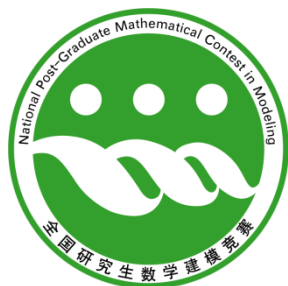


参赛密码 _____
(由组委会填写)



第十届华为杯全国研究生数学建模竞赛

题 目 PM2.5 演变评估模型及治理方案研究

摘 要：

本文针对日益严重的 PM2.5 大气污染问题，基于现有数据与相关研究，采用相关分析、逐步回归和有限元等方法，对 AQI 指标之间的相关性进行了定量分析，针对 PM2.5 成因、时空分布、演变规律和应急处理建立了数学模型，结合气象理论知识，进行了计算结果的定性与定量分析，并以武汉市为例给出了合理的 PM2.5 五年治理方案。

问题一：基于武汉与西安市现有的空气质量监测 AQI 指标数据，应用相关分析，分别求得两城市的 6 个 AQI 主要指标的相关矩阵，对两组数据中 PM2.5 与另 5 个 AQI 指标的相关性进行了定性与定量分析，得出了一致的结论，即 PM2.5 与另 5 个主要指标具有相关性，并与 SO₂、NO₂、PM10 及 CO 显著线性相关（相关系数均明显大于 0.7）。

问题二：首先应用回归分析方法，针对 2013 年仅有的 AQI 完整数据，完成了 PM2.5 关于 API 中前 3 个指标数据的二次模型拟合，拟合统计量表明该拟合是高度显著的。并基于大量 API 历史数据，计算每组 API 数据对应的 PM2.5 估算值。而后，利用 PM2.5 实测数据分析 PM2.5 浓度的时空分布特征，进而结合环境保护部新修订《环境空气质量标准》，对该地区分区进行了污染评估。综合多种气象因素，基于多元回归建立关于 PM2.5 演变规律的数学模型。结果表明，模型估计值与实测数据差值较小，模型参数与实际情况基本一致，验证了模型的合理性和有效性。针对浓度扩散问题，本文在适当简化实际问题的基础上，建立了描述 PM2.5 浓度扩散现象的偏微分方程，通过采用商用有限元软件 Comsol Multiphysics，建立了针对该问题的扩散预测与评估计算平台，并对 PM2.5 扩散现象进行了分析研究，得到了浓度突增等突发事件对 PM2.5 分布影响的几个重

要规律。主要包括速度和扩散系数的取值对扩散趋势的影响。并在此基础上，基于有限元仿真平台，运用自行推导的近似公式得到了扩散系数的估计值，对测点浓度突增现象进行了仿真计算，得到了 PM2.5 污染预测和评估结果。最后，从模型假设条件和模型主要结论两方面，分析了模型的合理性，并在前文研究的基础上对 PM2.5 的成因和演变规律进行了较系统的分析和探究。

问题三：首先应用逐步回归分析方法获得了 PM2.5 与 3 个既与综合治理又与专项治理有关因素 SO_2 、 NO_2 、PM10 的最优二次回归方程，在将治理基准年与五年后目标的 PM2.5 浓度折算为对应的 AQI 指标之后，结合数据分析计算出了基准年的 PM2.5、 SO_2 、 NO_2 、PM10 这 4 个基准指标值。然后按照每年这 4 个指标变化最小但仍能满足治理结果约束的原则，建立了非线性规划模型，并通过将非线性约束条件简化为线性约束，将模型转化为标准的非线性规划问题，进而求出了可行的每年全年年终平均治理指标。最后，结合 PM2.5 综合治理与专项治理的费用经验公式，按照五年治理投入总经费最小的原则建立了一个标准的非线性规划模型，并求解出了最少的五年总投入经费及每年经费投入。对于指导 PM2.5 的治理工作具有一定的实际意义。

关键词：PM2.5，相关分析，逐步回归，非线性规划，有限元

一、问题重述与分析

1.1 问题重述

随着空气污染的加剧,以及广大民众对生活环境与质量要求的日益提高,产生灰霾的主要因素——PM2.5 的浓度指标越来越受到广大民众与政府相关部门的重视,国内外的相关研究亦已逐步开展^[1-5]。

我国以 SO₂、NO₂ 与 PM10 这 3 个监测指标为主的 API 已于 2013 年升级为增加了 PM2.5 等 3 个指标共 6 个指标的 AQI。由于 2013 年之前我国并没有系统的 PM2.5 等监测数据,因此对 PM2.5 的成因、与其它污染指标间关系,以及 PM2.5 的预报、防治等问题仍有待于深入研究。本题就是关于上述问题建立相关的数学模型进行分析与解决^[3,5]。

问题一: PM2.5 相关因素分析。分析 AQI 中 6 个监测指标之间的相关与独立性进行定量分析,尤其是 PM2.5 与其它 5 个指标之间的相关分析。

问题二: 利用附件 2 与附件 3 中的数据,研究 PM2.5 的时空分布、演变及应急处理方案。

1) 描述西安市内 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;

2) 建立能够准确描述西安市 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,并进行定量与定性分析;

3) 假设西安市某监测点处的 PM2.5 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 小时的情况下,利用所建立的模型进行预测评估,给出重度污染和可能安全区域。

4) 采用适当方法检验所建立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。

问题三: 针对武汉市空气污染问题,采用综合治理和专项治理相结合的方式,制定每年的 PM2.5 治理指标方案,使之在五年内达到治理目标。

1) 该地区目前 PM2.5 的年平均浓度估计为 280 $\mu\text{g}/\text{m}^3$,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 35 $\mu\text{g}/\text{m}^3$,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) 若每减少一个 PM2.5 浓度单位,当年需投入一个百万元,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),设计西安市该监测点所在地区有效的专项治理计划,使之达到预定 PM2.5 减排计划,同时经费投入较为合理,要求给出五年投入总经费和逐年经费投入预算计划,并论述其的合理性。

1.2 问题分析

通过对所述问题,尤其是已有的相关研究结论、条件与数据的初步分析,针对三个主要问题可以分别进行研究与验证。

首先,PM2.5 成因与其它 5 个 AQI 指标,尤其是与前 3 个 API 指标,即 SO₂、NO₂ 与 PM10 这 3 个监测指标之间的相关性,可以应用已有的武汉市与西安市的相关数据进行相关统计分析与相关检验等,从而得以由监测数据分析的角度验证 PM2.5 的成因。

其次,关于西安市 PM2.5 时空分布模型的建立,由于西安市 PM2.5 仅有 2013 年首 4 个月的数据,不足以构造其全年的时空分布模型,故必须设法补充 2010 年 1 月至 2012 年 12 月的 PM2.5 数据;然后,通过分析西安市 PM2.5 的时空分布规律,进而构建 PM2.5 的发生和演变的数学模型,并在 PM2.5 浓度激增并持续数小时的突发情况下,建立污染扩散预测与评估的方法,最后检验所建模型与方法的合理性,以及探索 PM2.5 成因演变等的一般性规律。

最后,建立关于决定 PM2.5 浓度指标的相关可控因素的数学模型,并计算出五年内年均 PM2.5 浓度由 280 逐年降至 35 的可控因素指标的控制量,进而根据专项治理费用与 PM2.5 减少的经验公式计算出每年所需的费用金额,从而构建一个较为全面、合理的 PM2.5 治理方案。

二、模型的假设

根据有关研究与数理统计学等原理,在此对本文模型,尤其是对相关监测数据等做以下假设:

1) 所有 6 个 AQI 主要指标,以及西安市由于将 O₃ 在两个不同时间进行监测而形成的 7 个指标,其监测是相互独立的;

2) 所有监测数据的误差均相互独立且服从正态分布 $N(0,\sigma)$;

3) 监测数据是用统一的设备配置与标准获得的,即不同地点,尤其是同一城市中不同监测点的数据是一致的;

4) AQI 指标中的二氧化硫(SO₂),二氧化氮(NO₂)与一氧化碳(CO)在一定程度上可以决定 PM2.5 的形成;

5) PM10 与 PM2.5 的监测是独立的,但这两指标可能存在一定的相关性,有待统计检验;

6) 原 API 中的二氧化硫(SO₂),二氧化氮(NO₂)与 PM10 的监测数据,与 2013 年之后 AQI 中获得上述 3 个指标监测数据的设备配置与标准是相同的,即 2012 年及以前的 API 数据可用于 AQI 中 PM2.5 的数据分析;

7) PM2.5 在空气中的扩散问题,是一个复杂的物理化学过程,但从工程的角度而言,不可能把所有的因素均考虑在内,故作假设如下:

A. 气体流动是稳态的,且空气考虑为不可压缩流体;

B. 在选取的足够大的研究区域内,沿主风向的 2 个侧面近似为对称面;

C. 污染源近似为点源，污染源排放的污染物在短时间内没有变化；

D. 污染源为孤立源；

E. 在污染物的扩散过程中，不考虑污染物的干湿沉降，不考虑污染物扩散过程中的化学变化过程，只将其看作是一种污染物的单纯物理扩散过程；

F. 由于缺乏对近地面温度梯度分布的观测，温度对污染物垂直扩散速率的影响也不予考虑，且同样忽略地形等条件造成的空间扩散速率的各向异性差异，近似认为在各个方向上具有同样的扩散系数。

三、符号说明

本文模型涉及到概率论、多元统计分析、有限元等较多研究方向^[3-5,7]，涉及的专用符号较多，现将本文使用的主要符号列出如下：

σ : 方差

r : 相关系数

α : 显著性水平，一般设为 0.05 或更小，本文设为 0.01

R : 协方差矩阵

β : 回归方程的系数向量

y : PM2.5 的 AQI 指标值

x_1 : SO₂ 的 AQI 指标值

x_2 : NO₂ 的 AQI 指标值

x_3 : PM10 的 AQI 指标值

四、模型的建立与求解

4.1 PM2.5 的相关因素分析模型与分析结论

根据多元统计分析原理，两个随机向量 X 与 Y 的 Pearson 相关系数 r_{xy} 可由下式获得^[9]：

$$r = \frac{\sum XY - \frac{\sum X \sum Y}{N}}{\sqrt{(\sum X^2 - \frac{(\sum X)^2}{N})(\sum Y^2 - \frac{(\sum Y)^2}{N})}} \quad (1)$$

其中， N 是样本个数，所有求和亦为从 1 到 N ；相关系数 r 的值介于 -1 与 +1 之间，即 $-1 \leq r \leq +1$ 。其性质如下^[1-3]：

- 1) 当 $r > 0$ 时，表示两变量正相关， $r < 0$ 时，两变量为负相关；
- 2) 当 $|r| = 1$ 时，表示两变量为完全线性相关，即为函数关系；
- 3) 当 $r = 0$ 时，表示两变量间无线性相关关系；

- 4) 当 $0 < |r| < 1$ 时, 表示两变量存在一定程度的线性相关。且 $|r|$ 越接近 1, 两变量间线性关系越密切; $|r|$ 越接近于 0, 表示两变量的线性相关越弱。一般可按三级划分:

- A. $|r| < 0.4$ 为低度线性相关;
- B. $0.4 \leq |r| < 0.7$ 为显著性相关;
- C. $0.7 \leq |r| < 1$ 为高度线性相关。

4.1.1 基于武汉市一个监测点数据的 PM2.5 的相关因素分析

由于 PM2.5 的监测数据确实很难找到更完整的数据, 本文首先采用本题附件 1 中武汉市一个监测点的数据 (而后采用附件 2 中由 13 个监测点获得的西安市平均数据, 再做一次比较研究), 并经删除数据缺失行等数据预处理, 得到可用于统计分析的数据集, 如下表 1 所示 (由于篇幅限制, 仅示意性列出最前两次与最后两次观测值)。

表 1 武汉市一个监测点数据 (2013. 1. 1—2013. 8. 26)

SO ₂	NO ₂	PM10	CO	O ₃	PM2.5
53	89	76	19	30	90
47	83	88	29	8	143
.....
7	39	56	23	41	58
7	44	62	24	50	74

上述数据经去除不完整数据等预处理后, 由 Matlab 的 corrcoef 函数 (设显著性水平 $\alpha=0.01$) 求得武汉市该观测点数据中 SO₂ 与 PM2.5 等 6 种污染指标的相关矩阵如下所示:

$$R_1 = \begin{bmatrix} 1.0000 & 0.8051 & 0.6779 & 0.6570 & -0.1815 & 0.7241 \\ 0.8051 & 1.0000 & 0.7275 & 0.6245 & -0.0646 & 0.7324 \\ 0.6779 & 0.7275 & 1.0000 & 0.5860 & -0.0685 & 0.7787 \\ 0.6570 & 0.6245 & 0.5860 & 1.0000 & -0.3830 & 0.8216 \\ -0.1815 & -0.0646 & -0.0685 & -0.3830 & 1.0000 & -0.3549 \\ 0.7241 & 0.7324 & 0.7787 & 0.8216 & -0.3549 & 1.0000 \end{bmatrix} \quad (2)$$

即 PM2.5 与 SO₂、NO₂、PM10、CO 与 O₃ 的相关系数分别为:

$$r_{6\bullet} = (0.7241 \quad 0.7324 \quad 0.7787 \quad 0.8216 \quad -0.3549) \quad (3)$$

同时获得对应的显著性检验矩阵如下:

$$P_1 = \begin{bmatrix} 1.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0052 & 0.0000 \\ 0.0000 & 1.0000 & 0.0000 & 0.0000 & 0.3232 & 0.0000 \\ 0.0000 & 0.0000 & 1.0000 & 0.0000 & 0.2946 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 1.0000 & 0.0000 & 0.0000 \\ 0.0052 & 0.3232 & 0.2946 & 0.0000 & 1.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 1.0000 \end{bmatrix} \quad (4)$$

由于(3)式中前 4 个相关系数均大于 0.7，由相关系数的性质可知，PM2.5 与 SO₂、NO₂、PM10、CO 均为高度线性相关；PM2.5 与 O₃ 是低度线性相关，且为负相关。

由(4)式最后一行（除最后一个元素外）的元素均小于 0.00005 < α，根据 Matlab 及统计学原理，说明 PM2.5 与其它 5 个污染指标的线性相关性在显著性水平 α=0.01 上，均是显著的。

为了更形象而具体地分析 PM2.5 与其它指标之间，尤其是与 API 中亦包含的 SO₂、NO₂ 与 PM10 之间的关系，本文对上述指标分别进行了两两数据拟合，如图 1~3 所示，图 4 给出拟合的残差分析图，其中 y, x₁, x₂, x₃ 分别表示 PM2.5、SO₂、NO₂ 与 PM10 的 AQI 指标值（本文下同，不再赘述）。

尽管(3)式给出了的 PM2.5 与 SO₂、NO₂、PM10 及 CO 的显著线性相关性（相关系数都明显大于 0.7），但由这些图中可以看出，PM2.5 与这 4 个指标之间确实主要是线性关系，但亦与高次，至少与二次项仍具有一定关系，因此在进一步研究中不应完全忽略二次项的影响。

