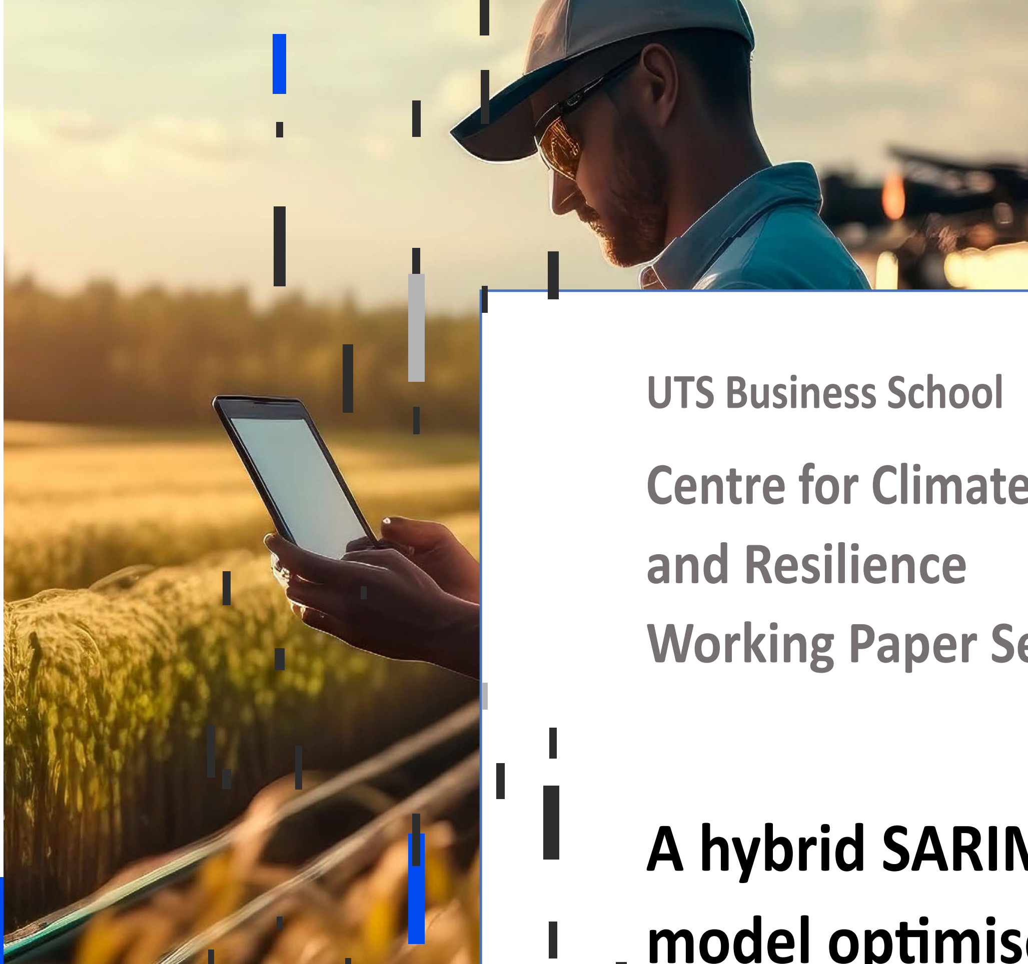


UTS Business School

Centre for Climate Risk
and Resilience
Working Paper Series

A hybrid SARIMAX-LSTM model optimised by ANN for near-term forecasting: An application to China's Natural Gas consumption

By Jinyong (Edward) Sung , Xunpeng (Roc) Shi, Sven Teske, Mengheng Li

Working Paper No. 2025/1

Series Editor: Martina Linnenluecke, CCRR Director

Connect with CCRR
Web https://www.uts.edu.au/about/uts-business-school/research/centre-climate-risk-and-resilience
Email ccrr@uts.edu.au

A hybrid SARIMAX-LSTM model optimised by ANN for near-term forecasting: An application to China's Natural Gas consumption

Jinyong (Edward) Sung [a], Xunpeng (Roc) Shi [b], Sven Teske [a], Mengheng Li [c]

[a] Institute for Sustainable Futures, University of Technology Sydney (UTS), Sydney, Australia

[b] Australia-China Relations Institute, University of Technology Sydney (UTS), Sydney, Australia

[c] Economics Discipline Group, UTS Business School, Sydney, Australia

Abstract: China's Natural Gas (NG) consumption is expected to undergo medium- and short-term growth in line with China's carbon neutrality policy. 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 need for more precise forecasting, we develop a hybrid model that optimally combines SARIMAX and LSTM through ANN for one-year-ahead forecasting of Chinese NG consumption, utilising eight identified key drivers of the NG market. This study evaluates the forecasting performance of all models – single SARIMAX, LSTM, and hybrid models – based on three metrics (R^2, RMSE, and MAE) and Cumulative Squared Forecast Errors (CSFEs). The results 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 different characteristics as inputs for hybrid models through PCA, but also identifies NG-related infrastructure factors as the most significant contributors to the forecasting. Our findings suggest policy, investment, and operational implications regarding near-term NG demand for government agencies or industry experts.

Keywords: Hybrid forecasting methods; SARIMAX; LSTM; ANN; Natural Gas; China

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 [1]. In this journey toward 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 [2]. 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. To address these uncertainties in the NG market, the importance of accurate forecasting of the NG consumption is rising. Thus, we aim to develop a high-performing one-year-ahead forecasting model using Chinese NG consumption as a case to study its merits. The forecasting model is deemed highly crucial for energy industry stakeholders in making near-term investment or operational decisions.

Traditional 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 [3-5]. 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 [5]. Representative traditional time series forecasting techniques include ARIMA-based models, while deep learning models encompass Artificial Neural Networks (ANN), Recurrent Neural Networks (RNN), 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 [6]. 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 US until 2025, comparing the performance of each model [7]. This study indicates an expected increase in US NG production (16%) and consumption (24%) by 2025. Vagropoulos et al. compare forecasting for grid-connected photovoltaic (PV) plants using SARIMA, SARIMAX, modified SARIMA (resulting from a posteriori modification of the SARIMA model), and ANN-based models [8]. 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 [9-14].

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 [4]. Cerit et al. employ SARIMAX, ANN, and hybrid SARIMAX-ANN models for seaborne transport forecasting, showing that the hybrid model outperforms the others [15]. 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 [16].

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 [17] as well as market instability caused by endogenous factors such as gas price regulations [18]. These factors can engender irrational patterns or substantial fluctuations in the NG consumption, which existing hybrid models cannot effectively capture [18]. 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 [19]. 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 reflecting China's current NG market conditions, to accurately predict China's NG consumption with significant volatility, a task beyond the capabilities of existing hybrid prediction models.

This research seeks to create a hybrid model optimally combining traditional forecasting 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 current 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 [20]. 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, we introduce a novel hybrid forecasting model that optimally amalgamates a traditional time-series forecasting model (SARIMAX) and a deep learning model (LSTM) through a deep learning approach (ANN) within the energy field. 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 the selected key drivers in China's NG market [20] as inputs. This innovative amalgamation can ultimately demonstrate its highest one-year-ahead predictive performance. Second, we not only demonstrate that these key drivers with different characteristics [20] 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-, and other energy-related 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 policy or investment implications regarding near-term NG demand for government agencies or industry experts.

The rest of this paper is organised as follows: Sections 2 and 3 discuss data and methods for the forecasting models, respectively. Section 4 presents the models' application as the results, and the final section concludes the study. The supplementary materials include terms and abbreviations, tables, and figures used throughout this paper.

2. Data

This study is conducted employing eight key driving factors as independent variables and Chinese NG consumption as the dependent variable.[1] 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 [20]. These eight factors are Price distortion, Temperature, HH gas price, piped NG (PNG) imports, Thermal coal price, Gas pipeline capa, 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 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[2], include economic developments, energy policies, environmental impacts, technological progress, and global market dynamics [20]. Particularly, these factors reflect NG's paradoxical role in achieving carbon neutrality – a short-term transitional fuel [21] and long-term phase-out [22]. Besides, these dynamics are influenced by the uncertainties addressed by these factors – the short-term uncertainties from NG market liberalisation [23, 24] and geopolitical developments [25, 26] and long-term uncertainties from electricity demand variations tied to carbon-neutral initiatives [27]. These eight factors are selected based on these shifting dynamics and uncertainties, considering their short- and long-term effects. Therefore, 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

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

[2] 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).

ADF test, a widely used unit root test, in Table 2 (see Appendix B). All variables except Price distortion are non-stationary at level and stationary at the first difference. SARIMAX can be applied to non-stationary variables [8, 28]. Deep learning models can learn complex patterns and dynamic relationships in data, allowing them to handle non-stationary variables [29]. Thus, the non-stationary key drivers can be directly utilised for hybrid SARIMAX-LSTM modelling. Detailed properties of the variables are in Table 2 (see Appendix B).

Table 1 Key driving factors

Factors	Description	Data source[3]
(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	External source[4]
(4) Climate (1 factor)		
Average national temperature (°C, Celsius)[5]	Monthly series	Bloomberg
(5) Price distortion (1 factor)		
Price distortion (City gate gas price[6] / Piped gas price[7], %)	Monthly series	Internal source[8], Bloomberg, and, Estimation[9]

3. Methods and Models

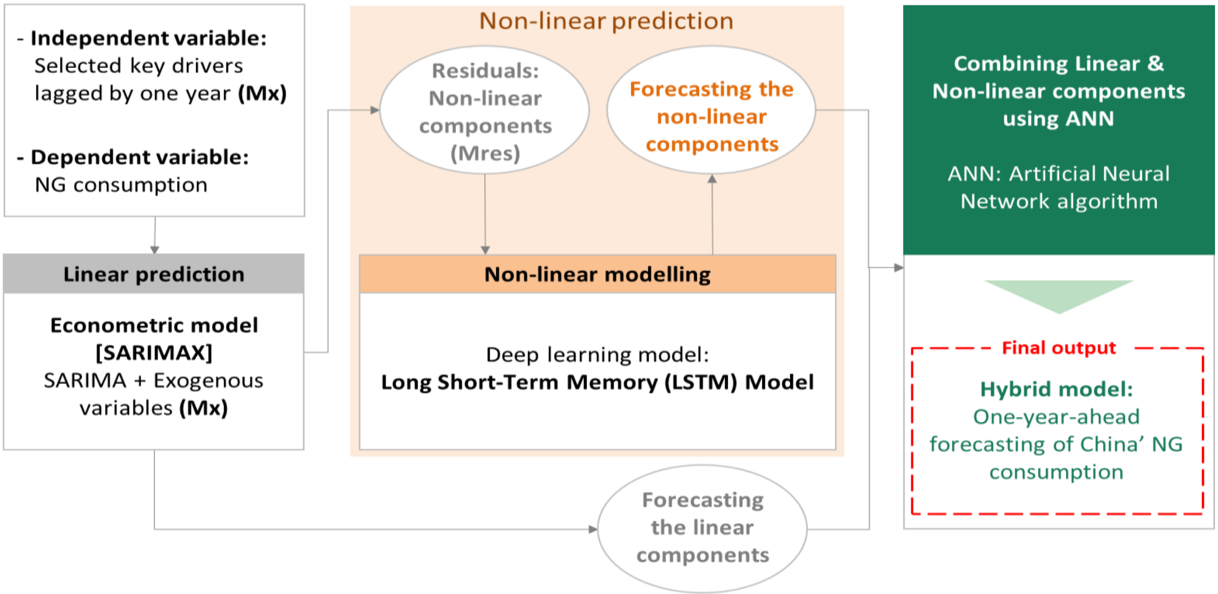


Figure 1 Hybrid model schematic

The hybrid model, comprising SARIMAX and LSTM through ANN, utilises a

[3] The monthly exchange rates, referenced from Bloomberg, are applied.
[4] This reference is taken from Chen et al.'s research [30].
[5] This factor is calculated as the average temperature of all provinces per month.
[6] The city gate gas price from December 2005 to June 2013 is based on Paltsev and Zhang's research [31]. 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.
[7] The piped gas price is referenced from Bloomberg.
[8] UTS ACRI provides the internal source.
[9] Price distortion variables are calculated using the formula: (City gate gas price / Piped gas price - 1) ×100.

methodology that sequentially calculates linear and non-linear predictions and then optimally combines these. As depicted in Figure 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 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. More details on the hybrid model are in the following sections.

