

城市空气污染数据的真实性判别及分析研究

张锡颖 曲红伟

(黑龙江八一农垦大学 信息技术学院 黑龙江 大庆 163319)

摘要：自我国推行空气质量数据监测后，一些地方不断爆出监测数据在改善，而环境质量却在持续恶化的新闻，此中蹊跷，引人深思。地方政府大气监测数据的准确性问题一度成为中央政府、环境保护部门和全国人民关注的热点问题。本文主要运用 MATLAB 软件对 AQI 值进行绘图分析，结合层次分析的方法对 AQI 值进行数学分析，建立了主成分分析模型，判别出地方政府上报的空气污染数据存在不真实现象，并且分析研究了数据不真实的类型及原因。

关键词：MATLAB；空气质量指数 AQI；层次分析法；主成分分析法

中图分类号：Q948.116 **文献标识码：**C **文章编号：**2096-1995(2016)11-0124-02

1 主成分分析模型的建立与求解

1.1 AQI 的空间分布特征

根据相关的天文知识和期刊文献^[5]可知：空气质量与气候有极大的相关性。因此我们在中国气象科学数据共享中心搜集了气象要素有：温度、降水量、风速和相对湿度。已知气温、降水量、风速和相对湿度的平均值按地理位置^[2]影响空气质量指数，因此可以利用层次分析法把京津冀、长三角和珠三角按照地理位置分为 3 个对象来处理。

1.2 按照空间特征分析 AQI 数据奠定建模基础

根据地方政府所给出的数据，可以观察到：

京津冀地区的北京、天津、石家庄、保定、廊坊、邢台、张家口、秦皇岛、衡水、承德和沧州这 11 个城市都缺少五天的空气质量指数数据，分别为：2014 年 1 月 23 日、3 月 24 日、8 月 8 日、8 月 22 日和 2015 年 1 月 1 日；在以上 5 天的基础上，唐山还缺少 2014 年 12 月 21 日的 AQI 数据，邯郸缺少 2014 年 12 月 14 日的 AQI 数据。

长三角地区的南京、杭州、无锡、常州、镇江、扬州、泰州、南通、湖州、宁波、台州和舟山这 12 个城市都缺少了五天的 AQI 数据，并且与北京等皆是相同的 5 天；在北京等城市的基础上，上海还缺少 2014 年 12 月 6 日和 12 月 7 日的 AQI 数据，绍兴缺少 2014 年 12 月 5 日和 12 月 6 日的 AQI 数据，共缺少 7 天，苏州还缺少 2015 年 1 月 2 日、3 日、4 日和 5 日的 AQI 数据，嘉兴缺少 2014 年 12 月 19 日、20 日、21 日和 22 日的 AQI 数据，共缺少 9 天。

珠三角地区的所有城市都缺少了五天的 AQI 数据，并且与北京等皆是相同的 5 天。

1.3 模型的建立与求解

(1) 主成分分析法模型的建立

随机抽取京津冀、长三角和珠三角中的 15 个不同城市，计算出这些城市空气的主要污染物：PM2.5、PM10、二氧化氮、二氧化硫和一氧化碳的年平均浓度，将影响空气质量优劣的指标看作不同的主体。建立观测数据矩阵为

$$X = \begin{bmatrix} x_{11} & x_{12} & L & x_{1p} \\ x_{21} & x_{22} & L & x_{2p} \\ M & M & M & M \\ x_{m1} & x_{m2} & L & x_{mp} \end{bmatrix} \quad (1)$$

(2) 对各影响因素原始数据进行标准化处理

我们可以用公式 2-4 求出各个指标，即因素的均值和方差，可以大体得到第 i 个城市的第 j 个因素 (PM2.5、PM10、二氧化碳、二氧化硫和一氧化碳) 对该城市空气质量影响的差别情况和对三个不同地区影响的平均水平，也就是说 越大，第 j 个因素对第 i 个城市的空气质量的影响越大。

$$x_j^* = \frac{x_{ij} - x_i}{\sqrt{\text{var}(x_j)}} \quad (i=1,2,L,15; j=1,2,L,5) \quad (2)$$

其中

$$\bar{x}_j = \frac{1}{n} \sum_{i=1}^m x_{ij} \quad (3)$$

$$\text{var}(x_j) = \frac{1}{n-1} \sum_{i=1}^m (x_{ij} - \bar{x}_j)^2 \quad (4)$$

(3) 计算样本相关系数矩阵

$$R = \begin{bmatrix} r_{11} & r_{12} & L & r_{1p} \\ r_{21} & r_{22} & L & r_{2p} \\ M & M & L & M \\ r_{m1} & r_{m2} & L & r_{mp} \end{bmatrix} \quad (5)$$

其中

$$r_{ij} = \frac{1}{m-1} \sum_{k=1}^m x_{ki} x_{kj} \quad (i=1,2,...,15; j=1,2,...,L...5) \quad (6)$$

对于每个城市不同地区主体因变量来说，样本的相关系数 r_{ij} 代表 i ($i=1, 2, \dots, 15$) 个不同地区与第 j 个因素的相关性关系，即 r_{ij} 越趋近于 1，则说明相关性越大，反之，则相关性越小。 (下转 P123)

