https://doi.org/10.31449/inf.v49i5.6556 Informatica **49** (2025) 19–36 **19**

Integrating ARIMA Model for Enhanced Financial and Tax Data Management and Accurate Departmental Budget Prediction

Mingyue Lv
School of Accounting, Changchun Finance College, Changchun 130000, China
E-mail: lvmingyue1990@126.com

Keywords: time series, financial and tax data, departmental budgets, leadership decision-making, data management

Received: July 4, 2024

Aiming at the problems of low timeliness and inaccurate budgets in the financial sector, the study constructs a model for fiscal and tax data management and departmental budget forecasting based on the time series model, using the autocorrelation function and partial correlation function to identify the time series characteristics. At the same time, the architecture of the fiscal and tax data management system is designed to realize data fusion and sharing and improve forecasting accuracy. The experimental results indicated that the departmental budget data of a province from 2015 to 2022 was a non-stationary series, which was converted to a stationary series after the smoothing process. The parameters of the time series model were determined by combining auto correlation function, partial auto correlation function and Akaike information criterion. The determined time series model was used to predict the departmental budget and the average error between the TV and PV of the departmental budget for the years 2020 to 2022 were 7.3%, 7.4% and 12.1%, respectively. The prediction errors of the research-designed model, the generalized autoregressive conditional heteroskedasticity model, the long and short-term memory network model and the generative adversarial network model in the static prediction were 12.1%, 10.9%, 28.1% and 28.9%, respectively. In the dynamic prediction, the average prediction errors of the four models were 11.3%, 15.4%, 6.3%, and 10.1%, with the accuracy of 88.7%, 84.6%, 93.7%, and 89.9%, respectively. Therefore, the prediction precision of the research prediction method is high and the error rate is low, which can effectively realize departmental budget prediction analysis and is of practical significance in guiding financial practice.

Povzetek: Raziskava uporablja model ARIMA za integracijo upravljanja finančnih in davčnih podatkov ter kvalitetno napovedovanje proračunov oddelkov, kar izboljša točnost in podpira odločanje vodilnih.

1 Introduction

Economic and social data are being collected due to the rapid development of computer science and computing, which makes it very easy for government agencies to implement intelligent management and services [1-2]. Financial work is related to all aspects of the economy and society, the use of economic and social data can improve the level of financial work. The traditional financial data processing method has problems such as slow processing speed and low efficiency. The use of Chinese customized reports to analyze data can neither meet the demand for data analysis nor provide effective assistance to leadership decision-making [3-5]. Effective analysis and processing of data in the financial business system is an important means to assist leadership decision-making and guide financial practice. Data mining technology is currently an important technology for processing big data and performing knowledge discovery [6]. The application of data mining technology to data warehouses represents a sophisticated process, entailing the selection of optimal analysis algorithms and the extraction or mining of pertinent information [7-8]. Analytical algorithms for data mining approaches that are frequently utilized include neural networks, fuzzy sets, rule-based reasoning, decision trees, example reasoning, applied statistical methods, and genetic algorithms [9]. Times series (TS) algorithm is one of the applied statistical methods to make predictions based on the temporal nature of data. TS algorithms are widely used in medical diagnosis, stock prediction, image analysis and diagnosis of electromechanical systems [10]. Peng et al. proposed a classification algorithm based on TS images to address the difficulty of mapping wetland degradation. The method utilized TS images to synthesize phenological features, combined with random forest algorithm and hierarchical decision tree to classify wetland types. Experimental results indicated that the algorithm has good precision [11]. To estimate the worldwide solar radiation, Mughal and colleagues suggested a prediction time model based on neural networks. To anticipate an appropriate choice of global solar radiation, the strategy integrated a nonlinear autoregressive model with statistical analysis to assess the outcomes. The experimental results indicated that the model has a high precision with a root mean square error of 1.28% [12]. Garai et al. proposed a hybrid and its learning algorithm based on empirical modal decomposition of adaptive noise-complete sets to address the problem of irregularities present in noisy TS data. The method utilized a combination of cryptographic algorithm and particle swarm algorithm to predict noisy TS data. The experimental results demonstrated that the algorithm was able to effectively predict the price series [13]. Li proposed dynamic time warping based on time weighting to address the problem of the influence of different time points. The method obtained matching time weights by measuring the distance between historical time points and the latest time points, combined with dynamic time warping. Experimental results demonstrated that the method was effectively used for similarity measurement in TS data mining [14]. Tan et al. proposed TS external regression algorithm in order to analyze the relationship between learning TS and continuous scalar variables. The method predicted the future values of TS through a temporal classification algorithm, combined with benchmarking. The experimental results indicated that the algorithm had high prediction precision [15].

Financial data management is the use of data warehousing techniques to enable intelligent management of the process of collecting, transforming, analyzing and distributing data. Shahid et al. proposed data warehouse for healthcare sensitive data applications to address the difficulty of monitoring changes in childhood obesity. The method collected children's data through different sensors and created an obesity prevalence model to monitor and store personal data on related behaviors. The outcomes demonstrated that this data warehouse could effectively enhance data precision and privacy [16]. Amo et al. constructed an educational data warehouse for educational data storage, analysis and access in order to improve educational data management performance. The approach utilized modular and extensible data systems to address data management and access complexity. The experimental results demonstrated the usefulness of this educational warehouse in assisting educational decision making and improving data security [17]. An IoT based real-time warehouse management system, suggested by Khan et al., separated the warehouse into several domains and combined architectural views to display the model in order to automate and digitize the administration of the manufacturing business. The approach collected and transmitted data through a generic prototype system. Experimental results indicated that the system was efficient in data transmission with low latency [18]. Kahn et al. migrated a research data warehouse to a public cloud in order to achieve innovation-driven clinical research data, thematically categorizing internal and public presentations and analysis documents before and after the transition. The experimental results showed that this migration process minimized the waste of resources and costs [19]. Tufano et al. proposed data warehouse management system to track data in order to mine logistics data to monitor its business processes, smooth warehousing processes and support strategic decision making. The method predicted benchmark metrics of the data storage system by training classifiers. Experimental results indicated that the warehouse system was effective in realizing data-driven and improving data prediction precision [20]. The results of the related research work are summarized in Table 1.

The application of data mining techniques can facilitate the classification and prediction of financial data, while data warehouse technology can be employed to transform financial data into knowledge. The combination of the aforementioned needs and technologies facilitates the acceleration of data management and mining in the financial sector. The study proposes a financial and tax data management and departmental budget prediction method that integrates the auto-regressive integrated moving average model (ARIMA), with the objective of assisting leaders in decision-making, improving financial sector management, and promoting financial innovation.

Table 1: Summary of related work

Literature number	Author(s)	Methodology	Data set	Results	Limitations
[11]	Peng K, Jiang W, Hou P, Wu Z, Ling Z, Wang X, Mao, D	A new algorithm for detailed wetland type classification combining k-fold random forests and hierarchical decision trees	North and South-Central Asia	Continental-scale wetland maps were obtained with an overall accuracy of 90.0 ± 0.5%.	Classification rules are not well set up, and there are some classification errors.
[12]	Mughal S N, Sood Y R, Jarial R K	Neural networks based on time series	/	The root-mean-square error was 1.28, which was lower than that of other time-series models; the hourly prediction model was more accurate than the daily prediction	/

				model.	
[13]	Garai S, Paul R K, Yeasin M, Paul A K	Hybrid and its learning algorithm based on adaptive noise-complete ensemble empirical modal decomposition; prediction of noisy time series data using encryption algorithm with particle swarm algorithm	Daily Wholesale Prices of Potatoes in Six Major Indian Markets	Optimization-based model combinations outperformed their individual counterparts	Insufficient prediction performance for data with different volatility levels
[14]	Li H	Dynamic time warping extension method based on time weights analysis	UCI data and financial stock exchange data	Different historical time points have different effects on the contribution to the minimum distance between two time series; the method facilitates similarity measures in time series data mining compared to existing methods	Model complexity needs to be simplified
[15]	Tan C W, Bergmeir C, Petitjean F, Webb G	Time series external regression algorithm	Time series external regression dataset	Rocket achieves the highest overall accuracy when adapting regression	Accuracy needs to be improved
[16]	Shahid A, Nguyen T A N, Kechadi M T	Three-tier flexible BigO data warehouse architecture, including backend, access control and controller layers	/	The data warehouse effectively improves data accuracy and privacy	Performance on distributed data needs to be improved
[17]	Amo D, Gómez P, Hernández-Ibáñez L，Fonseca D	Modular and scalable data system architecture	/	Addresses the complexity of data management and access	Model complexity needs to be simplified
[18]	Khan M G, Huda N U, Zaman U K U	An IoT-based real-time warehouse management architecture	A textile mill warehouse	Key performance parameters such as system resiliency, efficiency, and latency are effectively improved.	Architecture scalability, system security are not enough

[19]	Kahn M G, Mui J Y, Ames M J，Yamsani A K, Pozdeyev N, Rafaels N, Brooks I M	A thematic classification model for analyzing documents	/	Network, compute, and storage architectures that realize performance and cost advantages	Internal deployment network and system touch points add complexity
[20]	Tufano A, Accorsi R, Manzini R	Classifier-based warehousing system	Data from 16 different real companies in different industrial sectors and with different participants	Warehouse system is data-driven and improves data prediction accuracy	Insufficient research on predictive design of storage systems

2 Methods and materials

In light of the accelerated advancement of financial informatization technology, the accumulation of financial and tax data is becoming increasingly prevalent. To fully realize the value of these data and to conduct departmental budgeting in a scientific and rational manner, the study proposes the integration of the TS model for financial and tax data management and departmental budget prediction. This approach aims to anticipate financial and tax data trends and to inform the subsequent analysis and processing of these data. The method begins with the construction of the TS model, which is then utilized to analyze and process financial and tax data. This is followed by the mining of departmental historical budget data, which is then used to predict the departmental budget. The TS model serves as a guiding force throughout this process.