3.1 Linear prediction

3.1.1 SARIMAX

SARIMAX is a development of seasonal ARIMA (SARIMA) with exogenous factors (X), known as SARIMAX (p, d, q) (P, D, Q)^S [7]. SARIMA represents an enhancement of the ARIMA framework. It has been introduced to improve the effectiveness of auto-regressive (AR) integrated moving average (MA) models specifically tailored for modelling time series data with seasonal patterns [7]. SARIMAX is mathematically expressed as follows:

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

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.
- $\boldsymbol{\nabla^d}$ $(=\ 1 - B)^d$ 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.

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:

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

In the data pre-processing phase, all exogenous variables are first standardised within the range of -1 and 1 through MinMaxScaler[10], a famous scaling algorithm generally used in machine/deep learning [33]. The default range of MinMaxScaler is 0 to 1. Due to negative values in the data, we adjust the scaling range of MinMaxScaler to -1 and 1 [34, 35]. MinMaxScaler with the range of -1 and 1 is applied in the data pre-processing stage of all models in this study.

Subsequently, we split 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 [7, 36]. 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, we employ 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 [37]. To demonstrate the robustness of the previously identified optimal order, we 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.[11] To find methods for enhancing forecasting accuracy, this 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 (see Appendix C) [38]. The validated residuals are used for non-linear prediction through LSTM.

3.2 Non-linear prediction

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 Long Short-Term Memory (LSTM) model.

Long Short-Term Memory

[10] MinMaxScaler is within scikit-learn, a popular machine learning library in Python [32]. The code for all models in this study has been written in Python 3.7.13 environment using Jupyter Notebook, an open-source interactive editor.

[11] When the parameter is enabled, 'trend = ct' is set in SARIMAX, where 'ct' represents a constant multiplied by a linear trend with time.

The LSTM model is a specific type of RNN (see Appendix D) particularly designed to capture long-term dependencies and retain information over extended periods [10, 16]. RNN exhibits performance degradation in long-term sequences. LSTM has been introduced by S. Hochreiter and J. Schmidhuber in 1997 to overcome this degradation [39]. Figure 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/output, incorporating information from the current input/output and the hidden state from the previous time steps. The three gates enable LSTM to effectively capture long-term dependencies (see Appendix E).

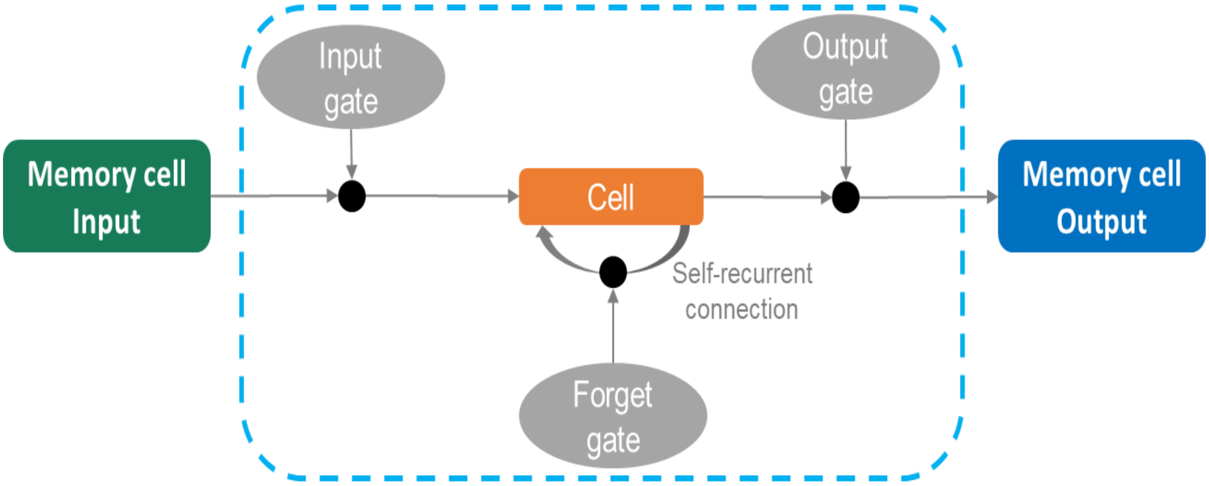


Figure 2 Schematic diagram of LSTM

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, China's NG consumption has continually increased since 2010 and has had cyclical fluctuations. There are apparent non-linear patterns that a linear trend 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.

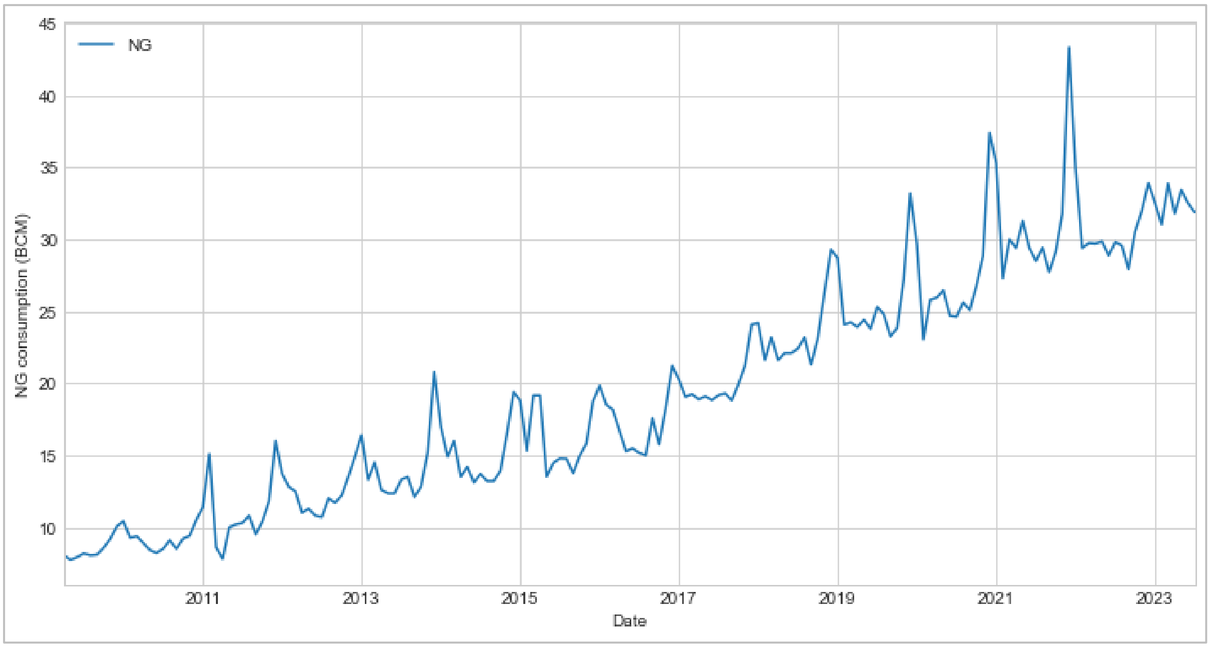


Figure 3 China's NG consumption

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 3.3.1). 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 [40]. 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 [41]. Therefore, the generalisation of the LSTM network is improved by employing a training method that involves shuffling the data rather than training it sequentially [42]. In a recent study, the best results on test data for LSTM are achieved by shuffling the input data [41].

<u>Phase 2: Model architecture</u>

This study employs the LSTM model structure, drawing inspiration from the machine and deep learning literature [33, 43] as well as recent relevant studies [44, 45]. 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 4 provides the specific model architectures of LSTM 1 and 2.

Table 4 Mode architectures of LSTM 1 and 2

Layer	LSTM 1	LSTM 2
LSTM layer	128 nodes with a 'tanh' activation function[12] and a 'dropout rate' of 20%[13]	
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	

<u>Phase 3: Model compilation and training</u>

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 and additional metric, respectively [47]. Furthermore, Adam, one of the most highly efficient optimization algorithms among practitioners, is selected as the optimiser [46, 48].

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[14] [46, 49]. In line with recent research findings, this study employs 100 epochs [15] [50]. Early stopping was implemented with a threshold of 5.[16]

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.

[12] The hyperbolic tangent function ('tanh') is a mathematical function that maps its input to the range between -1 and 1. In other words, the output of the tanh function is constrained within the interval [-1, 1] [46].
[13] Dropout is a regularisation technique for a neural network by randomly dropping out the nodes in the input and hidden layers [46].
[14] Batch size is the number of data samples used in a single forward and backward pass during each training iteration. [46].
[15] An epoch indicates one complete pass through the entire training dataset [46].
[16] Early stopping is the feature that halts training before overfitting occurs [46]. 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.

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 [51]. 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 integrity of its readings [52].

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 3.2.2).

<u>Addictive combination</u>

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 [4]. 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.

3.3.2 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:

<u>Artificial Neural Network</u>

An Artificial Neural Network (ANN) is a data-driven mathematical tool that demonstrates the capacity to address problems through machine-learning neurons [4, 8, 10]. ANN, as a general, flexible, and non-linear tool, can approximate any type of relationship between inputs and outputs [4]. This facilitates the optimal combination between the linear and non-linear predictions of China's NG consumption. Figure 4 provides a schematic diagram of ANN with six layers utilised in this study.