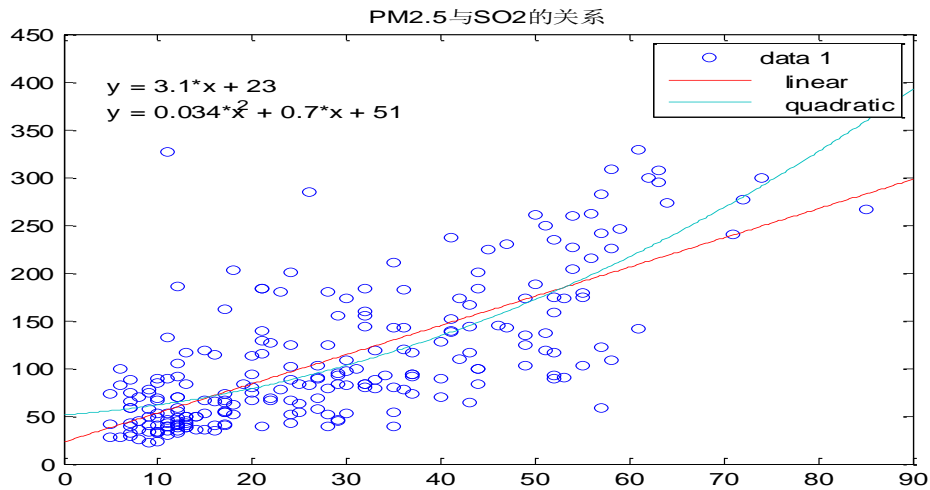


图 1 PM2.5 与 SO₂ 之间的拟合及关系图

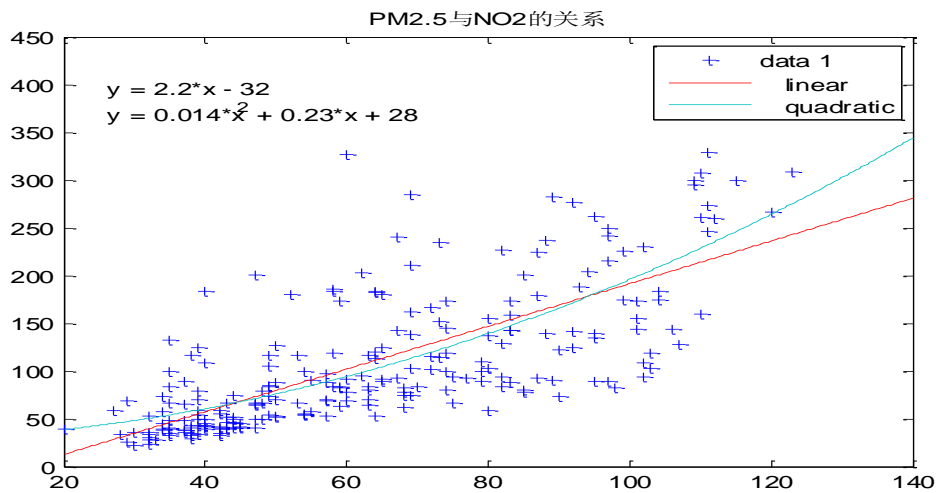


图 2 PM2.5 与 NO₂ 之间的拟合及关系图

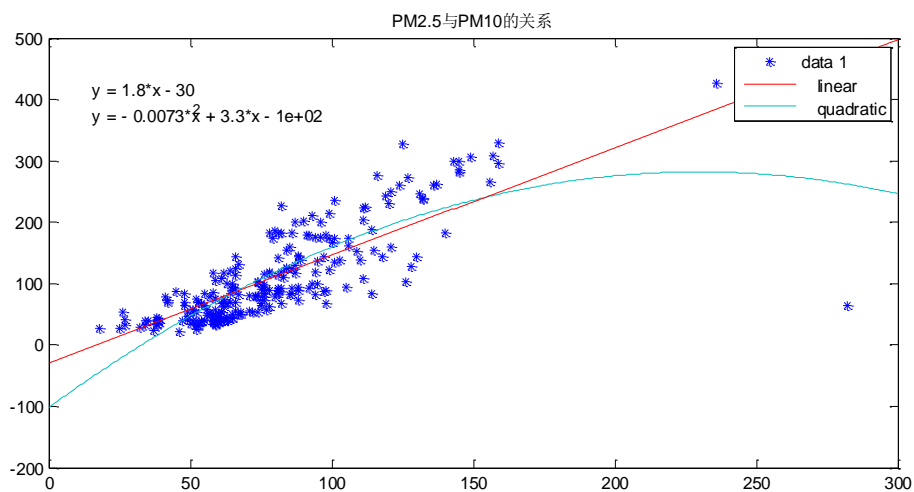


图 3 PM2.5 与 PM10 之间的拟合及关系图

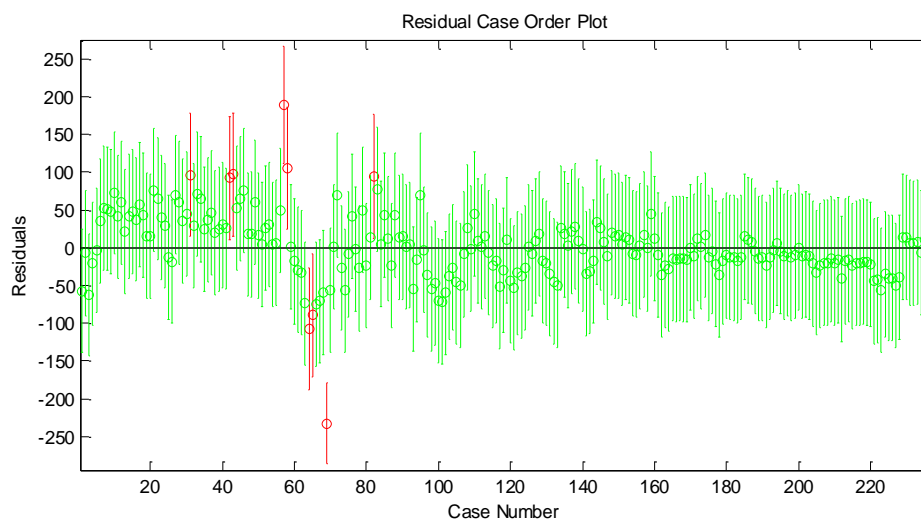


图 4 武汉市 PM2.5 拟合残差分析图

4.1.2 基于西安市13个监测点数据的PM2.5的相关因素分析

同于上述的分析方法，本文将本题附件2中由西安市13个监测点获得的“全市平均”数据，采用同与上文的模型，计算分析如下。

首先，对附件2中数据所含有7项AQI污染指标数据，即前117行，进行去除个别数据缺失行等预处理，得到完整的数据矩阵，限于篇幅，不在此一一罗列，请详见附件。

应用本节前述模型，由Matlab计算得到相关系数矩阵与显著性检验矩阵如下所示：

$$R_2 = \begin{bmatrix} 1.0000 & 0.4953 & 0.3853 & 0.6470 & -0.4538 & -0.5357 & 0.6741 \\ 0.4953 & 1.0000 & 0.3823 & 0.3761 & -0.1011 & -0.0936 & 0.5190 \\ 0.3853 & 0.3823 & 1.0000 & 0.2990 & -0.2881 & -0.2928 & 0.7433 \\ 0.6470 & 0.3761 & 0.2990 & 1.0000 & -0.4745 & -0.4548 & 0.7417 \\ -0.4538 & -0.1011 & -0.2881 & -0.4745 & 1.0000 & 0.9134 & -0.4421 \\ -0.5357 & -0.0936 & -0.2928 & -0.4548 & 0.9134 & 1.0000 & -0.4658 \\ 0.6741 & 0.5190 & 0.7433 & 0.7417 & -0.4421 & -0.4658 & 1.0000 \end{bmatrix} \quad (5)$$

由上式最后一行得到PM2.5与SO₂、NO₂、PM10、CO、O₃（1小时）与O₃（8小时）的相关系数分别为：

$$r_{\bullet} = (0.6741 \quad 0.5190 \quad 0.7433 \quad 0.7417 \quad -0.4421 \quad -0.4658) \quad (6)$$

同时获得对应的显著性检验矩阵如下：

$$P_2 = \begin{bmatrix} 1.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 1.0000 & 0.0000 & 0.0000 & 0.2822 & 0.3198 & 0.0000 \\ 0.0000 & 0.0000 & 1.0000 & 0.0012 & 0.0018 & 0.0015 & 0.0000 \\ 0.0000 & 0.0000 & 0.0012 & 1.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.2822 & 0.0018 & 0.0000 & 1.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.3198 & 0.0015 & 0.0000 & 0.0000 & 1.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 1.0000 \end{bmatrix} \quad (7)$$

由上述矩阵，可以同样分析出各因素与PM2.5的相关性，以及各因素之间的相关性，其结论与对武汉市1个监测点数据的分析几乎完全一致。仅有一点除外，即该附件2中数据的第5列与第6列均为O₃，于两个不同时间的监测值。此外，由上述相关矩阵明显表明该2个因素的相关系数高达0.9134，是非常显著的线性相关。

4.2 PM2.5的时空分布模型与演变及应急处理

首先，由于AQI今年刚刚在部分城市实施，关于PM2.5等的的数据由于未含在传统的API当中，而使PM2.5数据严重缺失。如在附件2中给的西安市13个监测点的大气污染指标数据中，除2013年的117组左右的数据包含了PM2.5之

外，后面的 2012 年 12 月 31 日之前的数据均没有 PM2.5，亦没有 CO 与 O₃，仅含 SO₂、NO₂ 与 PM10 这三组数据。而仅凭 2013 年这 117 组左右的数据又不可能完全反映 PM2.5 的时空分布，所以，必须采用适当的方法来处理大量 PM2.5 数据的缺失。

根据上一节的分析结论，我们知道 PM2.5 与 SO₂、NO₂ 与 PM10 这 3 个 API 指标都显著线性相关，而且，这 3 个指标之间以与另 2 个指标之间亦存在一定的相关性，甚至于亦有显著相关。而回归模型中的交叉乘积项，如 x_1x_3 恰能表示这两因素之间的交互关系。同时，由上一节中 PM2.5 与其它 3 个 API 指标的关系图亦不难看出，若用二次函数表达它们之间的关系应更为准确全面。因此，本文采用含多元一次、多元二次等模型的逐步回归来拟合这 3 个 API 指标与 PM2.5 的方程。当然，对于拟求解参数而言，这些回归模型仍是线性模型，如下式所示：

$$y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_3 + \beta_4x_1^2 + \beta_5x_1x_2 + \beta_6x_1x_3 + \beta_7x_2^2 + \beta_8x_2x_3 + \beta_9x_3^2 \quad (8)$$

其中， y 表示 PM2.5， x_1, x_2 与 x_3 分别表示 SO₂、NO₂ 与 PM10 这 3 个 API 指标。

设显著性水平 $\alpha=0.01$ ，应用 Matlab 的 regress(y,X, α)函数，对多个候选回归模型进行逐步回归。比如对一次与二次多项式模型，分别求得回归方程如下方程 (9)与方程(10)。

$$y = -61.9599 + 1.8783x_1 + 0.5380x_2 + 0.5761x_3 \quad (9)$$

同时求得上述方程(9)的可决系数 $r^2=0.73717$ ，统计量 $F=103.77512$ ，显著性概率 $p=0.00000$ ，模型方差 $\sigma^2=2864.62301$ 。

$$\begin{aligned} y = & -112.9848 + 1.1323x_1 + 2.1827x_2 + 0.9926x_3 - 0.0075x_1^2 \\ & - 0.0033x_1x_2 + 0.0097x_1x_3 - 0.0161x_2^2 + 0.0011x_2x_3 - 0.0019x_3^2 \end{aligned} \quad (10)$$

同样求得上述方程的可决系数 $r^2=0.79251$ ，统计量 $F=44.56124$ ，显著性概率 $p=0.00000$ ，模型方差 $\sigma^2=2390.67325$ 。说明该回归方程是显著的。该模型的残差分析图如下图所示，该图说明残差完全在可控且可接受的范围之内。

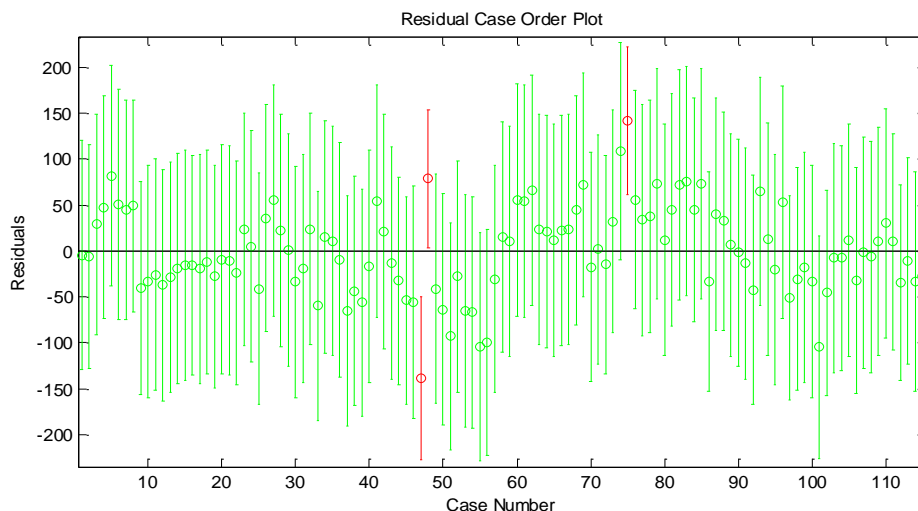


图 5 西安市 PM2.5 拟合残差分析图

根据回归原理与方法^[9]，由于二次模型方程(10)的可决系数 $r^2=0.79251$ 明显大于方程(9)的一次模型的可决系数 $r^2=0.73717$ ，并经与其它回归模型的可决系数等统计量比较（限于篇幅，在此不再罗列）后，可以确定方程(10)的二次模型是逐步回归所确定的最佳回归方程。

进而，采用上述的回归方程(10)，即可分别采用西安市 13 个观测站的第 118 行及以后的观测数据中的前 3 个 API 指标，分别计算出该时间与地点对应同一行观测数据的 PM2.5 估计值，限于篇幅，数据表不再此一一罗列，请详见附件。

4.2.1 PM2.5 的时空分布及其规律分析

1) PM2.5 时间分布及规律

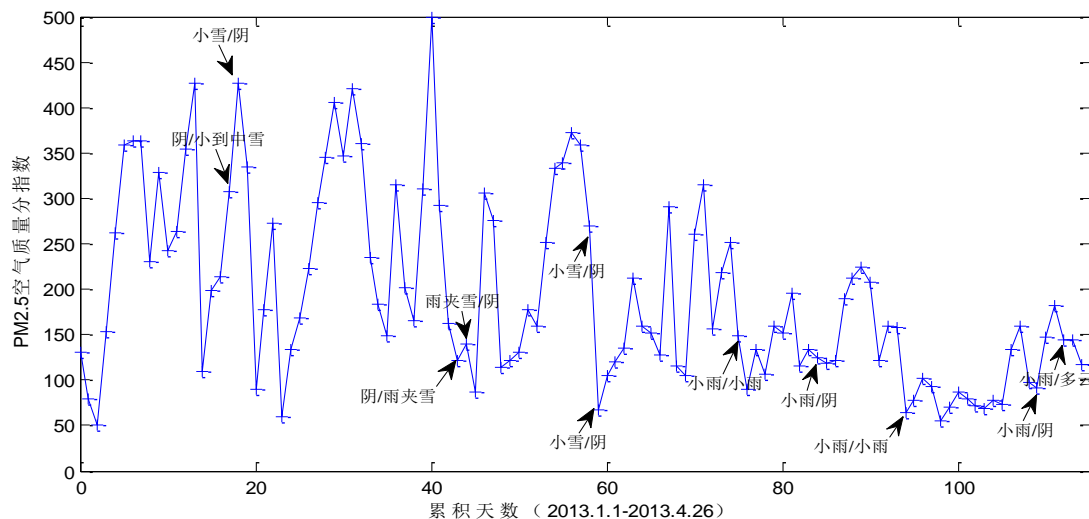


图 6 全市 2013.1.1-2013.4.26PM2.5 分指数演变轨迹

图 6 所示，是西安市 1 月到 4 月 PM2.5 空气质量分指数的演变轨迹。从图中可以观察到，1、2 月份 PM2.5 浓度较高，3、4 月逐渐降低。考虑到西安地处西北，分析气象资料得到以下结论：

A. 西安 1、2 月份的气温在 10 度以下，晚上气温一般在 0 摄氏度以下。

B. 西安 1、2 月份静风频率较高，风力一般小于 3 级，且无持续风向。

C. 西安 1、2 月份多阴天和雨雪天气，湿度比较大。

参照结论 A 并查阅资料可知，1、2 月份气温较低，为北方的采暖期，大量化石燃料的燃烧，导致 PM2.5 质量浓度处于较高水平。3、4 月份，西安气温回升，采暖需求减少，PM2.5 排放量减少。

参考结论 B，1、2 月份西安静风频率高，大气层结稳定，不利于污染物扩散，致使污染物容易积累，从而导致 PM2.5 浓度升高。

至于结论 C 的影响，需要结合演变轨迹的“尖劈”分析。上述的分析显示，1、2 月份的 PM2.5 浓度总体应该保持一个较高水平，但是演变轨迹显示，在 1、2 月份，PM2.5 浓度累积到一定程度，然后快速下降，衰减到一个较低水平。所以演变轨迹上呈现出一个个尖峰，本文将之称之为“尖劈”，以形容其快速上升和下降的特征。

结合图 6 和表 2 分析，可得到，雨雪天气能够引起 PM2.5 浓度的变化。但是这个影响不仅可能是正相关，也可能是负相关。多数情况下，降雨或降雪会导致空气中的微小颗粒聚集成大颗粒，并伴随雨雪沉降到地面，从而导致 PM2.5 浓度的显著下降，形成“尖劈”。

