

参赛编号：202406900554

## 2024 年第五届“大湾区杯”粤港澳 金融数学建模竞赛

题目：基于 PSTR 非线性回归模型的粤港澳大湾区经济预测与策略分析

### 摘要

在我国高质量发展的时代洪流下，粤港澳大湾区作为具有独特区位优势的新兴经济体，正在全国乃至世界范围内推动经济贸易的繁荣与发展。本文聚焦粤港澳大湾区经济发展的影响因子，将回归模型应用于大湾区的经济发展预测任务，力图寻找大湾区的最佳发展路径，实现未来的可持续发展。

针对任务一，本文从历史、人口、科技、产业、能源环境、经济状况和市场贸易等方面收集了 48 个细分指标，将年度数据采用三次样条插值法扩充为季度数据。接着对预处理后的数据进行因子分析，逐步检验因子与模型的适配度，实现特征因子的选择。由于后续需建立的时间序列回归模型，因此对数据进行相关性分析与平稳性检验，从而剔除 8 个与模型要求不符的特征因子，再对剩余的 40 个因子分级进行主成分分析，最终将所得 10 个主成分因子放入多元回归的动态面板模型进行显著性检验。本文将经济增长滞后期纳入了动态面板模型，从而提取出反映粤港澳大湾区经济动态发展模式的五大主成分因子，分别为教育规模因子、科技创新因子、产业产值因子、自然资源因子和经济状况因子。

针对任务二，基于任务一中挑选出的主成分因子，本文首先进行非线性检验，结果表明非线性回归模型中仅含有一个转移函数，接着通过面板平滑迁移函数的非线性作用对动态模型进行改写，采用 PSTR 模型估计转移函数的参数，得到非线性回归方程。随后采用 ARIMA (2, 1, 1) 预测五大主成分因子与核心解释变量未来 10 年的数值，代入非线性回归方程后，发现粤港澳大湾区的经济将随着核心解释变量的增长而持续发展，直至 2028 年面临门槛效应。因此本文提出“创新驱动，绿色转型，物价稳控”的大湾区经济发展三向协同策略，为粤港澳大湾区的可持续发展带来源泉动力。最后，本文通过替换解释变量对该非线性模型进行了稳健性分析，各解释变量与控制变量均在 5% 的水平下显著。

针对任务三，本文收集了旧金山、东京、纽约三大湾区在人口、科创、产业、环境、贸易这五大方面的相关特征指标，在经过一系列的因子分析与检验后，代入任务二中建立的非线性回归模型中，将所得结果标准化消除量纲后与粤港澳大湾区进行多方位对比。通过绘制四大湾区的人均 GDP 预测涨势图，宏观上分析粤港澳大湾区经济发展形势一片向好，将于 2025 年基本与东京湾区的人均 GDP 持平，2030 年看齐纽约湾区；通过绘制五大特征因子的雷达图，微观上分析粤港澳大湾区在不同方面与其余三大湾区之间的异同。

针对任务四，本文从建模依据、经济预测、发展建议三个方面进行详细分析，为粤港澳大湾区未来 10 年的经济发展提供科学、精确的决策。科技创新的持续驱动，第三产业的不断扩张，绿水青山的长效保育……粤港澳大湾区这一世界级城市群新星必将在科学的经济发展进程中冉冉升起。

关键词：因子分析、动态面板模型、PSTR 非线性回归、ARIMA

# 目录

1 绪论	1
1.1 问题背景	1
1.2 问题重述	1
1.3 文献综述	1
1.4 解决方案	2
2 模型假设	3
3 符号说明	3
4 数据收集与预处理	3
4.1 数据来源	3
4.2 数据可视化	3
4.3 数据预处理	5
5 任务一 湾区经济发展指标体系的构建	5
5.1 确定基础指标体系	5
5.2 特征因子关系的分析与检验	7
5.3 因子分析特征变换	9
5.4 动态面板特征选择	11
5.5 结果分析	13
6 任务二 粤港澳大湾区经济发展预测模型	14
6.1 非线性检验	14
6.2 PSTR 模型估计	15
6.3 运用 ARIMA 时间序列预测	17
6.4 通过因子预测制定大湾区宏观发展策略	18
6.5 稳健性分析	19
7 任务三 四大湾区发展异同分析	20
7.1 模型复现	20
7.2 湾区经济发展水平差异分析	22
8 任务四 粤港澳大湾区经济发展分析与展望简报	24
9 总结与展望	26
9.1 优势	26
9.2 不足	26
9.3 展望	26
参考文献	27
附录	28

# 1 绪论

## 1.1 问题背景

2014年，深圳市政府工作报告中首次提出了“发展湾区经济”的概念，表示要以“湾区经济”新发展构建对外开放新格局，加快推进粤港澳大湾区合作。在我国追求创新驱动发展，加快转变经济发展方式的当下，粤港澳大湾区已成为带动全球经济发展的重要增长极和引领技术变革的领头羊。然而，面对全球化和技术变革带来的新机遇与挑战，粤港澳大湾区如何在未来保持持续、稳定的经济增长，成为亟需研究的重要课题。湾区内部各城市间的合作与竞争关系、金融与产业的深度融合、科技创新与产业升级的路径选择、生态环境保护与可持续发展等问题，都是影响粤港澳大湾区未来经济发展的关键因素。

因此，本文旨在深入研究粤港澳大湾区的未来经济发展，探讨其在全球化背景下的战略定位、发展模式、产业升级路径以及区域合作机制等。通过综合分析和比较，提出具有前瞻性的政策建议，为粤港澳大湾区的经济社会发展提供科学依据与高质量发展策略与方案。

## 1.2 问题重述

任务一：粤港澳大湾区作为区域经济系统，其发展受多种因素共同影响。通过历史数据分析，我们确认人口、科技、金融等因素在塑造区域经中扮演关键角色。为量化这些因素的影响，我们将采用多元回归分析方法，重点研究人口结构、科技创新、金融规模等因素在未来10年对大湾区经济走势的潜在影响。

任务二：基于任务一的特征选择，构建回归模型预测未来10年粤港澳大湾区经济走势。根据预测结果设计具体可行的策略方案，以应对可能出现的经济挑战，促进区域快速发展。

任务三：为全面了解湾区经济发展，选取其他湾区进行比较分析。运用任务一中的模型，预测所选湾区未来10年的经济走势，并量化分析不同湾区之间的发展异同。

任务四：结合前三项任务的分析和预测结果，撰写1000字左右的简报，为决策部门提供有针对性的建议应对经济变化，促进湾区经济持续健康发展。

这些任务旨在研究粤港澳大湾区未来的经济发展并提出相应发展决策建议。

## 1.3 文献综述

近年来，学者们针对湾区经济的发展影响因素与发展时期差异进行了广泛而深入的研究。随着全球化和城市化的发展，湾区经济成为经济发展的主要载体，凭借其独特的优势，对全球发展具有重要意义。李立勋（2017）[1]指出，湾区在推动全球经济发展中具有重要的增长极地位，是科技创新的领导者。粤港澳大湾区重要的不是“湾区”，而是“粤港澳”，强调粤港澳合作关系的扩大和深化，维护港澳经济繁荣稳定，增强粤港澳在国民经济发展和开放中的地位。颜竹晗（2024）[2]在数字普惠金融与区域韧性性的研究中添加交互项来衡量各指标之间对经济发展的协同和抑制作用；马茹等学者（2019）对中国区域经济高质量发展评

价指标体系进行了完整的构建，从高质量供给、高质量需求、发展效率、经济运行和对外开放共五个方面对经济发展因素分层[3]。而相关研究的大部分学者对于不同区域所测度数据的异质性分析不足，针对该问题，本文将综合粤港澳大湾区“9+2”结构中城市间的数据异同来量化评价各因素对经济发展的影响，实现对大湾区总体经济发展状况的预测。

## 1.4 解决方案

在上述问题背景、问题重述和文献综述的基础之上，我们对收集到的数据进行相关性分析与平稳性检验，剔除一部分不符合模型要求的指标；接着进行主成分分析，将所得主成分因子投入动态面板回归模型中进行显著性检验，得到五大主成分因子作为建模指标体系；利用 ARIMA 预测未来 10 年粤港澳大湾区的五大主成分因子，代入 PSTR 非线性回归方程，实现对湾区人均 GDP 的预测。总流程图如下图 1 所示：



图 1 总流程图

## 2 模型假设

1. 假设本文数据来源均可靠有效并具有连续性；
2. 假设本文所研究的四大湾区在未来 10 年不会出现突发性灾难；
3. 假设 2017 年后因外界因素引起的变化均由粤港澳大湾区的设立而产生；
4. 假设满足使用三次样条插值法的条件。

## 3 符号说明

符号	说明
$F_i$	$i = 1, 2, \dots, 6$ , 分别表示人口、历史、科技、环境、经济状况、市场贸易因子
$F_{ij}$	$i = 1, 2, \dots, 6, j = 1, 2$ , 表示第 $i$ 大类下第 $j$ 个特征因子的转移函数估计参数, 分别表示教育规模、科技创新、产业产值、自然资源、经济状况因子的转移函数估计参数
$GDP_{t-1}$	去年人均 GDP
$GDP_t$	第 $t$ 年人均 GDP
$CPI_t$	第 $t$ 年居民消费价格指数
$\gamma$	解释变量转移参数

## 4 数据收集与预处理

### 4.1 数据来源

鉴于题目未直接给出所需数据, 为确保模型构建具备坚实的数据基础, 我们对各类数据来源进行了广泛的梳理与筛选, 最终锁定了一系列能够反映研究对象特征的数据资源。以下是对主要数据来源的汇总展示, 如下所示:

表 1 主要数据来源

数据集	数据来源
粤港澳大湾区广东城市数据	<a href="https://gdzd.stats.gov.cn/">https://gdzd.stats.gov.cn/</a>
香港数据	<a href="https://www.censtatd.gov.hk/sc/">https://www.censtatd.gov.hk/sc/</a>
澳门数据	<a href="https://www.dsec.gov.mo/zh-CN/Statistic/Database">https://www.dsec.gov.mo/zh-CN/Statistic/Database</a>
东京数据	<a href="https://www.toukei.metro.tokyo.lg.jp/">https://www.toukei.metro.tokyo.lg.jp/</a>
纽约、旧金山湾区数据	<a href="https://www.statista.com/">https://www.statista.com/</a>

### 4.2 数据可视化

2017 年 3 月, 李克强总理在十二届全国人大五次会议上作政府工作报告时, 明确提出“要推动内地与港澳深化合作, 研究制定粤港澳大湾区城市群发展规划”; 同年 7 月, 习近平

