

李学文 王宏洲 李炳照 编

数学建模优秀论文 精选与点评 (2011—2015)

清华大学出版社

李学文 王宏洲 李炳照 编

数学建模优秀论文 精选与点评 (2011—2015)

清华大学出版社
北京

内 容 简 介

本书精选 2011—2015 年全国大学生数学建模竞赛多篇优秀获奖论文,辅以指导教师的精彩点评。论文未作删节,所有细节和计算过程均予以保留。

本书非常适合初次参赛的高校学生查阅,也可供青年指导教师参考。

版权所有,侵权必究。侵权举报电话:010-62782989 13701121933

图书在版编目(CIP)数据

数学建模优秀论文精选与点评:2011—2015/李学文,王宏洲,李炳照编. —北京:清华大学出版社,2017

ISBN 978-7-302-46091-6

I. ①数… II. ①李… ②王… ③李… III. ①数学模型—文集 IV. ①O141.4-53

中国版本图书馆 CIP 数据核字(2017)第 006176 号

责任编辑:陈 明

封面设计:傅瑞学

责任校对:刘玉霞

责任印制:何 芊

出版发行:清华大学出版社

网 址: <http://www.tup.com.cn>, <http://www.wqbook.com>

地 址:北京清华大学学研大厦 A 座

邮 编:100084

社 总 机:010-62770175

邮 购:010-62786544

投稿与读者服务:010-62776969, c-service@tup.tsinghua.edu.cn

质量反馈:010-62772015, zhiliang@tup.tsinghua.edu.cn

印 刷 者:三河市君旺印务有限公司

装 订 者:三河市新茂装订有限公司

经 销:全国新华书店

开 本:185mm×260mm 印 张:21.75

字 数:523 千字

版 次:2017 年 1 月第 1 版

印 次:2017 年 1 月第 1 次印刷

印 数:1~3000

定 价:42.00 元

产品编号:068111-01

为了培养大批具有良好数学基础,能用数学思想、方法和工具解决各个领域的实际问题的具有创新能力的复合型人才,数学教学应在传授基础理论和基本技能的同时,培养学生数学建模能力,即分析实际问题、归结实际问题为数学问题,并用数学方法作为工具去解决实际问题的能力。以往我们的数学课程往往板着脸孔讲理论,学生学了一大堆定义、定理和公式却不知道这些究竟有什么用,但是如果能在教学中充分体现数学建模的思想,加强对学生数学建模的训练,对学生学习数学可以起到事半功倍的效果。目前数学建模进入课堂已经成为一种普遍现象,而且逐渐从本科院校推广到专科院校及职业学校甚至条件好的高中学校。为了进一步推动数学建模活动在各校的开展,针对数学建模的竞赛活动也越来越多。随着国际上各种数学建模竞赛活动的开展,中国从 1992 年开始举办全国大学生数学建模竞赛(CUMCM),至今参赛队数也已有一万多队,成为目前在大学校园中具有广泛影响力的赛事之一,对提高我国大学生的数学建模能力和数学应用水平起到了极大的推动作用。

本书编者均在北京理工大学从事数学建模课程的教学和组织指导数学建模竞赛工作多年,在课堂上和竞赛培训中亲眼见证了广大学生对数学建模活动的热情和喜爱。参加建模活动的学生不仅学到了大量的实用数学知识,而且各方面能力也有了很大提高。为了运用数学建模解决实际问题,学生们不仅要了解常见的数学模型及相应的求解方法,比如初等数学模型、微积分模型、概率统计模型、优化模型、微分方程模型、运筹学模型等,还要学会应用常见的数学软件,如 MATLAB、LINDO/LINGO、SPSS 等,除此之外他们还要学会对实际问题建立数学模型并分析求解,最后还要写成一篇规范的学术论文,这对大一、大二从未接触过学术论文写作的学生来说具有很强的挑战性,是一个非常辛苦但也有很大收获的长期过程。在这个过程中,他们各方面的能力有了很大提高,这些对他们来说无论是以后继续深造还是参加实际工作都是非常有帮助的。

