

参赛队号：（参赛队无须填写，参赛队号由大赛官网自动生成）

2023 年（第九届）全国大学生统计建模大赛

参 赛 作 品

参赛学校：

广东工业大学

论文题目：

基于电动汽车产业链发展的中国
式现代化统计测度的研究

参赛队员：

<已作匿名处理>

指导老师：

徐圣兵 童宁

基于电动汽车产业链发展的中国式现代化统计测度的研究

摘要

根据中国社会主义现代化国家基本特征，我国的现代化是人与自然和谐共生的现代化。注重同步推进经济建设和生态文明建设，不走先污染后治理的老路，而是要走节约资源、保护环境、绿色低碳的新型发展道路。要积极应对全球气候变化，力争 2030 年前实现碳达峰，努力争取 2060 年前实现碳中和，为全人类作出积极贡献。^[14]因此，本文主要研究电动汽车产业链和中国式现代化

本文将中国式“绿色现代化”测度拆解为三部分。

第一部分，中国的绿色现代化可以用各省份的现代化程度度量，本文使用 PCA 主成分分析对选取指标进行数据降维和相关性分析，然后利用 TOPSIS-熵权法进行赋权评价，最后按省份折算为现代化程度，得到中国国内现代化程度为 65.741%

第二部分，中国的现代化需和国外相比较，包括出口量、全球市场占有率等，其中我国的新能源汽车市场占有率在 2022 年达到了 60%以上，销量达一千万辆，连续多年占据全球第一，这表明国产汽车在新能源赛道上开始跟跑甚至领跑。

第三部分，对中国的“2030 碳达峰”承诺实证分析，利用改进后的 SIRD-NM 动态模型预测新能源汽车数量，同时与深度神经网络预测结果相对比，进而研究承诺实现的可行性。

最后，研究国家对新能源汽车鼓励政策的力度与新能源汽车普及率的关系，提出合理的政策建议。

关键词：2030 碳达峰；TOPSIS-熵权法；SIRD-NM 动态模型；深度神经网络预测

目录

摘要.....	2
一、引言.....	6
(一) 研究背景.....	6
(二) 研究目标.....	7
1. 为什么本文选取的指标能作为中国现代化指标?	7
2. 模型如何度量中国式现代化?	7
3. 指标的意义及结论?	8
(三) 研究的理论基础.....	9
1. 新能源汽车对绿色现代化和“2030 碳达峰”贡献度的研究现状	9
2. SIRD 模型研究现状	9
(四) 研究方法与技术路线.....	10
(五) 研究的创新点.....	12
二、基于箱根模型和布莱克标准的 TOPSIS-熵权法现代化评价模型	12
1. 主成分分析 PCA	13
2. K-means 聚类分析	14
3. 熵权法-TOPSIS 评价法	16
(三) 研究电动汽车产业链的现代化程度	18
三、基于马尔可夫状态转移矩阵预测电动汽车发展.....	20
(一) 构建马尔科夫状态转移矩阵.....	20
(二) 建立动态 SIRD-NM 模型.....	21
(三) 基于深度学习的新能源汽车数量预测模型.....	24
(四) 神经网络时间序列预测.....	27
四、基于环境库兹涅茨曲线对 2023 碳达峰的可实现性验证.....	28
(一) “倒 U 型”环境库兹涅茨曲线.....	28
1. 规模效应.....	28
2. 技术效应.....	28
3. 结构效应.....	28

(二) “2030” 碳达峰实证分析.....	30
五、结论与建议.....	32
(一) 建议政策.....	32
(二) 模型的不足.....	32
参考文献.....	33
致谢.....	35
附录.....	36

表格与插图清单

表 1. SIRD 模型输出和真实值对比

图 1.研究方法和技术路线

图 2.聚类中心数量和聚类效果

图 3.中国各省份聚类结果

图 4.全球各国家地区新能源汽车指标雷达图

图 5.2019-2022 年全球新能源汽车市场占有率柱状图

图 6.马尔科夫状态转移矩阵

图 7.改进后的 SIRD-NM 模型示意图

图 8.拟合结果

图 9.SIRD-NM 模型预测未来新能源汽车持有人群曲线图

图 10.两层多神经元深度学习预测新能源汽车数量网络结构示意图

图 11.拟合效果图一

图 12.拟合效果图二

图 13.均方误差图（左）

图 14.验证集收敛梯度（右）

图 15.最佳预测性能校验

图 16.时序响应误差

图 17.1999 年到 2019 年碳排放量与 GDP 关系散点图

图 18.碳排放量增速和时间的关系

图 19.单位 GDP 碳排放量

图 20.年均总碳排放量（百吨）

图 21.新能源汽车数量和政府鼓励力度 α 变化关系图

基于电动汽车产业链发展的中国式现代化统计测度的研究

一、引言

（一）研究背景

在绿色能源理念的指导下，中国经济和社会正面临着低碳能源转型的重要问题。为了应对这一挑战，国务院于 2021 年 2 月印发了《国家综合立体交通网规划纲要》，其中明确指出在交通领域中，应尽快达到二氧化碳排放的峰值^[1]。这一目标的实现需要采取一系列措施，包括推广新能源汽车、加强公共交通建设、提高绿色交通的比例等。通过这些措施的实施，可以有效减少交通领域的碳排放，为我国低碳能源转型进程做出重要贡献。同时，低碳能源转型也将推动中国经济和社会朝着可持续发展的方向迈进，实现经济、社会和环境多方面的共赢。

近年来，随着我国提出“2030 碳达峰、2060 碳中和”战略目标，环保问题已成为国家的重要议题。为了应对这一挑战，我国采取了一系列强有力的政策和措施，其中之一是大力推广新能源汽车。由于新能源汽车具有接近零排放和低碳绿色的特点，在“双碳”背景下，再度成为热门话题^[2]。此外，随着新能源汽车相关技术发展，消费者对新能源汽车的需求也在不断提升。预计在不久的将来，新能源汽车将成为汽车行业的主流产业，并超过传统燃油车的占有率。

随着碳中和计划的加速实施，新能源汽车产业得以快速发展。这一发展趋势备受关注，引起了相关学者的深入研究。他们探讨了新能源汽车产业的发展趋势、技术创新和政策支持等多个方面，并就此提出了很多有启发性的见解^[2]。例如，他们认为新能源汽车产业在技术方面的创新是实现可持续发展的关键，同时政府的政策支持和市场的需求也是新能源汽车产业发展的重要因素。李晓易、谭晓雨、吴睿、徐洪磊等人在对交通运输领域的绿色发展和碳排放现状进行系统分析后，提出了一个总体思路，即“分类施策、远近结合、先易后难、控增量调存量、积极稳妥推进、梯次有序达峰”，并根据 2060 年前的主要阶段指出了交通运输领域低碳发展的总体路径^[3]。这一规划框架旨在为中国在交通运输领域实现碳中和目标提供总体指导，并对实现低碳发展提供了具体的行动方案。刘建国、朱跃中、田智宇等人认为交通运输部门是中国能源需求及二氧化碳排放的重要领域之一，如

果中国交通运输部门延续西方发达国家及中国沿海发达地区的发展趋势，那么交通运输服务需求、千人汽车保有量、人均交通用能以及碳排放将快速增长。这将进一步加剧中国的能源消耗和碳排放，对环境和气候变化带来更大的挑战^[4]。

在各式新能源汽车中，电动汽车相对于氢能源、压缩空气等其他新能源汽车的产值、数量占主导地位，因此在本文主要利用中国的电动汽车产业链的各项数据研究中国现代化进程。

（二）研究目标

本文从以下三方面总结研究目标：

1.为什么本文选取的指标能作为中国现代化指标？

早在 1992 年，“两弹一星”的功臣钱学森便在给时任国务院副总理邹家华的信中，提出了这样一个建议：“我国汽车工业可以跳过柴油阶段，直接进入减少污染的新能源阶段！”^[11]

选取评价中国现代化程度的指标要求是覆盖面广、评价客观。新能源汽车的发展时间与中国现代化进程相近，中国新能源汽车产业始于 21 世纪初，而 2001 年，新能源汽车研究项目被列入国家“十五”期间的“863”重大科技课题、新中国在 1950-2018 年完成了第一次平均现代化、2000-2020 年完成全面建成小康社会。

在 2023 年 3 月的中国电动汽车百人会论坛中，党的二十大报告提出了中国式现代化，去年底的中央经济工作会议要求支持新能源汽车消费，今年两会政府工作报告更是把去年新能源汽车销量增长 93.4%写入数据亮点。^[10]

因此本文从“衣食住行”中的出行作为出发点，以新能源汽车为切入点研究中国现代化。截止目前，中国已有 41.67%的家庭拥有汽车，其中 31.1%的家庭拥有电动汽车，48.41%的家庭拥有传统的燃油汽车，数据量大，可以很好地降低异常值影响，得到准确的评价结果^[16]。而相较其他非常规新能源动力汽车（氢动力、空气动力），电动汽车技术较成熟，占据新能源汽车市场的绝对优势，因此选取电动汽车进行中国现代化评估可以使评价较为客观。^[12]

2.模型如何度量中国式现代化？

使用 TOPSIS-熵权法综合评价各省份的现代化程度；使用深度神经网络预测新能源汽车数量；使用神经网络时间序列预测计算新能源汽车数量；利用已有数据进行国内外电动汽车产业对比。

3.指标的意义及结论？

在 2019 年，中国原油对外依存度首次突破了 70%，已经连续十年超过了 50% 的国际安全警戒线。而汽车是消耗石油最多的大户，年消耗量占比接近 70%。发展新能源汽车，降低石油消耗量，是减少对外石油依赖最重要的途径。自 2001 年加入 WTO 以来，国家鼓励汽车进入家庭，2001-2019 年，汽车千人保有量从 50 台提升到了 170 台，全国汽车保有量自 2011 年首次突破 1 亿辆大关，到 2017 年突破 2 亿辆，发展到 2022 年 3 月已达 4.02 亿辆。随着汽车保有量快速提升，汽车对石油的消耗与日俱增。

相比中国原油对外依存度突破了 70% 更严峻的是，中国 70% 以上的进口石油都要经过马六甲海峡，运输通道集中而且单一，一旦国际局势不稳定，石油海上运输风险将激增。面对错综复杂的贸易环境，降低对进口石油的依赖，对当下的中国有着特别重大的意义。

