

Correlation Analysis between Air Quality and Meteorological Indicators in Qingdao

Lei Wang, Shuhe Lei, Shasha Ma

School of Mathematical Sciences, Ocean University of China, Qingdao Shandong

Email: wanglei_94@qq.com

Received: Oct. 17th, 2018; accepted: Oct. 30th, 2018; published: Nov. 6th, 2018

Abstract

In order to study the correlation between air quality and meteorological factors in Qingdao, the data of air quality and meteorological factors in Qingdao from January 2014 to December 2017 were collected. Canonical correlation analysis was conducted on the whole, four seasons and heating season data respectively. It shows that there is a correlation between air quality and meteorological factors, especially in autumn and winter. Air quality indicators affected by meteorological factors are different in different seasons. MSC and EMD method were used to observe the correlation of the band data. The relative humidity in spring and summer has great influence on PM2.5. On the 42 day periodic components, the increase of the relative humidity can reduce the concentration of PM2.5. In autumn and winter, O₃ are closely related to air temperature. As the component cycle becomes longer, the effect of temperature rise will be enhanced.

Keywords

Air Quality, Meteorological Factors, Canonical Correlation, MSC, EMD

青岛市空气质量指标与气象指标的相关性分析

王 蕾, 类淑河, 马莎莎

中国海洋大学数学科学学院, 山东 青岛

Email: wanglei_94@qq.com

收稿日期: 2018年10月17日; 录用日期: 2018年10月30日; 发布日期: 2018年11月6日

摘 要

为研究青岛市空气质量与气象指标之间的相关关系, 收集了青岛市2014年1月至2017年12月的空气质量

与气象指标数据。分别对全段、四季以及供暖季数据进行了典型相关分析得到, 空气质量与气象指标间存在相关性, 秋冬季尤为显著, 且不同季节受到气象指标影响的空气质量指标不同。利用凝聚谱分析和EMD方法观察各频段数据的相关性可知, 春夏季相对湿度对PM_{2.5}的影响较大, 且在42天左右的周期分量上, 相对湿度的增大有利于降低PM_{2.5}的浓度。全段数据以及秋冬季为O₃与气温密切相关, 且随着分量周期的变长, 气温升高对的促进作用增强。

关键词

空气质量, 气象指标, 典型相关, 凝聚谱, 经验模态分解

Copyright © 2018 by authors and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

影响空气质量的因素有许多, 如污染源的排放、气象因素的影响等, 其中, 当地污染物的排放造成的大气污染为主要长期因素。但随着越来越多人的关注以及国家关于生态文明建设政策的颁布, 各项污染源的排放量得到了一定程度的控制, 对空气质量的影响也趋于稳定。而气象因素处于不断的变化中, 属于相对不稳定因素, 故可认为气象因素为影响空气质量的主要短期因素, 研究气象指标对空气质量的影响十分必要。

近年来, 国内外学者已针对空气污染状况与气象因素的关系展开了大量研究, 并结合了乌鲁木齐、哈尔滨、北京、广州等城市的数据, 以AQI、API或是可吸入颗粒物为空气质量指标, 从相关系数角度和污染发生频率角度分析了空气质量与气象指标的相关性[1]-[10], 也有学者利用典型相关分析建立空气质量和气象指标的综合典型变量进行相关性分析[11][12], 均得出了空气质量与气象指标之间相关的结论。但这些研究通常是基于原始数据的相关系数得到的结论, 虽然直观, 但并不能表示出气象指标与空气质量的内部关联, 也有学者利用小波分析对PM₁₀等空气质量指标进行了特征分析[13][14][15], 但仅仅分析了污染物的特征, 并未去与气象条件相联系。

对于两组变量间相关性的刻画有多种不同的方法, 如相关系数、灰色相关系数、散点图比较、典型相关分析以及信号分解方法。其中信号分解有利于观察变量内部的特征, 故利用谱分析的方法对数据各频段进行研究, 讨论其潜在的相关关系。本文拟以PM_{2.5}、PM₁₀、SO₂、CO、NO₂以及O₃为空气质量指标, 以气温、气压、相对湿度、降雨/雪量、能见度、平均风速以及最大持续风速为气象指标, 通过典型相关分析选取相关性最强的空气质量指标和气象指标, 利用凝聚谱分析选出相关性较强的频段区间, 并利用EMD分解选出相应频段的数据分量, 从频域的角度分析空气质量与气象指标间的相关性。

