

# 空气质量数据的校准

## 摘要

本文采用灰关联分析、SPSS 多元统计分析及 RBF 神经网络等方法，对空气质量数据校准问题进行研究。

针对问题一，使用 SPSS 进行统计分析，对各个污染物分布特性规律有了整体把握，并通过 BDP 作图，可视化分析结果。得到如下结论：（1）CO 的均值为  $1.11919 \text{ mg/m}^3$ ，相对于其它气体的均值，其一天中变化范围最大，标准差达到了  $492093 \text{ mg/m}^3$ ，其次 PM10 和 O<sub>3</sub> 的变化较大。（2）6 个污染指标分布均正偏。（3）PM10 的峰度最高为 46.782，表明分布极其陡峭，大部分时刻集中在均值附近。（4）在长期变化趋势中，O<sub>3</sub>、PM10、SO<sub>2</sub>、PM2.5 具有相似的变化趋势，NO<sub>2</sub> 和 CO 的量较为稳定。一天内，污染物呈现先稳定不变再逐渐上升态势，PM2.5 和 PM10 均在下午 14 点后明显上升，污染最重的时候是在下午。

针对问题二，通过删除异常、重复数据，利用 matlab 数据标准化，并进行三次样条函数插值，保证了自建数据和国控数据处于同样的时间戳下。在此基础上，通过对两者的数据做差，将该差值与温度、湿度、压强、降水量、风速做灰色关联分析，得到以下结论：（1）风速与 PM2.5、PM10、NO<sub>2</sub>、SO<sub>2</sub> 的误差有明显的关联，灰色关联度均大于 0.9；（2）压强与 CO 的误差具有较高的灰色关联度；（3）降水量对 PM2.5、PM10、NO<sub>2</sub>、SO<sub>2</sub> 的误差的灰色关联度较大；（4）温度会对 PM2.5、PM10、NO<sub>2</sub>、SO<sub>2</sub> 的测量造成误差；（5）湿度只与 O<sub>3</sub> 的误差相关。

针对问题三，在模型二样条函数插值的基础上，分别采用加权平均模型和 RBF 神经网络两种模型对数据进行校准。一将问题二模型中经过插值后的国控点数据和自建点数据加权平均，将国控点数据的估计值引入自建点数据，修正自建点数据。二是通过建立 RBF 神经网络，预测每次的测量误差，从而根据误差和自建点数据预测最终误差；三是通过建立逐步多元线性回归模型，分析出其中的关系式，从而校准自测点数据。模型校准的结果见支撑材料中的校准结果文件夹。

**关键词：**灰关联分析 RBF 神经网络 样条函数插值 逐步多元线性回归

## 1 问题的重述

大气污染有害于生态环境和人类健康。对“两尘四气”的浓度进行实时监控，便于对空气质量进行控制，采取应对污染源的处理。虽然国家监控局（NCP）对“两尘四气”有监测数据，且较为准确，但国控点的布控较少，数据发布滞后，成本较高。本公司的自制迷你型空气质量检测器成本低，并能实时监测特定区域的空气质量，同时监测温度、湿度、风速、空气压、降雨量等气象参数。

由于所使用的电化学气体传感器在长期使用后会产生一定的零点漂移和量程漂移，因此存在常规气体污染物(气)浓度变化与气象因素对传感器的交叉干扰。在靠近国家控制点的自建点上，其收集的数据与国家控制点的数据之间存在一些差异。因此，需要使用国家控制点的每小时的数据校准近邻自建点的数据。

请确立研究以下问题的数学模型：

- 1.自我构筑点数据和国控点数据的探索的数据解析
- 2.请分析造成构筑点数据与国家控制点数据差异的原因。
- 3.使用自控点的数据建立了校准自己构筑点的数据的数学模型。

## 2 问题的分析

本题要求首先对自建点数据和国控点数据进行探索性数据分析，首先通过 SPSS 进行描述分析和统计分析，并采用 BDP 个人版进行数据可视化分析，据此通过可视化图标发现数据规律。之后自建点数据和国控点数据进行数据插值，将自建点数据与国控点数据的差值与风速、压强、降水量、温度、湿度的数据序列进行灰色关联分析，比较灰关联度大小，判断相关程度。最后通过建立多变量灰色预测模型，利用国控点数据、自建点数据、风速、压强、降水量、温度、湿度等数据进行数据校准，得到校准后的数据序列。

## 3 符号约定

$x_i$	自建点数据第 $i$ 个的时间戳
$y_i$	自建点数据第 $i$ 个的污染物数据
$u_i$	国控点数据第 $i$ 个的时间戳
$w_i$	国控点数据第 $i$ 个的污染物数据
$t$	新建立的时间戳
$A_i$	插值后的自建点数据
$B_i$	插值后的国控点数据

$\hat{A}_i$	插值后的自建点数据
$W_1, W_2$	权重值
$NPM_{2.5}$	自建点所测 $PM_{2.5}$ 的值
$NPM_{10}$	自建点所测 $PM_{10}$ 的值
$NSO_2$	自建点所测 $SO_2$ 的值
$NCO$	自建点所测 $CO$ 的值
$NNO_2$	自建点所测 $NO_2$ 的值
$NO_3$	自建点所测 $O_3$ 的值
$N$ 降水量	自建点所测降水量的值
$N$ 压强	自建点所测压强的值
$N$ 温度	自建点所测温度的值
$N$ 湿度	自建点所测湿度的值
$PM_{2.5}_{xz}$	自建点校准过后 $PM_{2.5}$ 的值
$PM_{10}_{xz}$	自建点校准过后 $PM_{10}$ 的值
$CO_{xz}$	自建点校准过后 $CO$ 的值
$NO_{2xz}$	自建点校准过后 $NO_2$ 的值
$SO_{2xz}$	自建点校准过后 $SO_2$ 的值
$O_{3xz}$	自建点校准过后 $O_3$ 的值

#### 4 模型的假设

- 1 假设题中给定的数据准确，没有错误；
- 2、假设与自建点产生误差的变量都已经考虑在数据中；
- 3、假设测试当天没有影响测量误差的其他特殊情况；
- 4、假设各个时刻国控点数据没有测量误差。

## 5 模型一的建立与求解

### 5.1 问题分析

题目要求对自建点数据和国控点数据进行探索性分析，挖掘内在规律。首先利用 SPSS，对两尘四气的数据序列进行统计分析，分析其统计规律。之后利用 BDP 个人版，把自建点数据和国控点的时间序列通过 BDP 进行时间可视化，分析“两尘四气”（PM2.5、PM10、CO、NO<sub>2</sub>、SO<sub>2</sub>、O<sub>3</sub>）浓度随时间变化序列的变化趋势。最后采用 SPSS 进行相关性分析，分析各个测量变量之间的相关关系。

### 5.2. 模型一的建立

#### 5.2.1 数据预处理

##### (1) 数据去重

由于对时间序列可视化需要将两尘四气表示为时间的函数，而函数的性质是对于每一个自变量，有且只能有一个因变量的数值与之对应。观察自建点数据，发现数据中存在较多的重复数据，多个重复数据保留重复数据部分第一个，进行数据去重，删除 4488 个重复数据。

##### (2) 去除缺失数据

检查国控点和自建点的缺失数据，发现两者均没有缺失数据。

##### (3) 删除异常数据

利用 SPSS 检查自建点时间序列的异常值，将异常值指标大于 3 的数据删除，共删除 732 条异常数据。

经过处理，自建点数据为 229497 条。国控点数据量为 4200 条。

#### 5.2.2 统计分析模型构建

将自建点和国控点数据导入 SPSS，进行统计分析，分别统计全距、极小值、极大值、均值、标准差、方差、偏度、峰度，查看自建点和国控点的统计规律。

表 1. 国控点各个统计量

		PM2.5	PM10	CO	NO <sub>2</sub>	SO <sub>2</sub>	O <sub>3</sub>
N	有效	4200	4200	4200	4200	4200	4200
	缺失	0	0	0	0	0	0
均值		56.73	83.82	1.11919	32.64	22.40	54.77
众数		38	50 <sup>a</sup>	.860	8	16	2
标准差		34.569	50.865	.492093	24.303	20.026	47.989
偏度		1.143	3.316	.944	1.233	2.064	1.385
偏度的标准误		.038	.038	.038	.038	.038	.038
峰度		1.524	46.782	1.955	1.335	4.496	2.094
峰度的标准误		.076	.076	.076	.076	.076	.076
全距		245	983	3.845	136	149	258
极小值		1	2	.050	5	1	1
极大值		246	985	3.895	141	150	259
和		238247	352054	4700.577	137101	94100	230018
百分位数	25	31.00	49.00	.78000	13.00	10.00	18.00

50	49.00	76.00	1.05000	26.00	15.00	45.00
75	76.00	110.00	1.40000	45.75	26.00	75.00

首先对自国控点的 PM2.5、PM10、CO、NO<sub>2</sub>、SO<sub>2</sub>、O<sub>3</sub> 6 个指标的时间序列进行统计分析。得到结果如下表所示。

从均值可以看出，CO 含量具有绝对优势，均值 1.11919 mg/m<sup>3</sup>，远高于第二位的 PM10（均值为 83.82 μg/m<sup>3</sup>），第三位 PM2.5 的（56.73 μg/m<sup>3</sup>），最后三位分别是 O<sub>3</sub>，NO<sub>2</sub>，SO<sub>2</sub>。可见该城市主要污染物是 CO 和颗粒物。

分析标准差可知，从变化数值上看，CO 在一天中变化范围最大，标准差达到了 492093 mg/m<sup>3</sup>，其次 PM10 和 O<sub>3</sub> 的变化较大，其标准差分别为 50.865μg/m<sup>3</sup> 和 47.989μg/m<sup>3</sup>，

分析偏度，发现 6 个污染指标的偏度均为正数，表示长尾均在右方，即全天污染物的分布物分布，大部分时刻处在低于均值的时刻，只有少数出现数值很大的污染。其中 PM10 的偏度明显偏大，表明 PM10 存在明显的正偏。PM2.5 偏度最小，表明，全天 PM2.5 在均值左右侧分布较为均匀。

分析峰度，PM10 的峰度为 46.782，远远高于其他几项污染物指标的峰度，表明 PM10 污染极端大和极端小的数值出现情况都很少。大部分情况集中在均值附近。

分析全距，表明 CO 是全天数值变化最大的污染源，全距达到了 3.845 mg/m<sup>3</sup>，而全距最小的为 NO<sub>2</sub>，全距仅为 136 g/m<sup>3</sup>。

各污染物的污染指数分布情况如图所示。

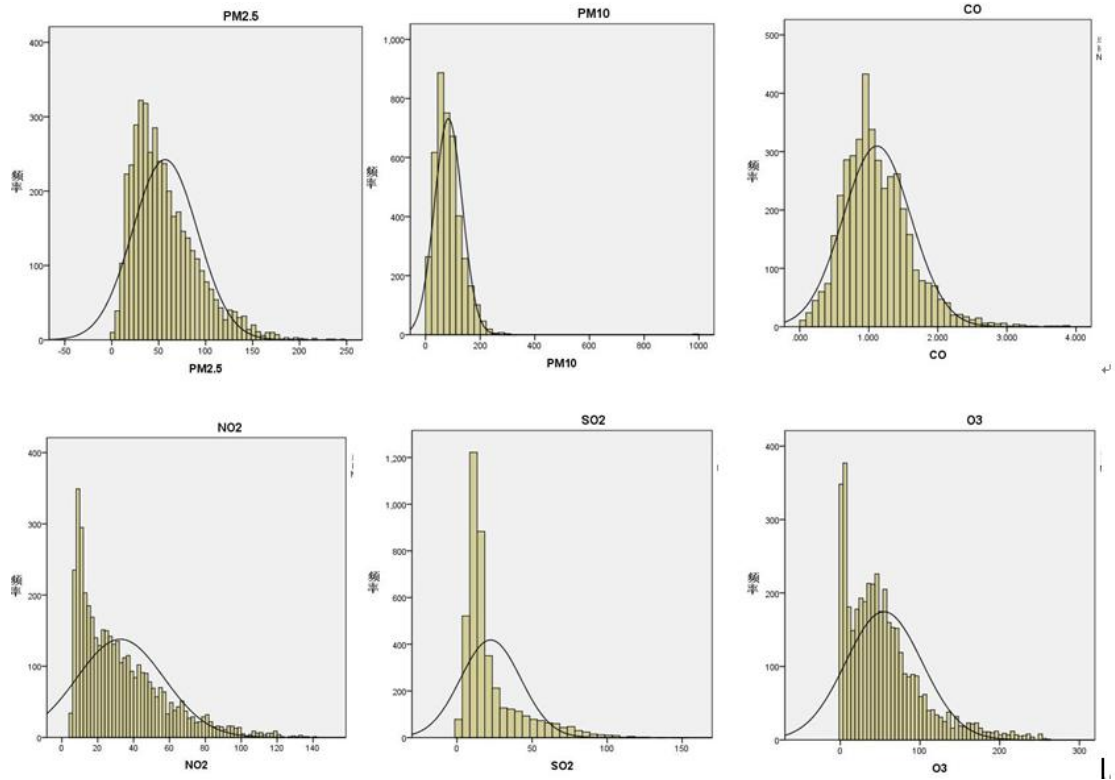


图 1 国控点各污染物分布情况

对国控点进行 PM<sub>2.5</sub>、PM<sub>10</sub>、CO、NO<sub>2</sub>、SO<sub>2</sub>、O<sub>3</sub> 6 个指标的时间序列进行统计分析。得到结果如下表所示。

