

吴千鹏,尹立,李兴荣,等.季风转换对深圳地区呼吸系统疾病的影响及预测研究[J].沙漠与绿洲气象,2023,17(6):32-40.

doi:10.12057/j.issn.1002-0799.2023.06.005

开放科学(资源服务)标识码(OSID):



季风转换对深圳地区呼吸系统疾病的影响及预测研究

吴千鹏¹,尹立²,李兴荣³,孙羽⁴,黄开龙⁵,苏春芳⁴,王式功^{1*}

(1.成都信息工程大学大气科学学院/环境气象与健康研究院,四川 成都 610225;

2.攀枝花市中心医院气象医学研究中心,四川 攀枝花 617000;3.深圳市气象局,广东 深圳 518040;

4.海南省第二人民医院气候医学研究中心,海南 五指山 572299;5.汕头市气象局,广东 汕头 515041)

摘要: 利用深圳地区2015—2016年呼吸系统疾病就诊人数资料及同期气象和污染物资料,并运用BP人工神经网络和LSTM网络构建呼吸系统疾病就诊人数预测模型。结果显示:每年9月开始,冬季风的冷胁迫效应使相关人群呼吸系统疾病发病人数波动式增加,直至次年冬季风向夏季风转换前的3月发病人数达到峰值;夏季风控制期间当地居民呼吸系统疾病发病人数呈波动式减少,比峰值期间减少35%;在不同季风控制期间不同呼吸系统疾病其主控因素也不相同;对比两种预测模型,总体上LSTM网络预报模型对深圳下呼吸道疾病风险预测准确率更高,可以满足健康气象预报服务业务需求。

关键词: 季风转换;呼吸系统疾病;气象与污染条件;人工神经网络;LSTM网络;预测模型

中图分类号: P49

文献标识码: A

文章编号: 1002-0799(2023)06-0032-09

呼吸系统疾病是人体最常见、多发的疾病之一,因其对天气气候变化有着较高的敏感度,致使人体极易患病影响健康,故又被叫做气象敏感性疾病。随着全球气候变化对人类健康产生的影响进一步加剧,呼吸系统等气候敏感性疾病发病率增加^[1-3]。联合国政府间气候变化专门委员会(Intergovernmental Panel on Climate Change,IPCC)第六次评估报告指出,天气气候变化导致热浪、干旱和洪水等极端气候事件增多、天气骤变加剧,相关气象敏感性疾病(呼

吸系统疾病等)发病率、死亡率逐年增加^[4]。在气候异常变化这一大背景下,天气气候因素很可能会对我国气候敏感人群特别是老人、儿童或患有呼吸系统疾病等人群构成巨大威胁。

天气气候变化是诱发呼吸系统疾病的主要原因之一,即呼吸系统疾病等气象敏感性疾病的发生发展与天气气候有着非常密切的关系^[5-11]。在温度、湿度、风速等气象要素相对适宜的情况下呼吸系统疾病发病或死亡人数最少,超出舒适范围人体均有不同程度的不良反应^[11-17]。另外,大气污染状况及其变化也会对呼吸系统疾病的发病率及死亡率产生影响^[17-21]。基于气象和环境数据与呼吸系统疾病数据间的密切相关性,如何利用这些数据构建较优的呼吸系统疾病预测模型,为呼吸系统疾病医疗气象服务提供一定的科学指导,非常重要。以往研究者倾向于使用回归分析来反映气象要素和敏感性疾病之间的关系^[5,22-33],其中被广泛使用的是多元回归和逐步回归。考虑到气象环境条件变化和人体疾病发病间的

收稿日期:2023-06-04;修回日期:2023-07-27

基金项目:攀枝花市科学技术局创新中心建设项目(2021ZX-5-1);2021年省级科技计划转移支付专项资金项目(21ZYZF-S-01);中国气象局西南区域气象中心创新团队基金(XNQYCXTD-202203);2021年度第二批攀枝花市市级科技计划项目(2021CY-S-4)

作者简介:吴千鹏(1996—),女,硕士研究生,主要从事气象环境与健康研究。E-mail:597639125@qq.com

通信作者:王式功(1955—),男,教授,主要从事气象环境与健康研究。E-mail:wangsg@cuit.edu.cn

非线性关系,且存在着一定的滞后效应,被最常应用的模型是广义相加模型(GAM)和分布滞后非线性模型(DLNM)^[24-31]。随着科技的进步,机器学习广泛应用于气象领域,目前主要应用于气象观测、数值天气预报、强对流天气识别预警以及卫星资料应用等方面^[34-38],同时也应用于对医疗就诊人数的预测^[9,10,45]。

深圳作为我国发展最快的一线城市,地处典型的东亚季风气候区,在一年之中,季风在每年4和9月转换方向^[39],季风预示着相关地区气候状况的大转折,对当地工农业生产和人民群众的日常生活产生重要影响。季风转换会对居民健康产生何种影响?目前尚未见到国内外相关报道。因而这是非常值得研究的问题。健康气象服务作为未来气象服务领域拓展的重点方向之一,已将其列入2021年11月发布的我国气象事业发展“十四五”规划中,提出要建立疾病发生发展风险预测模型。基于贯彻落实国家大健康政策及广大民众疾病预防的现实需求,研究构建呼吸系统疾病等气象敏感性疾病预测模型显得尤为迫切。开展季风转换对呼吸系统疾病的影响研究,既是践行国家全民大健康战略实施的具体行动,又有重要科学价值,意义非凡。本文利用深圳地区风向风速等资料研究季风及其转换对呼吸系统疾病的影响,基于不同季风影响期间内,下呼吸道感染疾病均对气象和环境条件敏感性更强,利用BP神经网络和LSTM网络方法尝试建立当地下呼吸道感染疾病发病风险预测模型,以期为深圳市呼吸系统疾病预防提供科学依据和技术支持,同时也可丰富季风及其转换背景下的医学气象学理论内涵。

