

参赛密码 \_\_\_\_\_  
(由组委会填写)

## 第十届华为杯全国研究生数学建模竞赛

学 校                      南京航空航天大学

---

参赛队号                      10287032

---

	1.	蒋雪峰
队员姓名	2.	李 洁
	3.	蒋 奎

---

参赛密码 \_\_\_\_\_  
(由组委会填写)



## 第十届华为杯全国研究生数学建模竞赛

题目 对空气中 PM2.5 问题的建模研究

### 摘 要:

本文针对空气中 PM2.5 的相关因素分析、分布与演变及应急处理和空气质量控制管理的问题,采用相关分析、回归分析、二维插值、微分方程、非线性规划等方法,建立了相关性分析模型、多元线性回归方程模型、Shepard 二维插值模型、偏微分方程模型和多目标非线性规划模型,运用最小二乘估计算法、数值插值算法等问题进行了求解与分析。

问题一中,要求对 PM2.5 进行相关因素分析。首先,运用相关性分析方法,建立 AQI 中 6 个基本监测指标间的相关性分析模型,利用 SPSS 软件进行求解,得到各指标间的相关性,如 PM2.5 与一氧化碳间相关系数为 0.822,呈显著正相关;然后,建立了 PM2.5 与其它 5 项分指标间的两两回归分析模型并分析;最后,运用回归分析的方法,建立 PM2.5 与其它 5 项分指标间的多元线性回归模型,采用最小二乘估计方法对回归系数进行估计,并对回归方程进行了残差检验,对异常点进行了逐次剔除最终得到拟合度达到 97.1%的多元线性回归方程。

问题二中,要求对 PM2.5 的分布与演变及应急处理进行研究。针对第 1 小问,首先,研究了 PM2.5 随时间的变化规律,建立了基于三次样条插值法的 PM2.5 随时间变化模型;然后,建立了 Shepard 二维插值模型,利用 MATLAB 软件编程求解,得到了 PM2.5 的空间分布规律,如 PM2.5 在高压开关厂的含量最大;最后,建立了分区污染评估模型,并得到了其评估结果,如属于一级区的草滩为中度污染,属于二级区的长安区为轻度污染。

针对问题二第 2 小问,首先,运用回归分析的方法,建立了 PM2.5 与气象因素间的多元线性回归模型来分析气象因素对 PM2.5 的影响,如在冬季时,PM2.5 浓度与湿度和温度呈正相关;然后,在建立 PM2.5 在边界层中扩散的物理模型和迁移、转化的基本模型的基础上建立了 PM2.5 扩散的偏微分方程模型,并利用 MATLAB 软件编程求解得到了 PM2.5 扩散的偏微分方程模型的解析解。

针对问题二第 3 小问, 首先, 建立了 PM2.5 在地面的浓度分布模型和 PM2.5 污染扩散的预测模型; 然后, 对 PM2.5 污染扩散预测模型进行了参数估计; 最后, 代入实例得到了各个监测点 PM2.5 浓度的预测评估结果, 如高新西区为重度污染区域, 广运潭为安全区域。针对问题二第 4 小问, 用残差检验与稳定性检验了模型合理性, 并总结已有研究成果给出了 PM2.5 的成因、演变等一般性规律。

问题三中, 要求对空气质量进行控制管理。针对第 1 小问, 引入了效用函数建立以满意度最大为目标的非线性规划模型, 利用了 LINGO 13.0 版优化软件进行了编程求解, 得到未来五年 PM2.5 每年的全年年终要求达标的年平均浓度分别为: 226.1835, 174.89086, 126.00372, 79.40928, 35 (单位为  $\mu\text{g}/\text{m}^3$ )。针对第 2 小问, 建立了以投入总费用最少和 PM2.5 减排计划实施满意度最大为目标的多目标非线性规划模型。在求解过程中, 利用了主要目标法将双目标简化为单目标; 然后利用了 LINGO 13.0 版优化软件进行了编程求解, 得出了五年投入的总费用最少为 64.13053 费用单位 (百万元), 并对方案的合理性进行了论述。

本文的特色在于, 在问题一中, 采用了最小二乘估计方法对回归系数进行估计, 并对回归方程进行了残差检验, 且对异常点进行了逐次剔除最终得到拟合度达到 97.1% 的多元线性回归方程; 在问题二中, 考虑了多方面情况, 建立了 Shepard 二维插值模型和偏微分方程模型; 在问题三中, 引入了效用函数建立以满意度最大为目标的非线性规划模型, 在对多目标非线性规划模型求解过程中, 利用了主要目标法将双目标简化为单目标。

**关键词:** 相关分析、回归分析、二维插值、最小二乘估计、残差检验、偏微分方程模型、多目标非线性规划模型

# 目 录

一 问题的提出.....	4
二 基本假设.....	6
三 符号说明.....	6
四 问题分析.....	7
五 模型的建立与求解.....	8
5.1 问题一：PM2.5 的相关因素分析.....	8
5.1.1 问题一的分析与数据预处理.....	8
5.1.1.1 问题一的分析.....	8
5.1.1.2 数据预处理.....	9
5.1.2 问题一模型的建立.....	9
5.1.2.1 AQI 中 6 个基本监测指标间的相关性分析模型的建立.....	9
5.1.2.2 PM2.5 与其它 5 项分指标间的两两回归分析模型的建立.....	10
5.1.2.3 PM2.5 与其它 5 项分指标间的多元线性回归模型的建立.....	10
5.1.3 问题一模型的求解.....	12
5.1.3.1 AQI 中 6 个基本监测指标间的相关性分析模型的求解.....	12
5.1.3.2 PM2.5 与其它 5 项分指标间的两两回归分析模型的求解.....	12
5.1.3.3 PM2.5 与其它 5 项分指标间的多元线性回归模型的求解.....	15
5.1.4 问题一结果的分析及验证.....	17
5.2 问题二：PM2.5 的分布与演变及应急处理.....	18
5.2.1 问题二第 1 小问模型的建立与求解.....	18
5.2.1.1 问题二第 1 小问的分析.....	18
5.2.1.2 基于三次样条插值法的 PM2.5 随时间变化模型.....	18
5.2.1.3 基于 Shepard 二维插值法的 PM2.5 空间分布模型.....	20
5.2.1.4 分区污染评估模型.....	23
5.2.2 问题二第 2 小问模型的建立与求解.....	25
5.2.2.1 问题二第 2 小问的分析.....	25
5.2.2.2 PM2.5 与气象因素间的多元回归模型的建立与求解.....	25
5.2.2.3 PM2.5 扩散模型的建立与求解.....	26
5.2.3 问题二第 3 小问模型的建立与求解.....	28
5.2.3.1 问题二第 3 小问的分析.....	28
5.2.3.2 PM2.5 污染扩散的预测与评估模型的建立.....	29
5.2.3.3 PM2.5 污染扩散的预测与评估模型的求解.....	30
5.2.4 问题二第 4 小问.....	33
5.3 问题三：空气质量的控制管理.....	34
5.3.1 问题三第 1 小问模型的建立与求解.....	34
5.3.1.1 问题三第 1 小问模型的建立.....	34
5.3.1.2 问题三第 1 小问模型的求解.....	36
5.3.2 问题三第 2 小问模型的建立与求解.....	37
5.3.2.1 问题三第 2 小问模型的分析.....	37
5.3.2.2 问题三第 2 小问模型的建立.....	38
5.3.2.3 问题三第 2 小问模型的求解.....	39
六 模型的评价与推广.....	41
七 参考文献.....	42
八 附录.....	42

## 一 问题的提出

大气为地球上生命的繁衍与人类的发展提供了理想的环境。它的状态和变化，直接影响着人类的生产、生活和生存。空气质量问题始终是政府、环境保护部门和全国人民关注的热点问题。

对空气质量监测，预报和控制等问题，国家和地方政府均制定了相应政策、法规和管理办法。2012年2月29日，环境保护部公布了新修订的《环境空气质量标准》(GB3095—2012)，本次修订的主要内容：调整了环境空气功能区分类，将三类区并入二类区；增设了颗粒物(粒径小于等于 $2.5\mu\text{m}$ )浓度限值和臭氧8小时平均浓度限值；调整了颗粒物(粒径小于等于 $10\mu\text{m}$ )、二氧化氮、铅和苯并(a)芘等的浓度限值；调整了数据统计的有效性规定。与新标准同步还实施了《环境空气质量指数(AQI)技术规定(试行)》(HJ633—2012)。

上述规定中，启用空气质量指数AQI作为空气质量监测指标，以代替原来的空气质量监测指标——空气污染指数API(Air Pollution Index)。原监测指标API为无量纲指数，它的分项监测指标为3个基本指标(二氧化硫 $\text{SO}_2$ 、二氧化氮 $\text{NO}_2$ 和可吸入颗粒物 $\text{PM}_{10}$ )。AQI也是无量纲指数，它的分项监测指标为6个基本监测指标(二氧化硫 $\text{SO}_2$ 、二氧化氮 $\text{NO}_2$ 、可吸入颗粒物 $\text{PM}_{10}$ 、细颗粒物 $\text{PM}_{2.5}$ 、臭氧 $\text{O}_3$ 和一氧化碳 $\text{CO}$ 等6项)。新标准中，首次将产生灰霾的主要因素——对人类健康危害极大的细颗粒物 $\text{PM}_{2.5}$ 的浓度指标作为空气质量监测指标。新监测标准的发布和实施，将会对空气质量的监测，改善生存环境起到重要的作用。

由于细颗粒物 $\text{PM}_{2.5}$ 进入公众视线的时间还很短，在学术界也是新课题，尤其是对细颗粒物 $\text{PM}_{2.5}$ 及相关的因素的统计数据还太少，对细颗粒物 $\text{PM}_{2.5}$ 的客观规律也了解得很不够。但是相关研究人员绝不能因此而放慢前进的脚步，不能“等”数据，因为全国人民等不起。我们必须千方百计利用现有的数据开展研究，同时新课题、探索性研究、“灰箱问题”也有可能成为数学建模爱好者的用武之地。请研究以下问题。

### 一、 $\text{PM}_{2.5}$ 的相关因素分析

$\text{PM}_{2.5}$ 的形成机理和过程比较复杂，主要来源有自然源(植物花粉和孢子、土壤扬尘、海盐、森林火灾、火山爆发等)和人为源(燃烧燃料、工业生产过程排放、交通运输排放等)，可以分为一次颗粒物(即由排放源直接排放到大气中

的颗粒物)和二次颗粒物(即通过与大气组成成分发生化学反应后生成的颗粒物)。PM<sub>2.5</sub>的成分主要由水溶性离子、颗粒有机物和微量元素等组成。有一种研究认为,AQI监测指标中的二氧化硫(SO<sub>2</sub>),二氧化氮(NO<sub>2</sub>),一氧化碳(CO)是在一定环境条件下形成PM<sub>2.5</sub>前的主要气态物体。请依据附件1或附件2中的数据或自行采集数据,利用或建立适当的数学模型,对AQI中6个基本监测指标的相关与独立性进行定量分析,尤其是对其中PM<sub>2.5</sub>(含量)与其它5项分指标及其对应污染物(含量)之间的相关性及其关系进行分析。如果你们进而发现AQI基本监测指标以外的、与PM<sub>2.5</sub>强相关的(可监测的)成分要素,请陈述你们的方法、定量分析结果、数据及来源。

## 二、PM<sub>2.5</sub>的分布与演变及应急处理

请依据附件2、附件3中的数据或自行采集某地区的数据,通过数学建模探索完成以下研究:

- 1、描述该地区内PM<sub>2.5</sub>的时空分布及其规律,并结合环境保护部新修订的《环境空气质量标准》分区进行污染评估。
- 2、建立能够刻画该地区PM<sub>2.5</sub>的发生和演变(扩散与衰减等)规律的数学模型,合理考虑风力、湿度等天气和季节因素的影响,并利用该地区的数据进行定量与定性分析。
- 3、假设该地区某监测点处的PM<sub>2.5</sub>的浓度突然增至数倍,且延续数小时,请建立针对这种突发情形的污染扩散预测与评估方法。并以该地区PM<sub>2.5</sub>监测数据最高的一天为例,在全地区PM<sub>2.5</sub>浓度最高点处的浓度增至2倍,持续2小时,利用你们的模型进行预测评估,给出重度污染和可能安全区域。
- 4、采用适当方法检验你们模型和方法的合理性,并根据已有研究成果探索PM<sub>2.5</sub>的成因、演变等一般性规律。

## 三、空气质量的控制管理

地方环境管理部门关心的重要问题之一是,为建设良好的人居环境,利用有限财力,制定本地区空气质量首要污染物PM<sub>2.5</sub>的减排治污可行规划。数据1所在地区的环境保护部门考虑治污达标的紧迫性和可行性,在未来五年内,拟采取综合治理和专项治理相结合的逐年达到治理目标的方案。请考虑以下问题:

- 1、该地区目前PM<sub>2.5</sub>的年平均浓度估计为280(单位为 ),要求未来五年内逐年减少PM<sub>2.5</sub>的年平均浓度,最终达到年终平均浓度统计指标35(单位为 ),请给出合理的治理计划,即给出每年的全年年终平均治理指标。

2、据估算，综合治理费用，每减少一个 PM2.5 浓度单位，当年需投入一个费用单位（百万元），专项治理投入费用是当年所减少 PM2.5 浓度平方的 0.005 倍（百万元）。请你为数据 1 所在地区设计有效的专项治理计划，使得既达到预定 PM2.5 减排计划，同时使经费投入较为合理，要求你给出五年投入总经费和逐年轻费投入预算计划，并论述该方案的合理性。

## 二 基本假设

- (1) 假设所给数据均真实可靠，具有统计意义；
- (2) 假设整个扩散过程 PM2.5 污染物质量守恒；
- (3) 假设整个研究空间内风速场是均匀稳定的；
- (4) 假设 PM2.5 污染源强是连续均匀的；
- (5) 假设某监测点处的 PM2.5 的浓度有可能突然增至数倍，且延续数小时；
- (6) 假设仅考虑 PM2.5 的专项治理和综合治理问题。

## 三 符号说明

$c_i$	第 $i$ 个监测站的 PM2.5 平均含量（分指数）
$x_{ij}$	第 $i$ 个监测站在第 $j$ 天的实测值；
$n_i$	第 $i$ 个监测站实测值个数。
$c(x, y, z)$	空间某一点 PM2.5 的浓度；
$u$	PM2.5 所处高度处的风速,单位 $m/s$
$\sigma_y$	垂直于主导风方向的横向扩散系数,单位 $m$
$\sigma_z$	铅直扩散系数,单位 $m$
$Q$	单位时间内的 PM2.5 排放量,单位 $mg/s$
$Z$	投入的总费用；
$Z_1$	综合治理投入的费用；
$Z_2$	专项治理投入的费用；
$x_i$	第 $i$ 年经综合治理 PM2.5 减少的年平均浓度；
$y_i$	第 $i$ 年经专项治理 PM2.5 减少的年平均浓度；
$u(x_i)$	第 $i$ 年综合治理的效用函数；

$u(y_i)$	第 $i$ 年专项治理的效用函数；
$F$	治理的满意度；

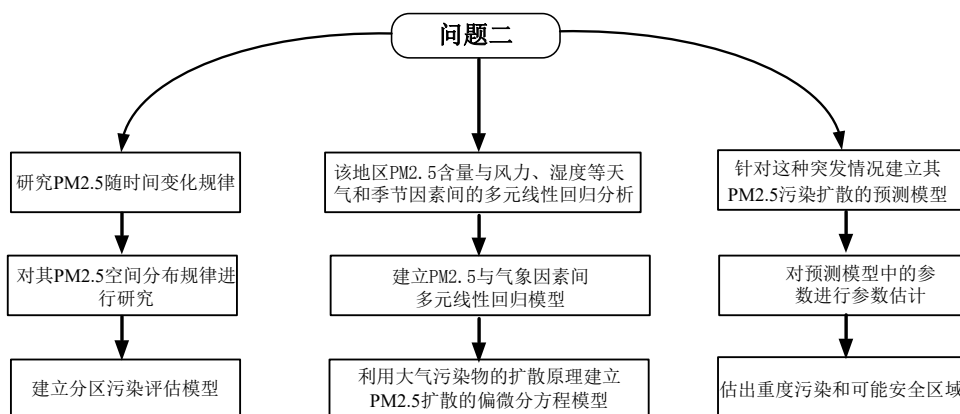
## 四 问题分析

本文是一个研究空气中 PM2.5 的相关因素分析、分布与演变及应急处理和空气质量控制管理的问题。

问题一要求依据附件 1 或附件 2 中的数据或自行采集数据，利用或建立适当的数学模型，对 AQI 中 6 个基本监测指标的相关与独立性进行定量分析，尤其是对其中 PM2.5（含量）与其它 5 项分指标及其对应污染物（含量）之间的相关性及其关系进行分析。根据问题一要求，对此，需首先对提供的附件数据进行预处理；然后，对其 AQI 中 6 个基本监测指标的相关与独立性进行定量分析，这里可建立其 6 个基本监测指标间的相关性分析模型，并用 SPSS 软件进行处理；其次，为探索发现 PM2.5 与其它 5 项分指标间两两间的相关性和关系，可对 PM2.5 与其它 5 项分指标间进行两两回归分析；最后，为了进一步探索发现 PM2.5 与所有的其它 5 项分指标间的关系，可对其建立 PM2.5 与其它 5 项分指标间的多元线性回归模型进行研究，其实现软件可用 MATLAB 软件对其进行编程求解，从而得到 PM2.5 与所有的其它 5 项分指标间的关系。