中国的能源结构国情是：煤炭储量丰富而轻质石油储量高但质量不足，同时考虑到军事资源储备，因此大量依赖进口石油，而交通行业消耗占据石油总消耗的 50% [17]。而通过将交通行业电能化，将能源消耗重心从成千上万台汽车的引擎转移到发电厂中，通过电能这一桥梁，中国能源结构重心可以从需要大量进口的石油转移到中国本土储量丰富的煤炭。在国家大力发展新能源汽车下，2022 年的石油进口量出现 20 年来首次下降。 [15]

同时，在外国的传统能源汽车技术成熟而新能源汽车技术却刚刚萌芽时，中国便开始大力支撑新能源发展，跳过碳排放严重且后期开支（加油）较高昂的燃油汽车时代，在新能源汽车发展赛道中将中国与外国拉回到同一起跑线。

中国的电动汽车产业链在过去几年中取得了快速发展，并成为了全球最大的电动汽车市场之一，中国的电动汽车产业链已经进入了快速发展阶段，成为了中国经济的重要组成部分。这些数据也反映了中国经济结构的转型和升级，以及中国现代化进程中科技创新和环境保护的重要性。

中国政府已经将电动汽车产业列为国家战略，并采取一系列政策措施，鼓励电动汽车的发展。这些政策措施包括财政支持、减免税收、充电基础设施建设等。同时，中国的电动汽车制造商和相关企业也在不断创新和提升产品质量，以满足

不断增长的市场需求。

因此，中国的电动汽车产业链能够作为中国现代化进程的指标，因为它不仅代表了中国经济结构的转型和升级，也代表了中国在环境保护和科技创新方面取得的重大成就。

此外，根据“倒 u 型”环境库兹涅茨曲线，新能源汽车还可以描述碳排放与经济发展 GDP 的关系，并判断其碳达峰情况。

（三）研究的理论基础

1. 新能源汽车对绿色现代化和“2030 碳达峰”贡献度的研究现状

当汽车实现电动化后，使用能源的清洁与否决定了减碳本质。“双碳”战略引导下，以新能源为主体的新型电力系统与新能源汽车协同发展成为必然。未来要下决心实现电动汽车主要使用清洁电力，实现汽车零排放。

近几年来，我国政府认识到交通活动产生的外部性影响，政策制度部门积极借鉴发达国家的经验，制定相关技术标准和财税政策，加快新能源汽车产业发展，替代传统燃油汽车，从而推进节能减排，加快大气污染治理。不同学者对新能源汽车领域的研究也各有不同。

崔冬（2019）综合整理了 2018 年我国出台的税收和财政补贴等重要的新能源汽车的相关政策。刘金周、方海峰和刘万祥（2018）基于地方政府新能源汽车政策体系，根据新能源汽车销量重点分析了补贴政策、交通政策以及基础设施等政策的直接效应，提出了相关建议。

2.SIRD 模型研究现状

SIRD 模型是一种常用的流行病学模型，用于描述传染病在人群中的传播过程。在 COVID-19 疫情中，SIRD 模型被广泛应用于估计疫情在不同地区和国家传播趋势，并预测未来的发展情况。SIRD 模型将人群分为四类：易感者(Susceptible)、感染者(Infected)、康复者(Recovered)和死亡者(Deceased)，并使用一组微分方程来描述它们之间的相互转化关系。SIRD 模型可以帮助研究者了解病毒传播的速度和规模，评估各种干预措施的效果，以及预测疫情的发展趋势。

J Fernández-Villaverde 和 CI Jones 反演 SIRD 模型，利用日死亡数据恢复时间变化的基本再生数，以捕捉不同地点在不同时期和强度下行为和政策变化的影响，在模拟未来结果时，允许从日死亡数据中得到反馈可以影响未来行为 [6]。

他们的模型还可以反推出感染的人口比例和感染者与易感人群之间的比例。该研究有助于人们了解 SIRD 模型在拟合观察数据方面的困难，并提出了提高模型的方法。

流行病时间演化的几种模型，包括集体模型和网络模型两大主要类型 [7]。集体模型特点是 4 参数数量相对较少，使用有限数量的集体变量来描述人群中的流行病传播。其中包括广义生长模型、Logistic 模型、Richards 模型、广义 Richards 模型、次生流行病波模型、可感染-感染-康复（SIR）模型、可感染-感染-康复-死亡（SIRD）模型、可感染-潜伏-感染-移除（SEIR）模型以及可感染-感染-诊断-患病-认识-威胁-康复-灭绝（SIDARTHE）模型等[7]。SIR、SIRD、SEIR、SIDARTHE 等模型都属于所谓的分隔模型。集体模型更适合于非专业操作员和公共卫生机构，并且在数据稀缺的情况下仍然可以提供简单但可靠的模型。因此本文采用 SIRD 模型做运算。Nisar K S, Ahmad S, Ullah A 等人证明了稳定性条件 [8]。

（四）研究方法与技术路线

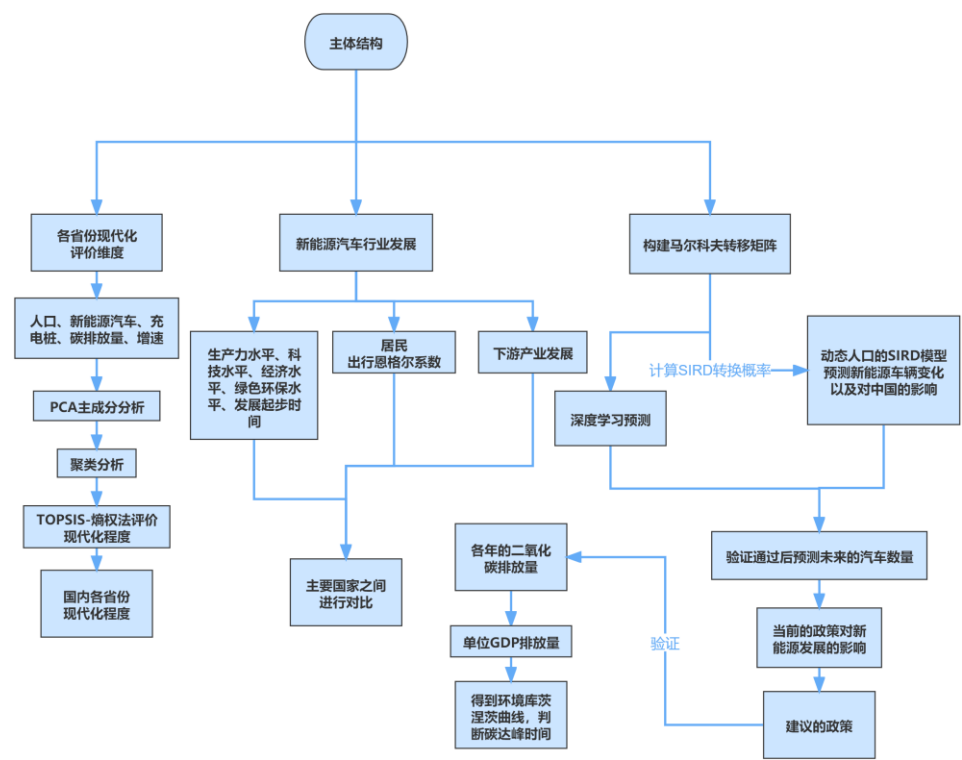


图 1 研究方法和技术路线

由于中国现代化和电动汽车现代化息息相关，进程和发展规模都高度重合，结合中国式现代化是人与自然和谐共生的现代化，本文将中国现代化解读

为各省份的新能源汽车现代化，通过箱根模型和布莱克标准提出八大因子，并通过 PCA 主成分分析降维并通过 K-means 聚类分析挑选省份代表，利用各省份代表的现代化程度综合评价中国现代化。而中国现代化需要与其他现代化程度高的国家相比较，研究中国在新能源汽车赛道上与国外的发展速度对比。

此外，为了预测新能源汽车未来发展趋势，以获取在不同政策影响下未来的各项指标变化情况，以便与国外对比现代化发展进程，同时对中国“2030”碳达峰承诺进行预测分析。

本文主要采用的研究方法主要有以下几种方法。

1.年鉴数据收集法

本文构建马尔科夫链需要交通运输、环境及宏观经济领域等方面的大量数据支持，这是一项复杂的工作。具体地，通过查阅如《中国统计年鉴》、《中国汽车工业年鉴》（2013-2017）等数据库网站和统计文献，构建马尔可夫状态转移矩阵。

2.文献查阅调研法

围绕本文的研究目的，需要查阅该领域大量的国内外研究文献，尤其是国内外相关研究现状、理论基础，分析已有研究的不足及可借鉴之处。

3.模型建立法

①TOPSIS-熵权法

基于 TOPSIS 优劣解距离法和熵权法的综合评价方法，熵权法可以获取指标的信息熵，进而计算信息丰富度，并为指标赋权，而 TOPSIS 优劣解距离法可以根据熵权计算样本的综合排名。

②SIRD-NM 动态总人口模型

模拟传染病模型，提出改进后的 SIRD-NM 动态总人口模型以分析未来新能源汽车发展变化。

③深度神经网络预测

利用多神经元的双层隐藏层 BP 神经网络预测新能源汽车数量变化。

4.模拟数值分析法

基于本文提出的 SIRD-NM 动态模型，以及深度神经网络预测未来的新能源汽车数量，并利用该模型模拟特定政策下的碳排放、新能源汽车产业链经济、国内外发展对比。

（五）研究的创新点

本文将新能源汽车产业链与中国现代化进程相结合，利用新能源汽车产业链现代化反映出中国现代化程度，并指出新能源汽车现代化对中国现代化的重要性，新能源汽车的普及对中国实现 2030 达到碳排放峰值承诺的贡献，模拟在特定政策的情况下新能源汽车的数量变化，以预测对能源结构、石油进口、新能源 GDP 的影响。本文基于 SIRD-NM 动态模型，提出适合中国国情的基于非封闭式系统动态总人口 M 模型，用以验证和预测中国未来新能源汽车现代化发展并为中国实现 2030 碳峰值承诺提出参考数据。本文将 TOPSIS 和熵权法综合得到评价模型，TOPSIS-熵权法是基于 TOPSIS 优劣解距离法和熵权法的综合评价方法，熵权法可以获取指标的信息熵，进而计算信息丰富度，并为指标赋权，而 TOPSIS 优劣解距离法可以根据熵权计算样本的综合排名。