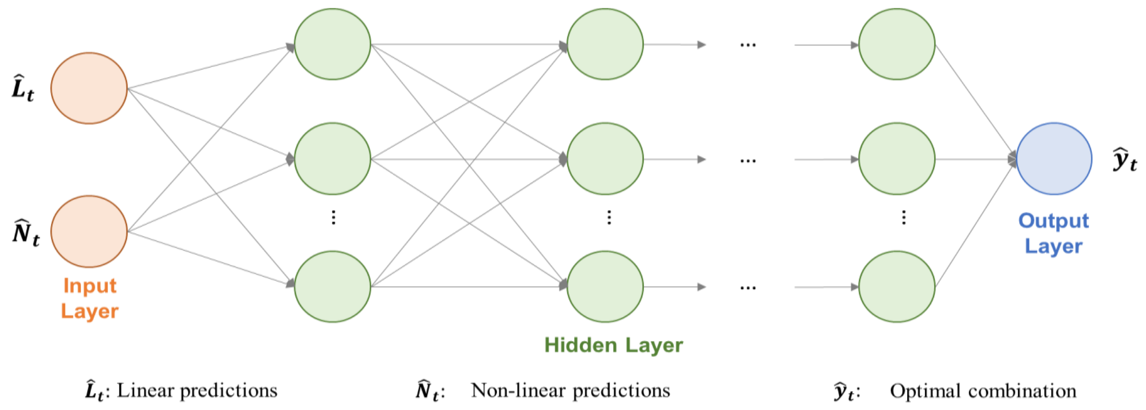


Figure 4 The structure of ANN

Optimal combination

The ANN-based optimal combination of Hybrid 2 is conducted using the linear and non-linear predictions as inputs. Three phases are involved in the optimal combination: data pre-processing, model architecture, and model compilation and training. The data pre-processing phase of ANN mirrors the data pre-processing in the LSTM model in Hybrid 1, except for the train-test split ratio – 80% as a training set and 20% as a test set. The model structure of the ANN consists of the input layer, followed by the first two dense layers with 'tanh' activation functions and the next two dense layers with 'elu' activation functions, and concludes with the output layer featuring a single node applying a linear transformation to its input. These choices for the activation functions stem from the research suggesting the efficacy of 'tanh' and 'elu' as continuous functions that can exhibit relatively high accuracy in the ANN-based continuous time series analysis [53]. The compilation and training phases of the ANN follow the same methods as introduced earlier for LSTM 1 and 2. Table 5 presents the detailed model architecture of the ANN.

Table 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[17]
2nd dense layer	64 nodes with a 'tanh' activation function
3rd dense layer	32 nodes with a 'elu' activation function[18]
4th dense layer	16 nodes with a 'elu' activation function
Output layer	1 node with a 'linear' activation function

[17] The hyperbolic tangent function ('tanh') is the same as that introduced for an LSTM model (See Section 3.2.2).
[18] 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 [53].

The linear and non-linear predictions of Hybrid 2 are the same as those in Hybrid 1. In addition, depending on the SARIMAX model – SARIMAX 1 or 2 – for the calculation of the liner predictions, Hybrid 2 is also categorised into a hybrid model combined with SARIMAX 1 (labelled as Hybrid 2-S1 hereafter) and another hybrid model combined with SARIMAX 2 (referred to as Hybrid 2-S2 hereafter), similar to Hybrid 1.

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 Mean Absolute Error (MAE) defined by equation (6) [54].

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 in various economic and econometric fields, including energy economics, computational economics and econometrics, and applied econometrics [55-57]. 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).

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

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

$$\boldsymbol{MAE = \frac{1}{n}\sum_{i=1}^{n}|\,y_i - \hat{y}_i|} \quad \textbf{(6)}$$

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

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

3.5 Principal Component Analysis (PCA)

As an additional study related to the hybrid forecasting model, this study explores the suitability of using the identified key drivers with different characteristics [20] as inputs for the forecasting model through Principal Component Analysis (PCA).

PCA is a popular machine learning algorithm for dimensionality reduction [46]. In machine learning, the dimensions of data are commonly referred to as features. To reduce the dimensions of the data, a PCA model is initially employed to find directions with significant variance in the data. A vector in the direction of maximum variance is termed the principal component (PC). Subsequently, the original data is projected onto this direction of maximum variance to decrease the data dimensions. In other words, the original data is effectively reduced in dimensionality using the PC. Since the PC represents the direction with the highest variance, the data transformed by projecting onto the PC best captures the characteristics of the original data.

Scikit-learn, a machine learning library, is utilised to perform PCA in this study. This PCA process involves first determining the appropriate number of dimensions and secondly, measuring the model performance with the transformed data using the chosen dimensions. In the initial step of the PCA process, a graph of explained variance ratio against the number of dimensions is created to decide on the appropriate number of dimensions to reduce. The explained variance ratio represents how well the PC captures the original data. By observing the graph of the explained variance ratio, one can identify dimensions where there is little significant change in explained variance, and these dimensions are considered appropriate for reduction. In the subsequent step, the data transformed by PCA is used to train the hybrid forecasting model. The performance of the model with the PCA-derived data is then compared with that of the model with the raw data to identify any differences.

4. An application to forecast near-term NG consumption in China

4.1 Model comparison

Model comparisons are initially conducted based on three metrics - R^2, RMSE, and MAE - in the following order: SARIMAX vs. LSTM, LSTM vs. Additively combined hybrid SARIMAX-LSTM models (Hybrid 1), and Hybrid 1 vs. ANN-based optimally combined hybrid SARIMAX-LSTM models (Hybrid 2). Additionally, the comparisons are also performed based on the CSFE. Consequently, to create a high-performing prediction model for one-year-ahead Chinese NG consumption using the key drivers, the best choice is the hybrid model (Hybrid 2-S2) that optimally merges SARIMAX with a trend capture function and LSTM via ANN. The detailed comparison results are as follows.

4.1.1 Comparison 1: SARIMAX vs. LSTM

SARIMAX 1 and SARIMAX 2 with the optimal order determined as (2, 1, 1) (0, 1, 1, 12) and (0, 0, 2) (1, 1, 0, 12), respectively, generate predictions from December 2020. The Ljung-Box and Jarque-Bera test results for two SARIMAX models indicate their good fit, as shown in Table 3 (see Appendix C). LSTM 1 and 2 models commence their predictions from September 2019. Forecasting performance for all the models is measured using three metrics: R^2, RMSE, and MAE. Figure 5 and Table 6 compare the models in terms of the metrics.

As shown in Figure 5 and Table 6, when comparing the SARIMAX models, SARIMAX 2 (R^2 = 21.25%, RMSE = 7.92) obtains a higher R^2 value and lower RMSE compared to SARIMAX 1 (R^2 = 12.55%, RMSE = 8.79), while SARIMAX 1 (MAE = 2.12) exhibits a lower MSE value compared to SARIMAX 2 (MAE = 2.21). Among the three metrics, SARIMAX 2 outperforms SARIMAX 1 in R^2 and RMSE, indicating that SARIMAX 2, with the enabled trend-capturing parameter, excels in forecasting ability.

In the comparison of the LSTM models, LSTM 2 (R^2 = 38.74%, RMSE = 8.94, MAE = 1.99) surpasses LSTM 1 (R^2 = 36.46%, RMSE = 9.27, MAE = 2.01) across the three metrics – achieving a higher R^2 value, lower RMSE, and lower MAE, thereby demonstrating superior forecasting ability in LSTM 2. Furthermore, when compared with SARIMAX models, the LSTM models exhibit higher R^2 and lower MAE values, showcasing their excellent forecasting ability. This aligns with findings in the literature review, which consistently highlight the superior forecasting performance of LSTM-based models.

Among SARIMAX and LSTM models, LSTM 2 stands out as the most superior, exhibiting the highest R^2 value and the lowest MAE. Therefore, for the non-linear (residual) predictions in Hybrid 1 and 2, LSTM 2 is employed. The graphs for SARIMAX and LSTM models are presented in Figure 6 (see Appendix F).

Figure 5 Comparison of SARIMAX, LSTM, and Hybrid models

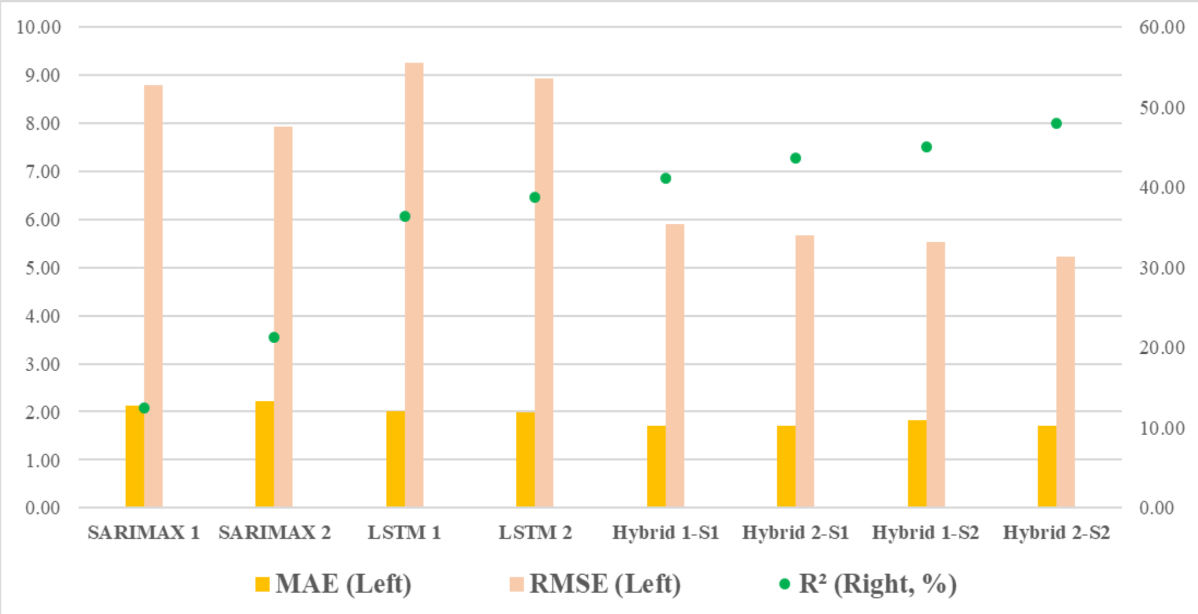


Table 6 Metrics of SARIMAX, LSTM, and Hybrid models

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: Results are rounded to two decimal places; Details in SARIMAX 1, 2, LSTM 1, and 2 are provided in Section 3.1 and 3.2; Details in Hybrid 1-S1, S2, Hybrid 2-S1, and S2 are provided in Section 3.3; S1 and S2 denote SARIMAX 1 and SARIMAX 2, respectively.

4.1.2 Comparison 2: LSTM vs. Hybrid 1

As shown in Table 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). 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. This indicates that Hybrid 1-S1 and S2 outperform LSTM 1 and 2 in terms of the three metrics: higher R^2, lower RMSE, and MAE. Consistent with findings in existing studies, these hybrid models demonstrate superior performance over the individual LSTM models. This also indicates that the hybrid models are superior to single SARIMAX models. In conclusion, these results are in line with discussions in the literature, where hybrid forecasting models demonstrate better accuracy compared to individual traditional time-series or deep learning forecasting models [4, 15, 16].

4.1.3 Comparison 3: Hybrid 1 vs. Hybrid 2

As indicated in Table 6, among the Hybrid 1 models – Hybrid 1-S1 and S2, Hybrid 1-S2 demonstrates superior performance as Hybrid 1-S2 (R^2 = 45.08%, RMSE = 5.52) obtains a higher R^2 value and lower RMSE compared to Hybrid 1-S1 (R^2 = 41.22%, RMSE = 5.91). Among the Hybrid 2 models – Hybrid 2-S1 and S2, Hybrid 2-S2 exhibits superior performance as Hybrid 2-S2 (R^2 = 47.94%, RMSE = 5.23, MAE = 1.70) achieve a higher R^2 value and lower RMSE and MAE compared to Hybrid 2-S1 (R^2 = 43.63%, RMSE = 5.67, MAE = 1.71). These results indicate that the hybrid models exhibit significantly higher performance, incorporating SARIMAX 2 (S2) with the enabled trend-capturing parameter. This underscores the crucial role of capturing the trend in linear components to enhance the forecasting ability of the hybrid SARIMAX-LSTM model.

