文章编号 1000-5269(2020)03-0111-08 DOI:10.15958/j.cnki.gdxbzrb.2020.03.19

基于函数型数据的气温与空气质量的相关性分析

杨 荣,戴家佳*

(贵州大学 数学与统计学院, 贵州 贵阳 550025)

摘 要:随着我国经济的快速发展,大气污染问题日益严重,旅游城市的空气质量问题也备受关注。本文基于 2018 年 30 座热门旅游城市的气温和空气质量数据,从气象条件的角度出发,研究了气温与空气质量之间的关系。采用函数型数据描述性分析的方法对空气质量和气温变化特征进行分析,用函数型回归分析研究气温对空气质量的影响。结果显示:函数型线性模型拟合优度 $R^2 = 0.898\ 1$, $F = 11.821\ 2$ 大于临界值 2.385 9,函数型检验统计量 $\tilde{F} = 8.723\ 3$ 大于 $\alpha = 0.05$ 水平下的临界值 1.705 8,说明气温与空气质量的函数型线性关系显著;气温对空气质量呈谐波正负影响交替出现,这与空气质量平均变化相似,在实际生活中是合理的。

关键词:函数型回归模型;基函数;最小二乘估计;气象因素;空气质量

中图分类号:X51;P423 **文献标识码**: A

2018 年联合国气候大会指出:“要将全球升温控制在 1.5 ℃内,如果能够坚持 1.5 ℃而不是 2 ℃的升温控制,受气候变化影响的人会减少 4.2 亿”。气候变化和经济社会发展带来的空气质量的恶化对人类健康和生活造成了一定影响。根据环境保护部发布的《环境空气质量指数(AQI)技术规定(试行)》[1],参与评价空气质量指数的污染物主要有细颗粒物($PM_{2.5}$)、可吸入颗粒物(PM_{10})、二氧化硫(SO_2)、二氧化氮(NO_2)、臭氧(O_3)、一氧化碳(CO)六项。

除了已纳入评价空气质量的这些污染物,气象因素也是影响空气质量水平的因素之一。2013 年陈松等[2]研究了驻马店城市温度变化与空气质量的关系;2018 年李育慧等[3]指出,在污染源排放相对稳定的情况下,对空气质量水平起主要作用的是气象因素;2018 年吕爱丽等[4]研究了大气污染与气象条件的相关性,研究表明 $PM_{2.5}$、PM_{10}、SO_2、NO_2浓度与气压存在正相关关系,与温度、降水、风速、日照存在负相关关系;2018 年盛永财等[5]应用主成分分析法研究了气象因素对乌鲁木齐市城区空气质量的影响;2018 年何佳璐等[6]通过相关分析方法研究了舟山市 $PM_{2.5}$浓度与气象因素的关系,结果表明大气 $PM_{2.5}$浓度随气温上升呈下降趋势,并且与相对湿度、降水量和风速等气象因素呈负相关关系;2018 年付伟等[7]研究了芜湖市气象因素对空气质量的影响,结果显示,芜湖市 6 种污染物表现出不同的季节特征,主要的污染物 $PM_{2.5}$ 冬季浓度最高,春、夏次之,秋季最低。由此可知气象因素会影响空气质量水平,我们研究气温对空气质量的影响是具有实际意义的。

1991 年 RAMSAY 和 DALZELL[8-9]正式提出了函数型数据分析(Functional Data Analysis,FDA),1997 年 RAMSAY[10]对一些传统统计分析方法进行改进使之适用于函数型数据分析,如主成分分析,线性回归分析和典型判别分析等,总结了函数型数据分析的理论和方法,全面地阐述了函数型数据的特点及其统计分析思想方法,极大地推动了函数型数据分析的发展。在函数型数据分析中,函数型回归模型是一个重要的数据分析工具,它主要用于分析函数型数据协变量与响应变量之间的关系。1999 年 CARDOT 等[11]使用光滑样条技术和函数型主成分(FPCA)得到了斜率函数的估计,并研究了估计的渐近性质;2006 年 CAI 和 HALL[12]用函

收稿日期:2019-11-29

基金项目:国家自然科学基金项目资助(11361015)

作者简介:杨 荣(1994-),女,在读硕士,研究方向:概率论与数理统计,Email:276602939@ qq.com.

* **通讯作者**:戴家佳,Email:sci.jjdai@ gzu.edu.cn.

