题 目 **空气中PM2.5问题的研究**

摘 要

先交代本文讨论的问题

就每个问题进行叙述，利用了什么方法，建立了什么模型，用什么算法进行求解，得到了什么结果

**关键词：**

一、问题重述

**1.1问题背景**

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。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》，启用空气质量指数AQI作为空气质量监测指标，新标准中首次将产生灰霾的主要因素细颗粒物PM2.5的浓度指标作为空气质量监测指标。然而，细颗粒物PM2.5进入公众视线的时间还很短，其相关的因素的统计数据还太少，人们对细颗粒物PM2.5的客观规律也了解得很不够。

**1.2问题的提出**

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。

描述该地区内PM2.5的时空分布及其规律，并结合《环境空气质量标准》分区进行污染评估。合理考虑风力、湿度等天气和季节因素的影响，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，并利用该地区的数据进行定量与定性分析。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，请建立针对这种突发情形的污染扩散预测与评估方法。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PM2.5 的成因、演变等一般性规律。

给出合理的治理计划，使得该地区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（*μg/m³*），即给出每年的全年年终平均治理指标。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经费投入预算计划，并论述该方案的合理性。

二、问题分析

**关于问题一：**

解题的大概思路、数据的初步处理

**关于问题二：**

三、符号说明

四、问题一

**4.1 模型假设**

**4.2 建模分析**

**题目要求判别检测指标之间的关系，利用相关性分析方法初步判别了pm2.5与其余指标的相关性，再利用三次回归分析方法分别求解pm2.5的指标与其余指标的关系。建立多元线性回归分析的模型，**

**4.3 模型建立与求解**

基本监测指标的相关性分析

Person相关分析就是研究变量间相关程度的一种统计分析方法，通过计算在一定的显著性水平下的相关系数来衡量不同检测指标之间的相关性。Person 相关系数取值在-1~1 之间，绝对值越大，说明相关性越显著。

根据附件1中武汉市一个检测点数据，对二氧化硫（SO2 ），二氧化氮（NO2 ），一氧化碳（CO）、可吸入颗粒物、臭氧、pm2.5的指标值做相关性分析，运用spss软件，得到的相关系数矩阵如下表

表4.1 基本监测指标的相关系数矩阵



即pm2.5与SO2、NO2、CO、O3 、pm10的相关系数分别为：

r=（0.726，0.734，0.822，-0.352，0.779）

式（）中有4个相关系数大于0.7，由相关系数的性质可知，由表4.1可知，pm2.5的指标与二氧化硫、二氧化氮、一氧化碳、可吸入颗粒物高度相关，呈正相关关系，相关系数分别为0.726、0.734、0.822、0.779，pm2.5的指标和臭氧低度相关，呈负相关关系，相关系数为-0.352。

为更好描述检测指标之间的关系，本文对pm2.5的指标值和其他检测指标进行三次回归拟合，给出拟合的结果和残差分析图，如图xx所示。

建立多元线性回归分析的模型，定量分析pm2.5的指标与其余指标的关系。令PM2.5为y，SO2为x1，NO2为x2，PM10为x3 ，CO为x4，O3为x5

Y=*β0+β1x1+β2x2+β3x3+β4x4+β5x5*

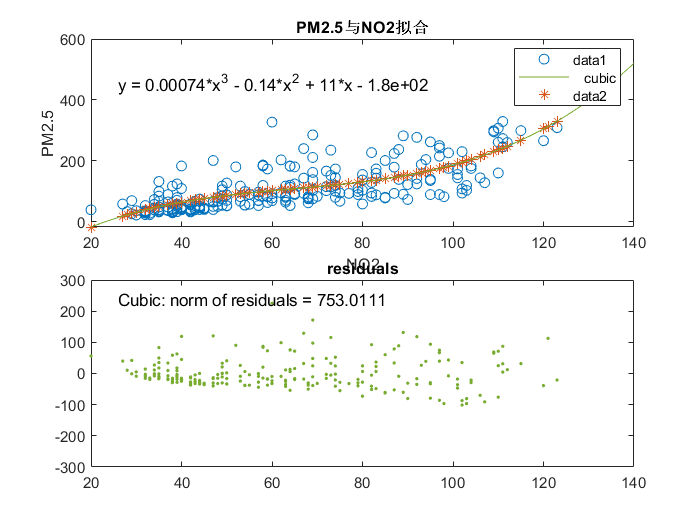
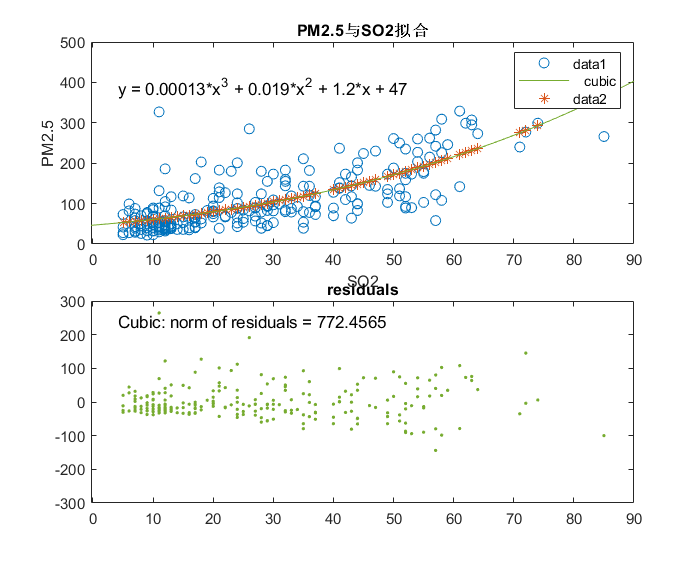
其中，β0，β1，β2，β3，β4，β5为回归系数，且都是与x1，x2，x3，x4，x5无关的参数。利用最小二乘法估计回归系数，利用matlab进行求解。

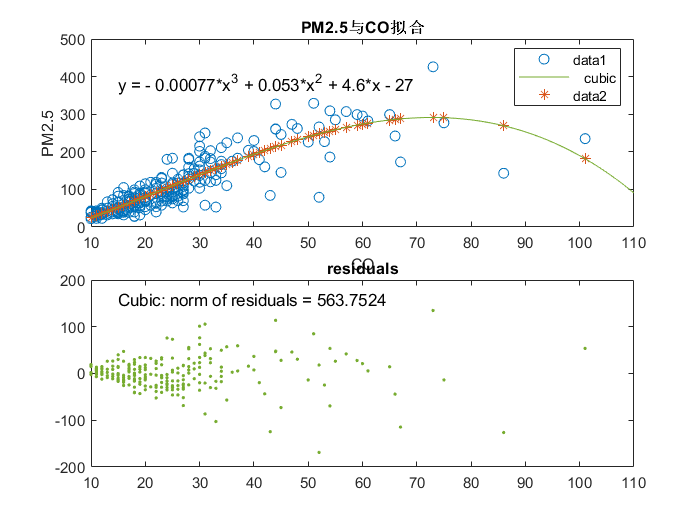
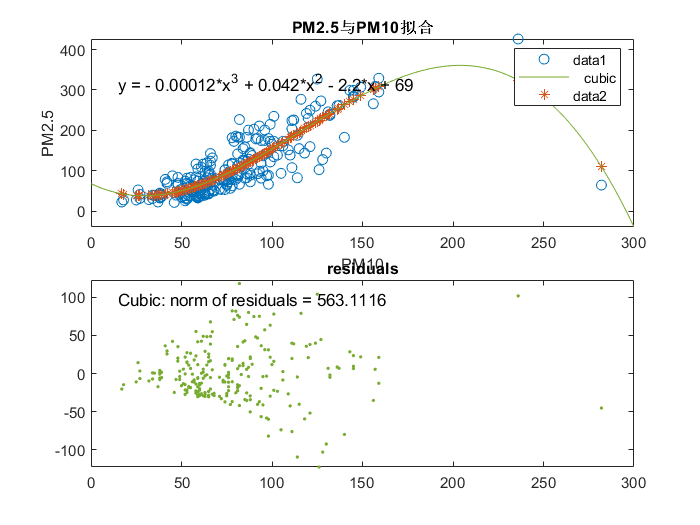
**4.4 模型结论**