随着数学建模活动的广泛开展,参与这项活动的学生越来越多,学生迫切希望阅读参考以往竞赛的优秀论文。编者已经于 2011 年出版了一本《数学建模优秀论文精选与点评(2005—2010)》,收录了 2005—2010 年十余篇我校参加全国大学生数学建模竞赛获全国一、二等奖的论文全文,受到了广大参赛同学的热烈欢迎。现在这本论文集可以看作其姊妹篇。与前一本论文集相同之处是:论文保持了文章原貌,所有细节得以保留;与上一本不同的是:我们不是每个题目只登载一篇论文,而是根据情况同一个题目可能会有两三篇优秀获奖论文,大家会对建模赛题的开放性有更深入的理解,同样一个题目可以有不同的侧重点和思考角度,大家可以更深入地了解问题及其建模过程和解决方法。

本书精选了 2011—2015 年北京理工大学获全国数学建模竞赛一、二等奖的部分参赛论文,并在保留全文的基础上对每篇论文进行了点评,指出了文章的特点。本书可供参加数学建模课程学习和竞赛的学生及教师参考,尤其适合初学者参考。为了方便读者阅读,本书在每一篇优秀论文前面附上了相应的数学建模竞赛原题目及解题思路分析。需要指出的是,由于很多题目都给出了大量数据,本书因篇幅限制,无法给出全部数据,读者可以到全国大学生数学建模竞赛组委会网站^①上的往年试题栏目中下载。部分论文后面也附有算法程序及计算结果数据,因篇幅原因我们也作了省略。

本书选编的论文都是我校学生多年来在数学建模小组的各位老师指导下完成的,他们是:孙华飞、徐厚宝、闫桂峰、曹鹏、姜海燕、李春辉、杨国孝、曹春雷、蔡亮、熊春光、董岩、满红英、温海瑞、李庆娜等,本书的出版应该归功于他们及所指导学生多年辛勤的努力付出。我们也要感谢多年来一直参加数学建模课程教学和竞赛指导的其他各位老师,以及所有参加了课程学习和各级竞赛的同学们!本书的出版还得到了北京理工大学数学学院及清华大学出版社的大力支持,借此机会也一并表示感谢!

限于编者才疏学浅,书中难免有错误及不妥之处,敬请各位专家、同行和广大读者批评指正!

编 者

2016 年 9 月

^① <http://www.mcm.edu.cn>.

第 1 章 城市表层土壤重金属污染分析(2011 A)	1
1.1 城市表层土壤重金属污染分析	1
1.2 问题分析与建模思路概述	3
1.3 获奖论文——城市表层土壤重金属污染分析	3
1.4 论文点评	24
1.5 获奖论文——基于对流-反应扩散方程的污染源分析模型	25
1.6 论文点评	39
第 2 章 交巡警服务平台的设置与调度(2011 B)	41
2.1 交巡警服务平台的设置与调度	41
2.2 问题分析与建模思路概述	45
2.3 获奖论文——交巡警服务平台的设置与调度模型	46
2.4 论文点评	62
2.5 获奖论文——交巡警服务平台的设置与调度	62
2.6 论文点评	79
第 3 章 葡萄酒的评价(2012 A)	80
3.1 葡萄酒的评价	80
3.2 问题分析与建模思路概述	80
3.3 获奖论文——葡萄酒评价模型	81
3.4 论文点评	95
3.5 获奖论文——葡萄酒质量与酿酒葡萄和葡萄酒理化指标的关系探究	96
3.6 论文点评	113
第 4 章 太阳能小屋的设计(2012 B)	115
4.1 太阳能小屋的设计	115
4.2 问题分析与建模思路概述	124
4.3 获奖论文——光伏电池的优化铺设问题	125

4.4	论文点评	142
4.5	获奖论文——太阳能小屋的设计	142
4.6	论文点评	161
第5章	车道被占用对城市道路通行能力的影响(2013 A)	162
5.1	车道被占用对城市道路通行能力的影响	162
5.2	问题分析与建模思路概述	163
5.3	获奖论文——车道被占用对城市道路通行能力的影响	164
5.4	论文点评	188
5.5	获奖论文——车道被占用对城市道路通行能力的影响	188
5.6	论文点评	209
第6章	碎纸片的拼接复原(2013 B)	211
6.1	碎纸片的拼接复原	211
6.2	问题分析与建模思路概述	212
6.3	获奖论文——基于边缘灰度信息的碎纸片半自动拼接	213
6.4	论文点评	226
第7章	嫦娥三号软着陆轨道设计与控制策略(2014 A)	227
7.1	嫦娥三号软着陆轨道设计与控制策略	227
7.2	问题分析与建模思路概述	232
7.3	获奖论文——嫦娥三号软着陆轨道设计与控制策略建模	233
7.4	论文点评	246
7.5	获奖论文——嫦娥三号软着陆轨道设计与控制策略	246
7.6	论文点评	268
第8章	创意平板折叠桌(2014 B)	269
8.1	创意平板折叠桌	269
8.2	问题分析与建模思路概述	270
8.3	获奖论文——关于一种新型折叠桌外形的数模分析	271
8.4	论文点评	284
第9章	太阳影子定位(2015 A)	285
9.1	太阳影子定位	285
9.2	问题分析与建模思路概述	287
9.3	获奖论文——基于非线性拟合的视频影长定位分析	288
9.4	论文点评	317

