

国家级地面自动站 A 文件质量 控制方法及软件开发

王新华^{1,2} 罗四维¹ 刘小宁² 任芝花² 孙化南²

(1. 北京交通大学, 北京 100044; 2. 国家气象信息中心)

提 要: 地面自动站 A 文件是以站月为单位的气象观测综合资料, 具有要素多、信息量大的特点。为了实现有效的质量控制, 研究了更为全面的国家级的质量控制方法并遵循软件工程理论, 采用模块化的设计方法完成了应用软件的开发, 这对保证我国地面自动站观测数据的质量起到了重要作用。

关键词: 国家级 地面自动站 质量控制

Quality Control and Software Development about A-format Data of National Basic Automatic Weather Station

Wang Xinhua^{1,2} Luo Siwei¹ Liu Xiaoning² Ren Zhihua² Sun Huanan²

(1. Beijing Jiaotong University, Beijing 100044; 2. National Meteorological Information Centre)

Abstract: National Basic Automatic Weather Station (AWS) data of A-format is a kind of integrated observation data, which includes kinds of elements and generous information. For realizing the effective quality control, more comprehensive quality control method is studied. According to software engineering theory, application software is developed with modularization design method. It is of great importance to the quality of National Basic Automatic Weather Station (AWS) data.

Key Words: automatic weather station (AWS) quality control software

引 言

随着我国地面气象自动站建设的发展,气象信息中心从 2005 年正式收集“地面气象观测数据文件”A 文件^[1]。A 文件是以站月为单位的综合观测资料,包括气压、气温、湿球温度、水汽压、露点温度、相对湿度、云量、云高、云状、能见度、降水量、天气现象、蒸发量、积雪、电线积冰、风、地温、冻土、日照等要素的定时或日值资料。近年来我国资料工作者对地面资料的质量控制流程及方法研究已有进展^[2]。我国自动站资料在报送国家级之前,台站和省级已经进行了质量控制,因此在国家级质量控制中与台站和省级即有区别又能发挥国家级的优势是十分重要的。为此研制了国家级地面自动站 A 文件质量控制方法和相应的软件,期望台站技术人员和气候研究者对我国自动站资料的质量控制状况有更进一步的了解。

1 质量控制方法

对自动站资料的质量控制,与一般的质量控制原理是一致的。其主要思路是:根据气象学、天气学及气候学原理,依据《地面气象观测规范》^[3]和数据文件格式的规定,进行气象观测数据序列的技术性、合理性检验。同时结合各站的气候背景,由程序自动对观测数据进行检查,找出不合理、不正常的错误或可疑记录,达到质量控制的目的。

1.1 台站参数检查

A 文件的台站参数必须保证完全准确,台站参数检查是检查几组参数在数据文件中与在最新台站参数文件中是否一致。通过检查的文件继续进行下一步的检查;未通过的文件需人工审核订正。

1.2 气候学界限值与要素允许值范围检查

气候学界限值检查是检查要素值是否在该要素的气候学界限值、要素允许值范围之内。气候学界限值是指从气候学角度不可能出现的临界值,要素允许值范围是指气象要素值允许出现有明确规定的范围。没有通过该项检查的数据被视为错误数据。

1.3 气候极值检查

气候极值检查是检查某要素值是否超过该站该要素该月历史上出现过的最大值和最小值。要素极值是随地理区域和季节的不同而有变化的,本系统的台站极值参数来源于各台站不同要素项目累年(1951~2000 年)各月最大值(MAX)或最小值(MIN)。需要注意的是,随着气候的变化,一些气候极值被突破,因此对检查超出气候极值范围的数据要在人工审查的基础上,判断该数据是否正确,并及时修改原极值参数。

1.4 内部一致性检查

要素项目间一致性检查是依据一定的气象学原理,对观测资料中某些物理特性关联的气象要素或项目之间是否符合一定规律进行的检测。由于每个单项检查都涉及多个要素或项目,检查发现问题时并不能确定是其中哪个要素或项目数据有误,因此,只能将相关的数据均作为可疑数据。内部一致性检查可分为两种:

第一种是同一时刻相同要素不同项目之间的一致性。例如:风向(dd)和风速(ff)的一致性检查,总云量必须大于等于低云量等。

第二种是同一时刻不同要素之间的一致性。各种要素从不同的侧面描述一个测站的天气气候特征,因此同一时刻不同要素之间存在不同程度的相关,例如:气温与湿球温

度一致性检查,气温应大于、等于湿球温度。

特别应当说明的是,对要素一致性相关联的参数范围的控制要合理设定,可根据实际数据进行调整,以保证检查出的数据在一个合理的范围之内。

1.5 时间一致性检查

时间一致性检查是指与要素时间变化规律性是否相符的检查。采用下述方法:

(1) 连续 24 小时定时观测数据相同,则有关数据可疑(缺测除外)。

(2) 要素日际变化分布界限值检验。

对于具有一定日变化规律的要素来说,各时次要素值的日际变化具有类似的分布特征。据此研究采用的方法是:用月内某时次前后各 1 个时次内的所有资料组成一个序列,计算统计序列的中值、加权均值和加权标准差,若被检数据与加权均值的绝对差值大于等于 3 倍加权标准差的值则认为可疑记录;若连续出现 5 次差值大于等于 3 倍的加权标准差的情况,则需考虑是否受到天气过程的影响。这是区别于台站级和省级的检查之一。

1.6 空间一致性检查

经过要素气候界限值、气候极值、内部一致性、时间一致性检查的有疑误的要素值进行空间一致性检查,对于检查出来的错误数据进行人工判断,一般可以通过显示等值线图等辅助方法确认数据正确与否。Madsen 方法是进行空间一致性检查的方法之一,也是与台站级和省级有区别的检查之一。

1.7 质量控制综合判断

对要素值进行气候界限值(要素允许值范围)检查、极值检查、内部一致性检查、时间一致性检查、空间一致性检查的结果,

依照未通过检查的影响程度和次数,经过质量控制综合判断法^[4]确定数据是正确、可疑或错误的结论意见。

2 软件设计开发

地面自动站 A 文件国家级质量控制软件开发遵循软件工程理论^[5],以工程化的原则和方法来指导软件的整个开发过程。

首先,根据以上介绍的方法制定了详细的质量控制方案和执行的业务流程即需求分析;其后进行了系统的详细设计;再进入编码过程;最后是软件的测试和结果分析。

2.1 质量检查软件的需求分析

在软件工程中,需求分析是一个相当重要的阶段,它准确回答了“系统需做什么”的问题。在需求分析中,确定了软件的数据流、功能要求、运行环境要求和控制方式,它们无疑都是下一步设计的基础,而且需求分析还能提供一份系统建成后评估其质量的依据。

