

闵晶晶,丁德平,李津,等. 2014. 北京急性脑血管疾病与气象要素的关系及预测. 气象, 40(1):108-113.

# 北京急性脑血管疾病与气象要素的关系及预测<sup>\* 1</sup>

闵晶晶<sup>1</sup> 丁德平<sup>1</sup> 李津<sup>1</sup> 张德山<sup>1</sup> 彭丽<sup>2,3</sup>

1 北京市气象服务中心,北京 100089

2 上海市城市环境气象中心,上海 200135

3 上海市气象与健康重点实验室,上海 200030

**提 要:** 基于 2006 年 1 月至 2010 年 12 月北京市 120 急救中心的逐日脑血管急症接诊病例数据资料,首先探讨北京市急性脑血管疾病与气象要素的关系,选取不同季节的影响因子,然后根据概率积分方法将发病人数划分为 4 个级别,并采用人工神经网络方法(artificial neural network, ANN)分别建立了北京市不同季节的急性脑血管疾病预测模型。研究表明:(1)急性脑血管疾病发病人数存在明显的季节性变化和日变化特征,冬春季发病人数高于夏、秋季,发病主要集中在早晨到中午的 09—14 时;(2)发病人数相对于气象要素存在明显的滞后效应,夏和冬秋季发病分别与高温高湿、冷空气活动有关;(3)脑血管疾病预测模型通过对新样本进行预报,除夏季外,完全准确率高于 30%,预报误差 $\leq \pm 1$  级的准确率高于 60%,研究成果对于预防急性脑血管疾病发病和调度 120 急救车辆等应急措施具有较好的科学参考价值。

**关键词:** 脑血管疾病,气象要素,预测模型,人工神经网络方法

**中图分类号:** P49

**文献标志码:** A

**doi:** 10.7519/j.issn.1000-0526.2014.01.013

## Relationship Between Acute Cerebrovascular Disease and Meteorological Factors in Beijing and Its Forecast

MIN Jingjing<sup>1</sup> DING Deping<sup>1</sup> LI Jin<sup>1</sup> ZHANG Deshan<sup>1</sup> PENG Li<sup>2,3</sup>

1 Beijing Meteorological Service Center, Beijing 100089

2 Shanghai Center for Urban Environmental Meteorology, Shanghai 200135

3 Shanghai Key Laboratory of Meteorology and Health, Shanghai 200030

**Abstract:** Based on daily emergency case data of cerebrovascular disease in Beijing during 2006—2010, which are obtained from the Beijing urgent care centre (120). The association between the daily meteorological factors and hospital emergency visits for cerebrovascular disease in different seasons in Beijing is explored. Then we choose meteorological factors with regression method so as to obtain the forecast factors which are finally used to build forecast models in different seasons based on the ANN (artificial neural network) method, and the daily hospital visit numbers are divided into four grades by using multiple regression probability grade analysis. The results show that: (1) There are obvious seasonal and diurnal variations in the number of acute cerebrovascular disease, the number of cases is significantly higher in spring and winter than in summer and autumn, and concentrate mainly in 09:00—14:00 BT. (2) The meteorological factors have obvious hysteresis to induce the recurrence of cerebrovascular disease. In addition, statistical results show that the condition of high temperature and humidity weather in summer or cold air activity in winter and autumn may aggravate disease. (3) The models in spring, autumn and winter are used

\* 上海市科学技术委员会 2011 新建上海市重点实验室项目号(11DZ2260900)、公益性行业(气象)科研专项(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

2012 年 12 月 10 日收稿; 2013 年 3 月 18 日收修定稿

第一作者: 闵晶晶, 主要从事数值预报产品释用和应用气象研究. Email: minjj06@163.com

to forecast daily disease grade of new samples, the test results show the complete accuracy exceeds 30%; If the difference between the forecasted grade and actual grade is no more than 1, the accuracy exceeds 60%. The research results offer scientific reference for preventing the development of cerebrovascular disease and scheduling such emergency measures as 120 emergency vehicles.

