International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

The Prediction of GDP by Aggregate Accounting Information. A Neural Network Model

Mu Sun[1*], Cristina Del Campo[2], Julián Chamizo-González[3], Elena Urquía-Grande[1]
[1] Faculty of Economics & Business, Department of Financial Administration and Accounting, Complutense University of Madrid, Madrid, 28223, Madrid, Spain
[2] Faculty of Economics & Business, Department of Financial and Actuarial Economics and Statistics, Complutense University of Madrid, Madrid, 28223, Madrid, Spain
[3] Faculty of Economics & Business, Department of Accounting, Autonomous University of Madrid , Madrid, 28049, Madrid, Spain
corresponding author*，Email：musun@ucm.es

Abstract

Currently, accounting information plays a key role in the economic world and bridges the gap between macroeconomics and microeconomics. The existing literature corroborates that aggregate-level accounting earnings data encapsulates information regarding GDP growth. Nonetheless, accounting data derived from GDP components have not been given due consideration. Consequently, this research proposes an aggregate-level model grounded in the four components of the GDP income-based method, intending to assess the predictive power of aggregate-level accounting data concerning GDP. Furthermore, this paper scrutinizes the forecasting performance of a neural network model built upon the conventional linear regression framework. Finally, a comparison of the outcomes derived from both models is conducted. The findings reveal that both the present value model and the value-added model corroborate the notion that accounting information derived from the four components of the income-based GDP accounting framework encapsulates data pertinent to future GDP. Furthermore, the model demonstrates heightened sensitivity towards the performance of the subsequent second quarter's GDP. Among the variables, Depreciation and Income exhibit the most substantial impact, while Salaries exhibit the least impact. Concurrently, this research noted an improvement in the fitting performance of the neural network model as compared to that of the traditional linear model.

Keywords: Aggregate accounting information, GDP, GDP forecast model, Backpropagation neural network.

1. INTRODUCTION

The GDP, as the core indicator of national economic accounting, is an indispensable metric for assessing a country's economic health. It serves as a crucial barometer for gauging a nation's wealth and the living standards of its citizens. Additionally, it offers valuable insights for policymakers to make informed decisions in response to evolving economic conditions (Henderson et al., 2012). In line with the United Nations' (UNDESA, 2022) guidelines, the computation of GDP should be rooted in economic theory and employ a comprehensive array of disciplines to evaluate a country's economic performance (or region) during a specific timeframe. The release of GDP data is typically coordinated by the statistical department of the respective country or region on a quarterly basis.

The estimation of GDP is crucial in the field of economics. Existing forecasting methodologies for GDP can be roughly categorized into either predicting and regressing GDP indicators via a linear regression forecasting model (Baffigi et al., 2004), or predicting GDP via non-linear models. The neural network model is a prime example of artificial intelligence (Jena et al., 2021; Teräsvirta et al., 2005; Tkacz, 2001).

Thus, researchers have been concentrating on novel approaches to enhance the precision of GDP projections. In recent times, the potential of micro-level accounting information to elucidate macroeconomic phenomena has garnered attention, initially sparked by observations that aggregate-level earnings can predict future market returns (Bailey & Lai, 2020; Gallo et al., 2016; Kothari et al., 2006; Patatoukas, 2014; Sun et al., 2022). Concurrently, multiple studies have probed the relationship between aggregate earnings and inflation (Cready & Gurun, 2010; Patatoukas, 2014; Shivakumar & Urcan, 2017). Existing research suggests that aggregate-level accounting earnings information contains valuable insights for GDP growth prediction, thereby facilitating more accurate future GDP projections (Abdalla & Carabias, 2022; Konchitchki & Patatoukas, 2014a, 2014b; Lalwani & Chakraborty, 2020; Sun et al., 2022).

Traditional econometric techniques impose stringent assumptions on time series data. When the data is non-stationary, it often requires adjustment to achieve stationarity. For instance, quarterly time series data may exhibit intercepts and trends, which can render traditional methods less suitable for forecasting. However, artificial neural network (ANN) models, which have gained prominence in recent years, can effectively avoid the limitations of linear regression (Feng & Zhang, 2014). ANN models are now extensively employed across various research domains, stemming from their robust parallel processing capabilities, fault tolerance, and high nonlinearity. Principally, they have been utilized for the prediction of macroeconomic indicators. Notable studies such as Panda and Narasimhan (2007) employed ANN to forecast the US dollar exchange rate, while Dunis et al. (2011) utilized a recurrent neural network to predict exchange rates for the euro and dollar. Nakamura (2005) assessed the efficacy of neural networks in forecasting inflation. Smalter Hall and Cook (2017) corroborated that neural network models can augment the accuracy of macroeconomic indicator forecasts. Numerous other existing studies have corroborated the efficacy of ANN models for macroeconomic forecasting indicators.

The rapid and widespread dissemination of COVID-19 has exerted an indeterminate influence on every aspect of the global economy. As of 2022, the pandemic continues to surge worldwide in successive waves. Furthermore, the escalating complexity of the situation in Ukraine in 2022 is anticipated to further impede the progress of global economic recuperation. According to projections by the International Monetary Fund (IMF) [1], global economic growth is projected to decelerate from 6.1% in 2021 to 3.6% in 2022 and 2023. Ambiguity is inherent in a fluctuating macroeconomic environment; however, the complexity of macroeconomic forecasting is amplified under the current circumstances.

Refining the relationship between company-level accounting information and the macroeconomy is crucial, as company-level activities form an integral part of the macro-economy. Examining the predictive power of future GDP embedded within aggregate accounting information can enhance GDP forecasting accuracy. Although most existing research is centered on earnings-level accounting information, there is a dearth of studies examining other aggregate-level accounting information and its relationship with GDP (Sun et al., 2022). Consequently, this research introduces a novel research concept. Drawing on the existing literature concerning aggregate-level accounting information and GDP forecasting, and considering the time series data employed in this research, it proposes the application of the backpropagation (BP) neural network model in addition to the traditional linear regression model, to test the predictive capacity of aggregate-level information on future GDP. Furthermore, this research compares the predictive performance of the traditional econometric model and the BP neural network model for GDP, thereby exploring the future GDP change information contained at the aggregate level from a fresh perspective.

This research makes the following contributions: Firstly, it innovatively constructs a GDP prediction model based on the income-based GDP accounting method providing a new approach for predicting GDP using overall accounting information. Secondly, this study enriches existing research and fills the gap in this topic. Lastly, by combining neural network algorithm models, this study offers feasibility for interdisciplinary research in this field.

