

温华洋,朱华亮,马文周,等,2021.基于 Bayes 判别法的结冰现象资料序列订正[J].气象,47(9):1113-1121. Wen H Y, Zhu H L, Ma W Z, et al, 2021. Data sequence correction of icing weather phenomena based on Bayes discriminant method[J]. Meteor Mon, 47(9): 1113-1121(in Chinese).

## 基于 Bayes 判别法的结冰现象资料序列订正<sup>\*</sup>

温华洋<sup>1</sup> 朱华亮<sup>1</sup> 马文周<sup>2</sup> 邱康俊<sup>1</sup> 张苗苗<sup>1</sup> 刘 壮<sup>2</sup>

<sup>1</sup> 安徽省气象信息中心,合肥 230031

<sup>2</sup> 合肥工业大学,合肥 230009

**提 要:** 针对安徽省 81 个国家级地面气象站 1961—2018 年结冰现象资料序列,采用要素一致性、内部一致性、空间一致性等方法进行数据质量控制,基于质量控制后的正常年份数据进行 Bayes 判别模型训练,应用训练所得模型完成异常年份结冰数据的订正。结果表明:安徽省共有 38 个台站累计 84 年的年结冰日数质量控制检查异常,年结冰日数异常年份主要集中在 1961—1970、1988—1999 和 2015—2017 年,造成年结冰日数异常的原因有部分台站历史观测任务简化、气象台站分类调整以及地面气象观测业务改革等。利用 Bayes 判别法构建了多个结冰现象判别模型,经检验发现,模型 1 和模型 3 具有较高的判别正确率、命中率、TS 评分以及较低的误警率。考虑计算的简便性,选用模型 1 对异常年份结冰数据进行逐日订正。通过六安站、太和站和无为站异常年份结冰现象订正结果对比发现,基于 Bayes 判别法的结冰现象判别模型,对不同时间段内、不同原因造成的结冰现象观测记录异常的订正均较为合理,订正后的年结冰日数变化趋势更符合实际情况,表明采用 Bayes 判别模型订正结冰现象是合理、可行的。

**关键词:** 结冰现象,序列订正,Bayes 判别法

**中图分类号:** P413

**文献标志码:** A

**DOI:** 10.7519/j.issn.1000-0526.2021.09.007

## Data Sequence Correction of Icing Weather Phenomena Based on Bayes Discriminant Method

WEN Huayang<sup>1</sup> ZHU Hualiang<sup>1</sup> MA Wenzhou<sup>2</sup> QIU Kangjun<sup>1</sup> ZHANG Miaomiao<sup>1</sup> LIU Zhuang<sup>2</sup>

<sup>1</sup> Anhui Meteorological Information Centre, Hefei 230031

<sup>2</sup> Hefei University of Technology, Hefei 230009

**Abstract:** Using the methods of element consistency, internal consistency and spatial consistency, the data qualities of 81 national surface meteorological stations in Anhui Province from 1961 to 2018 are controlled. The Bayes discriminant model is trained based on the normal year data after quality control, and the icing data of abnormal year are corrected by the model. The results show that total 84 years' annual icing days from 38 stations in Anhui Province are abnormal, mainly concentrated in 1961—1970, 1988—1999 and 2015—2017, and fail to pass the quality control. The main reasons for these phenomena are the simplification of historical observation tasks of some stations, the adjustment of meteorological station classification, the reform of ground meteorological observation operation and so on. Then, by using the Bayes discriminant method, several discrimination models of icing weather phenomena are constructed. After verification, it is found that Model 1 and Model 3 have higher recognition accuracy, hit rate and TS score but

<sup>\*</sup> 国家自然科学基金项目(41805023)和安徽省气象局研究性业务科技攻关项目(YJG202003)共同资助

2020 年 5 月 8 日收稿; 2021 年 7 月 16 日收修定稿

第一作者:温华洋,主要从事气象数据产品研发. E-mail:wenhy12@163.com

通讯作者:朱华亮,主要从事气象资料分析与评估. E-mail:hualiangzhu@126.com

lower false alarm rate. Considering the simplicity of calculation, Model 1 is selected to correct the icing data of abnormal years day by day. In addition, through the comparison of the icing weather phenomena correction results of Lu'an, Taihe and Wuwei Stations in abnormal years, we find that the model based on Bayes discriminant method is more reasonable for the correction of the icing weather phenomena caused by different reasons in different time periods, and the variation trend of annual icing days after correction is more in line with the actual situation, which shows that the Bayes discrimination model is reasonable and feasible to correct the icing weather phenomena.

**Key words:** icing weather phenomenon, sequence correction, Bayes discriminant method

## 引 言

结冰(又称冻结)指露天水面冻结成冰,包括器皿中的水冻结成冰,是日常生活中比较常见的自然现象,但其通常与低温冷害相伴随,对人类的生产生活、动植物的生长有较大的影响(中国气象局,2003;唐熠等,2019;刘洪兰等,2014;马树庆等,2015;李刚等,2020)。如在冬季的雨雪天气,高速公路、机场跑道上会出现积雪或结冰现象,由此引发的交通事故、交通延误等,已成为社会各界广泛关注的热点问题(Wang et al, 2008; 丁国香等, 2018; 蒲晓虎等, 2014)。因此,气象部门对结冰现象的监测和预报非常重视,高质量、长序列的结冰资料有利于气候变化分析、农业气象和道路交通的预报与服务,也是开展有效灾害预防、应急救援等活动的科学依据(张志富等,2015;张庆奎等,2020;舒斯等,2019;白永清等,2016;Toms et al,2017)。