第 10 章 “互联网+”时代的出租车资源配置(2015 B)	319
10.1 “互联网+”时代的出租车资源配置	319
10.2 问题分析与建模思路概述	319
10.3 获奖论文——“互联网+”时代的出租车资源合理配置问题	320
10.4 论文点评	336

第 1 章 城市表层土壤重金属污染分析

(2011 A)

1.1 城市表层土壤重金属污染分析

随着城市经济的快速发展和城市人口的不断增加,人类活动对城市环境质量的影响日益突出。对城市土壤地质环境异常的查证,以及如何应用查证获得的海量数据资料开展城市环境质量评价,研究人类活动影响下城市地质环境的演变模式,日益成为人们关注的焦点。

按照功能划分,城区一般可分为生活区、工业区、山区、交通区及公园绿地区等,分别记为 1 类区、2 类区、……、5 类区,不同的区域环境受人类活动影响的程度不同。

现对某城市城区土壤地质环境进行调查。为此,将所考察的城区划分为间距 1km 左右的网格子区域,按照每平方千米 1 个采样点对表层土(0~10cm 深度)进行取样、编号,并用 GPS 记录采样点的位置。应用专门仪器测试分析,获得了每个样本所含的多种化学元素的浓度数据。另一方面,按照 2km 的间距在那些远离人群及工业活动的自然区取样,将其作为该城区表层土壤中元素的背景值。

表 1-1 列出了采样点的位置、海拔高度及其所属功能区等信息,表 1-2 列出了 8 种主要重金属元素在采样点处的浓度,表 1-3 列出了 8 种主要重金属元素的背景值。

表 1-1 取样点位置及其所属功能区

编号	x/m	y/m	海拔/m	功能区
1	74	781	5	4
2	1 373	731	11	4
3	1 321	1 791	28	4
4	0	1 787	4	2
5	1 049	2 127	12	4
6	1 647	2 728	6	2
7	2 883	3 617	15	4
8	2 383	3 692	7	2
9	2 708	2 295	22	4
10	2 933	1 767	7	4
⋮	⋮	⋮	⋮	⋮
315	6 924	5 696	7	5
316	4 678	3 765	40	5

续表

编号	x/m	y/m	海拔/ m	功能区
317	6 182	2 005	25	5
318	5 985	2 567	44	4
319	7 653	1 952	48	5

表 1-2 8 种主要重金属元素的浓度

编号	As/($\mu g/g$)	Cd/(ng/g)	Cr/($\mu g/g$)	Cu/($\mu g/g$)	Hg/(ng/g)	Ni/($\mu g/g$)	Pb/($\mu g/g$)	Zn/($\mu g/g$)
1	7.84	153.80	44.31	20.56	266.00	18.20	35.38	72.35
2	5.93	146.20	45.05	22.51	86.00	17.20	36.18	94.59
3	4.90	439.20	29.07	64.56	109.00	10.60	74.32	218.37
4	6.56	223.90	40.08	25.17	950.00	15.40	32.28	117.35
5	6.35	525.20	59.35	117.53	800.00	20.20	169.96	726.02
6	14.08	1 092.90	67.96	308.61	1 040.00	28.20	434.80	966.73
7	8.94	269.80	95.83	44.81	121.00	17.80	62.91	166.73
8	9.62	1 066.20	285.58	2 528.48	13 500.00	41.70	381.64	1 417.86
9	7.41	1 123.90	88.17	151.64	16 000.00	25.80	172.36	926.84
10	8.72	267.10	65.56	29.65	63.00	21.70	36.94	100.41
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
315	6.47	197.00	38.18	21.09	64.00	18.60	40.18	168.05
316	6.47	100.70	36.19	13.31	42.00	11.50	34.34	56.23
317	4.79	119.10	35.76	19.71	44.00	9.90	39.66	67.06
318	7.56	63.50	33.65	21.90	60.00	12.50	41.29	60.50
319	9.35	156.00	57.36	31.06	59.00	25.80	51.03	95.90