总书记在香港亲自出席《深化粤港澳合作推进大湾区建设框架协议》签署仪式；2019年2月，中共中央国务院印发《粤港澳大湾区发展规划纲要》。粤港澳大湾区近15年来的经济发展进程广受海内外人民的关注，为全面且深入地评估其经济发展状况以及大湾区成立后实现经济飞跃的关键节点，本文选取了2010—2023年的大湾区数据展开分析，旨在捕捉大湾区成立前后的经济动态，为评估其政策效果和发展趋势提供充足的时间序列数据。为直观分析粤港澳大湾区在这段时间内的经济发展轨迹，在本节中选取若干重要指标绘制涨势图如下：

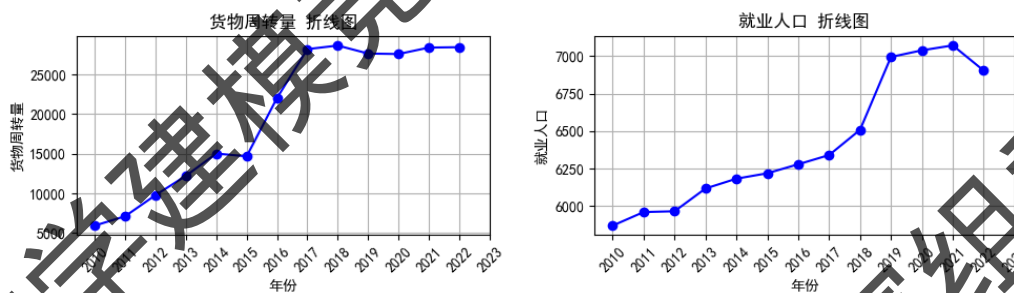


图2 货物周转量与就业人口

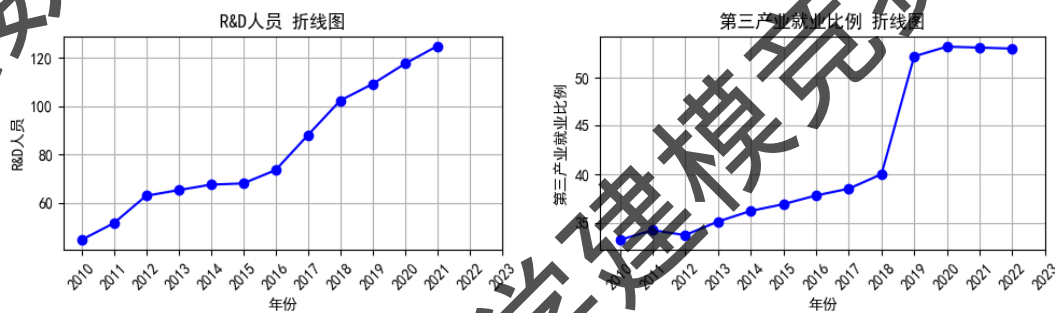


图3 R&D 人员与第三产业就业比例

由图可发现，在2017—2018年，货物周转量、就业人口以及R&D人员数量均出现了显著增长，表明大湾区的成立对区域经济发展产生了积极的推动作用。为了进一步了解产业就业比例，绘制第一、二、三产业就业比例的堆叠图如下：

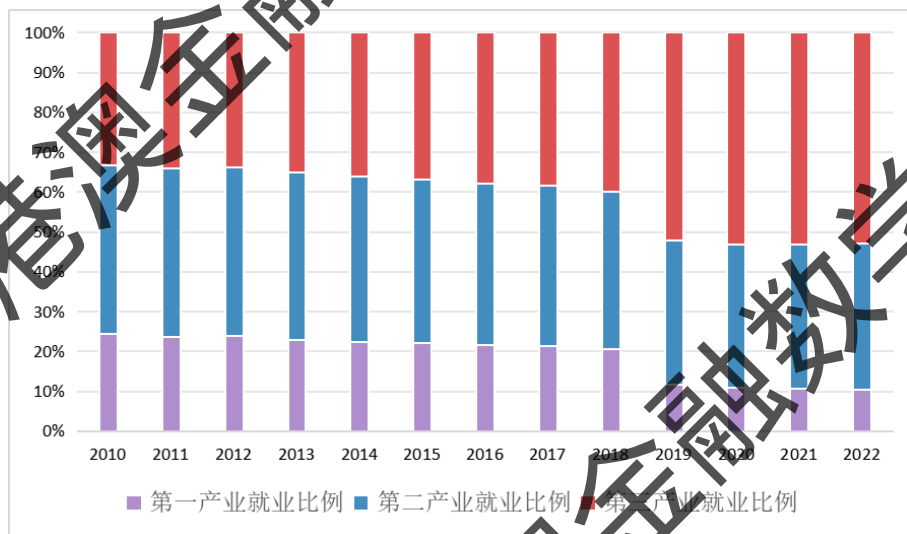


图4 第一、二、三产业就业比例堆叠图

在 2018 年，第三产业就业比例显著提升，第一、二产业就业比例的相应减少。这一变化趋势与大湾区成立后的发展战略紧密相关。2017 年大湾区概念提出后，政府积极推动产业转型升级，加大了对科技与高新技术产业的投资力度，一方面促进了传统产业的技术改造与升级，使得部分劳动力从第一、二产业向第三产业转移；另一方面新兴的高新技术产业迅速崛起，创造了大量的就业机会，吸引了众多高素质人才的流入，进一步推动了第三产业的发展，从而实现了产业结构的优化调整。本文将基于以上数据特征构建经济发展预测模型，为大湾区的未来发展提供更具前瞻性的决策建议。

### 4.3 数据预处理

首先，对广东统计年鉴数据集进行初筛：对于与大湾区经济发展直接或间接相关的变量，如产业生产总值等进行保留；对于与经济发展无关或缺失严重的变量，如行政区划、土地面积等进行剔除；对于直接相关变量的子变量，如第一产业生产总值子类下的子变量林业生产总值等也进行剔除。经过初步整理筛选后，得到与大湾区经济相关的 48 个特征变量。

由于数据没有被记录或发布，因此本文采用多项式插值进行缺失年份数据填充。随后，本文对每列数据进行标准化处理，以消除量纲影响。另外，考虑到样本数量较小且自变量较多的情况，会对后续模型拟合造成困难，即求解回归方程时方程个数大于未知数个数形成欠定方程组，本文采用三次样条插值法，将以年为单位的数据集扩大到以季度为单位，如下图：

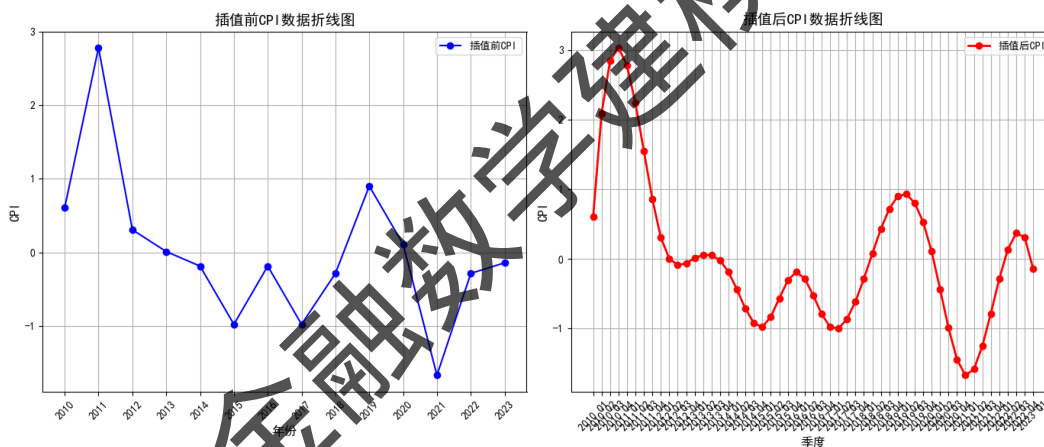


图 5 三次样条插值前后数据对比图

观察到插值后的数据符合原数据点的趋势走向，说明是有效扩大样本数据量，为后续回归和预测分析奠定基础。

## 5 任务一 湾区经济发展指标体系的构建

### 5.1 确定基础指标体系

**被解释变量 $y$ :** 本文选取粤港澳大湾区的人均 GDP 作为地区经济增长的度量指标，该指标同时兼顾了粤港澳大湾区地区经济增长的规模和效率。

**核心解释变量 $x$ :** 本文的核心解释变量为居民消费价格指数 (CPI)。从供给侧来看，居

民作为生产要素影响经济增长，从需求侧来看，作用通过消费影响经济发展。

**控制变量：**按照题目要求，从粤港澳大湾区的历史、人口、科技、产业、经济状况、环境、市场交易共 7 个方面收集数据，初步构建影响湾区经济发展的指标体系如下表：

表 2 湾区经济发展指标数据表

公共因子	主要特征变量	公共因子	主要特征变量
人口	总人口数 在校学生数 教育经费 高等教育毛入学率 就业人口 儿童人口比例 青壮年人口比例 老年人口比例	环境	碳排放 森林覆盖率 森林蓄积量 城市污水排放量 能源消费总量 第一产业能源消费总量 第二产业能源消费总量 第三产业能源消费总量 居民能源消费量 单位 GDP 能耗上升或下降
			经济状况
人文历史	文化遗产传承 文化艺术精品创作 文化产业体系 文化软实力提升 文化交流互鉴	市场贸易	商品交易市场数 小商品市场成交额 股票流通市值 总出口额 总进口额
科技	高技术产业企业数 高技术产业利润总额 专利授权量 R&D 人员 R&D 经费支出		
产业	第一产业生产总值 第二产业生产总值 第三产业生产总值 第一产业就业比例 第二产业就业比例 第三产业就业比例		

其中，人文历史类指标较为特殊。粤港澳大湾区作为国家重大战略引领下的世界级城市群汇聚要地，拥有雄厚的经济基础、悠久的人文历史和“一国两制”的制度优势，其人文历史对经济发展的影响无法直接找到对应的数值型指标进行衡量。因此，为全面分析粤港澳大湾区历史文化融合、文化产业发展与对外文化交流对经济发展的影响，以 2010—2023 年为时间跨度，搜集了 35 篇国家级、省级、市级和区级等不同级别政府部门发布的相关政策文件以及粤港澳大湾区具有影响力的两家综合性媒体《羊城晚报》《南方日报》的 87 篇媒体报道文本，构建文本关键词体系如下表：