**三次回归拟合的结果：**

**表4.2 Pm2.5与各项指标三次回归拟合**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | SO2 | NO2 | PM10 | CO | O3 |
| β3 | 0.001 | 0.007 | -0.0001 | -0.0008 | -0.0001 |
| β2 | 0.019 | -1.143 | 0.0423 | 0.0528 | 0.05 |
| β1 | 1.2288 | 10.5921 | -2.1597 | 4.626 | -5.0505 |
| β0 | 46.6202 | -177.9745 | 68.5221 | -26.6523 | 228.5385 |
| SSE（误差平方和） | 5.97E+05 | 5.67E+05 | 3.17E+05 | 3.18E+05 | 9.60E+05 |
| R-square（决定系数） | 0.5488 | 0.5712 | 0.7602 | 0.7597 | 0.2744 |
| Adjusted R-square | 0.543 | 0.5657 | 0.7571 | 0.7566 | 0.2651 |
| RMSE（均方根误差） | 50.5 | 49.23 | 36.81 | 36.85 | 64.03 |





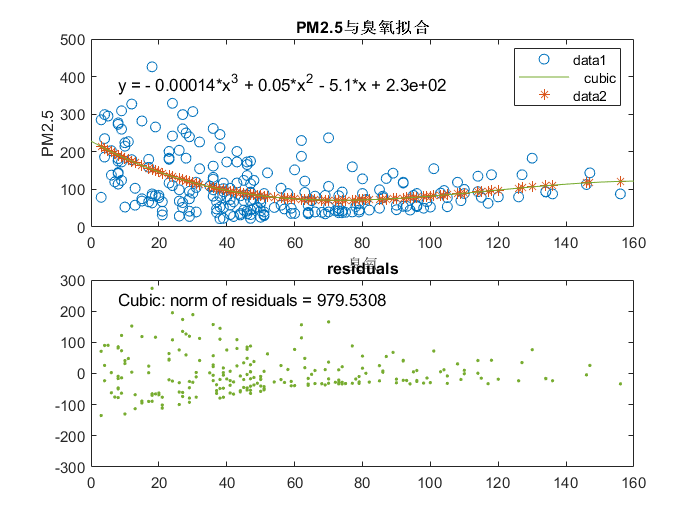


图4.1 Pm2.5与各项指标三次回归拟合及残差

残差分布散乱，不具有规律，表明拟合的结果合理。

多元线性回归的结果：

β0=-40.1429，β1=-0.1239，

β2=-0.1604，β3=1.5865，

β4=2.3662，β5=-0.3967

Y=-40.1429-0.1239x1-0.1604x2+1.5865x3+2.3662x4-0.3967x5

显著性概率p<0.0001，R2=0.8902,相关关系成立。

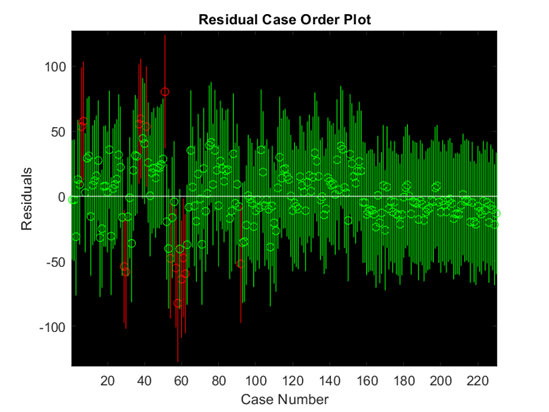


图4.2 **多元线性回归分析的残差分布图**

五、问题二

**5.1模型假设**

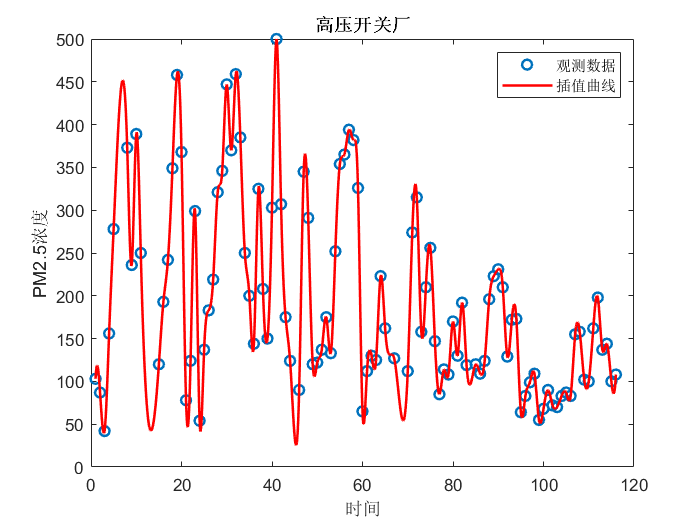
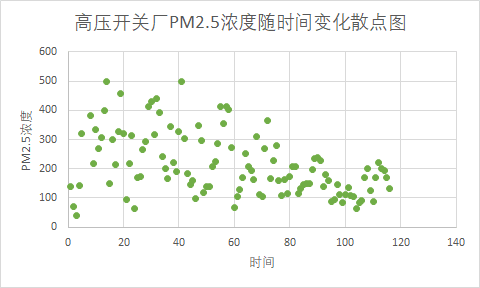
**5.2 建模分析**

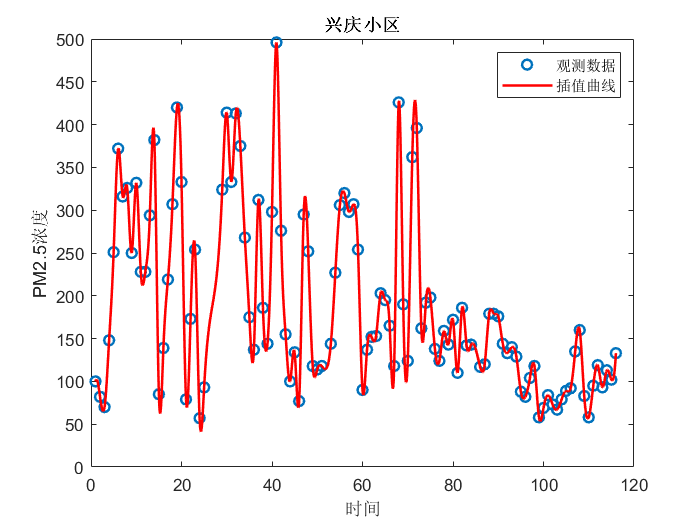
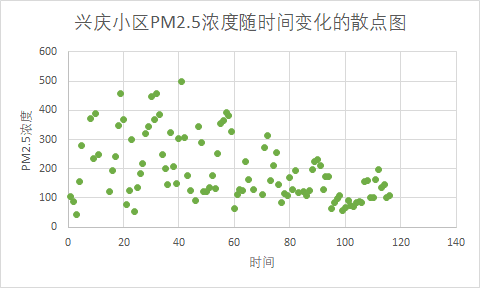
**5.3 模型建立与求解**

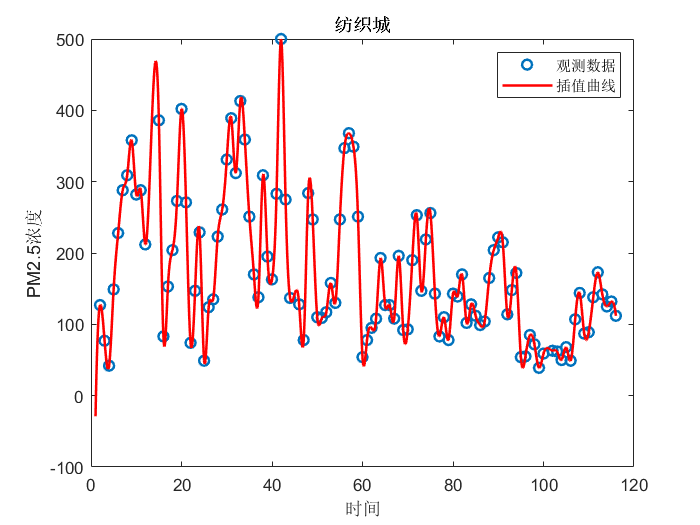
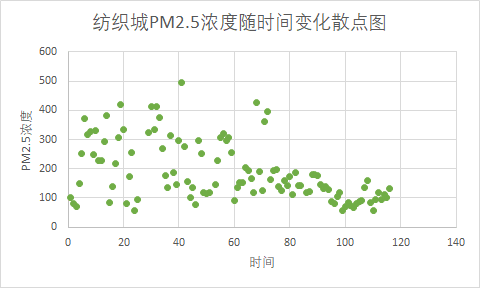
**5.3.1.1 PM2.5的时空分布**

