

---

**| RESEARCH ARTICLE**

## **Optimization of Real-Time Carbon Emission Monitoring on the Power Demand Side Based on System Dynamics**

**Yuqing Ye**<sup>†</sup>, Business School, Xi'an International Studies University, Xi'an, Shaanxi 710128, China

**Rongrong Li**<sup>†</sup>, Business School, Xi'an International Studies University, Xi'an, Shaanxi 710128, China

<sup>†</sup>These authors contributed equally to this work and share first authorship

---

**| ABSTRACT**

Against the backdrop of China's "dual carbon" strategic goals, the precise monitoring of carbon emissions on the grid's demand side is crucial for advancing the transformation of the energy structure and achieving low-carbon development. Current mainstream hourly-level monitoring methods have insufficient temporal resolution, coarse information granularity, and weak localization data support. These issues compromise the effectiveness of carbon reduction efforts and may even lead to secondary risks and economic losses. As a major province with high energy-consuming industries, Henan Province experiences frequent short-term high-energy consumption events in its heavy industry sector, where carbon emission characteristics are complex, making traditional monitoring methods inadequate. To address these challenges, this study focuses on the demand side of the power grid in Henan's heavy industry. It specifically analyzes the carbon emission characteristics of high-energy-consuming equipment and industrial electricity usage patterns during short-term high-energy events. Innovatively, a system dynamics model integrating STL, LightGBM, and a stream computing subsystem is constructed. This model not only achieves a technical upgrade from hourly to minute-level carbon emission monitoring but also fully considers the internal structures and coupling relationships among subsystems. Empirical research demonstrates that the model possesses the capability for minute-level data collection and dynamic updates of carbon emission factors, significantly improving the temporal resolution and response efficiency of the monitoring system. It enables high-frequency dynamic tracking of instantaneous load during industrial electricity consumption. These results provide robust data support for grid dispatch optimization and carbon market trading. The model enhances the accuracy of carbon emission predictions and offers reliable foundations and new optimization directions for exploring technological pathways toward energy conservation and emission reduction. Its application holds significant theoretical value and practical importance for promoting the modernization of regional energy governance, optimizing industrial structure, and formulating scientifically sound policies.

**| KEYWORDS**

Demand Side; Carbon emission monitoring; System dynamics; Minute-level monitoring

**| ARTICLE INFORMATION**

**ACCEPTED:** 23 March 2026

**PUBLISHED:** 07 April 2026

**DOI:** 10.32996/jbms.2026.8.6.2

---

### **1. Introduction**

At present, the global climate change situation is severe, and reducing carbon emissions has become an urgent task for the international community"[1].As the country with the largest global energy consumption and carbon emissions at present, China has clearly put forward the strategic goal of "carbon peaking and carbon neutrality""[2][3][4],,which fully shows that China adheres to a firm position in responding to global climate change and promoting green and low-carbon development. At present, the power industry is a key area of carbon emissions, and strengthening carbon emission monitoring is of key significance to achieving the "dual carbon" goal [5]. Carbon emission monitoring on the electricity side of the power grid is the core of the energy structure transformation, which can provide a scientific basis for the formulation of emission reduction policies and the optimization of energy structure [6]. Therefore, as an important industrial and energy consumption province in China, the monitoring of carbon emissions on the electricity side of the power grid in Henan Province is of great significance [7].

**Copyright:** © 2026 the Author(s). This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) 4.0 license (<https://creativecommons.org/licenses/by/4.0/>). Published by Al-Kindi Centre for Research and Development, London, United Kingdom.

However, there are still many difficulties and bottlenecks to achieve accurate carbon emission monitoring. With the rapid development of big data and artificial intelligence, big data and artificial intelligence have provided many new technical means for the monitoring of carbon emissions on the power grid. However, the traditional carbon emission accounting method still relies on statistical data and periodic reports. There are problems of insufficient timeliness and accuracy, which makes it difficult to meet refined real-time data monitoring and efficient dynamic management [8][9]. Accurate monitoring of carbon emissions on the power grid is conducive to promoting the sustainable development of the power industry and the transformation of the whole society to a low-carbon economy[10].

In the field of heavy industry in Henan, the frequency of short-term high-energy-consuming events is relatively high, and its carbon emission characteristics are highly nonlinear and sudden. At present, the existing mainstream technology still generally relies on hourly monitoring, resulting in a serious delay in data update. Hourly monitoring is difficult to accurately capture the frequent carbon emission characteristics of short-term high-energy-consuming events in heavy industry, and more refined monitoring means are urgently needed. The implementation of minute-level monitoring can account for carbon emissions more accurately, provide stronger data support for the formulation and implementation of carbon emission reduction measures, find more energy-saving opportunities, optimize electricity strategies, and quickly reduce carbon emissions on the electricity consumption side.

In view of the problem of monitoring electricity-side carbon emissions in short-term and high-energy-consuming scenarios in Henan heavy industry, we can achieve a breakthrough from traditional hourly monitoring to more accurate minute-level monitoring by constructing a system dynamics model. The system is mainly divided into three subsystems: STL, LightGBM and streaming computing, and realizes deep coupling between subsystems through the method of system dynamics[11] [12] [13] [14] [15] [16] [17]. Through data processing and analysis technology, it provides high-time-resolution carbon emission monitoring data, thus providing strong support for power grid scheduling, energy management and carbon trading.

First of all, we will conduct an in-depth analysis of the dynamic characteristics of carbon emissions of the power grid under the short-term high-energy consumption scenario in the heavy industry of Henan Province. Focusing on the heavy industry field in Henan Province, we use the system dynamics model method [11] [12] to establish a causal diagram and mathematical equations of the system, carry out simulation of the model, and analyze the dynamic monitoring of minute-level carbon emissions for the model dynamic simulation to provide basic data support[11] [12]. Then, a composite model framework based on the system dynamics framework is built, and the sub-model is built on the framework of the system dynamics model. Among them, the LightGBM subsystem builds a nonlinear response relationship of the key driving factors of carbon emissions through gradient improvement decision tree algorithm<sup>[14][15]</sup>. The adaptive control theory was first founded by Norbert Wiener and other scholars. The core of this theory is to enable the system to autonomously optimize the control strategy with the help of real-time monitoring and feedback regulation mechanism of the system operation status, so as to dynamically adapt to changes in the external environment. The STL subsystem uses time series decomposition technology to extract the time cycle characteristics and trend terms of minute-level carbon emissions[13] [14] [15]. Through dynamic optimization algorithms, complex optimization problems are decomposed into multiple sub-problems, and gradually solved through recursive methods to further analyze and process the decomposed features to improve the accuracy and efficiency of feature extraction. The streaming computing subsystem, built on the Flink platform, enables millisecond-level data stream computing and real-time updates to model parameters. [17] [18] [19]The three subsystems are organically coupled through the causal loop of system dynamics and the stock-flow structure [20], forming a whole-process monitoring chain of "data preprocessing - feature dynamic extraction - model adaptive training - real-time predictive feedback".

In summary, this study puts forward a new solution for the problem of monitoring carbon emissions on the power side of the power grid in typical short-term high-energy-consuming scenarios in Henan heavy industry. By constructing a system dynamics model and combining STL, LightGBM and streaming computing technology, the scheme realizes minute-level carbon emission monitoring, improving the monitoring accuracy and time resolution. The research results help to improve the operational efficiency and carbon emission management of the power industry, and can also provide certain references for policy formulation.

## 2.Review

Under the global climate crisis, carbon emission control has become a key international issue, and China has also formulated a "dual carbon" strategy to solve the global climate problem[1] [2] [5]. Under the severe situation of global warming, the power industry, as the "main force" of carbon emissions, is related to the realization of the global carbon neutrality goal. The traditional carbon emission reduction monitoring and accounting method mostly relies on manual statistics and periodic evaluation, which not only has poor timeliness, but also has obvious shortcomings in data accuracy, making it difficult to meet the current urgent needs of dynamic supervision of carbon emissions.

The in-depth application of advanced technologies such as big data and artificial intelligence is promoting the development of carbon emission reduction monitoring in the power industry in a more accurate and efficient direction. Based on the heavy industry field of Henan Province, this study deeply integrates cutting-edge technologies for short-term high-energy-consuming events to build a high-resolution carbon emission monitoring system. By capturing carbon emission data in real time and accurately quantifying the carbon footprint of each production link, it can not only provide a scientific basis for enterprise energy management, but also help the standardization and transparency of carbon market transactions.

In management practice, enterprises and power grid operators are actively exploring effective methods for carbon emission monitoring and management [7]. Wang Muhan studied the carbon emission monitoring system and accounting method of thermal power plants, put forward a carbon emission monitoring scheme based on actual production data, and provided practical guidance for carbon emission management in the power industry.

Traditional carbon emission monitoring methods (such as emission factor method and material balancing algorithm) rely heavily on macro-level statistics and regularly submitted periodic reports in the implementation process. This results in a relatively static perspective, making it difficult to reflect dynamic changes and failing to meet the requirements of precise control. Fang Xiaoling et al pointed out that traditional methods are difficult to accurately reflect actual emissions when dealing with complex industrial processes, and their timeliness is poor, failing to meet the needs of real-time monitoring.

