ITM Web of Conferences **70**, 04034 (2025) https://doi.org/10.1051/itmconf/20257004034
DAI 2024

The Relationship Between Consumption Power and GDP Growth in China and Machine Learning Prediction Analysis

Runsong Hu

Rosedale Global High School, 450000, Zhengzhou, China

Abstract. With the rapid development of China's economy, consumption has become a key driving force to boost domestic demand and promote economic structural transformation. By analyzing the current situation of China's economic development, this paper focuses on the relationship between residents' consumption capacity and Gross Domestic Product (GDP) growth. The paper examines the use of machine learning for predicting consumption capacity by integrating various economic models, including Support Vector Regression (SVR) and Long Short-Term Memory Network (LSTM). Moreover, by analyzing data on China's GDP and consumption capacity from 1991 to 2020, the study demonstrates a strong positive correlation between the two. This finding further highlights the significant impact of consumption capacity on economic growth. The article also applies univariate and bivariate analyses to calculate the concentration trend and degree of dispersion of GDP and consumption capacity, and it uses regression models to explore the relationship between the two further. This article can provide policymakers with a deeper understanding of regional consumption differences and economic development trends, thus providing a decision-making basis for optimizing economic policies.

1 Introduction

In recent decades, China's economy has maintained steady growth, making significant contributions to national development and the improvement of people's livelihoods. However, with the constant changes in the domestic and international situation and the restructuring of the economy, the Chinese economy is also facing new challenges and opportunities. Against this background, it is especially crucial to analyse the current consumption capacity of Chinese residents. From the perspective of the law of economic development, China's economic construction is characterized by domestic demand-led and internal sustainability. Consumption is the final demand, an important link, and an important engine to promote the smooth operation of the domestic economic cycle, and it has a lasting pulling effect on the economy. It is important to understand the population's consumption capacity, stabilize the consumption base, and promote a sustained rebound in consumption. This approach aims to enhance the quality and upgrade consumption practices, ultimately

Corresponding author: tiancu739@163.com

contributing to the high-quality development of the country's economy and accelerating the creation of a new economic framework. This not only helps this paper to better grasp the pulse of economic construction but also provides an important reference for the future development path.

With the development of society, the techniques for Gross Domestic Product (GDP) forecasting have gradually matured and achieved fruitful results. The Midas model of Liu Han [1] shows high accuracy in short-term forecasting. The variation model of Li Jun et al [2] has a small error in regional GDP forecasting, while the CDFM proposed by Glocker [3] combines macro and micro factors to improve the forecasting effect. Other innovative approaches, such as the grey system analysis model and the multilayer ANN model, also demonstrate superior forecasting performance [4-5]. These studies show that the accuracy and effectiveness of GDP forecasting have been continuously improved through the optimization and innovation of different models.

In addition, Pradyot et al. [6] systematically collected GDP data for 2020 from several countries in Asia, America, and Europe to construct a multi-layer ANN model. The model was validated to have a prediction error of less than 2%, demonstrating good prediction accuracy. However, although the Autoregressive Integrated Moving Average (ARIMA) model performs well in stable economic environments, in the face of China's current economic uncertainties, such as trade wars and global economic fluctuations, the Long Short-Term Memory Network (LSTM) model has a greater potential to reveal nonlinear trends. For example, Xiao Qiangyan et al. [7] compared the performance of the LSTM model and the Bayesian vector autoregressive model. The results of the study showed that the LSTM model performed well in revealing the nonlinear relationship between the research indicators and GDP, significantly improving the prediction accuracy. Zhao Junhao et al. [8] Aiming at the challenges of statistical data lag and internal complex relationships in macroeconomic forecasting, an innovative forecasting method is proposed

Therefore, this paper aims to reveal the current status and trend of China's economic development through a comprehensive analysis of the economy and a discussion of consumption capacity, so as to provide deeper insights and guidance for future economic construction.

2 Status of economic development in China

China's economy has achieved rapid development over the past few decades, especially since its reform and opening up, with a consistently high GDP growth rate, making it the second-largest economy in the world. From 1978 to 2023, China's GDP grew from RMB 367.9 billion to over RMB 120 trillion, demonstrating strong economic vitality. With the rapid development of the economy, the importance of consumption power in economic growth has gradually come to the fore, becoming a key driver of domestic demand and economic structural transformation. Consumption has been the main engine of China's economic growth for many consecutive years, and its contribution to GDP has been rising.