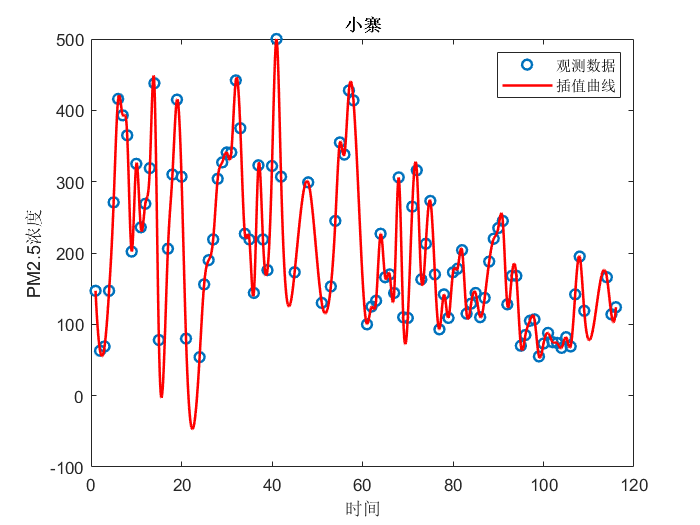
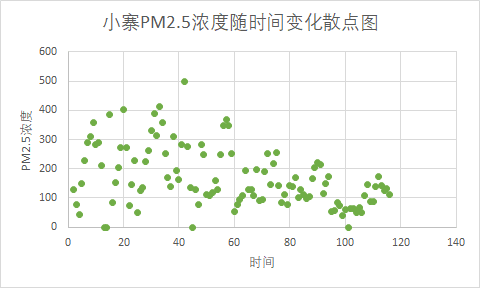
（1）时间分布

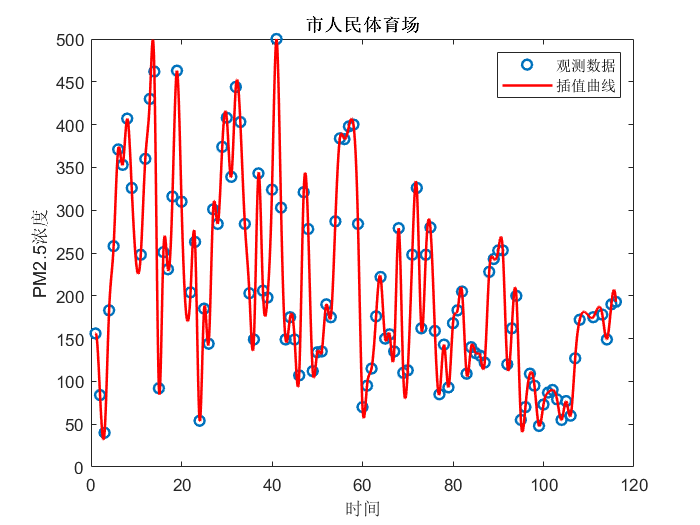
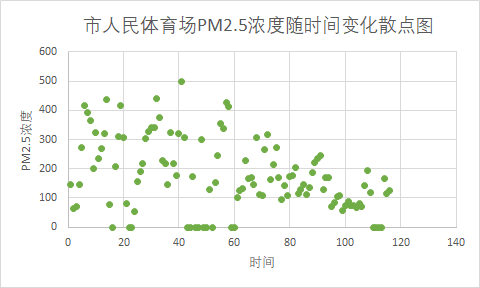
由附件2中西安市13个检测点的数据，利用一维线性插值方法拟合pm2.5的时间分布规律，得到不同区域的拟合曲线如下图所示

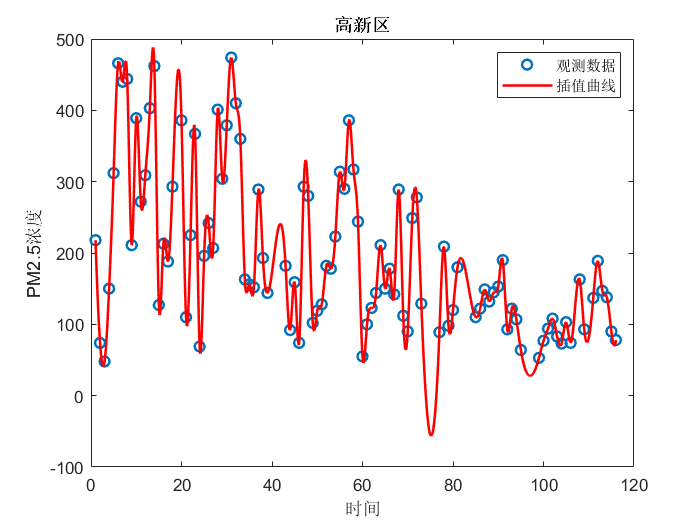
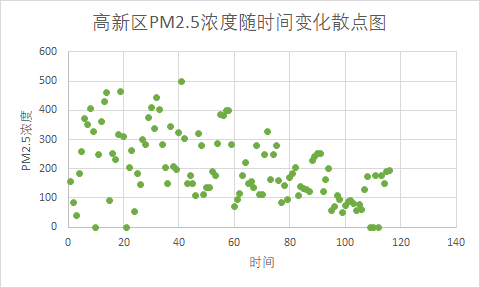


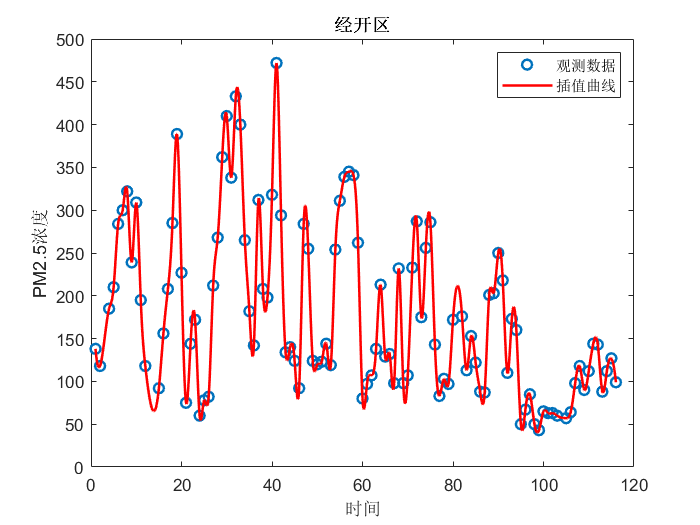
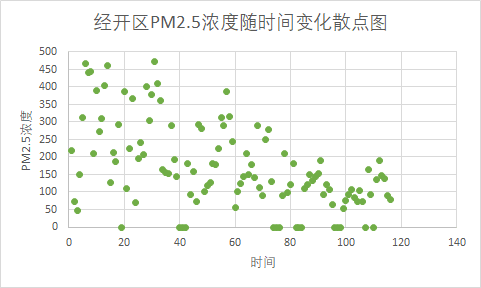