Table 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 shown in Table 7, the hybrid models demonstrate a synergistic effect between SARIMAX and LSTM models. The R^2 values for SARIMAX 1 and 2 are 12.55% and 21.25%, respectively, while the R^2 values for LSTM models, computed using residuals derived from SARIMAX 1 and 2, are 18.41% and 25.73%, respectively. In contrast, the hybrid models achieve a R^2 value of at least 41.22% (Hybrid 1-S1). This indicates that creating a hybrid model by combining the two has a synergistic effect.

As depicted in Table 6, in the comparison between Hybrid 1 and 2 models, where the same SARIMAX model (S1 or S2) is combined, Hybrid 2-S1 (R^2 = 43.63%, RMSE = 5.67) demonstrates a higher R^2 value and lower RMSE compared to Hybrid 1-S1 (R^2 = 41.22%, RMSE = 5.91), and Hybrid 2-S2 (R^2 = 47.94%, RMSE = 5.23, MAE = 1.70) obtains a higher R^2 value, lower MAE, and RMSE compared to Hybrid 1-S2 (R^2 = 45.08%, RMSE = 5.52, MAE = 1.82). These findings suggest that the Hybrid 2 models surpass the Hybrid 1 models. This implies that optimally combining SARIMAX and LSTM through ANN leads to significantly better predictions than simply adding them together.

In conclusion, to develop a high-performing forecasting model for one-year-ahead Chinese NG consumption using the key drivers, it is most appropriate to use Hybrid2-S2 that optimally combines SARIMAX with the trend capture function and LSTM through ANN. The graphs for all the hybrid models are presented in Figure 7 (see Appendix G).

4.1.4 Cumulative Squared Forecast Error (CSFE)

In addition to the three metrics, the performance of models is also compared using the CSFE. The CSFE of each model is calculated from December 2020 to July 2023, which is the same period as the test data for most models. Due to the window size of LSTM and the train-test split ratio, which differ from that of SARIMAX, the start date for the test data in LSTM 1 and 2 is September 2019. In contrast, for other models, the start date for the test data is December 2020. Therefore, for comparison among models, the start date for the test data used in CSFEs is set to December 2020. Figure 8 displays the CSFEs for all models.

The model that ultimately exhibits the smallest CSFE is considered the most superior, and evaluating whether it reduces CSFEs over time is to assess its predictive ability. A decrease in CSFEs indicates that the model becomes more accurate in its predictions as time progresses. As shown in Figure 8, the models exhibiting relatively lower CSFEs among all models are the Hybrid 1 and 2 models. Among all the Hybrid 1 and 2 models, the ones combined with SARIMAX 2 (S2) show relatively lower CSFEs than those combined with SARIMAX 1 (S1). This demonstrates that S2 incorporated with the trend-capturing function enhances predictive ability.

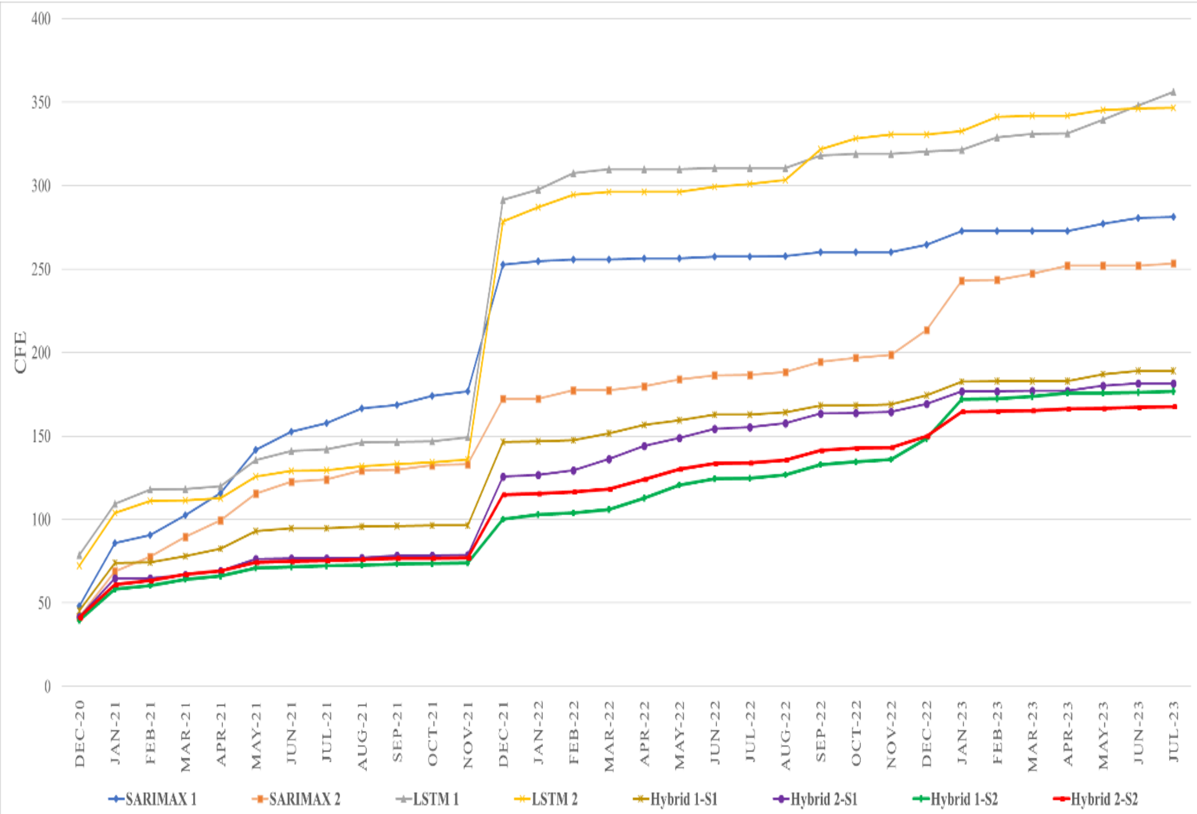
Similar to the comparison of the three metrics, in the comparison of CSFEs as well, hybrid 2-S2 (the red line) stands out as the most superior in predictive ability among all models. When comparing Hybrid 1 and 2 models with lower CSFEs – Hybrid 1-S2 and Hybrid 2-S2, initially, Hybrid 1-S2 exhibits lower CSFEs compared to Hybrid 2-S2, but, over time, Hybrid 2-S2 shows relatively lower CSFEs than Hybrid 1-S2. This indicates that Hybrid 2-S2 model's predictions become more accurate as time progresses. Moreover, Hybrid 2-S2 demonstrates excellent predictive capabilities for sudden changes. CSFE calculations are based on squared errors, which means they impose a significant penalty even for small errors. As shown in Figure 7 (see Appendix G), we can observe a sudden increase in Chinese NG consumption at the end of 2022. As illustrated in Figure 8, the CSFEs of Hybrid 2-S2 increase the least during this time. This indicates that Hybrid 2-S2 accurately predicts such abrupt changes in the NG consumption.

However, the LSTM models – LSTM 1 and 2 – do not accurately predict such sudden increases. In the comparison of the three metrics, the LSTM models show better performance than the SARIMAX models – S1 and 2. In the comparison of CSFEs as well, initially, the LSTM models exhibit lower CSFEs compared to the SARIMAX models. But, over time, the LSTM models increase their CSFEs, ultimately showing higher CSFEs than the SARIMAX models. As depicted in Figure 8, the point at which the CSFEs of the LSTM models more sharply increase than those of the SARIMAX models is from the end of 2021, when there is a relatively pronounced increase in the NG consumption. This indicates that the predictive ability of the LSTM model for sudden increases is relatively inferior compared to that of the SARIMAX model.

In the CSFE comparison with the SARIMAX models, they ultimately exhibit lower CSFEs than the LSTM models but show higher CSFEs than the hybrid models. Among the SARIMAX models, while the final CSFE of S1 is higher than that of S2, the CSFEs of S2 increase steeply over time. Particularly, the CSFEs of S2 rise most sharply at the point of the sudden increase in the NG consumption at the end of 2022. This indicates that when the NG consumption is predicted solely by S2 – the SARIMAX model with the trend-capturing function, S2 fails to accurately predict sudden changes in the NG consumption. However, as observed in Hybrid 2-S2, the SARIMAX models with the trend-capturing function, optimally combined through ANN in Hybrid models, enhance prediction accuracy.

In a nutshell, like the comparison of the three metrics, in the comparison of CSFEs, the optimal choice for creating an effective one-year-ahead prediction model of the NG consumption with key drivers is also Hybrid 2-S2, which optimally combines S2 and LSTM via ANN. Furthermore, the Ljung-Box and Jarque-Bera test results for Hybrid 2-S2 validate the robustness of its output, the ANN-optimised combination of linear SARIMAX and nonlinear LSTM predictions, as shown in Table 3 (see Appendix C).

Figure 8 Cumulative Squared Forecast Errors of Models



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 [20]. For this investigation, the PCA-transformed data is applied to hybrid models, and then their forecasting performances are compared.

First of all, the principal components (PCs) reflecting the significant features of the data are selected through a graph analysis of the explained variance ratio against the number of components [33, 46]. As indicated in Figure 9, it is evident that the first four PCs encapsulate the majority of the explained variance. Subsequent components demonstrate relatively smaller contributions to the variance. Thus, the number of PCs is chosen as four. The transformed data, obtained by applying these PCs, is then employed across all hybrid models, following the modelling approach introduced for Hybrid 1 and 2 models above. The hybrid models – Hybrid 1-S1 and S2 and Hybrid 2-S1 and S2 – incorporating PCA are denoted as Hybrid 1P-S1, Hybrid 1P-S2, Hybrid 2P-S1, and Hybrid 2P-S2, respectively.

Figure 9 Explained variance ratio

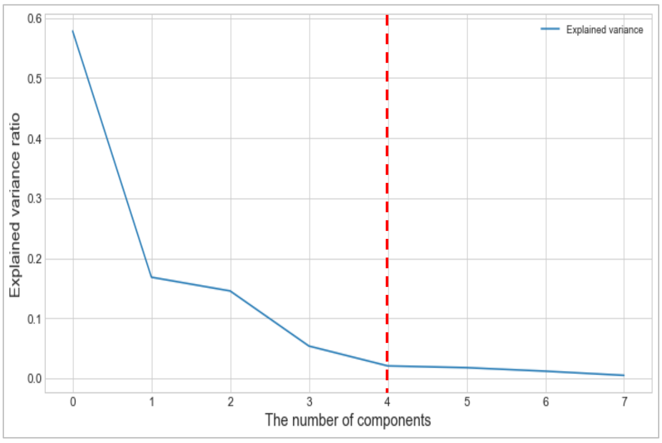


Table 8 Metrics of Hybrid models with PCA

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 Table 6 and 8, hybrid models with PCA achieve higher R^2, lower MAE, and RMSE than those without PCA. This implies that the PCs, which capture the significant features of the key drivers and remove their unnecessary noise, contribute to enhancing model performance. It also indicates that the key drivers possess important features for the hybrid models. Thus, the key drivers with different characteristics serve as valid inputs for the forecasting.

