



西南科技大学

Southwest University of Science and Technology

本科毕业设计（论文）

题目名称：中国碳核算可视分析系统设计与实现

学院名称	计算机科学与技术学院
专业名称	计算机科学与技术
学生姓名	田永强
学号	5120191028
指导教师	张晓蓉 讲师 王赋攀 讲师

二〇二三年六月

西南科技大学

本科毕业设计（论文）学术诚信声明

本人郑重声明：所呈交的毕业设计（论文），是本人在导师的指导下，独立进行研究工作所取得的成果。除文中已经注明引用的内容外，本论文不包含任何其他个人或集体已经发表或撰写过的作品成果。对本文的研究做出重要贡献的个人和集体，均已在文中以明确方式标明。本人完全意识到本声明的法律结果由本人承担。

作者签名：

日期： 年 月 日

西南科技大学

本科毕业设计（论文）版权使用授权书

本毕业设计（论文）作者同意学校保留并向国家有关部门或机构送交论文的复印件和电子版，允许论文被查阅和借阅。本人授权西南科技大学可以将本毕业设计（论文）的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存和汇编本毕业设计（论文）。

保密☐，在____年解密后适用本授权书。

本论文属于

不保密☐.

（请在以上方框内打“√”）

作者签名：

指导教师签名：

日期： 年 月 日

日期： 年 月 日

中国碳核算可视分析系统设计与实现

摘要: 面对日益严峻的温室气体排放问题,世界各国都希望对温室气体进行有效的治理。目前国内外的研究大部分是对宏观和微观角度进行研究,缺少对中观城市角度的分析。本系统依托中国碳核算数据库中提供的县级碳排放量、植被固碳量、分部门核算清单等数据来建立可视分析系统,旨在研究全国碳排放和植被固碳量时空分布状况和碳排放来源组成,以及各个城市的碳排放、植被固碳量特点和趋势。系统通过概览页面让用户快速了解碳核算的相关知识,为了充分利用数据信息进行可视化分析,将页面分为省级分析和市级分析页面。省级分析模块从空间和时间维度增强用户对数据的认知和对中国的能源结构的了解。通过分析碳排放量的变化,对了解我国碳排放方针和政策的落实效果有一定借鉴意义。通过市级分析并结合相关资料,观察城市群变迁规律,对城市群或个别城市的特征进行总结,城市区县进行对比,了解区县特点,有利于协同发展。系统采用前后端分离的设计,使用 D3 和 Echarts 进行可视化展示。通过改写的 K-Means 算法来实现城市群聚类,采用随机森林算法进行预测。通过可视化的方式,用户能够直观的了解近二十年间中国碳排放量的区域特点和来源组成关系,发现数据中蕴含的信息,为碳核算的研究提供了思路。

关键词: 碳核算; 聚类; 预测

Design and Implementation of China Carbon Accounting Visual Analysis System

Abstract: Faced with the increasingly severe issue of greenhouse gas emissions, countries around the world hope to effectively manage greenhouse gases. At present, most research both domestically and internationally focuses on macro and micro perspectives, lacking analysis from the perspective of meso cities. This system relies on county-level carbon emissions, vegetation carbon sequestration, and sub sector accounting lists provided in the China Carbon Accounting Database to establish a visual analysis system. The aim is to study the spatiotemporal distribution and source composition of national carbon emissions and vegetation carbon sequestration, as well as the characteristics and trends of carbon emissions and vegetation carbon sequestration in various cities. The system allows users to quickly understand the relevant knowledge of carbon accounting through the overview page. In order to fully utilize data information for visual analysis, the page is divided into provincial analysis and municipal analysis pages. The provincial analysis module enhances users' understanding of data and China's energy structure from both spatial and temporal dimensions. Analyzing the changes in carbon emissions has certain reference significance for understanding the implementation effect of China's carbon emission policies and policies. By analyzing at the city level and combining relevant data, observing the changes in urban agglomerations, summarizing the characteristics of urban agglomerations or individual cities, comparing urban districts and counties, and understanding the characteristics of districts and counties, it is beneficial for coordinated development. The system adopts a design of front and rear end separation, using D3 and Echarts for visual display, and the backend adopts. The urban agglomeration clustering is realized by rewriting K-Means algorithm, and the random forest algorithm is used for prediction. Through visualization, users can intuitively understand the regional characteristics and source composition relationships of China's carbon emissions over the past two decades, discover the information contained in the data, and provide ideas for research on carbon accounting.

Key words: carbon accounting; clustering; forecast

目 录

第 1 章 绪论	1
1.1 选题背景与研究意义	1
1.2 国内外研究现状	1
1.3 系统目标及内容	2
1.3.1 系统目标	2
1.3.2 系统内容	2
1.4 本文的主要工作	2
1.5 本章小结	3
第 2 章 中国碳核算可视分析系统需求分析	4
2.1 功能需求	4
2.1.1 概览模块	4
2.1.2 省级分析模块	5
2.1.3 市级分析模块	5
2.2 非功能需求	6
2.3 数据需求	6
2.4 可行性分析	9
2.5 本章小结	9
第 3 章 中国碳核算可视分析系统概要设计	10
3.1 系统结构设计	10
3.2 数据处理	10
3.3 数据存储设计	11
3.4 可视化布局设计	15
3.5 可视化编码	17
3.5.1 省级分析模块可视化编码	17
3.5.2 市级分析模块可视化编码	17
3.6 可视化交互方案设计	18
3.6.1 省级分析模块交互方案设计	18

3.6.2 市级分析交互方案设计·····	18
3.7 本章小结·····	18
第4章 中国碳核算可视分析系统详细设计与实现·····	19
4.1 系统功能详细设计与实现·····	19
4.1.1 概览模块详细设计与实现·····	19
4.1.2 省级分析模块详细设计与实现·····	19
4.1.3 市级分析模块详细设计与实现·····	21
4.2 聚类与预测算法设计与实现·····	23
4.2.1 聚类算法的设计与实现·····	23
4.2.2 预测算法的设计与实现·····	26
4.3 系统界面详细设计·····	28
4.3.1 概览模块界面·····	28
4.3.2 省级分析模块界面·····	29
4.3.3 市级分析模块界面·····	30
4.4 本章小结·····	31
第5章 中国碳核算可视分析系统测试与评估·····	32
5.1 系统功能测试·····	32
5.1.1 概览模块测试·····	32
5.1.2 省级分析模块测试·····	33
5.1.3 市级分析模块测试·····	35
5.2 系统性能测试·····	36
5.3 系统评估·····	38
5.3.1 省级分析模块评估·····	38
5.3.2 市级分析模块评估·····	41
5.4 本章小结·····	44
结论·····	45
致谢·····	46
参考文献·····	47

第 1 章 绪论

1.1 选题背景与研究意义

近年来,各国更加注重可持续发展,面对日益严重的温室效应,及时有效的对温室气体进行治理已经成为国际社会的共识。2015 年签订的《巴黎协定》在全球产生了巨大的影响,提出了“碳中和”的概念,已有一百多个国家做出碳中和承诺,极大的提振了气候治理的信心。2020 年,习总书记在联合国大会上承诺将力争二氧化碳排放于 2030 年前达到峰值,2060 年前实现碳中和。近年来,我国政府一直主张绿色低碳的发展理念,中国需要在探索中成长。

目前国内外的研究也存在一定的问题,如尺度关注不均衡,大部分是对宏观和微观角度进行研究,缺少了对中观的城市角度的研究。然而,在全球化背景下,中观的城市不可能独立的承担碳排放链,对城市角度的碳排放就显得难以界定。

在此背景下,本系统将借助中国碳核算数据库进行碳核算中观层次的研究,讨论城市之间的碳排放变化关系。将可视化技术引入到碳核算的分析中,通过经纬度聚类方法对中观城市层面进行分析,同时根据碳排放数据,对各个城市进行碳排放和植被固碳量的预测。通过可视化的相关技术,用户在中观上加深对碳核算的认识,更直观了解碳排放和植被固碳量的详细信息,为碳核算相关研究提供参考。

1.2 国内外研究现状

有关碳核算的研究在 1998 年政府间气候变化专门委员会(IPCC)成立后取得了一系列的成果。在数十年的研究中,通过数十份报告和指南,大致建立起各领域的碳核算估算框架。目前国内外的碳核算研究大体分为三类,分别从宏观、微观和中观角度入手。

宏观角度,主要以 IPCC 为主导,倡导和组织各国政府碳排放量核算与报告工作。Rogel^[1]等人,评估了各国提交的国家自主贡献计划(INDC)并与现行的政策进行了对比,并对 2100 年全球变暖的预计气温进行预测,最终得到各国需超额交付当前的 INDC,才能将全球变暖控制在 2 摄氏度以下。Chen^[2]等人通过研究第 26 届联合国气候变化缔约方会议的成果目标,提出了建立碳排放模型、推广脱碳技术、讨论碳交易和碳税等一系列有助于实现碳中和的经济战略。这些举措都有利于推进巴黎协定的落实。

微观上,主要在某一领域研究低碳技术。位楠楠^[3]采用碳同位素的技术来分析城市碳排放的来源,相比于现有的基于污染源台方清单统计估算的数据,碳同位素技术能够

有效减少大气二氧化碳来源误差。

而中观角度研究相对较少,研究主要与城市结合,包括城市碳循环、城市碳管理等领域。Zhou^[4]基于 IPCC 区域排放核算方法,对粤港澳大湾区(GBA)城市和周边进行了碳排放量的计算。该研究对 GBA 城市碳排放量增长速度,能源部门,城市类型等都进行了论述,为 GBA 地区制定低碳政策,监测减缓气候变化进度提供了参考。针对中观领域的不足,需要继续深入研究。

1.3 系统目标及内容

1.3.1 系统目标

本系统通过全国和城市碳排放和植被固碳量数据来建立可视分析,旨在研究全国碳排放碳吸收时间和空间分布状况,以及各个城市的碳排放来源和趋势。通过对城市进行归类,得出各个城市群的特点以及未来发展趋势,对城市未来碳排放和植被固碳量进行预测。

1.3.2 系统内容

将全国和部分城市的数据进行处理后,得到研究内容主要有以下四个方面:

- (1) 通过碳排放数据使用户了解全国碳核算的时间和空间分布特点。
- (2) 通过聚类确定不同的城市群及其特征,有助于提出针对性的发展战略。
- (3) 将城市碳排放和植被固碳量进行对比,了解城市特点,有利于补齐短板。
- (4) 通过预测算法对城市未来碳排放量进行预测,有利于及时制定积极的应对措施。

1.4 本文的主要工作

本文从中观尺度的碳核算入手,利用中国碳核算数据库的数据,研究省市两个维度碳排放和植被固碳量信息。系统采用前后端分离的模式,前端使用 Vue,后端采用 Springboot,数据库采用 Mysql,模型使用 Python。在省级分析模块中,将地图和折线图结合来展现碳排放的时空信息,通过河流图和饼图来展现部门类别信息。在市级分析模块中,通过 K-Means 进行聚类并使用随机森林算法进行碳排放和植被固碳量的预测。对于聚类的城市群进行总体分析和对单个城市进行分析,得出其经济发展和碳排放趋势。对每个具体城市,能够通过桑基图和折线图了解各区县碳排放和植被固碳量具体的占比

和趋势, 预测数据对区县的碳达峰的时间有帮助作用。

本文在组织结构方面, 第一章为绪论, 阐述系统的选题背景与研究意义、国内外研究现状、系统目标、本文的主要工作等内容。

第二章为中国碳核算可视分析系统的需求分析部分, 从功能需求、数据需求、可行性分析等方面入手, 通过功能结构图、用例图、数据流图等形式将系统需求逐步拆分为具体的模块, 方便后续的设计。

第三章为中国碳核算可视分析系统概要设计, 从系统结构设计、数据处理、数据存储设计、可视化布局设计、可视化编码、可视化交互方案设计等方面入手, 将需求进一步细化。

第四章为中国碳核算可视分析系统详细设计与实现, 包含系统功能详细设计与实现、聚类与预测算法设计与实现、系统界面详细设计与实现等基本分内容。该部分详细阐述了算法的流程、设计结果、关键代码、交互逻辑设计等内容。

第五部分为中国碳核算可视分析系统测试与评估部分, 主要分为功能测试、性能测试、系统案例评估等内容, 通过这三部分对系统进行全面评估。

1.5 本章小结

面对日益严重的温室效应, 各国积极投入对温室气体的治理中, 碳中和、碳达峰、碳核算等概念应运而生。狭义上的碳核算是指碳排放量的核算, 如碳计量、实测法等。广义上与碳有关的分析计算也是碳核算的范畴。而在相关的研究中, 国内外学者更多的是关注宏观和微观角度的碳排放碳循环, 如宏观上评估各国提交的国家自主贡献计划、微观上采用碳同位素的技术来分析城市碳排放的来源, 在中观的城市角度的研究偏少。本文利用中国碳核算数据库中的数据, 进行中观城市角度的可视化展示, 利用聚类算法和预测算法对城市及其区县的进行分析, 对合理政策的制定提供借鉴。

第2章 中国碳核算可视分析系统需求分析

2.1 功能需求

在碳核算研究领域,相关研究往往执着于具体的数据,而忽视了数据整体的特征与规律,这就导致数据的浪费。而可视化技术就可以解决这个问题,将数据的处理和分析后,可视化能够直白的展示出各省市的时空碳排放和植被固碳量数据,并且聚类、预测等算法也可以加入到可视化中,以此来智能的展示城市变迁和碳排放趋势。

因此,将功能需求进行细分,将中国碳核算系统大体分为概览模块、省级分析模块和市级分析模块。在概览模块,需要展示碳核算的意义、相关资讯、数据和地图来源等信息,使用户可以大体了解碳核算的意义。在省级分析模块,需要分析各个省的时间和空间的碳排放关系,同时,还要能够看到某个省的具体信息。最后,还要对各个行业部门的碳排放信息进行展示。在市级分析模块,要展示某市的区县的详细情况,并进行预测,得到未来的碳排放和植被固碳量趋势,此外,将城市聚类为城市群,分析各个城市群特点以及城市变迁规律。如图 2-1 所示。

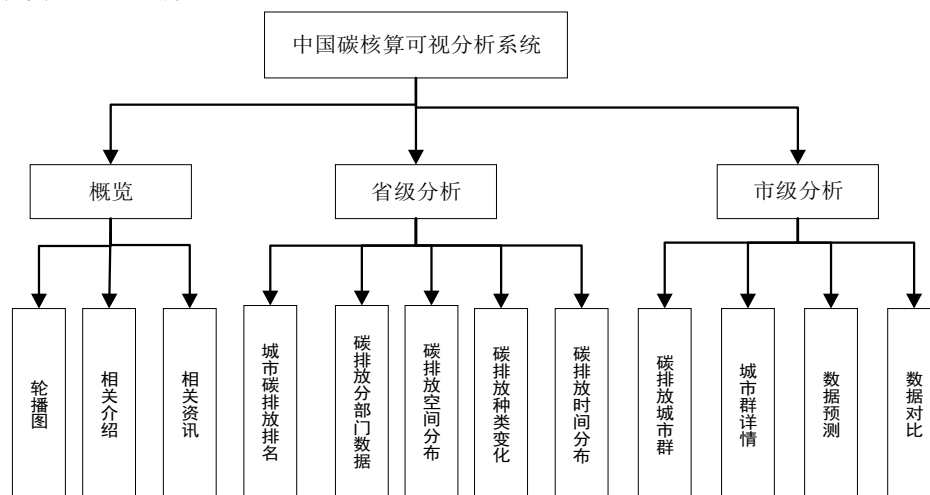


图 2-1 系统功能结构图

2.1.1 概览模块

在概览页面中,包括轮播图模块、相关介绍模块、相关资讯模块。轮播图模块用图片的形式生动形象的展示碳核算的意义。相关介绍模块简洁的介绍本系统的意义、目的等信息。相关资讯模块有一些碳核算相关介绍的链接。通过概览模块能让用户直观的了解相关的研究工作以及政策信息。该模块用例如图 2-2 所示。