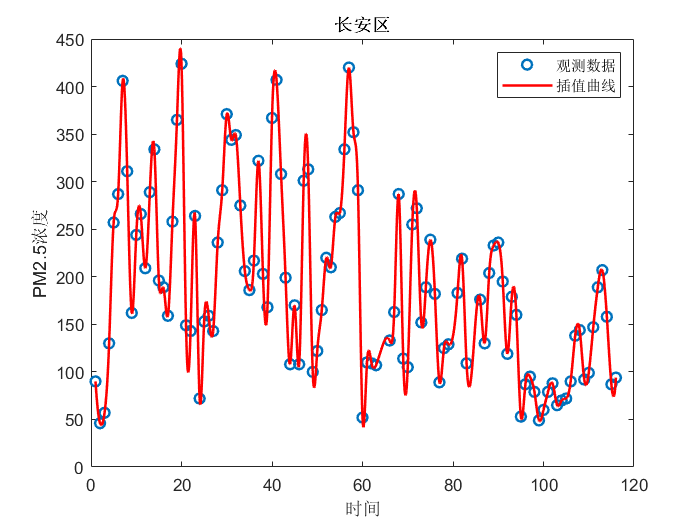
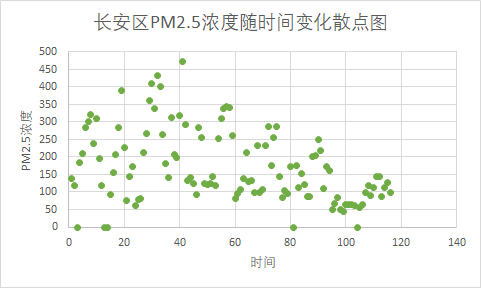


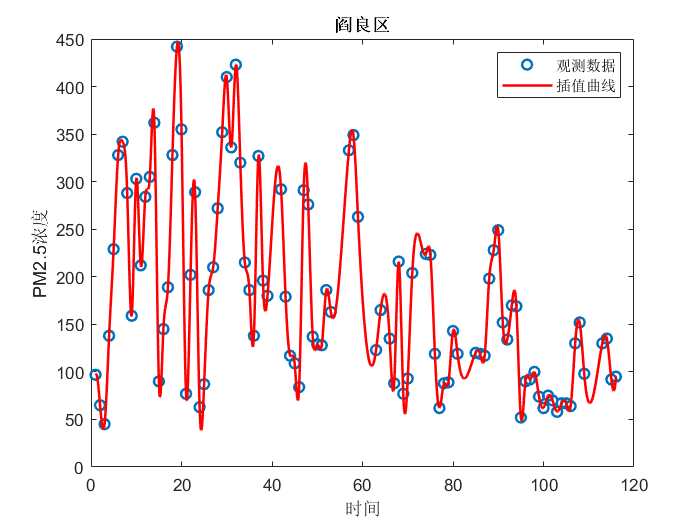
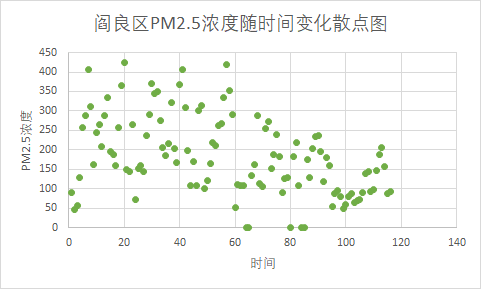


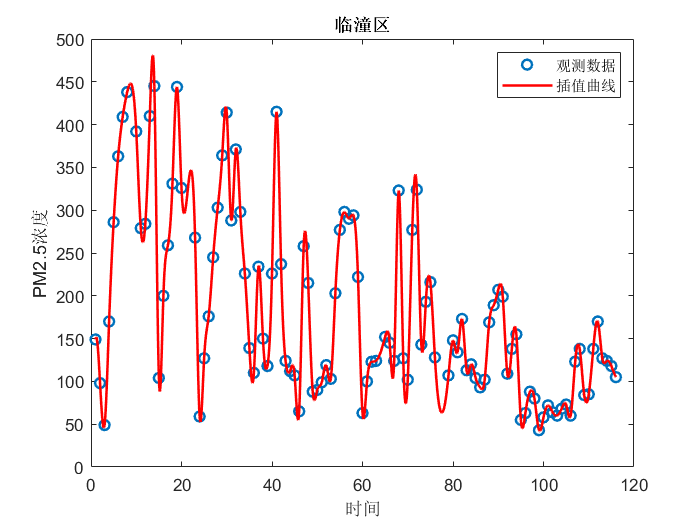
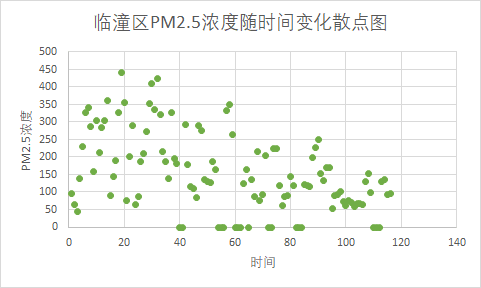


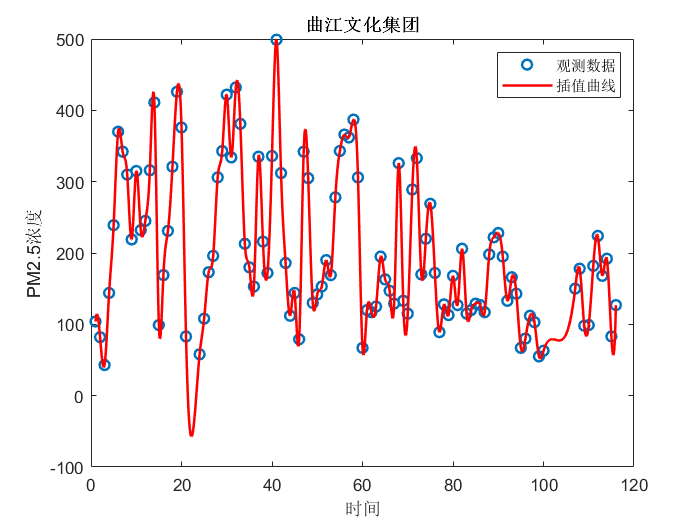
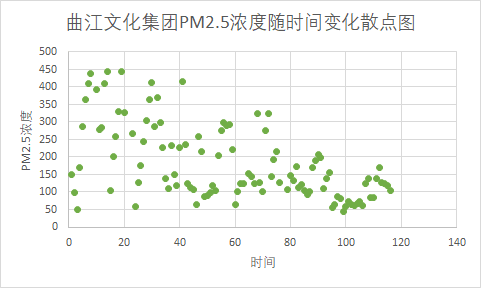


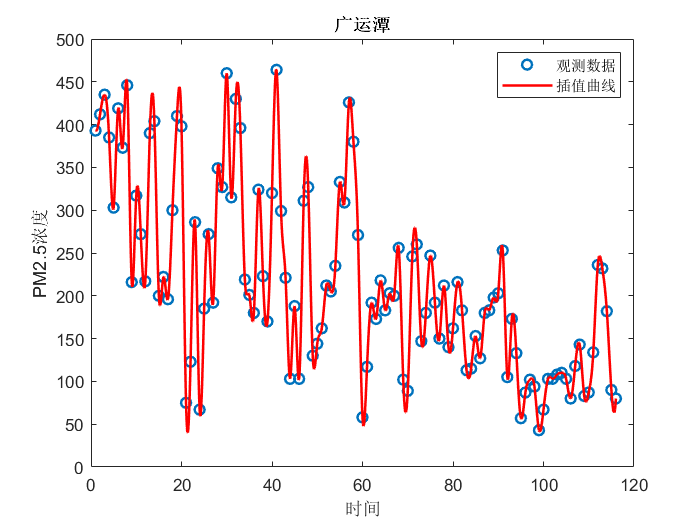
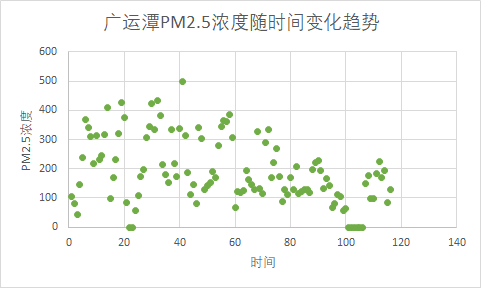