2. 资料与方法

2.1. 数据来源

山东省青岛市2014年1月~2017年12月空气质量指标的日均数据来自于PM_{2.5}历史数据网(<https://www.aqistudy.cn/historydata/>), 主要包括空气质量等级、AQI、PM_{2.5}、PM₁₀、SO₂、CO、NO₂、O₃等指标。同期的气象数据来源于天气网(<https://en.tutiempo.net/climate>), 数据包括平均气温、气压、相对湿度、降雨/降雪量、平均能见度、平均风速、最大持续风速等气象指标。

2.2. 青岛市空气质量基本情况

本根据《环境空气质量指数(AQI)技术规定(试行)》[15], 空气质量指数分为六级: 优、良、轻度污染、中度污染、重度污染、严重污染。统计各空气质量等级在 2014 年至 2017 年出现的天数占比可知, 良(65.85%)为青岛市最常见的空气质量情况, 其次是轻度污染(17.39%)和优(10.75%), 而中度污染(3.63%)、重度污染(2.19%)和严重污染(0.21%)出现的频率相对比较低。

逐月统计各空气质量等级出现的天数如图 1, 可知青岛市整体的空气质量较好, 每月的空气质量以“良”居多, 6~8 月空气质量较好, “优”出现的天数显著增多, 冬季的空气质量较差, “重度污染”和“严重污染”主要出现在 12 月和次年 1 月。这种现象一方面可归结为冬季采暖以及私家车出行增加带来的集中污染, 另一方面, 冬季气温较低, 不利于污染物扩散, 而夏季气温较高, 混合层高度提高, 有利于污染物的扩散, 且雨水较为丰沛, 有利于污染物溶入地表。

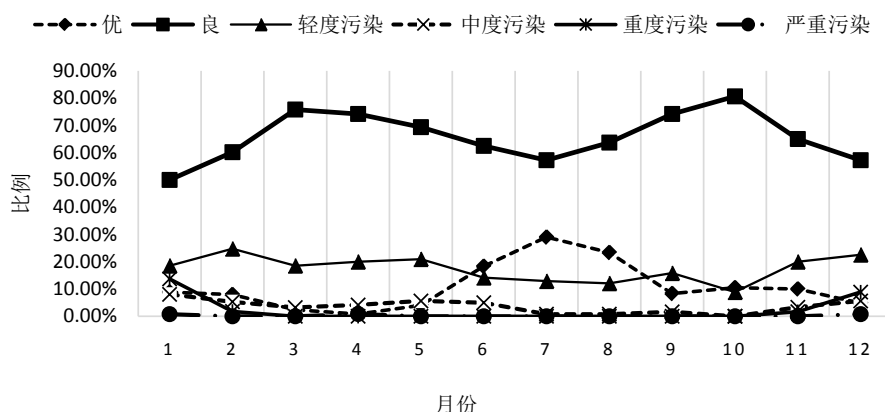


Figure 1. Air quality in Qingdao in every month
图 1. 青岛市各月的空气质量情况

2.3. 研究方法

典型相关分析[16]是利用综合变量之间的相关关系来反映两组指标之间的整体相关性, 分别对空气质量和气象指标内部变量进行线性组合构成综合变量, 通过综合变量间的相关系数大小, 选取若干具有代表性的综合变量, 进而可用综合变量对空气质量和气象指标间的关系进行分析。

功率谱分析[17] [18]可以从频域角度揭示一个平稳时间序列的内部结构, 凝聚谱在形式上类似于相关系数的计算, 利用凝聚谱可对数据进行谱分解, 并可计算空气质量和气象指标在各频段上的相关性, 进而选出显著相关的频率, 但无法输出数据在各频率上的分量。经验模态分解[19] [20] (Empirical Mode Decomposition, 简称 EMD)是一种循环迭代算法, 能够依据自身的时间尺度特征自适应地从时间序列中分解出一系列不同频率的本征模态函数(IMF)。利用 EMD 方法可以得到凝聚谱选出的频段上的数据分量, 进而对空气质量和气象因子的 IMF 分量进行相关性分析。