2.1.1 功能要求

(1) 按质量控制方法实现各项检查功能。

(2) 输出疑误报告作为人工审核的依据,并嵌套在微机端人机交互界面中。一条疑误信息作为一个记录,按要素代码、日期、时次、相关要素、疑误信息的顺序输出。

(3) 统计 A 文件基本项目缺测数、缺测率及疑误数、疑误率。

(4) 在 A 文件中被检查项目的规定位置标定最终质量控制码,并整理疑误信息,按省生成向省里下传的反馈文件。

2.1.2 运行环境要求

国家气象信息中心每月分三次收集全国各省上传的 A 文件,软件既需要在“国家

级存储检索系统”的 UNIX 服务器上实时运行；也要适应在微机上运行，待每月全部资料收集后，还需以人机结合的方式进行人工审核。因此对软件的设计提出了更高的要

求。

2.1.3 质量检查软件数据流程

图 1 以一个数据文件的质量检查为例，给出其数据流程。

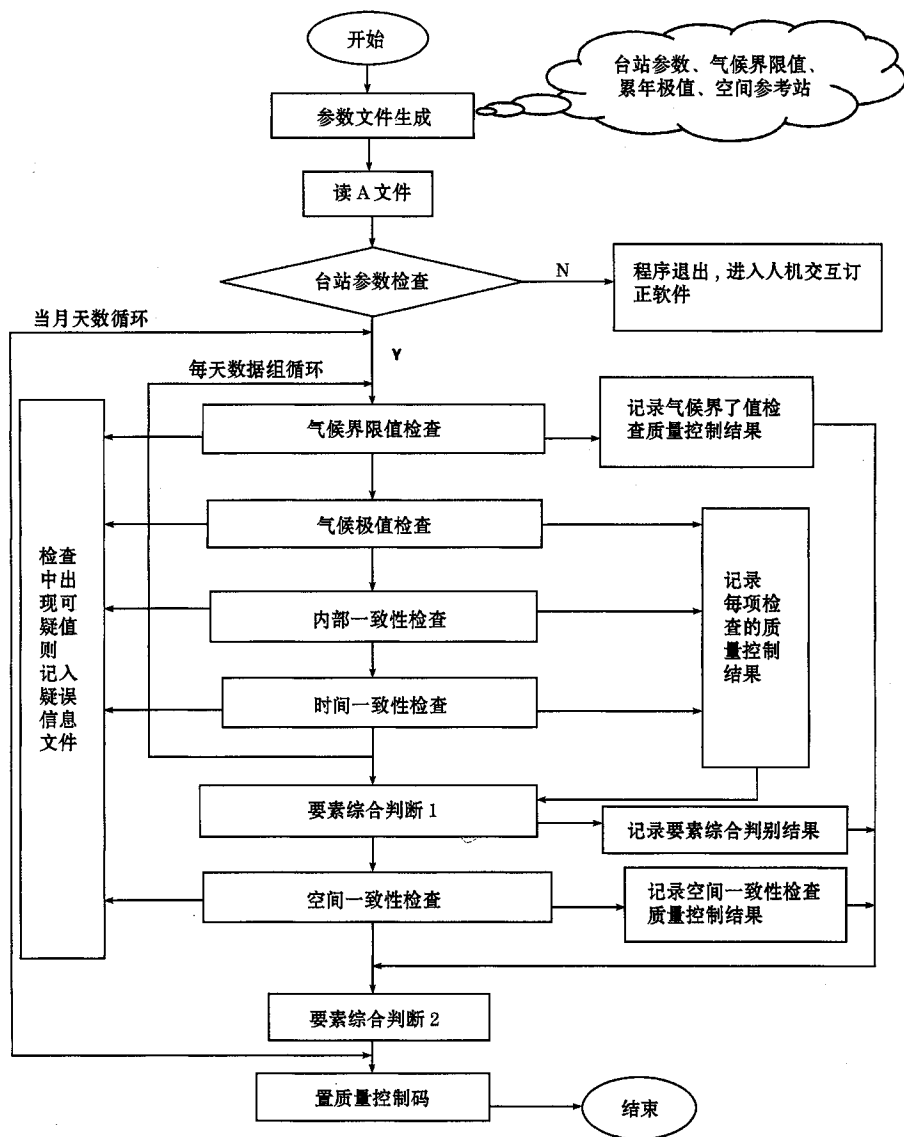


图 1 A 文件质量检查数据流程

2.1.4 控制流程

软件系统在运行控制上实现跨平台设

计，各省每月分三次上传 A 文件，在 HP UNIX 服务器上定时接收，并在固定时间内

对文件自动、批量地进行质量检查，确定质量控制码并生成疑误信息；在微机端安装有人机交互审核系统及经移植的质量检查执行模块，可以很方便地调用服务器上的疑误信息文件进行必要的人工审核，审核员还可以多次对资料数据进行修改和调用质量检查程序。

2.2 系统设计

系统设计主要包括系统结构设计和体系结构设计。

2.2.1 系统结构设计

采用分层模块化设计方法，将软件分成若干功能独立的模块（子模块），每个模块



图2 A文件质量检查控制流程

完成一个子功能。过程、函数、子程序等都可作为模块。模块集合起来组成一个整体，就可以完成系统的全部功能要求。模块化设计的显著优点是降低了系统的复杂性，使得系统更易于修改和实现，而且可以实现并行开发，从而提高了软件的生产效率。在模块划分过程中，需要具体分析和思考怎样划分更合理和便于集成实现，图3是质量检查软件系统的系统结构图。