**Key words:** cerebrovascular disease, meteorological factor, forecast model, artificial neural network (ANN)

## 引言

据统计(路凤等,2008),世界上三分之一的人患有心脑血管疾病,每年有1500万人被心脑血管疾病夺取生命,占死亡人数的60%以上,而我国每年因心脑血管疾病造成死亡的人数约260万(刘方等,2004)。近年来卫生部调查资料显示,呼吸道与心脑血管疾病的发病率位于前列,并有逐年升高的趋势,而脑血管疾病死亡率跃居首位。脑血管疾病的成因包括环境、情绪、生活习惯等多种因素,临床观测证明,在气象条件变化较大时,急性脑血管疾病人数增多。目前国内外许多研究也已证实(Shin et al,1990;卢爱梅等,1997;刘世玲等,1999),气象条件变化是脑血管疾病发病和死亡的主要诱因之一。

脑血管疾病的发生具有一定的确定性和随机性,根据气象条件变化对脑血管疾病的发生进行预测具有确定性和不确定性的双重特性。目前,国内针对脑血管疾病的预报研究多采用多元逐步回归(卢爱梅等,1997)、气象因子综合指标预报方法(张书余等,2010)、支持向量机(support vector machine,SVM)回归方法(毛清宇等,2010)和自动交互检测方法(automation interaction detector,AID)(韩建康等,2008)等,上述方法一般是根据气象因子和就诊人数或住院人数的统计关系建立预测模型。显然,根据传统的统计方法很难解决脑血管疾病预测中存在的多重性,而人工神经网络(artificial neural network,ANN)由于其具有自学习、自组织的非线性映射能力,适合于一些信息复杂、知识背景不清楚和推理规则不明确问题的建模。利用人工神经网络的自组织学习能力,可以从影响疾病的气象要素中发掘有用的信息,从而识别出发病人数与影响因子之间的非线性映射关系。由于人工神经网络模型对非线性样本具有比传统回归模型更好、更强的拟合和预报能力,用于受多种气象要素综合影响的医疗气象预报非常可行有效。因此本文采用一种改进

的ANN方法来建立气象条件变化与急性脑血管疾病发生的预测模型。

## 1 资料和方法

### 1.1 资料来源

收集北京市120急救中心2006年1月至2010年12月的急性脑血管疾病接诊病例9005例,其中脑出血3481例,脑梗塞5524例。病例信息包括年龄、性别、接诊时间和临床诊断。

同期逐日气象资料来源于北京市气象局,包括9个地面基本气象要素:日平均气温(单位:℃)、水汽压(单位:hPa)、相对湿度(单位:%)、风速(单位: $m \cdot s^{-1}$ )、海平面气压(单位:hPa)、日照时间(单位:h)、最高(最低)气温(单位:℃)和最低相对湿度(单位:%)。

### 1.2 资料处理

由于转折性天气对急性脑血管疾病有较大影响,为了分析它们之间的关系,根据气象台提供的9个基本要素,进一步加工派生出24h变温、24h变压和气温日较差3个因子。因气象条件对疾病的影响存在连续性和明显的滞后性,选取发病人数与各要素的时间派生因子( $X, X_{-1}, X_{-2}, X_{-3}, X - X_{-1}, X - X_{-2}, X - X_{-3}$ )分别进行相关分析。其中, $X$ 表示当天要素, $X_{-1}$ 表示前1天要素, $X_{-2}$ 前2天要素, $X_{-3}$ 前3天要素。进行研究分析的基本和派生气象因子共计 $12 \times 7$ 个。

### 1.3 人工神经网络方法

目前最常用的人工神经网络方法是BP算法,它对各种复杂信息的处理能力非常强。BP(Back Propagation)网络,是一种单向传播的多层前向网络,训练过程分为两个阶段:正向过程和反向过程。正向过程是把学习样本的输入值循环置入BP网络

的输入端,由前往后,依次计算网络的隐节点和输出节点的值;根据网络的输出值与期望输出值之间的误差,开始反向计算,即由后往前,按照训练的目标函数,依次调整网络各层节点之间的连接权值,直到目标函数接近极小值为止(董全等,2013;曾晓青,2013)。

在 BP 网络的学习算法中,权值修正一般采用梯度下降法,各节点神经元的激励函数常采用 Sigmoid 函数,其形式为:

$$f(x) = \frac{1}{1 + e^{-x}}$$

然而,传统 BP 算法在应用中存在许多亟待解决的问题,例如,网络训练过程中易陷入局部极小等,实际应用存在网络结构参数和学习训练参数难以确定的问题。为了方便实际应用,本文主要采用一种改进的 BP 算法进行预测模型的建立,该算法能有效解决传统算法中存在的问题,具体算法可参考文献(闵晶晶等,2010)。