二、基于箱根模型和布莱克标准的 TOPSIS-熵权法现代化评价模型

（一）模型简介

TOPSIS 法（Technique for Order Preference by Similarity to an Ideal Solution），又称优劣解距离法，是 C.L.Hwang 和 K.Yoon 于 1981 年首次提出，TOPSIS 法根据有限个评价对象与理想化目标的接近程度进行排序，是在现有的对象中进行相对优劣的评价。

熵权法是一种客观赋值方法。在具体使用的过程中,熵权法根据各指标的变异程度,利用信息熵计算出各指标的熵权,再通过熵权对各指标的权重进行修正,从而得到较为客观的指标权重。

（二）研究各省份电动汽车现代化程度

基于布莱克标准和箱根模型的六因子 TOPSIS-熵权法对中国各省份现代化程度评价。

本文研究了前人的现代化评价标准的布莱克标准和箱根模型，摒弃 GDP 经济现代化的西方资本主义陷阱^[9]，本文综合两者，提出选用：指示绿色环保等级的交通行业碳排放量 CETIF（Carbon emissions from the transportation industry factor）、指示科技水平的新能源汽车因子 NEVF（New energy vehicle factor）、指示基建配套水平的充电桩因子 CSF（Charging station factor）、指示发展潜力的人口因子 PF（Population factor），利用线性回归可得：

$$y = k_j x + b_j \quad (1)$$

拟合各自的数量 b 和增速 k ，共获得八项指标 CETIF_k, CETIF_b, NEVF_k, NEVF_b, CSF_k, CSF_b, PF_k, PF_b。其中截距 b 取同一时间（2020 年 1 月）对应的值。首先研究数据指标选取的合理性。由于部分省份缺失数据（山西、贵州、重庆、海南），因此本文将这些省份剔除，避免空白值影响结果。

1.主成分分析 PCA

对指标进行主成分分析（PCA）以进行数据降维并去除相关性显著的数据。

首先列出数据矩阵，其中共有 n 各省份和 p 个指标。

$$X = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1p} \\ x_{21} & x_{22} & \dots & x_{2p} \\ \dots & \dots & \dots & \dots \\ x_{n1} & x_{n2} & \dots & x_{np} \end{bmatrix} \quad (2)$$

对各类数据进行标准化处理，本文采用最大最小值标准化。

$$x_{ij} = \frac{x_{ij} - x_{jmin}}{x_{jmax} - x_{jmin}} \quad (3)$$

其中， x_{ij} 为第 i 个省份的第 j 项指标， x_{jmin} 是当前指标中数据最小的省份， x_{jmax} 是当前指标中数据最大的省份。

计算协方差 r_{ij}

$$r_{ij} = \frac{1}{n-1} \sum_{k=1}^n (x_{ki} - \bar{x}_i)(x_{kj} - \bar{x}_j) \quad (4)$$

得到协方差矩阵 R

$$R = \begin{bmatrix} r_{11} & r_{12} & \dots & r_{1p} \\ r_{21} & r_{22} & \dots & r_{2p} \\ \dots & \dots & \dots & \dots \\ r_{n1} & r_{n2} & \dots & r_{np} \end{bmatrix} \quad (5)$$

求解协方差矩阵 R 的特征向量 a_j 以及特征值 λ_j

$$a_j = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \dots \\ a_{pj} \end{bmatrix} \quad (6)$$

得到每个特征的贡献度 $rate_j$

$$rate_j = \frac{\lambda_j}{\sum_{k=1}^p (\lambda_k)} \quad (7)$$

计算第 j 个指标所有的样本贡献率之和 $total_rate_j$

$$total_rate_j = \frac{\sum_{k=1}^i (\lambda_k)}{\sum_{k=1}^p (\lambda_k)} \quad (8)$$

八类指标的贡献度分别为：人口数量 65.7%、人口增速 82.74%、电动汽车数量 88.47%、电动汽车数量增速 93.91%、充电桩数量 97.48%、充电桩数量增速 98.91%、碳排放量 99.62%、碳排放量增速 100%

取贡献度达到 85% 以上的指标：充电桩数量、充电桩数量增速、电动汽车数量、电动汽车数量增速、碳排放量、碳排放量增速。

可见，人口并非绿色现代化的主要构成部分，其影响可用其余指标衡量，如新能源汽车数量与人口呈正比，因此新能源汽车数量可以一定程度反映人口数量大小。

由于中国省份众多，为此，本文通过 K-means 聚类将省份按 PCA 得到的六类指标分类。

2.K-means 聚类分析

首先随机设定 k 个省份为数据中心，将所有数据样本 s_x 归类到第 c 个中心 $center_c$

$$class_x = \operatorname{argmin}_c ((center_c - s_x)^2), c \in (1, k)$$

$$s_x = \begin{bmatrix} s_{x1} \\ s_{x2} \\ \dots \\ s_{xp} \end{bmatrix} \quad (9)$$

其中 $class_x$ 为第 x 个样本 s_x 根据 p 个指标被归纳的类别，其归类标准是离当前点最近的聚类中心 $center_c$

更新类别中心点 $center_c$

$$center_c = \frac{\sum_{x=1}^{num_x} (s_x)}{num_x} \quad (10)$$

其中， num_x 为类别 x 中含有的样本数，重复计算，直至样本与数据中心的距离平分 $(center_c - s_x)^2$ 收敛为定值。

依次增大 k 的取值，并作图如下

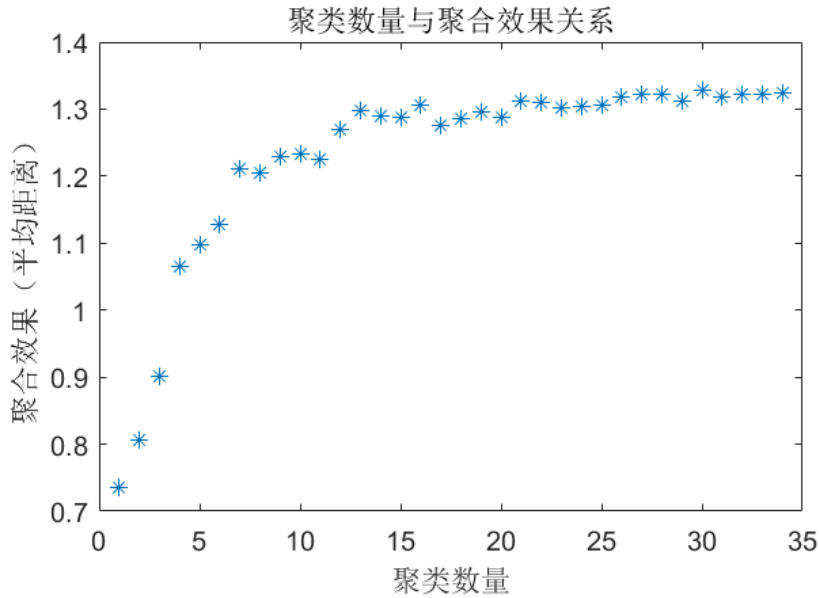


图 2 聚类中心数量和聚类效果图

其中，根据肘部法则确定 $k=4$ 时为聚类拐点，因此可将各省份分为 4 类。

湖南省为第 1 类。该类的特点是新能源汽车发展缓慢，普及低，配套设施（充电桩）数量不足。

云南、广西、江西、江苏、陕西、甘肃、内蒙古、辽宁、吉林、黑龙江为第 2 类。该类的特点是碳排放低，但同样地，充电桩数量也偏少，这可能是由于地

区经济不发达，处于环境库兹涅茨曲线左侧导致的。

四川、湖北、安徽、河南、河北为第 3 类。该类的特点是碳排放逐年增长，同时新能源汽车和充电桩数量也逐年增长，处于发展高峰期。

广东、福建、浙江、山东、北京、天津、宁夏、青海、西藏、新疆、上海为第 4 类。该类的特点是碳排放较高但逐年降低，新能源汽车和充电桩普及加快，处于越过环境库兹涅茨曲线顶峰（即实现了省内局部碳达峰）

山西、贵州、重庆、海南省、港澳台为第 5 类，共 7 个省份存在数据缺失，本文不作研究。



图 3 中国各省份聚类结果

3.熵权法-TOPSIS 评价法

数据标准化使用最大最小值法

$$X_{ij} = \frac{x_{ij} - x_{jmin}}{x_{jmax} - x_{jmin}} + 0.0001 \quad (11)$$

为避免最小项为 0 影响到后续的除法运算，因此在所有数据后加入极小项

0.0001 计算因子熵值 H_j ：

$$\begin{cases} H_j = -\frac{1}{\ln(n)} \sum_{i=1}^n f_{ij} \ln(f_{ij}) \\ f_{ij} = \frac{x_{ij}}{\sum_{i=1}^n x_{ij}} \end{cases} \quad (12)$$

计算各因子熵权 w_j :

$$w_j = \frac{1 - H_j}{\sum_{j=1}^p (1 - H_j)} \quad (13)$$

得到六个因子的权重分别为 $w_{CETIF_k}=0.182712$, $w_{CETIF_b}=0.154404$,

$w_{NEVF_k}=0.182429$, $w_{NEVF_b}=0.139092$, $w_{CSF_k}=0.159975$, $w_{CSF_b}=0.181387$

TOPSIS 标准化:

$$x_{ij} = \frac{x_{ij}}{\sqrt{\sum_{i=1}^n x_{ij}^2}} \quad (14)$$

构造加权规范阵:

$$\hat{x}_{ij} = w_j \cdot x_{ij} \quad (15)$$

计算解间距离 d_i^- 和 d_i^+ , 得到第 i 个类别的得分 S_i :

$$\begin{cases} d_i^- = \sqrt{\sum_{j=1}^p (x_{ij} - x_{jmin})^2} \\ d_i^+ = \sqrt{\sum_{j=1}^p (x_{ij} - x_{jmax})^2} \\ S_i = \frac{d_i^-}{d_i^- + d_i^+} \end{cases} \quad (16)$$

