# 论文标题

## 问题重述

### 问题背景

### 1.2 问题提出

#### 1.2.1 问题一

依据原题可知，问题一需要对区域碳排放量以及经济、人口、能源消费量进行现状分析，分析内容需包含三个部分：

1. 建立描述该区域和各部门碳排放状况的指标和指标体系；
2. 分析碳排放量的状况、碳排放的影响因素以及实现双碳目标的挑战；
3. 建立各指标之间的关联模型。

#### 1.2.2 问题二

依据原题可知，问题二需要建立区域碳排放量以及经济、人口、能源消费量的预测模型，由此可以进一步细化成两个小问：

1）基于人口与经济变化，建立一个能源消费量的预测模型；

2）建立一个与人口、GDP、能源消费量、能源消费以及供应部分相关联的区域碳排放量预测模型。

## 问题分析

### 2.1 问题一分析

### 2.2 问题二分析

问题二分为两个部分，第一个部分是进行区域能源消费量的预测，第二部分是进行区域碳排放量的预测的，其中其一部分是第二部分预测的前提与基础：

#### 2.2.1 区域能源消费量预测

由题意可知需要基于区域的人口与经济变化进而对区域能源消费量进行预测，因此首先对区域人口建立Logistic人口增长模型以及对区域经济（地区生产总值）建立时间序列预测模型（ARIMA）模型，分别预测出区域的人口与经济变化。基于人口与经济变化预测结果进行岭回归（相关分析存在多重共线性问题）预测出区域的能源消耗量。

#### 2.2.2 区域碳排放量预测

该小问基本思路同第一问相似，但需要进行细化，即针对不同部门及其对应的能源消耗品种分别进行碳排放量预测。首先需要利用表格提供的区域各部门的增加值变化量进行预测2021~2060年的各部门的增加值(短期灰色预测+长期ARIMA预测)，进而结合第一问预测出的人口进行回归预测出各部门的能源消耗量，其中能源部门需要进一步细化预测出能源消耗中发电、供热、其他转换以及损失这四部分的分量。第二步是预测各部门能源消费品种的结构，根据2010~2020年间的各部门能源消费品种比例进行短期灰色预测+长期ARIMA预测得到2021~2060年间的各部门能源消费品种结构，进而结合各部门能源消费量预测出各部门多对应各种能源的消费量。第三步是对各部门的各种能源碳排放因子预测。利用第二步中预测的各部门的能源消费品种结构预测比例分别对各部门每种碳排放因子进行岭回归预测，以体现消费品种结构变化对区域碳排放因子影响。最后将第二步得到的各部门各种能源消耗量与第三步各部门各种能源的碳排放因子相乘，计算得到区域的碳排放量预测结果。这个结果与2010~2020年实际结果进行对比，并进行适当的比例调整，以提高预测精度，得到最终的2021~2060年区域的碳排放量预测。

## 模型假设

### 3.1 相关性假设

对于回归模型，可以假设能源消耗量与人口、区域生产总值间存在关联关系，可以通过历史数据进行分析得出。

### 3.2 持续性假设

对于当前的数据所呈现的趋势，如人口、经济增加、能源消耗量与消费品种结构等，假设仍然可以在预测时间段内保持相同的发展趋势。

## 符号说明

## 问题一模型的建立和求解

## 问题二模型的建立和求解

### 6.1 基于人口和经济变化的能源消费量预测模型

#### 6.1.1 模型建立步骤

根据题目要求能源消费量的预测需要建立起能源消费量与人口与经济变化之间的关系，因此能源消费量预测主要分为三个步骤：

1）建立区域人口预测模型，预测2021~2060年间区域的人口变化；

2）建立区域的经济预测模型，即对地区2021~2060年间的生产总值进行预测；

3）根据区域的人口与经济变化预测，对地区2021~2060的能源消费量进行预测。

#### 6.1.2 人口预测模型建立

人口预测模型采用Logistic人口增长模型。记*x*(*t*)为*t*年的人口数量，设人口年增长率*r*(*x*)为*x*的线性函数，即。自然资源与环境条件所能容纳的最大人口数为，当时增长率，此时可以建立Logistic人口模型[10]:

 (6.1)

其解为:

 (6.2)

式中。上式模型中需要求解出未知参数与，因此需要利用提供的2010~2020年的人口数据对模型参数进行非线性最小二乘拟合，拟合效果如图6-1所示，得到模型如式6.3。