表 1-3 8 种主要重金属元素的背景值

元素	平均值	标准偏差	范围
As/($\mu g/g$)	3.6	0.9	1.8~5.4
Cd/(ng/g)	130	30	70~190
Cr/($\mu g/g$)	31	9	13~49
Cu/($\mu g/g$)	13.2	3.6	6.0~20.4
Hg/(ng/g)	35	8	19~51
Ni/($\mu g/g$)	12.3	3.8	4.7~19.9
Pb/($\mu g/g$)	31	6	19~43
Zn/($\mu g/g$)	69	14	41~97

注：题目及数据附件都可以到全国大学生数学建模竞赛官方网站 <http://www.mcm.edu.cn> 下载。

现要求你们通过数学建模来完成以下任务：

(1) 给出 8 种主要重金属元素在该城区的空间分布，并分析该城区内不同区域重金属的污染程度。

(2) 通过数据分析，说明重金属污染的主要原因。

(3) 分析重金属污染物的传播特征，由此建立模型，确定污染源的位置。

(4) 分析你所建立模型的优缺点,为更好地研究城市地质环境的演变模式,还应收集什么信息?有了这些信息,如何建立模型解决问题?

1.2 问题分析与建模思路概述

土壤重金属污染问题是要求在已有采样数据的基础上,用数学模型描述污染物在城市中的空间分布、分布特征,分析主要的污染原因和污染源。在我们的现实生活中,工业生产和日常生活造成的污染问题日益严重,并逐渐引起了各国政府和公众的高度重视,因此对污染物在不同介质中的扩散与追根溯源的研究成果也很多。此类问题有很成熟的数值模拟方法和偏微分方程模型可供利用,因此参赛队首先要对此类数学方法和算法有一定的了解。

对于问题一,可以选用插值拟合的方法给出污染物浓度的分布特征。如果对MATLAB等数学软件很熟悉,就可以很快算出结果并给出示意图,但应该注意对算法的数学原理予以明确说明。

对于问题二,首先需要通过文献检索了解各类重金属常见的来源,然后对各个区域的污染物浓度结合周边的工农业设施、道路和河流等地理特征进行相关性分析。通常某些重金属污染物会一起出现,这对我们了解污染原因、污染源是有利的,根据各个区域不同种类重金属的分布情况可以更容易确定其污染原因。要注意的是,不能仅根据文献中的结论就简单地认定污染原因,应该结合数据和模型来进行合理的论证,在此基础上给出判断。

对于问题三,要求建模确认污染源的位置,可选的方法很多。比如根据不同时间点各个采样点的污染数据,画出不同时间的污染范围,以此确定污染源位置;也可以根据同一时间点各个区域的污染数据和地理信息,近似画出污染物浓度的等高线,由此确定出污染源的位置。更复杂一些,还可以采用基于偏微分方程的扩散模型,以污染物浓度数据为初始值,反推其扩散过程,由此确定污染源。

对于问题四,可以结合前三个问题中建模和计算的体会,对现实中污染物数据收集提出一些建议。

1.3 获奖论文——城市表层土壤重金属污染分析

作者: 张雨 张阳 都春霞

指导教师: 闫桂峰

获奖情况: 2011 全国数学建模竞赛一等奖

摘要

随着城市经济的快速发展和城市人口的不断增加,人类活动对城市环境质量的影响日显突出。本文对某城市城区表层土壤重金属的污染进行了分析与评价,研究了人类活动影响下城市地质环境的演变模式。

对问题一,我们首先对所给数据进行了特异值检验与剔除,检验标准是根据表 1-3 中所给数据,发现 8 种主要重金属元素背景值的范围在平均值加减 2 倍标准偏差内,所以当检测

数据落在平均值加减 2 倍标准差外时,被认为是特异值,应将其剔除。然后,在新得到的共 283 个样本点数据的基础上,通过 MATLAB 函数制图绘制不同元素的空间分布图,给出 8 种主要重金属元素在该城区的空间分布。接着,我们先后采用了内梅罗综合污染指数法和根据重金属毒性响应系数修正污染因子权值后的模糊综合评价模型对该城区不同区域重金属的污染程度进行了评价,并讨论了这两种模型的优缺点,发现后者比前者更符合客观实际。