1 资料与方法

1.1 资料来源

呼吸系统疾病日就诊资料来源于深圳市龙华区观澜中心医院2015年1月1日—2016年12月31日每日就诊记录,该院为深圳市三级甲等医院,位于中心城区,具有一定的代表性。气象数据来自中国气象科学数据共享平台提供的国家基准站常规气象数据,包括气温、气压、相对湿度、水汽压等;污染数据来自中国监测网(<https://www.aqistudy.cn>),包括PM_{2.5}、PM₁₀、CO、SO₂、NO₂等污染物质量浓度。

1.2 研究方法

1.2.1 相关分析

不同环境要素和当地呼吸系统疾病的就诊人数之间有明确的相关性,描述变量间的相关程度,最常用的是Spearman秩相关系数 r_s ,本文利用Spearman

相关分析环境因子与当地医院就诊人数之间的相关程度。计算方法为:

$$r_s = 1 - \frac{6 \sum_{i=1}^n D_i^2}{n(n^2-1)} \quad (1)$$

式中, r_s 为Spearman秩相关系数, n 为样本量个数, $D_i = x_i^2 - Y_i^2$,其中假定原始数据 x_i (环境因子)、 Y_i (疾病就诊人数)已按时间顺序进行排列,记 x_i^2 、 Y_i^2 为原数据 x_i 、 Y_i 在排列后数据所在位置, x_i^2 、 Y_i^2 的秩次之差 $D_i = x_i^2 - Y_i^2$ 。

1.2.2 长短期记忆网络(LSTM网络)

气象要素对呼吸系统疾病发生的影响具有一定的滞后性^[24-31,40-44],大多数气象要素的滞后为1~3 d。LSTM网络(Long-Short Term Memory)由于其独特的设计结构,适合处理和预测时间序列。

LSTM网络是循环神经网络的一个变体,可以有效地解决简单循环神经网络(RNN)的梯度爆炸或消失问题^[45-46]。这种网络由输入层、输出层和之间若干层(一层或多层)隐含层构成,主要改进在2个方面:一是引入一个新的内部状态专门进行线性的循环信息传递,同时(非线性地)输出信息给隐藏层的外部状态。二是引入门控机制(Gating Mechanism)控制信息传递的路径,总共有三个“门”:输入门(控制当前时刻的候选状态有多少信息需要保存)、遗忘门(控制上一个时刻的内部状态需要遗忘多少信息)和输出门(控制当前时刻的内部状态有多少信息需要输出给外部状态)。

1.2.2.1 归一化处理

样本数据量纲、单位往往不同,进而会影响数据分析的结果。为消除样本数据之间的量纲影响、消除奇异样本数据导致的不良影响,且为满足节点函数要求,提高网络训练速度和收敛性,归一化处理公式为:

$$x_{ij} = 0.8 \times \frac{X_{ij} - x_{\min}}{x_{\max} - x_{\min}} + 0.1 \quad (2)$$

式中: x_{ij} 为归一化后的值, X_{ij} 为原始值, $x_{\min}$ 为自变量 X_i 中的最小值, $x_{\max}$ 为自变量 X_i 中的最大值; i 、 j 分别为自变量序号和样本序号。

1.2.2.2 隐藏层中的节点个数

由于学术界到目前为止没有明确的理论规定隐藏层中的节点个数,通常情况下使用以下公式确定节点个数:

$$m = \sqrt{n+l} + \alpha \quad (3)$$

式中: m 为隐藏层节点数, n 为输入层节点数, l 为输

出层节点数; α 为调节常数,为1~10。根据文中所选变量的个数,可以知道输入节点为11,输出节点为1,公式(3)可知隐藏节点为4~14。运用1层LSTM模型,试验并确定从4~14个节点所对应的误差率及误差平方和,并以此来判断模型的偏离程度。进而选取合适的隐藏层节点个数。

1.2.2.3 网络结构

学习率设置成0.001,权重 W^* 、 U^* 与偏移向量 b^* ,以上参数全部实现初始化随机应用。为了更新权重与偏差度,需要不断的训练序列来实现。选用平均绝对误差(MSE)表示所用LSTM模型中损失函数的修正误差,并使用Adam优化器,促使网络实现最优学习。

1.2.3 试预报模型的评价

在构建呼吸系统疾病预测模型、进行预测模型优度检验时,主要使用平均绝对误差(M_{AE})、平均绝对百分比误差(M_{APE})、均方根误差(R_{MSE})和预测准确度(P)等指标对预测模型的效果进行评价,其中 Y_i 是实际值, y_i 是预测值, n 是样本数,计算方法如下:

$$\left\{ \begin{aligned} M_{AE} &= \frac{1}{n} \sum_{i=1}^n |Y_i - y_i| \\ M_{APE} &= \frac{1}{n} \sum_{i=1}^n \left| \frac{Y_i - y_i}{Y_i} \right| \\ R_{MSE} &= \sqrt{\frac{1}{n} \sum_{i=1}^n (Y_i - y_i)^2} \\ P &= \left[1 - \frac{M_{AE}}{\frac{1}{n} \sum_{i=1}^n Y_i} \right] \times 100\% \end{aligned} \right. \quad (4)$$