问题二中，要求对 PM2.5 的分布与演变及应急处理进行研究。针对第 1 小问，该问题要求描述该地区内 PM2.5 的时空分布及其规律，并结合环境保护部新修订的《环境空气质量标准》分区进行污染评估。对此，我们可首先可研究 PM2.5 随时间变化规律，对其可先绘制散点图进行发现大体趋势，其次，可以对其 PM2.5 随时间变化进行插值；然后，对其 PM2.5 空间分布规律进行研究，这里可采用基于二维插值法进行研究；最后，可建立分区污染评估模型再对其分区进行污染评估。针对问题二第 2 小问，问题要求建立能够刻画该地区 PM2.5 的发生和演变（扩散与衰减等）规律的数学模型，合理考虑风力、湿度等天气和季节因素的影响，并利用该地区的数据进行定量与定性分析。对此，可先利用该地区的数据进行 PM2.5 与风力、湿度等天气和季节因素间的多元回归分析，即建立 PM2.5 与气象因素间的多元回归模型；然后，考虑 PM2.5 自身的扩散问题，可利用大气污染物的扩散原理建立 PM2.5 扩散的偏微分方程模型。针对问题二第 3 小问，问题要求假设该地区某监测点处的 PM2.5 的浓度突然增至数倍，且延续数小时时，建立针对这种突发情形的污染扩散预测与评估方法。并以该地区 PM2.5 监测数据最高的一天为例，在全地区 PM2.5 浓度最高点处的浓度增至 2 倍，持续 2 小时，并利用该模型进行预测评估，给出重度污染和可能安全区域。首先，我们针对这种情况建立其 PM2.5 污染扩散的预测模型；然后对预测模型中的参数进行参数估计；最后，利用 MATLAB 软件进行实例求解，并评估出重度污染和可能安全区域。





问题二的流程框图

问题三中，要求对空气质量进行控制管理。针对第 1 小问，该问题要求我们针对该地区目前 PM2.5 的年平均浓度估计为 280（单位为  $\mu\text{g}/\text{m}^3$ ）时，未来五年内逐年减少 PM2.5 的年平均浓度，最终达到年终平均浓度统计指标 35，给出合理的治理计划，即给出每年的全年年终平均治理指标。对此需建立其相应的规划模型。针对第 2 小问，该问题要求同时考虑综合治理和专项治理，为数据 1 所在地区设计有效的专项治理计划，使得既达到预定 PM2.5 减排计划,同时使经费投入较为合理，并要求给出五年投入总经费和逐年轻费投入预算计划。对此，可建立其相应的多目标规划模型。

## 五 模型的建立与求解

### 5.1 问题一：PM2.5 的相关因素分析

#### 5.1.1 问题一的分析与数据预处理

##### 5.1.1.1 问题一的分析

问题一要求依据附件 1 或附件 2 中的数据或自行采集数据，利用或建立适当的数学模型，对 AQI 中 6 个基本监测指标的相关与独立性进行定量分析，尤其是对其中 PM2.5（含量）与其它 5 项分指标及其对应污染物（含量）之间的相关性及其关系进行分析。

根据问题一要求，对此，需首先对提供的附件数据进行预处理；然后，对其 AQI 中 6 个基本监测指标的相关与独立性进行定量分析，这里可建立其 6 个基本监测指标间的相关性分析模型，并用 SPSS 软件<sup>[1]</sup>进行处理；其次，为探索发现 PM2.5 与其它 5 项分指标间两两间的相关性和关系，可对 PM2.5 与其它 5 项分指标间进行两两回归分析；最后，为了进一步探索发现 PM2.5 与所有的其它 5 项分指标间的关系，可对其建立 PM2.5 与其它 5 项分指标间的多元线性回归模型进行研究，其实现软件可用 MATLAB 软件对其进行编程求解，从而得到 PM2.5 与所有的其它 5 项分指标间的关系。

综上所述得出问题一的建模流程如图 1 所示：

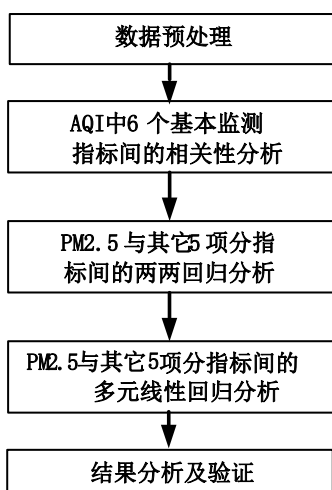


图 1 问题一流程框图

### 5.1.1.2 数据预处理

由于附件中收集的数据有缺漏或有个别本身没有测量，存在一些缺省值，所以在利用数据搭建模型之前，需在附件提供的数据库基础上进行数据预处理。

由题知数据中缺测指标的浓度和分指数均使用“NA”标识，此外缺漏值用“--”标识，对此，我们先将附件中的数据利用 SPSS 软件中的数据处理功能对其缺省值进行了处理，由于对于缺失数据不能全面的反映模型的变化规律，所以在分析时利用了 SPSS 进行了直接剔除。

### 5.1.2 问题一模型的建立

#### 5.1.2.1 AQI 中 6 个基本监测指标间的相关性分析模型的建立

根据相关性分析法理论基础<sup>[2]</sup>可知：设有随机变量  $X$  与  $Y$ ，对其进行了  $n$  次随机试验，得到的观测值分别为  $(X_i, Y_i) (i=1, 2, \dots, n)$ ， $\bar{X}$ 、 $\bar{Y}$  分别为各自的期望值， $\sqrt{DX}$  与  $\sqrt{DY}$  分别为  $X$  与  $Y$  的方差， $Cov(X, Y)$  为协方差， $r$  为相关系数， $R$  为随机变量  $X$  与  $Y$  对于样本  $(X_i, Y_i) (i=1, 2, \dots, n)$  的相关性系数，称之为样本相关系数。在实际中，常常用样本相关系数  $R$  作为相关系数  $r$  估计值。

根据以上理论知识可建立以下相关性系数模型：

$$\left\{ \begin{array}{l} r = \frac{Cov(X, Y)}{\sqrt{DX} \sqrt{DY}} \\ R = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2 \sum_{i=1}^n (Y_i - \bar{Y})^2}} \\ Cov(X, Y) = E[(X - \bar{X})(Y - \bar{Y})] \\ \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \\ \bar{Y} = \frac{1}{n} \sum_{i=1}^n Y_i \end{array} \right.$$

对有多个变量间的关系，其中随机变量  $X$  与  $Y$  可表示其中任意两个随机变量。

### 5.1.2.2 PM2.5 与其它 5 项分指标间的两两回归分析模型的建立

#### Step1: PM2.5 与其它 5 项分指标间的两两关系散点图分析

为了大致的分析细颗粒物 PM2.5 分别与与二氧化硫  $SO_2$ 、二氧化氮  $NO_2$ 、可吸入颗粒物 PM10、臭氧  $O_3$  和一氧化碳 CO 等 5 项指标的关系。我们利用 MATLAB 软件对附件 1. 数据 1.(武汉市一个监测点数据: 2013.01.01-2013.08.26)分别做出了 PM2.5 与  $SO_2$ 、 $NO_2$ 、PM10、 $O_3$  和 CO 等 5 项指标的散点图如图 2 所示(程序见附录 8.1.1)。

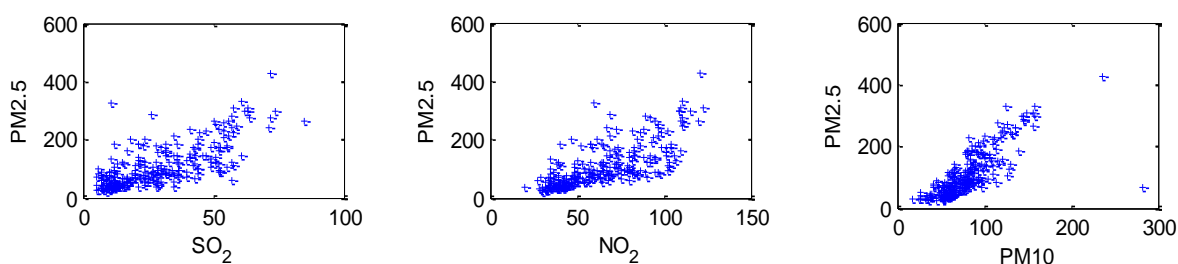


图 2(a) PM2.5 对  $SO_2$  的散点图 图 2(b) PM2.5 对  $NO_2$  的散点图 图 2(c) PM2.5 对 PM10 的散点图

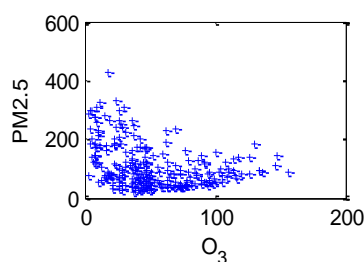


图 2(d) PM2.5 对  $O_3$  的散点图

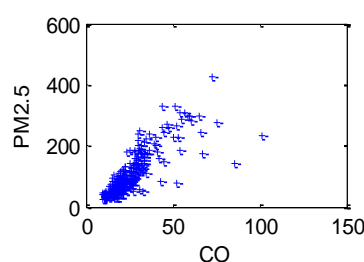


图 2(e) PM2.5 对 CO 的散点图

图 2 PM2.5 对 5 项指标的散点图

#### Step2: PM2.5 与其它 5 项分指标间的两两回归分析模型的建立

从图 2 中可以发现, PM2.5 的含量与其它 5 项分指标含量间大致呈增长趋势, 可拟合为常用的一次线性模型、二次模型或三次模型等。

其一次线性拟合模型为:  $y = \beta_0 + \beta_1 x$

二次模型为:  $y = \beta_0 + \beta_1 x + \beta_2 x^2$

三次模型为:  $y = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3$

其中,  $y$  为 PM2.5 的含量,  $x$  为与 PM2.5 进行两两分析的其它 5 项分指标之一的含量。

### 5.1.2.3 PM2.5 与其它 5 项分指标间的多元线性回归模型的建立

#### Step1: 建立多元线性回归的一般模型:

$$y = \beta_1 x_1(u) + \beta_2 x_2(u) + \beta_3 x_3(u) + \beta_4 x_4(u) + \beta_5 x_5(u) + \varepsilon \quad (1)$$

其中, 自变量  $x_1(u)$  为二氧化硫 ( $SO_2$ ) 的含量,  $x_2(u)$  为二氧化氮 ( $NO_2$ ) 的含量,  $x_3(u)$  为可吸入颗粒物 (PM10) 的含量,  $x_4(u)$  为一氧化碳 (CO) 的含量,  $x_5(u)$  为臭氧 ( $O_3$ ) 的含量, 因变量  $y$  为细颗粒物 (PM2.5) 的含量,  $\varepsilon$  为随

机误差，且服从于  $N(0, \sigma^2)$ 。

假设作了  $n$  次试验，得到  $n$  组观测数据为：

$$\begin{bmatrix} u_1 & y_1 \\ u_2 & y_2 \\ \vdots & \vdots \\ u_n & y_n \end{bmatrix} \quad (2)$$

带入 (1) 式可得

$$\begin{cases} y = \beta_1 x_1(u_i) + \beta_2 x_2(u_i) + \beta_3 x_3(u_i) + \beta_4 x_4(u_i) + \beta_5 x_5(u_i) + \varepsilon \\ i = 1, 2, \dots, n \end{cases} \quad (3)$$

其中， $\varepsilon_i$  为第  $i$  次试验的随机误差，且相互独立服从于  $N(0, \sigma^2)$ 。

### Step2: 回归系数 $\beta$ 的最小二乘估计

为了方便，引入矩阵记号：

$$\mathbf{Y} = \begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{bmatrix}, \mathbf{X} = \begin{bmatrix} x_1(u_1) & x_2(u_1) & \dots & x_m(u_1) \\ x_1(u_2) & x_2(u_2) & \dots & x_m(u_2) \\ \dots & \dots & \dots & \dots \\ x_1(u_n) & x_2(u_n) & \dots & x_m(u_n) \end{bmatrix}, \boldsymbol{\beta} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \dots \\ \beta_m \end{bmatrix} \quad (4)$$

选取  $\boldsymbol{\beta}$  的一个估计值  $\hat{\boldsymbol{\beta}}$ ，使得随机误差  $\boldsymbol{\varepsilon}$  的平方和达到最小，即

$$\min \boldsymbol{\varepsilon}' \boldsymbol{\varepsilon} = \min (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})' (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}) \quad (5)$$

写成分量形式

$$Q(\beta_1, \beta_2, \dots, \beta_m) = \sum_{i=1}^n [y_i - \beta_1 x_1(u_i) - \beta_2 x_2(u_i) - \dots - \beta_m x_m(u_i)]^2 \quad (6)$$

则

$$Q(\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_m) = \min Q(\beta_1, \beta_2, \dots, \beta_m) \quad (7)$$

注意到  $Q(\beta_1, \beta_2, \dots, \beta_m)$  是非负二次式，可微，由多元函数取得极值的必要条

件可得  $\frac{\partial Q}{\partial \beta_j} = 0$  ( $j = 1, 2, \dots, m$ )，即

$$\sum_{i=1}^n [y_i - \hat{\beta}_1 x_1(u_i) - \hat{\beta}_2 x_2(u_i) - \dots - \hat{\beta}_m x_m(u_i)] x_j(u_i) = 0 \quad j = 1, 2, \dots, m \quad (8)$$

整理得到

$$\left\{ \begin{array}{l} [\sum_{i=1}^n x_1^2(u_i)]\hat{\beta}_1 + [\sum_{i=1}^n x_1(u_i)x_2(u_i)]\hat{\beta}_2 + \dots + [\sum_{i=1}^n x_1(u_i)x_m(u_i)]\hat{\beta}_m \\ = \sum_{i=1}^n x_1(u_i)y_i \\ \dots \dots \dots \\ [\sum_{i=1}^n x_1(u_i)x_m(u_i)]\hat{\beta}_1 + [\sum_{i=1}^n x_2(u_i)x_m(u_i)]\hat{\beta}_2 + \dots + [\sum_{i=1}^n x_m^2(u_i)]\hat{\beta}_m \\ = \sum_{i=1}^n x_m(u_i)y_i \end{array} \right. \quad (9)$$

将  $\hat{\beta}$  带入模型(1)中,得模型的估计:  $\hat{Y} = X^T \hat{\beta}$ 。

### 5.1.3 问题一模型的求解

#### 5.1.3.1 AQI 中 6 个基本监测指标间的相关性分析模型的求解

利用 SPSS 软件<sup>[1]</sup>中的 Correlate 菜单能对变量间的关系进行很好的相关性分析,对此利用已知数据对 AQI 中 6 个基本监测指标间的关系进行相关性分析,可得到以下结果:

相关系数<sup>a</sup>

		PM2.5	二氧化硫	二氧化氮	可吸入颗粒物	一氧化碳	臭氧
PM2.5	Pearson Correlation	1	.724**	.732**	.779**	.822**	-.355**
	Sig. (2-tailed)		.000	.000	.000	.000	.000
二氧化硫	Pearson Correlation	.724**	1	.805**	.678**	.657**	-.182**
	Sig. (2-tailed)	.000		.000	.000	.000	.005
二氧化氮	Pearson Correlation	.732**	.805**	1	.727**	.625**	-.065
	Sig. (2-tailed)	.000	.000		.000	.000	.323
可吸入颗粒物	Pearson Correlation	.779**	.678**	.727**	1	.586**	-.069
	Sig. (2-tailed)	.000	.000	.000		.000	.295
一氧化碳	Pearson Correlation	.822**	.657**	.625**	.586**	1	-.383**
	Sig. (2-tailed)	.000	.000	.000	.000		.000
臭氧	Pearson Correlation	-.355**	-.182**	-.065	-.069	-.383**	1
	Sig. (2-tailed)	.000	.005	.323	.295	.000	

\*\* . Correlation is significant at the 0.01 level (2-tailed).

a. Listwise N=236

从分析结果表中,可以看出 PM2.5 与一氧化碳 CO 间相关系数为 0.822,呈显著正相关,且相关系数最大;PM2.5 与二氧化硫  $SO_2$  间相关系数为 0.724,呈显著正相关;PM2.5 与二氧化氮  $NO_2$  间相关系数为 0.732,呈显著正相关;PM2.5 与可吸入颗粒物 PM10 间相关系数为 0.779,也呈显著正相关;PM2.5 与臭氧  $O_3$  间相关系数为-0.355,呈负相关。

#### 5.1.3.2 PM2.5 与其它 5 项分指标间的两两回归分析模型的求解

##### Step1: 数据预处理

首先将缺省值运用 SPSS 软件进行一一剔除。

##### Step2: 运用 MATLAB 软件对其进行数据拟合分析(程序见附录 8.1.1)

分别将二氧化硫  $SO_2$ 、二氧化氮  $NO_2$ 、可吸入颗粒物 PM10、臭氧  $O_3$  和一氧化碳 CO 等 5 项指标作为自变量,PM2.5 作为因变量,利用 MATLAB 软件编程,