对问题二,我们利用了 SPSS 软件的主成分分析功能对 5 个功能区土壤重金属元素的 8 项指标进行主成分分析,并结合问题一中得出的结论以及该城区 5 个功能区产生重金属污染物的特征,最终确定出各个功能区重金属污染的主要原因。

问题三为本文关键,是研究重点。首先是应用改进的二维扩散模型,建立新的适合该城区的扩散方程。题目所给的较为分散的 319 个数据点(剔除后剩余 283 个),分布在长宽分别达两万米的广阔区域内。综合考虑到结果的准确度和 MATLAB 计算时间的长短,采取了内插 1000 个数据点,然后将数据代入方程,利用 MATLAB 编程,求出各个元素随 x, y 的变化分布图,由图中极大或最大值点得出污染源的具体坐标和分布区域。

在问题四中,考虑到重金属可能的传播途径,选取了汽车数量、生活垃圾、工业垃圾、大气沉降四个主要因素对扩散方程进行改进,并给出了需要测量的值和求解污染物浓度方程的步骤,从而可以研究城市每个坐标点地质环境随时间变化的情况。

关键词: 重金属污染 表层土壤 内梅罗综合污染指数法 模糊综合评价模型 改进的二维扩散模型

1.3.1 问题的背景和重述

在城市化的进程中,飞速发展的经济和急剧增加的人口使人类对城市环境质量的影响日益增长。土壤中的重金属污染由于其潜伏性和长期性尚未受到应有重视,但其严重性和危害性同样不容忽视,这也正日益成为人类对城市环境的质量评价以及城市地质环境在人类活动影响下的演变模式的研究中的重要因素。

本题用采样法研究某城市土壤地质环境受重金属污染的情况。考虑人类不同活动的影响,按功能将城区分为生活区、工业区、山区、交通区及公园绿地区等,各记为 1 类区、2 类区、……、5 类区,然后在整体被考察区域进行由 GPS 记录地点的网格采样,间距为 1km,样本为表层土(0~10cm 深度)。对其测试分析得其中多种化学物质的浓度数据。同时以相同方法在远离人类活动的自然区以 2km 为间隔采样并测试分析,将其数据作为该城区表层土壤中元素的背景值。

表 1-1~表 1-3 中给出 4 类数据:采样点的位置和海拔高度,其所属功能区,8 种主要重金属元素在采样点处的浓度,8 种主要重金属元素的背景值。

建模研究以下问题:

(1) 结合采样点的空间位置坐标和 8 种重金属元素在采样点的浓度,给出它们在该城区的空间分布,并观察各处浓度高低,分析其中不同区域重金属的污染程度;

(2) 分析附表中数据、上一问所得重金属分布情况、5 类区域分布情况,说明重金属污染的主要原因;

(3) 分析重金属污染物的传播特征,由此建立模型,确定污染源的位置;

(4) 分析上一问所建模型的优缺点,为更好地研究城市地质环境演变模式,还应收集什么信息?有了这些信息,如何建立模型解决问题?

1.3.2 基本假设

(1) 当地污染源与各处土壤的重金属元素浓度变化状态是经过长时间到达的动态平衡状态,短时间内其浓度无明显增减,且污染源数量不再增加;

(2) 少数数据异常点除外,各个采样点测得数据基本准确可信;

(3) 土壤中重金属以地表传播为主要传播途径,其他方式次之甚至可以忽略;

(4) 当地各处土壤 pH 在 6.5 附近。

1.3.3 符号约定

$\bar{a}$: 均值;

s : 标准差;

P_i : 重金属元素 i 的污染指数;

X_i : 重金属 i 的含量实测值;

S_i : 污染物 i 的评价标准,取值为该区的自然背景值的均值;

P_g : 采样点的综合污染指数;

P_g : 重金属污染因子;

W_{ki} : k 样品 i 元素的权重;

$F(x, y, z)$: 扩散方程中的源强;

u : 指定坐标点的浓度或浓度函数。

1.3.4 问题的分析

问题一的分析

该问题要求给出 8 种主要重金属元素在该城区的空间分布并分析不同区域内重金属的污染程度。考虑到原始数据可能存在异常值,因此首先需要对原始数据进行特异值的检验及剔除,然后在新得到的数据上进行分析。重金属的空间分布可通过作图来表示,设想是软件绘出该城区的三维地形图,然后绘出每一种重金属元素浓度在该城区二维平面上的三维分布图,再结合以颜色深浅表示其浓度的标有海拔等高线的平面分布图,就可以由图得到 8 种主要重金属元素在该城区的空间分布。而分析污染程度即为对土壤重金属污染程度的评价,可以采用多种评价方法,如内梅罗综合污染指数法、层次分析法、污染负荷指数法、神经网络法、模糊数学法(模糊综合评价模型)等,也可以多种评价方法结合使用。