数型主成分估计研究了函数型线性模型的估计与预测问题;2010 年 HALL 和 YANG[13] 分别用样条、主成分估计、偏最小二乘估计研究了函数型线性模型的估计与预测问题。在以往空气质量的研究中,一般采用传统的统计方法对空气质量的影响因素进行分析,但从时间尺度上来看,发现研究空气质量的影响因素的数据具有明显的函数特性,所以本文采用函数型数据分析(FDA)的方法来分析气温与空气质量的关系。函数型数据分析相对于传统的统计分析方法,它依赖的假设条件更少,将数据看为一个动态变化的过程,能挖掘出丰富的数据信息,给出更合理的解释。

本文首先对函数型数据回归模型的基本理论进行介绍,其次利用函数型回归分析讨论 2018 年 30 座热门旅游城市空气质量与气温变化之间的关系,并得出相应结论。

1 函数型回归模型的基本理论

在实际应用中,函数型回归模型是讨论函数型数据协变量与响应变量之间关系的一个重要方法。函数型回归模型一般分为以下几种类型[14]:1)协变量是函数,响应变量是标量;2)协变量是标量,响应变量是函数;3)协变量和响应变量都是函数;4)协变量是函数和标量的混合变量。本文主要研究第一种函数型线性模型,以 30 座热门旅游城市 2018 年全年气温数据为函数型协变量,记为 $x_i(t)$, $i=1,2,3,\cdots,30$,以空气质量指数为标量型响应变量,其中对每座城市全年空气质量指数求和再对其取对数,记为 y_i , $i=1,2,3,\cdots,30$。

1.1 函数型回归模型的估计

本文选用的模型为[9]:

$$y_i=\alpha_0+\int_T\beta(t)x_i(t)\mathrm{d}t+\varepsilon_i\ ,\ i=1,2,3,\cdots,n。\tag{1}$$

其中:$x_i(t)$ 为定义在紧区间 T 上的函数为型协变量,表示第 i 个城市的气温变化曲线;y_i 为第 i 个城市求和取对数后的空气质量指数;α 和 $\beta(t)$ 分别为模型的未知截距项和待估斜率函数;ε_i 为零均值方差有限的误差项且满足 $E(\varepsilon\mid X)=0$,与 $x_i(t)$ 独立, n 为样本量。

令

$$Y=\begin{pmatrix}y_1\\y_2\\\vdots\\y_n\end{pmatrix}\ ,\ X(t)=\begin{pmatrix}x_1(t)\\x_2(t)\\\vdots\\x_n(t)\end{pmatrix}\ ,\ \varepsilon=\begin{pmatrix}\varepsilon_1\\\varepsilon_2\\\vdots\\\varepsilon_n\end{pmatrix}。$$

则模型可写为:

$$Y=\alpha+\int_T\beta(t)X(t)\mathrm{d}t+\varepsilon,\ i=1,2,3,\cdots,n。\tag{2}$$

对模型(1)的估计一般是对 $x_i(t)$,$\beta(t)$ 进行基函数展开,即将 $x_i(t)$,$\beta(t)$ 表示成已知基函数的线性组合。对于基函数的选取,可以事先指定,一般具有周期性的数据用傅立叶(Fourier) 基展开;非周期性的数据用 B-样条基展开;当曲线中有剧烈跳动信号时,采用小波基展开。将选择的基函数设为 $\varphi_1(t)$,$\varphi_2(t)$,$\cdots$,对协变量和斜率函数进行基函数展开,可表示为:

$$x_i(t)=\sum_{k=1}^{K}c_{ik}\varphi_{ik}(t)=c'\varphi(t)\ ,\tag{3}$$

$$\beta(t)=\sum_{k=1}^{K}b_k\varphi_k(t)=b'\varphi(t)。\tag{4}$$

对模型(1)的估计,我们采用粗糙惩罚法来进行估计。则模型(1)的目标函数为:

$$PEN_\lambda(\alpha_0,\beta)=\sum\left[y_i-\alpha_0-\int\beta(t)x_i(t)\mathrm{d}t\right]^2+\lambda\int\left[L\beta(t)\right]^2\mathrm{d}t。\tag{5}$$

其中:λ 为惩罚参数,则协变量和惩罚矩阵可表示为:

$$Z=\begin{pmatrix}1&\int x_1(t)\Phi(t)\mathrm{d}t\\\vdots&\vdots\\1&\int x_n(t)\Phi(t)\mathrm{d}t\end{pmatrix}\ ,\ R(\lambda)=\begin{pmatrix}0&0\\0&\lambda R\end{pmatrix}。$$

其中:$\Phi(t)$ 为 $\beta(t)$ 的基函数展开, $R=\int\varphi(t)\varphi'(t)\mathrm{d}t$,则估计的系数为:

$$\hat{b}=(Z'Z+R(\lambda))^{-1}Z'y\ 。\tag{6}$$

将(6)式带入(4)、(5)式可以得到截距项和斜率函数的最小二乘估计。

对于粗糙惩罚参数选取,我们采用交叉验证法来进行选择。

$$C_{\mathrm{V}}(\lambda)=\sum_{i=1}^{N}\left(\frac{y_i-\hat{y}_i}{n-H_{ii}}\right)^2。$$

需要找出使得 C_{V} 达到最小的 λ 作为惩罚参数,λ 确定后可以计算出 $\hat{y}$:

$$\hat{y}=Z\left(Z'Z+R(\lambda)\right)^{-1}Z'y=Hy。$$

1.2 函数型回归模型的显著性检验

通常需要衡量线性回归关系对给定数据集的拟合优度[15],与传统线性回归模型一样采用复决

定系数来衡量,即:

$$R^2 = \frac{S_T - S_E}{S_T}。$$

其中:$S_E = \sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2$ 为残差平方和;$S_T = \sum_{i=1}^{n} (Y_i - \bar{Y})^2$ 为总离差平方和,R^2 越大,说明线性模型对数据拟合得越好。