2.1 Construction of the times series model

TS data are correlated data that are arranged in chronological order and dynamically change over time, which is a special kind of stochastic process [21-22]. TS algorithms are useful for analyzing and forecasting future data as well as for identifying patterns in previous data. The auto-regressive and moving average model (ARMA) is the most commonly used TS prediction method [23-24]. In ARMA (p, q), AR is autoregressive and p denotes the autoregressive terms. Q indicates the quantity of moving average terms, while MA stands for moving average. ARMA is commonly used to deal with smooth data series [25-26]. The AR model is predicted by linearly combining the current moment perturbation value with the previous moment observation. The calculation is shown in Equation (1).

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \cdots + \phi_p X_{t-p} \quad (1)$$

In Equation (1), X_t denotes the observed value at moment t. $\phi_i\left(i = 1, 2, \cdots, p\right)$ denotes the autoregressive model coefficients. $X_i\left(i = t-1, t-2, \cdots, t-q\right)$ is the observed values at different moments. The perturbation values of the present and preceding moments are combined linearly to predict the MA model. Equation (2) shows the result of the calculation.

$$X_t = \mu_t + \theta_1 \mu_{t-1} + \theta_2 \mu_{t-2} + \cdots + \theta_q \mu_{t-q} \quad (2)$$

In Equation (2), μ_t denotes the residuals and $\theta_i\left(i = 1, 2, \cdots, q\right)$ denotes the moving average model coefficients. The calculation of ARMA model is shown in Equation (3).

$$\begin{aligned} X_t &= \phi_1 X_{t-1} + \phi_2 X_{t-2} + \cdots + \phi_p X_{t-p} \\ &+ \mu_t + \theta_1 \mu_{t-1} + \theta_2 \mu_{t-2} + \cdots + \theta_q \mu_{t-q} \end{aligned} \quad (3)$$

When TS is non-smooth, it needs to be converted to a smooth series after finite difference operation. The ARIMA is developed on the basis of ARMA and can be used for modeling non-smooth TS [27-28]. The TS model prediction testing process is shown in Figure 1.

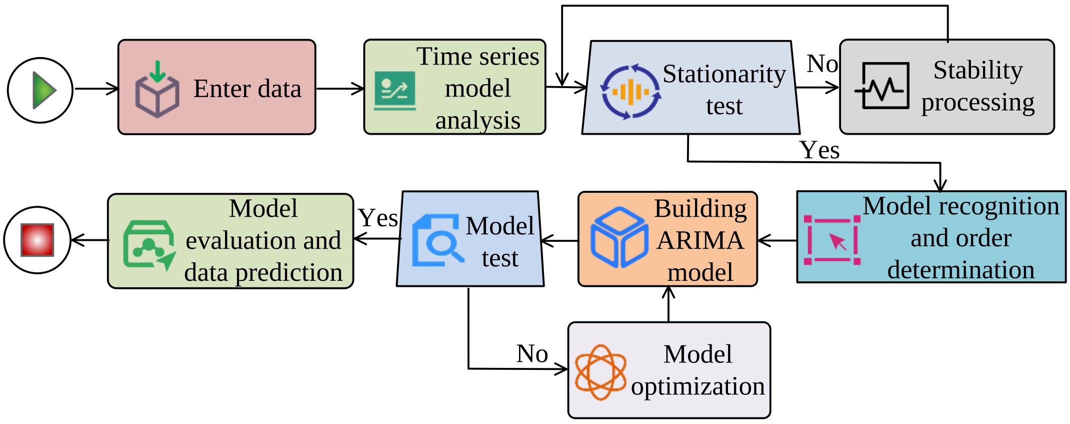


Figure 1: Time series model prediction flowchart

The objective of the stability test is to analyze and preprocess the raw data in order to obtain the model required for the study. This is followed by the preliminary stability test on the series, which is conducted in accordance with the time series diagram. After the stability test, if the data series is stationary, it is directly analyzed and processed. To acquire the differential order, differential processing must be performed if the data series is non-stationary. The data series after differential processing to get a differential series term, specifically as shown in Equation (4).

$$\nabla x_t = x_t - x_{t-1} \tag{4}$$

In Equation (4), ∇x_t denotes the first-order difference series term at moment t. x_t is the data series at moment t. x_{t-1} is the data series at the moment of $t-1$. Similarly, the second-order differential series term is obtained after the second-order difference (SOD) processing of the data series, as shown in Equation (5).

$$\nabla^2 x_t = \nabla x_t - \nabla x_{t-1} \tag{5}$$

In Equation (5), $\nabla^2 x_t$ denotes the second-order differential series term at moment t. ∇x_{t-1} denotes the first-order differential series term at moment $t-1$. Therefore, the higher-order differential series is calculated as shown in Equation (6).

$$\nabla^d x_t = \nabla^{d-1} x_t - \nabla^{d-1} x_{t-1} \tag{6}$$

In Equation (6), d denotes the difference order. $\nabla^d x_t$ and $\nabla^{d-1} x_t$ denote the d th and $d-1$ th differential series terms at the t moment, respectively. $\nabla^{d-1} x_{t-1}$ denotes the $d-1$ times difference series term of $t-1$ moment. If a non-stationary data series has a linear trend, the difference can be used to remove its linear trend and obtain a stationary series. The augmented Dickey-Fuller test (ADF) is used to further determine the stability of the differential TS. If the statistical P-value of the series test is less than 0.05, the differential series is determined to be stationary [29-30]. According to Equation (1), the nonzero characteristic root of the characteristic equation is obtained as x_1, x_2,..., x_p. A stationary series is indicated when the characteristic root's absolute value is smaller than 1. The characteristic root is equal to 1, which indicates that the series is nonstationary series. Model identification is based on the data auto correlation function (ACF) and partial auto correlation function (PACF), to identify the features of TS and construct the corresponding TS model. The data auto correlation is calculated as shown in Equation (7).

$$\hat{\gamma}_k = \frac{1}{n}\sum_{t=k+1}^{n} x_t x_{t-k}, k = 0,1,2,\cdots,n-1 \tag{7}$$

In Equation (7), $\hat{\gamma}_k$ denotes the data self-covariance and $x_i\left(i = 1,2,\cdots,t\right)$ denotes the data series. Equation (8) can be used to compute the self-covariance of the starting data.

$$\hat{\gamma}_0 = \frac{\sum_{t=1}^{n} x_t^2}{n} \tag{8}$$

In Equation (8), $\hat{\gamma}_0$ represents the initial data self-covariance. Therefore, the data ACF is shown in Equation (9).

$$\hat{\rho}_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0} = \frac{\sum_{t=k+1}^{n} x_t x_{t-k}}{\sum_{t=1}^{n} x_t^2} \tag{9}$$

In Equation (9), $\hat{\rho}_k$ denotes ACF. The selection of ARIMA model is determined based on the data ACF and PACF. Table 2 displays the assessment criteria for the ARIMA model.

