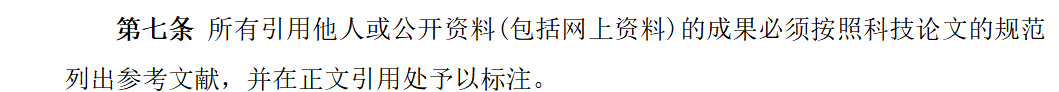
**使用须知：**该论文为直接可以提交的论文，会进行展示。很多人都会这一版本成品论文，直接提交一定会查重不过关。成品论文是按着半成品论文以及之前的解题思路写出的，如果不对外展示，在普通建模竞赛是可以直接提交100%获奖的文章，研赛二等奖的水平。进行展示的目的，及为了让辅助大家写论文，也让大家知道获奖的话应该要比展示的这篇成品论文要更好一些。

组委会要求重要格式规范如下所示，大家可以根据要求，以及1.1版本提供的优秀论文资料进行修改。以下是研赛论文要求，请认真阅读！！

* 每个参赛队可以从A、B、C、D、E、F题中任选一题完成论文。
* 论文题目、摘要和关键词写在论文摘要页上，摘要页的下一页开始论文正文。
* 论文从摘要页开始编写页码，页码必须位于每页页脚中部，用阿拉伯数字从“1 ”开始连续编号。
* 论文不能有页眉，论文中不能有任何可能显示答题人身份的标志。
* 论文题目用三号黑体字、一级标题用四号黑体字，并居中。论文中其他汉字一律采用小四号宋体字，行距用单倍行距。如题目中有要求，计算机结果和源程序需在规定时间内上传竞赛系统以备检查。
* 请大家注意：摘要应简明扼要，需包含：建模思路、主要方法、模型、结果与结论、创新点、关键词等，请认真书写（注意篇幅一般不超过两页，且无需译成英文）。评阅时对摘要和论文都会审阅。

使用、借鉴、抄写该论文是否违规：只要不超过查重率的20%，就不能算作违规。该论文属于公开发表的网上资料，进行借鉴、参考不能算作违规，切记不要抄的太过分，查重率一定要低于20%就可以。下图为最新的数模竞赛规定



**下面为成品论文 正文**

****    **** 

**中研研究生创新实践系列大赛**

**“华为杯”第二十届中国研究生**

**数学建模竞赛**

|  |  |
| --- | --- |
| **学 校** |  |
| **参赛队号** |  |
| **队员姓名** | 1. **B** |
| 1. **Z** |
| 1. **D** |

**中研研究生创新实践系列大赛**

**“华为杯”第二十届中国研究生**

**数学建模竞赛**

题 目： ­­­­\_\_\_\_区域双碳目标与路径规划研究

摘 要：

中国已经成为全球最大的能源生产国和消费国。中国在2020年的碳排放量占全球的31%。为了应对全球气候变暖和附带的一系列问题，中国提出了碳中和目标。本文将建立并运用数学模型，分析、评价和预测能效提升、产业（产品）升级、能源脱碳和能源消费电气化等重点工程对碳排放的影响。

对于题目给出的数据，首先进行数据预处理，包含降维、异常值、缺失值等操作。利用matlab的find函数寻找缺失值，对于无法解释的存在异常的缺失值使用插值进行填充处理。

对于问题一，对碳排放以及人口、经济、能源消费量的现状分析。首先，对于给出的指标构建指标评价体系。以经济、人口、能源消费量和碳排放量作为一级指标，根据给出的数据寻找其二级指标，完成对各部门的碳排放情况、描述指标之间的相互关系。利用处理后的数据绘制折线图，比较分析十二五、十三五的变化趋势，经济增长、人口增长、能源消费结构等方向出发，分析影响因素。引入person相关系数，对碳排放与经济、人口、能源消费量之间进行相关性分析，建立关联关系模型。

对于问题二，区域碳排放量以及经济、人口、能源消费量的预测模型。首先，对于基于人口和经济变化的能源消费量预测模型。利用数据处理之后的数据进行处理，建立回归预测模型。对于要求2、3，分别构建能源消费量与人口以及经济的线性回归模型进行预测，为了增加预测精度，引入LSTM、灰色预测模型再次进行预测。根据预测的结果，建立优化模型，以精度最小为目标函数，对三种预测方式进行加权，利用优化模型求解最优的权重。对于区域碳排放量预测模型，构建碳排放量与人口、GDP和能源消费量的多元线性回归模型进行预测，再利用第一小问的预测模型，提高预测精度。

对于问题三，以2030年达到碳达峰，2060年实现碳中和为确定时间节点，设计自然、基本、雄心三种不同的情景。分别为，没有人为干预、按时达到碳达峰碳中和、率先碳达峰与碳中和情况下的碳减排措施。对于各部门的能源消费量、能源消费品种及其碳排放量的预测方法，选取与问题二相同的预测方法。确定2025年、2030年、2035年、2050年和2060年的GDP、人口和能源消费量目标。将目标与提高能源利用效率和提高非化石能源消费比重的目标相关联，确保达到碳达峰和碳中和目标的路径与措施一致。对能效提升、产业（产品）升级、能源脱碳和能源消费电气化等措施进行定性和定量分析，以确定如何实现双碳目标。

最终，根据得出的结果规划各种情景下的路径，以实现双碳目标。这可能涉及政策制定、技术创新、产业结构调整等方面的措施。

关键词：数据预处理、预测模型、路径规划、双碳目标、相关关系

# 问题重述

## 问题背景

2020年9月22日，习近平主席在第七十五届联合国大会一般性辩论上提出中国将提高国家自主贡献力度，采取更加有力的政策和措施，二氧化碳排放力争于2030年前达到峰值，努力争取2060年前实现碳中和。

2021年9月22日，中共中央国务院正式发布《关于完整准确全面贯彻新发展理念做好碳达峰碳中和工作的意见》（以下简称《意见》），明确了中国双碳行动的顶层设计。

我国是世界上最大的发展中国家，为实现中华民族伟大复兴，规划了在2035年基本实现现代化、在2050年实现中国式现代化的经济社会发展目标。因此，实现2060年碳中和的目标，必须破解发展与碳减排之间的矛盾。其中，推动经济社会高质量发展是矛盾的主要方面。

经济增长与能源消费量以及能源消费量与碳排放量之间存在关联关系，破解发展与减排的矛盾，只能寻求经济增长与碳排放量的负相关变化，必须从提高能源利用效率和提高非化石能源消费比重两个方面入手。

提高能源利用效率（即降低单位GDP能耗），可实现经济增长与能源消费量增长的负相关变化。提高能源利用效率的主要途径：一是开展管理节能、技术节能和结构节能等能效工程，降低单位产品与服务的能耗；二是开展以科技创新为基础的产业升级工程，增加单位产品与服务的科技附加值。

提高非化石能源消费比重（即降低单位能耗碳排放），可实现能源消费量与碳排放量增长的负相关变化。提高非化石能源消费比重的主要途径有：一是开展新能源发电、火电脱碳与新型电网等能源脱碳工程，提升非化石能源发电占比；二是开展以电能替代化石能源为核心的能源消费电气化工程，提升电力消费比重。

本赛题需要建立并运用数学模型，分析、评价和预测能效提升、产业（产品）升级、能源脱碳和能源消费电气化等重点工程对碳排放的影响。

## 问题回顾

问题一：区域碳排放量以及经济、人口、能源消费量的现状分析

（1）建立指标与指标体系

（2）分析区域碳排放量以及经济、人口、能源消费量的现状

（3）区域碳排放量以及经济、人口、能源消费量各指标及其关联模型

问题二：区域碳排放量以及经济、人口、能源消费量的预测模型

（1）基于人口和经济变化的能源消费量预测模型

（2）区域碳排放量预测模型

问题三：区域双碳（碳达峰与碳中和）目标与路径规划方法

（1）情景设计（不少于三种情景，如无人为干预的自然情景、按时碳达峰与碳

中和的基准情景、率先碳达峰与碳中和的雄心情景等）。

（2）多情景下碳排放量核算方法

（3）确定双碳（碳达峰与碳中和）目标与路径

# 问题分析

## 2.1 数据分析

对于题目给出的数据，首先进行数据预处理，包含降维、异常值、缺失值等操作。利用matlab的find函数寻找缺失值，对于无法解释的存在异常的缺失值使用插值进行填充处理。