评价结果如下: 类别 1: 67.37904; 类别 2: 49.52626; 类别 3: 96.15002;
类别 4: 100。

第 1 类得分较低是由于新能源汽车发展缓慢, 普及低, 配套设施 (充电桩) 数量不足。

第 2 类得分最低, 这类的碳排放低, 但同样地, 充电桩数量也偏少, 这可能

是由于地区经济不发达，处于环境库兹涅茨曲线左侧导致的，没有达到现代化标准，只是达到了“绿色环保”的要求。

第 3 类特点是碳排放逐年增长，这表明该类正在靠近碳排峰值，同时新能源汽车和充电桩数量也逐年增长，处于发展高峰期。

第 4 类碳排放较高但将逐年降低，这是由于新能源汽车和充电桩普及加快，省份已越过环境库兹涅茨曲线顶峰(即实现了省内局部碳达峰)，正在向高 GDP、低碳排放的绿色现代化方向发展。

本文定义越过碳峰值或处于碳峰值的省份为现代化程度较高的省份，按得分依次赋予 100%、75%、50%、25%现代化程度于各个省份，计算中国国内现代化程度

$$modern = \frac{100\%num_4 + 75\%num_3 + 50\%num_1 + 25\%num_2}{34 - num_5} \quad (17)$$

由于中国共有 34 个省份而有 7 个省份数据缺失，因此 $num_5=7$

计算可得 $modern = 65.741\%$ ，这表明国内近 6 成以上的地区进入了“绿色现代化”。

(三) 研究电动汽车产业链的现代化程度

根据统计数据，每辆车平均每天要提供 9 千瓦时的能量，而一桶原油可提供 1700 千瓦时的能量，即平均每将一辆传统燃油汽车换成电动汽车，可减少 $\frac{9}{1700 \times 24} = 0.00028$ 桶每天，按照国际标准，一桶原油重 0.13664 吨，换算为吨每年即 $0.00028 \times 0.1364 \times 365 = 0.01396008$ 吨每年，根据全中国电动汽车统计数据，共 1310 万辆电动汽车正在被使用，即每年为中国减少 182,877.048 吨石油消耗。

根据 2022 年国际石油油价 99.04 美元/桶，即 699.133264 元/桶，将石油消耗量换算为进口金额为 17439480.308 元。

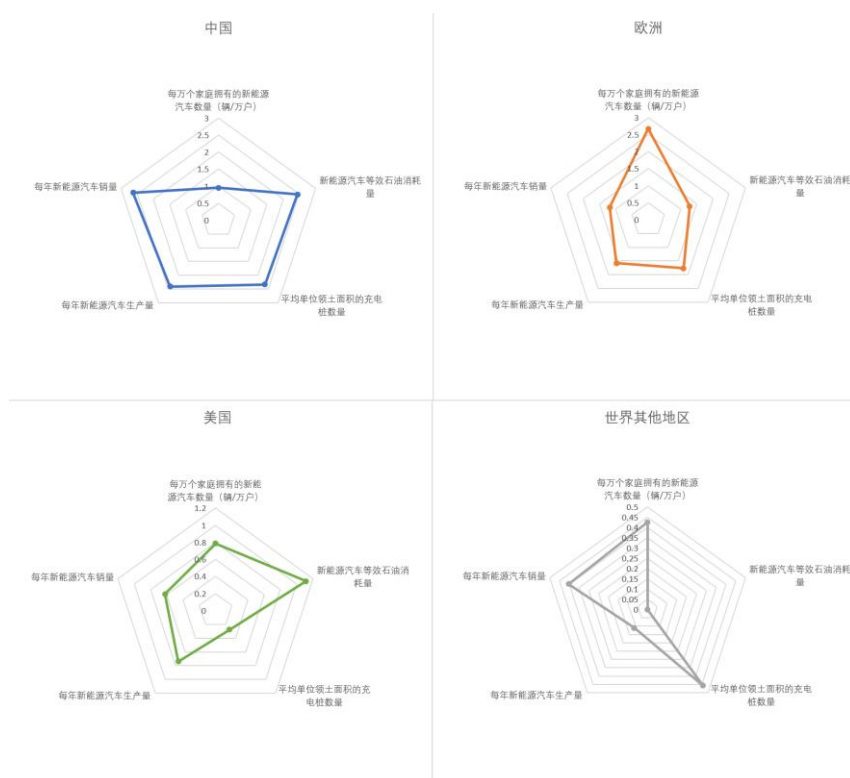


图 4 全球各国家地区新能源汽车指标雷达图

2022 全球电动汽车销量破 1000 万（含纯电+插混车型），国产新能源汽车占据总销量 60%，继续蝉联全球第一。新能源汽车行业为中国带来了 2100 亿元的收益，占总 GDP 的 0.173525%

国内外品牌全球市场新能源汽车市占率

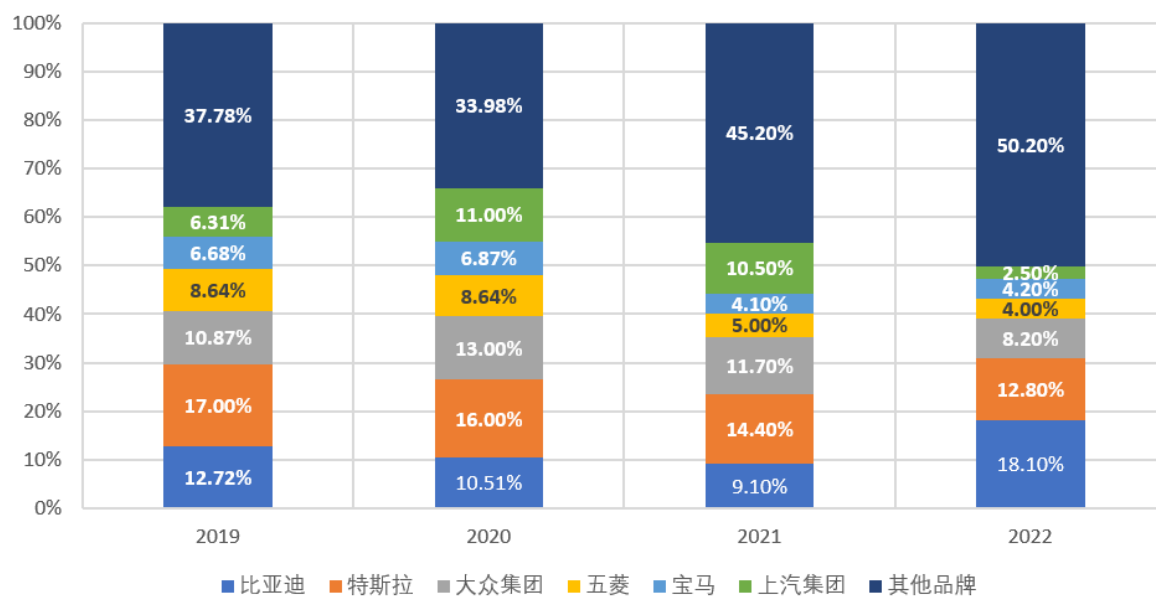


图 5 2019-2022 年全球新能源汽车市场占有率柱状图

由于 2020 和 2021 年，中国受疫情冲击，国产新能源汽车龙头企业比亚迪

的市占率受到一定程度的影响，但随着政策开放，在 2022 年迎头赶上，甚至超过了 2019 年的占有率达 6 个百分点。

中国的新能源汽车起步时间早，在钱学森院士的建议下，中国较早完成了布局，并颁布政策大力支撑新能源汽车发展，积极鼓励技术创新，国产新能源汽车在海外市场有了较高的认可度，在 2022 年售出 1065 万辆，占据了 63%的海外市场。

三、基于马尔可夫状态转移矩阵预测电动汽车发展

（一）构建马尔科夫状态转移矩阵

由于新能源汽车持有家庭数量大，在统计学上可认为频数约等于概率，得到马尔科夫状态转移矩阵

$$M = \begin{bmatrix} p_{11} & p_{12} & \dots & p_{1j} \\ p_{21} & p_{22} & \dots & p_{2j} \\ \dots & \dots & \dots & \dots \\ p_{i1} & p_{i2} & \dots & p_{ij} \end{bmatrix} \tag{108}$$

根据马尔可夫状态转移矩阵，可得到不同状态间的转移概率，进而得到 SIRD 模型的状态转移参数。基于数据库，本文得到 2022 年的马尔科夫状态转移矩阵如下所示

M	持有电动车I	未持有电动车S	曾经拥有电动车R	保有当前的电动车数量D	不曾也不会持有电动车N
持有电动车I	30.8%	/	/	69.2% γ	/
未持有电动车S	33% α	67%	/	/	/
曾经拥有电动车R	/	100%	/	/	/
保有当前的电动车数量D	35% ε	/	10.83% β	54.17%	/
不曾也不会持有电动车N	/	/	/	/	100%

图 6 马尔科夫状态转移矩阵

（二）建立动态 SIRD-NM 模型

由于新能源汽车的普及现象与病毒传染相近，本文将传染病模型和新能源汽车普及相结合，由于传统模型无法描述同一个病人同时患两种病毒，而居民可以同时拥有两辆汽车，因此本文基于传统模型提出了动态总人口下的 SIRD-NM 模型，对参数提出了新的定义：

S: 未购买新能源汽车的家庭，随时间变化有部分家庭会选择购入新能源汽车，进入状态 I；

I: 购买了新能源汽车的家庭（这段时间内的买车家庭有一段的推广效应时间，不会立刻买下一辆新能源汽车或立刻卖掉，在这段时间过后 I 类家庭会分为 RD 两类家庭）；