Table 2: Judgment criteria of ARIMA model

Model	ACF	PACF
$AR(p)$	Trailing	p-order truncation
$MA(q)$	q-order truncation	Trailing
$ARMA(p,q)$	Trailing	Trailing
Model not suitable	Truncation	Truncation

Trailing tails indicate that there are always non-zero values that vary randomly around zero after being greater than a certain order. Truncated tails indicate that the value tends to zero quickly after being greater than a certain order. The ARIMA model can be expressed as ARMA (p, d, q). Once the difference order d is determined, the ARIMA model becomes ARMA (p, q). A deformation of Equation (3) is performed as shown in Equation (10).

$$\begin{aligned}&\left(1-\phi_1 L-\phi_2 L^2-\cdots-\phi_p L^p\right) X_t \\ &=\left(1+\theta_1 L+\theta_2 L^2+\cdots+\theta_q L^q\right) \mu_t\end{aligned} \quad (10)$$

In Equation (10), $1-\phi_1 L-\phi_2 L^2-\cdots-\phi_p L^p$ and $1+\theta_1 L+\theta_2 L^2+\cdots+\theta_q L^q$ denote polynomials of order p and q about L, respectively. Using the delay operator, then the decentralized model is represented as shown in Equation (11).

$$\Phi(L) X_t=\Theta(L) \mu_t \quad (11)$$

The ARIMA model uses Akaike information criterion (AIC) for ordinalization. The AIC function is shown in Equation (12).

$$AIC=n \log \sigma^2+2(p+q) \quad (12)$$

In Equation (12), AIC denotes the function and σ^2 denotes the residual sum of squares. n denotes stationary series sample size. The AIC criterion fixed order calculation is shown in Equation (13).

$$\begin{aligned}&AIC(p,q)= \\ &\min_{k,l} AIC(k,l) \quad 0 \le k \le M, 0 \le l \le H\end{aligned} \quad (13)$$

In Equation (13), $AIC(p,q)$ denotes the AIC criterion fixed order function. M denotes the ARIMA model order maximum. H denotes the ARIMA model order maximum. In ARMA (p, d, q), $AR(p)$ model indicates the relationship between the current data and the data of the previous p period, $p \in \{0,1,2,3\}$. d indicates the number of differences of non-smooth data into smooth data series. The $MA(q)$ model represents the relationship between the current data and the error of the data in the previous q period, $q \in \{0,1,2,3\}$. White noise (WN) detection checks to see if all of the series' relevant data has been fully retrieved, or if the residual term is a WN series. If the series is WN, it indicates that the important data has been taken out of the series. The zero mean homoskedasticity of the WN series is calculated as shown in Equation (14).

$$E(\mu_t)=0 \quad (14)$$

In Equation (14), $E(\mu_t)$ denotes the WN series error. The WN series independent homogeneous distribution is calculated as shown in Equation (15).

$$E(\mu_t,\mu_s)=\begin{cases}\lambda^2, t=s \\ 0 \ \ , t \neq s\end{cases} \quad (15)$$

In Equation (15), $E(\mu_t,\mu_s)$ denotes the error of series μ_t and series μ_s . The model test is to detect whether the residual series of ARIMA model is WN. If it is not WN, then the residuals still contain valuable information, and the model needs to be optimized or more valuable information needs to be extracted. Model prediction is to predict the model that passes the test. The overall construction process of the ARIMA model is shown in Figure 2.

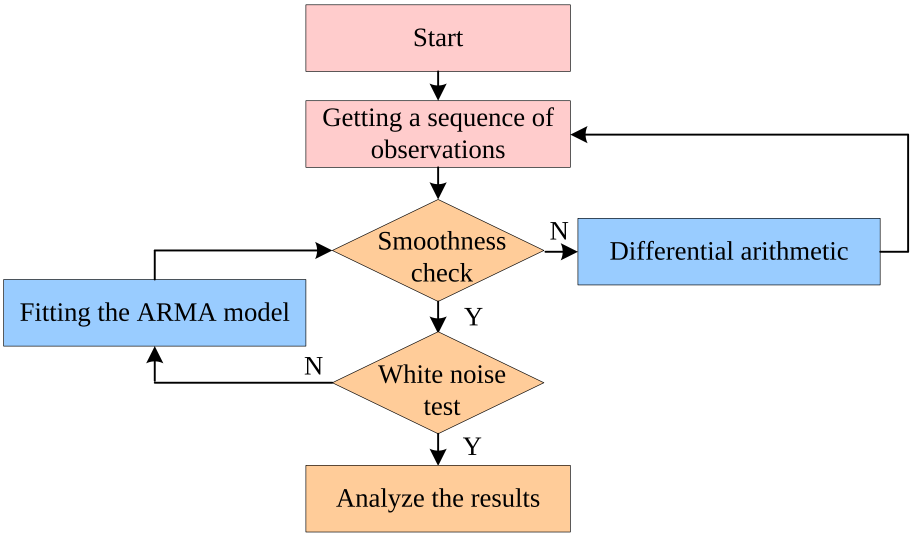


Figure 2: Flowchart of the overall construction of the ARIMA model

As illustrated in Figure 2, the ARIMA1 construction process initially entails the aggregation of time series data from disparate sources, including financial statements, sales records, market research, and so forth. This is followed by the application of linear interpolation to address any gaps in the data set and the utilization of a clustering algorithm to handle discrete values and anomalous data, thereby guaranteeing the completeness and precision of the data. The ADF smoothness test is performed on the processed data, and the results that do not pass the smoothness test must be differentiated until they pass the white noise test. If the sequence does not pass the white noise test, it is necessary to enter the model fitting. If the model passes all tests, the model is used to predict future values.

In addition, to ensure the robustness of the model, the study adopts the K-fold cross-validation method. The dataset is divided into K subsets, each of which is rotated as the test set, and the remaining subsets are used as the training set. The ARIMA model is trained on the training set, and the parameters of the ARIMA model are adjusted according to the characteristics of the data. Moreover, the error metric is used to measure the difference between the predicted and actual values of the model.

2.2 Design of a method for forecasting financial and tax departmental budget data based on ARIMA modeling

The effective management of financial and tax data facilitates the improvement of the precision and scientificity of the prediction of financial departmental budget data. The effective prediction of departmental budget data can be achieved by the comprehensive mining of departmental historical data through data mining technology, which is the financial and tax data analysis of the core content [31-33]. Therefore, the study proposes a prediction method for financial departmental budget data based on the ARIMA model to predict the financial departmental budget data and guide the analysis and processing of financial and tax data. At present, finance-related departments have generally established corresponding business application systems for daily departmental business processing. These systems do not realize data integration and sharing, resulting in lagging financial and tax basic data and business data, reducing the precision and timeliness of data analysis. The study suggests a financial and tax data management system to increase the efficiency and precision of data processing.

The architecture of the financial and tax data management system is displayed in Figure 3. The architecture consists of data collection, storage, application and release. The system integrates the fundamental financial and tax data within the business application framework, facilitating the integration of financial data and providing robust data support for informed decision-making at the executive level.

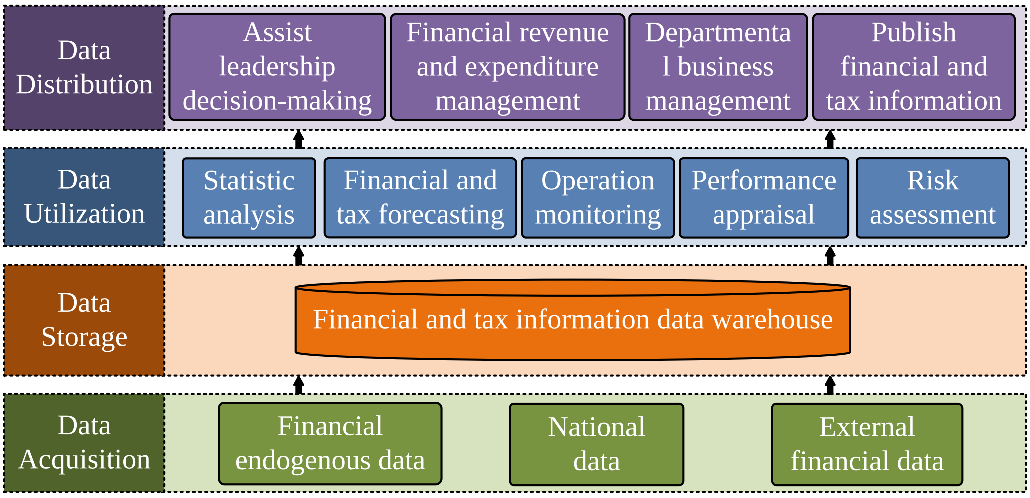


Figure 3: Architecture of financial and tax data management system

The business essential data is extracted, transformed, and loaded by the data collection layer before being uploaded to the data storage layer. The data storage layer analyzes and summarizes the data through the financial and tax information warehouse and prepares for data application. The data application layer performs statistical analysis, financial and tax prediction, and operation monitoring on the data in the warehouse. The data dissemination layer utilizes the results of the application layer for decision-making, management, and information dissemination. The financial and tax data warehouse adopts ERwin data modeling tools to construct the logical subject model of the financial and tax data warehouse and optimize the complex environment of financial and tax data management. The financial and tax data warehouse system divides financial and tax data into key taxpayer models, financial and tax revenue flow models and financial and tax revenue daily report models according to the analysis of financial and tax business data and user report requirements. According to the user requirements, the financial and tax revenue daily model dimension analysis is obtained. The details are shown in representation 3.

