

马小会,孙培源,闵晶晶,等.北京市海淀区腹泻疾病气象指数预报研究及应用[J].沙漠与绿洲气象,2014,8(1):70-73.  
doi:10.3969/j.issn.1002-0799.2014.01.014

# 北京市海淀区腹泻疾病气象指数 预报研究及应用

马小会<sup>1</sup>,孙培源<sup>2</sup>,闵晶晶<sup>1</sup>,宛公展<sup>3</sup>,张德山<sup>1</sup>

(1.北京市气象科技服务中心,北京 100089;2.北京市海淀区疾病预防控制中心,北京 100094;  
3.天津市气象科学研究所,天津 300081)

**摘要:**基于北京市海淀区2004—2007年夏季(6—9月)疾病预防控制中心的逐日腹泻疾病发病资料和同期地面气象观测资料,首先采用经典方差贡献分析技术分析了逐日发病人数与气象要素的相关性,选取方差贡献较大的要素作为预报因子,然后根据概率积分方法将发病人数划分为5个级别,采用回归方法分别建立了不同预报时效的预测模型,并检验历史样本的拟合结果。研究结果表明:(1)腹泻疾病发病人数与湿度、风速、气温均呈显著的正相关,其中,水汽压对腹泻疾病的方差贡献最大。(2)1~4 d的预报效果较好,其中1~2 d预报结果的完全正确率在40%以上,预报误差 $\leq \pm 1$ 级的准确率可达98%。

**关键词:**腹泻疾病;方差贡献;气象指数

**中图分类号:**P332.2

**文献标识码:**A

**文章编号:**1002-0799(2014)01-0070-04

北京夏季高温高湿,是肠道传染病高发季节。近些年,夏季的降水量呈明显下降趋势,主要是由于雨强减小造成<sup>[1]</sup>,这就造成了夏季高温高湿的环境增加,易引发肠道传染病。腹泻疾病是北京市海淀区夏、秋季节的常见流行性疾病或疫情<sup>[2-4]</sup>,因此,肠道传染病的防治、预测、控制工作显得非常重要。研究表明<sup>[2-7]</sup>,腹泻疾病的发生、发展及蔓延与环境气象条件密切相关。根据已有的历史病例和同期天气气候资料研究北京市海淀区腹泻疾病与天气气候的关系,建立腹泻病的预报模型,逐日滚动预报出未来1~2 d,1~3 d,……,1~7 d的腹泻疾病疫情的服务产品,为疾病预防控制中心和肠肛医院提供疾病可能流行的气象条件,而腹泻病预报成果在京津冀的推广,可提高百姓积极预防腹泻病疫情的能力和身体健康水平<sup>[8]</sup>。

## 1 资料搜集与因子相关分析

### 1.1 资料来源

本文采用的医疗数据来源于北京市海淀区疾病预防控制中心2004—2007年夏季(6—9月)腹泻疾病的逐日发病病例资料,监测天数为488 d(无缺测),发病人数共计33 089例。其中,4 a的病例涵盖人口为180~296万<sup>[9]</sup>。同期逐日气象资料由北京市气象台提供,包括9个地面基本气象要素(①平均气温、②极端最高气温、③极端最低气温、④平均相对湿度、⑤最小相对湿度、⑥平均水汽压、⑦日降水量、⑧平均风速、⑨平均海平面气压),因子样本数9×488个,没有缺测。

### 1.2 腹泻疾病发病的统计学特征

腹泻疾病一年四季都有发生,但以盛夏至秋季最甚,北京海淀区夏季逐日发病人数随时间的分布规律较明显(图1),根据2004—2007年夏季逐日发病资料统计分析了逐日发病人数的年平均分布,其中横坐标为研究日期的序号数,纵坐标为逐日每百万人口的平均发病人数(图中圆点曲线为腹泻病人数,光滑曲线为腹泻病谐波值)。发病人数在6—7月缓慢增长,到8月上旬快速增加,中旬以后逐渐下降。逐日发病人数的年平均统计参数值详见表1,统