最终，根据得出的结果规划各种情景下的路径，以实现双碳目标。这可能涉及政策制定、技术创新、产业结构调整等方面的措施。

## 2.2 问题一分析

对于问题一，对碳排放以及人口、经济、能源消费量的现状分析。首先，对于给出的指标构建指标评价体系。以经济、人口、能源消费量和碳排放量作为一级指标，根据给出的数据寻找其二级指标，完成对各部门的碳排放情况、描述指标之间的相互关系。利用处理后的数据绘制折线图，比较分析十二五、十三五的变化趋势，经济增长、人口增长、能源消费结构等方向出发，分析影响因素。引入person相关系数，对碳排放与经济、人口、能源消费量之间进行相关性分析，建立关联关系模型。

## 2.3 问题二三分析

对于问题二，区域碳排放量以及经济、人口、能源消费量的预测模型。首先，对于基于人口和经济变化的能源消费量预测模型。利用数据处理之后的数据进行处理，建立回归预测模型。对于要求2、3，分别构建能源消费量与人口以及经济的线性回归模型进行预测，为了增加预测精度，引入LSTM、灰色预测模型再次进行预测。根据预测的结果，建立优化模型，以精度最小为目标函数，对三种预测方式进行加权，利用优化模型求解最优的权重。对于区域碳排放量预测模型，构建碳排放量与人口、GDP和能源消费量的多元线性回归模型进行预测，再利用第一小问的预测模型，提高预测精度。

## 2.4 问题四分析

对于问题三，以2030年达到碳达峰，2060年实现碳中和为确定时间节点，设计自然、基本、雄心三种不同的情景。分别为，没有人为干预、按时达到碳达峰碳中和、率先碳达峰与碳中和情况下的碳减排措施。对于各部门的能源消费量、能源消费品种及其碳排放量的预测方法，选取与问题二相同的预测方法。确定2025年、2030年、2035年、2050年和2060年的GDP、人口和能源消费量目标。将目标与提高能源利用效率和提高非化石能源消费比重的目标相关联，确保达到碳达峰和碳中和目标的路径与措施一致。对能效提升、产业（产品）升级、能源脱碳和能源消费电气化等措施进行定性和定量分析，以确定如何实现双碳目标。

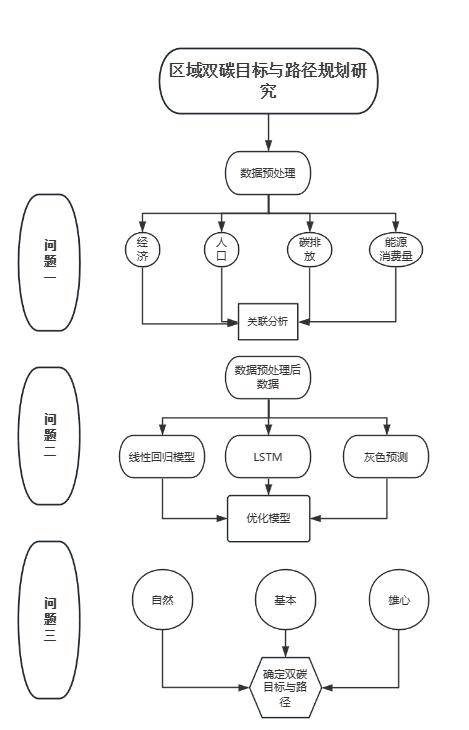


图1：思维导图

# 模型假设

为了更好的建立模型我们做出如下假设：

1. 经济增长假设：对于经济增长的模型，可以假设中国的经济将按照2020年为基期，到2035年翻一番，到2060年翻两番。这个假设可以根据中国的经济发展趋势和政策来制定。
2. 碳消纳量假设： 假设2035年和2060年的生态碳汇和工程碳汇或碳交易的碳消纳量分别为基期碳排放量的一定比例（例如10%）。这个假设可以基于森林增长、碳捕捉技术、碳交易市场等因素进行估算。
3. 能源消费结构假设： 在能源消费量预测模型中，可以假设在未来的情景下，能源消费结构会发生变化，非化石能源消费比重逐渐增加。这个假设可以基于政府政策和技术进步来制定。
4. 相关性假设： 对于相关性模型，可以假设碳排放量与人口、GDP和能源消费量之间存在线性或非线性的关联关系。这个假设可以基于历史数据分析来得出
5. 情景假设： 对于不同情景的设计，可以基于政府政策、技术发展和国际环境等因素，假设不同情景下的碳减排措施和能源消费结构调整程度。
6. 持续趋势假设： 假设历史趋势在未来一段时间内会持续，例如，过去的能源消费结构趋向非化石能源，未来也会继续朝着这个方向发展。

# 四、符号说明

为了方便我们模型的建立与求解过程 ，对使用到的关键符号进行以下说明：

|  |  |
| --- | --- |
| **符号** | **说明** |
|  | 为的一阶差分 |
| p | 熵权法计算所得到的权重 |
|  | 滞后阶数 |
|  | 第i个缺失项 |
| n | 缺失项个数 |
|  | 样本协方差 |
|  | 样本标准差 |
|  | 样本皮尔逊相关系数 |
|  | 为变量在t时刻和t-1时刻的值 |
|  | 服从正态分布 |
|  | 白噪声序列 |
|  | 各自变量系数 |
|  | 预测数据 |
|  | 真实数据 |
|  | 预测值 |
|  | 真实值 |

（注：这里只列出论文各部分通用符号，个别模型单独使用的符号在首次引用时会进行说明。）

# 五、问题一模型的建立与求解

## 5.1 数据预处理

首先对于给出的数据进行缺失值查找，将数据集导入MATLAB，通过MATLAB的find函数寻找空缺值。得到异常值都是表格中表示为“-”，在附件数据中对“-”进行描述为，数值栏中“-”表示该年该项能源未使用，无法计算实际的碳排放因子。将该异常值列举发现，存在部分“-”不符合实际情况，以2012年发电项目细分项热力为例，为例。如下所示，通过下表我们可以看出该数据题并不符合该年该项能源未使用，无法计算实际的碳排放因子。前后一直使用该种能源，理论上不会存在这种缺失情况，因此认定为异常缺失值。

表1：异常缺失值

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 年份 | **2010** | **2011** | **2012** | **2013** | **2014** | **2015** |
| 热力 | 2.906111839 | 2.916319046 | - | 2.966587506 | 2.997623473 | 3.106586243 |

对于该种缺失值，如果直接不再使用，一定会对结果带来一定的影响。因此这里使用插值填充的方式进行处理。对于缺失的数据中里面包含了数个缺失值异常值的节点，这正是我们在研究后续问题所需要的数据，因此对于这一部分缺失值，我们使用牛顿插值进行补充。[2]

计算公式为



式中，n为缺失项个数，表示第i个缺失项。

我们利用该公式，将数据导入MATLAB进行计算，得到部分结果如下所示。

表2：处理后结果

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **年份** | **2010** | **2011** | **2012** | **2013** | **2014** |
| 煤炭 | 2.66367346 | 2.66367346 | 2.6620884 | 2.652511856 | 2.66367346 |

## 5.2 指标评价体系的构建

为了更加直观的构建指标评价体系，首先根据给出的数据以及将要研究的问题选择主要指标，即一级评价指标，为经济指标、人口指标、能源消费量指标以及碳排放指标。

根据给出的数据，对于人口，选择常住人口总量作为其二级指标。

对于经济，选择GDP总量、第一产业、第二产业总量、第二产业能源供应部门、第二产业工业消费部门、第三产业总量、第三产业交通消费部门、第三产业建筑消费部门生产总值作为其二级指标。

