from: https://www.gov.scot/publications/cge-modelling-introduction/.
[230] Committee, A.T.A.a.P.A.S., *T4 Computable general equilibrium models in transport appraisal*, in *Australian Transport Assessment and Planning Guidelines* 2021, Australian Department of Infrastructure, Transport, Regional Development and Communications; Available from: https://www.atap.gov.au/sites/default/files/documents/t4-cge-models.pdf.
[231] Aguiar, A., et al., *The Global Trade Analysis Project (GTAP) Data Base: Version 11.* Journal of Global Economic Analysis, 2023. **7**(2).
[232] Scherer, L., A. de Koning, and A. Tukker, *BRIC and MINT countries' environmental impacts rising despite alleviative consumption patterns.* Science of The Total Environment, 2019. **665**: p. 52-60.
[233] IEA, *An energy sector roadmap to carbon neutrality in China*. 2021, International Energy Agency Paris.

[234] Klimek, P., S. Poledna, and S. Thurner, *Quantifying economic resilience from input–output susceptibility to improve predictions of economic growth and recovery.* Nature communications, 2019. **10**(1): p. 1677.
[235] SciPy.org. *SciPy documentation: solve_ivp*. 2025; Available from: https://docs.scipy.org/doc/scipy/reference/generated/scipy.integrate.solve_ivp.html#r179348322575-9.
[236] Do, T.N. and P.J. Burke, *Phasing out coal power in two major Southeast Asian thermal coal economies: Indonesia and Vietnam.* Energy for Sustainable Development, 2024. **80**: p. 101451.
[237] Ye, Q., *Commodity booms and their impacts on the Western Australian economy: The iron ore case.* Resources Policy, 2008. **33**(2): p. 83-101.
[238] Eikeland, P.O. and T.H.J. Inderberg, *Energy system transformation and long-term interest constellations in Denmark: can agency beat structure?* Energy Research & Social Science, 2016. **11**: p. 164-173.
[239] Zander, K.K., et al., *Preferences for and potential impacts of financial incentives to install residential rooftop solar photovoltaic systems in Australia.* Journal of Cleaner Production, 2019. **230**: p. 328-338.
[240] Li, J., J. Jiao, and Y. Tang, *An evolutionary analysis on the effect of government policies on electric vehicle diffusion in complex network.* Energy Policy, 2019. **129**: p. 1-12.
[241] Esmaeili, M., et al., *On the Role of Renewable Energy Policies and Electric Vehicle Deployment Incentives for a Greener Sector Coupling.* IEEE Access, 2022. **10**: p. 53873-53893.
[242] Jaegersberg, G. and J. Ure, *Renewable Energy Clusters: Recurring Barriers to Cluster Development in Eleven Countries*. 1 ed. Innovation, Technology, and Knowledge Management. 2017, Cham: Springer International Publishing AG. 1-276.
[243] OECD, *OECD Composite Leading Indicators: a Tool for Short-Term Analysis*. 1998.
[244] Tu, K.J. and S. Johnson-Reiser, *Understanding China's rising coal imports.* Policy outlook, 2012(Carnegie Endowment for International Peace): p. 16.