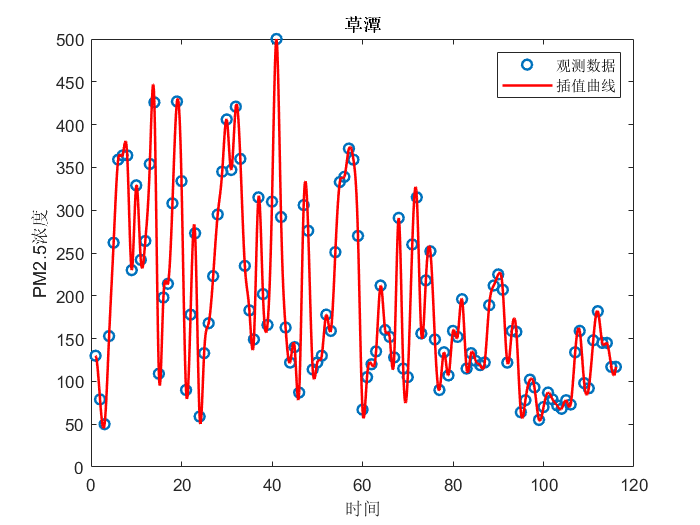
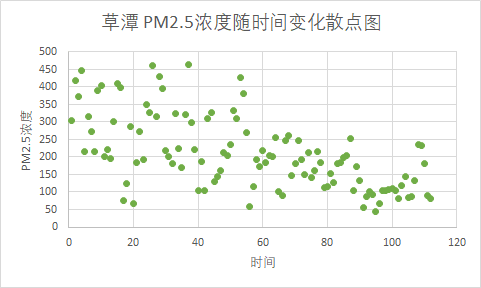












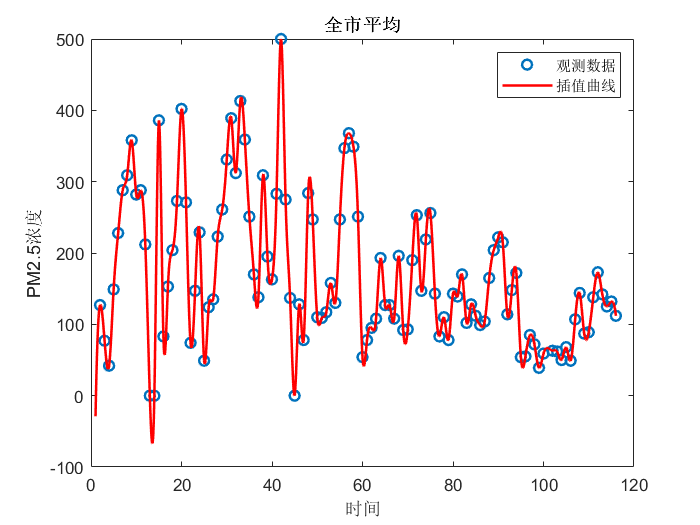
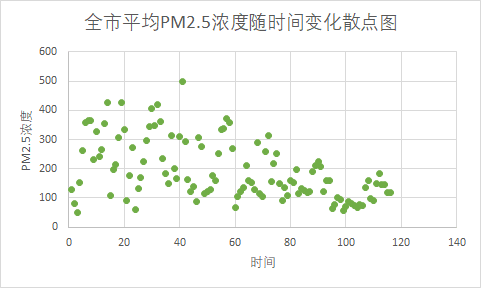


图5.1 pm2.5的时间变化规律

（2）pm2.5的空间分布

由 附件2中数据，首先计算2010年1月1日至2013年4月26日西安市13个监测点的浓度平均值，如图5.2所示

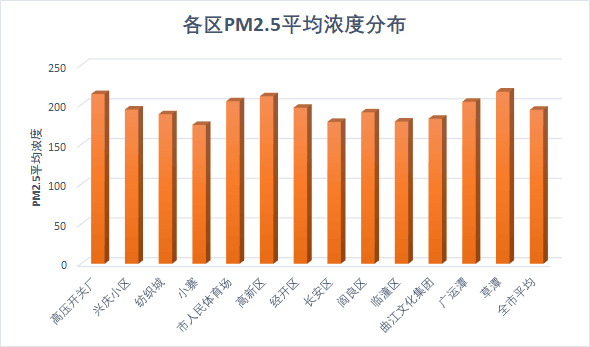


图5.2 西安市各监测点pm2.5的平均浓度分布

由谷歌地图实测得到13个监测点的坐标位置，如表5.1所示，其中x表示横坐标，y表示纵坐标，z表示pm2.5的平均浓度

表5.1 西安市13个监测点的pm2.5平均浓度

|  |  |  |  |
| --- | --- | --- | --- |
|  | 横坐标x | 纵坐标y | pm2.5平均浓度z |
| 高压开关厂 | -6.1 | -0.33 | 213.7544 |
| 兴庆小区 | 2.8 | -1.2 | 194.2075 |
| 纺织城 | 9.7 | -0.425 | 188.2793 |
| 小寨 | -0.7 | -5.2 | 174.9009 |
| 市人民体育场 | 0 | 0 | 204.63 |
| 高新区 | -6.2 | -4.3 | 210.9009 |
| 经开区 | -1.5 | 8.4 | 196.4554 |
| 长安区 | -4.4 | -12.4 | 178.6396 |
| 阎良区 | 25.1 | 43.6 | 190.5586 |
| 临潼区 | 23.8 | 10.8 | 179.1717 |
| 曲江广化区 | 2.1 | -6.7 | 182.6182 |
| 广运潭 | 9.1 | 5 | 204.0278 |
| 草潭 | -8.2 | 11 | 216.8017 |

利用shepard插值，曲面拟合方法求解pm2.5指标的空间分布情况，利用matlab求解，得到pm2.5指标的曲面图和等高线图像，如图5.3所示。

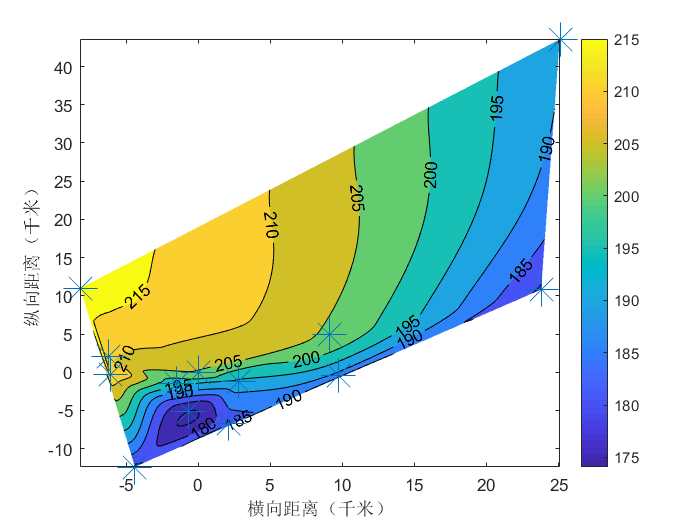
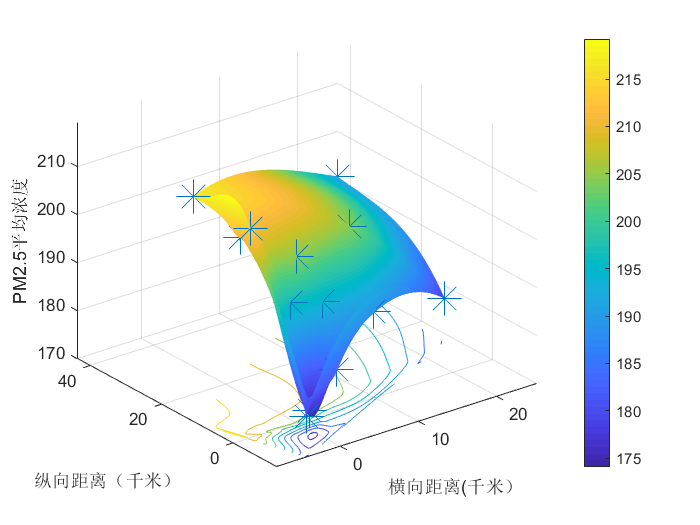