对于能源消费量，选择总量、化石能源消费和非化石能源消费。

对于碳排放量，选择排放总量、与各部分碳排（农林消费部门、工业消费部门、第三产业总量、第三产业交通消费部门、第三产业建筑消费部门、居民生活消费）作为二级指标。

对于化石能源消费和非化石能源消费二者的划分通过查阅相关文献得出：属于化石能源消费的包括：

煤炭：煤炭是一种化石燃料，通常用于发电、加热和工业过程中。

油品：包括石油、汽油、柴油和其他石油衍生产品。这些产品主要用于交通（例如汽车和飞机），以及工业用途和暖气等。

天然气：天然气虽然是天然资源，但也被归类为化石能源，因为它主要由甲烷组成，而甲烷是一种温室气体。天然气用于发电、供热、燃料以及工业过程中。

属于非化石能源消费的包括：

热力：非化石热力可以来自多种可再生能源，如太阳能热能、地热能等，也可以来自核能等非化石能源。这些能源通常用于供热或工业用途。

电力：电力本身可以是非化石能源，如来自太阳能、风能、水力等可再生能源的电力，或者来自核能等非化石能源的电力。电力用途广泛，包括发电、交通、家庭和工业用途。）

为了更加直观展示构建的指标评价体系，对体系进行可视化处理，如下所示。

表3：指标评价体系

|  |  |  |  |
| --- | --- | --- | --- |
| **一级指标** | **二级指标** | **一级指标** | **二级指标** |
| 人口 | 常驻人口 | 能源消费量 | 总量 |
| 经济 | GDP总量 | 化石能源消费 |
| 第一产业生产总值 | 非化石能源消费 |
| 第二产业生产总值 | 碳排放量 | 排放总量 |
| 第三产业生产总值 | 农林消费部门 |
| 第二产业能源供应部门 | 工业消费部门 |
| 第二产业工业消费部门 | 第三产业总量 |
| 第三产业交通消费部门 | 第三产业交通消费部门 |
| 第三产业建筑消费部门生产总值 | 第三产业建筑消费部门 |
|  | 居民生活消费 |

### 5.3 现状分析

为了更加直观的观察分析某区域十二五（2011-2015年）和十三五（2016-2020年）期间的碳排放量状况。将给出的数据集进行拆分，利用EXCEL绘制十二五、十三五时间下变化趋势。如图2所示，

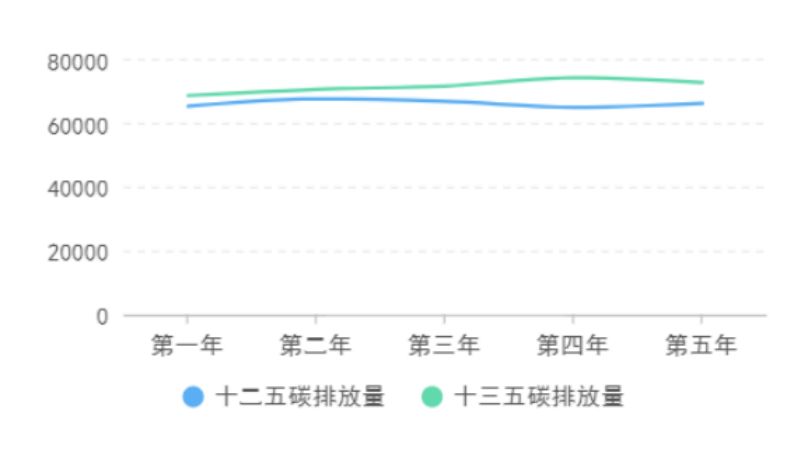


图2：十二五、十三五碳排放量变化趋势

通过折线图，可以较为清晰的看出，十三五总体的碳排放量略高于十二五。对于十二五而言，后几年有明显下降趋势。十三五后两天仍旧保持碳排放增势。对了增加直观的看出不同时期的增长率。计算十二五、十三五期间碳排放量的增长率绘制为柱状图如图3所示，

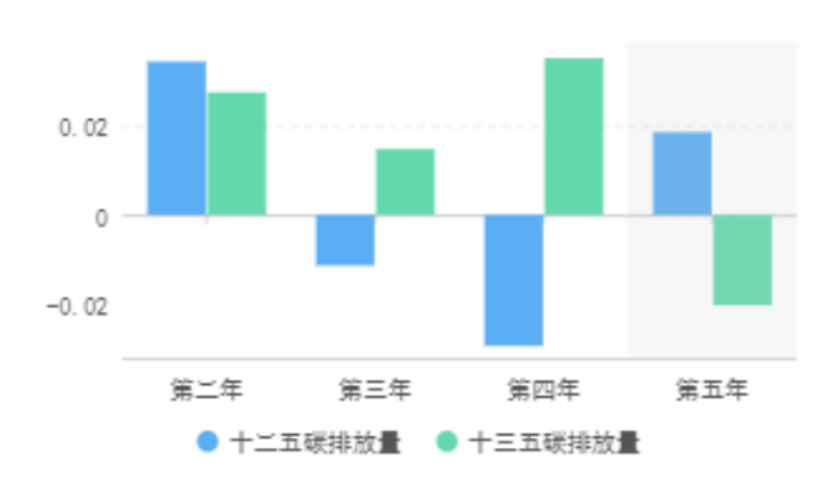


图3：十二五、十三五期间增长率

通过柱状图，可以清晰的看出十三五计划末期，碳排放已经开始降低，而十二五末期碳排放开始了增长趋势。



## 5.4 关联模型的建立与求解

为了更加直观的看各个指标之间的关系，首先绘制指标之间的散点图观察变化趋势。如下所示。

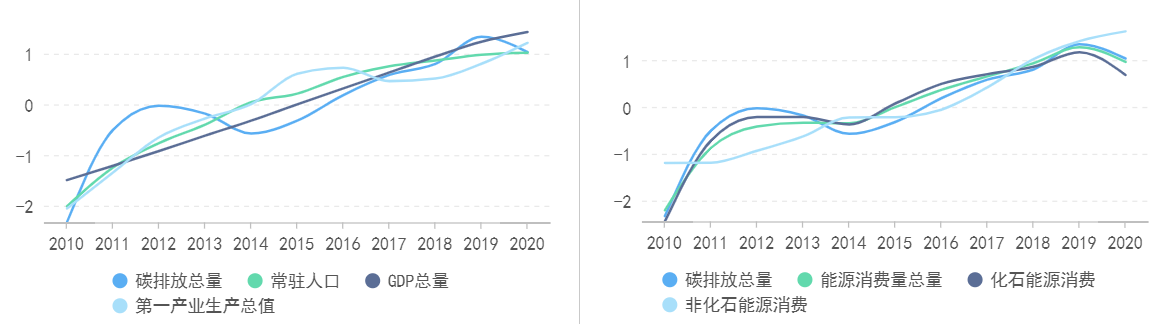


图4：碳排放量相关散点图

通过散点图，大致可以看出大部分指标都与碳排放量呈现一定的正相关。为了更加客观地判定结果，这里可以建立相关性分析模型，引入person相关性，进行判断指标之间的关系。

样本协方差：



样本标准差：



样本皮尔逊相关系数：



利用相关系数首先研究大部分指标都与碳排放量的相关关系，得到相关性分析结果如下表所示。这里首先展示二级指标之间的相关性

表4：二级指标相关性

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **GDP总量** | **第一产业生产总值** | **第二产业生产总值** | **能源供应部门** |
| GDP总量 | 1(0.000\*\*\*) | 0.919(0.000\*\*\*) | 0.998(0.000\*\*\*) | 0.985(0.000\*\*\*) |
| 第一产业生产总值 | 0.919(0.000\*\*\*) | 1(0.000\*\*\*) | 0.918(0.000\*\*\*) | 0.9(0.000\*\*\*) |
| 第二产业生产总值 | 0.998(0.000\*\*\*) | 0.918(0.000\*\*\*) | 1(0.000\*\*\*) | 0.987(0.000\*\*\*) |
| 能源供应部门 | 0.985(0.000\*\*\*) | 0.9(0.000\*\*\*) | 0.987(0.000\*\*\*) | 1(0.000\*\*\*) |
| 工业消费部门 | 0.998(0.000\*\*\*) | 0.918(0.000\*\*\*) | 1(0.000\*\*\*) | 0.985(0.000\*\*\*) |