GDP growth is directly related to the capacity for consumption. Economic theory explains that the consumption function illustrates how changes in income affect consumption expenditure. Additionally, the marginal propensity to consume measures the percentage of any increase in income that consumers decide to spend on consumption. As income levels rise, the capacity to consume increases further, thus fuelling further economic growth. In addition, the relationship between savings and investment cannot be ignored. While a high savings rate contributes to investment, it may also dampen consumption demand in the short term, and thus a reasonable guidance of the balance between consumption, savings, and investment is crucial for sustainable economic development.

In terms of structural factors, urban-rural differences, imbalances in regional development, upgrading of the industrial structure, and the income distribution gap all have a profound impact on the relationship between consumption capacity and GDP growth. The income gap between urban and rural residents limits the consumption potential of rural areas, while the imbalance in economic development between regions also creates regional differences in consumption capacity. In addition, with the transformation and upgrading of industrial structure, the phasing out of some low-value-added industries has had an impact on employment and income, which in turn affects residents' consumption capacity. Therefore, by accelerating urban-rural integration, promoting the development of central and western regions, and optimising income distribution and other policy measures, the overall consumption capacity can be effectively enhanced, contributing to the sustainable growth of GDP.

3 Application of machine learning in economic or consumption capacity analysis

With the development of big data technology, machine learning, as a powerful analysis tool, is being widely used in the field of economics and consumption capacity analysis. It can not only handle large-scale data but also be able to optimize prediction and decision-making through complex models. In this process, machine learning algorithms can effectively capture the hidden laws in economic data, thus providing new research paths for economic growth, consumption capacity forecasting and policy making.

3.1 Current state of application of machine learning algorithms in economic analysis

In the field of economics, researchers have explored a great deal about economic growth, consumer behavior, and market dynamics using machine learning techniques. For example, many scholars have applied machine learning to a variety of aspects such as GDP forecasting, inflation forecasting, income distribution, and consumption capacity analysis. Similarly, Nakamura et al. [9] analysed consumption data by means of a support vector machine (SVM) model and found that the model had an advantage over traditional economic models in predicting short-term consumption fluctuations.

In addition, the application of machine learning in the analysis of consumption capacity has also received widespread attention. The analysis of consumption capacity usually involves multi-dimensional data, including residents' income, consumption structure, savings rate, price level, etc. Although traditional econometric methods can explain the relationship between these factors, they still have limitations in analysing complex and non-linear consumption behaviours. The introduction of machine learning allows these nonlinear relationships to be captured more accurately by more sophisticated algorithms. Hassani et al. [10] used the random forest algorithm to model consumption data and found that the algorithm performs well in handling large-scale datasets and analyzing nonlinear features, and is able to effectively predict consumers' propensity to spend.

3.2 Main application scenarios of machine learning algorithms

3.2.1 Forecasting and Classification of Consumption Capacity

Machine learning algorithms are widely used to classify and predict the consumption patterns of residents in consumption capacity analysis. The K-means clustering algorithm analyses the consumption data of residents and classifies consumers into high-consumption, medium-consumption, and low-consumption groups on the basis of multidimensional features such as income, consumption patterns, and savings rates. This classification can help researchers identify the consumption characteristics of various groups more accurately, thus providing a basis for policy formulation.

In addition, algorithms such as decision trees and random forests can predict future consumer behavior and consumption trends based on historical consumption data. These predictions are important references for business decisions and macroeconomic policymaking.

3.2.2 Non-linear modelling of the relationship between income distribution and consumption

Uneven income distribution is a key factor affecting the ability to consume, and traditional econometric models, while capable of explaining some of the causal relationships, are deficient in dealing with complex non-linear relationships. Machine learning algorithms, especially deep learning models, can better mine the nonlinear relationship between income and consumption. Neural network-based models can capture the potential impact of changes in income distribution on consumption behaviour through layers of non-linear mapping and make predictions based on the consumption patterns of different income classes. This approach is particularly effective in coping with the impact of economic fluctuations and income differentiation on consumption.