## 2 结果与讨论

### 2.1 急性脑血管疾病发病特征

#### 2.1.1 发病人数的时间变化特征

北京地区急性脑血管 5 年总发病人数以冬季(12月至次年2月)最多,为 2470 人(27.4%);春季(3—5月)次之,为 2315 人(25.7%);秋季(9—11月)2176 人(24.2%);夏季(6—8月)最少,为 2044 人(22.7%)。这和南京(毛清宇等,2010)、吉林(张书余等,2010)等地区冬春季发病高于夏秋季的结论基本一致。这是因为冬季和春季温度较低,强冷空气容易造成气温骤降,气压升高,湿度降低,研究表明(刘世玲等,1999;叶殿秀等,2003),这些气象要素的变化是促使脑血管发病的有利条件。春、秋季是季节转换期,气温等要素变化较大,对自然环境变化适应性能力较差的群体,容易出现脑血管疾病的加重或复发。冬季,当冷锋过境或过境后,可使交感神经兴奋,引起血管收缩以及血管平滑肌痉挛,血管阻力增加,血压升高,儿茶酚胺分泌液增多,使血小板形成血栓,从而诱发脑血管疾病的加重或发生。从各月份看,1 和 11 月是全年发病最多的 2 个月,分别占总数的 9.78% 和 9.28%;9 和 6 月则是发病最少的 2 个月,分别占总数的 6.97% 和 7.20%。

急性脑血管疾病发病人数存在明显的日变化(图 1),发病时段主要集中在早晨和中午的 09:00—14:00,为 4829 人,占总人数的 53.6%,尤其在日升温最快的 09:00—11:00,急性脑血管疾病更易发生,这主要是与 24 h 血压位于 10:00、上午交感神经张力增高和外界刺激与体力活动增多有关(刘涛等,2006)。而在夜间到凌晨(23:00—05:00),发病人数较少且变化幅度不大,可能由于该时段为休息时间,受外界环境影响较小,不利于诱发脑血管疾病的发生。

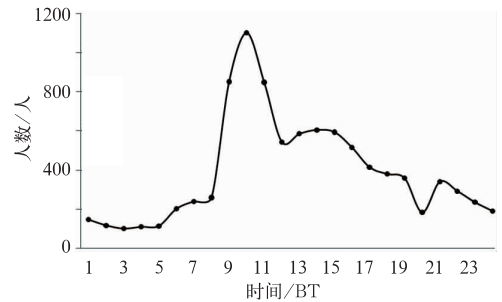


图 1 2006—2010 年北京急性脑血管发病人数的日变化

Fig. 1 The diurnal variation of acute cerebrovascular disease number in Beijing during 2006—2010

#### 2.1.2 发病人数的年龄与性别分布

2006—2010 年北京市急诊病例共 9005 例,其中男性 5267 例,女性 3738 例,男性较女性更容易出现急性脑血管疾病。

由 2006—2010 年不同性别的发病年龄分布曲线(图 2)得知,从小于 10 岁的儿童到大于 100 岁的老人,各年龄段均有急性脑血管疾病出现,其中病人

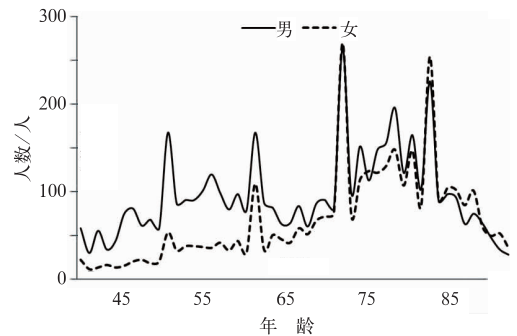


图 2 2006—2010 年北京急性脑血管发病人数的年龄分布

Fig. 2 The age distribution of acute cerebrovascular disease number in Beijing during 2006—2010

最小年龄 2 天,最大年龄 101 岁。40 岁之前发病人数较少,40 岁以后发病人数随着年龄增大而迅速增加,到 70—79 岁达到高峰值。同时,80 岁以下各年龄段男性发病人数明显多于女性,80 岁以后略少于女性。尤其在 30—59 岁之间,男女总发病人数比约为 3 : 1,远远大于其他年龄段。男女在 50、60、70 和 80 岁均存在明显的发病峰值。