表 2. 自建点各个统计量

		PM <sub>2.5</sub>	PM <sub>10</sub>	CO	NO <sub>2</sub>	SO <sub>2</sub>	O <sub>3</sub>
N	有效	234717	234717	234717	234717	234717	234717
	缺失	0	0	0	0	0	0
均值		69.35	113.47	.589	55.51	16.34	66.09
中值		59.00	93.00	.500	51.00	16.00	62.00
众数		40	72	.4	26	17	40
标准差		38.510	71.698	.2176	28.735	20.006	33.659
偏度		1.060	1.521	1.670	.687	37.003	.986
偏度的标准误		.005	.005	.005	.005	.005	.005
峰度		1.380	3.080	4.211	-.305	1503.600	1.188
峰度的标准误		.010	.010	.010	.010	.010	.010
全距		547	931	2.9	181	1101	249
极小值		1	2	.0	0	2	0
极大值		548	933	2.9	181	1103	249
和		16277696	26633136	138225.0	13028418	3836103	15511917
百分位数	25	41.00	66.00	.400	28.00	14.00	39.00
	50	59.00	93.00	.500	51.00	16.00	62.00
	75	92.00	144.00	.700	76.00	17.00	85.00

对自建点的数据进行统计分析后,发现均值与国控点的数据存在较为明显的差异。

PM10 的自建点数据均值 (113.46) 与国控点 (83.82) 数据均值存在明显差异, CO 自建点数据 (0.589) 同国控点数据 (1.119) 存在明显差异。 , 其余数据相差不大。对偏度数据进行分析,发现 SO<sub>2</sub> 的自建点偏度数据 (37.003) 明显大于国控点的偏度数据 (2.064), 峰度 (1503.600) 远远高于国控点数据, 表明自建点的统计分布存在明显的极端值, 有较大的偏差。国控点各个测试序列的均值、标准差与自建点各个序列标准差基本吻合, 表明自建点测试的数据与测试点的数据在分布上没有大的变动。

### 5.2.1 时间序列可视化分析

利用 BDP 可视化功能, 将国建点数据进行可视化观察数据趋势规律, 如图二所示。

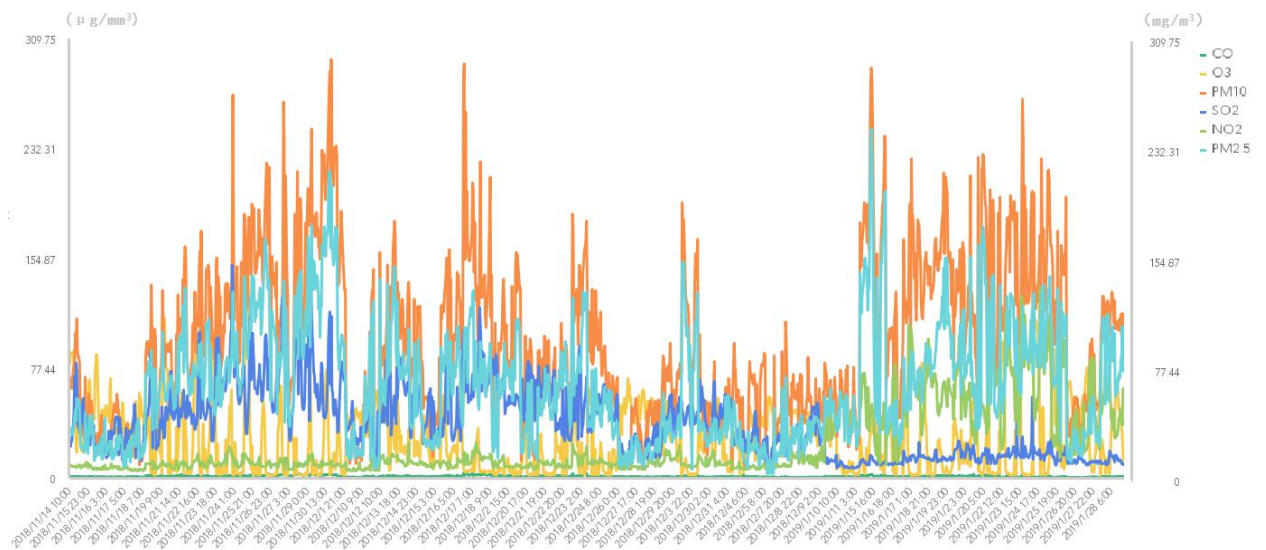


图 2. BDP 可视化国建点 3 个月时间序列数据图

从图中可以看出, O<sub>3</sub>、PM10、SO<sub>2</sub>、PM2.5 具有相似的变化趋势, 即波动趋势方面较为吻合, NO<sub>2</sub> 在前期始终保持较低的数值, 但是在 2019 年 1 月以后数值有明显的上升, 可能是因为春运导致的交通出行量剧增, 形成了较多含氮化合物尾气。CO 的量保持较为稳定的趋势。

3 个月时间跨度过长, 看不出污染物单日随时间变化趋势。因此随机抽取一天的数据作为样本, 做日污染物变化曲线, 不妨抽取 3 月 1 日的国建点数据作为样本点。观察当日数据变化趋势, 如图三所示。

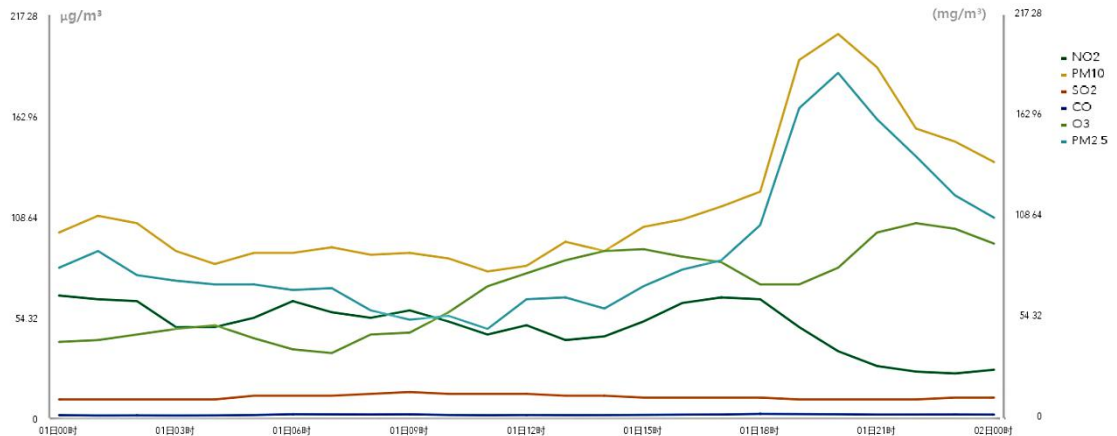


图 3. 三月一日的国建点数据变化趋势

从图 3 可以看出,污染源数据随一天内时间变化基本呈现先稳定不变再逐渐上升态势,在  $\text{SO}_2$  和  $\text{CO}$  在一天内数值较为稳定,而  $\text{PM}_{2.5}$  和  $\text{PM}_{10}$  均在下午 14 点以后呈现明显的上升态势在 20 点左右达到峰值,之后随着大气的自然沉淀作用逐渐下降。 $\text{O}_3$  和  $\text{NO}_2$  分别在 12 点和 18 点左右略微上升,之后回落。可见污染最重的时候在下午,在凌晨期间污染物随着大气的自然降解和沉淀作用逐渐下降。

## 6 问题二模型的建立与求解

### 6.1 问题分析

问题二要求对导致自建点数据与国控点数据造成差异的因素进行分析,自建点数据和国控点数据的数据序列不等长,并且数据序列的测量时间点不完全重合,不具备可比性。不妨采取插值的方法,利用 Matlab 对国控点数据和自建点数据插值,使得其数据序列和自建点序列等长,并且在测量时间点上——对应,之后对两个序列的差值进行分析,采用灰色关联模型对两序列差值、风速、压强、降水量、温度、湿度进行分析,通过比较灰色关联度大小探究造成误差的原因。

### 6.2 数据预处理

首先对日期数据进行处理。Matlab 中要求对时间数据要转换为 matlab 中规定的数值数据。因此不妨采用 matlab 的 `datenum` 函数<sup>[2]</sup>,将时间序列数据转换为数值型数据,以便于插值的进行。

对数据进行标准化处理。由于待分析的空气质量各个指标的数据量纲不同,直接进行灰色关联度分析会产生偏差,必须对数据进行标准化处理,消除量纲的影响。本文采用 `minmax` 标准化方法,可以利用 matlab 自带的函数 `mapminmax`



实现数据标准化。

### 6.3 样条函数插值模型的构建

#### 6.3.1 插值的定义

插值<sup>[3,4]</sup>，也叫“内插法”，即为根据目前已有的数据样本点，预测得到未知数据点的方法。本文中采用样条函数插值，利用 matlab 的 `interp1` 命令进行三次样条函数插值。

样条函数的定义：针对于 $[a, b]$ 上的一个区间分段

$$S : a = x_0 < x_1 < \dots < x_{n-1} < x_n = b$$

若  $f(x)$  符合以下的情形：

- (1) 对于每个区间  $[x_i, x_{i+1}]$ ,  $i = (0, 1, \dots, n-1)$ ,  $f(x)$  是次数小于或者等于  $k$  的多项式。
- (2) 并且  $f(x)$  在整个区间上都上有  $k-1$  阶导数，就可以称  $f(x)$  是定义在  $[a, b]$  上的  $k$  次多项式样条函数。

样条函数差值在实际应用中一般都取  $k=3$ ，即三次样条函数。

#### 6.3.2 数据插值

(1) 构建新的时间戳。将国控点数据和自建点数据的时间序列组合到一起，利用 SPSS 去除重复值，成为完整的时间序列。采用该时间序列，利用 `interp1` 函数分别对国控点数据和自建点数据进行三次样条函数插值。共去除 2444 个重复的时间戳数据，去重后时间戳总数 231253 条。三次样条函数插值后，国控点的时间戳和自建点的时间戳都变成了 231253 条。

(2) 插值。进行三次样条函数插值，国控点的数据成为了形状为  $[231253, 7]$  的矩阵，而自建点成为了  $[231253, 12]$  的矩阵序列。

### 6.4 灰色关联分析模型构建

#### 6.4.1 灰关联分析定义

灰色关联度分析<sup>[3,4]</sup>，是一种常用的统计分析方法，即一个灰色系统中，所关注变量受其他变量影响的相对强弱。采用灰色关联分析简便易行，但同时具有较强的主观性。

#### 6.4.2 灰关联分析

考虑影响到自建点测量误差的数据包含了风速、压强、降水量、温度、湿度，经过 6.2.3 插值处理后，时间戳得到了统一，在此基础上进行灰色关联分析。

分别对造成 PM2.5、PM10、CO、NO<sub>2</sub>、SO<sub>2</sub>、O<sub>3</sub> 测量误差的影响因素进行分析，比较各因素与测量误差之间的灰色关联度，得到各个测量误差与各影响因素之间的灰色关联度如下表所示。

表 3. 各测量误差与各影响因素的灰色关联度

灰关联度	PM2.5 的误差	PM10 的误差	CO 的误差	NO <sub>2</sub> 的误差	SO <sub>2</sub> 的误差	O <sub>3</sub> 的误差
风速	0.951083	0.994155	0.516806	0.997056	0.958080	0.344844
压强	0.521370	0.537394	0.923587	0.536017	0.522947	0.492007
降水量	0.948509	0.996057	0.517572	0.997728	0.955459	0.345186
温度	0.998644	0.944824	0.503078	0.949109	0.991003	0.338701
湿度	0.338721	0.345411	0.509104	0.344844	0.339110	0.998881

可以发现风速与 PM2.5、PM10、NO<sub>2</sub>、SO<sub>2</sub> 的误差有明显的关联，灰色关联度均大于 0.9；压强与 CO 的误差具有较高的灰色关联度；降水量对 PM2.5、PM10、NO<sub>2</sub>、SO<sub>2</sub> 的误差的灰色关联度较大。温度则会对 PM2.5、PM10、NO<sub>2</sub>、SO<sub>2</sub> 的测量造成误差。湿度只与 O<sub>3</sub> 的误差相关。