对回归关系的显著性检验我们用检验统计量 F 进行描述:

$$F = \frac{(S_T - S_E)/p}{S_E/(n - p - 1)}。$$

F 值越大,说明回归关系越显著。

由于函数型统计的性质[16],对于任何给定的检验统计量很难得到理论上的零分布,对函数型回归的统计检验,采用置换检验(Permutation Test)来检验响应变量与协变量之间的关系强度。置换检验是利用样本数据的随机排列进行统计推断的方法,它对总体分布自由,不需要依赖于分布假设,适用于总体分布未知以及某些难以用常规方法分析资料的假设检验问题。这里对响应变量进行随机排列,而协变量顺序保持不变,因为如果响应变量与协变量之间没有关系,那么随机地重新排列它们配对的方式,那么它应该不会产生差异,由此构造检验统计量。

$$\tilde{F} = \frac{V[\hat{y}]}{\frac{1}{n}\sum (y_i - \hat{y}_i)^2}。$$

通过重复计算统计量构造经验分布,如果观察到检验统计量处于该分布的尾部,则说明响应变量与协变量之间存在相关性。

2　实例分析

2.1　数据来源

从 2014 年起,国家环保部每天公布全国161 座城市的日空气质量指数,为空气质量研究提供了基础数据,每日空气质量指数(Air Quality Index,简称 AQI)数据是根据当天环保总站每小时数据计算求平均的结果;由于每个城市气温都有最高气温和最低气温,这里选取最高气温作为研究指标。本文主要研究 2018 年 30 座热门旅游城市(如北京、重庆、上海、广州等)的 AQI 和气温温度的相关性,其数据分别来源于 https://www.aqistudy.cn/historydata/和 http://tianqi.2345.com/wea_history/57816-.htm,部分数据见表 1。

表 1　部分研究数据展示

Tab.1　Partial research data display

日期	北京温度/℃	北京 AQI	重庆温度/℃	重庆 AQI	天津温度/℃	天津 AQI	成都温度/℃	成都 AQI
2018-01-01	3	60	11	89	4	100	10	137
2018-01-02	2	49	10	97	2	68	9	137
2018-01-03	2	29	9	57	1	40	5	60
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
2018-12-31	-2	70	7	63	0	88	4	72

2.2　空气质量与气温变化的特征分析

对 2018 年 30 座热门旅游城市的空气质量指数和气温变化进行分析,分析结果如下。

从图 1 可知,这些城市的平均空气质量指数在 40~150,总体上空气质量良好。空气质量指数的高峰值大概出现在 1、11 和 12 月,此时处于冬季,是取暖的高峰期,燃煤、天然气等化学物质大量排放,导致这一时段的空气污染物浓度加大,空气质量变差;此后,空气质量在春、夏、秋三个季节慢慢转好。