2 结果分析

2.1 呼吸系统疾病就诊人数逐月与季节分布

通过整理深圳地区2015年1月1日—2016年12月31日呼吸系统疾病数据资料,发现每年2月深圳地区呼吸系统疾病就诊人数最少,这与春节期间深圳务工人员返乡、外流人口较多有关联。因此在

分析深圳地区呼吸系统疾病就诊人数的逐月分布时,剔除了2月原始呼吸系统疾病数据,利用2月前2个月的呼吸系统疾病数据对2月的疾病数据进行订正。图1a为订正后2015—2016年深圳地区呼吸系统疾病就诊人数的年内逐月变化。深圳总呼吸系统疾病日均就诊人数3月最多,为288人/d,总呼吸系统疾病日均就诊人数8月最少(187人/d),3月下呼吸道感染就诊人数最多,6月上呼吸道感染就诊人数最多。虽然各类呼吸系统疾病就诊人数呈现一定的波动性,但也有相同之处:即2—3月总呼吸系统(ICD10编码J00—J99)、上呼吸道感染(简称“上感”,ICD10编码J00—J04,J06)和下呼吸道感染(简称“下感”,ICD10编码J10,J12,J15,J16等)就诊人数有明显的增多,可能与冬季风持续影响及其累积效应有关,季节交替就诊人数明显增加,由于季节变换时气温、湿度等气象条件出现突变,也可能是春节后深圳市人口增多的缘故。图1b中填色部分对应月(冬季1月、秋季11月、春季3月)分别为辽宁北票、北京和深圳呼吸系统疾病高发月,北票和北京对呼吸系统疾病为冷效应,深圳为热效应,2—3月北票和北京在季风转换前升温快,而深圳升温慢,回暖弱,所以北票和北京在秋、冬季呼吸系统疾病有峰值,深圳冬季风向夏季风转换前呼吸系统疾病发病出现峰值。

深圳地处东亚季风气候区,其风向会随着冬季风、夏季风的控制及其转换而变化。其中,由冬季风转换为夏季风的过渡期一般发生在4月,夏季风转换为冬季风的过渡期一般发生在9月^[39]。

本文综合2015—2016年各月深圳平均10 min风速风向特征和各月主导风控制天数(表1)研究发现,2015—2016年深圳地区季节转换月分别出现在4和9月,其中,5—8月主要受夏季风(偏南风)控制,10月—次年3月主要受冬季风(偏北风)控制。对照图1可看出,夏季风控制期间深圳地区总呼吸系统疾病就诊人数一直减少,反映出中医理论阐述

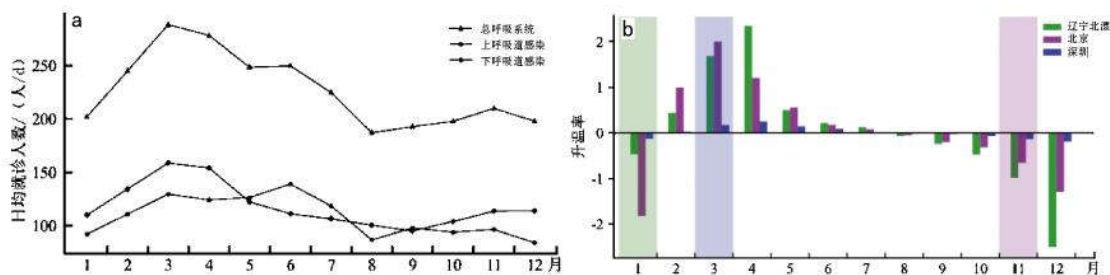


图1 深圳地区呼吸系统疾病日均就诊人数和升温率的逐月变化

(a为呼吸系统疾病日均就诊人数,b为升温率)

的“春夏为阳”的养生气候特征;相反,冬季风控制期间深圳地区总呼吸系统疾病就诊人数一直在增加,反映出中医理论中论述的“秋冬为阴”的气候特征,此时需加强疾病预防。另外,上感与下感对季节转换的响应略有差异。

表 1 深圳地区 2015—2016 年各月主导风控制天数占比

季风特征	月	偏北风控制 天数占比/%	偏南风控制 天数占比/%
冬季风主导	1	100	0
冬季风主导	2	98.25	1.75
冬季风主导	3	79.03	20.97
冬—夏季风转换	4	46.67	53.33
夏季风主导	5	35.48	64.52
夏季风主导	6	10.00	90.00
夏季风主导	7	16.13	83.87
夏季风主导	8	45.16	54.8
夏—冬季风转换	9	63.33	36.67
冬季风主导	10	90.32	9.68
冬季风主导	11	96.67	3.33
冬季风主导	12	98.39	1.61

综合图 1 和表 1 可知:(1)每年 9 月开始的冬季风对当地居民的冷胁迫效应会使相关人群总呼吸系统疾病发病人数波动式增加,直至次年冬季风为夏季风转换前的 3 月发病人数达到峰值。(2)每年 4 月冬季风开始转换为夏季风,整个夏季风控制期间对当地居民的温热效应使相关人群总呼吸系统疾病发病人数呈波动式减少态势,直至夏季风转换为冬季风之前的 8 月发病人数达到谷值,比峰值减少 35%。(3)去掉季风风向转换月(4 和 9 月),深圳地区夏季风主导期间总呼吸系统疾病日均就诊人数比冬季风主导期间多 7.26%。但上感与下感又有差异,