图6-1 Logistic人口模型拟合效果

 (6.3)

利用式6.3的模型预测得到2021~2060年该地区的人口如表6-1，从中不难发现人口变化逐渐趋于地区所能容纳的最大人口。

表6-1 地区人口预测

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 年份 | 人口(万人) | 年份 | 人口(万人) | 年份 | | 人口(万人) | 年份 | 人口(万人) | 年份 | 人口  (万人) |
| 2021 | 8497.43 | 2029 | 8542.05 | 2037 | 8548.62 | | 2045 | 8549.59 | 2053 | 8549.73 |
| 2022 | 8508.55 | 2030 | 8543.69 | 2038 | 8548.86 | | 2046 | 8549.62 | 2054 | 8549.73 |
| 2023 | 8517.31 | 2031 | 8544.99 | 2039 | 8549.05 | | 2047 | 8549.65 | 2055 | 8549.74 |
| 2024 | 8524.22 | 2032 | 8546.00 | 2040 | 8549.20 | | 2048 | 8549.67 | 2056 | 8549.74 |
| 2025 | 8529.66 | 2033 | 8546.80 | 2041 | 8549.32 | | 2049 | 8549.69 | 2057 | 8549.74 |
| 2026 | 8533.94 | 2034 | 8547.43 | 2042 | 8549.41 | | 2050 | 8549.70 | 2058 | 8549.74 |
| 2027 | 8537.31 | 2035 | 8547.93 | 2043 | 8549.48 | | 2051 | 8549.71 | 2059 | 8549.74 |
| 2028 | 8539.96 | 2036 | 8548.32 | 2044 | 8549.54 | | 2052 | 8549.72 | 2060 | 8549.75 |

#### 6.1.3 经济预测模型建立

地区的经济表现主要体现在区域的生产总值，因此对地区的经济预测可以通过对地区的生产总值预测实现。由于地区的生产总值是一个按年份变化的时间序列，因此自然想到利用时间序列分析对区域的生产总值进行预测，这里采用ARIMA(Autoregressive Integrated Moving Average Model, 自回归差分移动平均模型)模型进行预测。

记为*t*年地区生产总值，为*t*-*p*年地区生产总值，为*t*年与*t*-*q*年地区生产总值间的误差，根据ARIMA模型公式建立其地区经济预测模型如式6.4：



(6.4)

式中c为常数项，为AR（自回归）模型参数，用来描述当前GDP（生产总值）与过去*p*年GDP之间的关系，为MA（移动平均）模型参数，用来描述当前GDP与过去*q*年间GDP的误差关系，为*t*年的GDP误差项。由于利用该模型的前提是处理的时间序列是平稳的，因此当序列尚未处于平稳状态时需要进行差分处理，差分的阶数*d*由实际序列的平稳程度。

从上述模型公式来看，建立该地区GDP的ARIMA模型需要确定3个参数即自回归的长度*p*，移动平均的长度*q*以及差分阶数*d*。本文采用遍历参数法，并利用AIC (Akaike information criterion)与BIC (Bayesian Information Criterion)相结合的方法进行模型评估，综合选取最优的(*p*, *q*, *d*)参数，具体算法框图如图6-2所示。



图6-2 ARIMA最优参数确定算法

根据图6-2的模型参数确定算法对GDP预测模型的参数进行预测，得到的相关参数与结果如表6-2。

表6-2 ARIMA模型参数

|  |  |
| --- | --- |
| 参数名称 | 参数值 |
| *p* | 5 |
| *q* | 0 |
| *q* | 2 |
| 最小AIC | -24.9099 |
| 最小BIC | -24.4333 |

利用表6-2中的模型参数对地区GDP进行预测，首先选取70%数据作为训练集，30%数据作为测试集进行预测效果检验，结果如图6-3所示，图中2010至2016年为训练集数据，2018至2020年为测试集数据，不难发现模型预测结果与测试集数据基本一致，因此可以继续利用该模型对2021至2060年的地区GDP进行预测，预测结果如表6-3所示。