The iterative upgrades of the internet and information technology have driven innovation in data processing models. Offline batch processing technologies like Hadoop MapReduce, limited by their non-real-time computing mechanisms and high latency, are unable to meet the requirements of real-time data processing. To overcome this technical bottleneck, new streaming systems (such as Storm, Spark Streaming, and Flink) have gradually matured in the early 21st century [19] [20] [21]. With their millisecond-level response capabilities and high throughput characteristics, they have become key infrastructures in the field of real-time data analysis and play an important role in carbon monitoring scenarios. This type of real-time processing architecture effectively enhances the immediacy of carbon emission monitoring through continuous data stream computing mechanisms. Typical applications include the industrial data processing system built by Zhao Runfa's research group [19] using Flink technology, which can continuously track carbon emissions in the production process; and the distributed real-time database system developed by Guo Mengyu's team [21] [19] which optimizes the data partitioning strategy with the help of streaming computing, significantly improving the stability of monitoring results.

In terms of practical application, the integration of real-time processing technology with industry needs is becoming increasingly close. Zheng Zhenqiang [11] innovatively combined the Internet of Things sensor network with carbon metering algorithm to build a power carbon emission monitoring system with dynamic accounting capability. Its second-level data update capability strongly supports smart grid scheduling decision-making. Wu Chang's research group used the Flink engine to build an industrial data analysis platform, which shortened the carbon emission analysis response speed to the minute level. Liu Hui's team [20] [18] further improved the Flink task scheduling algorithm and compressed the data processing latency by 40%, which fully demonstrated the continuous optimization space of the streaming architecture. It is particularly important to point out that in the field of heavy industry in Henan, short-term high-energy-consuming events occur frequently, and its carbon emission characteristics show high nonlinearity and suddenness. This puts forward higher requirements for the dynamic resource allocation capability of the monitoring system, requiring the streaming processing system to have a fast response capability to cope with the changes in computing load caused by the operation of high-energy-consuming equipment.

In response to the accuracy limitations of traditional monitoring methods, the academic community has begun to explore the integration path of statistical modeling and intelligent algorithms. Yu Qingyun's project group [12] evaluated the interactive benefits by using the system dynamics method, and finally carried out a long-lasting systematic evaluation of the interactive benefits of the peak and valley time-sharing electricity price project through simulation examples. System dynamics is a discipline that explores the interaction between system behavior and internal structure. Its research focuses on the dynamic change characteristics and feedback regulation mechanism of the system. The theory was pioneered by Jay W. Forrester, a professor at Massachusetts Institute of Technology. He proposed that the specific behavior patterns presented by the system come from its inherent structural characteristics, which are manifested as key elements such as feedback loop mechanisms, time lag effects and nonlinear interaction relationships. When constructing a carbon emission monitoring model, the system dynamics causal feedback modeling method can be used to deeply analyze the dynamic characteristics and complex feedback relationships of the carbon emission system. The STL decomposition model developed by Liu Li's project group [15] successfully reduced the prediction error of electricity sales to less than 3% by separating the long-term trend, periodic fluctuation and random disturbance elements of the time series. The STL-LSTM hybrid model proposed by Shi Yutao [14] shows excellent feature recognition ability in water conservancy engineering and construction carbon emission monitoring scenarios. In the face of nonlinear data relationships, the LSTM and LightGBM integration model designed by Chen Weinan's team [16] improved the accuracy of load prediction by 15% through the collaborative training of deep learning networks and gradient lifting trees. The

empirical research of Ma Xiaojun's research group [17]<sup>[15]</sup> shows that the accuracy rate of the LightGBM algorithm in the TB-level carbon emission data classification task is 92%, which is significantly improved compared with the traditional regression model. In the process of model training, an adaptive control strategy is adopted to adjust the model parameters in real time according to the prediction error of the model to improve the prediction accuracy and stability of the model. This mechanism gives the system the efficiency of the independent optimization strategy based on real-time feedback information, so that it can implement intelligent response to environmental changes. It is understood that the core idea of dynamic optimization theory is to decompose complex optimization problems into multiple sub-problems and gradually solve them through recursive methods, so as to realize the dynamic optimization of the system. In carbon emission monitoring, the use of dynamic optimization algorithms can analyze and process the monitoring data in real time, optimize the monitoring strategy and resource allocation, and improve the efficiency and accuracy of the monitoring system.

Although the current carbon emission monitoring system has achieved stage results in terms of technical framework and policy coordination, its core performance still has key constraints. The most prominent problem is that the time resolution of the monitoring system is difficult to meet the needs of dynamic supervision in the field of heavy industry. Traditional accounting methods generally rely on monthly or annual data aggregation, resulting in the inability of hourly carbon emission detection to accurately capture the rapid changes in carbon emissions, resulting in certain distortions of carbon emission data observed in a relatively short time interval, and the actual emission situation cannot be accurately portrayed.

In view of the frequent occurrence of short-term high-energy-consuming events in the field of heavy industry in Henan, the existing monitoring system cannot accurately quantify the minute-level shock phenomenon of carbon emission intensity caused by the operation of high-energy-consuming equipment due to the lack of millisecond power fluctuation tracking mechanism. In order to realize the minute-level monitoring of carbon emissions on the power side of the power grid, it is necessary to focus on implementing three key links: First, build a system dynamics model that integrates STL, LightGBM and stream computing. In this system, the STL subsystem undertakes the task of accurately analyzing the characteristics of time series data, and can effectively separate and extract trend items, seasonal components and random residual components; the LightGBM subsystem makes efficient predictions based on these characteristics, significantly improving the prediction accuracy; the streaming computing module is based on the Flink distributed computing framework, which can complete the instant processing and continuous dynamic refresh of data. Second, through the deep coupling of these three subsystems, a breakthrough in carbon emission monitoring technology from hour to minute is achieved. Third, optimize the model parameters and algorithms to ensure the stability and accuracy of the system in high-frequency data processing, so as to provide a reference basis for power grid scheduling optimization, carbon market trading and energy saving and emission reduction decision-making.

### **3. Construction of minute-level monitoring system**

#### **3.1 Current situation and shortcomings of hourly monitoring**

Currently, hourly monitoring is one of the common methods for monitoring carbon emissions from the electricity consumption side. Specifically, it involves recording electricity consumption and carbon emissions hourly using smart meters or energy management systems, providing a relatively detailed reflection of energy consumption characteristics over different time periods. On the enterprise side, both businesses and grid operators have established relatively mature data management and analysis systems capable of statistically analyzing hourly data and generating energy consumption reports to provide a reference for carbon emission management. However, in Henan Province's heavy industry sector, due to the volatility and intermittent nature of this sector, hourly data often fails to accurately capture its rapidly changing characteristics, leading to a decrease in the accuracy of monitoring results.

Hourly monitoring cannot capture short-term high-energy-consumption events in some equipment, such as instantaneous starts and sudden malfunctions. Specifically, elevators may consume a large amount of electricity momentarily upon startup, as may air conditioning systems in commercial buildings. However, these short-term high-energy-consumption events may be averaged out when statistically analyzed hourly, making accurate identification and analysis impossible. For equipment that frequently starts and stops (such as air conditioners and escalators), hourly monitoring cannot accurately reflect the actual operating time and energy consumption. This may lead to misjudgments of equipment energy consumption and affect the formulation of energy-saving measures. Hourly monitoring data typically needs to be aggregated and analyzed at the end of each hour, failing to reflect current electricity consumption and carbon emissions in real time. In scenarios requiring real-time monitoring and adjustment of electricity consumption strategies, the insufficient real-time capability of hourly monitoring may lead to missed opportunities for optimal energy saving or cost optimization. In Henan Province's heavy industry sector, rapid fluctuations in power generation from heavy industry equipment can cause significant changes in carbon emissions within a short period, which hourly monitoring cannot capture in a timely manner, thus affecting the accurate assessment and timely adjustment of the power grid's operating status.

### **3.2 Construction of minute-level monitoring system**

The established minute-level monitoring system fully covers key aspects such as data collection, transmission, storage, processing, analysis, and visualization. Through efficient coordination and continuous optimization, each module has formed a full link from the starting point of minute-level data collection to real-time monitoring and management applications, which provides reliable technical support for accurate measurement and monitoring of carbon emissions on the electricity side.

**Data collection link.** Advanced metering equipment such as smart meters play a key role. This kind of monitoring equipment uses minutes as the time particle size for accurate electricity consumption data collection, which not only ensures the immediacy and content integrity of the original data in the time dimension, but also relies on the characteristics of high-frequency data collection to effectively capture the instantaneous fluctuation characteristics of the electricity load, so as to provide complete data support with fine granularity for subsequent monitoring and analysis. In order to achieve minute-level monitoring, the data acquisition equipment needs to have a high-precision time synchronization function to ensure that the collected data can accurately reflect the electricity consumption of each minute. At the same time, in view of the special attributes of short-term high-energy-consuming events, data acquisition equipment needs to configure flexible parameter adjustment and function expansion mechanisms to meet the differentiated monitoring data collection requirements in various high-energy-consuming application scenarios.