## 7 问题三模型的建立与求解

### 7.1 问题分析

问题三要求设根据国控点数据进行自建点的数据校准。有两种思路，一是可以将问题二模型中经过插值后的国控点数据和自建点数据进行加权平均，通过将国控点数据的估计值引入自建点数据中修正自建点数据；二是通过建立 RBF 神经网络，预测每次的测量误差，从而根据误差和自建点数据预测最终误差；三是通过建立 SPSS 多元线性回归<sup>[6]</sup>，来分析其相关关系，从而通过建立多元线性回归模型校准数据。

### 7.2 加权修正模型的建立

对于加权修正模型来说，一个合适的权重至关重要。由经验数据可得，可以将自建数据取 0.3 的权重，国控数据取 0.7 的权重，从而实现加权修正模型。

即：

$$\hat{A}_i = A_i * W_1 + B_i * W_2 \quad (1)$$

其中

$$A_i = \text{interp1}(x_i, y_i, t) \quad (2)$$

$$B_i = \text{interp1}(u_i, w_i, t) \quad (3)$$

$x_i, y_i$  表示自建点数据第  $i$  个的数据,  $u_i, w_i$  表示国控点数据第  $i$  个数据。  
 $t$  表示新建立的时间戳。 $A_i$  表示插值后的自建点数据,  $B_i$  表示插值后的国控点数据

$\hat{A}_i$  表示校准后的数据,  $W_1, W_2$  是权重值。

利用 matlab 可以方便的进行加权修正模型的构建与计算。部分计算结果如下所示。

时间	PM2.5 校准值	PM10 校准值	CO 校准值	NO <sub>2</sub> 校准值	SO <sub>2</sub> 校准值	O <sub>3</sub> 校准值
2018/11/14 10:00	21.05	40.75	0.3656	15.7	9.5	37.7
2018/11/14 10:02	22.35	44.75	0.3602	20.1	12.3	34.65
2018/11/14 10:06	26.65	50.2	0.3941	11.9	17.75	40.75
2018/11/14 10:09	33.8	57.2	0.41785	27.6	11	42.85
2018/11/14 10:10	19.55	37.5	0.35645	19.9	8.7	23.6
2018/11/14 10:14	32.45	54.65	0.3638	25.9	10.7	36.85
2018/11/14 10:16	29	47	0.3509	25.75	10.4	37.3
2018/11/14 10:18	15.9	30.45	0.34685	30.6	12.9	35.05
2018/11/14 10:19	15.6	30.25	0.3422	32	13.65	35.95
2018/11/14 10:23	15.6	31	0.33515	19.75	11.3	38.45
2018/11/14 10:24	9.7	16.8	0.3236	21.85	10.85	35.9
2018/11/14 10:28	17.45	33.4	0.33335	29.3	12.2	35.1
2018/11/14 10:29	19.45	36.35	0.329	29.65	10.95	35.85
2018/11/14 10:33	15.6	29.6	0.3245	25.85	10.05	35.95
2018/11/14 10:34	18.85	35.45	0.31685	15.25	7.15	22.2
2018/11/14 10:38	19	36.2	0.2856	17	9	21.3
2018/11/14 10:39	18.4	36.2	0.32915	28	10.2	37
2018/11/14 10:42	17.55	30.95	0.32555	20.65	9	35.55
2018/11/14 10:43	13.6	26.35	0.2862	20.3	9.35	27.3
2018/11/14 10:47	15.2	30.6	0.32945	20.65	8.55	30.55
2018/11/14 10:48	17.15	31.75	0.3344	18.9	8.1	33.7
2018/11/14 10:52	19.9	39.2	0.30885	21.55	9.2	34
2018/11/14 10:53	24.9	46.2	0.3087	19.8	9.8	34.35
2018/11/14 10:57	22.4	45.45	0.34265	23.45	8.75	35.7
2018/11/14 10:58	19.9	38.4	0.3458	17.65	8.1	28.4
2018/11/14 11:00	19.75	37.15	0.34775	17.5	7.95	31.05
2018/11/14 11:02	20.75	38.15	0.3467	19.8	9.05	35.35

表 4. 加权模型部分测量值校准数据

### 7.3 RBF 神经网络校准模型

RBF 神经网络是将 RBF 高斯核函数应用于神经网络的一种模型<sup>[5]</sup>,RBF 神经网络通常只有三层,第一层是输入的各个指标原始数值,即输入层;中间的隐含层是多个高斯核函数,每个高斯核函数都以一个样本点或者一个聚类中心作为高斯核函数的参数。经过隐含层,数据相当于经过了非线性的变化;之后在第二层和第三层之间采用线性输出,利用线性加权的方法将隐含层的数据输出到输出层,作为最终的预测结果。

将风速、压强、降水量、温度、湿度作为自变量,将上文预测得到的误差作为因变量,通过构建 RBF 神经网络,可以预测每次测量的误差值,进一步将误差值与自建点数据相加,得到最终预测数据。其中 MSE 降到了 109.2, RMSE 数值为 10.45。在一定范围内认为可以接受。训练 RBF 过程截图如图 4 所示。

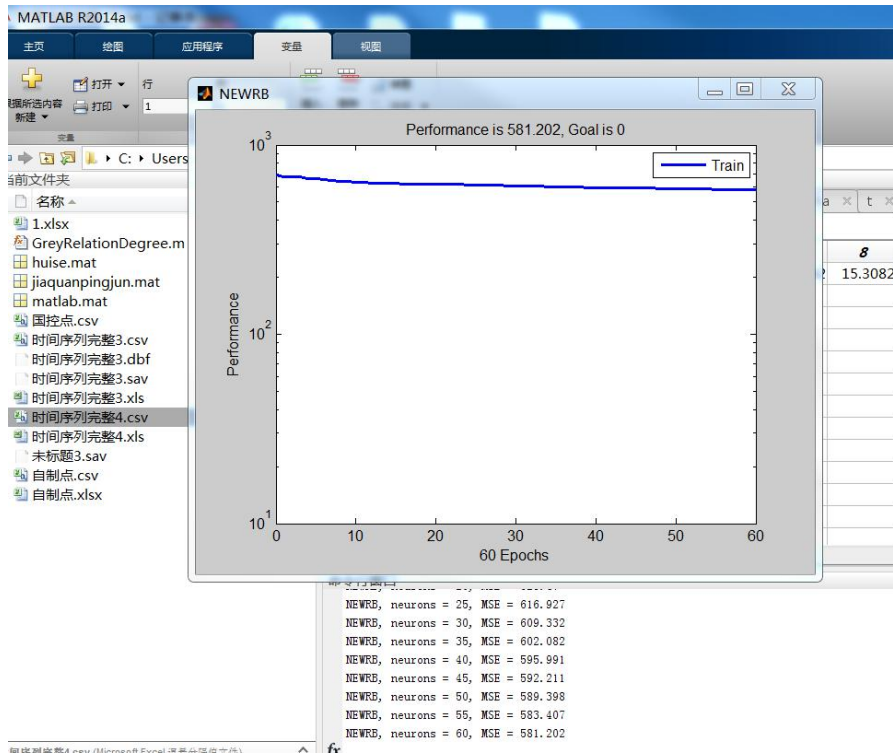


图 4. RBF 训练截图

训练结果中,以 PM2.5 前 500 条校准值和实际值为例,做折线图如下图所示:

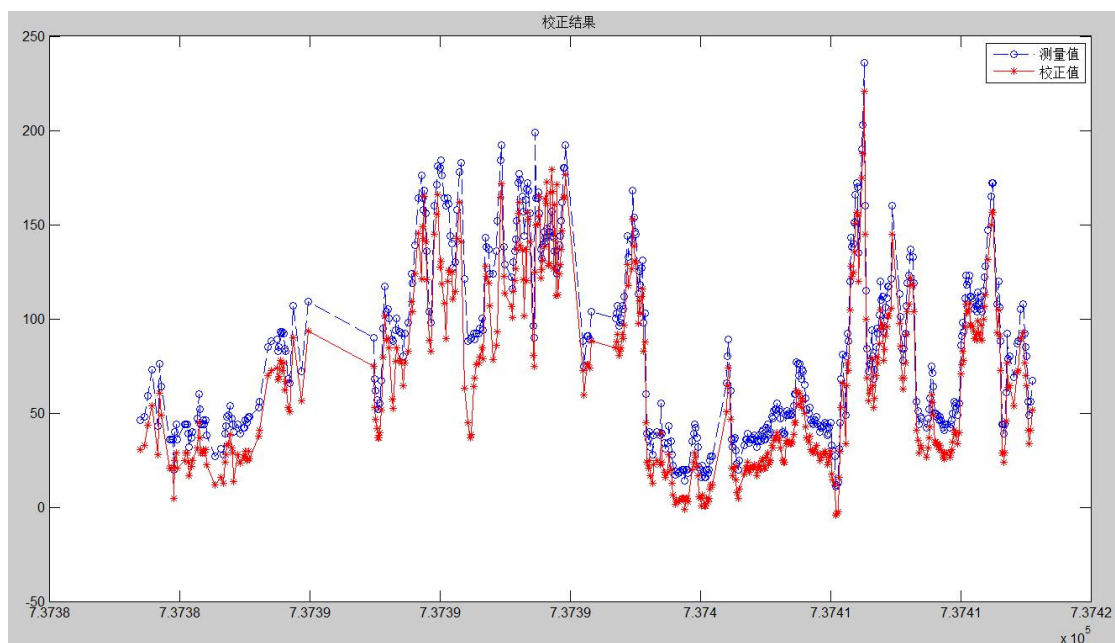


图 5. 采用神经网络进行数据校准举例

部分校准结果如下表所示，完整结果见支撑材料。

表 5. RBF 部分校准结果

时间	PM2.5	PM10	CO	NO <sub>2</sub>	SO <sub>2</sub>	O <sub>3</sub>
2018/11/14 10:00	30.69184	29.38671	1.085954	-3.98165	38.23162	33.72884
2018/11/14 10:02	32.69184	34.38671	1.085954	9.018353	37.23162	29.72884
2018/11/14 10:06	43.69184	44.38671	1.085954	-48.6953	38.23162	70.72884
2018/11/14 10:09	54.00516	67.38671	1.185954	10.09286	38.23162	61.72884
2018/11/14 10:10	27.69184	27.38671	1.085954	8.018232	37.23162	3.72884
2018/11/14 10:14	60.69184	73.38671	1.085954	26.01835	38.23162	52.72884
2018/11/14 10:16	48.69184	55.38671	1.085954	26.01833	38.23162	52.72884
2018/11/14 10:18	20.69184	12.38671	1.085954	-0.93364	37.23162	10.50235
2018/11/14 10:19	20.69184	11.38671	1.085954	2.115929	37.23162	9.129398
2018/11/14 10:23	20.69184	11.38671	1.085954	7.72292	38.23162	60.39163
2018/11/14 10:24	4.691837	-23.6133	1.085954	14.01835	38.23162	51.72884
2018/11/14 10:28	24.69184	23.38671	1.085954	31.54313	38.23162	53.24208
2018/11/14 10:29	28.69184	28.38671	1.085954	34.27333	37.23162	53.72884
2018/11/14 10:33	20.69184	10.38671	1.085954	25.01835	37.23162	52.72884
2018/11/14 10:34	24.60514	31.59475	1.085954	-3.92927	36.23162	5.728504
2018/11/14 10:38	28.59091	34.38672	0.985954	1.008817	37.23162	5.719131
2018/11/14 10:39	28.69145	31.38671	1.085954	32.01835	37.23162	52.72884
2018/11/14 10:42	23.69184	19.38671	1.085954	11.01835	37.23162	38.72884
2018/11/14 10:43	16.69184	5.386705	0.985954	10.01836	38.23162	11.72884
2018/11/14 10:47	21.69184	15.38671	1.085954	11.01835	37.23162	16.72884
2018/11/14 10:48	24.69184	20.38671	1.085954	6.018353	37.23162	22.72884
2018/11/14 10:52	31.69184	40.38671	0.985954	14.01835	38.23162	46.72884

2018/11/14 10:53	44.69184	57.38671	0.985954	9.018353	38.23162	47.72884
2018/11/14 10:57	36.69184	51.38671	1.085954	19.01835	38.23162	47.72884
2018/11/14 10:58	28.69184	33.38671	1.085954	2.018384	37.23162	24.72884
2018/11/14 11:00	28.69184	29.38671	1.085954	2.228208	37.23162	32.72884
2018/11/14 11:02	30.69184	34.38671	1.085954	12.19491	38.23162	46.72884

对比上述 RBF 部分校准结果与国控点所测数据后发现有些误差较大，所以我们考虑了采用逐步多元线性回归校准模型进行校准，结果表明逐步多元线性回归校准模型比 RBF 神经网络校准模型误差小。

