

海洋站观测资料的质量控制方法及其应用

杨扬, 苗庆生, 韦广昊, 董明媚, 东成

(国家海洋信息中心 天津 300171)

摘要: 文章阐明开展海洋站观测资料质量控制的必要性和重要性, 以海洋站观测延时(月报)资料为例, 详细论述其质量控制的对象、流程和参数以及包括日期检验、位置检验、格式检验在内的 19 种质量控制方法, 同时概述海洋站观测资料质量控制软件, 最后提出下一步开展海洋站观测资料的均一化检验和订正以及建立多模式检验系统和均一化数据集的研究重点, 以期为海洋和气候变化研究提供科学准确的第一手资料。

关键词: 海洋站; 海洋观测; 观测数据; 质量控制; 均一化

中图分类号: P71

文献标志码: A

文章编号: 1005-9857(2017)10-0109-05

Quality Control Methods and Application for the Oceanic Station Observed Data in the Delayed Mode

YANG Yang, MIAO Qingsheng, WEI Guanghao, DONG Mingmei, DONG Cheng

(National Marine Data and Information Service, Tianjin 300171, China)

Abstract: This paper expounded the necessity and importance of quality control for the oceanic station observed data. Taking the data of the delayed mode (monthly) of the oceanic station as an example, the object, the process and the parameter of the quality control were discussed in detail, including 19 tests in sum such as the date test, the position test, and the format test. At the same time, the paper analyzed the quality control software of the marine observed data. The next steps were put forward to carry out a variety of tests and correction methods combined with the use of multi-mode test system, which would make the results more reasonable and scientific, and also could be used to provide accurate first-hand information for the coastal climate change researches.

Key words: Oceanic station, Marine observation, Observed data, Quality control, Homogeneity

1 引言

海洋站观测获得数据资料的长期性、连续性和近岸性与其他观测方式无法比拟和替代的, 其实用

价值和应用意义不言而喻。尤其是在需要长时间序列资料时, 海洋站观测资料的重要性和宝贵性就更加明显。然而海洋站观测资料随时间的变化相

收稿日期: 2017-03-01; 修订日期: 2017-08-17

作者简介: 杨扬, 助理研究员, 硕士, 研究方向为水文气象资料的管理和质量控制

通信作者: 苗庆生, 工程师, 研究方向为水文气象资料的管理和质量控制

当显著且复杂,有些变化不是单独存在而是相互叠加,从而加大了观测要素随时间变化的显著性^[1]。因此,在对海洋站观测资料进行研究分析之前,必须对其进行必要的质量控制和人工审核,及时发现资料中的错误数据、可疑数据或异常数据信息,并根据提示对数据进行修改和查证处理,从而将资料中的错误减到最低,提高资料质量,确保数据的可靠性、代表性和可比性。

已有很多专家学者对海洋站观测资料的质量控制做过很多研究,但多数集中于地面气象资料^[2-7],对水文资料的研究鲜有报道,仅有李学坤等^[8]于1997年对海洋水文气象资料实时数据的质量控制做过分析研究。本研究以海洋站观测延时(月报)资料为例,详细论述其质量控制技术和方法,同时概述海洋站观测资料质量控制软件,并提出下一步研究重点。

2 质量控制技术

2.1 质量控制对象

海洋站观测月报资料共有4种类型10种文件。①温盐:T011定时温盐海发光和T012逐时温盐,共2种;②潮汐:T021逐时潮位和高低潮位潮时、T022 5 min潮位以及T023 1 min潮位,共3种;③波浪:T031定时(逐时)波浪,共1种;④气象:T051定时温压湿和逐时风等气象要素、T052逐时温压湿和逐时能见度、T053 10 min风速风向以及T054 1 min温压湿、风速风向和降水,共4种。

2.2 质量控制流程

海洋站观测月报资料的质量控制流程为:读入质控参数—检查站代码—检查资料的年月是否合法(按升序排列)—检查经纬度和表头各项观测参数—检查记录类型和观测日期—检查观测要素值是否合法(有无非法字符、是否在合理值范围内)—检查观测方法是否正确—检查仪器代码是否正确—检查质量符是否正确—观测要素范围检验—气候特性检验—相关性检验—极值检验—其他质量控制方法检验—检查说明记录是否合法—质量控制综合分析—数据质量标识。