分别拟合出其线性模型、二次模型、三次模型、复合模型、指数模型、幂模型等多种曲线，得到各种模型拟合结果和拟合曲线。

**Step3: 比较分析，确定最佳拟合模型**

由于回归方程的拟合优度  $R^2$  越接近 1，拟合效果越好，对此通过以上指标对各种模型拟合结果进行对比分析后，发现 PM2.5 与二氧化硫  $SO_2$ 、二氧化氮  $NO_2$ 、可吸入颗粒物 PM10、臭氧  $O_3$  和一氧化碳 CO 等 5 项指标间的两两回归模型拟合效果不是很好，其中较好的为 PM2.5 分别 PM10 和一氧化碳 CO 的拟合，而拟合出的三次模型其  $R^2 = 80\%$ ，为最佳拟合模型，这里给出最好的两组拟合，即 PM2.5 分别 PM10 和一氧化碳 CO 的拟合，即：

**(1) PM2.5 与 PM10 的拟合效果图、拟合系数、拟合模型及残差分析**

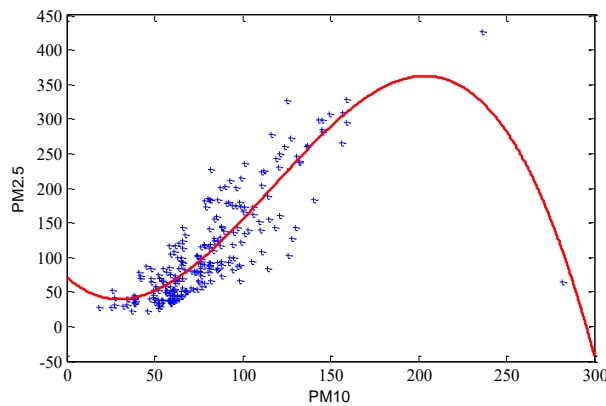


图 3 PM2.5 与 PM10 的拟合效果图

表 1 PM2.5 的含量随 PM10 含量变化的拟合方程系数估计结果

拟合系数	参数估计值	参数置信区间
$\beta_0$	72.0901	[29.2729 114.9073]
$\beta_1$	-2.2662	[-3.5294 -1.0029]
$\beta_2$	0.0431	[0.0323 0.0540]
$\beta_3$	-0.0001	[-0.0001 -0.0001]
$R^2 = 0.800, F = 244.4, p < 0.0001, S^2 = 1361.6$		

其中， $R$  表示相关系数， $F$  表示统计量值， $p$  表示统计量所对应的概率。由表中的结果可见， $R^2$  拟合优度为 80%。

PM2.5 与 PM10 的拟合模型为：

$$y = -0.0001x^3 + 0.0431x^2 - 2.2662x + 72.0901 \quad (10)$$

分析其残差，残差图如下图所示：

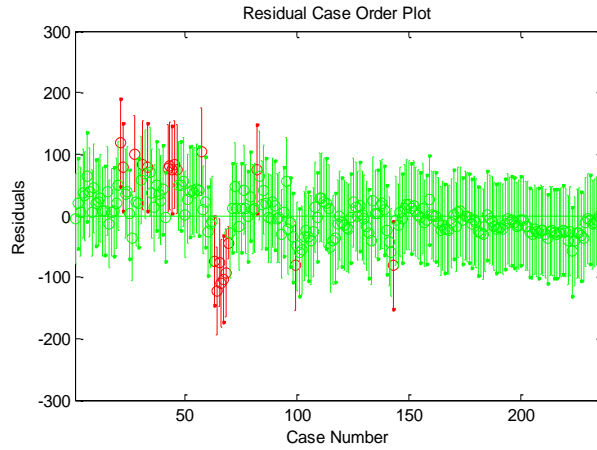


图 4 PM2.5 与 PM10 拟合的残差分析图

(2) PM2.5 与 CO 的拟合效果图、拟合系数、拟合模型及残差分析

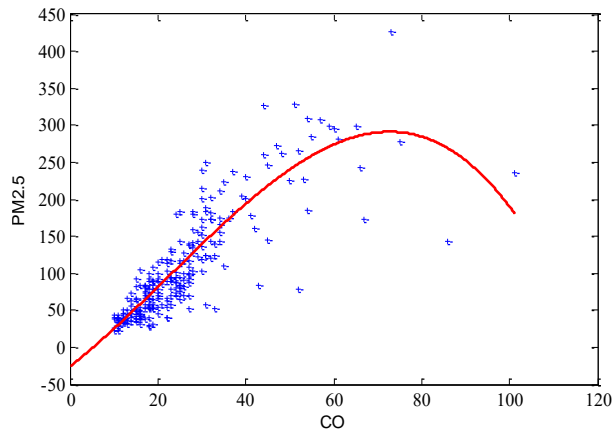


图 5 PM2.5 与 CO 的拟合效果图

表 2 PM2.5 的含量随 CO 含量变化的拟合方程系数估计结果

拟合系数	参数估计值	参数置信区间
$\beta_0$	-26.4116	[-59.3131 6.4900]
$\beta_1$	4.6251	[1.7723 7.4779]
$\beta_2$	0.0527	[-0.0171 0.1226]
$\beta_3$	-0.0008	[-0.0013 -0.0003]
$R^2 = 0.800$ $F = 243.1$ $p < 0.0001$ $S^2 = 1367.1$		

其中， $R$  表示相关系数， $F$  表示统计量值， $p$  表示统计量所对应的概率。由表中的结果可见， $R^2$  拟合优度为 80%。

PM2.5 与 CO 的拟合模型为：

$$y = -0.0008x^3 + 0.0527x^2 + 4.6251x - 26.4116 \quad (11)$$

分析其残差，残差图如下图所示：

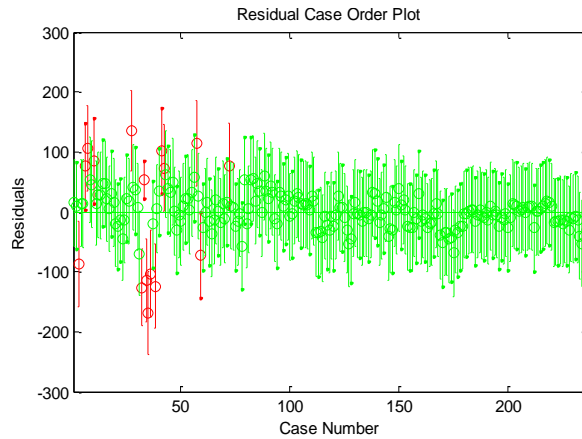


图 6 PM2.5 与 CO 拟合的残差分析图

### 5.1.3.3 PM2.5 与其它 5 项分指标间的多元线性回归模型的求解

**Step1:** 根据题目所给的数据，提取出回归变量和因变量，并用相应的字母表示。

**Step2:** 运用最小二乘估计法，利用 MATLAB 软件编程（程序见附录 8.1.1），计算回归系数如表所示。

表 3 PM2.5 与其它 5 项分指标间的多元线性回归方程系数估计结果

回归系数	参数估计值	参数置信区间
$\beta_0$	-32.9559	[-47.0906 -18.8213]
$\beta_1$	0.1733	[-0.2300 0.5765]
$\beta_2$	0.4417	[0.1369 0.7466]
$\beta_3$	0.8820	[0.7005 1.0634]
$\beta_4$	2.2071	[1.7865 2.6278]
$\beta_5$	-0.3339	[-0.4642 -0.2035]
$R^2 = 0.8366, F = 235.543, p < 0.0001, S^2 = 933.55$		

其中， $R$  表示相关系数， $F$  表示统计量值， $p$  表示统计量所对应的概率。由上表中的结果可见， $R^2$  不是很好，再分析其残差，残差图如图 7 所示。

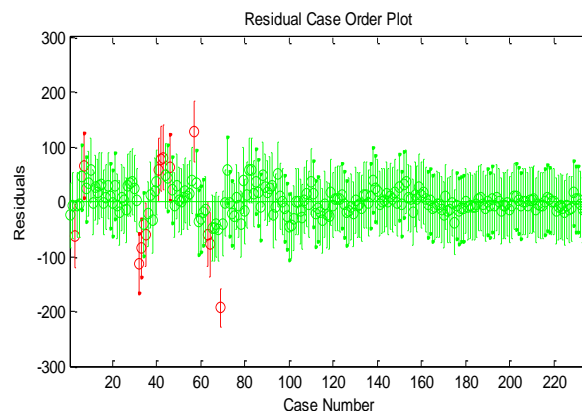


图 7 PM2.5 与其它 5 项分指标间多元线性回归方程的残差检验图

残差图可以看出，若所有数据的残差都包含零，且显著性概率  $P < 0.01$ ，回归效果显著。如果某个数据的残差不包含零，则常把它视为异常值，在回归中应把它剔除，再进行回归。从该残差图中可以看出，异常值较多，因此可将异常点剔除，具体的异常点如表 4 所示。



表 4 PM2.5 与其它 5 项分指标间回归方程检验所得异常点

3	7	32	33	35	41	42
43	46	57	63	64	69	

将表中的 13 个异常点数据剔除后,再次进行回归分析,利用 MATLAB 软件编程,可得回归分析结果如表所示。

表 5 第一次剔除异常点后回归方程系数估计结果表

回归系数	回归系数估计值	回归系数置信区间
$\beta_0$	-39.3207	[-48.8863 -29.7551]
$\beta_1$	-0.0313	[-0.3190 0.2565]
$\beta_2$	0.1022	[-0.1287 0.3331]
$\beta_3$	1.2876	[1.0955 1.4798]
$\beta_4$	2.3779	[1.9966 2.7592]
$\beta_5$	-0.3613	[-0.4521 -0.2705]
$R^2 = 0.9250 \quad F = 535.592 \quad p < 0.0001 \quad S^2 = 407.791$		

相应的残差图如图 8 所示

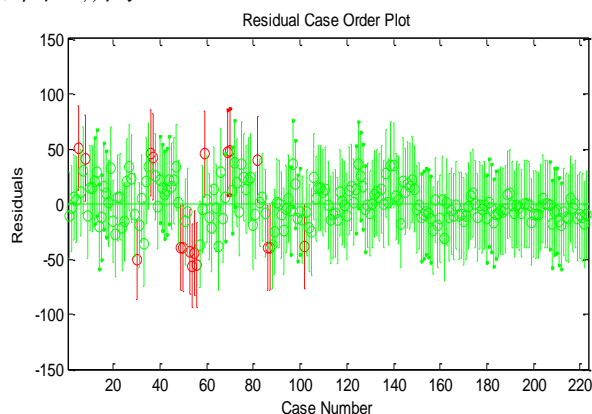


图 8 第一次剔除异常点后回归方程的残差检验图

由回归分析结果及残差图可看出,回归效果仍然不理想,还是有很多异常点,再次将这些异常点删除后,进行回归分析,仍然存在异常值,然后再删除,经过多次循环,得到最终回归系数估计值、置信区间、检验统计量。如表 6 所示

表 6 PM2.5 与其它 5 项分指标间最终回归方程系数

回归系数	参数估计值	参数置信区间
$\beta_0$	-45.3727	[-51.8182 -38.9271]
$\beta_1$	-0.1171	[-0.3130 0.0788]
$\beta_2$	0.3253	[0.1660 0.4847]
$\beta_3$	1.2216	[1.0818 1.3614]
$\beta_4$	2.3087	[1.9967 2.6208]
$\beta_5$	-0.3458	[-0.4091 -0.2824]
$R^2 = 0.971, F = 1199, p < 0.0001, S^2 = 151.72$		

最终残差图如图 9 所示

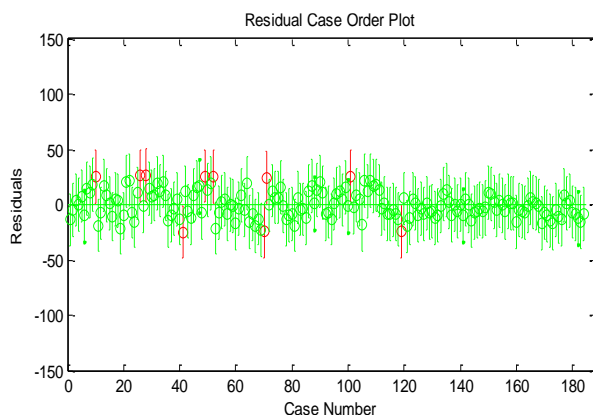


图 9 PM2.5 与其它 5 项分指标间多元线性回归方程最终残差检验图

#### 5.1.4 问题一结果的分析及验证

##### (1) AQI 中 6 个基本监测指标间的相关性分析结果

通过利用 SPSS 软件对 AQI 中 6 个基本监测指标间的相关性分析模型进行求解后,从其分析结果表中,可以看出 PM2.5 与一氧化碳 CO 间相关系数为 0.822,呈显著正相关,且相关系数最大;PM2.5 与二氧化硫  $SO_2$  间相关系数为 0.724,呈显著正相关;PM2.5 与二氧化氮  $NO_2$  间相关系数为 0.732,呈显著正相关;PM2.5 与可吸入颗粒物 PM10 间相关系数为 0.779,也呈显著正相关;PM2.5 与臭氧  $O_3$  间相关系数为-0.355,呈负相关。

##### (2) PM2.5 与其它 5 项分指标间的两两回归分析结果

通过运用 MATLAB 软件<sup>[7]</sup>对 PM2.5 与其它 5 项分指标间的两两回归分析模型的求解后,发现 PM2.5 与二氧化硫  $SO_2$ 、二氧化氮  $NO_2$ 、可吸入颗粒物 PM10、臭氧  $O_3$  和一氧化碳 CO 等 5 项指标间的两两回归模型拟合效果不是很好,其中较好的为 PM2.5 分别 PM10 和一氧化碳 CO 的拟合,而拟合出的三次模型其  $R^2 = 80%$ ,为最佳拟合模型。

	拟合模型	拟合优度
PM2.5 与 PM10 拟合	$y = -0.0001x^3 + 0.0431x^2 - 2.2662x + 72.0901$	80%
PM2.5 与 CO 拟合	$y = -0.0008x^3 + 0.0527x^2 + 4.6251x - 26.4116$	80%

##### (3) PM2.5 与其它 5 项分指标间的多元线性回归分析结果

通过运用最小二乘估计法,利用 MATLAB 软件编程(程序见附录 8.1.1),对 PM2.5 与其它 5 项分指标间的多元线性回归模型进行求解后,得到其对异常点逐次剔除后的 PM2.5 与其它 5 项分指标间多元线性回归方程为:

	对异常点逐次剔除后 PM2.5 与其它 5 项分指标间多元线性回归方程	拟合优度
原始	$y = 0.1733x_1 + 0.4417x_2 + 0.8820x_3 + 2.2071x_4 - 0.3339x_5 - 32.9559$	83.66%
首次剔除	$y = -0.0313x_1 + 0.1022x_2 + 1.2876x_3 + 2.3779x_4 - 0.3613x_5 - 39.3207$	92.5%
⋮	⋮	⋮
最终	$y = -0.1171x_1 + 0.3253x_2 + 1.2216x_3 + 2.3087x_4 - 0.3458x_5 - 45.3727$	97.1%

**(4) PM2.5 与其它 5 项分指标间的多元线性回归方程最终残差检验**

其 PM2.5 与其它 5 项分指标间多元线性回归方程最终残差检验图如图 10 所示，从残差检验图可以看出，所有数据的残差都基本包含零，且显著性概率  $p < 0.0001$ ，回归效果显著。

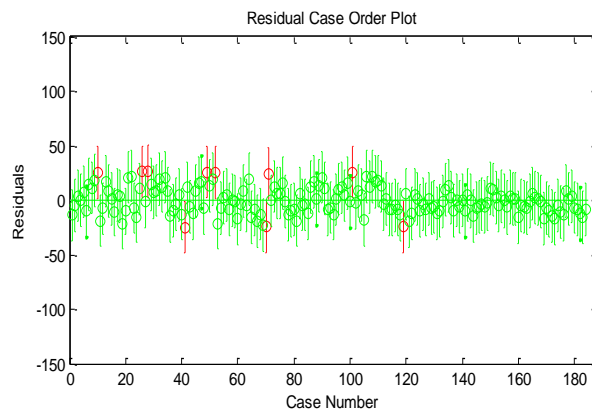


图 10 PM2.5 与其它 5 项分指标间多元线性回归方程最终残差检验图

**5.2 问题二：PM2.5 的分布与演变及应急处理**

**5.2.1 问题二第 1 小问模型的建立与求解**

**5.2.1.1 问题二第 1 小问的分析**

该问题要求描述该地区内 PM2.5 的时空分布及其规律，并结合环境保护部新修订的《环境空气质量标准》分区进行污染评估。对此，我们可首先可研究 PM2.5 随时间变化规律，对其可先绘制散点图进行发现大体趋势，其次，可以采用三次样条函数对其 PM2.5 随时间变化进行插值；然后，对其 PM2.5 空间分布规律进行研究，这里可采用基于 Shepard 二维插值法进行研究；最后，可建立分区污染评估模型再对其分区进行污染评估。

**5.2.1.2 基于三次样条插值法的 PM2.5 随时间变化模型**

**Step1: 各监测站 PM2.5 随时间变化的散点描述图**

附件 1 给出了西安市 13 个监测站及全市平均 PM2.5 含量（分指数）在一定时期内的数据，为了寻求该地区 PM2.5 含量的时空分布，首先绘制各监测站分别随时间变化的散点图，以寻找 PM2.5 的时间分布规律。