从 20 世纪 50 年代起,我国部分气象站就开始记录结冰现象,至今已形成了约 60 年的结冰现象观测资料,这些观测资料在农业、交通等领域的气象服务中发挥着重要作用,但同样存在着部分台站因历史观测任务简化、观测方式多次调整等诸多原因造成的数据质量不高的问题。余予等(2018)研究发现 1954—1979 年全国共有 662 个台站累计 6328 年实际未观测和记录结冰现象,1981 年前有 839 个台站累计 1453 年存在明显漏记结冰现象情况。如果直接使用这些原始观测资料而不进行有效地质量控制,对于结冰现象长期变化趋势分析等会得到不同的甚至相反的结论。因此,在使用结冰现象观测资料前需进行严格的质量控制和合理的数据订正。目前,国内外主要基于界限值、内部一致性和人工审核等方式对结冰现象观测数据进行检查(范邵华等,

2018;任芝花等,2007;2015;廖捷和周自江,2018;江益等,2018;刘小宁和任芝花,2005;陶士伟等,2009;闵锦忠等,2018),质量控制方法较为简单,如在历史资料的数字化过程中,仅对数据的完整性、规范性进行了检查,对出现的空白记录或者缺测记录无法进一步确认处理(范邵华等,2018);任芝花等(2015)设计的“台站-省级-国家级”三级质量控制只能对明显错误的数据进行缺测处理,无法确认某日是否出现结冰现象。在结冰现象观测数据订正方面,国内外相关研究较少,采用的线性回归法(余予等,2018)也仅能完成对年(或月)结冰日数的插补,较难认定某天是否发生结冰现象。为此,本文结合气温、地表温度等气象要素观测值对安徽省 81 个国家级地面气象站结冰现象观测资料进行了严格的质量控制,采用 Bayes 判别法构建了日结冰现象判别模型,并探讨了判别模型对结冰现象资料序列的逐日订正效果,旨在通过质量控制和数据订正形成高质量、长序列的结冰现象资料序列,为结冰现象的气候变化分析等提供科学依据。

## 1 资料与方法

### 1.1 资料与处理

本研究使用的资料包括安徽省 81 个国家级地面气象站 1961—2018 年的日结冰现象观测资料(当日是否出现)、日最低气温(单位:℃)、日最低地表温度(单位:℃),源自对应站点的地面气象月报表 A 文件。根据日最低气温和日最低地表温度观测值,对日结冰现象是否出现的资料序列进行初步质量控制。若某日出现结冰现象,且日最低气温和日最低地表温度均大于 10℃,则判定该日的结冰现象记录错误,将该日订正为无结冰现象;如果日最低气温

(或日最低地表温度)缺测,则在该日定时气温(或地表温度)中取最小值进行替代。依据《地面气象观测规范》(中国气象局,2003)统计得到安徽省 81 个国家级地面气象站 1961—2018 年的年结冰日数序列。

### 1.2 研究方法

为了更好地完成结冰现象资料序列的订正,采用要素一致性、内部一致性、空间一致性等方法对年结冰日数序列进行质量控制,方法和步骤如下:

第一步进行要素一致性检查。针对各站计算年结冰日数序列与年最低气温 $<0^{\circ}\text{C}$ 日数序列间的相关系数( $R$ )。选取相关系数 $R>0.9$ 且通过 0.05 的显著性水平检验的台站,利用这些台站的年结冰日数( $y$ )和年最低气温 $<0^{\circ}\text{C}$ 日数( $x$ ),求取两者间线性拟合公式。利用线性拟合公式估算各站每年的结冰日数,若某年观测值比估算值偏低 50%,则认为该年结冰现象数据异常,对其进行标记。

第二步进行内部一致性检查。基于完成要素一致性检查后的正确结冰现象数据,计算各站年结冰日数的均值( $\mu$ )和标准差( $\sigma$ ),若某年结冰日数位于区间 $[\mu-3\sigma, \mu+3\sigma]$ 之外,则认为该年结冰数据异常,对其进行标记。

第三步进行空间一致性检查。基于内部一致性检查后的正确结冰现象数据,为待检验站点(以下简称待检站)构建参考序列,参考站选取的步骤与标准为:(1)与待检站的海拔高度差不超过 200 m,地理环境相似;(2)与待检站的年平均气温序列相关系数 $R>0.6$ 且通过 0.05 的显著性水平检验;(3)与待检站的直线距离 $<100$  km,且距离最近的 5 个气象站。若选取的参考站数量不足 5 个,则以实际数量计算,若参考站数量为 0 个,则不进行空间一致性检查。计算待检站年结冰日数相对参考站平均年结冰日数的偏差,若某年的相对偏差超过 50%,则认为该年结冰数据异常,对其进行标记。

对经质量控制后的正常年份数据进行 Bayes 判别模型训练,应用所得的模型对异常年份结冰现象数据进行订正,其中 Bayes 判别模型训练方法和步骤(王斌会,2011;华连生等,2015)如下:

(1)计算训练样本集中日结冰现象不发生和发生的概率分别为  $P_0$  和  $P_1$ 。

(2)假设某日的相关气象要素观测值  $\mathbf{X}=(x_1,$

$x_2, \dots, x_n)'$ ,其中  $x_1, x_2, \dots, x_n$  表示日最低气温、日最低地表温度等。根据训练样本集计算该日结冰现象不发生和发生情况下观测到相应气象要素值  $\mathbf{X}$  的概率  $P(\mathbf{X}/Y_i), i=0,1$ ,其中  $Y_0$  代表日结冰现象不发生, $Y_1$  代表日结冰现象发生。

(3)根据 Bayes 定理,计算在相应气象要素观测值  $\mathbf{X}$  条件下该日结冰现象不发生或发生的概率  $P(Y_i/\mathbf{X})$ :

$$P(Y_i/\mathbf{X}) = \frac{P(\mathbf{X}/Y_i)P_i}{P(\mathbf{X}/Y_0)P_0 + P(\mathbf{X}/Y_1)P_1} \quad i = 0,1 \quad (1)$$