D: 保有当前数量的新能源汽车的家庭，并随时间变化有部分家庭会购买下一辆新能源汽车，回到状态 I；

R: 随时间变化有部分家庭会卖掉新能源汽车，回到未拥有新能源汽车的状态 S；

N: 不曾买也不会买新能源车的家庭，这部分家庭可能是由于家中已拥有燃油汽车或者无力购买新能源汽车等现实原因，保持在对模型无明显影响的状态 N；

M: 总家庭数量，被分为 SIRDN 五类家庭，且随时间变化，总家庭数将不断变化。

示意图如下：

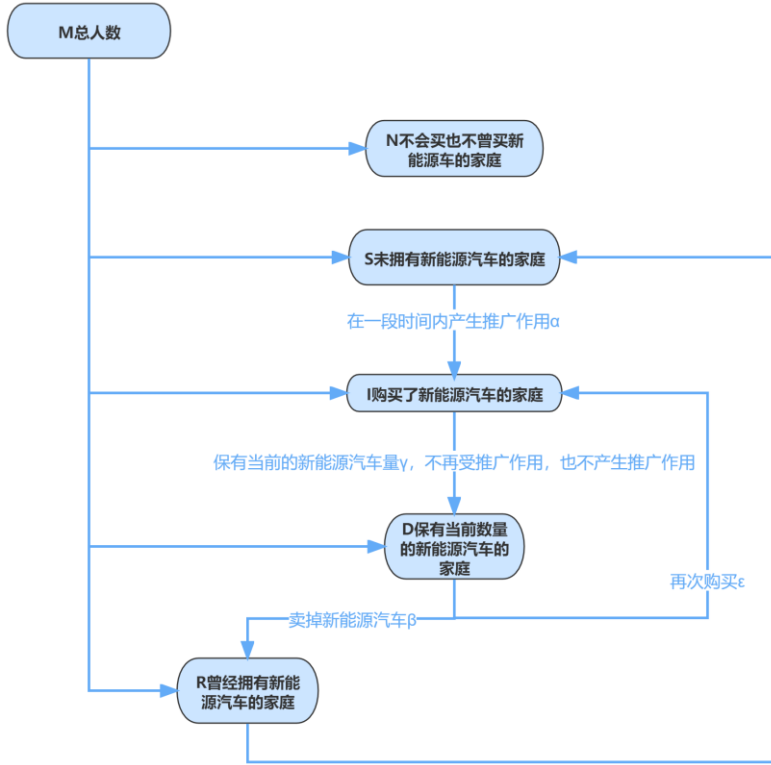


图 7 改进后的 SIRD-NM 模型示意图

根据上图可得到改进后的 SIRD 模型

$$\begin{cases}
 \hat{M} = M - N \\
 \frac{dS}{dt} = -\frac{\alpha I}{\hat{M}} \\
 \hat{S} = S + dR \\
 \frac{dI}{dt} = \frac{\alpha I}{\hat{M}} + \epsilon \hat{D} - \gamma I \\
 \frac{dD}{dt} = \gamma I - \epsilon \hat{D} \\
 R = \beta \hat{D} \\
 \hat{D} = D - dR
 \end{cases} \quad (19)$$

其中， \hat{M} 为去除不曾买也不会买新能源汽车的家庭 N 后的总家庭数。

由于中国每年的家庭数量一直在变化，因此作总家庭数量折线图如下：

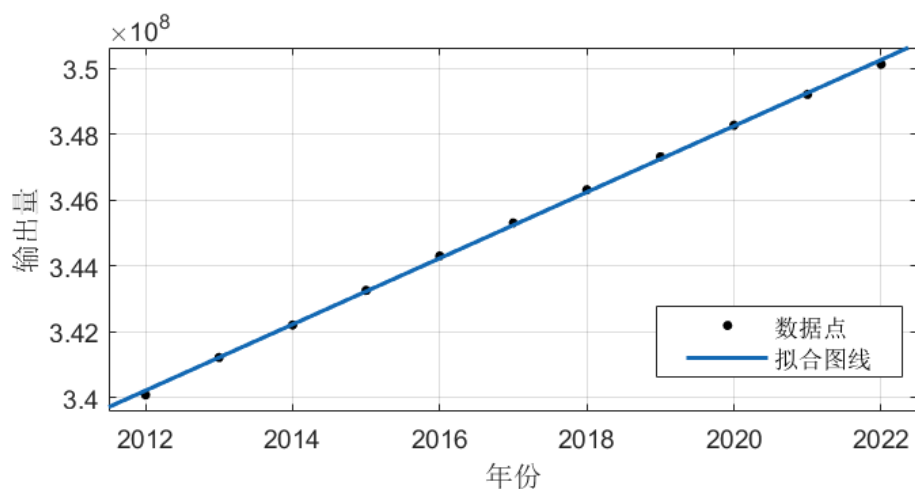


图 8 M 和 t 的拟合结果

可见，M 和 t 存在线性关系，利用线性回归：

$$M = kt + b \quad (11)$$

得到 $k=1.0043e+06$ ， $b=1.6804e+09$ ，方差 $R=0.9995$ 。

利用 SIRD-NM 模型计算新能源汽车保有量：

$$car = (1 + \epsilon)D + I \quad (21)$$

将模型代入 2022 年的初值，并用（19）替换 SIRD 中的总家庭数 M，可得 2030 年的新能源汽车数量数据。

表 1 SIRD 模型输出和真实值对比

年份	RIDS 模型输出	直线拟合输出	差值
2012	321400000	321384161	15840
2013	319400000	319386618	13382
2014	319700000	319674211	25789
2015	320000000	320003708	-3708
2016	320300000	320310705	-10705
2017	320600000	320594202	5798
2018	320900000	320895963	4037
2019	321200000	321204341	-4341
2020	321500000	321488993	11007

使用 Simulink 仿真得：

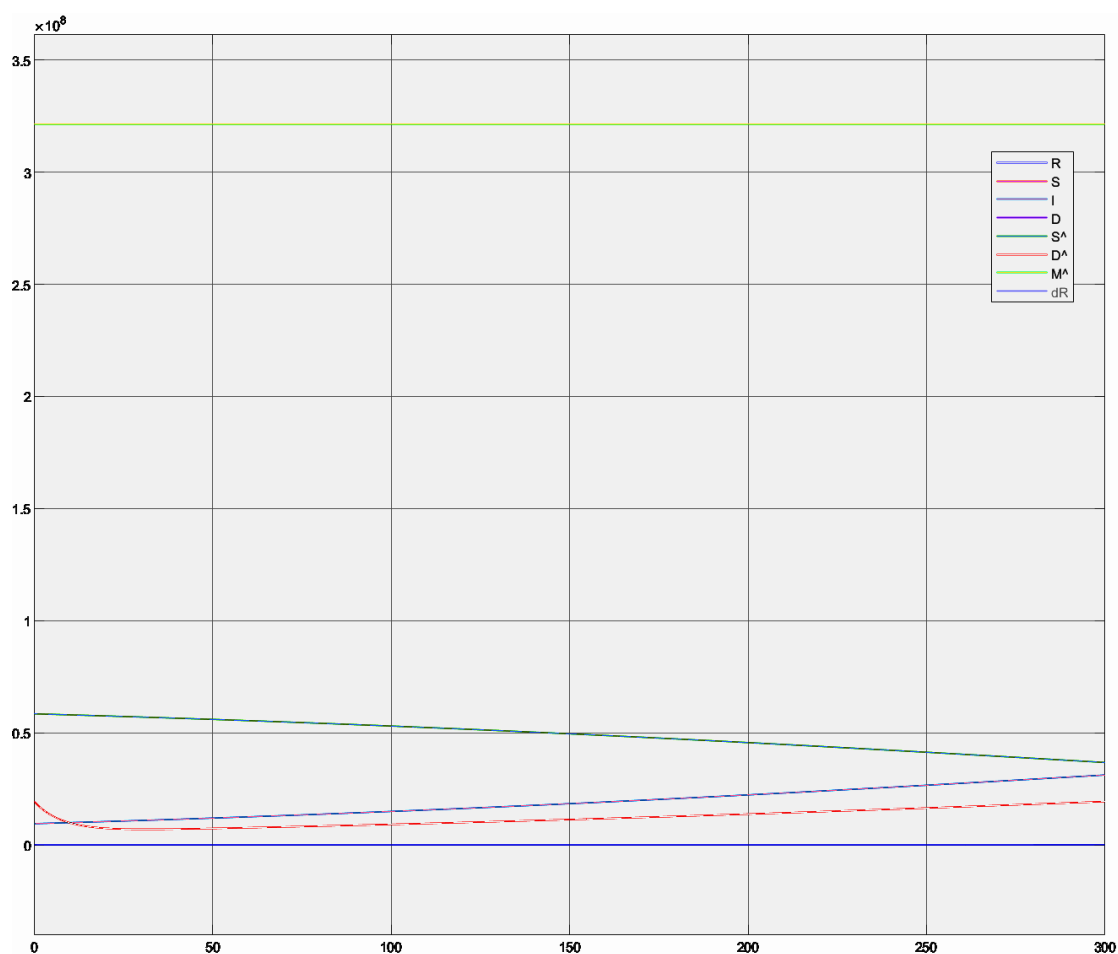


图 9 SIRD-NM 模型预测未来新能源汽车持有人群曲线图

可见，购车人群将逐步提升而非购车人群逐渐减少，减去不会购买新能源汽车的家庭 N ，随着时间发展，新能源将在 7~10 年内普及。

（三）基于深度学习的新能源汽车数量预测模型

由于新能源汽车数量与多种因素相关，难以通过单一的模型模拟，预测本文基于 2 层 BP 神经网络对数量进行预测，首先将数据按时间切片，利用 60% 的数据进行训练，再利用 20% 的数据进行验证，20% 的数据进行测试。网络结构示意图如下