问题二的分析

该问题要求通过数据分析说明重金属污染的主要原因。本题的要点在于数据分析,而这里可采用多种分析方法,如数据挖掘方法、主成分分析法等。以主成分分析法为例,分析过程可由 SPSS 软件完成,得到表明各个地区主成分的表格。在使用分析法分析数据的基础上,结合问题一得出的结论及相关重金属污染主要原因的文献,了解 8 种不同重金属污染

物主要的产生方式和原因,综合考虑该城区 5 个功能区的特征,即相关人类活动所容易产生重金属污染物,读表分析,即可最终得出重金属污染的主要原因。

问题三的分析

该问题要求通过分析重金属污染物的传播特征,建立模型,确定污染源位置。本问是重点也是难点,需准确把握其特征来建立模型。本题样本来自 0~10cm 的浅土层,而重金属具有稳定、难溶、难挥发等特性,综合重金属可能的传播途径,最终确定其以地表传播为主要方式。由此建立了改进的二维扩散模型以研究重金属污染物的传播特征。针对数据点少,较分散,采用插值法扩充数据样本点,在此基础上,对数据样本代入公式编程求解得污染源,并在图中读取坐标位置。最后结合所得结果分析产生结果可能的原因。

问题四的分析

该问题要求收集信息改进模型使之能更好地研究城市地质环境的演变模式。这既是对之前所建模型缺点的思考和反省,又充分弥补了该题中重金属元素传播途径单一有限的不足,能以开放的角度建立更为实际而完善的传播特征模型。解决该问题首先需要考察重金属的各种传播途径和影响因素,将其中主要部分纳入考虑进而修正方程,得到较为完善的重金属传播特性模型,比如考虑大气沉降等途径,降雨等天气的影响。而且上三问中假设该城区的重金属浓度值处在一个动态平衡的状态,而本题则应该将污染源产生重金属污染物的强度变化纳入考虑,建模模拟该城区各个点动态的重金属元素浓度,这样得到的计算结果会更接近实际。

1.3.5 模型的建立与求解

1. 模型的准备

对数据进行剔除异常值等处理工作。

1) 剔除特异值

特异值指观测数据中存在的过大或过小的值,可能是由于土壤采样、化验、分析过程中出现操作错误或条件改变导致。因此,有必要对原始数据进行特异值检验,并予以剔除。根据表 1-3 所给数据,可知 8 种重金属元素背景值的范围是在 $\bar{a} \pm 2s$ 之间($\bar{a}$ 为其均值, s 为其标准差),由此我们判断、剔除特异值的方法采用样本平均值 $\bar{a}$ 加减 2 倍标准差 s 法,认为数据在区间 $(\bar{a} - 2s, \bar{a} + 2s)$ 内为正常值,否则为特异值,给予剔除处理。

对表 1-2 中给出的 319 个样本点的数据使用 Excel 处理,结果如下:

As 测量值中特异值有 6 个,分别为样本点 6、29、30、41、84、178 处测得的数据; Cd 有 18 个,分别在样本点 6、8、9、16、22、34、35、40、90、95、143、152、163、179、192、223、310、313 处; Cr 有 5 个,分别在样本点 8、14、20、22、49 处; Cu 有 2 个,分别在样本点 8、22 处; Hg 有 4 个,分别在样本点 8、9、182、257 处; Ni 有 5 个,分别在样本点 8、22、61、128、135 处; Pb 有 12 个,分别在样本点 5、6、8、9、16、20、22、31、143、163、221、253 处; Zn 有 10 个,分别在样本点 6、8、9、20、22、30、36、61、143、178 处。共计剔除特异值样本点 36 个,筛选出 283 个样本点的数据。

