

AMDAR 观测资料分析及质量控制

陶士伟 郝 民 赵琳娜

(国家气象中心,北京 100081)

提 要: 随着中国气象局与民航总局 AMDAR(Aircraft Meteorological Data Relay) 合作项目的开展,气象部门获得的 AMDAR 观测资料逐渐增加。为获取高质量的 AMDAR 观测资料和充分合理地使用这些观测资料,必须对 AMDAR 观测资料时空分布、飞行状态信息、观测资料质量等信息进行充分调研分析,并开发设计出具有实际运用价值和能力的质量控制(QC)技术,使 AMDAR 观测资料在数值预报业务模式系统中发挥更大作用。AMDAR QC 方案在参考国际先进 AMDAR QC 方案的基础上,根据 AMDAR 资料的特点和数值天气预报模式对 AMDAR 资料应用的要求而设计,其目的是解决 AMDAR 观测资料在数值预报中应用的质量问题。文中方案采用多种质量控制技术,其中包括气候极值检查、要素间一致性检查、时间一致性检查、持续性检查、背景场一致性检查、空间一致性检查、综合决策算法、自动统计评估反馈技术,并将该方案应用于 2007 年 6—8 月的实际观测资料。试验结果表明:该方案能有效地识别观测资料中存在明显错误的资料,有效地为 AMDAR 资料在数值模式中的应用提供客观质量依据。

关键词: AMDAR 观测资料 质量控制 决策算法

Quality Control Procedures and Its Application in AMDAR Observation Data

Tao Shiwei Hao Min Zhao Linna

(National Meteorological Center, Beijing 100081)

Abstract: In order to make use of AMDAR data better, a feasible observation data quality control module should be designed to make sure AMDAR data used in the meteorological research and operation with a higher quality. The AMDAR QC module is brought forward consulting the characteristics of AMDAR. Many quality control techniques are integrated into the

项目资助: AMDAR 资料在区域数值模式中的应用(气法函[2006]92 号)课题、国家自然科学基金(基金号:40505021)和国际科技合作项目(基金号:2006DFA21530)

收稿日期: 2008 年 11 月 19 日; 修定稿日期: 2009 年 3 月 27 日

QC module, such as the station extreme value check, the interior consistency check, the consistency check with the first guess, the temporal consistency check, the space consistency check, the decision-making algorithm, the statistic and feedback technique. The AMDAR data quality control module has been applied into AMDAR data in August of 2007. Subjective verification result shows that error information in the observation can be distinguished effectively from the AMDAR data.

Key Words: AMDAR observation quality control decision-making algorithm

引 言

随着现代化技术和世界经济的发展,小尺度、高时间分辨率、特殊行业天气预报的需求日益高涨。气象预报及资料同化技术也日新月异。为了满足和适应日益高涨的气象服务需求,能获取描述中尺度、短时效天气现象的观测资料的现代化的观测手段是必不可少的。除卫星遥感、地面加密自动站外,AMDAR 观测资料也是中尺度、短时效天气预报重要的不可缺少的非常规气象资料来源之一。由于 AMDAR 观测资料具有较高的时空分辨率,弥补了常规探空观测时间间隔过长和空间分布过稀的不足,可为数值天气预报提供高时空分辨率的初始时刻大气实际状况,对提高数值天气预报效果起到很好的作用。

国际上,AMDAR 观测资料在数值天气预报中的应用工作已开展多年^[1-2],并取得明显效果,特别是在高层 AMDAR 资料的使用,使得整个斜压系统的结构发生了显著的改变,这种改变持续了至少 3 天,并从高层延伸到低层。在动力不稳定地区观测资料的选择对研究其结构具有重要的意义,因为在这些地区,轻微的扰动可以导致天气系统致命性的发展^[3-4]。

过去,由于我国 AMDAR 观测资料较少。并且同化系统能力有限,AMDAR 观测资料应用受到了限制。随着中国气象局与民航总局 AMDAR 合作项目的开展,气象部门

