

基于GM(1,1)的双碳预测模型

项 辉¹, 吴晨宇¹, 汪 扬¹, 马 凯², 农植标¹

¹上海理工大学机械工程学院, 上海

²上海理工大学出版印刷与艺术设计学院, 上海

收稿日期: 2023年5月23日; 录用日期: 2023年7月14日; 发布日期: 2023年7月21日

摘 要

通过建立数理分析模型, 利用源于生物学研究的描述统计法分析长三角地区新能源汽车产业在全国新能源汽车市场的地位及作用。通过建立基于“OLS + 稳健标准误”思想的回归分析模型, 使用Stata 16 SE进行回归分析, 得出新能源汽车产业与传统汽车产业之间存在竞争关系。利用GM(1,1)灰色预测模型, 处理原始时间序列数据后分别得到累加序列、紧邻均值生成序列, 定义灰色微分方程, 构建白化方程, 以最小二乘法的思想得到时间响应表达式, 实现未来3年长三角地区新能源汽车的市场保有量的预测。最后利用历年长三角区域煤炭、原油、天然气、电能的能源消耗量通过换算系数和碳排放系数的转换, 得到长三角地区碳排放总量, 再对四个能源关于时间的拟合曲线进行赋权, 得到组合拟合曲线, 画出碳排放的拟合曲线图, 再减去碳吸收量, 得到碳吸收拟合曲线, 得出长三角地区将在2029年实现碳达峰, 在2061年实现碳中和结论。

关键词

数理分析, GM(1,1)灰色预测模型, 回归分析模型, 组合拟合曲线

Carbon Peaking and Carbon Neutrality's Prediction Model Based on GM(1.1)

Hui Xiang¹, Chenyu Wu¹, Yang Wang¹, Kai Ma², Zhibiao Nong¹

¹School of Mechanical Engineering, University of Shanghai for Science and Technology, Shanghai

²College of Communication and Art Design, University of Shanghai for Science and Technology, Shanghai

Received: May 23rd, 2023; accepted: Jul. 14th, 2023; published: Jul. 21st, 2023

Abstract

Through establishment of mathematical analysis model, take advantage of statistical method de-

rived from biological research to analyze the Changjiang Delta's new energy car industry's position and role in nationwide new energy car market. By building a regression analysis model based on "OLS + ROBUST Standard Error"'s thoughts, take advantage of Stata 16 SE for regression analysis, concluded that there exist a competitive relationship between new energy car and traditional car. Take advantage of GM(1,1) grey prediction model, process the original time data to get cumulative sequence and the adjacent, mean the grey differential equation, establish the winterization equation. The time response expression was obtained by the least square method, to reach the prediction to the Changjiang Delta's market inventory of new energy car. At last, take advantage of Changjiang Delta's coal, crude oil, natural gas and electric energy's consumption, converting conversion coefficient and carbon emission coefficient, and then empower fitted curve which related to time from these four energy to get the combined fitted curve. It is concluded that the Changjiang Delta will achieve carbon peak in 2029 and carbon neutrality in 2061.

Keywords

Mathematical Statistics Method, GM(1,1) Grey Prediction Model, Regression Model, Combined Fitting Curve

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

现如今, 全球面临的最大挑战是气候变化问题。在现代化工业中, 煤炭、石油、天然气的大量使用产生了数量巨大的二氧化碳(CO₂)、二氧化硫(SO₂)、臭氧(O₃)等温室气体是造成全球气候变暖的主要原因。我国实施的“双碳”计划是新时代绿色发展理念的重要组成部分, 力争 2030 年前实现碳达峰, 2060 年前实现碳中和[1]。

学者们利用不同方法对碳排放量进行预测分析, 主要包括 STIRPAT 模型、情景分析法、IPAT 模型、LEAP 模型、蒙特卡洛动态模拟、中国能源环境综合政策评价模型、ARIMA 模型、BP 神经网络组合模型和系统动力学模型。黄蕊等使用 STIRPAT 模型和情景分析法预测了江苏省能源消费碳排放量发现, 当人口和经济低速增长, 技术高速增长时, 2020 年碳排放量预测值为 202.81 兆吨。长三角城市群是中国最大的城市群之一, 已跻身于国际公认的六大世界级城市群。2019 年, 长三角城市群的碳排放量占全国碳排放量的 13.6%, 是中国碳排放重要来源区域。赵成柏和毛春梅利用 ARIMA 模型和 BP 神经网络组合模型, 预测了中国碳排放强度的变化, 结果表明, 2020 年中国碳排放强度比 2005 年下降了 34%。结合长三角地区四省市的发展现状, 参照现行长三角地区各省市国民经济[2], 基于改进的线性回归函数, 利用 GM(1,1)模型, 在原有数据的基础上, 模拟长三角地区 2023~2025 年的碳排放量, 并预测长三角地区碳达峰和碳中和的时间节点。

2. 基于数理分析模型的发展现状研究

2.1. 描述统计分析模型的建立

描述统计源于生物学研究。严格意义上的描述统计学特指上世纪末至本世纪二十年代之间的以高尔顿为先导, 而以卡尔·皮尔逊为代表的用于对生物资料进行分析所提出的一系列统计方法。其后, 不少学者发现生物统计中的一些基本思想、方法和概念也适用于非生物学领域, 尤其在社会、经济及管理领

域中[3]。其通过图表或数学方法，对数据资料进行整理、分析，并对数据的分布状态、数字特征和随机变量之间关系进行估计和描述的方法。

假设第 t 年时，长三角区域内对应的新能源汽车产量为 O_{1t} ，全国范围内对应的新能源汽车产量为 O_{2t} ，定义一个变量：

$$P_t = \frac{O_{1t}}{O_{2t}}$$

其中， P_t 表示第 t 年长三角区域新能源汽车产量与全国范围内新能源汽车产量比值。

再定义一个变量：

$$GR_t = \frac{P_t - P_{t-1}}{\Delta t}, (\Delta t = 1, t = 2, 3, 4, 5, 6)$$

其中， GR_t 表示 t 时刻长三角区域新能源汽车产量在全国占比的增长率。

最后定义一个变量：

$$R_t = \frac{O_{1t} - O_{1t-1}}{|O_{2t} - O_{2t-1}|}, (t = 2, 3, 4, 5, 6)$$

其中， R_t 表示第 t 年长三角区域新能源汽车的生产量增长率相对于全国新能源汽车生产量增长率的比值。

$$\begin{cases} \text{当 } R_t > 1 \text{ 时, 表示长三角区域增长水平高于全国增长水平} \\ \text{当 } R_t = 1 \text{ 时, 表示长三角区域增长水平等于全国增长水平} \\ \text{当 } R_t < 1 \text{ 时, 表示长三角区域增长水平低于全国增长水平} \end{cases}$$

分母的绝对值作用为，当全国新能源汽车生产量增长率小于 0 时，求绝对值可以修正长三角区域与全国新能源汽车产业的生产量增长率比值。

2.2. 描述统计分析模型的求解

随不可再生能源的日渐枯竭以及环境污染的日益严重，世界各国逐渐将目光聚焦于新能源的研究与开发，其中卓有成效的新能源汽车领域，经过三十年的时间从规划到量产，我国的新能源汽车产业取得了不小的成就，从 2009 年数百台的销量到 2020 年 130 万余余台的销量，科技方面也有所突破，相关基础设施也在逐渐的完善过程中[1]。

在中国知网等数据统计网站采集了 2011~2022 年对关键词(平均续航里程、平均指导价、供油国石油储量等)的数据，进行建模分析，用 2011~2020 年的数据用来验证模型的精度，截取部分数据，得到表 1 所示内容：

Table 1. Production data of four provinces of Yangtze River Delta and the whole country in from 2017 to 2022
表 1. 2017~2022 年长三角四省份及全国产量数据

年份	上海(万辆)	浙江(万辆)	江苏(万辆)	安徽(万辆)	长三角地区(万辆)	全国(万辆)
2017	6.14	2.5	2.3	4.9	15.84	79.4
2018	8.8	4.6	5.2	8.4	27.0	127.0
2019	12.56	5.2	16.9	11.6	46.26	124.2
2020	23.86	7.7	10.08	10.5	52.14	130.0
2021	63.2	20.5	25.0	25.2	133.9	354.5
2022	99.0	58.15	68.7	52.7	278.55	705.8

