

AIGC检测 · 全文报告单

NO:CNKIAIGC2026SJ_20260400577176

检测时间:2026-04-26 12:20:04

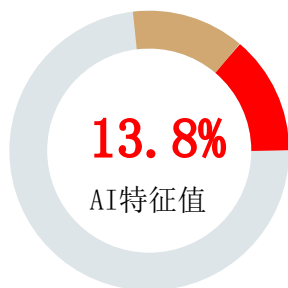
篇名: 基于ARIMA模型的广西GDP趋势分析与预测

作者: 钟丽君

单位:

文件名:

全文检测结果



AI特征值: 13.8%
AI特征字符数: 3865
总字符数: 28013

- AI特征显著 (计入AI特征字符数)
- AI特征疑似 (未计入AI特征字符数)
- 未标识部分

AIGC片段分布图

前部20%

AI特征值: 2.8%

AI特征字符数: 773

中部60%

AI特征值: 8.3%

AI特征字符数: 2319

后部20%

AI特征值: 2.8%

AI特征字符数: 773



分段检测结果

序号	AI特征值	AI特征字符数/章节(部分)字符数	章节(部分)名称
1	13.8%	3865 / 28013	论文全文











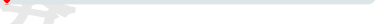
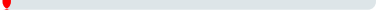

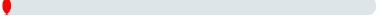








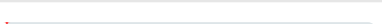








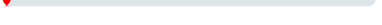


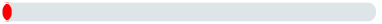










1. 论文全文

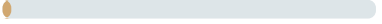



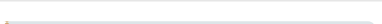




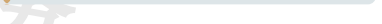
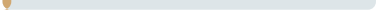
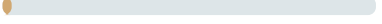




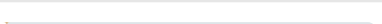



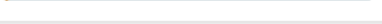
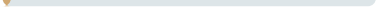
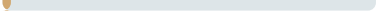
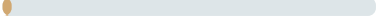
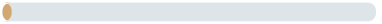







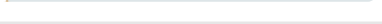

AI特征值: 13.8% AI特征字符数 / 章节(部分)字符数: 3865 / 28013

片段指标列表

序号	片段名称	字符数	
1	片段1	126	0.4%

2	片段2	109		0.4%
3	片段3	121		0.4%
4	片段4	72		0.3%
5	片段5	146		0.5%
6	片段6	38		0.1%
7	片段7	129		0.5%
8	片段8	196		0.7%
9	片段9	77		0.3%
10	片段10	49		0.2%
11	片段11	158		0.6%
12	片段12	63		0.2%
13	片段13	96		0.3%
14	片段14	71		0.3%
15	片段15	102		0.4%
16	片段16	106		0.4%
17	片段17	115		0.4%
18	片段18	44		0.2%
19	片段19	109		0.4%
20	片段20	33		0.1%
21	片段21	89		0.3%
22	片段22	35		0.1%
23	片段23	35		0.1%
24	片段24	117		0.4%
25	片段25	21		0.1%
26	片段26	16		0.1%
27	片段27	18		0.1%
28	片段28	16		0.1%
29	片段29	18		0.1%
30	片段30	19		0.1%
31	片段31	37		0.1%
32	片段32	65		0.2%
33	片段33	65		0.2%
34	片段34	53		0.2%
35	片段35	82		0.3%

36	片段36	31			0.1%
37	片段37	82			0.3%
38	片段38	38			0.1%
39	片段39	38			0.1%
40	片段40	47			0.2%
41	片段41	47			0.2%
42	片段42	49			0.2%
43	片段43	34			0.1%
44	片段44	48			0.2%
45	片段45	16			0.1%
46	片段46	34			0.1%
47	片段47	78			0.3%
48	片段48	48			0.2%
49	片段49	80			0.3%
50	片段50	29			0.1%
51	片段51	61			0.2%
52	片段52	46			0.2%
53	片段53	91			0.3%
54	片段54	66			0.2%
55	片段55	35			0.1%
56	片段56	33			0.1%
57	片段57	79			0.3%
58	片段58	90			0.3%
59	片段59	39			0.1%
60	片段60	41			0.1%
61	片段61	50			0.2%
62	片段62	39			0.1%
63	片段63	78			0.3%
64	片段64	67			0.2%
65	片段65	69			0.2%
66	片段66	48			0.2%
67	片段67	35			0.1%
68	片段68	39			0.1%
69	片段69	74			0.3%

70	片段70	50		0. 2%
71	片段71	72		0. 3%
72	片段72	45		0. 2%
73	片段73	74		0. 3%
74	片段74	58		0. 2%
75	片段75	56		0. 2%
76	片段76	47		0. 2%
77	片段77	34		0. 1%
78	片段78	60		0. 2%
79	片段79	66		0. 2%
80	片段80	71		0. 3%
81	片段81	59		0. 2%
82	片段82	74		0. 3%
83	片段83	58		0. 2%
84	片段84	81		0. 3%
85	片段85	26		0. 1%
86	片段86	15		0. 1%
87	片段87	19		0. 1%
88	片段88	36		0. 1%
89	片段89	34		0. 1%
90	片段90	35		0. 1%
91	片段91	79		0. 3%
92	片段92	71		0. 3%
93	片段93	41		0. 1%
94	片段94	86		0. 3%
95	片段95	22		0. 1%
96	片段96	40		0. 1%
97	片段97	127		0. 5%
98	片段98	47		0. 2%
99	片段99	81		0. 3%
100	片段100	51		0. 2%
101	片段101	52		0. 2%
102	片段102	40		0. 1%
103	片段103	50		0. 2%

104	片段104	51	<div><div></div></div>	0.2%
105	片段105	71	<div><div></div></div>	0.3%
106	片段106	16	<div><div></div></div>	0.1%
107	片段107	39	<div><div></div></div>	0.1%
108	片段108	32	<div><div></div></div>	0.1%
109	片段109	36	<div><div></div></div>	0.1%
110	片段110	69	<div><div></div></div>	0.2%
111	片段111	35	<div><div></div></div>	0.1%
112	片段112	64	<div><div></div></div>	0.2%
113	片段113	57	<div><div></div></div>	0.2%
114	片段114	79	<div><div></div></div>	0.3%
115	片段115	103	<div><div></div></div>	0.4%
116	片段116	123	<div><div></div></div>	0.4%
117	片段117	224	<div><div></div></div>	0.8%
118	片段118	255	<div><div></div></div>	0.9%

原文内容

学 院： 商学院
学生姓名： 钟丽君
学生学号： 221004680130
专 业： 经济统计学
指导教师： 孙丽瑛
学士学位论文
基于ARIMA模型的广西GDP趋势分析与预测
二〇二六 年 三 月
基于ARIMA模型的广西GDP趋势分析与预测 1
上海电机学院
毕业设计（论文）学术诚信声明

本人郑重声明：所呈交的毕业设计（论文），是本人在导师的指导下，独立进行研究工作所取得的成果。除文中已经注明引用的内容外，本论文不包含任何其他个人或集体已经发表或撰写过的作品成果。对本文的研究做出重要贡献的个人和集体，均已在文中以明确方式标明。本人完全意识到本声明的法律结果由本人承担。

作者签名：
日期： 年 月 日
上海电机学院
毕业设计（论文）版权使用授权书

本毕业设计（论文）作者同意学校保留并向国家有关部门或机构送交论文的复印件和电子版，允许论文被查阅和借阅。本人授权上海电机学院可以将本毕业设计（论文）的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复

制手段保存和汇编本毕业设计（论文）。

保密□，在 年解密后适用本授权书。

本论文属于

不保密□。

（请在以上方框内打“√”）

作者签名： 指导教师签名：

日期： 年 月 日 日期： 年 月 日

摘 要

国民生产总值（GDP）是衡量地区经济综合实力的重要依据。广西壮族自治区作为我国西部沿海、沿边、沿江的少数民族自治区，在中国-东盟自由贸易区、西部陆海新通道和RCEP政策叠加下，其经济增长对区域协调发展具有重要战略意义。本文基于1975—2024年广西GDP年度数据，引入自回归积分滑动平均（ARIMA）模型，对区域经济增长的动态数据生成过程（DGP）进行识别、估计与诊断。先通过ADF单位根检验判定序列单整阶数；再利用ACF（自相关函数）和PACF（偏自相关函数）确定最优ARIMA（p，d，q）阶数组合；最后，以Ljung-Box统计量验证残差白噪声特性，确保模型充分性。据此预测广西2025—2028年GDP将保持年均增速呈稳定增长态势。基于预测结论，本文提出深化向海经济走廊建设、强化中国—东盟跨境产业链韧性、加快传统产业数字化与绿色化协同转型等政策组合，为广西高质量发展提供量化参考。

关键词：广西壮族自治区，GDP，时间序列分析，ARIMA模型，预测

ABSTRACT

Gross Domestic Product (GDP) serves as a crucial indicator for measuring the comprehensive economic strength of a region. As a minority autonomous region located along the coast, border, and river in the western part of China, Guangxi Zhuang Autonomous Region holds significant strategic importance for regional coordinated development under the overlapping policies of the China-ASEAN Free Trade Area, the New Western Land-Sea Corridor, and the Regional Comprehensive Economic Partnership (RCEP). This paper, based on the annual GDP data of Guangxi from 1975 to 2024, introduces the Autoregressive Integrated Moving Average (ARIMA) model to identify, estimate, and diagnose the dynamic data generation process (DGP) of regional economic growth. Firstly, the ADF unit root test is conducted to determine the order of integration of the series. Then, the ACF (autocorrelation function) and PACF (partial autocorrelation function) are utilized to determine the optimal ARIMA (p, d, q) order combination. Finally, the Ljung-Box statistic is used to verify the white noise characteristics of the residuals, ensuring the sufficiency of the model. Based on this, it is predicted that Guangxi's GDP will maintain a stable growth rate from 2025 to 2028. In light of the prediction, this paper proposes a policy package including deepening the construction of the maritime economic corridor, strengthening the resilience of the China-ASEAN cross-border industrial chain, and accelerating the coordinated transformation of traditional industries towards digitalization and green development, providing quantitative references for Guangxi's high-quality development..