为了更加直观的看出相关性的数值大小，这里利用SPSSPRO绘制了矩阵热力图，如下所示。

# 相关系数热力图

图5：矩阵热力图

通过相关性分析图表的可视化结果，我们可以很直观看出经济指标与能源消费量之间具有较好的相关性。为了更加直观的展示几个一级指标之间的相关性，这里对几个多维一级指标进行降维处理，得到最终数据集如下所示。

表5：一级指标降维数据集

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **年份** | **常驻人口** | **经济** | **能源消费量** | **碳排放** |  |
| 2010 | 7869.34 | -1.621 | 0.56330178 | 1.382 |  |
| 2011 | 8022.99 | -1.243 | 0.39111999 | 1.177 |  |
| 2012 | 8119.81 | -0.812 | 0.32781364 | 1.016 |  |
| 2013 | 8192.44 | -0.579 | 0.34728297 | 0.627 |  |
| 2014 | 8281.09 | -0.26 | 0.40953106 | 0.222 |  |
| 2015 | 8315.11 | 0.05 | 0.3393682 | -0.195 |  |
| 2016 | 8381.47 | 0.319 | 0.29955465 | -0.371 |  |
| 2017 | 8423.5 | 0.602 | 0.34180494 | -0.414 |  |
| 2018 | 8446.19 | 0.913 | 0.40112017 | -0.64 |  |
| 2019 | 8469.09 | 1.238 | 0.41749169 | -0.882 |  |
| 2020 | 8477.26 | 1.391 | 0.46910766 | -1.922 |  |

将降维结果导入关联分析模型，得到下面的结果，如下所示



表6：相关性分析结果

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **常驻人口** | **经济** | **能源消费量** | **碳排放** |
| 常驻人口 | 1(0.000\*\*\*) | 0.971(0.000\*\*\*) | -0.351(0.290) | -0.92(0.000\*\*\*) |
| 经济 | 0.971(0.000\*\*\*) | 1(0.000\*\*\*) | -0.158(0.643) | -0.968(0.000\*\*\*) |
| 能源消费量 | -0.351(0.290) | -0.158(0.643) | 1(0.000\*\*\*) | 0.047(0.891) |
| 碳排放 | -0.92(0.000\*\*\*) | -0.968(0.000\*\*\*) | 0.047(0.891) | 1(0.000\*\*\*) |
| 注：\*\*\*、\*\*、\*分别代表1%、5%、10%的显著性水平 | | | | |

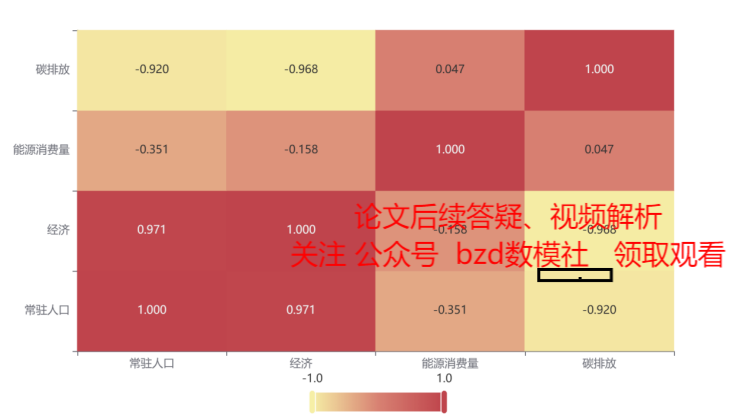


图6：相关性分析结果矩阵热力图

通过，图表表述我们可以清楚的得出，对于这四项一级指标大部分都具有较好的相关性。除了能源消费量与其他三个指标的相关性不太明显。

# 问题二模型的建立与求解

## 6.1基于人口的能源消费量预测模型

### 6.1.1关联关系

通过5.4小节中建立的关联模型可以清晰的看出常驻人口与能源消费量具有较为明显的线性关系。因此，这里建立线性回归模型，进行相关分析。如下所示



式中，服从正态分布，为自变量系数，为常数项系数。表示第i年的能源消费量，表示第i年该地的常驻人口数量。我们将收集的数据导入MATLAB，利用自带regress函数进行多元线性回归得到结果如下表所示。

表7：线性回归结果

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **非标准化系数** | | **标准化系数** | **t** | **R²** | **F** |
| B | 标准误 | Beta |
| 常数 | -70760.134 | 8593.918 | - | -8.234 | 0.937 | F=134.936 P=0.000\*\*\* |
| 常驻人口 | 12.064 | 1.039 | 0.968 | 11.616 |

F检验的结果分析可以得到，显著性P值为0.000\*\*\*，水平上呈现显著性，拒绝回归系数为0的原假设，因此模型基本满足要求。 对于变量共线性表现，VIF全部小于10，因此模型没有多重共线性问题，模型构建良好。

最终，二者的关系模型为



为了更加直观的观察二者的关系式，绘制了拟合示意图如下所示。

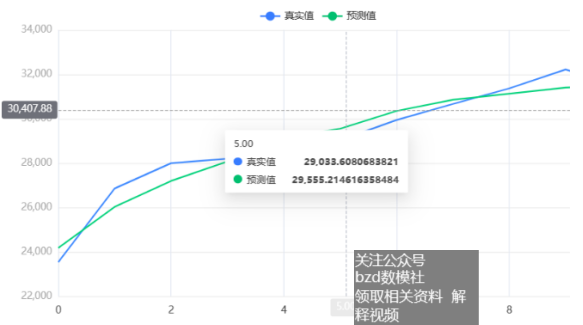


图7：拟合效果图

### 6.1.2 线性回归预测

为了对常驻人口进行预测，建立线性回归模型。根据第五节相关性分析的结果可以看出，二者具有较好的相关性。可以建立线性回归模型，用以后续的预测预报。建立线性回归模型，进行相关分析。如下所示



其中，服从正态分布，为自变量系数，为常数项系数。表示第i年该地的常驻人口数量，表示第i年。我们将收集的数据导入MATLAB，利用自带regress函数进行多元线性回归得到结果如下表所示。

表8：线性回归结果

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 非标准化系数 | | 标准化系数 | **t** | **P** | **VIF** | **R²** | **F** |
| B | 标准误 | Beta |
| 常数 | -108334.379 | 11712.369 | - | -9.25 | 0.000\*\*\* | - | 0.917 | F=99.12 P=0.000\*\*\* |
| 年份 | 57.869 | 5.813 | 0.957 | 9.956 | 0.000\*\*\* | 1 |
| 注：\*\*\*、\*\*、\*分别代表1%、5%、10%的显著性水平 | | | | | | | | |

对于得到的结果，F检验的结果分析可以得到，显著性P值为0.000\*\*\*，水平上呈现显著性，拒绝回归系数为0的原假设，因此模型基本满足要求。 对于变量共线性表现，VIF全部小于10，因此模型没有多重共线性问题，模型构建良好。因此，得到的结果为