**Data transmission link.** Adopt a high-speed and stable communication network to transmit the collected minute-level data to the data center in real time. The data transmission link focuses on strengthening security protection and transmission reliability, blocking the risk of data leakage and illegal tampering by deploying encryption technology, and ensuring that the transmitted data reaches the target node completely and accurately. In view of the typical short-term high-energy consumption scenario in the heavy industry field of Henan Province, the data transmission network needs to have greater bandwidth and stronger anti-interference ability to cope with the simultaneous transmission of a large amount of data and complex and changeable environmental conditions. In order to realize minute-level monitoring, the data transmission network needs to have the characteristics of low latency and high throughput to meet the needs of real-time data transmission. At the same time, it is also necessary to establish a redundancy mechanism for data transmission to ensure that data can still be transmitted reliably in the event of some network failures.

**Data storage:** A high-capacity, high-performance database platform is built to achieve efficient storage and scientific management of massive amounts of minute-level data. This database architecture is designed with flexible horizontal scalability to seamlessly handle continuous data growth and integrates a millisecond-level intelligent query engine to effectively support the dual needs of real-time monitoring and in-depth historical data analysis. Targeting typical short-term high-energy-consuming scenarios in Henan Province's heavy industry sector, the database system must be able to handle complex data types and data volume fluctuations in high-energy-consuming production processes, achieving unified data management and efficient utilization. A storage architecture integrating a data warehouse and a data lake is constructed. Through a layered storage mechanism for structured and unstructured data, it enables the scientific classification and correlation analysis of multi-source heterogeneous data, providing highly reliable, multi-dimensional decision-making data support for refined monitoring and energy efficiency optimization of energy consumption processes. To achieve minute-level monitoring, the database system needs rapid data writing and reading capabilities to meet the needs of real-time data storage and querying. Simultaneously, a data backup and recovery mechanism is established to ensure data security and reliability.

**Data processing stage.** Preprocessing techniques such as data cleaning, transformation, and aggregation are employed to systematically process the raw collected data: noise filtering, redundant data removal, and outlier detection eliminate data interference; format standardization is performed to adapt to analytical needs; and data aggregation and feature extraction are used to derive core indicator parameters, thus providing high-precision, low-noise, high-quality data resources for the model analysis stage. For typical short-term high-energy-consuming scenarios in Henan Province's heavy industry sector, the data processing process also needs to pay special attention to the complex data types and data volume fluctuations in high-energy-consuming production processes. For example, when processing electricity consumption data from high-energy-consuming industries such as steel, cement, and chemicals, the impact of production processes and equipment operating status on electricity load needs to be considered and incorporated into the data processing flow to improve data accuracy and reliability. To achieve minute-level monitoring, the data processing stage needs to have efficient data processing capabilities, capable of quickly completing data cleaning, transformation, and aggregation operations to meet the needs of real-time data processing. At the same time, it is also necessary to establish a data quality monitoring mechanism to detect and deal with data quality problems in time to ensure the quality and reliability of data.

**Data analysis and modeling links.** In the data analysis and modeling link of typical short-time high-energy-consuming scenarios in the heavy industry field of Henan Province, based on the STL-LightGBM-stream computing architecture, the pre-processed data is analyzed and modeled in depth. Based on the STL decomposition algorithm, the power load time series is

multi-dimensionally deconstructed, and the long-term trend component, cycle season component and random residual component in the sequence are accurately identified and extracted, and then the cyclical laws and core characteristics implicit in the load data are deeply revealed. Then use the LightGBM algorithm to build a prediction model, combined with the streaming computing architecture, to realize the rapid processing and predictive analysis of real-time data, so as to realize minute-level monitoring of the high energy consumption on the electricity side. In typical short-time high-energy consumption scenarios in the field of heavy industry in Henan Province, the model training process should fully consider the operating characteristics of high-energy-consuming production equipment and the complexity of the process. Because short-term high-energy-consuming events are often accompanied by sudden load changes and rapid adjustment of equipment operation status, we should focus on key factors such as equipment start-up and stop, production task scheduling, and raw material input in the selection of characteristic variables. For example, the operation status data of production equipment, production plan arrangement data, etc. can be introduced as the input features of the model, so that the model can better capture the occurrence and evolution trends of high-energy-consuming events, so as to achieve accurate prediction and real-time monitoring of high-energy consumption. At the same time, according to the characteristics of short-term high-energy-consuming scenarios, the model parameters are optimized and adjusted to improve the adaptability and prediction accuracy of the model when dealing with sudden high-energy-consuming events, and provide important support for energy management and energy conservation and emission reduction of enterprises.

**Results Presentation and Application:** Monitoring results will be presented to users and decision-makers in intuitive and easy-to-understand charts, reports, and other formats, enabling them to clearly understand the real-time operating status and energy consumption trends of high-energy-consuming equipment. Simultaneously, based on the monitoring results, decision support and management suggestions will be provided, such as optimizing production processes, developing energy-saving measures, and adjusting equipment operation strategies, providing a scientific basis for achieving efficient energy utilization and energy conservation and emission reduction in enterprises. In typical short-term high-energy-consuming scenarios in Henan Province's heavy industry sector, results presentation and application should also highlight key energy consumption nodes and energy-saving potential analysis in high-energy-consuming production processes. Through a visual interface, energy consumption data, operating efficiency, and corresponding energy-saving potential of high-energy-consuming equipment will be intuitively displayed, providing enterprise management and operators with clear directions for energy consumption management and energy-saving improvements. Furthermore, monitoring results can also provide data support for enterprise technological transformation, equipment upgrades, and the construction of energy management systems, promoting green and low-carbon development in the heavy industry sector.

**Technological Breakthroughs and Advantages.** In typical short-term high-energy-consumption scenarios in Henan Province's heavy industry, technological breakthroughs and advantages are reflected in several aspects. By extracting minute-level time-series features through STL decomposition and combining it with the Flink framework to achieve millisecond-level data processing, a full-process monitoring system covering data acquisition, dynamic prediction, and real-time early warning has been constructed. This system can effectively improve time resolution, achieving a monitoring capability of instantaneous peak industrial power consumption of 5 minutes/time, and an abnormal energy consumption event identification response time of <30 seconds, providing key technical support for enterprises' real-time energy management and scheduling. In typical short-term high-energy-consumption scenarios in Henan Province's heavy industry, this minute-level monitoring system can more promptly capture sudden changes in the operating status of high-energy-consuming equipment and drastic fluctuations in energy consumption, helping enterprise energy managers make quick decisions, optimize production processes, adjust equipment operating parameters, and reduce energy consumption. For example, when the monitoring system detects a sharp increase in energy consumption of a high-energy-consuming equipment due to the scheduling of production tasks, managers can quickly take measures, such as adjusting the production plan and optimizing the equipment operation mode, to ensure the efficient use of energy of the enterprise and reducing the cost increase that may be caused by energy waste. The difficulty of realizing minute-level monitoring mainly lies in how to avoid blockage in the process of detecting data flow. On the one hand, the complexity of the model increases and more data and characteristics need to be processed, which may lead to excessive consumption of computing resources and the extension of model training time. On the other hand, the batch processing of the computing architecture does not match the hardware equipment of the computing system, which may also become a bottleneck in the realization of minute-level monitoring. In order to solve these problems, it is necessary to adopt efficient algorithms and optimized computing architecture to improve the efficiency of data processing and model training. At the same time, it is also necessary to strengthen the performance of hardware equipment and improve the computing power and storage capacity to meet the needs of minute-level monitoring.

#### 4. Model construction

##### 4.1 Model overall architecture design

This research aims to build a system dynamics model to achieve a breakthrough in the monitoring of electricity-side carbon emissions from the hourly to the minute-level of the power grid. The model is composed of three subsystems: STL, LightGBM and stream computing, which realizes deep coupling between subsystems through system dynamics methods to form an efficient and coordinated monitoring system. The system integrates a number of advanced technologies to realize the whole process from data collection to carbon emission forecasting. The overall architecture includes data acquisition and preprocessing module STL, LightGBM time series decomposition module, prediction model module and streaming computing architecture module. Each module works closely together and works together. As shown in the overall structure diagram of the model in Figure1:

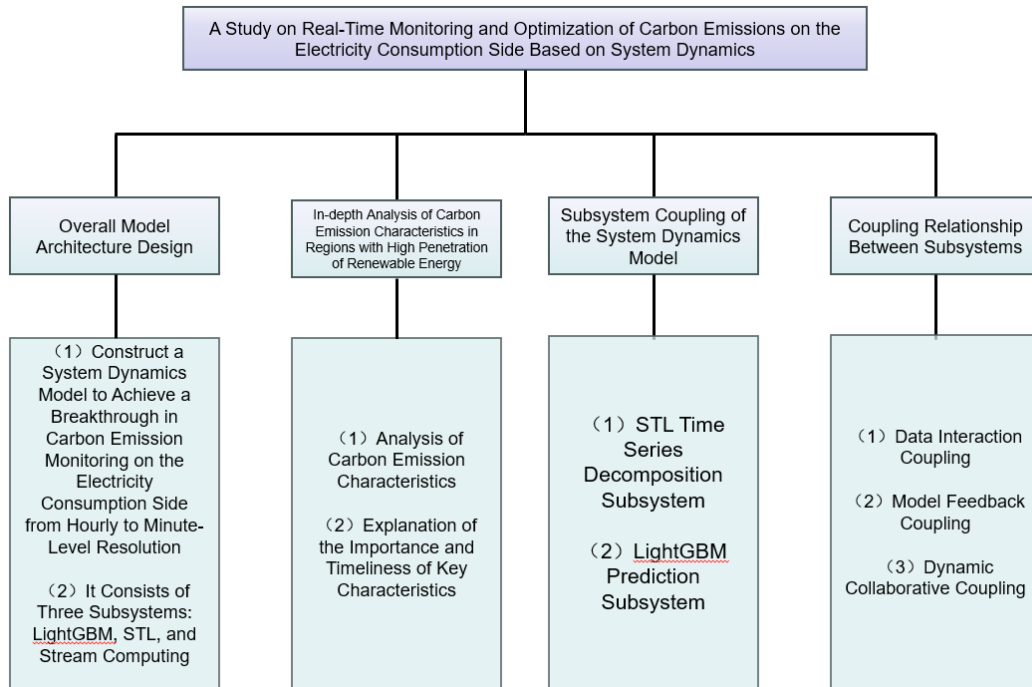


Figure 1 Model overall architecture diagram

##### 4.2 Coupling relationship between subsystems

**Data interaction coupling.** This system achieves tight connectivity between the streaming computing subsystem, the STL subsystem, and the LightGBM subsystem through data flow. The streaming computing subsystem is responsible for acquiring raw data in real time and preprocessing it, such as data cleaning, format conversion, and preliminary filtering to remove noise and outliers. The processed data is then passed to the STL subsystem. The STL subsystem extracts time-series features from this data, identifying key periodicity, trend, and seasonal characteristics. The extracted feature parameters serve as input variables for the LightGBM model, driving the construction and parameter optimization process of the prediction model. Specifically, in the short-term high-energy-consuming scenario of heavy industry in Henan, the real-time operating data of high-energy-consuming equipment collected by the streaming computing subsystem, after the periodic features of power fluctuations are extracted by the STL subsystem, provide accurate input to the LightGBM subsystem, enabling it to more accurately predict carbon emissions. This data interaction ensures efficient and accurate data transmission and sharing between subsystems, laying the foundation for the stable operation of the entire system.

**Model feedback coupling.** The system constructs a model feedback closed loop, triggering an adaptive coordination mechanism between subsystems through real-time interactive feedback signals, thereby achieving dynamic linkage and collaborative optimization between functional modules. The prediction results of the LightGBM subsystem will be fed back to the streaming calculation subsystem for real-time update of model parameters and adjustment of monitoring strategies. When the prediction results show that carbon emissions are about to fluctuate abnormally, the streaming computing subsystem will dynamically optimize the sampling strategy and processing process based on the real-time data load based on this prediction result, and realize the precise regulation of the monitoring process and the intelligent allocation of computing resources through

adaptive adjustment of the collection frequency and data processing methods. At the same time, the STL subsystem will also optimize and adjust the feature extraction process according to the feedback of the LightGBM subsystem. If the prediction accuracy does not meet expectations, the STL subsystem will re-evaluate the accuracy of feature extraction and adjust the feature extraction algorithm to better meet the needs of the prediction model, thus improving the prediction accuracy. The feedback mechanism drives the system to perceive data flow and predict output in real time. Through dynamic calibration of operating parameters and strategy configuration, it realizes the instant optimization of the system response speed and the continuous enhancement of scene adaptation ability.

Dynamic synergistic coupling. During the operation of the system, the three subsystems realize dynamic collaborative operation by sharing state variables and information flows under the framework of the system dynamics model. When the operating power of high-energy-consuming equipment fluctuates, the streaming computing subsystem can quickly perceive this change and transmit relevant data to the STL subsystem in real time. The STL subsystem quickly extracts the fluctuation characteristics and transmits the characteristics to the LightGBM subsystem. The LightGBM subsystem dynamically optimizes the model parameters based on the extracted characteristic parameters, and quickly iterates the prediction model through the incremental learning mechanism, which significantly improves the accuracy and response speed of carbon emission monitoring results. The synergistic effect formed by the subsystem and data preprocessing, feature engineering and other modules not only runs through the whole process of data cleaning, feature extraction to model training, but also reflects the overall performance of system resource scheduling and task execution. The close cooperation between subsystems enables the whole system to effectively respond to various dynamic changes and uncertainties in short-term high-energy-consuming scenarios in Henan heavy industry, such as the sudden operation and downtime of high-energy-consuming equipment. Through this dynamic coordination, the system has achieved a breakthrough in carbon emission monitoring from hour to minute, providing strong technical support for real-time scheduling, energy management and carbon trading of the power grid.

Relying on the highly coordinated coupling mechanism between subsystems, the system dynamics model effectively activates and integrates the core functional advantages of each submodule, realizes the organic combination of data processing, feature extraction, model prediction and real-time monitoring, and provides an effective solution for the minute-level monitoring of power grid electricity-side carbon emissions in typical short-term high-energy-consuming scenarios in Henan heavy industry.

### **4.3 Analysis of carbon emission characteristics**

In typical short-term high-energy-consuming scenarios within Henan's heavy industry sector, different types of high-energy-consuming equipment exhibit diverse operating modes and carbon emission characteristics. Production equipment in high-energy-consuming industries such as steel, cement, and chemicals typically has large electrical loads and complex operating modes. Its carbon emission characteristics are closely related to the production process, exhibiting significant suddenness and non-linearity. For example, short-term high-power operation of electric arc furnaces can lead to a sharp increase in carbon emissions within a short period, while the intermittent operation of rolling mill equipment can also cause fluctuations in carbon emissions. Furthermore, as some high-energy-consuming enterprises in Henan gradually introduce distributed energy, their electricity consumption behavior and carbon emission characteristics have become even more complex.

The sudden operation of high-energy-consuming equipment is an important factor affecting carbon emission monitoring. The start-up and shutdown of these devices are usually unpredictable, which may cause instantaneous changes in the grid load and thereby affect the real-time monitoring and calculation of carbon emissions. In the heavy industry sector of Henan Province, this phenomenon is more pronounced. Traditional monitoring methods are unable to meet the requirements of high precision and real-time performance, and more precise monitoring methods are needed to accurately capture the carbon emission characteristics of these equipment.

Currently, the extraction and analysis of carbon emission characteristics do not adopt the stream computing approach, which has obvious limitations. The traditional methods rely on fixed time intervals for data collection and batch processing, which cannot promptly capture the real-time dynamic changes of high-energy-consuming equipment operation. As a result, the carbon emission monitoring data is lagging and inaccurate. At the same time, the computational efficiency is low when dealing with large-scale data, making it difficult to meet the real-time processing requirements of the massive data in the heavy industry sector of Henan.

To achieve streaming computing, the key lies in designing an efficient streaming data processing architecture and selecting a lightweight model that meets real-time requirements. The streaming data processing architecture must meet the performance requirements of high concurrency throughput and real-time response, continuously receiving dynamic data streams and simultaneously completing processing and analysis tasks. The lightweight model must, while ensuring accuracy, possess fast

computing power and low resource consumption to adapt to the real-time requirements of streaming computing, thereby achieving real-time and accurate monitoring and analysis of carbon emission characteristics.

#### **4.4 System dynamics model subsystem coupling**

##### **4.4.1 STL time series decomposition subsystem**

STL is an analysis method developed by the Cleveland research team in 1990, which utilizes local weighted regression technology to decompose time series. Compared with conventional time series decomposition technology, the STL method shows more prominent advantages in applicability and accuracy when processing complex data. First, this method can handle any type of seasonal pattern. Whether the seasonal change is a fixed period or a non-fixed period, it can achieve effective identification and decomposition. Second, STL shows strong robustness to abnormal values, which can reduce the interference of abnormal data to the decomposition results, so as to ensure the accuracy of decomposition results. Third, the STL method can adapt time series data with multiple characteristics through the flexible configuration of parameters.

In this study, the STL method is used to decompose the carbon emission time series on the electricity side of the Henan power grid and split it into trend components, seasonal components and residual components. This decomposition process helps to deeply understand the long-term evolution trend of carbon emissions, periodic fluctuation laws and random fluctuation characteristics, provides key time series characteristics for subsequent carbon emission forecasts, and is an important basis for building an accurate forecasting model.