在质量控制过程中,可根据观测资料类别、观

测要素和用户需求的差异对上述流程环节进行增减。

2.3 质量控制参数

质量控制参数通常包括通用质量控制参数和单站质量控制参数。其中,通用质量控制参数是各海洋站在进行质量控制时均可使用的参数;单站质量控制参数是从各海洋站历年资料中经大量统计计算选取的、专门提供给对应海洋站进行质量控制的参数。

3 质量控制方法

依据海洋学、气象学、天气学和气候学原理以及《海滨观测规范》^[9]《地面气象观测资料质量控制》^[10]和数据文件格式的相关规定,进行水文气象观测数据序列的技术性和合理性检验;同时结合各海洋站的不同海洋气候背景,由程序自动对观测资料进行检查,找出不合理、不正常的错误或可疑记录,从而达到质量控制的目的^[11]。

目前采用的质量控制方法主要包括日期检验、位置检验、格式检验、观测要素范围检验、统计特性检验、气候特性检验、全等性检验、非法码检验、相关性检验、递增性检验、连续性检验、极值检验、尖峰检验、增减水位审核、一致性检验、变化连续性检验、相邻站的一致性检验、可视化图形绘制检验以及异常数据的判别和处理等,这些方法已经投入日常业务化工作中且取得了很好的效果。本研究分别介绍这些方法的具体内容。

(1)日期检验。海洋观测资料的观测日期应在合理取值范围内。如,年份的取值不大于当前年份,月份的取值范围为1~12,日期的取值范围为当月天数,小时的取值范围为0~23,分和秒的取值范围为0~59。

(2)位置检验。海洋观测资料的测站位置应在合理取值范围内。如,纬度的取值范围为 $-90^{\circ} \sim 90^{\circ}$,经度的取值范围为 $-180^{\circ} \sim 180^{\circ}$,固定观测站位漂移范围不超过5 km。

(3)格式检验。海洋观测资料均是按照规定格式记录的,如项目要素记录的起始位置和长度、数据记录的类型以及缺测值的填写等都有相应的要求,根据这些固定格式对资料进行检验。

(4)观测要素范围检验。简称范围检验,主要是对已有的国内外海洋观测资料数据进行统计分析,根据各项观测要素的自身特点和取值范围对资料进行检验;如果超出正常范围,则可认为该数据异常。

(5)统计特性检验。理论上海洋观测资料往往具有一定的概率统计特性,数据对应的随机变量和随机过程既相互独立又服从特定分布,时间序列资料对应的随机过程也是平稳的或周期性的;根据数据的这些特性对其进行检验,检验观测数据是否独立,独立的数据往往是异常值。

(6)气候特性检验。根据海域海洋环境气候要素的变化特点,检验其是否遵循相应的变化规律,如季节性变化和日变化等。具体检验方法为:从历史资料中统计测站位置要素数据累年逐月(累年逐季)平均值 Ann 和对应的均方差 S_d ,根据该测站要素变化的剧烈程度选取合理的倍数 m ;如果不能满足 $Ann - m \cdot S_d \leq \text{观测值} \leq Ann + m \cdot S_d$,则判定其为异常值,需进一步分析。

(7)全等性检验。主要针对观测记录中的某些要素项进行,如资料类型、浮标号、平台代码、海洋观测台站代码、观测方法、仪器名称、观测仪器海拔高度和观测点水深等,这些要素项的参数具有特定值且往往长期保持不变,记录值和约定值必须完全一致,否则视为错误。

(8)非法码检验。主要针对取值在有限代码范围内的海洋观测要素进行。如,《海滨观测规范》^[9]规定波型取值只能为 U 、 U/F 、 F/U 、 F 、 FU 、——、+++和空格的其中之一,不在此范围内的取值视为错误。

(9)相关性检验。根据海洋观测资料中要素间的相互关系检验数据是否异常。如,1 d 内各定时或逐时记录值是否超出日极值,最大波高必须大于或等于平均波高,最大周期必须大于或等于平均周期,高、低潮潮高和逐时潮高的关系,波型、波高和海况的关系,风速、波高和周期的关系以及海水盐度、温度和密度的关系等。

(10)递增性检验。检验递增量差值是否大于或等于某一确定值。具体检验方法为:假设当前观

测值为 $V(t)$,与其相邻的上一个正确值为 $V(t-1)$,检验值 $= V(t) - V(t-1)$ 。

(11)连续性检验。海洋观测要素在一定时空范围内具有连续性,即时间接近或位置邻近的观测要素的差值在一定范围内。具体检验方法为:假设当前观测值为 $V(t)$,与其相邻的上一个正确值为 $V(t-1)$,检验值 $= |V(t) - V(t-1)|$ 。