2. LITERATURE REVIEW

2.1 Aggregate-level Accounting Information Usefulness

Accounting information usefulness theory is based on the usefulness of accounting information for relevant decisions. The usefulness of accounting information at the aggregate-level is primarily reflected in the predictive value of aggregate earnings (AE) due to their extensive use in forecasting market returns. Pioneering studies, such as Kothari et al. (2006), investigated the correlation between AE and factors such as inflation or risk premium, which can impact future market returns. These authors aimed to explore the relationship between aggregate earnings and future market returns by examining these two metrics. Cready and Gurun (2010) highlighted a negative correlation between aggregate earnings news and market returns, regardless of the announcement period, indicating that the market does not necessarily appreciate the information conveyed by aggregate earnings. Patatoukas (2014) posited that aggregate earnings are intertwined with all components of stock market returns, and his findings revealed a positive correlation between the cash flow component in aggregate earnings and the discount rate.

Several studies have delved into the correlation between aggregate earnings and inflation, seeking to establish an indirect impact on market returns. Shivakumar and Urcan (2014) corroborated the notion that the expansion of aggregate earnings is prognostic of future inflation, a relationship that can be further elucidated by examining the interconnection between aggregate earnings and market returns. Gallo et al. (2016) uncovered a positive correlation between aggregate earnings and unexpected fluctuations in the U.S. funds target rate, thereby indicating a parallelism between aggregate earnings and risk premium.

[1]The International Monetary Fund:
https://www.imf.org/en/Publications/WEO

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

Recent research by Bailey and Lai (2020) employed predictive regression modeling of aggregate earnings, employing principal component analysis. Their findings revealed a negative correlation between aggregate earnings and expected market returns prior to 2000, which reversed upon an increase in volatility post-2000. Kim et al. (2020) delved into the reasons behind the variation of aggregate earnings growth with market returns, concluding that aggregate earnings exhibit a significant positive correlation with the growth of the U.S. Industrial Production Index. Moreover, they observed that the ability of aggregate earnings in the U.S. to explain market returns had improved in recent years, attributed to more accurate fair value estimates.

Previous research also examined the correlation between aggregate earnings and GDP to facilitate forecasting of GDP. Seminal studies, such as those conducted by Konchitchki and Patatoukas (2014a, 2014b), have tested the relationship between aggregate earnings and GDP. Their findings corroborate the predictive value of aggregate earnings for future GDP and highlight the potential of these earnings to improve the accuracy of GDP forecasts. Lalwani and Chakraborty (2020) found that professional forecasters demonstrated a poor response to aggregate earnings information based on past research involving data from various economies. They also discovered that changes in negative aggregate earnings conveyed more GDP-related information than did changes in positive aggregate earnings. Similarly, Gaertner et al. (2020) posited that accruals in aggregate earnings might better predict GDP, but their study failed to provide evidence supporting this hypothesis. They also found that only negative changes in aggregate earnings could predict future GDP growth, while positive changes did not exhibit such predictive power.

Finally, few studies have investigated the relationship between other aggregate-level accounting information and GDP. Abdalla and Carabias (2022) not only highlighted the information about future GDP embedded in aggregate earnings but also examined the GDP-related information content contained in special items at the aggregate level. Their research revealed that special items at the aggregate level could better capture changes in GDP.

2.2 Development of Artificial Neural Network and BP Neural Network Model

Reflecting the growing interest in artificial intelligence (AI) methods since the turn of the century, various research fields have witnessed an increasing adoption of artificial neural network (ANN) models, multi-case reasoning systems, and support vector machines. Among these, ANN, a representative of AI approaches, has been extensively utilized across diverse research domains (Wu & Feng, 2018).

The early ANN was initially referred to as the "Perceptron", a model introduced by (Rosenblatt, 1960). Building upon this foundation, Widrow and Hoff (1962) developed the Adaptive Linear Neuron or later Adaptive Linear Element (ADALINE) network model and employed the ANN for the first time in the investigation of practical issues. Before long, the ANN model experienced significant growth, including the subsequent development of discrete neural networks (Hopfield, 1982), Multi-layer Network Learning Algorithm models, Boltzmann machine models (Ackley et al., 1985), Error Backpropagation (BP) algorithms (Rumelhart et al., 1985), and Cellular Neural Network (CNN) models (Chua & Yang, 1988).

The BP algorithm (Rumelhart et al., 1985) possesses a significantly enhanced learning capability due to its error backpropagation mechanism and has become one of the most prominent neural network models. The architecture of a typical three-layer BP neural network is illustrated in Figure 1. This network model features no interconnections within the feedback layer.

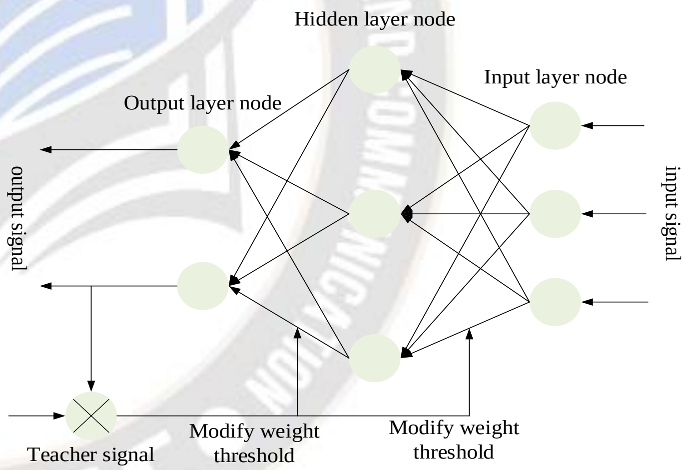


Figure 1. The topological structure of a three-layer BP neural network

The application of the BP neural network has been extensively explored in various research studies. For instance, in the financial sector, Mai et al. (2019) employed the BP neural network for bankruptcy prediction. Alameer et al. (2019) leveraged the BP neural network to forecast the future gold price. In the accounting domain, Geng et al. (2015) incorporated the BP neural network to predict the financial health of Chinese listed companies. Furthermore, in the business realm, Iftikhar and Khan (2020) employed the BP neural network to investigate social media requirements based on big data.

2.3 GDP Forecast and Artificial Neural Network Model

Forecasting GDP employing existing methodologies can be generally categorized into two types. The first involves employing a linear regression model to predict the GDP indicator. This approach benefits from extensive research and development by scholars, possessing a solid theoretical foundation (Baffigi et al., 2004). However, this method often necessitates stringent assumptions, such as determining parameters and adjusting research models, which may impact the quality of the forecast outcomes.

The second approach employs a non-linear model to estimate GDP, with the most prominent example utilizing Artificial Neural Networks (ANN). Previous research, such as Tkacz (2001), employed ANN models to enhance GDP forecasting precision and found that neural network models exhibited superior forecasting accuracy in his sample. Teräsvirta et al. (2005) employed both traditional econometric models and ANN models to compare and predict monthly macroeconomic variables in the G7 economies. Their findings indicated that ANN possessed a distinct advantage in predicting long-term sample data, but not in the short term. Furthermore, Jena et al. (2021) focused on GDP forecasting during the COVID-19 pandemic. They utilized ANN models to construct a GDP predictor, anticipating a significant decline in GDP for the sample countries from April to June 2020.