3.2.3 Analysis and forecasting of regional consumption differences

China's urban-rural gap and regional development imbalance issues directly affect consumption capacity, and machine learning techniques, such as the Support Vector Regression (SVR) algorithm, can help to analyse the structural factors behind such disparities. In specific applications, the SVR algorithm accurately predicts trends in consumption capacity across regions by combining consumption data with regional economic indicators. For example, the SVR model is able to use multi-dimensional economic data, such as residents' income, price level, infrastructure, etc., to identify the non-linear relationship between the consumption behaviors of different regions, so as to provide policymakers with a more accurate basis for economic decision-making. In the literature, studies such as Wang et al.'s [11] successfully applied the SVR model to analyse the consumption differences between Eastern and western China, concluding that the SVR model has high accuracy in predicting short-term consumption fluctuations. In addition, machine learning models based on geographic information (e.g., geographically weighted regression models, GWR) can combine spatial and economic data to analyse the differences in consumption patterns across regions and their impact on the overall economy in the spatial dimension. By taking spatial heterogeneity into account, the GWR model further enhances its ability to capture differences in the economic development of different regions and provides a stronger basis for the formulation of regional consumption promotion the policies.Stronger basis for the formulation of regional consumption promotion policies.

3.2.4 Sentiment calculation in consumer behaviour analysis

With the explosive growth of data from social media and e-commerce platforms, the relationship between consumer sentiment and consumption behaviour has become a hot research topic. Natural language processing (NLP) technology in machine learning is able to extract sentiment information from consumers' comments and social media speeches and combine it with consumption data, thus analyzing the impact of sentiment fluctuations on consumption behaviors. Xiang et al. [12] investigated the correlation between consumers' sentiment changes and consumption expenditures during major festivals by means of a sentiment analysis algorithm and found that positive sentiments are often accompanied by an increase in consumer spending. This kind of research provides a new basis for enterprises to formulate promotional strategies and the government to adjust consumption policies.

4 Case studies

In this study, China's GDP data is obtained from the official World Bank website, covering 30 years of data between 1991 and 2020, while the data on consumption capacity is obtained from the International Monetary Fund (IMF) statistics.

4.1 Data

In terms of data processing, the paper first calculates indicators of concentration trend (including mean, median, and plural) and measures of dispersion (including extreme deviation, interquartile deviation, variance, and standard deviation) for each of the two variables (GDP and consumption capacity). The basic description of these data can reveal the underlying distributional characteristics of the two variables. To further explore the structural nature of the data, this paper conducted a frequency analysis of the data through Excel and organized them into frequency tables by interval. Pie charts, bar charts, box plots, and line graphs were then generated using Excel's plotting functions to visualize the trends of GDP and spending power over time.

Table 1 Response variable: the power of Chinese consumption

Measures of Central Tendency	Measures of Spread
Mean: $x = \frac{\sum x}{n}3$ = 12902.87	Range:
Median: 8775.03	Max-min= 56064.09-1810 25=54253.84
Mode: N/A	Quartiles Q1:4539. 83, Q2: 8775 03
Outlier lower end=Q1-1.51QR=-33587 .34	Q3: 29957.94.IQR=Q3-Q1=25418 11
Upper end=Q3+1.51QR=68085.11	Variance: 310345041.5
X<33587.34 0r X 68085.11	Standard Deviation: 17616.61
There is no outlier on both end	

Table 1 demonstrates the concentration trend finger. As can be seen from the figure, the mean (Mean) is 12902.87, indicating that China's average consumption capacity was high during the statistical period, showing that consumer spending continues to drive economic development. The median (Median) is 8775.03, indicating that half of the data points have a consumption capacity below this value and the other half above it. The median is slightly lower than the mean, indicating that the data may be influenced by higher values with a somewhat skewed rightward distribution. The outlier analysis mentions that there are no significant outliers (outliers), which means that there are no extreme highs or lows in the analysed dataset, which helps to indicate that the change in spending power is relatively smooth.