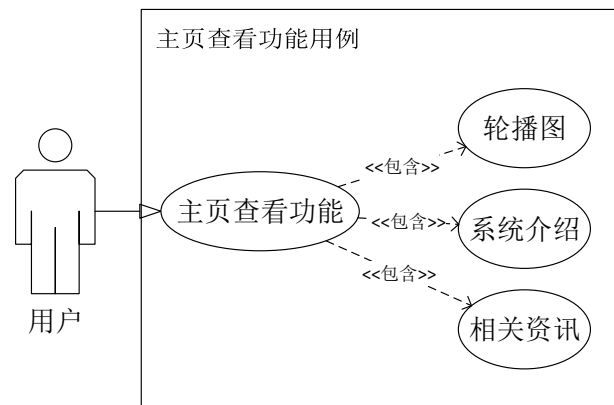


图 2-2 主页查看功能用例图

2.1.2 省级分析模块

省级分析模块细分为城市碳排放排名模块、碳排放分部门数据模块、碳排放空间分布模块、碳排放种类变化模块、碳排放时间分布模块。用户在该模块需要分析各个省份碳排放量的时间和空间分布特征以及碳排放的行业部门，用户不光需要全国的总体概览信息，也存在查看某个省份的需求。行业部门包含查看大类部门和部门详情两部分。该模块用例如图 2-3 所示。

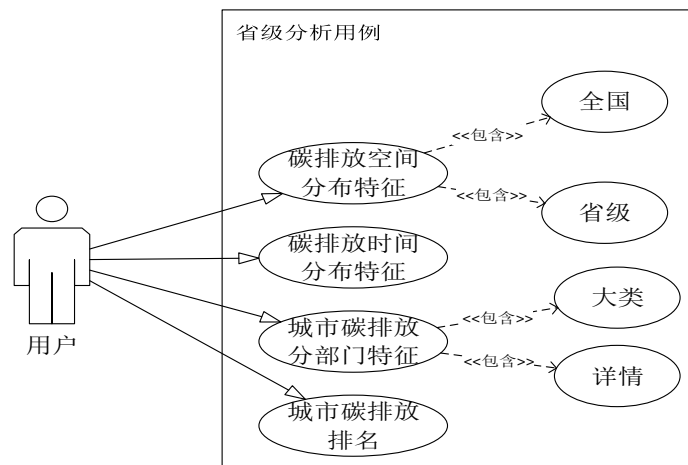


图 2-3 省级分析用例图

2.1.3 市级分析模块

市级分析模块细分为碳排放城市群模块、城市详情模块、数据预测模块、数据对比模块。用户在该模块需要进行城市区县的对比以及查看聚类的城市群信息。每个用例都需要筛选范围和年份。该模块用例如图 2-4 所示。

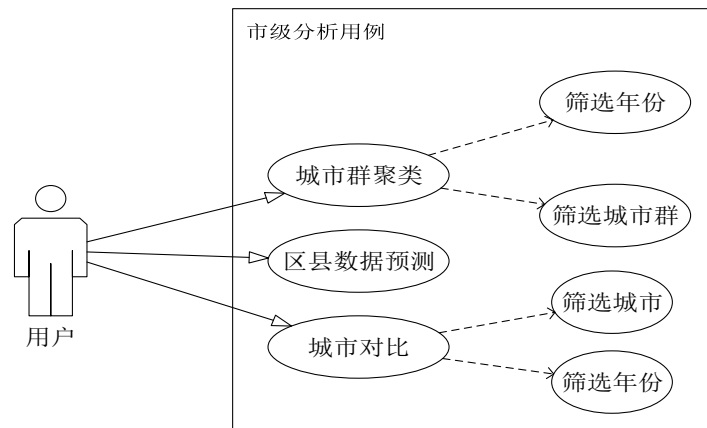


图 2-4 市级分析用例

2.2 非功能需求

虽然功能性需求是人们普遍关注的,但也不能忽略非功能性需求的分析。系统的非功能需求主要从以下四个方面进行分析,分别是可靠性、真实性、易用性、维护性。

首先,在可靠性上,要求系统能够承受较大的并发量,能够流畅运行,基本无卡顿现象,并且要求系统稳定。在可靠性上需要对系统进行功能和性能测试来检验。

其次,真实性既要求数据真实可靠,又要求经过一系列处理后最终界面呈现的结果真实。系统初始数据来源在中国碳核算数据库中,数据来源真实有效。在数据预处理中对缺失值、重复值、异常值都要进行一系列处理。在后续的处理中也应当确定数据计算无误。

另外,在易用性方面,对于刚接触系统的用户,希望系统学习成本小,易上手。因此需要编写系统相关的介绍页面和案例的分析。通过浏览系统介绍,用户可以快速了解碳核算的相关工作,通过案例分析部分,能了解系统的使用方法。

最后,在维护性方面,系统采用主流的 Vue 和 Springboot 的前后端分离的方式来编写。无论是前端的组件式开发还是后端的分层开发,都便于维护。

2.3 数据需求

根据功能需求,要求有如下数据,首先有省市县三级的碳排放和植被固碳量各年的数据。其次,需要相关的省市地图的数据以及城市聚类数据。最后,需要有各个行业部门的碳排放数据。而中国碳核算数据库中提供的“全国区县碳排放数据”、“全国区县植被固碳量数据”、“中国分部门核算碳排放清单”、“中国能源清单以及省级碳排放

清单”等数据可以满足数据需求。

根据系统的功能需求，设计出系统的零层数据流图如图 2-5 所示，零层主要为用户与三个某块进行数据的传输。在系统零层数据流图中，用户作为外部实体，向可视化系统传入参数。浏览概览模块、查看省级模块、查看市级模块三个模块接收输入并进行加工，最后将输出通过更新图表展现出来。而数据存储部分，系统介绍信息和相关资讯信息为固定数据存放在前端，不参与图表数据的处理和绘制。其余数据根据实际需求，以 csv、json 格式存储在前端或数据库中。

由于零层数据流图主要是概览作用，现对其进行细化，设计一层数据流图。系统的一层数据流图中，因为概览模块数据和变化较少且为固定数据，省级分析模块与市级分析模块又相似，所以选择图 2-6 省级分析模块来进行展示。

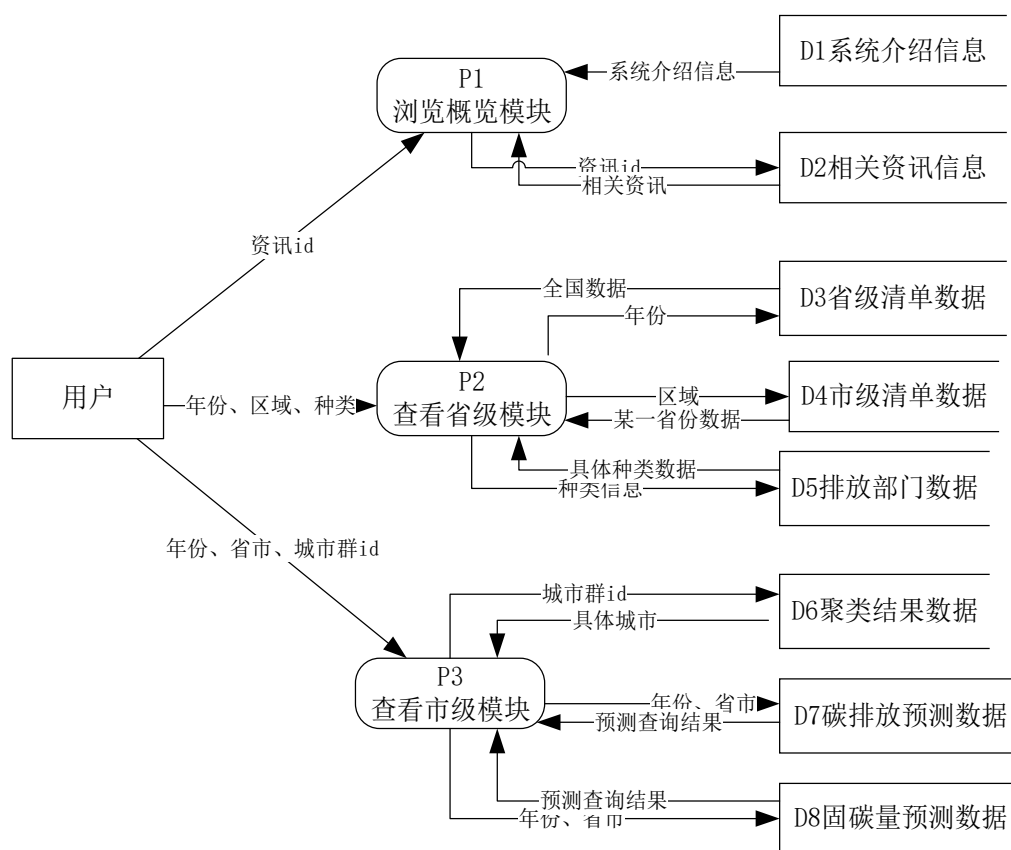


图 2-5 系统零层数据流图

在一层省级模块数据流图中，用户可分别对年份、区域、碳排放种类信息进行选择。之后系统会根据输入信息，进行对应的处理，在到数据存储位置查询符合条件的数据，最后将对应的数据处理成图表所需要的格式并更新绘制具体图表。对于该部分数据的详细信息，数据字典如表 2-1，2-2 所示。

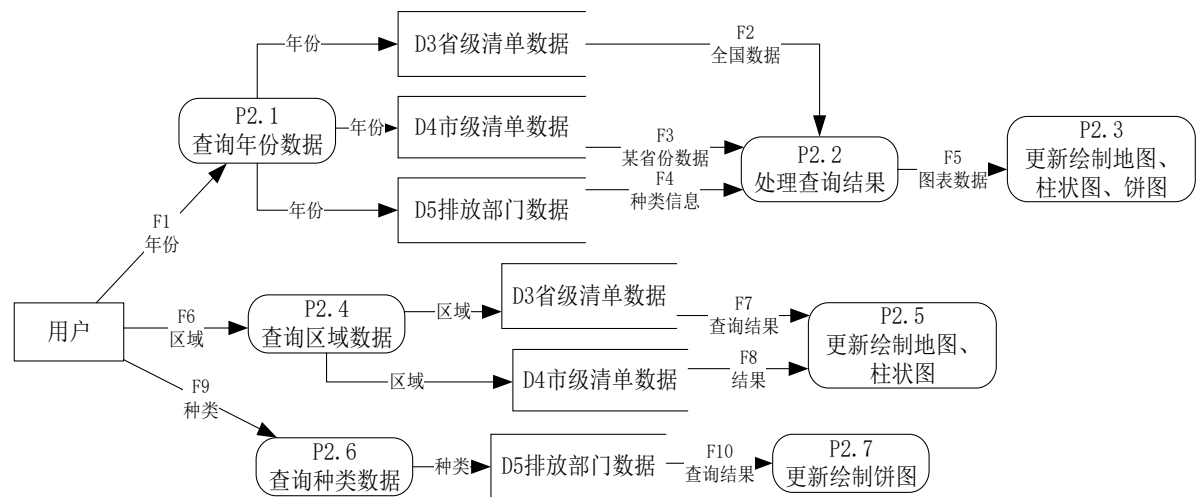


图 2-6 系统一层查看省级模块数据流程图

表 2-1 数据存储字典

编号	名称	输入数据流	输出数据流	内容	说明
D3	省级清单数据	F1,F6	F2,F7	省名+年份+石油类+煤炭类+可燃气体类+其他类+总数	30 个省的碳排放清单
D4	市级清单数据	F1,F6	F3,F8	年份+省名+城市名+总数	城市的碳排放清单
D5	排放部门数据	F1,F9	F4,F10	年份+18 种小类数据	18 种部门的碳排放数据

表 2-2 数据处理字典

编号	名称	来源	处理逻辑概括	输出数据流	说明
P2.1	查询年份数据	F1	用户选择年份，到文件中查询	F2,F3,F4	数据存储在前端文件中
P2.2	处理查询结果	F2,F3,F4	将查询结果处理	F5	前端的 JS 处理函数
P2.3	更新绘制地图、柱状图、饼图	F5	将数据更新图中重新绘制	—	图使用 D3 绘制
P2.4	查询区域数据	F6	根据区域在文件中查询	F7,F8	数据存储在前端文件中
P2.5	更新绘制地图、柱	F7,F8	将数据更新图中重	—	图使用 D3 绘制

	状图		新绘制		
P2.6	查询种类数据	F9	根据碳排放种类在 文件中查询	F10	数据存储在前 端文件中
P2.7	更新绘制饼图	F10	将数据更新图中重 新绘制	—	图使用 D3 绘制

2.4 可行性分析

技术可行性：系统通过前后端分离的形式来进行设计，技术成熟，代码便于管理更新，实现低耦合高内聚，使得层次设计更加合理。前端在 Vue 框架下，使用 Html、Css、JavaScript 能够在网页中实现预期效果，Echarts 和 D3 作为主流的可视化工具，能够满足可视化的需求。后端借助 Springboot，既可以实现对 MySQL 数据库的操作，又可以将处理好的数据发送到前端。对于回归预测算法和聚类算法，目前也有许多优秀的论文可以参考，对 Python 的 Sklearn 库中的聚类和预测模型进行修改，也可以达到预期效果。综上，本系统在技术上有可行性。

经济可行性：本系统的开发软件如 VS Code 等工具都为开源免费软件，并且个人电脑就可以满足要求，不需要有其他的费用。另外，数据采用开放数据库中的数据，也不需要支付额外的费用。因此，该系统在经济上完全可行。

2.5 本章小结

本章对系统需求进行分析，系统要将全国省市县的碳排放、植被固碳量和各行业部门的碳排放数据进行可视化分析展示。在此需求上，细分为省市两个大的模块，每个模块又细分为小的模块。每个小模块都明确了数据的流向，输入输出、具体功能职责等。在数据需求方面，从原始数据进行分析和预处理，根据模块功能设计出合理的数据结构。最后，从技术和经济方面分析可行性，技术上采用了成熟的技术和算法，经济上采用开源软件且不需要过高的硬件，系统有可行性。

第3章 中国碳核算可视分析系统概要设计

3.1 系统结构设计

中国碳核算可视分析系统根据实际需要，主要采用 Vue、Springboot 和 mysql 的前后端分离模式。在系统架构方面，为了实现低耦合，决定采用 MVC 的软件架构模式。该模式将程序划分为模型层（Model）、视图层（View）和控制层（Controller）三层。这样的分层结构对程序的复杂度进行了简化，使得程序更加直观，也赋予了各个部分对应的功能，增强了代码的逻辑性和可维护性。项目开发中使用 SpringBoot 作为 MVC 思想的体现，系统将模型层分为数据访问层（Dao）和业务逻辑层（Service），负责处理数据的逻辑和业务规则，业务逻辑层通过 Dao 接口来调用数据访问层的模块。数据访问层可以对数据库进行存取等操作。视图层是浏览器，用户能够看到的交互页面。控制层能够接受用户从浏览器中的输入，并通过 Service 接口去调用模型层的模块。如图 3-1 所示。

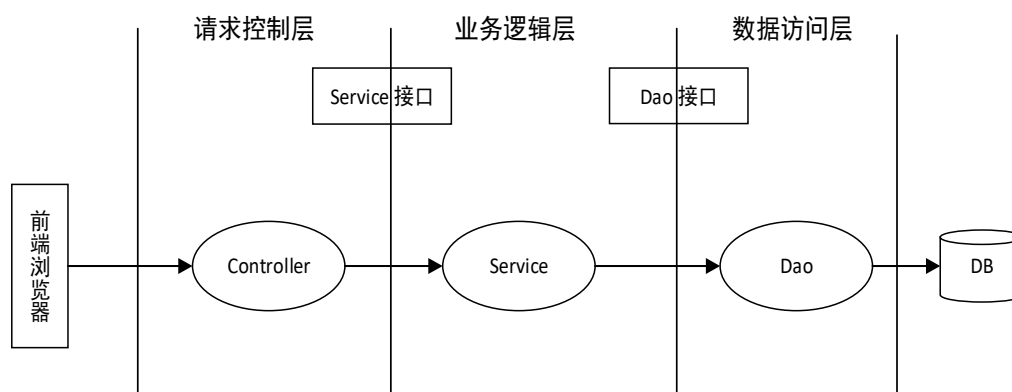


图 3-1 系统分层关系

3.2 数据处理

首先，在已经获得的碳排放量数据中，各行业部门的数据每个年份为一个表格，表格内是各省的具体行业类别的碳排放量。这种多个表格的形式不利于信息的读取，因为信息读取时只能一行一行读入，而转换成 json 格式的数据就可以快速读出，非常适合高维度的数据。在省市县三级的碳排放和植被固碳量的各年数据中，因为前端要进行一些城市群类型的计算、绘制地图、展现各个年份碳排放的信息，所以将省市区县和年份作为主键，省去格式转换和按年份查询不同表的时间。