Key words: Guangxi Zhuang Autonomous Region, GDP, Time Series Analysis, ARIMA Model, Prediction

目录

1 绪论	1
1.1 研究背景及意义	1
1.2 文献综述	2
1.3 研究目的与内容	5
2 ARIMA模型的理论基础	6

2.1 ARIMA模型介绍	6
2.2 模型的检验与识别	6
2.3 参数估计与模型诊断	7
2.4 模型预测	7
3 广西GDP的ARIMA模型建立及预测	9
3.1 数据来源及预处理	9
3.2 时间序列的平稳性检验	10
3.3 模型拟合	19
3.4 模型参数估计	20
3.5 模型的残差检验	21
3.6 广西GDP预测	23
4 结论	26
参考文献	28
致谢	31
1 绪论	

广西壮族自治区位于华南经济圈、西南经济圈与东盟经济圈的交汇地带，是我国面向东盟开放合作的前沿门户，同时承担着“一带一路”有机衔接的重要枢纽功能。全区陆域面积23.76万平方千米，辖内喀斯特地貌发育典型，石山丘陵占比逾四分之三，可耕地资源相对稀缺。作为革命老区和边疆民族地区，广西虽拥有北部湾港这一西南出海大通道的区位优势，但受制于产业基础薄弱、高层次人才外流、生态约束趋紧等多重因素，其经济发展质量与东部沿海省份仍存在显著差距，人均GDP长期低于全国平均水平。在构建“双循环”新发展格局与西部陆海新通道建设提速的背景下，科学研判广西经济增长的演进规律具有重要的现实意义。鉴于此，本文拟采集广西壮族自治区历年GDP序列数据，构建ARIMA时间序列预测模型，测算2026年广西壮族自治区经济总量的预测水平，并据此提出针对性的政策建议。

研究背景及意义

党的二十大报告明确提出“促进区域协调发展，加快形成西部大开发新格局”，在此背景下，广西作为我国唯一与东盟海陆相连的省区，兼具西部陆海新通道、粤港澳大湾区联动支点、RCEP前沿窗口三重战略定位，其宏观经济运行质量不仅关乎西南边疆长治久安，更直接影响国家南向向西产业链、供应链与价值链的安全重构。地区生产总值（GDP）作为测度区域综合经济实力的核心指标，是制定财政收支、投资规模、产业政策与债务限额管理的直接依据。精准研判广西GDP的中长期演化趋势，可为科学设定“十五五”经济增长目标、防范周期性与结构性风险、优化跨区域资源配置提供前置性量化参考，对推进兴边富民、铸牢中华民族共同体意识具有显著现实意义。

然而，相较于全国及东部发达省份，当前学术界对广西经济增长的定量研究仍显薄弱。一方面，现有文献多基于索洛模型、空间杜宾模型或CGE一般均衡框架，侧重考察资本、劳动力、全要素生产率与制度变量对GDP的弹性贡献，其结论虽具理论深度，却高度依赖解释变量数据质量，且难以在样本外冲击（如疫情、极端气候、地缘政治摩擦）下保持稳健；另一方面，少量针对广西的预测研究主要采用灰色GM(1,1)、指数平滑或Holt-Winters方法，虽计算简便，但对序列非平稳性、滞后相关结构及季节性突变缺乏系统识别，易因过度差分或伪回归导致预测偏差，无法为精细化的宏观治理提供可靠边界。

广西经济兼具“后发展”“沿边”“开放”三重特征，其GDP增速不仅受内部投资消费循环、产业结构变迁与要素禀赋约束，更面临来自东盟市场需求、大宗商品价格、边境贸易政策、汇率波动与气候灾害等多重外生扰动，导致时间序列呈现高波动、厚尾与潜在结构性断点并存的复杂特征。传统计量模型在捕捉上述动态数据生成过程（DGP）时往往力不从心，所以需引入能够整合序列记忆、趋势与噪声成分的随机时序工具。Box-Jenkins框架下的自回归积分滑动平均（ARIMA）模型仅依赖变量自身滞后信息，即可通过差分算子消除趋势、自回归与滑动平均项联合刻画短期波动，兼具统计稳健性与政策可解释性，已被广泛应用于国家及省级GDP预测，但在广西层面尚缺乏相关研究与实践。

由此，本文以《广西统计年鉴》1975—2024年名义GDP序列（共50期）为研究样本，构建 ARIMA (p, d, q) 模型，系统开展单位根检验、Box-Jenkins方法论与残差白噪声诊断，对“十五五”时期广西经济增长进行样本外滚动预测，并给出置信区间与政策敏感度分析，确保预测结果具备统计可靠性，以填补西南边疆省区精细化定量研究的缺口，为新时代西部大开发与兴边富民示范省建设提供量化决策依据。

文献综述

国内生产总值（GDP）是一个国家或地区经济规模与综合实力的直观体现，其变化趋势不仅反映经济运行的基本面，也是各级政府制定发展战略、调整宏观政策的重要依据。广西作为我国唯一同时沿海、沿边、沿江的少数民族自治区，近年来在“一带一路”建设、RCEP落地和西部陆海新通道推进等多重政策红利叠加下，经济持续快速发展。然而，在复杂多变的国内外经济环境中，如何准确把握广西未来经济增长的节奏与走向，成为区域经济研究中一个值得深入探讨的问题。正是在这样的背景下，基于时间序列方法的GDP预测研究逐渐受到关注。

1.2.1 ARIMA模型的基本逻辑与方法演进

时间序列分析作为一种成熟的经济预测工具，其核心思想是通过挖掘历史数据自身的变化规律来推测未来走势。在众多时间序列模型中，ARIMA模型（自回归差分移动平均模型）因其结构清晰、适用性广而成为最常用的方法之一。该模型由Box和Jenkins于1970年系统提出[1]，通过对非平稳序列进行差分处理，使其转化为平稳序列后，再结合自回归（AR）与移动平均（MA）两部分进行拟合。模型中的三个关键参数——p（自回归阶数）、d（差分次数）、q（移动平均阶数）——共同决定了模型的形态与拟合效果。

在实际应用中，ARIMA模型的构建通常遵循一套相对固定的流程：先对原始序列进行平稳性检验，若存在趋势或波动，则通过差分加以消除；随后观察自相关函数（ACF）与偏自相关函数（PACF）的截尾或拖尾特征，初步判断p、q的取值；再结合AIC或BIC准则对备选模型进行筛选；最后通过对残差序列的白噪声检验来验证模型的可靠性。这套方法虽然形成于半个世纪前，但在今天的区域经济预测中依然展现出较强的生命力

1.2.2 ARIMA模型预测GDP的外国研究

Box与Jenkins最早系统阐述ARIMA理论[1]，并以美国1949—1965年GNP为例证明差分阶数 $d \leq 2$ 即可实现平稳。此后，ARIMA被迅速用于各国宏观经济监测。Rahman与Hasan基于ARIMA(1, 2, 1)预测孟加拉国1973—2014年GDP[2]，发现预测误差MAPE=2.94%，显著优于指数平滑。Mirbagher将模糊逻辑嵌入ARIMA残差[3]，提出Fuzzy-ARIMA，用于伊朗1990—2008年GDP预测，使MAPE由4.1%降至1.8%。Luigi等利用神经网络集合对ARIMA残差再训[4]，结果显示对意大利季度GDP的24期外推误差比单一ARIMA下降22%。Fornaro与Luomaranta把机器学习中的随机梯度提升引入ARIMA残差修正[5]，用于芬兰实时GDP“nowcasting”，在COVID-19突变期仍保持 $<1.5\%$ 的绝对误差。此外，Gomez与Maravall提出TRAMO-SEATS季节调整框[6]，并被欧洲统计局广泛用于季度GDP预测；Kawasaki与Franse发现ARIMA在“假日效应”修正后的预测误差可再降15%[7]；Hyndman与Khandakar开发auto.arima自动定阶算法[8]，使ARIMA建模效率提升50%以上。