另一方面，从风力角度分析，如图 7 和表 2 所示，多数刮风天气，由于空气流动快，污染物不易聚集，PM2.5 的浓度比较低。而且，下雨或下雪天气往往伴随着刮风，两种因素叠加，对 PM2.5 的清除作用最为显著，如图 7 中的点(59, 67)、(75, 149)、(94, 64)、(109, 92)。

表 2 1-4 月份雨雪天气列表

编号	日期	天气	最高/低温度	风力和风向
17	2013/1/19	阴 /小到中雪	5℃ / -1℃	南风 ≤3 级 /南风 ≤3 级
18	2013/1/20	小雪 /阴	3℃ / -2℃	南风 ≤3 级 /西南风 ≤3 级
43	2012/2/14	阴 /雨夹雪	7℃ / 1℃	无持续风向 ≤3 级 /无持续风向 ≤3 级
44	2012/2/15	雨夹雪 /阴	6℃ / 0℃	无持续风向 ≤3 级 /无持续风向 ≤3 级
59	2013/3/1	小雪 /阴	8℃ / 2℃	北风 3-4 级 /北风 ≤3 级
70	2013/3/11	阴 /小雨	17℃ / 6℃	东风 ≤3 级 /西南风 ≤3 级
75	2013/3/16	小雨 /小雨	13℃ / 9℃	西南风 ≤3 级 /西北风 3-4 级

84	2013/3/25	小雨 /阴	17℃ / 7℃	西南风 ≤3 级 /东风 ≤3 级
94	2013/4/4	小雨 /小雨	18℃ / 9℃	西南风 ≤3 级 /西风 3-4 级
109	2013/4/19	小雨 /阴	14℃ / 6℃	东风 5-6 级 /东风 ≤3 级
112	2013/4/22	小雨 /多云	15℃ / 9℃	南风 ≤3 级 /西风 ≤3 级

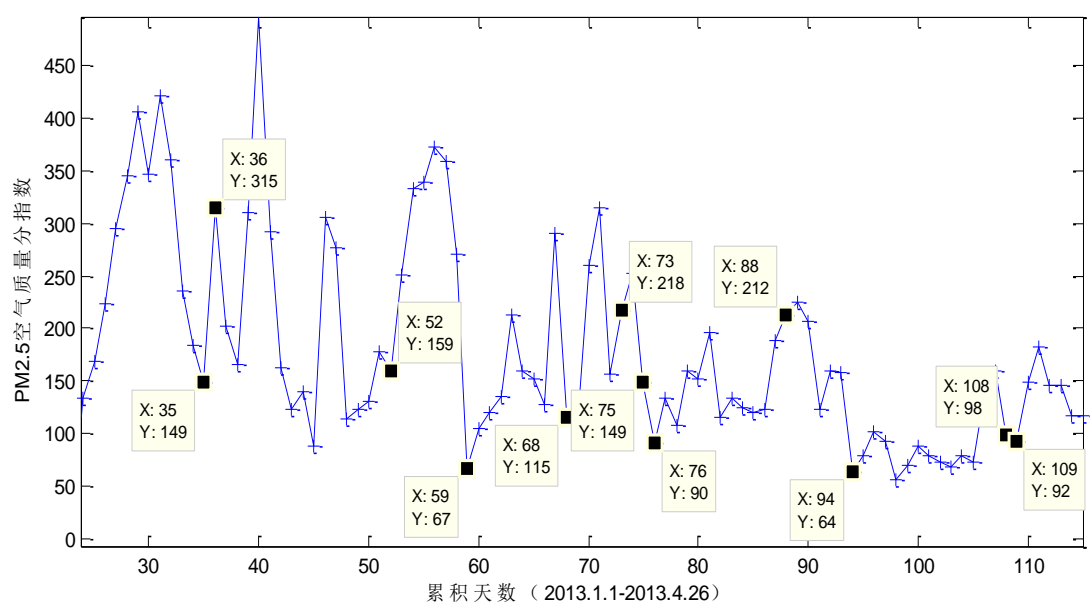


图 7 PM2.5 演变轨迹+标记三级以上风力

表 3 三级以上风力列表

序号	日期	天气	温度	风力和风向
35	2012/2/6	多云 /晴	4℃ / -6℃	北风 3-4 级 /北风 3-4 级
36	2012/2/7	晴 /晴	2℃ / -7℃	北风 3-4 级 /无持续风向 ≤3 级
52	2012/2/23	阴 /阴	8℃ / 1℃	无持续风向 ≤3 级 /东风 3-4 级
59	2013/3/1	小雪 /阴	8℃ / 2℃	北风 3-4 级 /北风 ≤3 级
68	2013/3/9	多云 /浮尘	22℃ / 9℃	西北风 3-4 级 /北风 3-4 级
73	2013/3/14	多云 /多云	16℃ / 7℃	东风 3-4 级 /东风 ≤3 级
75	2013/3/16	小雨 /小雨	13℃ / 9℃	西南风 ≤3 级 /西北风 3-4 级

76	2013/3/17	多云 /晴	21℃ / 8℃	西风 3-4 级 /西北风 3-4 级
88	2013/3/29	阴 /多云	23℃ / 10℃	南风 3-4 级 /东风 ≤3 级
94	2013/4/4	小雨 /小雨	18℃ / 9℃	西南风 ≤3 级 /西风 3-4 级
108	2013/4/18	阴 /阴	24℃ / 12℃	北风 3-4 级 /东风 5-6 级
109	2013/4/19	小雨 /阴	14℃ / 6℃	东风 5-6 级 /东风 ≤3 级

借鉴 HJ633-2012 环境空气质量指数(AQI)技术规定（试行）中表 3 中空气质量指数级别划分，对 13 年西安市区前 4 个月的 PM2.5 分指数分别按月统计，统计结果如图 8 所示。

从各月 PM2.5 分布可以看出，1 月 PM2.5 浓度超三级的天数超过当月有效统计天数的 75%，2 月 PM2.5 浓度超三级的天数接近当月有效统计天数的 75%，3 月超三级天数少于当月有效统计天数的 50%，4 月超三级天数少于当月有效统计天数的 25%。单就 PM2.5 分指数而言，四月空气质量最好，三月次之，一月最差。

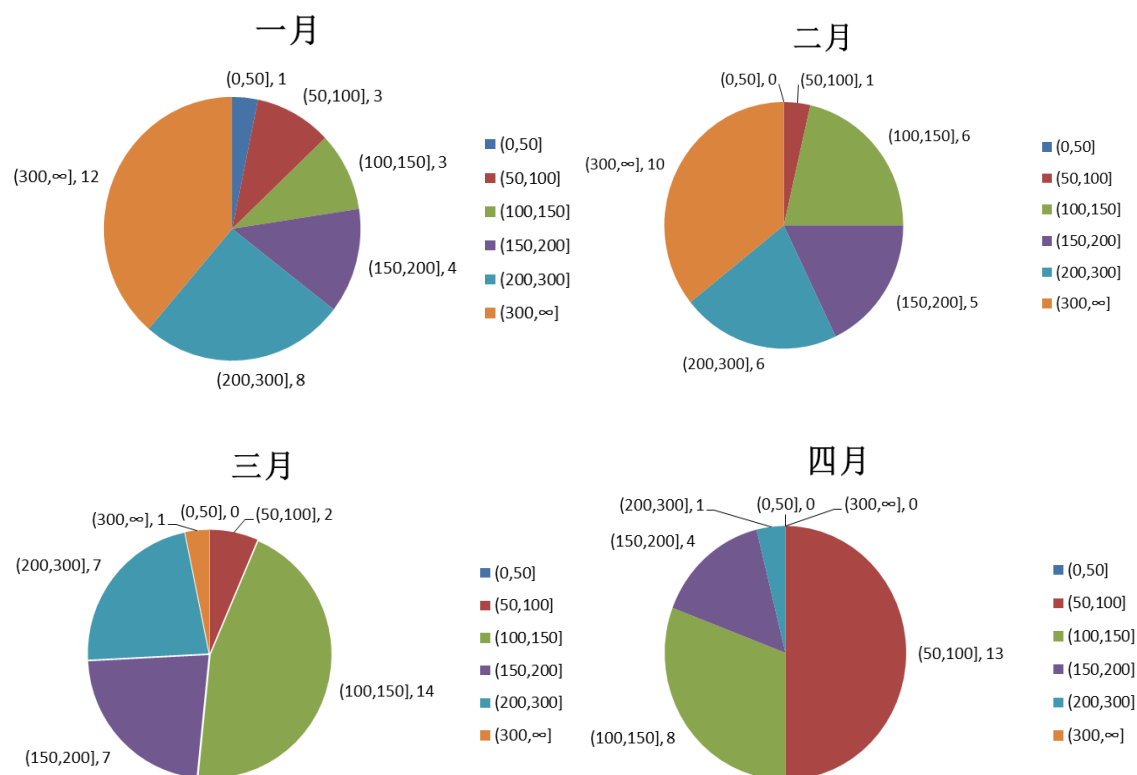


图 8 2013 年西安市 PM2.5 前四个月变化情况

2) PM2.5 空间分布及规律

根据各个测点的 PM2.5 数据绘制 PM2.5 浓度随时间演变的轨迹图，可以得到这样的结论：各个测量虽然位置不同，但是各测点 PM2.5 浓度随时间演变的趋势几乎一致。各个测点之间的监测数据存在较大的相似性。全市 PM2.5 浓度整体变化趋势是相近的。

图 9 显示的是全市平均、长安区与高压开关厂的 PM_{2.5} 监测数据。以全市平均值作为对照，长安区代表了对环境要求较高的风景区，高压开关厂代表工业区。从图中可以看到，高压开关厂的监测数据高于全市平均值，长安区监测值总是略低于全市平均值。最关键的是，三条演变曲线的发展变化趋势几乎一致。

借鉴 HJ633-2012 环境空气质量指数(AQI)技术规定（试行）中表 2 中空气质量指数级别划分，根据 13 个监测站 1-4 月份的 PM_{2.5} 数据，统计每个站点 1-4 月中，各级别 PM_{2.5} 浓度出现的天数，作图。图中纵坐标表示天数。每个色块的长度代表对应等级 PM_{2.5} 出现的天数。从图中可以看到，全市平均与高压开关厂之间存在明显的比例关系，类似的还有兴庆小区与纺织城、市人民体育场与高新西区等。

为了定量揭示这种相似性，本文对 13 个监测点 PM_{2.5} 数据作相关分析。除小寨外，其余 12 个监测站之间的相关系数都大于 0.8，且大部分情况下大于 0.9。小寨与其余各子站的相关系数均小于 0.8。

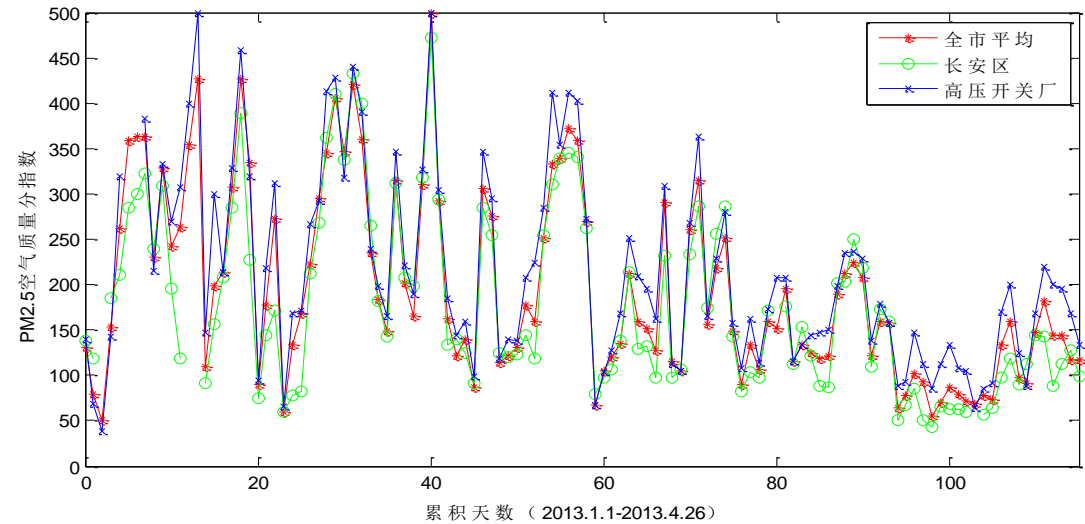


图 9 长安区和高压开关厂 P2.5 演变轨迹对比

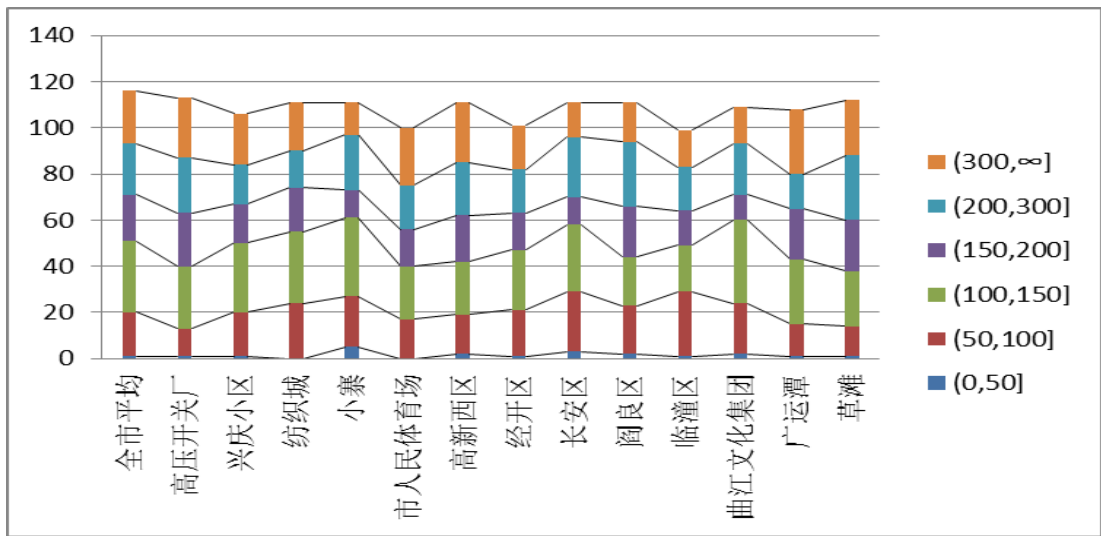


图 10 监测站各级 PM_{2.5} 发生天数统计图(2013.1.1-4.26)

3) 分区污染评估

按照环境保护部修订的《环境空气质量标准》，对 2013 年 1 月 1 日至 2013 年 4 月 26 这段时间内的各项污染物浓度限值是否符合二级标准进行统计。统计结果请详见下表。

表 4 各监测点环境污染物达标率

观测点	SO ₂	NO ₂	CO	O ₃ (h)	O ₃ (8h)	PM10	PM2.5
高压开关厂	0.59	0.29	0.59	0.92	0.92	0.13	0.11
兴庆小区	0.59	0.55	0.92	0.93	0.92	0.19	0.17
纺织城	0.91	0.48	0.91	0.95	0.95	0.18	0.21
小寨	0.79	0.81	0.65	0.95	0.94	0.12	0.23
市人民体育场	0.77	0.58	0.59	0.90	0.91	0.08	0.15
高新西区	0.89	0.54	0.60	0.96	0.95	0.10	0.16
经开区	0.86	0.65	0.66	0.88	0.88	0.16	0.18
长安区	0.89	0.96	0.57	0.94	0.91	0.16	0.25
阎良区	0.89	0.78	0.73	0.93	0.91	0.12	0.20
临潼区	0.83	0.52	0.62	0.84	0.84	0.24	0.25
曲江文化集团	0.78	0.47	0.59	0.66	0.69	0.22	0.21
广运潭	0.92	0.47	0.59	0.94	0.94	0.16	0.13
草滩	0.91	0.93	0.41	0.97	0.96	0.15	0.12

通过分析监测站各项污染物达标率表格，可以得到如下评估结论：

- A. 各区 PM10、PM2.5 达标率都非常低，说明全市颗粒物污染较为严重；
- B. 高压开关厂除颗粒物污染严重外，NO₂、SO₂、CO 三个项目的达标率均不高，特别是 NO₂，其达标率低于 30%；
- C. 兴庆小区 NO₂、SO₂ 达标率在 50%~60%，存在氮化物和硫化物污染；
- D. 纺织城 NO₂ 达标率不足 50%，存在氮化物污染；
- E. 小寨存在轻微的硫化物、氮化物和一氧化碳污染；
- F. 市人民体育场硫化物和一氧化碳污染，PM10 污染在所有监测站中最严重
- G. 高新西区存在较严重的氮化物和一氧化碳污染；
- H. 经开区存在氮化物和一氧化碳污染，还伴有轻微的 SO₂ 污染和臭氧含量过高；
- I. 长安区存在较严重的一氧化碳污染物，轻微的硫化污染；
- J. 阎良区存在轻微的 NO₂、SO₂、CO 污染；