收稿日期:2013-07-22;修回日期:2013-08-25

基金项目:国家公益性行业(气象)专项项目(GYHY201106034)和国家人口与健康科学数据共享平台“气象环境与健康”专题服务建设项目共同资助。

作者简介:马小会(1979-),女,工程师,现从事气象服务工作。

E-mail:Salinamxh@sohu.com

计结果表明,日平均发病人数每百万人口约为 13.4,中位数为 8.4,由于中位数远小于平均值,概率分布显然不属于正态分布类型。

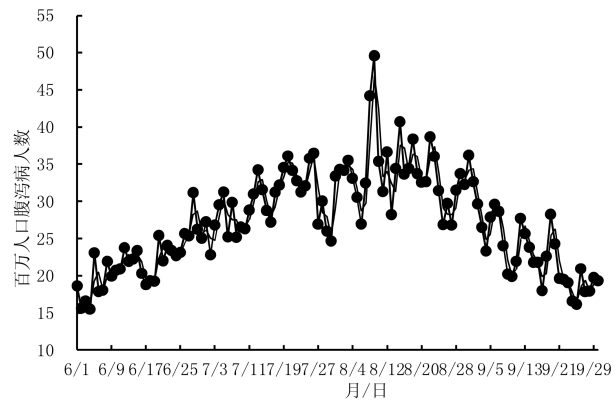


图 1 2004—2007 百万人口腹泻疾病发病年度变化曲线

表 1 北京市海淀区夏季腹泻疾病逐日发病的统计特征值(单位:人数/每百万人口)

| 特征分类 | 最大值        | 次大值        | 第三名        | 最小值        | 平均值  | 标准差 |
|------|------------|------------|------------|------------|------|-----|
| 统计参数 | 72.2       | 61.4       | 60.3       | 6.3        | 27.2 | 6.6 |
| 出现日期 | 2006-08-09 | 2006-08-08 | 2004-07-26 | 2006-06-03 |      |     |

1.3 腹泻疾病发病与环境气象因子的相关分析

所谓环境气象因子是指根据北京市观象台 9 个地面基本气象要素按一定方法加工派生而成 [10-11]。其中,因子 1、4、5、6、7、8、9 是因子源值的等权或半衰权重的平均值。因子 2 和 10 是过去 7 d 逐日的气温、气压变化倾向值,其计算方法是将因子源值先用直线方程拟合,求解得出斜率值再乘以方程相关系数的绝对值。因子 3 是日较差。为了分析腹泻疾病发病与环境气象因子之间的相关关系,采用经典方差贡献分析技术。现将部分计算结果列于表 2。因子  $X_i$  的方差贡献值  $PV_i$ ,即因子的偏回归平方和,是因子重要性的度量方法之一,常用计算式为:

$$PV_i = L_{yy}^{-1} B_i^2 / C_{ii}, \tag{1}$$

公式(1)中,  $B_i^2$  为因子  $X_i$  的偏回归平方和,  $L_{yy}$  为腹泻疾病发病的原方差,  $C_{ii}$  为因子协方差逆矩阵的对角线元素,  $i=1,2,3,\cdots,m, m=10, y=1$ 。为便于不同因子相互比较,表 2 中已将其处理为百分比。

根据表 2 中不同气象因子的方差贡献结果表明:在 24~168 h 的预报中,平均水汽压在腹泻病发病的过程中贡献最大,方差贡献值达到 48%以上;平均风速次之;日平均气温和平均相对湿度的方差贡献也较为明显。分析中所有因子的复相关系数平

