

地面气象自动站数据审核规则库设计

赵洪武¹ 高杰² 张蕊³ 方晓²

(1. 辽宁省水利工程技术审核中心 沈阳 110006; 2. 辽宁省气象台 沈阳 110016; 3. 沈阳观象台 沈阳 110168)

摘要 基于地面气象自动站资料的存储格式,结合地面气象资料的质量控制规范,充分考虑气象审核人员的具体业务需求,分析了气象要素内部特征与各种要素间的约束关系,编制了基于数据库的质量控制审核规则体系结构,并据此开发了地面气象数据审核规则库管理系统和质量控制系统。

关键词 自动站 数据审核 规则库

随着地面气象自动站投入业务运行,多种气象要素实现了在线连续观测,气象资料的采集频率发生了很大变化。中国气象局根据2004年版的《地面气象观测规范》^[1]制定了新版“地面气象观测数据文件格式”(简称A文件)^[2],增加了质量控制信息。对新模式下的地面气象资料审核工作也提出了新的要求。原有的气象资料审核方法,没有使用其他变量的任何信息,忽略了同一观测变量在不同时间测量值之间的相关性(序列内相关),也忽略了观测变量之间的相关性(序列间相关),同时也忽略了不同测站观测变量间的相关性(空间相关),对故障识别缺乏灵敏性,已不能满足质量控制工作的需要。紧密结合地面气象资料的质量控制规范,设计基于关系数据库的审核规则库结构,是资料审核人员的工作需要,是提高气象资料服务质量的基础。

1 A文件的存储格式和数据类型划分

A文件是由地面气象观测中的台站参数、观测数据、质量控制、附加信息4个部分构成。

台站参数由区站号、纬度、经度、观测场海拔高度、气压感应器海拔高度、风速感应器距地(平台)高度、观测平台距地高度等12组数据构成,是台站属性数据库的重要组成部分。

观测数据是A文件的主要内容,由20个地面要素构成。从数据类型上可以划分为:定时值(根据方式位不同,可记录3,4,24测次的数值),日极大(高)值,日最小(低)值,日合计值,日极值时间5种类型;从数值类型上可以划分为数值数据(如气压、气温等)和现象类数据(文本数据,如云状、天气现象等)。

质量控制部分位于观测数据之后,分为质量控制码段和更正数据段。质量控制码表示数据质量的状况。根据数据质量控制流程,将其分为3个级别:台站级、省(地区)级和国家级。更正数据段是订正和修改数据的更正情况记录,更正数据段记录个数不限,每次订正或修改均添加到最后一条记录后面,不必考虑要素顺序。

附加信息部分由“月报封面”、“纪要”、“本月天气气候概况”、“备注”4个数据段组成,其数据是台站属性数据库的重要组成部分。

2 A文件质量控制的主要内容

根据《地面气象观测规范》要求,A文件的质量控制主要包括以下内容:日极值与定时值的比较检查,如日最低气压小于等于定时气压或大于等于日最高气压;要素的相关性检查,主要完成当一要素出现某一现象或值超过某一范围时,对与其相关的要素

进行审核;项目的逻辑检查,主要完成观测要素的数值逻辑审核及某些特定符号(如云状)的约定与内容检查;项目之间的差值检查,主要进行相关要素间变化协同性的检查,如5 cm与10 cm地温各定时记录的差值小于15.0℃等;气候极值比较检查,主要完成观测要素的数值定义域的强制检查,如日最高本站气压小于1 050.0 hPa变化趋势审核。随着观测频率的增加,气象要素变化过程的连续性、周期性、自相关特征非常明显,通过审核对变化异常点做出提示;附加信息的格式及信息变更审核,主要完成必报信息审核、数据格式审核及内容的模糊审核,同时附加信息数据是台站属性数据库的重要组成部分,通过与数据库信息的对比完成项目变更审核工作。

3 审核结果的信息等级

根据地面气象观测规范的要求,结合审核内容及约束条件的置信度水平,系统将质量审核结果的信息等级按照重要程度由低到高依次划分为3个级别:提示、警告和错误。用户可以根据自己对约束条件的理解,在设置审核规则时灵活设定其信息等级。