图6-3 ARIMA模型GDP预测效果检验

表6-3 2021至2060年地区GDP预测

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 年份 | GDP  (亿元) | 年份 | GDP  (亿元) | 年份 | | GDP  (亿元) | 年份 | GDP  (亿元) |
| 2021 | 95582.25 | 2031 | 144308.97 | 2041 | 191931.56 | | 2051 | 238785.09 |
| 2022 | 100082.31 | 2032 | 148954.81 | 2042 | 196882.73 | | 2052 | 243862.57 |
| 2023 | 104848.56 | 2033 | 153679.04 | 2043 | 201658.28 | | 2053 | 248670.38 |
| 2024 | 109398.29 | 2034 | 158296.52 | 2044 | 206533.51 | | 2054 | 253699.60 |
| 2025 | 114140.89 | 2035 | 162701.59 | 2045 | 210977.74 | | 2055 | 258469.10 |
| 2026 | 119134.30 | 2036 | 167448.28 | 2046 | 215540.10 | | 2056 | 263028.32 |
| 2027 | 123977.39 | 2037 | 172042.42 | 2047 | 220071.93 | | 2057 | 267702.65 |
| 2028 | 129188.74 | 2038 | 176852.23 | 2048 | 224459.42 | | 2058 | 271993.27 |
| 2029 | 134259.90 | 2039 | 181911.28 | 2049 | 229248.74 | | 2059 | 276465.97 |
| 2030 | 139222.31 | 2040 | 186772.62 | 2050 | 233917.74 | | 2060 | 280989.19 |

#### 6.1.4 能源消费量预测模型

依据题目要求地区能源消费量的预测需要与人口预测以及经济预测相关联，因此拟采用回归预测利用前两小节预测的人口与GDP变化对地区能源消费量进行预测。

最初打算采用多元线性回归，回归自变量为地区的人口与GDP。然而对自变量进行相关系数检验发现二者相关系数高达0.9588，这表明变量间存在多重共线性问题，故最终采用岭回归对线性回归进行正则化，实现对区域能源消耗量的预测。记*F*(*t*)为第*t*年的能源消耗量，*x*(t)和*E*(t)分别为第*t*年的人口与GDP，*b*0为常量，建立能源消耗岭回归预测模型如式6.5：

 (6.5)

利用MATLAB的岭回归预测函数进行模型预测时需要输入一个岭回归参数*k*，这里通过遍历查找最优的参数，最优判断准则是预测出的数据与原始数据具有最小的相对误差，判定结果如图6-4的子图1，可以看出在*k*=1附近，相对误差存在极小值，因此选取该点对应的*k*值最为输入参数。在确定参数*k*后，将2010~2020年的能源消耗量作为岭回归预测的因变量，人口与GDP数据作为自变量，进行岭回归预测，得到岭回归系数分别为-32364、6.9702与0.0568，回归预测效果如图6-4的子图2，可以看出在2010~2020年间岭回归拟合数据与原始数据的趋势与数值均较为吻合，故可以利用该模型进一步预测出2021至2060年的区域能源消费总量如表6-4所示。



图6-4 岭回归预测能源消费量

表 6-4 2021至2060年地区能源消费量预测

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 年份 | 能源  消费量  (万tce) | 年份 | 能源  消费量  (万tce) | 年份 | | 能源  消费量  (万tce) | 年份 | 能源  消费量  (万tce) |
| 2021 | 32316.94 | 2031 | 35418.28 | 2041 | 38155.58 | | 2051 | 40821.70 |
| 2022 | 32650.24 | 2032 | 35689.45 | 2042 | 38437.67 | | 2052 | 41110.39 |
| 2023 | 32982.26 | 2033 | 35963.58 | 2043 | 38709.64 | | 2053 | 41383.73 |
| 2024 | 33289.02 | 2034 | 36230.44 | 2044 | 38987.17 | | 2054 | 41669.65 |
| 2025 | 33596.52 | 2035 | 36484.30 | 2045 | 39240.11 | | 2055 | 41940.81 |
| 2026 | 33910.22 | 2036 | 36756.84 | 2046 | 39499.71 | | 2056 | 42200.00 |
| 2027 | 34209.03 | 2037 | 37020.13 | 2047 | 39757.51 | | 2057 | 42465.73 |
| 2028 | 34523.76 | 2038 | 37295.22 | 2048 | 40007.07 | | 2058 | 42709.64 |
| 2029 | 34826.58 | 2039 | 37584.12 | 2049 | 40279.44 | | 2059 | 42963.90 |
| 2030 | 35120.12 | 2040 | 37861.50 | 2050 | 40544.94 | | 2060 | 43221.03 |