K. 临潼区 NO_2 污染较严重，还有轻微的 CO 污染；

L. 曲江文化集团臭氧超量在所有监测站中最为严重，还存在较为严重的 NO_2 、CO 污染物和轻微的 SO_2 污染；

M. 广运潭存在较严重的 NO_2 、CO 污染；

N. 草滩 CO 达标率较低， NO_2 、 SO_2 、臭氧的达标率均较高，环境比较好。

4.2.2 $\text{PM}_{2.5}$ 的时空分布的发生与演变模型

1) $\text{PM}_{2.5}$ 演变规律多元回归模型的建立

A. 回归变量的初步挑选

通过阅读相关文献，发现当日 $\text{PM}_{2.5}$ 浓度可能受到前日 $\text{PM}_{2.5}$ 浓度、前日气象因素、当日气象因素等的影响。本题附件中给出了 1 至 4 月份的 $\text{PM}_{2.5}$ 浓度值，11 年到 13 年的气象资料。气象资料中的气象数据比较粗略，仅给出了三个指标：天气状况（阴晴雨雪变化），当日最高/最低气温，风力等级和风向。需从气象数据中提炼出回归变量。

首先，天气状况是用描述性的语言给出的，必须量化才能代入回归方程。本文将天气状况分成 5 个等级，并用数值量化，见表 5。

表 5 天气状况量化

天气状况描述	量化值
晴	0
多云	0.5
阴	1
雪	1.5
小雨（雨夹雪）	2

同样，将风力按照对应等级的风速进行量化，并对日间和晚间风速取平均，得到全天风速值。

最终，本文总结出 15 个可能的回归变量，列表如下。

表 6 初选回归变量

初选回归变量	初选回归变量
前日 $\text{PM}_{2.5}$ 浓度	前日温差
天气状况	当日日间风速
当日最高温度	当日夜间风速
当日最低温度	当日平均风速

当日温差	前日日间风速
前日最高温度	前日夜间风速
前日最低温度	前日平均风速

B. 相关系数精选回归变量

采用相关系数筛选对当日 **PM2.5** 浓度具有较大关联的变量。本文对 1 到 4 月全部数据进行分析, 然后分季节进行分析 (1、2 月份为冬季, 3、4 月为春季)。

1 到 4 月数据整体做相关分析, 得到相关系数, 见表 7。

表 7 初选回归变量及相关系数

符号	初选回归变量	相关系数	符号	初选回归变量	相关系数
X_1	前日 PM2.5 浓度	0.626814	X_8	前日温差	-0.39131
X_2	天气状况	0.200685	X_9	当日日间风速	-0.09783
X_3	当日最高温度	-0.5099	X_{10}	当日夜间风速	-0.02637
X_4	当日最低温度	-0.47063	X_{11}	当日平均风速	-0.09607
X_5	当日温差	-0.43334	X_{12}	前日日间风速	-0.12658
X_6	前日最高温度	-0.48837	X_{13}	前日夜间风速	-0.1185
X_7	前日最低温度	-0.46087	X_{14}	前日平均风速	-0.15426

另外, 当日最高温度与前日最高温度、当日最低温度与前日最低温度、当日最高温度与当日最低温度的相关系数分别为 0.93、0.95、0.95, 相关性很高。为了降低维数, 这 4 个初选变量只取当日最高温度作为精选回归变量。其余相关系数小于 0.2 的初选变量舍弃。最后得到 5 个精选回归变量: 前日 **PM2.5** 浓度 X_1 、天气状况 X_2 、当日最高温度 X_3 、当日温差 X_5 和前日温差 X_8 。

冬季数据做相关分析, 相关系数, 见表 8。

表 8 初选回归变量及相关系数

符号	初选回归变量	相关系数	符号	初选回归变量	相关系数
X_1	前日 PM2.5 浓度	0.510967	X_8	前日温差	-0.12554

X_2	天气状况	0.212101	X_9	当日日间风速	-0.02966
X_3	当日最高温度	-0.0924	X_{10}	当日夜间风速	-0.12281
X_4	当日最低温度	0.019287	X_{11}	当日平均风速	-0.11391
X_5	当日温差	-0.1426	X_{12}	前日日间风速	0.016812
X_6	前日最高温度	-0.10196	X_{13}	前日夜间风速	0.08123
X_7	前日最低温度	-0.01108	X_{14}	前日平均风速	0.044485

选择相关系数大于 0.1 的作为精选变量,得到 7 个精选回归变量:前日 PM2.5 浓度 X_1 、天气状况 X_2 、当日温差 X_5 、前日最高温度 X_6 、前日温差 X_8 、当日夜间风速 X_{10} 、当日平均风速 X_{11} 。

春季数据做相关分析,相关系数,见表 9。

表 9 初选回归变量及相关系数

符号	初选回归变量	相关系数	符号	初选回归变量	相关系数
X_1	前日 PM2.5 浓度	0.399405	X_8	前日温差	-0.13832
X_2	天气状况	0.233939	X_9	当日日间风速	-0.03768
X_3	当日最高温度	-0.16317	X_{10}	当日夜间风速	0.158748
X_4	当日最低温度	-0.06027	X_{11}	当日平均风速	0.072637
X_5	当日温差	-0.21587	X_{12}	前日日间风速	-0.13346
X_6	前日最高温度	-0.05871	X_{13}	前日夜间风速	-0.14456
X_7	前日最低温度	0.028274	X_{14}	前日平均风速	-0.16083

选择相关系数大于 0.1 的作为精选变量,得到 9 个精选回归变量:前日 PM2.5 浓度 X_1 、天气状况 X_2 、当日最高温度 X_3 、当日温差 X_5 、前日温差 X_8 、当日夜间

风速 X_{10} 、前日日间风速 X_{12} 、前日夜间风速 X_{13} 、前日平均风速 X_{14} 。

C. 1 到 4 月监测数据整体回归分析

上文精选得到 5 个回归变量，然后构建回归模型

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_5 + \beta_5 x_8 + \varepsilon \quad (11)$$

设显著性水平 $\alpha=0.01$ ，应用 Matlab 的 $\text{regress}(y, X, \alpha)$ 函数，我们可以求得：

$$\beta_0 = 100.5504, \beta_1 = 0.4956, \beta_2 = 12.7727$$

$$\beta_3 = -3.7962, \beta_4 = 3.5042, \beta_5 = 0.1066$$

同时求得上述方程的可决系数 $r^2=0.4471$ ，统计量 $F=17.6255$ ，显著性概率 $p=9.16^{-13}<0.01$ ，模型方差 $\sigma^2=6135.2398$ ，说明该回归方程是显著的。该模型的残差分析图如图 11 所示，该图说明残差完全在可控且可接受的范围内。图 12 所示是模型估计的 PM2.5 的值与实测数据较好的吻合，总体变化趋势基本一致。

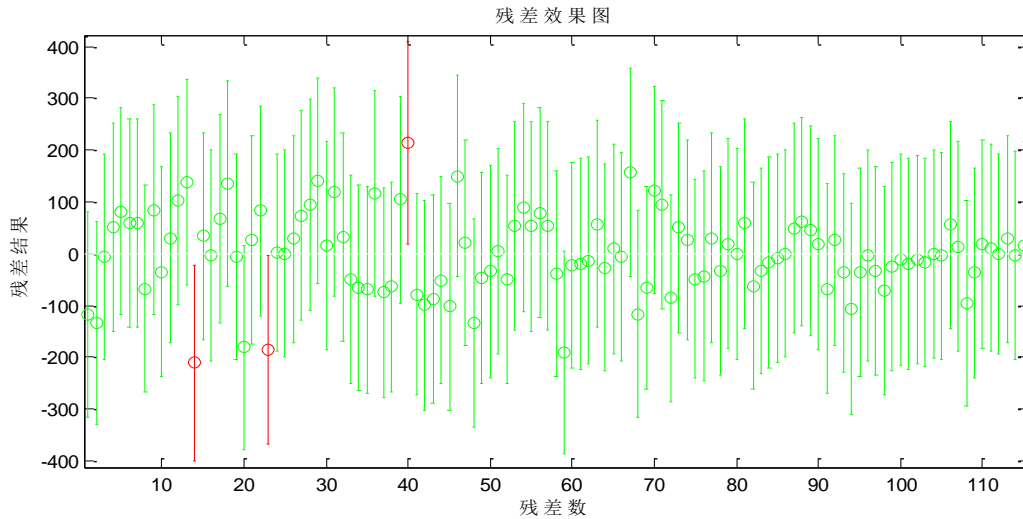


图 11 全体数据建模残差效果图

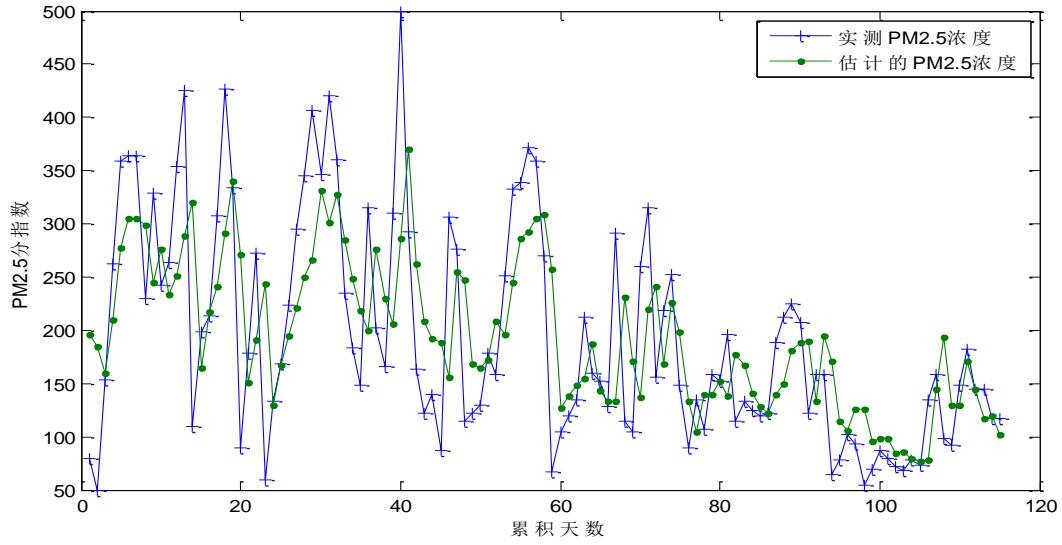


图 12 全体数据建模时，PM2.5 演变模型估计值与实测数据比较

D. 冬季节数据回归分析

根据回归变量构建回归模型

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_5 + \beta_4 x_6 + \beta_5 x_8 + \beta_6 x_{10} + \beta_7 x_{11} + \varepsilon \quad (12)$$

设显著性水平 $\alpha=0.01$ ，应用 Matlab 的 regress(y,X, α)函数，可以求得：

$$\beta_0 = 295.6826, \beta_1 = 0.5092, \beta_2 = 35.6491$$

$$\beta_3 = -125.449, \beta_4 = 66.3015, \beta_5 = -5.6569$$

$$\beta_6 = 1.4746, \beta_7 = 8.6706$$

同时求得上述方程的可决系数 $r^2=0.3186$ ，统计量 $F=3.3401$ ，显著性概率 $p=0.005<0.01$ ，模型方差 $\sigma^2=9178.353$ 。说明该回归方程是显著的。该模型的残差分析图如图 13 所示，该图说明残差完全在可控且可接受的范围内。图 14 所示是模型估计的 PM2.5 值与实测数据较好的吻合，总体变化趋势基本一致。

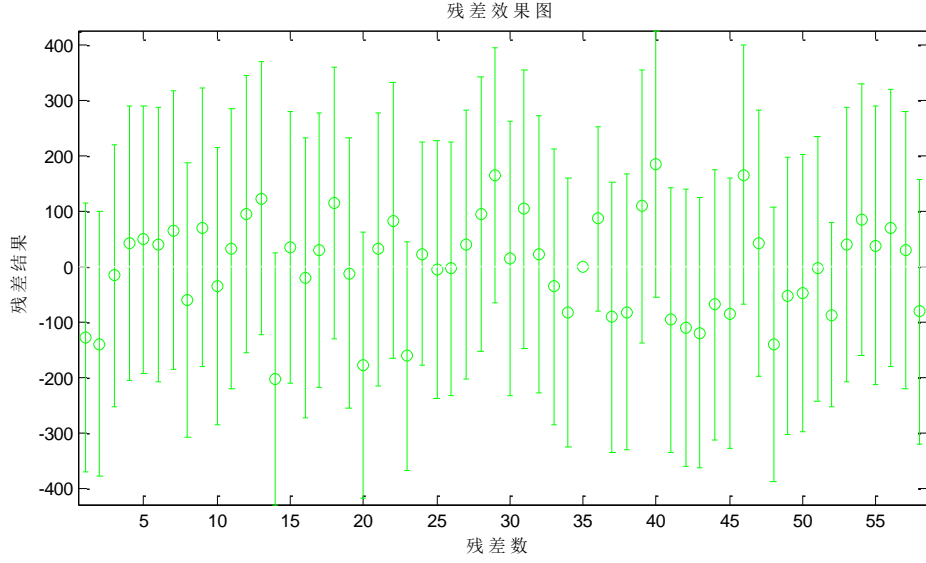


图 13 冬季数据建模残差效果图

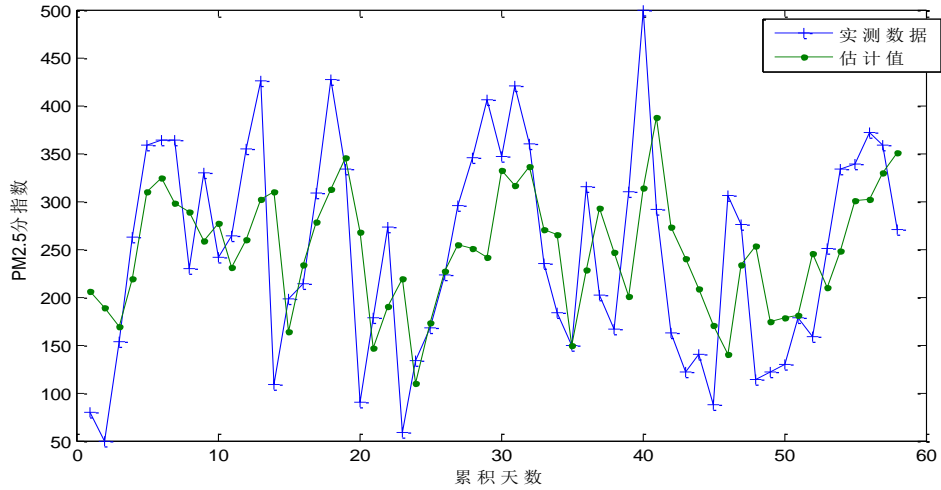


图 14 冬季数据建模时，PM2.5 演变模型估计值与实测数据比较

E. 春季节数据回归分析

根据回归变量构建回归模型

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_5 + \beta_5 x_8 + \beta_6 x_{10} + \beta_7 x_{12} + \beta_8 x_{13} + \beta_9 x_{14} + \varepsilon \quad (13)$$

设显著性水平 $\alpha=0.01$ ，应用 Matlab 的 `regress(y,X, α)` 函数，可求得：

$$\begin{aligned} \beta_0 &= 188.7501, \beta_1 = 0.4906, \beta_2 = 14.3369 \\ \beta_3 &= -3.6999, \beta_4 = 2.0447, \beta_5 = -27.1941 \\ \beta_6 &= -37.0409, \beta_7 = 42.8711, \beta_8 = 3.5260 \\ \beta_9 &= -0.2452 \end{aligned}$$

同时求得上述方程的可决系数 $r^2=0.4622$, 统计量 $F=10.1161$, 显著性概率 $p=4.46^{-11}<0.01$, 模型方差 $\sigma^2=6169.117$ 。说明该回归方程是显著的。该模型的残差分析图如图 15 所示, 该图说明残差完全在可控且可接受的范围内。图 16 所示是模型估计的 PM2.5 的值与实测数据较好的吻合, 总体变化趋势基本一致。但是图中实测数据的变化剧烈的连续多个峰值附近, 模型逼近效果不是特别理想, 偏差较大。