However, with the PCA-derived data alone, discerning the factors that significantly influence the forecasting is challenging. Therefore, an investigation is additionally conducted to identify the factors among the key drivers that contribute the most to the forecasting. The graphs for the hybrid models with PCA are presented in Figure 10 (see Appendix H).

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. The model utilises data lagged by one year and is designed to incorporate trends, implying that changes in the input may influence the positive or negative direction of predicted results, which are affected by lagged elements and trends. However, in identifying contributing factors, the study considers only the absolute value of the changes in the output, excluding the direction of changes.

As shown in Figure 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 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 rankings across all multiples of input. Apart from NG-related infrastructure factors, upon examining the factors highlighted in green (3rd place) and blue (4th place), PNG imports, LNG imports, and temperature can be identified as factors influencing output changes in 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 [5].

The remaining three factors, Thermal coal price, Price distortion, and HH gas price factors, have a relatively smaller impact on output changes than other factors and also exhibit similar output changes. Therefore, it is difficult to categorise them as contributing factors. From this analysis, it is evident that among the key driving factors, NG-related infrastructure factors have the most significant impact on the forecasting.

Figure 11 The output changes for each multiple of input

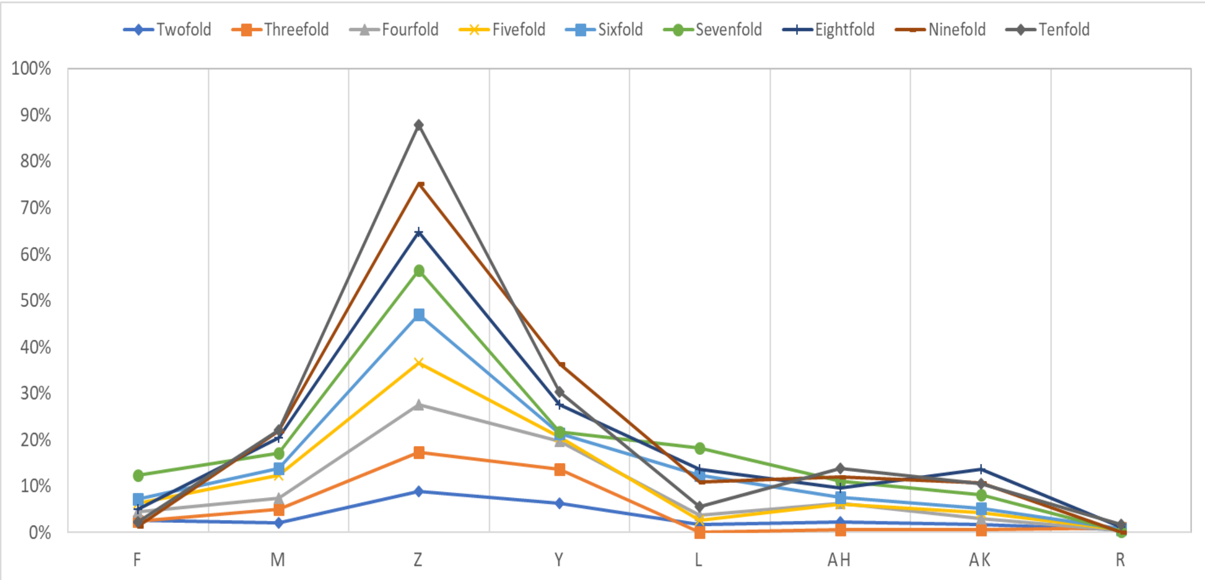


Table 9 The ranks of the output changes

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.

Practical application: Enhancing forecasting and gas networks via telemetry data

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 through the telemetry control system of the gas transportation system [58]. Gas pipeline capacity and Accumulated LNG terminal capacity can be estimated in real-time through the telemetry control system, based on pressure and flow rate [59], and storage volume and real-time send-out capacity [60], respectively. Besides, PNG imports can be measured in real time with pipeline flow sensors, LNG imports inside transport tanks, and temperature via pipeline or storage facility monitors [58, 61].

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 significant in short-term forecasting than the latest historical data, as it better detects anomalies, thus improving forecasting performance [62, 63]. Therefore, the real-time factors can enhance the high-performing hybrid forecasting model (Hybrid 2-S2) in this study. This improved model, in conjunction with the telemetry control system, can be used to design an efficient system that prepares safety measures in advance. For example, if the predicted NG consumption increases sharply, the system can automatically adjust pipeline pressure, flow rate, terminal storage volume, and send-out rate through the telemetry control system, improving the system's stability and overall efficiency.

The telemetry control system and Hybrid 2-S2 can be integrated through an Application Programming Interface (API) and a data pipeline, ultimately connecting with gas distribution networks. An API enables data exchange between software and systems [64], and a data pipeline collects, transforms, and transfers data to storage, thereby connecting the model with the telemetry control system [65]. Developed in Python, an open-source programming language, the Hybrid 2-S2 model can be integrated with a web-based API, REST API [64], and an ETL data pipeline [65]. 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. Consequently, this system will not only automate its safety and operations but also optimise the efficiency of the gas networks. Further details on this practical application are left for future research.

5. Conclusions

This study explores the optimal combination of a traditional time-series prediction model and a deep learning model. The hybrid model, which optimally combines SARIMAX and LSTM through ANN, indicates superior performance compared to single SARIMAX, LSTM, and additively combined hybrid SARIMAX-LSTM models. This is shown by comparing the forecasting performance of single and hybrid models for one-year-ahead Chinese NG consumption, using the identified key drivers from April 2009 to July 2023 as inputs [20]. Moreover, as extensive research on the key drivers, the study not only conducts a validity analysis of these inputs – the key drivers with different characteristics – for the hybrid model through PCA, but also delves into the factors most significantly affecting the forecasting among the drivers.

The first significant finding is that optimally combined hybrid forecasting models through ANN (Hybrid 2) perform better than traditional time-series (SARIMAX), deep learning (LSTM), and additively combined hybrid forecasting models (Hybrid 1). To demonstrate this, the study compares all the models' forecasting performance based on three metrics – R^2, RMSE, and MAE – as well as CSFEs. It initially compares individual models before comparing them with hybrid models. In comparing individual SARIMAX or LSTM models with hybrid models, it is demonstrated that hybrid models achieve higher R^2 values, lower MAE, and RMSE compared to SARIMAX or LSTM models, indicating superior predictive performance of the hybrid models. In the comparison of CSFEs as well, hybrid models exhibit the lowest CSFEs, indicating an improvement in accuracy. This aligns with the literature, where hybrid forecasting models exhibit superior accuracy compared to individual traditional time-series or deep learning forecasting models [4, 15, 16]. Among the hybrid models, Hybrid 2 models demonstrate superior performance compared to Hybrid 1 models by achieving higher R^2 values, lower MAE, RMSE, and CSFEs. This indicates that the ANN-based optimal combination significantly enhances hybrid models' predictive ability.

The second substantial finding is that SARIMAX with a trend-capturing function (SARIMAX 2) can significantly improve forecasting accuracy. This suggests that in devising models that handle complex near-term time series predictions, such as one-year-ahead Chinese NG consumption, the trend-capturing function plays a vital role in adequately reflecting the structural characteristics of temporal data. Among SARIMAX models, SARIMAX 2 (S2) demonstrates superior performance compared to SARIMAX 1 (S1), SARIMAX without a trend-capturing function; S2 not only obtains a higher R^2 value of 21.25% and a lower RMSE of 7.92 but also exhibits lower CSFEs. In comparing the hybrid models, the hybrid models, including S2 (Hybrid 1-S2 and Hybrid 2-S2), demonstrate enhanced forecasting accuracy. Overall, the ANN-optimised hybrid model with S2 (Hybrid 2-S2) emerges as the highest-performing model, achieving the highest R^2 value of 47.94%, the lowest RMSE of 5.23, and the lowest MAE of 1.70. This is robustly anchored by the results in comparing CSFEs, where the final CSFE of the hybrid 2-S2 also shows the lowest value.

The third major finding is that the study not only demonstrates the key drivers as valid inputs but also identifies NG-related infrastructure factors as the most significant contributors to the near-term forecasting. For the validity analysis of the inputs, PCA-transformed data, which capture essential features and remove unnecessary noise, are used in the hybrid models, and their predictive performances are compared. The hybrid models using the PCA-transformed data achieve higher R^2 values and lower MAE and RMSE than those using the key drivers (raw data). This implies that the key drivers themselves contain important features for the forecasting. Furthermore, to analyse the factors significantly affecting the forecasting, the study investigates changes in the output (Chinese NG consumption prediction) in response to increases in a specific input (a key driver). This analysis identifies NG-related infrastructure factors as the most influential in the forecasting.

The hybrid model's improved accuracy can aid governmental agencies in evidence-based short-term NG policy decisions or support energy companies in proper near-term NG investment or operational decisions. Besides, its superior prediction accuracy indicates potential application not only in other energy fields, such as oil consumption, solar, or wind energy production, where near-term forecasting is crucial, but also in other prediction fields, such as tourism, exports, or imports. Several implication could be drawn from these foundings.

First, hybrid forecasting approaches with advanced pattern-capturing techniques should be promoted to enhance safety, operations, or investment decisions regarding NG demand. This study validates the use of the ANN-optimised hybrid model with the trend-capturing function for complex near-term forecasting tasks. The ANN-optimised hybrid model, which outperforms other singular or hybrid models, can further enhance forecasting accuracy, especially for volatile resources like NG. Policymakers could prioritise the use of such an advanced and above-peer hybrid forecasting model to enhance the accuracy of energy demand forecasts, which could support more effective energy planning and allocation of resources. For NG companies, especially in production and distribution, this hybrid forecasting model provides a reliable tool for short-term demand forecasting. Integrating the model with the telemetry control system of the gas transportation system could improve supply chain and resource allocation, helping to enhance safety management and operational efficiency. For the financial institutions involved in energy-related investments, the model allow their stakeholders to more precisely forecast NG demand that makes their better-informed decisions, optimising their risk-return trade-offs. Further research could explore integrating cutting-edge deep learning algorithms or advanced trend-capturing techniques to advance the model.

*Second, NG-related infrastructure should be prioritised***.** As per the finding that such infrastructure is the most crucial predictor of NG consumption, policymakers should prioritise developing NG-related infrastructure, such as pipelines, storage facilities, and import terminals, in response to the anticipated increase in demand. Additionally, the emphasis on NG infrastructure as a significant predictor of consumption highlights the need in targeted investments in the infrastructure by energy companies and investors. Such investments can strengthen supply resilience and support demand growth, particularly in high-consumption markets like China.