Table 3: Key taxpayer model

Dimension	Content
Date	Ten-day serial number; Month; Quarterly; Year; Monthly week series number; Weekly serial number within the year
Standard region	Region code; Region name; Regional level; Superior region code; Name of superior region
Tax level	Tax level code; Description of tax level
Levy item	Levy item code; The name of levy item; Levy item level; Superior levy item code; Name of superior levy item
Levy institution	Tax source code; Description of tax sources

In addition, the reports of the financial and tax revenue daily model contain monthly financial and tax revenues, cumulative annual revenues, monthly revenues for the same period of the previous year, and cumulative annual revenues for the same period of the previous year. On top of the financial and tax revenue daily model dimensions, the focused taxpayer model adds a standard industry dimension. The report contains the cumulative annual completion amount and the completion amount for the same period of the previous year. The financial and tax revenue flow model adds the local tax and national tax dimensions. The report contains the annual cumulative completion amount, monthly completion amount, completion amount in the same period of the previous year, and monthly completion amount in the same period of the previous year. According to the above logical subject model, ETL tools are used to process the raw financial and tax data from the data collection layer to provide basic data for the financial and tax data warehouse. The financial and tax business source data is stored in the Oracle database system. The Trinity tool is utilized to update the database in real time, and finally the data is imported into the Teradata all-in-one machine. Automatic data extraction (ADE) extracts basic information from the Oracle database and generates data files. The data files are downloaded by the Trinity tool, saved to its local directory, processed, and imported into the Teradata all-in-one PC. ADE is able to export csv data files from Oracle database at regular intervals and put the csv data files into the directory of the specified file transfer protocol server. ADE can realize database connection parameter modification, manual export of historical data, automatic export of new data, and other functions. ADE workflow is shown in Figure 4.

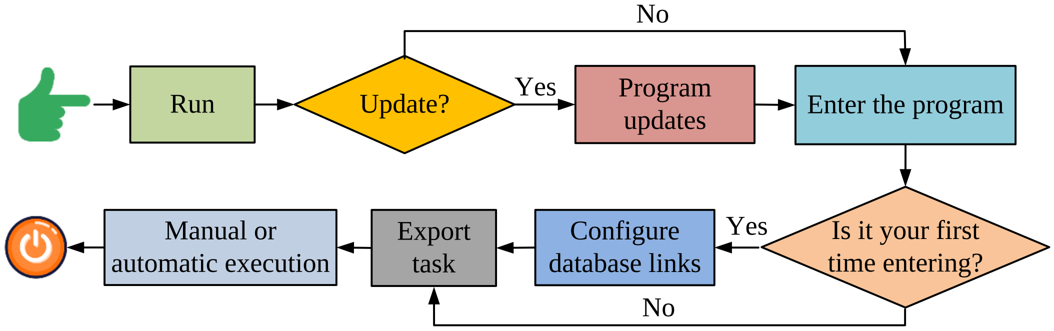


Figure 4: ADE workflow

Teradata Data Warehouse invokes Fastload script to read the data files under Trinity server into the database, completing the process of financial and tax data from the business system to the data warehouse. Similarly, all financial and tax data will be entered into the database to realize data fusion, which is a solid foundation for data analysis and processing. Teradata data warehouse consists of three levels of databases: source database, intermediate database and application database. The source database stores the raw data, the intermediate database stores the data processing results and intermediate tables, and the application database stores the final results of data processing. The analysis and visualization of financial and tax data is implemented by IBM Cognos software. The framework manager component is first used to construct a multidimensional data source and analyze the relationships between the above logical topic models and dimensions. Then the Report Studio reporting component is used for data visualization. Framework manager utilizes the relationship between the fact table and dimension table to generate data packages. The workflow is shown in Figure 5.

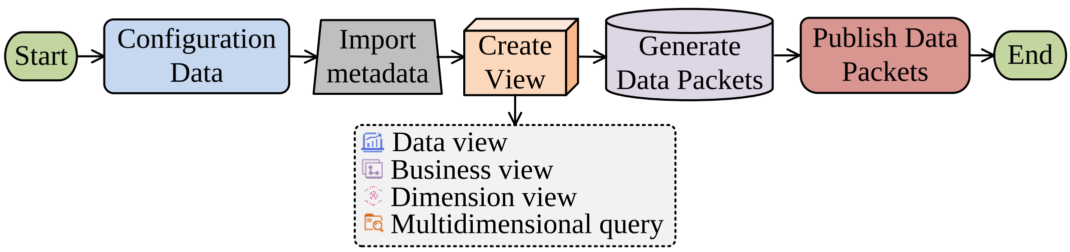


Figure 5: Framework manager workflow

Data view means that the fact table has correspondence in the dimension table. Business view is to change the name of the data table in the view to Chinese explanation. Dimension view is to create multidimension according to business. Multidimensional query is to put the fact table and dimension table into the corresponding topic. After the four-layer view is built, the data package is sent to the development server for financial and tax data analysis and visualization. The Report Studio reporting component is utilized to be able to build smart reports for chart federation. Different reports are linked to each other to enhance the flexibility of reports. Users are able to analyze the data from different dimensions, comprehensively discover the characteristics and connections of the data, and realize the intelligent visualization of the data. Once the fundamental financial and tax data housed within the financial sector's business application system have been aggregated and synthesized, the study proceeds to construct a processed data set in the form of a TS and devise a departmental budget prediction algorithm based on this TS. This approach allows for the observation of the year-by-year change rule governing the departmental budget data. According to the above TS model to predict, the specific prediction process is shown in Figure 6.

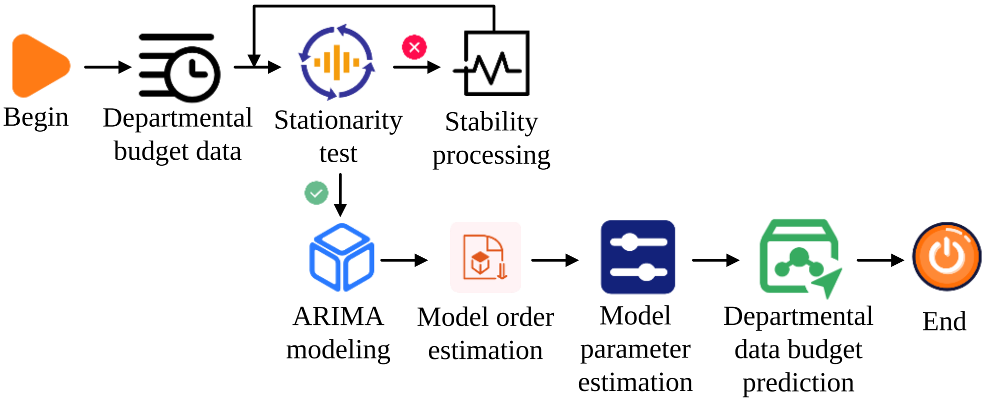


Figure 6: Method for predicting budget data in the financial department based on ARIMA

Stability tests are performed on the departmental budget data. If it is a smooth series, then go directly to ARIMA modeling, otherwise perform difference operations until the data is a smooth series. After the data is stable, ARIMA modeling is carried out and the ACF and PACF are used to determine model approximate ranges of orders p and q. The AIC criterion is then used to determine the specific orders p and q. The model difference orders are determined by the difference operation used in the smoothing process. Once all the parameters of the ARIMA model are obtained, predicts can be made for the departmental budget data.

3 Results

One of the important elements of financial work is the departmental budget indicator. In order to verify the effectiveness of the financial and tax data management and departmental budget prediction method that integrates ARIMA model, the study selects the total number of departmental budget of a province from 2015 to 2022 as the data test set. Through the data stability test and stable process, the departmental budget in 2023 is quantitatively predicted to provide data support for assisting the leadership in decision-making and implementing the departmental budget work in a scientific and effective way.

3.1 Experiments and analysis of the times series model

To verify the effectiveness of the financial and tax data management and departmental budget prediction method that integrates ARIMA model, the study adopts the error rate (Error) and precision rate (Precision) as the experimental evaluation indexes. The total amount of departmental budget of a province from 2015 to 2022 is taken as the experimental dataset. The experimental parameters are configured as 750GB of ROM, 8GB of RAM, and an Inter Core i7-quad-core processor. The experimental dataset collects departmental budget data on a quarterly basis, specifically on January 1, April 1, July 1, and October 1 of each year (in case of holidays, it is automatically postponed by one day). The departmental budget data TS for a province from 2015 to 2022 is shown in Figure 7.