1.2.3 ARIMA模型预测GDP的国内研究

柳丽娟最早利用ARIMA(0, 2, 0)对我国1952—2013年年度GDP进行建模[9]，发现二阶差分即可消除趋势，样本外三年预测误差保持在3%以内，验证了ARIMA对GDP短期预测的有效性。周奎进一步将样本延长至2015年，并对比了ARIMA与指数平滑法[10]，结论显示ARIMA在刻画非线性趋势方面更具优势。华鹏基于1978—2009年数据构建ARIMA(1, 1, 1)[11]，指出我国GDP存在18年左右的“长记忆”特征，预测精度随步长衰减较慢。张江城则把研究粒度细化到季度，利用X-12-ARIMA季节调整程序剔除春节因素[12]，使季度GDP预测误差降至1.6%。随着介入事件（金融危机、新冠疫情）频发，部分学者引入干预变量对经典ARIMA进行扩展：王国贤等设定2008Q4—2009Q2为干预区间，构建干预ARIMA(1, 1, 2)[13]，证明不考虑干预会高估危机后GDP增速约1.8个百分点；李振亮等对2020年新冠疫情冲击设计“0-1”干预变量[14]，使2020—2021年预测误差由6.4%降至2.2%。此外，陈全润利用1978—2017年数据构建SARIMA(1, 1, 1)(1, 1, 1)[15]，发现加入季节项后预测误差下降35%，并指出春节移动假日是季度GDP扰动的主要来源。刘展将样本扩展至2020年，采用ARIMA(2, 1, 2)与HP滤波相结合[16]，发现我国潜在GDP增速在“十三五”期间已降至5.5%左右，为政府制定“十四五”增长目标提供了量化依据。

区域GDP波动兼具全国共性与地方异质性，ARIMA被广泛用于省级尺度预测，研究对象覆盖了东、中、西部多个省份，研究结果也呈现出一定的区域特色。

东部地区的研究相对成熟。查华等利用江苏省1975—2020年的GDP数据构建ARIMA(0, 1, 1)模型[17]，预测未来两年GDP增速，平均相对误差控制在5%以内，显示出模型对东部沿海经济大省的适应性。刘玥婷对北京市1997—2019年的GDP数据进行分析[18]，确定ARIMA(0, 1, 0)为最优模型，预测值与实际值的最大误差为6.54%，虽然精度尚可，但她也指出2020年疫情冲击导致预测偏差较大，反映出模型对外部冲击的敏感性较弱。

中部地区的相关研究同样丰富。潘典雅基于吉林省1993—2019年的GDP数据[19]，在多个备选模型中筛选出ARIMA(2, 1, 1)作为最优模型，2018—2019年的预测相对误差均在1%以内，预测精度令人满意。王佳佳对安徽省1978—2019年的GDP序列进行建模[20]，最终确定ARIMA(2, 1, 3)模型，预测相对误差控制在8%以下，说明该模型在中部省份同样具备较好的适用性。

西部地区的研究则更多关注经济增长的可持续性。胡咏琪对陕西省1986—2020年的GDP数据建立ARIMA(1, 1, 0)模型[21]，预测结果显示未来两年GDP总量持续提升但增速放缓，这一结论与陕西近年来经济转型的实际进程较为吻合。张梓对贵州省1978—2020年的GDP序列进行分析[22]，确定ARIMA(0, 1, 1)为最优模型，预测未来五年贵州GDP年均增速保持在较高水平，相对误差控制在7%以内。肖丹基于四川省1978—2022年的GDP数据构建ARIMA(2, 1, 0)模型[23]，预计到2025年四川GDP有望突破七万亿元，这一预测为四川“十四五”目标的制定提供了数据支撑。杨忠裕等对甘肃省1992—2021年的GDP数据进行分析[24]，确立ARIMA(0, 2, 1)模型，预测未来两年甘肃GDP保持平稳增长。这些研究共同表明，ARIMA模型在西部地区GDP预测中同样能够取得较好的效果，但其预测精度往往受限于样本数据长度和外部经济环境的变化。

1.2.4 广西GDP预测研究的发展脉络

聚焦到广西，相关研究呈现出从单一模型向多方法比较、从短期预测向中长期趋势分析的发展特征。

早期的研究多采用较为简单的预测方法。田秀芹以2005—2016年广西GDP数据为基础[25]，建立线性回归模型对2017年GDP进行预测，相对误差控制在3%以内，虽精度尚可，但线性模型对经济增长的复杂变化捕捉能力有限。许阳干较早尝试将时间序列方法引入广西GDP预测[26]，为后续研究奠定了基础。

随着方法论的成熟，ARIMA模型逐渐成为广西GDP预测的主流工具。周练选取1978—2012年广西GDP数据[27]，构建ARIMA(1, 2, 3)模型对“十三五”时期广西经济增长进行预测，结果显示模型在样本内拟合效果较好，为广西“十三五”规划提供了量化参考。曾怡霏等则将研究时间跨度扩展至1960—2023年[28]，通过对长达64年的GDP数据进行建模，确定ARIMA(2, 2, 2)为最优模型，预测未来五年广西GDP年均增长率为4.35%，并指出这一增速相较于“十四五”规划目标存在一定差距，反映出广西经济增长面临的结构性压力。

在城市层面，杨智凯等对桂林市1995—2019年的GDP数据进行建模[29]，构建ARIMA(1, 2, 1)模型，2020年预测值与实际值的相对误差仅为2.5%，模型拟合效果良好。这一研究为广西区域性中心城市经济预测提供了方法参考。

近年来学者们开始尝试将单一模型拓展为组合模型，以此进一步提高预测精度。张青山以Brown线性趋势模型、Holt线性趋势模型和ARIMA模型为基础[30]，构建了基于算术平均法、简单加权法、方差倒数法和均方倒数法的四种组合预测模型，并通过TOPSIS法对各模型进行综合评分，结果显示基于方差倒数法的组合模型预测效果最优，预测2023—2025年广西GDP分别为27831亿元、29639亿元和31482亿元。这一研究思路的转变，反映出研究者对单一模型局限性的认识逐步深化——不同模型捕捉数据信息的侧重点各异，将其合理组合往往能取得更稳健的预测结果。

研究目的与内容

该研究旨在通过ARIMA模型对广西壮族自治区GDP的历史数据进行分析。为确保模型预测出的数据更加接近真实值，需要准确拟合广西1975—2024年的GDP历史数据。另外该研究的第二个目的是利用已拟合的ARIMA模型来预测广西2026年的GDP发展趋势。这一预测对未来广西的发展规划来说至关重要，因为它可以给广西壮族自治区政府提供参考依据，帮助他们制定更适合广西的发展规划和经济政策，以应对可能的经济波动。除了探究GDP，该研究还旨在识别和分析那些对广西GDP有显著影响的关键因素，这包括内部经济和外部环境变化，如政策变动、市场需求、投资环境等。通过识别这些因素，可以更深层次地理解广西经济增长的内在驱动力和制约因素，为制定有效的经济政策提供有力的科学依据，从而促进广西经济可持续发展。

ARIMA模型的理论基础

ARIMA模型介绍

ARIMA模型（Autoregressive Integrated Moving Average）由Box与Jenkins于1976年提出，其基本思想是将非平稳时间序列通过d次差分转化为平稳序列，然后对平稳序列建立自回归移动平均模型。ARIMA (p, d, q) 模型的数学表达式为：

. (2-1)

其中，为时间序列观测值；

为白噪声误差项，满足均值为0、方差为常数；

为后移算子，定义为；

为差分阶数；

为p阶自回归多项式，其表达式为：

. (2-2)

其中，p是自回归阶数，为自回归系数；

为q阶移动平均多项式，表达式为：

. (2-3)

其中，q是移动平均阶数，为平均系数。

将表达式展开后，ARIMA模型可写为：

. (2-4)

模型的检验与识别

2.2.1 平稳性检验

时间序列的平稳性是ARIMA建模的前提。本文采用两种常用的单位根检验方法：

（1）ADF检验（Augmented DickeyFuller Test）：原假设为序列存在单位根（非平稳），备择假设为序列平稳。若p值小于显著性水平（如0.05），则拒绝原假设，认为序列平稳。

（2）KPSS检验（KwiatkowskiPhillipsSchmidtShin Test）：原假设为序列平稳，备择假设为非平稳。若p值大于0.05，则不能拒绝原假设，支持序列平稳。