### 6.2 区域碳排放量预测模型

#### 6.2.1 模型建立步骤

依据题目要求地区碳排放量预测需要与多个变量相关联，包括地区人口、经济、各部门能源消耗量以及消费品种结构等，故需要首先对这些变量进行预测，然后建立变量与地区碳排放量间的关联，建立步骤如下：

1）建立消费部门能源消费量预测模型，采用岭回归预测，预测的变量为地区人口与该部门的增加值。部门的增加值预测采用灰色预测与ARIMA预测相结合的方式，先利用灰色预测进行短期预测进行数据量扩充，提升ARIMA预测的可靠性，然后利用ARIMA模型进行长期预测完成。

2）建立能源供应部门能源消费量预测模型。基本思路与消费部门的能源消费量预测一致，但能源供应部门需要进一步对能源消费量细化为发电、供热、其他转换与损失四种能源供应形式，预测模型与消费部门预测模型相同。

3）部门能源消费品种结构预测模型。消费品种结构可以用每年各能源消费比例来体现，因此需要针对不同部门各种能源消费量进行预测，然后计算出每年对应能源的比例。各种能源的消费量预测仍然采用短期灰色预测与长期ARIMA模型预测相结合的方式。在得到了部门的能源消费品种结构后，可以利用已经预测出的部门能源消费量乘上对应的消费品种比例，可获得对应品种的消费量。

4）碳排放因子预测模型。碳排放因子的预测需要与各部门的能源消费品种比例建立关联，因此利用已经预测出的消费品种比例作为回归变量对部门的各种能源的碳排放因子进行岭回归预测，得到各种碳排放因子预测结果

5）碳排放量预测模型。地区的碳排放量预测为各部门的碳排放量预测结果之和，各部门的碳排放量预测为部门所消耗的各种能源消费量预测与对应能源的碳排放因子相乘，因此结合此前的预测可以最终计算出碳排放量预测结果。

6）模型验证与调整。由于预测的碳排放量是根据相关联的量进行综合预测的结果，因此需要将碳排放量预测值与2010至2020年实际碳排放量进行变化趋势与实际大小的对比。倘若二者趋势一致就验证了预测的模型是可靠的，此时只需要根据实际碳排放量适当调整碳排放量预测结果，使得预测结果更加贴近实际碳排放量即可。

#### 6.2.3 消费部门能源消费量预测

各消费部门的能源消费量采用岭回归预测，回归预测的自变量为地区的人口以及对应部门的增加值。在问题二的第一问中已经建立了人口预测模型，因此现在需要建立各部门的经济预测模型。

考虑到各部门的GDP数据量很少且波动较大，在进行直接ARIMA模型预测过程中出现了明显的预测误差，故采用短期灰色预测结合长期ARIMA预测的方法进行综合预测。利用灰色系统预测模型中的GM(1,1)模型建模过程如下：

1）记消费部门的增加值为*A*，2010至2020年某消费部门的增加值为。为弱化原始增加值序列的波动性与随机性对数据进行累加，得到新的数列，其中 ；

2）生成的邻均值等权数列，其中 ；

3）根据灰色理论对建立关于时间t的一阶微分方程：

 (6.6)

式中a,u均为微分方程的待定参数，记a,u构成的矩阵为灰参数矩阵。将灰参数求出后即可通过式6.6求出，进而求出的预测值。

1. 将累加生成数据做均值生成*B*与常数项向量如式6.7：

 (6.7)

5)利用最小二乘法求解出灰参数矩阵。将灰参数代入式6.6进行求解得：

 (6.8)

6)将式6.8的累加结果累减还原得到实际预测值：

 (6.9)

式6.9中前11个数据为2010至2020年间的原始数据，后m个数据为打算预测的数据。本文中统一设置m为10，即利用GM(1,1)模型预测2021至2030年的数据。

#### 6.2.4 各部门GDP预测

#### 6.2.5 各部门各品种能耗和碳排放因子预测

#### 6.2.7 区域碳排放量预测结果

## 模型的评价

## 参考文献

[10] 司守奎, 孙玺菁. Python数学实验与建模[M]. 北京: 科学出版社, 2020.

[11] 木羽Cheney. 时间序列模型(四)：ARIMA模型[EB/OL]. (2023-06-02)[2024-01-14]. <https://zhuanlan.zhihu.com/p/634120397>.

## 附录

### 9.1 问题1代码

### 9.2 问题2代码