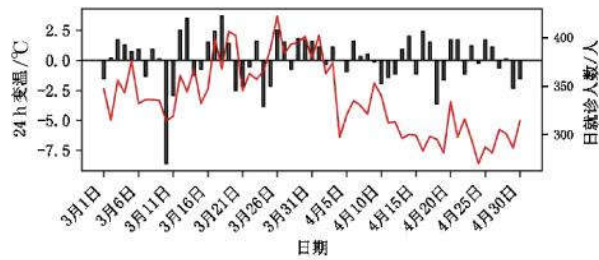


图 2 2016 年 3—4 月深圳地区冬季风向夏季风转换期间 24 h 变温与呼吸系统疾病就诊人数逐日变化

其中夏季风主导期间上感日均就诊人数比冬季风主导期间多 22.93%,下感就诊人数夏季风主导期间比冬季风主导期间少 5.58%。

深圳地区 3 月就诊人数表现出较为明显的阶段性爆发,图 2 为 2016 年 3—4 月深圳地区 24 h 变温与呼吸系统疾病就诊人数逐日变化。深圳地区 24 h 变温与总呼吸系统疾病就诊人数逐日变化关系密切,相邻 2 d 内气温变化趋势相反(第一天升温第二天降温或第一天降温第二天升温)呼吸系统疾病就诊人数都会增加。

2.2 深圳地区呼吸系统就诊人数与气象和环境要素的相关分析

表 2 列出了 2015—2016 年不同季风影响期间深圳地区呼吸系统就诊人数与各气象环境因子变量之间的相关系数。气象要素对呼吸系统疾病的发生有较为显著的影响,整体来看,下感就诊人数与气象环境因子变量间的相关性更强。气压和气温与下感患病人数数的相关性最好,水汽压与下感患病人数相

表 2 深圳地区呼吸系统发病人数与主要环境因子的 Spearman 相关

变量	夏季风控制期间			冬季风控制期间		
	总呼吸	上感	下感	总呼吸	上感	下感
平均气温/℃	-0.273***	-0.025	-0.464***	-0.015	0.119*	-0.179***
最高气温/℃	-0.306***	-0.085	-0.452***	0.026	0.148**	-0.135*
最低气温/℃	-0.265***	-0.022	-0.464***	-0.031	0.11*	-0.196***
平均气压/hPa	0.366***	0.18**	0.478***	0.047	-0.107*	0.207***
最高气压/hPa	0.384***	0.182**	0.511***	0.048	-0.103*	0.206***
最低气压/hPa	0.346***	0.182**	0.442***	0.039	-0.115*	0.199***
平均相对湿度/%	0.005	-0.039	0.042	0.091	0.158**	0.017
最小相对湿度/%	0.07	0.019	0.091	0.048	0.113*	-0.01
水汽压/hPa	-0.255**	-0.037	-0.42***	0.024	0.168***	-0.138**
平均风速/(m/s)	0.059	0.123*	-0.054	-0.108*	-0.124*	-0.056
最大风速/(m/s)	-0.024	0.093	-0.19***	-0.181***	-0.185***	-0.14**
极大风速/(m/s)	0.073	0.189***	-0.126*	-0.063***	-0.078	-0.031
PM _{2.5} /(ug/m ³)	0.119*	0.107	0.289***	0.159***	0.051	0.213***
PM ₁₀ /(μg/m ³)	0.112*	-0.037	0.245***	0.166***	0.051	0.219***
SO ₂ /(μg/m ³)	0.151**	0.107	0.157**	0.174***	0.094	0.204***
NO ₂ /(μg/m ³)	-0.126*	-0.211***	0.015	0.231***	0.126**	0.286***
CO(mg/m ³)	-0.082	-0.143*	0.031	0.053	-0.062	0.139**

注:*** ** 和 * 分别表示通过了 $\alpha=0.001$ 、 $\alpha=0.01$ 、 $\alpha=0.05$ 的显著性检验,下同。

关性次之。下感就诊人数与气压呈显著正相关,与气温、水汽压和风速呈显著负相关。下感发病人数还与各污染物浓度呈显著正相关。

2.3 深圳地区下感风险预测模型

因研究区域下感就诊人数与环境气象要素间的相关性更强,即下感发病对气象和环境条件更敏感,且已有较多新的方法构建过上感风险预测模型^[10,26],所以本文着重研究下感风险预测模型。夏季风控制期间内下感就诊人数预报方程使用的气象和环境要素为:平均气温、平均最高气温、平均最低气温、平均气压、平均最高气温、平均最低气温、水汽压、最大风速、极大风速、 $PM_{2.5}$ 、 PM_{10} 、 SO_2 ,冬季风控制期间内下感疾病就诊人数预报方程使用的气象和环境要素为:平均气温、平均最高气温、平均最低气温、平均气压、平均最高气温、平均最低气温、水汽压、最大风速、 $PM_{2.5}$ 、 PM_{10} 、 SO_2 、 NO_2 、 CO 。气象要素对呼吸系统疾病发生的影响具有一定的滞后性^[28,31-33],大多数气象要素滞后为1~3 d,对疾病数据和气象数据进行滑动平均处理,因空气污染因子对呼吸系统疾病发生的影响是即时效应^[15],所以未对污染数据作平滑处理。