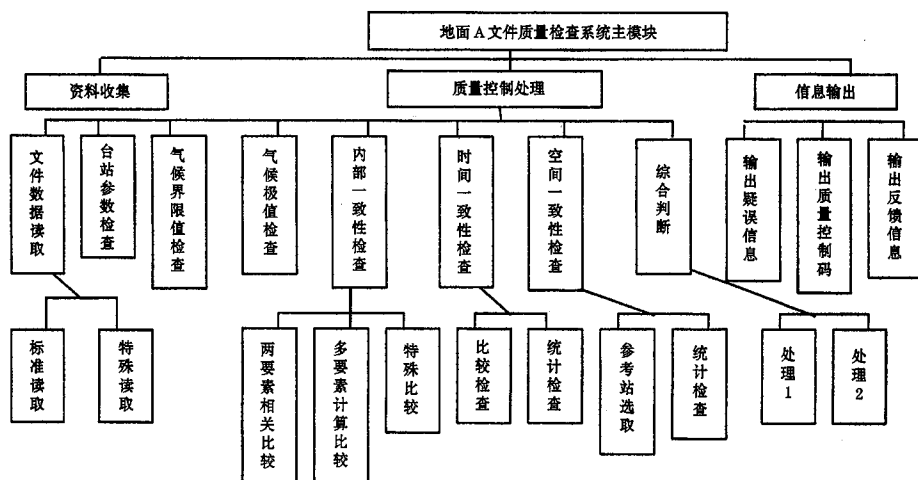


图3 质量检查软件系统结构

资料收集模块：定时自动接收各省上传的 A 文件，存入相应磁盘目录。

文件数据标准读取模块：实现组数固定的要素读取。

文件数据特殊读取模块：实现不定长的要素读取，如云状、天气现象。

台站参数检查模块：完成台站参数检查。

气候界限值检查模块：完成气候界限值

检查。

气候极值检查模块：完成气候极值检查。

两要素相关比较模块：在内部一致性检查中实现两要素或同一要素两项目的相关检查。

多要素计算比较模块：在内部一致性检查中实现要素与其计算值之间的检查。

特殊比较模块：在内部一致性检查中实

现云状及天气现象同其它要素的相关检查。

时间一致性比较检查模块：实现前后时段比较和全月缺测是否合理的检查。

时间一致性统计检查模块：按统计方法实现时间一致性的检查。

空间一致性选站模块：按规则对每个台站整理出与其相关的参考台站。

空间一致性统计检查模块：采用 Madsen 方法进行统计检查。

综合判定模块 1：对极值、内部、时间一致性检查结果进行综合分析和运算处理。

综合判定模块 2：对气候界限值、综合判别 1、空间一致性检查的结果进行综合分析，得到各要素最终的质量控制码。

输出疑误信息：以上检查中的每个疑误情况作为一条记录输出，可直接进入人工审核界面。

输出质量控制码：将最终质量控制码写入 A 文件，替换初值 9。

输出反馈信息：对各台站的疑误信息以省为单位进行整理，并向各省上传反馈。

2.2.2 体系结构设计

采用客户机/服务器结构，服务器端完成数据收集和质量检查，客户端完成人工审核。

2.2.3 开发平台

软件开发平台是 UNIX 服务器，操作系统是 hp-ux 11，采用标准 C 编译环境；软件同时可以移植到微机平台，VC++ 编译环境。

2.3 软件测试和结果分析

在编码基本结束后对软件运行结果进行了具体的测试和分析。我国地理环境差异较

大，因此针对于一些处理结果进行了更深入的研究和比较，部分修改质量检查方案。如在一致性检查中涉及的参数，有些控制范围较小，造成可疑记录多，经具体资料分析，修改了原方案中的参数，使数据的准确度控制在一个合理的范围内。并同时完善软件，最终形成一套更合理、科学、实用性强的软件系统。

3 小 结

地面气象自动站 A 文件国家级质量控制系統全面地对观测数据进行了质量控制。遵循软件工程理论，软件采用结构化模块设计，层次清晰，便于软件的管理、维护；软件功能扩充能力较强；而且采用标准编译环境，便于软件的移植。此软件生成的疑误信息文件可以再进入人机交互系统，进行必要的人工审核。该软件的开发和使用，全面实现了计划方案的质量控制要求，对保证我国地面自动站观测数据的质量起到了重要作用。

参考文献

- 1 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，2005：18~46.
- 2 刘小宁，任芝花. 地面气象资料质量控制方法研究概述. 气象科技，2005，33（3）：199~203.
- 3 中国气象局. 地面气象观测规范. 北京：气象出版社，2003：133~134，137~138.
- 4 王伯民. 基本气象资料质量控制综合判别法的研究. 应用气象学报，2004，15（增刊）：50~59.
- 5 王庆育. 软件工程. 北京：清华大学出版社，2004：169~283.