获得的 AMDAR 观测资料逐渐增加。目前,国内的 AMDAR 观测资料每天平均有 25000 份左右,这些资料为进一步提高我国数值天气预报效果提供了可能。

但是,AMDAR 观测资料不同于普通的常规气象观测数据,其观测要素单一,虽然在航线上时空分辨率很高,但其分布极不均匀,这些特点又增加了 AMDAR 观测资料的使用难度。为获取高质量的 AMDAR 观测资料和充分合理地使用这些观测资料,使之在数值预报业务系统模式中发挥作用,必须对 AMDAR 观测资料时空分布、飞行状态信息、观测质量等进行充分调研,并开发设计具有实际运用价值和能力的 QC(Quality Control Procedures)技术。

本文描述的 AMDAR QC 方案的特点是尽可能多地运用各种 QC 技术^[5-10],如气象要素间一致性检查(又称内部一致性检查)、时间一致性检查、持续性检查、背景场一致性检查、空间一致性检查和决策算法,并同时统计 AMDAR 观测资料时空分布、飞行状态信息、观测质量等信息。其中背景场一致性检查可避免观测资料和背景场偏差太大,造成模式初值不协调。空间一致性检查的插值对象是背景场与观测值的差(简称观测余差,下同),即精度较高的小量插值,并运用了二次迭代过程,消除了错误观测资料对其周围观测的影响。将上述 QC 方案运用于 2007 年 6—8 月实际 AMDAR 观测资料,结果表明:该方案能有效地判别因各种原因产生的错误资料。

本文第 1 节主要分析 AMDAR 观测资料分布的特点;第 2 节介绍 AMDAR 观测资料 QC 方案结构;第 3 节分析该方案应用效果;最后是简要小结。

1 AMDAR 观测资料特点分析

AMDAR 资料具有在航线上高时空分

辨率的特点,水平距离几百米,垂直高度几十米,时间间隔几秒,十几秒。可详细地反映出某些中小尺度天气系统的动力和热力特征及其变化。

图 1 是 2008 年 1 月 1—10 日 3 小时时间间隔各时次的观测资料份数。如图所示,平均每 3 小时可以得到 7000 份左右的 AMDAR 观测资料,且白天的资料多于夜间。

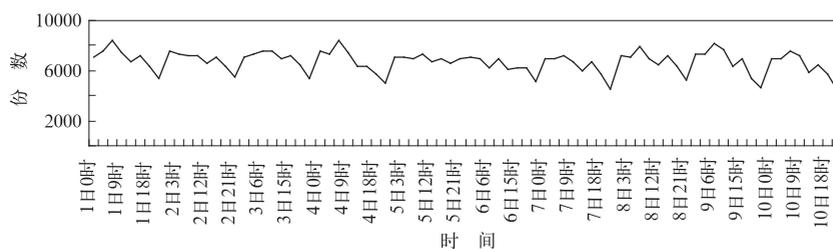


图 1 2008 年 1 月 1—10 日 3 小时时间间隔的观测资料份数

图 2 是 2008 年 1 月 1—10 日 3 小时时间间隔各时次的飞机航班数。平均每 3 小时的航

班数在 320~550 之间,同时表明白天航班数多于夜间。

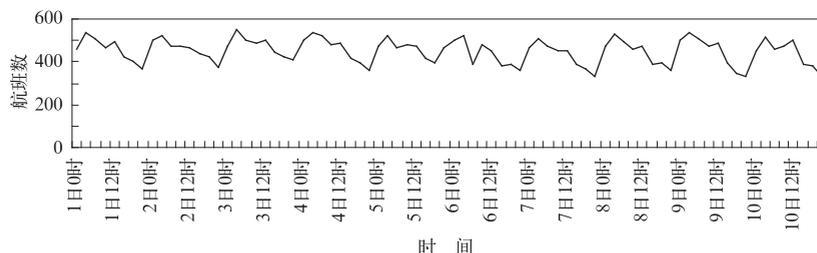


图 2 2008 年 1 月 1—10 日 3 小时时间间隔的飞机航班数

此外,飞机在不同飞行状态下观测资料数也是不同的。图 3 给出了 2008 年 1 月 1—10 日 3 小时时间间隔各时次不同飞行状态下的观测数。飞机在起飞上升、飞行巡航和着陆下降这三种飞行状态下的观测数基本相