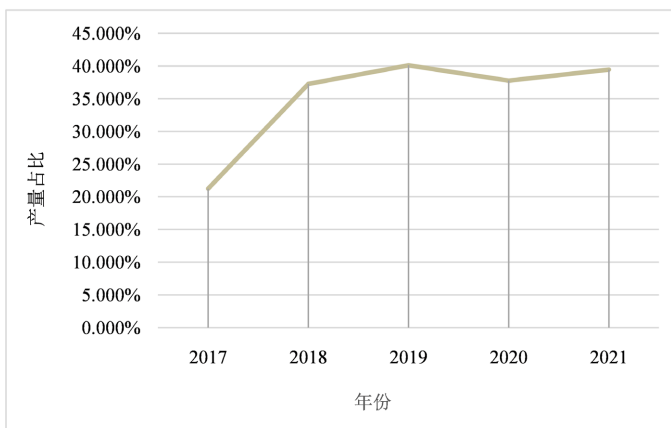


Figure 1. Chart of the ratio between the output of Changjiang Delta new energy car and nationwide new energy car

图 1. 长三角地区新能源汽车产量与全国范围内新能源汽车产量历年比值图

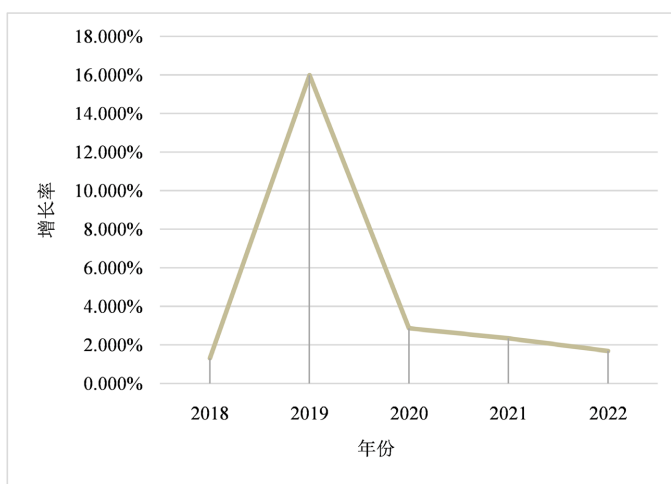


Figure 2. Annual growth rate chart of national proportion of new energy vehicle production in Changjiang Delta

图 2. 长三角地区新能源汽车产量全国占比历年增长率图

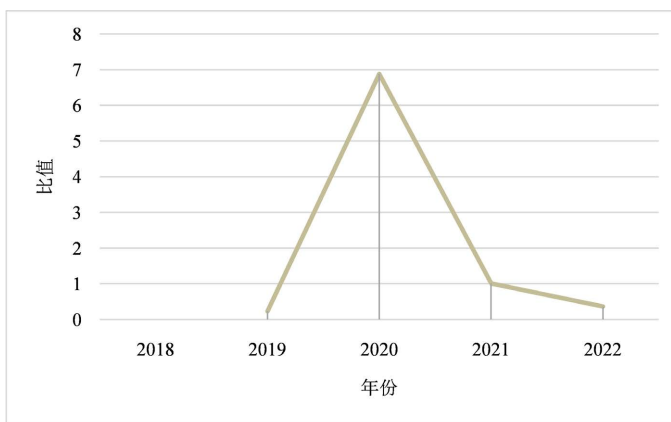


Figure 3. Ratio chart of output of new energy car between Changjiang Delta and the whole country

图 3. 长三角地区与全国新能源汽车的生产量增长率历年比值图

由图 1 可知,长三角地区新能源汽车产量占全国范围内新能源汽车产量比重总体呈现逐年增长趋势;由图 2 可知,历年产量占比增长率呈正数;由图 3 可知,每年长三角地区新能源汽车产业较全国其他地区增长率更高、比重更大、发展速度更快。

2.3. 基于多元线性回归的竞争模型的建立

新能源汽车是我国汽车产业高质量发展的战略选择,也是助力我国实现“双碳”目标的重要举措。新能源汽车行业的快速发展,给传统燃油汽车带来了巨大的挑战。接下来通过建立各因素与新能源汽车产业和传统汽车产业的函数关系,研究我国新能源汽车与传统燃油汽车的市场竞争关系。

分析可知,在理想状态下若只存在一种类型的汽车,则汽车的销量增长量与汽车的生产量,汽车的需求量有关,则得到如下的关系式:

$$d(t) = f(X_{\text{生产量}}, Q_{\text{需求量}})$$

其中, $d(t)$ 表示汽车的销售增长量, $X_{\text{生产量}}$ 表示 t 时刻汽车的生产量, $Q_{\text{需求量}}$ 表示 t 时刻汽车的市场需求量。

进一步,如果同一个市场中存在着两种类型的汽车,则另一个汽车的生产量对第一种类型的汽车增长量有一定的影响,得到新能源汽车增长量的表达式如下所示:

$$d_1(t) = f(f_1(t) + f_2(t) + g(t))$$

其中, $d_1(t)$ 表示新能源汽车的增长量, $f_1(t)$ 表示新能源汽车的生产值, $f_2(t)$ 表示市场汽车需求量, $g(t)$ 表示传统汽车的生产量。

同理,得到传统汽车汽车增长量 $d_2(t)$ 的表达式如下所示:

$$d_2(t) = f(f_1(t) + f_2(t) + g(t))$$

回归模型(linear regression model)是一种特殊的数学回归模型,能够确定变量之间的相关性并且可以针对某一变量进行目的性预测。由于汽车保有量受到多个因素的影响,故建立多元回归模型,具体表达式如下所示:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \mu$$

其中, β_0 为常数项, β_1 、 β_2 、 β_3 称为回归系数, μ 为随机误差项, x_1 、 x_2 、 x_3 为非随机的自变量, y 为随机的自变量[4]。

2.4. 基于多元线性回归的竞争模型的求解

将收集到的 2010~2022 年新能源汽车与传统汽车生产量与销售量的数据如表 2 所示:

Table 2. Table of production and sales data of new energy vehicles and traditional vehicles from 2010 to 2022
表 2. 2010~2022 年新能源汽车与传统汽车产销数据表

年份	新能源汽车产量 (万辆)	传统汽车产量 (万辆)	新能源汽车销量 (万辆)	传统汽车销量 (万辆)	市场汽车需求量 (万辆)
2010	0.5093	1838.9	0.5015	1800.7	1801.20
2011	0.8159	1841.1	0.8368	1849.1	1849.94
2012	1.2552	1925.9	1.2791	1929.3	1930.58
2013	1.7533	2209.9	1.7642	2196.6	2198.36

Continued

2014	6.25	2366	6.25	2342.94	2349.19
2015	34.3	2449.5	31.25	2470	2501.25
2016	53.1	2758.58	50.0	2752.8	2802.80
2017	78.1	2823.4	75.0	2812.89	2887.89
2018	126.2	2654.7	121.5	2686.6	2808.10
2019	121.2	2447.9	120.2	2456.3	2576.50
2020	136.7	2385.8	136.3	2394.8	2531.10
2021	353.8	2254.4	348.7	2278.8	2627.50
2022	707.7	1994.4	697.4	1988.8	2686.20

分析表 2 中关于新能源汽车产量、传统汽车产量和市场汽车的需求量的数据，通过作出新能源汽车产量、传统汽车产量、市场汽车需求量这三个数据关于年份的散点图。在改进拟合前，本文使用万辆作为数据的单位，并全部做二阶多项式拟合得到的三个散点图和拟合曲线如图 4~图 6 所示：

上面 3 个二阶散点图中拟合曲线的拟合优度 R^2 分别为 0.8498、0.9156、0.8706 均大于 0.84，拟合优度大，自变量对因变量的解释程度高，自变量引起的变动占总变动的百分比高。在改进拟合前，直接使用 Stata 16 SE 将这三个变量作为自变量，以新能源汽车的保有量为因变量，考虑到存在的异方差或自相关的情形，使用“OLS + 稳健标准误”的基本思想进行三元线性回归得出的结果如表 3 所示。

得到的结果拟合优度 R^2 为 0.9477，观测值拟合的效果好；F 值为 236.96 很大，说明十分符合线性关系；P 值为 $0.0000 < 0.0001$ ，故置信区间达到 99.99% 以上，说明本次拟合结果好。以新能源汽车的生产量 x_1 、传统汽车的生产量 x_2 、传统汽车的需求量 x_3 为自变量，新能源汽车的增长量 $d_1(t)$ 为因变量，得