表 3 人文历史文本关键词

人文历史二级指标	细分领域	关键词
历史文化融合	文化遗产传承	普查、保护、传承、弘扬、技艺、非遗、民俗、古迹
	文化艺术精品创作	扶持、人才、文艺、艺术、繁荣、作品质量、推广
历史文化产业发展	文化产业体系	布局、高质发展、科技、创新驱动、产业链条、消费
	特色旅游产品开发	特色产品、文旅融合、新型业态、度假区、体验
	文化软实力提升	软实力、挖掘、影响、输出、文化自信、认同感
历史文化交流与合作	文化交流互鉴	“一带一路”、多元、交流互鉴、民心相通、共享
	区域协同发展	协同规划、整体、政策、功能互补、区域合作
	开放合作机制	合作机制、开放、环境、互利共赢、模式、优势

基于该关键词体系的粤港澳大湾区人文历史发展指数 $I$ 的计算公式如下：

$$I = \sum (TF(t_i) \times IDF(t_i) \times f_i) \quad (1)$$

其中， $TF(t_i)$ 是关键词 $t_i$ 在特定文档中的出现频率， $IDF(t_i)$ 是逆文档频率，用来减少在多数文档中普遍出现的词的权重； $TF(t_i) \times IDF(t_i)$ 即为第 $i$ 个关键词的权重 $w_i$ ； $f_i$ 表示该关键词在文本中出现的频率。

## 5.2 特征因子关系的分析与检验

### 5.2.1 相关性分析

为初步了解各一级指标下二级指标间的大致关联，本文先对粤港澳大湾区 2010 年至 2023 年的人口、科技、产业、经济状况、环境、运输和市场贸易共 8 个指标下的细分数据进行相关性分析，观察各细分因子与人均 GDP 的相关性，若相关系数较低，则说明该指标与人均 GDP 在实际经济意义上的相关度较低，应将其剔除指标体系。

由于篇幅有限，此处仅列举“市场贸易”指标下的因子相关性分析。由表 4：市场贸易各相关指标与人均 GDP 之间均有着较强的相关性，但小商品市场成交额与人均 GDP 的相关系数低于 0.50，因此将该指标剔除经济发展指标体系。在其他大类指标下，我们也发现了一些与人均 GDP 相关度较低的指标，包括森林覆盖率、旅客周转量、小商品市场成交额与第一产业能源消费总量共 4 个指标，剔除后湾区经济发展的指标体系剩余 44 个指标。

表 4 粤港澳大湾区市场贸易产业相关系数

相关系数	人均 GDP	商品交易市场数	小商品市场成交额	股票流通市值	总出口额	总进口额
人均 GDP	1					
商品交易市场数	-0.987	1				
小商品市场成交额	0.181	-0.127	1			
股票流通市值	0.969	-0.974	0.201	1		
总出口额	0.937	-0.945	0.011	0.911	1	
总进口额	-0.758	0.694	0.012	-0.732	-0.518	1

### 5.2.2 平稳性检验

一般来说，只有满足数据序列均值与时间无关联和数据序列拥有无时变性的有限方程两个条件，才能够证明该数据序列具有平稳性。否则，即使数据序列之间无任何关系，也可能因为某种强劲的趋势而反映出一定的关系，这样的相关即称为虚假相关，产生“伪回归”现象，导致数据序列分析出现错误[4]。因此，为保证后续回归分析的稳健型，本节将对各大类指标下的细分指标通过 ADF 检验进行平稳性检验，并删去非平稳的指标序列。

由于指标较多，因此此处仅以“科研”这一大类指标下的细分指标为例，ADF 检验结果如下表所示：

表 5 ADF 检验结果

变量	延迟阶数	模型结构	P 值	平稳性
高技术产业 企业数	0	$y_t = \alpha + \delta t + \mu_t$	0.9924	非平稳
	1	$y_t = \phi_1 x_{t-1} + \mu_t$	0.7648	非平稳
	2	$y_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \mu_t$	0.1306	非平稳
高技术产业 利润总额	0	$y_t = \alpha + \delta t + \mu_t$	0.3114	非平稳
	1	$y_t = \phi_1 x_{t-1} + \mu_t$	0.0206	平稳
专利授权量	0	$y_t = \alpha + \delta t + \mu_t$	0.9547	非平稳
	1	$y_t = \phi_1 x_{t-1} + \mu_t$	0.0183	平稳
R&D 人员	0	$y_t = \alpha + \mu_t$	0.9979	非平稳
	1	$y_t = \phi_1 x_{t-1} + \mu_t$	0.4144	非平稳
	2	$y_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \mu_t$	0.1846	非平稳
R&D 经费支出	0	$y_t = \alpha + \delta t + \mu_t$	0.9547	非平稳
	1	$y_t = \alpha + \delta t + \phi_1 x_{t-1} + \mu_t$	0.0183	平稳

从检验结果可以得出，高技术产业企业数、高技术产业利润总额、专利授权量、R&D 人员数、R&D 经费支出在 5% 的显著性水平下均存在单位根，均为非平稳时间序列，P 值均大于 0.05；经过一阶差分后，高技术产业利润总额、专利授权量、R&D 经费支出在 5% 的显著性水平下不存在单位根，为平稳时间序列，P 值均小于 0.05；但是高技术产业企业数与 R&D 人员在 5% 的显著性水平下仍存在单位根，为非平稳序列。因此可以得到高技术产业利润总额、专利授权量、R&D 经费支出三个变量为一阶平稳时间序列，即满足同阶单整，可用于后续回归分析，而二阶差分后仍不满足非平稳时间序列的高技术产业企业数和 R&D 人员则不能用于后续回归分析，此处将其剔除指标体系。

对剔除相关性较低的 44 个二级指标作平稳性检验后，发现高技术产业企业数、R&D 人员数、第三产业生产总值、文化遗产传承这 4 个指标数据为非平稳序列，因此将其剔除指标体系，剩余指标共 40 个。

### 5.3 因子分析特征变换

因子分析是一种研究因素之间相互影响并处理的统计手段。因为经济发展的各指标之间具有某种程度上的关联性和交集性，因子分析采用降维手段把很多有强关联性的指标总结成了可以覆盖大多数经济发展的历史、人口、科技、产业、经济状况、环境、市场交易等信息的少数几个弱相关性的公因子。除此之外，每个公因子所占的比例都是通过分析与计算得到，避免了主观因素造成的错误。

从粤港澳大湾区的历史、人口、科技、产业、环境、市场贸易等信息对其进行评估，产生影响的指标多且关系复杂。因此先基于因子分析对指标进行降维，具体过程如下：

**Step1** 首先对原始数据进行标准化处理，采用 Min-Max 标准化方法：

$$X = \frac{X_{ij} - m_{ij}}{M_{ij} - m_{ij}} \quad (2)$$

其中  $M_{ij}$  为  $X$  的最大值， $m_{ij}$  为  $X$  的最小值。

**Step2** 进行适应性检验。本文利用巴特利特球度检验和 KMO 检验判断选区的指标数据能否进行因子分析，由此产生的结果如表 6（仅展示人口、产业、历史大类）。

表 6 KMO 检验和巴特利特检验

大类指标	KMO 检验值	巴特利特球度检验		
		近似卡方	自由度	显著性
人口	0.726	9491.265	28	0.000
产业	0.851	2550.760	10	0.000
历史	0.367	441.087	6	0.000

在 KMO 检验中，若其检验值低于 0.5，则表明其不适合进行主成分因子分析。由上表可见，历史大类指标的 KMO 检验值为 0.367，不适宜进行下一步的因子分析，因此将其剔除。

**Step3** 确定公共因子数。本文采用通过计算累计方差贡献率大小来确定因子个数，以碎石图将其可视化，其中累计方差贡献率体现了因子对原始变量的信息保留度，人口与产业的碎石图与总方差解释结果如下图：

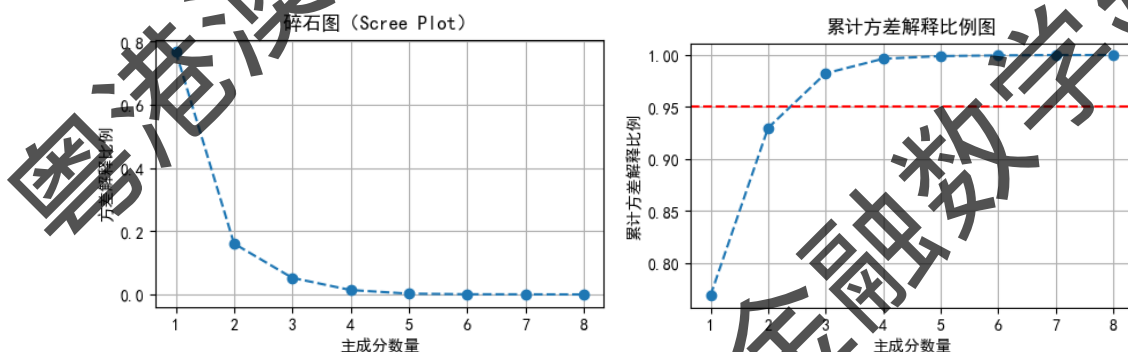


图 6 人口指标主成分个数检验

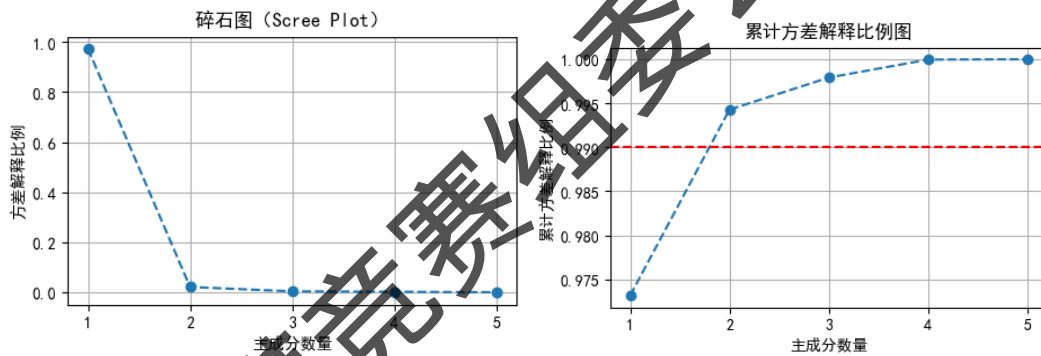


图7 产业指标主成分个数检验

将总方差解释表中的信息作为分析基础，将具有一个以上特征的因素视为公共因子，并以方差百分数来表达其所代表的信息量，所提取的公共因子的累计信息保存度应超过85%，这时说明可以用这些公共因子表示原始数据的大部分信息，而忽略剩余的不重要的方差信息。