用 SPSS 对剔除前后的样本分别作统计分析,得到的描述统计量如表 1-4、表 1-5 所示。

表 1-4

	描述统计量					
	N 统计量	极小值 统计量	极大值 统计量	均值		标准差 统计量
				统计量	标准误	
As	319	1.61	30.13	5.676 5	0.169 33	3.024 29
Cd	319	40.00	1 619.80	302.396 2	12.596 89	224.987 60
Cr	319	15.32	920.84	53.509 7	3.919 35	70.001 79
Cu	319	2.29	2 528.48	55.0167	9.121 49	162.915 10
Hg	319	8.57	16 000.00	299.711 3	91.236 71	1 629.539 78
Ni	319	4.27	142.50	17.261 8	0.556 61	9.941 42
Pb	319	19.68	472.48	61.740 9	2.802 70	50.057 76
Zn	319	32.86	3760.82	201.202 6	18.993 38	339.232 54
有效的 N (列表状态)	319					

表 1-5

	描述统计量					
	N 统计量	极小值 统计量	极大值 统计量	均值		标准差 统计量
				统计量	标准误	
As	283	1.61	10.99	5.281 3	0.118 13	1.987 28
Cd	283	40.00	751.20	249.962 9	8.396 82	141.256 33
Cr	283	15.32	172.29	43.919 8	1.159 89	19.512 32
Cu	283	2.29	277.82	38.165 4	2.200 81	37.023 44
Hg	283	8.57	1801.00	95.903 0	11.280 47	189.766 83
Ni	283	4.27	36.00	15.861 4	0.325 73	5.479 58
Pb	283	19.68	140.62	52.317 3	1.540 89	25.921 80
Zn	283	32.86	572.96	135.191 3	5.718 92	96.207 11
有效的 N (列表状态)	283					

对比分析可知,剔除 36 个特异值后样本的标准差和均值的标准误差都显著减小,说明剔除特异值确实对计算误差的减小有助益,故之后的统计计算均采用剔除特异值后 283 个样本点的数据。

2) 数据的标准化处理

在实际问题中,不同的数据可能有不同的性质和不同的量纲,为使原始数据能够符合模糊聚类的要求,需将原始数据(即表 1-2 中剔除特异值后的各采样点 8 种重金属元素的浓度)列成 283×8 的矩阵,并作标准化处理。本文中,我们通过极差变换

$$x'_{ij} = \frac{x_{ij} - \min_{1 \leq i \leq n} \{x_{ij}\}}{\max_{1 \leq i \leq n} \{x_{ij}\} - \min_{1 \leq i \leq n} \{x_{ij}\}}$$

(1)

将原始数据标准化,由于将 283 个样本点的数据极差化后所得表格过大,所以收在附录中,表 1-6 仅列出 5 个功能区的极差化后的数据(其中每个功能区每种重金属元素用于极差标准化的值是该区所有样本点的均值,见表 1-7,以下所有关于 5 个功能区的数据处理计算都基于均值)。

表 1-6

	As	Cd	Cr	Cu	Hg	Ni	Pb	Zn
生活区	0.436 326 767 09	0.363 606 244 38	0.207 708 184 36	0.143 288 934 43	0.140 464 219 33	0.390 203 449 8	0.274 851 162 79	0.259 115 231 14
工业区	0.468 670 921 93	0.441 625 728 59	0.376 057 074 91	0.266 235 210 84	0.108 039 976 8	0.414 510 362 22	0.355 483 042 37	0.394 019 988 34
山区	0.250 952 1218 7	0.306 374 966	0.262 628 983 8	0.220 224 277 37	0.156 682 699 66	0.322 283 273 72	0.178 592 289 72	0.270 383 985 47
交通区	0.487 886 644 02	0.355 864 553 78	0.218 289 744 26	0.140 056 760 99	0.051 858 728 624	0.472 110 323 5	0.318 615 352 33	0.246 685 976 1
公园绿地区	0.415 914 518 15	0.281 092 012 13	0.314 331 584 02	0.534 657 244 43	0.075 830 961 899	0.385 483 870 97	0.216 866 205 47	0.335 131 289 34

表 1-7 均值

	As ($\mu\text{g/g}$)	Cd (ng/g)	Cr ($\mu\text{g/g}$)	Cu ($\mu\text{g/g}$)	Hg (ng/g)	Ni ($\mu\text{g/g}$)	Pb ($\mu\text{g/g}$)	Zn ($\mu\text{g/g}$)
生活区	6.105 5	248.677 5	50.411 75	43.993 25	87.569 75	17.712 5	53.976 5	143.367 75
工业区	5.888 965 517 2	307.313 793 1	42.664 137 931	44.475 172 414	205.096 206 9	17.422 413 793	64.072 413 793	153.546 551 72
山区	4.106 25	152.623 437 5	36.306 875	16.994 375	41.432 343 75	14.069 843 75	36.496 25	70.624 531 25
交通区	5.278 907 563	299.596 638 66	46.782 100 84	49.522 268 908	97.026 470 588	16.146 806 723	57.602 521 008	172.166 806 72
公园绿地区	6.084 838 709 7	208.4	41.447 096 774	24.855 161 29	112.649 677 42	14.615 806 452	51.554 193 548	98.830 967 742