运用 MATLAB 软件可绘制出各监测站及全市 PM2.5 含量分别随时间变化的

散点图（程序见附录 8.1.2），由于其大致规律一致，这里选取了草滩、长安区、高压开关厂和全市平均的 PM2.5 含量分别随时间变化的散点图如图 11 所示。

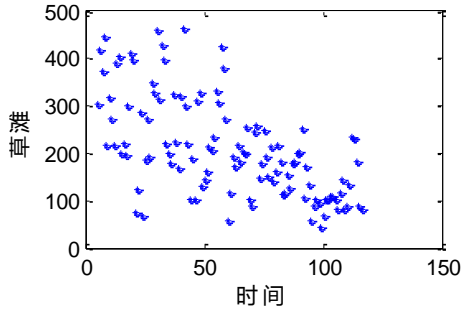


图 11(a) 草滩 PM2.5 含量与时间散点图

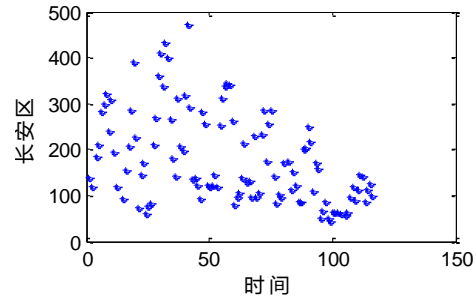


图 11(b) 长安区 PM2.5 含量与时间散点图

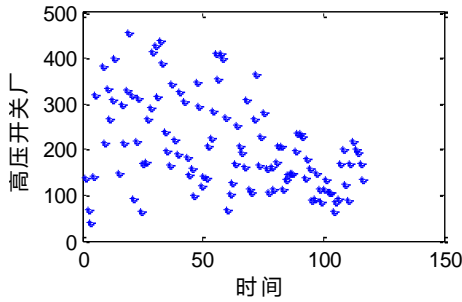


图 11(c) 高压开关厂 PM2.5 含量与时间散点图

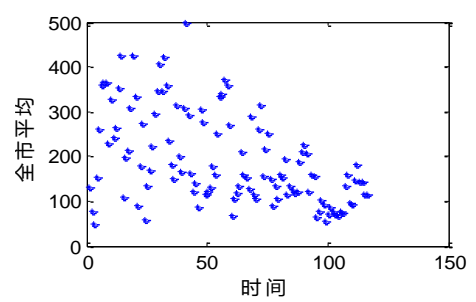


图 11(d) 全市平均 PM2.5 含量与时间散点图

图 11 各监测站及全市平均 PM2.5 含量与时间的散点图

从图 11 中可以看出，西安市各个监测站 PM2.5 的含量（分指数）随时间变化趋势大致相同，即都为随着时间 PM2.5 含量大致呈减少趋势，而且都与全市 PM2.5 平均含量（分指数）大体一致，因此可用全市 PM2.5 平均含量（分指数）随时间变化的规律描述各个监测站 PM2.5 含量（分指数）的时间分布。

### Step2: 基于三次样条函数的全市 PM2.5 随时间变化的插值曲线

由前述分析可知，全市 PM2.5 平均含量（分指数）随时间的变化即描述了各监测站 PM2.5 的时间分布规律。最理想的情况是寻找一个解析函数表达式精确的刻画 PM2.5 含量的变化规律，但实际上寻找这样一个拟合精度非常高的函数是很困难的，因此本文采用一维数值插值方法描述全市平均 PM2.5 含量的时间分布。

插值是由已知离散因变量的值来估计未知的中间值的方法，其基本思想就是构造一个简单函数  $y = s(x)$  作为  $f(x)$  的近似表达式，以  $s(x)$  的值作为  $f(x)$  的近似值，而且要求  $s(x)$  在给定点  $x_i$  的取值相同，即  $s(x_i) = f(x_i)$ ，通常称  $s(x)$  为  $f(x)$  的插值函数。

考虑到 PM2.5 应是随时间连续变化的，为了提高精度，我们采用三次样条函数进行插值，对于  $[a, b]$  上的划分  $\Delta: a = x_0 < x_1 < \dots < x_n = b$ ，则有三次样条函数

$$\begin{cases} s(x) = \alpha_0 + \alpha_1 x + \frac{\alpha_2}{2!} x^2 + \frac{\alpha_3}{3!} x^3 + \sum_{j=1}^{n-1} \frac{\beta_j}{3!} (x - x_j)_+^3 \in S_p(\Delta, 3) \\ (x - x_j)_+^3 = \begin{cases} (x - x_j)^3, & x \geq x_j \\ 0, & x < x_j \end{cases} \quad (j = 1, 2, \dots, n-1) \end{cases} \quad (12)$$

其中,  $S_p(\Delta, k)$  为  $k$  次样条函数空间。

利用 MATLAB 软件编程 (程序见附件 8.1.2), 对全市平均 PM2.5 含量 (分指数) 的实测数据进行插值, 并绘出全市平均 PM2.5 含量 (分指数) 随时间变化的一维插值图, 如图 12 所示。

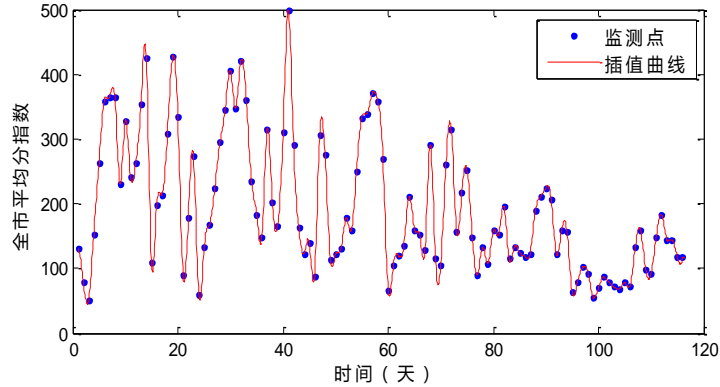


图 12 全市平均 PM2.5 含量 (分指数) 插值曲线

### 5.2.1.3 基于 Shepard 二维插值法的 PM2.5 空间分布模型

#### Step1: 数据预处理及定性分析

首先对全市 13 个监测点 PM2.5 含量在 2013 年 1 月 1 日~2013 年 4 月 26 日时间段内求平均值得到各监测点 PM2.5 平均含量, 即有

$$c_i = \frac{1}{n_i} \sum_{j=1}^{n_i} x_{ij} \quad (i=1, 2, \dots, 13; j=1, 2, \dots, n_i) \quad (13)$$

其中,  $c_i$  —— 第  $i$  个监测站的 PM2.5 平均含量 (分指数);

$x_{ij}$  —— 第  $i$  个监测站在第  $j$  天的实测值;

$n_i$  —— 第  $i$  个监测站实测值个数。

用 excel 软件按上式计算, 可得各个监测站的 PM2.5 平均含量, 如图 13 所示。

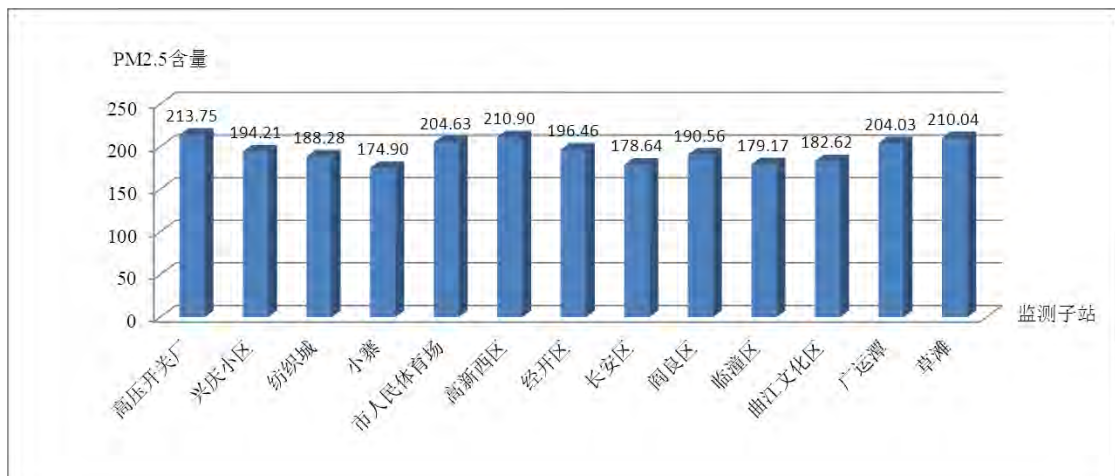


图 13 各监测站 PM2.5 平均含量柱形图

从图可直观看出, 高压开关厂的 PM2.5 含量是 13 个监测站中最大的, 市人民体育场、广运潭、草滩 PM2.5 含量也较高; 小寨、临潼区的 PM2.5 含量较低。

### Step2: 确定各监测站平面坐标

为了建立三维空间坐标系，可以 13 个监测站中的任意一个为  $xoy$  平面的坐标原点，本文设市人民体育场的位置坐标为  $xoy$  坐标原点，即  $(0,0)$ ，而  $z$  轴表示 PM2.5 平均含量。通过 Google 地图的测距功能，可得该空间坐标系下的三维坐标，如表所示。

表 8 各监测站空间坐标值

监测站	坐标 (单位: km)	监测站	坐标 (单位: km)
高压开关厂	(-6.1, -0.33, 213.754)	长安区	(-4.4, -12.4, 178.64)
兴庆小区	(2.8, -1.2, 194.208)	阎良区	(25.1, 43.6, 190.559)
纺织城	(9.7, -0.425, 188.279)	临潼区	(23.8, 10.8, 179.172)
小寨	(-0.7, -5.2, 174.901)	曲江文化区	(2.1, -6.7, 182.618)
市人民体育场	(0, 0, 204.63)	广运潭	(9.1, 5, 204.028)
高新西区	(-6.2, -4.3, 210.901)	草滩	(-8.2, 11, 210.036)
经开区	(-1.5, 8.4, 196.455)		

### Step3: 建立 Shepard 二维插值模型

根据题干与附件的相关数据，当给定城区内一点空间坐标  $(x,y,z)$ ，即可唯一确定该点 PM2.5 含量。故本问题应该建立 PM2.5 浓度（分指数） $C$  与空间坐标点  $(x,y,z)$  的对应函数关系。即

$$C = f(x,y,z) \quad (14)$$

又依据题目，当给出空间一点的平面坐标，便可以唯一的确定该点下地表的海拔高度，因此可以建立海拔高度  $z$  关于空间坐标的函数关系

$$z = H(x,y) \quad (15)$$

综合上述二式，最终可以建立空间任意点平面坐标和该空间点 PM2.5 浓度（分指数）的一一对应的函数关系

$$C = f(x,y,z(x,y)) = f(x,y) \quad (16)$$

在实际推断上述解析表达式时，由于三维曲面拟合相当复杂，很难求得函数的精确解析表达式，故本文利用数值插值的方法，定量地给出 PM2.5 浓度（分指数）的空间分布。

在本问中，由于要建立 PM2.5 浓度  $C$ （分指数）与水平面坐标  $x$ 、 $y$  两个变量之间的函数关系，故需采用二维插值法对其进行求解。

在此，我们应用 Shepard 方法进行插值计算。Shepard 方法又称反距离加权平均法，该方法基本思想为，在非给定数据的点处，定义其函数值由已知数据点与该点距离作加权平均决定，即：

$$\begin{cases} r_k = \sqrt{(x-x_k)^2 + (y-y_k)^2} \\ F(x,y) = \begin{cases} f_k & r_k = 0 \\ \sum_{k=1}^N W_k(x,y) f_k & r_k \neq 0 \end{cases} \\ W_k = \frac{1}{r_k^2} / \sum_{k=1}^N \frac{1}{r_k^2} \end{cases} \quad (17)$$

其中， $r_k$ ——已知数据点与选定点距离；

$f_k$ ——点 $(x_k, y_k)$ 对应的数据值；

$W_k$ ——数据点与选定点的加权平均值。

#### Step4: PM2.5 空间分布规律

根据表的数据利用 MATLAB 软件编程（程序见附件），可绘出各监测站 PM2.5 含量随水平位置  $xy$  变化的三维空间分布图，如图 14 所示，相应的等高线图如图 15 所示。

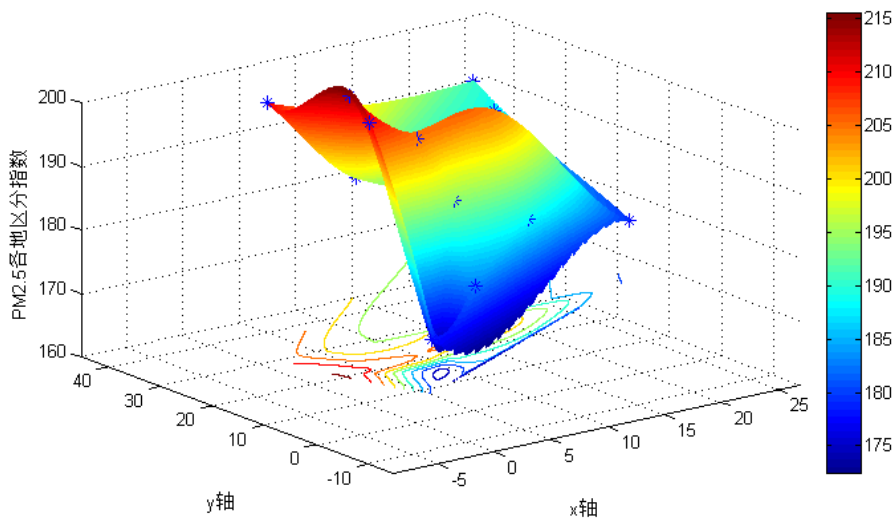


图 14 各监测站 PM2.5 含量二维插值曲面

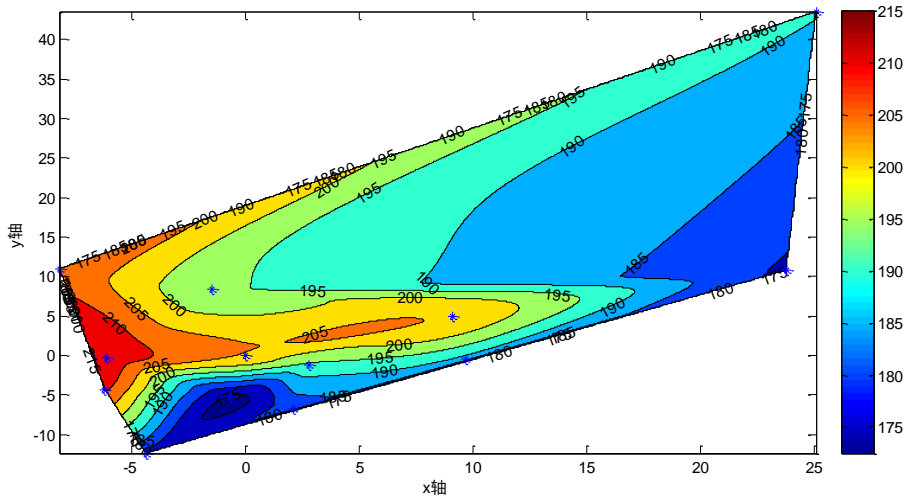


图 15 PM2.5 含量空间分布等高线

从图可直观看出，高压开关厂的 PM2.5 含量是 13 个监测站中最大的，市人民体育场、广运潭、草滩 PM2.5 含量也较高；小寨、临潼区的 PM2.5 含量较低。

#### 5.2.1.4 分区污染评估模型

##### Step1: 对 13 个监测点进行环境空气功能分区

根据环境保护部新修订的《环境空气质量标准》，将附件 2 中西安地区的 13 个监测点进行环境空气功能区分区如下表所示：

表 9 13 个监测点按环境空气功能的分区

一级区	曲江文化集团、广运潭、草滩
二级区	高压开关厂、兴庆小区、纺织城、小寨、高新西区、市人民体育场、经开区、长安区、阎良区、临潼区

##### Step2: 对 13 个监测点分别进行按区评估

##### 方法一：基于 PM2.5 期望值进行评估

根据《环境空气质量标准》中有关环境空气污染物基本浓度限值的说明，PM2.5 在一级区的 24 小时平均浓度限制值为  $35\mu\text{g}/\text{m}^3$ ，在二级区的 24 小时平均浓度限制值为  $75\mu\text{g}/\text{m}^3$ 。结合空气质量分指数计算公式，如下式所示。

$$IAQI_p = \frac{IAQI_{Hi} - IAQI_{Lo}}{BP_{Hi} - BP_{Lo}} (C_p - BP_{Lo}) + IAQI_{Lo} \quad (18)$$