#### 7.4 逐步多元线性回归校准模型

逐步多元线性回归中确定一个因变量，再选择几个自变量，分析出因变量与自变量之间的关系，在本文中就是需要校准的数据与原数据之间的关系。多元线性回归分析情况如下：

(1) 将  $PM10_{xz}$  作为因变量，自测点的各项数值作为自变量，使用 SPSS 进行逐步多元线性回归建模<sup>[6]</sup>，结果见下表。

表 6. 将  $PM10_{xz}$  作为因变量的逐步多元线性回归系数表

模型	非标准化系数		标准系数	t	Sig.	B 的 95.0% 置信区间	
	B	标准 误差	试用版			下限	上限
(常量)	1492.989	144.415		10.338	.000	1209.859	1776.119
NPM25	.708	.049	.577	14.541	.000	.612	.803
N 湿度	-1.112	.032	-.479	-34.278	.000	-1.176	-1.048
NNO2	.307	.020	.167	15.705	.000	.269	.346
N 降水量	-.064	.006	-.110	-10.663	.000	-.076	-.053
NCO	33.305	2.979	.140	11.178	.000	27.464	39.146
N 压强	-1.384	.139	-.241	-9.959	.000	-1.657	-1.112
N 温度	-1.463	.165	-.247	-8.890	.000	-1.785	-1.140
NPM10	.106	.027	.167	3.898	.000	.053	.160

可得到多元线性回归模型如下

$$PM10_{xz} = b_0 + b_1 * NPM25 + b_2 * N湿度 + b_3 * NNO_2 + b_4 * N降水量 + b_5 * NCO + b_6 * N压强 + b_7 * N温度 + b_8 * NPM10$$

其中:

$$b_0 = 1492.989, b_1 = 0.708, b_2 = -1.112, b_3 = 0.307, b_4 = -0.064,$$

$$b_5 = 33.305, b_6 = -1.384, b_7 = -1.463, b_8 = 0.106$$

(2) 将  $PM2.5_{xz}$  作为因变量，自测点的各项数值作为自变量，使用 SPSS 进行逐步多元线性回归建模，结果见下表。

表 7. 将 PM2.5<sub>xz</sub> 作为因变量的逐步多元线性回归系数表

模型	非标准化系数		标准系数	t	Sig.	B 的 95.0% 置信区间	
	B	标准 误差	试用版			下限	上限
(常量)	560.233	59.768		9.374	.000	443.057	677.409
NPM25	.730	.020	.876	36.130	.000	.690	.770
N 湿度	-.337	.013	-.213	-24.986	.000	-.363	-.310
N 压强	-.531	.058	-.136	-9.220	.000	-.643	-.418
NNO2	.085	.008	.068	10.315	.000	.069	.101
N 降水量	-.028	.003	-.069	-10.907	.000	-.032	-.023
NCO	11.020	1.272	.068	8.665	.000	8.526	13.513
N 温度	-.353	.068	-.088	-5.155	.000	-.487	-.218
NPM10	.031	.011	.071	2.722	.007	.009	.053
N 风速	-1.058	.413	-.015	-2.560	.011	-1.868	-.248
NSO2	-.019	.009	-.013	-2.284	.022	-.036	-.003

可得到多元线性回归模型如下

$$PM2.5_{xz} = b_0 + b_1 * NPM25 + b_2 * N湿度 + b_3 * N压强 + b_4 * NNO2 + b_5 * N降水量 + b_6 * NCO + b_7 * N温度 + b_8 * NPM10 + b_9 * N风速 + b_{10} * NSO_2$$

其中：

$$b_0 = 560.233, b_1 = 0.730, b_3 = -0.337, b_4 = 0.085, b_5 = -0.028$$

$$b_6 = 11.020, b_7 = -0.353, b_8 = 0.031, b_9 = -1.058, b_{10} = -0.019$$

(3) 将 CO<sub>xz</sub> 作为因变量，自测点的各项数值作为自变量，使用 SPSS 进行逐步多元线性回归建模，结果见下表：

表 8. 将 CO<sub>xz</sub> 作为因变量的逐步多元线性回归系数表

模型	非标准化系数		标准系数	t	Sig.	B 的 95.0% 置信区间	
	B	标准 误差	试用版			下限	上限
(常量)	27.225	1.676		16.244	.000	23.939	30.511
NPM25	.006	.000	.543	40.085	.000	.006	.007
N 压强	-.026	.002	-.469	-16.109	.000	-.029	-.023
NNO2	.002	.000	.128	9.608	.000	.002	.003
NCO	.400	.032	.174	12.658	.000	.338	.462
N 温度	-.020	.002	-.357	-10.463	.000	-.024	-.017
N 湿度	-.003	.000	-.145	-8.435	.000	-.004	-.002
N 降水量	.000	.000	.082	6.434	.000	.000	.001
N 风速	-.063	.012	-.064	-5.319	.000	-.087	-.040

可得到多元线性回归模型如下

$$CO_{XZ} = b_0 + b_1 * NPM25 + b_2 * N压强 + b_3 * NNO_2 + b_4 * NCO + b_5 * N温度 + b_6 * N湿度 + b_7 * N降水量 + b_8 * N风速$$

其中：

$$b_0 = 27.225, b_1 = 0.006, b_2 = -0.026, b_3 = 0.002, b_4 = 0.400, b_5 = -0.020$$

$$b_6 = -0.003, b_7 = 0.000, b_8 = -0.063$$

(4) 将  $NO_{2XZ}$  作为因变量，自测点的各项数值作为自变量，使用 SPSS 进行逐步多元线性回归建模，结果见下表。

表 9. 将  $NO_{2XZ}$  作为因变量的逐步多元线性回归系数表

模型	非标准化系数		标准系数	t	Sig.	B 的 95.0% 置信区间	
	B	标准 误差	试用版			下限	上限
(常量)	1589.198	86.230		18.430	.000	1420.142	1758.254
NNO2	.384	.012	.438	31.992	.000	.361	.408
N 降水量	-.033	.004	-.117	-8.724	.000	-.040	-.025
N 湿度	-.642	.020	-.579	-32.206	.000	-.682	-.603
NO3	-.073	.013	-.095	-5.705	.000	-.098	-.048
NPM25	.430	.031	.733	13.979	.000	.369	.490
NPM10	-.186	.018	-.614	-10.594	.000	-.221	-.152
N 风速	-7.882	.596	-.161	-13.235	.000	-9.049	-6.714
N 温度	-1.845	.101	-.652	-18.312	.000	-2.042	-1.647
N 压强	-1.475	.083	-.538	-17.760	.000	-1.638	-1.312
NCO	-5.794	1.894	-.051	-3.059	.002	-9.506	-2.081
NSO2	.024	.012	.024	1.998	.046	.000	.049

可得到多元线性回归模型如下

$$NO_{2XZ} = b_0 + b_1 * NNO_2 + b_2 * N降水量 + b_3 * N湿度 + b_4 * NO_3 + b_5 * NPM25 + b_6 * NPM10 + b_7 * N风速 + b_8 * N温度 + b_9 * N压强 + b_{10} * NCO + b_{11} * NSO_2$$

其中：

$$b_0 = 1589.198, b_1 = 0.384, b_2 = -0.033, b_3 = -0.642, b_4 = -0.073, b_5 = 0.430$$

$$b_6 = -0.186, b_7 = -7.882, b_8 = -1.845, b_9 = -1.475, b_{10} = -5.794, b_{11} = 0.024$$

(5) 将  $O_{3XZ}$  作为因变量，自测点的各项数值作为自变量，使用 SPSS 进行逐步多元线性回归建模，结果见下表。



表 10. 将 O<sub>3XZ</sub> 作为因变量的逐步多元线性回归系数表

模型	非标准化系数		标准系数	t	Sig.	B 的 95.0% 置信区间	
	B	标准 误差	试用版			下限	上限
(常量)	-1091.908	114.446		-9.541	.000	-1316.282	-867.533
N 温度	3.025	.127	.542	23.769	.000	2.775	3.274
NNO2	-.545	.016	-.315	-34.563	.000	-.576	-.515
N 湿度	-.246	.026	-.112	-9.402	.000	-.297	-.195
NO3	.418	.017	.277	25.130	.000	.385	.450
N 风速	9.396	.801	.097	11.734	.000	7.826	10.965
NPM10	-.575	.022	-.959	-25.631	.000	-.619	-.531
NPM25	.951	.040	.823	23.805	.000	.873	1.030
N 压强	1.104	.110	.204	10.013	.000	.887	1.320
N 降水量	.012	.005	.022	2.451	.014	.002	.022

可得到多元线性回归模型如下

$$O_{3XZ} = b_0 + b_1 * N\text{温度} + b_2 * NNO_2 + b_3 * N\text{湿度} + b_4 * NO_3 + b_5 * N\text{风速} + b_6 * NPM10 \\ + b_7 * NPM25 + b_8 * N\text{压强} + b_9 * N\text{降水量}$$

其中：

$$b_0 = -1091.908, b_1 = 3.025, b_2 = -0.545, b_3 = -0.246, b_4 = 0.418, b_5 = 9.396$$

$$b_6 = -0.575, b_7 = 0.951, b_8 = 1.104, b_9 = 0.012$$

(6) 将 SO<sub>2XZ</sub> 作为因变量，自测点的各项数值作为自变量，使用 SPSS 进行逐步多元线性回归建模，结果见下表。

表 11. 将 SO<sub>2XZ</sub> 作为因变量的逐步多元线性回归系数表

模型	非标准化系数		标准系数	t	Sig.	B 的 95.0% 置信区间	
	B	标准 误差	试用版			下限	上限
(常量)	-363.904	36.933		-9.853	.000	-436.313	-291.496
NCO	33.890	1.560	.363	21.723	.000	30.832	36.949
N 压强	.346	.036	.153	9.650	.000	.276	.416
N 降水量	.022	.003	.098	6.865	.000	.016	.029
NPM10	.099	.015	.397	6.471	.000	.069	.129
NPM25	-.146	.027	-.302	-5.406	.000	-.198	-.093
N 风速	-2.990	.522	-.074	-5.724	.000	-4.014	-1.966
NO3	.079	.011	.126	7.242	.000	.058	.101
NNO2	.066	.010	.091	6.250	.000	.045	.086
NSO2	-.060	.011	-.071	-5.584	.000	-.081	-.039
N 湿度	.031	.014	.034	2.191	.029	.003	.058

可得到多元线性回归模型如下

$$SO_{2xz} = b_0 + b_1 * NCO + b_2 * N压强 + b_3 * N降水量 + b_4 * NPM10 + b_5 * NPM25 \\ + b_6 * N风速 + b_7 * NO_3 + b_8 * NNO_2 + b_9 * NSO_2 + b_{10} * N湿度$$

其中：

$$b_0 = -363.904, b_1 = 33.890, b_2 = 0.346, b_3 = 0.022, b_4 = 0.099, b_5 = -0.146$$

$$b_6 = -2.990, b_7 = 0.079, b_8 = 0.066, b_9 = -0.060, b_{10} = 0.031$$

利用这些关系式对自测点的数据进行校准，部分校准结果如下图所示，完整校准结果见支称材料

表 12.逐步多元线性回归部分校准结果

时间	PM2.5	PM10	CO	NO <sub>2</sub>	SO <sub>2</sub>	O <sub>3</sub>
2018/11/14 10:02	41.39	71.9724	0.9069	26.9504	28.0592	46.633
2018/11/14 10:06	38.8327	68.7165	0.7691	13.6125	21.0868	61.5088
2018/11/14 10:09	39.493	67.6059	0.8559	26.3588	24.5602	50.512
2018/11/14 10:10	38.5638	67.7009	0.8012	19.948	21.8162	59.9914
2018/11/14 10:14	38.0408	67.0457	0.7927	19.5692	22.45	56.291
2018/11/14 10:16	38.0378	66.9297	0.7877	18.4562	22.392	59.971
2018/11/14 10:18	39.3397	67.3751	0.8616	27.5849	25.7914	49.398
2018/11/14 10:19	41.1398	70.4817	0.8157	20.7822	21.979	62.11
2018/11/14 10:23	39.1526	69.9422	0.834	19.077	25.941	61.5924
2018/11/14 10:24	37.4056	66.1987	0.7605	15.3714	21.05	68.2464
2018/11/14 10:28	34.9378	66.5755	0.6907	7.9244	18.2088	74.2246
2018/11/14 10:29	39.4838	69.5245	0.8072	20.2664	22.5188	62.5976
2018/11/14 10:33	34.8097	65.4319	0.7101	12.6281	19.4962	72.2906
2018/11/14 10:34	36.1136	66.5555	0.7885	21.6376	22.5538	59.95
2018/11/14 10:38	34.7776	65.5395	0.7175	11.4096	19.9718	72.884
2018/11/14 10:39	38.4535	66.6569	0.8189	23.0553	24.5802	61.893
2018/11/14 10:42	34.9984	65.1393	0.7763	21.553	23.3376	65.781
2018/11/14 10:43	36.6559	66.5339	0.7963	22.8599	23.8502	63.3232
2018/11/14 10:47	36.4834	68.4093	0.7598	17.958	22.0546	67.766
2018/11/14 10:48	34.3423	64.2627	0.7772	21.6747	23.682	64.325
2018/11/14 10:52	37.6598	65.9603	0.8062	21.8866	24.0246	66.2132
2018/11/14 10:53	34.2457	62.8427	0.7596	17.8323	23.061	68.7752
2018/11/14 10:57	36.1084	67.1713	0.7853	22.131	23.4566	66.664
2018/11/14 10:58	34.5347	68.8002	0.7466	13.1079	23.241	76.0332
2018/11/14 11:02	33.7503	61.6595	0.7727	20.3569	23.0668	70.2426
2018/11/14 11:07	34.6808	68.6984	0.7832	17.1206	25.0852	72.2484