我们对每个大类指标下的细分因子都进行了个数检验，建议保留1个主成分因子的大类指标有：科技、市场贸易，其余大类指标均建议保留1个主成分因子，共保留了10个主成分因子进入后续回归模型，这些公共因子均可以代表该大类指标数据信息的99%以上。

**Step4 公共因子解释。**本文采用最大方差正交化轮换的方法，轮换后的因素分量矩阵是各个变量和公共因子之间的相关系数，绝对值越大，二者之间的关联度就越高，表明这个原始变量对构造公共因子的贡献度。人口、环境大类指标的主成分因素分量矩阵如下：

表7 旋转后的成分矩阵

人口大类指标	成分		环境大类指标	成分	
	$F_{11}$	$F_{12}$		$F_{51}$	$F_{52}$
总人口数	0.379	-0.217	碳排放	0.356	0.050
在校学生数	0.086	0.831	森林蓄积量	0.303	-0.475
教育经费	0.046	0.399	城市污水排放量	0.363	-0.046
高等教育毛入学率	0.110	0.396	能源消费总量	0.366	-0.072
就业人口	0.389	-0.020	第一产业能源消费总量	0.245	0.619
儿童人口比例	0.332	0.284	第二产业能源消费总量	0.364	0.002
青壮年人口比例	-0.388	-0.034	第三产业能源消费总量	0.350	-0.174
老年人口比例	0.347	-0.370	居民能源消费量	0.356	-0.207
—	—	—	单位GDP能耗上升或下降	0.272	0.555

根据表格所示，旋转后的因子成分矩阵中各指标载荷值大小得出：主成分因子 $F_{11}$ 在总人口数、就业人口、儿童人口比例以及老年人口比例上均有较大的正载荷，分别为0.3793、0.3985、0.3324和0.3472，但在校学生数的载荷相对较小0.0860，因此将 $F_{11}$ 命名为人口结构因子。主成分因子 $F_{12}$ 的在校学生数、教育经费和高等教育毛入学率上具有极高的正载荷0.8306、0.399、0.396，因此将 $F_{12}$ 命名为教育规模因子。

主成分因子 $F_{51}$ 在多个与能源消耗相关的指标上均有较大的正载荷,包括碳排放(0.3564)、森林蓄积量(0.3031)、城市污水排放量(0.3626)、能源消费总量(0.3656),因此将主成分 $F_{51}$ 命名为**能源效率因子**。而主成分因子 $F_{52}$ 表现出与自然资源较强的相关性:在第一产业能源消费总量(0.6191)和单位 GDP 能耗上升或下降(0.5547)上具有较大的正载荷,同时与森林蓄积量(-0.4754)呈较大的负载荷,因此将 $F_{52}$ 命名为**自然资源因子**。

公共因子具体命名与主要特征对应情况如下表所示:

表 8 公共因子对应的特征变量

公共因子	名称	主要特征变量
$F_{11}$	人口结构因子	总人口数、就业人口、儿童人口比例、老年人口比例
$F_{12}$	教育规模因子	在校学生数、教育经费、高等教育毛入学率
$F_{31}$	科技创新因子	高技术产业利润总额、专利授权量和 R&D 经费支出
$F_{41}$	产业转型因子	第一、第二、第三产业就业比例
$F_{42}$	产业产值因子	第二、第三产业生产总值
$F_{51}$	能源效率因子	碳排放、森林蓄积量、城市污水排放量、能源消费总量
$F_{52}$	自然资源因子	第一产业能源消费总量、单位 GDP 能耗上升或下降
$F_{61}$	城市运输因子	城市公共交通工具运营数、港口货物吞吐量
$F_{62}$	经济状况因子	财政收入、财政支出、城镇、非城镇居民人均可支配收入
$F_{71}$	市场外贸因子	商品交易市场数、股票流通市值、总出口额、总进口额

## 5.4 动态面板特征选择

为预测未来 5—10 年的粤港澳大湾区的经济走势,我们计划建立回归模型来完成这一预测任务。但回归分析包括线性、多元线性和非线性等多种模型,如何选择匹配任务的数据和模型是建模的挑战。在建模之前,我们也需要借助显著性检验来筛选用于回归的特征因子。

回归模型根据数据类型的不同,也对应分为截面数据的回归模型、时间序列数据的回归模型和面板模型。对于面板模型而言,其可以均衡的反应个体维度与时间维度两方面的差异,包含的信息内容更多,因此本文也选择面板模型对粤港澳大湾区的区域经济模式进行分析。首先聚焦面板模型中的动态模型,通过尝试动态回归模型来寻找通过显著性检验的主成分因子来作为最终衡量粤港澳大湾区经济发展的指标体系。

$$Y_{it} = \beta_0 + \sum_{i=1}^n \beta_i X_{it} + \varepsilon_{it} \quad (3)$$

$$Y_t = \alpha + \beta_1 Y_{t-1} + \sum_{i=1}^p \gamma_i X_{it} + \varepsilon_{it} \quad (4)$$

(3) 式为静态面板模型的结构,(4) 式为动态面板模型。观察两式不难发现,静态模

型是将每一年或每一季度的数据进行单独分析，而忽略了时间序列数据的自相关性；而动态模型则假设因变量与其滞后期呈自相关关系，并通过纳入这样的滞后因子完成动态视角下的模型估计。但与此同时，动态模型也存在一个关键问题，那就是因变量的滞后期与干扰项相关，这种强相关性造成了严重的模型内生性，因此 OLS 估计结果不具备一致性。为解决这一难题，本文采用一阶差分 GMM 方法对因变量滞后期存在的内生性进行剔除，从而得到可靠的估计结果。

下面，我们也将基于动态视角，考察纳入经济增长滞后期后，各变量对经济增长的影响是否仍存，据此全面分析粤港澳大湾区经济模式的动态成因。使用一阶差分 GMM 对动态面板方程进行参数估计，下表 9 给出了动态面板方程的参数估计结果。模型的  $R^2 = 0.936$ ， $MSE = 450640.083$ ，常数项、核心解释变量（CPI）、 $F_{12}$ 、 $F_{31}$ 、 $F_{42}$ 、 $F_{42}$ 、 $F_{52}$ 、 $F_{62}$ 、 $Y_{t-1}$  的 P 值均小于 0.05，通过了 5% 显著性的检验，模型总体拟合效果较好。

表 9 动态面板方程参数估计结果

Variable	Coef	t	P
截距	48931.01	151.358	0.0001
居民消费价格指数 (CPI)	5272.77	0.9392	0.0354
$F_{11}$	-8125.01	-0.2725	0.0787
$F_{12}$	-3322.75	-0.6981	0.0489
$F_{31}$	28062.93	0.8486	0.0401
$F_{41}$	4301.694	0.4108	0.0684
$F_{42}$	-15396	-0.8688	0.039
$F_{51}$	-10597.2	-0.5945	0.0556
$F_{52}$	1859.954	0.6874	0.0496
$F_{61}$	6172.423	0.5066	0.0615
$F_{62}$	-8156.31	-0.754	0.0455
$F_{71}$	13931.48	0.3029	0.0764
$Y_{t-1}$	0.4538	0.4449	0.0659

观察动态模型回归结果，可以发现：人均 GDP 的滞后一阶项系数为正，强调了经济活动连续性和惯性对大湾区经济增长的重要性。科技创新因子对人均 GDP 有显著的正向影响，强调了技术创新在推动大湾区经济增长中的核心作用；同时，自然资源因子也表现出正向影响，说明自然资源的有效利用对大湾区经济发展具有积极影响。然而，教育规模因子的系数为负，反映了当前教育资源分配对经济增长的推动作用尚未充分发挥。另外，产业产值因子的系数为负，这意味着产业结构或产值效率方面存在的问题制约了人均 GDP 的增长。人口结构因子对经济增长的影响接近显著水平且为负，暗示了人口老龄化等人口结构变化带来的挑战，与我们的实际预期相吻合。此外，能源效率因子系数为负，表明当前能源利用效率的提升尚未充分转化为经济增长的动力；城市运输因子、产业转型因子和市场外贸因子均对人均 GDP 有正向影响，但未通过显著性检验，说明这些领域的发展对大湾区经济实力的提升已生效甚微。

基于以上动态面板的回归分析，我们发现动态面板方程不仅纳入了更多实际经济活动中可能出现的滞后现象，还利用了一阶差分的 GMM 估计来减少时序模型易出现的自相关问题，因此，我们采用动态模型中的通过显著性检验的因子来进行后续的预测任务，整理如下：

表 10 最终湾区经济发展指标体系

公共因子	名称	主要特征变量	隶属类别
$F_{12}$	人口规模因子	总人口数、就业人口、在校学生数、教育经费、 高等教育毛入学率	人口
$F_{31}$	科技创新因子	高技术产业利润总额、专利授权量、 R&D 经费支出	科技
$F_{42}$	产业产值因子	第二、第三产业生产总值	产业
$F_{52}$	自然资源因子	碳排放、森林蓄积量、第一产业能源消费总量、 单位 GDP 能耗上升或下降	环境
$F_{62}$	经济状况因子	财政收入、财政支出、 城镇、非城镇居民人均可支配收入	经济状况

## 5.5 结果分析

在动态模型中筛选的主要特征变量，均承载着深刻的经济含义：

人口规模因子显示，其动态回归系数为负，意味着人口过度增长已对粤港澳大湾区经济构成压力。目前劳动力市场接近饱和，非就业人口增加会导致社会需抚养人口增多，减缓经济发展。为优化人口红利，政府需合理配置劳动力资源和提供经济政策支持，如改革养老保险、提升教育水平和建设高效要素市场[5]。

科技创新因子系数显著为正，表明科技发展对粤港澳大湾区经济具有强大推动力。根据柯布-道格拉斯生产方程，科技进步与经济增长紧密相连，技术创新通过改变供需结构、优化资源配置和调整产业结构[6]，持续推动大湾区经济发展。

产业产值因子中，第二、三产业生产总值综合回归方程与主成分因子系数为正数，表明这些产业的发展极大促进了经济。改革开放后，第二产业扩张带动劳动力需求扩大，第三产业的蓬勃发展使产业结构得到优化，进而推动城市化进程[7]，加速粤港澳大湾区经济发展。

自然资源因子系数虽小但为正，说明自然资源开发有助于湾区经济发展，主要体现在物质资本和资源开发投入上。然而，过度开发可能削弱科技创新和教育投入[8]，减缓经济增长。