其中， $IAQI_p$ ——污染物项目 P 的空气质量分指数；

$C_p$ ——污染物项目 P 的质量浓度指数；

换算出 PM2.5 在一级区的 24 小时平均空气质量分指数限制值为 50，在二级区的 24 小时平均空气质量分指数限制值为 100。

一级区污染评估模型：

$$n_1 = \frac{IAQI_{PM2.5}}{50} \quad (19)$$

二级区污染评估模型：

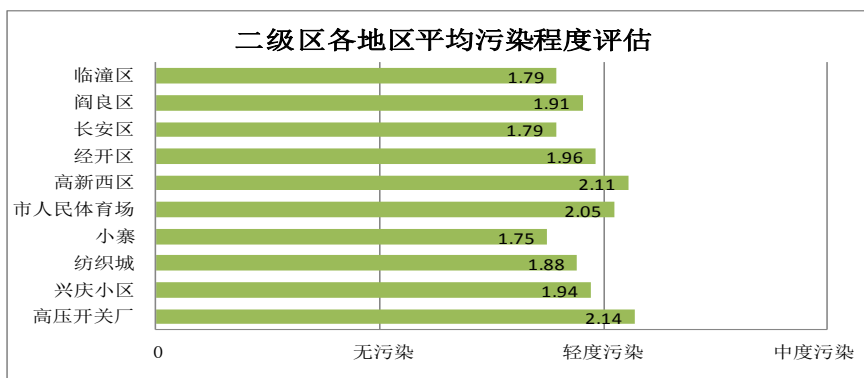
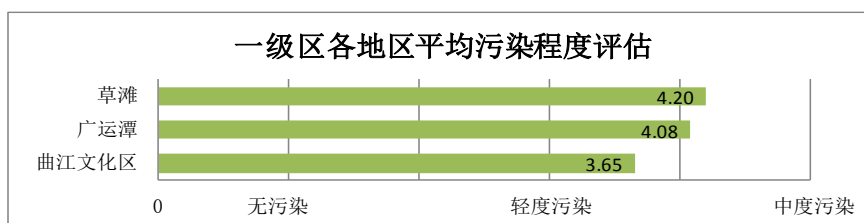
$$n_2 = \frac{IAQI_{PM2.5}}{100} \quad (20)$$



在此基础上，以计算得到的  $n$  值为依据，并结合环境空气质量评估标准给出分区进行污染评估的标准如下表：

一级区污染等级划分 ( $n_1$ )		二级区污染等级划分 ( $n_2$ )	
0—1	无污染	0—1	无污染
1—3	轻度污染	1—2	轻度污染
3—5	中度污染	2—3	中度污染
5—7	重度污染	3—4	重度污染
7 以上	严重污染	4 以上	严重污染

其 13 个监测点评估结果如下图所示



### 方法二：基于 PM<sub>2.5</sub> 出现频数进行评估

此外，在总体评估的基础上，为了更好地评估各监测点的实况，我们还对其 PM<sub>2.5</sub> 在不同等级下出现的频数进行了统计分析，得到了一级区和二级区的评估结果（见附录 ），并进一步统计得到了 13 个监测点在各自的环境空气功能区的整体污染评估结果况如下图所示：

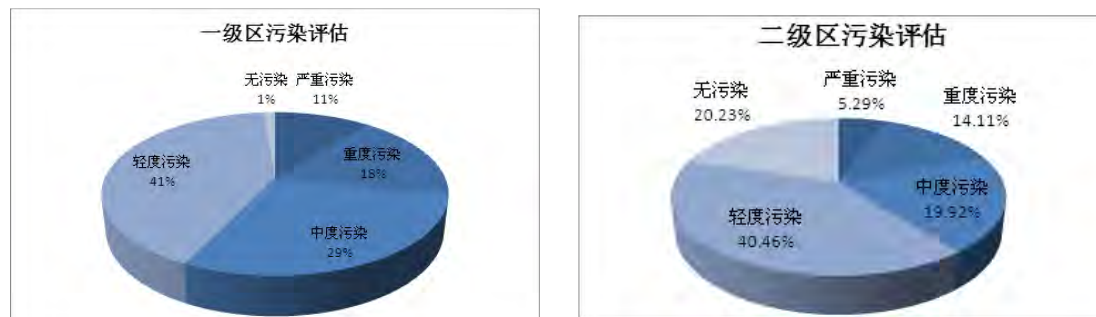


图 16 两类区域监测点在各自的环境空气功能区的整体污染评估结果

## 5.2.2 问题二第 2 小问模型的建立与求解

### 5.2.2.1 问题二第 2 小问的分析

问题要求建立能够刻画该地区 PM2.5 的发生和演变（扩散与衰减等）规律的数学模型，合理考虑风力、湿度等天气和季节因素的影响，并利用该地区的数据进行定量与定性分析。对此，可先利用该地区的数据进行 PM2.5 与风力、湿度等天气和季节因素间的多元回归分析，即建立 PM2.5 浓度的估算模型；然后，考虑 PM2.5 自身的扩散问题，可利用大气污染物的扩散原理建立 PM2.5 扩散的偏微分方程模型。

### 5.2.2.2 PM2.5 与气象因素间的多元回归模型的建立与求解

#### Step1: 数据预处理及相关性分析

PM2.5 扩散时，除了自身扩散机理外，其还受一些气象因素的影响，如空气湿度、温度、风速、风压等天气和季节因素的影响<sup>[5]</sup>。对于季节，我国对季节的划分：将一年分为四个季节，即春、夏、秋、冬。农历一至三月为春季；四至六月为夏季；七至九月为秋季；十至十二月为冬季。

对于湿度，与天气状况有着密切的关系，本文在查阅气象局信息与大量文献后，将数据中的天气状况近似做了如表所示的转换。

表 10 天气状况与相对湿度转换表

天气状况	晴	多云	阴	小雪	雨夹雪	小雨
相对湿度	30%	40%	50%	60%	70%	80%

对于风速，根据风速与风级的对照表，由于基本都为小于等于三级风，若为平均值或等级值将为定常数不具有统计作用，对此我们另外在相关气息官网上查阅到的该地区风速情况的具体数据以及历史数据，对此为风速数据。

对于风压，可以利用风压的贝努利公式得到：

$$w = \frac{1}{2} \rho v^2 = \frac{\gamma}{2g} v^2 = \frac{v^2}{1630} (kN/m^2).$$

其中  $w$  为风压， $v$  为风速， $\rho = 1.25 kg/m^3$  为标准的空气密度， $\gamma = 0.012018 kN/m^3$  为空气单位体积的重力， $g = 9.8 m/s^2$  重力加速度。

对此，可首先绘制 PM2.5 含量与各气象因素的散点图，包括白天和夜晚，发现白天和夜晚的变化情况基本一致，这里给出其白天的 PM2.5 含量与各气象因素的散点图如下图所示。

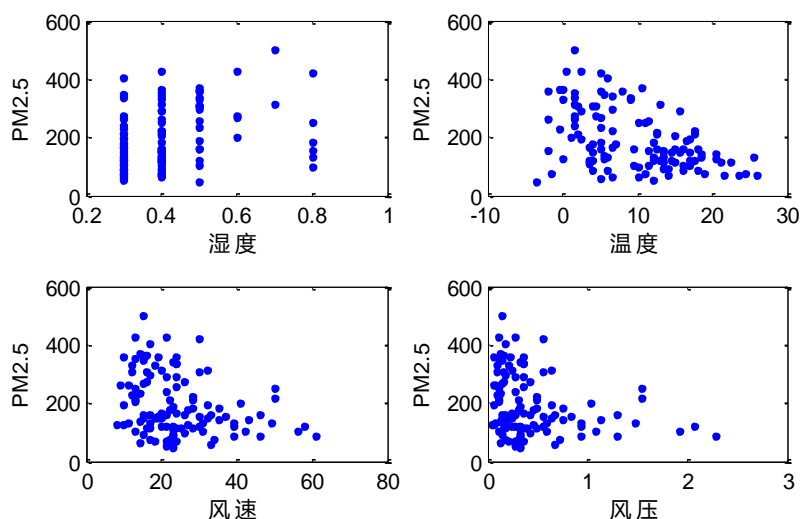


图 17 白天时 PM2.5 含量与各气象因素的散点图

**Step2: PM2.5 与气象因素间多元线性回归模型的建立**

从上图中可看出, PM2.5 扩散时, 其含量与湿度、温度、风速、风压等因素均有关。因此, 由 2013 年 1 月 1 日~2013 年 4 月 26 日所给数据, 分为冬季白天、冬季夜晚、春季白天、春季夜晚四种情形进行多元线性回归分析。

假设其多元线性回归模型为:

$$y = \beta_1 x_1(u) + \beta_2 x_2(u) + \beta_3 x_3(u) + \beta_4 x_4(u) + \varepsilon \quad (21)$$

其中, 自变量  $x_1(u)$  为湿度,  $x_2(u)$  为温度,  $x_3(u)$  为风速,  $x_4(u)$  为风压, 因变量  $y$  为 PM2.5 的浓度 (分指数),  $\varepsilon$  为随机误差, 且服从于  $N(0, \sigma^2)$ 。

**Step3: PM2.5 与气象因素间多元线性回归模型的求解**

采用最小二乘估计法, 通过 MATLAB 编程 (程序见附录 8.1.2), 可计算出回归系数, 从而可得冬季白天、冬季夜晚、春季白天、春季夜晚四种情形的多元回归方程, 如下表所示。

表 11 PM2.5 与各气象因素间的回归方程表

季节昼夜	PM2.5 浓度 (分指数) 与各气象因素间的回归方程	R 统计量
冬季白天	$y = 34.434 + 567.6704x_1 + 6.9481x_2 + 2.6728x_3 - 305.0341x_4$	0.666
冬季夜晚	$y = 136.3687 + 197.5192x_1 + 1.9748x_2 + 6.8363x_3 - 361.636x_4$	0.3783
春季白天	$y = 147.9484 + 123.5419x_1 - 4.1541x_2 + 1.4811x_3 - 45.9246x_4$	0.416
春季夜晚	$y = 140.5836 + 128.1092x_1 - 4.1558x_2 + 2.0595x_3 - 64.1105x_4$	0.4096

从求解结果可知, PM2.5 浓度 (分指数) 是会受气象影响的。在冬季时, PM2.5 浓度 (分指数) 与湿度呈正相关, 与温度也为正相关, 且湿度比温度的影响更大, 与风压呈负相关。在春季时, PM2.5 浓度 (分指数) 与湿度任然呈正相关, 而与温度呈负相关, 与风压也呈负相关。

**5.2.2.3 PM2.5 扩散模型的建立与求解**

**Step1: 建立 PM2.5 在边界层中扩散的物理模型**

由于 PM2.5 是大气污染物中的一种, 对此它适用于所有大气污染物的扩散模型, 我们知道, 从地面到约 1000 米高度的大气边界层中发生的主要物理过程是地面与大气之间的质量, 动量及能量的交换, 其基本运动方程如下<sup>[6]</sup>:

质量守恒: 
$$\frac{\partial \rho}{\partial t} + \frac{(\rho u_i)}{\partial x_i} = 0$$

动量守恒: 
$$\frac{\partial u_i}{\partial x_i} + u_j \frac{\partial u_i}{\partial x_j} + 2\varepsilon_{ijk} \omega_j u_k = -\frac{1}{\rho} \frac{\partial \rho}{\partial x_i} - \frac{\Delta T}{T_0} g \delta_{i3} + \nu \frac{\partial^2 u_i}{\partial x_k \partial x_k} + \frac{\partial(-u'_j u'_j)}{\partial x_j}$$

$$\text{能量守恒: } \frac{\partial T}{\partial t} + u \frac{\partial T}{\partial x_i} = \left( \frac{k}{\rho c_p} \right) \frac{\partial^2 T}{\partial x_k \partial x_k} + \frac{\partial(-\theta u'_i)}{\partial x_i} + \frac{\phi}{\rho c_p}$$

### Step2: 建立 PM2.5 迁移、转化的基本模型

考虑到环境介质中会由于物理、化学、生物等作用引起 PM2.5 的增减, 在公式中附加一个源-汇项  $S(x, y, z, c, t)$  则可得到 PM2.5 迁移、转化的基本模型为:

$$\begin{aligned} & \frac{\partial c}{\partial t} + u_x \frac{\partial c}{\partial x} + u_y \frac{\partial c}{\partial y} + u_z \frac{\partial c}{\partial z} \\ & = \frac{\partial}{\partial x} \left( D_x \frac{\partial c}{\partial x} \right) + \frac{\partial}{\partial y} \left( D_y \frac{\partial c}{\partial y} \right) + \frac{\partial}{\partial z} \left( D_z \frac{\partial c}{\partial z} \right) + S(x, y, z, c, t) \end{aligned} \quad (22)$$

上式与环境流体介质的运动方程进行耦合运用就可模拟 PM2.5 在环境介质中的迁移、转化过程。虽然在理论上说模型是完整的, 但求解这些方程还很困难, 需要掌握不同时间各空间位置上的许多参数和模型系数的资料, 而这些资料同时又包含很多复杂的时间和空间函数关系。因此, 在不影响求解精度的情况下, 实际工作中需要根据现实条件和 PM2.5 特点作一些简化。

### Step3: 建立 PM2.5 扩散的偏微分方程模型

在无边界大气条件下, 可把 PM2.5 扩散模型看为一个稳定污染点源连续排放的扩散模型, 其可通过 PM2.5 迁移、转化的基本模型简化得到。

对此需有几个基本假设:

- (1) PM2.5 在烟羽或烟团的各端面上呈正态分布(高斯分布);
- (2) 整个研究空间内风速场是均匀稳定的;
- (3) PM2.5 污染源强是连续均匀的;
- (4) 整个扩散过程 PM2.5 污染物质质量守恒。

此外, 需将 PM2.5 视为定常连续点源, 对定常连续点源, 又有以下假设:

- (1) 流动是稳定的, 且有主导方向;
- (2) PM2.5 在大气中只有物理运动, 没有发生化学和生物变化;
- (3) 在所监测的范围内没有其他同类污染源和汇。

即有:

$$\begin{cases} u = u_x = \text{const}, & u_y = 0, & u_z = 0 \\ \frac{\partial c}{\partial t} = 0 \\ S(x, y, z, c, t) = 0 \end{cases} \quad (23)$$

则 PM2.5 扩散的偏微分方程模型可由 PM2.5 迁移、转化的基本模型变为:

$$u \frac{\partial c}{\partial x} = D_y \frac{\partial^2 c}{\partial y^2} + D_z \frac{\partial^2 c}{\partial z^2} \quad (24)$$

其已知边界条件为:

$$\begin{cases} x = y = z \rightarrow 0 \text{ 时}, & c \rightarrow \infty \\ x = y = z \rightarrow \infty \text{ 时}, & c \rightarrow 0 \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} u c u dy dz = Q \end{cases} \quad (25)$$

#### Step4: 求解 PM2.5 扩散的偏微分方程模型

利用 MATLAB 软件编程求解 PM2.5 扩散的偏微分方程模型可得浓度的解析解为:

$$c(x, y, z) = \frac{Q}{4\pi u(D_y D_z)^{1/2}} \exp\left[-\frac{u}{4x} \left(\frac{y^2}{D_y} + \frac{z^2}{D_z}\right)\right] \quad (26)$$

设:  $x = ut$ , 令  $2D_y t = \sigma_y^2, 2D_z t = \sigma_z^2$

则上述浓度的解析解可表达为:

$$c(x, y, z) = \frac{Q}{4\pi u \sigma_y \sigma_z} \exp\left[-\frac{1}{2} \left(\frac{y^2}{\sigma_y^2} + \frac{z^2}{\sigma_z^2}\right)\right] \quad (27)$$

其中,  $c(x, y, z)$ ——空间某一点 PM2.5 的浓度;

$u$ ——PM2.5 所处高度处的风速,单位  $m/s$

$\sigma_y$ ——垂直于主导风方向的横向扩散系数,单位  $m$

$\sigma_z$ ——铅直扩散系数,单位  $m$

$Q$ ——单位时间内的 PM2.5 排放量,单位  $mg/s$

### 5.2.3 问题二第 3 小问模型的建立与求解

#### 5.2.3.1 问题二第 3 小问的分析

问题要求假设该地区某监测点处的 PM2.5 的浓度突然增至数倍,且延续数小时,建立针对这种突发情形的污染扩散预测与评估方法。并以该地区 PM2.5 监测数据最高的一天为例,在全地区 PM2.5 浓度最高点处的浓度增至 2 倍,持续 2 小时,并利用该模型进行预测评估,给出重度污染和可能安全区域。对此,首先,我们针对这种情况建立其 PM2.5 污染扩散的预测模型;然后对预测模型中的参数进行参数估计;最后,利用 MATLAB 软件进行实例求解,并评估出重度污染和可能安全区域。

### 5.2.3.2 PM2.5 污染扩散的预测与评估模型的建立

#### Step1: 建立 PM2.5 在地面的浓度分布模型

PM2.5 在地面的浓度分布是实际大气影响评价所关心的问题的重要方面，其分布规律可由第二小问在  $z=0$  和 PM2.5 有效高度  $H_e$  时的情况下得到：

$$c(x, y, z, H_e) = \frac{Q}{4\pi u \sigma_y \sigma_z} \exp\left(-\frac{y^2}{2\sigma_y^2} - \frac{H_e^2}{2\sigma_z^2}\right) \quad (x > 0) \quad (28)$$

地面轴线是从 PM2.5 原点向下风向延伸的方向,此时在上式中令  $y=0$ , 可得轴线上的地面浓度为:

$$c(x, 0, 0, H_e) = \frac{Q}{\pi u \sigma_y \sigma_z} \exp\left(-\frac{H_e^2}{2\sigma_z^2}\right) \quad (29)$$

#### Step2: 建立 PM2.5 污染扩散的预测模型

最大落地浓度公式是最常用的公式之一，排放标准中的允许排放量和环境评价中需要预测的 1 h 浓度，通常都是利用最大落地浓度公式计算的。最大落地浓度发生在  $x$  轴线上,由  $x=ut$ ，令  $2D_y t = \sigma_y^2, 2D_z t = \sigma_z^2$ 。结合浓度分布模型可得：

$$c(x, 0, 0, H_e) = \frac{Q}{2\pi x (D_y D_z)^{1/2}} \exp\left(-\frac{u H_e^2}{4x D_z}\right) \quad (30)$$