公共供油管,通过公共供油管压力的调节实现最后的高压燃油喷射。高压共轨电喷技术可以使喷油压力不受发动机转速的干扰,实现更好的燃油经济性。近年来,随着柴油车的功率要求不断提高和排放法规的日益严格,柴油发动机需要不断提高功率系数并降低自身油耗。这就要求柴油喷射系统的压力不断增加,控制方案更加高效。柴油共轨电喷技术因其上述优势得到了全球柴油发动机生产商的青睐。

4 底盘节能与环保技术

底盘作为承载整车重量,传递动力的重要部分,在节能和环保方面也拥有着较大的发展潜力。目前底盘节能与环保技术主要集中在自动变速器技术中。汽车变速器可以改变汽车动力传递的传动比,使汽车适应不同的路面条件。传统的手动变速器对驾驶员的技术要求性很高,换挡操作复杂。因此,越来越多的消费者更加青睐自动变速器,它操作简单,但是油耗较手动变速器略有增加。

目前主流的自动变速器有有级式自动变速器(AMT),无极式自动变速器(CVT)以及综合式自动变速器(AT)。其中AMT的性价比最高,目前在AMT基础上开发的双离合自动变速器(DCT)已经有了较快的发展。DCT即双离合自动变速器,它拥有两个独立的离合器,可以保证换挡过程动力不中断,换挡过程更加平稳,驾驶者感受不到换挡的顿挫感,并且DCT的油耗较其他几种自动变速器都要少,它的油耗更加接近手动变速器的水准。除此以外,DCT还可以利用AMT的零部件在相同的生产线上进行生产,降低了生产的成本。DCT作为未来先进自动变速器的发展方向,必将提高动力传递效率,增加换挡舒适性的同时进一步减少油耗。

5 车身节能与环保技术

车身节能与环保技术主要集中在车身的轻量化制造以及采用更加环保的制造工艺和材料方面。

随着汽车轻量化技术的成熟,一些轻质材料开始被越来越多的使用。类似镁合金,铝合金,钛合金等轻质材料用来制造车身,可以使整车的质量大大降低。相应的汽车在减少惯性,提高操纵灵活性的同时更有效降低了油耗,提升了燃油经济性。与此同时,铝合金,镁合金等还可以实现百分之百的循环再制造,减少了材料浪费,实现了材料的循环利用。

采用更加先进的制造工艺例如激光拼焊技术,液压成型技术以及热成型技术等可以实现更加复杂的工艺结构并减少制造时间。这些先进的制造工艺不仅可以生产出轻量化的零件,还减少了生产过程中能源消耗和材料浪费。

6 结束语

汽车工业同时面临着巨大的市场和严峻的挑战,能源消耗,环境污染等问题制约着汽车工业的进一步发展。发展新能源汽车虽然可以彻底的解决上述问题,但其转型的技术成本和社会成本较大,近期内传统内燃机汽车仍会占据市场绝大部分份额。因此我们必须大力发展传统汽车的节能与环保技术来谋求汽车工业的可持续发展。

参考文献:

- [1] 李骏,曲卫东,高巍.汽车动力总成节能环保先进技术分析[J].汽车技术,2009(05):1-7.
- [2] 张宗法,吴家年,韦文求.浅谈未来车用发动机发展趋势[J].汽车工业研究,2012(08):9-13.

(上接 P124)

(4) 计算特征值和特征向量

用 MATLAB 程序编程求相关系数矩阵 R 的特征值 $\lambda_1, \lambda_2, \dots, \lambda_p$ 和相应的特征向量 L_i

$$L_i = [l_{1i} \ l_{2i} \ l_{3i}]^T, \quad i=1,2,\dots,15 \quad (7)$$

(5) 计算选择 t 个主成分,计算综合评价

将特征值从大到小排列,设排列顺序为 $\lambda_1, \lambda_2, \dots, \lambda_5$ 找出 t, 计算其贡献率

$$Y_{ik} = \frac{\sum_{j=1}^t \lambda_j}{\sum_{j=1}^p r_{jj}} \quad (8)$$

为主成分 Y_{kj} 的信息贡献率

$$\alpha_t = \frac{\sum_{j=1}^t \lambda_j}{\sum_{j=1}^p r_{jj}} \quad (9)$$

为主成分 Y_t 的累计贡献率。当主成分 α_t 的累积贡献率接近于 1 ($\alpha_t = 0.58, 0.90, 0.95$) 时,则选择前 t 个

指标变量 Y_1, Y_2, \dots, Y_t 作为 t 个主成分,从而可对 t 个主成分进行综合分析。即 t 个空气质量优劣有影响的主要因素,令

$$Z_i = \sum_{j=1}^p l_{ji} Y_j, \quad (i=1,2,\dots,t) \quad (10)$$

再计算前 t 个按方差比例大小进行的排列主成分的值

$$z_i = \sum_{j=1}^p l_{ji} Y_j, \quad (i=1,2,\dots,15; j=1,2,\dots,t) \quad (11)$$

其中 Y_{ik} 为第 j 个主成分的信息贡献率,根据综合得分值就可以进行评价。

参考文献:

- [1] HJ633-2012, 中华人民共和国国家环境保护标准[S].
- [2] 周兆媛,张时煌,高庆先,等.京津冀地区气象要素对空气质量的影响及未来变化趋势分析[J].资源科学,2014(1):191-199.
- [3] 滕延芹.基于主成分分析的空气质量指数评价方法优化[J].资源科学,2016(3):9-24.
- [4] 林艺滨.Excel 软件在计算空气质量指数的应用[J].能源与环境,2012(05):31-35.