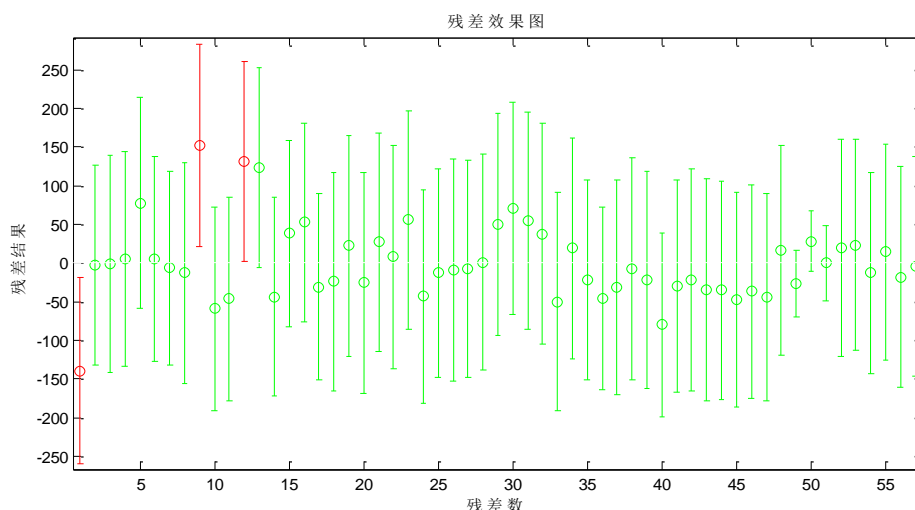


图 15 春季数据建模残差效果图

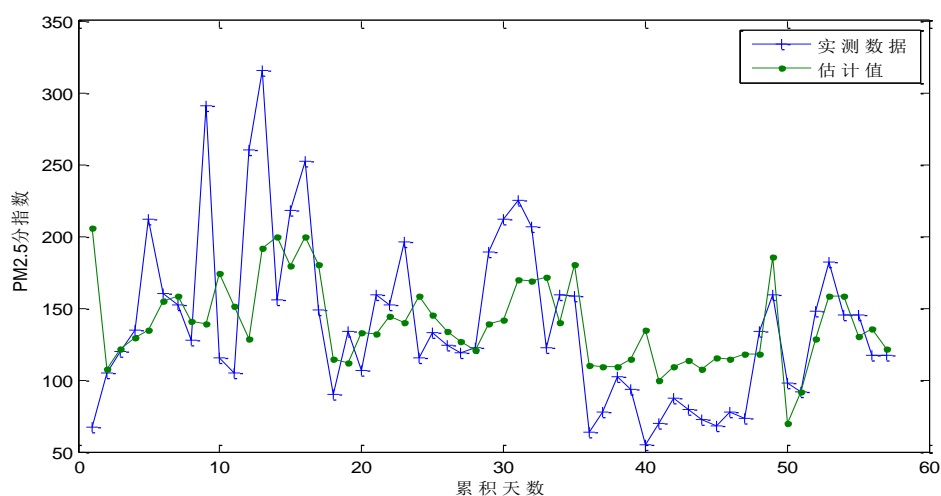


图 16 春季数据建模时, PM2.5 演变模型估计值与实测数据比较

分析 PM2.5 与气象和季节因素的影响关系。

首先, 每天的 PM2.5 是连续变化的, 因此前后两天之间的 PM2.5 浓度必然有一个正的传递关系, 否则, 前后两天的 PM2.5 的值必然在某时刻要发生突变, 这显然不符合实际。

参考文献[11]指出, 日均相对湿度与 PM2.5 的浓度呈负相关, 并解释原因是, 空气湿度大时, 污染物容易吸附到水滴上造成空气污染加重。但是参考上文 PM2.5 时间分布特征, 下雨、下雪等相对湿度较大的天气情况下, 往往伴随着 PM2.5 的浓度有一个较大的下降, 而不是浓度增加。另一个常识是, 雨后、雪后

往往大家感觉空气更加清新，空气质量变好。笔者认为应该存在一个湿度值得界限，在湿度值低于这个阈值时，污染物会发生像参考文献[]提出来的变化；在湿度值大于这个阈值时，空气中的小水滴相当于凝结核，会促进小颗粒凝聚成大颗粒，自然沉降或者随雨雪降落。因此，笔者认为，该点有必要进一步探讨。

PM2.5 的浓度与气温 呈正相关气温升高时，大气粒子动能大，速度快，有利于污染物的扩散，所以 PM2.5 的浓度与当日温度呈正相关，该结论与文献[12]一致。

温差过大，易于形成逆温现象。地表温度低，高空气温高，低空空气流动性差，抑制了污染物向周边和向上输送，从而导致污染物积累。

风速对于颗粒物的输送具有很大影响，风速的影响分两类：风可能把其他高污染地区的污染物搬运过来，也可能将本地污染物吹散。所以，风对 PM2.5 的影响要视情况而定。大多数情况下，风发挥的是吹散本地污染物的效用。

4.2.3 PM2.5 浓度突发时的污染扩散预测与评估

1) 模型的建立

A. 抛物型方程的导出

设 $u(x, y, z, t)$ 是 t 时刻点 (x, y, z) 处一种污染物的浓度。任取一个闭曲面 S ，它所围的区域是 Ω ，由于扩散，从 t 到 $t + \Delta t$ 时刻这段时间内，通过 S 流入 Ω 的质量为

$$M_1 = \int_t^{t+\Delta t} \iint_S (D_x \frac{\partial u}{\partial x} \cos \alpha + D_y \frac{\partial u}{\partial y} \cos \beta + D_z \frac{\partial u}{\partial z} \cos \gamma) dS dt \quad (14)$$

由高斯公式得

$$M_1 = \int_t^{t+\Delta t} \iiint_{\Omega} (D_x \frac{\partial^2 u}{\partial x^2} + D_y \frac{\partial^2 u}{\partial y^2} + D_z \frac{\partial^2 u}{\partial z^2}) d x d y d z \quad (15)$$

其中， D_x, D_y, D_z 分别是沿 x, y, z 方向的扩散系数。

由于衰减(例如吸收、代谢等)， Ω 内的质量减少为

$$M_2 = \int_t^{t+\Delta t} \iiint_{\Omega} k^2 u dx dy dz dt \quad (16)$$

其中 k^2 是衰减系数。

考虑该区域存在一个污染排放源，排放速率为 $A(x, y, z, t)$ ，则由于污染源排放引起的 Ω 内的质量增加为

$$M_3 = \int_t^{t+\Delta t} \iiint_{\Omega} A(x, y, z, t) dx dy dz dt \quad (17)$$

其中，为了简化分析，排放速率 $A(x, y, z, t)$ 的空间分布只在一个污染源处(不妨取坐标 $(0, 0, 0)$ 点处)有常数值，并持续一段时间。即

$$A(x, y, z, t) = \begin{cases} a & x = y = z = 0, t \in [0, \tau] \\ 0 & \text{其它} \end{cases} \quad (18)$$

由质量守恒定律, 在 Ω 内由于扩散与衰减的合作用, 积存于 Ω 内的质量为 $M_1 - M_2 + M_3$ 。

换一种角度看, Ω 内由于深度之变化引起的质量增加为

$$M_4 = \iiint_{\Omega} [u(x, y, z, t + \Delta t) - u(x, y, z, t)] dx dy dz = \int_t^{t+\Delta t} \iiint_{\Omega} \frac{\partial u}{\partial t} dx dy dz dt \quad (19)$$

由于空气流场对污染物扩散有很大的影响, 因此实际扩散模型中还需考虑风场对浓度的影响。由于风场的作用引起的质量增加为

$$M_5 = \int_t^{t+\Delta t} \iiint_{\Omega} (v_x \frac{\partial u}{\partial x} + v_y \frac{\partial u}{\partial y} + v_z \frac{\partial u}{\partial z}) dx dy dz dt \quad (20)$$

显然, $M_4 + M_5 = M_1 - M_2 + M_3$, 即

$$\begin{aligned} & \int_t^{t+\Delta t} \iiint_{\Omega} \frac{\partial u}{\partial t} dx dy dz dt + \int_t^{t+\Delta t} \iiint_{\Omega} (v_x \frac{\partial u}{\partial x} + v_y \frac{\partial u}{\partial y} + v_z \frac{\partial u}{\partial z}) dx dy dz dt \\ &= \int_t^{t+\Delta t} \iiint_{\Omega} (D_x \frac{\partial^2 u}{\partial x^2} + D_y \frac{\partial^2 u}{\partial y^2} + D_z \frac{\partial^2 u}{\partial z^2} - k^2 u + A(x, y, z, t)) dx dy dz \end{aligned} \quad (21)$$

由 $\Delta t, t, \Omega$ 之任意性得

$$\frac{\partial u}{\partial t} + v_x \frac{\partial u}{\partial x} + v_y \frac{\partial u}{\partial y} + v_z \frac{\partial u}{\partial z} = D_x \frac{\partial^2 u}{\partial x^2} + D_y \frac{\partial^2 u}{\partial y^2} + D_z \frac{\partial^2 u}{\partial z^2} - k^2 u + A(x, y, z, t) \quad (22)$$

如果近似认为扩散到地面的污染物不会再出现回升现象, 则将整个污染扩散问题等效为一个三维自由扩散场, 从而将扩散过程中的衰减项考虑到传播过程中, 因此式中可不显含衰减项。则方程可简化为

$$\frac{\partial u}{\partial t} + v_x \frac{\partial u}{\partial x} + v_y \frac{\partial u}{\partial y} + v_z \frac{\partial u}{\partial z} = D_x \frac{\partial^2 u}{\partial x^2} + D_y \frac{\partial^2 u}{\partial y^2} + D_z \frac{\partial^2 u}{\partial z^2} + A(x, y, z, t) \quad (23)$$

另外, 由假设分析, 受目前相关气象和地理信息的数据限制, 不能得到垂直方向上的逆温, 层结, 沉降等现象的数据, 且所能得到的风向也同样不具备垂直方向的信息, 因此, 拟忽略各测量区的高度差异, 将该扩散问题, 简化为一个二维平面对流扩散问题。则方程可进一步简化为:

$$\frac{\partial u}{\partial t} + v_x \frac{\partial u}{\partial x} + v_y \frac{\partial u}{\partial y} = D_x \frac{\partial^2 u}{\partial x^2} + D_y \frac{\partial^2 u}{\partial y^2} + A(x, y, t) \quad (24)$$

对于具体问题, 为了求解方程需要给出适当的定解条件(初始条件与边值条件等)。可以认为污染物扩散问题, 是关于空间上的边值问题, 关于时间上的初值问题。

B. 定解条件的确定^[13]

偏微分方程常用的定解问题主要有三种：偏微分方程+初始条件+边界条件（初边值问题）；偏微分方程+初始条件（初值问题，也称柯西问题）；偏微分方程+边界条件（边值问题）。

a) 初始条件确定：

由问题分析可知，考虑某监测点浓度突变，则各点存在一个初始浓度 $\varphi_0(x, y, t)|_{t=0}$ ，通过对该地区 PM2.5 污染日平均的空间分布规律可以知道，该地区 PM2.5 浓度值的空间差异较小，于是为了简化工作量，近似取初始时刻该地区浓度分布为该地区平均值，即有 $\varphi_0(x, y, t)|_{t=0} = \varphi_0$ ，其中 φ_0 为浓度突变前该地区浓度平均值。

b) 边界条件确定：

在自由扩散假设条件下，可以将问题近似成一个柯西问题。

所以，对该地区浓度突变情况问题的描述，可简化为求如下的定解问题：

$$\begin{cases} \frac{\partial u}{\partial t} + v_x \frac{\partial u}{\partial x} + v_y \frac{\partial u}{\partial y} = D_x \frac{\partial^2 u}{\partial x^2} + D_y \frac{\partial^2 u}{\partial y^2} + A(x, y, t) & x, y \in (-\infty, +\infty), t \geq 0 \\ u(x, y, 0) = \varphi_0 & x, y \in (-\infty, +\infty) \end{cases} \quad (25)$$

C. 扩散系数的确定：

颗粒物扩散系数是衡量以 PM2.5 为代表的微小颗粒污染物扩散能力的重要参数，通过查阅相关研究文献，得到扩散系数的计算公式如下^[14]：

$$D = 1.233 \times 10^{-19} \left[\frac{T^{1/9}}{d} + \frac{6.215 \times 10^{-10} T^{10/9}}{d^2} \right] (m^2 / s) \quad (26)$$

其中，d 为颗粒物直径，m，T 为绝对温度，K。从计算式中可以发现，扩散系数的计算主要考虑了各向同性介质中温度和颗粒大小的影响，为简化问题，可忽略了局部地形和空间温度梯度的影响，认为扩散系数近似为一个恒定的常数，且为了简化研究，不妨取 x，y 方向扩散系数相同，即 $D_x = D_y = D$ 。其中，d 取 PM2.5 的最大值为 2.5um，T 取为当天的平均温度，污染排放率 A 取为当地监测值浓度的 n 倍。

对于上述形式的非齐次抛物型方程，要求得解析解通常是不容易的，工程上常用到的是偏微分方程的数值解法，主要包括时域有限差分方法和有限元方法。

基本思想是：把连续的定解区域用有限个离散点构成的网格来代替。这些离散点称作网格的节点。把连续定解区域上的连续变量的函数用在网格上定义的离散变量函数来近似，把原方程和定解条件中的微商用差商来近似，积分用积分和来近似，于是原微分方程和定解条件就近似地代之以代数方程组，即有限差分方程组。解此方程组就可以得到原问题在离散点上的近似解。然后再利用插值方法便可以从离散解得到定解问题在整个区域上的近似解。

有限元方法的基本思想同有限差分法相似，都是通过离散化网格把原微分方程，用节点值的插值来近似表示整个有限单元内的值。区别在于，有限元方法需

要找到偏微分方程相应变分问题的泛函，然后将插值关系带入到泛函中去，在每个单元上得到关于节点值为未知量的方程，将所有单元对应的方程组组合起来，得到一个大的方程组，然后根据初边值条件，适当处理方程组，并求出其解^[15]。

相比较而言，有限差分法最大的优点是比较直观、理论也比较成熟、精度可选，便于编程计算，但当区域不规则时，求解会变得困难；而有限元法的网格划分可以随意设置，能够对复杂区域进行计算，缺点是编程实现较复杂，但目前成熟的商业有限元软件很多，可以直接利用有限元软件进行数值计算，从而避免复杂的编程过程。在该问题的研究中，拟采用功能强大的多物理场仿真软件 **Comsol Multiphysics**，借助其强大的计算内核和友好的人机界面进行仿真计算，分析 PM2.5 浓度的演变规律。**Comsol Multiphysics** 软件的操作界面如图 17 所示。

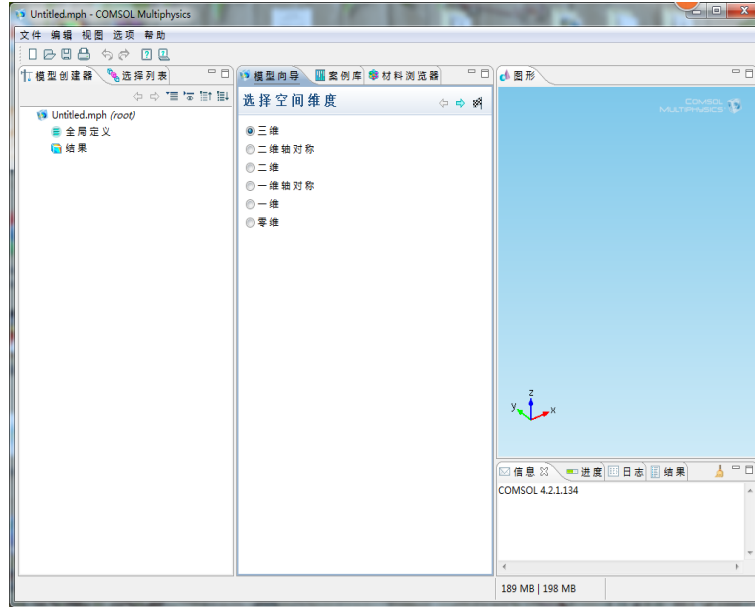


图 17 Comsol Multiphysics 操作界面

运用该软件的 **PDE** 计算模块可以很好地对偏微分方程进行计算和结果显示。在该模块中，提供了对经典偏微分方程的求解模块。我们主要选用‘对流—扩散方程’模块进行计算。