表 2 不同预报时效中气象因子与腹泻疾病发病的方差贡献/%

| 气象因子                                                     | 24 h | 48 h | 72 h | 96 h | 120 h | 144 h | 168 h |
|----------------------------------------------------------|------|------|------|------|-------|-------|-------|
| 1.日平均气温/ $^{\circ}\text{C}$                              | 7.5  | 7.3  | 9.2  | 10.8 | 9.4   | 7.1   | 4.6   |
| 2.气温变化倾向 $\Delta T/^{\circ}\text{C}$                     | 0.4  | 0.1  | 0    | 0.1  | 0.2   | 0.5   | 0.3   |
| 3.气温日较差 $T_{\text{max}}-T_{\text{min}}/^{\circ}\text{C}$ | 0.1  | 0.2  | 0.5  | 0.5  | 0.8   | 1.7   | 2.1   |
| 4.平均相对湿度 $TH/\%$                                         | 8.2  | 9.5  | 10.1 | 10.7 | 10.4  | 11.1  | 12.8  |
| 5.最小相对湿度 $R_{\text{min}}/\%$                             | 9.1  | 8.2  | 8.1  | 7.4  | 6.4   | 5.1   | 3.8   |
| 6.平均水汽压 $V_p/\text{hPa}$                                 | 48.8 | 47.9 | 48.2 | 48.8 | 49.1  | 48.7  | 49.1  |
| 7.日降水量 $P_{\text{24}}/\text{mm}$                         | 0.1  | 0.1  | 0.3  | 0.5  | 0.7   | 0.6   | 0.2   |
| 8.平均风速 $W_{\text{avg}}/\text{m}\cdot\text{s}$            | 25.9 | 26.4 | 22.8 | 20.3 | 21.8  | 24.1  | 26.3  |
| 9.平均海平面气压 $AP/\text{hPa}$                                | 0.1  | 0.1  | 0.2  | 0.4  | 0.8   | 0.8   | 0.6   |
| 10.气压变化倾向 $\Delta H_{\text{ps}}$                         | 0.1  | 0.2  | 0.7  | 0.6  | 0.3   | 0.2   | 0.1   |

均可达 0.766 4,统计量  $F$  值高达 76.07,远远超过了通过  $F$  检验所需的下限阈值( $\alpha=0.001, F_{\alpha}=2.96$ )。研究结果表明腹泻疾病发病与气象因子群之间是少见的高度相关,利用环境气象因子的变化预测海淀区未来 24~168 h 腹泻疾病发病趋势是完全可行的。

根据方差贡献分析看出,湿度、风速、气温对腹泻疾病的发病有很大影响,呈明显正相关。湿热季节环境非常有利于腹泻病菌的滋生、繁殖和疫情传播。夏季的闷热(高温高湿)天气条件,往往会加剧人体的不舒服感 [12-13],此季节人们也喜欢大量食用一些生、冷食品,从而使胃肠道防御功能降低导致腹泻的发病流行。

2 建模思路与方法

腹泻疾病气象预报采用 5 档分级的方法,即按概率密度函数标准将疾病发病人数分级。预报模型主要通过以下两个步骤完成。

2.1 组建逐日发病人数滚动预报的多元回归模型

应用表 2 中的 10 个预报因子,分别与未来 7 个时效的平均腹泻病病例进行多元统计分析,建立滚动预报模型为:

$$Y = B_0 + \sum B_j \times X_j, \tag{2}$$

(2)式中,  $Y$  表示每百万人口的平均日发病人数的回归拟合值,  $X_j$  表示进入方程的预报因子,  $B_j$  ( $j=1,2,\cdots,m$ ) 表示相应的回归系数,  $B_0$  表示常数系数。

列出未来 24~168 h 腹泻疾病的回归预测模型:

24 h 预报:  $Y_{24}=128.031\ 8-2.852\ 9X_1+0.959\ 0X_2-0.102\ 9X_3-0.988\ 4X_4-0.743\ 5X_5+6.581\ 6X_6-0.044\ 8X_7-14.165\ 0X_8-0.001\ 4X_9-0.081\ 6X_{10}$  (3)