当,上升比下降和巡航稍少一些,其变化趋势也基本相同。除图 3 给出的三种飞行状态外,另外还有约总观测数的四分之一飞行状态不明。

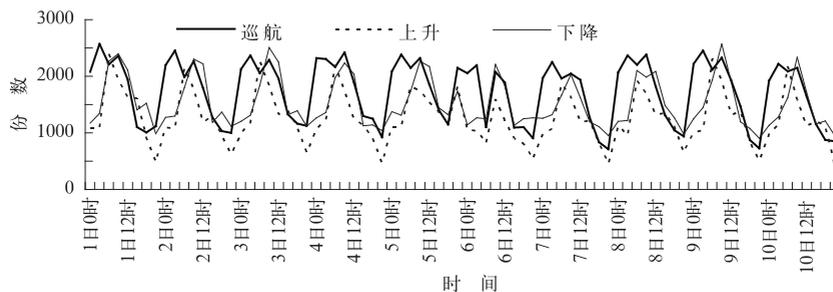


图 3 2008 年 1 月 1—10 日 3 小时时间间隔不同飞行状态观测数

AMDAR 观测资料在水平方向上分布也有比较明显的特点。图 4 给出的是 2008 年 8 月 10 日 AMDAR 观测资料的水平分布图。从图 4 中可见,AMDAR 观测资料的水平分布有两个特点:(1)观测分布与航线有关,在航线上十分密集;(2)观测资料主要集中在四个区域:北美、东亚、欧洲、大洋洲,其中在北美和欧洲分布最为密集,这和各地区的航线分布有关。

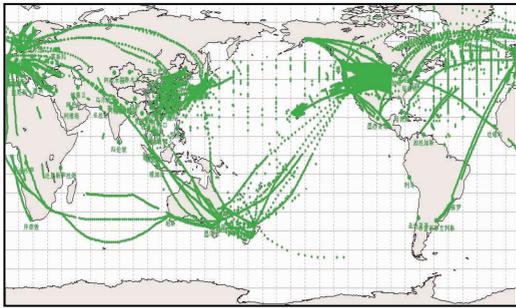


图 4 2008 年 8 月 10 日 AMDAR 观测资料的水平分布

2 自动资料处理和 QC 方案设计

要设计开发 AMDAR 观测资料的 QC 方案,详细了解其特点是非常有必要的。飞机上的观测仪器是在无人操作和实时管理的情况下自动进行实时采集数据、形成观测报告、生成和传输数据文件,并需要传感器、数据采集、数据通讯传输诸系统,实时性很强,错误数据源于多种途径,不易被发现。因此,需要应用多种 QC 技术从不同角度发现和识别其中的错误观测资料。本文设计的自动 AMDAR 的资料质量控制(QC)方案是在数值预报同化系统中实现的 QC 过程。它包括资料的时间一致性检查、内部一致性检查、持续性检查、背景场一致性检查、空间一致性检查和决策算法。AMDAR 观测资料只有温度和风两种要素,除内部一致性检查只用于