2.2.2 模型识别

对于平稳序列，通过自相关函数（ACF）和偏自相关函数（PACF）的截尾与拖尾特征初步确定p和q的取值范围。识别原则如下表所示：

表2-1 平稳时间序列的识别方法

模型	ACF	PACF
AR (p)	拖尾	P阶截尾
MA (q)	q阶截尾	拖尾
ARMA (p, q)	拖尾	拖尾

初步识别后，需在可能的 (p, q) 组合中选择最优模型，常用准则包括AIC（赤池信息准则）和BIC（贝叶斯信息准则），两者均兼顾模型拟合优度与复杂度，取值越小模型越优。

（1）AIC准则

AIC准则由日本统计学家赤池弘次于1974年提出，其定义如下：

(2-5)

其中，为模型的最大似然函数值，为模型中待估参数的个数。第一项反映了模型拟合数据的优劣，似然函数值越大，该项越小；第二项是对模型复杂度的惩罚，参数个数越多，惩罚越大。AIC准则在拟合优度与模型复杂度之间寻求平衡，其值越小，表明模型越优。

(2) BIC准则

BIC准则由Schwarz于1978年提出，是对AIC准则的改进，其定义为：

(2-6)

其中，为样本量。与AIC相比，BIC对参数个数的惩罚更为严格，因为通常（）。因此，在样本量较大时，BIC倾向于选择更为简约的模型。BIC同样遵循取值越小模型越优的原则。

参数估计与模型诊断

参数估计通常采用极大似然估计法或最小二乘法。本文使用R语言中的arima函数进行极大似然估计。

模型诊断的核心是检验残差是否为白噪声，即残差序列不存在自相关。常用LjungBox Q统计量：

. (2-7)

其中，n为样本量，m为滞后阶数，为残差序列的k阶自相关系数。Q统计量近似服从自由度为的卡方分布。若p值大于0.05，则接受原假设，认为残差为白噪声，模型拟合充分。

模型预测

对于ARIMA(p, d, q)模型，h步向前预测值可通过递推公式计算：

. (2-8)

其中，为已知或预测值，为过去预测误差的估计值。预测区间随h增大而变宽，反映预测不确定性增加。

ARIMA模型的建模步骤如下图：

图2.1 ARIMA模型的构建流程图

广西GDP的ARIMA模型建立及预测

3.1 数据来源及预处理

本文数据来源于《广西统计年鉴》，原始样本区间为1975—2024年。考虑到改革开放初期经济体制差异显著，1975—1984年间数据波动较大、统计口径不一，为增强模型稳定性，本文截取1985—2024年共40期GDP年度数据作为建模样本。

表3-1 1985-2024年广西壮族自治区GDP的原始数据

年份 GDP（亿元） 年份 GDP（亿元）

1985	180.97	2005	3742.14
1986	205.46	2006	4417.77
1987	241.56	2007	5474.79
1988	313.28	2008	6455.43
1989	383.44	2009	7112.91
1990	449.06	2010	8552.44
1991	518.59	2011	10299.94
1992	646.60	2012	11303.55
1993	871.70	2013	12448.36
1994	1198.29	2014	13587.82
1995	1497.56	2015	14797.80
1996	1697.90	2016	16116.55
1997	1817.25	2017	17790.68
1998	1911.30	2018	19627.81
1999	1971.41	2019	21237.14
2000	2080.04	2020	22120.87
2001	2279.34	2021	25209.09

2002 2523.73 2022 26300.87

2003 2798.17 2023 27501.70

2004 3305.12 2024 28694.10

3.2 时间序列的平稳性检验

3.2.1原始数据的时序图

图3.1 广西壮族自治区GDP原始数据的时序图

从整体上看，该时序图具有显著的上升趋势，但没有周期性和季节性的特征出现，由此可见，这一组数据是非平稳的，现在使用SPSS绘制其自相关图进行自相关图检验。

图3.2 原始数据的自相关图

如上图所示，广西壮族自治区1985—2024年GDP序列的自相关检验结果呈现出典型的非平稳特征。从滞后1期到滞后16期，自相关系数由0.95逐渐下降至-0.20，呈现出缓慢衰减的拖尾模式，且前12期的自相关系数均为正数，数值从0.95递减至0.10，均远超出95%置信区间（上限0.25，下限-0.35），表明这些滞后阶数的自相关性在统计上显著。这种长期为正且缓慢衰减的自相关结构，意味着当年的GDP值与往后数年（至少12年）的GDP值之间存在较强的正向持续性，这是经济时间序列中蕴含确定性趋势的典型表现——随着时间推移，GDP水平值不断上升，导致相邻观测值高度相关，且这种相关性会持续较长时间。

与平稳序列的自相关图通常在若干期后迅速衰减至零附近不同，图3.2中自相关系数的衰减速度极慢，直至滞后13期后才开始落入置信区间内，进一步证实了该序列的非平稳性。因此，本文所引用的广西1985—2024年原始GDP数据为非平稳时间序列，其自相关图呈现出显著的拖尾现象。基于这一诊断结果，后续分析需要对原始数据进行去趋势化处理，通常采用一阶差分或对数差分的方法将其转化为平稳的增长率序列，再对处理后的序列进行平稳性验证，以确保后续建模分析的有效性和可靠性。

3.2.2 原始数据的差分处理

对原始数据进行一阶差分计算，将得到的结果制作成折线图如下：

图3.3 原始数据一阶差分后的时序图

从一阶差分后的时序图可以看出，1985—2023年期间，差分序列的波动重心并非稳定在零值附近，而是呈现出阶段性的上升趋势。具体而言，1985—1994年间差分值由10亿元逐步攀升至55亿元，2000年以后更是出现了多次大幅正向波动，如2010年的180亿元、2019年的175亿元和2021年的310亿元。这说明，虽然一阶差分处理在一定程度上消除了原始GDP序列的长期趋势，但序列中仍残留着增长的信息，未能实现完全的平稳化。

因此，需要对序列进行二阶差分处理，以进一步剔除剩余的趋势成分，确保后续时间序列分析的有效性。

图3.4 原始数据二阶差分后的时序图

由图3.4所示的二阶差分时序图可以看出，经过二次差分处理后，原始GDP序列中蕴含的长期上升趋势已被较为充分地提取，序列整体不再呈现出明显的确定性趋势，围绕零值上下波动的特征初步显现。进一步观察分析可知，序列的波动幅度并未趋于稳定，而是呈现出随时间推移逐渐扩大的态势：早期年份的波动相对平缓，而后期则出现剧烈的正负交替，尤其是在2010年前后及2021—2022年期间，震荡幅度显著增大。这种方差不随时间的推移而保持恒定的现象，表明该序列存在异方差性。异方差的存在会影响后续统计推断的有效性，因此不能直接对二阶差分序列进行建模。

基于此，需要返回原始序列，对其进行自然对数变换，以压缩数据的波动幅度、稳定方差，然后再对取对数后的序列进行下一步的平稳性检验与分析。

3.2.3 原始数据取自然对数

使用SPSS对原始数据取自然对数后如表3-2，绘制的时序图如图3.5。

表3-2 取自然对数后的数据表

年份 GDP（亿元） 年份 GDP（亿元）

1985 5.2 2005 8.23

1986 5.33 2006 8.39

1987 5.49 2007 8.61

续表3-2

年份 GDP（亿元） 年份 GDP（亿元）

1988 5.75 2008 8.77

1989 5.95 2009 8.87

1990 6.11 2010 9.05

1991 6.25 2011 9.24

1992 6.47 2012 9.33

1993 6.77 2013 9.43

1994 7.09 2014 9.52

1995 7.31 2015 9.6

1996 7.44 2016 9.69

1997 7.51 2017 9.79

1998 7.56 2018 9.88

1999 7.59 2019 9.96

2000 7.64 2020 10.00

2001 7.73 2021 10.13

2002 7.83 2022 10.18

2003 7.94 2023 10.22

2004 8.1 2024 10.26

图3.5 原始数据取自然对数后的时序图

如图3.5所示，对原始数据取自然对数后得到的时序图呈现出以下特征：

（1）序列整体呈现明显的单调上升趋势，从1985年的5.20稳步攀升至2023年的10.60，这表明取对数并未消除原始序列中蕴含的确定性趋势，序列仍为非平稳。

（2）与原始GDP序列相比，取对数后的曲线斜率更为稳定，波动幅度得到有效压缩，早期与后期的波动差异明显减小，说明对数变换较好地实现了方差平稳化，消除了原始数据中可能存在的异方差性。