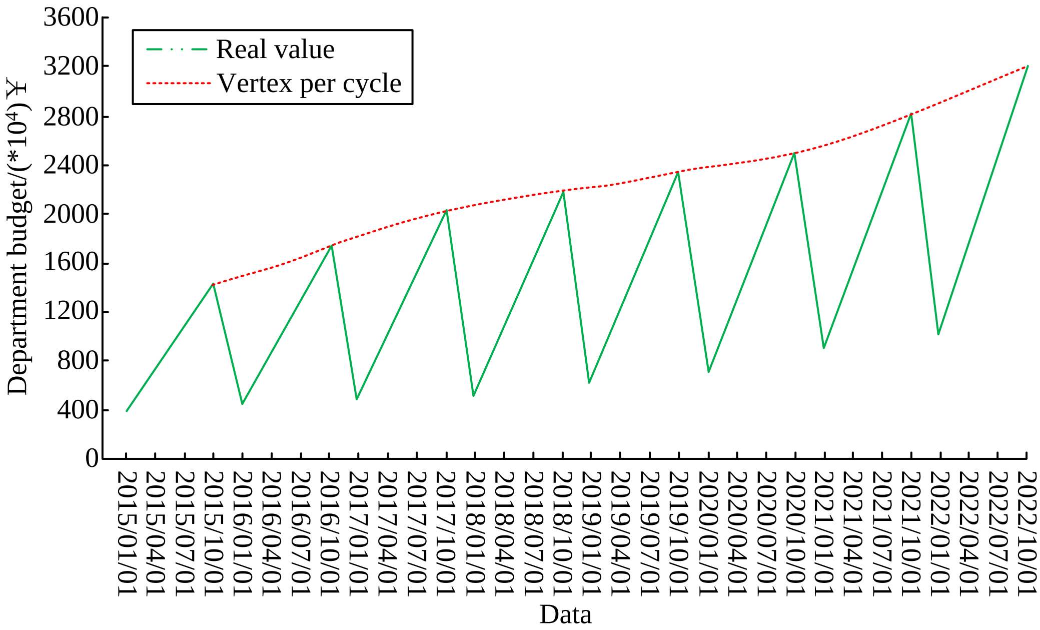
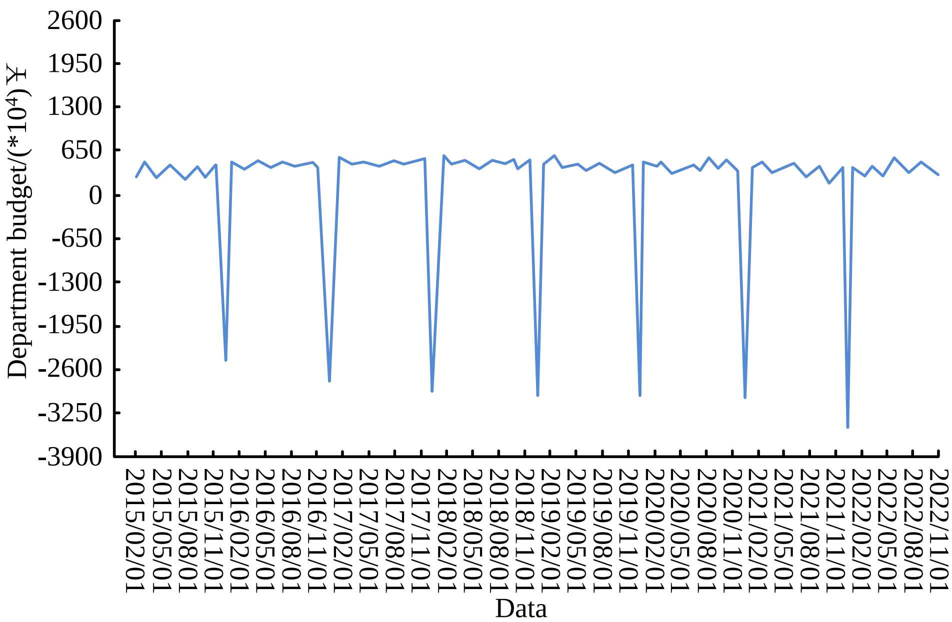
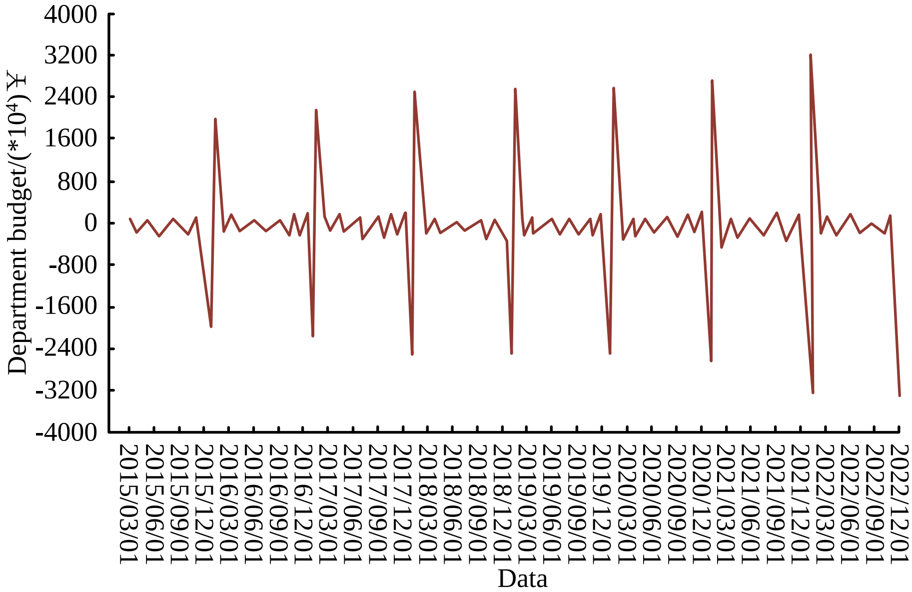


Figure 7: Time series diagram of department's original budget data

In Figure 7, the 2015-2022 departmental budget data is clearly cyclical and shows an upward trend at the top of each cycle. Therefore, the 2015-2022 historical departmental data is non-stationary and cannot be used directly for predicting analysis. It is necessary to convert the non-smooth historical departmental data into stationary series. The study uses difference operation to process the above non-smooth data. Since the above data has a non-linear trend, 2 difference operations are performed to smooth the data. The results are shown in Figure 8.



(a) First order differential sequence diagram



(b) Second order differential sequence diagram

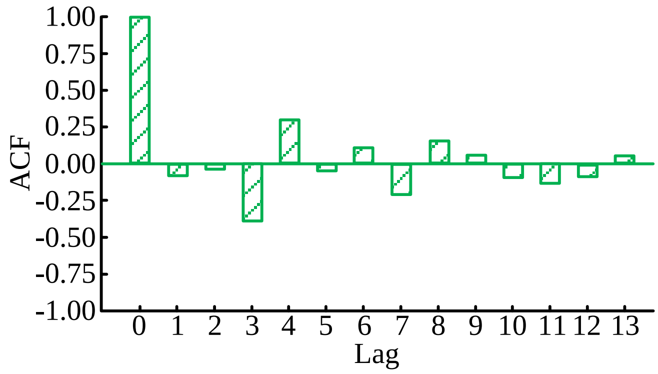
Figure 8: Method for stabilizing departmental budget data

In Figure 8(a), the first-order differential series did not fluctuate randomly up and down around the zero level, indicating that this budget data series did not reach a smooth state. To facilitate the progression of the series to a harmonious conclusion, it is necessary to continue differentiating the first-order differential series. In Figure 8(b), the second-order differential series is basically floating up and down regularly around the zero level line, with no obvious upward or downward trend characteristics, and the series is stable after the SOD. The 2015-2022 departmental budget data meets the requirements of the TS-based departmental budget prediction algorithm, so the historical departmental budget data can be used in the prediction algorithm to predict the departmental budget. To verify that the the SOD of the departmental budget data series has really reached a smooth state, the study continues to carry out the ADF test on the series. The results of ADF test are shown in Table 4.

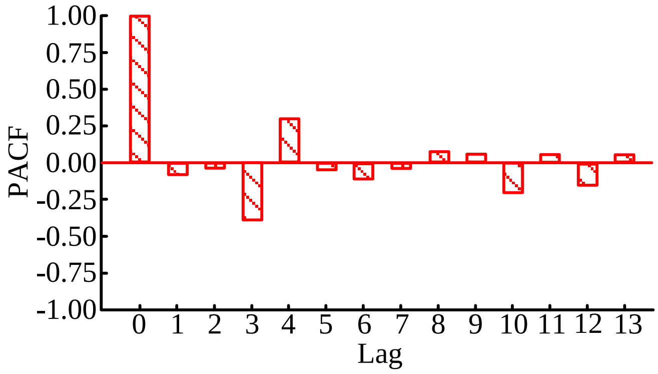
Table 4: ADF test result table

d	*t*	*P*	Critical value		
			1%	5%	10%
0	-0.812	0.913	-3.678	-2.863	-2.584
1	-2.139	0.229	-3.678	-2.863	-2.584
2	-4.325	0.000	-3.716	-2.871	-2.592

In Table 4, the t-value of the ADF test of the departmental budget data series without difference treatment is -0.812, $P=0.913>0.05$, which cannot reject the meta-hypothesis, indicating that the original TS is not stable. The t-value of the ADF test of the data after first-order difference for the departmental budget data series is -2.139, $P=0.229>0.05$, which cannot reject the meta-hypothesis, indicating that the series is not smooth. After the series is second-order differenced, the t-value of the ADF test of the data is -4.325, $P=0.000<0.05$, which rejects the original hypothesis, indicating stationary series. Therefore, it can be confirmed that the SOD of the departmental budget data series achieves stability. The departmental budget data series, subjected to a WN test using Python, returned a P-value of 0.000, which is less than 0.05. This result rejects the original hypothesis that the TS is a pure random series. Therefore the second-order differential series is not a WN series and is of research interest. The ARMA (p, q) model orders p and q of the stable non-WN series are initially determined by ACF and PACF. The ACF and PACF of the 2015-2022 departmental budget data series after SOD are shown in Figure 9.