Indicator of degree of dispersion:

The variance (Variance) of 310345041.5 and the standard deviation (Standard Deviation) of 17616.61 show that there is some volatility in consumption power, especially in some areas where higher levels of consumption lead to larger standard deviations.

The interquartile range (IQR) of 25418.11 indicates that the middle 50 percent of the data on spending power is relatively widely distributed, suggesting that there are significant differences in spending power between different regions or groups.

The extreme variance (Range) is 54253.84, which further reflects the variability of consumption power, with the maximum and minimum values of consumption power differing considerably in different intervals of the dataset, reflecting the impact of China's urban-rural and regional development imbalance on consumption power.

Next, a linear regression analysis of GDP and consumption power was conducted using the least squares method, resulting in a scatter plot and its line of best fit, and the coefficient of determination and correlation coefficients were calculated so that the relationship between the two variables could be further analyzed.

4.2 Methods of Analysis

In the univariate analysis of China's GDP and consumption power, the results show that the concentration trend and dispersion of China's GDP indicate that the mean value of GDP is $47,063.27 billion, and the median value is $2,213,679 billion. No significant outliers were found during the data analysis, indicating that the data are relatively normal overall. The standard deviation of GDP is 46857.64, reflecting the high volatility of GDP in the dataset, with significant upward and downward fluctuations.

Analyses of consumption capacity show that the mean value of consumption capacity is $1,290,287 million and the median is $877,503 million, indicating a more pronounced difference in the level of consumption among different groups or regions. The standard deviation of this data is 17,616.61, further indicating a wide distribution of consumption power and large differences in consumption power.

In order to explore the relationship between GDP and spending power, a linear regression model was used. By calculating the correlation coefficient (r=0.9981) and the coefficient of determination (R^2=0.9962), the results show a very strong positive linear correlation between the two. In addition, the slope of the regression line is 0.3752 indicating that for every US$10 billion increase in GDP, there is an increase in spending power of US$3,752 million. The intercept is $24.29 billion, and although the intercept value is not of practical significance, it indicates the initial level of consumption in the theoretical model.

4.3 Analysis of results

By analysing the data, the results verify the strong positive correlation between GDP and consumption capacity. From 1991 to 2020, with the continuous growth of China's GDP, the consumption capacity has increased accordingly. Especially in the period of rapid economic development, such as 2001-2010, the trend of simultaneous growth of consumption capacity and GDP is more obvious. This trend reflects the fact that economic development has led to an increase in residents' incomes, which in turn has boosted consumer spending.

Fig. 1 illustrates the trend of GDP and consumption capacity between 1991 and 2020. It can be seen that the growth in GDP has been more significant, especially after the 2008 financial crisis, when China's economy recovered faster and its consumption capacity increased significantly.

Fig. 1 Box plot of China's gross domestic product (GDP) (in billions) (Photo/Picture credit: Original)

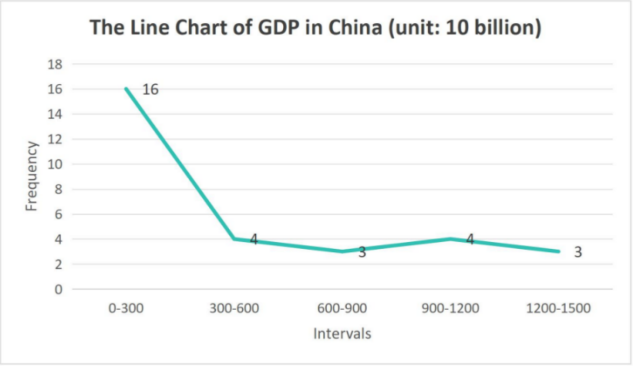

Fig. 2 Line graph of China's GDP (in billions) (Photo/Picture credit: Original)

The line in Fig. 2 shows the trajectory of GDP and consumption capacity over the 30-year period. The distribution of GDP shows greater volatility, especially after 2000, when GDP increased rapidly, suggesting that China's economy made leaps and bounds during this period. In contrast, consumption capacity is less volatile, but generally shows an upward trend, especially after 2000, when growth accelerated markedly.