Its decomposition process: STL decomposes the time series  $Y_t$  into trend components  $T_t$ , seasonal components  $S_t$  and residual components  $R_t$ , Its mathematical expression is:

$$Y_t = T_t + S_t + R_t \quad (1)$$

Among them,  $Y_t$  represents the original carbon emission time series and contains all the information about carbon emissions;  $T_t$  reflects the long-term change trend of the time series, reflecting the long-term impact of factors such as energy restructuring, technological progress and other factors on carbon emissions;  $S_t$  represents the seasonal component, reflecting the periodic fluctuations in the time series, such as the periodic changes in carbon emissions caused by changes in electricity demand caused by the change of seasons;  $R_t$  represents the residual item after the decomposition of the time series, specifically refers to the remaining fluctuation part after excluding the trend component and the periodic seasonal component, which usually originates from unpredictable random disturbance effects, such as sudden electricity consumption abnormal events, temporary policy adjustments, etc. on carbon emissions. This decomposition method can clearly analyze the multi-dimensional characteristics of carbon emission time series and provide key data support for subsequent analysis and forecasting. We have plotted the area chart representing the decomposition result of the time series data, with the X-axis indicating time (in months) and the Y-axis indicating carbon emissions (in tons). The area chart of the decomposition result of the time series data is shown in Figure 2:

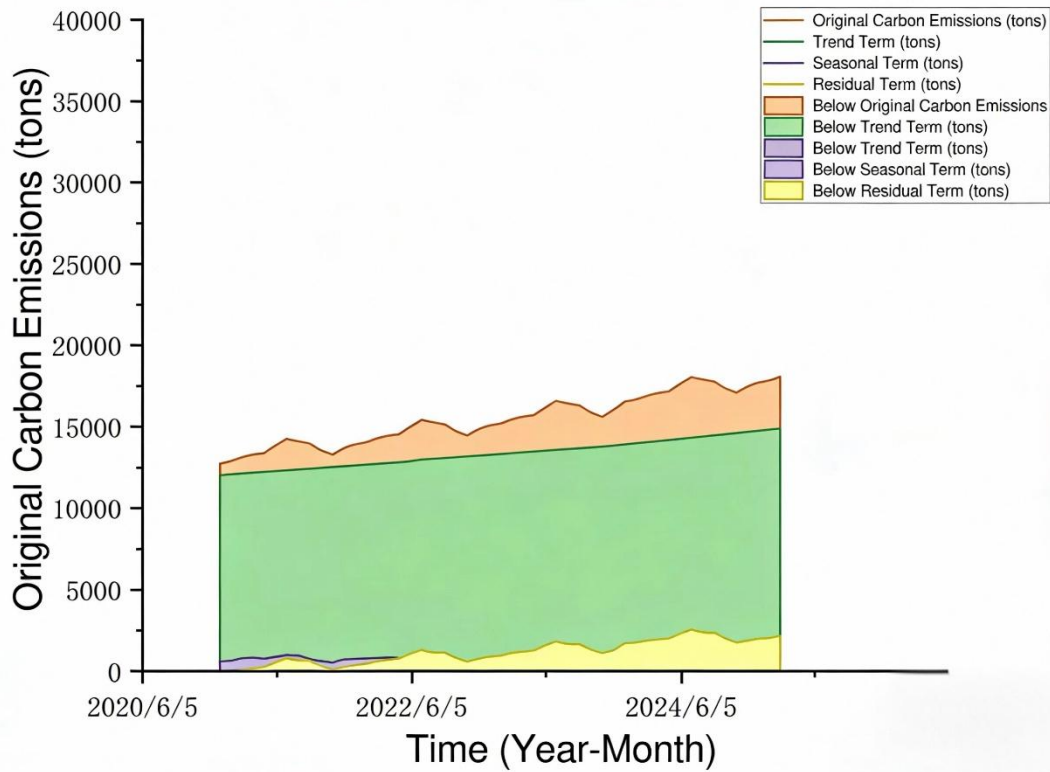


Figure 2: Area chart showing the decomposition result of the time series data

The decomposition steps are:

First, go to the trend. First, remove the trend component  $T_t$  from the original time series to get the detrend sequence  $Y_t - T_t$ . In practice, the trend component  $T_t$  is often estimated with the help of moving average method, polynomial fitting and other techniques. The moving average method realizes data smoothing based on the data mean calculation of a fixed time span and effectively extracts long-term trend characteristics; polynomial fitting technology accurately describes the trend change law in the original data by constructing the optimal fitting polynomial function. Through these methods, the trend component can be removed to achieve the preliminary separation of the long-term change trend of the time series.

Second, seasonal extraction. Seasonal feature extraction is carried out for the data sequence of the removed trend components, so as to separate the seasonal components. This processing link is usually realized by data smoothing methods such as moving average method and Fourier transform technology. The moving average highlights the periodic fluctuation characteristics by calculating the detrend sequence in a specific seasonal cycle; the Fourier transform is based on the combination of trigonometric functions to fit the periodic changes of the data, which can more accurately extract the seasonal components, so as to determine the seasonal change law of carbon emissions.

Third, residual calculation. Calculate the residual component  $R_t$ , that is:

$$R_t = Y_t - T_t - S_t \tag{2}$$

Through this calculation process, the pure random fluctuation data after excluding trend items and seasonal characteristics can be separated. The residual component reflects random changes in the time series that cannot be explained by trends and seasonal factors, helping to analyze the impact of sudden and unpredictable factors on carbon emissions.

Fourth, iterative optimization. The STL method continuously improves the accuracy of decomposition by iteratively optimizing the above steps. In each iteration, the estimates of trends, seasons and residual components are updated according to the results of the previous iterations. After many iterations, the decomposition results are closer to the real situation and provide more reliable basic data for subsequent predictive analysis.

The specific carbon emission trend breakdown chart is shown in Figure 3 2021-2025 carbon emission trend breakdown chart:

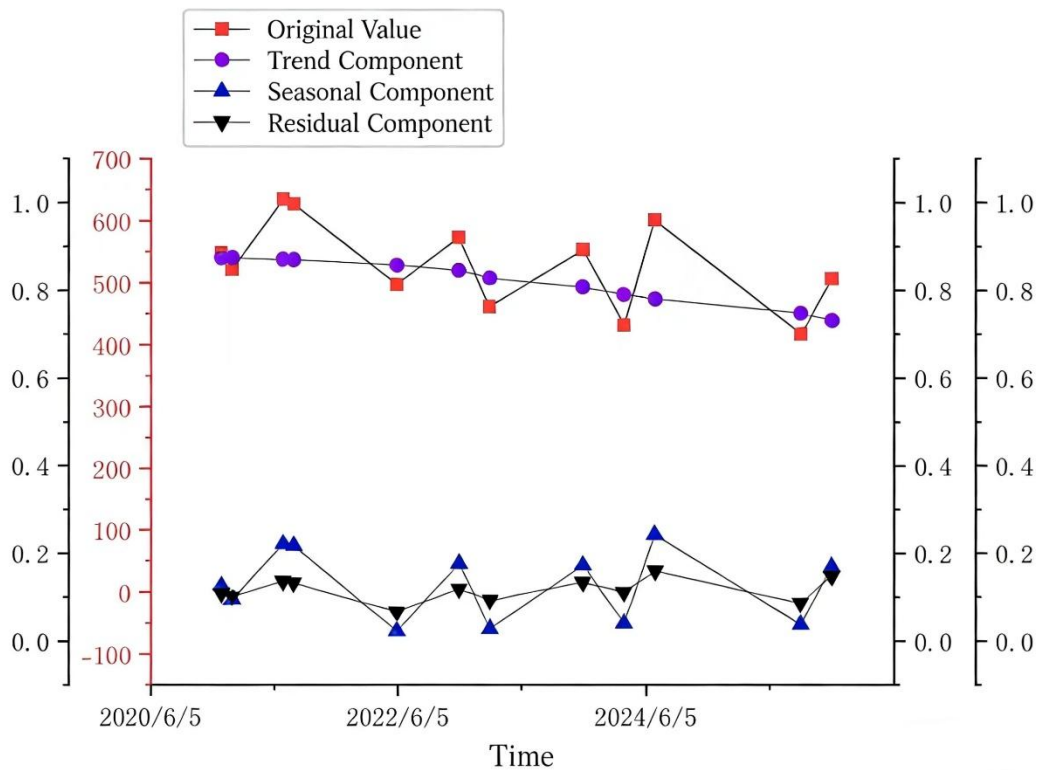


Figure 3 Decomposition Chart of Carbon Emission Trends from 2021 to 2025

#### 4.4.2 LightGBM Prediction Subsystem

LightGBM is a gradient enhancement framework developed by Microsoft. Because of its unique optimization strategy, it shows excellent performance in the fields of data mining and predictive modeling. Compared with the traditional gradient enhancement algorithm, LightGBM shows significant advantages in model efficiency and accuracy: it innovatively applies histogram algorithms to achieve feature discrete processing, effectively reducing memory consumption and computational complexity; integrates core technical optimization means such as unilateral gradient sampling (GOSS) and mutually exclusive feature bundling (EFB) to greatly improve training efficiency; at the same time, leaf-wise tree growth strategy (leaf-wise) replaces the traditional hierarchical growth strategy (level-wise), which achieves a lower error rate under the same number of iterations conditions and significantly enhances the model prediction accuracy.

Based on the three core components of STL decomposition output - trend terms, seasonal characteristics and random fluctuation components, this study builds an adaptive LightGBM prediction model respectively. In the process of model construction, carefully select the appropriate feature set for each time series component. For trend components, select characteristics related to long-term energy policies, industrial restructuring, etc., which can reflect the factors affecting long-term changes in carbon emissions; for seasonal components, select characteristics related to seasonal changes, holiday laws, etc. to capture the seasonal fluctuation laws of carbon emissions; for residual components, select characteristics related to sudden electricity consumption events, short-term policy changes, etc. to assess the impact of random factors on carbon emissions.