3. 数据分析结果

3.1. 典型变量

由于青岛市 2014 年 1 月~2017 年 12 月的气象指标数据存在少量缺失, 故在数据分析前需要对于缺失数据进行补足, 采取的方式为, 取该日前后两天的均值作为当天数据的补值。在进行相关性分析前, 对数据进行标准化处理, 可去除数据的量纲影响。

青岛市的四季[19]可大致划分为：春季 4~5 月份、夏季 6~9 月份、秋季 10~11 月份、冬季 12 月至次年 3 月份。青岛市的供暖时间为 11 月中旬至次年 3 月中旬，故可将 11 月至次年 3 月作为供暖季数据，4 月至 10 月作为非供暖季数据。

分别对青岛市的空气质量指标与气象指标的全段数据、四季数据以及供暖与否数据进行典型相关分析，观察其相关程度。由于第一对典型相关变量间的相关性最强，故取各段数据的第一对典型相关变量进行分析讨论，可得到表 1。

Table 1. Results of canonical correlation analysis in different periods of Qingdao
表 1. 青岛市不同时期的典型相关分析结果

空气质量		全段	春	夏	秋	冬	供暖	非供暖
气象指标								
空气质量指标	PM2.5	-0.84	6.53	-6.09	-0.42	1.90	1.35	-0.35
	PM10	0.60	-5.25	0.27	0.18	-1.98	-1.76	-2.34
	SO ₂	-0.87	-3.46	-1.29	4.61	0.45	0.17	-1.21
	CO	-0.19	2.50	-1.43	-0.09	-0.47	-0.04	0.17
	NO ₂	-0.79	-4.42	-1.14	3.68	-2.69	-2.91	-1.37
	O ₃	1.29	-3.59	-0.91	-2.79	-7.11	-5.61	-0.72
气象指标	T	2.36	-2.32	-1.34	-7.26	-9.56	-5.61	-0.07
	SLP	-0.02	2.26	-0.32	-1.64	-0.81	0.21	-0.17
	H	0.31	6.63	6.05	-5.58	1.72	2.72	4.21
	PP	-0.08	-0.20	-0.06	0.83	0.71	1.29	-0.07
	VV	0.71	1.96	5.11	-3.71	1.13	2.31	2.87
	V	0.80	0.51	0.61	-2.44	0.36	0.95	-0.22
	VM	0.04	-0.61	0.63	0.43	0.70	-0.80	0.41
	相关系数	0.78	0.77	0.79	0.81	0.80	0.74	0.79

注：T 为平均温度(°C)，SLP 为气压(hPa)，H 为相对湿度，PP 为降雨/降雪量(mm)，VV 为平均能见度(Km)，V 为平均风速(Km/h)，VM 为最大持续风速(Km/h)。

对于全段数据而言，臭氧受到温度的影响较大，且温度越高，臭氧浓度越高。不同时期，受到气象指标影响较大的空气质量指标并不相同，气象指标对于空气质量指标的影响程度也不相同，且在秋冬两季，空气质量与气象指标间的相关性更强。春夏两季相对湿度极大地影响了空气中 PM2.5 的含量，湿度越大会加剧春季 PM2.5 的污染，但却有利于降低夏季空气中 PM2.5 的浓度。秋季气温对 SO₂ 的影响较显著，且随着气温的降低，SO₂ 的污染加剧。冬季与供暖季仅相差一个月，故结果相似，且与全段数据的结果相仿，均为温度与臭氧的浓度成正比。非供暖季时期，受气象指标影响较大的是 PM10，而相对湿度为关键气象指标，且湿度的增大可以减轻 PM10 的污染。

综合考虑不同时期的典型相关结果，可知，PM2.5 和 O₃ 受到气象指标的影响最为显著，其中，相对湿度对 PM2.5 的浓度有制约作用，气温与 O₃ 的浓度密切相关。