2. 模型的建立与求解

1) 问题一

使用上一步中处理过的数据,将其输入 MATLAB 中作图,依照输入不同依次得到该城区的地形图、等高线地形图、8 种元素关于该城区平面的浓度三维分布图和平面浓度分布图,如图 1-1~图 1-10 所示。

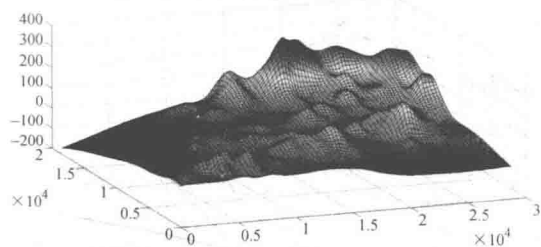


图 1-1 三维地形图

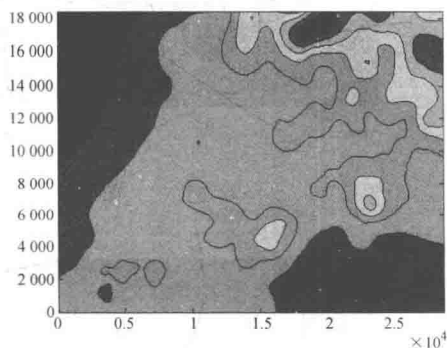


图 1-2 等高线地形图

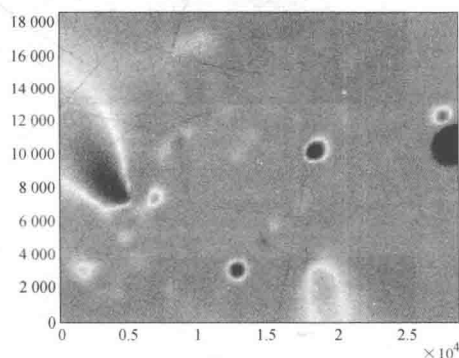
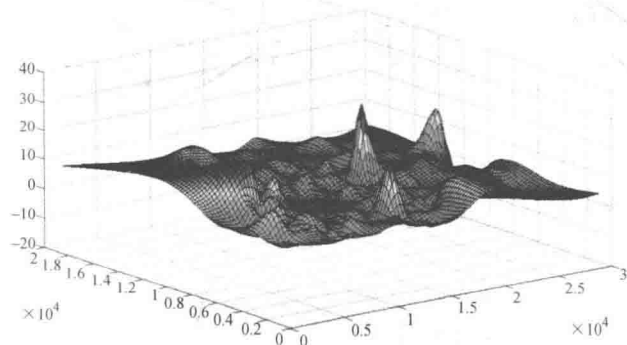


图 1-3 As 三维浓度分布图和平面浓度分布

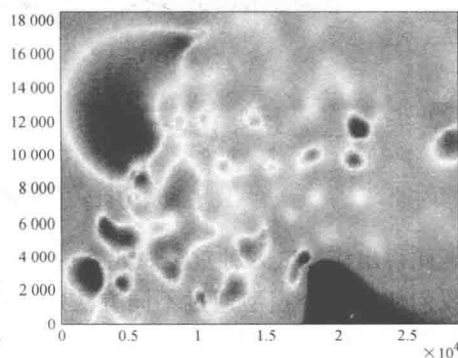
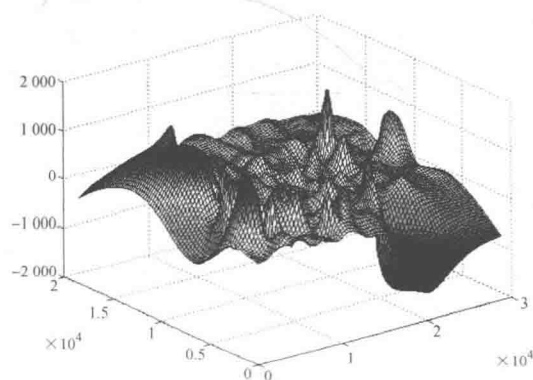


图 1-4 Cd 三维浓度分布图和平面浓度分布