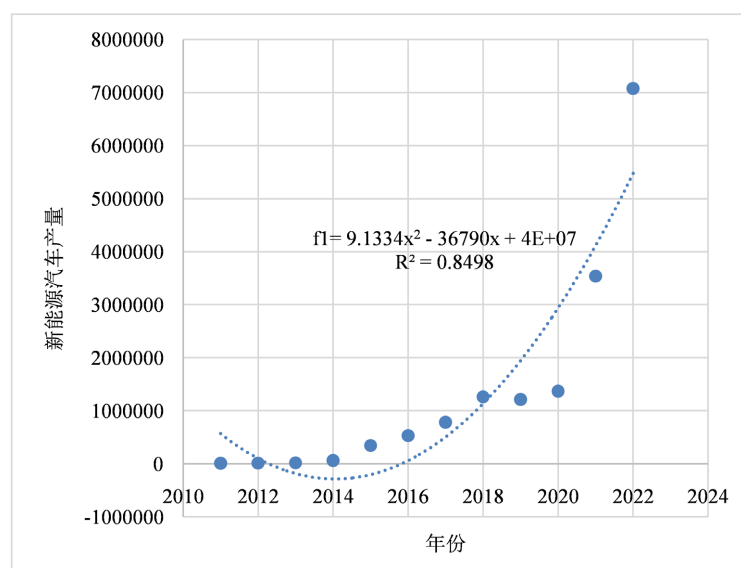


Figure 4. The fitting graph of the relationship between production and year of new energy vehicles

图 4. 新能源汽车产量与年份关系拟合图

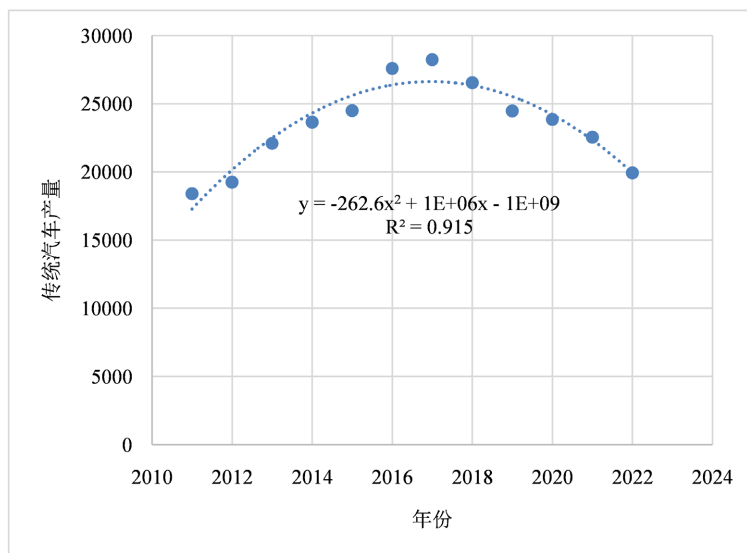


Figure 5. The fitting graph of the relationship between the production of new energy car and time

图 5. 传统汽车产量与年份关系拟合图

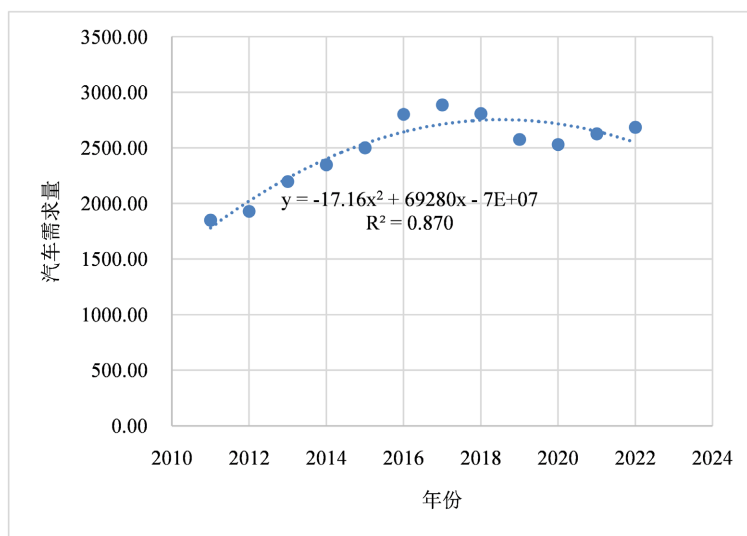


Figure 6. The fitting graph of car demand and time

图 6. 汽车需求量与年份关系拟合图

Table 3. New energy vehicle sales growth rate of three linear regression results

表 3. 新能源汽车销量增长率三元线性回归得到的结果

因变量	拟合优度 R^2	F 值	P 值	均方根误差 RMSE
新能源销量增长率	0.9477	236.96	0.0000	29.187

到表达式为:

$$d_1(t) = a_1x_1 + a_2x_2 + a_3x_3 + con$$

使用 SPSS 进行多元线性回归, 得到各变量的系数如表 4 所示:

Table 4. New energy vehicle sales growth rate of ternary linear regression coefficient**表 4.** 新能源汽车销量增长率三元线性回归得到的系数

自变量	新能源产量	传统产量	需求量	常数
系数	0.5174	-0.0101	-0.0111	39.2882

首先将得到的系数代入拟合关系式 $d_1(t) = a_1x_1 + a_2x_2 + a_3x_3 + con$ ，得到具体表达式如下所示：

$$d_1(t) = 0.5174f_1(t) - 0.1001g(t) - 0.0111f_2(t) + 39.2882$$

再将图 6~图 8 中的三个拟合表达式代入到回归得出表达式中，得到如下新能源汽车与新能源汽车的生产量 x_1 、传统汽车的生产量 x_2 、传统汽车的需求量 x_3 的关系式：

$$d_1(t) = 2.8955x^2 - 30394.11x - 1.307 \times 10^7$$

同上述模型求解过程，三元线性回归得到的结果如表 5 所示：

Table 5. The traditional car sales growth rate is the result of three linear regression**表 5.** 传统汽车销量增长率三元线性回归得到的结果

因变量	拟合优度 R^2	F 值	P 值	均方根误差 RMSE
传统销量增长率	0.6659	23.78	0.0002	124.1

得到结果的拟合优度 R^2 为 0.6659，观测值拟合效果较好；F 值为 23.78 较大，表明比较符合线性关系；P 值为 0.0002，故置信区间达到 99.98% 以上，表明拟合结果好。以新能源汽车的生产量 x_1 、传统汽车的生产量 x_2 、传统汽车的需求量 x_3 为自变量，传统汽车的增长量 $d_2(t)$ 为因变量，得到表达式为：

$$d_2(t) = a'_1x_1 + a'_2x_2 + a'_3x_3 + con'$$

各变量的系数如表 6 所示：

Table 6. The coefficient of traditional car sales growth rate by ternary linear regression**表 6.** 传统销量汽车增长率三元线性回归得到的系数

自变量	新能源产量	传统产量	需求量	常数
系数	3.7457	4.4162	-4.4018	76.5589

首先将得到的系数代入拟合关系式 $d_2(t) = a'_1x_1 + a'_2x_2 + a'_3x_3 + con'$ ，得到：

$$d_2(t) = 3.7456f_1(t) + 4.4162g(t) - 4.4018f_2(t) + 76.5589$$

再将图 6~图 8 中的三个拟合表达式代入到回归得出表达式中，得到如下新能源汽车与新能源汽车的生产量 x_1 、传统汽车的生产量 x_2 、传统汽车的需求量 x_3 的关系式：

$$d_2(t) = 15.809x^2 - 2.12 \times 10^5x - 6.11 \times 10^7$$

将 $d_1(t)$ 和 $d_2(t)$ 的图像在 Matlab R2021a 中在同一个图窗画出如图 7 所示。

由上图可知，拟合改进前得出的新能源汽车和传统汽车的保有量明显可看出其拟合出来的曲线是不符合实际情况的，拟合存在问题，故在后续需要对对操作步骤进行改进。

在改进得到拟合曲线的过程中，操作步骤与拟合改进前的操作步骤一致，只是进行如下改进。首先

在改进拟合后，不是使用万辆作为数据的单位，而是选用辆作为数据的单位。其次，使用指数拟合做出新能源汽车产量的拟合图；使用二阶多项式做出传统汽车产量、市场汽车需求量关于年份的拟合图，最终得到的折线图和拟合曲线中，其他的拟合图以及拟合曲线都没有变化，仅新能源汽车产量与年份的拟合图有所变化，如图 8 所示。