深圳地区季风风向由冬季风(偏北风)转为夏季风(偏南风)后,深圳地区受夏季风(偏南风)影响期间下感就诊人数大体上随时间一直减少,深圳地区季风风向向夏季风(偏南风)转为冬季风(偏北风)后,深圳地区受冬季风(偏北风)影响期间(剔除2月的数据)的下感就诊人数随时间推移一直增多,因此本文构建下呼吸道感染风险预测模型时,大致划分为2个时间段:夏季风影响期间(4—8月)和冬季风影响期间(9—12月、1和3月)。每年2月呼吸系统疾病就诊人数为最小值,这很可能与春节期间外来人口返乡、加之流动人口也大幅度减少等有关,因此后续在构建预测模型时将2月相关数据剔除,以便更好地排除其他人为因素干扰,抓取典型的自然气候影响特征分析主要趋势,利于提高模型的学习效果、改善其预报准确度。

2.3.1 夏季风影响期间下感风险预测模型

2.3.1.1 BP 人工神经网络预测模型

将总样本归一化处理后的80%作为输入、对应的就诊人数作为输出建立BP神经网络预测模型,输入层神经元个数为12,输出层神经元个数为1,传递函数选择Purelin。隐含层为1层,通过经验公式和试凑法等确定隐含层神经元数目为6,传递函数为Logsig。深圳下感疾病就诊人数的神经网络预测

模型结构为12-6-1,训练精度为0.001。此时网络稳定性达到最好,试预报和拟合效果也较好。经BP神经网络算法进行网络学习训练,建立下感人数与空气污染物浓度的关系。把总样本归一化后剩余的20%样本作为验证样本输入到已经训练好的网络输入层中,对验证样本进行仿真验证。对仿真结果进行反归一化,得到下感就诊人数验证样本的仿真值(图3)。对试预报结果进行统计分析,检验试预报方程效果:平均绝对误差MAE为14.72,平均绝对百分比误差MAPE为12.04%,均方根误差RMSE为21.93,预测准确率P为87.89%。

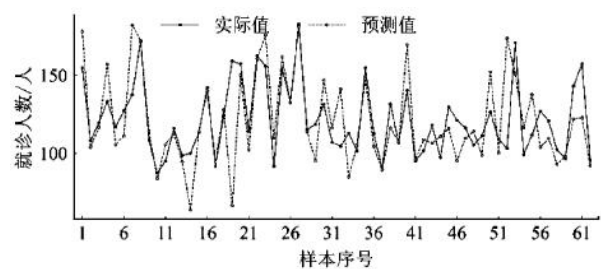


图3 深圳地区夏季风影响期间下感疾病就诊人数的模拟值与实际值拟合曲线(BP网络)

2.3.1.2 LSTM 网络(长短期记忆神经网络)

将总样本归一化处理后的80%作为输入、对应的就诊人数作为输出建立LSTM神经网络预测模型,夏季风期间输入层神经元个数为12,输出层神经元个数为1,隐含层为1层,运用1层LSTM模型,试验并确定从4~14个节点所对应的误差率及误差平方和,并以此判断模型的偏离程度,进而确定隐藏层节点个数为9。经LSTM神经网络算法进行网络学习训练,建立下感疾病就诊人数与环境要素的关系。把总样本归一化后剩余的20%样本作为验证样本输入到已经训练好的网络输入层中,对验证样本进行仿真验证。对仿真结果进行反归一化,得到下感疾病就诊人数验证样本的仿真值(图4)。对试预报结果进行统计分析,检验试预报方程效果为:平

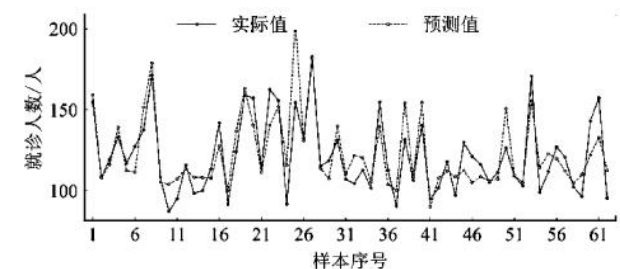


图4 深圳地区夏季风影响期间下感疾病就诊人数的模拟值与实际值拟合曲线(LSTM网络)

均绝对误差 MAE 为 10.27, 平均绝对百分比误差 $MAPE$ 为 8.47%, 均方根误差 $RMSE$ 为 13.21, 预测准确率 P 为 91.56%。

2.3.2 冬季风影响期间下感风险预测模型

2.3.2.1 BP 人工神经网络预测模型

输入层神经元个数为 13, 输出层神经元个数为 1, 传递函数选择 Tansig。隐含层为 1 层, 通过经验公式和试凑法等确定隐含层神经元数目为 6, 传递函数为 Logsig。深圳下感疾病就诊人数的神经网络预测模型结构为 13-6-1, 训练精度为 0.001。此时网络达到最好, 试预报和拟合效果也较好。经 BP 神经网络算法进行网络学习训练, 建立下感人数与环境要素的关系。把总样本归一化后剩余的 20% 样本作为验证样本输入到已经训练好的网络输入层中, 对验证样本进行仿真验证。对仿真结果进行反归一化, 得到下感就诊人数验证样本的仿真值(图 5)。对试预报结果进行统计分析, 检验试预报方程效果为: 平均绝对误差 MAE 为 20.55, 平均绝对百分比误差 $MAPE$ 为 16.36%, 均方根误差 $RMSE$ 为 29.84, 预测准确率 P 为 82.33%。