若  $P(Y_0/\mathbf{X}) = \max_{i \in \{0,1\}} P(Y_i/\mathbf{X})$ ,即结冰现象不发生的概率大于发生的概率,则认为该日无结冰现象;否则,认为该日有结冰现象。假设结冰未发生条件下气象要素的协方差矩阵和结冰发生条件下气象要素的协方差矩阵相同,且都服从多元正态分布  $N(\mu_i, \Sigma)$ ,则有:

$$\begin{cases} P(\mathbf{X}/Y_0) = (2\pi)^{-n/2} |\Sigma|^{-1/2} \cdot \\ \exp[-\frac{1}{2}(\mathbf{X}-\mu_0)'\Sigma^{-1}(\mathbf{X}-\mu_0)] \\ P(\mathbf{X}/Y_1) = (2\pi)^{-n/2} |\Sigma|^{-1/2} \cdot \\ \exp[-\frac{1}{2}(\mathbf{X}-\mu_1)'\Sigma^{-1}(\mathbf{X}-\mu_1)] \end{cases} \quad (2)$$

式中: $\mu_i=(\mu_{i1}, \mu_{i2}, \dots, \mu_{in})', i=0,1$ ;其中  $\mu_{i1}, \mu_{i2}, \dots, \mu_{in}$  表示第  $i$  个总体中各变量的均值。协方差矩阵  $\Sigma=(\sigma_{jl})_{n \times n}$ ,协方差逆矩阵  $\Sigma^{-1}=(\sigma^{jl})_{n \times n}$ ,其中  $\sigma_{jl}$  和  $\sigma^{jl}$  分别表示协方差矩阵及其逆矩阵中的第  $j$  行第  $l$  列元素, $j=1,2, \dots, n, l=1,2, \dots, n$ 。因为

$$\begin{aligned} \max_{i \in \{0,1\}} P(Y_i/\mathbf{X}) &\Leftrightarrow \max_{i \in \{0,1\}} [P(\mathbf{X}/Y_i)P_i] \\ &\Leftrightarrow \max_{i \in \{0,1\}} \ln[P(\mathbf{X}/Y_i)P_i] \end{aligned}$$

式中“ $\Leftrightarrow$ ”表示等价的意思,则有:

$$\begin{aligned} \ln[P(\mathbf{X}/Y_i)P_i] &= \ln P_i - \frac{n}{2} \ln(2\pi) - \\ &\frac{1}{2} \ln |\Sigma| - \frac{1}{2} \mathbf{X}'\Sigma^{-1}\mathbf{X} - \\ &\frac{1}{2} \mu_i'\Sigma^{-1}\mu_i + \mathbf{X}'\Sigma^{-1}\mu_i \end{aligned} \quad (3)$$

略去式(3)中与  $i$  无关的项,则可写为如下线性判别函数:

$$\begin{cases} Y_0 = \ln P_0 + a_{00} + a_{01}x_1 + \dots + a_{0n}x_n \\ Y_1 = \ln P_1 + a_{10} + a_{11}x_1 + \dots + a_{1n}x_n \end{cases} \quad (4)$$

式中:

$$a_{i0} = -\frac{1}{2} \boldsymbol{\mu}_i' \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}_i$$

$$a_{ij} = \sum_{l=1}^n \sigma^{il} \mu_{il}$$

$$i = 0, 1, j = 1, 2, \dots, n$$

若  $Y_0 > Y_1$ , 认为该日无结冰现象; 反之, 认为该日有结冰现象。根据每日相关气象要素的观测值(日最低气温、日最低地表温度等), 利用式(4)可以完成逐日结冰现象的订正。

### 1.3 Bayes 判识模型评价指标

为检验基于 Bayes 判别法建立的结冰判别模型判识效果, 采用准确率(PC)、命中率(POD)、误警率(FAR)和 TS 评分(TS)等指标进行评估, 具体计算公式如下:

$$PC = \frac{TP + TN}{TN + FN + FP + TP} \times 100\% \quad (5)$$

$$POD = \frac{TP}{FN + TP} \times 100\% \quad (6)$$

$$FAR = \frac{FP}{FN + FP} \times 100\% \quad (7)$$

$$TS = \frac{TP}{FP + FN + TP} \times 100\% \quad (8)$$

式中:  $TN$  表示人工观测和模型判识都无结冰现象的日数,  $TP$  表示人工观测和模型判识都有结冰现象的日数,  $FN$  表示人工观测有结冰现象但模型判识无结冰现象的日数,  $FP$  表示人工观测无结冰现象但模型判识有结冰现象的日数。

## 2 数据质量控制检查结果分析

计算各台站年结冰日数与年最低气温  $< 0^\circ\text{C}$  日数间的相关系数, 其中共有 39 个台站的相关系数  $R > 0.9$ , 且通过 0.05 的显著性水平检验。选用这 39 个台站的数据计算年结冰日数( $y$ )和年最低气温  $< 0^\circ\text{C}$  日数( $x$ )间的线性拟合公式为:

$$\hat{y} = 3.55 + 0.965x \quad (9)$$

式中  $\hat{y}$  表示年结冰日数估算值。式(9)及其回归系数均通过 0.05 的显著性水平检验, 利用建模数据进行回算发现, 估算值与观测值之间的平均绝对偏差为 4.2 d, 均方根误差为 5.5 d, 两者误差较小, 表明式(9)对年结冰日数的拟合效果较好, 故可采用该式对各站的年结冰日数进行要素一致性检查。通过式(9)的要素一致性检查并辅以人工校对发现, 安徽省共有 23 个台站累计 52 年的年结冰日数异常。在

内部一致性检查方面, 检查结果显示共有 2 个台站累计 2 年的年结冰日数异常。在空间一致性检查方面, 天柱山、岳西、九华山和黄山等 4 个台站因海拔高度较高, 没有符合参考站选取标准的相邻站, 未进行空间一致性检查, 对其余台站进行空间一致性检查并辅以人工校对发现, 安徽省共有 17 个台站累计 30 年的年结冰日数异常。质量控制结果表明, 1961—2018 年安徽省共有 38 个台站累计 84 年的年结冰日数异常。