将上述三种方法的校准结果与国控数据对比之后,发现逐步多元线性回归模型校准的结果误差最小,采用该模型得到的校准数据与国控数据的误差分析结果见下所示,由此可见逐步多元线性回归校准模型是简单、实用、可靠的校准模型。

表 13.校准数据与国控数据之间的误差分析

	PM25	PM10	CO	NO2	SO2	O3
数据之间误差的标准差	12.4562853	30.19392415	0.361734234	17.94346805	15.75096246	24.15714565
数据之间误差的均值	-0.475222446	0.069381626	-0.07100922	0.304541139	-0.356725765	0.640619656
数据之间绝对误差的标准差	9.544020855	25.77405167	0.24908357	11.49880581	11.79364236	15.35081887
数据之间绝对误差的均值	8.017554471	15.72632245	0.272218647	13.81996337	10.41426456	18.55974054

## 8 模型的结论

对于问题一,使用SPSS进行统计分析,对各个污染物分布特性规律有了整体把握,并通过BDP作图,可视化分析结果。可知该城市主要污染物是CO和颗粒物,污染最重的时候是在下午,在凌晨期间污染物随着大气的自然降解和沉淀作用逐渐下降。

对于问题二,通过删除缺失、重复数据进行数据预处理,并利用 matlab 进行数据标准化,之后利用 matlab 进行三次样条函数插值,保证了自建数据和国控数据处于同样的时间戳下。通过对两者的数据做差,并将该差值与温度、湿度、压强、降水量、风速做灰色关联分析,得到几个变量之间的灰色关联度,从而得到以下结论(1)风速与 PM2.5、PM10、NO<sub>2</sub>、SO<sub>2</sub> 的误差有明显的关联,灰色关联度均大于 0.9;(2)压强与 CO 的误差具有较高的灰色关联度;(3)降水量对 PM2.5、PM10、NO<sub>2</sub>、SO<sub>2</sub> 的误差的灰色关联度较大。(4)温度会对 PM2.5、PM10、NO<sub>2</sub>、SO<sub>2</sub> 的测量造成误差。(5)湿度只与 O<sub>3</sub> 的误差相关。

对于问题三,在模型二样条函数插值的基础上,分别采用加权平均模型和 RBF 神经网络两种模型对数据进行校准。一将问题二模型中经过插值后的国控点数据和自建点数据加权平均,将国控点数据的估计值引入自建点数据,修正自建点数据。二是通过建立 RBF 神经网络,预测每次的测量误差,从而根据误差和自建点数据预测最终误差。三是通过逐步多元线性回归分析出其中的关系式,从而通过计算,校准自测点的数据。模型校准的结果见支撑材料。

## 9 模型的优缺点分析与推广

模型一利用 spss 进行统计分析, 分析数据的统计分布规律, 并利用 BDP 可视化直观显示数据的趋势性变化规律和全天各个时刻的变化规律。操作简单方便迅速, 可以迅速得到数据的分布规律。但是若数据过长过多的时候, 可视化方法将会操作实施困难, 需要对数据进行采样来进行可视化分析。

模型二在数据预处理和数据标准化的基础上采用三次样条函数插值, 统一时间戳, 便于后续的数据处理。之后才用灰色关联分析的方法, 针对各个污染物测量误差与各个可能的影响因素, 建立灰关联分析模型, 通过灰色关联度, 比较各个影响因素与误差之间的相关程度。有效利用了灰色关联分析模型针对序列分析的优势。其中通过插值统一时间戳的方法可以推广到其他时间戳不相等的场所, 从而加快数据校准操作。

模型三在模型二样条函数插值的基础上使用了加权平均模型和 RBF 神经网络模型, 加权平均模型易操作, 好实现, 但是准确率较低。RBF 神经网络利用 RBF 的强拟合能力, 对误差进行了有效的拟合, 从而可以通过预测误差的方法, 间接校准模型, 但是操作费时费力, 面对大规模数据时无能为力。采用 SPSS 逐步多元线性回归的方法利用分析所得各数据之间的关系式, 从而计算出校准之后的自测点数据, 该方法误差更小, 且简单方便, 便于推广。

### 参考文献

- 【1】韩中庚, 数学建模方法及其应用, 北京: 高等教育出版社, 2017.
- 【2】李昕 编著, MATLAB 数学建模, 北京: 清华大学出版社, 2017.
- 【3】姜启源、谢金星、叶俊 编著, 数学模型, 北京: 高等教育出版社, 2018.
- 【4】袁震东, 蒋鲁敏, 束金龙 编著, 数学建模简明教程, 上海: 华东师范大学出版社, 2002 年出版.
- 【5】刘金琨 编著, RBF 神经网络自适应控制及 Matlab 仿真, 北京: 清华大学出版社, 2019.
- 【6】张文彤 编著, SPSS 统计分析基础课程, 北京: 高等教育出版社, 2017.

## 附录一

### Matlab 代码程序

```
load('matlab.mat')
guo_2= interp1(guo(:,7),guo(:,1:6),time,'spline');
zizhi_2= interp1(zizhi(:,12),zizhi(:,1:11),time,'spline');
wucha=zizhi_2(:,1:6)-guo_2(:,1:6);

stats=[wucha(:,1),zizhi_2(:,2),zizhi_2(:,3),zizhi_2(:,4),zizhi_2(:,5),zizhi_2(:,6)];
stats=stats';
T=GreyRelationDegree(stats)
T1=mean(T,2);

stats=[wucha(:,2),zizhi_2(:,2),zizhi_2(:,3),zizhi_2(:,4),zizhi_2(:,5),zizhi_2(:,6)];
stats=stats';
T=GreyRelationDegree(stats)
T2=mean(T,2);

stats=[wucha(:,3),zizhi_2(:,2),zizhi_2(:,3),zizhi_2(:,4),zizhi_2(:,5),zizhi_2(:,6)];
stats=stats';
T=GreyRelationDegree(stats)
T3=mean(T,2);

stats=[wucha(:,4),zizhi_2(:,2),zizhi_2(:,3),zizhi_2(:,4),zizhi_2(:,5),zizhi_2(:,6)];
stats=stats';
T=GreyRelationDegree(stats)
T4=mean(T,2);

stats=[wucha(:,5),zizhi_2(:,2),zizhi_2(:,3),zizhi_2(:,4),zizhi_2(:,5),zizhi_2(:,6)];
stats=stats';
T=GreyRelationDegree(stats)
T5=mean(T,2);

stats=[wucha(:,6),zizhi_2(:,2),zizhi_2(:,3),zizhi_2(:,4),zizhi_2(:,5),zizhi_2(:,6)];
stats=stats';
T=GreyRelationDegree(stats)
T6=mean(T,2);

%训练 RBF 神经网络，选择其目标变量为误差，自变量为前面相关的几个变量
net=newrbf(zizhi_2(1:10000,7:11)',wucha(1:10000,1)',0,0.5,85,5)
%用训练好的 RBF 拟合残差
```

```

t=sim(net,zizhi_2(:,7:11));
%得到新的预测数据
G2=zizhi_2(:,1)-t';
%画出预测数据和实际数据的图
plot(time(1:500),zizhi_2(1:500,1),'bo--');
hold on;
plot(time(1:500),G2(1:500,1),'r*-');
title('校准结果');
legend('测量值','校准值');

%训练 RBF 神经网络，选择其目标变量为误差，自变量为前面相关的几个变量
net=newrb(zizhi_2(1:10000,7:11)',wucha(1:10000,2)',0,0.5,85,5)
%用训练好的 RBF 拟合残差
t=sim(net,zizhi_2(:,7:11));
%得到新的预测数据
G3=zizhi_2(:,2)-t';

%训练 RBF 神经网络，选择其目标变量为误差，自变量为前面相关的几个变量
net=newrb(zizhi_2(1:10000,7:11)',wucha(1:10000,3)',0,0.5,85,5)
%用训练好的 RBF 拟合残差
t=sim(net,zizhi_2(:,7:11));
%得到新的预测数据
G4=zizhi_2(:,3)-t';

%训练 RBF 神经网络，选择其目标变量为误差，自变量为前面相关的几个变量
net=newrb(zizhi_2(1:10000,7:11)',wucha(1:10000,4)',0,0.5,85,5)
%用训练好的 RBF 拟合残差
t=sim(net,zizhi_2(:,7:11));
%得到新的预测数据
G5=zizhi_2(:,4)-t';

%训练 RBF 神经网络，选择其目标变量为误差，自变量为前面相关的几个变量
net=newrb(zizhi_2(1:10000,7:11)',wucha(1:10000,5)',0,0.5,85,5)
%用训练好的 RBF 拟合残差
t=sim(net,zizhi_2(:,7:11));
%得到新的预测数据
G6=zizhi_2(:,5)-t';

%训练 RBF 神经网络，选择其目标变量为误差，自变量为前面相关的几个变量
net=newrb(zizhi_2(1:10000,7:11)',wucha(1:10000,6)',0,0.5,85,5)
%用训练好的 RBF 拟合残差
t=sim(net,zizhi_2(:,7:11));
%得到新的预测数据
G7=zizhi_2(:,6)-t';

```

## 附录二

```

GET DATA
  /TYPE=TXT
  /FILE="C:\Users\Administrator\Desktop\CUMCM2019Problems22\D-2019 中文\附件 1.csv"
  /DELCASE=LINE
  /DELIMITERS=","
  /ARRANGEMENT=DELIMITED
  /FIRSTCASE=2
  /IMPORTCASE=ALL
  /VARIABLES=
  PM2.5 F3.0
  PM10 F3.0
  CO F5.3
  NO2 F2.0
  SO2 F3.0
  O3 F3.0
  时间 A16.
CACHE.
EXECUTE.
DATASET NAME 数据集 1 WINDOW=FRONT.
DESCRIPTIVES VARIABLES=PM2.5 PM10 CO NO2 SO2 O3
  /SAVE
  /STATISTICS=MEAN SUM STDDEV VARIANCE RANGE MIN MAX SEMEAN
KURTOSIS SKEWNESS.

```

### 描述

附注		12-SEP-2019 22:45:34
创建的输出		
注释		
	数据	C:\Users\Administrator\Desktop\CUMCM2019Problems22\D-2019 中文\附件 1.csv
输入	活动的数据集	数据集 1
	过滤器	<none>
	权重	<none>

缺失值处理	拆分文件	<none>
	工作数据文件中的 N 行	4200
	对缺失的定义	用户定义的缺失值作为缺失数据对待。
	使用的案例	使用所有非缺失数据。
语法		DESCRIPTIVES VARIABLES=PM2.5 PM10 CO NO2 SO2 O3  /SAVE  /STATISTICS=MEAN SUM STDDEV VARIANCE RANGE MIN MAX SEMEAN KURTOSIS SKEWNESS.
	资源	
	处理器时间	00:00:00.05
	已用时间	00:00:00.02
已?建或修改的?量	ZPM2.5	Zscore(PM2.5)
	ZPM10	Zscore(PM10)
	ZCO	Zscore(CO)
	ZNO2	Zscore(NO2)
	ZSO2	Zscore(SO2)
	ZO3	Zscore(O3)

[数据集 1]

描述统计量

	N	全距	极小值	极大值	和	均值
	统计量	统计量	统计量	统计量	统计量	统计量
PM2.5	4200	245	1	246	238247	56.73
PM10	4200	983	2	985	352054	83.82
CO	4200	3.845	.050	3.895	4700.577	1.11919
NO2	4200	136	5	141	137101	32.64
SO2	4200	149	1	150	94100	22.40
O3	4200	258	1	259	230018	54.77
有效的 N (列表状态)	4200					

描述统计量

	均值	标准差	方差	偏度		峰度
	标准误	统计量	统计量	统计量	标准误	统计量
PM2.5	.533	34.569	1195.004	1.143	.038	1.524
PM10	.785	50.865	2587.273	3.316	.038	46.782



CO	.007593	.492093	.242	.944	.038	1.955
NO2	.375	24.303	590.657	1.233	.038	1.335
SO2	.309	20.026	401.039	2.064	.038	4.496
O3	.740	47.989	2302.969	1.385	.038	2.094
有效的 N (列表状态)						

描述统计量

	峰度
	标准误
PM2.5	.076
PM10	.076
CO	.076
NO2	.076
SO2	.076
O3	.076
有效的 N (列表状态)	

DESCRIPTIVES VARIABLES=PM2.5 PM10 CO NO2 SO2 O3

/SAVE

/STATISTICS=MEAN SUM STDDEV VARIANCE RANGE MIN MAX KURTOSIS  
SKEWNESS.