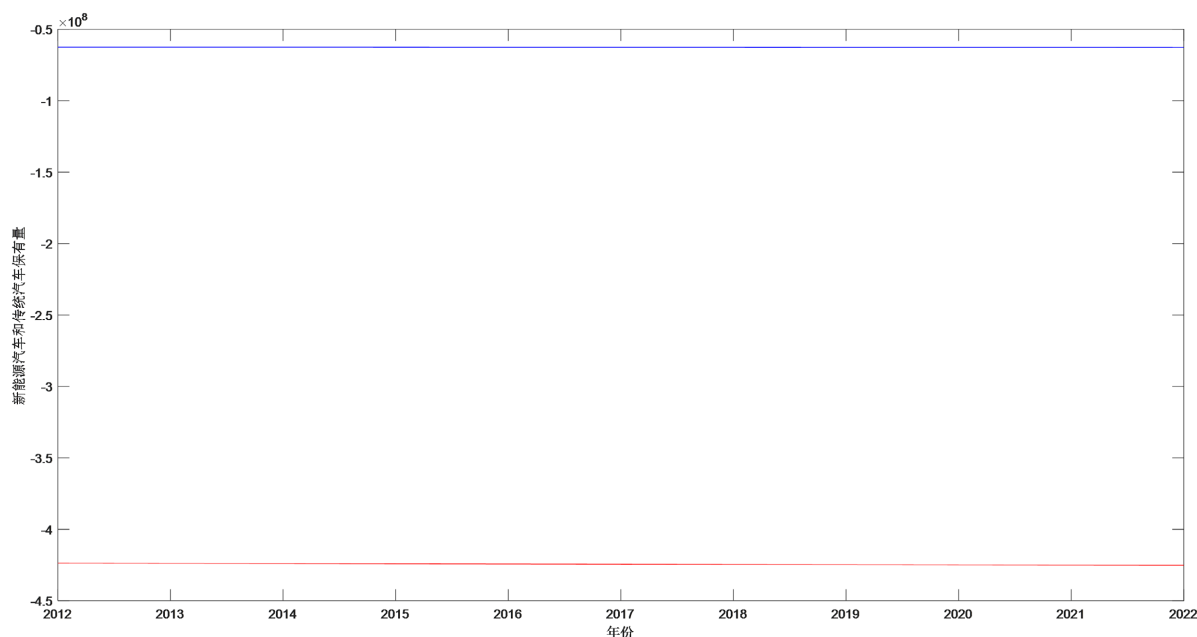


Figure 7. The previous $d_1(t)$ and $d_2(t)$'s fitting curve images were improved

图 7. 改进前 $d_1(t)$ 和 $d_2(t)$ 的拟合曲线图像

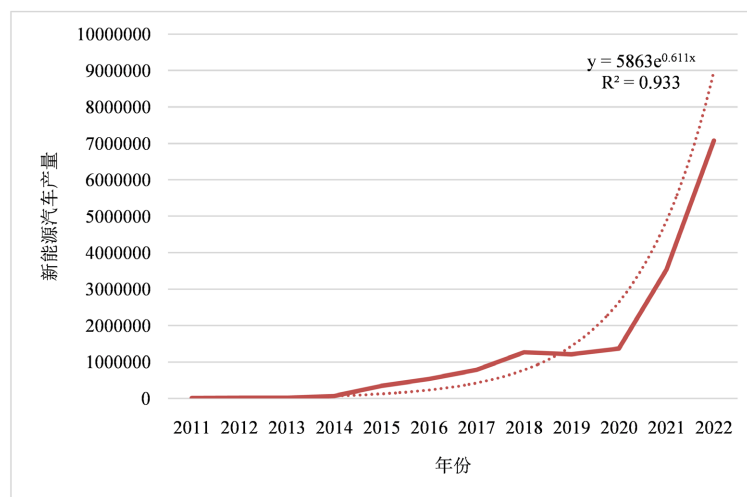


Figure 8. The fitting graph of the relationship between production of new energy vehicles and time

图 8. 新能源汽车产量与年份关系拟合图

在改进当使用 Stata 16 SE 进行回归分析时，不直接使用三元线性回归模型，而是改成使用三次一元线性模型得到新能源汽车销量的拟合结果如表 7 所示：

Table 7. New energy car sales growth after the improvement of fitting analysis results
表 7. 新能源汽车销量增长拟合改进后得到的分析结果

自变量	拟合优度 R^2	F 值	P 值
新能源产量	0.9441	168.85	0.0000
传统产量	0.0740	0.80	0.3922
需求量	0.1126	2.77	0.1268

以此得到的新能源产量、传统产量、需求量这三个变量的拟合优度 R^2 分别为 0.9441、0.0740、0.1126，可知新能源产量与新能源销量增长的相关程度最大；F 值分别为 168.85、0.80、2.77 主要变量的 F 值很大，说明十分符合线性关系；P 值分别为 0.0000、0.3922、0.1268，主要变量的 P 值为 $0.0000 < 0.0001$ ，故置信区间达到 99.99% 以上，说明本次拟合结果好。得到的各变量系数如表 8 所示：

Table 8. The improved new energy car sales growth rate of the ternary linear regression coefficient
表 8. 改进后新能源汽车销量增长率三元线性回归得到的系数

自变量	新能源产量	传统产量	需求量	常数
系数	0.5135	-0.0937	0.1083	55.8746

同理，在使用线性回归求解传统汽车销量的拟合结果时，也把三个变量求解，得到的结果如表 9 所示：

Table 9. The analysis results are obtained after the improvement of traditional car sales growth fitting
表 9. 传统汽车销量增长拟合改进后得到的分析结果

自变量	拟合优度 R^2	F 值	P 值
新能源产量	0.5175	34.62	0.0002
传统产量	0.0328	0.42	0.5326
需求量	0.0817	1.38	0.2680

得到的新能源产量、传统产量、需求量的拟合优度 R^2 分别为 0.5175、0.0328、0.0817，可知新能源产量与传统汽车销量增长的相关程度最大；F 值分别为 34.62、0.42、1.38 主要变量的 F 值较大，说明比较符合线性关系；P 值分别为 0.0002、0.5326、0.2680，主要变量的 P 值为 $0.0002 < 0.0003$ ，故置信区间达到 99.97% 以上，说明本次拟合结果好。得到的各变量系数如表 10 所示：

Table 10. The improved new energy car sales growth rate of the ternary linear regression coefficient
表 10. 改进后新能源汽车销量增长率三元线性回归得到的系数

自变量	新能源产量	传统产量	需求量	常数
系数	-0.6396	0.1049	-0.1552	2723808

最后，将得到的两组数据分别结合三个拟合曲线，得到如下两个新的表达式：

$$d_1'(t) = -3749e^{0.6112x} + 4282x^2 - 150900x + 22723800$$

$$d_2'(t) = 3010e^{0.6112x} + 38550x^2 - 202000x + 10^7$$

因较大数据计算后会超过 Matlab 所允许的数据范围，故在使用 Matlab R2021a 画新能源汽车产业和传统汽车产业的增长率对比图时需要将横坐标年份的数据减去 2010 以修正，使得区间变为[0, 16]，在软件中画出改进后的新能源汽车和传统汽车增长率对比图，如图 9 所示：

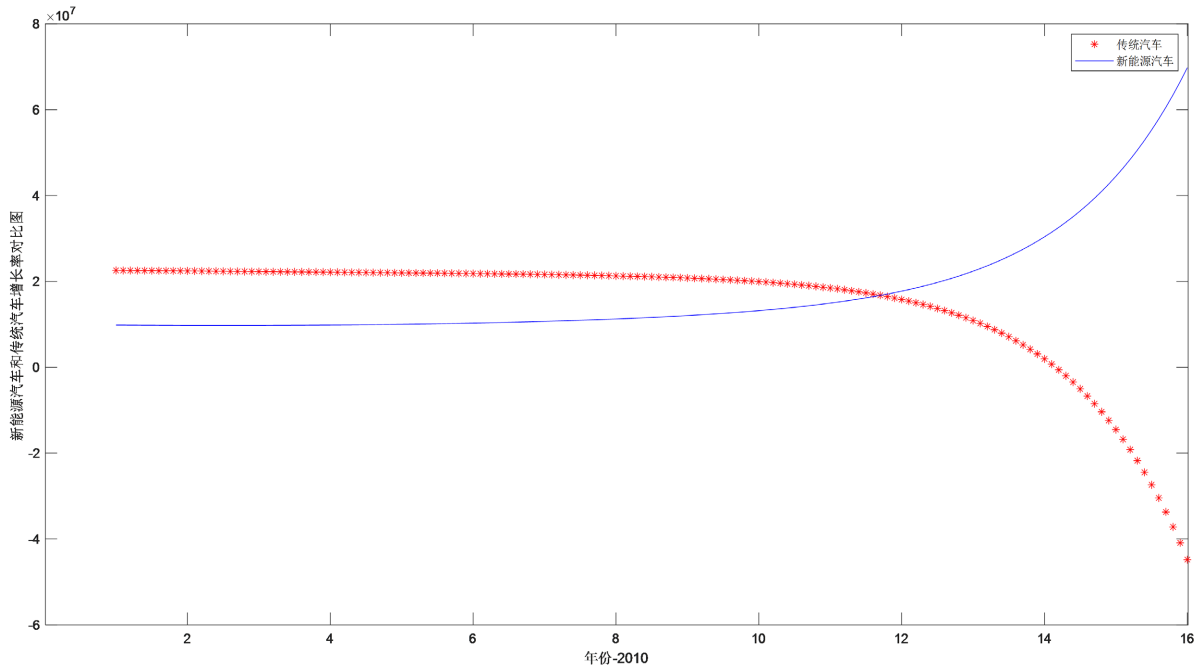


Figure 9. The improved $d_1(t)$ and $d_2(t)$ fitting curve images
图 9. 改进后 $d_1(t)$ 和 $d_2(t)$ 的拟合曲线图像