图 1 给出了 1961—2018 年结冰日数质量控制异常台站数逐年分布, 从图中可以看出, 年结冰日数异常年份主要集中在 1961—1970、1988—1999 和 2015—2017 年。这主要是由于 20 世纪 60 年代早期对结冰现象的观测要求较为简单(中国气象局, 1961), 导致这类台站的结冰现象记录异常偏少, 甚至为零。随着《地面气象观测规范》的修改和逐步完善(宁惠芳等, 2014), 1970 年之后这一问题得到了明显地扭转。1988 年中国气象局制定实施的《地基气象探测系统发展方案》, 将部分国家级地面气象观测站(如太和站、怀远站和怀宁站)划分为观测辅助站(吴增祥, 2006), 对气压、天气现象等气象要素不做观测要求, 造成记录的结冰现象显著偏少或无记录, 1999 年后观测辅助站重新归类为国家一般气象站(以下简称一般站), 所以 1988—1999 年有 3~4 个台站的年结冰日数出现异常。2014 年中国气象局实施《地面气象观测业务改革方案》(中国气象局综合观测司, 2013), 将国家基准气候站(以下简称基准站)和国家基本气象站(以下简称基本站)夜间天气现象连续观测(20 时至次日 08 时)调整为 4 次整点(20、23、02、05 时)观测, 同时取消一般站夜间值守任务, 这就造成夜间出现短时间结冰现象而漏记的情况, 2018 年安徽省加强了对月报表 A 文件中结冰现象的人工质量控制工作, 因此, 2015—2017 年出现较多台站年结冰日数异常偏少现象, 而 2018 年未见数据异常现象。图 2 给出了安徽省年结冰日数检查异常的台站分布, 从图中可以看出, 在全省年结冰日数异常的 38 个台站中, 基准站和基本站共有 10 个站累计 13 年的年结冰日数异常, 一般站共有 28 个台站累计 69 年的年结冰日数异常, 异常站点及累计异常年数中一般站的占比都比较大, 这与一般站观测方式的多次调整有关。在空间分布上, 年结冰日数记录异常并无明显特征。

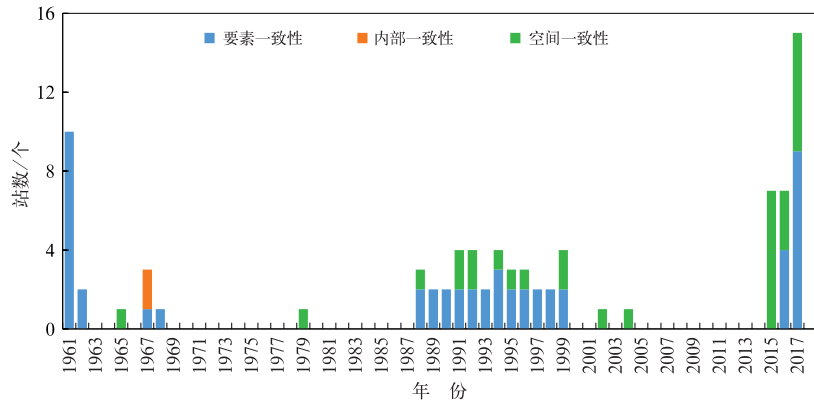


图 1 1961—2018 年安徽省年结冰日数异常台站数逐年分布  
 Fig. 1 Distribution of stations with abnormal annual icing days in Anhui Province from 1961 to 2018

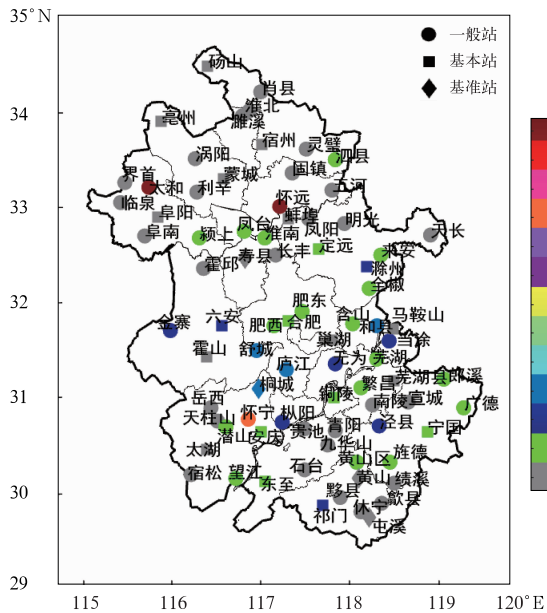


图 2 1961—2018 年安徽省各台站累计异常年数分布(单位:a)

Fig. 2 Distribution of cumulative abnormal years of stations in Anhui Province from 1961 to 2018 (unit: a)

### 3 结冰现象资料序列订正

#### 3.1 Bayes 判别模型建立

针对经过质量控制后的正确年份结冰现象数据,剔除日最低气温和日最低地表温度均大于 10℃ 的日数据,得到较易发生结冰现象的日数据,共有数据 1203734 d,随机选取 80% 作为模型训练数据集,剩余 20% 作为模型验证数据集。训练数据集中无结冰日数占比  $P_0 = 0.66$ ,有结冰日数占比  $P_1 =$

0.34。鉴于温度对结冰现象的影响,将日最低气温 ( $T_{min}$ )、日最低地表温度 ( $D_{min}$ ) 依次代入式(4)中建模,得到结冰现象判别模型 1、模型 2 和模型 3 如下:

模型 1(日最低气温判别模型):

$$\begin{cases} Y_0 = 0.6319 + 0.2726T_{min} \\ Y_1 = 0.5670 - 0.1496T_{min} \end{cases} \quad (10)$$