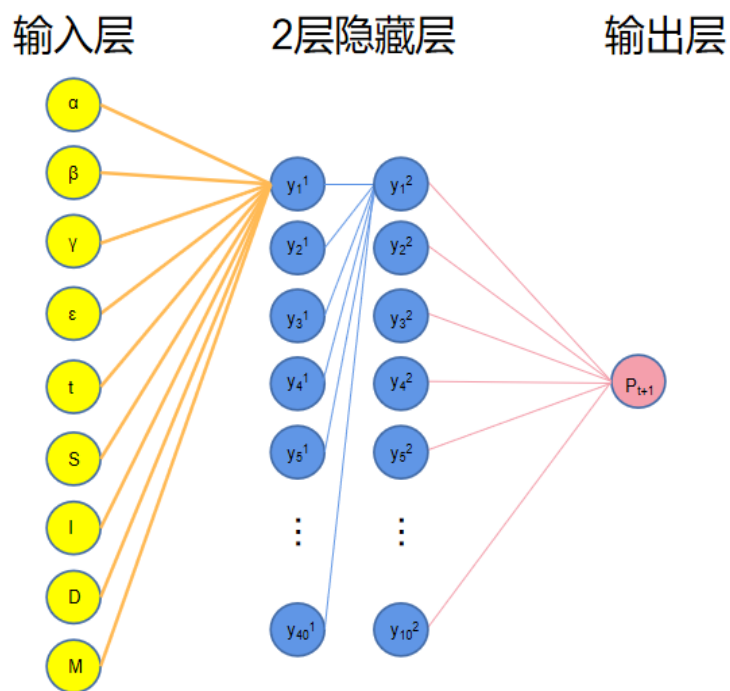


图 10 两层多神经元深度学习预测新能源汽车数量网络结构示意图

输入状态转移概率以及时间、各类人群初始值，通过一层有 40 个神经元的隐藏层和一个有 10 个神经元的隐藏层，得到相应年份 t 的新能源汽车数量。

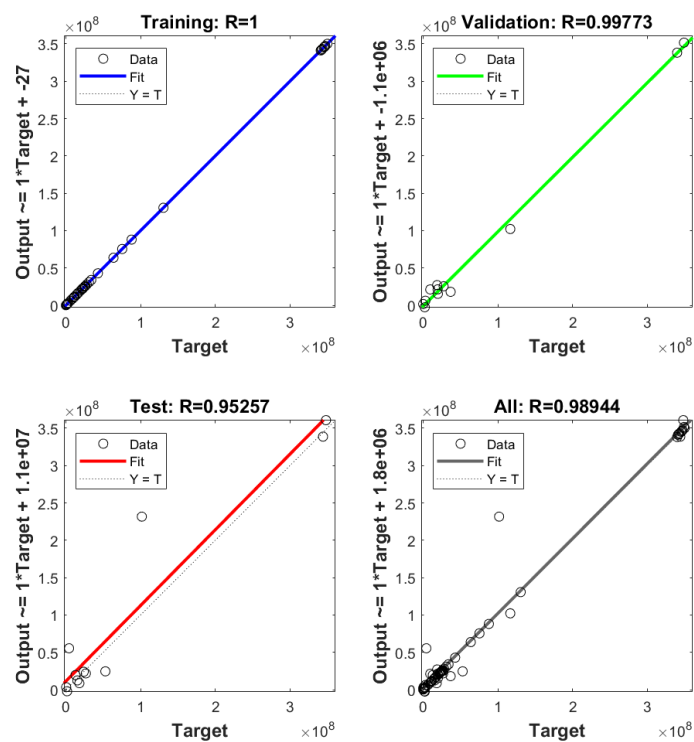


图 11 拟合效果图一

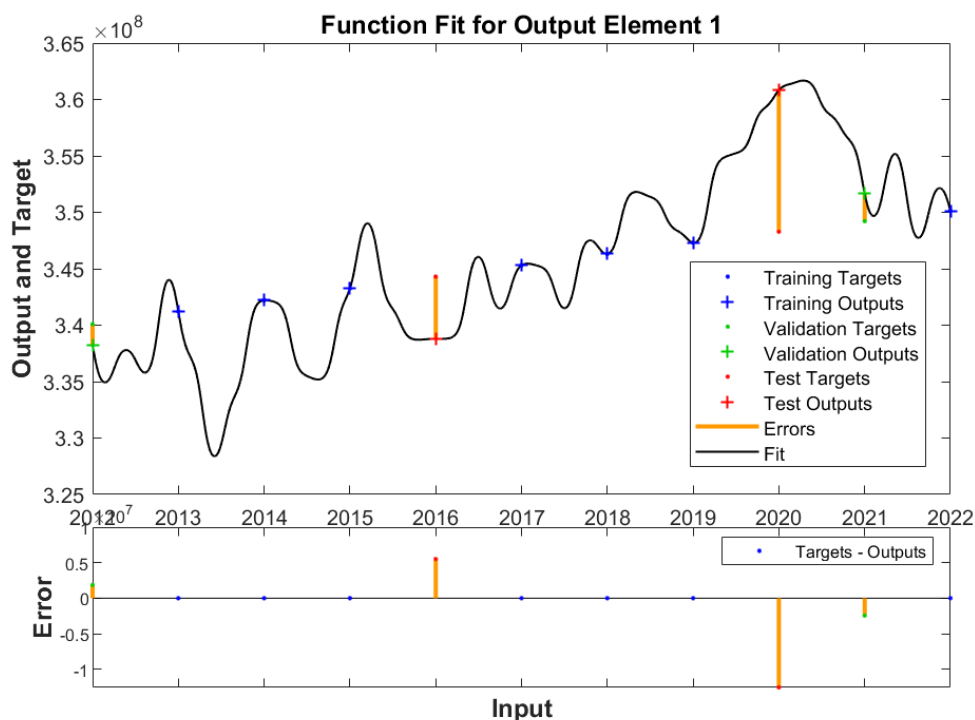


图 12 拟合效果图二

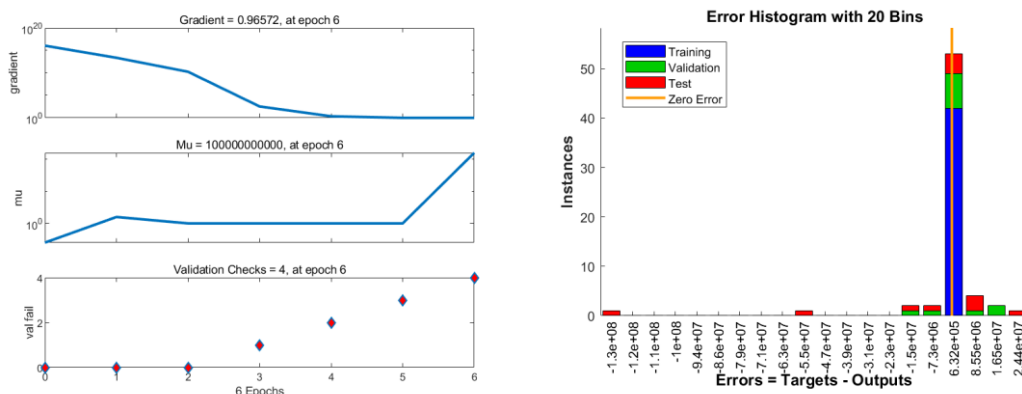


图 13 均方误差图（左）

图 14 验证集收敛梯度（右）

由于中国家庭数量在百万数量级，所以均方误差除以 10^{10} 后为标准化后的真实误差，最大方差 $R_{max}=1.5$ 。

在神经元数量为 40 时，效果比 30 和 50 个神经元要好，在数量过少时，神经网络无法捕捉并拟合影响新能源汽车数量的隐含因子，而神经元数量过多时会导致过拟合使得神经网络适应度差，无法适应输入不同的情况。

（四）神经网络时间序列预测

基于泛用性广但训练速度较慢的贝叶斯正则化法训练，直接得到新能源汽车数量，其误差如下：

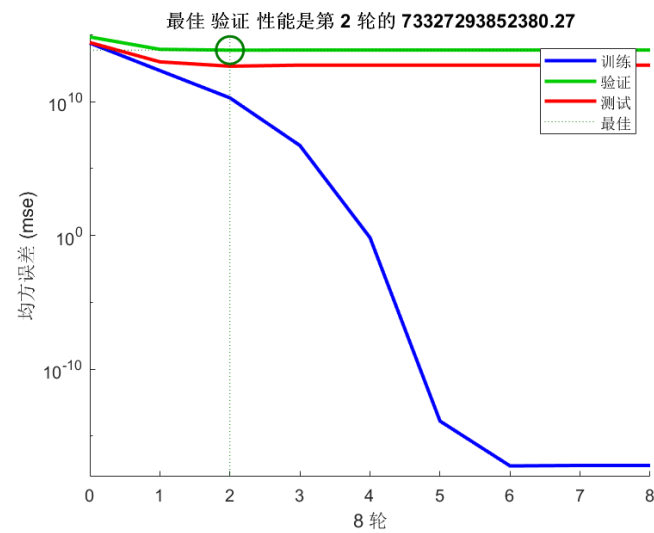


图 15 最佳预测性能校验

在训练时误差下降较快，而对验证集和测试集表现均较差。

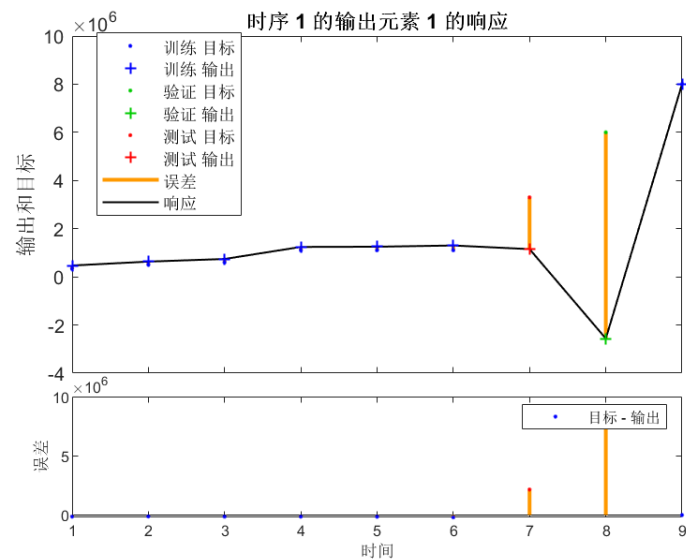


图 16 时序响应误差

可见，在时间接近 2019 时误差过大，而在 2019 年以前误差几乎为 0，可以判断该神经网络出现了过拟合现象，因此，本文采用深度神经网络进行新能源汽车数量的预测。

四、基于环境库兹涅茨曲线对 2023 碳达峰的可实现性验证

（一）“倒 U 型”环境库兹涅茨曲线

Grossman 和 Krueger 认为“倒 U 型”结构的形成会受到规模效应、技术效应与结构效应三种途径的影响，具体来说：

1. 规模效应

根据环境经济学的相关理论，新能源产业会从两个方向对生态环境质量产生影响。例如，在经济发展初期，普遍借助于大量化石能源燃烧的方式，以实现经济的高速增长，最终导致能源被大量消耗，生态环境被破坏；另一方面，随着生产不断提速，经济发展对资源的消耗量也日渐增加，进而导致生态环境面临更大威胁。