经济状况因子系数为负，揭示财政收入和支出扩张对粤港澳大湾区经济的抑制作用。短期内，财政支出扩大对经济增长作用有限，长期亦存局限。政府财政支持可能使企业产生依赖，抑制内生动力[9]，总体上对经济增长产生负面效应。

这些主要特征变量共同描绘了大湾区经济发展的复杂图景，为政策制定提供了重要参考。

## 6 任务二 粤港澳大湾区经济发展预测模型

### 6.1 非线性检验

通过上述的动态面板回归模型，我们挑选出对人均 GDP 发展有显著性影响的特征因子，但这些因子之间的经济关系还是有待检验。例如：核心变量 CPI 对湾区经济的支持作用是否保持不变？教育规模与科技创新因子的系数是否存在相互作用而产生的波动？上述动态回归面板的特征选择虽然反应了一般规律，但若是认为他们会恒定不变，显然是有悖经济直觉的。为此，本节对基准动态模型进行非线性扩展，以期最真实地还原变量间的本质作用关系，选择面板平滑迁移 PSTR 模型进行非线性面板估计，以核心变量 CPI 作为门槛转移变量，以教育规模因子、科技创新因子、产业产值因子、自然资源因子、经济状况因子作为解释变量。

#### 6.1.1 非线性特征的检验

面板平滑迁移 PSTR 模型可以依据某个连续变量（如本研究中的居民消费指数 CPI），灵活且平滑地描述变量之间关系的动态变化，使得它能够更好地拟合经济数据中隐藏的复杂结构。基于此，本文选用面板平滑迁移 PSTR 模型开展非线性估计。本文选取居民消费指数（CPI）作为门槛转移变量，将经过任务一筛选后的五大主成分以及滞后一期 GDP 作为解释变量。在进行模型估计前，对模型是否存在非线性特征进行检验，结果如表 11 所示：

表 11 PSTR 模型非线性拟合结果

统计量	结果取值	P 值
LM	21.664	0.001
LMF	46.650	0.008
LRT	32.493	0.000

LM 检验通过检验模型残差中是否存在未被线性形式充分解释的部分，来识别非线性关系的存在。LMF 检验则进一步考虑了固定效应对模型的影响。LRT 检验通过比较嵌套模型的似然比，来评估添加非线性项后模型拟合度的提升情况。结果显示，三种统计检验方法均在 1% 的显著性水平上一致地拒绝了模型为线性的原假设，这不仅证明了即使考虑了固定效应，模型中仍然存在着极为显著的非线性关系，而且揭示了添加非线性项后，模型拟合度得到了显著提升，从而验证了选用 PSTR 模型进行非线性分析的科学性和合理性。

#### 6.1.2 残余非线性的检验

为了确保模型的准确性和完整性，我们进一步对 PSTR 模型进行了残余非线性检验，旨在判断模型中转移函数的最佳数量。如表 12 所示给出了 PSTR 模型残余非线性检验的结果。

表 12 PSTR 模型的残余非线性检验结果

统计量	不存在两个及以上的转移函数	P 值
LM	26.982	0.921
LMF	-1.684	1.000
LRT	0.442	0.998

结果显示，三种检验方法 P 值均接近 1，接受了仅存在一个转移函数的原假设。证明模型在设置以个转移函数时已经能够有效地捕捉数据中的非线性特征，额外增加转移函数（两个及以上）可能会导致模型过度拟合或引入不必要的复杂性。

## 6.2 PSTR 模型估计

在动态面板的特征选择后，考虑到线性模型可能难以清晰地实现粤港澳大湾区区域经济模式成因的特征识别，特别是难以反映这些非线性因素的动态变化趋势以及规模门槛特征。[10]鉴于此，我们使用面板平滑迁移函数的非线性检验作用对动态模型（4）式进行改写，具体形式如下：

$$\begin{cases} Y_{it} = \beta_0 + \sum_{i=1}^n \beta_i X_{it} + \sum_{k=1}^k (\beta_i^k X_{it}) \Gamma^k(q_{it}; \gamma, \bar{q}_h) + \alpha_1 Y_{t-1} + \varepsilon_{it} \\ \Gamma^k(q_{it}; \gamma, \bar{q}_h) = \left[ 1 + \exp(-\gamma^k \prod_{h=1}^{H_k} (q_{it} - \bar{q}_h)) \right]^{-1} \end{cases} \quad (5)$$

其中，最核心的部分就是转移函数 $\Gamma$ ，该模型可以利用平滑转移函数刻画粤港澳大湾区经济模式成因的关键决定要素。指数转移函数的值将在 0—1 之间浮动，这说明面板平滑迁移模型很好地实现了非线性的平滑迁移，倘若转移函数处在变化过程中，而非是极为接近 0 或 1，那么就表明各因素对粤港澳大湾区的经济发展影响也处于动态变迁过程中；若转移函数值稳定在 0 或者 1，那么就意味着各要素对湾区经济增长的影响机制已经达到了稳态，要么是处于最低效状态，要么是处于效率瓶颈。

单独讨论转移函数 $\Gamma$ ，在前一节中，原模型已通过了残余非线性检验，则 $H_k$ 取 1，刻画逻辑转移函数的显形表达

$$\Gamma^k(q_{it}; \gamma, \bar{q}_h) = \left[ 1 + \exp(-\gamma^k \prod_{h=1}^{H_k} (q_{it} - \bar{q}_h)) \right]^{-1} \quad (6)$$

通过（6）式不难发现，参数 $\gamma$ 决定了面板平滑迁移模型在两区制内的转移速度，当 $\gamma$ 取值较小时，机制迁移较为缓慢，可能会导致整个模型在样本期间内都体现出较高的时变特性；而当 $\gamma$ 取值较大时，机制迁移速度较快，样本区间可能被分为两个静态子样本，这是面板平滑迁移模型的估计结果可以用面板门限模型进行检验。

观察表 12 可知，教育规模因子、科技创新因子、产业产值因子、自然资源因子、经济状况因子对经济增长的作用具有转移特征，所得模型估计结果如下表达式：

$$\begin{aligned} \ln GDP_{it} = & -1596.87 F_{12} + 1096.195 F_{31} - 1028.84 F_{42} + 217.67 F_{52} \\ & + 1833.41 F_{62} + [1 + \exp(-7.1131(CPI_{it} - 0.8514))]^{-1} (2166.83 F_{12} \\ & - 2188.06 F_{32} + 981.10 F_{42} - 2580.15 F_{52} - 2057.27 F_{62}) \end{aligned} \quad (7)$$

表 13 非线性面板平滑迁移模型的参数估计

参数	$F_{12}$	$F_{31}$	$F_{42}$	$F_{52}$	$F_{62}$	$\bar{c}$
估计值	1596.87***	1096.19**	1028.84	-217.67***	1833.41**	0.8514
参数	$F_{12}^1$	$F_{31}^1$	$F_{42}^1$	$F_{52}^1$	$F_{62}^1$	$\gamma$
估计值	1026.91**	995.62**	-1709.94*	60.39	-676.14*	7.1131

(注: \*表示  $P \leq 0.05$ , \*\*表示  $P \leq 0.01$ , \*\*\*表示  $P \leq 0.001$ )

同样地,还可以从估计结果中得出位置参数值为 0.8514, 门槛估计值较小, 表明多数样本处于我们选取的样本与转移函数取 1 时, 模型刻画的状态更为接近。为进一步说明, 我们列示出转移函数分别取 0 和 1 时, 整个系统的两个稳态方程:

$$\ln GDP_{it} = 1596.87 F_{12} + 1096.195 F_{31} + 1028.84 F_{42} - 217.67 F_{52} + 1833.41 F_{62} \quad (8)$$

$$\ln GDP_{it} = 369.96 F_{12} + 2091.81 F_{31} - 681.10 F_{42} - 157.28 F_{52} + 1157.27 F_{62} \quad (9)$$

式(8)与式(9)给出了区制 1 与区制 2 对应的估计结果, 其中, 当  $c$  低于 0.8514 时, 实际经济增长与各湾区要素的关系更接近区制 1 的结果; 当  $c$  高于 0.8514 时, 则更接近区制 2 的结果。根据转移函数特征可知, 区制 1 下转移函数取 0, 意味着居民消费价格指数很大; 而区制 2 下转移函数为 1, 代表着居民消费价格指数较低, 也即经济早期发展阶段。式(8)和(9)中, 教育规模的系数分别为 1596.87 和 369.96, 表明大湾区早期的经济发展非常依赖于教育资源, 教育资源具有经济增长加速效应, 而当居民消费价格指数达到一定程度后, 通货膨胀开始影响正常的消费者需求, 此时出现下沉市场效应, 教育规模、产业产值的提升对经济增长的正向促进作用明显减弱, 这也深刻地体现出, 湾区经济增长正由传统的人口密集型粗放增长方式向资本深化的集约型生产方向转变。

此外, 对于自然资源因子  $F_{52}$  而言, 尽管从早期逐渐发展至成熟的进程中, 自然资源被占用得越多, 经济发展得越快, 即自然资源规模与经济发展呈反向关系。但受国家政策“金山银山就是绿水青山”政策的影响, 自然资源因子的系数由 -217.67 变为 -157.28, 可见大湾区已经在尽可能减少通过损耗自然资源促进的 GDP 增长。

同理, 科技创新因子、产业产值因子、经济状况因子都出现了不同程度的效率下降, 这服从于边际效用递减规律, 表明模型估计结果可信。同时也深刻地说明, 大湾区的经济发展现在已进入了效率饱和阶段, 需要寻求更好地方式予以破解。

考虑到湾区经济发展存在着效率瓶颈, 本文进一步测算了效率瓶颈出现的节点, 根据中转移函数的测算结果, 当转移变量规模达到 5.6 以后, 转移函数取值 0.99, 即湾区城市 CPI 高于 8.7 后, 现阶段各变量对经济增长的支撑作用将触发瓶颈。由此可见, 居民消费价格指数绝非是越高越好, 还要考虑通货膨胀带来的负面影响, 寻求可持续的重要增长点, 如培养高新科技项目等。从长远来看, 尚未饱和的湾区城市仍可以通过政府宏观调控适当刺激经济增长后发赶超, 而稳态化增长的城市则可以保持中高速平稳健康发展, 这有利于整个大湾区经济的协调稳定发展。