模型 2(日最低地表温度判别模型):

$$\begin{cases} Y_0 = 1.1136 + 0.1546D_{min} \\ Y_1 = 0.3349 - 0.1965D_{min} \end{cases} \quad (11)$$

模型 3(日最低气温和日最低地表温度判别模型):

$$\begin{cases} Y_0 = 23.5844 + 3.6126T_{min} + 0.7133D_{min} \\ Y_1 = 18.3872 - 6.6013T_{min} - 1.3021D_{min} \end{cases} \quad (12)$$

为考察模型 1~模型 3 对结冰现象的判别能力,利用验证数据集对各模型进行检验,其中验证数据集中无结冰日为 168179 d,有结冰日为 72432 d,共计 240611 d。表 1 给出了各模型的正确率、命中率、误警率、TS 评分等检验评价指标,模型 1 和模型 3 具有较高的判别正确率、命中率和 TS 评分,且模型的误警率较低,模型 2 的判别效果不如模型 1 和模型 3,表明可用模型 1 和模型 3 对日结冰现象进行判别,但两者判别效果差异不大,且模型 1 比模型 3 少一元变量,更便于计算,因此,选用模型 1 对质量控制检查异常年份结冰数据进行逐日订正。

#### 3.2 基于 Bayes 判别模型的结冰序列订正

针对数据质量控制标记的结冰现象记录异常年

份数据,利用模型 1 进行逐日判别。如果模型判别结果与人工记录结果不一致,则将该日结冰现象订正为模型判别结果;否则不进行订正。如果某日最

低气温缺测,则在该日定时气温中取最小值进行替代。

表 1 基于验证数据集的各模型判别效果

Table 1 Discrimination effect of each model based on validation data set

类型	TN/d	FP/d	FN/d	TP/d	PC/%	POD/%	FAR/%	TS
模型 1	164 848	9 533	3 331	62 899	94.7	95.0	4.6	83.0
模型 2	162 730	19 574	5 449	52 858	89.6	90.7	7.5	67.9
模型 3	165 380	9 462	2 799	62 970	94.9	95.7	3.9	83.7

通过上述方法对安徽省 38 个台站累计 84 年的异常年份数据进行订正,共订正了 2 949 d 的结冰现象记录,订正数据量占异常年份总数据量的 9.5%,占比较少。基于订正后的结冰现象资料序列重新进行要素一致性、内部一致性、空间一致性等检查,发现仍有 2 个台站累计 2 年的年结冰日数异常(表 2)。从表 2 中可以看出,望江站模型判别结冰日数相对参考站平均结冰日数明显偏少,导致该站 2017 年订正后的年结冰日数未通过空间一致性检验。经统计发现,人工记录结冰日数明显少于年最低气温  $<0^{\circ}\text{C}$  日数和年最低地表温度  $<-1^{\circ}\text{C}$  日数,而模型判别结冰日数更接近于年最低气温  $<0^{\circ}\text{C}$  日数和年最低地表温度  $<-1^{\circ}\text{C}$  日数,表明望江站 2017 年人

工观测结冰现象的记录中存在漏记的可能。淮南站 1999 年人工记录结冰日数和模型判别结冰日数相对于参考站平均结冰日数都明显偏少,所以该年结冰日数也未通过空间一致性检验,但人工观测结冰日数、模型判别结冰日数和年最低气温  $<0^{\circ}\text{C}$  日数三者间都比较接近,故认为当年结冰现象人工记录正常。总体而言,经 Bayes 判别模型订正后,38 个台站累计 84 年的异常年份数据中有 36 个台站累计 82 年通过质量控制检查,虽有 2 个台站累计 2 年的年结冰日数未通过检验,但模型判别结果比较符合台站自身的实际情况,表明经 Bayes 判别模型订正后结冰现象数据质量得到明显提升。

表 2 异常站点年结冰日数的质量控制结果

Table 2 The quality control results of revised annual icing days at abnormal stations

站名	异常年份	检查方法	人工记录结冰日数/d	模型判别结冰日数/d	年最低气温 $<0^{\circ}\text{C}$ 日数/d	年最低地表温度 $<-1^{\circ}\text{C}$ 日数/d	参考站平均结冰日数/d
望江	2017	空间一致性	1	5	5	10	17.6
淮南	1999	空间一致性	21	19	19	—	56

注:“—”表示缺测。

Note: “—” represents no data.

### 3.3 订正结果对比

以六安站、太和站和无为站为例,考察 Bayes 判别模型对质量控制检查异常年份结冰现象的订正效果,并与参考序列进行对比,其中参考序列的构建选用通过质量控制的站点数据,具体参见空间一致性检验参考站选取的方法与步骤。六安站的年结冰日数异常年份出现在 1967 年和 1968 年,这两年人工观测的年结冰日数分别为 23 d 和 8 d,利用模型 1 订正后的年结冰日数分别为 69 d 和 56 d,模型订正值与人工观测值差异较大,但与参考序列的年结冰日数更为接近(图 3a)。在年结冰日数变化趋势上,订正序列的趋势与参考序列的趋势更为接近,且均

为下降趋势,表明 Bayes 判别模型对六安站 1967 年和 1968 年的结冰现象订正效果较好。太和站在 1988—1999 年为观测辅助站,对结冰现象不做观测要求,造成 1988—1999 年的结冰现象未做记录,通过模型 1 订正后,1988—1999 年订正序列平均年结冰日数为 63.8 d,最低年结冰日数为 42 d,最高年结冰日数为 82 d,与参考序列的平均绝对差值为 6.5 d,最大绝对差值为 12.2 d,而 1961—1987 年年结冰日数原始序列与参考序列的平均绝对差值为 5.6 d,最大绝对差值为 15.4 d,2000—2018 年年结冰日数原始序列与参考序列的平均绝对差值为 7.1 d,最大绝对差值为 12.4 d,表明 1988—1999 年年结冰日数订正序列与参考序列的误差在合理范围

