



東華理工大學

EAST CHINA UNIVERSITY OF TECHNOLOGY

数学建模作业论文

姓名 1: 黄泽利 学号: 201520070315

专业: 计算机科学与技术

姓名 2: 万超静 学号: 201520320108

专业: 信息与计算科学

姓名 3: 胡秋秋 学号: 201520320125

专业: 信息与计算科学

2017 年 7 月 6 日

中国主要城市PM2.5数据分析与预测

摘要

近年来，空气污染问题日益凸显，全国各大城市PM2.5污染问题不容乐观，已经严重的影响了人的正常生活和身体健康。当PM2.5浓度较高时，会降低大气能见度，易造成交通事故可以作为水汽凝结核，可与空气中的二氧化硫结合，形成酸雾，腐蚀建筑物也可以直接或者间接吸收反射太阳辐射，影响气候变化及其附着的污染物还可深达肺泡并沉积，进而进入血液循环，严重影响人类健康。包括无机离子、重金属和有机污染物等。因此掌握PM2.5污染的变化特征对控制其浓度显得极为重要。

针对问题一，运用时间序列模型通过2014-2016年PM2.5的浓度变化从而对2017年至以后的PM2.5的浓度进行预测。采用指数平滑法通过对数据的分析，从不同的方面研究PM2.5的变化特征，首先以月为单位，分析2014-2016年每个月PM2.5的变化趋势；将一年划分为四个季度，分析每个季度PM2.5的浓度变化趋势，为了了解全国PM2.5的浓度有无缓解，分析各年的PM2.5总体浓度的变化趋势，从而对2017年的PM2.5的总体浓度进行预测；由于经济发展的程度以及地理位置的不同，每个城市PM2.5的浓度也会有所差异，通过分析可得出近三年来各城市PM2.5的浓度变化。

针对问题二，研究全国城市PM2.5污染的分区结构特征，根据2014-2016年的数据计算各城市的平均浓度，绘制浓度分布图，结合浓度图的特点，并对城市进行K均值聚类分析，将聚类之后的各结果作为城市的分区，针对具体的分区，进行多方面的研究，得到分区的结构特征。

针对问题三，研究城市之间PM2.5污染的因果关系，对于城市之间的因果关系，建立格兰杰因果关系模型，利用Eviews软件进行求解，根据显著性水平判断各城市之间的因果关系，对因果关系进行分析，并结合问题二中的结论，提出个性化的PM2.5污染治理的综合方案。

关键词：时间序列 聚类分析 格兰杰因果关系

一、问题重述

2013 年，“雾霾”成为年度关键词，雾霾笼罩着全国各地，影响了全国人民的健康，雾霾天气是一种大气污染状态，是对大气中各种悬浮颗粒物含量超标的笼统表述。 $PM_{2.5}$ 被认为是造成雾霾天气的元凶之一。 $PM_{2.5}$ 是指大气中直径小于或等于 2.5 的微米的颗粒物，属于细颗粒物。2013 年之前，我国在大气污染颗粒物方面的检测对象主要是 PM_{10} ，属于可吸入颗粒物。2013 年 1 月 1 日，我国环保部开始正式将 $PM_{2.5}$ 列入空气检测指标中， $PM_{2.5}$ 成为了一个重要的检测空气污染程度的指数。 $PM_{2.5}$ 来源主要是自然和人为两种。自然过程如尘土、森林火灾、细菌等均可产生 $PM_{2.5}$ 。而其绝大部分来自人类的生产生活过程，化石燃料的燃烧、机动车尾气的排放、生物质燃烧等均会产生 $PM_{2.5}$ 。 $PM_{2.5}$ 不仅会影响环境，造成大气污染，更重要的是长期吸入 $PM_{2.5}$ 会对人体造成极大的危害，甚至会增加致病和死亡的概率。附录中数据集包括除香港、澳门、台湾之外的 22 个省会城市、4 个直辖市、5 个自治区首府，共 31 个城市的 $PM_{2.5}$ 浓度数据。对这些数据进行分析处理，挖掘数据背后隐含规律，对我国关于 $PM_{2.5}$ 的研究、预防与治理有着重要的意义。

研究的主要问题

试基于 31 个省会城市 2014 年至 2016 年数据分析 $PM_{2.5}$ 的分布。你们小组完成以下几个任务：

1. 对数据进行动态分析及预测，揭示 $PM_{2.5}$ 污染的变化特征。
2. 分析全国城市 $PM_{2.5}$ 污染的分区结构特征。
3. 确定城市之间 $PM_{2.5}$ 污染的因果关系并给出合理的 $PM_{2.5}$ 污染治理的具体方案。

二、问题的分析

自雾霾进入人们的视野以来，就一直备受人们的关注，雾霾对人们的生活和健康都要造成了严重的影响，而造成雾霾主要元凶之一的 $PM_{2.5}$ ，更是受到了各相关部门的高度重视。雾霾和 $PM_{2.5}$ 的研究对空气污染的治理具有重要的意义。