其次，对于地图数据，除了直接从阿里的 DataV 网站中下载中国地图数据，还要有

各省的数据以及全部市的数据。最终使用爬虫，对于省级地图，将省名称列表遍历爬取并保存。而对于全部市的数据，对各个省的数据进行提取，把省下面的所有市的地图数据保存起来，最后将全国地图下面的数据替换，就得到了可以绘制城市群聚类的地图数据。对于城市群聚类数据，因为是根据不同年份聚类，所以会采用省、市、年份和城市群聚类类别四个维度属性。

最后，对于各行业部门的数据，因为河流图总是展示全国各年大类数据，所以可以将各行业详细数据进行求和计算作为一个表格，不用每次初始化时都要重新计算。对于各行业的原始数据，由于过于混乱，需要进行清洗和整理得到正确的格式。

3.3 数据存储设计

数据根据实际需要，在进行完数据处理后，分别以 json 和 csv 格式进行存储。对于数据量较大且要进行查询等操作的数据，存放到数据库中，后端进行查询操作后将查询结果返回给前端。而对于较大的地图数据，为了更快的加载地图，则直接存放到前端。

对需求分析部分内容进行展开，数据存储设计包括概念设计、逻辑设计和物理设计三部分。

在概念设计部分，因为图表和属性太多，一个 E-R 图太复杂，所以绘制了图 3-2 系统总体 E-R 图和图 3-3 系统详细 E-R 图。在图 3-2 中，用户虽然没有专门的表，但作为系统的使用者，也加入其中。查询关系都是一对多的关系，表中的主键用下划线标出。

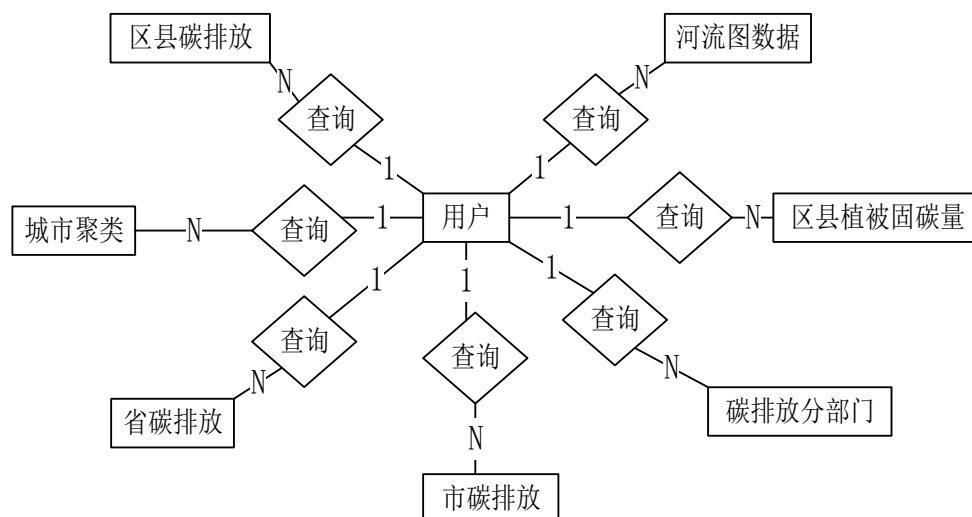


图 3-2 系统总体 E-R 图

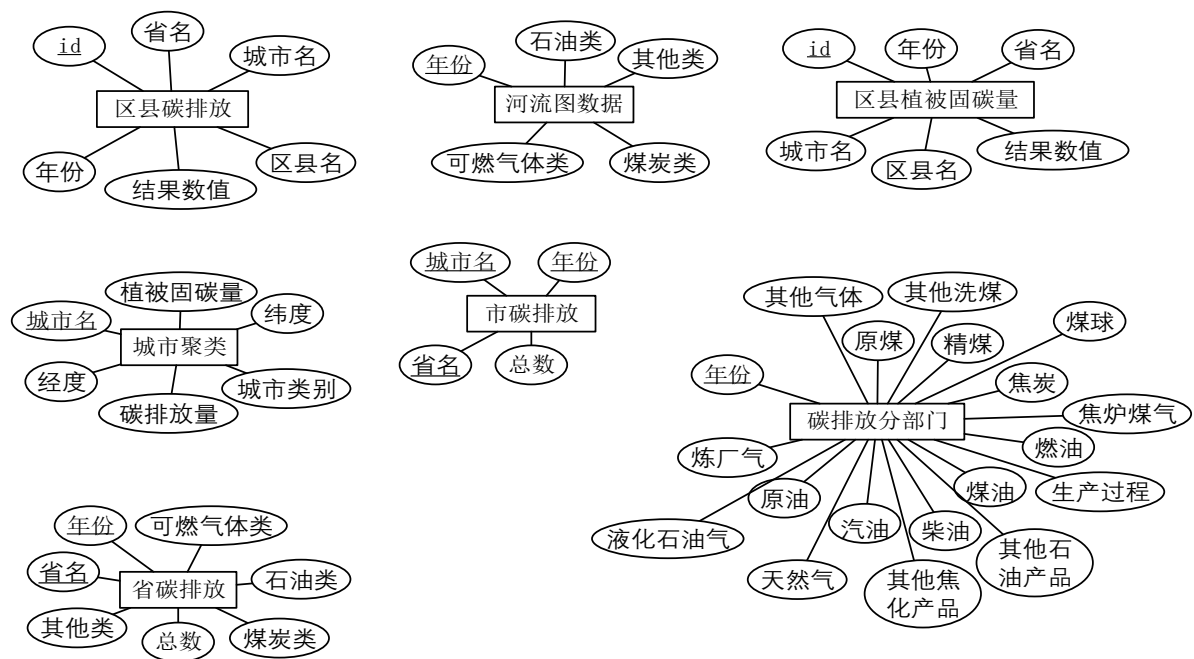


图 3-3 系统详细 E-R 图

在逻辑设计部分，在系统设计和 E-R 图的基础上，进行数据表的设计。分别将 E-R 图中的区县碳排放、区县植被固碳量、河流图数据、城市聚类、省碳排放、市碳排放、碳排放分部门转换为 carbon、plant、river、kmeans、data、city_data、detail 七个关系模型，因为用户没有单独建表，所以图中的一对多关系就没有体现。

data	
PK year	varchar
PK province	varchar
Raw coal total	float
Crude oil total	float
Natural gas total	float
Cement	float
total	float

carbon	
PK id	int
province	varchar
city	varchar
county	varchar
year	varchar
res	float

plant	
PK id	int
province	varchar
city	varchar
county	varchar
year	varchar
res	float

detail	
PK year	varchar
Raw_Coal	float
Cleaned_Coal	float
Other_Washed_coal	float
Briquettes	float
Coke	float
Coke_Oven_Gas	float
Other_Gas	float
LPG	float
Refinery_Gas	float
Natural_Gas	float
Crude_Oil	float
Gasoline	float
Kerosene	float
Diesel_Oil	float
Fuel_Oil	float
Other_Coking_Products	float
Process	float
Other_Petroleum_Products	float

city_data	
PK province	varchar
PK year	varchar
PK city	varchar
total	float

kmeans	
PK City	varchar
lon	float
lat	float
carbon	float
plant	float
class	int

river	
PK year	varchar
coal	float
gas	float
oil	float
other	float

图 3-4 数据表设计图

在物理设计部分，主要对后端数据库和前端数据文件的字段进行说明。

在后端数据库中，主要有区县碳排放量预测表和区县植被固碳量预测表两个表。两个表中的字段名都相同，所以只用一个表 3-1 来表示。每个表共有 70000 多条数据，包含 30 个省级行政单位的 2735 个区县。数据包含 2000 年到 2025 年的碳排放量或植被固碳量数据，年份在 2018 年以前为真实数据，2018 年及以后的为预测数据。后端查询时根据省市和年份来查询。

表 3-1 区县碳排放和植被固碳量预测信息表

字段名	字段类型	小数点	不为空	主键	注释
id	int	0	√	√	id 序号
province	varchar	0	—	—	省名
city	varchar	0	—	—	城市名
county	varchar	0	—	—	区县名
res	float	7	—	—	结果数值
year	varchar	0	—	—	年份

前端数据存储在文件夹中，既有 csv 文件，也有 json 文件。csv 文件有河流图数据、分部门数据和聚类数据，其中每年的聚类数据保存为一个带年份信息的 csv 文件，格式如表 3-2、表 3-3 和表 3-4 所示。json 格式文件为省碳排放清单和市碳排放清单，分别如表 3-5 和表 3-6 所示。

表 3-2 河流图数据信息表

字段名	字段类型	小数点	不为空	主键	注释
year	varchar	—	√	√	年份
coal	float	2	√	—	煤炭类
gas	float	2	√	—	可燃气体类
oil	float	2	√	—	石油类
other	float	2	√	—	其他类

表 3-3 碳排放分部门数据信息表

字段名	字段类型	小数点	不为空	主键	注释
year	varchar	—	√	√	年份
Raw_Coal	float	3	√	—	原煤

Cleaned_Coal	float	3	√	—	精煤
Other_Washed_Coal	float	3	√	—	其他洗煤
Briquettes	float	3	√	—	煤球
Coke	float	3	√	—	焦炭
Coke_Oven_Gas	float	3	√	—	焦炉煤气
Other_Gas	float	3	√	—	其他气体
LPG	float	3	√	—	液化石油气
Refinery_Gas	float	3	√	—	炼厂气
Natural_Gas	float	3	√	—	天然气
Crude_Oil	float	3	√	—	原油
Gasoline	float	3	√	—	汽油
Kerosene	float	3	√	—	煤油
Diesel_Oil	float	3	√	—	柴油
Fuel_Oil	float	3	√	—	燃油
Other_Coking_Products	float	3	√	—	其他焦化产品
Process	float	3	√	—	生产过程
Other_Petroleum_Products	float	3	√	—	其他石油产品

表 3-4 城市聚类信息表

字段名	字段类型	小数点	不为空	主键	注释
City	varchar	—	√	√	城市名
lon	float	4	√	—	经度
lat	float	4	√	—	纬度
carbon	float	10	√	—	碳排放量
plant	float	10	√	—	植被固碳量
class	int	—	√	—	城市类别

表 3-5 省碳排放清单信息表

层级名	层级数	字段类型	小数点	不为空	注释
year	1	varchar	—	√	年份
province	2	varchar	—	√	省份名称
Raw coal total	3	float	15	√	煤炭类

Crude oil total	3	float	15	√	石油类
Natural gas total	3	float	15	√	其他类
Cement	3	float	15	√	可燃气体类
total	3	float	15	√	总数

表 3-6 市碳排放清单信息表

层级名	层级数	字段类型	小数点	不为空	注释
province	1	varchar	—	√	省份名称
year	2	varchar	—	√	年份
city	3	varchar	—	√	城市名称
total	4	float	15	√	总数

3.4 可视化布局设计

中国碳核算可视分析系统共概览模块、省级分析模块、市级分析模块三个 web 页面，页面采用栅格布局来设置各个视图区域的位置和大小。省级、市级分析两个模块页面右下角有提示按钮可以查看案例，方便读者理解系统的逻辑。系统在页面顶部有头部导航栏，负责各个页面的切换，如图 3-5，图 3-6，图 3-7 中的 A，在各个模块中不再赘述。

在概览模块中，如图 3-5，B 为轮播图，C 为系统介绍，D 为相关资讯模块。轮播图模块因为页面宽度方向空间空余，但高度方向空间匮乏，所以使用卡片化走马灯。系统介绍模块，采用手风琴效果的折叠面板，每次只能展开一个面板，这样设计既可以使页面布局合理，也可以使内容主次分明。相关资讯模块，采用列表的形式展示碳核算的相关信息，点击即可跳转到对应页面，同时，也标注出系统使用的数据和地图的来源。

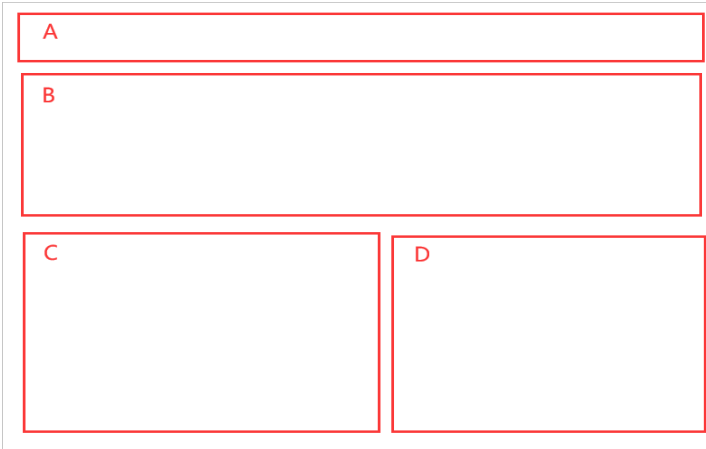


图 3-5 概览页面布局

在省级分析模块,如图 3-6,B 为碳排放城市排名柱状图,会根据年份、地域信息展示不同的城市信息排名,不同的颜色代表不同的碳排放情况,绿色表示良好,红色表示排放过多。C 为分部门平均碳排放饼图,饼图有两种模式,分别是四饼图和单饼图。四饼图模式时分别显示不同年份煤炭类、石油类、可燃气体类、其他类四大类的详细信息,单饼图模式时只展示四大类中的一类,该模式可看作将四饼图中的某个图进行放大。D 为碳排放地图,有全国地图和省级地图两种模式。地图会根据年份展示省份的碳排放信息,颜色代表碳排放量的多少,与碳排放城市排名柱状图的颜色设置相同。E 为碳排放折线图,有全国和省级两张模式,展示 1997 年至 2017 年 20 年间的碳排放趋势。F 为历年碳排放量河流图,河流图将 2000 年到 2015 年碳排放的四大类来源进行展示,可以直观地发现我国历年的碳排放来源的占比与发展趋势。

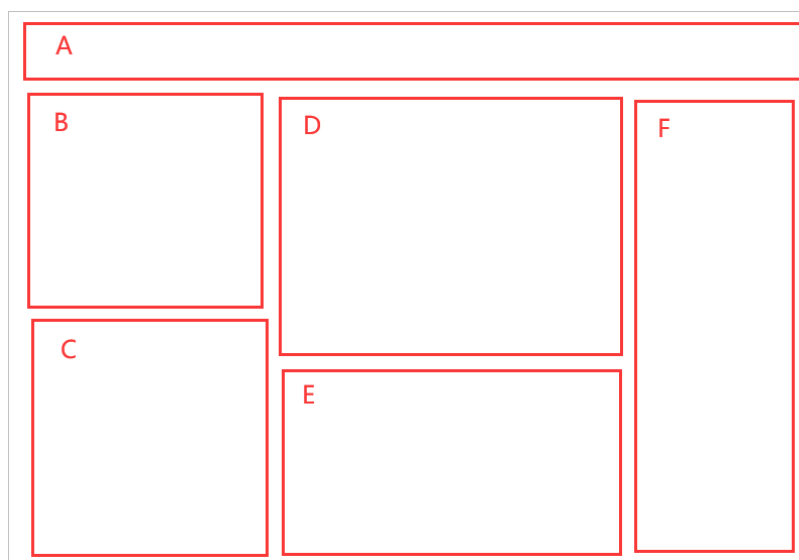


图 3-6 省级分析模块布局

在市级分析模块,如图 3-7,B 为城市群饼图和词云图,二者上下排列。饼图展示该城市群某年植被固碳量的占比,通过饼图能够直观地发现植被固碳量较多或极少的城市。词云图展示该城市群某年碳排放量的情况,词云图中词云的大小对应碳排放量的多少。C 为某年的碳排放和植被固碳量的城市聚类地图,该部分将全国的市聚类为 17 个城市群,下方有历年全国碳排放热力图,负责切换年份。D 为区县碳排放和植被固碳量桑基图,需要选择区县和年份信息,红色桑基图表示碳排放,绿色桑基图表示植被固碳量。E 为区县碳排放和植被固碳量预测折线图,二者横向排列。展示 2000 年到 2025 年的信息,其中 2018 年到 2025 年为预测数据。