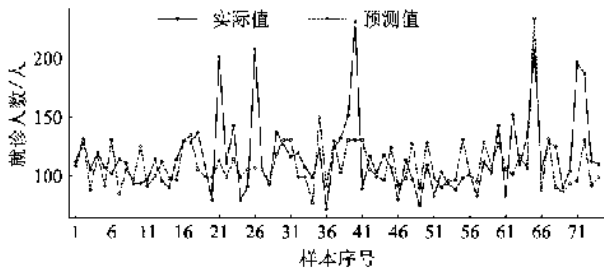


图 5 深圳地区冬季风影响期间下感疾病就诊人数的模拟值与实际值拟合曲线(BP 神经网络)

2.3.2.2 LSTM 网络(长短时记忆神经网络)

将总样本归一化处理后的 80% 作为输入、对应的就诊人数作为输出建立 LSTM 神经网络预测模型, 输入层神经元个数为 13, 输出层神经元个数为 1, 隐含层为 1 层, 运用 1 层 LSTM 模型, 试验并确定从 4~14 个节点所对应的误差率及误差平方和, 并以此来判断模型的偏离程度, 进而确定隐藏层节点个数为 6 个。经 LSTM 神经网络算法进行网络学习训练, 建立下感疾病就诊人数与环境要素的关系。把总样本归一化后剩余的 20% 样本作为验证样本输入到已经训练好的网络输入层中, 对验证样本进行仿真验证。对仿真结果进行反归一化, 得到下感就诊人数验证样本的仿真值(图 6)。对试预报结果进行统计分析, 检验试预报方程效果为: 平均绝对误差

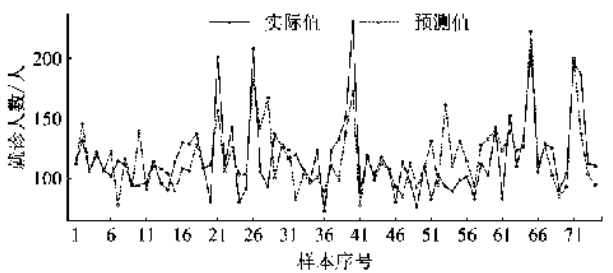


图 6 冬季风影响期间下感疾病就诊人数的模拟值与实际值拟合曲线(LSTM 网络)

MAE 为 17.87, 平均绝对百分比误差 $MAPE$ 为 16.22%, 均方根误差 $RMSE$ 为 24.31, 预测准确率 P 为 84.64%。

深圳地区受夏季风控制期间 2 种预报模型的预报结果均好于受冬季风控制时段, 这与冬季风控制时气象因子和就诊人数变化幅度较大有关。

2.3.3 两种模型的对比与评价

为客观评价 2 种模型对深圳地区下感疾病就诊人数的预测能力, 将两种模型的预报效果进行比较(表 3), 与 BP 神经网络对比, LSTM 网络的平均绝对误差和平均绝对百分比误差更小, 预测准确度更高, 预测模型精度略胜一筹, 总体上 LSTM 网络预测模型表现更优。

表 3 2 种预测模型对比

	模式方法	MAE	$MAPE/\%$	$RMSE$	$P/\%$
夏季风影响期间	BP 神经网络	14.72	12.04	21.93	87.89
	LSTM 网络	10.27	8.47	13.21	91.56
冬季风影响期间	BP 神经网络	20.55	16.36	29.84	82.33
	LSTM 网络	17.87	16.22	24.31	84.64
总	BP 神经网络	17.89	14.39	26.53	84.93
	LSTM 网络	14.4	12.69	20.03	87.87

3 结论

(1) 2015—2016 年深圳总呼吸系统疾病日均就诊人数 3 月最多(288 人/d), 8 月最少(187 人/d), 3 月下感就诊人数最多, 6 月上感就诊人数最多。深圳地区春、夏季为呼吸系统疾病发病高峰期; 春季(特别是冬季风向夏季风转换之前的 3 月)发病人数最多, 初秋时段发病人数最少。上感和下感疾病发病第一高峰期在春季, 上感疾病因夏季天气炎热导致高发也不可忽视。

(2) 每年 9 月开始的冬季风对当地居民的冷胁

迫效应使相关人群总呼吸系统疾病发病人数波动式增加,直至次年冬季风向夏季风转换之前的3月发病人数到达峰值;整个夏季风控制期间对当地居民的热胁迫效应使相关人群总呼吸系统疾病发病人数呈波动式减少态势,直至转换为冬季风主导前的8月发病人数达到谷值,比峰值减少35%。深圳地区夏季风主导期间总呼吸系统疾病日均就诊人数比冬季风主导期间多7.26%,但是上感与下感有差异,其中夏季风主导期间上感日均就诊人数比冬季风主导期间多22.93%,下感日均就诊人数在夏季风主导期间比冬季风主导时段少5.58%。已有研究表明,3月我国北方地区城市如北京、辽宁北票气温回暖快,发病人数减少得快;本研究表明,深圳3月气温升温慢且在冬季风向夏季风转换之期,当地发病人数达到峰值,之后开始下降,此现象与北方城市明显不一样,值得对更多南方城市呼吸系统疾病发病情况的年变化进行研究。