These trends not only reflect the overall prosperity of the Chinese economy, but also indicate that a combination of factors, including economic policies, higher levels of residents' income, and the optimisation of the consumption structure, are contributing to the rise in consumption capacity. By analysing these changes, the papers show that the focus of future policies should continue to be on promoting residents' income and consumption levels in order to maintain sustained economic growth.

This analytical methodology and graphical presentation can provide decision support to governments and businesses, helping them to better understand the dynamics of consumption in economic development and to formulate more targeted economic policies and business strategies.

5 Conclusion

Although machine learning has shown strong potential for economic and spending power analyses, its application still faces a number of challenges. First, the quality and source of data are key factors that affect the accuracy of machine learning models. Economic data is often characterized by problems such as noise and missing values, which can affect the stability and predictive effectiveness of models. Second, the "black box" nature of machine learning models makes their decision-making process difficult to interpret, especially in the policymaking process, and this opacity may lead to increased decision-making risk.

In the future, with the continuous development of technology, the application of machine learning in economic analyses will become more widespread. By combining big data, cloud computing and other technologies, machine learning will be able to handle larger-scale economic data and provide stronger support for policymaking. At the same time, the development of enhanced explanatory machine learning will help to improve the transparency of the models and make them more useful in economics research and practical applications.

In conclusion, the introduction of machine learning has brought new ideas and tools to the analysis of the economy and consumption capacity. It can not only improve the accuracy of prediction but also help researchers to dig deeper into the non-linear relationships that are difficult to capture by traditional economics methods, which provides an important reference for the decision-making of the government and enterprises. With the continuous optimization of algorithms and the improvement of data quality, the application of machine learning in economics will have a broader prospect.

Reference

1. H. Liu, J. Liu, Real-time forecasting and short-term forecasting of China's macroeconomic aggregates—an empirical study based on a mixed-frequency data forecasting model. Econ. Res. **46**(3), 4-17 (2011)
2. J. Li, H. Li, M. Liu, et al., Research on GDP forecasting under multi-factor coupling based on time series—Taking Yili Prefecture as an example. Pract. Underst. Math. **48**(2), 269-278 (2018)
3. C. Glocker, Kaniovski, et al., Macroeconometric forecasting using a cluster of dynamic factor models. Empir. Econ. **63**(1), 1-49 (2021)
4. Z.C. Tian, M. Liu, Empirical evidence of GDP forecasting based on improved grey GM(1,1) model. Stat. Decis. Mak. **34**(11), 83-85 (2018)
5. P. Wu, S.B. Qiu, Macroeconomic development forecasting based on optimal multidimensional grey model. Stat. Decis. Mak. **36**(3), 42-45 (2020).
6. P.J. Ranjan, M. Ritanjali, K. Rajesh, et al., Impact of COVID-19 on GDP of major economies: application of the artificial neural network forecaster. Econ. Anal. Policy **69**, 324-339 (2021)
7. Q.Y. Xiao, L.J. Liu, T.R. Zhao, et al., Can deep learning neural networks improve GDP forecasting ability? Econ. Manag. Res. **41**(7), 3-17 (2020).
8. J.H. Zhao, Y.H. Li, L. Huo, et al., A macroeconomic forecasting method integrating microblog sentiment analysis and deep learning. Comput. Appl. **38**(11), 3057-3062 (2018)
9. D. Komura, H. Nakamura, S. Tsutsumi, H. Aburatani, S. Ihara, Multidimensional support vector machines for visualization of gene expression data, in Proceedings of the 2004 ACM Symposium on Applied Computing 175-179 (2004)
10. I.E. Hassani, C.E. Mazgualdi, T. Masrour, Artificial intelligence and machine learning to predict and improve efficiency in manufacturing industry. arXiv preprint arXiv:1901.02256 (2019)
11. A. Wang, J.C. Lam, S. Song, V.O. Li, P. Guo, Can smart energy information interventions help householders save electricity? An SVR machine learning approach. Environ. Sci. Policy **112**, 381-393 (2020)
12. R. Xiang, J. Li, M. Wan, J. Gu, Q. Lu, W. Li, C.R. Huang, Affective awareness in neural sentiment analysis. Knowl.-Based Syst. **226**, 107137 (2021)