描述

附注

创建的输出	12-SEP-2019 22:49:06
注释	
数据	C:\Users\Administrator\Desktop\CUMC M2019Problems22\D-2019 中文\附件 1.csv
输入	数据源 1
过滤器	<none>
权重	<none>
拆分文件	<none>
工作数据文件中的 N 行	4200
对缺失的定义	用户定义的缺失值作为缺失数据对待。
缺失值处理	使用的案例 使用所有非缺失数据。

		DESCRIPTIVES VARIABLES=PM2.5
		PM10 CO NO2 SO2 O3
		/SAVE
		/STATISTICS=MEAN SUM
		STDDEV VARIANCE RANGE MIN
		MAX KURTOSIS SKEWNESS.
语法		
资源	处理器时间	00:00:00.03
	已用时间	00:00:00.03
	ZSco01	Zscore(PM2.5)
	ZSco02	Zscore(PM10)
已?建或修改的?量	ZSco03	Zscore(CO)
	ZSco04	Zscore(NO2)
	ZSco05	Zscore(SO2)
	ZSco06	Zscore(O3)

[数据集 1]

描述统计量

	N	全距	极小值	极大值	和	均值
	统计量	统计量	统计量	统计量	统计量	统计量
PM2.5	4200	245	1	246	238247	56.73
PM10	4200	983	2	985	352054	83.82
CO	4200	3.845	.050	3.895	4700.577	1.11919
NO2	4200	136	5	141	137101	32.64
SO2	4200	149	1	150	94100	22.40
O3	4200	258	1	259	230018	54.77
有效的 N (列表状态)	4200					

描述统计量

	标准差	方差	偏度		峰度	
	统计量	统计量	统计量	标准误	统计量	标准误
PM2.5	34.569	1195.004	1.143	.038	1.524	.076
PM10	50.865	2587.273	3.316	.038	46.782	.076
CO	.492093	.242	.944	.038	1.955	.076
NO2	24.303	590.657	1.233	.038	1.335	.076
SO2	20.026	401.039	2.064	.038	4.496	.076
O3	47.989	2302.969	1.385	.038	2.094	.076
有效的 N (列表状态)						

附注

创建的输出	12-SEP-2019 23:14:44
-------	----------------------

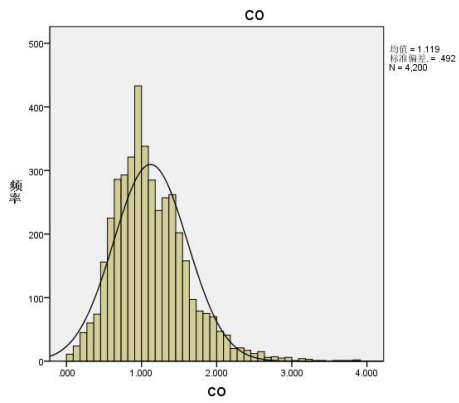
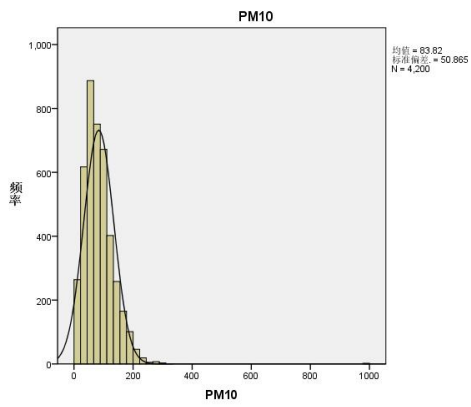
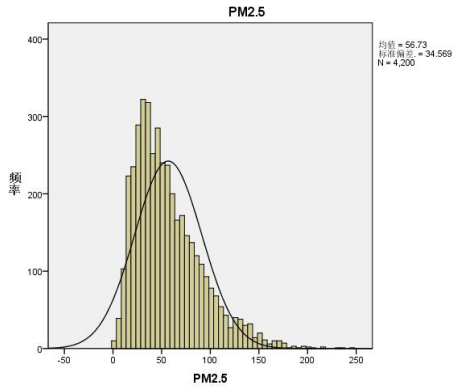
注释		C:\Users\Administrator\Desktop\CUMC
	数据	M2019Problems22\D-2019 中文\附件
		2.csv
输入	活动的数据集	数据集 2
	过滤器	<none>
	权重	<none>
	拆分文件	<none>
	工作数据文件中的 N 行	234717
	对缺失的定义	用户定义的丢失值作为丢失对待。
缺失值处理	使用的案例	统计量的计算将基于所有包含有效数据的案例。
		FREQUENCIES VARIABLES=PM2.5
		PM10 CO NO2 SO2 O3
		/NTILES=4
语法		/STATISTICS=STDDEV RANGE
		MINIMUM MAXIMUM MEAN
		MEDIAN MODE SUM SKEWNESS
		SESKEW KURTOSIS SEKURT
		/ORDER=ANALYSIS.
资源	处理器时间	00:00:00.41
	已用时间	00:00:00.50

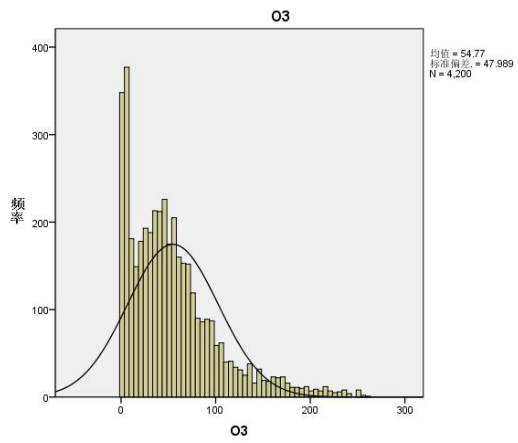
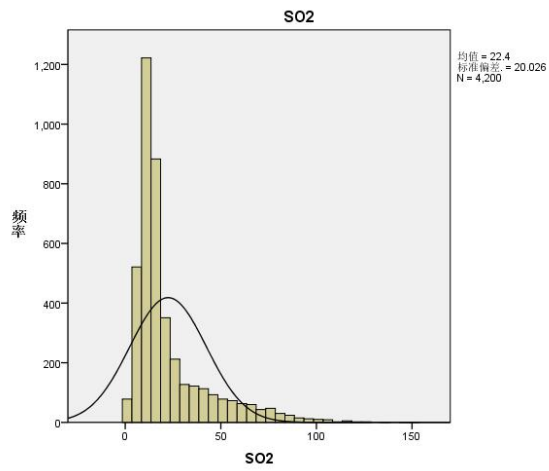
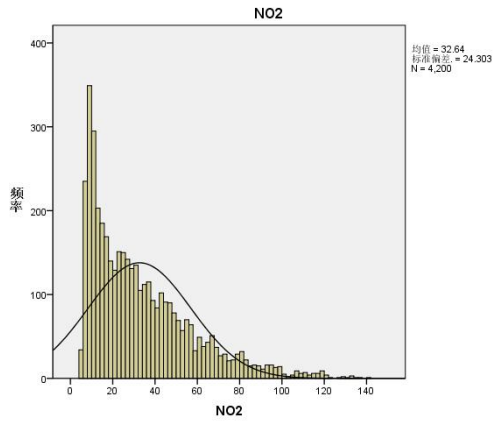
[数据集 2]

		统计量					
		PM2.5	PM10	CO	NO2	SO2	O3
N	有效	234717	234717	234717	234717	234717	234717
	缺失	0	0	0	0	0	0
	均值	69.35	113.47	.589	55.51	16.34	66.09
	中值	59.00	93.00	.500	51.00	16.00	62.00
	众数	40	72	.4	26	17	40
	标准差	38.510	71.698	.2176	28.735	20.006	33.659
	偏度	1.060	1.521	1.670	.687	37.003	.986
	偏度的标准误	.005	.005	.005	.005	.005	.005
	峰度	1.380	3.080	4.211	-.305	1503.600	1.188
	峰度的标准误	.010	.010	.010	.010	.010	.010
	全距	547	931	2.9	181	1101	249
	极小值	1	2	.0	0	2	0
	极大值	548	933	2.9	181	1103	249
	和	16277696	26633136	138225.0	13028418	3836103	15511917

	25	41.00	66.00	.400	28.00	14.00	39.00
百分位数	50	59.00	93.00	.500	51.00	16.00	62.00
	75	92.00	144.00	.700	76.00	17.00	85.00

### 直方图





### 附录三

#### SPSS 多元逐步线性回归分析

系数<sup>a</sup>

模型		非标准化系数		标准系数	t	Sig.	B 的 95.0% 置信区间	
		B	标准 误差	试用版			下限	上限
1	(常量)	22.389	1.168		19.165	.000	20.099	24.679
	NPM25	.836	.014	.682	60.390	.000	.809	.863
2	(常量)	70.564	1.695		41.622	.000	67.240	73.887
	NPM25	1.006	.013	.821	77.194	.000	.981	1.032
	N 湿度	-.880	.025	-.379	-35.647	.000	-.929	-.832
3	(常量)	61.266	1.741		35.198	.000	57.854	64.679
	NPM25	.958	.013	.782	73.769	.000	.933	.984
	N 湿度	-.951	.024	-.409	-39.059	.000	-.998	-.903
	NNO2	.306	.019	.167	16.289	.000	.269	.343
4	(常量)	66.622	1.776		37.514	.000	63.140	70.103
	NPM25	.937	.013	.764	72.529	.000	.912	.962
	N 湿度	-.939	.024	-.404	-39.153	.000	-.986	-.892
	NNO2	.380	.020	.207	19.398	.000	.341	.418
	N 降水量	-.067	.006	-.116	-11.514	.000	-.079	-.056
5	(常量)	52.511	2.078		25.264	.000	48.436	56.586
	NPM25	.901	.013	.735	69.196	.000	.875	.926
	N 湿度	-.905	.024	-.390	-38.154	.000	-.951	-.858
	NNO2	.335	.020	.183	17.152	.000	.297	.374
	N 降水量	-.075	.006	-.129	-12.983	.000	-.086	-.064
	NCO	29.398	2.356	.124	12.480	.000	24.780	34.017
6	(常量)	217.500	58.792		3.699	.000	102.235	332.764
	NPM25	.911	.013	.743	67.574	.000	.884	.937
	N 湿度	-.902	.024	-.388	-38.002	.000	-.948	-.855
	NNO2	.332	.020	.181	16.942	.000	.293	.370
	N 降水量	-.071	.006	-.121	-11.725	.000	-.082	-.059
	NCO	27.962	2.409	.118	11.609	.000	23.240	32.684
	N 压强	-.162	.058	-.028	-2.808	.005	-.276	-.049
7	(常量)	1369.379	141.128		9.703	.000	1092.692	1646.065
	NPM25	.890	.014	.726	65.718	.000	.864	.917
	N 湿度	-1.101	.032	-.474	-34.011	.000	-1.165	-1.038
	NNO2	.306	.020	.167	15.619	.000	.268	.345
	N 降水量	-.062	.006	-.107	-10.356	.000	-.074	-.051
	NCO	38.550	2.663	.162	14.478	.000	33.330	43.770
	N 压强	-1.267	.136	-.221	-9.320	.000	-1.533	-1.000
	N 温度	-1.477	.165	-.249	-8.961	.000	-1.800	-1.153

8	(常量)	1492.989	144.415		10.338	.000	1209.859	1776.119
	NPM25	.708	.049	.577	14.541	.000	.612	.803
	N 湿度	-1.112	.032	-.479	-34.278	.000	-1.176	-1.048
	NNO2	.307	.020	.167	15.705	.000	.269	.346
	N 降水量	-.064	.006	-.110	-10.663	.000	-.076	-.053
	NCO	33.305	2.979	.140	11.178	.000	27.464	39.146
	N 压强	-1.384	.139	-.241	-9.959	.000	-1.657	-1.112
	N 温度	-1.463	.165	-.247	-8.890	.000	-1.785	-1.140
	NPM10	.106	.027	.167	3.898	.000	.053	.160

a. 因变量: PM10

$$PM10 = b_0 + b_1 * NPM25 + b_2 * N \text{ 湿度} + b_3 * NNO2 + b_4 * N \text{ 降水量} + b_5 * NCO + b_6 * N \text{ 压强} + b_7 * N \text{ 温度} + b_8 * NPM10;$$