(12)极值检验。一般情况下,定点定时要素观测值的取值应在该点该要素多年极值的范围内,即 $\text{Min}(X) \leq \text{观测值} \leq \text{Max}(X)$;否则判定其为异常值,需进一步分析。

(13)尖峰检验。经检验如出现较大突变,则判定其为异常值,需进一步分析,具体参数视观测要素而定;具体检验方法为:假设当前观测值为 $V(t)$,与其相邻的第一个正确值分别为 $V(t-1)$ 和 $V(t+1)$,检验值 $= |V(t) - [V(t+1) + V(t-1)]/2| - |[V(t+1) - V(t-1)]/2|$ 。

(14)增减水位审核。当浅水潮较大区域的实际潮位数据和增减水位同时超出合理范围时,或当其他区域的实际潮位数据和增减水位其中之一超出合理范围时,则认为观测数据是可疑的。

(15)一致性检验。主要包括内部一致性检验和时间一致性检验。其中,内部一致性检验是对同一时间观测的水文气象要素之间的关系是否符合一定物理联系的检验,时间一致性检验是对水文气象要素在一定时间范围内的变化是否具有特定规律的检验。各要素观测记录应符合的关系主要包括:干球温度大于(等于)湿球温度(湿球结冰时除外)、气温大于(等于)露点温度、10 min 平均风速小于(等于)最大风速、2 min 平均风速小于(等于)极大风速、能见度小于(等于)4 km 时应有雾出现、极大风速大于(等于)17.0 m/s 时应有大风现象以及静稳时风速应小于 0.2 m/s。不符合上述关系的记录中至少有 1 个视为错误数据。

(16)变化连续性检验。1 min 间隔数据的变化范围和 24 h 内气温气压的变化范围应满足的各项要求分别如表 1 和表 2 所示,超出范围则视为可疑数据。

表 1 1 min 间隔和给定时间范围数据变化界限

要素	1 min 间隔变化 允许的最大值	变化幅度的最小值 (过去 60 min 内)
气压/ hPa	1	0.1
气温/℃	3	0.1
露点温度/℃	2	0.1
地面温度/℃	5	0.1
相对湿度/%	10	1.0
平均风速(2 min)/ (m·s ⁻¹)	20	0.5

表 2 24 h 内气温气压的变化范围

要素	变化范围
气压/ hPa	±50
气温/℃	±50

(17)相邻站的一致性检验。①物理高斯权重插值法。从统计学和气象学相关性的角度,为周围 3 个相关站点的某要素分配不同的权重,通过周围要素值并根据物理高斯权重插值法的原理计算出欲测站点某要素的插值。②24 h 变量场法+插值法。用前 24 h 的资料求得变量场,再使用插值法进行检验;如果没有前 24 h 的资料,则不进行此项检验。系统在默认情况下首选使用最优的 24 h 变量场法+插值法,最终选择哪种方法由系统根据报文类型决定。

(18)可视化图形绘制检验。由于观测要素的变化是连续的,通过绘制可视化图形可直观地判断出异常值。如,绘制要素的时间序列过程线,显示的尖峰值即视为异常值(图 1)。

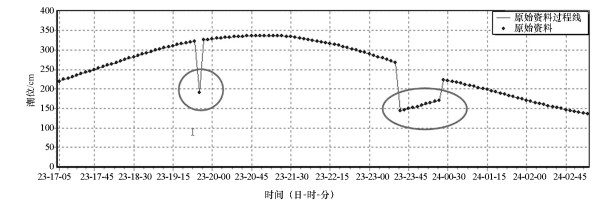


图 1 5 min 潮位数据的时间序列

(19)异常数据的判别和处理。异常数据的判别和处理是在海洋站观测资料质量控制中需解决的重要问题。异常数据主要包括 2 类:①正确的异

常值,即海况急剧变化的真实记录,如台风过境时风速和水位观测数据的异常增大等;②含有过失误差的异常值,即由于仪器失灵、外界干扰或观测人员失误造成的错误记录,应在资料质量控制中加以标识或删除。