综合上文分析不难发现，新能源汽车保有量的相关影响因素众多，指标之间的重复性会导致输入样本数据信息的大量重叠。中国新能源汽车市场形成了传统车企与新势力车企共同投资建设及相互竞争的格局。新能源汽车产业保有量逐年升高，且每年增长量都有所增高，保持较高的增长速度；虽然传统汽车产业保有量也逐年升高，但每年增长量增高不大，保持较低甚至负的增长速度，故新能源汽车和传统汽车之间存在竞争关系。

3. 基于 GM(1,1)模型长三角地区新能源汽车保有量预测

3.1. GM(1,1)灰色预测模型的建立

灰色预测是一种对含有不确定因素的系统进行预测的方法，有助于解决样本数量少、数据信息少和信息不完整等不确定的问题[5]。灰色预测 GM(1,1)模型首先是利用随机原始离散非负的数据列，通过一次累加的方式得到新的离散数据列，建立相应的微分方程模型，用一阶线性微分方程的解来逼近时间累加后形成的新时间序列所呈现的规律。它通过一个序列即可进行计算，是一阶线性的动态模型[6]。建立 GM(1,1)灰色预测模型的具体过程如下：

首先，假设 $X^{(0)}$ 为非负序列，并根据已有的数据构建最初的数据序列：

$$X^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n)\}$$

对于上述表达式, 有 $X^{(0)}(i) \geq 0, i = 1, 2, \dots, n$ 。

再将建立的数列进行一次累加, 生成如下所示的累加数据序列:

$$X^{(1)} = \{x^{(1)}(1), x^{(1)}(2), \dots, x^{(1)}(n)\}$$

其中, $X^{(1)}(k) = \sum_{k=1}^n X^{(0)}(k), k = 1, 2, \dots, n$ 。

其次, 对累加生成的序列进行处理, 得出的紧邻均值生成序列为:

$$Z^{(1)} = (Z^{(1)}(2), Z^{(1)}(3), \dots, Z^{(1)}(n))$$

式中, $Z^{(1)}(k) = \frac{1}{2}x^{(1)}(k) + \frac{1}{2}x^{(1)}(k-1), k = 2, 3, \dots, n$ 。

定义灰色微分方程(近似微分方程)的表达式如下:

$$x^{(0)}(k) + ax^{(1)}k = b$$

构建白化方程, 表达式如下:

$$\frac{dx}{dt} + ax^{(1)} = u$$

使用最小二乘法的思想, 对参数 a 和参数 b 进行估计, 得到:

$$\hat{a} = [a, b] = (B^T B)^{-1} B^T Y_n$$

得到最终的 GM(1,1)模型的时间响应表达式如下:

$$\hat{x}^{(1)} = \left(x^{(1)}(0) - \frac{b}{a} \right) e^{-a(k-1)} + \frac{b}{a}$$

其中, $x^{(1)}(0) = x^{(0)}(1), k = 1, 2, \dots, n$ 。

还原值为:

$$\hat{x}^{(0)}(k) = \hat{x}^{(1)}(k) - \hat{x}^{(1)}(k-1)$$

该模型的预测结果是否可取需要进行下述检验, 计算公式如下:

绝对残差: $\varepsilon^{(0)}(k) = x^{(0)}(k) - \hat{x}^{(0)}(k), k = 2, 3, \dots, n$

相对残差: $w^{(0)}(k) = \left| \frac{x^{(0)}(k) - \hat{x}^{(0)}(k)}{x^{(0)}(k)} \right| * 100\%, k = 2, 3, \dots, n$

平均相对残差: $\bar{w} = \frac{1}{n} \sum_{k=1}^n |w_r^{(0)}(k)|$

精度: $p^{(0)} = (1 - \bar{w}) * 100\%$

原始序列均方差: $\bar{X}^{(0)} = \frac{1}{n} \sum_{k=0}^n X^{(0)}(i)$

$$S_0^2 = \sum_{k=1}^n (X^{(0)} - \bar{X}^{(0)})^2$$

$$S_0 = \sqrt{\frac{S_0^2}{n-1}}$$

残差序列均方差: $\bar{\varepsilon}^{(0)} = \frac{1}{n} \sum_{k=0}^n \varepsilon^{(0)}(i)$

$$S_1^2 = \sum_{k=1}^n (\varepsilon^{(0)} - \bar{\varepsilon}^{(0)})^2$$

$$S_1 = \sqrt{\frac{S_1^2}{n-1}}$$

方差比: $C = \frac{S_1}{S_0}$

3.2. 利用 GM(1,1)模型进行预测

仅有 8 个数据样本, 且预测未来三年长三角地区新能源汽车保有量为短期预测, 而灰色预测模型正好适合样本少, 短期预测问题的运用。公安部发布的长三角地区历年新能源汽车保有量的数据如表 11 所示:

Table 11. New energy car ownership scale of Yangtze River Delta from 2015 to 2022

表 11. 2015 年~2022 年长三角地区新能源汽车保有量表

时间	2015	2016	2017	2018	2019	2020	2021	2022
长三角新能源车保有量(万辆)	10.617	26.264	46.4	70.778	98.4	136.56	252.84	372.8

根据 3.1 中建立的 GM(1,1)预测模型, 利用最小二乘法, 使用 PyCharm Community Edition 软件求解得到:

$$\hat{a} = [a, b] = (B^T B)^{-1} B^T Y_n = (-0.4257, 52.2575)$$

得到的长三角地区新能源汽车保有量数据序列 X 的预测模型为:

$$\hat{x}^{(0)}(k) = \hat{x}^{(1)}(k) - \hat{x}^{(1)}(k-1) = 146.756 * e^{(k-1)} + 122.757, k = 2, 3, \dots, n$$

最终得到长三角地区新能源汽车保有量的模拟数据, 以及模拟数据和实际保有量的比较值, 如表 12 所示:

Table 12. GM(1,1) prediction model compares the simulated data and the actual data of the Yangtze River Delta new energy car ownership

表 12. GM(1,1)模型对长三角地区新能源汽车保有量模拟数据与实际数据对比

年份	实际数据(万辆)	模拟数据(万辆)	残差	相对残差
2015	10.617	10.617	0.00	0.00%
2016	26.264	22.219	4.04	15.4%
2017	46.4	34.414	11.99	25.8%
2018	70.778	53.301	17.48	24.6%
2019	98.4	82.554	15.85	16.1%
2020	136.56	127.862	8.69	6.3%
2021	252.84	198.036	54.8	21.67%
2022	372.8	306.723	66.08	17.7%

下面对预测的结果进行精度检验, 得到的计算结果如下所示:

$$\text{平均相对残差: } \bar{w} = \frac{1}{n} \sum_{i=1}^n |w^{(0)}(i)| = 15.9\%$$

$$\text{精度: } p^{(0)} = (1 - \bar{w}) * 100\% = (1 - 0.159) * 100\% = 84.03\%$$

$$\text{方差比: } C = \frac{S_1}{S_2} = 0.0376$$

该模型的精度大于 84%，因此预测的结果比较合理。通过实际统计数据和模拟数据的对比处理可以看出，新能源汽车保有量的模拟数据和实际数据最大偏差不超过 26%，精准度可达 84%以上，方差比预测精度为一级，充分说明 GM(1,1)预测模型预测精度较好，能够应用于新能源汽车保有量预测。

将 $k = 9, 10, 11$ 代入预测模型 $\hat{x}^{(0)}(k) = 146.756 * e^{(k-1)} + 122.757$, $k = 9, 10, 11$ ，得到全国范围内未来 3 年的新能源汽车保有量，如表 13 所示：

Table 13. Predicted number of Yangtze River Delta new-energy vehicle ownership in 2023~2025