针对问题一，采用动态分析对 2014-2016 年 $PM_{2.5}$ 的浓度变化进行分析，时

间序列法对 2017 年至以后的 PM2.5 的浓度进行预测。通过对数据的分析，从不同的方面研究 PM2.5 的变化特征，首先以月为单位，分析 2014-2016 年每个月 PM2.5 的变化趋势；将一年划分为四个季度，分析每个季度 PM2.5 的浓度变化趋势，为了了解全国 PM2.5 的浓度有无缓解，分析各年的 PM2.5 总体浓度的变化趋势，从而对 2017 年的 PM2.5 的总体浓度进行预测；由于经济发展的程度以及地理位置的不同，每个城市 PM2.5 的浓度也会有所差异，通过分析可得出近三年来各城市 PM2.5 的浓度变化。

针对问题二，研究全国城市 PM2.5 污染的分区结构特征，根据 2014 年 PM2.5 浓度，运用 SPSS 中的 K 均值聚类分析方法，将各城市进行分区，然后根据分区的城市来研究 PM2.5 的特征。

针对问题三，运用 EViews 软件中的格兰杰因果检验法，确定各城市之间 PM2.5 污染的因果关系，通过问题一、问题二的结论给出合理的 PM2.5 污染治理的方案。

三、模型假设

通过对本问题的分析，提出以下四点假设：

- ① 假设题中所给的数据准确无误
- ② 假设所给数据中缺失的数据和均值相等
- ③ 不考虑香港、澳门和台湾的 PM2.5 的变化特征结构特征及因果关系

四、符号说明

各具体符号说明见表 1：

表 1 符号说明

符号	含义
S_t	表示时间的平滑值
a	平滑常数
y_t	时间的 t 实际值
S_{t-1}	时间的 t-1 实际值
sig	概率 p 值

五、模型的建立与求解

5. 1问题一的求解

根据所给数据，对 2014-2015 年的 31 个城市 PM.5 的浓度进行初步分布分析，观察全国的各月 PM2.5 的平均浓度，建立时间序列模型进行分析。

5. 1. 1PM2. 5的年变化趋势分析

通过对 2014-2016 年的各城市的数据进行统计，得到 2014-2016 年的总月平均 PM2.5 的浓度值，如下表：

表 2 2014-2016 年月平均 PM2.5 浓度表

年份\月	1	2	3	4	5	6	7	8	9	10	11	12
2014	108.33	82.62	65.79	57.02	51.38	48.04	43.91	39.91	41.62	74.63	73.20	73.95
2015	88.23	73.95	55.44	50.66	44.18	37.43	38.56	36.46	36.27	53.35	68.21	91.19
2016	76.63	62.74	62.87	45.21	39.90	33.28	32.66	30.21	40.55	42.16	69.82	93.96

通过表 2，绘制 2014-2016 年的变化趋势折线图，得到三年变化趋势，如下：

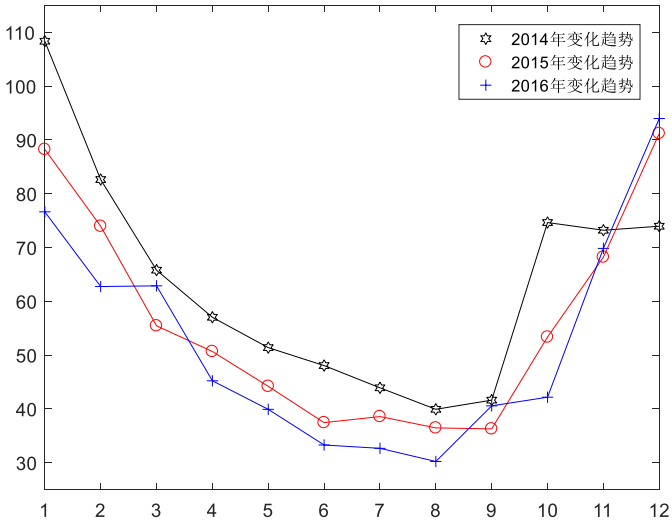


图 1 2014-2016 年月平均 pm2.5 浓度折线图

据图可知，PM2.5 的年变化呈现先下降，后上升的趋势，11 月至 1 月之间达到峰值，6-9 月的值较低，可知 PM2.5 的浓度变化与月份之间存在较大的相关性，故推测 PM2.5 均浓度的变化具有季节性，呈现 U 型特征。

5.1.2 PM2.5 的季节性变化趋势

根据表 2 中的数据，按季节进行统计，得到各年的季节平均数据，

表 3 各年季度均值

年份\季节	1	2	3	4
2014	85.58	52.15	41.81	73.92
2015	72.54	44.209	37.1	70.92
2016	67.41	39.46	34.47	68.65