In parallel, Loermann and Maas (2019) employed an ANN model to project the quarterly GDP of the United States. Their feed-forward neural network approach exhibited forecast precision on par with that of professional forecasters. Jahn (2020) examined the predictive performance of ANN models in dealing with non-linear structured panel data, highlighting that the ANN model's advantage lied in its lack of temporal trend restrictions, thereby enabling more accurate GDP predictions.

More recent studies have demonstrated the predictive power of machine learning algorithms in economic modeling. For instance, Richardson et al. (2021) utilized an ANN model to estimate New Zealand's annual GDP, highlighting the significant enhancement that these algorithms can bring to traditional econometric methods. Building upon this, Longo et al. (2022) proposed a joint machine-learning approach, which combined a recurrent neural network model, a dynamic factor model, and a generalized autoregressive score. Their ensemble model demonstrated robust predictive capabilities, even for depressive data in the context of the COVID-19 pandemic.

Powered by their versatile applications, ANN models are extensively employed to forecast macroeconomic indicators, including exchange rates and inflation. Notably, Haider and Hanif (2009)) concentrated on ANN-driven predictions of macro indicators such as inflation, GDP, and circulation. By leveraging Pakistan's monthly macro data, they established that the ANN model outperforms traditional econometric models in terms of accuracy. Choudhary and Haider (2012) further examined monthly inflation indicators for 28 OECD countries. Their findings corroborated the superiority of the ANN model, which boasted a 45% accuracy rate compared to the 23% accuracy rate of traditional econometric models.

Recent studies, such as that conducted by Chuku et al. (2019), have compared the predictive capabilities of traditional econometric models and Artificial Neural Network (ANN) models for commodity prices, trade, inflation, and interest rates. Their findings corroborate the superiority of the ANN model over traditional regression models, with the latter being approximately 150 basis points more accurate. In recent years, macroeconomics has witnessed the emergence of new research frontiers, such as energy economics. Building upon this, Sharma et al. (2021) focused on natural gas within the realm of energy economics and employed ANN models to capture and forecast fluctuations in international natural gas prices. Their research indicates that an ANN model integrating multiple modules demonstrates a strong predictive capacity for global natural gas prices.

Based on the previous review of existing research, the present research identifies the following research questions:

RQ1: Can other aggregate-level accounting information based on the GDP income method model predict China's GDP?

RQ2: Can the neural network model optimize the GDP forecasting model proposed in this research?

3. METHODOLOGY, INSTRUMENTS, AND SAMPLE

3.1 Sample

The present research aims to examine the predictive capacity of aggregate-level data derived from GDP composition based on income for GDP growth. It utilizes the quarterly GDP data released by the National Bureau of Statistics of China (NBS) and the quarterly reports of all Chinese A-share listed companies to achieve this objective. The release dates of the listed company data are consistent, and the data records are comprehensive. However, obtaining financial statement data of Chinese small and medium-sized enterprises (SMEs) remains challenging as it involves a variety of factors, including legal policies implemented by SMEs, introverted corporate culture, immature accounting technology and market competition (Li et al., 2017).

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

It should be highlighted that from 2007 to 2008, a significant reform of China's accounting system occurred with the implementation of the “New Accounting Standards for Business Enterprises” (MOF China, 2006). The reform marks China's adoption of standards closer to international accounting standards to improve the quality and transparency of financial reporting. There are significant differences between the old accounting standards and the new CAS, so financial statements for different periods have significantly different accounting treatments and disclosures. As a result, rendering pre-2008 data non-comparable. Consequently, the sample period for this research spans from 2009 Q1 to 2022 Q2, consisting of 54 quarters of comprehensive financial reports.

The present research procured quarterly financial statement data from China's Stock Market and Accounting Research Database (CSMAR) [2]. CSMAR is amongst the most prominent financial and economic databases in China, boasting high data integrity. To avoid data redundancy, only the consolidated statements of the parent company were retained. Furthermore, the cumulative data and quarterly GDP in all earnings reports were necessitated to be adjusted to the current added value.

3.2 Variables and Methodology

The present research is based on the UNDESA (1993) system of national accounts and the definition of the GDP income-based method by the "China 2019 Statistical Yearbook" (NBS, 2022). Four components make up the GDP: compensation of employees, net taxes on production, depreciation of fixed assets, and operating surplus, as shown in Model (1).

$$GDP = compensation of employees + net taxes on production + depreciation of assets + operating surplus \quad (1)$$

These four components can be identified with the four financial accounts in the financial statements. The “compensation of employees” is approximately equal to the current added value of the “employee salaries payable” item on the balance sheet plus the current value of "cash paid to employees and for employees" in the cash flow statement. The "net taxes on production" is roughly equivalent to the "operating tax surcharge" account on the income statement. The “depreciation of fixed asset” directly finds the same account in the cash flow statement, and the "operating surplus" is approximately equal to the “income from operations” account. Hence, the baseline model of this research is shown in Model (2).

$$GDP_q = \alpha + \beta_1 salaries_q + \beta_2 tax_q + \beta_3 depreciation_q + \beta_4 income_q + \varepsilon \ (2)$$

Therefore, this research uses the changed values of the four variables in the q quarter to predict the GDP growth in the $q+1$ quarter ΔGDP_{q+1}.

$$\Delta GDP_{q+1} = \alpha + \beta_1 \Delta salaries_q + \beta_2 \Delta tax_q + \beta_3 \Delta depreciation_q + \beta_4 \Delta income_q + \varepsilon_{q+k} \ (3)$$

All variables are at the aggregate level, including accounting information in the quarterly reports of all Chinese A-share listed companies. Table 1 shows this research's variable definitions, interpretations, and sources.

Abbreviation	Variable	Definition Explanation	Source
gdp	Nominal GDP	China's quarterly nominal GDP, unit: yuan	NBS China
ln_gdp_sa	GDP seasonally adjusted	Eviews X-12 seasonally adjusted original GDP and take its natural logarithm	Calculated in this paper
ln_gdp_sa_c	GDP growth	ln_gdp_sa changes between the current quarter and the previous quarter	Calculated in this paper
salaries	Compensation of employees	The present value of the item "employee compensation payable" plus the present value of the cash flow statement "cash paid to and for employees" from the quarterly financial reports of Chinese A-share companies. unit: yuan	CSMAR
ln_ salaries_sa	Compensation of	Eviews X-12	Calculated