图5.3 PM2.5的空间分布情况

**5.3.1.2区域污染程度评估**

根据参考文献[1],对环境空气功能区分类:

一类：曲江文化集团、广运潭、草滩

二类：高压开关厂、兴庆小区、纺织城、小寨、市人民体育场、高新区、经开区、长安区、阎良区、临潼区

空气质量分指数计算方法：

污染物项目P的空气质量分指数按式（1）计算



其中，IAQIP为污染物项目P的空气质量分指数，Cp为污染物项目P的质量浓度值，BPHi为空气质量分指数及对应的污染物项目浓度限值中与Cp相近的污染物浓度限值的高位值，BPL0为空气质量分指数及对应的污染物项目浓度限值中与Cp相近的污染物浓度限值的低位值，IAQIHi为空气质量分指数对应的污染物项目浓度限值中与BPHi对应的空气质量分指数，IAQIL0为空气质量分指数及对应的污染物项目浓度限值中与BPL0对应的空气质量分指数。

空气质量指数计算方法：

AQI=max{IAQI1,IAQI2,...IAQIn}

其中，IAQI为空气质量分指数，n为污染物项目。

由上，计算出西安市13个检测点的pm2.5的空气质量分指数，如图5.4所示

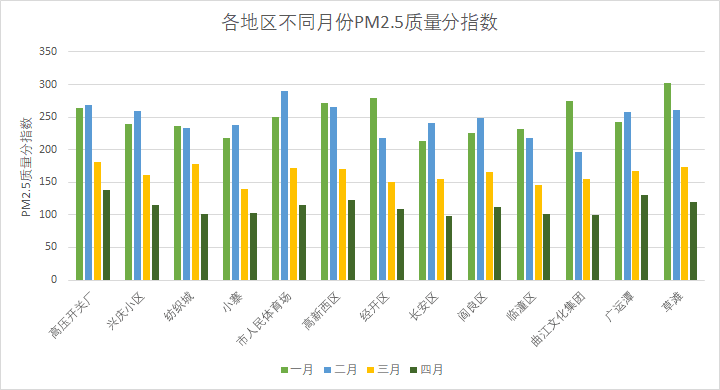
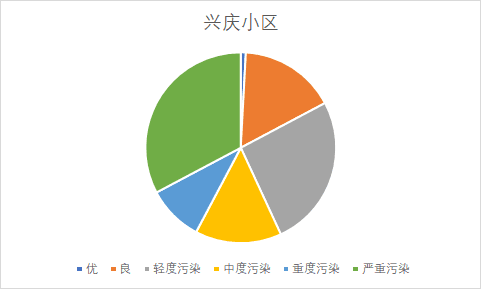
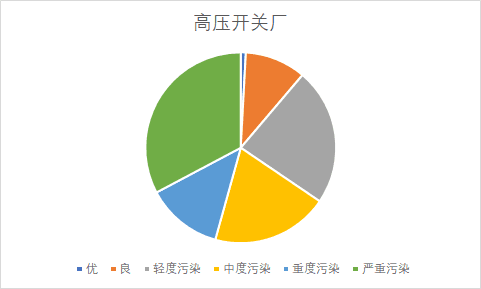


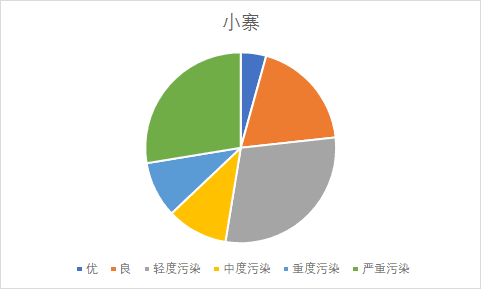
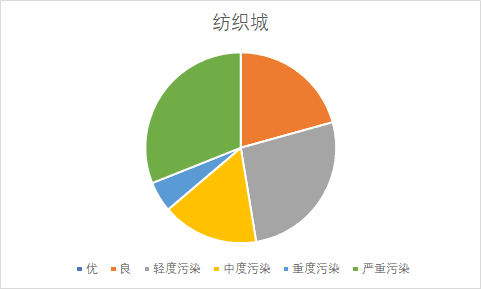
图5.4 西安市13个监测点pm2.5的空气质量分指数

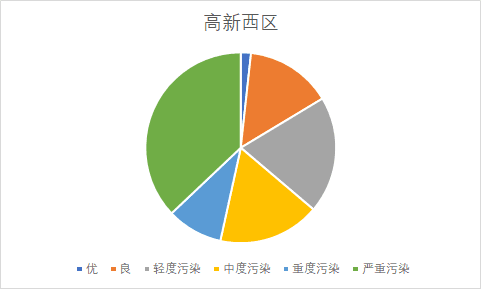
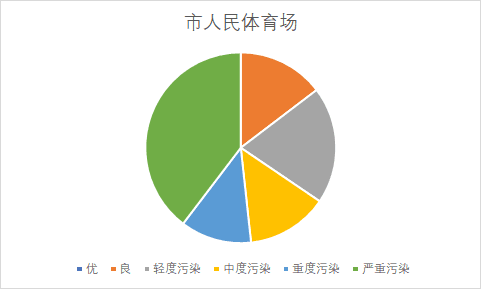
为更直观的反应出西安市13个监测点不同污染程度的天数，将数据放入图和表，如表5.2和图 5.5所示

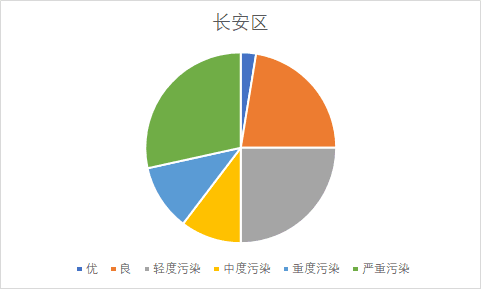
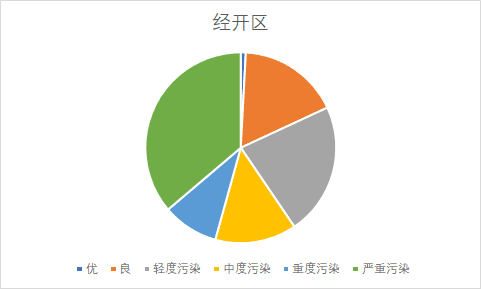
表5.2 西安市13个监测点的污染程度

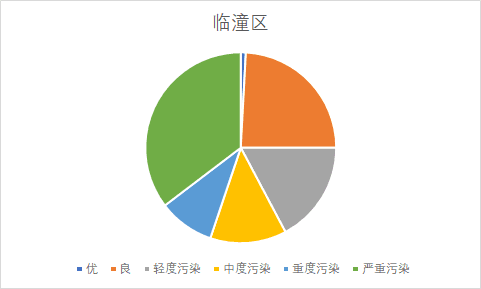
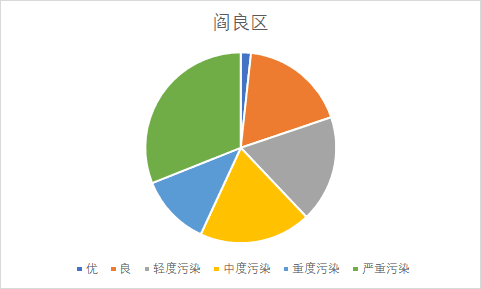
|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | 优 | 良 | 轻度污染 | 中度污染 | 重度污染 | 严重污染 |
| 高压开关厂 | 1 | 12 | 27 | 23 | 15 | 38 |
| 兴庆小区 | 1 | 19 | 30 | 17 | 11 | 38 |
| 纺织城 | 0 | 24 | 31 | 19 | 6 | 36 |
| 小寨 | 5 | 22 | 34 | 12 | 11 | 32 |
| 市人民体育场 | 0 | 17 | 23 | 16 | 14 | 46 |
| 高新区 | 2 | 17 | 23 | 20 | 11 | 43 |
| 经开区 | 1 | 20 | 26 | 16 | 11 | 42 |
| 长安区 | 3 | 26 | 29 | 12 | 13 | 33 |
| 阎良区 | 2 | 21 | 21 | 22 | 14 | 36 |
| 临潼区 | 1 | 28 | 20 | 15 | 11 | 41 |
| 曲江广化区 | 2 | 22 | 36 | 11 | 10 | 35 |
| 广运潭 | 1 | 14 | 28 | 22 | 12 | 39 |
| 草潭 | 1 | 13 | 24 | 22 | 19 | 37 |