48 h 预报:  $Y_{48}=69.977\ 7-1.137\ 0X_1+0.217\ 0X_2-$

$$0.076\ 5X_3-0.428\ 3X_4-0.302\ 0X_5+2.686\ 8X_6-0.034\ 1X_7-6.024\ 2X_8-0.013\ 7X_9-0.174\ 6X_{10} \quad (4)$$

$$72\text{ h 预报: }Y_{72}=102.345\ 4-1.310\ 0X_1-0.049\ 1X_2-0.1092X_3-0.4491X_4-3.208X_5+2.8336X_6-0.0672X_7-5.9040X_8-0.0421X_9-0.3838X_{10} \quad (5)$$

$$96\text{ h 预报: }Y_{96}=128.8607-1.4071X_1-0.1789X_2-0.113\ 6X_3-0.4640X_4-0.321\ 2X_5+2.912\ 2X_6-0.093\ 5X_7-5.779\ 0X_8-0.066\ 9X_9-0.345\ 6X_{10} \quad (6)$$

$$120\text{ h 预报: }Y_{120}=143.0723-1.1817X_1-0.3374X_2-0.1446X_3-0.4250X_4-0.2850X_5+2.6780X_6-0.1135X_7-5.7155X_8-0.0856X_9-0.2487X_{10} \quad (7)$$

$$144\text{ h 预报: }Y_{144}=137.463\ 4-0.919\ 0X_1-0.457\ 4X_2-0.193\ 5X_3-0.407\ 3X_4-0.236\ 4X_5+2.433\ 3X_6-0.097\ 5X_7-5.643\ 3X_8-0.084\ 4X_9-0.1755X_{10} \quad (8)$$

$$168\text{ h 预报: }Y_{168}=111.878\ 7-0.650\ 6X_1-0.345\ 4X_2-0.201\ 1X_3-0.393\ 2X_4-0.185\ 3X_5+2.201\ 4X_6-0.056\ 6X_7-5.400\ 0X_8-0.064\ 6X_9-0.137\ 3X_{10} \quad (9)$$

(3)~(9)式中,  $Y_{24}, Y_{48}, Y_{72}, \dots, Y_{144}, Y_{168}$  分别表示腹泻疾病的未来 24~48, 24~72, 24~48, …… , 24~144, 24~168 h 百万人口发病人数的预报结果,  $X_1, X_2, X_3, \dots, X_{10}$  依次对应表 2 中的对应时次的 10 个气象因子。

## 2.2 腹泻病气象指数预报分级标准

采用经验概率密度分级方法, 将腹泻疾病气象指数划分为医学常用的概率密度分布函数: 0.035 94、0.238 31、0.451 5、0.238 31、0.035 94。5 个级别依次表示为: 5 级, 发病人数很多; 4 级, 发病人数较多; 3 级, 发病人数偏多; 2 级, 发病人数较少; 1 级, 发病人数很少。按照腹泻疾病病例的概率区间划分出腹泻疾病气象指数等级标准<sup>[14]</sup>(表 3)。由表 3 可知, 腹泻病等级越高, 腹泻病病例越多。

表 3 腹泻病气象指数概率分级标准

| 气象等级    | 1           | 2         | 3         | 4         | 5           |
|---------|-------------|-----------|-----------|-----------|-------------|
| 概率分布/%  | 3.6         | 23.8      | 45.2      | 23.8      | 3.6         |
| 概率区间/%  | $\geq 96.4$ | 72.6~96.4 | 27.4~72.6 | 3.6~27.4  | $\leq 3.6$  |
| 腹泻病病例等级 | $\leq 12.5$ | 12.6~20.0 | 20.1~34.3 | 34.4~41.8 | $\geq 41.9$ |