表 13. 2023~2025 年长三角地区新能源汽车保有量预测值

年份	2023	2024	2025
保有量(万辆)	475.06	735.79	1139.61

画出的 2015 年到 2025 年的长三角地区新能源汽车保有量曲线图如图 10 所示：

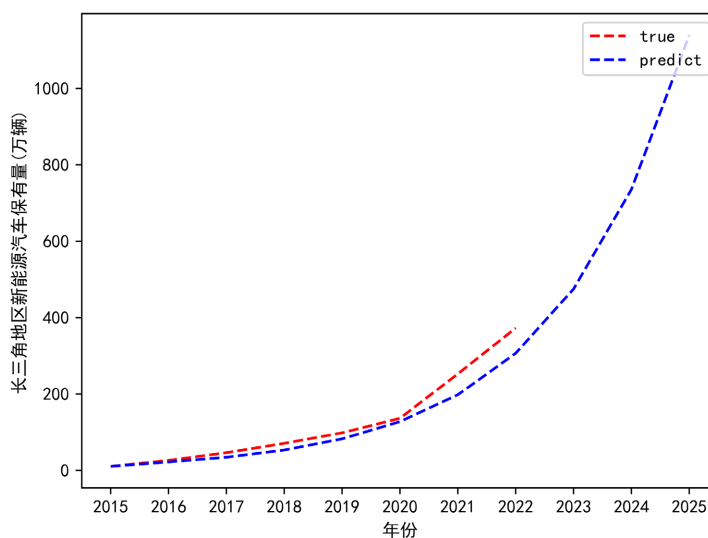


Figure 10. A forecast figure of new-energy car ownership in the Yangtze River Delta from 2015 to 2025

图 10. 2015 年到 2025 年的长三角地区新能源汽车保有量预测图

从上图可以看出，长三角地区新能源汽车的发展从初期的缓慢增长逐步迈入快速发展时期，且增速逐渐加快。根据现有的预测结果充分说明长三角地区新能源汽车产业有巨大的发展空间[7]。

4. 基于赋权改进的拟合曲线实现双碳时间预测

碳达峰是指我国承诺在 2030 年以前，不再增长二氧化碳的排放量，达到排放峰值之后再逐步降低。碳中和是指企业、团体或个人测算在一定时间内直接或间接产生的温室气体排放总量，然后通过植物造

树造林、节能减排等形式，抵消自身产生的二氧化碳排放量，实现二氧化碳“零排放”。在查询相关文献后，分析可知碳排放量与各能源消耗量以及该地区森林覆盖面积均有关，收集到的各能源消耗量、森林覆盖面积如表 14 所示：

Table 14. Table of data on energy consumption and forest cover

表 14. 各能源消耗量和森林覆盖面积数据表

年份	煤炭(万吨)	原油(万吨)	天然气(亿立方米)	电能(亿千瓦时)
2004	34601.29	5990.91	14.30	4576.99
2005	40123.53	6759.29	35.44	5339.89
2006	43735.72	6695.67	70.39	6131.31
2007	48019.19	6873.38	94.48	6982.87
2008	50618.66	6978.37	118.00	7438.29
2009	52250.75	7558.55	126.02	7891.11
2010	56301.56	8438.03	160.87	9059.08
2011	62820.00	8539.66	213.19	9959.34
2012	62543.00	8312.51	250.50	10506.00
2013	63453.60	9411.86	281.89	11348.40
2014	61419.74	9234.50	312.75	11473.14
2015	61434.64	9886.74	357.61	11713.94
2016	62350.92	9772.83	378.73	12613.14
2017	61544.53	10149.07	470.22	13448.77
2018	60681.04	9875.21	557.62	14362.82
2019	59516.97	10849.65	594.30	14839.84

在计算碳排放量时，必须转化成标准统计量，参照《中国能源统计年鉴》给出的具体换算系数[7]，如表 15 所示：

Table 15. Table of conversion coefficient between each energy source and standard coal

表 15. 各能源与标准煤换算系数表

能源种类	煤炭(t 标煤/t)	原油(t 标煤/t)	天然气(t 标煤/万 m ³)	电能(t 标煤/万 kWh)
换算系数	714.3	1428.6	0.4479	2.2132

各种能源的碳排放系数，如表 16 所示：

Table 16. Table of carbon emission coefficient of each energy source

表 16. 各能源碳排放系数表

能源种类	煤炭(t 碳/t 标煤)	原油(t 碳/t 标煤)	天然气(t 碳/t 标煤)	电能(t 碳/t 标煤)
碳排放系数	0.7476	0.5854	0.4479	2.2132

设煤炭、石油、天然气、电能这四种能源的消耗量分别用 CS_1 、 CS_2 、 CS_3 、 CS_4 表示，其对应的系数(换算系数*碳排放系数)为 λ_1 、 λ_2 、 λ_3 、 λ_4 ，则碳排放总量的计算公式为：

$$W_{OUT,t} = \sum_{i=1}^5 \lambda_{i,t} CS_{i,t}$$

利用能源消耗量、能源与标准煤换算系数、碳排放系数得到表 17 所示的长三角地区碳排放量数据：

Table 17. Data table of carbon emissions and total carbon emissions from energy sources in Yangtze River Delta
表 17. 长三角地区各能源碳排放量及总体碳排放量数据表

年份	煤炭(万吨)	原油(万吨)	天然气(万吨)	电能(万吨)
2004	18477.4584	5010.212651	8.5186101	12449.51716
2005	21426.3935	5652.81072	21.11185608	14524.62255
2006	23355.3416	5599.605158	41.93181573	16677.30299
2007	25642.7603	5748.224465	56.28239736	18993.56561
2008	27030.905	5836.027858	70.293426	20232.31839
2009	27902.4585	6321.233808	75.07099614	21463.99912
2010	30065.6343	7056.745078	95.83138509	24640.90415
2011	33546.5509	7141.738495	126.9987753	27089.63187
2012	33398.63	6951.772396	149.2246035	28576.55954
2013	33884.9001	7871.161483	167.9238462	30867.90674
2014	32798.7971	7722.834882	186.3073643	31207.20239
2015	32806.7539	8268.304786	213.0307803	31862.18388
2016	33296.0572	8173.041575	225.6121121	34308.02838
2017	32865.4363	8487.692005	280.1133455	36580.96103
2018	32404.3234	8258.662219	332.1781373	39067.19787
2019	31782.6976	9073.588768	354.0286701	40364.70315

分别作出原油、天然气、电能等关于年份的拟合图先如下图 12~图 14 所示：

从图 11 可知，原油碳排放量表达式与时间拟合曲线为：

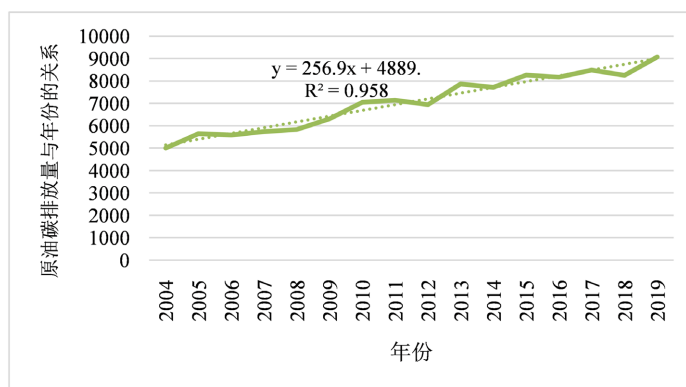


Figure 11. The fitting diagram of crude oil carbon emission and time in Yangtze River Delta

图 11. 长三角地区原油碳排放量与时间拟合图

$$y_1 = 256.92x + 4889.5$$

从图 12 可知，天然气碳排放量表达式与时间拟合曲线：

$$y_2 = 0.9528x^2 + 6.1037x + 9.315$$

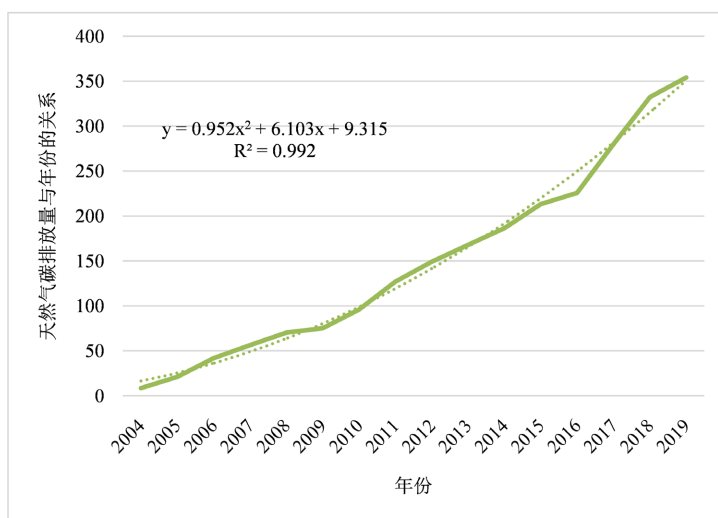


Figure 12. The fitting diagram of natural gas carbon emission and time in Yangtze River Delta