(a) ACF of second-order differential sequence



(b) PACF of second-order differential sequence

Figure 9: ACF and PACF of second-order differential series

In Figure 9(a), the ACF fluctuation process around the zero-value deletion is slower, which is due to the ACF trailing off. In Figure 9(b), the PACF fluctuation process around the zero-value censoring is continuous, which is due to the PACF trailing off. After the PACF and ACF are roughly determined for q and p, the final model order is determined by the AIC criterion for different values of p and q at different values. Table 5 displays the AIC function values for the ARMA model.

Table 5: ARMA model AIC function values

ARMA	AIC	ARMA	AIC
ARMA (0, 0)	23976.23	ARMA(2, 0)	23970.45
ARMA(0, 1)	23974.54	ARMA(2, 1)	23954.32
ARMA(0, 2)	23972.33	ARMA(2, 2)	23932.58
ARMA(0, 3)	23982.62	ARMA(2, 3)	23924.74
ARMA(1, 0)	239.82.30	ARMA(3, 0)	23974.63
ARMA(1, 1)	23978.65	ARMA(3, 1)	23974.50
ARMA(1, 2)	23974.21	ARMA(3, 2)	23920.56
ARMA(1, 3)	23985.75	ARMA(3, 3)	23926.35

In Table 5, the AIC function value of ARMA (3, 2) is minimized when P=3 and q=2. The model fits the data better the smaller the data, according to the AIC criterion. Therefore ARIMA (3, 2, 2) model is the best fitting model for departmental budget data. The residual WN test is employed to ascertain whether the pertinent information in the departmental budget data series has been fully extracted. The resulting value of P=0.827 is found to be greater than 0.05, thereby suggesting that the TS is a WN series. This finding implies that departmental budget prediction can benefit from the application of the ARIMA (3, 2, 2) model.

3.2 Experimentation and analysis of departmental budget prediction method based on ARIMA modeling

To verify the feasibility of the prediction method of financial departmental budget data based on ARIMA (3, 2, 2) model, the study takes the historical data of departmental budget from 2015 to 2019 as the data test set. The effect of departmental budget prediction from 2020 to

2022 is evaluated by error rate and precision rate indicators. The annual budget, monthly budget, and daily budget of the financial department are predicted and analyzed. Error rate=|True value (TV)-Predicted value (PV)|/TV. Precision = 1-Error. The results of departmental budget prediction for the years 2020 to 2022 are shown in Figure 10.

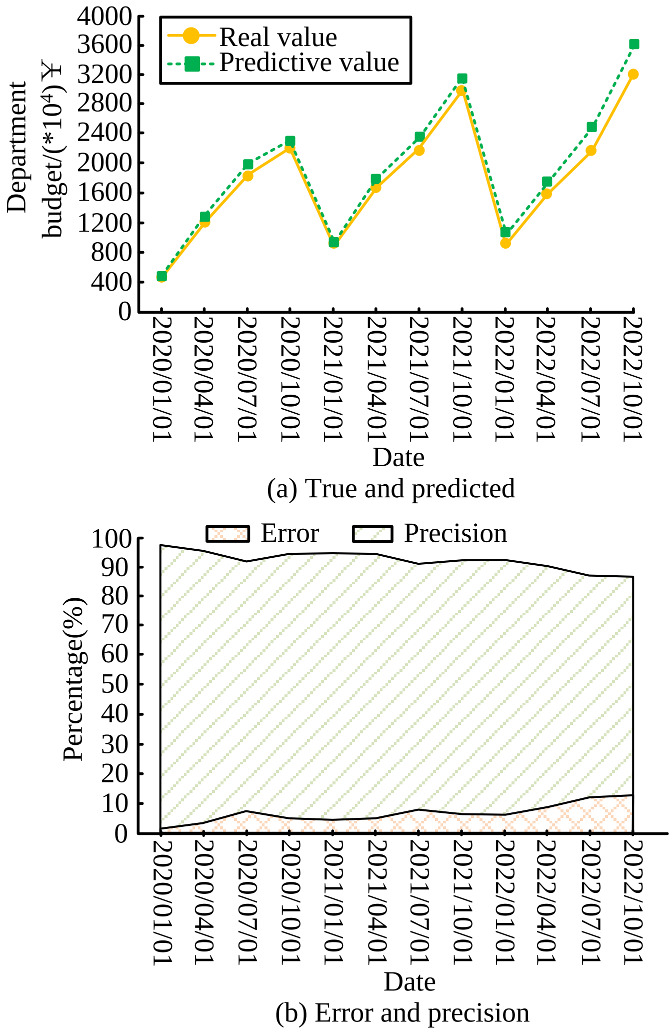
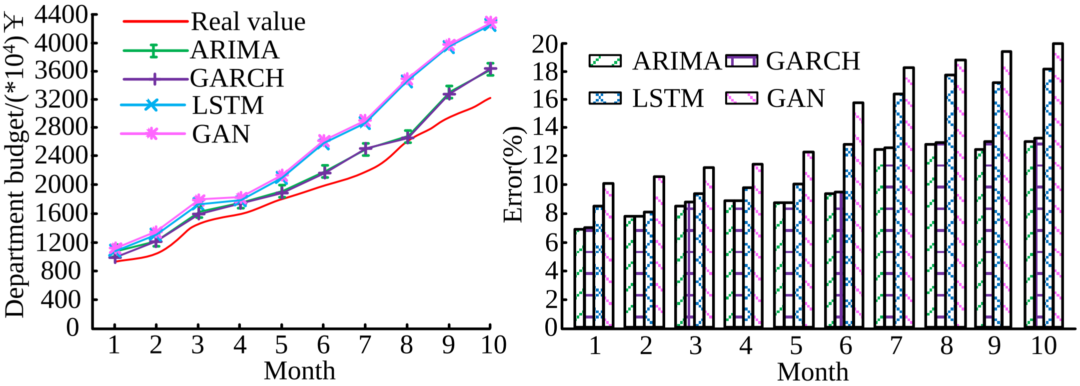


Figure 10: Departmental budget results

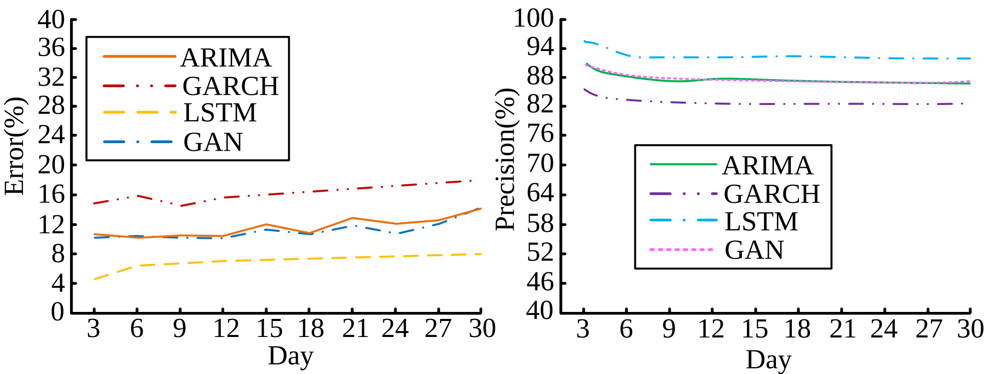
In Figure 10(a), the annual projected values of departmental budget are 182.6 million yuan, 247.25 million yuan, and 266.7 million yuan for the years 2020 through 2022, respectively. The departmental budget annual TV are 17.01 million yuan, 230.25 million yuan, and 237.9 million yuan, respectively. The error between the TV and PV for the years 2020 through 2022 is 7.3%, 7.4%, and 12.1%, respectively. The trend of the curve change between the true value and the PV is consistent. In Figure 10(b), the prediction error displays an upward trend as the year increases. This is due to the high precision of ARIMA (3, 2, 2) model in short-term predicting. When the financial departmental budget is affected by major emergencies, the model prediction precision decreases. Combining Figures 10(a)-(b), the prediction accuracies of the departmental budget data from 2020 to 2022 are all above 85%, indicating that the ARIMA (3, 2, 2) model predicts better. To further test the effectiveness of ARIMA(3, 2, 2) model for fiscal departmental budget data prediction method, the study selects generalized autoregressive conditional heteroskedasticity (GARCH), long short term memory (LSTM) and generative adversarial networks (GAN) with ARIMA (3, 2, 2) model for departmental budget prediction. Based on the state of time, the methods are categorized into static and dynamic prediction. The static prediction interval is January October 2022. The results of the static prediction of fiscal departmental budget are shown in Figure 11.