3.2. PM2.5 与相对湿度的频域相关性分析

利用 ADF 检验对 PM2.5 和相对湿度进行平稳性检验如表 2，显然两个序列均为平稳序列，可进行凝聚谱分析。

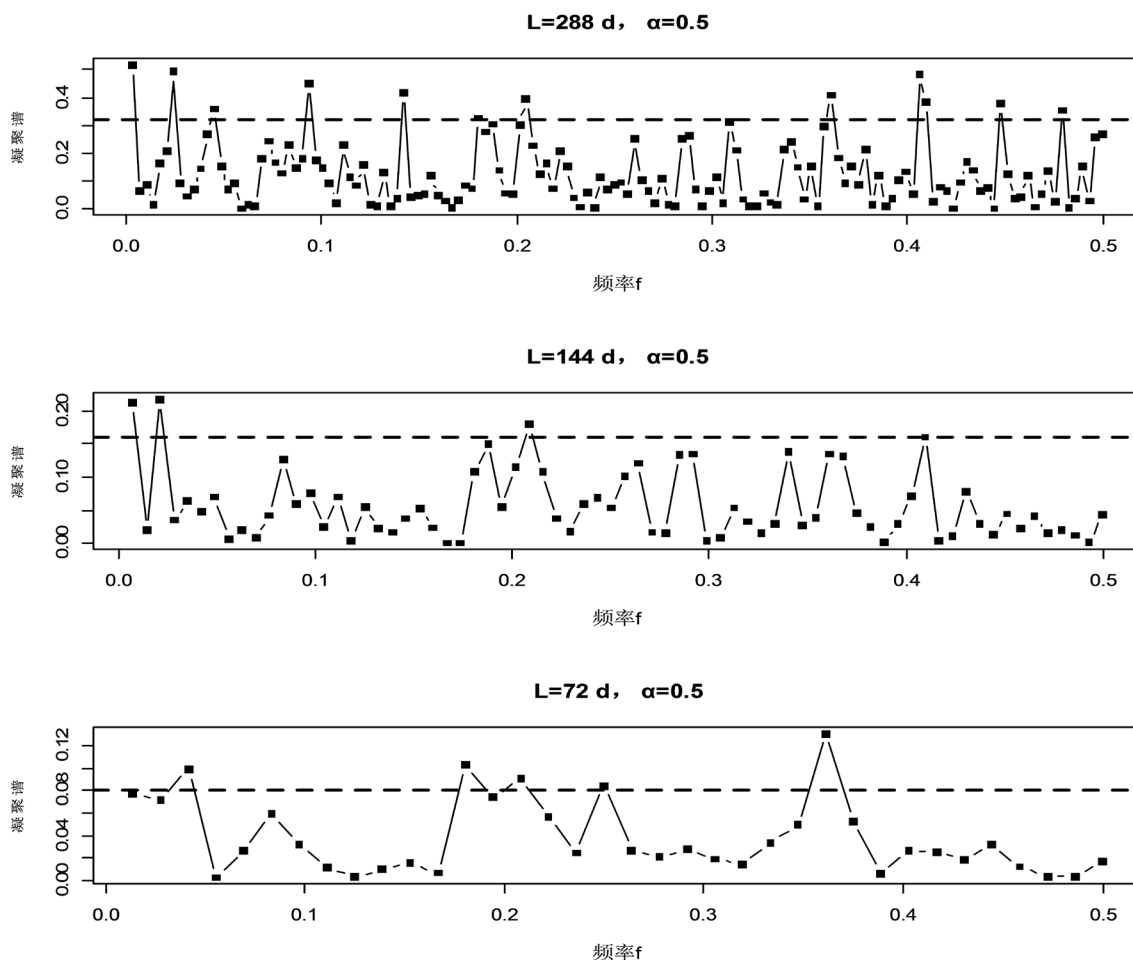
为便于计算，取自 2013 年 1 月 1 日起前 1440 行的数据进行凝聚谱分析，在 50%重叠率下，依次选择段长为 288、144 和 72，可得到各频段上的凝聚谱如图 2，图中垂直于 y 轴的虚线为凝聚谱在该段长下