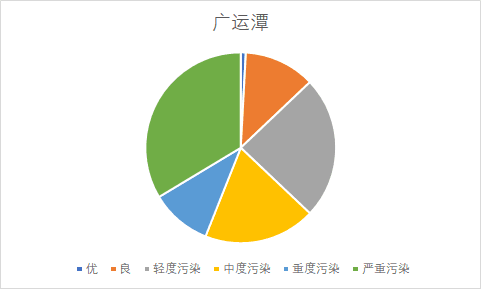
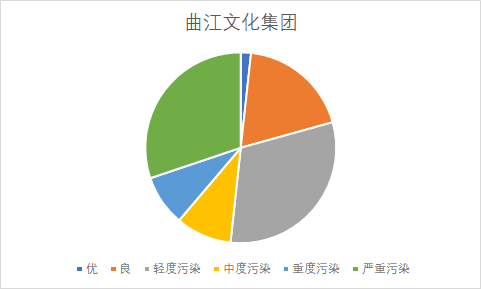












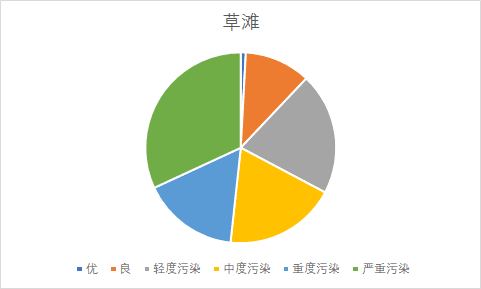


图5.5 西安市13个监测点的污染程度

评价模型：污染指数*n*



其中，*x*为污染程度权重，*Pi*为污染程度所占天数，*i*=1-6，分别指代优、良、轻度污染、中度污染、重度污染、严重污染。经计算，西安市13个监测点的污染指数如表5.3所示

表5.3 西安市13个监测点的污染指数

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 高压开关厂 | 兴庆小区 | 纺织城 | 小寨 | 市人民体育场 | 高新区 | 经开区 | 长安区 | 阎良区 | 临潼区 | 曲江广化区 | 广运潭 | 草潭 |
| 5.01 | 4.8 | 4.63 | 4.46 | 5.13 | 4.98 | 4.9 | 4.53 | 4.81 | 4.78 | 4.58 | 4.95 | 5.04 |

表5.4 污染评判标准

|  |  |  |
| --- | --- | --- |
| 评判机制 | 不包含区间头部（包含尾部）污染指数范围 |  |
| 优 | <4.5 | 小寨 |
| 良 | 4.5-4.65 | 纺织城、长安区、曲江文化集团 |
| 轻度污染 | 4.65-4.8 | 兴庆小区、临潼区 |
| 中度污染 | 4.8-4.95 | 经开区、阎良区、广运潭 |
| 重度污染 | 4.95-5.1 | 高新区、高压开关厂、草潭 |
| 严重污染 | >5.1 | 市人民体育场 |

**5.3.2 PM2.5的发生与演变**

天气与相对湿度的转换情况如下表所示

**表5.4 天气状况与相对湿度转换表**

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

建立多元回归分析模型

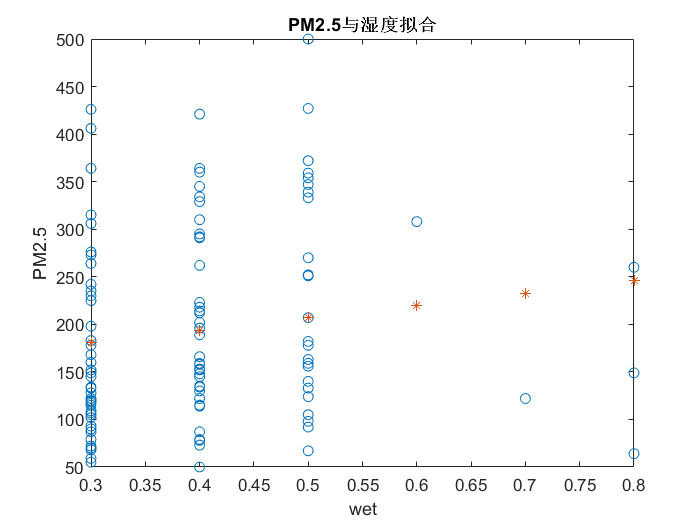
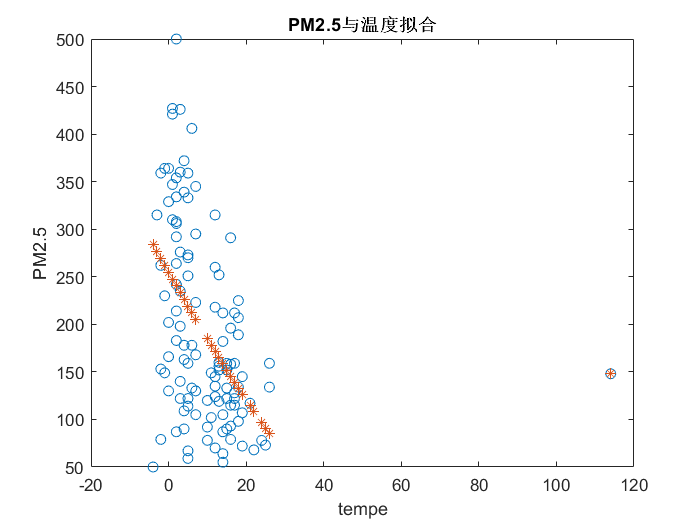
y=β4x4+β3x3+β2x2+β1x1+β0

*x0*=0，*x* 1=0.1351，*x* 2=86.1147，*x* 3=147.1688, *x* 4=51.0547

x1表示温度，x2表示风力，x3表示湿度，x4表示季节，y表示pm2.5。

利用matlab求解，得

y=-51.0547x4+147.1688x3+86.1147x2+0.1351x1



(a)

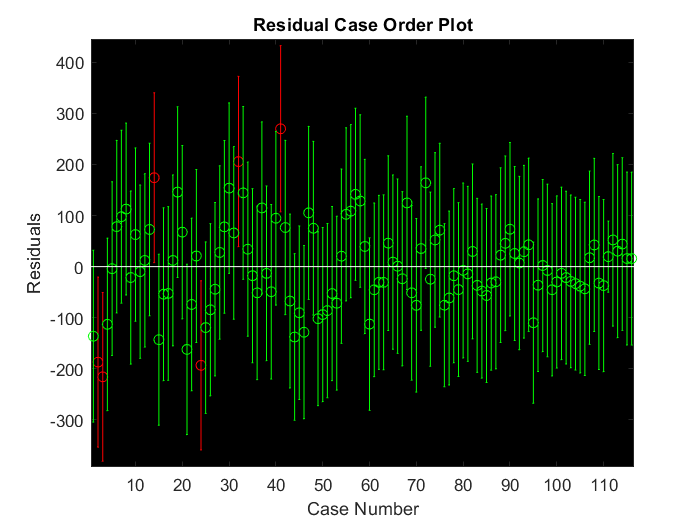
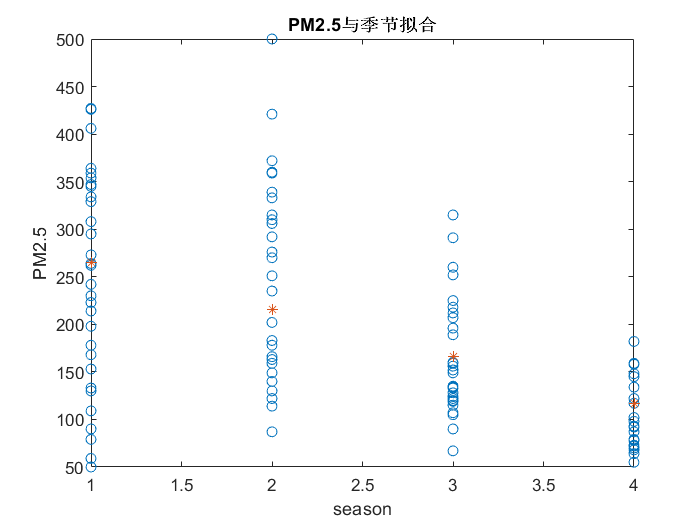