根据表 3 的数据，分析季节之间的变化特征，为了反映出季节的曲线变化趋

势，因此采用三分端样条插值法，根据表中的数据进行插值，并运用插值后的数据进行三分端样条拟合，得到三年的季节变化曲线，如下图：

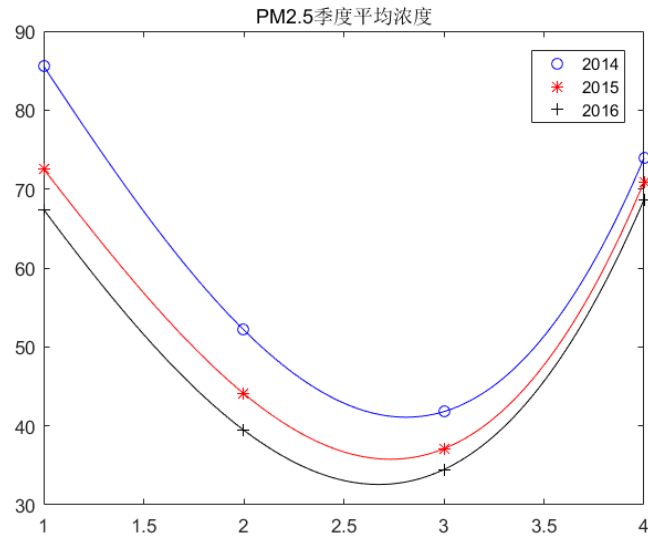


图 2 PM2.5 季度变化曲线

通过 PM2.5 季度变化曲线，可以看出 2014-2016 年 PM2.5 的浓度逐年下降，PM2.5 的浓度会随着季度的变化发生改变，呈现先减后增的趋势，在第 1、4 季度的浓度很高。

5.1.3 年平均PM2.5的浓度变化趋势

PM2.5 的危害严重威胁到了人们的健康，为了减轻这种伤害，国家采取了一些措施，例如实施可持续发展战略，坚持保护环境和节约资源等，为了了解这些方法是否有效，统计 2014-2016 年 PM2.5 的浓度数据，通过数据画出柱形及折线图，如下：

表 4 年平均 PM2.5 的浓度

年份	2014	2015	2016
PM2.5 浓度	63.67	56.16	52.5

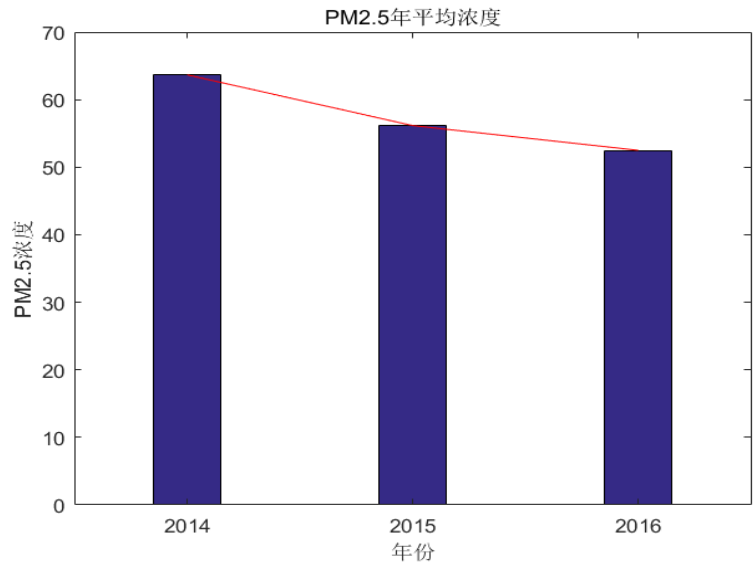


图 3 各年平均浓度

通过对 2014-2016 年 PM2.5 的年平均浓度分析, 可以发现全国 PM2.5 年平均浓度逐年下降, 2015 年 PM2.5 的浓度下降速度比 2016 年的快。通过各方面的了解, 发现在 2015 年环保部开始调控城市机动车出行量以及投资 3500 亿元治理空气, 这些措施使得 2015 年 PM2.5 的浓度大幅度减低。

5.1.4 各城市年度PM2.5浓度变化趋势

由于各城市的经济发展程度以及地理位置的不同, PM2.5 的浓度也存在差异; 通过数据整理, 画出 31 个城市(名称代号)在 2014-2016 年 PM2.5 的浓度的柱形图:

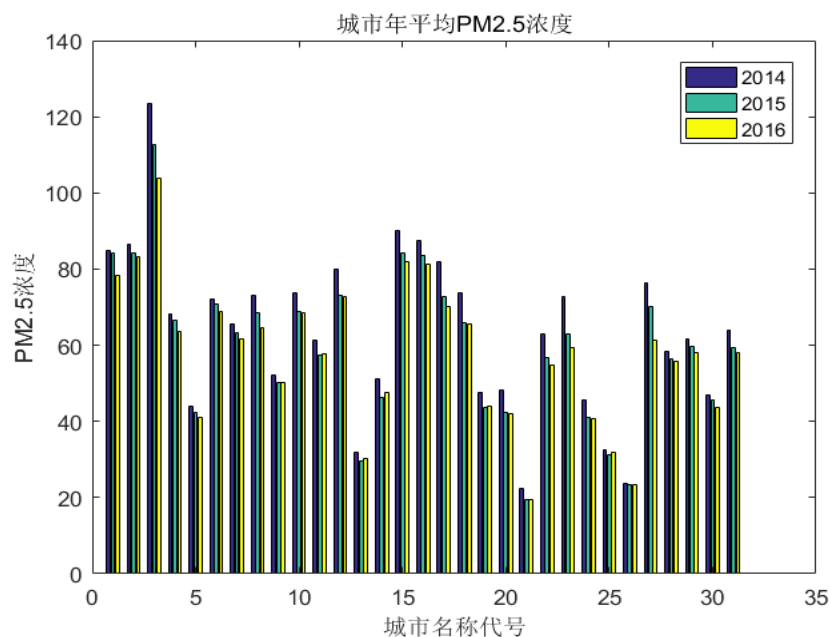


图 4 各城市年平均 PM2.5 浓度

由上图可以看出, 石家庄的 PM2.5 的浓度位于榜首, 海口的 PM2.5 的浓度最低; 在 2014-2016 年, 除福州、南昌、昆明外, 其他城市的 PM2.5 的浓度逐年降低; 拉萨的 PM2.5 的浓度保持稳定水平。

5.1.5 PM2.5的预测

5.1.5.1 预测方法

指数平滑法是一种常用的时间序列分析方法, 指数平滑法的估计是非线性的, 其目标是使预测值和实测值间的均方差 (MSE) 最小。指数平滑模型是在移动平均模型基础上发展起来的一种时间序列分析预测法, 其原理是任一期的指数平滑值都是本期实际观察值与前一期指数平滑值的加权平均。指数平滑模型的思想是对过去值和当前值进行加权平均、以及对当前的权数进行调整以前抵消统计数值的摇摆影响, 得到平滑的时间序列。指数平滑法不舍弃过去的的数据, 但是对过去的的数据给予逐渐减弱的影响程度 (权重)。

5.1.5.2 模型类型

为了消除月度季节的影响, 这里采用季节性的 Winters 可加性, 该模型适用于具有线性趋势且不依赖于序列水平的季节性效应的序列, 其平滑参数是水平、趋势和季节。通过相应的拟合图形可以得出因变量无需转换。从而选择指数平滑法中的季节性的 Winters 可加性。

5.1.5.3 模型的建立

指数平滑法的基本公式:

$$S_t = ay_t + (1-a)S_{t-1}$$

S_t 表示时间的平滑值, y_t 表示时间 t 的实际值, a 表示平滑常数, 取值范围为 $[0,1]$, S_{t-1} 表示时间 $t-1$ 的实际值。

由该公式可知:

- 1. S_t 是 y_{t-1} 和 S_{t-1} 的加权算数平均数, 随着 a 取值的大小变化, 决定 y_{t-1} 和 S_{t-1} 对 S_t 的影响程度, 当 a 取 1 时, $S_t=y_{t-1}$; 当 a 取 0 时, $S_t=S_{t-1}$ 。
- 2. S_t 具有逐期追溯性质, 可探源至 S_{t-t+1} 为止, 包括全部数据。其过程中, 平滑常数以指数形式递减, 故称之为指数平滑法。指数平滑常数取值至关重要。平滑常数决定了平滑水平以及对预测值与实际结果之间差异的响应速度。平滑常数 a 越接近于 1, 远期实际值对本期平滑值的下降越迅速; 平滑常数 a 越接近于 0, 远期实际值对本期平滑值影响程度的下降越缓慢。由此, 当时间数列相对平稳时, 可取较大的 a ; 当时间数列波动较大时, 应取较小的 a , 以不忽略远期实际值的影响。
- 3. 尽管 S_t 包含有全期数据的影响, 但实际计算时, 仅需要两个数值, 即 y_{t-1} 和 S_{t-1} , 再加上一个常数 a , 这就使指数滑动平均具有逐期递推性质, 从而给预测带来了极大的方便。

5.1.5.4 模型的求解

根据 SPSS 中的分析预测创建时间序列模型, 应用指数平滑法得到 2017 年 PM2.5 的浓度, 如下表所示:

表 5 预测结果

模型		一月 2017	二月 2017	三月 2017	四月 2017	五月 2017
pm2.5-模型_1	预测	78.66	25.08	28.94	74.45	34.31
	UCL	91.10	37.89	41.80	87.48	46.96
	LCL	66.23	12.26	16.07	61.43	21.66

对于每个模型，预测都在请求的预测时间段范围内的最后一个非缺失值之后开始，在所有预测值的非缺失值都可用的最后一个时间段或请求预测时间段的结束日期（以较早者为准）结束。