At the same time, the cross-verification method is used to optimize the model parameters. By dividing the training set and the verification set many times, the performance of the model under different parameter combinations is tested, and then the optimal combination of key hyperparameters such as learning rate and tree depth is determined through parameter optimization. The model optimizes the loss function between the predicted value and the actual observation. In practice, the mean square error (MSE) or the average absolute error (MAE) is often used as the loss function of quantitative prediction deviation. Take MSE and MAE as examples, the calculation formula is:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (3)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (4)$$

Among them  $n$  is the number of samples,  $y_i$  is the actual value, and  $\hat{y}_i$  is the predicted value. By minimizing the mean square error (MSE), the deviation between the model forecast results and the actual carbon emission monitoring value is minimized, thus improving the accuracy and credibility of the prediction results.

### 3. Stream Computing Subsystem

To achieve real-time monitoring of carbon emissions on the power grid's user side, the research's stream computing subsystem is based on the Flink framework. This system architecture is responsible for real-time data processing and model inference tasks, ensuring the timeliness and execution efficiency of the monitoring process. The overall architecture consists of four parts: the data acquisition layer, the data processing layer, the model inference layer, and the result output layer.

**Data Acquisition Layer.** As the front-end data collection module, it is dedicated to obtaining electricity consumption data and related external information in real time from various heterogeneous data sources. The data sources include smart electricity meter terminals, meteorological data service interfaces, macroeconomic statistics platforms, etc. The collected data is transmitted through the high-throughput distributed messaging middleware Kafka. This component is designed based on a distributed architecture, providing reliable message persistence and high-speed transmission capabilities, ensuring that the entire data flow process is loss-free and delivering a stable data stream to the downstream computing link.

**Data processing layer.** Based on the stream computing engine Flink, real-time data in the Kafka message queue is processed through cleaning, transformation, and aggregation operations. In the data cleaning stage, predefined rules are applied to identify and eliminate noise interference, duplicate records, and outliers, thereby improving data purity; in the transformation stage, the original data format is standardized to a form suitable for the model, such as normalizing numerical variables to eliminate dimensional differences to ensure feature comparability; in the aggregation module, key feature indicators are extracted through statistical calculations, such as the average electricity load of users and the cumulative carbon emissions per day, providing high-precision input data for model reasoning.

**Model inference layer.** As the core analysis module, it conducts real-time predictive analysis on the preprocessed data stream by invoking the already trained LightGBM model. The powerful computing capability of Flink enables the model to quickly process the input data and obtain carbon emission prediction values based on the training results of the model, meeting the strict timeliness requirements of real-time monitoring.

**The results output layer** is responsible for storing, distributing, and visualizing the predicted data. This layer writes the predicted results generated by model inference into the database system, supporting diverse storage solutions such as relational or distributed databases to meet the needs of long-term data archiving and efficient retrieval. Simultaneously, it provides data services to downstream systems through standardized API interfaces, supporting real-time access to the predicted results by the front-end display system and other business applications. After receiving the data, the front-end display module uses visualization components such as charts and reports to transform the carbon emission prediction information into an intuitive graphical interface, enabling graphical presentation and interactive access to the monitoring results. Other business systems can then utilize these prediction results for business operations such as carbon emission early warning and power grid dispatching decisions, according to their own needs.

The core advantage of streaming computing architecture is that it can process massive data in real time and effectively ensure the timeliness of carbon emission monitoring. Flink's built-in fault-tolerant mechanism and precise one-time processing semantics effectively ensure the accuracy and execution reliability of data processing. This feature allows it to maintain stable data processing quality in complex network topology and massive data concurrent processing scenarios, and also ensure that data is not lost and processed repeatedly, providing stable and accurate data support for carbon emission monitoring. Its functions include:

**Data acquisition preprocessing.** Implement the real-time acquisition and standardized processing process of multi-source timing data: the system captures the minute-level monitoring data generated by heterogeneous data sources such as smart meters and meteorological monitoring equipment in real time through the Internet of Things interface, and then implements data purification (including noise filtering, duplicate record elimination, abnormal value detection), format conversion (unified into analysis-ready data structure) and feature engineering (including data aggregation operation and core indicator extraction)

based on the distributed flow processing framework, and finally outputs a standardized data set that meets the analysis requirements, and provides high-quality real-time data input for the STL subsystem and LightGBM subsystem.

Model reasoning and dynamic update module. Perform real-time prediction and adaptive optimization: The system implements online inference on the preprocessed data flow based on the training convergence LightGBM model. Through the high concurrent processing capacity of the Flink flow computing engine, it analyzes input data in real time and outputs carbon emission prediction results to build a dynamic monitoring system for carbon emissions on the power side of the power grid. The module integrates the online learning mechanism synchronously. Based on real-time data flow feedback and predictive deviation analysis, it dynamically adjusts the model parameters and monitoring rule library to ensure that the monitoring model continues to maintain the optimal performance state with the changes in the operation conditions of the power grid and the evolution of the dynamic characteristics of the system.

Output and display of results. The forecast results are displayed to users and decision-makers in the form of intuitive and easy-to-understand charts, reports, etc., so that they can clearly understand the real-time status and changing trends of carbon emissions on the electricity side. Provide decision-making support and management suggestions based on monitoring results, such as optimizing scheduling strategies, formulating energy-saving measures, carrying out carbon trading, etc., to provide a scientific basis for the low-carbon and efficient operation of the power grid. In addition, the streaming subsystem is also responsible for storing the monitoring results in the database for long-term data storage and historical data analysis.

As shown in Figure 4, the streaming computing architecture diagram:

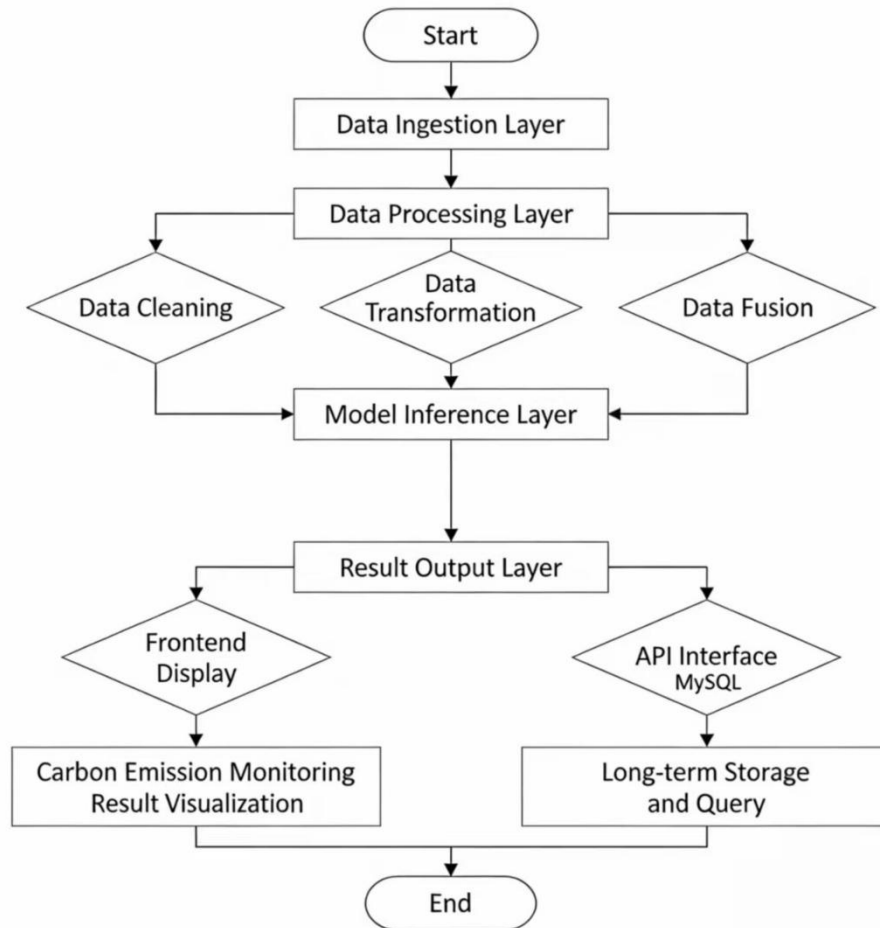


Figure 4: Stream Computing Architecture Diagram

The traditional batch processing mode requires accumulating data within a fixed time window and then processing it centrally, resulting in delays ranging from minutes to hours, which cannot meet the second-level data update requirements of the power grid. This study designs a streaming computing architecture based on the event-time semantics and sliding window

technology of Flink, enabling real-time processing of data one by one, with the delay controlled at the millisecond level. Compared to batch processing, the key technological breakthroughs lie in:

First, the state management mechanism stores intermediate computation states through RocksDB, and combined with Flink's Exactly-Once semantics, ensures that data is not lost or duplicated during fault recovery.

Secondly, the parallel processing capability utilizes the TaskManager parallel thread pool to split the data and parallelly feed it into the model inference layer. The throughput is improved by 10 to 20 times compared to batch processing. For example, when there is a sudden change in industrial load, the streaming architecture can capture the carbon emission transient in real time, providing a second-level response support for the dynamic adjustment of the power grid's unit combination. In contrast, the batch processing mode requires waiting for the next computing cycle to output the results.