内。从变化趋势上看,订正序列的变化趋势与参考序列更接近(图 3b),且订正序列线性趋势线的拟合优度  $R^2$  要明显大于原始序列趋势线的  $R^2$ ,表明利用订正序列得到的趋势线更可靠。总体看来,Bayes 判别模型对太和站 1988—1999 年的结冰现象订正较为合理。图 3c 给出了无为站订正前后 1961—2018 年年结冰日数序列,无为站 2016 年和 2017 年人工观测的结冰日数分别为 12 d 和 8 d,相较于往年观测值和参考序列都明显偏少,订正后的结冰日

数分别为 26 d 和 21 d,与参考序列更为接近,也更符合序列变化趋势。

通过六安站、太和站和无为站异常年份结冰现象订正结果对比发现,基于 Bayes 判别法的结冰现象判别模型,对不同时间段、不同原因造成的典型的结冰观测记录异常的订正均较为合理,订正后的年结冰日数变化趋势更符合实际情况,表明采用 Bayes 判别模型订正结冰现象是合理、可行的。

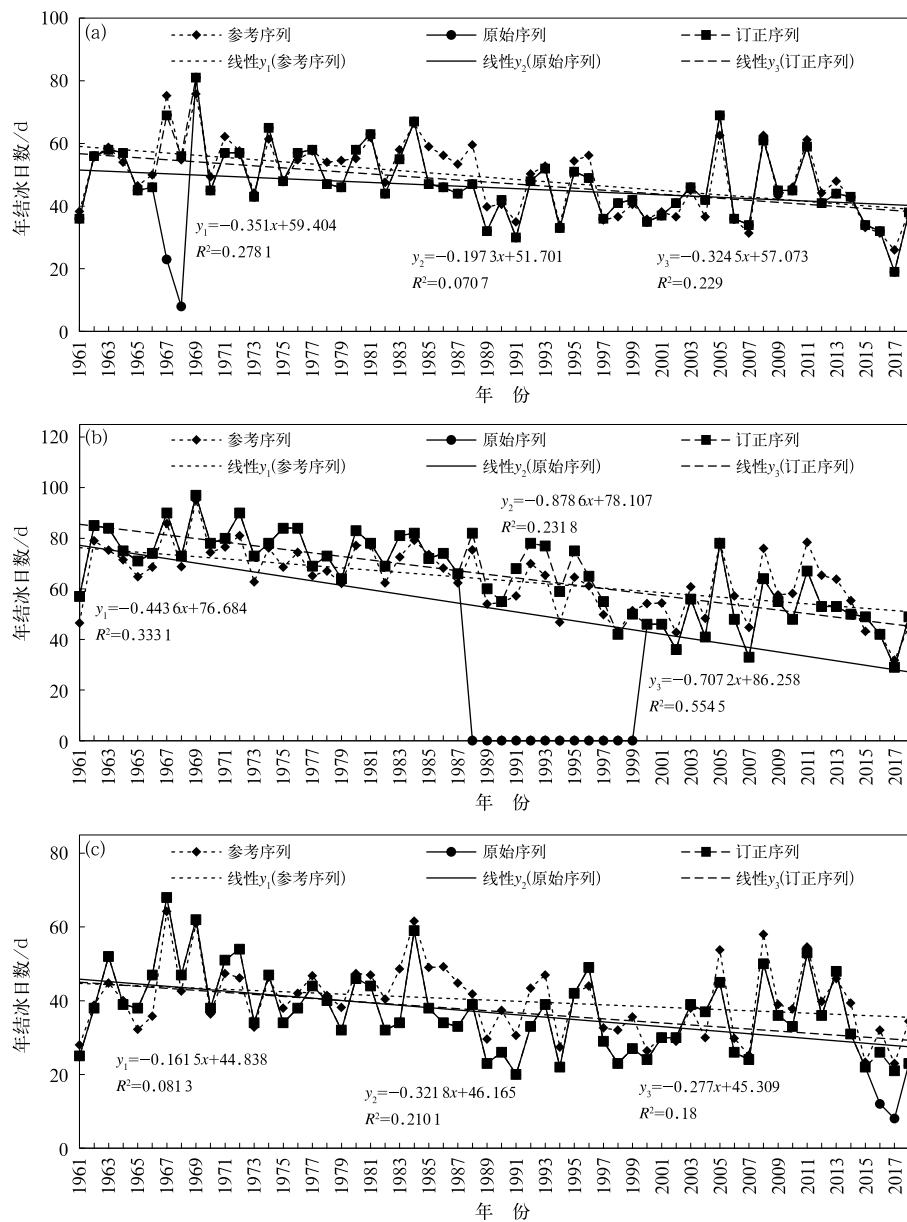


图 3 1961—2018 年六安站(a)、太和站(b)和无为站(c)年结冰日数序列  
 Fig. 3 Annual icing days series of Lu'an Station (a), Taihe Station (b) and Wuwei Station (c) from 1961 to 2018

## 4 结论与讨论

(1)本研究采用要素一致性、内部一致性、空间一致性等方法对安徽省 81 个台站 1961—2018 年年结冰日数序列进行更细致的质量控制,与余予等(2018)中质量控制方法相比,本研究所采用的方法更为严格,主要体现在:①本研究内部一致性检查时采用 3 倍标准差,而余予等(2018)采用 5 倍标准差;②本研究采用了空间一致性检查,而余予等(2018)未进行空间一致性检查。

(2)年结冰日数质量控制结果表明,安徽省共有 38 个台站累计 84 年的年结冰日数质量控制检查异常,年结冰日数异常年份主要集中在 1961—1970、1988—1999 和 2015—2017 年,造成年结冰日数异常的原因有部分台站历史观测任务简化、气象台站分类调整以及地面气象观测业务改革等。余予等(2018)对安徽省 1981 年之前的年结冰日数质量控制发现,共有 1 站累计 1 年未记录结冰现象,有 8 站存在 1~2 年的年结冰日数显著偏低。从图 1 可以看出,本研究发现的 1981 年之前的年结冰日数异常站数和累计年数要明显多于余予等(2018)的质量控制结果,这也佐证了本研究质量控制标准比余予等(2018)的质量控制标准更为严格。