要求  $x$  的轴线浓度最大值，将上式对  $x$  求导数，令其等于零，得地面最大落地浓度点距离为： $x_{cm} = \frac{u H_e^2}{4 D_z}$ 。将  $x=ut, 2D_y t = \sigma_y^2, 2D_z t = \sigma_z^2$  代入此公式中，就可

以计算出当  $x=x_{cm}$  时：

$$\sigma_z \Big|_{x=x_{cm}} = \frac{H_e}{\sqrt{2}} \quad (31)$$

再将扩散系数  $\sigma_y, \sigma_z$  表示成经验函数形式  $\sigma_y = \gamma_1 x^{\alpha_1}, \sigma_z = \gamma_1 x^{\alpha_2}$  并代入 PM2.5 浓度分布模型可得 PM2.5 地面浓度最大值为：

$$c(x, 0, 0, H_e) = \frac{2Q}{\pi e H_e^2 \sigma_y} = \frac{0.117Q}{u \sigma_y \sigma_z} \quad (32)$$

此时， $x_{cm} = \left(\frac{H_e^2}{4D_z}\right)^{\alpha_2}$

对于 PM2.5 在地面连续排放点源模型中，当有效源高  $H_e=0$  时，则有：

$$c(x, y, z, 0) = \frac{Q}{\pi u \sigma_y \sigma_z} \exp\left[-\frac{1}{2}\left(\frac{y^2}{\sigma_y^2} + \frac{z^2}{\sigma_z^2}\right)\right] \quad (33)$$

由此可见, PM2.5 地面源产生的浓度恰好是无限空间连续点源所产生浓度的二倍。而地面连续排放点源在地面上 PM2.5 任一点产生的浓度则为:

$$c(x, y, 0, 0) = \frac{Q}{\pi u \sigma_y \sigma_z} \exp\left(-\frac{y^2}{\sigma_y^2}\right) \quad (34)$$

### Step3: PM2.5 污染扩散预测模型的参数估计

PM2.5 在大气中的扩散与浓度分布是在湍流作用下形成的, 而湍流统计量与采样时间长短有关。由于大气湍流作用, 风速大小和方向随时在变化, 在短时间内风向变化角度小, 烟云散布范围窄,  $\sigma$  值小; 反之亦然。我们查阅《大气环境质量标准》发现“一次最大允许浓度”所对应的采样时间为 0.5h。

#### (1) 0.5h 取样时间时 $\sigma_y$ 、 $\sigma_z$ 的确定

PM2.5 扩散参数  $\sigma_y$ 、 $\sigma_z$  与水平距离  $x$  的关系为:  $\sigma_y = \gamma_1 x^{\alpha_1}$ ,  $\sigma_z = \gamma_1 x^{\alpha_2}$ 。对应于 0.5h 的取样时间,  $\sigma_y$ 、 $\sigma_z$  的确定按照表 (2 个) (见附录) 执行。根据中华人民共和国《制定地方大气污染物排放标准的技术方法》中有关扩散参数 ( $\sigma_y$ ,  $\sigma_z$ ) 的确定 (参考件), 这里我们选择稳定度等级为 C-D。

#### (2) 大于或等于 0.5h 取样时间时 $\sigma_y$ 、 $\sigma_z$ 的确定

当取样时间大于 0.5h 时需要对扩散参数与水平距离  $x$  之间的关系式进行修正, 对铅直方向扩散参数来说,  $\sigma_z$  不变, 而对  $\sigma_y$  来说, 回归指数  $\alpha_1$  保持不变,

而回归系数  $\gamma$  需满足公式  $\gamma_{1\tau_2} = \gamma_{1\tau_1} \left(\frac{\tau_2}{\tau_1}\right)^q$ , 因此横向扩散系数  $\sigma_y$  需要满足公式:

$$\sigma_{y\tau_2} = \sigma_{y\tau_1} \left(\frac{\tau_2}{\tau_1}\right)^q \quad (35)$$

其中,  $\tau_1$ 、 $\tau_2$  为取样时间

$q$  为时间稀释指数,  $t$  在 [0.5, 1] 时,  $q = 0.2$  在 [1, 100] 时  $q = 0.3$

$\sigma_{y\tau_1}$ 、 $\sigma_{y\tau_2}$  分别为对应取样时间为  $\tau_1$ 、 $\tau_2$  时的横向扩散系数

$\alpha_1$ 、 $\alpha_2$  分别为对应取样时间为  $\tau_1$ 、 $\tau_2$  时的横向扩散参数回归系数

利用上述估计模型和国标计算取样时间大于或等于 0.5h 的  $\sigma_{y\tau_2}$  或  $\gamma_{1\tau_2}$  时, 先根据取样时间计算 0.5h 取样时间的  $\sigma_y$  和  $\gamma_1$ , 以此作为  $\sigma_{y\tau_1}$  和  $\gamma_{1\tau_1}$  计算  $\sigma_{y\tau_2}$ 、 $\gamma_{1\tau_2}$ 。

### 5.2.3.3 PM2.5 污染扩散的预测与评估模型的求解

#### Step1: 确定 PM2.5 监测数据最高点

首先通过 MATLAB 软件可找出该地区 PM2.5 监测数据最高的监测点及监测日期, 如表 14 所示。

表 14 PM2.5 浓度最高的监测点、日期

最大监测点	高压开关厂	兴庆小区	小寨	人民体育场	高新西区
-------	-------	------	----	-------	------

监测日期	2013-1-14; 2013-2-10	2013-2-10	2013-2-10	2013-2-10	2013-2-10
------	-------------------------	-----------	-----------	-----------	-----------

从表中可看出，该市各监测点 PM2.5 浓度最高的日期均在 2013 年 2 月 10 日，其中，高压开关厂还有 2013 年 1 月 14 日这一天，因此，以 2013 年 2 月 10 日为准。在这些监测点中，高压开关厂 PM2.5 的 IAQI 平均值最大，因此，最终以高压开关厂在 2013 年 2 月 10 日这一天为例。

### Step2: 以扩散源为原点建立三维空间坐标系

将高压开关厂所在位置作为扩散源，并将其定为三维空间坐标的原点，利用坐标变换公式可将 5.2.1.3 中建立的  $xoy$  平面坐标系转换到新的坐标系中，即

$$\begin{cases} x' = x - a \\ y' = y - b \\ z' = z \end{cases} \quad (36)$$

其中， $(x, y, z)$  为旧空间坐标， $(x', y', z')$  为各监测点新空间坐标， $(a, b, z)$  为旧坐标系中高压开关厂所在位置坐标。各个监测点在新的坐标系里的位置坐标如表所示。

表 15 新坐标系中各监测点位置坐标

监测站	坐标 (单位: km)	监测站	坐标 (单位: km)
高压开关厂	(0, 0)	长安区	(1.7, -12.07)
兴庆小区	(8.9, -0.87)	阎良区	(31.2, 43.93)
纺织城	(15.8, -0.095)	临潼区	(29.9, 11.13)
小寨	(5.4, -4.87)	曲江文化区	(8.2, -6.37)
市人民体育场	(6.1, 0.33)	广运潭	(15.2, 5.33)
高新西区	(-0.1, -3.97)	草滩	(-2.1, 11.33)
经开区	(4.6, 8.73)		

### Step3: 对各个监测点 PM2.5 浓度的预测评估结果

由题意可知，高压开关厂 PM2.5 的初始浓度为  $1000 \mu\text{g}/\text{m}^3$ ，也即该点的浓度瞬时达到  $1000 \mu\text{g}/\text{m}^3$ ，因此，单位时间内 PM2.5 排放量  $Q$  可近似为  $Q=1000 \mu\text{g}/\text{s}$ 。将各监测点坐标代入扩散模型，通过 MATLAB 编程（程序见附录 8.1.2）可预测出各个监测点经过 2 小时后 PM2.5 的浓度值，并据此判断其污染程度，各个监测点 PM2.5 浓度预测值如表所示。

表 16 各监测点 PM2.5 浓度预测及污染状况

	高压开关厂	兴庆小区	纺织城	小寨	市人民体育场	高新西区	经开区
PM2.5 浓度 ( $\mu\text{g}/\text{m}^3$ )	1000.000	72.425	27.511	102.952	138.456	287.985	61.259
污染状况	严重污染	良	优	轻度污染	中度污染	严重污染	良
监测区	长安区	阎良区	临潼区	曲江文化区	广运潭	草滩	
PM2.5 浓度 ( $\mu\text{g}/\text{m}^3$ )	42.770	3.416	8.328	56.175	26.625	47.059	



污染状况	良	优	优	良	优	良	
------	---	---	---	---	---	---	--

从表中可以看出,污染源高压开关厂和高新西区属于严重污染区,兴庆小区、经开区、长安区、曲江文化区、草滩污染状况均为良,而纺织城、阎良区、临潼区、广运潭污染状况达到优。

#### Step4: 绘制 PM2.5 的扩散曲面及等值线

由西安市各监测点位置坐标及扩散 2 小时后的浓度值,通过 MATLAB 软件(程序见附录)可绘出高压开关厂 PM2.5 扩散三维图,如图 17 所示。

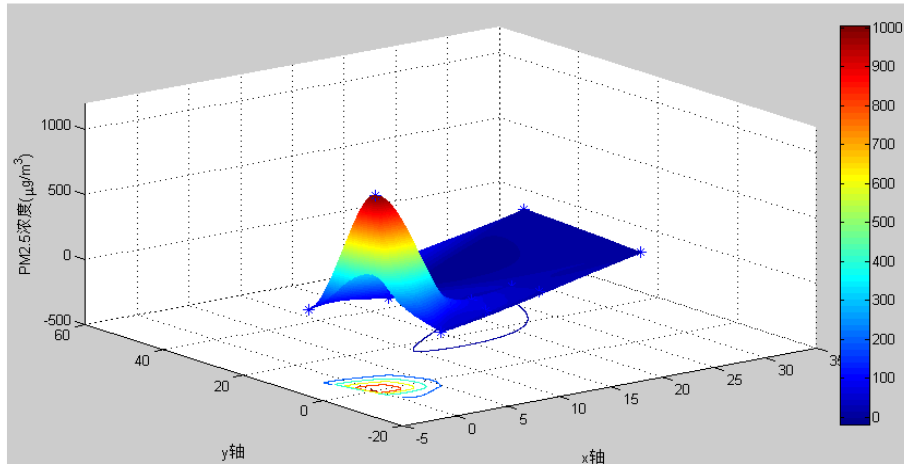


图 17 PM2.5 扩散曲面

从图中可以看出,在污染源高压开关厂处,PM2.5 浓度最高,并随着 XOY 平面上位置点的变化,PM2.5 的浓度值跟随着变化,且距原点越远的点,其 PM2.5 浓度越低,当经过一段时间后,各位置点 PM2.5 的浓度逐渐趋于稳定。

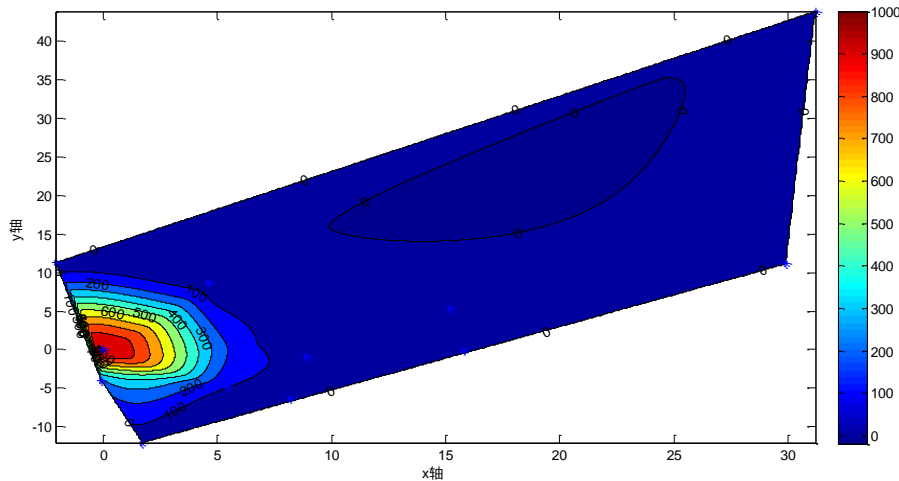


图 18(a) PM2.5 浓度等值曲线整体图

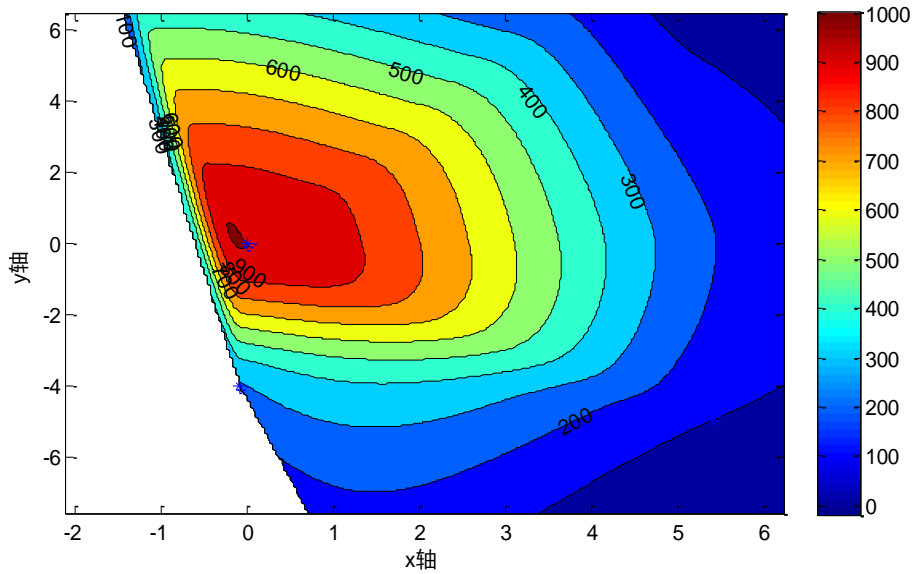


图 18 (b) PM2.5 浓度等值曲线局部放大图

图 18 PM2.5 浓度等值曲线图

图 18 为 PM2.5 扩散时其等值浓度曲线在 XOY 平面上的投影，直观地展现了 PM2.5 扩散时其浓度变化过程。

## 5.2.4 问题二第 4 小问

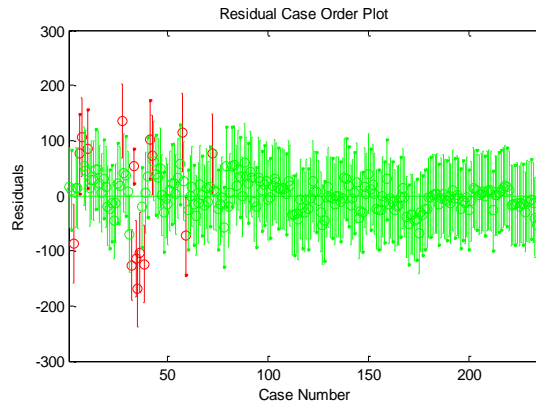
### 5.2.4.1 合理性检验

针对于第一小问，我们通过对各个监测点 PM2.5 随时间变化的分析，发现西安全市平均 PM2.5 含量可以大致反映该地区 PM2.5 含量的时间分布，并通过三次样条插值得到 PM2.5 的时间分布规律；进而，计算出各监测点 PM2.5 平均浓度，通过 Shepard 二维插值描述 PM2.5 的空间分布规律；最后，建立基于 PM2.5 期望值的分区评估模型。通过其插值图可以检验该方法是合理的。

此外，在对 13 个监测点分别进行按区评估时，用基于 PM2.5 期望值进行评估和基于 PM2.5 出现频数进行评估结果进行比较也可检验评估结果的合理性。

第二小问中，PM2.5 扩散时，除了自身扩散机理外，其还受一些气象因素的影响，如空气湿度、温度、风速、风压等天气和季节因素的影响。此问中，通过先考虑 PM2.5 扩散时受各种气象因素的影响，建立多元线性回归模型，得出 PM2.5 扩散时会受到一定气象因素的影响；然后通过机理分析，推导了 PM2.5 考虑风速影响的扩散方程。

对于多元线性回归模型，可用残差对其检验，若所有数据的残差都包含零，且显著性概率  $P < 0.01$ ，回归效果则显著。



残差检验图

第三小问中，全地区浓度最高处即高压开关厂的浓度突增至 2 倍，结合第二小问中的模型，预测出了 XOY 平面内任一点处 PM2.5 浓度值，进而找出重度污染区与可能安全区域。

对于其微分方程模型，可利用稳定性原理来判断其合理性，在求解过程中发现是合理的。

#### 5.2.4.2 PM2.5 的成因、演变等一般性规律

由 5.2.1 可知，PM2.5 浓度（分指数）是会受气象影响的。在冬季时，PM2.5 浓度（分指数）与湿度呈正相关，与温度也为正相关，且湿度比温度的影响更大，与风压呈负相关。在春季时，PM2.5 浓度（分指数）与湿度任然呈正相关，而与温度呈负相关，与风压也呈负相关。

此外，通过对本问的探索研究，可知 PM2.5 的形成与 SO2、NO2、CO 等具有很强的相关性，其扩散原理的基础是湍流扩散理论，而研究的方法也有多种，如梯度输送理论、相似理论和统计理论。