a. 根据平稳 R 方的最佳拟合模型（值越大表示拟合越好）。

具体图形如下：

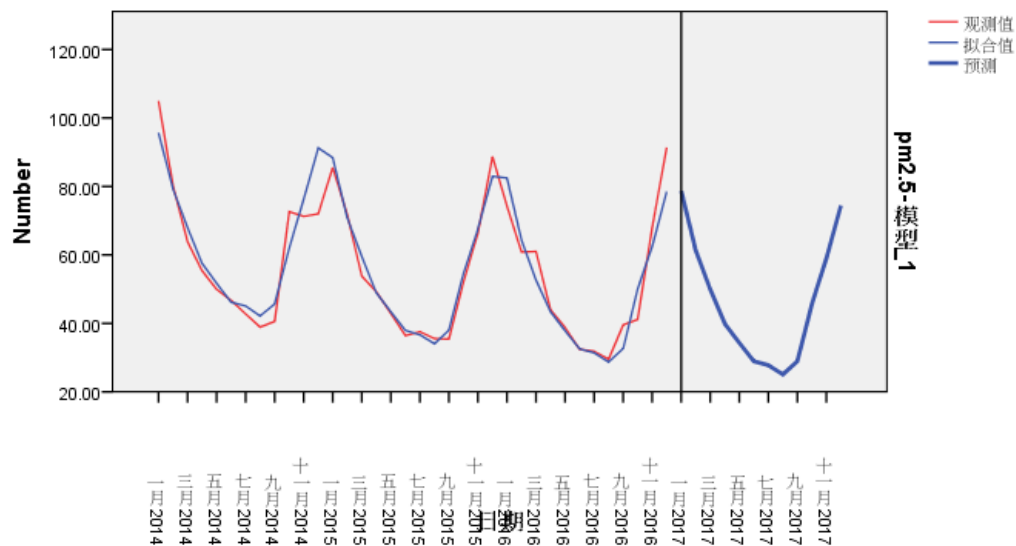


图 5 2017 年的预测结果图

5.1.5.5 模型的分析

(1) 从模型统计量中表格中可以得到以下信息：

本实验拟合的指数平滑模型的水平，即平稳的 R 方为 0.804，Ljung-BoxQ 统计量值为 21.684，显著水平为 0.116，因此接受残差序列为独立序列的原假设，故采用平滑指数法。

表 6 模型统计量

模型统计量 a					
模型	模型拟合统计量	Ljung-Box Q(18)			离群值数
	平稳的 R 方	RMSE	统计量	Sig.	
pm2.5	0.804	6.113	21.684	0.116	0

(2) 从指数平滑法模型参数表格中可以得到以下信息：

本实验拟合的指数平滑模型的水平 Alpha 值为 0.094，P 值为 0.12，不仅作用不大而且不显著。而季节 Delta 值为 0，该值特别小而且没有显著性，因此该序列几乎没有任何季节性特征。

表 7 指数平滑法模型参数

模型		估计	SE	t	Sig.
pm2.5-模型无转换_1	Alpha (水平)	.094	.059	1.594	.120
	Gamma (趋势)	1.033E-006	.004	.000	1.000
	Delta (季节)	.000	.076	.002	.998

a. 根据平稳 R 方的最佳拟合模型（值越大表示拟合越好）。

5.2问题二的求解

5.2.1 城市聚类

5.2.1.1 k-均值聚类原理

K 均值聚类是先随机选取 K 个对象作为初始的聚类中心。然后计算每个对象与各个种子聚类中心之间的距离，把每个对象分配给距离它最近的聚类中心。聚类中心以及分配给它们的对象就代表一个聚类。一旦全部对象都被分配了，每个聚类的聚类中心会根据聚类中现有的对象被重新计算。这个过程将不断重复直到满足某个终止条件。终止条件可以是没有（或最小数目）对象被重新分配给不同的聚类，没有（或最小数目）聚类中心再发生变化，误差平方和局部最小。

本文是用两者的欧几里得距离来作为相异度，欧几里得距离的定义如下：

$$d(X,Y)=\sqrt{(x_1-y_1)^2+(x_2-y_2)^2+\dots+(x_n-y_n)^2}$$

5.2.1.2 k-均值聚类结果

通过 SPSS 的相关操作，得出如下表格：

表 8 聚类结果表

第一类	福州 海口 昆明 拉萨
第二类	北京 天津 合肥 济南 郑州 武汉
第三类	太原 沈阳 长春 哈尔滨 南京 杭州 长沙 重庆 成都 西安 西宁 乌鲁木齐
第四类	呼和浩特 上海 南昌 广州 南宁 贵阳 兰州 银川
第五类	石家庄

我们将这 31 个地区分成五类，按照 PM2.5 的浓度聚类中心逐次递增，第一类有 4 个地区；其次有 6 个地区；中间类别有最多的地区，有 12 个；第四类地区有 8 个；最后一类仅有一个被聚为一类，那就是石家庄。