在该模块中方程的一般形式为

$$d_a \frac{\partial u}{\partial t} + \nabla \cdot (-c \nabla u) + \beta \cdot \nabla u = f$$

$$\nabla = \left[\frac{\partial}{\partial x}, \frac{\partial}{\partial y} \right]$$
(27)

其中， d_a 为质量系数（一般取 1）， \mathbf{c} 为扩散因子矢量(对应原方程中的 \mathbf{D})， β 为风速矢量(对应原方程中的 \mathbf{v})， f 为源项(对应原方程中的 $A(x, y, t)$)。

由于篇幅所限，在此不对操作步骤进行详细叙述，主要步骤为：定义求解区域；设定初始边界条件；输入模型参数；划分网格；求解。通过上述步骤我们建立了如图 18 所示的有限元计算仿真模型。其中，中心圆域代表一个近似点源，圆形边界代表一个无穷远边界(满足所谓 **Dirichlet** 条件)。

结合具体问题的点源位置坐标和各测量点之间的相对位置关系,通过在模型中设置相应的探针来监测其它测点的浓度变化,并考虑具体的天气等因素的影响,计算出模型参数并输入,划分网格后便可进行求解和结果显示与分析。

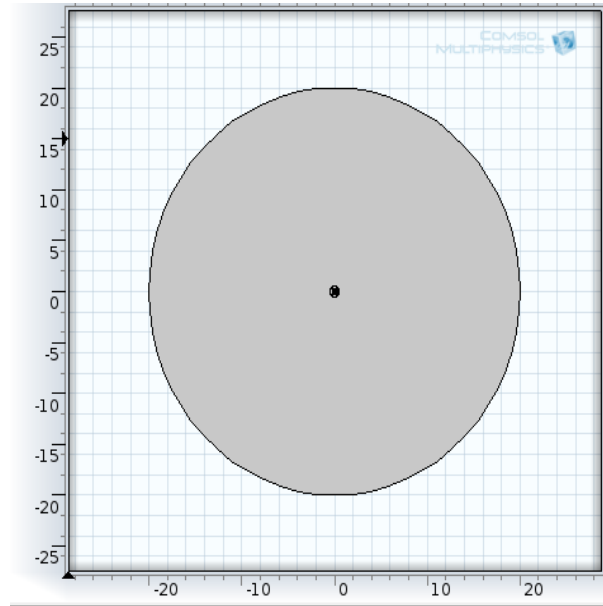


图 18 二维有限元仿真模型图

5) 污染扩散规律分析

为了对浓度突增等突发状况可能的污染情况进行预测和评估,需要先对污染扩散规律做一个必要的分析,下面运用已经建好的仿真分析模型对可能存在的突发情况进行分析,得到污染扩散问题的规律性认识,并总结出合理的评估方法。分析过程中,取圆域中心为污染源,该区域初始浓度为 $500\mu\text{g}/\text{m}^3$,当浓度增加至 2 倍,并保持 2 小时,分析该区域的随时间的污染扩散情况。

由前文提到的 PM2.5 扩散系数计算公式,取温度 T 为 298K(相当于常温条件 25°C),取 PM2.5 颗粒的平均直径为 $d=1\mu\text{m}$,计算可得到 PM2.5 粒子的扩散系数 $D=2.75\times 10^{-13}\text{m}^2/\text{s}$,若取值可信,则说明在 PM2.5 自由扩散条件下,浓度扩散特别缓慢,在一段时期内可以认为几乎不发生扩散。

A. 无风条件下的扩散现象

首先考虑在无风条件下,有限元仿真结果如图 19 所示,由结果分析可知,在无风条件下,污染源的扩散速度非常慢,可以认为经过 2 小时,基本不存在明显的扩散。因此,可以认为无风条件下的浓度突增,短时间内并不会对其它区域造成明显的污染危害,而是会因为排放累积,导致该地区污染持续加重。

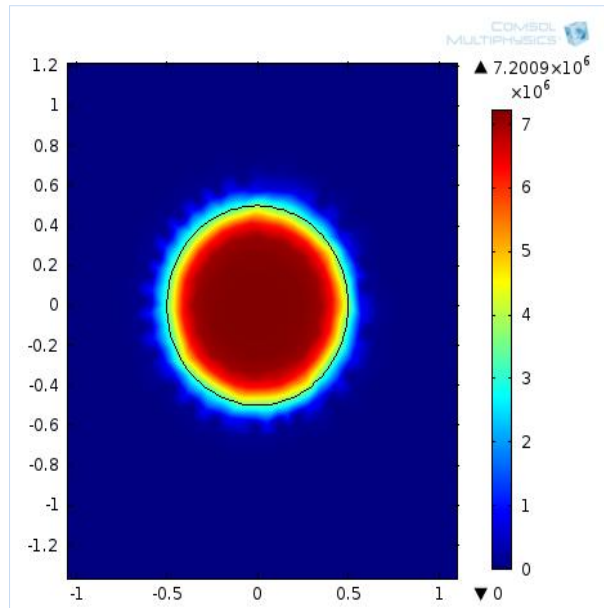


图 19 无风条件下的扩散现象

B. 有风条件下的扩散

在流动的流体中，物质的扩散应按移流扩散方程(在紊流则按紊流扩散方程)分析^[16]。为了简化研究，主要考虑单向均匀流动，即考虑 $v_x = v \cos \theta, v_y = v \sin \theta, v_z = 0$ ，(这只是一种理想的流动，这种假设的流动便于理论分析，可对有些实际紊流情况进行近似应用。)其中 θ 为水平面内速度矢量与 x 轴夹角。通过对数据的分析，取出现较多的3级风速为研究参数，通过查表得到三级风速的定义范围为： $3.4 \sim 5.4 m/s$ ，不妨取最大速度 $5.4 m/s$ 来讨论浓度突变可能导致的最大的影响。方向西安地区的常年的主风向东北风，近似取 $\theta = 45^\circ$ 。然而，在流动流体中，考虑流体的影响，扩散系数 D 需要考虑成对应流动状态的扩散系数，而该参数通常需要通过参数估计得到。出于简化问题的考虑，选定紊流条件下扩散系数为 $D' = 1 m^2/s$ 。

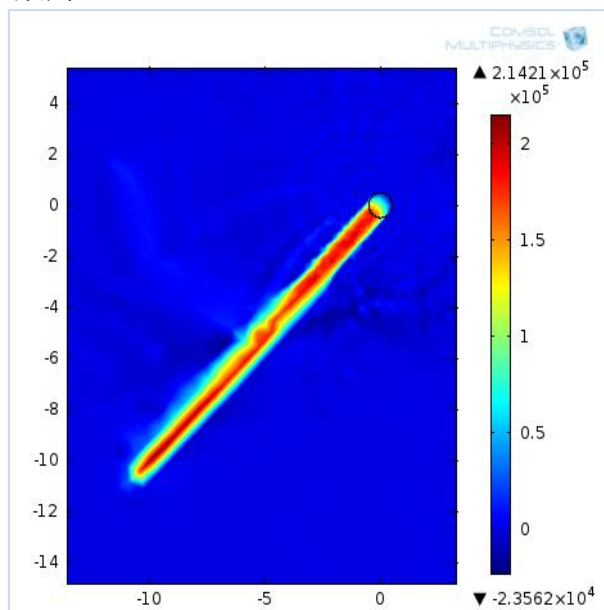


图 20 有风条件下的扩散现象

仿真分析结果如图 20 所示。从仿真结果可以看出，由于定向风的影响，整个污染浓度分布呈现为长长的‘水射流状’，在定向风方向，颗粒群呈现一个整体性地移动，并随着射线的延伸，浓度向横向方向的上延伸。但由于扩散系数的值相比于风速为较小，使得横向延伸较纵向延伸来说不那么明显。

C. 扩散系数 D 的取值对污染扩散情况的影响

扩散系数对污染扩散情况起着重要的作用，通过考虑不同紊流条件下扩散系数的不同，来比较分析可能带来的污染扩散分布区域差异。其在不同 D 取值条件下的仿真结果如图 21 所示。通过图形的分析可知， D 的不同取值对 $PM_{2.5}$ 污染物的横向扩展有重要影响。当 $D_2 = 51m^2/s$ 时，污染向两边的扩散明显增强。因此，应该可以得出这样的结论，在不同的紊流条件下，污染源的横向扩展能力不同：当扩散系数较小时，污染物的扩散主要受风向和风速的影响，若将污染源看作面源，则浓度扩散的主要影响区域可以看作是在风速作用下整个颗粒团扫过的面积；当扩散系数较大时，污染物的影响区域可以认为是如图 22 所示的辐射扇形区域。

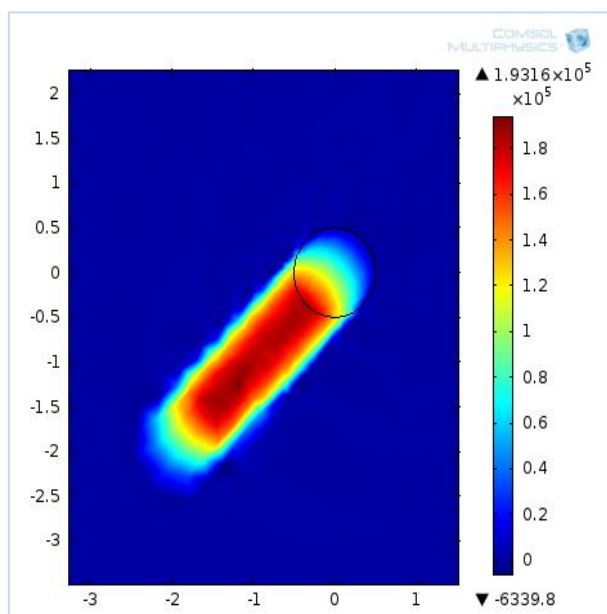


图 21 扩散系数 $D = 1m^2/s$ 时的瞬态扩散图

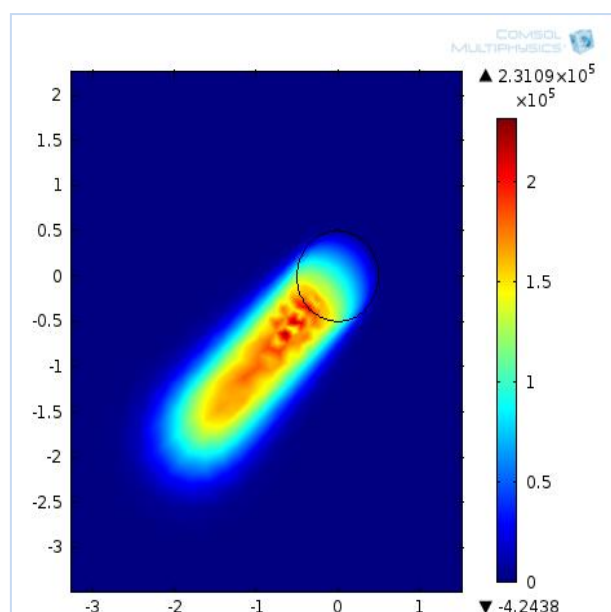


图 22 扩散系数 $D=51m^2/s$ 时的瞬态扩散图

D. 对于无持续风向的天气情况的讨论

由上文的分析可知，在紊流场中，污染源扩散很大程度上受到扩散系数的影响，因而对于无持续风向的天气情况，不妨将风场的影响直接考虑到扩散系数的大小中，将无持续风向的情况简化为一个考虑紊流的静止风场，而只是将扩散系数取一个较大值。具体的参数需要根据实测的数据经行估算。出于时间和各方面条件的限制，目前难以得到一个可信的数值，为了对扩散规律经行定性的分析，不妨就取 $D_2 = 51m^2/s$ ，通过有限元仿真分析得到如下传播扩散效果。其中考虑了两小时后，污染源不再排放的情况。其中，通过图 23 与图 19 的比较可以发现，当扩散系数增大后，污染源的扩散速率明显加快；图 24 表示当污染源在两小时后停止排放，第 10 小时，PM2.5 的浓度分布图。

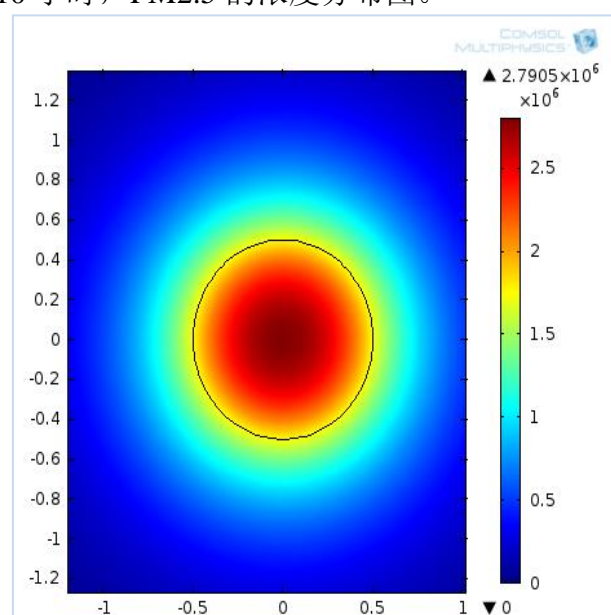


图 23 浓度突增两小时后的扩散图

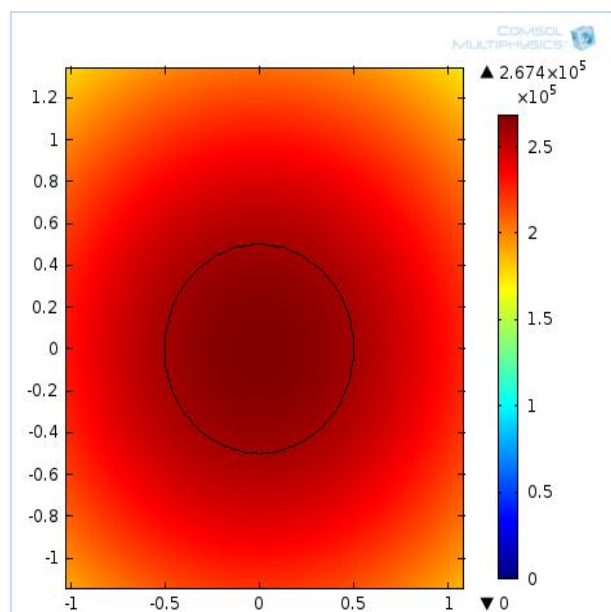


图 24 污染源停止排放后第 10 小时的瞬态扩散图

6) 基于模型的扩散预测与评估

原题要求以该地区 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，利用建立的模型进行预测和评估。由数据分析可知，在所测得的数据中，PM2.5 浓度最大值为 2013 年 2 月 10 日，浓度达到 $u = 500 \mu\text{g} / \text{m}^3$ ，而在这一天浓度达到最大值的监测点有多个，为了简化分析，暂不考虑多点扩散的问题，于是选择取西安市整段观测中 PM2.5 浓度一直较大的高压开关厂为污染源。并将其视为模型中心，通过谷歌地图得到其它各测量点相对高压开关厂的距离。当天的天气条件为：多云 / 多云， $4^\circ\text{C} / -3^\circ\text{C}$ ，无持续风向 ≤ 3 级 / 无持续风向 ≤ 3 级。其中主要用到的是风速，将无持续风向近似认为是紊流场条件，该因素认为已考虑在扩散系数 D 中，故不再加入定向风项，风速取为 3 级风的最大值 $v = 5.4 \text{m} / \text{s}$ 。

由上述的扩散规律分析可知，扩散系数的确定对污染源的扩散分布有很大的影响，因此要利用模型进行准确的扩散预测与评估，需要得到较为准确的扩散系数 D 。

通过查阅相关文献资料^[17]，可得到扩散系数如下的近似公式 $2Dt = \sigma_y^2 \approx (0.6x^{0.85})^2 = (0.6(vt)^{0.85})^2 = 0.36(vt)^{1.7}$ ，从而可以得到扩散系数的随时间变化的近似公式： $D = 0.18 \times v^{1.7} \times t^{0.7}$ ，为了简化分析，考虑取 D 为 $[0, 2]h$ 内的时间平均，即

$$\bar{D} = \frac{1}{7200} \int_0^{7200} 0.18 \times v^{1.7} \times t^{0.7} dt = 0.18 \times \frac{1}{7200} \times v^{1.7} \left. \frac{1}{1.7} t^{1.7} \right|_{t=7200} = 933.3 \text{m}^2 / \text{s}。 \quad (28)$$