[2] CSMAR Database: https://cn.gtadata.com/

	employees seasonally adjusted	Seasonally adjusted original salaries values and take their natural logarithm	in this paper
ln_ salaries_sa_c	Compensation of employees growth	ln_ salaries _sa changes between the current quarter and the previous quarter	Calculated in this paper
tax	Net taxes on production	"Operating tax surcharge" on the income statement of Chinese A-share companies. unit: yuan	CSMAR
ln_ tax _sa	Net taxes on production seasonally adjusted	Eviews X-12 Seasonally adjusted original tax values and take their natural logarithm	Calculated in this paper
ln_ tax _sa_c	Net taxes on production growth	ln_ tax _sa changes between the current quarter and the previous quarter	Calculated in this paper
depreciation	Depreciation of fixed assets	"Depreciation of fixed asset" on the cash flow statement of Chinese A-share companies. unit: yuan	CSMAR
ln_ depreciation _sa	Depreciation of fixed assets seasonally adjusted	Eviews X-12 Seasonally adjusted original depreciation values and take their natural logarithm	Calculated in this paper
ln_ depreciation _sa_c	Depreciation of fixed assets growth	ln_ depreciation _sa changes	Calculated in this paper
		between the current quarter and the previous quarter	
income	Operating surplus	"Income from operations" on the income statement of Chinese A-share companies. unit: yuan	CSMAR
ln_ income_sa	Operating surplus seasonally adjusted	Eviews X-12 Seasonally adjusted original income values and take their natural logarithm	Calculated in this paper
ln_ income_sa_c	Operating surplus growth	ln_ income _sa changes between the current quarter and the previous quarter	Calculated in this paper

Table 1. Variable definitions

In addition, for the method of processing aggregated accounting information, the general calculation method, such as the market value weighting method(Konchitchki & Patatoukas, 2014a; Kothari et al., 2006), is based on the company's profit rate and the weight of its market value in the total market value, as shown in formula 4:

$$Aggregateearnings_{i,t} = \sum \frac{P_{i,t}}{R_{i,t}} \times W_{i,t} \quad (4)$$

where P is the company's net profit, R is the total revenue, E is the profit margin, and W is the market value weight. However, this research found in previous tests that weighting the four aggregate accounting information variables with market value as the weight does not increase the model's explanatory performance for GDP. Therefore, it uses direct arithmetic aggregate values of the original data.

It must be pointed out that, according to the explanation of “China Accounting Standards for Business Enterprises No. 4: Fixed Assets (MOF China, 2006)”, the depreciation of fixed assets is only disclosed twice a year, namely semi-annual report, and annual report. The present research adjusts them to four quarters by weighting the corresponding quarterly fixed assets.

In addition, the present research also tests whether the accuracy of GDP forecasts could be optimized using the BP

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

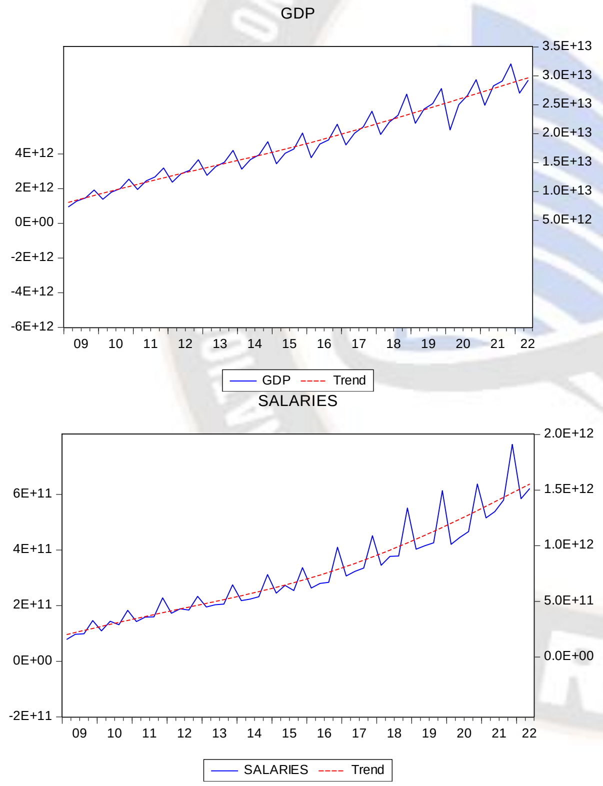
neural network model. It considers pre-training methods to provide objective and accurate initial parameters for subsequent model training. The BP neural network model established in this research corresponds to this paper's traditional linear regression model (3). The input layer has four neurons corresponding to the four independent variables, one hidden layer, and one neuron in the output layer corresponding to the dependent variable ΔGDP_{q+1}.

The linear time series regression part of this research was carried out in Eviews10, and the BP neural network was implemented in MATLAB (version R2021a).

4. RESULTS AND DISCUSSION

4.1 Linear Regression Analysis

First, this research used the Hodrick-Prescott filter method to analyze trends in the raw data. Setting the value of Lambda to the default 1600 yields the result shown in Figure 2.



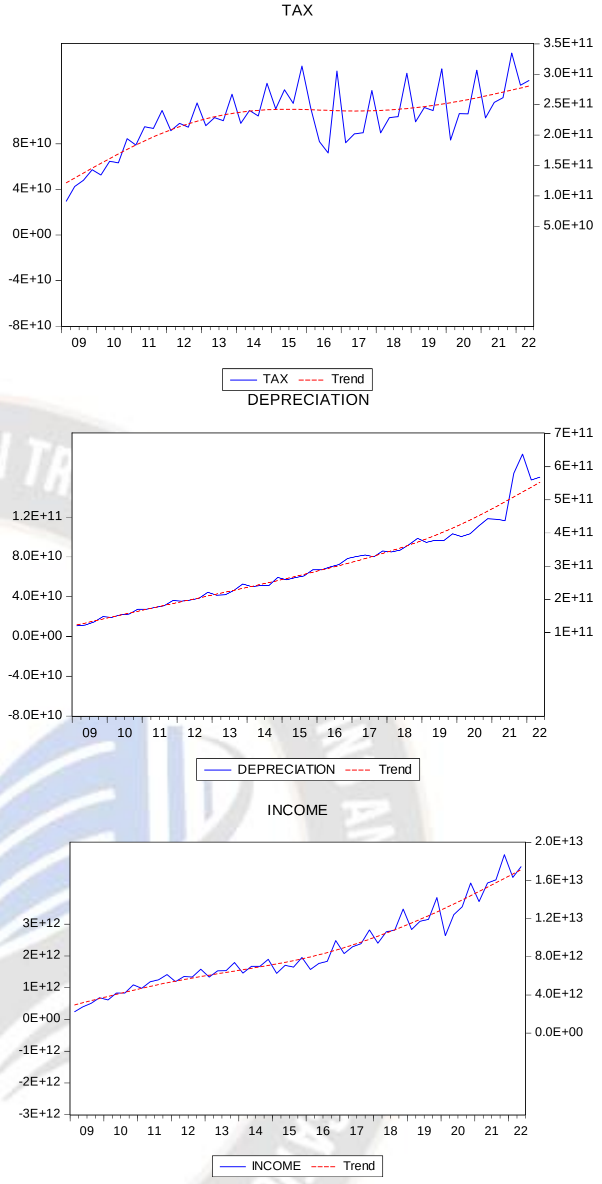


Figure 2. Variables Trend

As illustrated in Figure 2, it can be observed that, in general, all variables exhibit a consistent upward trend, with notable seasonal fluctuations. Both GDP and salaries demonstrate a similar steady escalation in volatility. The tax curve, however, displays relatively significant seasonal fluctuations from 2015 to 2022, with the upward trend of the curve decelerating. This outcome is attributed to the tax reduction policy implemented by the Chinese government in 2015 (Chinadaily, 2015). The most consistent performance is demonstrated by the depreciation of fixed assets. Excluding the abrupt short-term fluctuations in 2021-2022, the other periods are comparatively stable. It should be noted that, following the outbreak of the COVID-19 pandemic in 2020, the global economy was severely affected, resulting in a corresponding downward trend in China's real economy and fixed asset depreciation curve. However, China's epidemic control measures in early 2021 enabled the nation's