对 30 座热门旅游城市的 AQI 利用 B-样条基函数函数化,并进行拟合,部分拟合结果如图2所

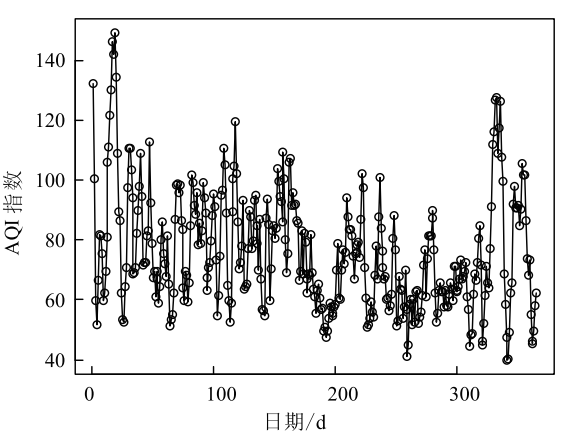
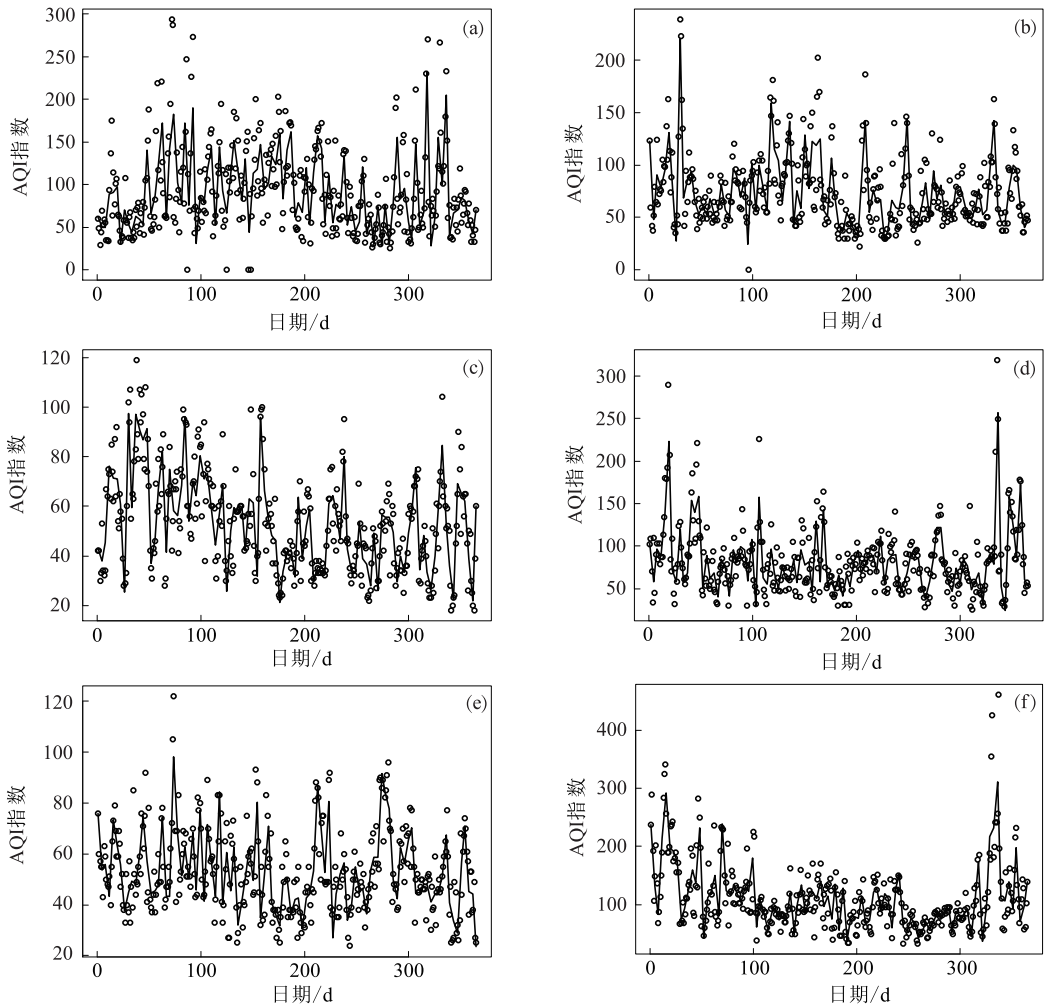


图 1　2018 年 30 座热门旅游城市 AQI 平均变化图

Fig.1　Average AQI changes of 30 popular tourist cities in 2018



(a)北京;(b)上海;(c)贵阳;(d)长沙;(e)厦门;(f)西安。

图 2 部分城市的 AQI 拟合图

Fig.2 AQI fitting graphs of some cities

示。结果显示,从总体上看,30 座城市的 AQI 走势与其平均变化走势大致一致。从城市等级分布来看,二线城市空气质量要优于一线城市;从地域分布来看,华东地区的空气质量要优于华北地区,上海的空气质量优于北京;西南地区又优于华中地区,贵阳要优于长沙;沿海地区比内陆城市的空气质量好,厦门要优于西安。根据这 30 座旅游城市的空气质量情况,可分为三个梯度,华北地区空气质量最差,华中地区良好,西南地区最好。呈现这样的梯度变化可能是由于城市发展水平越高,工业越发达,人流量越大,汽车尾气排放量就越大,造成空气污染严重。另一方面有可能是由华北到西南地区,地势由平缓逐步变崎岖,华北、华中地区城市建筑群密集,城市植被覆盖率不高,不利于大气污染物扩散,尤其是颗粒物长期留在空气中或附着在建筑物上;相较而言,西南地区城市建筑群比较稀疏,城市绿植覆盖率较高,而植被对空气质量具有一定地调节作用;从时间维度来看,城市的 AQI 水平冬季高,春、秋次之,夏季最低。

从 30 座热门旅游城市的平均气温变化可以看出,这些热门旅游城市总体气温是比较适宜的,平均最低温度 2.5 ℃,平均最高温度 34 ℃。将图 1 和图 3 结合分析,当气温比较低时,AQI 值比较高,空气质量比较差;当气温比较高时,AQI 值比较低,空气质量比较好。为了详细地分析气温与空气质量的关系,对 30 座城市的气温进行拟合,分析结果如下。

将 AQI 拟合图(图 2)与气温拟合图(图 4)对应分析,无论是从地域分布还是时间维度上来看,总体上气温与 AQI 值呈负相关关系,气温越低,AQI值越高,空气质量越差;气温越高,AQI值越低,