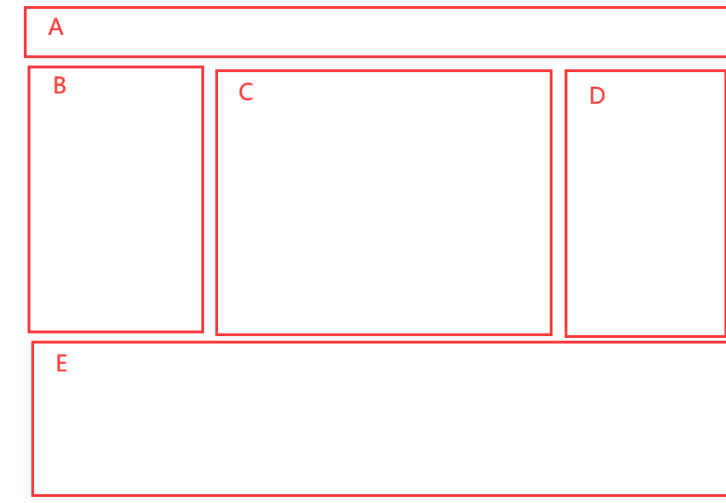


图 3-7 市级分析模块布局

3.5 可视化编码

3.5.1 省级分析模块可视化编码

在省级分析模块，编码采用 D3。在图 3-6 的 B 中，城市碳排放量的大小映射为矩形的长度以及颜色，颜色从 0 到最大值分别对应绿色和红色。C 图在每个大部门类别中，碳排放量映射为饼图的角度。D 图中，全国碳排放量时，碳排放数值 0 到 1500 映射到绿色至红色的范围区间。某一省份碳排放量时，碳排放量数值变为 0 到 100。E 图中，每一年的碳排放量数据映射为一个点，为了便于观察点之间的关系，用线将点进行连接。F 图中，对于四大类碳排放源，采用堆叠的河流图进行展示，每一年的数据映射为河流的宽度。

3.5.2 市级分析模块可视化编码

在市级分析模块，编码采用 Echarts 实现。在图 3-7 的 B 中，将聚类结果按照城市群存放到不同的数组中，根据当前的城市群序号选择对应的数据并将数据处理成对应的数据形式。将碳排放量映射成词云的大小，植被固碳量映射成饼图的角度。在 C 图中，地图中根据聚类的 17 个类别，映射成对应的颜色。地图下方的热力图，则把全国碳排放量从 0 到最大值映射成白色到红色的范围。D 图中两个桑基图分别为碳排放量图和植被固碳量图。该城市的总量映射为桑基图的起点，而下辖各区县的碳排放量或植被固碳量映射为终点。E 图中有两个折线图，分别为碳排放量图和植被固碳量图，每个区县映射为一条折线。

3.6 可视化交互方案设计

中国碳核算可视分析系统交互方案包括鼠标点击、鼠标滚动、鼠标悬停、筛选等操作,通过这些鼠标事件可以实现高亮显示、多图联动等效果。

3.6.1 省级分析模块交互方案设计

图 3-6 省级分析模块布局图中, B、C、D、E 图鼠标悬停, 都可以看到碳排放量数据。B 图为横向排序柱状图, 鼠标滚动可以查看隐藏的城市。C 图鼠标悬停饼图会高亮显示。D 碳排放量地图和 E 折线图作为主操作视图, 分别显示碳排放的空间和时间信息。地图中, 点击某一个省份, 地图下钻, 由原来的全国地图变为某一省份地图, 同时显现该省份碳排放信息, B 图, E 图会更新到该省份的信息。折线图中, 点击圆点, B、C、D 图会更新到对应年份的数据。F 作为辅助视图, 点击每一大类, C 图中的饼图会变为该大类的单饼图模式。点击纵轴的年份, B、C、D、E 图都会更新显示该年份的数据。

3.6.2 市级分析交互方案设计

图 3-7 市级分析模块布局图中, B、C、D、E 图鼠标悬停, 都可以看到对应城市或区县的碳排放量或植被固碳量信息。B 图、C 图都显示城市群相关数据。B 图中通过选择器来选择某一个城市群, 词云和饼图更新为该城市群数据。C 图在地图下方还有热力图, 点击热力图中的年份, B 图、C 图就会更新为该年的城市群聚类数据。D 图、E 图显示城市下面所有区县的数据。通过 D 中的选择框选择城市和年份, D 图、E 图会更新为该城市对应年份的数据。

3.7 本章小结

中国碳核算可视分析系统采用 Vue、Springboot、Mysql 的前后端分离的形式, 需要查询且数据量较大的数据存储于数据库中, 后端进行查询和处理后发送给前端, 而数据量较小或者地图数据存储在前端。根据上一章的需求分析, 本章进行了合理的可视化布局、编码和交互方案设计。根据第二章的需求, 每个模块和视图都从数据处理到与可视化映射编码做了全流程分析, 设计出具有可行性的可视化系统。

第4章 中国碳核算可视分析系统详细设计与实现

4.1 系统功能详细设计与实现

4.1.1 概览模块详细设计与实现

在进行该模块的设计前,先完成顶部栏的设计,顶部栏除了系统图标和系统名以外,还有导航菜单负责页面的切换。导航菜单采用 Element UI 中的 NavMenu 导航菜单,设置为水平模式,每个 el-menu-item 上绑定好页面的路由,点击后回跳转到对应页面。而“系统主页”、“省级分析”、“市级分析”的路由都为主页的孩子路由,所以点击后顶部导航栏不变,下方内容改变。

而系统主页部分,分别使用 el-carousel、el-collapse、el-table 来实现走马灯、折叠面板、信息列表的效果。在走马灯模块,设计了 setSize()来动态的改变图片和容器的大小,初始时调用该函数,当窗口大小发生改变时,再调用一次该函数。在相关资讯模块,绑定了 rowHandleclick()点击事件,点击后打开该条数据中的网址。而最下方的地图和数据来源,则直接采用 href 链接的方式来完成。

4.1.2 省级分析模块详细设计与实现

省级分析模块采用栅格布局,整体分为左中右三部分,其中左、中又都分为上下两部分,通过调整各部分的比例来实现图 3-3 的布局效果。

案例分析界面使用 el-drawer 来实现,有一个状态变量,当变量为真时右侧展示案例信息,当为假时关闭侧边抽屉。点击案例提示按钮会改变抽屉的状态。

可视化部分使用 D3,其中,一种方式为对象式的编码,将饼图和河流图单独写成一个 js 文件,使用时调用类创建一个对象。另一种方式是函数式的编码,如地图、折线图和条形图。在 JS 部分通过一个个的函数来实现对应的内容。两种方式各有优劣。当每个图代码量达到四五百行时,在 js 文件中设计类的方法就会显得很条理。在函数式的编程中,想要调用其他类的方法、数据时都要容易。不需要使用 window 方法来调用相关属性、方法和数据。

系统函数主要分为加载数据、处理数据、绘制图形三大类。加载数据类函数 loadData()和 loadCityData()采用异步函数加载的方法,在不阻塞主线程的同时,又可以暂停执行。处理类函数如 getBarChartData()主要负责数据形式的转换,将加载的数据转换成绘制图表需要的形式。绘制图形类函数如 renderLineChart()负责绘制图形和处理交互事件,其

中各图表的交互除了调用绘制函数外，还使用了回调函数和将图表对象作为函数传参的方法。

在实际的编码中，绘制一大一小两个地图用于解决全国地图下钻到省份地图后如何返回的问题。当展示区域为全国时，大地图显示除南海外全国的地图数据，小地图显示南海地图数据，当展示区域为某一省份时，大地图显示某省份，小地图显示全国地图，点击小地图即可返回显示全国数据。

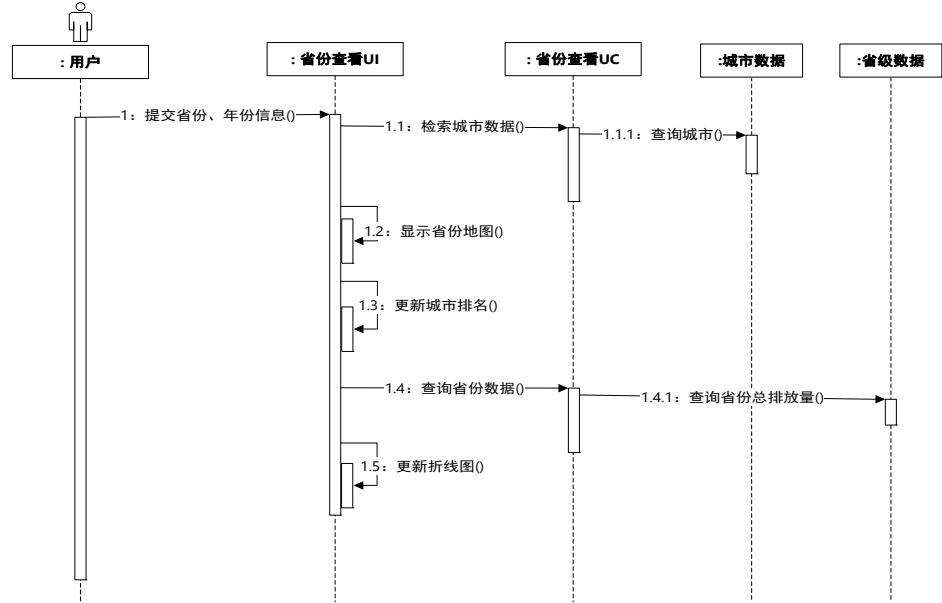


图 4-1 省级分析省份查看 UML 顺序图

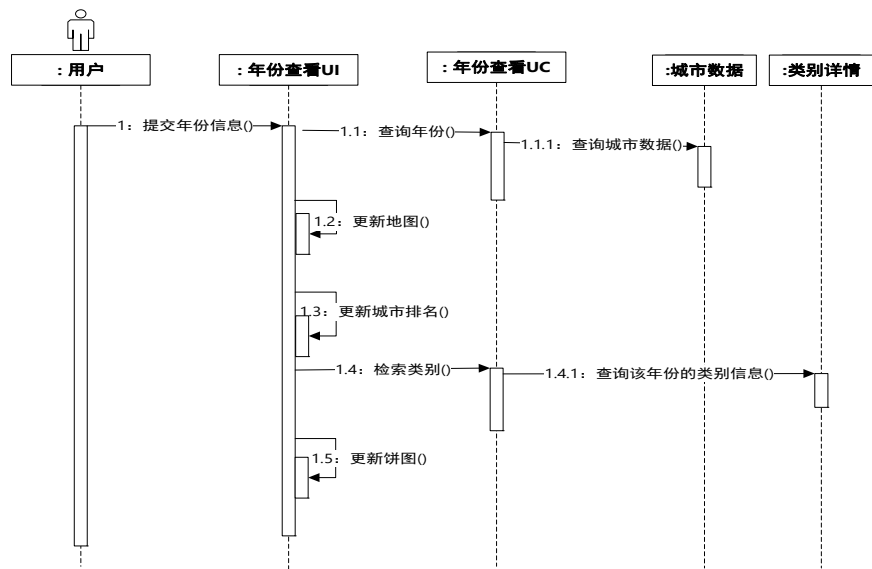


图 4-2 省级分析年份查看 UML 顺序图

最后, 省级分析设计过程如图 4-1, 图 4-2 所示, 分别对查看省份和查看年份的用例设计 UML 顺序图。两者大体过程都是根据用户在 UI 界面中提交的信息进行查询检索, 将查询数据结果返回再进行视图的更新。

4.1.3 市级分析模块详细设计与实现

市级分析模块布局也采用栅格布局, 整体上分为上下两部分, 上面部分又分为左中右三部分, 最终达到图 3-4 的布局效果。

页面图表全部采用 Echarts 完成。案例分析部分代码与省级分析模块的案例相同。为了统一页面风格, 所有选择框、按钮都继续使用 Element UI 中的组件, 大小都设置为“mini”, 并绑定好点击事件。城市群选择框的候选项为 17 个城市群, 地区候选框为引入的“element-china-area-data”包的地区, 点击会出现两级菜单, 第一级为省级, 第二级为市级。年份候选框内为 2000 年到 2017 年。

在数据方面, 页面的全局变量主要有三大类, 第一类是 Echarts 图表变量, 用于绘制图。第二类是数据, 用于存储从后端接收的数据或者图表的配置项数据, 第三类是状态量, 如当前年份、当前城市群序号等。

本模块函数也大体分为三类, 分别是与数据相关、图表设置、交互事件的函数。第一类主要是接收数据和处理数据的功能, 如异步加载函数的 `getData()`, 和处理数据函数 `getLinks()`、`getSangOptionData()`等。第二类主要是 Echarts 图表的初始化、绘制、更新等一系列操作函数, 如 `initLineChart()`、`updateLine1Chart()`等。第三类主要是点击等交互事件, 如 `handleChangeCluCity()`、`handleChangeYear()`、`handleChangeCity()`等。

在市级分析中, 数据来自后端, 后端 UML 类图如图 4-3 所示。其中 `treePredict` 与 `carbonPredict` 为实体类, 即 Springboot 中的 Entity 层。`carPreRepository` 和 `treePreRepository` 是 Dao 层中的两个接口, 两者都继承自 `JpaRepository`。Dao 层中主要对实体层进行增删查改等操作。`PreController` 位于 Controller 层, 根据前端的请求信息执行对应的函数, 该层会把 Dao 层的接口作为私有对象进行调用。

在市级分析页面的设计过程中, 也通过 UML 图对用例进行了设计。年份改变用例如图 4-4 所示。当用户提交了年份时, 系统会根据年份信息找到该年份的聚类数据、碳排放数据和植被固碳量数据。之后将返回数据用来更新视图。