风观测要素外,以上检查方法都用于这两种要素。各种检查过程和决策算法都赋予相应的 QC 标识。其程序结构如图 5。

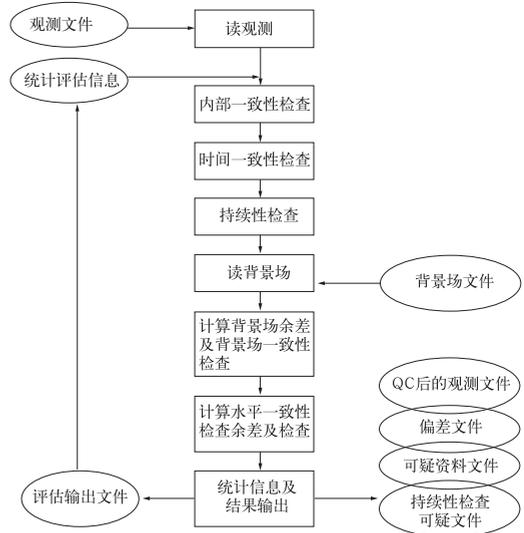


图 5 AMDAR 观测资料 QC 结构图

2.1 时间一致性检查

时间一致性检查是从时间变化率的角度检验观测信息或观测要素,识别出不理想的突然变化。时间一致性检查方法是用当前的观测信息与相邻时次的观测信息相比,如果它们的差别超过给定的允许范围,当前的观测信息被认为是可疑的。时间一致性检查的检查判据与样本的时间分辨率有关。时间一致性检查适用于高时间分辨率,因为相邻样本的相关性随着时间分辨率增加而加强。飞机是在三维空间移动的观测平台,一个要素不同时段观测值的差别主要包括观测要素随时间、垂直高度和水平位置的变化引起的。因此,时间一致性检查判据 $\text{Max}(DA)$ 为:

$$\begin{aligned} \text{Max}(DA) = & \text{Max}(\delta a / \delta z) \Delta z + \\ & \text{Max}(\delta a / \delta \varphi) \Delta \varphi + \\ & \text{Max}(\delta a / \delta t) \Delta t \end{aligned}$$

式中 $\delta a / \delta z$ 、 $\delta a / \delta \varphi$ 、 $\delta a / \delta t$ 分别代表观测要素

随高度、纬度和时间的变化率, Δz 、 $\Delta \varphi$ 、 Δt 分别代表两个观测的高度差、纬度差和时间差, Max 表示最大值。Max($\delta a / \delta z$)、Max($\delta a / \delta \varphi$)、Max($\delta a / \delta t$) 的取值如表 1。

表 1 观测要素随高度、纬度和时间的最大值变化率

	Max($\delta a / \delta z$)	Max($\delta a / \delta \varphi$)	Max($\delta a / \delta t$)
温度	4°C/100m	4°C/1°	5°C/1h
风	10m·s ⁻¹ /100m	10m·s ⁻¹ /1°	20m·s ⁻¹ /1h

2.2 持续性检查

持续性检查的主要功能是检查仪器失灵(如结冰、通讯问题等)造成的观测参数变化太小或不变。与时间一致性检查相比,持续性检查是从另一个角度检查观测要素时间变化率的合理性。持续性检查有两种方法,其一是标准差方法,它要求给定时间范围内观测资料的标准差大于给定的判据。其二是相邻观测时次资料差方法。它要求给定时间范围内所有相邻观测时次资料差的最大值大于给定判据。这里采用方法二。检查结果作用于给定时间窗内所有资料。温度和风持续性检查判据分别为 0.1°C 和 0.5m·s⁻¹。

若给定时间范围内所有相邻观测时次观测资料差值的最大值小于给定判据,则没有通过持续性检查。

2.3 背景场一致性检查

背景场一致性检查是资料同化中一种重要的 QC 方法,可避免观测资料和背景场偏离太大,造成模式初值不协调。背景场一致性检查是根据观测资料和观测点处背景场的差(简称观测余差)是否超出允许范围而判别观测资料可疑与否。一般取观测余差 3、4 倍标准差作为背景场一致性检查的可疑、异常判据。