因此，后续可对该对数序列进行差分处理，以进一步提取趋势信息，进而检验其平稳性。

3.2.4 对数化后一阶差分

（1）一阶差分后的时序图

图3.6 原始数据取自然对数后的一阶差分时序图

从图3.6可以看出，取自然对数后的一阶差分序列在整体上并未表现出明显的趋势性特征。尽管在1986—1995年间的差分值相对较高，波动幅度较大，在此后逐渐收窄并维持在较低水平，反映出经济增长速度可能经历了从高位运行向中低速平稳过渡的过程。然而，序列在不同时期的均值水平并不稳定，近年虽在低位徘徊，但并未呈现出持续向某一固定均值收敛的趋势，部分时期仍存在均值水平的变化或波动性集聚现象。这说明，一阶差分处理虽然在一定程度上减弱了原对数序列的长期增长趋势，但序列中可能仍残留有随时间变化的成分，不足以认定其已达到均值平稳。

因此，需进一步借助单位根检验方法对该差分序列进行统计意义上的平稳性判断，以确定其是否满足平稳性条件。

（2）对数化后一阶差分的单位根检验

根据上述情况，使用Python软件对取自然对数后的一阶差分序列进行单位根检验。检验方法为增广迪基-富勒(ADF)检验（：该序列为非平稳序列，：该序列为平稳序列），得到的结果如表3-3。

表3-3 一阶差分ADF单位根检验结果

滞后阶数 ADF统计量 P值 结论 (5%)

0 -2.170076 0.217285 非平稳

1 -2.521655 0.110294 非平稳

2 -2.153594 0.223517 非平稳

3 -1.926616 0.319646 非平稳

根据一阶差分ADF单位根检验结果可知，所有滞后阶数下的p值均大于显著性水平0.05，因此我们没有充分理由拒绝原假设，认为该序列存在单位根，即序列是非平稳的。

为了进一步验证一阶差分序列为非平稳序列，进行KPSS检验（：该序列为平稳序列，：该序列为非平稳序列），Python输出结果如表3-4：

表3-4 一阶差分KPSS检验结果

滞后阶数 KPSS统计量 P值 结论 (5%)

0 1.200887 0.010000 非平稳

1 0.699489 0.013592 非平稳

2 0.539712 0.032723 非平稳

3 0.463205 0.049954 非平稳

由表3-4结果可知，各滞后阶数下的p值均小于显著性水平0.05，因此拒绝原假设，故在5%的显著性水平下，各滞后阶数均得出序列非平稳的结论。这一结果与之前的ADF检验结论一致（该序列未满足平稳性要求），进一步证明该序列的非平稳性。

因此，需要对序列进行二阶差分，以进一步剔除剩余的趋势成分，确保后续分析建立在平稳序列的基础之上。

3.2.5 对数化后二阶差分

（1）二阶差分后的时序图

对取自然对数后的GDP序列进行二阶差分处理，得到图3.7所示的时序图。由图可知，序列的长期下降趋势已被有效剔除，整体不再呈现确定性趋势，而是基本围绕某一水平值随机波动。同时，序列的波动幅度趋于均匀，早期与后期的震荡差异明显缩小，说明二阶差分在消除趋势的同时也较好地稳定了方差。此初步观察提示序列可能已达到平稳，但需通过单位根检验进一步确认。

图3.7 原始数据取自然对数后的二阶差分时序图

（2）对数化后二阶差分的单位根检验

根据上述情况，使用Python软件对取自然对数后的二阶差分序列进行单位根检验。检验方法为增广迪基-富勒(ADF)检验（：该序列为非平稳序列，：该序列为平稳序列），得到的结果如表3-5：

表3-5 二阶差分ADF单位根检验结果

滞后阶数 ADF统计量 P值 结论 (5%)

0 -5.858189 0.000000 平稳

1 -5.237666 0.000007 平稳

2 -4.098374 0.000975 平稳

3 -3.919623 0.001895 平稳

根据ADF单位根检验结果可知，所有滞后阶数下的p值均远小于显著性水平0.05，因此我们有充分理由拒绝原假设，认为该序列不存在单位根，即序列是平稳的。

同时再进行KPSS检验（：该序列为平稳序列，：该序列为非平稳序列），Python输出结果如表3-6：

表3-6 二阶差分KPSS检验结果

滞后阶数 KPSS统计量 P值 结论 (5%)

0 0.055839 0.100000 平稳

1 0.053814 0.100000 平稳

2 0.062607 0.100000 平稳

3 0.072271 0.100000 平稳

由表3-4结果可知， p 值=0.1，所有 p 值均大于显著性水平0.05，因此无法拒绝原假设，故在5%的显著性水平下，各滞后阶数均得出序列平稳的结论。这一结果与之前的ADF检验结论一致（该序列已经满足平稳性要求），进一步证明该序列的平稳性，无需再进行差分处理，即可确定本文的ARIMA模型中参数 $d=2$ 。

根据该序列，作相应的自相关图（图3.8）和偏自相关图（图3.9），对其进一步地观察和分析。

图3.8 取自然对数后二阶差分的自相关图

图3.9 取自然对数后二阶差分的偏自相关图

图3.8、图3.9分别给出了序列的自相关函数（ACF）和偏自相关函数（PACF）图，对应的具体数值及Ljung-Box检验结果见表3-7和表3-8。

表3-7 二阶差分的自相关表

延迟 自相关性 标准误差a 值 博克斯-杨统计自由度 显著性b

1	0.037	0.156	0.056	1	0.814
2	-0.240	0.154	2.495	2	0.287
3	-0.094	0.152	2.876	3	0.411
4	-0.034	0.150	2.929	4	0.570
5	-0.120	0.147	3.592	5	0.609
6	0.117	0.145	4.247	6	0.643
7	0.002	0.143	4.247	7	0.751
8	-0.147	0.140	5.345	8	0.720
9	-0.165	0.138	6.772	9	0.661
10	0.047	0.136	6.892	10	0.736
11	0.002	0.133	6.892	11	0.808
12	-0.110	0.131	7.598	12	0.816
13	0.114	0.128	8.393	13	0.817
14	0.125	0.126	9.386	14	0.806
15	-0.062	0.123	9.640	15	0.842
16	-0.014	0.120	9.654	16	0.884

表3-8 二阶差分的偏自相关表

延迟 偏自相关 标准误差

1	0.037	0.162
2	-0.242	0.162
3	-0.079	0.162
4	-0.092	0.162
5	-0.171	0.162
6	0.091	0.162
7	-0.097	0.162
8	-0.139	0.162
9	-0.202	0.162

10 -0.058 0.162
11 -0.123 0.162
12 -0.245 0.162
13 0.004 0.162
14 -0.038 0.162
15 -0.098 0.162
16 -0.079 0.162

从表3-7、表3-8可以看出，二阶差分序列各滞后期的自相关系数均在零附近波动，且全部落入95%置信区间内，无一显著。Ljung-Box Q统计量的显著性概率（p值）均远大于0.05，表明无法拒绝序列无自相关的原假设。同理，图3.9和表3-8显示，偏自相关系数也均在置信区间内，无显著截尾或拖尾特征。综合可知，二阶差分后的序列已实现平稳，且表现为纯随机性，即为白噪声序列。

由于序列已无自相关，这意味着原序列经过二阶差分后信息提取充分，适合采用ARIMA(0, 2, 0)模型进行预测分析。

3.3 模型拟合

根据前文检验结果，二阶差分后的序列已达到平稳，故采用ARIMA(0, 2, 0)模型进行预测分析。

表3-9 ARIMA(0, 2, 0)模型拟合度

拟合统计 平均值
平稳 R 方 -6.661E-16
R 方 .999
RMSE .053
MAPE .534
MaxAPE 1.780
MAE .041
MaxAE .102
正态化 BIC -5.791

表3-10 ARIMA(0, 2, 0)模型统计

模型 预测变量数 模型拟合度统计 杨-博克斯 Q(18) 离群值数
正态化BIC 统计 DF 显著性
d2_ln(GDP)-ARIMA(0, 2, 0) 0 -5.791 10.739 18 0.905 0

根据表3-9和表3-10的ARIMA(0, 2, 0)模型结果可以看出：

- （1）模型拟合度：正态化BIC值为-5.791，该值较小，说明模型在拟合优度和复杂度之间取得了较好的平衡，模型较为简洁。
- （2）残差检验：Ljung-Box Q检验的显著性为0.905，远大于0.05，说明残差不存在自相关性，模型拟合良好，残差为白噪声。
- （3）误差指标：RMSE为0.053，MAPE为0.534%，MAE为0.041，各项误差指标均处于较低水平，表明模型预测精度较高，误差较小。

综上，ARIMA(0, 2, 0)模型在保持较高预测精度的前提下，具有较小的BIC值和理想的残差白噪声性质，模型简洁有效。

3.4 模型参数估计

表3-11 ARIMA(0, 2, 0)模型参数

估算 标准误差 t 显著性
ARIMA (0, 2, 0) 模型 ln_GDP 不转换 常量 -0.002 0.009 -0.277 0.783

表3-11为ARIMA模型的参数估计结果。模型拟合过程表明，经过二阶差分后的ln(GDP)序列已为白噪声，因此选择 $p=0$ 、 $d=2$ 、 $q=0$ 的模型形式。