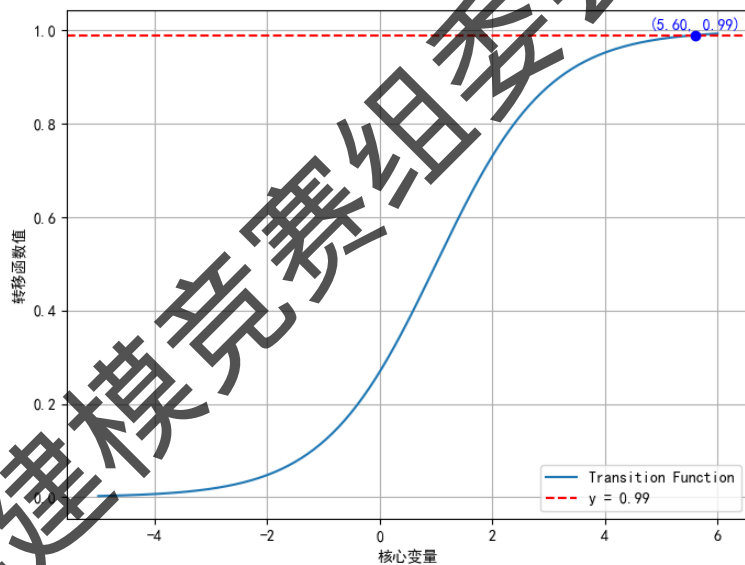


图 8 平滑迁移函数图

### 6.3 运用 ARIMA 时间序列预测

ARIMA 模型全称为自回归差分移动平均模型，其基本思想是用数据历史信息预测未来。一个时间点的标签值受过去标签值和偶然事件影响，即 ARIMA  $(p, d, q)$  模型试图通过数据自相关性和差分提取时间序列模式以预测未来。

本文通过前期的平稳性分析，发现特征指标在经过一阶差分处理后平稳，因此确定参数  $d$  的值为 1。为了确定  $p$  和  $q$  的数值，对所有模型指标绘制 ACF 图和 PACF 图。由于篇幅限制，所有指标在图中表现出相近的特征，在此只展示 CPI 和  $Y_{t-1}$  的 ACF 图和 PACF 图。

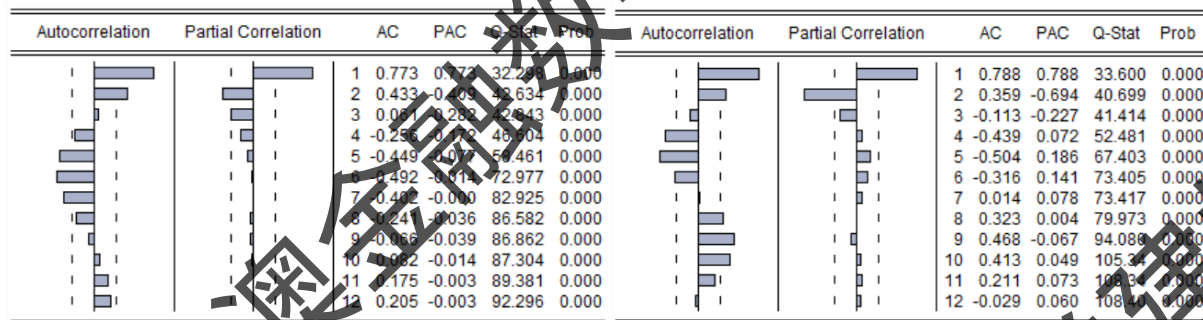


图 9 CPI (左图) 和  $Y_{t-1}$  (右图) 的自相关与偏自相关图

如图 9 所示，自相关图在一阶后衰减趋于零，偏自相关图在二阶后衰减向零，因此确定  $p$  值为 2， $q$  的值为 1。综合以上分析，采用 ARIMA(2,1,1) 对因子  $F_{12}$ 、 $F_{31}$ 、 $F_{42}$ 、 $F_{52}$ 、 $F_{62}$ 、 $Y_{t-1}$  和居民消费价格指数 (CPI) 等指标未来十年的数值进行预测，预测结果如下表所示：

表 14 各标准化指标 10 年预测表

年份	CPI	$F_{12}$	$F_{31}$	$F_{42}$	$F_{52}$	$F_{62}$	$Y_{t-1}$
2024	-1.604	2.861	8.605	1.127	2.883	1.284	139283.069
2025	-0.693	2.930	14.159	0.764	1.047	1.228	145906.297
2026	-0.854	2.359	15.675	0.122	1.849	1.045	152528.101

2027	-1.065	2.092	12.450	0.511	2.160	1.047	159148.483
2028	-0.844	2.281	7.516	0.878	1.410	1.153	165767.441
2029	-0.936	2.497	4.874	0.530	1.928	1.169	172384.976
2030	-0.947	2.486	6.296	0.346	1.878	1.111	179001.089
2031	-0.905	2.369	10.217	0.628	1.626	1.092	185615.780
2032	-0.933	2.327	13.328	0.698	1.889	1.122	192229.049
2033	-0.927	2.372	12.323	0.485	1.790	1.138	198840.897

由表 14 可得，CPI 在预测前期波动较大，随后逐渐趋于稳定，这与宏观经济环境的动态变化以及市场供需关系的调整密切相关； $F_{12}$  整体呈稳步上升态势，反映出大湾区在教育领域持续投入的积极成效，为人才培养提供了坚实支撑； $F_{31}$  前期增长较快，但在中期出现增速放缓迹象，表明技术创新在达到一定阶段后面临突破瓶颈的挑战； $F_{42}$  在波动中上升，显示产业发展在面临挑战的同时仍具备增长潜力； $F_{52}$  受环保政策等因素影响，数值变化体现出大湾区在经济发展与生态保护间的权衡； $F_{62}$  与  $Y_{t-1}$  总体稳定增长，表明大湾区经济发展的惯性与韧性。

将预测出的因子带入 PSTR 模型后，得到最终回归后的人均 GDP 预测值，如下图所示：

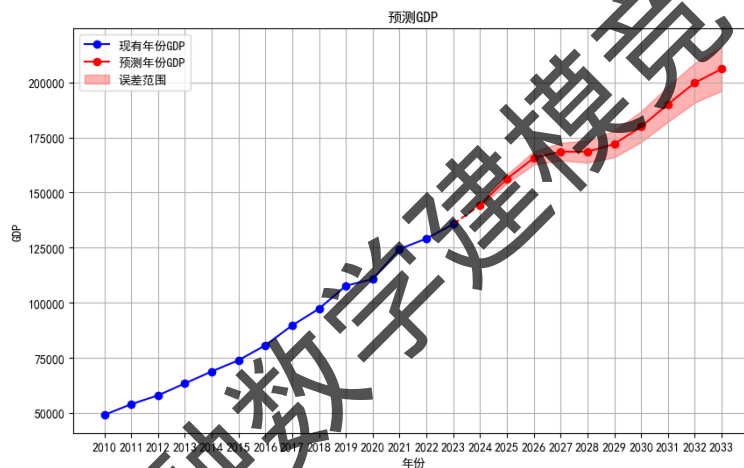


图 10 未来 10 年人均 GDP 走势

如图 10 所示，可以看到未来十年粤港澳大湾区的人均 GDP 呈现出稳步提升的良好态势，只在 2028 年出现一定程度的减少现象。结合 6.2 节提到的门槛效应，对照上述 ARIMA 预测模型，发现 2028 年时大湾区的居民消费价格指数预测值已超过 8.7（标准化后），则此时各变量对经济增长的支撑作用将触发瓶颈，导致增速放缓的同时出现经济短暂停滞的问题。

## 6.4 通过因子预测制定大湾区宏观发展策略

为了使粤港澳大湾区顺利通过 2026—2028 年的增速平缓期，实现经济的稳健、可持续发展，运用 ARIMA(2,1,1)模型预测未来十年因子变化，见表 14。基于 ARIMA 预测的未来十年因子变化，提出“创新驱动，绿色转型，物价稳控”的大湾区经济发展三向协同策略：

(1) 创新驱动：强化科技创新引领与产业升级。面对科技创新因子  $F_{31}$  和和产业产值因子  $F_{42}$  在未来发展中的增速放缓趋势，加大科技创新投入力度并制定系统性、针对性优惠政策，

加强与高校、科研院所的产学研合作，形成协同创新网络，加速科技成果转化为现实生产力。吸引高端科技人才汇集，建立国际化人才招聘平台。培育创新型企业梯队，打造高新技术产业集群，推动传统产业向高端化、智能化、绿色化转型。

(2) 绿色转型：践行可持续发展理念与生态经济模式。为了化解自然资源因子 $F_{52}$ 对经济增长产生的抑制影响，在“绿水青山就是金山银山”理念的指引下，推广节能减排技术并提高资源利用效率，积极发展绿色产业，推动绿色金融创新，探索生态产品价值实现机制，将生态优势转化为经济优势。加强环境保护与生态修复，严守生态保护红线，实施森林质量精准提升工程、海洋生态保护修复工程等，推进生态城市建设，提高城市生态宜居水平。

(3) 物价稳控：加强宏观调控与民生保障。CPI 波动显著，对大湾区经济稳定和居民生活产生重要影响。政府应强化宏观经济检测，加强货币政策、财政政策与产业政策的协同配合，保障民生领域供应，建立重要民生商品储备制度，完善应急保供机制，稳定居民生活成本。在物价稳定的基础上，完善社会保障体系，增强居民消费信心，发展消费新业态新模式，实现消费与投资的良性互动和平衡发展。

通过“创新驱动，绿色转型，物价稳控”的大湾区经济发展三向协同策略，大湾区有望在应对科技创新、自然资源、物价波动等多重挑战时，充分发挥自身优势，挖掘发展潜力，实现经济的长期稳定增长，成为具有高度创新力、竞争力和可持续发展能力的世界级大湾区。

## 6.5 稳健性分析

为深入剖析经济增长相关因素的效应及其稳定性，本文进行稳健性分析，通过替换解释变量的方式来重新考察经济关系。财政支出对经济增长具有直接作用，能够拉动消费、刺激需求、优化经济结构，并促进重点产业与新兴产业的成长，基于此，本文选取 2010—2023 年大湾区财政支出数据作为新的核心解释变量，运用到 PSTR 模型中进行回归估计。估计对比结果如下表 15 所示：

表 15 替换核心解释变量后的非线性参数估计

参数	$F_{12}$	$F_{31}$	$F_{42}$	$F_{52}$	$F_{62}$	$\bar{c}$
估计值	1056.35 *	1503.98*	591.89 *	-524.20 *	2105.72 **	5218.76 ***
参数	$F_{12}^1$	$F_{31}^1$	$F_{42}^1$	$F_{52}^1$	$F_{62}^1$	$\gamma$
估计值	598.36*	2839.56 **	3811.03 **	-402.98 **	1286.68 *	3470.00 ***