模式提供的背景场是在规则的三维模式网格点上,观测资料是在不规则的观测站点

上。观测余差的计算需要把背景场从规则的三维模式网格点转换到不规则的观测站点上。

2.4 空间一致性检查

空间一致性检查是利用客观分析方法由周围观测资料计算被检查观测资料的分析值作为估计值。并要求观测值和估计值的差小于给定的判据,否则,认为被检查观测资料异常或可疑,即没有通过空间一致性检查。可用于空间一致性检查的客观分析方法有很多。本方案运用 Barnes 客观分析方法^[11]。该方法采用的插值对象是观测余差,即精度较高的小量插值。并运用了二次迭代过程,消除了错误观测资料对其周围观测的影响。既可以满足计算观测资料估计值的要求,又可节省多次迭代的计算机资源。

Barnes 客观分析方法计算的估计值是周围观测资料的加权平均,权重函数是随距离呈指数衰减的函数。权重函数的衰减系数取决于观测网的分辨率。对于飞机观测网,空间一致性检查的影响半径可设为 100km。空间插值余差(估计值与观测值之差)标准差的 3 倍、4 倍分别作为空间一致性检查的可疑、异常判据。

Barnes 客观分析方法利用观测余差替代观测值,不仅能减小插值误差,又起到消除观测资料受飞行高度的影响。

空间一致性检查公式:

$$x' = x_j - \frac{\sum_{i=1}^N \omega_i x_i}{\sum_{i=1}^N \omega_i}$$

$$\omega_i = 1 / (1 + 4r_i^2)$$

$$r = 6371 \times \arccos[\sin \varphi_1 \sin \varphi_2 + \cos \varphi_1 \cos \varphi_2 \cos(\theta_1 - \theta_2)]$$

式中, x' 为空间插值余差, x_j 为被检查点观测余差, x_i 为周围点观测余差, ω_i 插值权重,

r 为 j 点与 i 点之间的距离。

2.5 内部一致性检查

内部一致性检查的基本算法是基于两观测要素的气象学相关性,检查它们之间是否存在矛盾现象。当被检查的两个观测资料矛盾时,则被认为是可疑的,需要用其他方法再进一步确定究竟哪一个资料是有问题的。在本方案中,进行内部一致性检查的观测资料只有风向(dd)与风速(ff)^[6]。当下列情况发生时,则认为风向与风速矛盾。

dd 缺测,FF 非缺测;或 dd 非缺测,FF 缺测;或 $dd < 0$ 或 $dd > 360$;或 $dd \neq 0, FF = 0$;或 $dd = \text{可变的}, FF > 6\text{m} \cdot \text{s}^{-1}$ 。

2.6 决策算法及评估反馈

决策算法是综合分析各种自动 QC 方法的结果,再进一步确定观测资料的质量标识(正确、可疑、可能错误、错误)。比如在时间一致性检查、或内部一致性检查中同时被标为是可疑的两个资料,在空间一致性检查中如果其中一个资料没有通过检查,则可认为该资料是错误的,提高它的质量标识可疑等级,另一个可认为是正确的,降低它的质量标识可疑等级。如果一个资料在两个以上的检查方法中没有通过,在决策算法中可相应提高它的质量标识可疑等级。只在一种检查中没有通过,有可能是局地逆温、中小天气系统的影响,相应降低它的质量标识可疑等级。

本方案的另一个主要特点是包含各个 QC 分量效果统计评估技术及统计评估信息自动反馈到 QC 方案的各个分量,这些技术的主要作用是:(1)它是观测资料用户了解各类观测资料质量信息的主要来源,以便正确使用观测资料,如在资料同化系统中合理设置观测资料可用性和观测误差;(2)观测系统

管理维护人员借助统计评估信息可发现观测系统或观测仪器异常,及时对观测仪器进行改进和维护;(3)观测资料质量控制方案设计开发人员也需要了解自动 QC 方案的各个分量的性能,不断地改进质量控制方案以减少质量控制错判率;并实现自动化的逐渐逼近完善的质量控制方案。统计评估信息包括按测站、区域评估各观测要素及各种质量控制方法检查余差的平均值、均方差、最大值、最小值、错误率、可疑率等。