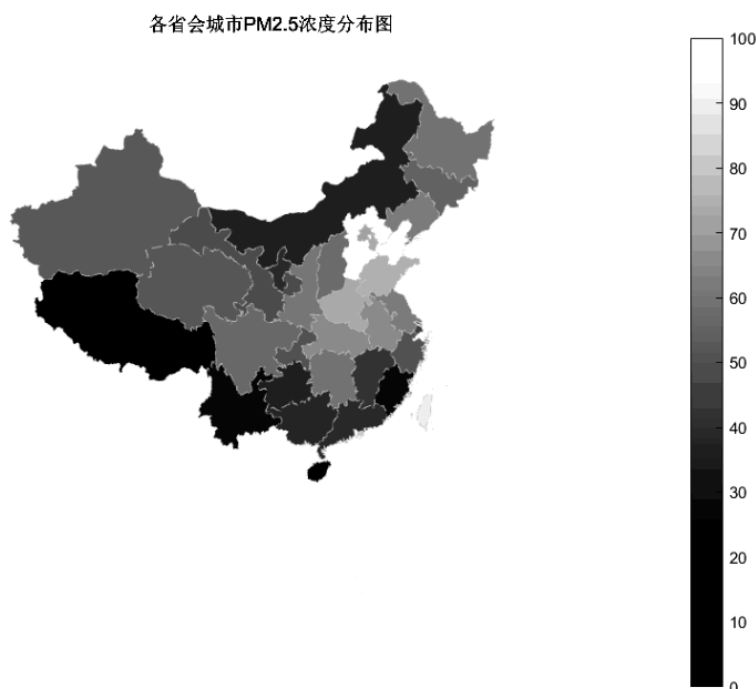


图 6 PM2.5 的浓度分布图

5.2.2 城市结构特征

由于第三类地区分布较散乱且数量较多，现对第三类各大地区逐次分析，而其他地区分布较为集中，直接用概括性的语言描述。

第一类有四大地区，分别是海口、福州、昆明和拉萨。对于海口和福州，两者都处于沿海城市，而沿海城市气候扩散条件好，空气中的 PM2.5 再海风的作用不易聚集，因而两者 pm.5 都较低。而昆明与拉萨正好都处于内陆地区，工业覆盖面积对于绿化面积较少。而昆明的少量 PM2.5 主要来源为建筑施工扬尘和汽车尾气。

第二类有八大地区，分别是呼和浩特、上海、南昌、广州、南宁、贵阳、兰州和银川。其中广州和上海主要为第三产业，且两者都作为大都市，对于环境的治理力度大，对于各项指标的要求都较为严格，因而两者的 pm2.5 的浓度含量较低。而其他的六个城市中，主要为二线城市，重工业的分布不较少，且不发达，周边的总工业、基地较少，不容易受到周边城市的影响，因而 PM.5 的含量较少。

第三类有十二大地区，分别是太原、沈阳、长春、哈尔滨、南京、杭州、长沙、重庆、成都、西安、西宁、乌鲁木齐。其中，我国重要的煤炭重工业、综合性工业、中国最大的汽车研发生产基地都在这些地区，由于自然环境以及特殊的工业结构、燃料结构等多种因素制约，使得这些地区 PM2.5 的浓度较高。由问题

(1) 可知 PM2.5 的浓度具有季节性，现已南京市作为研究，南京秋季污染水平低于冬春两季的结果是一致的，由于冬春季节风速大于夏秋季节，导致地面扬尘增加，细颗粒物浓度升高。秋季秋高气爽，大气扩散条件好，初秋繁茂的植被对颗

颗粒物有良好的吸纳作用。而且冬春季节温度较低,大气层结较稳定,易出现辐射逆温现象,这些都不利于 PM2.5 迅速的扩散迁移,使其浓度发生积累而增大。这也就导致了南京市整体的 PM2.5 处于中等水平。

第四类有六大地区,分别是北京、天津、合肥、济南、郑州、武汉。它们主要位于西北地区,四季分明,光照充足,冬季寒冷干燥且较长。冬季狂风四起,造成尘土飞扬,因而此地区 PM2.5 的浓度位居第二。此类地区经济发展较为迅速,城市中的建筑物很高,阻挡和摩擦作用使风流经城区时明显减弱。静风现象增多,不利于大气中悬浮微粒的扩散稀释,容易在城区和近郊区周边积累。垂直方向上出现逆温。逆温层好比一个锅盖覆盖在城市上空,这种高空的气温比低空气温更高的逆温现象。空气中悬浮颗粒物的增加。随着城市人口的增长和工业发展直接导致了能见度降低。

第五类只有石家庄,由于 PM2.5 浓度远高于其他地区,故自成一类。石家庄地区西高东低,呈‘避风港’式地形,PM2.5 没有扩散的条件。然而,比自然因素更重要的,还是巨大的污染排放量。此地区产业结构偏重,钢铁、建材、石化、电力等“两高”行业集中,其中,钢铁粗钢产量超全国总量的 1/4;能源结构不尽合理,能源消费居全国第二位,单位 GDP 能耗比全国水平高近 60%。石家庄有 23 家热电联产企业,加上 7 座冬季供热站年耗煤高达 2390 万吨,超过了北京全年的煤耗总量。这样的产业与能源结构,给了环境巨大的压力。其中,氮氧化物、烟(粉)尘排放量居全国第一位,因此它的污染程度最大。在石家庄大片裸露的土壤,重型卡车经过,一片尘土飞扬,让人不得不掩面而逃,而旁侧的建筑垃圾山上已经长出野草。

总结:通过对选取的地区进行聚类分析可知,PM2.5 浓度除了与工业的发展等有关,还与一些气候有关,比如温度、湿度、风速、季节、降水,温度对 PM2.5 值的影响比较复杂,受到各种人为干扰因素,但总体还是表现为温度相对较低的情况下,PM2.5 值相对较大。还与一些地理环境有关,如沿海城市 PM2.5 浓度可能相对于其他非沿海就略低一些。PM2.5 的浓度同时受环境因素与人为因素有关,在环境因素无法改变的前提下,我们应最大化的将人为因素对 PM2.5 浓度的影响降到最低。