(3)针对质量控制后的正确年份结冰现象数据集,采用 Bayes 判别法建立模型 1、模型 2 和模型 3。经检验发现,模型 1 和模型 3 各项评价指标均较好。考虑计算的简便性,选用模型 1 对质量控制检查异常年份结冰数据进行逐日订正。通过订正结果对比发现,基于 Bayes 判别法的结冰现象判别模型,对不同时间段内,不同原因造成的结冰现象观测记录异常的订正均较为合理,订正后的年结冰日数变化趋势更符合实际情况,表明采用 Bayes 判别模型订正结冰现象是合理、可行的。

(4)本研究发现部分台站历史观测任务简化、气象台站分类调整、地面气象观测业务改革等引发了安徽省部分台站的年结冰日数异常问题,这些问题势必在全国其他地区同样存在,余予等(2018)仅对 1981 年前的年结冰日数进行质量控制,就发现全国 1954—1979 年共有 662 个台站累计 6 328 年实际未观测和记录结冰现象,1981 年前有 839 个台站累计 1 453 年存在明显漏记结冰现象情况。因此,在使用全国结冰资料进行气候变化分析时,还需进行有效

的质量控制和序列订正。本研究所用方法对全国范围内的结冰资料的质量控制和订正提供了一种参考,但其质量控制和订正效果还有待于进一步探讨。

## 参考文献

- 白永清,陈城,何明琼,2016. 利用 Logistic 回归方法的高速桥(路)面结冰风险等级预警模型[J]. 气象科技,44(2):336-340. Bai Y Q, Chen C, He M Q, 2016. Warning model of risk degree for bridge/road surface icing of expressway based on Logistic regression[J]. Meteor Sci Technol, 44(2):336-340(in Chinese).
- 丁国香,刘承晓,周建平,等,2018. 气象条件对安徽高速公路运营的影响[J]. 干旱气象,36(5):858-863. Ding G X, Liu C X, Zhou J P, et al, 2018. Impact of meteorological conditions on expressway operation in Anhui Province[J]. J Arid Meteor, 36(5):858-863(in Chinese).
- 范邵华,余予,鞠晓慧,等,2018. 气象资料数字化进展及应用[J]. 气象科技进展,8(1):64-70,78. Fan S H, Yu Y, Ju X H, et al, 2018. A review for the progresses and applications of historical meteorological data rescuing and digitizing over China[J]. Adv Meteor Sci Technol, 8(1):64-70,78(in Chinese).
- 华连生,温华洋,朱华亮,等,2015. 基于 Bayes 判别法的霜生自动化观测模型探讨[J]. 气象,41(8):964-969. Hua L S, Wen H Y, Zhu H L, et al, 2015. Automated observation model for frost based on Bayes discriminant method[J]. Meteor Mon, 41(8):964-969(in Chinese).
- 江益,王立俊,羊清雯,等,2018. 南海区域站气象资料质量控制方法改进[J]. 气象科技,46(3):474-478,496. Jiang Y, Wang L J, Yang Q W, et al, 2018. Improvement of quality control method for real time meteorological data from regional automatic stations on South China Sea[J]. Meteor Sci Technol, 46(3):474-478,496(in Chinese).
- 李刚,马继望,梁湘三,2020. 2008 年 1 月中国南方低温雨雪期间异常阻塞高压事件的多尺度动力过程分析[J]. 气象学报,78(1):18-32. Li G, Ma J W, Liang X S, 2020. A study of the multiscale dynamical processes underlying the blocking high that caused the January 2008 freezing rain and snow storm in southern China[J]. Acta Meteor Sin, 78(1):18-32(in Chinese).
- 廖捷,周自江,2018. 全球常规气象观测资料质量控制研究进展与展望[J]. 气象科技进展,8(1):56-63. Liao J, Zhou Z J, 2018. Quality control of the global conventional meteorological observations: progresses and prospects[J]. Adv Meteor Sci Technol, 8(1):56-63(in Chinese).
- 刘洪兰,张俊国,阙龙凯,等,2014. 基于 BP 神经网络的张掖国家湿地公园水域结冰厚度预报模型[J]. 高原气象,33(3):832-837. Liu H L, Zhang J G, Que L K, et al, 2014. Forecasting model for ice thickness in Zhangye National Wetland Park watershed based on BP neural network[J]. Plateau Meteor, 33(3):832-837(in Chinese).
- 刘小宁,任芝花,2005. 地面气象资料质量控制方法研究概述[J]. 气象科技,33(3):199-203. Liu X N, Ren Z H, 2005. Progress in