3 QC 方案应用效果分析

将该方案应用于 2007 年 6—8 月的实际 AMDAR 观测资料,背景场为国家气象中心 T213 数值业务预报系统提供的 6 小时预报。在试验中,为了防止误判,对观测余差和空间插值余差的标准差设置一个最小限制,如表 2。

表 2 观测余差、空间插值余差标准差最小值

要素名	观测余差	空间插值余差
温度(度)	3.0	2.5
$u(\text{m} \cdot \text{s}^{-1})$	4.0	3.0
$v(\text{m} \cdot \text{s}^{-1})$	4.0	3.0

3.1 个例分析

我们对大量的质量控制结果进行了主观分析,结果表明该方案能有效地识别观测资料中存在的明显错误资料。下面是从众多的试验结果中随机选取的几个个例,以说明本质量控制方案和检查结果的合理性。

个例 1 是 2007 年 8 月 31 日 12 时对 AFZA22 航班的质量控制结果。该航班共观测 21 次,观测开始时间为 11 时 39 分,对应观测位置为(8.96°S、6.44°E),观测结束时间为 13 时 24 分,对应观测位置(18.85°S、16.91°E),在约 2 个小时的观测中,飞机飞越了约 10 个纬距和 10 个经距,但是观测的高

度(11186m)和温度(223.47K)一直保持不变,这显然是观测仪器异常造成的错误资料。本方案的持续性检查识别出了这种错误,如果不使用持续性检查,这种错误是不容易识别的。

个例 2 是对 2007 年 8 月 24 日 00 时 AAL5 航班的温度记录。如图 6。

在图 6 中, AAL5 航班在 (30.6°N、133.4°W) 处的温度记录为 218K, 气压层为 227hPa, 其观测和空间插值余差分别为 -14.3 和 -10.7K, 综合决策算法判别为错误资料。根据周围观测的温度和气压进行的主观判断, 认为该温度记录的确偏低 10K 左右。综合决策算法的结论是可信的。

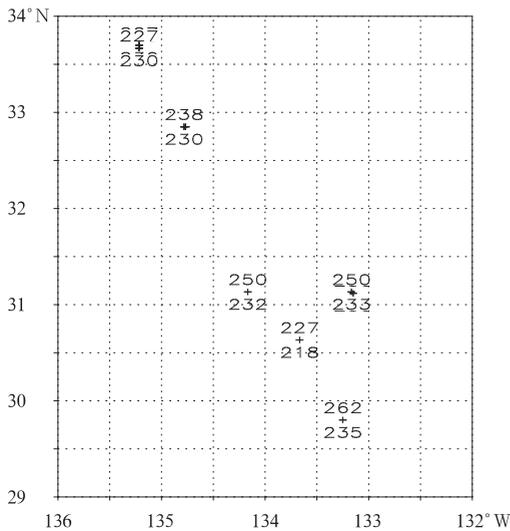


图 6 2007 年 8 月 24 日 00 时 250hPa 附近 AAL5 航班周围的温度记录

个例 3 是 2007 年 8 月 24 日 00 时 PAL02 航班的风记录(如图 7)。

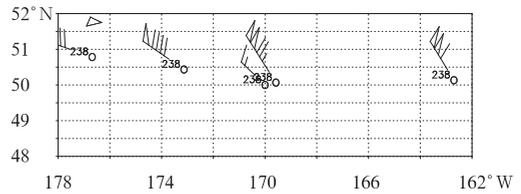


图 7 2007 年 8 月 24 日 00 时 250hPa 附近 PAL102 周围的风记录

在图 7 中, PAL5 航班在 (50.6°N、169.4°W) 处的风记录为(风向 310°, 风速 5.9 m · s⁻¹), 气压层为 238hPa。其 U、V 的观测余差分别为 -20.87m · s⁻¹ 和 30.31m · s⁻¹; 因周围观测太少, 没有进行空间一致性检查; 综合决策算法判别为错误资料。根据周围观测的风记录(40m · s⁻¹ 左右), 主观判断, 该记录风速偏慢 35m · s⁻¹ 左右。综合决策算法的结论是可信的。