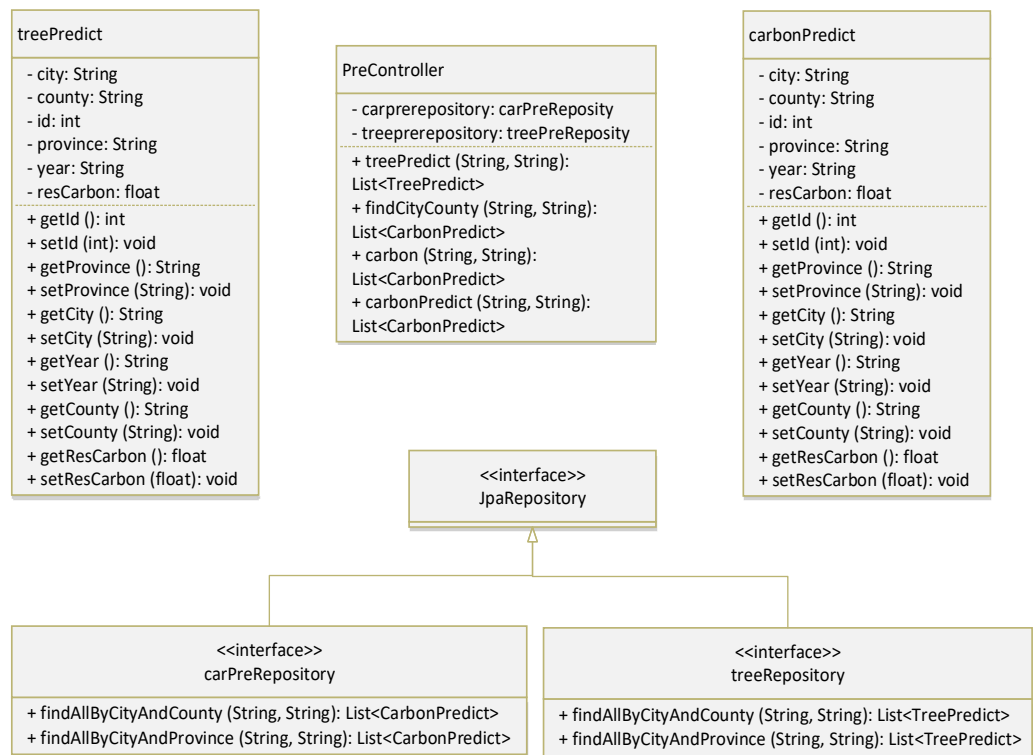


图 4-3 市级分析 UML 类图

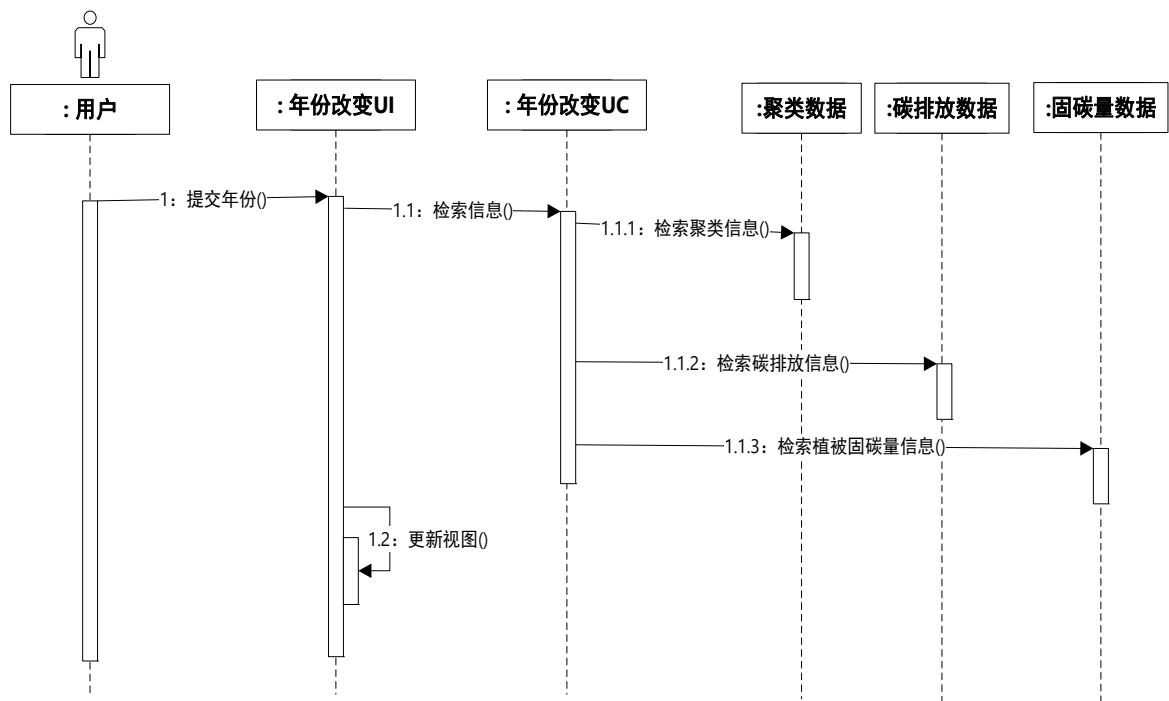


图 4-4 市级分析年份改变 UML 顺序图

4.2 聚类与预测算法设计与实现

4.2.1 聚类算法的设计与实现

市级分析模块中的聚类是探究市级维度上碳排放和植被固碳量的特点。聚类分析流程如图 4-5 所示，将每一年的碳排放量数据和植被固碳量数据进行聚类，通过历年的聚类结果发现城市群变迁规律。

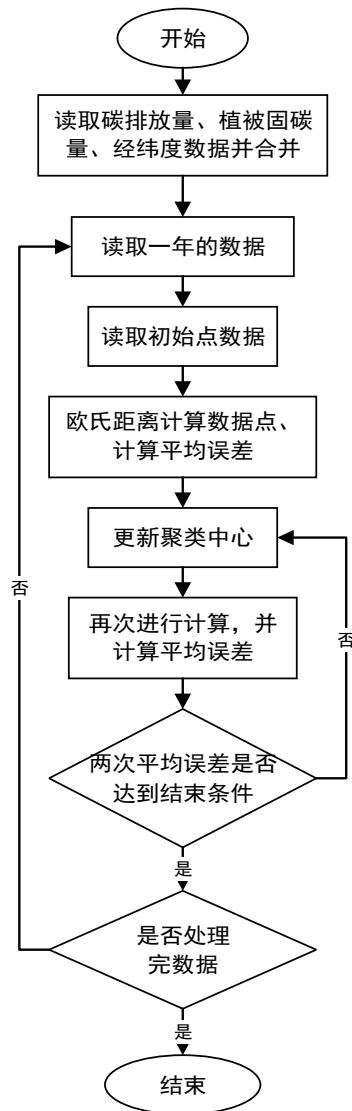


图 4-5 经纬度聚类城市群流程图

但在直接调用 Sklearn 库中的 K-Means 函数时，出现了不同类型城市之间相互混在一起的情况，见图 4-6。分析原因，是输入参数为经度、纬度、碳排放量和植被固碳量，而 K-Means 的距离计算公式会把这四个维度平等处理。同时还会发现，上海、重庆、天津等因为体量较大，则直接占据了一个类别，这不能体现对周围城市的辐射带动作用，

不符合需求，所以需要对算法进行改写，即进行加权处理。另外，通过查阅资料，将城市群的聚类个数确定为 17 个，为了确保聚类结果的可靠性，将每个城市群的中心城市的经纬度以及碳排放数据作为聚类的初始值，而不是采用 K-Means 原本的随机初始值。

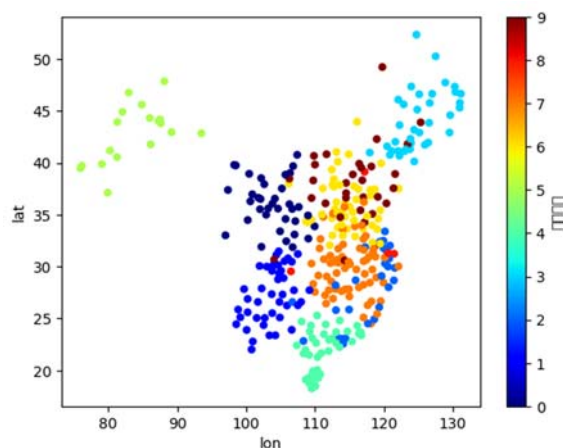


图 4-6 Sklearn 中的 K-Means 聚类结果

K-Means 的主体思想就是先从样本集中随机选取 K 个样本作为簇中心，并计算所有样本与这 K 个“簇中心”的距离，对于每一个样本，将其划分到与其距离最近的“簇中心”所在的簇中，对于新的簇计算各个簇的新的“簇中心”。在设计过程中，需要在代码上进行修改而不是直接调用，首先将 17 个城市群的中心城市的信息作为簇中心进行初始化，之后使用 haversine 公式来计算两地的经纬度距离，在进行一系列映射之后，再跟碳排放量进行加权求和作为某城市与簇中心城市“距离”，将城市划分到“距离”最小的“簇中心”的一类。经过不断地重新计算质心，两次平均误差达到结束条件后，得到最终的效果。将经纬度距离的权重设置大一些，可以确保城市群会连成一片，最终结果如图 4-7 所示。

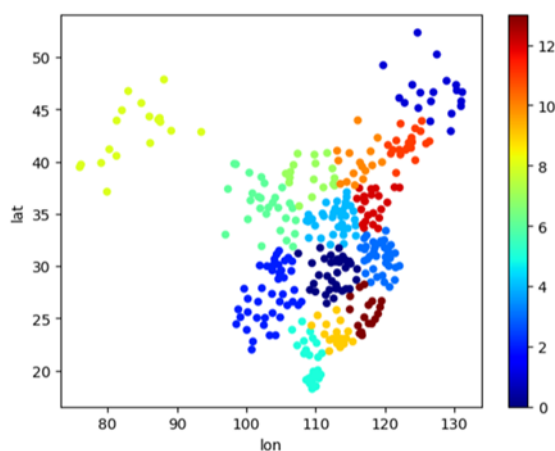


图 4-7 改写后的 K-Means 聚类结果

其中，对于关键问题如何度量城市之间的距离，采用式 4-1 进行计算：

$$D = w_1 \times \text{dist}(c_i \times c_j) + w_2 \times r \quad (\text{式 4-1})$$

其中， $\text{dist}(c_i, c_j)$ 表示两个城市之间的地理距离，最后再进行归一化处理，如式 4-2，使其范围在[0,1]内。 r 为聚类的四个维度，使用皮尔逊相关系数进行计算，但地理位置越近的城市，相关系数越大，城市之间的度量越小，越不相似，这与聚类需求相反。所以需要进行式 4-3 所示的修正以满足要求，其中 $\text{abs}()$ 返回数字的绝对值，经过修正最终范围也在[0,1]。参数 w_1 和 w_2 是权重，需要人工进行设置。经过多次比较后，选取 w_1 为 5， w_2 为 1，此时可以获得较好的聚类效果。

$$\text{dist}(c_i, c_j) = \frac{d_{\max} - \text{dist}(c_i, c_j)}{d_{\max}} \quad (\text{式 4-2})$$

$$r = 1 - \text{abs}(r) \quad (\text{式 4-3})$$

代码 4-1 K-Means 聚类关键代码

```
# 计算两个城市间距离
def haversine(lon1, lat1, lon2, lat2):
    lon1, lat1, lon2, lat2 = map(radians, [lon1, lat1, lon2, lat2])
    # haversine 公式
    dlon = lon2 - lon1
    dlat = lat2 - lat1
    a = sin(dlat/2)**2 + cos(lat1) * cos(lat2) * sin(dlon/2)**2
    c = 2 * asin(sqrt(a))
    r = 6371
    return c * r * 1000
#将距离映射为[0,1]之间
def dist_normalization(distance):
    dmin = 0.0 # 5791.349304739649
    dmax = 4559369.499100928
    return (distance - dmin) / (dmax - dmin)
# 计算不同城市之间的度量距离
def my_Distance(city1, city2):
```



```

geo_distance = haversine(city1.iat[0,0], city1.iat[0,1], city2.iat[0,0], city2.iat[0,1])
geo_distance = dist_normalization(geo_distance)
x_str=[]
y_str = []
x_str.append(city1.iat[0,0])
y_str.append(city2.iat[0,0])
x_str.append(city1.iat[0,1])
y_str.append(city2.iat[0,1])
x_str.append(city1.iat[0,2])
y_str.append(city2.iat[0,2])
x_str.append(city1.iat[0,3])
y_str.append(city2.iat[0,3])
p = fabs(pearsonr(x_str, y_str)[0])
if np.isnan(p) == False:
    r = 1 - p
else:
    r = 0
return w1 * geo_distance + w2 * r

```

4.2.2 预测算法的设计与实现

设计数据预测模块的原因主要有以下两点：一是在省级分析页面中数据截止到 2017 年，近几年的数据缺失不利于分析。二是希望通过预测模块可以对未来区县的碳排放和植被固碳量趋势有一个清晰的认识。本模块通过对线性回归、支持向量回归、岭回归、KNN 回归、随机森林回归等回归预测算法进行分析，确定了最适合本数据集的回归预测算法以及最优的训练集测试集占比。

总体上，模型设计的主要步骤主要有以下几个部分：数据处理、确定模型参数、模型预测、保存。具体详见图 4-8。

数据处理主要有读取数据、数据预处理（处理异常值、缺失值、重复值等）、数据拆分重组、创建城市字典并完成替换。确定模型参数主要包括确定训练集占比和具体模型。在经过训练后得到结果如图 4-9 所示。图纵坐标为不同的模型的准确率，横坐标为

不同的训练集占比。其中，线性回归和岭回归因为原理相似，所以重合为一条折线。

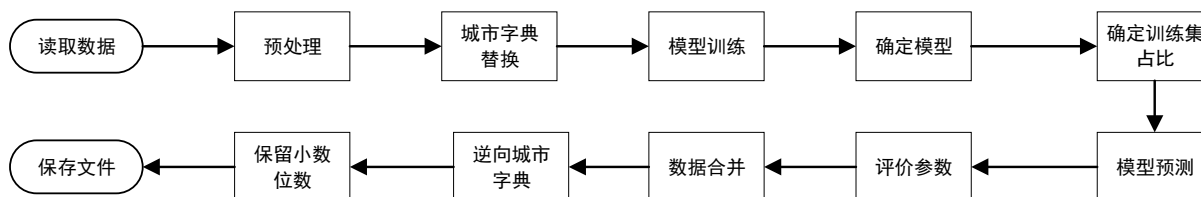


图 4-8 区县碳排放、植被固碳量预测流程图

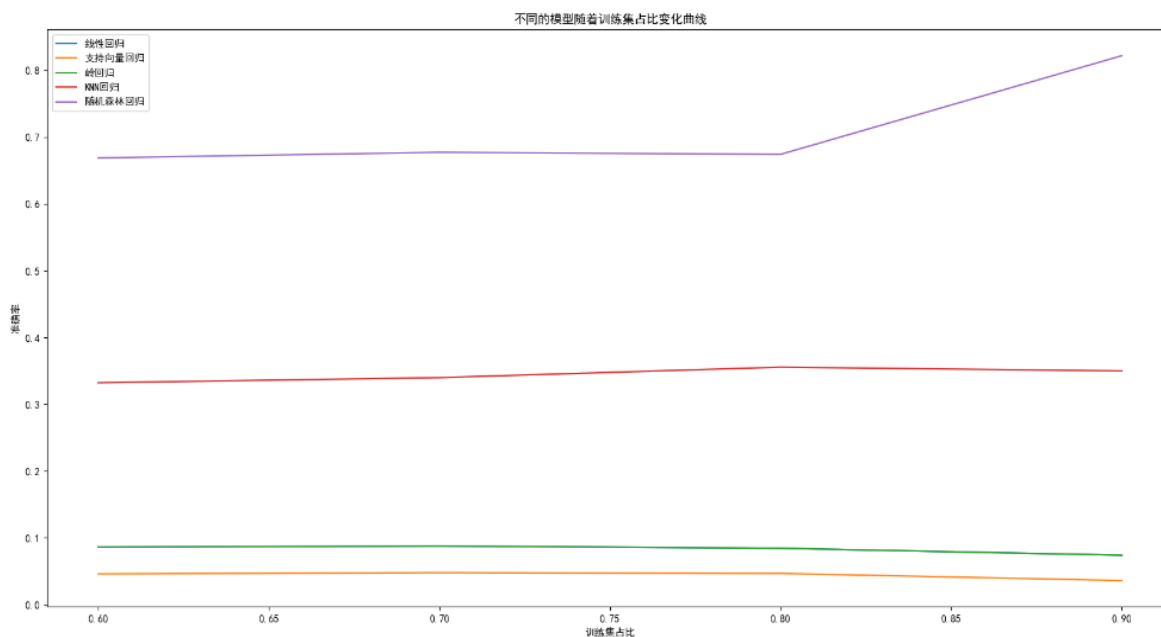


图 4-9 不同模型随训练集占比变化曲线

通过分析比对训练结果，选取训练集占比为 0.9 的随机森林模型作为最终要使用的模型，该模型的平均绝对误差、均方误差、可解释方差等都达到了预期的要求。在数据预测完成后，对数据进行合并，通过逆向字典转化成省份和城市名称，保留小数位数后进行保存。

整个数据预测模块，通过对比多种训练集占比情况下的不同模型，按照数据集特点，选择了最合适的训练模型，实现了数据预测的功能。

代码 4-2 预测关键代码

```

# 不同的训练集占比训练不同的模型
for i in range(0,4):
    train_x, test_x, train_y, test_y = train_test_split(feature , target,train_size=size[i],random_state=1)
    train_x = standard.fit_transform(train_x)
    test_x = standard.fit_transform(test_x)
  