## 3 腹泻病气象指数预报的历史拟合准确率分析

### 3.1 历史拟合率检验效果

基于表 2 中选取的气象因子和预报模型公式(2), 利用 2004—2007 年夏季的总样本分别建立未来 1~7 d 的百万人口的腹泻病人数预报模型, 然后将历史资料进行回代, 得到的预报拟合结果用表 3 列出的标准进行分级, 然后分别计算不同预报时效

的历史拟合准确率。拟合结果见表 4。

根据表 4 的统计分析结果得知: 预测模型对腹泻病气象指数的预报效果随着预报时效的增长而变差, 其中 1~2 d 的预报效果较好, 完全正确拟合率在 40% 以上; 而 3~4 d 的拟合效果稍差, 完全正确拟合率在 30% 以上, 也具有一定的可信度。如果腹泻病气象指数预报等级相差  $\pm 1$  级, 未来 7 个预报时效的拟合率均可达  $\geq 97.5\%$ 。

表 4 预报未来 7 个时效的腹泻病气象指数拟合率

| 预报时效/d | 完全正确/% | 误差 $\pm 1$ 级/% | 误差 $\leq 1$ 级/% | 误差 $> 1$ 级/% |
|--------|--------|----------------|-----------------|--------------|
| 1      | 42.8   | 53.1           | 95.9            | 4.1          |
| 2      | 40.9   | 56.6           | 97.5            | 2.5          |
| 3      | 35.5   | 62.1           | 97.6            | 2.4          |
| 4      | 31.2   | 67.2           | 98.4            | 1.6          |
| 5      | 28.7   | 70.2           | 98.9            | 1.1          |
| 6      | 30.4   | 68.3           | 98.7            | 1.3          |
| 7      | 32.5   | 65.6           | 98.1            | 1.9          |

### 3.2 2009—2011 年业务应用结果分析

以预报未来 120 h 腹泻病为例, 给出 2009—2011 年逐日预报腹泻病结果(图 2)。2009 年(圆点曲线)和 2011 年(三角形曲线)的腹泻病预报趋势相近, 2009 年 6—7 月和 9 月上旬腹泻病预报值低于 2011 年。2010 年的腹泻病预报值最大, 其中高峰值 43 人, 出现在 2010 年 8 月 1 日。3 a 夏季总的预报趋势与图 1 非常近似。

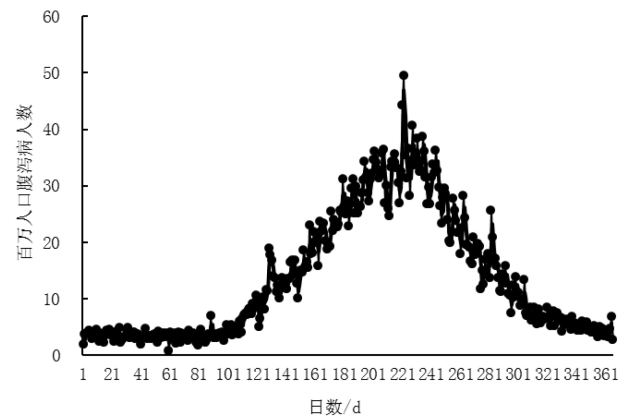


图 2 2009—2011 年逐日滚动预报未来 120 h 腹泻病人数曲线分布

## 4 结语

腹泻疾病发病与气象因素关系密切。用于预报腹泻病的 10 个预报气象因子中, 贡献最大的是日平均水汽压(贡献率占 47.9%~49.1%), 其次是平均风速(贡献率占 20.3%~26.4%)。

建立了北京市海淀区夏季未来 24~48 h、24~72

h,……,24~144 h,24~168 h 预报模型,历史拟合率达 90%以上(误差 $\leq 1$  级统计)。

研究成果已经应用到北京市专业气象台“医疗气象和公共安全系统”中,为北京市预防腹泻病提供了科学依据。在京津冀具有很好的实际推广价值,但是仍然需要继续积累腹泻病的病例样本,不断完善腹泻病的预报模型,使其成果具有稳定的规律性。