提示。该类约束不是质量控制规范中严格要求的,可以是用户根据某些气象要素的统计规律或经验确定的约束条件,如气象要素的变化趋势约束。该类约束的审核结果提醒审核人员引以注意。

警告。该类约束是对规范中相应级别限制的实现,如“极大风速大于等于17.0 m/s时,应有大风现象”。对于该类约束的审核结果,审核人员应该引起充分注意,并采用各种质量控制方法对该信息进行确认。

错误。该类约束是对A文件存储模式及必报项目等的严格审核,如字符长度、天气现象编码约束,最大值、最小值、日合计值与日定时值的逻辑约束等。该类约束的错误不仅影响系统对当前字段的解析,而且可能影响系统对此之后整个A文件的审核,审核人员必须按照约束的反馈信息对该数据进行相应修改。

为了使质量控制系统反馈信息清楚、明了,系统对每一疑误数据应只提示一条审核信息。同时为了突出重点,系统根据审核人员对该疑误数据的修改情况优先反馈最高约束级别的审核信息。

4 审核规则库涉及的主要内容

根据前面A文件数据类型划分、文件审核主要内容、审核结果的信息等级分析,审核规则数据库中主要应该包括如下内容。

收稿日期:2005-04-19;修订日期:2005-08-30

4.1 气象要素名称

通过下拉列表选择欲建立约束的气象要素名称。该名称应该是惟一表示该要素的数据段数及数据类型。在一条约束条件中可以包含2个气象要素名称,以完成相关要素间的约束审核。

4.2 约束信息

为一数学表达式,通过数学解析可以确定为一个约束信息集合。约束信息字段中可以输入逻辑运算符及约束值。逻辑运算符可为>、>=、<、<=、或=,约束值为一整数或为云状及天气现象代码,如“>=5”、“42,31,39”等。

4.3 启用状态标识

通过下拉列表确定该条审核约束是否启用。

4.4 约束类型

通过鼠标确定该条审核约束的信息显示类型(用于审核结果显示窗口)。

5 审核规则划分与审核规则库的建立

综合质量控制规范要求、A文件气象要素之间的联系以及质量审核规则约束类型,将审核规则划分为如下5类,对于每种类型的审核规则分别设计成表格结构。数据库结构如图1。

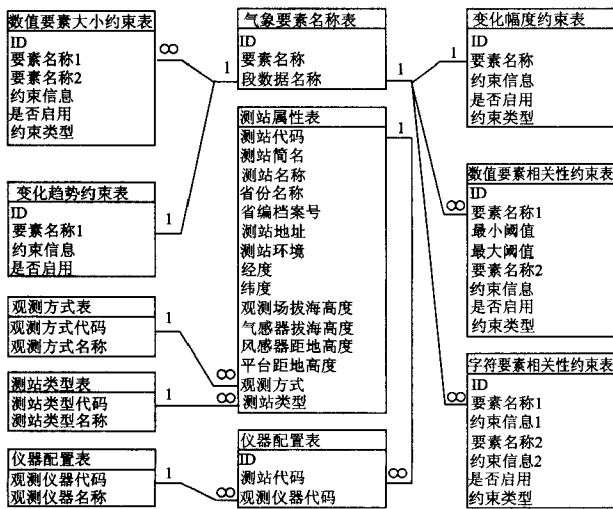


图1 审核规则数据库结构

5.1 站属属性数据审核

站属属性数据主要包括A文件首行的台站参数数据和附加信息数据。为了保证站属属性数据的正确性,系统在数据格式审核的同时还需要对数据的一致性进行审核,并据此掌握测站的沿革变化。站属属性数据审核主要通过数据库属性对照审核来完成。