为了更加直观的观察二者的关系式，绘制了拟合示意图如下所示。

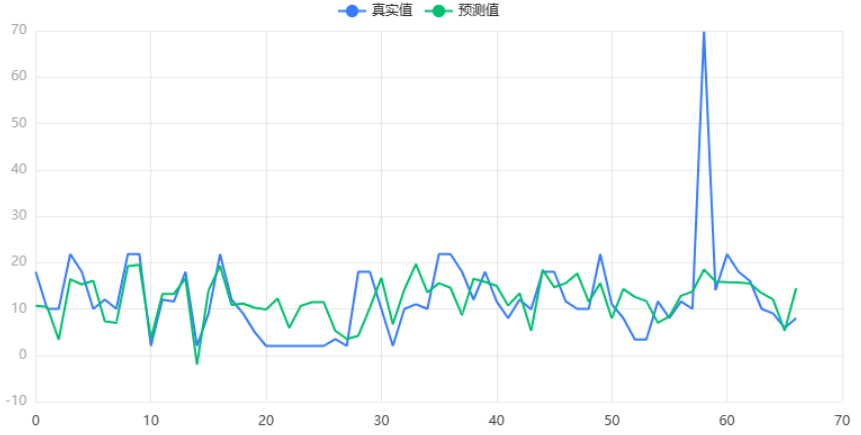


图8：拟合效果图

### 6.1.3 灰色预测

这里我们会建立GM（1,1）模型对选区的数据进行预测。

**数据的检验与处理**

为了保证模型的可行性，我们需要对进行预测的数据进行必要的检验。首先就是求序列的级比，看看是否所有的级比都落在可容覆盖内。

设原始数据为计算数列的级比。

如果所有的级比都落在可容覆盖区间内，则数列可以建立GM（1,1）模型进行灰色预测。否则，我们将选择平滑指数预测、回归预测，或者进行适当的变换处理，如平移变换：

取适合的c使得级比落在可容覆盖内。

**建立模型**

我们选取的数据已经满足要求时，以它为数据建立GM（1,1）模型。

利用回归分析求得a,b的估计值，得到白化模型为

解为

于是得到预测值

从而相应地得到预测值：

最终，根据预测结果绘制散点图图下所示

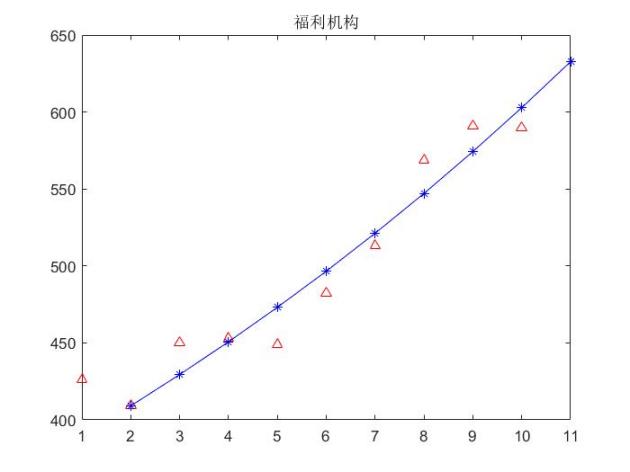


图9：拟合效果图

### 6.1.4 LSTM预测模型的建立

**数据归一化处理**

为了消除不同场景由于不同因素造成的影响，采用最大/最小归一化方法对样本数据进行归一化处理，公式如下：



式中：为原始样本数据；为原始样本数据的最小值；样本数据的最大值；x为归一化处理后的样本数据。

**网络输入与输出变量的选择**

选用三个核心指标的数据作为样本数据，确定网络输入变量的个数。模型示意图如下图所示



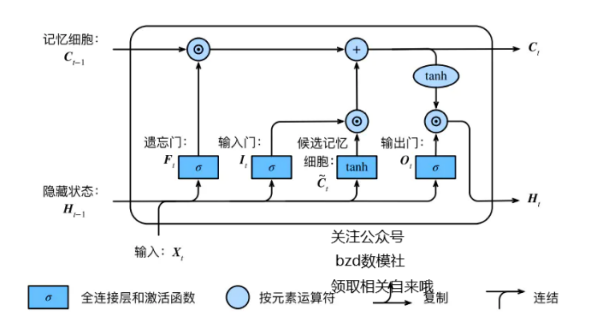


图8：LSTM模型示意图[5]

其中，输入门：



遗忘门：



输出门：



候选记忆细胞计算：



我们将数据导入Matlab，利用编写的程序进行预测。其训练进度如下图所示

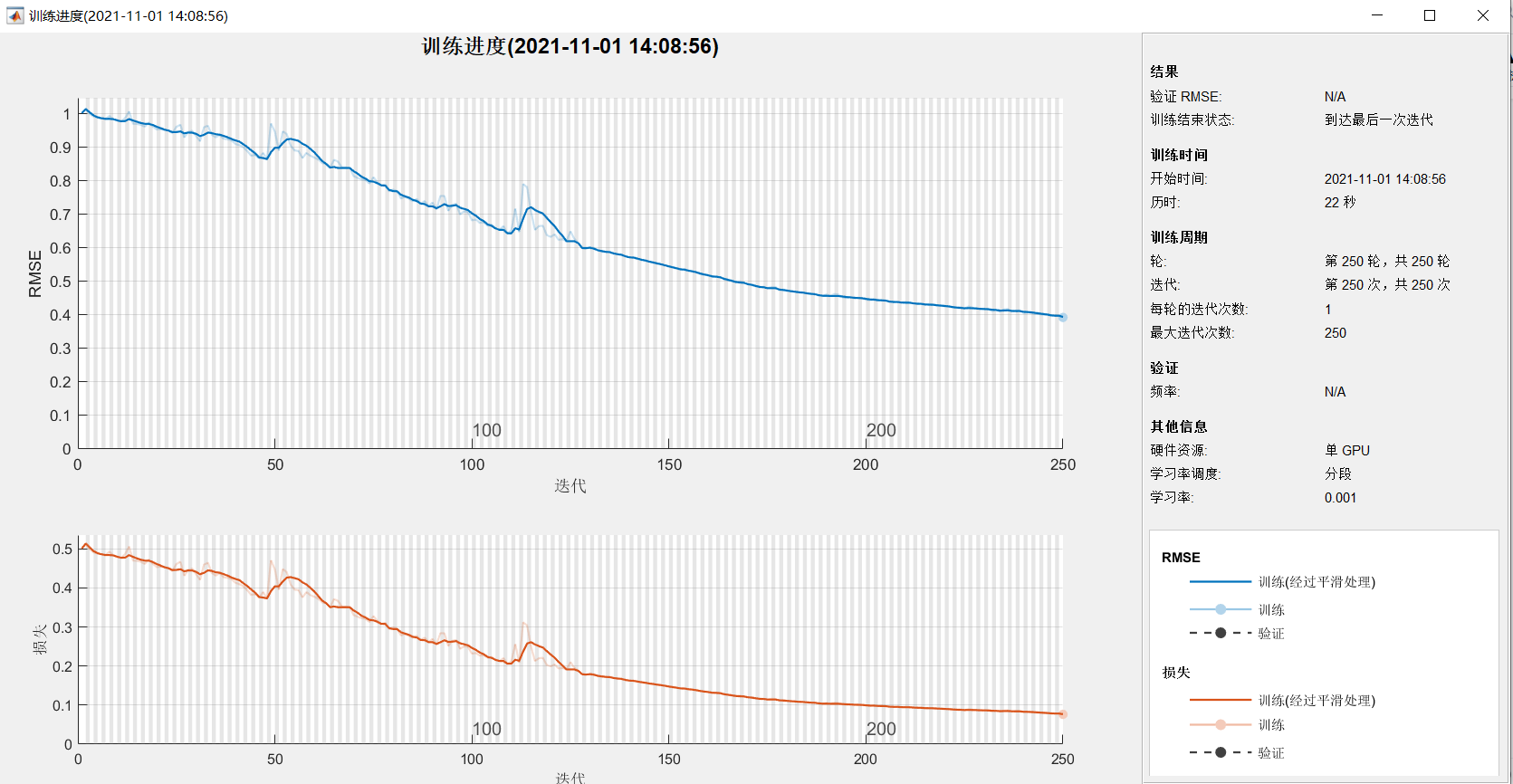


图9：训练进度

### 6.1.5 加权预测模型

对于人口建立预测模型，这里的预测模型，为了提高精度将选择线性回归模型、灰色预测、LSTM神经网络预测三种预测模型，进行预测。对于预测结果构建加权优化模型，以提高模型精度。