4 质量控制软件

根据上述质量控制的原则和方法,国家海洋信息中心研发了海洋站观测资料质量控制系统,为海洋站观测资料的质量控制工作提供了良好的工具,是质量控制技术的最终体现。目前该系统已更新为 2.03 版本,包含上述各种质量控制方法,并利用可视化质量控制技术实现交互操作,极大地提高了质量控制的工作效率。此外,还建立了海洋站观测资料质量控制的审核制度,从海洋站、中心站、各海区信息中心到国家海洋信息中心,每级至少有 1 位审核员负责观测资料的质量控制工作,并在形成的文件说明中记录审核人、审核日期和相关信息。审核员都是海洋观测一线工作人员,最为了解当时当地的天气、环境和仪器设备状况,因此他们对观测资料的审核极为重要也最为有效。

5 结语

开展海洋站观测资料质量控制的主要目的是确定正确记录,找出缺测记录、错误记录和可疑记录,并对其做出标识或使用尽可能准确的值代替,以确保提供应用的观测资料符合各种要求^[12]。我国现有的观测资料质量控制技术和方法与发达国家相比差距很大^[13],而观测资料的质量直接影响其应用效果。此外,由于在观测记录过程中不可避免地受到诸多主观或客观方面的影响^[14-15](如台站迁移、仪器故障或变更、系统更新升级、观测人员更换或水平不一、观测频次和时间改变以及统计分析方法改变等),观测数据在记录真实环境气候变化之外还隐藏了一些非均一的变化,这些变化将不利于资料同化系统的发展以及预报模式等业务和研究工作水平的提高^[16],因此仅对观测数据本身进行质量控制是远远不够的,还应加强对观测数据非均一性的研究。本研究下一步的工作即在不断研发新的质量控制方法的基础上,重点研究海洋站观测资料(尤其是海洋水文资料)的均一化检验和订正,选

取最适合中国海洋环境水文和气候资料特点的均一性检验方法,努力建立多种均一性检验和订正方法联合使用和研究的多模式检验系统,构建高质量的中国海洋站海洋水文气象资料均一化数据集,为沿海气候变化研究提供准确的第一手资料。

参考文献

[1] 张铁艳,王化仁,杨鲲,等.海洋调查观测资料的质量控制[J].水道港口,2006,27(1):48—50.

[2] 熊安元.北欧气象观测资料的质量控制[J].气象科技,2003,31(5):314—320.

[3] 刘小宁,任芝花.地面气象资料质量控制方法研究概述[J].气象科技,2005,33(3):199—203.

[4] 王海军,杨志彪,杨代才,等.自动气象站实时资料自动质量控制方法及其应用[J].气象,2007,33(10):102—109.

[5] 窦以文,屈玉贵,陶士伟,等.北京自动气象站实时数据质量控制应用[J].气象,2008,34(8):77—81.

[6] 尹嫦姣,江志红,吴息,等.空间差值检验方法在地面气象资料质量控制中的应用[J].气候与环境研究,2010,15(3):229—236.

[7] 刘小宁,鞠晓慧,范邵华,等.空间回归检验方法在气象资料质量检验中的应用[J].应用气象学报,2006,17(1):37—43.

[8] 李学坤,李凤金,等.海洋水文气象实时数据质量控制[J].海洋预报,1997,14(3):71—78.

[9] 中华人民共和国国家技术监督局.海滨观测规范:GB/T 14914—2006[S].北京:中国标准出版社,2006.

[10] 中华人民共和国国家技术监督局.地面气象观测资料质量控制:QX/T 118—2010[S].北京:气象出版社,2010.

[11] 何健,王潜梅,钱光明,等.广东省区域自动气象站资料的质量控制与评估[J].广东气象,2011,33(3):37—40.

[12] 中华人民共和国国家技术监督局.地面气象观测规范 第 22 部分:观测记录质量控制:QX/T 66—2007[S].北京:气象出版社,2007.

[13] 王英,苑跃,王小兰.气象资料的质量控制[J].高原山地气象研究,2007,27(3):39—40.

[14] 李庆祥,刘小宁,张洪政,等.定点观测气候序列的均一性研究[J].气象科技,2003,31(1):3—10.

[15] 李琰,牟林,王国松,等.环渤海沿岸海表温度资料的均一性检验与订正[J].海洋学报,2016,38(3):27—39.

[16] 张高杰,何金海,周自江,等.RHtest 方法对我国降水资料的均一性检验试验[J].气象科技,2012,40(6):915—921.