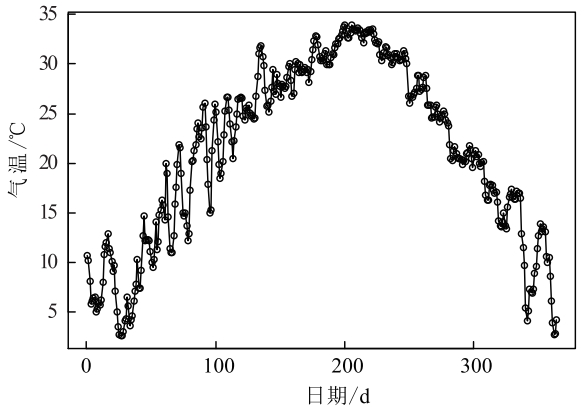


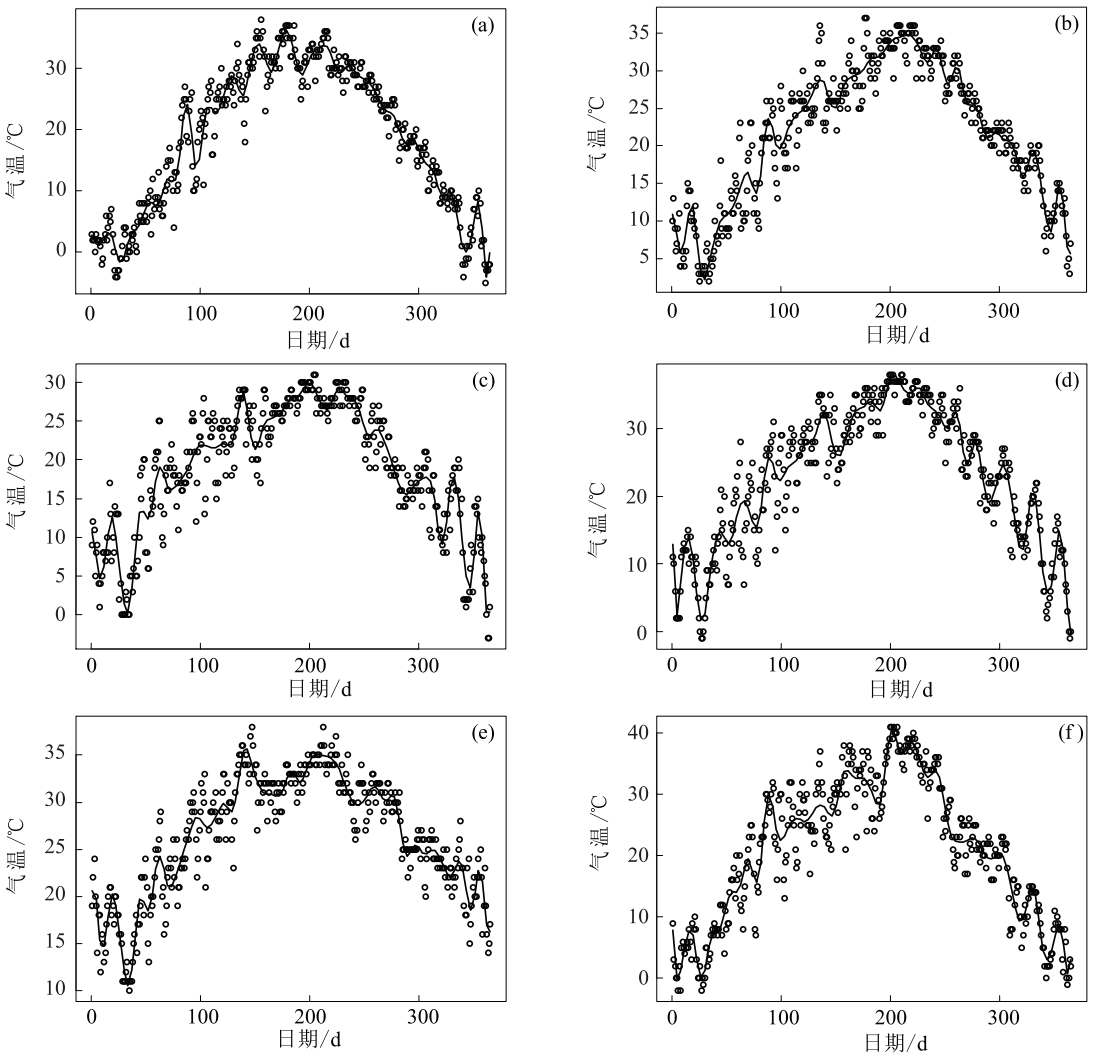
图 3　2018 年 30 座热门旅游城市的气温平均变化图

Fig.3　Average temperature changes of 30 popular tourist cities in 2018

空气质量越好。具体来看,冬季时,华北、华中地区总体气温比西南地区低,其空气质量也比西南地区差;夏季时,华北、华中地区的总体气温都比西南地区高,但其空气质量还是比西南地区差,这就需要进一步探索气温与空气质量水平之间的影响关系。