设 是同一个问题通过m种预测方法（每种方法都通过各自的检验）所得到的预测结果，是第种预测方法对原始数据  的模拟值。则这m种方法的加权预测值为：



其中，

表示第i种预测方法与第j个历史数据的模拟值之间的误差，记





因此，以误差最小为目标函数，权重和为1为约束条件，构建优化模型。由于我们建立的规划模型直接进行求解，无论利用lingo还是matlab都会面对巨大的运算量，耗费大量的时间。因此，对于我们建立的多元规划模型，我们将利用多目标粒子群优化算法进行求解。

粒子群算法最早在1995由Eberhart和Kennedy提出，其特点时简单易行，收敛速度块，设置参数少。这些特点正好巧妙地弥补了我们建立模型的缺陷。我们基于传统的粒子群算法，结合我们的这个题目进行改进。通过我们的规划模型，每个粒子确定每个粒子个体的最优解并从这些个体最优解找到一个全局最优值，这里涉及到了我们粒子的位置和距离公式。[6]



其中，为0到1之间随机数。



为惯性因子，其值非负。数值越大，则寻找最优解的能力越强，局部寻找最优解的能力越弱，反之结论正好相反。我们这里采用线性递减权值策略。



为最大迭代次数，为初始惯性权值，迭代到最大进化数时的惯性权值。

为了方便大家更好的理解该算法，绘制了思维导图，如下所示

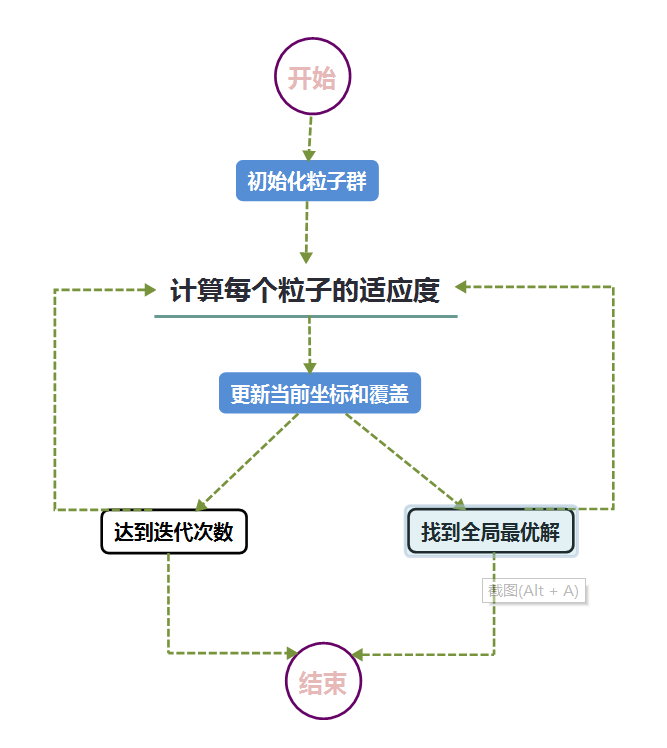


图10：优化算法粒子群

最终，得到预测结果如下所示

表9：问题二第一小问最终结果

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **年份** | **2021** | **2022** | **2023** | **2024** | **2025** |
| 人口 | 8591.621 | 8643.532 | 8695.756 | 8748.296 | 8801.153 |
| 人口-能源消费量 | 32889.18174 | 33515.43605 | 34145.46638 | 34779.30894 | 35416.97579 |
| 年份 | 2056 | 2057 | 2058 | 2059 | 2060 |
| 人口 | 7951.225784 | 7830.246394 | 7694.467128 | 7546.078094 | 7385.96663 |
| 人口-能源消费量 | 25163.45386 | 23703.9585 | 22065.91743 | 20275.75213 | 18344.16742 |

### 6.1.6 模型评估

我们这里将利用平均绝对百分比误差 MAPE对模型进行评估。例如，如果 MAPE 为 8，则表示预测结果较真实结果平均偏离8 %。

MAPE公式为：



我们利用给出的数据对2021年-2060年的人口数据进行预测，得到预测结果，作为，而我们将题目给出的当年数据作为。

最终，我们得到问题二四种模型模型的MAPE进行汇总得到结果如下表所示：

表10：模型评估结果

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **方法** | **线性回归** | **灰色预测** | **LSTM** | **加权预测** |
| 第一段预测 MAPE值 | 8.4 | 4.5 | 6.4 | 4.1 |
| 第一段预测 MAPE值 | 15.2 | 29.3 | 8.4 | 7.8 |

## 6.2 基于经济变化的能源消费量预测模型

建立与6.1节相同的模型，这里不做过多步骤展示，直接展示模型结果。

表11：线性回归模型结果

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **非标准化系数** | | **标准化系数** | **t** | **P** | **VIF** | **R²** | **F** |
|  | B | 标准误 | Beta |  |  |  |  |  |
| 常数 | 19513.993 | 1171.039 | - | 16.664 | 0.000\*\*\* | - | 0.886 | F=69.864 P=0.000\*\*\* |
| GDP总量 | 0.145 | 0.017 | 0.941 | 8.358 | 0.000\*\*\* | 1 |  |  |
| 因变量：能源消费量总量 | | | | | | | | |
| 注：\*\*\*、\*\*、\*分别代表1%、5%、10%的显著性水平 | | | | | | | | |

得到结果，如下所示



根据关系式建立相同的预测模型，得到结果如下所示

表12：问题二第二小问结果

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **年份** | **2021** | **2022** | **2023** | **2024** | **2025** |
| 经济 | 94782.85648 | 99656.16243 | 104529.4684 | 109402.7743 | 114276.0803 |
| 经济-能源消费量 | 33257.50719 | 33964.13655 | 34670.76591 | 35377.39528 | 36084.02464 |
| 年份 | 2056 | 2057 | 2058 | 2059 | 2060 |
| 经济 | 265348.5647 | 270221.8706 | 275095.1766 | 279968.4825 | 284841.7884 |
| 经济-能源消费量 | 57989.53488 | 58696.16424 | 59402.7936 | 60109.42296 | 60816.05232 |

## 6.3 区域碳排放量预测模型

### 6.3.1区域碳排放量与人口、GDP和能源消费量关联模型构建



由于对于区域碳排放量与人口、GDP和能源消费量建立多元线性回归得到的结果并不理想。因此，这里引入岭回归。在一般线性回归的基础上加入正则项，保证保证最佳拟合误差的同时，参数尽可能地简单。是的模型的泛化能力变强。一般采用一，二范数作为正则项加入回归，使得模型更具泛化性，解决线性回归中不可逆的情况。对于二范数的岭回归，如下所示：



其迭代优化函数如下所示：

## 

从最小二乘的角度来看，通过引入二范数正则项，也使得主对角线元素来强制矩阵可逆。



利用SPSSPRO进行求解，得到结果如下所示

表13：岭回归结果

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **K=0.147** | **非标准化系数** | **标准化系数** | **R²** | **调整R²** | **F** |
| B | Beta |
| 常数 | 8515.373 | - | 0.983 | 1.167 | -5.351(NaN) |
| 常驻人口 | 1.736 | 0.072 |
| GDP总量 | 0.021 | 0.07 |
| 能源消费量总量 | 0.387 | 0.199 |
| 非化石能源消费比重 | 9721.063 | 0.038 |
| 非化石能源发电比重 | 89650.488 | 0.14 |
| 农林消费部门 | 1.123 | 0.009 |
| 工业消费部门 | 0.363 | 0.115 |
| 交通消费部门 | 0.736 | 0.058 |
| 建筑消费部门 | 4.632 | 0.156 |
| 居民生活消费 | 1.128 | 0.082 |
| 能源供应部门 | 2.308 | 0.34 |
| 因变量：碳排放总量 | | | | | |
| 注：\*\*\*、\*\*、\*分别代表1%、5%、10%的显著性水平 | | | | | |