(a) Static prediction results (b) Different algorithm errors

Figure 11: Static prediction results using different algorithms

In Figure 11(a), the average monthly budget of the financial sector from January to October 2022 is 19.825 million yuan. In the static prediction, the ARIMA (3, 2, 2) model, the GARCH model, the LSTM model, and the GAN model have monthly PV of 2,225,000-yuan, 22,000,000-yuan, 25,396,000 yuan, and 25,561,000 yuan, respectively. The errors between the PV and the TV obtained from the ARIMA (3, 2, 2) model, GARCH model, LSTM model and GAN model are 12.1%, 10.9%, 28.1% and 28.9%, respectively. The ARIMA (3, 2, 2) model reduces the prediction error by 56.9% and 58.1% than the LSTM and GAN models and improves the prediction error by about 10% than the GARCH model. In Figure 11(b), the errors of ARIMA (3, 2, 2) model and GARCH model are basically the same. Overall, the RIMA (3, 2, 2) model has a smaller prediction error. The GAN model has the largest prediction error and the worst prediction effect, which is due to the impact of the amount of data in the GAN model, resulting in a decrease in its prediction precision. The dynamic prediction interval is from October 1 to 30, 2022. The results of the dynamic prediction of the financial departmental budget are shown in Figure 12.



(a) Different algorithm errors (b) Different algorithm precisions

Figure 12: Dynamic prediction results using different algorithms

In Figure 12(a), the average prediction errors obtained from the ARIMA(3, 2, 2) model, GARCH model, LSTM model, and GAN model are 11.3%, 15.4%, 6.3%, and 10.1%, respectively, in the dynamic prediction. The LSTM model has the smallest prediction error. The ARIMA (3, 2, 2) model and the GAN model show a small difference in the prediction error. This is due to the fact that the LSTM model and GAN model are deep learning prediction models with strong learning fitting ability and better prediction. The GARCH model has the largest prediction error and the worst prediction effect. In Figure 12(b), the average prediction precision obtained from ARIMA (3, 2, 2) model, GARCH model, LSTM model and GAN model are 88.7%, 84.6%, 93.7% and 89.9%, respectively. The GARCH model has the worst prediction results. The prediction precision of ARIMA (3, 2, 2) model, LSTM model and GAN model are above 85%, which indicates that the prediction result of ARIMA (3, 2, 2) model is good and the prediction is valid. Therefore, according to the predicted data, the total number of departmental budget is judged. Moreover, the budget is allocated to the priority sub-sections of each department under the premise of not exceeding the budget, which provides powerful data support to assist the leadership in decision-making.

Finally, the sensitivity analysis results of the ARIMA model are shown in Table 6. There are significant differences in the sensitivity of different factors to the model prediction performance. The sensitivity of autoregressive order and differential integration order is as high as 83.24% and 81.06%. The sensitivity of moving average order is slightly lower, only 65.16%. Accordingly, the results of the sensitivity analysis enable the study to select the optimal parameter combinations for enhancing the predictive performance of the model.

Table 6: Sensitivity analysis of ARIMA model

Dimension	Influence factor	Management elements	Absolute sensitivity value
Parameters	Autoregressive order	ARIMA	6.16%
	Differential integration order	ARIMA	89.16%
	Moving average order	ARIMA	79.58%

4 Discussion

Under the background of big data and information age, economic and financial work is facing unprecedented opportunities and challenges. The conventional budget forecasting techniques employed by finance and taxation departments predominantly rely on the incremental budget method, zero-based budget method, fixed budget method, and so forth. These techniques are primarily based on the budget implementation of the previous year, coupled with adjustments made to align with changes in relevant factors, with the aim of preparing a budget that accurately reflects the anticipated comprehensive balance of economic activities. However, the effectiveness of these techniques is often limited, particularly in terms of their accuracy and the extent to which they facilitate effective budget management. In recent years, the development of computer information technology has accelerated the progress of reform of financial wisdom management and wisdom services, including data processing platforms such as distributed storage and cloud computing technology. At the same time, advanced data analysis algorithms such as deep learning and time series analysis have been gradually applied in the field of economy and finance. However, the traditional time series analysis still has the shortcomings of higher complexity and lower accuracy in the process of analyzing the trends and patterns of financial data. Meanwhile, in recent years, various research scholars have begun to introduce advanced time series analysis algorithms into various fields. Literature [12] Mughal S N et al. introduced time series based neural network for solar radiation forecasting with root mean square error of 1.28. Literature [13] Garai S et al. utilized a hybrid and its learning algorithms based on adaptive noise-complete set empirical modal decomposition to achieve the prediction of the market price. Literature [14] Li H et al. based on the dynamic time distortion expansion method of time weights analysis realized the prediction of financial stock trading. In this regard, the study innovatively selected ARIMA model to construct a departmental budget forecasting model. The experimental results showed that in static forecasting the prediction errors of ARIMA (3, 2, 2) model, GARCH model, LSTM model, and GAN model were 12.1%, 10.9%, 28.1%, and 28.9%, respectively. Moreover, the maximum reduction of prediction error was 58.1%. In dynamic prediction, the average prediction errors of ARIMA(3, 2, 2) model, GARCH model, LSTM model, and GAN model were 11.3%, 15.4%, 6.3%, and 10.1%, respectively. The average prediction accuracy were 88.7%, 84.6%, 93.7%, and 89.9%, respectively. The prediction of GARCH model was the worst, and the prediction of ARIMA (3, 2, 2) model, LSTM model and GAN model all have prediction accuracy above 85%. The ARIMA model was effective in static forecasting, as it was capable of capturing the linear relationship of time series data with greater precision. However, its predictive accuracy was slightly inferior to that of the GARCH model. This was attributed to the ARIMA model's comparatively limited capacity to cope with volatile aggregations in the data. Whereas the LSTM and GAN models had slightly worse prediction performance in static prediction due to insufficient data features. The slightly lower prediction accuracy of ARIMA model in dynamic forecasting compared to LSTM model was due to better sensitivity of LSTM model in dynamic environment.

The study introduces advanced time series analysis algorithms to the field of finance, and the ARIMA model can more accurately capture the complex patterns and trends in fiscal data, thus improving the accuracy of budget forecasting. Meanwhile, the ARIMA model is able to continuously update and adjust the forecasting results with the addition of new data. Moreover, it can take into account the impact of multiple factors on fiscal data, thus realizing an automated and intelligent forecasting process. The research advances the comprehensive advancement of the domain of financial data forecasting and the implementation and incorporation of financial data forecasting technology across a broader range of industries. This can serve as a crucial foundation for the formulation of fiscal policies and economic plans.

5 Conclusion

The amount of economic and social data has increased significantly as a result of computer science advancements. The data may be evaluated and extracted in an efficient manner through the utilization of data mining and warehouse technologies. The field of finance is intertwined with all aspects of economic activity. The mining and analysis of financial data can therefore be an effective means of enhancing the decision-making abilities of those in positions of leadership, as well as facilitating the delivery of efficient and intelligent management and services. By analyzing financial and tax data and predicting departmental budget data in depth, the study proposed a financial and tax data management and departmental budget prediction methodology incorporating the ARIMA model. The method enabled quantitative prediction of future data for the financial sector. The method first analyzed and evaluated the order of the ARMA model. For the non-smooth financial and tax data, the ARIMA model was obtained by adding difference processing to the ARMA model. The non-stationary data were also converted into stationary data and then analyzed and processed. Then the data warehouse was constructed to integrate and summarize the financial and tax data, which provides powerful data support to assist the leadership in decision-making. The outcomes indicated that the departmental budget data of a province from 2015 to 2022 was non-stationary series. It was transformed into stationary series by SOD processing, and the order P=3 and q=2 were determined by using ACF, PACF, and AIC criterion to obtain ARIMA (3, 2, 2) model to predict the departmental budget. The error rates for 2021 and 2022 were essentially within 15%, and the precision rates were both above 85%. In static prediction, the RIMA (3, 2, 2) model had the smallest prediction error. In dynamic prediction, ARIMA (3, 2, 2) model was good with a prediction precision of 88.7%. Accordingly, the research method is an effective means of predicting departmental budget data, with a low error rate, and provides robust data support for leadership decision-making. The study of natural disasters and other factors affecting the financial departmental budget is not yet comprehensive. In the future, it will be expanded to include natural disasters and other factors affecting the economy.

Fundings

The research is supported by Project of the 2024 Teaching Reform Research on Vocational Education and Adult Education in Jilin Province: Research on the Practice of Teaching Reform of the Course "Tax Calculation and Declaration" Enabled by Digital Technology in Higher Vocational Colleges Project Approval Number: 2024ZCZ016.