Third, accurate near-term forecasting models should be feasible and supported for China's journey toward carbon neutrality. Precise near-term predictions of NG consumption could be utilised to set and monitor near-term carbon emission targets based on NG usage predictions. This, in turn, effectively aids both the government and businesses in implementing near-term carbon-neutral action plans as they work toward achieving carbon neutrality.

Lastly, the high predictive accuracy of this hybrid 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 hybrid 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 energy sectors.

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 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 [22]. 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 in addition to 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 [20] can enhance the realism of these models [66, 67]. We leave this for future research.

Acknowledgements

This research is supported by an Australian Government Research Training Program Stipend (UTS Research Excellence) and Fee Offset Scholarship. The views expressed in this paper are solely those of the authors and do not necessarily reflect the views of the supporting agencies and authors' affiliations.

Appendix

A. Terms and abbreviations

TERM	DESCRIPTION
Accum.	Accumulative
ACRI	Australia-China Relations Institute at UTS
ADF	Augmented Dickey-Fuller
AIC	Akaike Information Criterion
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
BCM	Billion cubic metres
Capa	Capacity
ELU	Exponential Linear Unit
ETL	Extract, Transform, Load
IEA	International Energy Agency
KEEI	Korea Energy Economics Institute
KPSS	Kwiatkowski-Phillips-Schmidt-Shin
LNG	Liquefied Natural Gas
LSTM	Long Short-Term Memory
MAE	Mean Absolute Error
MMBTU	Metric Million British Thermal Unit
MMT	Million Metric Tons
MSE	Mean Square Error
MT	Metric Tons
N/A	Not Available
PCA	Principal Component Analysis
PNG	Piped Natural Gas
RMSE	Root Mean Square 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
UTS	University of Technology Sydney

B. Table 2 ADF test result

Table 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.

C. Table 3 Ljung-Box and Jarque-Bera test results

Table 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.

D. 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 [10, 12]. 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 12.

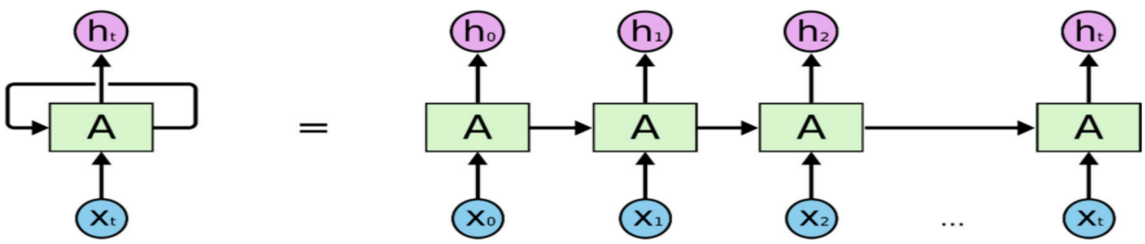


Figure 12 The structure of RNN

E. 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 [10, 16]. 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 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.

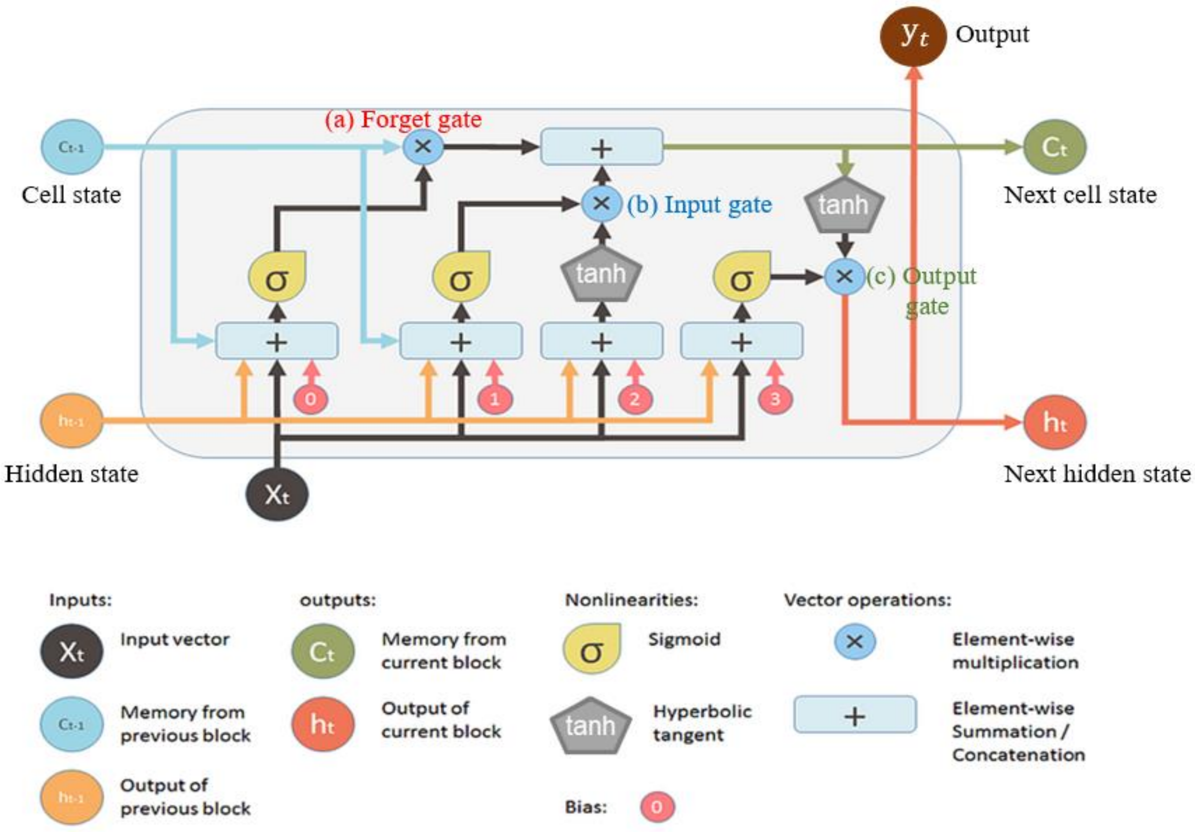
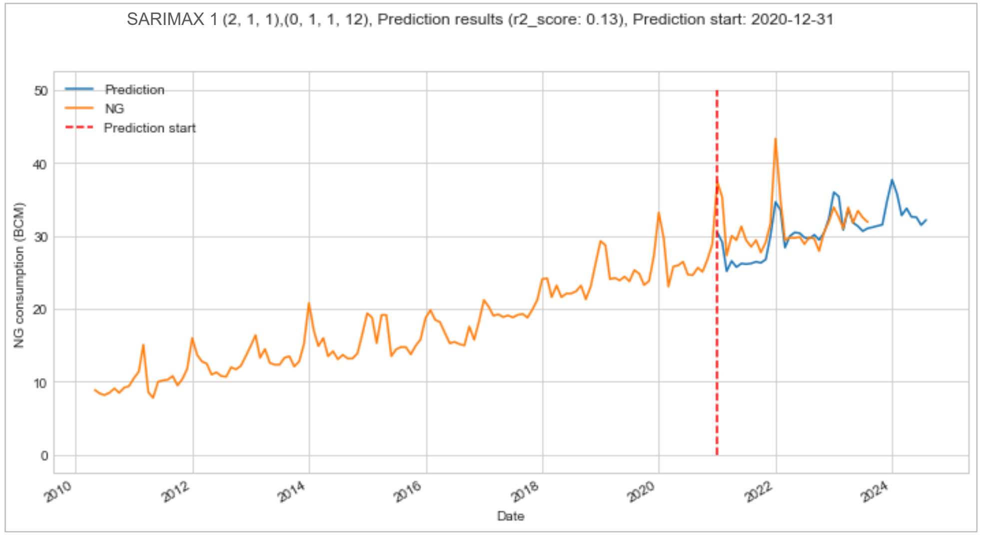


Figure 13 The structure of LSTM. Derived from Le et al [10]

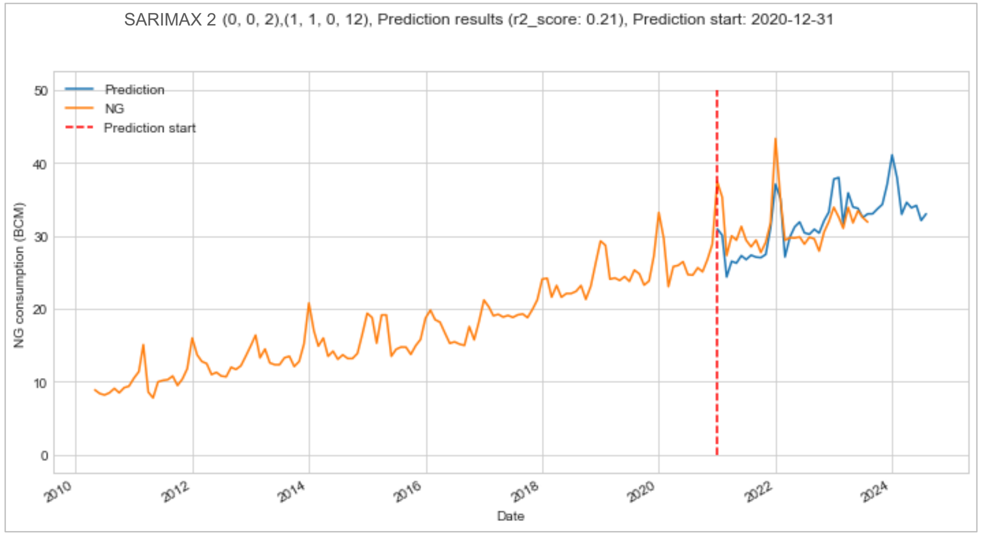
F. Figure 6 SARIMAX vs. LSTM

Actual values for Chinese NG consumption: Orange line / Predicted values for Chinese NG consumption: Blue line

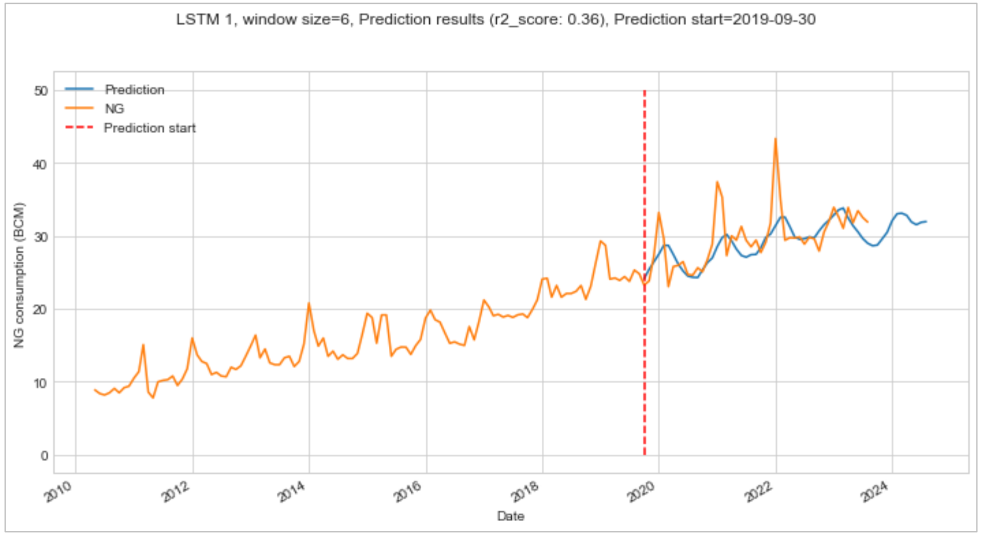
SARIMAX 1 (2,1,1),(0,1,1,12) - Prediction start: Dec-31-2020