5.3 问题三的求解

5.3.1 格兰杰因果检验法的原理

格兰杰因果关系检验假设了有关 y 和 x 每一变量的预测的信息全部包含在这些变量的时间序列之中。检验要求估计以下的回归:

$$(1) \quad y_t = \sum_{i=1}^q a_i x_{t-i} + \sum_{j=1}^q \beta_j y_{t-j} + u_{1t}$$

$$(2) \quad x_t = \sum_{i=1}^q \lambda_i x_{t-i} + \sum_{j=1}^q \delta_j y_{t-j} + u_{2t}$$

其中白噪音 u_{1t} 和 u_{2t} 假定为不相关的。

式 (1) 假定当前 y 与 y 自身以及 x 的过去值有关，而式 (2) 对 x 也假定了类似的行为。

对式 (1) 而言，其零假设 $H_0: a_1 = a_2 = \cdots = a_q = 0$

对式 (2) 而言，其零假设 $H_0: \delta_1 = \delta_2 = \cdots = \delta_q = 0$

分四种情形讨论：

(1) x 是引起 y 变化的原因，即存在由 x 到 y 的单向因果关系。若式 (1) 中滞后的 x 的系数估计值在统计上整体的显著不为零，同时式 (2) 中滞后的 y 的系数估计值在统计上整体的显著为零，则称 x 是引起 y 变化的原因。

(2) y 是引起 x 变化的原因，即存在由 y 到 x 的单向因果关系。若式 (2) 中滞后的 y 的系数估计值在统计上整体的显著不为零，同时式 (1) 中滞后的 x 的系数估计值在统计上整体的显著为零，则称 y 是引起 x 变化的原因。

(3) x 和 y 互为因果关系，即存在由 x 到 y 的单向因果关系，同时也存在由 y 到 x 的单向因果关系。若式 (1) 中滞后的 x 的系数估计值在统计上整体的显著不为零，同时式 (2) 中滞后的 y 的系数估计值在统计上整体的显著不为零，则称 x 和 y 间存在反馈关系，或者双向因果关系。

(4) x 和 y 是独立的，或 x 与 y 间不存在因果关系。若式 (1) 中滞后的 x 的系数估计值在统计上整体的显著为零，同时式 (2) 中滞后的 y 的系数估计值在统计上整体的显著为零，则称 x 和 y 间不存在因果关系。

格兰杰因果关系检验的结论只是统计意义上的因果性，而不一定是真正的因果关系。虽然可以作为真正的因果关系的一种支持，但不能作为肯定或否定因果关系的最终根据。当然，即使格兰杰因果关系不等于实际因果关系，也并不妨碍其参考价值。

5.3.2 各城市 PM2.5 浓度的因果关系

通过格兰杰因果检验得出如下因果表格，其中数字 1-31 分别代表 31 个城市：北京、天津、石家庄、太原、呼和浩特、沈阳、长春、哈尔滨、上海、南京、杭州、合肥、福州、南昌、济南、郑州、武汉、长沙、广州、南宁、海口、重庆、成都、贵阳、昆明、拉萨、西安、兰州、西宁、银川和乌鲁木齐。既有单向因果，也有双向因果。若存在城市无因果关系，则未出现在表格上。

表 9 部分城市因果关系表

因	果
1	2 3 9 10 11 16 23 27 28 30 31
2	10 11 31
3	5 8 11 27 30 31
4	1 2 3 5 8 9 10 11 12 15 16 17 22 23 26 27 28 29 31
5	3 9 10 11 12 16 17 22 23 25 27 31
6	2 3 4 9 10 11 12 15 16 17 22 23 27 28 31
7	2 3 5 9 10 11 12 15 16 22 23 26 28 30 31

表格中的因列数字为城市的序号，全部城市详见附录

5.3.3 PM2.5 污染治理的方案

随着我国经济社会的快速发展，全国各大城市和地区雾霾天气频发，而 PM2.5 作为影响雾霾天气的主要因素，为了减轻 PM2.5 对人呼吸系统和心血管系统造成的伤害，减少 PM2.5 的浓度成为一项不容忽视的任务。

对于 PM2.5 污染治理，我们总结了一些方法：

(1) 整体性治理

不同的行政区有不同的执法部门，对于跨地区 PM2.5 污染问题，若职能有交叉各执法部门相互推脱；各地区只会针对本地区进行管理，缺乏整体性的治理，因而 PM2.5 的治理没有明显的改善。整体性治理是以协调、整合、责任为治理机制，对治理层级、功能、部门关系及信息系统等碎片化问题进行有机的协调与整合，从分散走向集中、从部分走向整体、从碎片走向整合的。空气具有全球性，因而 PM2.5 也具有全球性。由各城市之间的 PM2.5 的污染因果关系可知，整体性治理的方案能够促进城市之间的 PM2.5 的治理。