2.3　函数型线性回归估计结果

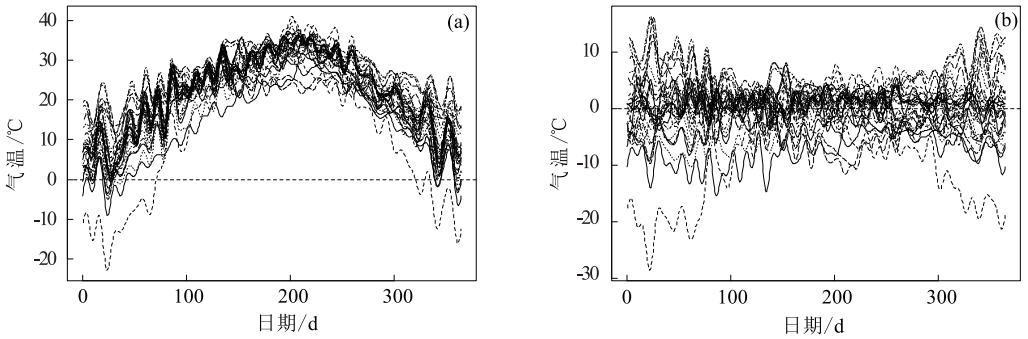
本文重点研究是气温对空气质量的影响,将空气质量指数(AQI)作为响应变量,气温温度作为函数型协变量。气温和空气质量指数数据都是日度数据。对于空气质量指数,将每个城市 2018 年的 AQI 指数求和后再取对数,这样有利于计算;对于气温数据,它是周期数据,所以采用傅立叶基函数将 30 个城市的气温数据进行函数平滑并中心化,结果如图 5 所示。



(a)北京;(b)上海;(c)贵阳;(d)长沙;(e)厦门;(f)西安。

图 4　部分城市的气温拟合图

Fig.4　Temperature Fitting in Some Cities



(a)β(t)的估计;(b)β(t)的 95%置信区间。

图 5 2018 年 30 座热门旅游城市的气温平滑曲线图

Fig.5 Smooth temperature graphs of 30 popular tourist cities in 2018

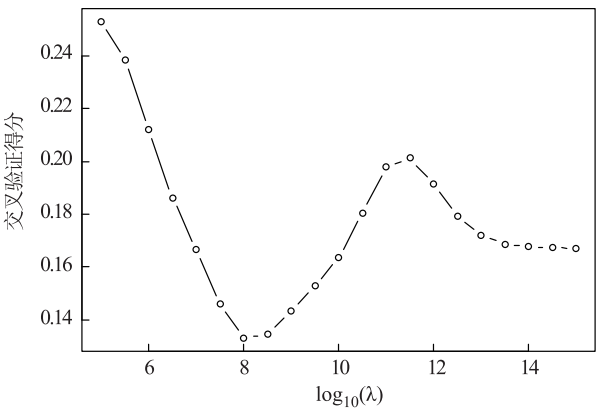
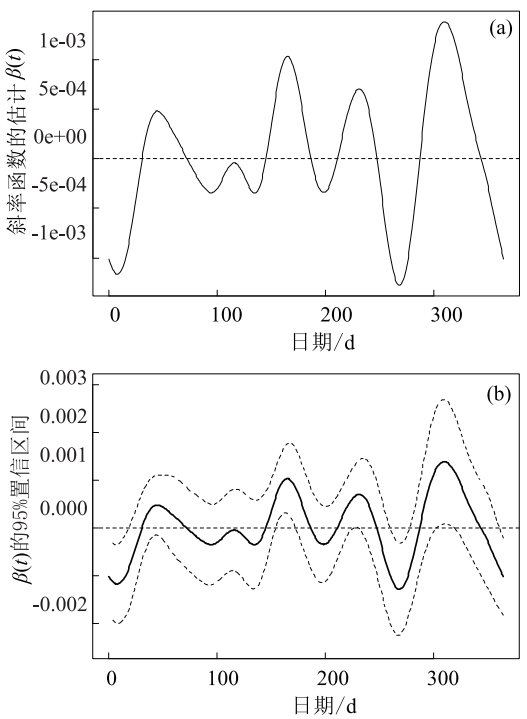


图 6 光滑惩罚参数选择

Fig.6 Smooth penalty parameter selection



(a)$\beta(t)$的估计;(b)$\beta(t)$的 95%置信区间。

图 7 斜率函数 $\beta(t)$ 的估计和其 95%置信区间

Fig.7 Estimation of the slope function $\beta(t)$ and its confidence interval of 95%

本文采用 2.1 节介绍的粗糙惩罚法进行函数型线性回归,并对斜率函数 $\beta(t)$ 进行估计。选取惩罚参数 λ 。

由图 6 可知使得 C_V 达到最小的 $\lambda = 10^8$,用这个参数进行惩罚回归,截距项 $\alpha = 0.012\ 14$,斜率函数估计结果和其 95% 的置信区间如图 7 所示。

模型(1)可表示为:

$$y_i = 0.012\ 14 + \int_T \beta(t) X_i(t)\,\mathrm{d}t + \varepsilon\ ,i = 1,2,3,\cdots,n。 \tag{7}$$