(3)下感与气温及水汽压呈负相关关系,与气压呈正相关关系,与气温相关性最好,与气压、水汽压相关性次之,还与污染要素呈显著正相关。与BP神经网络模型对比,LSTM网络模型表现更为出色,具有更高的预测精度和泛化能力,对呼吸系统疾病就诊人数预测的准确性更好,可以作为一种新的实用方法为呼吸道疾病发病风险预测提供技术支持。

参考文献:

- [1] MCMICHAEL AJ, WOODRUFF RE, HALES S. Climate change and human health: present and future risks[J]. Lancet, 2006, 367(9513): 859-69.
- [2] KAN H, CHEN R, TONG S. Ambient air pollution, climate change, and population health in China[J]. Environment International, 2012, 42(none): 10-19.
- [3] SEMENZA J C, SUK J E, ESTEVEZ V, et al. Mapping climate change vulnerabilities to infectious diseases in Europe[J]. Environmental Health Perspectives, 2012, 120(3): 385-392.
- [4] Climate change 2022—the physical science basis: Working group III contribution to the sixth assessment report of the IPCC[M]. London Cambridge University Press, 2022.
- [5] FALAGAS M E, THEOCHARIS G, SPANOS A, et al. Effect of meteorological London variables on the incidence of respiratory tract infections[J]. Respiratory Medicine, 2008, 102(5): 733-737.
- [6] 李娟, 张志薇, 于庚康, 等. 气象要素对南京市呼吸系统疾病的影响研究[J]. 气象科学, 2017, 37(3): 409-415.
- [7] 赵金琦, 金银龙. 气候变化对人类环境与健康影响[J]. 环境与健康杂志, 2010, 27(5): 462-465.
- [8] 马盼, 王式功, 尚可政, 等. 气象舒适条件对呼吸系统疾病的影响[J]. 中国环境科学, 2018, 38(1): 374-382.
- [9] 马盼, 王馨梓, 张莉, 等. 深圳流感发病的气象诱因及预测建模研究[J]. 气象学报, 2022, 80(3): 421-432.
- [10] 郑甲炜, 王式功. 基于机器学习的两个代表城市上呼吸道感染与气象要素关系及其发病风险预测研究[J]. 沙漠与绿洲气象, 2023, 17(4): 162-170.
- [11] 乐满, 王式功, 谢佳君, 等. 环境条件对遵义市呼吸系统疾病的影响及预测研究[J]. 中国环境科学, 2018, 38(11): 4334-4347.
- [12] WENFANG G, YI L, WANG P, et al. Assessing the effects of meteorological factors on daily children's respiratory disease hospitalizations: A retrospective study[J]. Heliyon, 2020, 6(8): e04657.
- [13] 汪庆安. 2010—2015年宁夏气温变化对人群死亡的影响[D]. 银川: 宁夏医科大学, 2018.
- [14] 张越, 阚海东, 彭丽, 等. 日均气温与呼吸系统疾病日入院人次相关性的时间序列分析[J]. 中华预防医学杂志, 2014(9): 795-799.
- [15] 莫运政, 郑亚安, 陶辉, 等. 日均气温与呼吸系统疾病急诊人次相关性的时间序列分析[J]. 北京大学学报(医学版), 2012, 44(3): 416-420.
- [16] WANG Z, ZHOU Y, LUO M, et al. Association of diurnal temperature range with daily hospitalization for exacerbation of chronic respiratory diseases in 21 cities, China[J]. Respiratory Research, 2020, 21(1): 1-10.
- [17] 周建丁. 北京市大气污染和气温对呼吸系统疾病的影响研究[D]. 兰州: 兰州大学, 2019.
- [18] 张莹, 邵毅, 王式功, 等. 北京市空气污染物对呼吸系统疾病门诊人数的影响[J]. 中国环境科学, 2014, 34(9): 2401-2407.
- [19] 吴彧, 张衍桑, 刘玲, 等. 郑州市空气污染对每日门诊人次影响的滞后效应和收获效应[J]. 环境与健康杂志, 2011, 28(3): 233-237.
- [20] 陶燕, 羊德容, 兰岚, 等. 兰州市空气污染对呼吸系统疾病入院人数的影响[J]. 中国环境科学, 2013, 33(1): 175-180.
- [21] 孔祥意, 齐麟, 滕聪. 空气状况、气象条件、舒适度指数与人体健康关系及预报反演[J]. 测绘与空间地理信息, 2020, 43(10): 148-152, 155.
- [22] CHEN S, LIU C, LIN G, et al. The role of absolute humidity in respiratory mortality in Guangzhou, a hot and wet city of South China[J]. Environmental Health and Preventive Medicine, 2021, 26(1): 1-11.
- [23] PEDDER H, KAPWATA T, Howard G, et al. Lagged association between climate variables and hospital admissions for pneumonia in South Africa[J]. International Journal of Environmental Research and Public Health, 2021, 18(12): 6191.