图 12. 长三角地区天然气碳排放量与时间拟合图

从图 13 可知，电力碳排放量与时间拟合曲线：

$$y_3 = 1830x + 11246$$

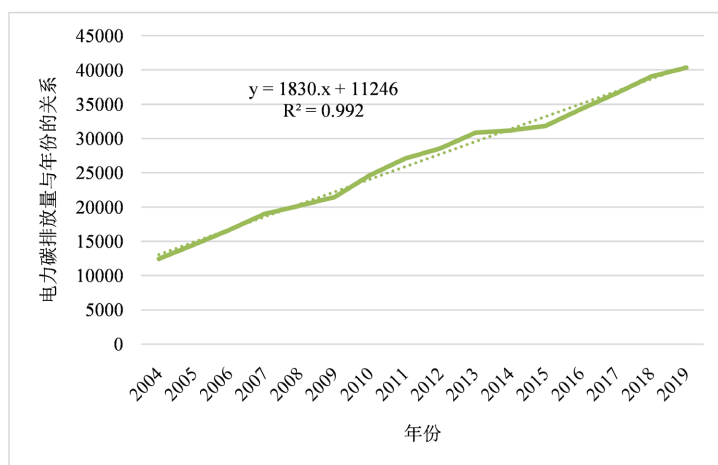


Figure 13. Power carbon emission and time fitting graph in Yangtze River Delta

图 13. 长三角地区电力碳排放量与时间拟合图

由于煤炭的折线图中存在着一个向上较为突出的孤立点，所以在此处引入拟合修正因子，分别作出线性拟合曲线和二阶多项式拟合曲线，结果如图 14 和图 15 所示：

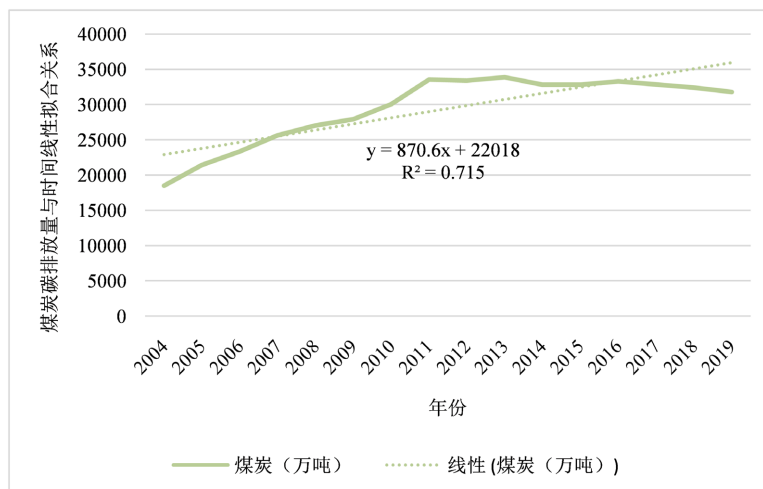


Figure 14. Linear fitting diagram of coal carbon emission and time in Yangtze River Delta

图 14. 长三角地区煤炭碳排放量与时间线性拟合图

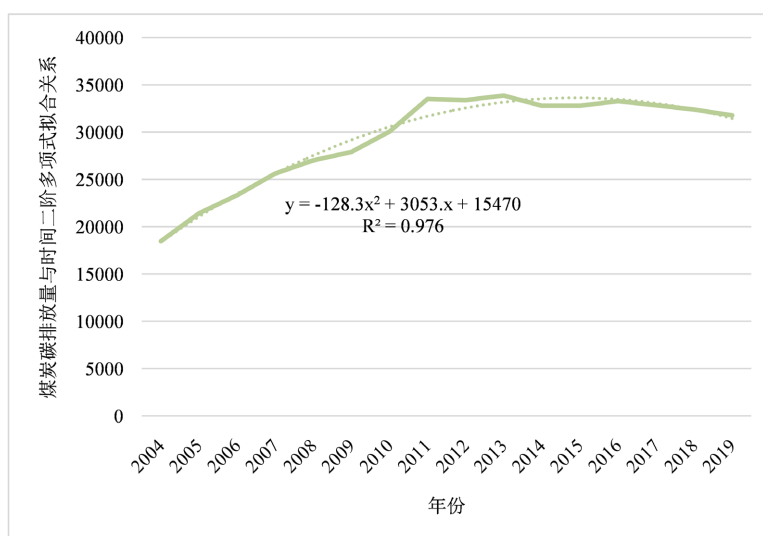


Figure 15. Second order polynomial fitting diagram of coal carbon emission and time in Yangtze River Delta

图 15. 长三角地区煤炭碳排放量与时间二阶多项式拟合图

经过分析可知，线性拟合曲线拟合优度为 0.7155，提供向上的修正量；二阶多项式拟合曲线拟合优度为 0.9769，提供向下的修正量。由于折线图上存在一个向上突出的偏差点，所以需要向上修正，才能得到最精确的拟合曲线。提供分析，已经综合考虑其他影响因素，最终给线性拟合曲线赋权值 0.3，给二阶多项式拟合曲线赋权值 0.7，得到的最终拟合曲线为：

$$\begin{aligned}
 y_4 &= 0.3*(870.6x + 22018) + 0.7*(-128.39x^2 + 3053.2x + 15470) \\
 &= -89.873x^2 + 2398.42x + 17434.4
 \end{aligned}$$

将 y_1 、 y_2 、 y_3 、 y_4 四个表达式相加，使用 Matlab R2021a 画出碳排放量关于时间的拟合曲线如图 16 所示：

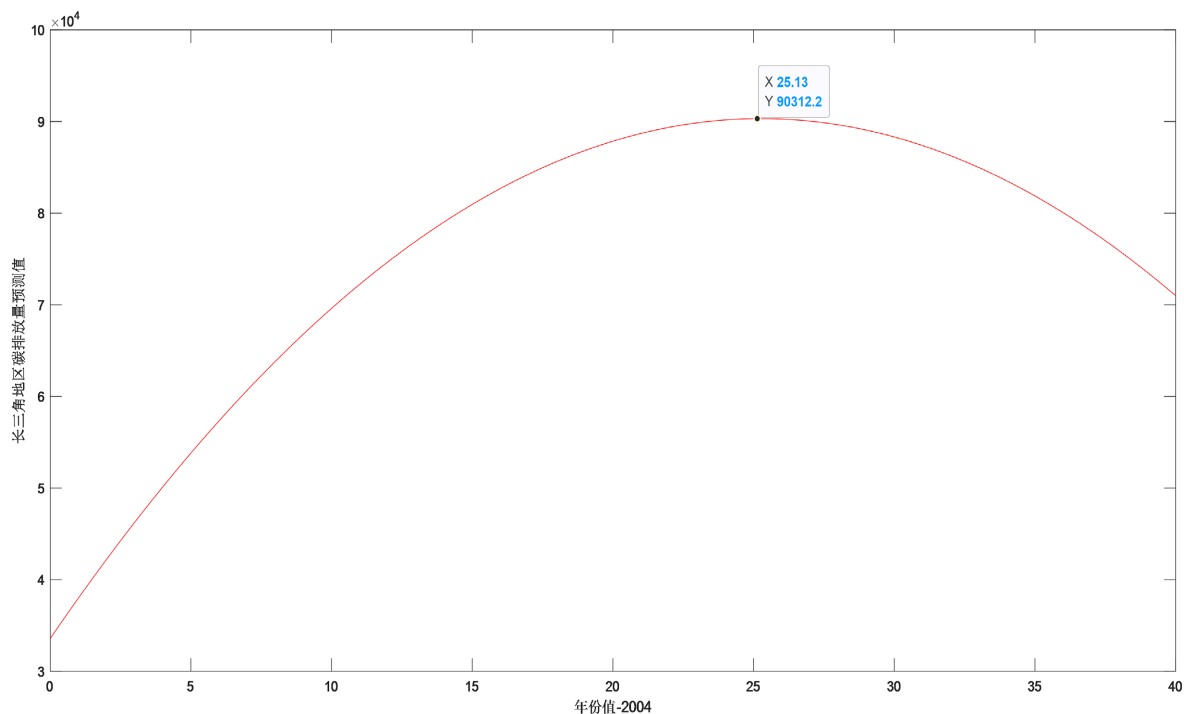


Figure 16. Prediction of fitting curve of carbon emission in Yangtze River Delta
图 16. 长三角地区碳排放量拟合曲线预测图