```

```
model = LinearRegression()
model.fit( train_x , train_y )
scorelist[0].append(model.score(test_x , test_y ))

model = SVR()
model.fit( train_x , train_y )
scorelist[1].append(model.score(test_x , test_y ))

model = Ridge()
model.fit( train_x , train_y )
scorelist[2].append(model.score(test_x , test_y ))

model = KNeighborsRegressor()
model.fit( train_x , train_y )
scorelist[3].append(model.score(test_x , test_y ))

model = RandomForestRegressor()
model.fit( train_x , train_y )
scorelist[4].append(model.score(test_x , test_y ))
```

4.3 系统界面详细设计

4.3.1 概览模块界面

概览模块界面主要有三部分组成，顶部轮播图，左下角系统介绍，右下角相关资讯，如图 4-10 所示。

在轮播图板块，因为考虑到轮播图模块页面宽度方向空间空余，但高度方向空间匮乏，所以采用了 Element UI 中的卡片风格的 Carousel 走马灯。使用该组件，可以实现点击图片查看，或者点击左右两侧的箭头进行切换，或者鼠标悬浮在底部指示器显示不同位置的图片。在图片的选择上，选取“碳达峰”、“碳中和”的意义、碳核算的意义、双碳五个维度、落实双碳的方法等内容，能够让读者直接准确的了解相关工作的意义。

在系统介绍板块，采用 Element UI 中手风琴效果的折叠面板，手风琴效果为每次只能打开一个面板，这样在实现效果的同时，又能节约页面，使系统介绍和相关资讯两个板块更协调。在内容方面，该板块选择了“背景”、“国内外研究现状”、“系统目标”、“主要工作”、“创新点”五个方面进行介绍，通过简短的语言，让用户了解碳核算的背景以及本系统的工作。

模块的参考案例，使用的是 Element UI 中的 Drawer 抽屉来实现的。

省级分析模块的交互设计如图 4-12 所示，地图改变区域大小，实现折线图、柱状图的多图联动。折线图控制年份信息，年份的改变可导致地图、柱状图、饼图等信息的更新。河流图点击面积可以使饼图进行更新，点击纵轴年份可以使地图和柱状图更新。

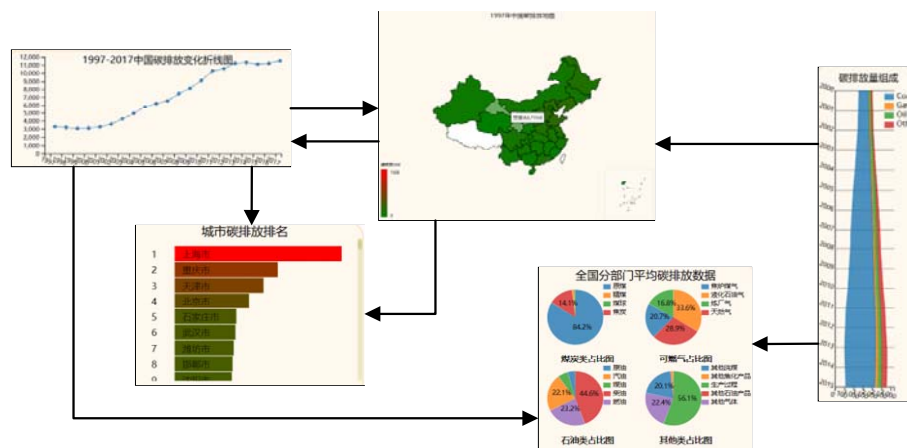


图 4-12 省级分析模块交互设计图

4.3.3 市级分析模块界面

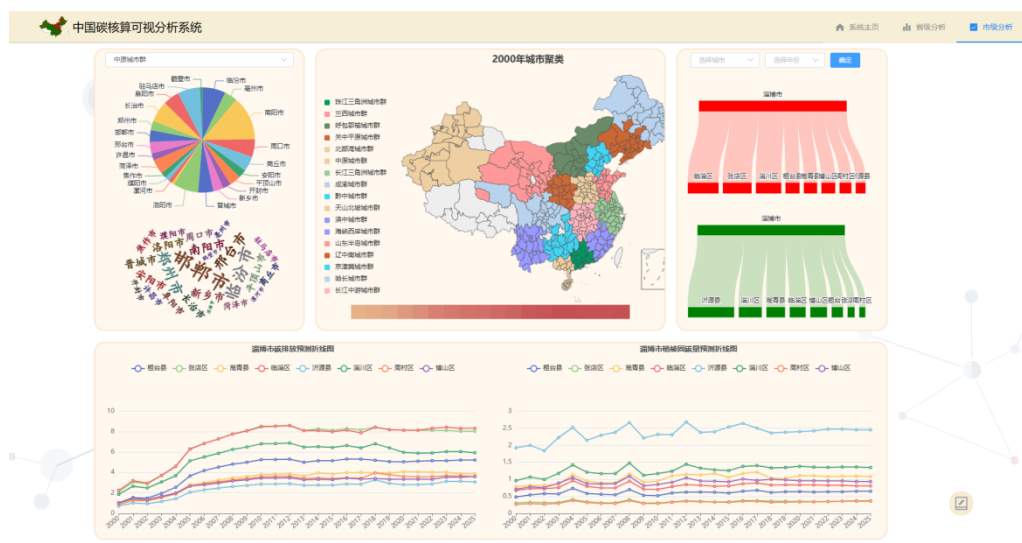


图 4-13 市级分析模块界面图

市级分析模块界面如图 4-13 所示，地图、词云、饼图显示城市数据，桑基图和折线图显示某一城市的区县的碳排放和植被固碳量的数据。在第三章的页面布局和可视化编码中已经进行了详细的描述。另外，右下角的图书按钮点击后会从侧边弹出市级分析案例，为用户提供使用思路。

本模块的主要交互方案是选择框和点击热力图，通过热力图切换不同的年份，可以实现多图联动，而选择框范围相对较小，只能使部分图更新数据。通过这些交互，可以从城市群、市、区县的空间维度和时间维度展示碳排放和植被固碳量信息。

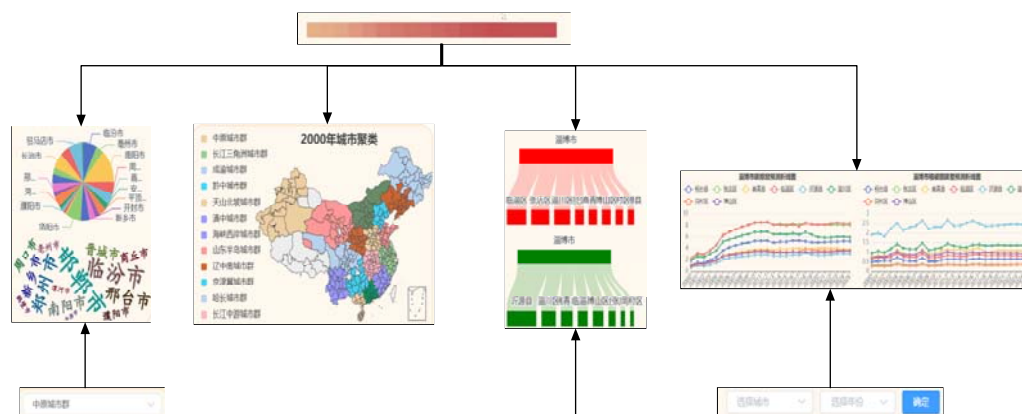


图 4-14 市级分析模块交互设计图

4.4 本章小结

本章在前两章需求分析、概要设计的基础上，对系统进行详细设计与实现，主要分为系统功能详细设计与实现、聚类与预测算法设计与实现、系统界面详细设计三部分。本章详细的说明了可视化的功能的实现逻辑和编程思想。系统功能详细设计与实现部分既梳理了前端页面所用的组件，又概括了图表的功能和交互逻辑，对底层函数也做了对应的说明。算法部分共使用了聚类和预测算法，对于聚类算法，针对 Sklearn 库中的 K-Means 中存在的不能定义初始点和聚类城市群混杂的两个问题，通过改写聚类算法，重新设置城市间距离“度量”公式，实现了符合要求的聚类功能。对于预测算法，进行数据处理后，将不同的训练集占比数据放入五种回归预测模型中进行预测，根据准确率确定当训练集占比为 0.9，模型类型为随机森林时，准确率最高，预测并处理后得到最终的预测数据。系统界面详细设计部分以 3.4 各模块的可视化布局设计为框架，结合 3.5 和 3.6 可视化编码和可视化交互方案设计，最终得到了完整的系统。

第 5 章 中国碳核算可视分析系统测试与评估

本章主要有系统功能测试、系统性能测试、系统评估三个部分。

测试的软硬件环境和测试工具如表 5-1、表 5-2 所示，其中功能测试采用黑盒测试，性能测试采用 Jmeter。

表 5-1 硬件测试环境信息表

机型（配置）			操作系统	用途
联想小新 14	20+1T	i5	Win11	运行系统客户端，功能测试，性能测试

表 5-2 软件测试环境及测试工具信息表

软件需求/测试工具	用途
Firefox	录制操作用于自动化测试
Jmeter	性能测试
Chrome、IE	进行软件兼容测试和界面测试
Katalon	录制脚本
Jmeter	自动化测试脚本并进行性能测试

5.1 系统功能测试

系统功能测试是对系统的所有功能分模块进行测试。将系统分为概览模块测试、省级分析模块测试、市级分析模块测试三部分，测试内容包括前端 UI 界面功能是否完整可用、可视化交互是否符合要求，后端是否能正确接收和发送数据。

5.1.1 概览模块测试

概览模块功能测试主要测试轮播图是否能够正常切换、折叠面板是否能够展开、相关资讯是否能够正常跳转。

表 5-3 “用户查看概览信息” 功能测试

序号	输入	预期输出/响应	测试结果	状态
1-1	点击导航栏	跳转到对应子页面	与预期的输出相同	正确
1-2	点击轮播图切换箭头	会切换到上（下）一张	与预期的输出相同	正确
1-3	鼠标悬浮到图片	图片停留，不继续切换	与预期的输出相同	正确

1-4	不对轮播图进行操作	轮播图自动切换	与预期的输出相同	正确
1-5	点击轮播图的底部触发器	轮播图切换到对应位置 图片	与预期的输出相同	正确
1-6	点击折叠面板	点击部分展开，且只展 开这一个面板	与预期的输出相同	正确
1-7	点击相关资讯	新建对应的资讯页面	与预期的输出相同	正确
1-8	点击数据来源	跳转到对应网站	与预期的输出相同	正确

5.1.2 省级分析模块测试

省级分析模块主要对单个图表功能、多图联动功能、案例分析功能进行测试。由于省级分析和市级分析的案例分析功能基本相同，所有在本模块与市级分析模块的案例分析一起进行功能测试。对于数据部分，则功能测试时在前后端控制台输出，测试完成后再把输出信息注释掉。

表 5-4 “省级分析模块单个图表” 功能测试

序号	输入	预期输出/响应	测试结果	状态
2-1	鼠标悬浮在柱状图上	显示城市碳排放信息	与预期的输出相同	正确
2-2	鼠标在柱状图内滚动	滚动查看其他城市信息	与预期的输出相同	正确
2-3	鼠标悬浮在饼图上	饼图突出高亮显示	与预期的输出相同	正确
2-4	鼠标悬浮在地图上	对应区域高亮显示且显 示碳排放量信息	与预期的输出相同	正确
2-5	点击某一省份	为该省份地图，小地图 切换为全国地图	与预期的输出相同	正确
2-6	点击全国小地图	大地图返回为全国地 图，小地图为南海地图 点突出显示，并通过提	与预期的输出相同	正确
2-7	鼠标悬浮在折线图的点上	示框输出对应年份碳排 放量信息	与预期的输出相同	正确

2-8	点击折线图的点	控制台输出年份	与预期的输出相同	正确
2-9	点击河流图的分类	控制台输出该分类信息	与预期的输出相同	正确
2-10	点击河流图的年份	控制台输出年份信息	与预期的输出相同	正确
2-11	读取全国碳排放量数据	控制台输出数据信息	与预期的输出相同	正确
2-12	读取全国分部门碳排放数据	控制台输出数据信息	与预期的输出相同	正确
2-13	读取各省份碳排放来源大类信息	控制台输出数据信息	与预期的输出相同	正确

表 5-5 “省级分析模块多图联动”功能测试

序号	输入	预期输出/响应	测试结果	状态
3-1	点击地图省份	折线图、柱状图更新为该省份信息	与预期的输出相同	正确
3-2	点击折线图年份	地图、柱状图、饼图更新为该年份信息	与预期的输出相同	正确
3-3	点击河流图某一类别	饼图突出显示该类别	与预期的输出相同	正确
3-4	点击河流图某一年份	地图、饼图、折线图、柱状图更新年份信息	与预期的输出相同	正确

表 5-6 “案例分析”功能测试

序号	输入	预期输出/响应	测试结果	状态
4-1	点击省级分析模块案例分析按钮	右侧显示省级案例分析内容	与预期的输出相同	正确
4-2	关闭省级分析模块案例分析界面	回到省级分析页面	与预期的输出相同	正确
4-3	点击市级分析模块案例分析按钮	右侧显示市级案例分析内容	与预期的输出相同	正确
4-4	关闭市级分析模块案例分析界面	回到市级分析页面	与预期的输出相同	正确

5.1.3 市级分析模块测试

表 5-7 “市级分析模块单个图表”功能测试

序号	输入	预期输出/响应	测试结果	状态
5-1	城市群选择框选择城市群	饼图和词云能够更新为正确的城市群	与预期的输出相同	正确
5-2	鼠标在饼图上悬浮	饼图突出显示以及提示植被固碳量信息	与预期的输出相同	正确
5-3	鼠标在词云上悬浮	词云突出显示以及提示碳排放量信息	与预期的输出相同	正确
5-4	鼠标悬浮在城市群图例上	高亮显示改城市群城市区域和名称	与预期的输出相同	正确
5-5	点击热力图	地图更新对应年份的城市群聚类	与预期的输出相同	正确
5-6	单选框选择年份和地域	控制台输出年份和地域信息	与预期的输出相同	正确
5-7	点击确定按钮	控制台输出后端查询数据	与预期的输出相同	正确
5-8	鼠标拖动桑基图	桑基图随鼠标移动	与预期的输出相同	正确
5-9	鼠标点击折线图图例	折线图隐藏/显示该条信息	与预期的输出相同	正确
5-10	获取后端某年份的碳排放预测数据	控制台输出数据信息	与预期的输出相同	正确
5-11	获取后端某年份的植被固碳量预测数据	控制台输出数据信息	与预期的输出相同	正确

表 5-8 “市级分析模块多图联动”功能测试

序号	输入	预期输出/响应	测试结果	状态
6-1	点击热力图	页面所有图表更新为该年份数据	与预期的输出相同	正确

5.2 系统性能测试

系统性能测试根据实际应用场景，用户为碳核算领域团队内部人员使用，网站不对外开放，所以并发用户数不会多于 100 人。如表 5-9 显示系统性能测试用例，图 5-1 显示 JMeter 的测试结构，图 5-2 到图 5-5 分别显示并发用户数在 5、20、50、100 时的情况。

表 5-9 系统性能测试用例表

序号	场景说明	并发用户数说明	预期结果	测试结果
1	用户浏览概览模块	5、20、50、100	响应时间不超过 3 秒，且不出现错误	测试通过
2	用户加载省级分析界面	5、20、50、100	响应时间不超过 3 秒，且不出现错误	测试通过
3	用户加载市级分析界面	5、20、50、100	响应时间不超过 3 秒，且不出现错误	测试通过
4	用户在市级分析界面点击确定按钮请求后端数据	5、20、50、100	响应时间不超过 3 秒，且不出现错误	测试通过