的阈值[21], 凝聚谱超过阈值的频段为 PM2.5 与相对湿度相关性较强的频段。显然, 选择不同段长, 满足要求的频段数与频段位置不同, 但也存在相近部分, 频率在 0~0.05、0.2 左右的凝聚谱在三种段长下均超过了阈值, 故可认为 PM2.5 和相对湿度在这三种频率上的数据分量相关性较强。

Table 2. Results of ADF test for PM2.5 and relative humidity

表 2. PM2.5 和相对湿度的 ADF 检验结果

	Dickey-Fuller	Lag order	p-value	平稳性
PM 2.5	-5.5699	11	<0.01	平稳
相对湿度	-5.5699	11	<0.01	平稳



注: L 为数据分段的段长, α 为数据重叠率。

Figure 2. The MSC of PM2.5 and relative humidity under different lengths of three segments

图 2. PM2.5 和相对湿度在三种段长下各频段的凝聚谱

利用 EMD 方法对 PM2.5 和相对湿度进行本征模态函数的分解, 可得到各 IMF 分量的频率如表 3, 将凝聚谱超过阈值的频率与各模态的频率相互对应, 可选出 PM2.5 和相对湿度的部分 IMF 分量如表 4。

除了 IMF2 分量之外, 其他分量间的相关系数与原始数据相比均有提高, 且值得注意的是, 仅有 42 天周期上两变量间的关系表示了 PM2.5 与相对湿度负相关, 其他分量间均表示为两者正相关。这与典型相关分析时的结果相近, 春夏季相对湿度对 PM2.5 的影响不一致, 说明相对湿度对 PM2.5 的影响存在两

面性，即相对湿度增大并不一定能降低的浓度，且从周期来看，较长或较短周期上的相对湿度的增大会导致污染加剧，在 40 天左右的周期上相对湿度的增大可以降低的 PM2.5 浓度。

Table 3. The frequency of each IMF component of PM2.5 and relative humidity

表 3. PM2.5 和相对湿度各 IMF 分量的频率 f

	PM2.5	相对湿度		PM2.5	相对湿度
IMF1	0.593	0.595	IMF6	0.025	0.023
IMF2	0.292	0.284	IMF7	0.015	0.011
IMF3	0.149	0.151	IMF8	0.010	0.006
IMF4	0.077	0.078	IMF9	0.005	0.004
IMF5	0.044	0.042	IMF10	0.002	0.005

Table 4. The correlation of IMF components with similar periodicity of PM2.5 and relative humidity

表 4. PM2.5 和相对湿度相近周期的 IMF 分量的相关系数

PM2.5	相对湿度	周期 t (d)	相关系数
原始数据	原始数据	—	-0.08
IMF2	IMF2	4	0.03
IMF5	IMF5	23	0.18
IMF6	IMF6	42	-0.09
IMF8	IMF7	98	0.09
IMF9	IMF10	208	0.13

注：周期 $t = 1/f$ ，且由于 PM2.5 和相对湿度的各 IMF 分量的频率存在一些偏差，故周期取的是两个频率相近分量的周期平均值。

3.3. O₃ 与平均气温的频域相关性分析

对利用 ADF 检验对 O₃ 和平均气温进行平稳性检验如表 5，臭氧序列平稳，但平均气温序列并不平稳，故对平均气温序列进行差分，即可得到气温增量序列，经过 ADF 检验，可知气温增量序列为平稳序列，故可对 O₃ 和气温增量进行凝聚谱分析。

Table 5. ADF test results of O₃, mean temperature and air temperature increment

表 5. O₃、平均气温、气温增量的 ADF 检验结果

	Dickey-Fuller	Lag order	p-value	平稳性
O ₃	-4.1609	11	<0.01	平稳
平均气温	-1.7943	11	0.6654	不平稳
气温增量	-9.1984	10	<0.01	平稳