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

economy to recover rapidly (World Bank, 2021). Consequently, from 2021 to 2022, the quantity of fixed assets held by Chinese enterprises escalated rapidly, leading to a corresponding rapid increase in depreciation expenses. The income curve, generally, exhibits a continuous upward trend, with the negative impact of the COVID-19 pandemic in 2020 on the income curve being evident.

For time series regression models, it is crucial to ensure the stationarity of all variables included in the analysis. Consequently, based on the findings presented in Figure 2, all independent and dependent variables series exhibit intercepts and trends and need adjustments for stationarity. Therefore, EViews X12 seasonal adjustment module was employed to smooth these variables. Additionally, to mitigate the impact of data scale, we applied natural logarithm processing to all seasonally adjusted data. The results of the ADF unit root test for time series variables across all models are presented in Table 2.

Variable	Exogenous	t-Statistic	Prob.
LN_GDP_SA	Constant, Linear Trend	-2.821715	0.0068
LN_GDP_SA_C	Constant, Linear Trend	-9.250467	0.0000
LN_SALARIES_SA	Constant, Linear Trend	-5.62684	0.0001
LN_SALARIES_SA_C	Constant, Linear Trend	-9.404729	0.0000
LN_TAX_SA	Constant, Linear Trend	-3.709841	0.0302
LN_SALARIES_SA_C	Constant, Linear Trend	-9.404729	0.0000
LN_DEPRECIATION_SA	Constant, Linear Trend	-2.874467	0.0059
LN_DEPRECIATION_SA_C	Constant, Linear Trend	-7.253022	0.0000
LN_INCOME_SA	Constant, Linear Trend	-4.433257	0.0044
LN_INCOME_SA_C	Constant, Linear Trend	-6.54086	0.0000

Table 2. Augmented Dickey-Fuller test statistic

The null hypothesis is rejected for all sequences, suggesting that the sequence variables can pass the ADF test after seasonal adjustment and natural logarithm processing, indicating their stability. Model (2) was employed via linear regression analysis. The outcomes of the three tests are included in Table 3. These tests investigate the impacts of the four independent variables on GDP for the current, subsequent, and quarter after that, respectively. To consider the robustness of the test, a residual autocorrelation testing (Breusch-Godfrey Serial Correlation LM Test) and heteroskedasticity testing (Heteroskedasticity Test: White) were conducted on the three regression test results.

	Test 1	Test 2	Test 3
Variable	Dependent Variable: LN_GDP_SA Sample: 2009Q1 2022Q2	Dependent Variable: LN_GDP_SA(1) Sample (adjusted): 2009Q1 2022Q1	Dependent Variable: LN_GDP_SA(2) Sample (adjusted): 2009Q1 2021Q4
C	12.77534**	12.85886**	13.84662**
LN_SALARIES_SA	0.597649**	0.532601**	0.544352**
LN_DEPRECIATION_SA	0.090722	0.186811**	0.172795**
LN_TAX_SA	0.021501	0.041759	0.076936**
LN_INCOME_SA	0.012937	0.022684	0.02251*
R-squared	0.903255	0.902443	0.912965
Adjusted R-squared	0.902704	0.901813	0.912366
Heteroskedasticity Test: White	Prob. F(7,46): 0.0130	Prob. F(7,46): 0.0751	Prob. F(7,46): 0.1038
Breusch-Godfrey Serial Correlation LM Test:	Prob. F(2,47): 0.0113	Prob. F(2,47): 0.1275	Prob. F(2,47): 0.1417

**: Prob. <0.05, *: Prob. <0.1

Table 3. Model (2) linear regression test results

Overall, the adjusted R-squared values for the three tests are all above 0.9, with the third test demonstrating the highest degree of fit, reaching an adjusted R-squared of 0.9123 (see Table 3). These results indicate that the time series model regression fit is rather satisfactory. For test 1, only salaries with a p-value less than 0.05 reject the null hypothesis that it can significantly positively affect the current GDP. These robustness test results of Test 1 reject the null hypothesis, indicating residual autocorrelation and heteroscedasticity in the regression.

The results of Test 2 demonstrate that both salaries and depreciation can have a significant positive effect on GDP in the subsequent quarter, with all its robustness tests accepting the null hypothesis, suggesting no residual autocorrelation and heteroscedasticity in the model.

The findings of Test 3 reveal that all four independent variables, except for income, can have a significant positive impact on the future second-quarter GDP. However, the p-value of income only exhibits a positive impact on the subsequent second-quarter GDP within the confidence interval of 0.1. All robustness tests of Test 3 accept the null hypothesis, indicating the absence of residual autocorrelation and heteroscedasticity in the model.

In summary, considering the results of the three tests, Test 3 contains more statistically significant variables, making it more suitable for Model (2) proposed in this research. This demonstrates that accounting data at the aggregate level possesses an excellent ability to explain the GDP of the subsequent second quarter in the future.

Now the results of the linear regression analysis based on model (3) are presented. Analogous to the findings in Table 2, Table 4 also comprises the results of the three tests. As per model (3), these tests scrutinize the impact of the change values of the four independent variables on the GDP change in the present quarter, the subsequent quarter, and the second quarter thereafter.