#### 参考文献:

- [1] 尤焕苓,任国玉,刘伟东.1961—2010 年北京地区降水变化特征[J].沙漠与绿洲气象,2012,6(4):13-20
- [2] Hayashi T. Impact of meteorological elements on diarrhea diseases in Bangladesh [J]. Preprints Meteor Soc, 2004, 86: 257.
- [3] 谢旭,任金马.深圳市肠道传染病与气象及媒介因素的贝叶斯正规化 BP 神经网络模型 [J]. 医学动物防治, 2007, 23(7): 485-488.
- [4] 高婷,刘桂荣,黎新宇,等.北京市近 14 年细菌性痢疾的疫情分析[J].中华预防医学杂志,2007,41(1):54-57.
- [5] 黄成钢,金如锋.我国某地区痢疾发病率与气象因素的关系及其预测模型[J].现代预防医学,2009,36(7):1207-1210.
- [6] 谈荣梅,陈坤,屠春雨.气象因素与细菌性痢疾发病关系的探讨[J].浙江预防医学,2003,15(3):7-9.
- [7] Reena BK S, Simon H, Neil de W, et al. The Influence of Climate Variation and Change on Diarrheal Disease in the Pacific Islands. Environ Health Perspect [J]. 2001, 109: 155-159.
- [8] 孙凤华,胡伟,白乐生.辽宁省气象与健康城市专业气象服务业务系统[J].气象与环境学报,1999,15(3):48-50.
- [9] 北京统计局.北京统计年鉴,2004—2006,北京:统计出版社,2002—2005.
- [10] 屠其璞,王俊德,丁玉国,等.气象应用概率统计学[M].北京:气象出版社,1984.
- [11] 宛公展,王宝成,弥明祥.PC-1500 袖珍计算机 BASIC 语言应用程序集[M].北京:气象出版社,1988,132.
- [12] 张德山,邓长菊,尤焕苓,等.海淀区中暑气象指数预报与服务[J].气象科技,2005,33(6):574-575.
- [13] 于杰,孟巧玲.安阳市 1990—2003 年肠道传染病疫情分析[J].现代预防医学,2005,32(9):1121-1122.
- [14] 张姝丽,张德山,冯涛,等.北京地区细菌性痢疾医疗气象预报模式及流行风险水平评估[J].气候与环境研究, 2009, 14(6): 561-566.

## Research and Application of Meteorological Index Forecast of Diarrhea Disease in Haidian District , Beijing

MA Xiaohui<sup>1</sup>, SUN Peiyuan<sup>2</sup>, MIN Jingjing<sup>1</sup>, WAN Gongzhan<sup>3</sup>, ZHANG Shuli<sup>1</sup>

(1.Beijing Meteorological Science and Technology Service Centre, Beijing 100089, China;

2.Haidian Center for Disease Prevention and Control, Beijing 100094, China;

3.Tianjin Meteorological Institute, Tianjing 300081, China)

**Abstract** Based on the synchronization data of diarrhea disease and the meteorological observation in Haidian District of Beijing from June to September during 2004–2007, the relationship between the incidence of diarrhea disease and weather conditions was analyzed in terms of Classical Variance Contribution Analysis technology. The meteorological factors which have significant variance contribution were selected as forecast factors. The daily incident number of disease was divided into five grades by using multiple regression probability grade analysis, and the forecast models for different forecast periods were established with Regression Analysis method and were examined with historical independent sample tests. The results show as follows. (1) There was significant correlation between incident number of diarrhea disease and meteorological condition, especially for the factors of humidity, wind speed and temperature, and the proportion of variance contribution of vapor pressure was the highest. (2) The forecasting for 1~4 days had a good effect, and the exact correct rate for 1~2 days forecast was more than 40%, if the difference between the forecasted grade and actual grade was no more than 1, the mean exact correct rate was 98%.

**Key words** diarrhea disease; variance contribution; meteorological index