并利用SPSSPRO绘制了岭迹图，如下所示。同时，模型的拟合优度R²为0.983，模型表现为较为较为优秀。

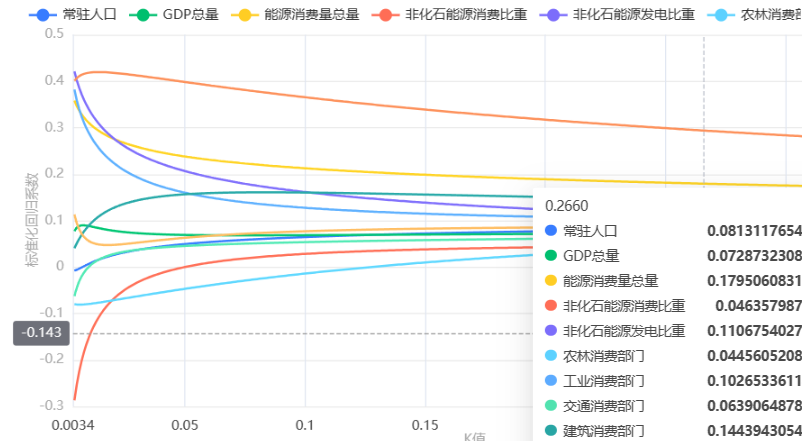


图11：岭迹图

建立得到，碳排放量与人口、GDP和能源消费量预测相关联关系式为：

 最终，根据关联式采用加权预测模型，得到结果如下所示。

表14：问题二最终结果

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **年份** | **2021** | **2022** | **2023** | **2024** | **2025** |
| 碳排放量 | 78473.416 | 76147.756 | 77751.046 | 74155.613 | 75245.457 |
| 年份 | 2056 | 2057 | 2058 | 2059 | 2060 |
| 碳排放量 | 67078.84462 | 66974.95138 | 66731.62992 | 66415.67403 | 65589.38462 |

# 问题三模型的建立与求解

7.1 情景设计

1.自然情景：这是基准情景，假设没有特别的政策干预，未来发展按照当前趋势进行。碳排放和经济增长不受特别政策影响。

2.基准情景：这个情景按照国家碳达峰和碳中和的时间节点（例如，2030年碳达峰，2060年碳中和）设定。这个情景下，制定政策和措施来确保在规定的时间内达到这些目标，模拟达到目标的路径。

3.雄心情景：这个情景可以设定更积极的碳达峰和碳中和目标，以反映更高的雄心和更积极的政策干预。这可能包括更大幅度的能源结构调整、更高效的技术创新等。

7.2 自然情景下碳排放量核算

题目给出假设如下所示

假设1、2035年的GDP比基期（2020年）翻一番；2060年比基期翻两番；

假设2、2060年生态碳汇的碳消纳量为基期碳排放量的10%；

假设3、2060年工程碳汇或碳交易的碳消纳量为基期碳排放量10%。

### 7.2.1 计算人均GDP

对于自然情景，根据假设的经济增长率，计算不同时间点的人均GDP以2035年的GDP比基期（2020年）翻一番；2060年比基期翻两番；为依据对2020年-2060年GDP数据进行插值计算。由于题设规定采取与问题二相同的条件。因此，计算问题二计算得到的关系模型，计算在没有任何干预的情况下即自然情景下的人口、能源消费量。这里使用matlab进行线性插值处理，得到2020-2060年每年的GDP数据，具体如下所示。



图12：线性插值结果

利用问题二得到，GDP与能源消费关联关系模型、人口与能源消费关联关系模型，进行计算得出，在自然情境下，常住人口与能源消费量的相关数据。





则人均GDP可以表示为



### 7.2.2 估算单位GDP能耗

利用7.2.1小节计算得到的各项数值进行计算即可。



因此

 其中



最终，得到

利用该公式，进行计算单位GDP能耗。得到部分结果如下所示，

表15：单位GDP能耗部分结果

|  |  |
| --- | --- |
| **年份** | **单位**GDP**能耗** |
| 2021 | 0.351288964 |
| 2022 | 0.339154319 |
| 2023 | 0.328367968 |
| 2024 | 0.318717022 |
| 2025 | 0.310031171 |
| 2026 | 0.302172544 |
| 2027 | 0.295028337 |

### 7.2.3 利用Kaya模型计算碳排放量

对于单位能耗二氧化碳排放量，利用给出数据集中的能源消费的碳排放因子进行判定即可。最终，引入Kaya模型，用来分析区域碳排放量和该区域人口、社会经济发展水平、能源利用效率以及碳排放因子的关系。模型主要表示为：



利用7.2.1-7.2.2得到的数据集进行计算在自然情境下的碳排放量。

表16：自然情景下碳排放量

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **年份** | **人口** | **单位GDP能耗** | **人均GDP** | **单位能耗二氧化碳排放量** | **碳排放量** |
| 2021 | 8619.89922 | 0.351288964 | 10.97407597 | 2.165263796 | 71952.43091 |
| 2022 | 8690.959489 | 0.339154319 | 11.56461992 | 2.170118533 | 73974.13526 |
| 2023 | 8762.019757 | 0.328367968 | 12.14558521 | 2.148444048 | 75077.103 |
| 2024 | 8833.080025 | 0.318717022 | 12.71720301 | 2.173086307 | 77801.14763 |
| 2025 | 8904.140293 | 0.310031171 | 13.27969712 | 2.080365754 | 76264.99046 |

对于利用Kaya模型计算碳排放量于沿用问题二求解思路求解得到的各项部门碳排放量求和，得到的结果近似相同。即，判定区域碳排放与各部门碳排放量的总和服从一致性。

## 7.3 其他情景碳排放量核算

### 7.3.1 基准情景碳排放量核算

基于假设1的前提，提前确定各个时间点的GDP含量。基于Kaya模型进行求解运算。



对等式两边同时对时间t求导，得到



 利用2035年时，2060年时，对2035、2060年的碳排放增长率进行求解。利用数年的碳排放数据以及GDP数据利用问题二两个等式，对常住人口、以及能源消耗量进行计算。得出在基准情景下的各指标的对应数值，计算各部门碳排放量。结果如下所示。

表17：各部门碳排放量以及总碳排放部分数据

|  |  |  |  |
| --- | --- | --- | --- |
| **年份** | **各部门之和** | **碳排放量** | **变化率** |
| 2021 | 76275.78591 | 76847.11159 | 0.007434576 |
| 2022 | 77533.04694 | 78491.5012 | 0.01221093 |
| 2023 | 79869.6888 | 79935.8908 | 0.000828189 |
| 2024 | 81141.52649 | 81380.28041 | 0.002933806 |
| 2025 | 82413.36417 | 82824.67002 | 0.004965982 |
| 2026 | 78491.5012 | 78790.30797 | 0.003792431 |
| 2027 | 79935.8908 | 80047.569 | 0.001395148 |

可以看出绝大部分结果的变化了率均小于0.01。因此，可以认定为服从一致性。

### 7.3.2 雄心情景碳排放量核算

对于雄心情景，我们以基准情景获得的各指标数据为样本，计算每年的变换率，求解40年的平均变化率。改变分母即可加速实现碳中和。以常住人口为例，我们利用基准情景下2021-2060年每年的常主人口数量计算平均增长率。可以表示为



其中，为基准情景下的平均增长率，为了提前实现双碳目标。这里对于雄心情景下的人口平均增长率选择为



这样可以达到更积极的碳达峰和碳中和目标，以反映更高的雄心和更积极的政策干预。这可能包括更大幅度的能源结构调整、更高效的技术创新等。

## 7.4 确定双碳（碳达峰与碳中和）目标与路径

对于最终确定双碳（碳达峰与碳中和）目标与路径。基于问题二三得出的等式，进行汇总，进行求解即可。最终建立的双碳（碳达峰与碳中和）目标与路径模型为：



最终结果如下所示，

表18：最终结果

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **年份** | **2025** | **2030** | **2035** | **2050** | **2060** |
| GDP | 123565.279 | 169089.3292 | 202907.195 | 304360.7926 | 374597.8985 |
| 人口 | 8391.426024 | 8863.976807 | 9215.01453 | 10106.1103 | 10997.20605 |
| 能源消费量 | 36766.35186 | 40126.85447 | 46343.7843 | 58441.59371 | 68523.10156 |
| 非化石能源消费 比重 | 21.17% | 26.57% | 38.66% | 71.63% | 89.17% |
| 单位GDP能耗 (利用效率) | 0.018069345 | 0.01463616 | 0.01156249 | 0.00624372 | 0.00405846 |