由表3-11可知，常数项的估计值为-0.002，标准误为0.009， t 统计量为-0.277，对应的 p 值为0.783。该 p 值远大于常用的显著性水平0.05，说明常数项在统计上不显著，即二阶差分序列的均值与0无显著差异。这一结果与前期序列平稳性检验中“二阶差分后序列表现为白噪声”的结论一致。尽管常数项不显著，但其绝对值极小，保留该参数对模型预测结果的影响微乎其微，且SPSS软件在包含常数项时能够正常输出预测值，因此后续预测仍基于包含常数项的ARIMA(0, 2, 0)模型进行。

结合表3-9、表3-10可知，残差的LjungBox Q统计量为10.739， p 值为0.905，远大于0.05，表明残差序列不存在自相关，为白噪声，模型拟合充分。同时，模型的平稳 R 方接近0（ $-6.661E-16$ ）， R 方为0.999，RMSE为0.053，MAPE为0.534%，各项拟合优度指标均表明模型具有较高的预测精度。

故，

. (3-1)

其中，为广西壮族自治区第 t 年的地区生产总值（亿元）；

对进行二阶差分后得到平稳序列：

. (3-2)

ARIMA（0, 2, 0）模型的形式为：

. (3-3)

式中，为常数项，为白噪声序列。根据表3-9，常数项估计值=-0.002。由于该常数项的 p 值为0.783，在5%水平下不显著，因此模型实际等价于

. (3-4)

但与SPSS软件输出一致并便于后续预测，仍保留常数项，模型可写为：

. (3-5)

原始GDP序列的表达式为：

. (3-6)

3.5 模型的残差检验

3.5.1 残差的ACF和PACF图检验

为进一步验证模型的合理性，对ARIMA（0, 2, 0）模型进行残差检验。用SPSS绘制残差的ACF和PACF图进行观察。

从图3.10残差的ACF图可以看出，在延迟期数为1、2、3等各阶滞后位置上，自相关系数均非常接近于0，并且全部落在两倍标准误差（即95%置信区间）的范围之内。这一结果表明，残差序列中不存在显著的自相关性，即当前模型已经充分提取了时间序列中的动态依赖关系，拟合效果较为理想。

从图3.10残差的PACF图同样可以观察到，在延迟期数为1、2、3等各阶滞后位置上，偏自相关系数也均接近于0，且绝大多数位于两倍标准误差以内。尽管个别滞后阶数的偏自相关系数略微超出了置信区间，但超出幅度很小，在实际分析中通常可以忽略不计。这进一步印证了残差序列不存在显著自相关性的结论。

综上，该残差检验图显示，模型的残差序列未表现出显著的自相关性，即残差为白噪声序列。这说明所建立的模型已经成功地捕捉到了时间序列中的主要信息，残差中不再含有可识别的模式，模型拟合效果较好。为进一步验证模型的充分性，后续可结合Ljung-Box（LB）检验对残差进行统计诊断，以增强结论的可靠性。

图3.10 残差检验ACF和PACF图

3.5.2 LB检验

（1）使用SPSS对残差序列进行LB检验（原假设为残差在指定的滞后阶数内没有显著的自相关性，备择假设为残差在指定的滞后阶数内存在显著的自相关性），得到的结果如下表3-12：

表3-12 残差检验结果

检验统计量 自由度 P值

10.739 18 0.905

(2) 使用Python得出结果如表3-13:

表3-13 Python输出的残差结果

滞后阶数 检验统计量 P值 滞后阶数 检验统计量 P值

1 6.401 0.011 10 6.443 0.776

2 6.401 0.040 11 6.450 0.841

3 6.419 0.092 12 6.456 0.891

4 6.421 0.169 13 6.457 0.927

5 6.422 0.267 14 6.459 0.953

6 6.426 0.377 15 6.459 0.970

7 6.428 0.490 16 6.460 0.982

8 6.433 0.598 17 6.460 0.989

9 6.438 0.695 18 6.462 0.993

图3.11 Python输出的残差ACF和PACF图

为了检验ARIMA(0,2,0)模型残差的白噪声性，进行Ljung-Box检验。SPSS输出结果中（表3-10），滞后1~18阶的联合检验统计量为 $Q(18)=10.739$ ， $p=0.905>0.05$ ，在5%的显著性水平下不能拒绝原假设，表明残差在整体上不存在显著自相关，即残差表现为白噪声，模型通过诊断检验。

在Python输出的结果中（表3-11），各滞后阶数的单独检验结果发现，滞后1和2的p值分别为0.011和0.041，均小于0.05，且残差的ACF和PACF图（图3.11）中，滞后1和2的相关系数略超出95%置信区间，暗示残差可能存在微弱的短期自相关。

考虑到多阶联合检验能够有效控制累积第一类错误，且具有更高的统计功效，其结果比逐次检验更为可靠。个别低阶滞后的显著性可能是由抽样波动或有限样本特性所致，并未形成系统性模式，不足以推翻联合检验的结论。

综上，尽管残差在个别低阶滞后上存在轻微相关迹象，但联合检验结果表明残差整体上符合白噪声假定，ARIMA(0,2,0)模型能够较好地拟合数据，可用于后续分析与预测。

3.6 广西GDP预测

对模型的残差进行白噪声检验并得到证明之后，接下来对广西壮族自治区GDP进行预测，预测结果如表3-14:

表3-14 预测结果

年份 预测值(亿元) 上置信区间(亿元) 下置信区间(亿元) 同比(%)

2025 29871.51 26955.99 33102.37 4.10

2026 31027.97 24661.56 39037.88 3.87

2027 32157.41 21897.58 47224.36 3.64

2028 33253.73 18947.23 58362.65 3.41

2029 34310.83 16019.71 73486.52 3.18

2030 35322.67 13261.07 94086.76 2.95

图3.12 广西GDP预测

图3.13 广西GDP的动态预测图

模型结果显示，ARIMA(0,2,0)模型对1985—2020年广西GDP的历史数据具有优异的拟合效果。从图3.12可见，实际值与拟合值的曲线高度重合，表明模型有效捕捉了区域经济发展的长期趋势与波动特征。该模型通过二阶差分处理，成功消除了原始时间序列的非平稳性，使序列趋于稳定，为后续预测奠定了可靠基础。同时，残差白噪声检验通过，说明模型已充分提取序列中

的主要信息成分，未留下显著的可预测结构，进一步验证了模型设定的合理性与有效性。

在确保历史拟合精度的基础上，模型对2025—2030年广西GDP进行了外推预测。结果显示，未来六年经济总量将延续增长态势，呈平稳线性上升趋势。具体而言，预测值从2025年的29871.51亿元逐步增至2030年的35322.67亿元，累计增长约18.4%，年均复合增速约为3.4%。这一增长路径契合ARIMA(0,2,0)模型的线性外推特性，也与广西近年来依托西部陆海新通道、平陆运河等重大战略项目推动经济稳步发展的现实逻辑基本一致，反映出区域经济具备较强的内生增长韧性与持续发展能力。

值得注意的是，预测增速呈现逐年放缓的趋势，由2026年的3.87%逐步下降至2030年的2.95%。这一变化符合经济规模扩大后增速自然趋缓的一般规律，也反映出结构性转型期的增长压力。从动态预测机制看，模型在递推过程中引入了2025年广西实际GDP（29727.45亿元）作为基期值，使预测结果更贴近近期经济运行实际，因而略高于静态预测值，增强了预测的现实参考价值。

在区间估计方面，随着预测期的延长，95%置信区间的宽度显著扩大，反映出模型长期预测的不确定性逐步累积。这一现象提示，单纯依赖趋势外推存在局限，需结合外部环境变化进行动态修正。置信区间的上下限分别对应乐观与保守两种情景，为政策制定提供了重要的风险边界参考。例如，重大战略项目推进超预期，实际GDP有望突破乐观情景；外部冲击加剧，经济增长可能面临保守情景的压力测试。

从结构性因素看，预测期内增速放缓与广西产业结构转型升级密切相关。传统资源型产业占比逐步下降，而新兴产业尚处于培育阶段，短期内可能形成增长压力。然而，平陆运河通航、中国—东盟产业合作区建设等重大工程的产能释放虽有时滞，预计将在“十五五”中后期逐步显现对经济的拉动作用，成为新的增长引擎。

综上所述，本文提出以下政策建议：

（1）以人工智能赋能产业升级，培育新质生产力

广西2025年高技术制造业增加值同比增长23.7%，人工智能赋能效应持续释放。建议沿循“北上广研发+广西集成+东盟应用”发展路径，加快建设中国—东盟人工智能合作中心，深入实施“人工智能+”行动。重点推动汽车、机械、钢铁等传统优势产业与人工智能深度融合，支持柳州具身智能产业集群发展，培育壮大汽车用锂离子动力电池、服务机器人等高技术产品产能。完善人工智能政策体系和人才支撑，推动科技创新与产业创新深度融合，加快构建人工智能及新一代信息技术、有色金属、先进钢铁等制造业十大现代化支柱产业。