图5.6 pm2.5与温度、湿度、季节拟合情况

5.3.3污染扩散预测与评估方法

高斯烟羽模型：

由正态分布假设可以导出下风向任意一点x(x,y,z)处泄露气体浓度的函数为：

X（x，y，z）=A（x）*e-ay2e-bz2*

由概率统计理论可以写出方差的表达式为



源强的积分公式



式中，σy，σz为泄露气体在y，z方向分布的标准差，单位为m，X（x，y，z）为任一点处泄露气体的浓度，单位为kg/m，u为平均风速，单位为m/s，Q为源强（即泄露速度），单位为kg/s。

将（1）式代入（2）式，积分可得







上式为误解空间连续点源扩散的高斯模型公式，然而在实际中，由于地面的存在，烟雨的扩散是有界的。根据假设可以把地面看作一镜面，对泄露气体起全反射作用，并采用像源法处理，把任意一点p处的浓度看作两部分的贡献之和：一部分是不存在地面时造成的泄露物浓度，一部分时由于地面反射作用增加的泄露物浓度，该处的泄露物浓度相当于不存在地面时由位于（0，0，H）处的实源，和位于（0，0，-H）的像源在p点处所造成的泄露物浓度之和。

其中，实源的贡献为



其中，像源的贡献为



该处的实际浓度为



由上，可得高架连续点源扩散的高斯烟雨模型公式为：



其中，X(x,y,z)为下风向x米，横向y米，地面上方z米处的扩散气体浓度，单位为kg/m，Q为源强，即释放速率，单位为kg/s，u为平均风速，单位为m/s，t为泄露后的时间，单位是s，H为泄漏源有效高度，单位为m，y为横向距离，单位为m，z为垂直方向距离，单位为m。

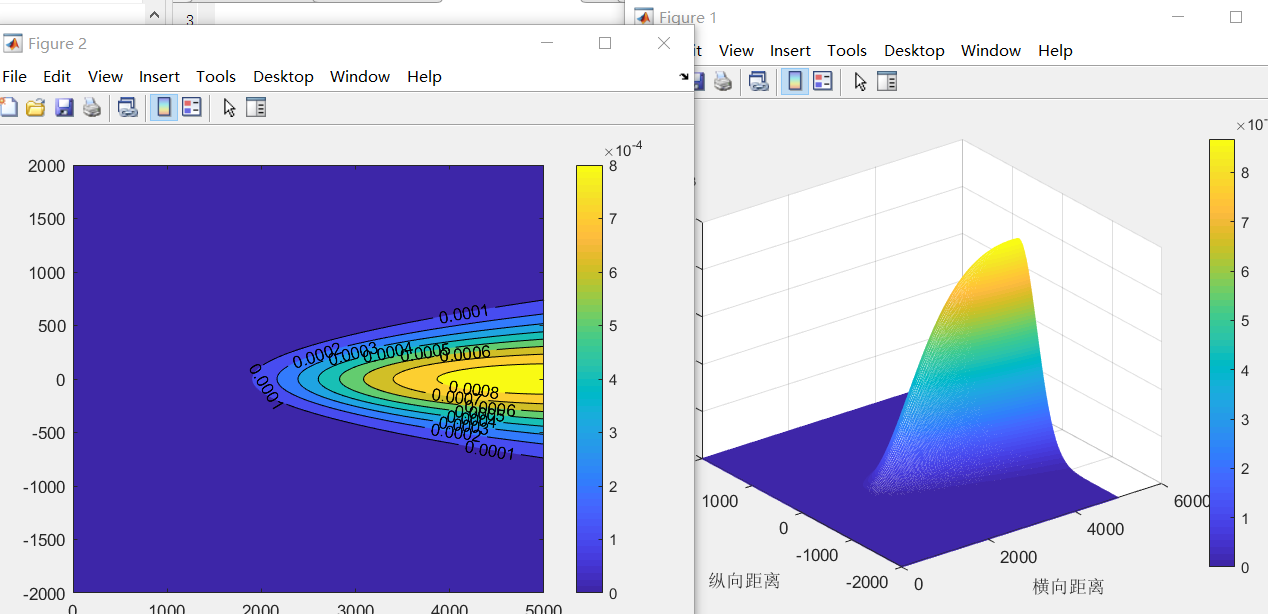


图5.7

5.3.4模型的检验

六．问题三

**6.1模型假设**

1. 五年每年减少的pm2.5浓度量成等比数列分布；
2. 每年都能按预期完成治理任务；
3. 排除题设中其他因素对pm2.5浓度的影响；
4. 按年平均浓度计算，不考虑单位浓度值的浮动。

**6.2 建模分析**

**6.3 模型的的建立与求解**

**6.3.1 pm2.5的治理计划及指标**

效用函数：使减少pm2.5浓度的措施效率尽可能高

u(xi)=xi/(xi+a)

其中，xi为第i年减少的平均浓度，u（xi）为第i年降低浓度措施的效率，a为随机函数，取0.2。

满意度：



其中，F为治理满意度。

约束条件:

（1）五年内，pm2.5的浓度降低至35以内；

（2）每年减少的年平均浓度以等比数列逐年减少；



（3）每年的效用函数值大于0.99.

u(xi)≥0.99，i=1,2,3,4,5

规划模型：

目标函数：

max 

约束条件：



**6.3.2 专项治理计划及其合理性**

投入总费用：z=z1+z2，其中，z为总费用，z1为综合治理费用，z2为专项治理费。

治理满意度：



其中，F1表示综合治理和专项治理共同作用的满意度，u（xi）表示第i年综合治理效用函数，xi表示第i年综合治理降低量，u（yi）表示第i年专项治理效用函数，yi表示第i年专项治理降低量。

约束条件：

（1）五年内pm2.5浓度降低至35，此处取武汉市pm2.5浓度的峰值426



（2）投入费用逐年成等比排列



（3）效用函数：

u(xi)≥0.9

u(yi)≥0.9

i=1,2,3,4,5

建立规划模型：

目标函数：min z=z1+z2，max F=



**6.4模型结论**

**6.4.1**

表6.1 pm2.5的治理指标

|  |  |  |  |
| --- | --- | --- | --- |
|  | 减少量 | 剩余量 | 效用率 |
| 第一年 | 74.72465 | 205.27535 | 99.70% |
| 第二年 | 58.82321 | 146.45214 | 99.70% |
| 第三年 | 46.3056 | 100.14654 | 99.70% |
| 第四年 | 36.45175 | 63.69479 | 99.60% |
| 第五年 | 28.6948 | 34.99999 | 99.50% |

**6.4.2**

表6.2 治理计划方案

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | 综合治理减少量 | 综合效用率 | 专项治理减少量 | 专项效用率 | 剩余量 | 治理费 |
| 第一年 | 19.8 | 99% | 64.56668 | 99.69% | 341.6333 | 55.38869 |
| 第二年 | 19.8 | 99% | 64.56668 | 99.69% | 257.2666 | 55.38869 |
| 第三年 | 19.8 | 99% | 64.56668 | 99.69% | 172.8999 | 55.38869 |
| 第四年 | 19.8 | 99% | 39.74944 | 99.50% | 113.3505 | 37.53068 |
| 第五年 | 14.51522 | 99% | 63.83529 | 99.69% | 35.0000 | 45.20924 |

七．模型总结与评价

（误差分析；

模型的进一步推广）

参考文献