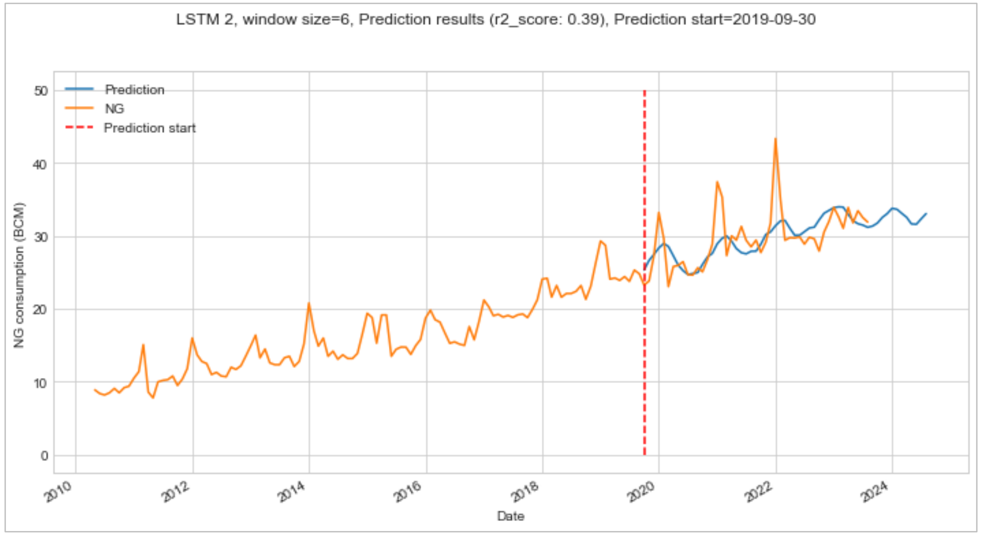
SARIMAX 2 (0,0,2),(1,1,0,12) - Prediction start: Dec-31-2020



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



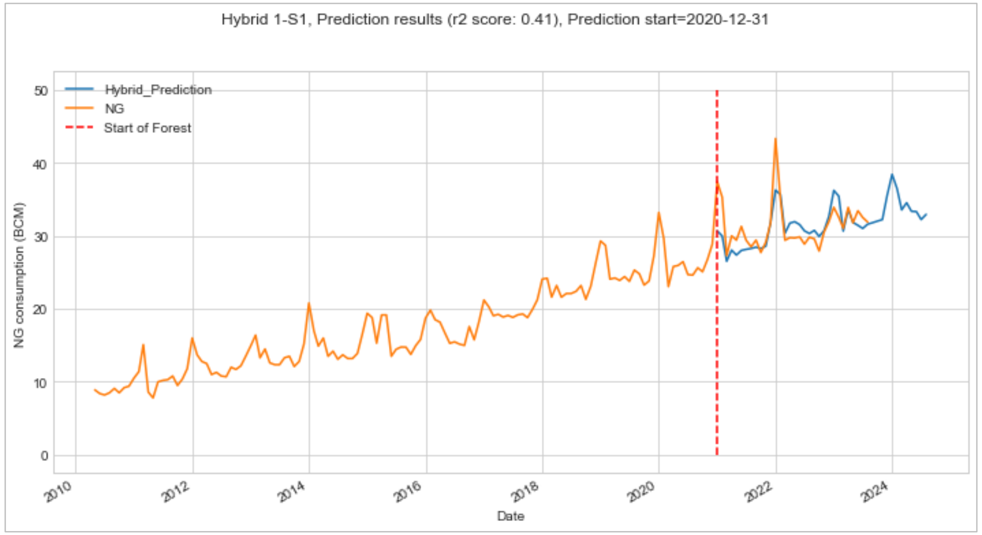
LSTM 2 (Window size=6) - Prediction start: Sep-30-2019



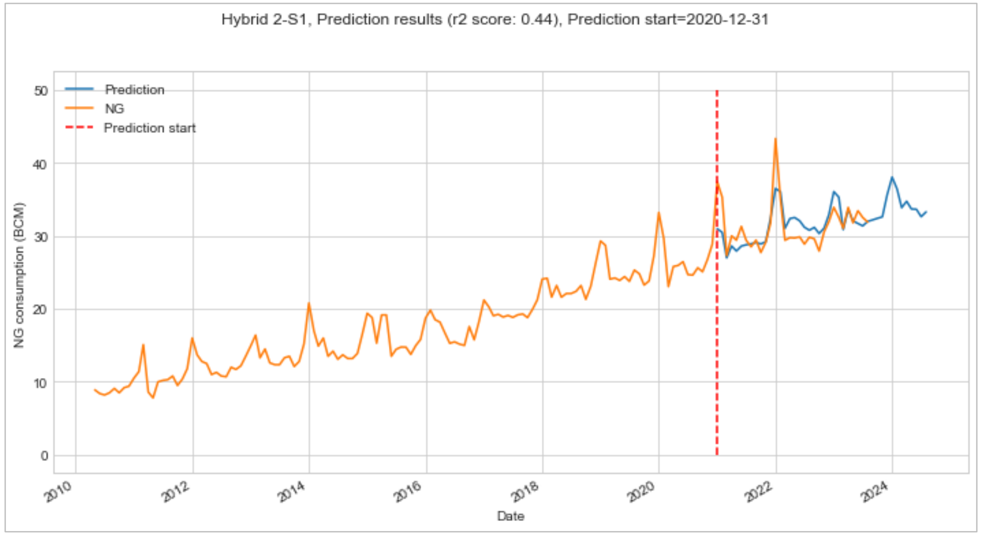
G. Figure 7 Hybrid 1 vs. Hybrid 2

Actual values for Chinese NG consumption: Orange line / Predicted values for Chinese NG consumption: Blue line

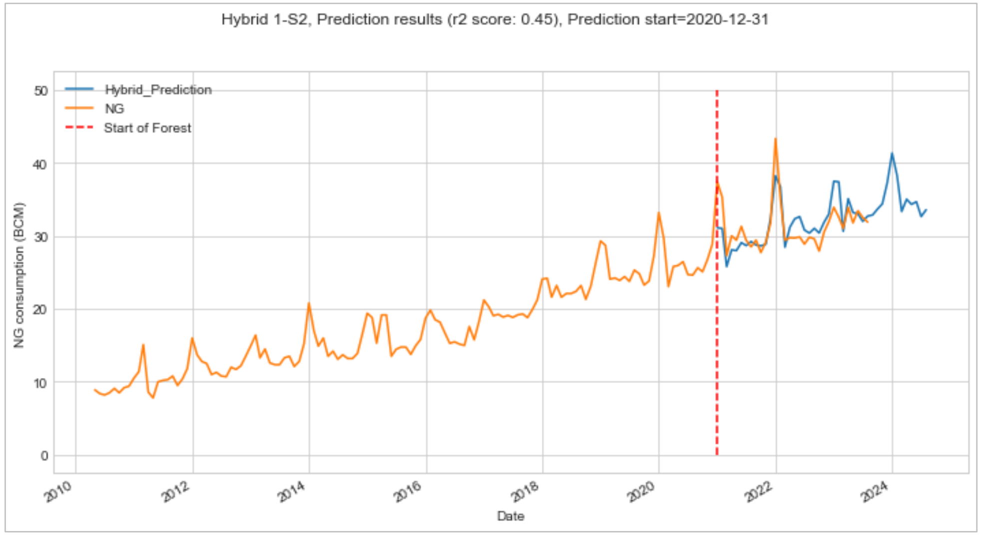
Hybrid 1-S1 (Prediction start: Dec-31-2020)



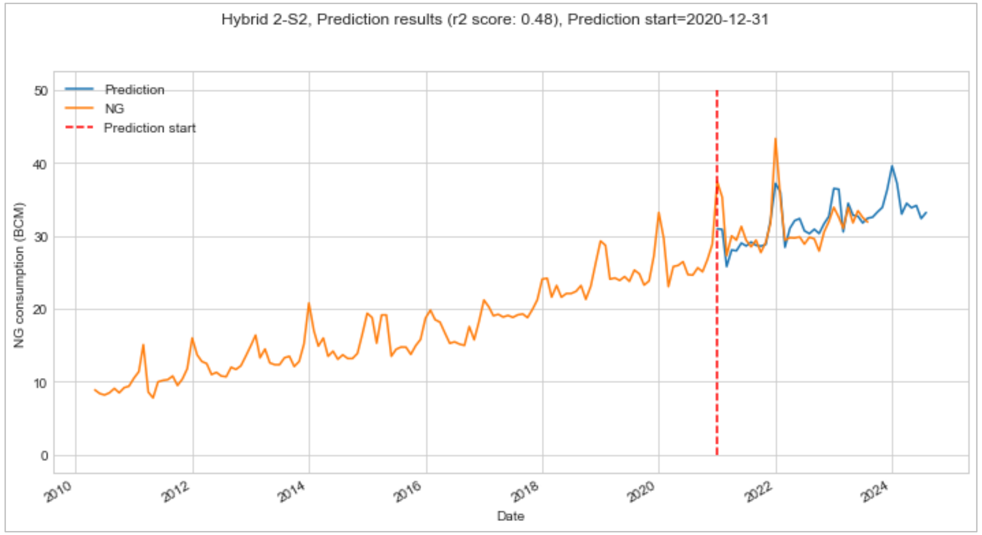
Hybrid 2-S1 (Prediction start: Dec-31-2020)



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



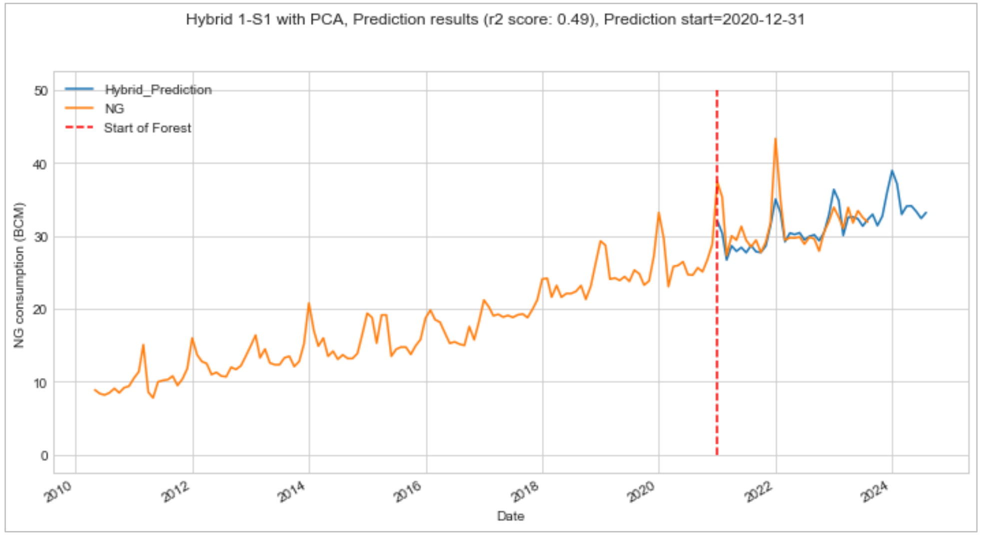
Hybrid 2-S2 (Prediction start: Dec-31-2020)