(注：\*表示  $P \leq 0.1$ ，\*\*表示  $P \leq 0.05$ ，\*\*\*表示  $P \leq 0.01$ )

由表 15 不难发现，替换核心解释变量后，财政支出对 GDP 的影响在 1%置信区间内显著为正，即在较高的置信水平下，财政支出的增加对大湾区 GDP 的增长具有积极的推动作用。各指标系数在 5%的置信区间内显著，进一步证明了模型的稳健性和可行性。

## 7 任务三 四大湾区发展异同分析

### 7.1 模型复现

不同湾区的经济发展有自身特点，为进一步探究湾区之间的经济发展异同，本文选取日本——东京湾区、美国西海岸——旧金山湾区、美国东北部——纽约湾区与上文分析的中国——粤港澳大湾区进行对比，分别使用上述非线性回归模型预测各湾区未来 5—10 年的经济走势，并通过量化各湾区间经济发展水平差异与时空差异来对比不同湾区间的异同。

为尽可能消除各湾区间数据异质性的问题，我们所选取的指标隶属大类均与粤港澳大湾区的指标体系——人口与教育因子、科技与创新因子、环境与能源因子、产业与产值因子、经济与市场因子。东京湾、纽约湾、旧金山湾区衡量经济发展的指标体系如下：

表 16 三大湾区经济发展指标体系

变量因子	主要组成特征
被解释变量	人均 GDP
核心解释变量	居民消费价格指数
控制变量	人口与教育大类 总人口数、就业人口
	科技与创新大类 高技术产业利润总额
	环境与能源大类 碳排放量
	产业与产值大类 第一、二、三产业生产总值、运输业 GDP
	经济与市场大类 财政收入、财政支出、城镇居民人均收入、总出口额、总进口额

#### 7.1.1 旧金山湾区

在确定旧金山湾区衡量经济发展的指标体系后，进行一系列数据处理与插值变换，重复任务一中的因子分析，对各大类特征进行主成分分析，并放进动态回归面板模型中初步观测，保留通过显著性检验的因子；重复任务二中的非线性检验与回归分析，并寻找最优 ARIMA(p, d, q) 模型预测未来 10 年的主成分因子，所得非线性预测 10 年人均 GDP 值如下。

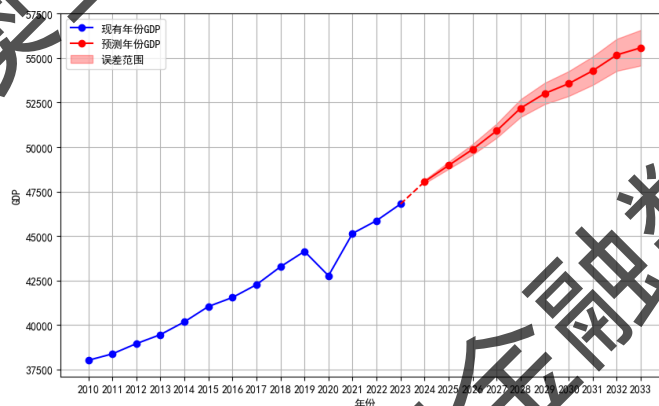


图 11 旧金山湾区未来 10 年人均 GDP 拟合效果图

旧金山湾区是全球最富裕的城市群之一，拥有强大的经济总量和高度集中的财富，在高新技术产业方面，旧金山湾区作为全球闻名的高科技研发基地，硅谷就位于湾区南部。通过研究发现，该湾区以科学技术为基础，依靠信息产业来推动金融、旅游业和其他服务的发展和扩展，成为世界级的城市群。

纵观旧金山湾区近 15 年来的人均 GDP 历史数据，也能看出其超然于其他湾区的傲然实力，尽管 2020 年受到新冠疫情的影响，但依托于多样化的高新产业与金融产业，该湾区具备强大的自我修复能力，在未来 10 年内，其非线性预测结果仍保持强劲的上升态势。

### 7.1.2 东京湾区

东京湾区位于日本本州中东部中部海岸的太平洋入口，是日本最大的工业城市群、交通运输中心、国际金融中心、贸易中心和消费中心，主要包括东京、川崎和横滨等城市，是日本经济最发达的区域。通过研究发现，该湾区的突出优势体现在港口、人口集聚、对外开放和规划的先导性，这些是东京湾区的区域经济效率提升起到重要基础作用。

对于收集到的东京湾数据，其处理与分析方法同理可得如下预测：

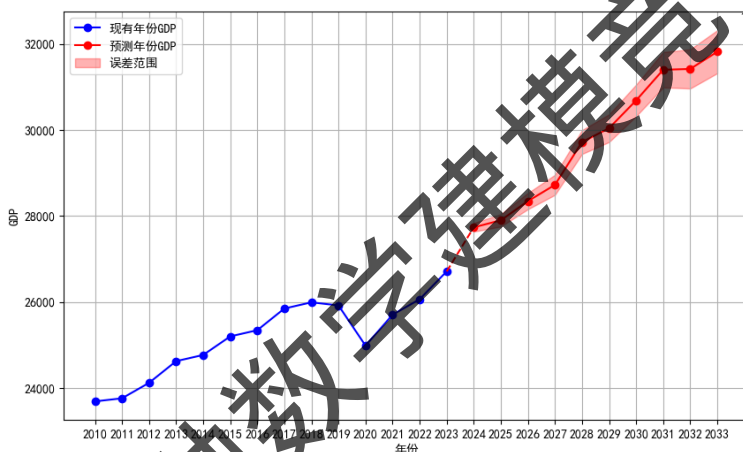


图 12 东京湾区未来 10 年人均 GDP 拟合效果图

图中蓝色线条为历史过往年份的人均 GDP 值，红色线条为未来预测年份的人均 GDP 值，红色阴影范围为预测值的 95% 置信区间。由图可知，尽管受到 2020 年新冠疫情影响，导致东京湾经济出现短暂的停滞与下行，但未来 10 年仍保持经济上升趋势，发展态势良好。

### 7.1.3 纽约湾区

纽约湾区是“金融湾区”，是世界金融中心，其金融业、奢侈品业和城市文化都具有全球影响力，世界金融、期货、保险、外贸等多家组织和机构总部也位于纽约湾区[11]。除此之外，纽约大都会地区的高等教育网络也十分发达。通过研究发现，该湾区内自然地理优势、基础设施、产业链等都为区域的经济效率提供了良好的指导。

对于收集到的纽约湾数据，其处理与分析方法同理可得如下预测：

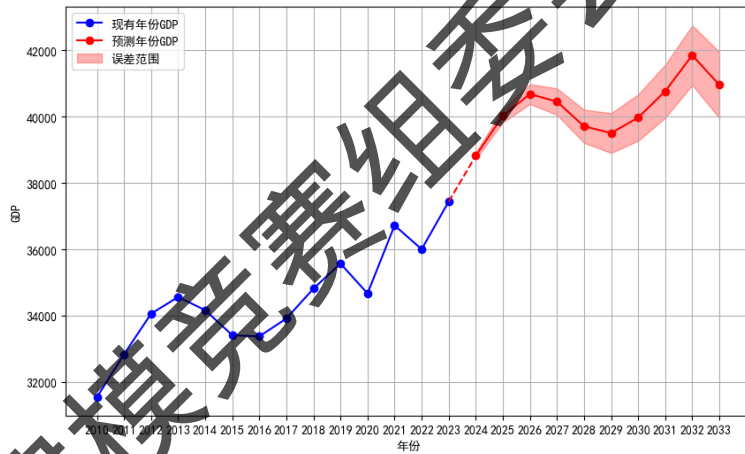


图 13 纽约湾区未来 10 年人均 GDP 拟合效果图

由图可知，纽约湾从 2010 到 2015 年的经济发展波动较大，后又受到 2020 年新冠疫情的影响，导致纽约湾经济在近 15 年历史数据中表现出方差较大，均值较小的强波动特点。由于过往数据的不稳定，非线性模型所预测出的未来经济发展也表现出波动中上升的态势。

## 7.2 湾区经济发展水平差异分析

通过以上回归模型预测得到 2024 年——2033 年四大湾区的经济发展水平，对比如图：

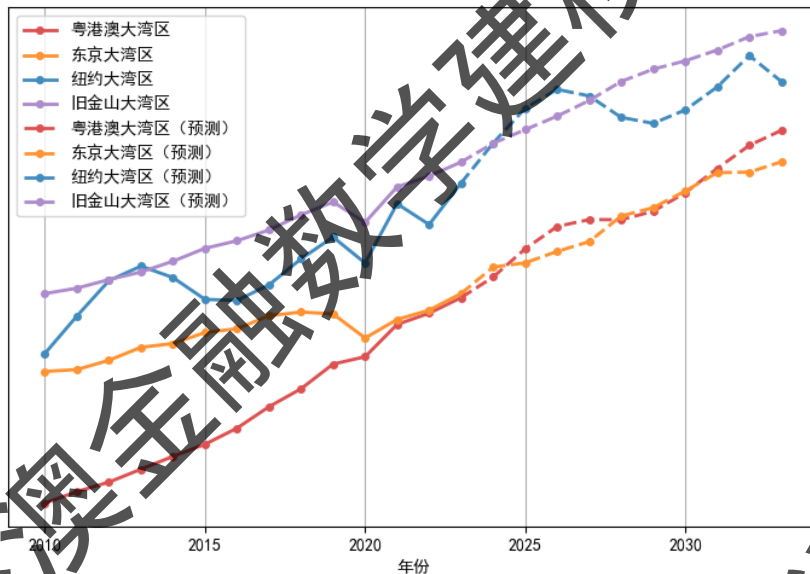


图 14 四大湾区预测人均 GDP 对比图

与当前全球影响力最大的纽约湾区、旧金山湾区和东京湾区相比，粤港澳大湾区正在实现金融、科技等领域的全面追赶。相较于其他三大湾区，粤港澳大湾区的人均 GDP 与其他三大湾区相比仍存有不小差距。在我们的预测模型中，粤港澳大湾区可能于 2025 年实现东京湾区人均 GDP 的超越，但要看齐纽约湾和旧金山湾区，发展之路仍然任重道远。