	Test 1	Test 2	Test 3
	Dependent Variable: LN_GDP_SA_C	Dependent Variable: LN_GDP_SA_C(1)	Dependent Variable: LN_GDP_SA_C(2)
Variable	Sample: 2009Q1 2022Q2	Sample (adjusted): 2009Q1 2022Q1	Sample (adjusted): 2009Q1 2021Q4
C	0.07630	0.01328	-0.01893
LN_SALARIES_SA_C	0.161217	0.071872	0.129020*
LN_DEPRECIATION_SA_C	0.203101*	0.120911	0.593401**
LN_TAX_SA_C	0.976223**	0.190110	0.264475**
LN_INCOME_SA_C	0.058301**	0.247478	0.433401**
R-squared	0.054457	0.096674	0.156670
Adjusted R-squared	0.049271	0.075182	0.083367
Heteroskedasticity Test: White	Prob. F(7,46): 0.0000	Prob. F(7,46): 0.9617	Prob. F(7,46): 0.3232
Breusch-Godfrey Serial Correlation LM Test:	Prob. F(2,47): 0.0000	Prob. F(2,47): 0.1575	Prob. F(2,47): 0.2432

**: Prob. <0.05, *: Prob. <0.1

Table 4. Model (3) linear regression test results

In Table 4, overall, the adjusted R-squared value for the change in the four independent variables' effect on the GDP change in the second quarter next year is the highest (0.08), followed by 0.075 for the subsequent quarter and 0.049 for the current quarter. Only Tax and Income demonstrate a significant positive correlation in the test of current GDP changes, whereas no independent variable exhibits statistical significance in the test of future GDP changes. In the examination of GDP changes in the subsequent second quarter, all four independent variables, excluding Salaries, have a significant and considerable positive influence on the dependent variable, and the p-value of Salaries can also positively impact the dependent variable within the confidence interval of 0.1. Thus, model (3) exhibits greater sensitivity to changes in GDP in the subsequent second quarter, and its explanatory power for GDP changes is deemed acceptable.

Moreover, to assess the robustness of the analysis, this study additionally conducted the LM Test and the Heteroskedasticity Test on the three regression outcomes. The second and third tests passed these robustness checks, indicating that the model exhibits no residual autocorrelation or heteroscedasticity. Overall, the findings from models (2) and (3) corroborate that RQ1 of this research, namely, the accounting information derived from the four components of the income-based GDP method, contains predictive information about future GDP. Specifically, both research models demonstrate a positive impact of the four types of accounting information on GDP responsiveness in the subsequent quarter. The four accounting variables exhibit strong predictive power for forecasting future GDP values, as the changes in these accounting variables also possess explanatory power for future GDP growth fluctuations.

4.2 BP neural network regression analysis

In this subsection, a BP neural network model for the GDP growth part of the subsequent second quarter in the model (3) is constructed, aiming to investigate the RQ2 proposed in this research. The model used is shown in Equation 5.

$$\Delta GDP_{q+2} = \alpha + \beta_1 \Delta salaries_q + \beta_2 \Delta tax_q + \beta_3 \Delta depreciation_q + \beta_4 \Delta income_q + \varepsilon_{q+2} \quad (5)$$

According to Equation 5, the input layer of the BP neural network consists of four independent variable neurons, with a single dependent variable neuron in the output layer. In this research model, a preliminary set of 3, 6, 8, and 11 hidden layer nodes was established for testing purposes. Subsequently, these nodes with superior fitting performance were selected based on error analysis in various scenarios. MATLAB terminated the training process when the network training iterations and training result errors of different hidden layer nodes were sorted and are presented in Table 5.

Number of hidden layer nodes	Number of iterations	Mean square error (10^{-4})
3	5000	52.81
6	815	36.14
8	820	12.08
11	2281	22.19

Table 5. Comparison of nodes in different hidden layers

The optimal number of hidden layer nodes is determined to be eight, as it yields the smallest mean squared error (refer to Table 5). Concurrently, the activation function for the hidden layer is selected as the hyperbolic tangent function, while the activation function for the output layer is the identity function. In order to abbreviate the learning time and preclude the established backpropagation (BP) neural network from converging to an optimal local solution, this study employs three enhanced functions from the MATLAB toolbox, namely trainlm (Levenberg-Marquardt algorithm), trainrp (Resilient backpropagation algorithm), and trainbfg (BFGS quasi-Newton backpropagation algorithm) for comparative analysis. Trainlm generally performs well on moderately sized neural networks, with fast convergence, but may become slower for large networks. Trainrp is generally able to converge faster and is particularly suitable for training deep neural networks because it can overcome the vanishing and exploding gradient problems. Trainbfg generally performs well on small to medium sized neural networks. The functional properties, performance, and characteristics of these three functions are specified in Table 6.

ID	Function	Fit the Problem	Performance	Training Times (sec)	Mean Square Error
1	trainlm	Function fitting	Fast convergence and small error	1410	0.0136
2	trainrp	Simulation classification	Fastest	1726	0.0208
3	trainbfg	Function fitting	Fast convergence	2103	0.0171

Table 6. Comparison of three training functions

This research selects the optimal BP neural network training function according to the training results' mean square error and speed. According to the training results of the network training function, it can be found that compared with other training functions, the training speed of the trainlm function is the fastest, and the mean square error is smaller than 0.0136 (see Table 6). Therefore, this research finally chooses trainlm as the training function of the model (3).

Furthermore, according to the results in Table 7, it can be found that the independent variable importance results in the BP neural network model are similar to the linear regression results. Among them, the highest positive correlation between Depreciation and Income is higher than 0.4, the importance of Tax is generally 0.1970, and the lowest importance of Salaries is 0.0716. These results show that in the test of GDP growth in the subsequent second quarter, the growth of the Depreciation and the Income has the highest positive impact, Tax has a moderately positive impact, and Salaries has the smallest positive impact.

Feature	Importance
LN_SALARIES_SA_C	0.0716
LN_DEPRECIATION_SA_C	0.4310
LN_TAX_SA_C	0.1970
LN_INCOME_SA_C	0.4502

Table 7. Importance of features

In order to verify the RQ2 and test whether the BP neural network can optimize the prediction ability of the traditional linear regression model for GDP growth, the present research compares the regression residuals of the BP neural network model and the traditional linear model for the subsequent second-period GDP forecast.