### 2.2 急性脑血管疾病与气象要素的相关性

在分析气象因子与急性脑血管发病人数的相关性发现,不同季节的影响因子有所差异,故本文针对不同季节分析两者的关系。同时相关研究(Capon et al,1992)通过分析脑血管患者发病前气象要素的动态变化,即发病时的某一气象要素与病前 24 h 的同一气象要素的差值,通过判别分析得知气象要素动态变化的因子贡献最大,即气压、温度和湿度的剧烈变化均是导致脑血管发病的重要因子。对此,本文主要选取各要素的时间动态变化进行相关分析,图 3 列出了脑血管发病人数与气象因子的相关系数(由于篇幅有限,列出主要因子),其中横坐标中  $P_i$  ( $i=1, \dots, 7$ ) 分别表示 1.2 节中描述的各要素的时

间派生因子  $X, X_{-1}, X_{-2}, X_{-3}, X-X_{-1}, X-X_{-2}$  和  $X-X_{-3}$ , 而  $X$  表示当天要素,  $X_{-1}$  表示前 1 天要素,  $X_{-2}$  前 2 天要素,  $X_{-3}$  前 3 天要素;  $R(0.05)$  和  $R(0.01)$  分别表示相关系数通过  $\alpha=0.05$  和  $\alpha=0.01$  显著性检验的临界值。

春季发病人数主要与当天和前 1~3 d 的要素关系密切。发病人数仅与风速、海平面气压存在显著的正相关,与其他要素的相关性较差。

夏季发病人数主要与发病当天和前 1~3 d 的要素关系密切。发病人数与温度(气温、最高气温和最低气温)、风速、海平面气压呈明显的正相关,与水汽压和相对湿度也有较好的正相关。

秋季发病人数主要与发病当天的气象要素关系密切。发病人数与温度(气温、最低气温)、水汽压呈显著负相关,与风速、海平面气压、气温日较差存在明显的正相关。

冬季发病人数主要与发病当天与前 1~2 d 的要素关系密切。发病人数与温度(气温、最低气温、24 h 变温)呈显著负相关,与水汽压、风速、海平面气压、气温日较差呈显著正相关。

根据四个季节的逐日发病人数与各气象要素的

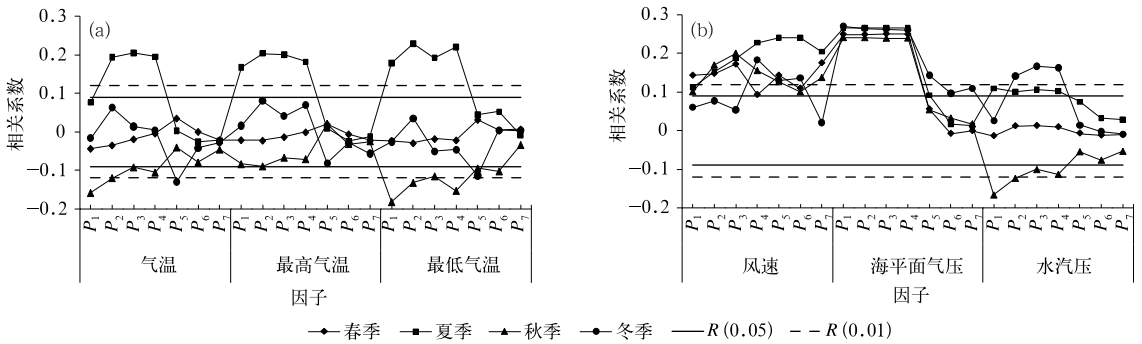


图 3 北京市急性脑血管疾病逐日发病人数与气象因子的相关系数  
Fig. 3 The correlation coefficient between the daily number of acute cerebrovascular disease and meteorological factors in Beijing

相关性分析结果可看出:不同季节中,气温对脑血管疾病都有一定程度的影响,且各因子对诱发脑血管疾病的发生存在滞后效应。其中,春季脑血管疾病的影响条件不明显;夏季,高温高湿天气容易影响脑血管疾病的发生;秋、冬季,当冷空气活动时,会导致大幅度降温,气压升高,带来较强的风,这些气象要素会影响到脑血管疾病的诱发。同时,夏季最高气温对发病人数也有明显的影响,而其他季节不存在这种相关,表明夏季日最高气温的增加可能是心脑