(2) 倡导公民绿色出行，控制汽车尾气排放

PM2.5 的污染有很大一部分是来自于汽车尾气的排放。因此，全民要积极投身到环保事业中，倡导人们低碳出行，增强人们的环保意识，开展全民节能减排活动。现在越来越多的人意识到低碳环保的重要性，倡导人们少开车，对于近距离的路程选择公共交通工具，减少尾气的排放，来缓解空气污染的壓力。近年来，共享单车在很多城市风靡，它的出现不仅减少了交通压力，还减少了空气的污染。对汽车尾气排放限制方法：可以控制城市车辆量、淘汰一批汽车尾气排放不合格的车辆、加快公共交通工具的建设，提高公共交通工具的运行速度，建立现代化的交通体系。同时鼓励市民购买新能源汽车，减少 PM2.5 的浓度。

(3) 加强绿化建设，提高植被覆盖率

PM2.5 指的是空气中直径小于或等于 2.5 微米的可被吸入肺中的微小颗粒物。对于某些植被覆盖率较小的城市，冬季干燥且风大，容易照成尘土飞扬，这样会提高 PM2.5 污染。为了防止 PM2.5 的污染还可以加强城市绿化管理，增强城市绿色植被种植，尽量减少城市裸露的空地，提高城市绿色植被的覆盖率，在城

市道路两旁设立防护带，种植吸附性较强的灌木植物，加强道路的打扫，保持道路细润度，并在城市的周边建立防护森林等

(4) 加强管理力度，完善监控体制

政府应当加强对大气污染的监管力度，对规模较大的工业进行风险排查，关闭一些污染严重、排放不符合要求的工业园区，同时要对新建的工业园区采取从严审批政策。对于监管部门，应当建立一套完整的体系来管理这些工业园区，从根本上减少 PM2.5 的浓度。

六、模型的评价

模型的优缺点

优点：

- (1) 问题一中采取了时间序列模型，运用指数平滑法能够较高精度的预测下一个时期 PM2.5 的浓度并发现每年浓度变化的趋势
- (2) 问题二中运用 K 均值聚类法，依据 PM2.5 的浓度将 31 个城市进行很好的聚类，使得 5 类划分达到平方误差最小
- (3) 运用插值法画图，能够根据曲线图直观的发现 PM2.5 的浓度变化及预测的 PM2.5 浓度在未来的变化趋势

缺点：

- (1) 指数平滑法突出时间序列不考虑其他外在因素影响，因而存在预测误差的缺陷
- (2) 基于时间序列的预测模型运算过程较为麻烦，数据多，运算过程庞大。

七、参考文献

[1] 姜启源，谢金星，叶俊，数学模型(第三版)[M]，北京:高等教育出版社， 2003。
[2] 余胜威，Matlab 优化算法分析与应用[M]，北京：清华大学出版社，2014。
[3] 司守奎，孙玺菁，数学建模算法与应用[M]，北京：国防工业出版社，2013。
[4] 韩娇. 浅谈细颗粒物 PM2.5[J]. 生物技术世界，2013（02）。

八、附录

城市序号	城市名
1	北京
2	天津

3	石家庄
4	太原
5	呼和浩特
6	沈阳
7	长春
城市序号	城市名
8	哈尔滨
9	上海
10	南京
11	杭州
12	合肥
13	福州
14	南昌
城市序号	城市名
15	济南
16	郑州
17	武汉
18	长沙
19	广州
20	南宁
21	海口
城市序号	城市名
22	重庆
23	成都
24	贵阳
25	昆明
26	拉萨
27	西安
28	兰州
城市序号	城市名
29	西宁
30	银川
31	乌鲁木齐

绘制地图代码：

```

clc
clear
sheng=shaperead('bou2_4p.shp', 'UseGeoCoords', true);
load chinese_name.mat
for i=1:length(sheng)
    sheng(i).NAME=sheng_chinese_name{i};
end

```



```

unique(sheng_chinese_name)
length(sheng)
d=importdata('gaodu.txt');
data=d.data;
textdata=d.textdata;
k=128;
mycolormap=summer(k);
geoname={sheng.NAME}';
max_data = max(data);
n=length(data);
mysymbolspec=cell(1,n);
for i=1:n
    count=data(i);
    mycoloridx=floor( k * count / max_data );
    mycoloridx(mycoloridx<1)=1;
    myprovince=textdata{i};
    geoidx=strmatch(myprovince, geoname);
    if numel(geoidx) > 0
        province_name=geoname( geoidx(1) );
        mysymbolspec{i} = {'NAME', char(province_name), 'FaceColor',
mycolormap( mycoloridx, :) };
    end
end

```

```

figure
ax=worldmap('china');
setm(ax,'grid','off')
setm(ax,'frame','off')
setm(ax,'parallellabel','off')
setm(ax,'meridianlabel','off')
symbols=makesymbolspec('Polygon',{'default','FaceColor',[0.9 0.9
0.8],...
'LineStyle','--','LineWidth',0.2,...
'EdgeColor',[0.8 0.9 0.9]},...
mysymbolspec{:}...
);
geoshow(sheng,'SymbolSpec',symbols);
colormap(summer(k))
hcb=colorbar('EastOutside');
step=round(max_data/11);
set(hcb,'YTick',(0:1:1))
set(hcb,'YTickLabel',num2cell(0:step:max_data))
title('中国人口密度分布图')

```