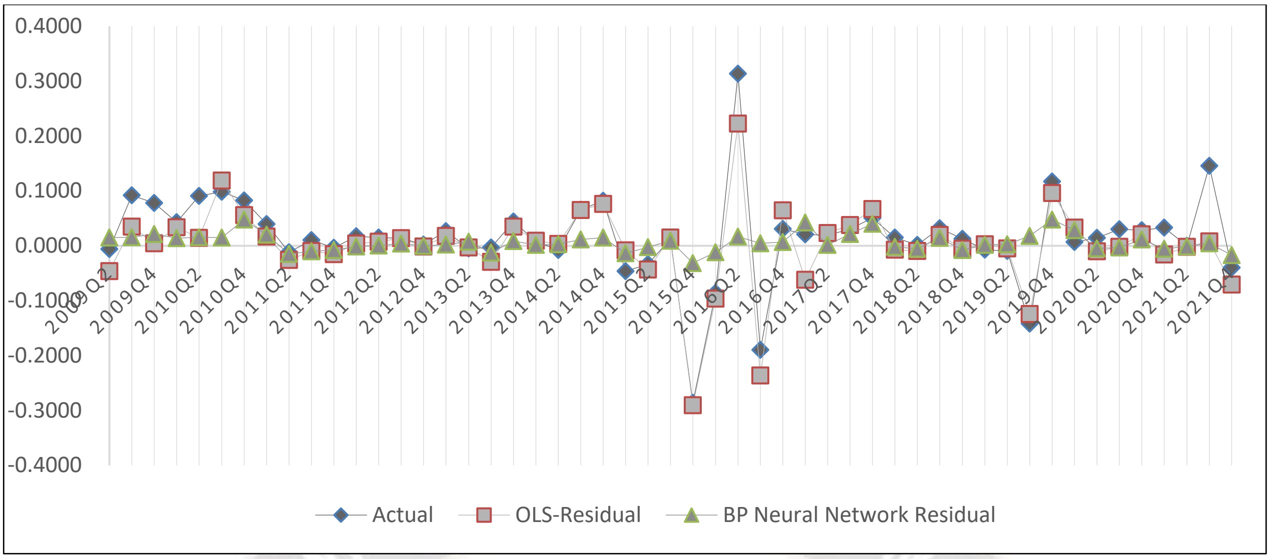


Figure 3. Comparison of forecast results

Figure 3 shows a graph of the residuals and actual values of the two models, showing the predictive effect of the two models. The closer the residual is to the horizontal axis, the better the fitting effect. It can be seen that the residual performance of the BP neural network model has remained stable near the horizontal axis, while the residual of the traditional linear regression has a significant fluctuation between 2015Q4-2016Q3. Although the considerable fluctuation of the actual GDP growth value in the 2015Q4-2016Q3 interval is the reason for the significant prediction error of the linear regression, the non-linear regression fitting results of the BP neural network remain stable. Hence, the relevant result of the BP neural network is better than the traditional linear regression model.

5. CONCLUSIONS

Scholars have predominantly concentrated on exploring the utility of accounting information. Given that scholars have corroborated that aggregate earnings encompass incremental information regarding GDP, research on the aggregate-level usefulness of accounting information offers a novel perspective on the topic. In recent years, the integration of artificial neural networks in various fields has also yielded novel concepts and enhancements to existing research.

The objective of this research was to investigate the incremental GDP information embedded in other accounting data at the aggregate level, despite the current research being limited. Consequently, it reviewed pertinent research on aggregate earnings and discovered that the existing studies predominantly focus on market returns and macroeconomic indicators. In recent years, research concerning the information content of GDP growth in aggregate earnings has gained prominence. This study also examined the development of the ANN model and its research involving GDP, concluding that, overall, the ANN model demonstrated superior performance compared to the traditional linear autocorrelation model in the direct prediction of GDP.

This research presented two research questions, derived from the review of the existing literature. A GDP influencing factor model, incorporating accounting information at four aggregate levels as independent variables, was developed based on the four components of the GDP income-based method. Results revealed that both the current value of the variable (Model 2) and the added value of the variable (Model 3) contained in the four types of accounting information based on the income-based GDP method's components provide information about future GDP. The model demonstrated greater sensitivity to the performance of GDP in the subsequent second quarter. Among the four aggregate-level variables, Depreciation and Income had the most significant impact, followed by Tax and Salaries. The findings indicate that the Depreciation of fixed assets and Income are the most critical components in China's income-based GDP composition. Additionally, by comparing the performance of the BP neural network model and the traditional linear model in forecasting the future second quarter GDP growth, this study found that the fitting effect of the BP neural network was superior to that of the traditional linear regression model. In addition, this study also confirms that COVID-19 causes variables to fluctuate abnormally, which increase the difficulty of macroeconomic forecasts.

Future features and developments for current research. Firstly, it solely addresses the accounting information pertinent to the four components of the income-based GDP method. Consequently, future research should endeavor to explore the incremental GDP information embedded in more aggregate-level accounting information. Secondly, the sample of this study is confined to listed companies. Acknowledging the significant contribution of non-listed companies to a nation's GDP, unlisted companies should also be incorporated into the research sample in order to further investigate the relationship between aggregate-level accounting information and economic development. Finally, the current study only used one neural network algorithm model for the analysis. It is hoped that the inclusion of more interdisciplinary methodologies can better the study of this topic in the future.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Data availability statement

The data supporting this research's findings are available from the corresponding author upon reasonable request.

Funding

The author(s) received no financial support for this article's research, authorship, and/or publication.

REFERENCES

[1] Abdalla, A. M., & Carabias, J. M. (2022). From accounting to economics: the role of aggregate special items in gauging the state of the economy. *The Accounting Review*, *97*(1), 1–27.

[2] Abiodun, O. I., Jantan, A., Omolara, A. E., Dada, K. V., Mohamed, N. A., & Arshad, H. (2018). State-of-the-art in artificial neural network applications: A survey. *Heliyon*, *4*(11), e00938.

[3] Ackley, D. H., Hinton, G. E., & Sejnowski, T. J. (1985). A learning algorithm for Boltzmann machines. *Cognitive Science*, *9*(1), 147–169.

[4] Alameer, Z., Abd Elaziz, M., Ewees, A. A., Ye, H., & Jianhua, Z. (2019). Forecasting gold price fluctuations using improved multilayer perceptron neural network and whale optimization algorithm. *Resources Policy*, *61*(C), 250–260.

[5] Baffigi, A., Golinelli, R., & Parigi, G. (2004). Bridge models to forecast the euro area GDP. *International Journal of Forecasting*, *20*(3), 447–460.

[6] Bailey, W., & Lai, H. (2020). On the Expected Earnings Hypothesis Explanation of the Aggregate Returns–Earnings Association Puzzle. *Journal of Financial and Quantitative Analysis*, *55*(8), 2732–2763.

[7] Chinadaily. (2015). *China's import tax cut to reduce revenue, aid consumption*. https://www.chinadaily.com.cn/business/2015-05/30/content_20863873.htm

[8] Choudhary, M. A., & Haider, A. (2012). Neural network models for inflation forecasting: an appraisal. *Applied Economics*, *44*(20), 2631–2635.

[9] Chua, L. O., & Yang, L. (1988). Cellular neural networks: Applications. *IEEE Transactions on Circuits and Systems*, *35*(10), 1273–1290.

[10] Chuku, C., Simpasa, A., & Oduor, J. (2019). Intelligent forecasting of economic growth for developing economies. *International Economics*, *159*(C), 74–93.

[11] Cready, W. M., & Gurun, U. G. (2010). Aggregate market reaction to earnings announcements. *Journal of Accounting Research*, *48*(2), 289–334.

[12] Dunis, C. L., Laws, J., & Sermpinis, G. (2011). Higher order and recurrent neural architectures for trading the EUR/USD exchange rate. *Quantitative Finance*, *11*(4), 615–629.

[13] Feng, L., & Zhang, J. (2014). Application of artificial neural networks in tendency forecasting of economic growth. *Economic Modelling*, *40*(C), 76–80.