其中:y_i 为第 i 个城市取对数后的空气质量指数;$X_i(t)$ 为第 i 个城市的气温拟合函数。

对观测值与拟合值进行拟合,结果如图 8 所示:

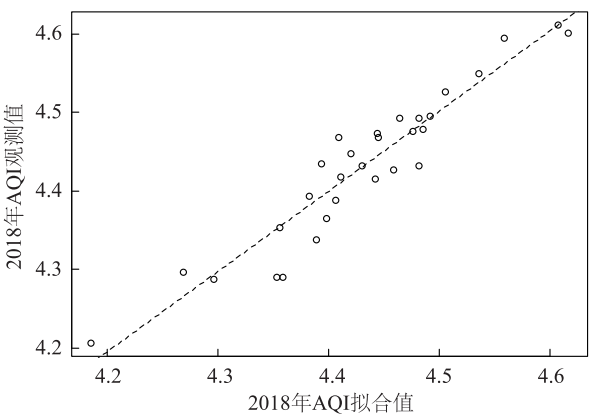


图 8 观测值与拟合值

Fig.8 Observations and fitted values

对模型(1)的显著性可以通过构造针对响应变量为标量,协变量为函数型变量的 $\tilde{F}$-统计量来检验,同时还可以用传统统计方法中的复决定系数 R^2 和 F 检验来进行。在给定显著性水平 $\alpha = 0.05$ 下,函数型 $\tilde{F}$-统计量检验结果如下。

由图 9 可知,在 95% 的置信水平下, $\tilde{F}$ = 8.723 3 大于 α = 0.05 水平下的临界值1.705 8,对模型(1)的检验是显著的,即说明气温对空气质量的影响是显著的;同时计算出复决定系数 R^2 = 0.898 1, F = 11.821 2 大于临界值 2.385 9, 这说明气温对空气质量的函数型线性关系很强。

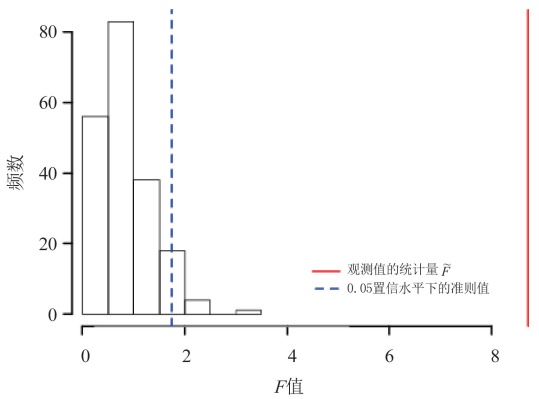


图 9　气温变化对空气质量影响的显著性检验结果

Fig.9　Significant test results of the effect of temperature changes on air quality

2.4　气温对空气质量的影响分析

将图 1 和图 7 结合分析,估计的斜率函数 $\beta(t)$ 的函数曲线振幅变化与 AQI 值平均变化相似,并且 $\beta(t)$ 的曲线振幅变化越大,说明气温对空气质量的影响越强。图 8 中观测值与拟合值大概在一条直线上,且它们的相关系数为0.898 6,说明用函数型回归模型探索气温与空气质量之间的响应关系是比较合理的。

从图 7 中发现,在冬季,气温对空气质量指数有负影响,气温越低,AQI 值越高,空气质量越差。在春、夏、秋三季气温对空气质量的影响呈正负交替出现。出现这样的情况可能是这 30 座热门旅游城市处于我国的季风区,冬季陆地近地层受蒙古西伯利亚高压系统控制,盛行偏北风,气候寒冷干燥,造成气压高,温度低,风速小,空气内部气流下沉,低层空气对流较弱,不利于污染物扩散,并且春节期间在冬季,人们出行频繁,走亲戚或出门旅游,汽车尾气排放增加,燃放烟花爆竹等增加污染物浓度,所以冬季空气质量较差。

春、夏、秋三季的空气质量比冬季好,这可能是由于春季以后,气温逐步回升,到了夏季陆地近地层受印度低压控制,盛行偏南风,气候湿热,造成气压低,温度高,风速大,气流做垂直上升运动,近地层大气对流强,有利于污染物的扩散,再加上这三个季节比较多雨,降雨对污染物浓度有一定的稀释作用;在图 7 中,发现在春、夏、秋三季,气温对空气质量指数的影响呈正负交替出现,这可能是由其他气象条件(例如降雨时会降温)或其他因素影响导致气温对空气质量呈谐波波动影响。例如,在春季,气温不稳定,忽高忽低,气温对 AQI 值的影响有正有负;在夏季,总体气温较高,降雨量也较大,并且当温度过高时,人们外出减少,空气污染减少,空气质量良好;在秋季气温对 AQI 值有较强的正影响,这可能由于是国庆小长假,天气越好,外出旅游的人流量越大,汽车尾气排放量越大,空气污染增加,造成气温越高空气质量越差的现象。