从国家统计局的《中国统计年鉴》中收集关于长三角地区三省一市和总体的碳吸收量，数据如表 18 所示：

Table 18. Table of carbon absorption coefficient in Yangtze River Delta region
表 18. 长三角地区碳吸收量系数表

年份	上海(万吨)	江苏(万吨)	浙江(万吨)	安徽(万吨)	总计(万吨)
2016	4.15391517	103.31936	380.144985	235.274133	722.8923932
2017	4.15391517	103.31936	380.144985	235.274133	722.8923932
2018	4.15391517	103.31936	380.144985	235.274133	722.8923932
2019	5.43025782	99.39584	382.461765	244.845765	732.1336278
2020	5.43025782	99.39584	382.461765	244.845765	732.1336278
2021	5.43025782	99.39584	382.461765	244.845765	732.1336278
2022	5.43025782	99.39584	382.461765	244.845765	732.1336278

作出长三角地区碳吸收总量关于年份的散点图，如图 17 所示。

因长三角地区碳吸收量关于年份变化不大，加上考虑到实际情况下森林等碳吸收量变化不大，为了简化模型，取均值 727.5 万吨为每年的碳吸收量。得到的长三角地区碳净排放量如图 18 所示。

所以由图 16 和图 18 可知，长三角地区预测在 2029 年实现碳达峰，在 2061 年实现碳中和。

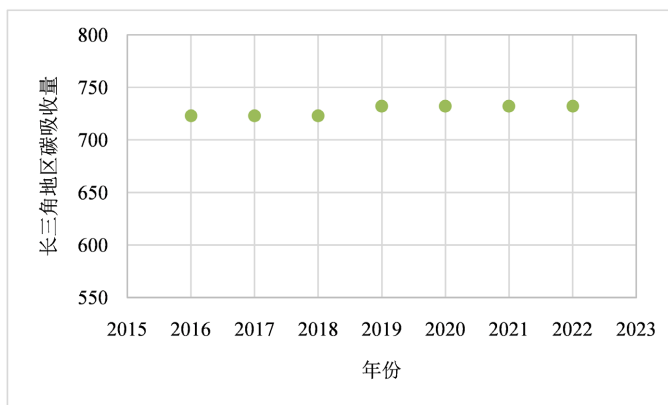


Figure 17. Scatter plot of total carbon uptake in Yangtze River Delta region with respect to year

图 17. 长三角地区碳吸收总量关于年份的散点图

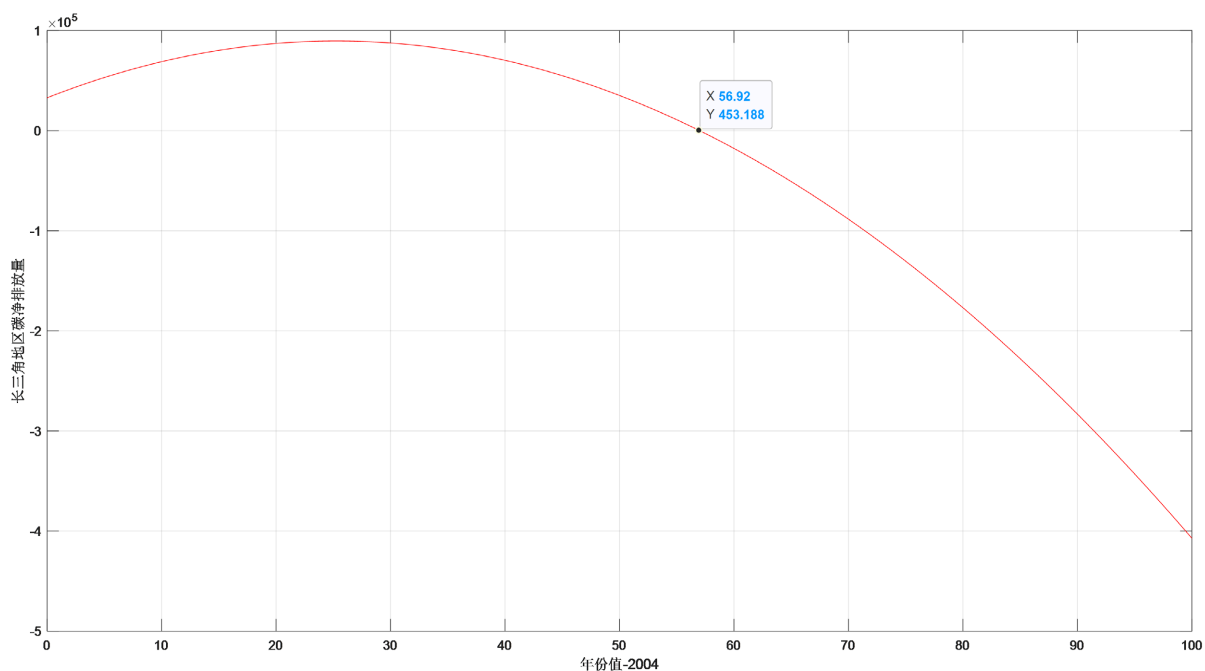


Figure 18. Fitting curve of net carbon emission in Yangtze River Delta

图 18. 长三角地区碳净排放量拟合曲线图

5. 结语

使用描述统计分析模型,利用 2011~2020 年长三角地区及全国新能源汽车产量数据,画出逐年对比图,可知近 5 年来长三角区域新能源汽车产量占全国范围内新能源汽车产量的比重总体呈现逐年增长的趋势,说明长三角地区新能源汽车产业在全国的影响和作用越来越大;每一年长三角区域新能源汽车比全国新能源汽车的增长率都要高,说明长三角区域新能源汽车产业总体发展速度比全国新能源汽车产业的发展速度要快。

通过理想状态的假设,逐渐增加影响因素进行推导,分析 2010~2022 年长三角地区新能源汽车与传统汽车的生产量和销售量,以及市场的汽车需求量,使用二阶线性回归得到新能源汽车产量、传统汽车

产量和汽车需求量关于年份的拟合表达式。首先直接使用 SPSS 进行三元线性回归, 得到新能源汽车销量与新能源汽车产量、传统汽车产量、汽车需求量这三个自变量的关系, 再代入其关于年份的拟合式, 得到初版的竞争模型, 使用三次一元拟合回归分别得到新能源汽车销量与新能源汽车产量、传统汽车产量、汽车需求量的三个拟合关系式, 并将三个自变量与年份的拟合式代入, 从优化后新能源汽车和传统汽车的拟合曲线, 得到中国新能源汽车市场形成了传统车企与新势力车企共同投资建设及相互竞争的关系。

建立 GM(1,1)灰色预测模型, 根据 2015~2022 年长三角地区新能源汽车保有量的数据, 预测 2023~2025 年长三角地区新能源汽车保有量。通过对煤炭、原油、天然气、电能这四种能源对年份的拟合进行赋权改进, 得到长三角地区将于 2029 年实现碳达峰, 2061 年实现碳中和。

参考文献

- [1] 沈占胜. “双碳”计划的价值导向与实施路径研究——以习近平新时代绿色发展理念为指导[J]. 上海节能, 2022(11): 1375-1379. <https://doi.org/10.13770/j.cnki.issn2095-705x.2022.11.002>
- [2] 李建豹, 黄贤金, 揣小伟, 孙树臣. “双碳”背景下长三角地区碳排放情景模拟研究[J]. 生态经济, 2022, 38(11): 13-18+35.
- [3] 刘定平. 描述统计与推断统计的比较分析[J]. 经济经纬, 1997(4): 83+92.
- [4] 蒋景明. 多元线性回归模型中异方差问题的研究[D]: [硕士学位论文]. 桂林: 桂林理工大学, 2022. <https://doi.org/10.27050/d.cnki.gglgc.2022.000006>
- [5] 程晓丽. 基于灰色模型的中国纯电动汽车销量预测[J]. 汽车实用技术, 2022, 47(9): 28-31. <https://doi.org/10.16638/j.cnki.1671-7988.2022.009.006>
- [6] 李功伟. GM(1,1)灰色预测模型在高速铁路项目施工成本动态管理中的应用[J]. 价值工程, 2019, 38(12): 30-32. <https://doi.org/10.14018/j.cnki.cn13-1085/n.2019.12.009>
- [7] 朱曼宇, 张峥. 基于 GM(1,1)模型的新能源汽车销量预测研究[J]. 物流工程与管理, 2023, 45(2): 138-141.