2. 技术效应

通常来说，居民的高收入与高环保技术水平呈正比例关系。例如，在国家整体技术进步提高后，会产生两方面的影响：一方面，技术水平提高将会提高能源利用效率，减少单位产出所需的能源消耗，从而大幅度降低生产对自然和环境的影响；另一方面，清洁能源技术的开发和使用，将大量替代污染严重的技术，从而降低单位产出的污染排放量。

3. 结构效应

伴随居民收入水平的提高，整体产业结构和投入结构也会相应产生改变。比如，在经济发展早期阶段，经济结构重心在农业，后来随着工业革命的推进，工业的流水线式生产在扩大了生产效率的同时，也加重了对环境的污染。此后，经济进一步发展，向第三产业的服务业和知识密集型产业转移，单位污染排放下降，环境质量开始改观。

根据国家统计局发布的数据，根据 1999 年-2019 年的 GDP 数据，以及相应年份的碳排放量数据，可作图如下

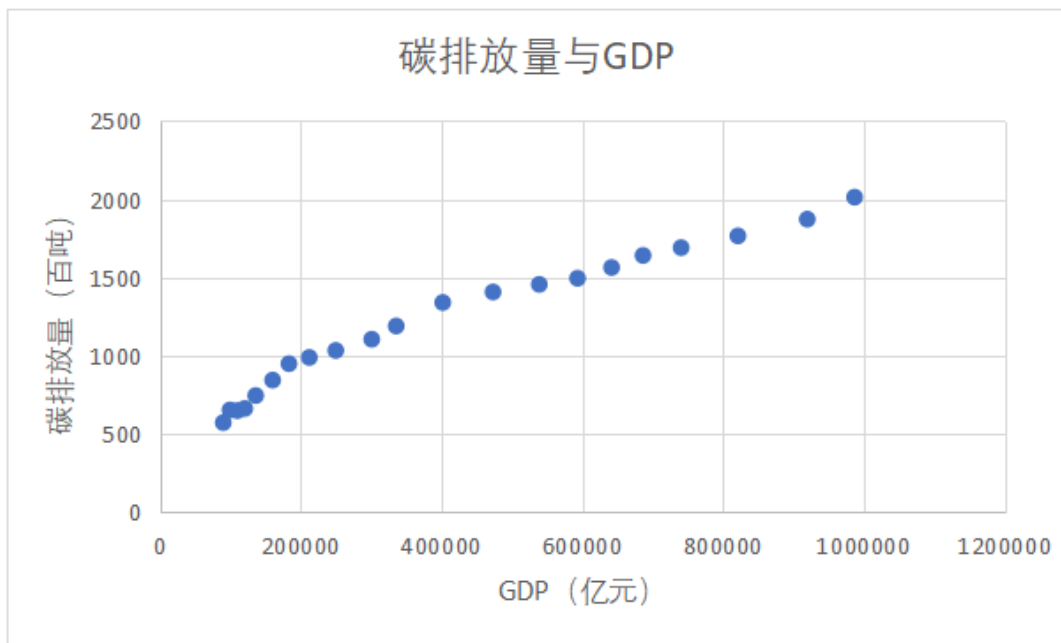


图 17 1999 年到 2019 年碳排放量与 GDP 关系散点图

可见，截止 2019 年中国仍然处于环境库兹涅茨曲线峰值左侧，是经济高速发展同时也是碳排放最严重的时候，为了越过碳排放峰值同时尽量提升 GDP，新能源汽车的普及和推广是必要的。

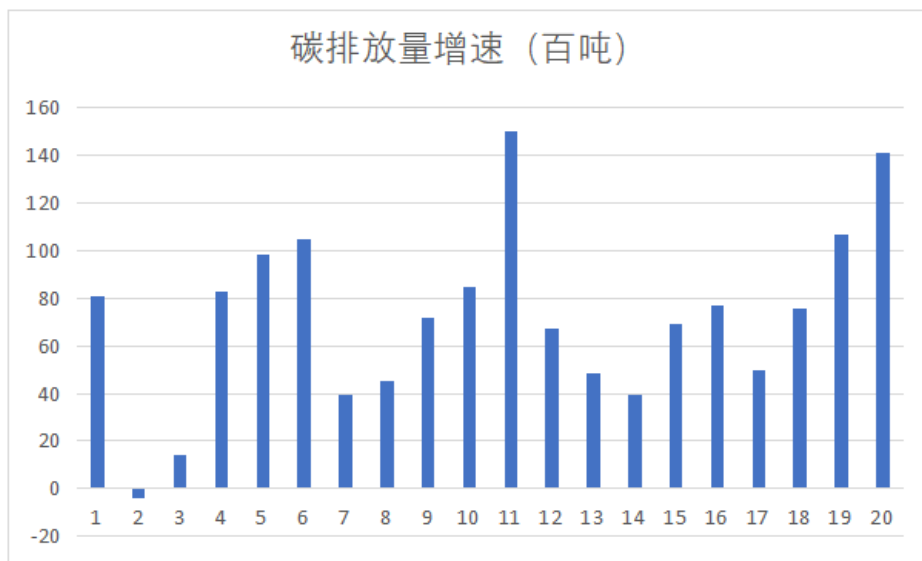


图 18 碳排放量增速和时间的关系

可见，仅在 2002 年出现了碳排降低，其余年份的碳排放量均在持续升高，可见，我国在 2019 年仍处于高速经济发展中，仍未达到碳排峰值 GDP，因此，提升碳吸收等级和降低碳排放等级是必要的。

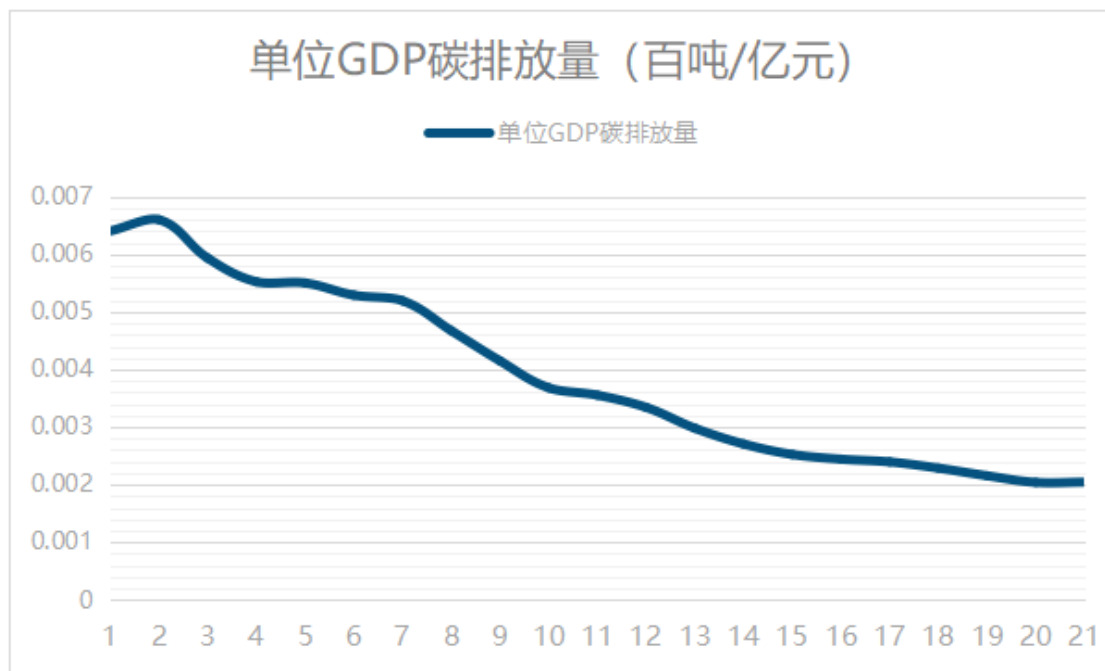


图 19 单位 GDP 碳排放量

从上图可见，若保持碳吸收等级不变，中国继续发展下去，单位 GDP 碳排放量将达到定值，若加强植树造林等活动提高碳吸收等级，单位 GDP 碳排放量将进一步降低。

（二）“2030”碳达峰实证分析

本文对 2030 年的新能源汽车数量进行预测，以检验“碳达峰”实现度，同时作保守假设：国内碳吸收量保持不变，即碳达峰仅取决于碳排放量。

根据碳排放趋势线，若新能源汽车数量始终不变，可得 2030 年预计碳排放量约为 2580.14 百吨，而在 2030 年根据 SIRD 模型仿真得到有 15929801 辆新能源汽车，而深度学习神经网络预测得到有 17276748 辆，两者结果相差不大，可认为结果较为可信。

而平均每辆车每年可减少碳排放 780.1 kg，则 2030 年的新能源汽车数量可降低交通行业的碳排放量至 0，而火力发电厂的碳排放量增多，但由于火力发电厂的处理设施完善，因此可以减少发电产生的二氧化碳逸散到大气。中国的“2030 碳达峰”目标是可行的，但前提是其他条件保持不变甚至减少碳排放。