同 2.3，为便于计算，在 50%重叠率下，依次选择段长为 288、144 和 72，可得到各频段上的凝聚谱如图 3，图中垂直于 y 轴的虚线为凝聚谱在该段长下的阈值，凝聚谱超过阈值的频段即为 O₃ 与气温增量相关性较强的频段。当选择段长为 72 天时，几乎所有频段的凝聚谱均超过了阈值，随着段长的增大，频率的分割变细，发现尽管并不是所有的频段都满足要求，但 O₃ 和气温增量的凝聚谱超过阈值的频段涉及到了大部分频率区间。

利用 EMD 方法对 O₃ 和气温增量进行本征模态函数的分解，可得到各 IMF 分量的频率如表 6，将 O₃ 和平均气温各模态的频率一一对应，并与凝聚谱超过阈值的频率进行对照，可选出 O₃ 和气温增量的部分 IMF 分量如表 7。

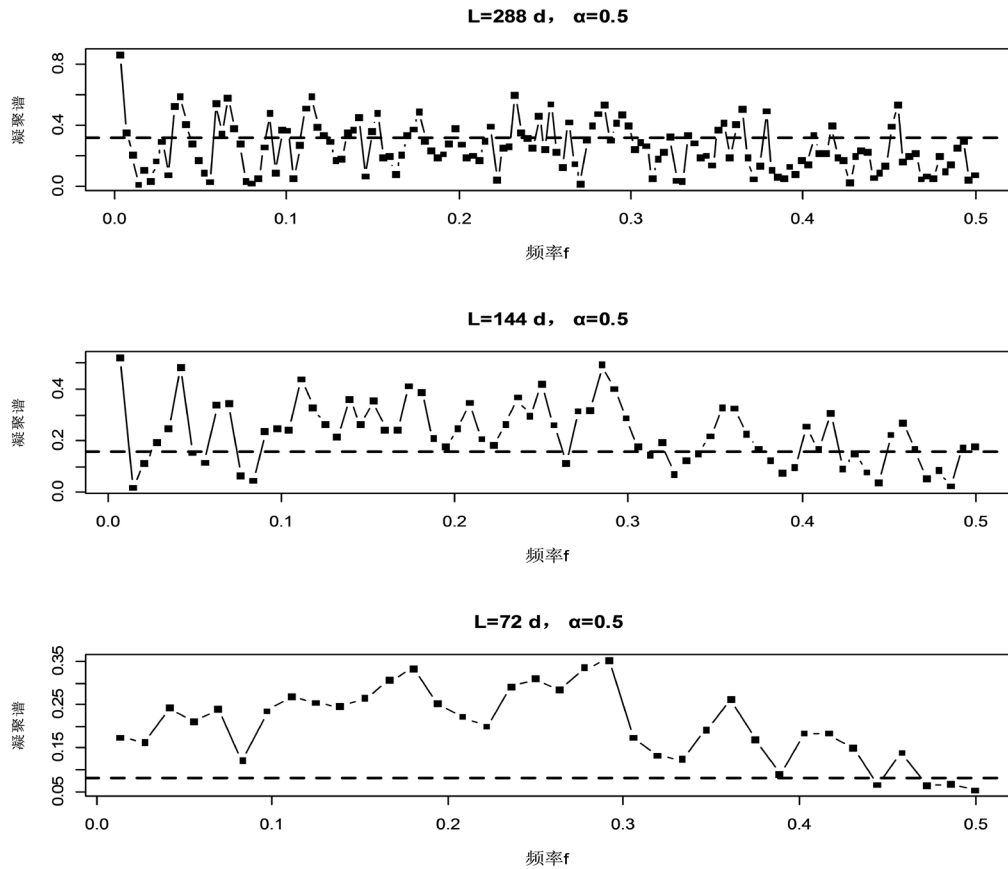


Figure 3. The MSC of O₃ and temperature increments under different lengths of three segments
图 3. O₃ 和气温增量在三种段长下各频段数据的凝聚谱

Table 6. Frequency of each IMF component of O₃ and temperature increment
表 6. O₃ 和气温增量各 IMF 分量的频率 f

	O ₃	气温增量		O ₃	气温增量
IMF1	0.604	0.645	IMF6	0.023	0.034
IMF2	0.303	0.329	IMF7	0.012	0.015
IMF3	0.158	0.188	IMF8	0.007	0.006
IMF4	0.082	0.101	IMF9	0.004	0.004
IMF5	0.047	0.060	IMF10	0.003	0.002