**Model application**

To systematically test the comprehensive effectiveness of the system dynamics model constructed in the scenario of predicting carbon emissions on the electricity consumption side of the power grid, this study designed and carried out a series of model effectiveness testing experiments from multiple dimensions, mainly covering prediction accuracy comparison, technical breakthroughs, real-time monitoring performance evaluation, and model stability testing.

**5.1 Comparison of prediction accuracy**

**5.1.1 Experimental data and preprocessing**

This study is based on the measured data of a typical short-term high-energy consumption scenario in a heavy industry sector of the Henan power grid. The data includes the user-side electricity monitoring data, environmental meteorological parameters, and the operation status data of high-energy-consuming equipment. The time span of the data is one year, covering different seasons, different production loads, and the sudden operational changes of high-energy-consuming equipment. It is highly representative. In the data preprocessing stage, the initial processing steps focus on purifying the original dataset: systematically eliminating random noise interference and statistical outliers in the collected data to ensure the quality of the data for subsequent analysis. For data missing values, the time series interpolation algorithm is used to reconstruct the missing values to guarantee the completeness of the dataset; subsequently, the STL decomposition method is employed to decompose the carbon emission time series data into three structured features: trend component, seasonal component, and residual term, which are used as input variables for the model to enhance the algorithm's ability to identify potential patterns in the data.

**5.1.2 Selection of Comparison Models**

Among them, the ARIMA model, as a representative time series prediction method, is suitable for the analysis and prediction scenarios of steady time series data; SVM is a commonly used prediction algorithm in the field of machine learning, which has certain advantages in small sample and nonlinear data processing. The purpose of selecting these two models is to comprehensively highlight the advantages of this research model in responding to the task of carbon emission monitoring under the short-term high-energy consumption scenario in the field of heavy industry in Henan from different technical routes.

**5.1.3 Determination of Evaluation Indicators**

In order to comprehensively evaluate the performance of the STL-LightGBM-stream computing model proposed in this paper, we have selected the following commonly used evaluation indicators in Table 1:

<b>Evaluation Criteria</b>	<b>Specific uses</b>
Mean Absolute Percentage Error (MAPE)	The relative error between the predicted value and the actual value
Root Mean Square Error (RMSE)	The absolute error between the predicted value and the actual value
Determination Coefficient (R <sup>2</sup> )	The degree to which the model fits the data
Mean Absolute Deviation (MAE)	The average absolute error between the predicted value and the actual value

The calculation formulas for these indicators are as follows:

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100\% \tag{6}$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \tag{7}$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n y_i^2} \tag{8}$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \tag{9}$$

Among them,  $y_i$  represents the actual value,  $\hat{y}_i$  represents the predicted value,  $\bar{y}$  represents the average of the actual values, and  $n$  represents the sample size.

### 5.2 Model selection comparison

In order to scientifically evaluate the performance advantages of the system dynamics model integrating STL decomposition, LightGBM prediction and flow calculation in carbon emission monitoring, we have selected the following commonly used prediction models for comparison, see Table 2 prediction model comparison table:

Table 2: Comparison of Prediction Models

Model	Comparison Result
ARIMA Model	The traditional time series prediction model is applicable to stationary time series data
LSTM Model	Long Short-Term memory network, a deep learning model, is capable of handling long and short-term dependencies in time series data
LightGBM Mode	A single LightGBM model directly predicts the original carbon emission time series
STL-LightGBM Mode	Combine STL decomposition with the LightGBM model, but use batch processing for prediction

### 5.3 Comparison of experimental results and analysis

The system dynamics model proposed in this paper is experimented with the above-mentioned comparison model on the same data set, and their predictive performance is compared. The experimental results are found in Table 3 compared with the model to predict the experimental data:

Table 3 Compares the experimental data from the model prediction experiments:

Model	MAPE(%)	RMSE(Ten thousand tons CO <sub>2</sub> )	R <sup>2</sup>	MAE(Ten thousand tons CO <sub>2</sub> )
ARIMA	12.4	0.87	0.76	0.68
LSTM	8.9	0.62	0.84	0.49
LightGBM	7.3	0.51	0.87	0.41
STL-LightGBM	6.2	0.43	0.90	0.35
LightGBM-STL-stream computing computation	4.7	0.32	0.94	0.26

The experimental results show that the system dynamics model constructed in this study shows significant advantages in various evaluation indicators: compared with the traditional ARIMA model, the MAPE of the model is reduced by 62.1% and the RMSE is reduced by 63.2%; compared with the single LightGBM prediction model, the MAPE is reduced by 35.6% and the RMSE is reduced by 37.3%. This shows that the trend, season and residual component of the time series can be extracted through STL decomposition, and the real-time forecasting of the flow computing architecture can significantly improve the accuracy of carbon emission forecasting.

Specifically, the prediction accuracy of the ARIMA model is limited in dealing with nonlinear characteristics and non-steady time series data. The high MAPE and RMSE values reflect that it is difficult for the model to accurately capture the complex change law of carbon emissions on the electricity side of Henan power grid. Although the LSTM model can learn the long-term and short-term dependencies of data to a certain extent, its prediction accuracy still has room for improvement in the face of high-dimensional, multi-source and complex fluctuation data such as power grid data. The single LightGBM model directly predicts the original carbon emission time series and fails to fully exploit the trends, seasons and other characteristics within the data, so the prediction effect is not as good as the LightGBM model after STL decomposition and preprocessing. Although the STL-LightGBM model uses STL decomposition, it adopts batch processing and cannot reflect the dynamic changes of data in real time. In contrast, the STL-LightGBM-stream computing model in this article is more advantageous in real-time and accuracy.

In addition, we also tested the real-time nature of the model. Experimental results show that the system dynamics model can complete the processing and prediction of new data in 1 second to meet the needs of real-time monitoring of power grid carbon emissions. In practical application scenarios, fast response speed means that it can timely capture changes in carbon emissions on the power side of the power grid, provide timely data support for decision-making such as power grid scheduling and energy management, avoid decision-making errors caused by data processing delays, thereby enhancing the safety, stability and operational economy of power grid operation, and effectively promoting the realization of low-carbon emission reduction goals. As shown in the comparison chart of model prediction performance in Figure 5.

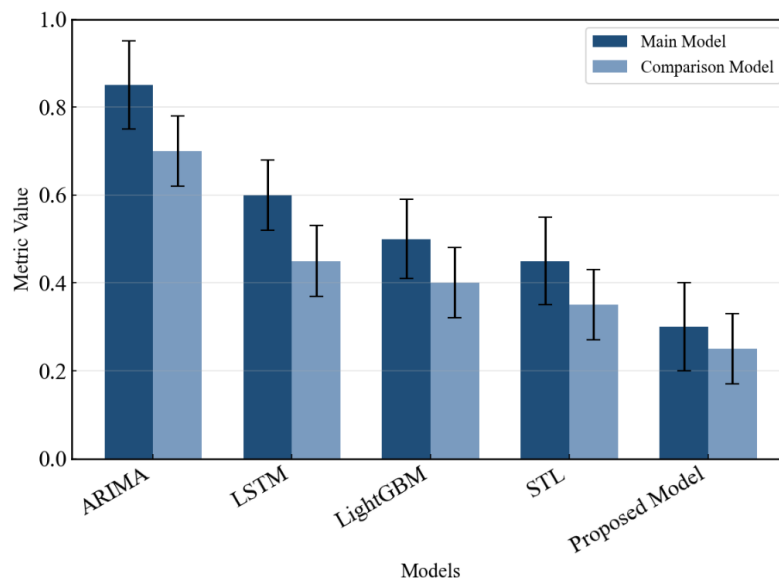


Figure 5 Comparison Chart of Model Prediction Performance

## 5.4 Performance testing and analysis

### 5.4.1 Real-time testing and dynamic prediction

In order to verify the real-time processing ability of the system dynamics model based on STL-LightGBM-stream computing coupling, this study builds a simulation test environment and conducts an in-depth test of the real-time nature of the system. Test data confirms that the system shows excellent performance in real-time data flow processing. The average delay of the whole process from data acquisition to the output of the final forecast results is controlled within 1 second. Compared with the monitoring method of the traditional batch mode, the delay is shortened by two orders of magnitude. This efficient real-time processing capability provides a solid technical guarantee for the minute-level monitoring of carbon emissions on the power side of the power grid.

In the traditional batch processing mode, the model predicts batch data based on fixed time intervals and cannot capture the dynamic changes of the data in time. The system architecture of this study collects and preprocesses data in real time through the streaming computing subsystem. The STL subsystem quickly extracts time series features, and the LightGBM subsystem uses these features for real-time prediction. Flink's powerful computing power supports the rapid processing of data and the real-time loading of models, so that the system can quickly process input data and obtain carbon emission forecast values, meet the timeliness requirements of real-time monitoring, timely reflect changes in power grid carbon emissions, and provide data support for power grid scheduling and energy management decision-making.