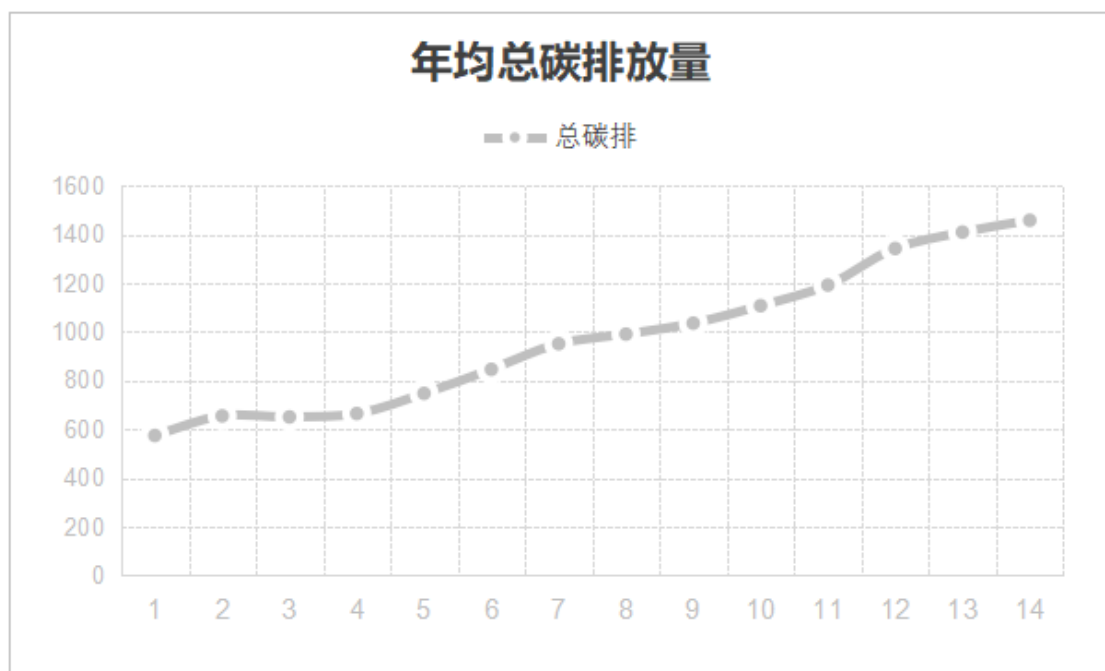


图 20 年均总碳排放量（百吨）

假设我国对新能源汽车产业进行补贴、并对高碳排车辆收取碳税，执行力度可以提高转移概率 α ，可得到新政策情况下的 2030 年碳排放情况。

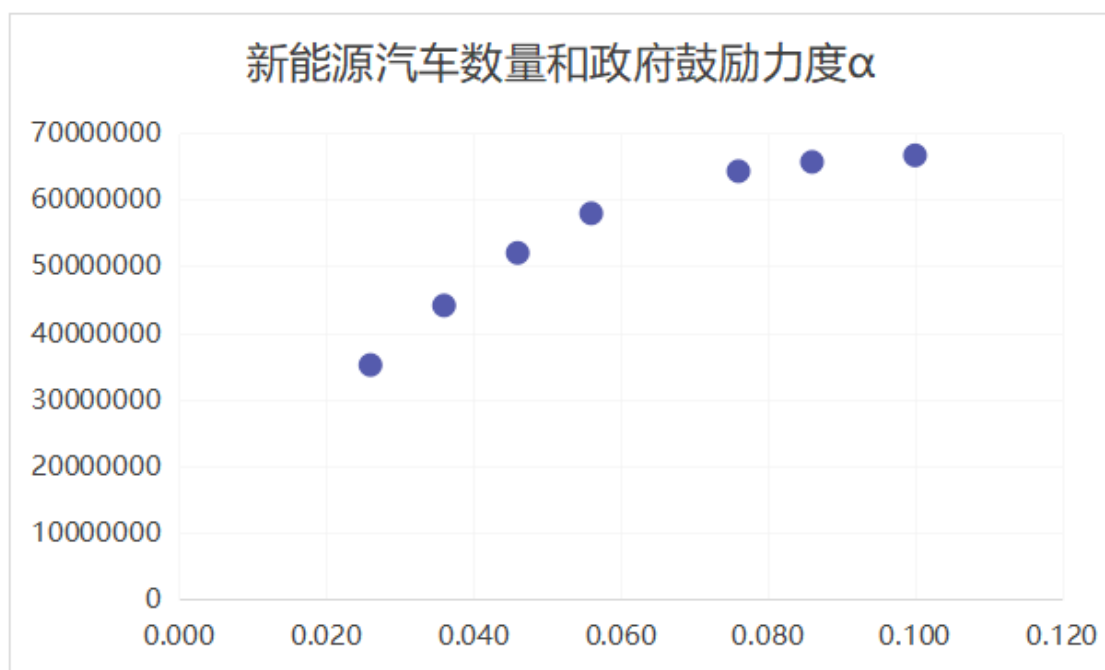


图 21 新能源汽车数量和政府鼓励力度 α 变化关系图

五、结论与建议

（一）建议政策

提高新能源汽车补贴，降低新能源购置税，以提高新能源汽车的市场竞争力，减少高能量密度电池和先进电动机的成本负担，在引进国外先进技术的同时支持国产自主创新，提高海外市场竞争实力。

政府部门应继续加大对新能源汽车产业的支持力度，加大补贴力度，同时加强对新能源汽车的技术研发和市场推广，推动新能源汽车产业的发展。提高碳税，减少碳排放量同时提高新能源汽车普及率。

鼓励企业加强创新和技术研发，提高新能源汽车的技术含量和品质，提高产品性能和市场竞争力。

优化新能源汽车的技术体系和产业链，加强新能源汽车的生产和销售环节的协调，提高供应链的效率和灵活性。

促进新能源汽车的推广和普及，加强对新能源汽车消费者的教育和宣传，鼓励消费者购买新能源汽车，提高新能源汽车的普及率。

加强新能源汽车产业的国际合作和交流，促进技术和市场的共享，提高新能源汽车产业的国际竞争力。

（二）模型的不足

TOPSIS-熵权法综合评价的主观性和权重的确定难度；

神经网络模型需要大量数据进行训练，并可能存在过拟合问题；

时间序列预测模型需要考虑多种因素，且预测精度可能受到多种因素的影响；

产业对比需要考虑多种因素，如政策环境、市场需求、技术水平等，需要进行合理的数据选择和加工，以及进行科学的分析和比较。

参考文献

- [1] 魏文强.碳中和背景下新能源汽车产业发展研究[J].时代汽车,2023,No.402(06):87-89.
- [2] 张振丽.“双碳”背景下新能源汽车的机遇与挑战[J].专用汽车,2023,No.310(03):10-12.DOI:10.19999/j.cnki.1004-0226.2023.03.004.
- [3] 李晓易,谭晓雨,吴睿,徐洪磊,钟志华,李悦,郑超蕙,王人洁,乔英俊.交通运输领域碳达峰、碳中和路径研究[J].中国工程科学,2021,23(06):15-21.
- [4] 刘建国,朱跃中,田智宇.“碳中和”目标下 我国交通脱碳路径研究[J].中国能源,2021,43(05):6-12+37.
- [5] 袁博.碳中和目标下中国新能源汽车产业发展[J].管理工程师,2022,27(05):5-10.DOI:10.19327/j.cnki.zuaxb.1007-1199.2022.05.001.
- [6] Fernández-Villaverde J, Jones C I. Estimating and simulating a SIRD model of COVID-19 for many countries, states, and cities[J]. Journal of Economic Dynamics and Control, 2022, 140: 104318.
- [7] Calafiore G C, Novara C, Possieri C. A time-varying SIRD model for the COVID-19 contagion in Italy[J]. Annual reviews in control, 2020, 50: 361-372.
- [8] Nisar K S, Ahmad S, Ullah A, et al. Mathematical analysis of SIRD model of COVID-19 with Caputo fractional derivative based on real data[J]. Results in Physics, 2021, 21: 103772.
- [9] 陈柳钦, 国内外现代化指标体系和标准概述, 天津社会科学院城市经济研究所, 天津 300191
- [10] https://www.sohu.com/a/663838103_121474359 比亚迪股份有限公司董事长兼总裁王传福发言
- [11] https://roll.sohu.com/a/622352916_121622941 钱学森院士写给时任国务院副总理邹家华的信
- [12] http://tj.nmg.gov.cn/tjdt/fbyjd_11654/202209/t20220908_2128980.html
- [13] http://k.sina.com.cn/article_7031151442_1a316db5200100oebp.html
- [14] <https://zhidao.baidu.com/question/1869441046182904067.html>

- [15] <https://baijiahao.baidu.com/s?id=1725640141334665315&wfr=spider&for=pc>
- [16] [2021 中国用户新能源车消费决策和态度调研报告 腾讯新闻 \(qq.com\)](#)
- [17] [2020 年交通运输行业发展统计公报 部门政务 中国政府网 \(www.gov.cn\)](#)

致谢

作为大二学生，我们很荣幸能够参加全国大学生统计建模大赛，也很高兴能够与我的队员们一起完成这篇论文。在此，我要由衷地感谢我的队员们，没有他们的支持和帮助，这篇论文是无法完成的。在论文的选题、数据收集、实证模拟以及最终的论文撰写和修改过程中，我们不断地讨论、分析、相互协作，每个人都尽最大的努力，共同完善这篇论文。正是有了我们的集体智慧、彼此的鼓励和支持，这篇论文才得以顺利完成。在此，我要向我的队员们表示最真挚的感谢！

同时，我们也要感谢指导老师。老师在选题、梳理框架等环节中，为我们提供了宝贵的意见和指导，帮助我们在论文编写的大方向上更加明确。在实证研究中遇到问题时，老师也给予了我们耐心指导和帮助。在此，我要向指导老师表示最衷心的感谢！这次比赛让我深刻体会到了团队合作和指导老师的重要性，我会倍加珍惜这次经历，不断提高自己的能力。

在这次比赛中，我们遇到了很多困难和挑战，但是通过团队的合作和努力，成功地克服了这些困难。在这个过程中，我也学到了很多新的知识和技能，这将对我的学习和未来的发展产生积极的影响。

最后，我们还要感谢比赛的组织者和评委们，感谢他们为我们提供了这个学习和交流的平台。比赛不仅是一次展示自己才华的机会，更是一次学习和成长的机会。我相信，通过这次比赛，我和我的队员们都会更加努力地学习和提高自己，为未来的发展打下坚实的基础。

2023 年 5 月 26 日

附录

附录 1

介绍：支撑材料的文件列表

/part1/TopsisMethod.m: Topsis 评价核心代码

/part1/KMeans.mlx: K-Means 聚类核心代码

/part1/k1secs_repaired.mat: 原始数据和被处理的数据

/part1/EntropyMethod.m: 熵权法核心算法

/part1/china_elevation_draw: 中国地图热力图画图代码

/part1/kmeans: K-Means 聚类代码库文件

/part1/pca: PCA 主成分分析代码

/part2/RIDS_function: RIDS 算法核心代码

/part2/opt_SIRD.slx: 改进 RDIS 算法的 Simulink 模型

/part2/TimeSeries_function.m: 时间序列预测神经网络模型函数

数据来源: <https://data.stats.gov.cn/>