Table 7. Correlation coefficient of IMF with similar period component of O₃ and temperature increment
表 7. O₃ 和气温增量相近周期的 IMF 分量的相关系数

O ₃	气温增量	周期 t(d)	相关系数
原始数据	原始数据	—	0.63
IMF2	IMF2	3	-0.12
IMF3	IMF3	6	-0.04
IMF5	IMF5	19	-0.02
IMF7	IMF7	73	0.09
IMF8	IMF8	154	0.06
IMF9	IMF9	243	0.24
IMF10	IMF10	365	0.76

注：周期 $t = 1/f$ ，且由于 O₃ 和气温增量的各 IMF 分量的频率存在一些偏差，故周期取的是两个频率相近分量的周期平均值。

已知大部分频段的凝聚谱超过了阈值,可知气温增量在各个频段上均对 O_3 有较大的影响,但不同模式分量间的相关系数正负不一,且随着 IMF 分量周期的变长,相关系数由负变正。可推测,气温短时间的突发性升高可能会导致 O_3 浓度的降低,气温在长周期上的增长有利于 O_3 浓度的升高。在一年左右的周期上, O_3 与气温增量的相关性达到最大且超过原始数据的相关性,可认为气温增长以年度增长的幅度对空气质量造成较大影响。

4. 结论

论文利用计算频数频率的方法分析了青岛市 2014 年至 2017 年的空气质量整体状况,得到了青岛市空气质量整体较好,“良”为主要空气质量等级,空气质量的变化存在周期性,且污染发生的高峰期出现在每年的 12 月、1 月左右,推测这与冬季煤炭燃烧量增加有关,也可能冬季气温整体较低,不利于污染物扩散有关。

因空气质量和气象指标较多,故利用典型相关分析对青岛市不同时间段的空气质量和气象指标建立了综合变量,通过综合变量刻画两者间的相关关系。结果表明,尽管不同时间段上的空气质量和气象指标的相关模式不尽相同,但相关系数均可达到 0.7 以上,且均通过了相关性检验,即空气质量与气象指标间存在显著相关关系,且秋冬季相关性尤为显著。此外,不同季节受到气象指标显著影响的空气质量指标不完全相同,春夏两季以及非供暖季中,颗粒物污染与气象指标的关系密切,主要受到相对湿度的影响;秋冬两季与全年的结果相类似,臭氧受到气象因素的影响较大,且主要气象指标为气温。

为了对空气质量与气象指标的相关关系进行深入分析,利用凝聚谱和 EMD 方法选出 $PM_{2.5}$ 与相对湿度以及臭氧与气温间显著相关的频段,并分析相应的数据分量间的相关关系。结果表明,不同周期的相对湿度分量对 $PM_{2.5}$ 的影响不同,在 42 天的周期分量上,相对湿度有利于降低 $PM_{2.5}$ 的浓度,可认为湿度的月度变化对 $PM_{2.5}$ 的影响较大。秋冬两季与全年的结果相类似,均极大地受到了气温的影响,不同周期的气温分量对臭氧的影响也不相同,周期越长,正面影响越大,且气温的年度间变化对臭氧的影响十分显著。