通过模型仿真和探针值提取，得到各个测点浓度值如下表 10 所示：

表 10 各测点坐标及浓度值

气象测量点	坐标	u(t=2h)	u(t=12h)
高压开关厂	(0, 0)	340000	15264
兴庆小区	(9.5,-0.6)	995	7584
纺织城	(15.7,-1)	501	2394
小寨	(4.7,-5)	4887	10664
市人民体育场	(6,0.2)	8638	11619
高新西区	(-2.1,-6)	6844	11232
经开区	(3.8,9)	895	7302
长安区	(2.2,-16)	500	2195
阎良区	(31,43)	500	500
临潼区	(29,12)	500	431
曲江文化集团	(7,-7)	849	7157
广运潭	(15,6)	500	2192
草滩	(-2.6,11.5)	555	5210

从仿真结果图 25 可以发现,在 $t \in [0,2]h$ 范围内,污染源保持恒定的排放率,污染源的影响范围在两小时内逐渐扩大到红色区域所示范围,当 $t \in [2,12]h$ 时,污染源停止排放,此时相当于一个瞬态面源的扩散过程,随着时间的推移,污染物影响范围逐渐扩大到如下图黄色区域所示。显然,由于没有考虑风速的影响,PM2.5 扩散规律为较规则的圆形;同时可以发现,由于考虑的是污染源排放量在前一段时间内恒定,在污染源处出现了累积现象,和原题要求存在一定的出入。



图 25 浓度突变问题的主要影响区域

4.2.4 PM2.5 的时空分布模型的合理性检验

1)模型的合理性检验

A. 模型简化假设的合理性分析

a) 关于扩散系数各向同性的合理性

在忽略局部地形地貌等的影响条件下，考虑在较高的空间研究浓度的分布，这样的假设是同很多文献相一致的，在定性分析扩散特性地问题中具有一定的合理性。

b) 关于排放率为恒定值的合理性

题目要求污染源在扩散过程中，保持一定的浓度不变，若将扩散源看作一个点源，则可以认为改点在恒定排放源的输入下不会存在浓度积累，但实际有限元建模仿真时，为了求解的需要，建立了一个小圆域，由仿真结果发现，在该区域出现了明显的累积效应。但通过分析可以发现，在多数实际情况中，当扩散系数较小时，在排放污染源处确实会出现浓度累积的现象，而且排放率恒定的假定，更接近实际情况。即使排放源存在一定的差异，当并不影响对扩散规律的分析，所以，采用排放率为恒定值的方式应该是合理的。

c) 关于等效成二维模型的合理性

因为难以得到西安市各气象测点的详细高度信息，且垂直方向的扩散系数难以确定，为了简化模型，考虑将三维问题投影到二维。可以想象，在水平面风场的作用下，三维扩散结果应该是一个旋转抛物面，且由于气体横向扩散速度较风速作用较小，则最大浓度面和最大扩散速度面应该是在水平面内，二维模型可以近似反映三维情况，对于分析扩散规律没有太大的不合理性。

为了进一步验证二维假设的合理性，采用有限元软件进行了三维扩散仿真，

具体仿真模型和仿真结果如图 26~27 所示。从图中可以发现，在 xy 切面上，浓度扩散分布同二维相似，这进一步证明了取近似二维模型进行分析的合理性。

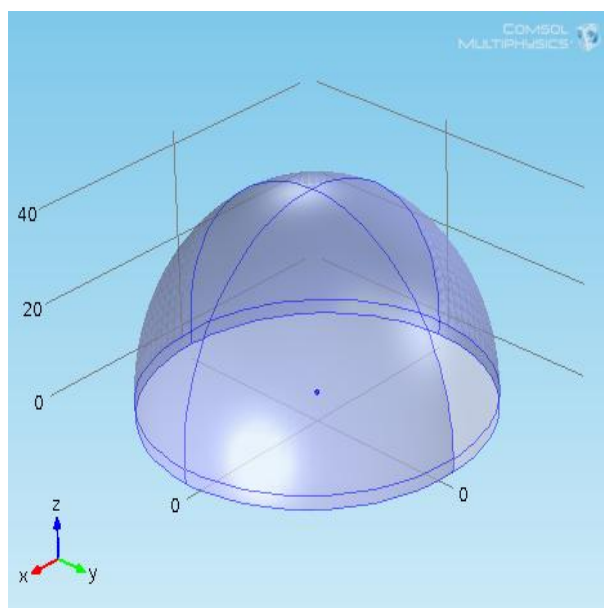


图 26 三维扩散仿真模型图

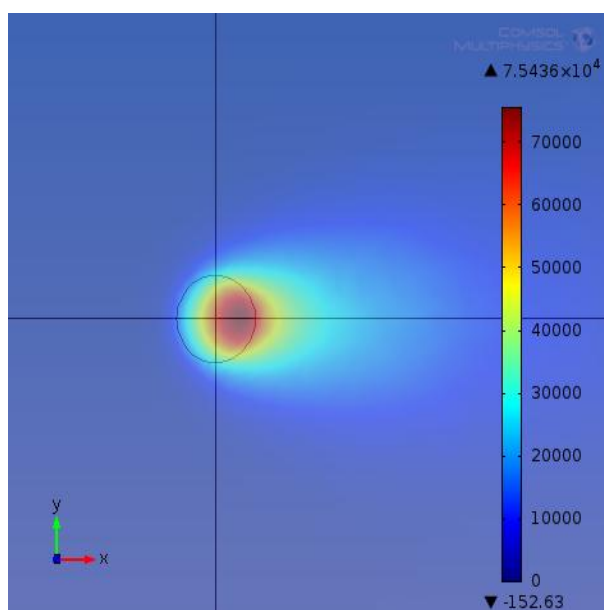


图 27 xy 切面扩散图

B. 模型相应结论的合理性验证

通过使用模型仿真分析，得到了浓度扩散的几个重要规律：风速大小和风的方向对浓度扩散问题具有重要影响；扩散系数的选取对浓度分布有重要影响，在湍流条件下的扩散系数比静止流场中的扩散系数大很多，且随着扩散系数的增大，浓度的扩散速度越快。这些观点能够和现有文献的表述相吻合，从而可以证明模型的合理性。

2) PM2.5 来成因与演变规律分析

A. PM2.5 的主要来源与成因分析^[18]

PM2.5 不是一种单一成分的空气污染物,而是由来自多种污染源的大量不同化学组分组成的一种复杂而可变的大气污染物。成分的复杂性,也决定了其来源和形成原因的复杂性,总体说来,其主要来源大致可分为以下几个方面:

a) 化石燃料(主要是石油和煤炭)和生物质燃料的燃烧。从第一题对 PM2.5 与其它空气污染源的相关性分析可知,PM2.5 与 SO₂、NO₂ 和 CO 等化石燃料中产生的气体具有较好的相关性,则可以认为它们具有同样的来源。从其微观机理上分析,高温状态下,PM2.5 可以气态形式排出,在烟羽的稀释和冷却过程中凝结而成的固态颗粒物,主要由半挥发性有机物组成。同样,其也可以是各污染源排出的气态前体物如经冷凝或复杂的大气化学反应转化而生成的二次粒子(气-粒反应)。

b) 人类生产活动。人类生产活动主要包括由于人类活动引起的道路扬尘,工业生产过程中诸如矿物加工、精炼过程等活动。以西安市为例,西安市是典型的北方城市,冬季干燥且寒冷,从每年 11 月 15 日开始供暖至次年 3 月 15 停止,历时 4 个月,因此供暖燃煤量剧增、机动车冷启动产生的更多尾气以及由于低温所造成汽车燃料的不完全燃烧等多种因素,导致冬季的 PM2.5 质量浓度远远高于其它季节。另外,在分区进行污染评估时,草滩 PM2.5、PM10、CO 达标率较低,NO₂、SO₂ 的达标率较高,说明草滩地区 PM2.5 的来源不是化石燃料。通过网上查找资料,发现草滩属于农村地区。联系 CO 浓度超标,本文认为,农村较多燃烧秸秆、稻草煮饭供暖等,会产生较多烟尘,因此也是 PM2.5 的一个重要成因,特别是农忙季节,一些地区农民有就地焚烧秸秆沤肥的习俗。可能这个时期内,PM2.5 的浓度会急剧升高。

c) 沙尘暴、风化和森林火灾等自然灾害的影响。同样是西安市,其春季的 PM2.5 突变的原因,很多时候是受到了沙尘暴、扬沙及浮尘天气的影响^[19]。

B. 浓度演变规律分析

通过第二题中对扩散模型的建立和扩散规律的分析可知,影响一个城市污染物浓度的因素主要包括两个方面:一是污染源的排放及分布状况;二是大气对污染的疏散能力。

a) 源排放是影响 PM2.5 污染浓度的直接来源。排放量的多少直接决定了一段时期内该区域 PM2.5 污染总量的大小。不同的排放规律和排放分布对整个地区的浓度分布都有着密切的影响。根据前文对西安市的季节性分析可知,冬季和春季浓度差异较大的一个重要原因就是冬季进行的供暖燃煤,导致这段时期的源排放相对春季较多,从而使得整体上冬季的 PM2.5 浓度较春季高出许多。

b) 大气对 PM2.5 污染的疏散能力是影响该地区污染程度的重要因素。可以认为源排放及分布在一段时间内是相对不变的,而大气的扩散能力却变化很大。其主要受到天气,季节性气候和地形等的影响,具体包括:风向风速,湿度,温度等。

① 风向和风速^[20]

由前面的回归分析和有限元仿真计算可知大气风场的存在对大气的污染扩散能力有着重要的影响。其中,风向的影响主要体现在风对污染物的水平输送的作用上,高污染区域总是在污染源的下风向方向,由前面的讨论结果可知,当横

向扩散系数较小时, $PM_{2.5}$ 的浓度分布主要取决于污染源沿风的方向扫过的面积。而风速的影响, 与大气的扩散能力直接对应, 风速越大, 可认为污染扩散的越快。由前文数据分析可知, 静风和无风条件, 导致扩散能力较弱, 也是西安市冬季浓度较高的一个可能原因。

② 相对湿度

由前文的回归分析可知, $PM_{2.5}$ 的浓度与相对湿度呈很好的正相关。因为湿度较高时容易形成雾、阴霾等稳定天气, 造成污染物累积, 同时高湿稳定的天气条件也有利于二次 $PM_{2.5}$ 的形成由于 $PM_{2.5}$ 的吸附性, 易附着在水汽中, 在不发生沉降的情况下悬浮在大气中难于扩散, 从而造成 $PM_{2.5}$ 的高浓度污染。

③ 大气湍流扩散

由之前对扩散系数的讨论分析可知, 如果大气中只有风而无湍流运动, 污染物的横向扩散速率会很慢, 则污染物除了被直接整体冲淡稀释外, 颗粒污染物只能靠分子扩散, 扩散速率很慢。而实际上, 大气中几乎时时处处存在着不同尺度的湍流运动。在湍流的影响下, 污染物与流场介质充分混合, 大大加快了污染物的扩散速度, 从而在湍流作用下的扩散系数相对较大。这是与之前的分析相符合的。

另外, 空间温度的分布以及地形等因素也会对扩散速率产生重要影响。

总体说来, 表现到扩散方程模型上, 大气扩散能力主要体现在两个量值的选取上, 即风矢量 v 和扩散系数 D 。通过仿真分析发现, 这两者对浓度场的分布有着决定性的影响, 这也反过来验证了仿真结论的合理性。

4.3 $PM_{2.5}$ 治理规划

4.3.1 武汉市 $PM_{2.5}$ 与 5 个污染指标之间的模型建立

针对武汉市的 $PM_{2.5}$ 的治理规划, 首先应根据已有数据建立一个 $PM_{2.5}$ 与可控污染因素之间关系的模型。根据有关研究^[10]: $PM_{2.5}$ 主要由一次微粒与二次微粒组成; 其中, 一次微粒可控的来源主要由植物及矿物燃料燃烧产生的碳黑粒子组成; 二次微粒的形成过程是大气中的 SO_2 、 NO_x 、 NH_3 、 VOC_2 发生通过冷凝或发生化学反应而生成的。而目前我们所能得到的数据中一次微粒直接相关就是 PM_{10} , 二次微粒相关可控指标只有 SO_2 与 NO_2 。而且由本章 4.1 节的相关分析中得知 $PM_{2.5}$ 与 SO_2 、 NO_2 及 PM_{10} 这 3 个污染因素指标是高度相关的, 所以本文采用多重线性多项式模型来拟合 $PM_{2.5}$ 与这 3 个污染指标之间的关系回归方程。

由于武汉市有 2013 年 1 至 8 月份 8 个月的含 $PM_{2.5}$ 等共 6 个指标的较完整的数据, 若采用二次多重线性多项式模型拟合 $PM_{2.5}$ 与个指标之间的关系, 则除常数项外, 另有 3 个一次项与 6 个二次项, 因而项数过多, 而且这 3 个指标间亦存在相关性, 因而在去除头面缺失数据等数据预处理后, 直接采用 Matlab 的逐步回归函数 `stepwise()` 进行逐步回归, 来保证回归方程更加简单明了, 计算得到回归方程如下:

$$y = -40.6496 + 2.2801x_2 + 0.0487x_1^2 - 0.0299x_1x_3 - 0.0466x_2^2 + 0.0485x_2x_3 - 0.0036x_3^2 \quad (29)$$

上式中相关统计量为：可决系数 $r^2=0.77675$ ，统计量 $F=132.79618$ ，显著性概率 $p=0.00000$ ，模型方差 $\sigma^2=1281.15095$ ，说明该回归方程是显著的，尤其是其可决系数已经很接近 1 了，说明回归拟合非常接近实际。逐步回归的结果数据与残差图如下两图所示。

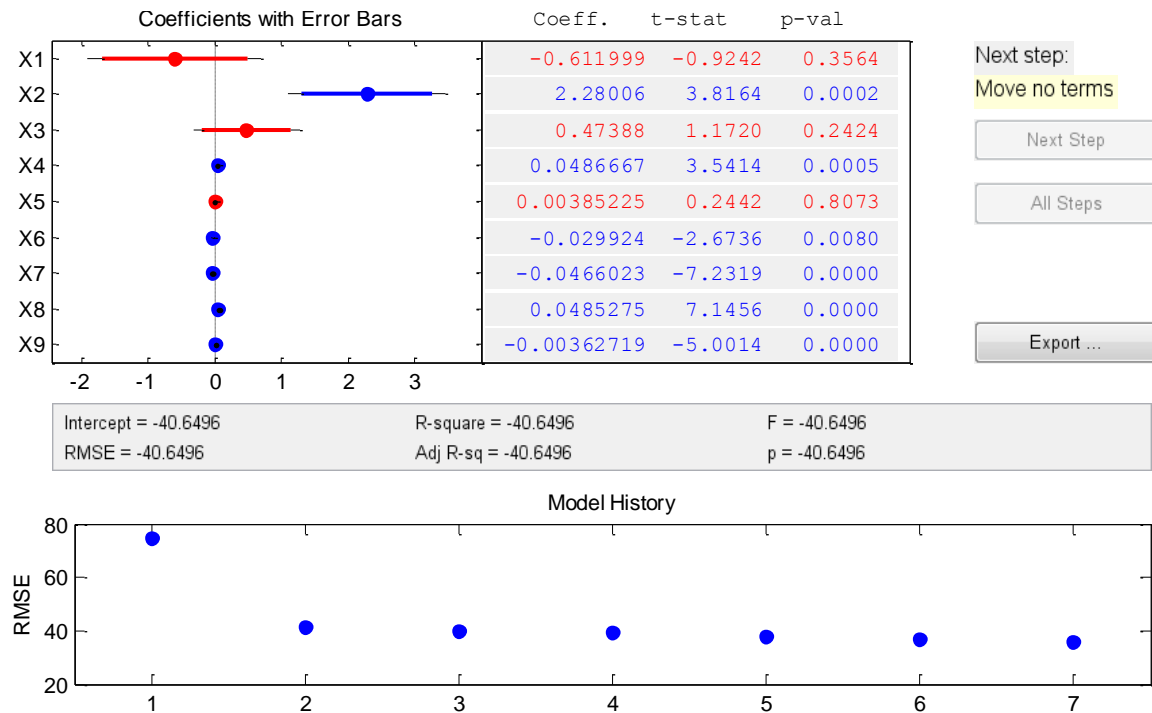


图 28 逐步回归分析的结果数据与统计分析

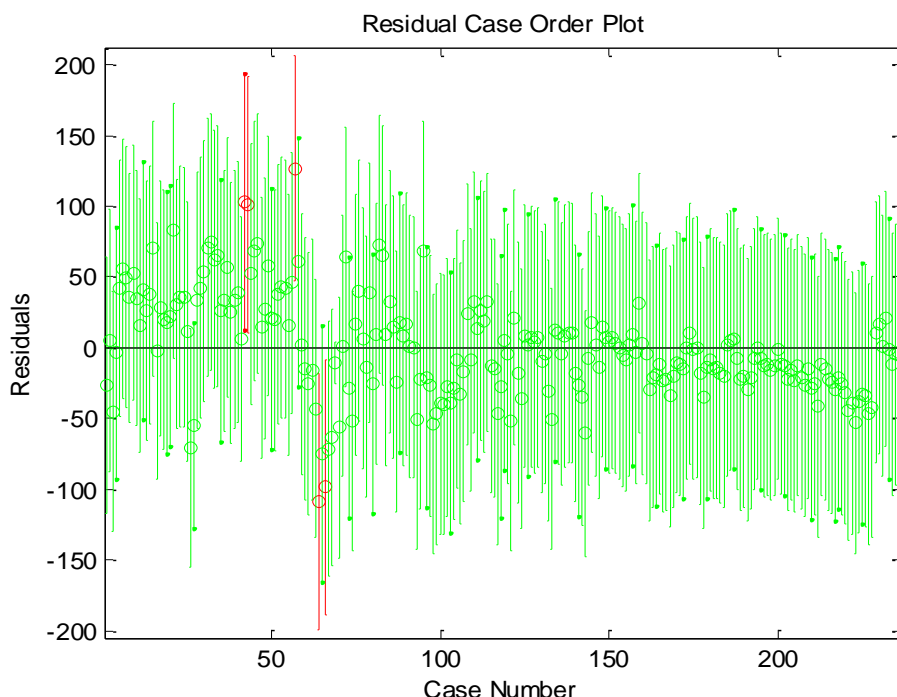


图 29 回归方程残差分析图

4.3.2 PM2.5 浓度与其对应 AQI 指标的转换

由于一般监测 PM2.5 所直接得到的是 PM2.5 的浓度, 因此必须根据环境空气质量指数 (AQI) 技术规定 (试行) [6], 来计算出其对应的 PM2.5 指标值, 反之亦然。设 PM2.5 的浓度值为 p , 则按照 AQI 技术规定 [6], 在 Excel 中其对应的 PM2.5 的指标值可用正式计算 [7]:

$$\begin{aligned} & \text{IF}(p < 0.035, 50 * p / 0.035, \text{IF}(p < 0.075, 50 * (p - 0.035) / 0.04 + 50, \text{IF}(p < 0.115, \\ & 50 * (p - 0.075) / 0.04 + 100, \text{IF}(p < 0.15, 50 * (p - 0.115) / 0.035 + 150, \text{IF}(p < 0.25, \\ & 100 * (p - 0.15) / 0.1 + 200, \text{IF}(p < 0.35, 100 * (p - 0.25) / 0.1 + 300, \text{IF}(p < 0.5, 100 * \\ & (p - 0.35) / 0.15 + 400, 500)))))) \end{aligned} \quad (30)$$

上式中的 PM2.5 浓度的单位为 mg/m^3 , 在 Excel 中应用将上述公式, 将本题中所设定的武汉市当前 PM2.5 的年平均浓度 $280\mu\text{g}/\text{m}^3$ 与五年后的最终要求达到的年终平均浓度统计指标 $35\mu\text{g}/\text{m}^3$, 折算为 AQI 标准的 PM2.5 指标分别为: 330 与 50。反之, 亦可进行相关计算。

4.3.3 武汉市 PM2.5 治理基准年相关指标的估算

在大气污染治理规划中, 首先要制定基准年的相关指标 [10], 由于 AQI 刚刚实施, 本文是对含 PM2.5 等指标的基准年估计的初步尝试。

由于按题目要求假设当年的 PM2.5 的浓度为 $280\mu\text{g}/\text{m}^3$, 折算为 AQI 的 PM2.5 指标为 330, 应用本章 4.3.1 节所得的回归方程, 其它 3 个主要指标应该

满足：

$$330 = -40.6496 + 2.2801x_2 + 0.0487x_1^2 - 0.0299x_1x_3 - 0.0466x_2^2 + 0.0485x_2x_3 - 0.0036x_3^2 \quad (31)$$

由于上述方程再加上这 3 个主要指标的上下限约束，仍有无穷多个解，为求得其中一个合理的可行解，我们先在西安市全市平均观测数据中 PM2.5 指标最接近 330 的 5 行数据如下：

表 11 武汉市 PM2.5 治理基准数据中最接近上限 330 对应的数据

点位名称	SO ₂	NO ₂	PM ₁₀	PM _{2.5}	首要污染物	时间
全市平均	77	103	230	339	严重	2013-2-25
全市平均	75	104	230	333	严重	2013-2-24
全市平均	73	84	181	315	严重	2013-2-6
全市平均	74	100	230	334	严重	2013-1-20
全市平均	101	114	275	329	严重	2013-1-10

计算其中 SO₂ 与 NO₂ 的平均值，分别为 80，101。代入到本节上述的约束方程中，即可求得 PM10 的指标的一个解为 156.5。至此，五年治理前的 PM2.5、SO₂NO₂ 与 PM10 的指标的基准年估计指标(y, x₁₀, x₂₀, x₃₀)均已得到，即：(y, x₁₀, x₂₀, x₃₀)=(330, 80, 101, 156.5)。

4.3.4 武汉市 PM2.5 五年治理方案

设 y₁, y₂, y₃, y₄, y₅ 分别为 PM2.5 治理第一至第五年每年终平均治理指标，x_{ji} 分别为第一到第三个指标在治理第一至第五年每年终平均治理指标。则由 4.3.1 的回归方程我们得到下面的约束，其中，i=1,2,3,4,5。

$$y_i = -40.6496 + 2.2801x_{2i} + 0.0487x_{1i}^2 - 0.0299x_{1i}x_{3i} - 0.0466x_{2i}^2 + 0.0485x_{2i}x_{3i} - 0.0036x_{3i}^2 \quad (32)$$

从而，合理的治理规划转化为在上式及这 4 个指标的上、下界及逐年递减的约束条件下，这 4 个指标每年递减之和最小的最优解问题，综合考虑到“专项治理投入费用是当年所减少 PM2.5 浓度平方的 0.005 倍”，问题转化为关于下式的最优解问题：

$$\min f = \sum_{i=1}^3 \sum_{j=1}^5 (x_{i(j-1)} - x_{ij}) + 0.005 \sum_{j=1}^5 (y_{j-1} - y_j)^2 \quad (33)$$

上式可以简化为：

$$\min f = \sum_{i=1}^3 (x_{i0} - x_{i5}) + 0.005 \sum_{j=1}^5 (y_{j-1} - y_j)^2 \quad (34)$$

尽管上述的最优化问题的目标函数是简单二次函数，但其约束除了上、下界及逐年递减的约束是线性之外，PM2.5 与其它 3 个指标之间的回归方程约束却是

非线性的，这给该问题的求解带来极大的困难。限于时间，本文采用重新用一次线性回归附件 1 中的数据，求得简化的做为约束的线性模型如下：

$$y = -40.7541 + 1.0907x_1 + 0.5528x_2 + 1.0650x_3 \quad (35)$$

上式中相关统计量为：可决系数 $r^2=0.68718$ ，统计量 $F=169.88325$ 仍说明该回归方程是显著的，回归方程是完全可以接受的，虽然可决系数不是非常接近 1。

从而，非线性约束转化为如下的线性约束，其中， $i=1,2,3,4,5$ 。

$$y_i = -40.7541 + 1.0907x_{1i} + 0.5528x_{2i} + 1.0650x_{3i} \quad (36)$$

当然其目标函数仍是二次的，将以上整理得出标准的关于指标变量 y_i 与 x_{ji} 标准的有约束的非线性规划问题如下：

$$\begin{aligned} \min f = & \sum_{i=1}^3 (x_{i0} - x_{i5}) + 0.005 \sum_{j=1}^5 (y_{j-1} - y_j)^2 \\ s.t. \quad & \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{pmatrix} = \begin{bmatrix} 1 & x_{11} & x_{21} & x_{31} \\ 1 & x_{12} & x_{22} & x_{32} \\ 1 & x_{13} & x_{23} & x_{33} \\ 1 & x_{14} & x_{24} & x_{34} \\ 1 & x_{15} & x_{25} & x_{35} \end{bmatrix} \begin{pmatrix} -40.7541 \\ 1.0907 \\ 0.5528 \\ 1.0650 \end{pmatrix} \\ & y_i \geq 20 \\ & y_i < y_0 \\ & y_i > y_{i-1} \\ & y_5 = 50 \\ & x_{ji} \geq 5 \\ & x_{ji} < x_{j0} \\ & x_{ji} > x_{j(i-1)} \end{aligned} \quad (37)$$

其中， $i=1, 2, 3, 4, 5$ ； $j=1, 2, 3$ 。

将上式中 15 个变量 x_{ji} 按字典顺序排列，加上 4 个 y_1, y_2, y_3, y_4 ，共计 14 个自变量，并将等式与上下界等约束条件表达为标准的线性组合表达(限于篇幅，不在此一一列出，详见附件的 Matlab 程序"MinCost.m")，应用 Matlab 的 fmincon() 函数，并设定这 14 个变量的初始值为：

$$x0=[80,60,40,20,20,100,80,60,40,20, 260,210,160,110,60, 350,300,200,100] \quad (38)$$

求解上式一个尽可能逐年分布较均匀的可行解得到 PM2.5 逐年的全年年终平均治理指标 y_1 至 y_4 分别为：(209.6688, 169.7519, 129.8343, 89.9170)，加上按要求已知的 $y_5=50$ ，将其折算为相应的浓度(单位： $\mu\text{g}/\text{m}^3$ ，折算公式详见附件“PM25 浓度与指标转换.xlsx”，已经得到 PM2.5 五年内每年的全年年终平均治理指标浓度为：

$$(y_1, y_2, y_3, y_4, y_5) = (159.6688, 128.82633, 101.85087, 66.9336, 30) \quad (39)$$

同时亦得到另外 3 个指标的五年分内每年的控制指标, 如 PM10 的第一至第五年的控制指标分别为: 126.0, 109.029, 75.945, 40.637, 5.0。

可将 PM2.5 的指标控制看成为综合治理与专项治理的共同结果。

4.3.5 武汉市 PM2.5 五年治理经费投入方案

设第一至第五年的由专项治理所得到的 PM2.5 浓度的减少为 $x_i, i=1,2,3,4,5$, 则控制总经费投入最少条件下每年合理经费投入的计算转化为求解下面的非线性规划问题:

$$\begin{aligned} \min & \left(\sum_{i=1}^5 (y_{i-1} - y_i - x_i) + 0.005 \sum_{i=1}^5 (x_i^2) \right) \\ \text{s.t.} & \quad y_{i-1} - y_i \geq x_i \geq 0 \end{aligned} \quad (40)$$

其中, $y_0=280$, $(y_1, y_2, y_3, y_4, y_5)$ 如上小节中最后一式所示。

上式仅存在最简单的上、下界约束, 应用 Matlab 的 fmincon() 函数可以直接求解 (数据整理与算法详见附件 MinCost2.m), 解得当

$$(x_1, x_2, x_3, x_4, x_5) = (99.8988, 30.8425, 26.9755, 34.9173, 36.9336)$$

时, 最少的五年投入总经费为 91.6424 万元。每年投入的经费可由下式计算:

$$f_i = (y_{i-1} - y_i - x_i) + x_i^2 \quad (41)$$

其中, $i=1,2,3,4,5$, 第一个括号内为第 i 年的综合治理费用, 乘以 0.005 的二次项为当年的专项治理经费。计算得出第一至第五年每年的经费预算如下表所示:

表 12 武汉市 PM2.5 治理第一至第五年每年的经费预算

第一年经费 总额	第二年经费 总额	第三年经费 总额	第四年经费 总额	第五年经费 总额	累计
70.3312	4.7563	3.6384	6.0961	6.8205	91.6424

五、小结与建议

本文首先分析了 AQI 中 6 个主要指标的相关性, 并得出了 PM2.5 与另 5 个主要 AQI 指标之间的都是线性相关的, 尤其是与 SO₂、NO₂、PM10 及 CO 显著线性相关的结论。然后, 应用回归方程估计了 2013 年之前的 PM2.5 的指标之后, 应用有限元分析方法建立了发生与时空演变模型, 并综合考虑风力与温度等天气和季节因素影响下对西安市的相关数据进行了定量与定性分析。亦对 PM2.5 激增等严重污染突发事件进行了扩散预测与评估。最后给出了西安市在未来五年内

PM_{2.5} 浓度由 280 降为 35 的分年度治理计划, 以及合理的所需投入总经费与逐年轻费投入计划。所建模型与所得结论基本合理可行。

不过, 本文研究还有很多需要改进之处。如, 由于西安市的 PM_{2.5} 仅有 2013 年首不足 4 个月的相关数据, 而武汉市虽然仅有一个监测点, 但其 PM_{2.5} 却是 2013 年 1 至 8 月份近 8 个月的数据。在武汉市这一个监测点的数据是准确的条件下, 若能将武汉市数据的分析与多元回归拟合结果应用于武汉市这一数据极其缺乏的相关多元回归拟合的整体趋势调整之上, 西安市的 2010 年 1 月至 2012 年 12 月的 PM_{2.5} 估计等可以更加科学与合理。

参考文献

- [1] Brett J Tunno, Kyra Naumoff Shields, Paul Lioy *et al*, Understanding intra-neighborhood patterns in PM_{2.5} and PM₁₀ using mobile monitoring in Braddock, PA, *Environ Health*, 33(1): 11-76, 2012.
- [2] Pinto JP, Lefohn AS, Shadwick DS, Spatial variability of PM_{2.5} in urban areas in the United States, *Journal of Air Waste Management Assoc.*, 54(4): 440-449, 2004.
- [3] 董志龙, 兰州大气环境质量影响因素相关分析, *干旱区资源与环境*, 23(12): 49—53, 2009.
- [4] 李小飞, 张明军, 等, 中国空气污染指数变化特征及影响因素, *环境科学*, 33(6): 1936—1943, 2012.
- [5] 郭建斌, 北京市空气污染季节变化规律研究及污染控制建议, *生态环境学报*, 18(3): 952-956, 2009.
- [6] 环境空气质量指数(AQI)技术规定(试行), 北京: 中华人民共和国国家环境保护标准, HJ633-2-12, 2012.
- [7] 姚玉刚, 朱燕玲, 邹强, 环境空气质量指数 AQI 的 EXCEL 实现研究, *环境科学与管理*, 38(6): 140-142, 2013.
- [8] Curtis F.Gerald, Patrick O.Wheatley. *Applied Numerical Analysis*, 北京: 机械工业出版社, 2006, P434~452.
- [9] 张尧庭, 方开泰, 多元统计分析引论, 北京: 科学出版社, 1982.
- [10] 张婷婷, PM_{2.5} 污染危害分析及防控措施研究, *中国环境管理*, 4(3): 19-23, 2012.
- [11] 宁海文, 西安市大气污染气象条件分析及空气质量预报方法研究, 33-54.
- [12] 李文杰, 张时煌等, 京津石三市空气污染指数(API)的时空分布特征及其与气象要素的关系, *资源科学*, 34(8):1392-1400, 2012.
- [13] 王明新, 数学物理方程(第 2 版), 北京: 清华大学出版社, 2009, P13~15.
- [14] 李艳强, 杨炯照, 周勇, 矿井粉尘扩散系数及其规律, *矿业工程研究*, 27(2): P49~51, 2012.
- [15] 潘树龙, 孙维夫, 抛物型偏微分方程的多步有限差分法计算方法, *烟台职业学院学报*, 17(1): 76-80, 2011.
- [16] 李明, 吴超, 粉尘点污染扩散模型的可视化研究, *中国科技论文在线*, <http://www.paper.edu.cn>, 2013-9-24.
- [17] 尹凤, 大气污染物扩散的理论和试验研究, 中国海洋大学, 2006.
- [18] 张碧云, 西安市南郊大气 PM_{2.5} 及其重金属元素污染特征研究, 西安建筑科技大学, 2011.

- [19] 宁海文, 西安市大气污染气象条件分析及空气质量预报方法研究, 南京信息工程大学, 2006。
- [20] 孙铁, 城市大气污染物扩散过程的数值模拟, 内蒙古科技大学, 2006。