（2）高水平共建西部陆海新通道，释放平陆运河经济动能

平陆运河将于2026年建成通航，将深刻改变广西区域经济地理格局。建议抢抓运河通航战略机遇，加快平陆运河经济带规划建设，推动南宁从内陆首府向江海联动门户枢纽跃升。大力发展临港工业与向海经济，优化用海产业结构布局，谋划建设一批“港口+园区+新城”综合体。深化与北部湾港务集团、平陆运河集团等企业合作，完善江海联运体系，提升北部湾国际门户港集装箱吞吐能力，拓展面向东盟的航运网络。积极争取中央预算内投资、超长期特别国债等政策支持，强化重大项目要素保障。

（3）深化面向东盟开放合作，拓展跨境产业链空间

广西对东盟进出口连续26年保持最大贸易伙伴地位，2025年同比增长8%。建议充分发挥面向东盟开放合作前沿门户优势，深入实施自由贸易试验区提升战略，稳步扩大制度型开放。加快建设中国—东盟产业合作区协同发展区，推动跨境产业合作向价值链高端延伸。加强与粤港澳大湾区、长江经济带等重点区域联动发展，深化粤桂合作特别试验区建设，积极承接产业转移。拓展“东盟出产+广西集散+中国畅销”新链路，推动内外贸一体化发展，提升广西在区域产业链供应链中的地位。

（4）优化区域经济布局，促进城乡区域协调发展

广西县域经济发展仍不均衡，城镇化率58.09%仍有提升空间。建议深入实施县域经济高质量发展三年行动方案，统筹推进新型城镇化和乡村全面振兴。支持经济大市挑大梁，鼓励各市发挥比较优势、各展所长。扎实推进新时代兴边富民行动，推动边境地区高质量发展。强化国土空间规划和“三区三线”实施管理，促进城乡要素平等交换、双向流动。加大民生领域投入，持续将财政支出八成左右投向民生领域，补齐教育、医疗、养老等短板。

（5）统筹风险防范与动态调整，增强政策响应能力

预测结果显示置信区间随预测期延长而扩大，提示经济增长面临不确定性累积。建议建立健全宏观经济预测动态更新机制，根据每年实际GDP数据回测模型精度并适时调整预测。加强财源建设，兜牢基层“三保”底线。统筹推进房地产、地方政府债务、中小金融机构等风险化解，加强防风险与促发展政策协同。强化宏观政策取向一致性评估，发挥存量政策与增量政策集成效应，增强经济韧性和应对外部冲击能力。

总体而言，基于ARIMA(0, 2, 0)模型的预测结果不仅揭示了广西经济未来发展的基本趋势，也为政策制定提供了量化依据与风险预警。在复杂多变的内外部环境，唯有坚持动态评估、科学调整、精准施策，方能确保区域经济在高质量发展轨道上行稳致远。

结论

本文基于广西壮族自治区1985—2024年GDP年度数据，运用ARIMA模型对区域经济增长趋势进行系统分析与预测。通过自然对数变换消除异方差，经二阶差分处理后，ADF与KPSS单位根检验一致表明序列平稳，ACF与PACF图及Ljung-Box检验（ $Q(18)=10.739$, $p=0.905$ ）证实序列为白噪声，据此确定ARIMA(0, 2, 0)为最优模型。模型拟合优度较高（ $R^2=0.999$, $RMSE=0.053$, $MAPE=0.534\%$ ），残差检验通过，表明模型信息提取充分，能够较好地拟合广西GDP历史演进规律。

预测结果显示，2025—2030年广西GDP呈平稳增长态势，预测值由2025年的29871.51亿元增至2030年的35322.69亿元，年均增速约3.4%。从增速变化看，由2026年的3.87%逐步放缓至2030年的2.95%，这一趋势符合经济规模扩大后增速收敛的一般规律。从区间估计看，95%置信区间宽度随预测期延长显著扩大——上、下区间差值由2025年的6146.38亿元增至2030年的80825.69亿元，反映模型长期预测的不确定性累积效应。值得注意的是，预测结果与广西近期经济表现高度吻合：2025年广西实际GDP达29727.45亿元（同比增长5.1%），高技术制造业增长23.7%，人工智能赋能效应持续释放，印证了模型对区域经济增长态势的有效捕捉。

本研究在方法论层面验证了ARIMA模型对边疆民族地区GDP序列的适用性，通过严格遵循Box-Jenkins建模流程，系统展示从平稳性检验、模型识别到诊断检验的完整技术路线，为同类研究提供了可复制的分析范式。研究证实，对于呈现长期趋势但无显著周期波动的GDP序列，二阶差分结合白噪声特征的ARIMA(0, 2, 0)模型能够以简洁形式实现高精度拟合，为简约模型在宏观经济预测中的有效性提供了实证支持。

预测结果为广西“十五五”规划及年度经济目标制定提供了量化参考。2025—2030年GDP年均增速约3.4%的预期，与广西2026年政府工作报告提出的“5%左右增长目标”及“一区两地一园一通道”战略定位相衔接，有助于合理配置财政资源、优化产业布局。置信区间为政策制定提供了风险边界——上、下区间分别对应乐观与保守情景，提示在平陆运河建设、西部陆海新通道推进等重大项目执行中需建立动态调整机制，增强经济韧性和应对外部冲击的能力。同时，预测趋势与高技术制造业快速增长、人工智能赋能实体经济等积极因素相呼应，为培育新质生产力、深化面向东盟开放合作等政策方向提供了数据支撑。

需要指出的是，本研究存在一定局限性：样本区间仅覆盖1985—2024年共40期年度数据，样本量相对有限，可能影响模型对长期波动特征的捕捉能力；ARIMA模型仅依赖序列自身滞后信息，未纳入投资、消费、进出口、政策干预等外生变量，难以全面反映复杂经济系统的多因素驱动机制；预测区间随步长延长快速扩宽，长期预测精度下降，模型对结构性突变（如疫情、国际形势变化）的响应能力较弱；研究未对不同产业GDP进行分层建模，无法揭示产业结构变动对总量增长的具体贡献。

基于上述分析，未来研究可从以下方向深化：一是引入干预变量或虚拟变量，刻画重大政策（如RCEP实施、平陆运河通航）和外部冲击对经济增长的阶段性影响；二是构建组合预测模型，融合ARIMA与机器学习方法（如LSTM、BP神经网络），捕捉线性趋势与非线性特征的互补信息；三是拓展至产业层面，分别对三次产业GDP建立ARIMA模型，解析产业结构演进对总量增长的影响路径；四是纳入更多经济变量，构建VAR或状态空间模型，探究投资、消费、进出口等与GDP的动态关联；五是利用季度或月度数据开展高频预测，提升政策响应的时效性。通过上述改进，可进一步增强预测的科学性与政策参考价值，为广西经济高质量发展提供更为坚实的数据支撑。

参考文献

Box G E P, Jenkins G. Time Series Analysis: Forecasting and Control[M]. San Francisco: Holden-Day, 1976.