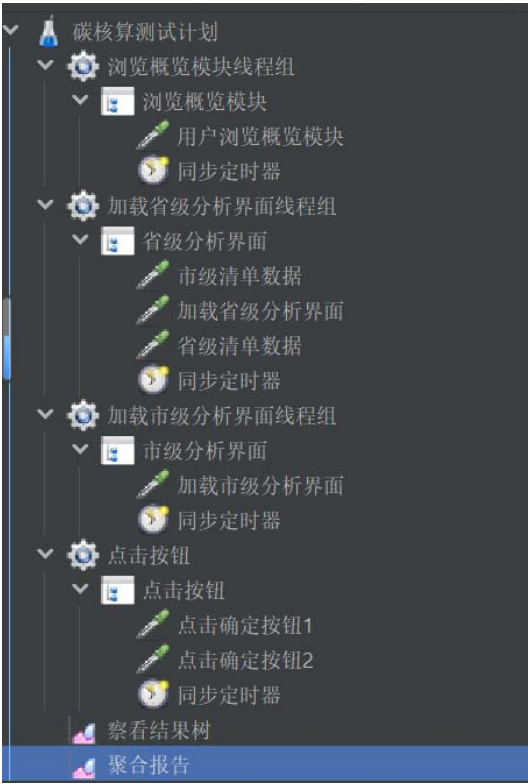


图 5-1 JMeter 测试计划结构图

JMeter 的测试目录结构如图 5-1 所示，测试计划中包含线程组、同步定时器、事务控制器、HTTP 请求、查看结果树、聚合报告等部分。根据测试用例分成了不同的线程组，每个线程组又一个同步定时器跟线程组的线程数相结合模拟并发用户的情况。

当并发数为 5 时，各个场景的响应时间在 200ms 以内，并且未出现错误情况，响应时间较短，测试通过。

Label	# 样本	平均值	中位数	90%	百分位95%	百分位99%	百分位最小值	最大值	异常 %	吞吐量	接收 KB/s	发送 KB/sec
用户浏览概览模块	5	10	10	12	13	13	6	13	0.00%	384.61538	261.04	43.57
浏览概览模块	5	10	10	12	13	13	6	13	0.00%	6.14251	4.17	0.7
市级清单数据	5	29	33	34	35	35	19	35	0.00%	142.85714	96136.02	19.81
加载市级分析界面	5	36	36	38	41	41	32	41	0.00%	119.04762	80.8	13.49
市级分析界面	5	36	36	38	41	41	32	41	0.00%	5.91716	4.02	0.67
加载省级分析界面	5	12	11	14	16	16	9	16	0.00%	312.5	212.1	35.4
省级清单数据	5	6	7	9	9	9	3	9	0.00%	500	76259.77	66.89
省级分析界面	5	48	53	53	56	56	33	56	0.00%	5.78035	4775.44	2.23
点击确定按钮1	5	119	119	120	120	120	118	120	0.00%	41.66667	958.5	8.14
点击确定按钮2	5	66	66	66	70	70	63	70	0.00%	71.42857	1645.37	13.88
点击按钮	5	185	185	185	188	188	183	188	0.00%	5.03018	231.58	1.96
总体	55	50	33	120	184	185	3	188	0.00%	55.1655	8761.75	10.05

图 5-2 并发数为 5 时的聚合报告

当并发数为 20 时，各个场景的响应时间不超过 600ms，并且未出现错误情况，响应时间可以接受，测试通过。

Label	# 样本	平均值	中位数	90%	百分位95%	百分位99%	百分位最小值	最大值	异常 %	吞吐量	接收 KB/s	发送 KB/sec
用户浏览概览模块	20	35	38	50	51	54	13	54	0.00%	370.37037	251.37	41.96
浏览概览模块	20	35	38	50	51	54	13	54	0.00%	19.92032	13.52	2.26
市级清单数据	20	118	106	178	187	190	57	190	0.00%	103.09278	69376.51	14.3
点击确定按钮1	20	236	213	302	306	325	167	325	0.00%	60.60606	1394.18	11.84
加载市级分析界面	20	215	204	251	251	252	189	252	0.00%	78.125	53.02	8.85
市级分析界面	20	215	204	251	251	252	189	252	0.00%	16.57001	11.25	1.88
加载省级分析界面	20	80	78	93	94	95	70	95	0.00%	208.33333	141.4	23.6
省级清单数据	20	33	35	51	57	59	7	59	0.00%	338.98305	51701.54	45.35
省级分析界面	20	233	218	280	301	303	168	303	0.00%	15.34919	12680.74	5.92
点击确定按钮2	20	156	164	210	212	215	101	215	0.00%	92.59259	2132.88	17.99
点击按钮	20	392	373	507	509	512	269	512	0.00%	13.34223	614.26	5.2
总体	220	159	168	298	372	507	7	512	0.00%	145.69536	23140.31	26.54

图 5-3 并发数为 20 时的聚合报告

当并发数为 50 时，各个场景的响应时间在 2000ms 以内，并且未出现错误情况，响应时间在可接受范围内，测试通过。

Label	# 样本	平均值	中位数	90%	百分位95%	百分位99%	百分位最小值	最大值	异常 %	吞吐量	接收 KB/s	发送 KB/sec
用户浏览概览模块	50	130	122	178	190	193	74	193	0.00%	243.90244	165.54	27.63
浏览概览模块	50	130	122	178	190	193	74	193	0.00%	42.15852	28.61	4.78
市级清单数据	50	646	555	1056	1095	1127	196	1127	0.00%	43.14064	29031.59	5.98
点击确定按钮1	50	566	598	833	850	899	172	899	0.00%	54.88474	1262.56	10.72
加载市级分析界面	50	843	907	1109	1113	1140	287	1140	0.00%	42.73504	29	4.84
市级分析界面	50	843	907	1109	1113	1140	287	1140	0.00%	23.22341	15.76	2.63
点击确定按钮2	50	589	514	981	1039	1083	239	1083	0.00%	45.95588	1058.6	8.93
点击按钮	50	1156	1084	1801	1829	1929	423	1929	0.00%	16.6334	765.79	6.48
加载省级分析界面	50	163	153	249	251	265	30	265	0.00%	185.87361	126.15	21.06
省级清单数据	50	448	544	621	629	640	12	640	0.00%	77.7605	11859.99	10.4
省级分析界面	50	1258	1261	1893	1932	1996	307	1996	0.00%	16.34521	13503.6	6.31
总体	550	616	547	1127	1614	1929	12	1996	0.00%	179.38682	28491.42	32.68

图 5-4 并发数为 50 时的聚合报告

当并发数为 100 时，各个场景的响应时间在 3000ms 以内，并且未出现错误情况，

响应时间达到要求，与预期结果相符，测试通过。

Label	# 样本	平均值	中位数	90%	百分位95%	百分位99%	百分位最小值	最大值	异常 %	吞吐量	接收 KB/s	发送 KB/sec
用户浏览概览模块	100	295	256	478	494	530	71	581	0.00%	164.47368	111.63	18.63
浏览概览模块	100	295	256	478	494	530	71	581	0.00%	62.57822	42.47	7.09
点击确定按钮1	100	904	935	1381	1438	1457	288	1463	0.00%	65.65988	1510.43	12.82
市级清单数据	100	1259	1188	1869	2134	2331	537	2337	0.00%	40.76641	27433.84	5.65
加载市级分析界面	100	1358	1365	1935	1976	2401	584	2403	0.00%	39.72984	26.97	4.5
市级分析界面	100	1358	1365	1935	1976	2401	584	2403	0.00%	28.36075	19.25	3.21
点击确定按钮2	100	819	810	1364	1428	1509	227	1516	0.00%	65.87615	1517.47	12.8
点击按钮	100	1723	1874	2602	2678	2777	521	2920	0.00%	24.61236	1133.13	9.59
加载省级分析界面	100	573	612	675	687	691	210	691	0.00%	144.09222	97.8	16.32
省级清单数据	100	171	149	332	346	357	7	360	0.00%	273.9726	41786.17	36.65
省级分析界面	100	2005	1960	2879	3038	3357	800	3382	0.00%	22.03614	18205.17	8.5
总体	1100	978	802	1974	2371	2879	7	3382	0.00%	241.70512	38389.23	44.03

图 5-5 并发数为 100 时的聚合报告

5.3 系统评估

系统评估部分因为概览模块内容较少都内容展示为主，所以只对省级分析模块和市级分析模块进行分析。结合系统目标，通过实际的案例对系统的实用性进行检验，评估是否符合预期，达到设计要求。

5.3.1 省级分析模块评估

在省级分析模块，先通过点击折线图来发现全国时间维度的碳排放变化关系。

发现在柱状图中，上海、重庆、天津三个直辖市在 1997 年到 2017 年二十年间碳排放排名一直占据前三位，并且碳排放量由原来的上海远大于其他城市到后来差距比例慢慢变小。其中北京由一开始的碳排放量第四名到 2005 年被苏州市超越保持在第五位，直到 2011 年北京碳排放量突然降到 13 名，2013 年到全国 20 名，2017 年更是达到了 32 名，如图 5-6 所示。



图 5-6 北京市全国碳排放排名变化图

通过横向柱状图可以掌握某一城市的名次变化信息，而碳排放量的名次变化经常与经济发展、高耗能产业相关联，这种名次的突变是尤其引人关注。查阅相关资料，造成这种现象的原因可能与北京市的产业转移、京津冀协同发展等政策有关。2005 年，北京市首钢搬迁到河北，从此拉开了北京产业转移的序幕。北京在 2010 年底，河北省为了

承接北京的产业转移,曾提出建设“环首都绿色经济圈”的计划。2014年,疏解北京非首都功能、推进京津冀协同发展更是上升为国家战略。通过一系列举措,北京市的高耗能企业减少,碳排放量降低,更符合城市定位,有利于持续健康发展。

同样在时间维度上,如图 5-7,通过河流图可以发现,2000 年到 2015 年期间,各部分的总量都在增长,但从 2013 年开始,总量基本保持稳定,增长缓慢,煤炭类甚至开始减少。其中煤炭占比远大于其他,其中原煤在煤炭类中的占比稳定在 80%左右。观察详细信息,发现原煤所占比例不断下降,从 2000 年占煤炭类的 84%到 2015 年的 77%。而天然气在可燃气类中的比例不断上升,从 2000 年的 29%上升到 2015 年的 48%。查阅资料,近 20 年间,我国已初步建成了一个包括西气东输在内的全国性的天然气管道网络,并且大力推进“煤改气”等方案的落实,碳排放量组成的变化从侧面说明了我国天然气基础设施建设与运营取得较好的成绩。

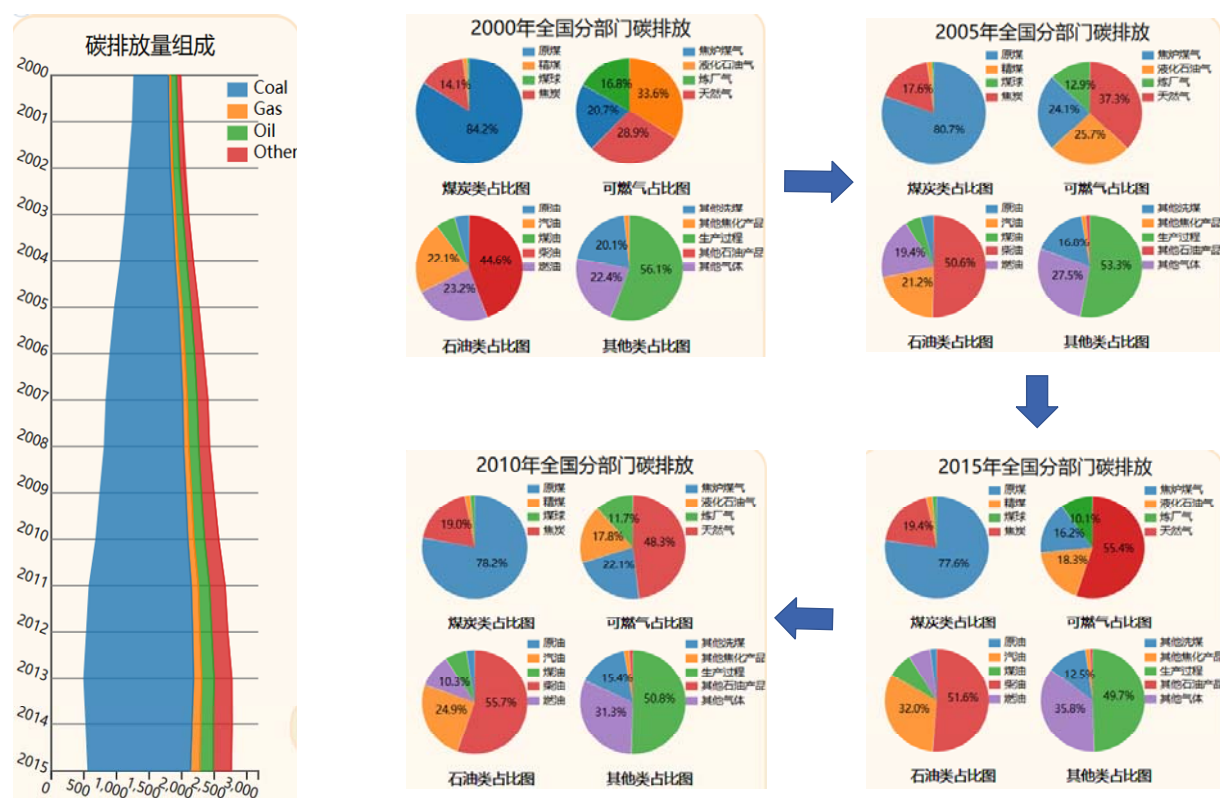


图 5-7 全国碳排放分部门组成变化图

通过不同年份的变化,如图 5-8 发现山西省的碳排放量 2013 年以后特别高,单纯的时间维度已经不能满足,于是进行空间维度的探索。

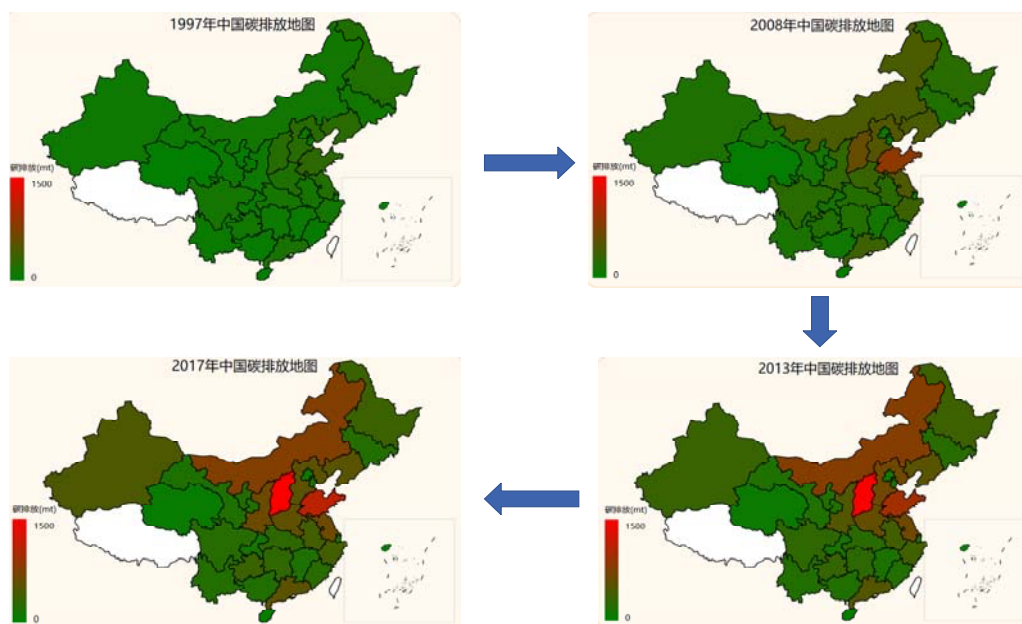


图 5-8 全国碳排放量空间变化

点击山西省，地图切换到山西省界面，如图 5-9 在时间维度上通过折线图发现山西省 2013 年碳排放量激增，空间维度上通过地图和城市碳排放柱状图发现碳排放较多的是临汾市、吕梁市。