血管疾病发病的一个危险性因素;而其他季节,日最高气温的增加反而作为一种保护性因素,能够减少心脑血管疾病的发生(王佳佳等,2009)。

### 2.3 急性脑血管预报模型

#### 2.3.1 急性脑血管发病人数等级划分

同步评估疾病流行性风险水平与预报结果的气象分级,一般应用概率积分的方法进行评估与分级。本文根据概率分布将北京急性脑血管疾病划分为 4

级,各级所占概率分别为  $P_1$ 、 $P_2$ 、 $P_3$  和  $P_4$ , 其中  $\sum P_i=1(i=1,2,3,4)$ 。其中 1~4 级分别表示日发病人数较少、偏多、较多和很多。

通过分析北京急性脑血管疾病的日发病人数的概率分布(图4),四个季节中日发病人数的概率分

布基本一致,多数集中在 3~7 人之间。根据概率分布方法,将  $P_1$ 、 $P_2$ 、 $P_3$  和  $P_4$  分别定为 40%、30%、20%和 10%。由于不同季节的日发病人数的概率分布有一定差异,故利用规定的概率积分值进行不同季节的发病等级划分,划分结果如表 1。

2.3.2 预报结果检验

将相关系数通过  $\alpha=0.05$  显著性水平检验的气象因子作为预报因子,日发病等级作为预报量。在预报模型建立前,先对各个因子进行归一化处理,有利于避免各个因子间的量级差异,然后将归一化的因子作为 BP 网络输入层的节点输入,而预报量作为预报输出。按照上述思路,采用 2006—2009 年的逐日发病人数资料分别建立春、夏、秋、冬脑血管日发病等级的预报模型并进行反算,然后利用预报模型对 2010 年逐日发病等级进行预报并进行检验。

在对历史样本的拟合结果和独立样本的预报结果进行检验时,主要采用两种方法:(1)若预报值和实际值处于同一级别算完全正确,否则算错;(2)若两者处于同一级或差一级算对,否则算错。

通过预测模型对 2006—2009 年逐日发病人数等级(历史样本)的拟合及 2010 年逐日发病人数等级(独立样本)的预报且进行检验。历史建模样本的检验结果(表 2)表明:夏季拟合效果最好,秋季和冬季次之,而春季最差。独立新样本的预报结果(表 3)表明:在发病人数较多的春、秋和冬季,预报结果的完全准确率都超过 30%,相差 1 级的准确率

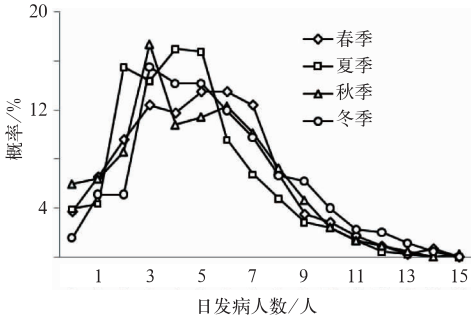


图 4 2006—2010 年北京急性脑血管疾病日平均发病人数的概率分布

Fig. 4 The probability distribution of the average daily number of acute cerebrovascular disease in Beijing during 2006—2010

表 1 北京急性脑血管发病人数分级表  
Table 1 The grade classification table of acute cerebrovascular disease in Beijing

等级	就诊人数	发病人数范围/人			
		春季	夏季	秋季	冬季
1	较少	≤4	≤3	≤3	≤4
2	偏多	5~6	4~5	4~6	5~6
3	较多	7~8	5~8	7~8	7~9
4	很多	≥9	≥9	≥9	≥10

表 2 2006—2009 年历史建模样本的拟合结果

Table 2 The fitting result of history modeling samples during 2006—2009

等级	完全正确/%				误差±1 个等级/%			
	春季	夏季	秋季	冬季	春季	夏季	秋季	冬季
1	43.7	89.5	92.0	83.4	83.9	92.8	96.9	92.7
2	43.3	56.1	44.9	26.8	89.7	98.4	93.2	94.8
3	37.7	66.7	46.6	37.3	82.0	89.9	82.8	85.3
4	40.6	85.0	81.0	82.4	53.1	95.0	81.0	82.4
平均	41.3	74.3	66.1	57.5	77.2	94.0	88.5	88.8

表 3 2010 年独立新样本的预报结果

Table 3 The forecast result of independent new samples in 2010

等级	完全正确/%				误差±1 个等级/%			
	春季	夏季	秋季	冬季	春季	夏季	秋季	冬季
1	40.7	66.7	81.8	62.9	55.6	71.4	81.8	74.3
2	32.0	12.5	10.5	21.1	88.0	96.9	84.2	89.5
3	29.6	7.7	16.7	7.7	63.0	30.8	61.1	34.6
4	23.1	7.7	20.8	30.0	53.8	7.7	33.3	50.0
平均	31.4	23.6	32.5	30.4	65.1	51.7	65.1	62.1