基金项目

国家自然科学基金(U1706226)。

参考文献

- [1] 王露云. 中国 31 个主要城市空气质量评价及主要污染物浓度预测[D]: [硕士学位论文]. 重庆: 重庆师范大学, 2014.
- [2] 罗琳. 我国 31 个省会城市空气质量指数的分布特征及其与气象因素的关系[D]: [硕士学位论文]. 重庆: 重庆医科大学, 2017.
- [3] 邓伟年, 李刚. 2013 年郑州经济技术开发区环境空气质量状况分析[J]. 河南科技, 2015, 564(5): 138-140.
- [4] 赵敬国, 王式功, 王嘉媛, 闭建荣, 史晋森, 王天河, 张天宇. 兰州市空气污染与气象条件关系分析[J]. 兰州大学学报(自然科学版), 2013, 49(4): 45-52.
- [5] 徐丽, 李艳红. 乌鲁木齐市空气质量状况与常规气象条件的关系[J]. 农业灾害研究, 2013, 3(1): 41-44.
- [6] 刘宇飞. 哈尔滨市空气质量与气象条件的关系分析. 中国气象学会. 第 34 届中国气象学会年会 S1 灾害天气监测、分析与预报. 中国气象学会: 2017: 9 全国文献工作标准化技术委员会第六分委员会. CB6447-S6 文摘编写规则[S]. 北京: 标准出版社, 1986.
- [7] 杨旭, 康延臻, 王式功, 等. 郑州市大气污染特征及其气象条件的关系[J]. 兰州大学学报: 自然科学版, 2017, 53(3): 348-354.
- [8] 申占营, 熊杰伟, 陈东, 秦世广. 郑州市区 PM_{10} 污染状况及相关气象条件分析[J]. 河南气象, 2005(1): 28-29.
- [9] 连东英, 林长城, 王宏, 吴德辉. 三明市大气空气质量状况与气象条件的关系[J]. 环境与健康, 2008, 25(10):

880-884.

- [10] 胡琳, 曹红利, 张文静, 等. 西安市环境空气质量变化特征及其与气象条件的关系[J]. 气象与环境学报, 2013, 29(6): 150-153.
- [11] 沈家芬, 张凌, 莫测辉, 冯建军. 广州市空气污染物和气象要素的主成分与典型相关分析[J]. 生态环境, 2006, 15(5): 1018-1023.
- [12] 李晓红, 李万伟, 徐东群. 西安市空气质量与气象因素的典型相关分析[J]. 环境卫生学杂志, 2017, 7(3): 203-208.
- [13] Tian, G.J., Qiao, Z. and Xu, X.L. (2014) Characteristics of Particulate Matter (PM10) and Its Relationship with Meteorological Factors during 2001-2012 in Beijing. *Environmental Pollution*, **193**, 266-274.
- [14] Li, L., Qian, J., Ou, C.-Q., Zhou, Y.-X., Guo, C. and Guo, Y.M. (2014) Spatial and Temporal Analysis of Air Pollution Index and Its Timescale-Dependent Relationship with Meteorological Factors in Guangzhou, China, 2001-2011. *Environmental Pollution*, **190**, 75-81. <https://doi.org/10.1016/j.envpol.2014.03.020>
- [15] Tiwari, S., Dahiya, A. and Kumar, N. (2015) Investigation into Relationships among NO, NO₂, NO_x, O₃, and CO at an Urban Background Site in Delhi, India. *Atmospheric Research*, **157**, 119-126. <https://doi.org/10.1016/j.atmosres.2015.01.008>
- [16] 高惠璇. 应用多元统计分析[M]. 北京: 北京大学出版社, 2014.
- [17] 丁裕国, 江志红, 编. 气象数据时间序列信号处理[M]. 北京: 气象出版社, 1998.
- [18] 林振山, 邓自旺, 著. 子波气候诊断技术的研究[M]. 北京: 气象出版社, 1999.
- [19] 熊学军, 郭炳火, 胡筱敏, 刘建军. EMD方法和 Hilbert 谱分析法的应用于探讨[J]. 黄渤海海洋, 2002, 20(2): 13-21.
- [20] 方勇, 孔祥星, 著. 金融市场情绪演化机制与量化投资[M]. 上海: 上海财经大学出版社, 2016: 113-114.
- [21] 孙丰霖. 基于 Welch 方法的凝聚谱阈值估计问题及其应用[D]: [硕士学位论文]. 青岛: 中国海洋大学, 2017.

知网检索的两种方式:

1. 打开知网页面 <http://kns.cnki.net/kns/brief/result.aspx?dbPrefix=WWJD>
下拉列表框选择: [ISSN], 输入期刊 ISSN: 2168-5711, 即可查询
2. 打开知网首页 <http://cnki.net/>
左侧“国际文献总库”进入, 输入文章标题, 即可查询

投稿请点击: <http://www.hanspub.org/Submission.aspx>

期刊邮箱: ccrl@hanspub.org