References

[1] D. Bergemann, A. Bonatti, and T Gan, "The economics of social data," The RAND Journal of Economics, vol. 53, no. 2, pp. 263-296, 2022. https://doi.org/10.1111/1756-2171.12407

[2] J. Sheng, J. Amankwah-Amoah, Z. Khan, and X Wang, "COVID-19 pandemic in the new era of big data analytics: methodological innovations and future research directions," British Journal of Management, vol. 32, no. 4, pp. 1164-1183, 2021. https://doi.org/10.1111/1467-8551.12441

[3] S. Duleba, F. K. Gündoğdu, and S. Moslem, "Interval-valued spherical fuzzy analytic hierarchy process method to evaluate public transportation development," Informatica, vol. 32, no. 4, pp. 661-686, 2021. https://doi.org/10.15388/21-INFOR451

[4] M. Cipriani, L. S. Goldberg, and G. L. Spada, "Financial sanctions, SWIFT, and the architecture of the international payment system," Journal of Economic Perspectives, vol. 37, no. 1, pp. 31-52, 2023. https://doi.org/10.1257/jep.37.1.31

[5] Y. Han and T. Wang, "Semi-supervised clustering for financial risk analysis," Neural Processing Letters, vol. 53, no. 5, pp. 3561-3572, 2021. https://doi.org/10.1007/s11063-021-10564-0

[6] G. Mehdi, H. Hooman, Y. Liu, S. Peyman, R. Arif, J. Ahmad, R. Hamidreza, and A. Rashid, "Data mining techniques for web mining: a survey" Artificial Intelligence and Applications, vol. 1, no. 1, pp. 3-10, 2022. https://doi.org/10.47852/bonviewAIA2202290

[7] K. Bhosle and V. Musande, "Evaluation of deep learning CNN model for recognition of devanagari digit," Artificial Intelligence and Applications, vol. 1, no. 2, pp. 114-118, 2023. https://doi.org/10.47852/bonviewAIA3202441

[8] U. Hewage, R. Sinha, and M. A. Naeem, "Privacy-preserving data (stream) mining techniques and their impact on data mining precision: a systematic literature review," Artificial Intelligence Review, vol. 56, no. 9, pp. 10427-10464, 2023. https://doi.org/10.1007/s10462-023-10425-3

[9] M. Douiba, S. Benkirane, A. Guezzaz, and M. Azrour, "An improved anomaly detection model for IoT security using decision tree and gradient boosting," The Journal of Supercomputing, vol. 79, no. 3, pp. 3392-3411, 2023. https://doi.org/10.1007/s11227-022-04783-y

[10] W. Lu, J. Li, J. Wang, amd L. Qin, "A CNN-BiLSTM-AM method for stock price prediction," Neural Computing and Applications, vol. 33, no. 10, pp. 4741-4753, 2021. https://doi.org/10.1007/s00521-020-05532-z

[11] K. Peng, W. Jiang, P. Hou, Z. Wu, Z. Ling, X. Wang, Z. Niu, and D. Mao, "Continental-scale wetland mapping: a novel algorithm for detailed wetland types classification based on time series Sentinel-1/2 images," Ecological Indicators, vol. 148, no. 4, pp. 110113-110128, 2023. https://doi.org/10.1016/j.ecolind.2023.110113

[12] S. N. Mughal, Y. R. Sood, and R. K. Jarial, "A neural network-based time-series model for predicting global solar radiations," IETE Journal of Research, vol. 69, no. 6, pp. 3418-3430, 2023. https://doi.org/10.1080/03772063.2021.1934576

[13] S. Garai, R. K. Paul, M. Yeasin, and A. K. Paul, "CEEMDAN-based hybrid machine learning models for time series forecasting using MARS algorithm and PSO-optimization," Neural Processing Letters,

vol. 56, no. 2, pp. 92-115, 2024. https://doi.org/10.1007/s11063-024-11552-w

[14] H. Li, "Time works well: dynamic time warping based on time weighting for time series data mining," Information Sciences, vol. 547, no. 2, pp. 592-608, 2021. https://doi.org/10.1016/j.ins.2020.08.089

[15] C. W. Tan, C. Bergmeir, F. Petitjean, and G. I. Webb, "Time series extrinsic regression: predicting numeric values from time series data," Data Mining and Knowledge Discovery, vol. 35, no. 3, pp. 1032-1060, 2021. https://doi.org/10.1007/s10618-021-00745-9

[16] A. Shahid, T. A. N. Nguyen, and M. T. Kechadi, "Big data warehouse for healthcare-sensitive data applications," Sensors, vol. 21, no. 7, pp. 2353-2380, 2021. https://doi.org/10.3390/s21072353

[17] D. Amo, P. Gómez, L. Hernández-Ibáñez, and D. Fonseca, "Educational warehouse: modular, private and secure cloudable architecture system for educational data storage, analysis and access," Applied Sciences, vol. 11, no. 2, pp. 806-823, 2021. https://doi.org/10.3390/app11020806

[18] M. G. Khan, N. U. Huda, and U. K. U. Zaman, "Smart warehouse management system: architecture, real-time implementation and prototype design," Machines, vol. 10, no. 2, pp. 150-170, 2022. https://doi.org/10.3390/machines10020150

[19] M. G. Kahn, J. Y. Mui, M. J. Ames, A. K. Yamsani, N. Pozdeyev, N. Rafaels, and I. M. Brooks, "Migrating a research data warehouse to a public cloud: challenges and opportunities," Journal of the American Medical Informatics Association, vol. 29, no. 4, pp. 592-600, 2022. https://doi.org/10.1093/jamia/ocab278

[20] A. Tufano, R. Accorsi, and R. Manzini, "A machine learning approach for predictive warehouse design," The International Journal of Advanced Manufacturing Technology, vol. 119, no. 3, pp. 2369-2392, 2022. https://doi.org/10.1007/s00170-021-08035-w.

[21] R. P. Masini, M. C. Medeiros, and E. F. Mendes, "Machine learning advances for time series forecasting," Journal of Economic Surveys, vol. 37,no. 1, pp. 76-111, 2023. https://doi.org/10.1111/joes.12429

[22] T. Wen, H. Chen, and K. H. Cheong, "Visibility graph for time series prediction and image classification: a review," Nonlinear Dynamics, vol. 110, no. 4, pp. 2979-2999, 2022. https://doi.org/10.1016/j.jobe.2021.102537

[23] Y. Gao, K. M. Mosalam, Y. Chen, W. Wang, and Y. Chen, "Auto-regressive integrated moving-average machine learning for damage identification of steel frames," Applied Sciences, vol. 11, no. 13, pp. 6084-6110, 2021. https://doi.org/10.3390/app11136084

[24] L. You and M. Zhu, "Digital twin simulation for deep learning framework for predicting solar energy market load in Trade-By-Trade data," Solar Energy, vol. 250, no. 1, pp. 388-397, 2023. https://doi.org/10.1016/j.solener.2023.01.004

[25] W. Grzesiak, D. Zaborski, I. Szatkowska, and K. Królaczyk, "Lactation milk yield prediction in primiparous cows on a farm using the seasonal auto-regressive integrated moving average model, nonlinear autoregressive exogenous artificial neural networks and Wood's model," Animal Bioscience, vol. 34, no. 4, pp. 770-782, 2021. https://doi.org/10.5713/ajas.19.0939

[26] P. M. Almeida-Junior and A. D. C. Nascimento, "ARMA process for speckled data," Journal of Statistical Computation and Simulation, vol. 91, no. 15, pp. 3125-3153, 2021. https://doi.org/10.1080/00949655.2021.1922688

[27] B. Ma, P. Li, X. Guo, H. Zhao, and Y. Chen, "A novel online prediction method for vehicle velocity and road gradient based on a flexible-structure auto-regressive integrated moving average model," Sustainability, vol. 15, no. 21, pp. 15639-15656, 2023. https://doi.org/10.3390/su152115639

[28] B. Rele, C. Hogan, S. Kandanaarachchi, and C. Leigh, "Short-term prediction of stream turbidity using surrogate data and a meta-model approach: A case study," Hydrological Processes, vol. 37, no. 4, pp. 14857-14869, 2023. https://doi.org/10.1002/hyp.14857

[29] B. Yang, X. Liu, W. Long, and L. Peng, "A unified unit root test regardless of intercept," Econometric Reviews, vol. 42, no. 6, pp. 540-555, 2023. https://doi.org/10.1080/07474938.2023.2217077

[30] S. Erdogan and S. A. Solarin, "Stochastic convergence in carbon emissions based on a new fourier-based wavelet unit root test," Environmental Science and Pollution Research, vol. 28, no. 17, pp. 21887-21899, 2021. https://doi.org/10.1007/s11356-020-12033-y

[31] K. Yamazaki, V. K. Vo-Ho, D. Bulsara, and N. Le. "Spiking neural networks and their applications: a review," Brain Sciences, vol. 12, no. 7, pp. 863-892, 2022. https://doi.org/10.3390/brainsci12070863

[32] K. M. Hamdia, X. Zhuang, and T. Rabczuk, "An efficient optimization approach for designing machine learning models based on genetic algorithm," Neural Computing and Applications, vol. 33, no. 6, pp. 1923-1933, 2021. https://doi.org/10.1007/s00521-020-05035-x

[33] J. F. Torres, D. Hadjout, A. Sebaa, F. Martínez-Álvarez, and A. Troncoso, "Deep learning for time series forecasting: a survey," Big Data, vol. 9, no. 1, pp. 3-21, 2021. https://doi.org/10.1089/big.2020.0159