- Rahman A, Hasan M M. Modeling and Forecasting of Carbon Dioxide Emissions in Bangladesh Using ARIMA Models[J]. Open Journal of Statistics, 2017, 7(4): 560-566.
- Mirbagheri M. Fuzzy-logic and Neural Network Fuzzy Forecasting of Iran GDP Growth[J]. African Journal of Business Management, 2010, 4(6): 925-929.
- Luigi L, Massimo R, Armando R. A Neural Network Ensemble Approach for GDP Forecasting[J]. Journal of Economic Dynamics and Control, 2022, 134: 104278.
- Fornaro P, Luomaranta H. Nowcasting Finnish Real Economic Activity: A Machine Learning Approach[J]. Empirical Economics, 2020, 58(1): 55-71.
- Gomez V, Maravall A. Seasonal Adjustment and Signal Extraction in Economic Time Series[C] Proceedings of the Business and Economic Statistics Section, American Statistical Association, 2001:1-10.
- Kawasaki Y, Franses P H. Do Holiday Effects Remove Predictable Seasonality in GDP?[J]. Econometric Institute Report, 2004, 4(1): 1-15.
- Hyndman R J, Khandakar Y. Automatic Time Series Forecasting: The Forecast Package for R[J]. Journal of Statistical Software, 2008, 27(3): 1-22.
- 柳丽娴. ARIMA模型在中国GDP预测中的应用[J]. 商, 2016(22): 182.
- 周奎. ARIMA模型在我国GDP预测中的应用[J]. 广西职业技术学院学报, 2016, 9(1): 19-23.
- 华鹏. ARIMA模型在广东省GDP预测中的应用[J]. 统计与决策, 2010(12): 166-167.
- 张江城. 基于时间序列的我国GDP的短期预测[J]. 商, 2016(5): 204-206.
- 王国贤, 范英兵. 基于干预ARIMA模型的中国GDP趋势分析[J]. 河南建材, 2020(3): 33-35.
- 李振亮, 乐听雨. 基于ARIMA模型的北京市GDP分析与预测[J]. 中小企业管理与科技, 2023(1): 112-115.
- 陈全润. 基于季节ARIMA模型的中国季度GDP预测[J]. 统计研究, 2019, 36(2): 33-42.
- 刘展. 基于ARIMA与HP滤波的中国潜在GDP估算[J]. 数量经济技术经济研究, 2021, 38(5): 89-105.
- 查华, 石舫. 基于ARIMA模型对江苏省GDP的预测[J]. 兰州文理学院学报(自然科学版), 2022, 36(03): 33-36.
- 刘玥婷. 基于ARIMA模型的北京市GDP分析及预测研究[J]. 价值工程, 2022, 41(12): 31-33.
- 潘典雅. 基于ARIMA模型的吉林省GDP分析及预测[J]. 中国集体经济, 2021, (27): 15-16.
- 王佳佳. 基于ARIMA模型的安徽省GDP的分析与预测[J]. 枣庄学院学报, 2020, 37(05): 38-44.
- 胡咏琪. ARIMA模型在陕西省GDP预测中的应用研究[J]. 中国商论, 2021, (13): 174-176.
- 张梓. 基于ARIMA模型的贵州省GDP分析与预测[J]. 国土与自然资源研究, 2022, (05): 39-41.
- 肖丹. 基于ARIMA模型的四川省GDP分析与预测[J]. 生产力研究, 2023, (10): 62-66.
- 杨忠裕, 薛紫玥. 基于ARIMA模型的甘肃省GDP的分析与预测[J]. 中国市场, 2023, (06): 1-4.
- 田秀芹. 广西GDP预测与分析[J]. 现代商贸工业, 2017, (17): 24-25.
- 许阳干. 广西GDP的时间序列分析与预测模型[J]. 沿海企业与科技, 2010(07): 54-57.
- 周练. 基于ARIMA模型的“十三五”期间广西经济增长预测分析[J]. 商业经济研究, 2015, (15): 137-139.
- 曾怡霏, 石虹. ARIMA模型在广西壮族自治区GDP预测研究中的应用[J]. 时代经贸, 2025, 22(01): 129-133.
- 杨智凯, 范彦勤, 印海廷, 等. 基于ARIMA模型对桂林市GDP的预测研究[J]. 桂林航天工业学院学报, 2021, 26(04): 477-483.
- 张青山. 基于组合预测的广西GDP预测研究[D]. 广西师范大学, 2023.
- 刘嘉成. 基于ARIMA模型对江苏省总产出的预测[J]. 大众投资指南, 2025, (32): 127-129.
- Wu D. A Comparative Analysis of ARIMA Models for Forecasting China's GDP[J]. Advances in Operation Research and Production Management, 2025, 4(2): 1-11. DOI:10.54254/3029-0880/2025.26616.
- 骆俊蓉. ARIMA模型在兰州市GDP预测中的应用[J]. 中国电子商情, 2025, 31(11): 7-9.

- 马涛. 基于ARIMA模型和BP神经网络的河北省GDP预测研究[D]. 沈阳师范大学, 2025. DOI:10.27328/d.cnki.gshsc.2025.0003 10.
- 陈满丽, 张慧娟, 焦楠楠, 等. 基于MATLAB的ARIMA模型对我国GDP预测的研究[J]. 中国市场, 2024, (12):1-4. DOI:10.13939/j.cnki.zgsc.2024.12.001.
- 赵珂一. 基于ARIMA模型的云南省第三产业发展预测[J]. 价值工程, 2023, 42(33):44-48.
- 田美雪. ARIMA模型在山东省GDP预测中的应用[J]. 中国管理信息化, 2023, 26(01):139-141.
- 夏如玉, 王梓桥. 基于ARIMA模型对重庆市GDP预测分析[J]. 中国储运, 2022, (08):93-94. DOI:10.16301/j.cnki.cn12-1204/f.2022.08.044.
- 蓝广荣, 郭景福. 边境贸易促进广西经济增长效应的实证研究[J]. 广西职业技术学院学报, 2021, 14(04):60-66.
- 张鹏. 改进的ARIMA-GM-SVR组合预测模型及应用[J]. 统计与决策, 2019, 35(13):82-84. DOI:10.13546/j.cnki.tjyjc.2019.13.019.
- 魏新宇, 李旭芳. 基于ARIMA模型的浙江省GDP预测模型研究[J]. 建模与仿真, 2023, 12(6):5696-5705.
- Tong Y, Nie J, Cheng X. Guangxi GDP Prediction Model Based on Principal Component Analysis and SSA-SVM[J]. Computational Economics, 2025, 66(2):1191-1213.
- Khoso W, Daudpoto M R. Forecasting GDP Growth Using ARIMA: Long-term Projections for China, India, and the USA (2023 - 2045) [J]. African Research Reports, 2025, 1(4):155-166.
- Peng M, Liu Z. Forecasting Analysis of Guangzhou's Gross Domestic Product Based on ARIMA Model[C] Proceedings of the 3rd International Conference on Management Science and Software Engineering (ICMSSE 2023). Atlantis Press, 2023:376-383.
- Lin W, Shi Y. A Study on the Development of China's Financial Leasing Industry Based on Principal Component Analysis and ARIMA Model[J]. Sustainability, 2023, 15(13): 9913.
- 王燕. 应用时间序列分析[M]. 北京:中国人民大学出版社, 2022.
- 翟静, 曹俊. 基于时间序列ARIMA与BP神经网络的组合预测模型[J]. 统计与决策, 2016(4):29-32.
- 徐卫泽. 预测经济和金融时间序列: ARIMA与LSTM模型的比较[J]. 山西农经, 2020(3):36-37.
- 谭梓怡, 何文. 基于ARIMA模型的上海市GDP分析与预测[J]. 理论数学, 2024, 14(1): 919-928.
- 陈诚, 徐加成, 蒋文. 基于R语言的我国GDP的建模及其预测[J]. 理论数学, 2025, 15(10):2624-2636.

致谢

旧游无处不堪寻，无寻处，唯有少年心。

行文至此落笔为终，始于初秋终于盛夏。总以为来日方长，殊不知时光匆匆，当敲下论文最后一个句点，来到了这篇论文的终篇——致谢，同时也意味着我的大学生涯接近尾声，是结束，也是开始。

四年的上海电机之行，无论是喜悦还是悲伤，所有经历于我而言都是馈赠，所有相遇于我而言都是幸运。在此，我想感谢每一位曾经给予我支持和帮助的人，让我感受大学时光里难忘的瞬间。

过往如云烟，恩情最难忘。感谢我的父母，虽然他们很平凡，但是他们把最好的爱与所有全都给予了我。他们没有太高的文化水平，但却一直支持我去看更好的风景，永远在背后支持我的每个选择，是我最坚强的后盾。祝愿我的家人们身体健康，万事顺遂。

教诲如春风，师恩似海深。感谢我的导师孙丽瑛老师，在她的指导下，论文的选题以及写作中都给予了我悉心的指导。同时，我还要感谢大学期间的所有专业课老师以及辅导员老师，你们的谆谆教诲令我受益匪浅，祝愿老师们万事胜意，平安喜乐。

山河不足重，重在遇知己。感谢我的每一位朋友，在青春的道路上因为你们我才不会孤独，愿我们都能在自己的领域闪闪发光，成为更好的自己。愿我们在未来的生活中，砥砺前行，不被现实磨灭棱角，将理想照进现实，未来可期。同时你们也希望遇到困难时不缺乏再来的勇气。祝愿你们平安喜乐，岁月无忧。

道阻且长，行则将至。最后我想感谢平凡的自己，虽然平凡，但不甘于平庸，纵使生活中有诸多的不如意，但我仍保持着乐观善良，希望未来的自己能永怀赤城，勇敢且自信，保持自己的热爱，坚定不移的走下去。怀着理想走出自己的人生。

千言万语终觉词不达意，感谢经历，感谢相遇。山水相逢，终有一别，落笔于此，行而不止。将热烈怀揣继续前行，续写人生新的篇章。

谨以此文，献给我的二十三岁，献给这段热烈而真挚的青春。

说明：

- 1、支持中、英文内容检测；
- 2、AI特征值=AI特征字符数/总字符数；
- 3、红色代表AI特征显著部分，计入AI特征字符数；
- 4、棕色代表AI特征疑似部分，未计入AI特征字符数；
- 5、检测结果仅供参考，最终判定是否存在学术不端行为时，需结合人工复核、机构审查以及具体学术政策的综合应用进行审慎判断。



关注微信公众号