模型		非标准化系数		标准系数	t	Sig.	B 的 95.0% 置信区间	
		B	标准 误差	试用版			下限	上限
1	(常量)	1.195	.457		2.614	.009	.299	2.091
	NPM25	.756	.005	.907	139.524	.000	.745	.766
2	(常量)	16.880	.694		24.339	.000	15.520	18.239
	NPM25	.811	.005	.974	152.089	.000	.801	.822
	N 湿度	-.287	.010	-.182	-28.372	.000	-.306	-.267
3	(常量)	373.894	23.585		15.853	.000	327.656	420.133
	NPM25	.825	.005	.991	156.355	.000	.815	.836
	N 湿度	-.275	.010	-.174	-27.857	.000	-.294	-.256
	N 压强	-.352	.023	-.090	-15.144	.000	-.398	-.307
4	(常量)	366.320	23.306		15.718	.000	320.629	412.011
	NPM25	.813	.005	.976	151.767	.000	.802	.823
	N 湿度	-.293	.010	-.186	-29.593	.000	-.313	-.274
	N 压强	-.347	.023	-.089	-15.107	.000	-.392	-.302
	NNO2	.079	.008	.063	10.313	.000	.064	.094
5	(常量)	307.843	23.727		12.974	.000	261.324	354.361
	NPM25	.803	.005	.963	148.953	.000	.792	.813
	N 湿度	-.291	.010	-.184	-29.722	.000	-.310	-.272
	N 压强	-.287	.023	-.074	-12.263	.000	-.333	-.242
	NNO2	.107	.008	.085	13.292	.000	.091	.122
	N 降水量	-.025	.002	-.063	-10.201	.000	-.030	-.020
6	(常量)	253.409	24.117		10.508	.000	206.127	300.690
	NPM25	.788	.006	.946	142.575	.000	.777	.799

	N 湿度	-281	.010	-.178	-28.893	.000	-.300	-.262
	N 压强	-.238	.024	-.061	-10.043	.000	-.285	-.192
	NNO2	.094	.008	.075	11.664	.000	.078	.109
	N 降水量	-.029	.002	-.073	-11.729	.000	-.034	-.024
	NCO	9.656	.988	.060	9.773	.000	7.718	11.593
7	(常量)	522.906	58.263		8.975	.000	408.679	637.134
	NPM25	.783	.006	.940	140.086	.000	.772	.794
	N 湿度	-.328	.013	-.208	-24.530	.000	-.354	-.302
	N 压强	-.497	.056	-.127	-8.851	.000	-.607	-.387
	NNO2	.088	.008	.070	10.837	.000	.072	.104
	N 降水量	-.027	.002	-.068	-10.868	.000	-.032	-.022
	NCO	12.133	1.099	.075	11.037	.000	9.978	14.288
	N 温度	-.345	.068	-.086	-5.078	.000	-.479	-.212
8	(常量)	563.207	59.661		9.440	.000	446.240	680.174
	NPM25	.724	.020	.869	35.995	.000	.684	.763
	N 湿度	-.331	.013	-.210	-24.728	.000	-.358	-.305
	N 压强	-.535	.057	-.137	-9.316	.000	-.647	-.422
	NNO2	.088	.008	.071	10.895	.000	.072	.104
	N 降水量	-.028	.002	-.070	-11.098	.000	-.033	-.023
	NCO	10.423	1.231	.065	8.468	.000	8.010	12.836
	N 温度	-.341	.068	-.085	-5.016	.000	-.474	-.208
	NPM10	.035	.011	.080	3.076	.002	.013	.057
9	(常量)	553.860	59.732		9.272	.000	436.753	670.968
	NPM25	.725	.020	.870	36.070	.000	.686	.765
	N 湿度	-.335	.013	-.212	-24.877	.000	-.361	-.308
	N 压强	-.525	.058	-.135	-9.120	.000	-.637	-.412
	NNO2	.084	.008	.067	10.196	.000	.068	.100
	N 降水量	-.027	.003	-.068	-10.720	.000	-.032	-.022
	NCO	10.287	1.231	.064	8.355	.000	7.873	12.701
	N 温度	-.334	.068	-.083	-4.920	.000	-.468	-.201
	NPM10	.034	.011	.078	3.000	.003	.012	.056
	N 风速	-1.061	.414	-.015	-2.567	.010	-1.872	-.251
10	(常量)	560.233	59.768		9.374	.000	443.057	677.409
	NPM25	.730	.020	.876	36.130	.000	.690	.770
	N 湿度	-.337	.013	-.213	-24.986	.000	-.363	-.310
	N 压强	-.531	.058	-.136	-9.220	.000	-.643	-.418
	NNO2	.085	.008	.068	10.315	.000	.069	.101
	N 降水量	-.028	.003	-.069	-10.907	.000	-.032	-.023
	NCO	11.020	1.272	.068	8.665	.000	8.526	13.513
	N 温度	-.353	.068	-.088	-5.155	.000	-.487	-.218



NPM10	.031	.011	.071	2.722	.007	.009	.053
N 风速	-1.058	.413	-.015	-2.560	.011	-1.868	-.248
NSO2	-.019	.009	-.013	-2.284	.022	-.036	-.003

a.因变量 PM2.5

$PM2.5 = b_0 + b_1 * NPM25 + b_2 * N \text{ 湿度} + b_3 * N \text{ 压强} + b_4 * NNO2 + b_5 * N \text{ 降水量} + b_6 * NCO + b_7 * N \text{ 温度} + b_8 * NPM10 + b_9 * N \text{ 风速} + b_{10} * NSO2;$

模型		非标准化系数		标准系数	t	Sig.	B 的 95.0% 置信区间	
		B	标准 误差	试用版			下限	上限
1	(常量)	.599	.012		48.308	.000	.574	.623
	NPM25	.007	.000	.597	48.267	.000	.007	.007
2	(常量)	12.286	.692		17.744	.000	10.929	13.644
	NPM25	.008	.000	.643	52.370	.000	.007	.008
	N 压强	-.012	.001	-.207	-16.883	.000	-.013	-.010
3	(常量)	12.049	.671		17.966	.000	10.734	13.364
	NPM25	.007	.000	.581	46.653	.000	.007	.007
	N 压强	-.011	.001	-.206	-17.312	.000	-.013	-.010
	NNO2	.004	.000	.204	16.720	.000	.003	.004
4	(常量)	10.677	.677		15.769	.000	9.349	12.004
	NPM25	.007	.000	.553	43.803	.000	.006	.007
	N 压强	-.010	.001	-.184	-15.362	.000	-.012	-.009
	NNO2	.003	.000	.178	14.425	.000	.003	.004
	NCO	.283	.028	.123	9.978	.000	.227	.338
5	(常量)	18.150	1.302		13.936	.000	15.596	20.703
	NPM25	.006	.000	.521	38.832	.000	.006	.006
	N 压强	-.017	.001	-.314	-13.784	.000	-.020	-.015
	NNO2	.003	.000	.160	12.784	.000	.002	.003
	NCO	.367	.031	.160	11.897	.000	.307	.428
	N 温度	-.010	.001	-.168	-6.704	.000	-.012	-.007
6	(常量)	25.310	1.633		15.495	.000	22.108	28.513
	NPM25	.006	.000	.529	39.513	.000	.006	.007
	N 压强	-.024	.002	-.436	-15.407	.000	-.027	-.021
	NNO2	.003	.000	.166	13.342	.000	.003	.003
	NCO	.425	.032	.185	13.389	.000	.362	.487
	N 温度	-.019	.002	-.334	-9.835	.000	-.023	-.015
	N 湿度	-.003	.000	-.123	-7.189	.000	-.004	-.002
7	(常量)	27.723	1.679		16.513	.000	24.431	31.014
	NPM25	.006	.000	.544	40.054	.000	.006	.007
	N 压强	-.027	.002	-.479	-16.435	.000	-.030	-.023

	NNO2	.003	.000	.142	10.808	.000	.002	.003
	NCO	.411	.032	.179	12.978	.000	.349	.473
	N 温度	-.021	.002	-.364	-10.637	.000	-.025	-.017
	N 湿度	-.003	.000	-.136	-7.916	.000	-.004	-.002
	N 降水量	.000	.000	.074	5.832	.000	.000	.001
8	(常量)	27.225	1.676		16.244	.000	23.939	30.511
	NPM25	.006	.000	.543	40.085	.000	.006	.007
	N 压强	-.026	.002	-.469	-16.109	.000	-.029	-.023
	NNO2	.002	.000	.128	9.608	.000	.002	.003
	NCO	.400	.032	.174	12.658	.000	.338	.462
	N 温度	-.020	.002	-.357	-10.463	.000	-.024	-.017
	N 湿度	-.003	.000	-.145	-8.435	.000	-.004	-.002
	N 降水量	.000	.000	.082	6.434	.000	.000	.001
	N 风速	-.063	.012	-.064	-5.319	.000	-.087	-.040

a.因变量 co

$CO=b_0+b_1*NPM25+b_2*N \text{ 压强}+b_3*NNO2+b_4*NCO+b_5*N \text{ 温度}+b_6*N \text{ 湿度}+b_7*N \text{ 降水量}+b_8*N \text{ 风速};$

模型	非标准化系数		标准系数	t	Sig.	B 的 95.0% 置信区间		
	B	标准 误差	试用版			下限	上限	
1	(常量)	14.202	.809		17.560	.000	12.617	15.788
	NNO2	.319	.013	.364	25.289	.000	.294	.344
2	(常量)	20.225	.828		24.417	.000	18.601	21.849
	NNO2	.399	.013	.455	31.454	.000	.374	.424
	N 降水量	-.081	.004	-.291	-20.135	.000	-.089	-.073
3	(常量)	35.760	1.177		30.382	.000	33.452	38.067
	NNO2	.457	.013	.521	36.154	.000	.433	.482
	N 降水量	-.081	.004	-.292	-20.946	.000	-.089	-.074
	N 湿度	-.274	.015	-.247	-17.961	.000	-.304	-.244
4	(常量)	53.408	1.503		35.528	.000	50.461	56.355
	NNO2	.459	.012	.523	37.599	.000	.435	.483
	N 降水量	-.063	.004	-.228	-16.396	.000	-.071	-.056
	N 湿度	-.373	.016	-.337	-23.767	.000	-.404	-.343
	NO3	-.192	.011	-.252	-17.896	.000	-.213	-.171
5	(常量)	54.015	1.473		36.673	.000	51.127	56.903
	NNO2	.415	.012	.473	33.485	.000	.390	.439
	N 降水量	-.054	.004	-.192	-13.854	.000	-.061	-.046
	N 湿度	-.454	.017	-.409	-27.484	.000	-.486	-.422
	NO3	-.222	.011	-.291	-20.651	.000	-.243	-.201
	NPM25	.112	.008	.191	13.412	.000	.096	.128

6	(常量)	42.683	1.661		25.691	.000	39.426	45.941
	NNO2	.423	.012	.482	34.856	.000	.399	.447
	N 降水量	-.047	.004	-.169	-12.373	.000	-.055	-.040
	N 湿度	-.393	.017	-.354	-23.462	.000	-.426	-.360
	NO3	-.157	.012	-.206	-13.621	.000	-.180	-.134
	NPM25	.490	.029	.836	17.034	.000	.434	.546
	NPM10	-.212	.015	-.697	-13.708	.000	-.242	-.181
7	(常量)	52.126	1.783		29.228	.000	48.629	55.622
	NNO2	.388	.012	.442	31.798	.000	.364	.412
	N 降水量	-.040	.004	-.144	-10.603	.000	-.047	-.033
	N 湿度	-.427	.017	-.385	-25.657	.000	-.459	-.394
	NO3	-.161	.011	-.211	-14.269	.000	-.184	-.139
	NPM25	.489	.028	.835	17.350	.000	.434	.545
	NPM10	-.212	.015	-.699	-14.010	.000	-.242	-.182
N 风速	-8.093	.622	-.165	-13.009	.000	-9.313	-6.873	
8	(常量)	57.117	1.856		30.770	.000	53.477	60.756
	NNO2	.399	.012	.455	32.832	.000	.375	.423
	N 降水量	-.046	.004	-.166	-12.140	.000	-.054	-.039
	N 湿度	-.457	.017	-.412	-27.146	.000	-.490	-.424
	NO3	-.107	.013	-.140	-8.383	.000	-.132	-.082
	NPM25	.586	.030	1.000	19.510	.000	.527	.645
	NPM10	-.279	.017	-.919	-16.586	.000	-.312	-.246
	N 风速	-8.492	.618	-.173	-13.737	.000	-9.704	-7.280
	N 温度	-.414	.047	-.146	-8.796	.000	-.506	-.322
9	(常量)	1633.359	85.119		19.189	.000	1466.480	1800.237
	NNO2	.379	.012	.432	32.272	.000	.356	.402
	N 降水量	-.033	.004	-.119	-8.864	.000	-.040	-.026
	N 湿度	-.657	.019	-.592	-33.765	.000	-.695	-.619
	NO3	-.082	.012	-.108	-6.655	.000	-.107	-.058
	NPM25	.455	.030	.776	15.293	.000	.396	.513
	NPM10	-.204	.017	-.671	-12.222	.000	-.236	-.171
	N 风速	-7.806	.596	-.159	-13.107	.000	-8.973	-6.638
	N 温度	-1.954	.095	-.691	-20.642	.000	-2.139	-1.768
	N 压强	-1.518	.082	-.554	-18.522	.000	-1.679	-1.357
10	(常量)	1596.883	86.175		18.531	.000	1427.935	1765.831
	NNO2	.385	.012	.439	32.124	.000	.362	.409
	N 降水量	-.033	.004	-.120	-8.948	.000	-.041	-.026
	N 湿度	-.645	.020	-.582	-32.427	.000	-.684	-.606
	NO3	-.074	.013	-.097	-5.780	.000	-.099	-.049
	NPM25	.435	.031	.742	14.201	.000	.375	.495