基于湍流统计理论体系的高斯扩散模型应用较多，本问即在此基础上考虑 PM2.5 特有的性质，并受其他气象因素影响时的扩散模型。其中的扩散参数  $\sigma_y$ 、 $\sigma_z$  是表示扩散范围及速率大小的特征量，也即正态分布函数的标准差。而  $\sigma_z$  的估计值与温度层结的关系很大，适用于近地源的 1km 以内的扩散。因此，污染物扩散参数的准确定量描述也是需要深入研究的课题。

### 5.3 问题三：空气质量的控制管理

#### 5.3.1 问题三第 1 小问模型的建立与求解

##### 5.3.1.1 问题三第 1 小问模型的建立

该问题要求我们针对该地区目前 PM2.5 的年平均浓度估计为 280(单位为 ) 时，要求未来五年内逐年减少 PM2.5 的年平均浓度，最终达到年终平均浓度统计指标 35，给出合理的治理计划，即给出每年的全年年终平均治理指标。

对此我们建立其相应的规划模型，由于要求未来五年内“逐年”减少 PM2.5 的年平均浓度，假设未来五年每年各自减少的平均浓度为  $x_1$ 、 $x_2$ 、 $x_3$ 、 $x_4$ 、 $x_5$ ，对此可假设其每年各自减少的年平均浓度以一个等比数列的方式“逐年”减少，其等比数列的公比为  $q$  ( $q \neq 0$ )。

**目标分析：**

**满意度最大（引入效用函数）**

本问题为规划与决策分析问题，对此可引入效用函数来刻画规划决策的满意度，由于每年减少 PM2.5 的年平均浓度越多，决策部门越满意，但又得考虑减少其 PM2.5 的年平均浓度太多，付出的人力物力也就越多，对此可取以下效用函数：

**效用函数：**

$$u(x_i) = \frac{x_i}{x_i + a} \quad (37)$$

其中， $x_i$  为第  $i$  年减少的年平均浓度

对于参数  $a$ ，利用“高通滤波系统中下限截止频率”的方法可以确定为  $a = 0.2$ ，其效用函数曲线如图 11 所示：

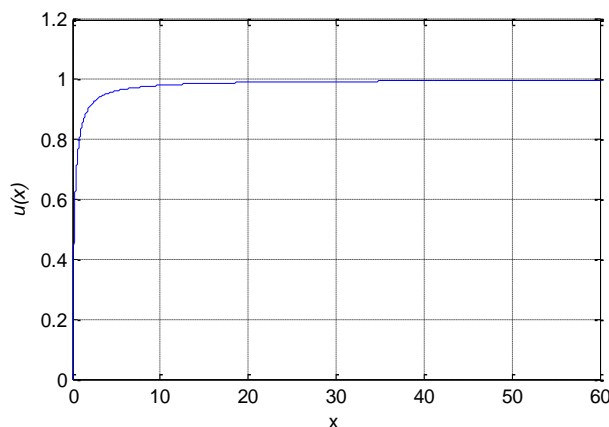


图11 效用函数曲线

**满意度最大：**

假设治理满意度为  $F$ ，治理满意度越高，说明治理计划越合理，则未来五年的综合满意度最大可表示为：

$$\max F = \sum_{i=1}^5 u(x_i) / 5 \quad (38)$$

其中： $F$  为治理满意度；

$u(x_i)$  为第  $i$  年的效用函数；

$x_i$  为第  $i$  年 PM2.5 减少的年平均浓度；

**约束分析：**

(1) 五年内 PM2.5 需减少的年平均浓度总和的约束

由问题知，该地区目前 PM2.5 的年平均浓度估计为 280（单位为  $\mu\text{g}/\text{m}^3$ ），要求未来五年内逐年减少 PM2.5 的年平均浓度，最终达到年终平均浓度统计指标 35，对此有约束：

$$\sum_{i=1}^5 x_i = 280 - 35 \quad (39)$$

(2) 逐年减少 PM2.5 的年平均浓度的约束

由于问题中，要求进行逐年减少，则不能一年就减少完，也不能某一年一点都不减少，而是一个逐步减少的过程，对此假设其每年各自减少的年平均浓度以一个等比数列的方式“逐年”减少，其等比数列的公比为 $q$  ( $q \neq 0$ )，则有约束：

$$\begin{cases} x_i = qx_{i-1} \\ x_i > 0 \quad (i=1,2,3,4,5) \end{cases} \quad (40)$$

(3) 每年效用函数的约束

由于每年的效用函数随该年 PM2.5 的年平均浓度减少量而决定，对此有约束：

$$u(x_i) = \frac{x_i}{x_i + 0.2} \quad (41)$$

于是可得到规划模型如下：

$$\begin{aligned} \max \quad & F = \sum_{i=1}^5 u(x_i) / 5 \\ \text{s.t.} \quad & \begin{cases} u(x_i) = \frac{x_i}{x_i + 0.2} \\ \sum_{i=1}^5 x_i = 280 - 35 \\ x_i = qx_{i-1} \\ q \neq 0 \\ x_i > 0 \quad (i=1,2,3,4,5) \end{cases} \end{aligned} \quad (42)$$

模型说明：

目标函数 F 为治理满意度最大；

$u(x_i)$  为第 i 年的效用函数；

$x_i$  为第 i 年 PM2.5 减少的年平均浓度；

$q$  表示每年 PM2.5 各自减少的年平均浓度以一个等比数列的方式“逐年”减少， $q$  为其等比数列的公比。

### 5.3.1.2 问题三第 1 小问模型的求解

对于这个规划模型，运用 LINGO 13.0 版本优化软件对其进行编程求解（其程序见附录 8.1.3）

求解最优结果为：

$$F = 0.9959 \quad q = 0.9531$$

$$x_1 = 53.81650 \quad x_2 = 51.29264 \quad x_3 = 48.88714 \quad x_4 = 46.59445 \quad x_5 = 44.40928$$

即最大满意度为 99.59%，逐年在上一年减少量的基础上减少上年减少量的 95.31%，未来五年 PM2.5 年平均浓度减排治污的具体治理计划如表 17 所示

表 17 未来五年 PM2.5 年平均浓度减排治污的具体治理计划

年份	该年 PM2.5 需减少的年平均浓度 (单位为 $\mu\text{g}/\text{m}^3$ )	该年 PM2.5 全年年终要求达标的年平均浓度 (单位为 $\mu\text{g}/\text{m}^3$ )
第一年	53.8165	226.1835
第二年	51.29264	174.89086
第三年	48.88714	126.00372
第四年	46.59445	79.40928
第五年	44.40928	35

### 5.3.2 问题三第 2 小问模型的建立与求解

#### 5.3.2.1 问题三第 2 小问模型的分析

该问题要求同时考虑综合治理和专项治理,为数据 1 所在地区设计有效的专项治理计划,使得既达到预定 PM2.5 减排计划,同时使经费投入较为合理,并要求给出五年投入总经费和逐年轻费投入预算计划。对此,对此可建立其相应的模型,以满足如下两个需求因素:

目标 1: 投入总费用最少

目标 2: PM2.5 减排计划实施满意度最大

目标分析:

#### 目标1: 投入总费用最少

该问题要求同时考虑综合治理和专项治理下,达到预定 PM2.5 减排计划的同时希望经费投入较为合理,当然经费投入越少越好,问题中给出了综合治理费用,每减少一个 PM2.5 浓度单位,当年需投入一个费用单位(百万元),专项治理投入费用是当年所减少 PM2.5 浓度平方的 0.005 倍(百万元)。

由于问题本身还可能存在其他对 PM2.5 不确定影响的复杂性,这里假设仅考虑 PM2.5 的专项治理和综合治理问题,设  $Z$  为投入的总费用,  $Z_1$  为综合治理投入的费用,  $Z_2$  为专项治理投入的费用,则投入的总费用最少可表示为:

$$\min Z = Z_1 + Z_2 \quad (43)$$

其中,  $Z_1 = \sum_{i=1}^5 x_i$ ,  $Z_2 = \sum_{i=1}^5 0.005(x_i + y_i)^2$

这里,  $Z$  表示投入的总费用;

$Z_1$  表示综合治理投入的费用;

$Z_2$  表示专项治理投入的费用;

$x_i$  为第  $i$  年经综合治理 PM2.5 减少的年平均浓度;

$y_i$  为第  $i$  年经专项治理 PM2.5 减少的年平均浓度;

#### 目标2: PM2.5减排计划实施满意度最大

在经费投入时,也希望 PM2.5 减排计划得到很好的实施与实现,能很好地达到预定 PM2.5 减排计划,对此存在 PM2.5 减排计划实施满意度的因素,假设综合治理和专项治理下的总满意度为  $F$ ,总满意度越高,说明整体治理计划越合理,则未来五年的 PM2.5 减排计划实施满意度最大可表示为:

$$\max F = \frac{1}{2} \left[ \sum_{i=1}^5 u(x_i) / 5 + \sum_{i=1}^5 u(y_i) / 5 \right] \quad (44)$$

其中：F 表示综合治理和专项治理下的总满意度；

$u(x_i)$  表示第 i 年综合治理的效用函数；

$u(y_i)$  表示第 i 年专项治理的效用函数；

$\sum_{i=1}^5 u(x_i) / 5$  表示未来五年综合治理下的满意度；

$\sum_{i=1}^5 u(y_i) / 5$  表示未来五年专项治理下的满意度；

**约束分析：**

(1) 五年内 PM2.5 需减少的年平均浓度总和的约束

由问题知，该地区目前 PM2.5 的年平均浓度估计为 280（单位为 ），要求未来五年内逐年减少 PM2.5 的年平均浓度，最终达到年终平均浓度统计指标 35，对此有约束：

$$\sum_{i=1}^5 x_i + \sum_{i=1}^5 y_i = 280 - 35 \quad (45)$$

(2) 逐年轻费投入的约束

由于问题中，要求进行逐年轻费投入，则不能一年就投完，也不能某一年一点都不投，而是一个逐步投入经费的过程，对此假设其每年各自投入总经费以一个等比数列的方式“逐年”投入，其等比数列的公比为  $q$  ( $q \neq 0$ )，则有约束：

$$x_i + 0.005(x_i + y_i)^2 = q[x_{i-1} + 0.005(x_{i-1} + y_{i-1})^2] \quad (46)$$

(3) 每年综合治理的效用函数的约束

由于每年综合治理的效用函数随该年 PM2.5 的年平均浓度减少量而决定，对此有约束：

$$u(x_i) = \frac{x_i}{x_i + 0.2} \quad (47)$$

(4) 每年专项治理的效用函数的约束

由于每年专项治理的效用函数随该年 PM2.5 的年平均浓度减少量而决定，对此有约束：

$$u(y_i) = \frac{y_i}{y_i + 0.2} \quad (48)$$

### 5.3.2.2 问题三第 2 小问模型的建立

针对这个多目标非线性规划问题，基于 5.3.2.1 的分析，建立多目标非线性规划模型如下：

$$\begin{aligned}
& \min Z = Z_1 + Z_2 \\
& \max F = \frac{1}{2} \left[ \sum_{i=1}^5 u(x_i) / 5 + \sum_{i=1}^5 u(y_i) / 5 \right] \\
& \left. \begin{aligned}
& Z_1 = \sum_{i=1}^5 x_i \\
& Z_2 = \sum_{i=1}^5 0.005(x_i + y_i)^2 \\
& \sum_{i=1}^5 x_i + \sum_{i=1}^5 y_i = 280 - 35 \\
& x_i + 0.005(x_i + y_i)^2 = q[x_{i-1} + 0.005(x_{i-1} + y_{i-1})^2] \\
& u(x_i) = \frac{x_i}{x_i + 0.2} \\
& u(y_i) = \frac{y_i}{y_i + 0.2} \\
& q \neq 0 \\
& x_i > 0 \quad (i=1,2,3,4,5) \\
& y_i > 0 \quad (i=1,2,3,4,5)
\end{aligned} \right\} \text{s.t.} \tag{49}
\end{aligned}$$

模型说明：

目标 1 为投入总费用最少；

目标 2 为 PM2.5 减排计划实施满意度最大；

$Z_1$  表示综合治理投入的费用；

$Z_2$  表示专项治理投入的费用；

$x_i$  为第  $i$  年经综合治理 PM2.5 减少的年平均浓度；

$y_i$  为第  $i$  年经专项治理 PM2.5 减少的年平均浓度；

$u(x_i)$  表示第  $i$  年综合治理的效用函数；

$u(y_i)$  表示第  $i$  年专项治理的效用函数；

### 5.3.2.3 问题三第 2 小问模型的求解

#### (1) 模型的简化

此目标规划是多目标规划，为了使此规划问题简单易行，对此采用主要目标法<sup>[2]</sup>将多目标规划转化为单目标规划。

若有  $P$  个目标函数，假设  $f_1(x)$  为主要目标，而其余  $P-1$  个目标  $f_j(x)$  ( $j=2,3,\dots,P$ ) 有一组允许的界限值  $a_j$ ，即希望满足要求：

$$f_i(x) \leq a_j \quad (j=2,3,\dots,p) \tag{50}$$

这样就可将它转化成



$$\min_{x \in \tilde{R}} u(x) = f_1(x) \quad \tilde{R} = \{x \in R \mid f_j(x) \leq a_j, j = 2, 3, \dots, p\}$$

对于界限值  $a_j$ ，一般取  $a_j \geq f_j = \min_{x \in R} f_j(x) (j = 2, 3, \dots, p)$

本题中有两个目标函数，设为  $f_1(x)$ ， $f_2(x)$ ； $f_1(x)$  为投入总费用最少的函数  $Z$ ； $f_2(x)$  为 PM2.5 减排计划实施满意度最大的函数  $F$ 。

令  $f_1(x)$  为主要目标函数即以投入总费用最少为主要目标函数， $f_2(x)$  的界限值为  $a_2$ ，通过 Lingo 软件<sup>[4]</sup>求解得  $\min_{x \in R} f_2(x) = 0$ ，则  $a_2 \geq 0$ ，我们暂时令界限值为  $a_2$ 。

则模型的表达式可简化为如下：

$$\begin{aligned} \min \quad & Z = Z_1 + Z_2 \\ \text{s.t.} \quad & \left\{ \begin{aligned} & F = \frac{1}{2} \left[ \sum_{i=1}^5 u(x_i) / 5 + \sum_{i=1}^5 u(y_i) / 5 \right] \geq a_2 \\ & Z_1 = \sum_{i=1}^5 x_i \\ & Z_2 = \sum_{i=1}^5 0.005(x_i + y_i)^2 \\ & \sum_{i=1}^5 x_i + \sum_{i=1}^5 y_i = 280 - 35 \\ & x_i + 0.005(x_i + y_i)^2 = q[x_{i-1} + 0.005(x_{i-1} + y_{i-1})^2] \\ & u(x_i) = \frac{x_i}{x_i + 0.2} \\ & u(y_i) = \frac{y_i}{y_i + 0.2} \\ & q \neq 0 \\ & x_i > 0 \quad (i = 1, 2, 3, 4, 5) \\ & y_i > 0 \quad (i = 1, 2, 3, 4, 5) \end{aligned} \right. \end{aligned} \quad (51)$$

## (2) 模型的求解结果

对此我们运用 LINGO 13.0 版本软件对其进行编程求解，可得到满意度  $F$  不同限制值  $a_2$  时，对应的最小投入总费用，综合其满意度、最小投入总费用和其变化情况，最终选取了比较适合满意度  $a_2 = 90\%$ ，五年投入的总费用最少为 64.13053 费用单位（百万单位）， $q$  为 1。具体每年综合和专项的计划为：

表 18 未来五年综合治理与专项治理经费投入的具体预算计划

年份	综合治理投入经费（百万单位）	专项治理投入经费（百万单位）
第一年	0.8211	12.005
第二年	0.8211	12.005
第三年	0.8211	12.005
第四年	0.8211	12.005
第五年	0.8211	12.005
总费用	4.1055	60.025

### (3) 方案合理性的论述

在上述模型求解中，运用 LINGO 13.0 版本软件对其进行编程求解，可得到满意度 F 不同限制值  $a_2$  时，对应的最小投入总费用，如下表所示：

表 19 满意度 F 不同限制值下的最小投入总费用

满意度 F 的限制值 $a_2$	$\min Z$ （百万单位）	$\Delta Z/\Delta a_2$ 的变化率
0.99	128.65250	——
0.97	76.99973	2582.6385
0.95	69.46639	376.6670
0.9	64.13053	106.7172
0.85	62.40457	34.5192
0.8	61.55083	17.0748

由表知，随着满意度的减小，最小投入总费用也减小，且变化率越来越小，综合其满意度、最小投入总费用和其变化情况，最终选取了的满意度  $a_2=90\%$ ，五年投入的总费用最少为 64.13053 费用单位（百万单位）是比较适中的，比较适合决策规划者。

对于专项治理投入经费为什么比综合治理投入经费多，是因为由专项治理投入经费和综合治理投入经费，在投入相同经费下，专项治理投入经费减少 PM2.5 含量多的原因所致。

但若单一全投入专项治理，那样综合治理的满意度就为 0，从而影响到总的满意度，对此该方案综合满意度和经费投入上总体上是比较好的。

模型求解出的每年综合治理和专项治理减少的 PM2.5 年平均浓度如下：

$$x_1 = x_2 = x_3 = x_4 = x_5 = 0.8211$$

$$y_1 = y_2 = y_3 = y_4 = y_5 = 48.17889$$

通过所得数据反推可知五年减少的 PM2.5 年平均浓度为 245，最终能达到最终的 35，该方案既达到了预定 PM2.5 减排计划，同时也达到了使经费投入较为合理的目的，可见该方案是合理的。

## 六 模型的评价与推广

### (1) 模型的评价

本文针对空气中 PM2.5 的相关因素分析、分布与演变及应急处理和空气质量控制管理的问题，采用了相关分析、回归分析、二维插值、微分方程、非线性规划等方法，建立了相关性分析模型、多元线性回归方程模型、Shepard 二维插值模型、偏微分方程模型和多目标非线性规划模型，运用最小二乘估计算法、数值插值和搜索算法等对问题进行了较好地求解与分析。

在问题一中,采用了最小二乘估计方法对回归系数进行估计,并对回归方程进行了残差检验,且对异常点进行了逐次剔除最终得到拟合度达到 97.1%的多元线性回归方程;在问题二中,考虑了多方面情况,建立了 Shepard 二维插值模型和偏微分方程模型;在问题三中,引入了效用函数建立以满意度最大为目标的非线性规划模型,在对多目标非线性规划模型求解过程中,利用了主要目标法将双目标简化为单目标。

## (2) 模型推广

对于问题中,我们进行了很多假设,简化了问题的复杂化,但也过于理想化,如在第二问中假设了整个扩散过程 PM2.5 污染物质量守恒,整个研究空间内风速场是均匀稳定的,第三问中仅考虑了 PM2.5 的专项治理和综合治理问题,没考虑其他影响。对此,这些还有待于以后进一步深入研究。

## 七 参考文献

- [1] 章文波,陈红艳,实用数据统计分析及 SPSS12.0 应用[M],人民邮电出版社,2006.
- [2] 韩中庚.数学建模方法及其应用[M],高等教育出版社,2005.
- [3] 姜启源等.《数学模型》(第三版)[M],高等教育出版社,2003.8.
- [4] 谢金星,薛毅.优化建模与 LINDO/LINGO 软件[M],北京:清华大学出版社,2005.
- [5] 余学春,北京 PM<sub>2.5</sub> 水溶性物种污染特征及来源解析[D],清华大学 2004
- [6] 韩明山,城市大气污染实时监测与控制系统中污染物的扩散特性研究[D],中国海洋大学,2004.
- [7] 徐瑞,黄兆东等.MATLAB2007 科学计算与工程分析[M],科学出版社,2008.