# 模型评价

## 8.1 模型优点

1. 综合性分析：对问题进行了全面的综合性分析，包括建立数学模型、数据预处理、多种预测方法的比较和分析、情景设计、路径规划等多个方面，确保了问题的全面解决。

2. 数据处理：对原始数据进行了预处理，包括降维、处理异常值和缺失值等，确保了模型的数据质量。

3. 多模型比较：不仅使用了多元线性回归模型，还引入了LSTM、灰色预测等多种方法，以提高预测精度，并采用加权平均的方法综合考虑各模型的结果，这有助于减小模型预测的误差。

4. 情景设计：您在问题三中设计了不同情景，考虑了不同的政策和发展路径，这有助于更全面地评估碳达峰和碳中和目标的可行性。

5. 模型假设明确：您明确列出了模型中的各种假设，这有助于透明地展示了模型的基础和前提条件。

6. 路径规划：您提出了路径规划，包括政策、技术和产业结构等方面的措施，以实现碳达峰和碳中和目标，这是解决碳减排问题的关键步骤。

7. 关键词总结： 您在最后提供了关键词总结，清晰地概括了您的求解过程，有助于理解和传达您的分析方法。

总体而言，求解过程考虑了问题的多个方面，采用了多种方法和技术，以全面、系统地解决问题，这是一个非常有价值的分析和建模过程。

## 8.2 模型的缺点

1.模型的复杂性：在模型中引入了多个预测方法和加权平均，这增加了模型的复杂性。复杂的模型可能难以解释和调试，而且需要更多的计算资源。在选择模型时，需要权衡模型的复杂性与性能之间的关系。

2.假设的不确定性：模型中的假设可能会受到未来变化和不确定性的影响。对于一些关键假设，建议进行敏感性分析，以了解它们对模型结果的敏感程度，并考虑不同的假设情景。

3.数据可用性：提到了数据的预处理，但在实际应用中，数据的质量和可用性可能会受到限制。如果数据质量不高或存在大量缺失值，可能需要更复杂的数据处理方法或考虑使用替代数据源。

4.情景设计的复杂性：在问题三中，情景设计是一个复杂的任务，涉及政策、技术和发展路径的多个因素。在设计情景时，需要确保这些情景是合理的、可行的，并且能够提供有用的政策建议。

5.模型验证与校准： 尽管建立了多个模型，但对于这些模型的验证和校准过程未明确提及。在实际应用中，需要验证模型的准确性，并可能需要进行参数校准以提高模型的性能。

6.不确定性传播： 模型中的不确定性可能会传播到最终的路径规划和政策建议中。考虑使用蒙特卡洛模拟等方法来估计不确定性对决策的影响。

7.细节和实施： 在路径规划和政策建议中，需要更多关注细节和实施的问题。具体的政策和技术措施如何实施、监测和评估等方面需要更详细的讨论。

总之，分析方法和建模过程非常有价值，但需要在模型的复杂性、数据质量、假设不确定性和模型验证等方面进行更深入的考虑和改进。这将有助于提高模型的准确性和实用性，以支持更可靠的决策制定。

## 8.3 模型的推广

1.企业管理和可持续性战略：企业可以使用模型来制定碳减排战略和可持续性计划。模型可以帮助企业评估不同决策对其碳足迹的影响，例如产品设计、供应链管理和能源消耗优化。

2. 城市规划和可持续城市发展：城市可以使用模型来规划碳中和城市或碳减排目标，以改善城市的气候韧性和环境可持续性。模型可以帮助城市评估交通、建筑、能源和废弃物管理等领域的碳减排潜力。

3. 新能源项目和能源市场分析：模型可以应用于评估新能源项目（如风电、太阳能、储能等）的可行性，并分析能源市场的发展趋势。这对于投资者和能源公司来说尤其有用。

4. 碳中和认证和标准遵循：模型可以帮助组织和企业评估其碳中和认证和标准遵循的程度，以确保他们达到所承诺的碳减排目标。

5. 气候政策评估：政府和政策制定者可以使用模型来评估不同气候政策和法规的影响，以支持决策制定和政策分析。

6. 生态系统管理和碳汇管理：模型可以用于评估生态系统的碳消纳能力，以支持生态保护和可持续土地管理。

7. 能源效率项目和节能计划：组织可以使用模型来评估能源效率项目的效益，并制定节能计划，以减少碳排放和能源成本。

8. 气候变化教育和宣传：模型可以用于教育和宣传，帮助公众和决策者更好地理解气候变化问题和碳减排解决方案。

最终，推广模型需要考虑不同领域的需求、数据可用性和政策背景。成功的推广需要与潜在用户、决策者和利益相关者进行密切合作，以确保模型满足他们的需求，并在实际应用中产生有意义的影响。

# 参考文献

1. 张一林，郁芸君，陈珠明.人工智能、中小企业融资与银行数字化转型[J].中国工业经济，2021（12）：69-87.DOI：10.19581/j.cnki.ciejournal.2021.12.003.
2. 杨寅,刘勤,吕晓雷.中国企业智能财务应用现状及发展趋势分析——基于2021年调查问卷数据的例证[J].会计之友,2022(20):111-117.
3. 刘薇,宁晓红.护士安宁疗护知信行及培训需求调查问卷的编制及信效度检验[J].实用老年医学,2022,36(11):1170-1173+1184.
4. 史枚翎.移动APP对图书馆社会化阅读推广服务的影响因素研究——基于调查问卷的实证分析[J].图书馆研究与工作,2022(06):16-20.
5. 贺宇,赵建航,孔朝阳.中国产品在东盟市场的印象研究——基于印尼、泰国、老挝三国调查问卷的分析[J].科学决策,2022(10):125-137.
6. 吕承超,崔悦.乡村振兴发展:指标评价体系、地区差距与空间极化[J].农业经济问题,2021(05):20-32.DOI:10.13246/j.cnki.iae.2021.05.004.
7. 裴征，朱晓伟，龚超.基于层次分析法和熵值法组合的DRG指标评价体系权重赋值研究[J].中国医院管理，2020，40（11）：69-72+83.
8. 崔明明,聂常虹.基于指标评价体系的我国粮食安全演变研究[J].中国科学院院刊,2019,34(08):910-919.DOI:10.16418/j.issn.1000-3045.2019.08.009.
9. 张宝友,杨玉香,孟丽君.物流服务质量评价模型与方法研究综述[J].中国流通经济,2021,35(02):49-60.DOI:10.14089/j.cnki.cn11-3664/f.2021.02.005.
10. 谢幼如,常亚洁.绩效导向的教育信息化评价模型的构建[J].中国电化教育,2015(01):56-61+92.
11. 刘晋,程彦斌,齐东川等.基于支持向量机的化工工艺安全评价模型构建及优化 研究[J].中国安全生产科学技术,2022,18(12):154-161.

# 附录

相关性分析代码

矩阵热力图代码

MAPE代码

1. 相关性分析代码

import numpy as np

from scipy.stats import pearsonr

# 计算Pearson相关系数和p-value

correlation\_coefficient, p\_value = pearsonr(x, y)

# 打印Pearson相关系数

print(f'Pearson相关系数：{correlation\_coefficient}')

print(f'p-value：{p\_value}')

# 判断相关性的强度

if correlation\_coefficient > 0.7:

print('强正相关')

elif correlation\_coefficient < -0.7:

print('强负相关')

elif correlation\_coefficient > 0.3:

print('正相关')

elif correlation\_coefficient < -0.3:

print('负相关')

else:

print('无相关性')

1. 矩阵热力图代码

% 绘制热力图

heatmap(matrix, 'Colormap', cool, 'ColorbarVisible', 'on', 'CellLabelColor', 'none', 'FontSize', 12);

% 添加标题

title('矩阵热力图');

% 显示热力图

1. MAPE评价指标代码

a=[];

a1=mean(a);

mape=mean(abs((a1-a)./a));