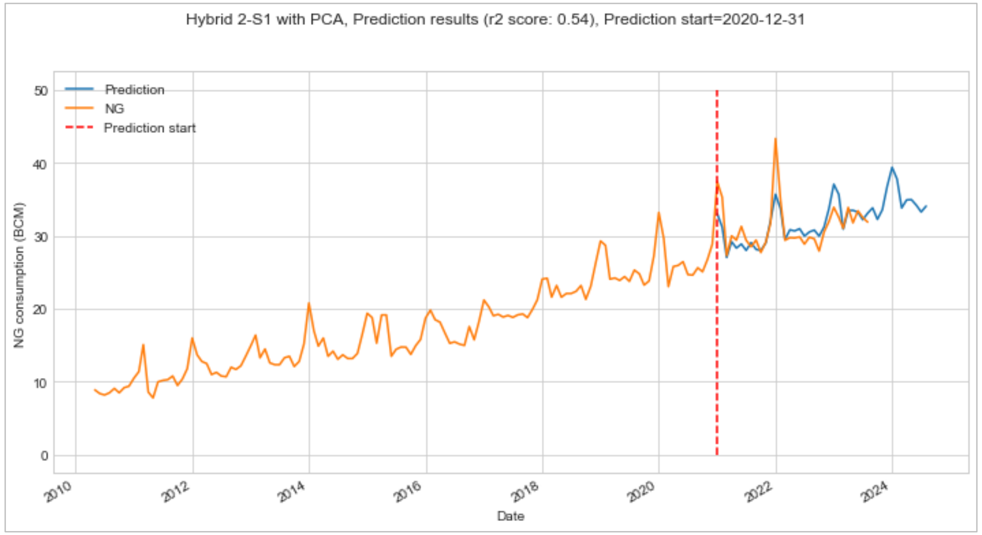
H. Figure 10 Hybrid models with PCA: Hybrid 1P-S1, Hybrid 1P-S2, Hybrid 2P-S1, and Hybrid 2P-S2

Actual values for Chinese NG consumption: Orange line / Predicted values for Chinese NG consumption: Blue line

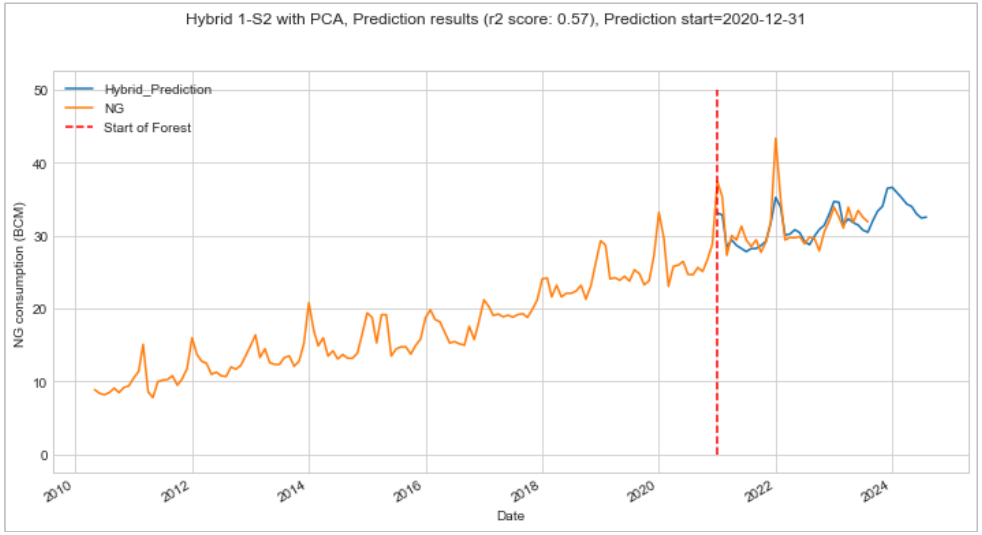
Hybrid 1P-S1 (Prediction start: Dec-31-2020)



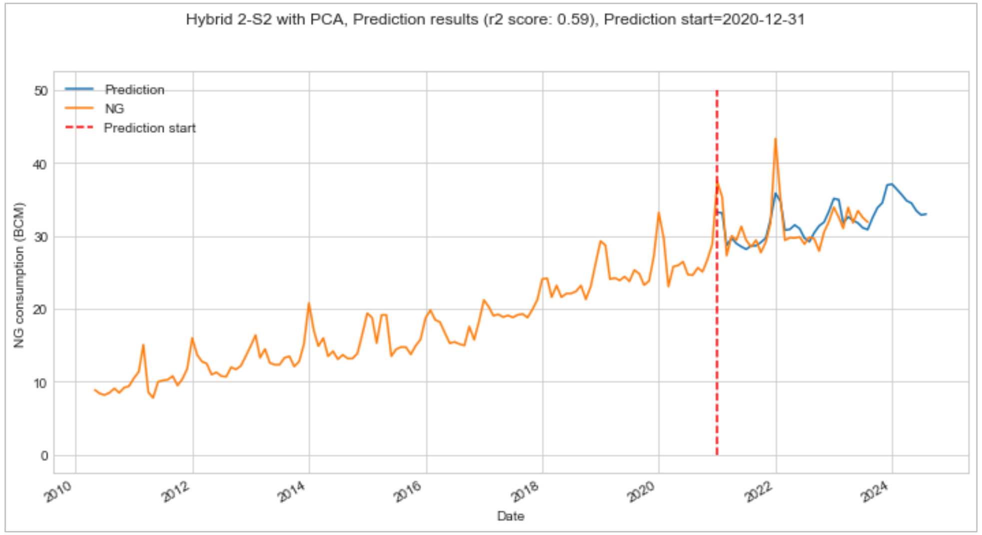
Hybrid 2P-S1 (Prediction start: Dec-31-2020)



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



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



Reference

[1] 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.

[2] *Shell LNG Outlook*. 2024, Shell.

[3] 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.

[4] 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.

[5] 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.

[6] Elamin, N. and M. Fukushige, *Modeling and forecasting hourly electricity demand by SARIMAX with interactions.* Energy, 2018. **165**: p. 257-268.

[7] Manigandan, P., et al., *Forecasting Natural Gas Production and Consumption in United States-Evidence from SARIMA and SARIMAX Models, Energies, 14 (19)*. 2021.

[8] 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.

[9] Muzaffar, S. and A. Afshari, *Short-Term Load Forecasts Using LSTM Networks.* Energy Procedia, 2019. **158**: p. 2922-2927.

[10] Le, X.-H., et al., *Application of long short-term memory (LSTM) neural network for flood forecasting.* Water, 2019. **11**(7): p. 1387.

[11] Chen, X., L. Wei, and J. Xu, *House price prediction using LSTM.* arXiv preprint arXiv:1709.08432, 2017.

[12] 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.

[13] SEO, S.H., *Forecasting Korean LNG import price using ARIMAX, VECM, LSTM and hybrid models.* 2021.

[14] 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.

[15] 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.

[16] Dave, E., et al., *Forecasting Indonesia Exports using a Hybrid Model ARIMA-LSTM.* Procedia Computer Science, 2021. **179**: p. 480-487.

[17] Szoplik, J., *Forecasting of natural gas consumption with artificial neural networks.* Energy, 2015. **85**: p. 208-220.

[18] 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.

[19] Lee, J. and Y. Cho, *National-scale electricity peak load forecasting: Traditional, machine learning, or hybrid model?* Energy, 2022. **239**: p. 122366.

[20] Sung, J., et al., *Short- and Long-term effects of key drivers in China's Natural Gas market (Working Paper).* 2023.

[21] 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.

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

[23] 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.

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

[25] 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.

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

[28] Shadkam, A., *Using SARIMAX to forecast electricity demand and consumption in university buildings*. 2020, University of British Columbia.

[29] 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.

[30] 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.

[31] Paltsev, S. and D. Zhang, *Natural gas pricing reform in China: Getting closer to a market system?* Energy Policy, 2015. **86**: p. 43-56.

[32] 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.

[33] Albon, C., *Machine Learning with Python Cookbook: Practical Solutions from Preprocessing to Deep Learning*. 2018, Sebastopol: O'Reilly Media, Incorporated.

[34] Kishan, S.A., et al., *Handwritten character recognition using cnn.* Published in, 2018. **5**(3).

[35] Nugroho, R.A., et al. *Covid-19 classification using K-nearest neighbor with various distances*. in *AIP Conference Proceedings*. 2024. AIP Publishing.

[36] 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.

[37] 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.

[38] 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.

[39] Hochreiter, S. and J. Schmidhuber, *Long short-term memory.* Neural computation, 1997. **9**(8): p. 1735-1780.

[40] 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.

[41] Kaushik, S., et al., *AI in healthcare: time-series forecasting using statistical, neural, and ensemble architectures.* Frontiers in big data, 2020. **3**: p. 4.

[42] 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.

[43] Chollet, F.o., *Deep learning with Python*. Second edition. ed. 2021, Shelter Island, New York: Manning Publications Co. LLC.

[44] 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.

[45] 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.

[46] 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.

[47] 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.

[48] 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.

[49] 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.

[50] 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.
[51] 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.
[52] Elhalwagy, A. and T. Kalganova, *Hybridization of Capsule and LSTM Networks for unsupervised anomaly detection on multivariate data.* arXiv preprint arXiv:2202.05538, 2022.
[53] 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.
[54] 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.
[55] Zhang, Y., et al., *Forecasting the prices of crude oil: An iterated combination approach.* Energy Economics, 2018. **70**: p. 472-483.
[56] Reschenhofer, E., *Forecasting volatility: double averaging and weighted medians.* International Journal of Computational Economics and Econometrics, 2010. **1**(3-4): p. 317-326.
[57] Pettenuzzo, D. and F. Ravazzolo, *Optimal portfolio choice under decision‐based model combinations.* Journal of Applied Econometrics, 2016. **31**(7): p. 1312-1332.
[58] Rosa, A., *Innovative handling of telemetry enhances gas system automation.* Pipeline & gas journal, 2005. **232**(11): p. 54-57.
[59] 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.
[60] 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.
[61] 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.
[62] Raiyn, J. and T. Toledo, *Real-time short-term forecasting based on information management.* Journal of Transportation Technologies, 2014. **4**(1): p. 11-21.
[63] Stark, T. and D. Croushore, *Forecasting with a real-time data set for macroeconomists.* Journal of Macroeconomics, 2002. **24**(4): p. 507-531.
[64] 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.
[65] 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.
[66] Schoemaker, P.J., *How historical analysis can enrich scenario planning.* Futures & Foresight Science, 2020. **2**(3-4): p. e35.
[67] 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.