- quality control of surface meteorological data[J]. Meteor Sci Technol, 33(3):199-203. (in Chinese).
- 马树庆, 袁祝香, 马力文, 等, 2015. 北方水稻低温冷害指标持续适用性检验与比较[J]. 气象, 41(6):778-785. Ma S Q, Xi Z X, Ma L W, et al, 2015. Test and comparison of sustainable suitability of meteorological indicators for rice cold damages in the North of China[J]. Meteor Mon, 41(6):778-785(in Chinese).
- 闵锦忠, 王晨珏, 贾瑞怡, 2018. 苏皖地面自动站资料的质量控制及结果分析[J]. 大气科学学报, 41(5):637-646. Min J Z, Wang C J, Jia R Y, 2018. Quality control and result analysis for surface AWS data in Jiangsu and Anhui Provinces[J]. Trans Atmos Sci, 41(5):637-646(in Chinese).
- 宁惠芳, 张志富, 范邵华, 2014. 浅析《地面气象观测规范》的演变及对资料质量控制的影响[J]. 青海气象, (3):53-56. Ning H F, Zhang Z F, Fan S H, 2014. Analysis of the evolution of "ground meteorological observation standard" and its influence on data quality control[J]. J Qinghai Meteor, (3):53-56(in Chinese).
- 蒲晓虎, 张春波, 马舒庆, 等, 2014. 地面气象观测中结冰自动化观测的研究[J]. 气象, 40(2):234-237. Pu X H, Zhang C B, Ma S Q, et al, 2014. Research on icing automation observation in ground meteorological observation[J]. Meteor Mon, 40(2):234-237(in Chinese).
- 任芝花, 熊安元, 邹凤玲, 2007. 中国地面月气候资料质量控制方法的研究[J]. 应用气象学报, 18(4):516-523. Ren Z H, Xiong A Y, Zou F L, 2007. The quality control of surface monthly climate data in China[J]. J Appl Meteor Sci, 18(4):516-523(in Chinese).
- 任芝花, 张志富, 孙超, 等, 2015. 全国自动气象站实时观测资料三级质量控制系统研制[J]. 气象, 41(10):1268-1277. Ren Z H, Zhang Z F, Sun C, et al, 2015. Development of three-step quality control system of real-time observation data from AWS in China[J]. Meteor Mon, 41(10):1268-1277(in Chinese).
- 舒斯, 熊守权, 陈英英, 等, 2019. 湖北省高速公路道路结冰预警模型[J]. 气象, 45(11):1589-1599. Shu S, Xiong S Q, Chen Y Y, et al, 2019. Warning model of road surface icing of expressway in Hubei Province[J]. Meteor Mon, 45(11):1589-1599(in Chinese).
- 唐熠, 周秀华, 郑传新, 等, 2019. 广西重大低温雨雪冰冻过程500 hPa信号场异常特征分析[J]. 气象, 45(10):1446-1456. Tang Y, Zhou X H, Zheng C X, et al, 2019. Analysis of abnormal characteristics at 500 hPa signal field in major cryogenic freezing rain and snow processes in Guangxi[J]. Meteor Mon, 45(10):1446-1456(in Chinese).
- 陶士伟, 仲跻芹, 徐枝芳, 等, 2009. 地面自动站资料质量控制方案及应用[J]. 高原气象, 28(5):1202-1209. Tao S W, Zhong J Q, Xu Z F, et al, 2009. Quality control schemes and its application to automatic surface weather observation system[J]. Plateau Meteor, 28(5):1202-1209(in Chinese).
- 王斌会, 2011. 多元统计分析及 R 语言建模: 第 2 版[M]. 广州: 暨南大学出版社: 113-153. Wang B H, 2011. Multivariate Statistical Analysis and Modeling for R Language[M]. 2nd ed. Guangzhou: Jinan University Press: 113-153(in Chinese).
- 吴增祥, 2006. 中国地面气象台站(1950—2004)沿革情况概述[C]//中国气象学会 2006 年年会“气象史志研究进展”分会场论文集. 成都: 中国气象学会: 87-97. Wu Z X, 2006. An overview of the evolution of China's surface meteorological stations (1950—2004)[C]//Proceedings of the "Research Progress of Meteorological History" branch of the 2006 Annual Meeting of China Meteorological Society. Chengdu: China Meteorological Society: 87-97(in Chinese).
- 余予, 任芝花, 孟晓艳, 2018. 中国结冰现象序列的建立及气候变化分析[J]. 高原气象, 37(2):553-559. Yu Y, Ren Z H, Meng X Y, 2018. Construction and climate variation analysis of icing weather phenomenon series over China[J]. Plateau Meteor, 37(2):553-559(in Chinese).
- 张庆奎, 项阳, 吉振明, 等, 2020. 近 55 年安徽省结冰现象气候特征及趋势分析[J]. 自然灾害学报, 29(6):218-226. Zhang Q K, Xiang Y, Ji Z M, et al, 2020. Analysis of climatic characteristics and trends for icing weather phenomenon in Anhui Province in recent 55 years[J]. J Nat Dis, 29(6):218-226(in Chinese).
- 张志富, 希爽, 余予, 等, 2015. 1961—2012 年中国 5 类主要冰冻天气的气候及变化特征[J]. 冰川冻土, 37(6):1435-1442. Zhang Z F, Xi S, Yu Y, et al, 2015. Climatic characteristics and variations of the gelivation weathers in China during 1961—2011[J]. J Glaciol Geocryol, 37(6):1435-1442(in Chinese).
- 中国气象局, 1961. 地面气象观测规范[M]. 北京: 气象出版社: 83-98. China Meteorological Administration, 1961. Specification for Ground Meteorological Observation[M]. Beijing: China Meteorological Press: 83-98(in Chinese).
- 中国气象局, 2003. 地面气象观测规范[M]. 北京: 气象出版社: 21-27. China Meteorological Administration, 2003. Specification for Ground Meteorological Observation[M]. Beijing: China Meteorological Press: 21-27(in Chinese).
- 中国气象局综合观测司, 2013. (2013-12-05) [2014-01-08]. 关于做好全国地面气象观测业务调整工作的通知: 气测函[2013]321号 [A/OL]. <https://wenku.baidu.com/view/55b1110bf01dc281e53af0c6.html>. Department of Integrated Observations, China Meteorological Administration, 2013. (2013-12-05) [2014-01-08]. Notice on the adjustment of national surface meteorological observation[A/OL]. <https://wenku.baidu.com/view/55b1110bf01dc281e53af0c6.html>(in Chinese).
- Toms B A, Basara J B, Hong Y, 2017. Usage of existing meteorological data networks for parameterized road ice formation modeling[J]. J Appl Meteor Climatol, 56(7):1959-1976.
- Wang Y G, Zhou Z J, Zhang Q, et al, 2008. Main meteorological disasters and their impacts on the economic and societal developments in China[J]. Int J Risk Assess Manage, 8(4):384-394.