## 八 附录

### 8.1 附录清单

#### 8.1.1 问题一相关程序

- 1、 PM2.5 与其余 5 项分指标间的散点图程序
- 2、 PM2.5 与 5 项指标间的回归方程程序

#### 8.1.2 问题二相关程序

- 1、 各监测地区 PM2.5 含量随时间变化散点图程序
- 2、 西安地区 PM2.5 含量随时间变化的一维插值程序
- 3、 各监测地区 PM2.5 含量分布二维插值程序
- 4、 西安市 PM2.5 扩散受风力、湿度等影响的散点图程序
- 5、 PM2.5 扩散受其他因素影响的回归模型程序
- 6、 PM2.5 点源扩散模型求解程序

#### 8.1.3 问题三相关程序

- 3、 全年年终平均治理指标计算 LINGO 程序
- 4、 求投入总经费和逐年轻费投入预算计划的 LINGO 程序

### 8.2 附录内容

#### 8.2.1 问题一相关程序

## 1、PM2.5 与其余 5 项分指标间的散点图程序

```
clear;
a=[];% a为PM2.5与其余5项指标数据矩阵，因数据太多，此处数据为附上；
for i=1:5;
    figure(i)
    plot(a(:,i),a(:,end),'+')
end
```

## 2、PM2.5 与 5 项指标间的回归方程程序

```
x1=[a(:,1:5)];
x=[ones(size(a(:,3))),x1];
[b,bint,r,rint,stats]=regress(a(:,end),x);
figure
b,bint,stats,rcoplot(r,rint)
% 剔除一些点
a([3,7,32,33,35,41,42,43,46,57,63,64,69],:)=[];
x1=[a(:,1:5)];
x=[ones(size(a(:,3))),x1];
[b1,bint1,r1,rint1,stats1]=regress(a(:,end),x);
figure
b1,bint1,stats1,rcoplot(r1,rint1)
% 剔除点
a([5,8,30,36,37,49,50,53,54,55,56,59,69,70,82,86,87,102],:)=[];
x1=[a(:,1:5)];
x=[ones(size(a(:,3))),x1];
[b2,bint2,r2,rint2,stats2]=regress(a(:,end),x);
figure
b2,bint2,stats2,rcoplot(r2,rint2)
% 剔除点
a([17,30,32,42,46,53,58,61,71,72,107,121,122],:)=[];
x1=[a(:,1:5)];
x=[ones(size(a(:,3))),x1];
[b3,bint3,r3,rint3,stats3]=regress(a(:,end),x);
figure
b3,bint3,stats3,rcoplot(r3,rint3)
% 剔除点
a([6,18,24,30,48,66,70,131],:)=[];
x1=[a(:,1:5)];
x=[ones(size(a(:,3))),x1];
[b4,bint4,r4,rint4,stats4]=regress(a(:,end),x);
figure
b4,bint4,stats4,rcoplot(r4,rint4)
```

### 8.2.2 问题二相关程序

#### 1、各监测地区 PM2.5 含量随时间变化散点图程序

```
figure
plot(gaoya(:,1),gaoya(:,2),'*','MarkerSize',4)
```

```

xlabel('时间');ylabel('高压开关厂')
figure
plot(xingqin(:,1),xingqin(:,2),'*','MarkerSize',4)
xlabel('时间');ylabel('兴庆小区')
figure
plot(fangzhi(:,1),fangzhi(:,2),'*','MarkerSize',4)
xlabel('时间');ylabel('纺织城')
figure
plot(xiaozhai(:,1),xiaozhai(:,2),'*','MarkerSize',4)
xlabel('时间');ylabel('小寨')
figure
plot(cityPE(:,1),cityPE(:,2),'*','MarkerSize',4)
xlabel('时间');ylabel('市人民体育场')
figure
plot(gaoxing(:,1),gaoxing(:,2),'*','MarkerSize',4)
xlabel('时间');ylabel('高新西区')
figure
plot(jingkaiqu(:,1),jingkaiqu(:,2),'*','MarkerSize',4)
xlabel('时间');ylabel('经开区')
figure
plot(changanqu(:,1),changanqu(:,2),'*','MarkerSize',4)
xlabel('时间');ylabel('长安区')
figure
plot(yanliangqu(:,1),yanliangqu(:,2),'*','MarkerSize',4)
xlabel('时间');ylabel('阎良区')
figure
plot(lintongqu(:,1),lintongqu(:,2),'*','MarkerSize',4)
xlabel('时间');ylabel('临潼区')
figure
plot(qujiangwenhua(:,1),qujiangwenhua(:,2),'*','MarkerSize',4)
xlabel('时间');ylabel('曲江文化集团')
figure
plot(guangyuntan(:,1),guangyuntan(:,2),'*','MarkerSize',4)
xlabel('时间');ylabel('广运潭')
figure
plot(caotan(:,1),caotan(:,2),'*','MarkerSize',4)
xlabel('时间');ylabel('草滩')
figure
plot(totalcity(:,1),totalcity(:,2),'*','MarkerSize',4)
xlabel('时间');ylabel('全市平均')

```

**2、西安地区 PM2.5 含量随时间变化的一维插值程序**

```

x=totalcity(:,1);y=totalcity(:,2);
i=1:0.1:size(x);
a=interp1(x,y,i,'spline')
plot(x,y,'.',i,a,'r')

```

```
legend('监测点','插值曲线')
```

### 3、各监测地区 PM2.5 含量分布二维插值程序

```
clear;
x=[-6.1 2.8 9.7 -0.7 0 -6.2 -1.5 -4.4 25.1 23.8 2.1 9.1 -8.2];
y=[-0.33 -1.2 -0.425 -5.2 0 -4.3 8.4 -12.4 43.6 10.8 -6.7 5 11];
z=[213.754386 194.2075472 188.2792793 174.9009009 204.63
210.9009009 196.4554455 178.6396396 190.5585586
179.1717172 182.6181818 204.0277778 210.0357143];
ti=-15:0.01:45;
[xi,yi]=meshgrid(linspace(min(x),max(x),1000),linspace(min(y),max(y),1000));
zi=griddata(x,y,z,xi,yi,'cubic');
figure
meshc(xi,yi,zi)
colorbar
% figure
% surfc(xi,yi,zi)
% colorbar
shading interp
xlabel('x 轴')
ylabel('y 轴')
zlabel('PM2.5 各地区分指数')
hold on
plot3(x,y,z,'*','MarkerSize',8)
figure
[c,h]=contourf(xi,yi,zi);
clabel(c,h)
colorbar
% grid on;
hold on
plot(x,y,'*','MarkerSize',6)
xlabel('x 轴')
ylabel('y 轴')
```

### 4、西安市 PM2.5 扩散受风力、湿度等影响的散点图程序

```
figure
for i=1:4
subplot(2,2,i)
plot(data_day(:,i+1),data_day(:,1),'.')
end
figure
for j=1:4
subplot(2,2,j)
plot(data_nigh(:,j+1),data_nigh(:,1),'.')
end
plot(data_day(:,3),data_day(:,4),'');
```

### 5、PM2.5 扩散受其他因素影响的回归模型程序

```

x1=[data_day(1:41,2:4)];
% x1=[shidu data_day(1:41,3) feng(1:41) data_day(1:41,5)];
x1=[ones(size(data_day(1:41,3))),x1];
[b1,bint1,r1,rint1,stats1]=regress(data_day(1:41,1),x1);
figure
b1,bint1,stats1,rcoplot(r1,rint1)

```

```

x2=[data_nigh(1:41,2:4)];
% x1=[shidu data_day(1:41,3) feng(1:41) data_day(1:41,5)];
x2=[ones(size(data_nigh(1:41,3))),x2];
[b2,bint2,r2,rint2,stats2]=regress(data_nigh(1:41,1),x2);
figure
b2,bint2,stats2,rcoplot(r2,rint2)

```

```

x3=[data_day(42:end,2:4)];
% x1=[shidu data_day(1:41,3) feng(1:41) data_day(1:41,5)];
x3=[ones(size(data_day(42:end,3))),x3];
[b3,bint3,r3,rint3,stats3]=regress(data_day(42:end,1),x3);
figure
b3,bint3,stats3,rcoplot(r3,rint3)

```

```

x4=[data_nigh(42:end,2:4)];
% x1=[shidu data_day(1:41,3) feng(1:41) data_day(1:41,5)];
x4=[ones(size(data_nigh(42:end,3))),x4];
[b4,bint4,r4,rint4,stats4]=regress(data_day(42:end,1),x4);
figure
b4,bint4,stats4,rcoplot(r4,rint4)
r=sqrt([stats1(1) stats2(1) stats3(1) stats4(1)])

```

## 6、PM2.5 点源扩散模型求解程序

```

clear;
x=[-6.1 2.8 9.7 -0.7 0 -6.2 -1.5 -4.4 25.1 23.8 2.1 9.1 -8.2];
y=[-0.33 -1.2 -0.425 -5.2 0 -4.3 8.4 -12.4 43.6 10.8 -6.7 5 11];
x1=x+6.1;
y1=y+0.33;
s=sqrt(x1.^2+y1.^2);
Q=500*2;
u=2.7;
% 求 sigma_y
% 稳定度等级=C-D
% 0.5h 取样时间
apha1=0.884946;
gama1=0.189396;
sigma_y1=gama1*s.^apha1;
% 2h 取样时间
q=0.3;

```

```

tao1=0.5;
tao2=2;
sigma_y2=sigma_y1*(tao2/tao1)^q;
% 求 sigma_z
% 稳定度等级=C-D
% 0.5h 取样时间
apha2=0.815575;
gama2=0.136659;
sigma_z1=gama2*s.^apha2;
% 2h 取样时间
q=0.3;
tao1=0.5;
tao2=2;
sigma_z2=sigma_z1*(tao2/tao1)^q;
% 地面连续排放点源在地面上任一点产生浓度
%
c1=abs(Q./(pi*u*sigma_y2(2:end).*sigma_z2(2:end)).*exp(-y1.^2./sigma_y2.^2(2:en
d)))
c1=zeros(1,13);
c1(1)=Q;
c1(2:end)=abs(Q./(pi*u*sigma_y2(2:end).*sigma_z1(2:end)));

[xi,yi]=meshgrid(linspace(min(x1),max(x1),1000),linspace(min(y1),max(y1),100
0));
ci=griddata(x1,y1,c1,xi,yi,'cubic');
figure
meshc(xi,yi,ci)
colorbar
% figure
% surfc(xi,yi,zi)
% colorbar
shading interp
xlabel('x 轴')
ylabel('y 轴')
zlabel('PM2.5 浓度')
hold on
plot3(x1,y1,c1,'*','MarkerSize',8)
figure
[c,h]=contourf(xi,yi,ci);
clabel(c,h)
colorbar
% grid on;
hold on
plot(x1,y1,'*','MarkerSize',6)
xlabel('x 轴')

```



ylabel('y 轴')

### 8.2.3 问题三相关程序

#### 1、全年年终平均治理指标计算 LINGO 程序

model:

```
max=(u1+u2+u3+u4+u5)/5;
u1=x1/(x1+0.2);
u2=x2/(x2+0.2);
u3=x2/(x2+0.2);
u4=x4/(x4+0.2);
u5=x5/(x5+0.2);
x1+x2+x3+x4+x5=245;
x2=p*x1;
x3=p*x2;
x4=p*x3;
x5=p*x4;
x2<x1;
x3<x2;
x4<x3;
x5<x4;
p>0;
end
```

Local optimal solution found.

Objective value:	0.9959544
Infeasibilities:	0.000000
Total solver iterations:	15

Model Class:	NLP
--------------	-----

Total variables:	11
Nonlinear variables:	6
Integer variables:	0
Total constraints:	16
Nonlinear constraints:	9
Total nonzeros:	41
Nonlinear nonzeros:	13

Variable	Value	Reduced Cost
U1	0.9962974	0.000000
U2	0.9961159	0.000000
U3	0.9961159	0.000000
U4	0.9957260	0.000000

U5	0.9955166	0.000000
X1	53.81650	0.000000
X2	51.29264	0.000000
X4	46.59445	0.000000
X5	44.40928	0.000000
X3	48.88714	0.000000
P	0.9531024	0.000000

## 2、求投入总经费和逐年轻费投入预算计划的 LINGO 程序

model:

$\min = x1 + x2 + x3 + x4 + x5 + 0.005 * (x1 + y1)^2 + 0.005 * (x2 + y2)^2 + 0.005 * (x3 + y3)^2 + 0.005 * (x4 + y4)^2 + 0.005 * (x5 + y5)^2;$

$F = (u1 + u2 + u3 + u4 + u5 + u6 + u7 + u8 + u9 + u10) / 10;$

$F > 0.9;$

$x1 + x2 + x3 + x4 + x5 + y1 + y2 + y3 + y4 + y5 = 245;$

$u1 = x1 / (x1 + 0.2);$

$u2 = x2 / (x2 + 0.2);$

$u3 = x3 / (x3 + 0.2);$

$u4 = x4 / (x4 + 0.2);$

$u5 = x5 / (x5 + 0.2);$

$u6 = y1 / (y1 + 0.2);$

$u7 = y2 / (y2 + 0.2);$

$u8 = y3 / (y3 + 0.2);$

$u9 = y4 / (y4 + 0.2);$

$u10 = y5 / (y5 + 0.2);$

$x2 + 0.005 * (x2 + y2)^2 = p * (x1 + 0.005 * (x1 + y1)^2);$

$x3 + 0.005 * (x3 + y3)^2 = p * (x2 + 0.005 * (x2 + y2)^2);$

$x4 + 0.005 * (x4 + y4)^2 = p * (x3 + 0.005 * (x3 + y3)^2);$

$x5 + 0.005 * (x5 + y5)^2 = p * (x4 + 0.005 * (x4 + y4)^2);$

$p > 0;$

end

Local optimal solution found.

Objective value:	64.13053
Infeasibilities:	0.000000
Total solver iterations:	65

Model Class:	NLP
--------------	-----

Total variables:	22
Nonlinear variables:	11
Integer variables:	0

Total constraints:	19
Nonlinear constraints:	15

Total nonzeros: 73  
Nonlinear nonzeros: 40

Variable	Value	Reduced Cost
X1	0.8211065	0.6046537E-07
X2	0.8211064	-0.1373528E-07
X3	0.8211065	0.3463184E-07
X4	0.8211065	0.3903357E-07
X5	0.8211064	0.000000
Y1	48.17889	0.000000
Y2	48.17889	0.000000
Y3	48.17889	0.000000
Y4	48.17889	0.000000
Y5	48.17889	0.000000
F	0.9000000	0.000000
U1	0.8041340	0.000000
U2	0.8041340	0.000000
U3	0.8041340	0.000000
U4	0.8041340	0.000000
U5	0.8041340	0.000000
U6	0.9958660	0.000000
U7	0.9958660	0.000000
U8	0.9958660	0.000000
U9	0.9958660	0.000000
U10	0.9958660	0.000000
P	1.000000	-0.7991577E-07