细分四大湾区的经济发展的五大维度水平：人口、科技、产业、环境与经济状况，我们将预测所得的主成分因子进行对比，并绘制雷达图如下：

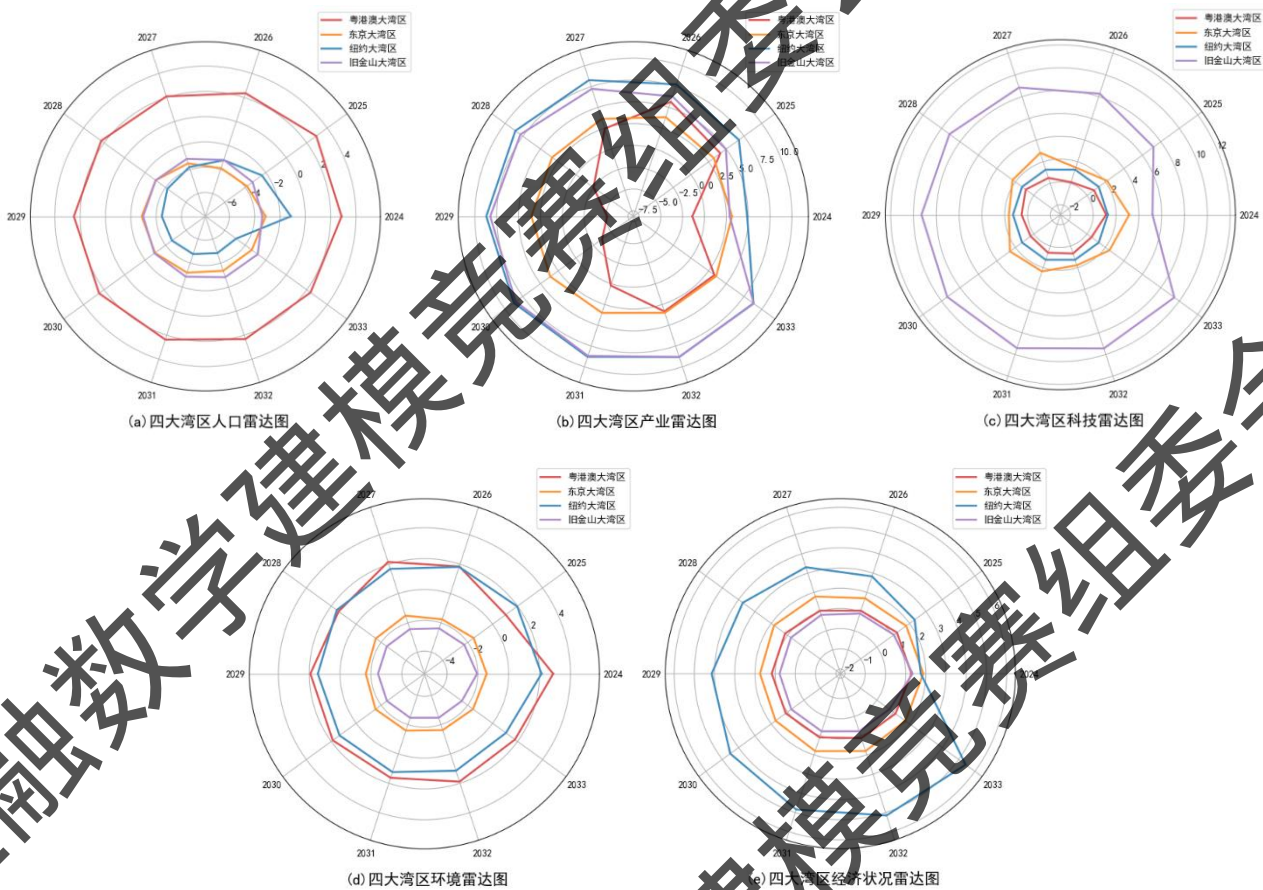


图 15 四大湾区经济发展五大维度雷达图

由图中可以观察到，粤港澳大湾区的人口规模远超其他湾区；在科技方面，旧金山湾区作为全球高技术科技研发中心之一，其科技因子值在未来 10 年内将持续领先，与此同时，受到“创新是第一生产力”号召的粤港澳大湾区也表现出科技因子的高增速趋势；在产业方面（在本文中该因子主要侧重三产规模与就业比例），预测所得东京湾、纽约湾、旧金山湾区的该因子值较为接近，粤港澳大湾区也将在未来 5 年持续追赶，保持增速；在环境方面，受我国政策影响，粤港澳大湾区的各类能源消耗表现出逐年递减的趋势，而东京湾则呈现出较为严重的自然资源损耗；在金融状况方面，纽约湾作为全球第一大金融中心，在未来 10 年内的金融因子值也表现出良好的增长态势，该因子在粤港澳大湾区上也呈现出稳步增长。

在全球经济一体化的今天，湾区城市群作为推动经济发展的重要力量，其重要性愈发凸显。而在众多湾区中，纽约湾区、旧金山湾区和东京湾区以其卓越的经济水平和发展独特的产业结构，被誉为世界三大湾区。这三大湾区城市群不仅在全球经济版图中占据重要地位，而且其发展历程、产业结构特征以及发展经验，对于粤港澳大湾区的发展具有重要的借鉴意义。因此，在下一节中，我们将撰写一份简报给决策部门建言，包括前文的建模依据、我们对未来 10 年湾区经济的预测以及采取措施对大湾区经济变化的影响。

# 粤港澳大湾区经济发展分析与展望

Analysis and prospect of economic development in GBA



粤港澳大湾区作为我国经济发展的重要引擎，其经济发展态势不仅关乎区域繁荣，更对全国乃至全球经济具有深远影响。本简报旨在基于深入的数据分析与建模，为粤港澳大湾区未来10年的经济发展提供科学、精准的决策依据，助力大湾区持续繁荣。

## 内容概要

### 1 建模依据

### 2 经济预测

### 3 发展建议

1 本研究从多个维度广泛收集数据，筛选出40个二级指标。运用因子分析进行降维处理，提取出如人口规模因子、科技创新因子、金融规模因子、产业产值因子等关键公共因子，以评估大湾区经济发展的潜在方向。通过静态与动态面板模型综合筛选，构建了一套完备的预测指标体系。

## 2 经济预测

运用 PSTR 模型预测大湾区经济走势，发现教育与科技创新等因子对经济增长作用呈显著转移特征。分析大湾区未来10年的趋势，早期教育和科技推动大湾区经济高速发展，但到2028年将面临门槛效应，增速放缓。此时政府需发挥主导作用，运用宏观调控推动产业转型、增加居民收入与消费指数，助力大湾区平稳过渡，实现新腾飞，下文给出发展建议。



## 3 发展建议

- 1.加大教育投入，改善设施以培育人才，鼓励企业与高校、科研机构深度合作，借鉴旧金山湾经验发展高新技术产业，推动科技成果转化。
  - 2.依据大湾区发展实际，合理规划产业布局。参考世界三大湾区经验避免城市间恶性竞争，推动传统产业数字化转型，积极发展高端制造业、现代服务业等，提高产值效率。
  - 3.在发展中牢固树立环保理念，推广清洁能源与循环经济，实现资源合理开发利用。同时政府应灵活运用财政政策、货币政策等宏观调控手段，依经济形势适时调整政策方向和力度。如经济下行时刺激增长，过热时防范通胀，保障经济稳定。
- 相信通过这些举措，粤港澳大湾区将在全球经济格局中更具影响力。

## 9 总结与展望

### 9.1 优势

- 精确量化影响因素：通过构建湾区经济发展指标体系，运用如因子分析、静态和动态面板筛选等方法，能够精确地分析评价影响粤港澳大湾区经济发展的众多因素，并量化各因素对经济发展的影响。这种量化分析有助于深入理解经济发展的内在机制，为决策提供科学依据。
- 多模型综合预测：采用多种数学模型如 PSTR 模型估计、ARIMA 时间序列预测等，对粤港澳大湾区未来 5—10 年的经济走势进行预测。多种模型的综合运用提高了预测的准确性和可靠性，考虑到了不同模型对不同经济现象的适应性。
- 全面的湾区对比：在分析不同湾区的经济发展时，通过熵值法等手段，不仅能够预测所选湾区未来的经济走势，还能量化分析不同湾区之间发展的异同。这有助于借鉴其他湾区的发展经验，为粤港澳大湾区的发展提供更广阔的视角。

### 9.2 不足

- 模型复杂性与资源消耗：模型构建过程涉及多种复杂方法和大量数据处理，如因子分析和多种模型融合，导致计算成本增加、资源消耗较大，在一定程度上限制了模型快速运算和频繁使用的效率。
- 数据与现实适应性：模型依赖于历史数据，对于突发的、未在历史数据中体现的新因素或重大变化适应性较差，可能无法及时准确反映现实经济发展中的新趋势，影响预测的准确性和策略的有效性。

### 9.3 展望

- 模型改进与融合：未来的研究可以致力于改进现有的数学模型，或者探索将不同模型进行更有效的融合，以提高对湾区经济发展预测的准确性。例如，研究新型的混合模型，使其能够更好地适应湾区经济复杂多变的特性。
- 实时数据整合：为了应对数据时效性的问题，可以探索建立实时数据收集与整合机制，及时将新的数据纳入到模型中，使模型能够动态地反映湾区经济的最新变化，从而提高预测的精准度和决策的有效性。

## 参考文献

- [1] 李立勋:《关于“粤港澳大湾区”的若干思考》,《热带地理》2017年第06期,第757~761页。
- [2] 颜竹晗.数字普惠金融、区域经济韧性与金融风险[D].广西大学,2024.DOI:10.27034/d.cnki.ggxu.2024.001006.
- [3] 马茹,罗晖,王宏伟,等.中国区域经济高质量发展评价指标体系及测度研究[J].中国软科学,2019,(07):60-67.
- [4] 叶妍琬.广东省物流业与地区产业发展的相互关系研究[D].广东工业大学,2015.
- [5] 郭柯.经济增长中的“比较优势陷阱”分析[D].中共中央党校,2011.
- [6] 李春彦.财政投入、科技创新对经济高质量发展的影响研究[D].兰州财经大学,2023.DOI:10.27732/d.cnki.gnzsx.2023.000517.
- [7] 田菁.中国开放经济下的最低工资制度和二元结构整合[D].复旦大学,2009.
- [8] 郝娟娟.中国自然资源禀赋与经济发展关系研究[D].中央财经大学,2015.
- [9] 张雅琼.金融科技对实体经济的影响[D].北京邮电大学,2021.DOI:10.26969/d.cnki.gbydu.2021.001403.
- [10] 张羽,赵鑫.农村金融发展拉动了农村经济增长吗——基于面板平滑迁移模型的经验证据[J].社会科学战线,2015,(10):257-261.
- [11] 邓志新:《粤港澳大湾区与世界著名湾区经济的比较分析》[J],《对外经贸实务》2018年第351(04)期,第94~97页。