#### **5.4.2 System Fault Tolerance and Reliability**

The Stream Computing Subsystem Of This Study Has A Strong Fault Tolerance Mechanism And Exactly-Once Semantics. Flink's fault tolerance mechanism can ensure that data is not lost or processed repeatedly in complex network environments and large-scale data processing scenarios. This reliability guarantee enables the carbon emission monitoring system to maintain stable operation in the process of long-term and high-frequency data processing, which provides a solid technical foundation for the accurate monitoring of carbon emissions on the power side of the power grid under typical short-term high-energy-consuming scenarios in Henan heavy industry. In order to evaluate the stability of the model under long-term operating conditions, this study carried out a 30-day continuous operation test. The results show that there was no failure or stuttering in the system during the whole operation, and the performance indicators were always stable. The model prediction results were coherent and reasonable, and there were no abnormal fluctuations or mutations. Even in the extreme case of data loss rate of 10%, the model can still accurately restore the change trend of carbon emissions based on existing data, and the attenuation of the prediction accuracy is controlled within a reasonable range.

#### **5.4.3 Performance Summary**

In summary, the system dynamics model based on STL-LightGBM-flow computing coupling proposed in this study shows significant technical advantages in the monitoring of electricity-side carbon emissions in power grids in typical short-term high-energy-consuming scenarios in the field of heavy industry in Henan. Compared with the traditional batch processing mode, the system has made a major breakthrough in data processing timeliness and model prediction dynamics. First of all, the breakthrough of data processing timeliness enables the system to process data in real time, and the delay from data collection to the output of prediction results is greatly shortened, meeting the needs of minute-level monitoring. Secondly, the dynamic breakthrough of model prediction enables the system to capture the dynamic changes of data in real time, reflect the changing trend of carbon emissions in time, and provide timely and accurate data support for power grid scheduling and energy management. These technological breakthroughs not only improve the real-time and reliability of the system, but also provide effective technical means for the monitoring of power grid carbon emissions in the field of heavy industry in Henan, and provide key support for the power industry to promote low-carbon transformation and achieve sustainable development goals.

#### **Conclusion**

Driven by the global dual-carbon goal, this study focuses on the carbon emission characteristics of the power grid in the field of heavy industry in Henan Province, especially the carbon emission characteristics of high-energy-consuming equipment and industrial electricity mode in a specific period of time. In order to more accurately analyze the carbon emission characteristics under short-term high-energy consumption, a new system dynamics model is proposed. This model combines STL, LightGBM and streaming computing technology to realize minute-level carbon emission monitoring of the power grid, so as to break through the limitations of the traditional hourly monitoring method and respond to the changes in carbon emissions on the power side of the power grid. This study conducts an in-depth analysis of the carbon emission characteristics of short-term high-energy consumption scenarios in Henan heavy industry, and deeply considers the operation mode of high-energy-consuming equipment and the suddenness and non-linearity of short-time high-energy consumption events. Based on this, this study constructs an STL-LightGBM-stream computing system. The system consists of three core sub-modules of collaborative work: the STL decomposition module accurately identifies the trend characteristics, seasonal laws and cycle change patterns in the time series data; the LightGBM modeling module carries out model training and parameter optimization based on the extracted feature set to effectively improve the prediction accuracy; the streaming processing module relies on the Flink distributed computing framework to complete the efficient processing of real-time data flow and the dynamic update of model parameters. Through the close coupling and coordination of the three subsystems, the system can more effectively improve the time resolution and accuracy of carbon emission monitoring, and also solve the shortcomings of traditional monitoring methods in terms of timeliness and accuracy.

After experimental verification, the model has fast data acquisition ability and can update the carbon emission factor in real time within minutes, thus significantly improving the detection efficiency of carbon emission data. The model can also reduce the instantaneous peak monitoring frequency of industrial electricity to once every few minutes, and shorten the identification

response time of abnormal carbon emission events to less than 30 seconds. Compared with ordinary methods, the new model performs excellently in various evaluation indicators, which can greatly improve the accuracy of carbon emission forecasting, which will provide solid technical support for real-time scheduling and carbon trading of the power grid. In addition, the results of this study provide innovative technical means for low-carbon transformation and sustainable development in the field of heavy industry in Henan. Policymakers make wiser decisions based on our data model, which helps to improve the operational efficiency and carbon emission management of the power industry. Our efforts are not only reflected in technological innovation, but also in contributing to energy development and environmental management.

**This research was supported by the Innovation and Entrepreneurship Training State Grid Henan Electric Power Company Program for College Students of Xi'an International Studies University. Project Number: 202510724002**

**Conflicts of Interest:** The authors declare no conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers.

## References

- [1] Yu, B., Zhao, G., An, R., Chen, J., Tan, J., & Li, X. (2021). Research on China's CO<sub>2</sub> emission pathway under carbon neutral target. *Journal of Beijing Institute of Technology (Social Sciences Edition)*, 23(2), 17-24.
- [2] Hu, A. (2021). China's goal of achieving carbon peak by 2030 and its main approaches. *Journal of Beijing University of Technology (Social Sciences Edition)*, 21(3), 1-15.
- [3] Zhang, Z., & Kang, C. (2022). Challenges and prospects for constructing the new-type power system towards a carbon neutrality future. *Proceedings of the CSEE*, 42(8), 2806-2819.
- [4] Jiang, S., Du, H., & Xu, M. (2024). Dual transition of energy and industrial structure under the carbon peaking and neutrality goals. *J. Quant. Technol. Econ*, 41, 109-130.
- [5] Yinbiao, S., Liying, Z., Yunzhou, Z., Yaohua, W., Gang, L., & Bo, Y. (2021). Carbon peak and carbon neutrality path for China's power industry. *Strategic Study of CAE*, 23(6), 1-14.
- [6] Hao, W., Xin, R., & Junjie, Z. (2023). Current situation and review of carbon dioxide emission monitoring technology in power generation industry. *Thermal Power Generation*, 52(7), 1-13.
- [7] Mingda, L. I. U., Jijun, M. E. N. G., & Bihan, L. I. U. (2014). Progress in the studies of carbon emission estimation. *Tropical Geography*, 34(2), 248-258.
- [8] Fang, X. L., Liu, Y. P., Li, L., Lei, Q. X., & Qin, X. M. (2025). Analysis of carbon emission accounting methods in industrial parks—Emission factor method. *Energy Conservation & Environmental Protection*, (3), 2-9.
- [9] Zheng, Z. Q. (2023). Research on carbon emission measurement and monitoring system based on IoT power carbon measurement model. *China Measurement*, (10), 95-99.
- [10] Xu, G. Q., & Zou, J. (2006). System dynamics method: Principles, characteristics and recent progress. *Journal of Harbin Institute of Technology (Social Sciences Edition)*, (4), 72-77.
- [11] Yu, Q. Y., Dai, X. M., Shu, Y. H., & Peng, F. (2022). Comprehensive evaluation model of power supply and consumption interaction benefit based on system dynamics. *Electric Power Science and Engineering*, 38(06), 9-17.
- [12] Wu, C., Cui, J. Y., Huang, F., Song, C. Y., Zhang, S., Zhao, P. B., & Zhang, H. P. (2025). Rapid salinity prediction of Chenhang Reservoir based on STL-MIKE-LSTM joint modeling. *Netwater Technology*, 44(5), 57-68, 186.
- [13] Shi, Y. T. (2025). Research on surface runoff prediction in Tangshan City based on STL-LSTM. *Shaanxi Water Resources*, (04), 26-28, 31.
- [14] Liu, L., Wang, Y. B., Pang, X. F., & Geng, H. N. (2020). Comprehensive monthly electricity sales forecasting method based on STL model. *Control Engineering of China*, 27(11), 1930-1936.
- [15] Chen, W. N., Hu, Z. J., Yue, J. P., Du, Y. X., & Qi, Q. (2021). Short-term load forecasting based on long short-term memory network and LightGBM combination model. *Automation of Electric Power Systems*, 45(4), 91-97.
- [16] Ma, X. J., Sha, J. L., & Niu, X. Q. (2018). Design and application of P2P project credit rating model based on LightGBM algorithm. *The Journal of Quantitative & Technical Economics*, 35(5), 144-160.
- [17] Li, X. L., & Chen, L. L. (2025). Economic analysis and optimization strategies of renewable energy power generation systems. *Popular Standardization*, (10), 140-142.
- [18] Zhao, R. F., Lou, Y. S., Ye, F., & Shi, H. (2022). Research and application of industrial big data platform based on Flink. *Computer Engineering and Design*, 43(3), 886-894.
- [19] Liu, H., & Chen, G. (2021). Real-time industrial big data analysis platform based on Flink. *Electronic Technology & Software Engineering*, (6), 185-187.
- [20] Guo, M. Y., Kang, H., & Yuan, X. J. (2017). Real-time database partitioning system based on stream computing framework. *Computer Engineering*, 43(11), 8-15.
- [21] Zhang, L. B., Hu, Y. N., Jin, J. L., Wu, C. G., Zhou, Y. L., & Cui, Y. (2021). Dynamic prediction and regulation of water resources carrying capacity in Chaohu Lake Basin based on system dynamics. *Journal of Lake Sciences*, 33(1), 242-254.
- [22] Wang, H. G., Xiao, L. X., & Liao, B. (2022). Simulation of carbon emission reduction paths in China based on system dynamics. *Journal of Natural Resources*, 37(5), 1352-1369.