超过60%,而在夏季准确率略低于这三个季节。

可以看出,发病较为集中的春、秋和冬季的历史拟合效果和独立样本的预报效果较为一致,而夏季的拟合结果较好,但预报效果不理想。根据预报值和真实值的对比(图略),预报值能较好地反映真实值的变化趋势,可见该方法具有较好的预报性。但由于模型对发病极值的预测水平较弱,导致预报准确性随着发病等级的提高而降低,同时,模型在夏季的预报推广能力较差,这是因为预报效果除了与网络模型的拓扑结构及各参数有关外,其本质上还取决于预报系统的预报因子与预报量的内在关系。而影响脑血管疾病发生的要素较复杂,本文仅考虑了气象因子,而一些非气象要素也是直接诱发脑血管疾病的重要原因,而在预报模型中无法得到体现。

### 3 结 论

本文通过相关分析分别选取不同季节的急性脑血管影响气象因子,并采用人工神经网络方法构建了北京市急性脑血管疾病逐日发病等级的预测模型,主要结论如下:

(1) 北京急性脑血管发病人数具有明显的季节和日变化特征,冬春季发病高于夏秋季,发病人数主要集中在每天的09:00—14:00。40岁以后发病人数随着年龄增大而迅速增加,70—79岁达到峰值,且男性明显多于女性。

(2) 春季脑血管疾病的影响条件不明显;夏季,高温高湿天气容易影响脑血管疾病的发生;秋和冬季,当冷空气活动时,会导致大幅度降温,气压升高,带来较强的风,这些气象要素会影响到脑血管疾病的诱发。同时,夏季最高气温对发病人数也有很不利的影响,而其他季节不存在这种相关。

(3) 根据概率积分方法,将发病人数划分为4

个等级,利用人工神经网络方法建立的预报模型,对2010年独立样本本具有较好的预测效果,表明该方法具有一定的实际业务应用价值。

### 参考文献

- 董全,黄小玉,宗志平.2013.人工神经网络法和线性回归法对降水相态的预报效果对比.气象,39(3):324-332.
- 韩建康,刘小琦,顾志伟,等.2008.湖州市心脑血管疾病与气象因素的关系分析及预报研究.浙江预防医学,20(12):8-10.
- 刘方,张金良,陆晨.2004.我国气象因素与心脑血管疾病研究现状.气象科技,32(6):425-437.
- 刘世玲,刘济跃,李志莉,等.1999.脑血管发病与气象条件的关系.临床神经学杂志,12(2):76-78.
- 刘涛,范利,李冬云.2006.组织多普勒成像评价血压昼夜节律变异对左心室舒张功能的影响.中华老年心脑血管病杂志,8:370-372.
- 路风,金银龙,陈义斌.2008.气象因素与心脑血管疾病关系的研究进展.国外医学卫生学分册,35(2):83-87.
- 卢爱梅,徐红梅,高谨,等.1997.齐齐哈尔市心、脑血管疾病发病与气象要素的关系及其预测预报研究.中国慢性病预防与控制,5(2):61-63.
- 毛清宇,尹东屏,孙宁,等.2010.南京市心脑血管疾病的医疗气象预报研究.气象,36(11):82-87.
- 闵晶晶,孙景荣,刘还珠,等.2010.一种改进的BP算法及在降水预报中的应用.应用气象学报,21(1):55-62.
- 王佳佳,郭玉明,李国兴,等.2009.日最高气温与医院心脑血管疾病急诊人次关系的病例交叉研究.环境与健康杂志,26(12):1073-1076.
- 叶殿秀,杨贤为,吴桂贤.2003.京、沪两地脑卒中发病率及其预测模型.气象科技,31(6):381-384.
- 曾晓青.2013.BP神经网络在建模中的参数优化问题研究.气象,39(3):333-339.
- 张书余,王宝鉴,谢静芳,等.2010.吉林省心脑血管疾病与气象条件关系分析和预报研究.气象,36(9):106-110.
- Capon A, Demeurisse G, Zheng L.1992. Seasonal variation of cerebral hemorrhage in 236 consecutive cases Brussels. Stroke, 23(1):24-27.
- Shin kawa A, Ueda K, Hasuo Y, et al.1990. Seasonal variation in stroke incidence in Hisayama, Japan. Stroke, 21(9):1262.