[14] Gaertner, F. B., Kausar, A., & Steele, L. B. (2020). Negative accounting earnings and gross domestic product. *Review of Accounting Studies*, *25*(4), 1382–1409.

[15] Gallo, L. A., Hann, R. N., & Li, C. (2016). Aggregate earnings surprises, monetary policy, and stock returns. *Journal of Accounting and Economics*, *62*(1), 103–120.

[16] Geng, R., Bose, I., & Chen, X. (2015). Prediction of financial distress: An empirical study of listed Chinese companies using data mining. *European Journal of Operational Research*, *241*(1), 236–247.

[17] Haider, A., & Hanif, M. N. (2009). Inflation forecasting in Pakistan using artificial neural networks. *Pakistan Economic and Social Review*, *47*(1), 123–138.

[18] Henderson, J. V., Storeygard, A., & Weil, D. N. (2012). Measuring economic growth from outer space. *American Economic Review*, *102*(2), 994–1028.

[19] Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences*, *79*(8), 2554–2558.

[20] Iftikhar, R., & Khan, M. S. (2020). Social media big data analytics for demand forecasting: development and case implementation of an innovative framework. *Journal of Global Information Management*, *28*(1), 103–120.

[21] Jahn, M. (2020). Artificial neural network regression models in a panel setting: Predicting economic growth. *Economic Modelling*, *91*(C), 148–154.

[22] Jena, P. R., Majhi, R., Kalli, R., Managi, S., & Majhi, B. (2021). Impact of COVID-19 on GDP of major economies: Application of the artificial neural network forecaster. *Economic Analysis and Policy*, *69*(C), 324–339.

[23] Kim, J., Schonberger, B., Wasley, C., & Land, H. (2020). Intertemporal variation in the information content of aggregate earnings and its effect on the aggregate earnings-return relation. *Review of Accounting Studies*, *25*(4), 1410–1443.

[24] Konchitchki, Y., & Patatoukas, P. N. (2014a). Accounting earnings and gross domestic product. *Journal of Accounting and Economics*, *57*(1), 76–88.

[25] Konchitchki, Y., & Patatoukas, P. N. (2014b). Taking the pulse of the real economy using financial statement analysis: Implications for macro forecasting and stock valuation. *Accounting Review*, *89*(2), 669–694.

[26] Kothari, S. P., Lewellen, J., & Warner, J. B. (2006). Stock returns, aggregate earnings surprises, and behavioral finance. *Journal of Financial Economics*, *79*(3), 537–568.

[27] Lalwani, V., & Chakraborty, M. (2020). Aggregate earnings and gross domestic product: International evidence. *Applied Economics*, *52*(1), 68–84.

[28] Li, Y., Zhang, A., & Jing, L. (2017). SME Financial Report: Challenges from the New Normal in China. *2017 International Conference on Education Science and Economic Management (ICESEM 2017)*, 363–367.

[29] Loermann, J., & Maas, B. (2019). *Nowcasting US GDP with artificial neural networks*.

[30] Longo, L., Riccaboni, M., & Rungi, A. (2022). A neural network ensemble approach for GDP forecasting. *Journal of Economic Dynamics and Control*, *134*(C), 104278–104278.

[31] Mai, F., Tian, S., Lee, C., & Ma, L. (2019). Deep learning models for bankruptcy prediction using textual disclosures. *European Journal of Operational Research*, *274*(2), 743–758.

[32] MOF China. (2006). *China Accounting Standards for Business Enterprises No. 4: Fixed Assets*. https://www.gov.cn/gongbao/content/2002/content_61664.htm

[33] Nakamura, E. (2005). Inflation forecasting using a neural network. *Economics Letters*, *86*(3), 373–378.

[34] NBS. (2022). *China 2022 Statistical Yearbook*. http://www.stats.gov.cn/tjsj/ndsj/2022/indexeh.htm

[35] Panda, C., & Narasimhan, V. (2007). Forecasting exchange rate better with artificial neural network. *Journal of Policy Modeling*, *29*(2), 227–236.

[36] Patatoukas, P. N. (2014). Detecting news in aggregate accounting earnings: Implications for stock market valuation. *Review of Accounting Studies*, *19*(1), 134–160.

[37] Richardson, A., van Florenstein Mulder, T., & Vehbi, T. (2021). Nowcasting GDP using machine-learning algorithms: A real-time assessment. *International Journal of Forecasting*, *37*(2), 941–948.

[38] Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1985). Learning internal representations by error propagation. In *Parallel Distributed Processing* (Vol. 1). California Univ San Diego La Jolla Inst for Cognitive Science. https://apps.dtic.mil/sti/citations/ADA164453

[39] Sharma, V., Cali, Ü., Sardana, B., Kuzlu, M., Banga, D., & Pipattanasomporn, M. (2021). Data-driven short-term natural gas demand forecasting with machine learning techniques. *Journal of Petroleum Science and Engineering*, *206*(C), 108979–108979.

[40] Shivakumar, L., & Urcan, O. (2014). Why Does Aggregate Earnings Shocks Predict Future Inflation Shocks. In *Seminal paper submitted at Harvard Business School*.

[41] Shivakumar, L., & Urcan, O. (2017). Why does aggregate earnings growth reflect information about future inflation? *The Accounting Review*, *92*(6), 247–276.

[42] Smalter Hall, A., & Cook, T. R. (2017). Macroeconomic indicator forecasting with deep neural networks. *Federal Reserve Bank of Kansas City Working Paper*, 11–17.

[43] Sun, M., Urquía-Grande, E., Chamizo-González, J., & Del Campo, C. (2022). Accounting Earnings and Economic Growth, Trends, and Challenges: A Bibliometric Approach. *Computational Intelligence and Neuroscience*, *2022*(C), 7352160.

[44] Teräsvirta, T., Van Dijk, D., & Medeiros, M. C. (2005). Linear models, smooth transition autoregressions, and neural networks for forecasting macroeconomic time series: A re-examination. *International Journal of Forecasting*, *21*(4), 755–774.

[45] Tkacz, G. (2001). Neural network forecasting of Canadian GDP growth. *International Journal of Forecasting*, *17*(1), 57–69.

[46] UNDESA. (1993). *System of National Accounts*. https://unstats.un.org/unsd/nationalaccount/sna1993.asp

[47] UNDESA. (2022). *System of National Accounts*. https://unstats.un.org/unsd/nationalaccount/sna.asp

[48] Widrow, B., & Hoff, M. E. (1962). Associative storage and retrieval of digital information in networks of adaptive "neurons." In *Biological Prototypes and Synthetic Systems* (p. 160). Springer.

[49] World Bank. (2021). *China Economic Update, December 2021: Rebalancing Act - From Recovery to High-Quality Growth*. http://hdl.handle.net/10986/36828

[50] Wu, Y., & Feng, J. (2018). Development and application of artificial neural network. *Wireless Personal Communications*, *102*(2), 1645–1656.