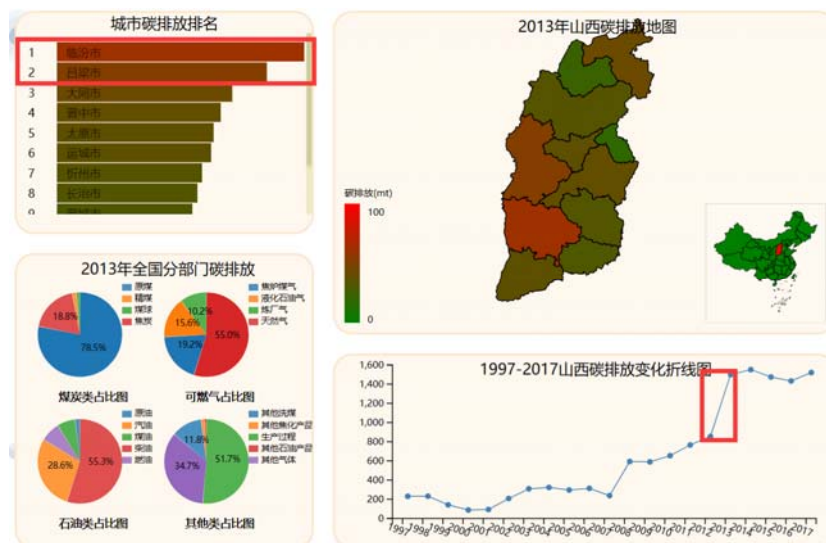


图 5-9 山西省 2013 年碳排放情况

于是，查阅相关资料，发现 2013 年山西煤炭产量创历史新高。煤炭产量与碳排放量有一定的相关性。山西省的煤炭除去在省内生产使用外，还有一部分会运输给其他省份，大部分是用来发电或者从事其他高耗能生产活动，这必然会造成一定的污染。于是从另一角度雾霾入手，发现“2013 年中国中东部严重雾霾事件”、“2013 年全国遭史上

最严重雾霾天气，创 52 年之最”、“国务院发布《大气污染防治行动计划》十条措施”、“国务院办公厅关于印发 2014-2015 年节能减排低碳发展行动方案的通知”等信息。

通过上述案例分析，猜测山西省的煤炭产量创历史新高导致碳排放量突增，同时煤炭产品参与其他省份的生产活动，叠加其他各类因素后对 2013 年的雾霾事件有一定的相关性。并且在此之后，国家出台了一系列如“大气十条”等节能减排条例和措施，2013 年之后碳排放增长开始减缓，煤炭所占比例稳步降低。

在省级分析模块，系统从空间和时间维度增强了用户对数据的认知。如在本模块中，一般的产煤区如山西、内蒙古等和人口密集地区如山东、河北、河南等碳排放量都比较高。在时间上，发达地区和不发达地区排名变化不明显但差值一直增大，如上海排放量一直比较高，而青海的所属市排放量一直较低，因为青海地处三江源，所以一直以保护生态为重。另外，也对中国的产业结构有了更明显的认知。在近二十年，原煤是我国的主要能源，占 70%以上。而 2021 年煤炭产量减少导致部分城市大面积停电也提醒我们要尽快改变我国的能源结构。最后，通过案例分析，也可以一定程度上了解政府政策的落实程度和我国社会能源部门结构的变化。综上，省级分析模块满足了设计需求，展示了时间和空间上我国的省级碳排放基本情况，发现了我国近二十年碳排放方针和政策的调整。

5.3.2 市级分析模块评估

市级分析模块又分为两部分，第一部分，饼图、词云图、地图和热力图是与城市群相关的内容，而桑基图、折线图是与某个城市有关的内容。相比于省级分析模块，市级分析模块数据加入了植被固碳量数据，能够更加综合的体现一个城市的减排固碳成绩。

饼图和词云图模块，根据饼图角度的大小和词云字体的大小，能够直观的发现植被固碳量和碳排放量较为突出的城市。如图 5-10，在 2000 年中原城市群植被固碳量饼图中，可以看到河南省南阳市占比较大，南阳市的森林覆盖率约 40%，同时是河南省面积最大的市，有“中州粮仓”之称，除森林外还有众多的农作物，植被固碳量高。碳排放量词云图中，河北邯郸市、山西临汾市、河南郑州市、河北邢台市字体较大。邯郸市有“钢城煤都”之称，临汾市含煤面积占全市总面积的 76%，煤炭在临汾占有重要地位，郑州市重工业比重占比下降缓慢，化石能源占比高。邢台市是以钢铁、机械、化工、纺织为主的轻重工业配合的城市，碳排放量较多。



图 5-10 2000 年中原城市群占比图

在时间维度上，各城市群中，如图 5-11 所示，天山北坡城市群、哈长城市群、兰西城市群较为稳定。长春大部分年份归属于辽中南城市群，究其原因，是黑龙江北部城市纬度差别较大，哈长城市群的中心向北偏移造成，北部湾城市群同理，因为海南而导致城市中心偏移。从城市类群的变迁中，能够大体看出某些城市的发展变化，如甘孜藏族自治州最开始隶属于成渝城市群，后来绝大部分年份属于滇中城市群，可以看出川渝城市群经济发展加快，能源消耗强度增高，而甘孜州则与云南的较低排放特征相似。同理，京津冀城市群也因为发展迅速导致碳排放量与植被固碳量的比值迅速上升，导致周边的承德市、秦皇岛市等城市相继加入其他城市群。

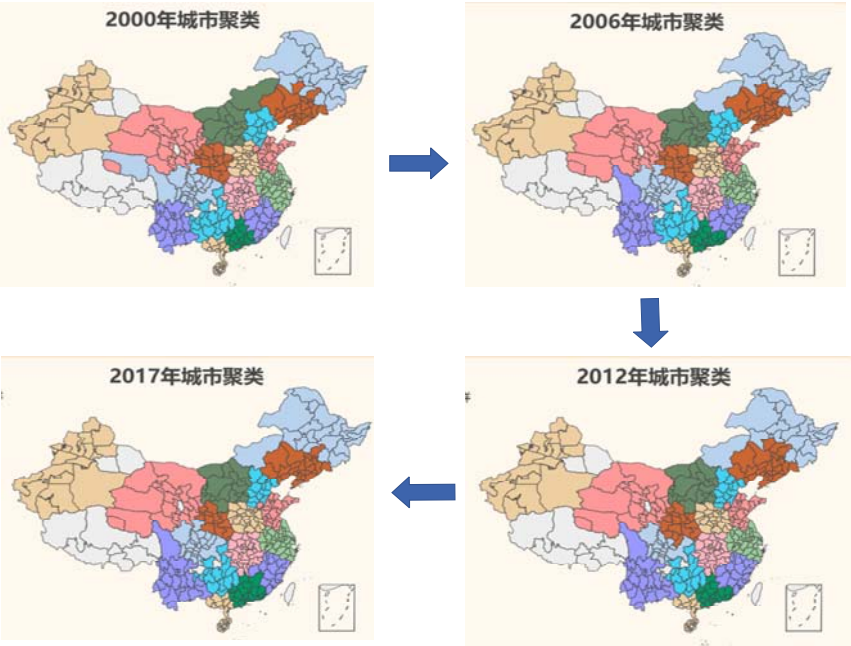


图 5-11 城市群变化图

在桑基图和折线图部分，选取山东省淄博市来进行案例分析。桑基图能直观的显示各个区县的占比情况，从桑基图中的数据可以看到，张店区、临淄区碳排放量稳居前二，沂源县最低。而植被固碳量部分，沂源县第一，临淄区、张店区排名靠后。从折线图的可以看到历年排名信息基本与桑基图相同。查阅资料可知，沂源县森立覆盖率 62%，位于山东丘陵，有山东屋脊之称，处于五市交界处，交通相对受限，经济发展相对较弱。而张店区、临淄区作为百年老工业城市，化工、钢铁等传统高耗能企业在碳排放上占比较大。通过折线图可以看到，淄博市从 2010 年开始碳排放量就处于稳定状态，仅 2009 年，淄博市就采取众多节能减排措施，如整治 2200 多家污染企业，关停 1200 多家土小企业，一系列措施导致碳排放几乎维持在稳定水平，并且观察预测数据发现也维持在稳定水平。



图 5-12 2000 年淄博市碳排放量植被固碳量组成图



图 5-13 淄博市碳排放量固碳量预测折线图

在本模块中，几个案例分析说明，在城市群中，碳排放量与植被固碳量差值大的以能源、钢铁、重工业为主，人口密度较大，经济较发达，这部分城市面临的碳达峰压力较大，污染水平也可能较大。而植被固碳量较大的城市都是产业结构以农林为主。这些城市容易实现碳达峰的目标。在查阅资料的过程中，几乎所有的资源型城市和钢铁等重工业城市的政府工作报告中都不断提及“淘汰落后产能”、“转型”、“节能减排”等字眼，通过桑基图和折线图也可以看出一些城市确实取得了不错的效果。综上，市级分

析模块达到了预期，符合设计要求。

5.4 本章小结

本章作为最后一个章节，主要设计软件的功能测试、性能测试和系统评估。功能测试根据模块采用了 6 个测试用例进行测试，测试内容包含单个图表的功能和多图联动的功能以及按钮、链接等页面元素。性能测试，根据实际需求，分别对系统进行加载数据和页面操作的并发测试，最终都达到预期。最后通过案例分析进行系统评估，通过对省级、市级模块的典型案例分析，达到了展示全国碳排放时空分布特点的目标，总结了城市群或个别城市的特征，将城市碳排放量和植被固碳量的对比，能够了解城市区县的特点，有利于补齐短板。

结论

本文利用中国碳核算数据库中的碳排放、植被固碳量、分部门核算清单等数据,设计实现了中国碳核算可视分析系统。可视化系统从省级市级角度进行可视化展示,另外还有概览模块和案例分析用于帮助用户来了解熟悉系统。经过评估,系统能够达到基本的要求,基本可以完成可视化分析任务,能够完成分析目标。

在技术层面,系统采用合理的前后端分离架构,前端采用 Vue,后端采用 Springboot,数据经过预处理后根据性能和空间需求分别存储在前端和 Mysql 数据库中。系统设计中,数据的预处理和模型的调试都采用 Python,数据预处理主要是将原始数据转化成适合的形式,聚类模型采用 K-Means,为了符合要求重写了距离“度量”公式,预测模型经过比较后选择了训练集占比 0.9 的随机森林模型进行预测。系统采用栅格布局进行布局设计。页面组件元素采用 Element UI 中的组件。可视化绘制分别采用 D3 和 Echarts 来完成,并且都设置了多图联动效果,能够进行更高层次的分析。

在功能层面,省级分析从时间、空间、碳排放量组成三个方面进行展示。通过图表间的交互,用户可以发现某一省份碳排放量的变化情况以及各部门碳排放趋势。通过变化发现我国煤炭、钢铁、重工业、人口密集区域碳排放量较高,以农林为主的区域植被固碳量较大。在产业结构上,原煤的产量逐年降低,天然气等清洁能源的占比不断增高。

“大气十条”等相关政策出台后,我国碳排放量增速放缓,表明节能减排取得一定效果。市级分析页面从城市群和城市下辖区县入手,将城市群碳排放量和植被固碳量占比进行展示,在时间维度上又能够通过聚类城市群显示城市变迁规律,能够细化到区县进行分析,了解城市组成特点。通过回归预测算法也能够大体预测出未来的碳排放量、植被固碳量的发展趋势。中国碳核算可视分析系统在功能上利用可视化展示大数据,多图交互合理,能够清晰准确的展示用户所需信息,研究内容都得到了很好的解决。

同时,系统还有几点不足,需要进行后续的改进和优化。首先是聚类算法方面,距离度量公式中的两个权重 w_1 和 w_2 需要手动设置,聚类的结果需要人工进行判断,后续可以进一步细化两个权重的结果,寻找到更合适的权重。其次是用户权限方面,目前系统不需要身份验证,对用户权限没有进行划分。后续可以添加登录功能,设置不同的权限等级。最后,系统只有对单个省份或城市的展示,后续可以对城市添加对比功能,用户可以选择两个或多个城市进行碳排放和植被固碳量的对比。

致谢

岁月如梭，转眼已经临近毕业，大学四年里，也还算是收获良多。在西南科技大学，我认识了一群亦师亦友的老师，不断地向师兄师姐们请教，结交了各具特点的优秀同学们，也对实验室的学弟学妹们进行过交流指导。

能够走到今天这一步，不光是我个人的努力，更离不开周围人的鼓励和帮助。首先要感谢张晓蓉老师、彭莉娟老师和王赋攀老师，张老师在我大二下学期就进行了论文的选题工作，后续又进行了多次指导，这给予了我很大的帮助。相比于代码，我认为更困难是如何设计好一个系统，张老师的指导总是会给予我正确的方向，在经历了多个版本的设计后，才有了现在系统的页面。另外，感谢华清远见的刘静老师，刘老师在实习期间给予了我很多的帮助，让我有了技术的积累。

其次，要感谢周阳师兄和鄢移凌师姐，虽然周师兄早已毕业一年，但在我大二大三的学习生活中总是不厌其烦的回答我的问题，解决我的疑惑，后来在我的毕业设计的布局 and 交互问题上提出了有用的建议，使我能更好的完成毕业设计。鄢师姐为我提供了一些参考文献，帮我理清了聚类算法的度量公式。

另外，还要感谢我的室友孔晨辉，在数据可视化、数据分析等课程中我们曾一起组队，而这些课程的学习，也让我积累了相关的项目经验，能够在毕业设计中实现这部分的内容。然后是实验室的学妹。在备战考研期间，虽然我没有写过系统的代码，但刘珊珊等几位学妹为我验证了一些技术方面可行性的猜想，这使得后来我在完成毕业设计时会有事半功倍的效果。

最后，是感谢我的父母，是他们一路上的支持，才能让我有一个无忧无虑的大学生活。

纸短情长，致谢的原因不仅是在毕业设计上给予了我帮助，更是在整个大学生涯中一段段美好的记忆和珍贵的感情。在虚拟现实与可视化实验室的三年间，老师的教诲、师兄的帮助、同学间的探讨，早已印入我的脑海。身边的老师和同学都在学习和生活上给了我很多建议和帮助，所以很荣幸研究生期间能够继续在虚拟现实与可视化实验室学习，在未来的日子里，一定做积极努力的自己，继续加油。

参考文献

- [1] Rogelj J, den Elzen M, Höhne N, et al. Paris Agreement climate proposals need a boost to keep warming well below 2 ° C[J]. Nature, 2016, 534(7609):631-639
- [2] Chen L, Msigwa G, Yang M, et al. Strategies to achieve a carbon neutral society: a review[J]. Environ Chem Lett, 2022, 20(4):2277-2310.
- [3] 位楠楠. 用碳同位素比技术定量分析城市碳排放的来源[D]. 衡阳: 南华大学, 2011.
- [4] Ya Zhou, Yuli Shan, Guosheng Liu, et al. Emissions and low-carbon development in Guangdong-Hong Kong-Macao Greater Bay Area cities and their surroundings[J]. Applied Energy, 2018, 228:1683-1692.
- [5] 张浩楠, 申融容, 张兴平, 等. 中国碳中和目标内涵与实现路径综述[J]. 气候变化研究进展, 2022, 18(02): 240-252.
- [6] 王灿, 张雅欣. 碳中和愿景的实现路径与政策体系[J]. 中国环境管理, 2020, 12(06):58-64.
- [7] 孙国道, 胡亚娟, 蒋莉, 等. 基于城市群的空气质量数据的可视分析方法研究[J]. 计算机辅助设计与图形学学报, 2017, 29(01):17-26.
- [8] 张忠华, 王智超, 胡杰. 全国碳排放及基于 K-means 聚类算法的省级碳排放研究[J]. 中外能源, 2022, 27(09):8-15.
- [9] 宋杰鲲. 基于支持向量回归机的中国碳排放预测模型[J]. 中国石油大学学报(自然科学版), 2012, 36(01):182-187.
- [10] 李立. 中国国家及区域碳排放分析[D]. 南京: 南京大学, 2017.