3　结论

本文通过对 2018 年 30 座热门旅游城市的气温和空气质量变化特征进行分析,发现这些城市的空气质量变化具有季节性和地域性特征,冬季最差,春秋次之,夏季最好;空气质量由华北向西南地区逐渐变好,并且气温对空气质量有显著性影响。

通过利用函数型回归分析的方法对气温与空气质量之间的响应关系进行探讨,发现气温对空气质量具有显著性影响。经回归诊断, R^2 = 0.933 8, F = 11.284 7 大于临界值 2.521 5, 函数型回归诊断的 $\tilde{F}$ = 14.084 55 大于 α = 0.05 水平下的临界值 2.927 7, 充分说明了气温与空气质量的函数型线性关系是显著的。根据斜率函数 $\beta(t)$ 的估计曲线,发现气温对空气质量是呈正负交替影响的,并且 $\beta(t)$ 的振幅变化越强,气温对空气质量的影响越大。由此,气温不仅可以作为人们出行的晴雨表,也可以作为判断空气质量好坏的晴雨表。

参考文献:

[1]环境空气质量指数(AQI)技术规定(试行)[J].中国环境管理干部学报,2012,22(1):44.

[2]陈松,张俊波,房稳静. 驻马店城市温度变化与空气质量的关系分析[J].河南科学,2013,31(6):855-859.

[3]李育慧,张家斌,李育乔. 漳州市大气颗粒物时空分布特征及其与气象条件的关系[C]//第 35 届中国气象学会年会 S12 大气成分与天气、气候变化与环境影响暨环境气象预报及影响评估. 合肥:中国气象学会,2018:305-311.

[4]吕爱丽,王会荣,赵晓丽. 临汾市大气污染变化特征及其与气象条件的相关分析[C]//第 35 届中国气象学会年会 S12 大气成分与天气、气候变化与环境影响暨环境气象预报及影响评估. 合肥:中国气象学会,2018:390-396.

[5]盛永财,孜比布拉·司马义,李颖. 气象因素对乌鲁木齐市城区空气质量的影响[J].环境工程,2018,36(5):78-83.
[6]何佳璐,李鹏,张永利,等. 舟山市 PM2.5 浓度与气象因素的关系研究[J].预防医学,2018,30(12):1245-1247.
[7]付伟,司红君,卢尧,等. 芜湖市空气质量特征及其受气象要素的影响分析[C]//第 35 届中国气象学会年会 S12 大气成分与天气、气候变化与环境影响暨环境气象预报及影响评估. 合肥:中国气象学会,2018:128-129.
[8]RAMSAY J O. When the data are functions[J]. Psychometrika, 1982, 47(4):379-396.
[9]RAMSAY J O, DALZELL C J. Some tools for functional data analysis[J]. Journal of the Royal Statistical Society, 1991, 53(3):539-572.
[10]RAMSAY J O, SILVERMAN B W. Functional data analysis[M]. New York: Springer, 2005.
[11]CARDOT H, FERRATY F, SARDA P. Functional linear model [J]. Statistics & Probability Letters, 1999, 45(1):11-22.
[12]CAI T T, HALL P. Prediction in functional linear regression[J]. Annals of Statistics, 2006, 34(5):2159-2179.
[13]HALL P, YANG Y J. Ordering and selecting components in multivariate or functional data linear prediction[J]. Journal of the Royal Statistical Society, 2010, 72(1):93-110.
[14]丁辉,许文超,朱汉兵,等. 函数型数据回归分析综述[J]. 应用概率统计,2018,34(6):630-654.
[15]梅长林,王宁. 近代回归分析方法[M]. 北京:科学出版社,2012.
[16]RAMSAY J, HOOKER G, GRAVES S. Functional data analysis with r and MATLAB[M]. New York: Springer, 2009.

(责任编辑:于慧梅)

Correlation Analysis of Temperature and Air Quality Based on Functional Data

YANG Rong, DAI Jiajia*
(School of Mathematics and Statistics, Guizhou University, Guiyang 550025, China)

Abstract: With the rapid development of China's economy, the problem of air pollution has become increasingly serious, and the air quality problems in tourist cities have also received much attention. Based on the temperature and air quality data of 30 popular tourist cities in 2018, this paper studies the relationship between temperature and air quality from the perspective of meteorological conditions. Descriptive analysis of functional data was used to analyze the characteristics of air quality and temperature changes, and functional regression analysis was used to study how temperature affects changes in air quality. The analysis results show that the functional goodness of the linear model is $R^2 = 0.898\ 1$, $F = 11.821\ 2$ is greater than the critical value of 2.385 9, and the functional test statistic $\tilde{F} = 8.723\ 3$ is greater than the lower critical value of 1.705 8 when $\alpha = 0.05$, it indicates that the linear relationship between temperature and air quality is significant, and the temperature has an alternating positive and negative effect on the air quality. This is similar to the average change in air quality and is reasonable in real life.
Key words: functional regression model; basis function; least squares estimation; meteorological factors; air quality.