5.2 数值要素(变化幅度)约束

有些数值类气象要素的变化过程是连续且平稳的,如地温的变化过程,因此,确定该类数据的变化幅度约束是数值连续性审核的一个主要方法。根据对历史数据的统计分析,深层地温1.6m的变化率应该小于0.4℃/h,深层地温3.2m的变化率应该小于0.2℃/h。数值要素(变化幅度)约束的数学表达式:

$$|X_{i,j} - X_{i,j-1}| \leq C \quad (1)$$

式(1)中: $X_{i,j}$ 为第*i*天第*j*个测次的要素值; $i = 1, 2, \dots, 31$; $j = 1, 2, 3, \dots, 24$;C为由数理统计方法确定的气象要素X的允许变化率集合。

5.3 气象要素(相关性)约束

气象要素间的变化是相关的,各气象要素间的变化具有

很明显的协同性,如风速与天气现象、云状与云量间等都存在紧密的联系。传统的阈值检测法在确定每个观测变量的阈值时,并没有使用其他变量的任何信息,忽略了观测变量之间的相关性。气象要素(相关性)约束可理解为:当气象要素出现某些现象或者变化到达一定程度时,与之相关联气象要素应该存在协同性变化。该约束的数学表达式:

$$\left. \begin{aligned} X_{i,j} &= a \\ h &= f_1(j) \\ k &= f_2(h) \\ Y_{i,j} &\subseteq C \end{aligned} \right\} \quad (2)$$

式(2)中: $X_{i,j}$ 为第一个气象要素第*i*天第*j*个测次的要素值; a 为气象要素 $X_{i,j}$ 的变化集合; h 为气象要素一第*j*个测次的时间; k 为气象要素二*h*时对应的测次序号; $Y_{i,k}$ 为与 $X_{i,j}$ 关联要素在第*i*天第*k*个测次附近一定时间内的要素值集合;C为约束信息集合。

5.4 数值要素间(大小关系)约束

有些数值类气象要素的变化过程间具有很强的协同性,并表现为对应时刻数值间的大小关系约束,如本站气压与海平面气压、不同深度地温的变化过程等。根据对历史数据的统计分析,就可以确定其约束信息。数值要素(大小关系)约束的数学表达式:

$$\left. \begin{aligned} X_{i,j} - Y_{i,k} &\subseteq C \\ h &= f_1(j) \\ k &= f_2(h) \end{aligned} \right\} \quad (3)$$

式(3)中: $X_{i,j}$ 为气象要素X第*i*天第*j*个测次的要素值; h 为气象要素一第*j*个测次的时间; k 为气象要素二*h*时对应的测次序号; $Y_{i,k}$ 为气象要素Y在第*i*天第*k*个测次的要素值;C为约束信息集合。

5.5 要素(变化趋势)约束

A文件可以比较真实地反映出气象要素的变化过程。由于受季节及日夜交替影响,使气象要素的变化过程具有显著的周期性和自相关性。传统检测方法忽略了同一观测变量在不同时间测量值之间的相关性,同时由于周期性因素的影响,对故障缺乏灵敏性。系统采用单变量多尺度小波变换(DWT)^[3]气象要素的变化趋势审核。约束条件式:

$$\left. \begin{aligned} X &\stackrel{DWT}{\Rightarrow} Y \\ Y_{i,j} &\subseteq f(Y) \end{aligned} \right\} \quad (4)$$

式(4)中: X 为气象要素系列,在A文件中为月观测系列;Y为气象要素X经DWT变换后的系列; $f(Y)$ 为根据变换后的系列Y确定的变化趋势阈值。

6 结语

根据前述的审核规则库体系,我们编制了地面气象数据审核规则库管理系统和质量控制系统。系统采用ACCESS数据库存储审核规则,实现了审核规则维护、数据质量审核控制、缺测信息统计、质量控制信息自动维护、平行观测数据评估等功能。系统于2004年在辽宁、吉林等几个省份开始试运行,明显提高了工作效率,保证了审核工作的正常进行,取得了较好的效果。

参考文献

- 1 中国气象局.地面气象观测规范.北京:气象出版社,2003.
- 2 中国气象局.地面气象观测数据文件和记录簿表格式.北京:气象出版社,2004.
- 3 程正兴.小波分析算法与应用.西安:西安交通大学出版社,1998.