	NPM10	-.190	.018	-.624	-10.819	.000	-.224	-.155
	N 风速	-7.877	.596	-.161	-13.222	.000	-9.045	-6.709
	N 温度	-1.866	.100	-.660	-18.621	.000	-2.063	-1.670
	N 压强	-1.482	.083	-.541	-17.857	.000	-1.645	-1.320
	NCO	-4.835	1.833	-.043	-2.638	.008	-8.428	-1.242
11	(常量)	1589.198	86.230		18.430	.000	1420.142	1758.254
	NNO2	.384	.012	.438	31.992	.000	.361	.408
	N 降水量	-.033	.004	-.117	-8.724	.000	-.040	-.025
	N 湿度	-.642	.020	-.579	-32.206	.000	-.682	-.603
	NO3	-.073	.013	-.095	-5.705	.000	-.098	-.048
	NPM25	.430	.031	.733	13.979	.000	.369	.490
	NPM10	-.186	.018	-.614	-10.594	.000	-.221	-.152
	N 风速	-7.882	.596	-.161	-13.235	.000	-9.049	-6.714
	N 温度	-1.845	.101	-.652	-18.312	.000	-2.042	-1.647
	N 压强	-1.475	.083	-.538	-17.760	.000	-1.638	-1.312
	NCO	-5.794	1.894	-.051	-3.059	.002	-9.506	-2.081
	NSO2	.024	.012	.024	1.998	.046	.000	.049

a.因变量 NO2

$$NO_2 = b_0 + b_1 * NNO_2 + b_2 * N \text{ 降水量} + b_3 * N \text{ 湿度} + b_4 * NNO_3 + b_5 * NPM_{25} + b_6 * NPM_{10} + b_7 * N \text{ 风速} + b_8 * N \text{ 温度} + b_9 * N \text{ 压强} + b_{10} * NCO + b_{11} * NSO_2;$$

模型	非标准化系数		标准系数	t	Sig.	B 的 95.0% 置信区间	
	B	标准 误差	试用版			下限	上限
1 (常量)	-6.343	.809		-7.843	.000	-7.929	-4.757
NCO	46.981	1.247	.503	37.661	.000	44.535	49.427
2 (常量)	-577.346	29.753		-19.404	.000	-635.678	-519.014
NCO	49.683	1.204	.531	41.250	.000	47.322	52.045
N 压强	.559	.029	.247	19.198	.000	.502	.616
3 (常量)	-495.902	30.151		-16.447	.000	-555.015	-436.789
NCO	47.155	1.207	.504	39.084	.000	44.790	49.521
N 压强	.476	.030	.211	16.098	.000	.418	.534
N 降水量	.035	.003	.151	11.459	.000	.029	.040
4 (常量)	-373.141	32.524		-11.473	.000	-436.905	-309.377
NCO	41.988	1.312	.449	31.993	.000	39.415	44.561
N 压强	.354	.032	.157	11.074	.000	.291	.417
N 降水量	.038	.003	.165	12.583	.000	.032	.044
NPM10	.034	.004	.137	9.483	.000	.027	.041

5	(常量)	-242.622	35.787		-6.780	.000	-312.784	-172.460
	NCO	37.264	1.417	.399	26.291	.000	34.485	40.043
	N 压强	.230	.035	.102	6.596	.000	.162	.299
	N 降水量	.035	.003	.151	11.520	.000	.029	.041
	NPM10	.151	.014	.605	10.547	.000	.123	.179
	NPM25	-.217	.026	-.449	-8.420	.000	-.267	-.166
6	(常量)	-261.201	35.722		-7.312	.000	-331.235	-191.166
	NCO	36.958	1.411	.395	26.191	.000	34.192	39.725
	N 压强	.251	.035	.111	7.203	.000	.183	.320
	N 降水量	.036	.003	.155	11.879	.000	.030	.041
	NPM10	.146	.014	.585	10.226	.000	.118	.174
	NPM25	-.214	.026	-.444	-8.361	.000	-.265	-.164
	N 风速	-3.327	.508	-.082	-6.554	.000	-4.323	-2.332
7	(常量)	-310.485	36.452		-8.518	.000	-381.951	-239.019
	NCO	33.694	1.501	.360	22.443	.000	30.751	36.638
	N 压强	.298	.036	.132	8.383	.000	.228	.368
	N 降水量	.033	.003	.142	10.802	.000	.027	.039
	NPM10	.121	.015	.484	8.171	.000	.092	.150
	NPM25	-.168	.027	-.348	-6.317	.000	-.220	-.116
	N 风速	-3.784	.511	-.094	-7.407	.000	-4.785	-2.782
	NO3	.059	.010	.094	6.166	.000	.040	.078
8	(常量)	-330.763	36.459		-9.072	.000	-402.242	-259.283
	NCO	31.952	1.523	.342	20.977	.000	28.966	34.938
	N 压强	.316	.036	.140	8.888	.000	.246	.386
	N 降水量	.026	.003	.111	7.915	.000	.019	.032
	NPM10	.116	.015	.463	7.832	.000	.087	.145
	NPM25	-.169	.026	-.350	-6.381	.000	-.221	-.117
	N 风速	-3.156	.519	-.078	-6.075	.000	-4.174	-2.137
	NO3	.072	.010	.115	7.336	.000	.053	.091
	NNO2	.063	.011	.087	5.989	.000	.042	.084
9	(常量)	-351.543	36.516		-9.627	.000	-423.134	-279.952
	NCO	33.982	1.560	.364	21.780	.000	30.923	37.041
	N 压强	.336	.036	.149	9.447	.000	.267	.406
	N 降水量	.024	.003	.103	7.303	.000	.017	.030
	NPM10	.108	.015	.433	7.336	.000	.079	.137
	NPM25	-.156	.026	-.324	-5.901	.000	-.208	-.104
	N 风速	-3.149	.518	-.078	-6.085	.000	-4.164	-2.134
	NO3	.069	.010	.109	7.003	.000	.049	.088
	NNO2	.067	.010	.092	6.342	.000	.046	.087
	NSO2	-.060	.011	-.071	-5.609	.000	-.081	-.039

10	(常量)	-363.904	36.933		-9.853	.000	-436.313	-291.496
	NCO	33.890	1.560	.363	21.723	.000	30.832	36.949
	N 压强	.346	.036	.153	9.650	.000	.276	.416
	N 降水量	.022	.003	.098	6.865	.000	.016	.029
	NPM10	.099	.015	.397	6.471	.000	.069	.129
	NPM25	-.146	.027	-.302	-5.406	.000	-.198	-.093
	N 风速	-2.990	.522	-.074	-5.724	.000	-4.014	-1.966
	NO3	.079	.011	.126	7.242	.000	.058	.101
	NNO2	.066	.010	.091	6.250	.000	.045	.086
	NSO2	-.060	.011	-.071	-5.584	.000	-.081	-.039
	N 湿度	.031	.014	.034	2.191	.029	.003	.058

a.因变量 SO2

$$SO_2 = b_0 + b_1 * NCO + b_2 * N \text{ 压强} + b_3 * N \text{ 降水量} + b_4 * NPM10 + b_5 * NPM25 + b_6 * N \text{ 风速} + b_7 * NNO_3 + b_8 * NNO_2 + b_9 * NSO_2 + b_{10} * N \text{ 湿度};$$

模型		非标准化系数		标准系数	t	Sig.	B 的 95.0% 置信区间	
		B	标准 误差	试用版			下限	上限
1	(常量)	10.038	.929		10.802	.000	8.216	11.860
	N 温度	3.776	.064	.676	59.458	.000	3.652	3.901
2	(常量)	54.386	1.321		41.175	.000	51.797	56.976
	N 温度	3.440	.054	.616	63.671	.000	3.334	3.546
	NNO2	-.698	.017	-.403	-41.653	.000	-.731	-.665
3	(常量)	99.416	2.081		47.768	.000	95.336	103.497
	N 温度	2.722	.057	.487	47.980	.000	2.611	2.833
	NNO2	-.602	.016	-.347	-37.829	.000	-.633	-.571
	N 湿度	-.610	.023	-.279	-26.720	.000	-.655	-.566
4	(常量)	81.688	2.287		35.726	.000	77.205	86.170
	N 温度	2.515	.056	.450	44.579	.000	2.404	2.625
	NNO2	-.626	.015	-.361	-40.402	.000	-.656	-.596
	N 湿度	-.529	.023	-.241	-23.317	.000	-.573	-.484
	NO3	.233	.014	.154	16.469	.000	.205	.260
5	(常量)	67.357	2.494		27.006	.000	62.467	72.247
	N 温度	2.574	.055	.461	46.401	.000	2.466	2.683
	NNO2	-.585	.016	-.338	-37.748	.000	-.616	-.555
	N 湿度	-.475	.023	-.217	-20.992	.000	-.519	-.430
	NO3	.229	.014	.152	16.528	.000	.202	.256
	N 风速	11.189	.854	.116	13.099	.000	9.514	12.864
6	(常量)	68.288	2.467		27.681	.000	63.451	73.124
	N 温度	2.362	.059	.423	40.167	.000	2.247	2.477

	NNO2	-551	.016	-318	-35.096	.000	-.582	-.520
	N 湿度	-.416	.023	-.190	-18.012	.000	-.462	-.371
	NO3	.284	.015	.188	19.242	.000	.255	.313
	N 风速	10.592	.846	.110	12.514	.000	8.933	12.251
	NPM10	-.062	.006	-.104	-9.997	.000	-.075	-.050
7	(常量)	54.448	2.429		22.413	.000	49.686	59.211
	N 温度	1.856	.061	.332	30.608	.000	1.737	1.974
	NNO2	-.536	.015	-.309	-35.904	.000	-.565	-.506
	N 湿度	-.382	.022	-.174	-17.376	.000	-.425	-.339
	NO3	.456	.016	.302	28.252	.000	.424	.487
	N 风速	10.349	.804	.107	12.878	.000	8.774	11.925
	NPM10	-.518	.022	-.864	-23.489	.000	-.561	-.475
	NPM25	.841	.039	.727	21.448	.000	.764	.918
8	(常量)	-1145.024	112.442		-10.183	.000	-1365.469	-924.579
	N 温度	3.050	.127	.546	24.031	.000	2.801	3.299
	NNO2	-.532	.015	-.307	-36.087	.000	-.560	-.503
	N 湿度	-.234	.026	-.107	-9.106	.000	-.285	-.184
	NO3	.428	.016	.284	26.510	.000	.396	.459
	N 风速	9.619	.796	.100	12.084	.000	8.058	11.180
	NPM10	-.576	.022	-.961	-25.684	.000	-.620	-.532
	NPM25	.948	.040	.819	23.714	.000	.869	1.026
	N 压强	1.155	.108	.213	10.670	.000	.943	1.367
9	(常量)	-1091.908	114.446		-9.541	.000	-1316.282	-867.533
	N 温度	3.025	.127	.542	23.769	.000	2.775	3.274
	NNO2	-.545	.016	-.315	-34.563	.000	-.576	-.515
	N 湿度	-.246	.026	-.112	-9.402	.000	-.297	-.195
	NO3	.418	.017	.277	25.130	.000	.385	.450
	N 风速	9.396	.801	.097	11.734	.000	7.826	10.965
	NPM10	-.575	.022	-.959	-25.631	.000	-.619	-.531
	NPM25	.951	.040	.823	23.805	.000	.873	1.030
	N 压强	1.104	.110	.204	10.013	.000	.887	1.320
	N 降水量	.012	.005	.022	2.451	.014	.002	.022

$O_3 = b_0 + b_1 * N \text{ 温度} + b_2 * NNO_2 + b_3 * N \text{ 湿度} + b_4 * NNO_3 + b_5 * N \text{ 风速} + b_6 * NPM_{10} + b_7 * NPM_{25} + b_8 * N \text{ 压强}$   
 $+ b_9 * N \text{ 降水量};$