- [24] 刘娟娟,张涵.呼吸系统疾病与气象要素的关系及预测分析[J].农业科技与信息,2016(11):55.
- [25] 刘新超,郭洁,陈婷,等.儿童呼吸系统疾病与气象要素的关系及其预测[J].高原山地气象研究,2015,35(1):92-96.
- [26] 黄开龙.深圳市上呼吸道感染与气象因子的关系及疾病预测模型研究[D].成都:成都信息工程大学,2020.
- [27] 李瑞盈,张一博,杨佳,等.秦皇岛气象因素对儿童下呼吸道感染疾病就诊人数影响及预测研究[J].干旱气象,2019,37(3):460-466.
- [28] 谭玉龙.中国东西部不同地区气象要素对心脑血管疾病的影响及其预测[D].成都:成都信息工程大学,2019.
- [29] 刘娟娟,张涵.呼吸系统疾病与气象要素的关系及预测分析[J].农业科技与信息,2016(11):55-55.
- [30] 李雪源,景元书,吴凡,等.南京市呼吸系统疾病死亡率与气象要素的关系及预测[J].气象与环境学报,2012,28(5):46-48.
- [31] 宋文燕,刘新华.呼吸与循环系统疾病的发病与气象要素关系的分析与预测[J].中国医院统计,2002,9(2):81-83.
- [32] 田依洁.北京市呼吸系统疾病与气象因素及空气质量关系的研究[D].兰州:兰州大学,2018.
- [33] 马盼,李若麟,赵笑颜,等.北京市空气质量指数与气象条件的关系及其对呼吸系统疾病的滞后影响[J].兰州大学学报(自然科学版),2017,53(3):388-393.
- [34] MCGOVERN A,ELMORE K L,GAGNE D J,et al.Using artificial intelligence to improve real-time decision making for high-impact weather[J].Bulletin of the American Meteorological Society,2017,98(10).
- [35] RUIZ A,VILLA N.Storms prediction:Logistic regression vs random forest for unbalanced data[J].arXiv preprint arXiv2008.0804.0650.
- [36] GREGORY R H,RUSS S S.Money doesn't grow on trees, but forecasts do:forecasting extreme precipitation with tandem gores[J].Monthly Weather Review,2018,146(5).
- [37] 周康辉,郑永光,王婷波.利用深度学习融合 NWP 和多源观测数据的闪电落区短时预报方法[J].气象学报,2021,79(1):1-14.
- [38] 任萍,陈明轩,曹伟华,等.基于机器学习的复杂地形下短期数值天气预报误差分析与订正[J].气象学报,2020,78(6):1002-1020.
- [39] 刘程.深圳近地面风的时空特征研究[D].深圳:中国科学院大学(中国科学院深圳先进技术研究院),2020.
- [40] 刘新超,郭洁,陈婷,等.儿童呼吸系统疾病与气象要素的关系及其预测[J].高原山地气象研究,2015,35(1):92-96.
- [41] 张晓云,王德征,蔡子颖,等.天津市气象要素和大气污染物对呼吸道疾病影响分析[C]//第 32 届中国气象学会年会 S13.气候环境变化与人体健康,2015:290-299.
- [42] 张德山,刘燕,丁德平,等.京津地区儿童呼吸系统疾病医疗气象预报初步研究[J].气候与环境研究,2007(6):804-810.
- [43] 程一帆,张莹,王式功,等.人工神经网络在呼吸系统疾病急诊就诊人数预报中的应用[J].兰州大学学报(自然科学版),2014,50(1):75-79.
- [44] 黄开龙,林锦春,马盼,等.气象条件对深圳市罗湖区上呼吸道感染就诊人数的影响[J].干旱气象,2021,39(6):995-1005.
- [45] 贾子舟,张钰嘉,荣书玲,等.LSTM 神经网络模型在冠心病月度入院人数预测中的研究[J].中西医结合心脑血管病杂志,2021,19(18):3145-3148.
- [46] 查博.基于 LSTM 神经网络的创新创业发展预测与区域发展比较——以陕西与四川为对象[J].科技管理研究,2021,41(19):76-85.

Study of the Influence of Monsoon Change on Respiratory Patients and the Prediction Model in Shenzhen Area

WU Qianpeng¹, YIN Li², LI Xingrong³, SUN Yu⁴, HUANG Kailong⁵, SU Chunfang⁴,
WANG Shigong¹

(1.School of Atmospheric Sciences/Institute of Environmental Meteorology and Health, Chengdu University
of Information Technology, Chengdu 610225, China;

2.Meteorological Medical Research Center, Panzhihua Central Hospital, Panzhihua 617000, China;

3.Shenzhen Meteorological Bureau, Shenzhen 518040, China;

4.Climate Medicine Research Center, Second People's Hospital in Hainan Province,
Wuzhi Shan 572299, China;

5.Bureau of Meteorology of Shantou City, Shantou 515041, China;

6.Public Meteorological Service Center of China Meteorological Administration, Beijing 10081, China)

Abstract This paper used the data on the number of patients with respiratory disease, meteorological factors and pollutant concentration from 2015 to 2016 to construct a prediction model for the admission visits from respiratory disease by using BP artificial neural network and LSTM network. The results showed that the cold stress effect of winter monsoon would increase the incidence of respiratory diseases in related populations fluctuating from September to the peak of March before the transition from winter wind to summer wind in the following year. During the summer wind control period, the incidence of respiratory diseases among local residents fluctuated and decreased, which was 35% lower than that during the peak period. In addition, the main control factors of respiratory diseases were different. Compared with the two prediction models, on the whole, the LSTM network forecasting model has a higher accuracy rate in predicting the risk of respiratory diseases in Shenzhen, which can meet the business needs of health weather forecasting services.

Key words monsoon transition; respiratory diseases; meteorological and pollution conditions; artificial neural network; LSTM network; prediction model