3.2 统计结果分析

我们将该方案应用于 2007 年 6—8 月的实际 AMDAR 观测资料, 每月统计信息基本一致, 即稳定性很好。表 3 为 2007 年 8 月各要素在各种检查方法的平均偏差和均方差。

在表 3 中, 各要素的平均偏差都很小, 这说明 AMDAR 观测资料的系统性偏差很小; 温度观测余差的均方差约 2.2°, U、V 分量观测余差的均方差约 3.6m · s⁻¹, 这个量级说

表 3 2007 年 8 月各要素在各种检查方法的平均偏差和均方差

	温度资料				U 分量			V 分量	
	o-b*	o-a*	o1-o2*	o-b	o-a	o1-o2	o-b	o-a	o1-o2
平均偏差	-0.15	0.01	0.09	0.26	0.02	-0.02	-0.07	-0.01	0.01
均方差	2.19	1.10	2.50	3.67	2.47	2.70	3.65	2.57	2.71

注: * : o-b 表示观测值减背景场; o-a 表示观测值减空间插值; o1-o2 表示观测值减上一次观测值

明 MDAR 观测资料的精度是较高的。各要素时间一致性检查的均方差和观测余差的基

本相当, 各要素空间一致性检查的均方差都小于观测余差和时间一致性检查的均方差,

这说明两个问题:(1) AMDAR 观测资料的空间一致性较好;(2)空间一致性的敏感性更好。表 4 是各检查方法没有通过检查资料的比率。

表 4 2007 年 8 月各检查方法没有通过检查资料的百分率(%)

	fgschk	barchk	timchk	intchk	prschk	DMA
温度	0.37	0.38	0.02	0.00	0.26	0.43
风	0.42	0.50	0.01	0.06	0.01	0.52

注:fgschk、barchk、timchk、intchk、prschk 和 DMA 分别代表背景场检查、空间一致性检查、时间一致性检查、内部一致性检查、持续性检查和综合决策算法

从表 4 可见,在背景场检查、空间一致性检查中,没有通过检查资料的百分率是比较合理的。内部一致性检查只对风资料进行,没有通过检查资料的比率很小。时间一致性检查中没有通过检查资料的比率也很小,原因可能是给定的判据偏大。对于温度资料,持续性检查的作用也是很明显的;最后的综合决策算法判别的错误比率在 0.4%~0.5%左右。

从以上的个例分析和统计信息分析,我们认为针对 AMDAR 观测资料设计的本 QC 方案具有较强的错误资料识别能力。

4 小结

AMDAR 观测资料不同于普通的常规气象观测数据,其观测要素单一、时空分辨率极不均匀,为获取高质量的 AMDAR 观测资料和充分合理地使用这些观测资料,使之在数值预报业务系统模式中发挥作用,我们对 AMDAR 观测资料时空分布、飞行状态信息进行了分析;并设计了应用于数值预报同化系统中的 QC 方案。将该方案应用于 2007 年 6—8 月的 3 个月的实际 AMDAR 观测资料进行了试验,它包括资料的时间一致性检

查、内部一致性检查、持续性检查、背景场一致性检查、空间一致性检查和决策算法。

资料分析和质量控制的初步结论如下:

(1) AMDAR 观测资料平均每 3 小时在 7000 份左右,白天多于夜间。观测资料的水平分布在航线上十分密集;并集中在四个区域:北美、东亚、欧洲、大洋洲,其中在北美和欧洲分布最为密。

(2) QC 方案尽可能多地运用了各种 QC 技术和决策算法,以增加 QC 结果的准确性。

(3) 该 QC 方案应用于实际的观测资料,结果证明其功能达到预期的效果。从统计信息分析,各种 QC 方法的偏差、均方差和错误资料比率都是合理的;随机选取的个例分析也说明 QC 方案对错误资料的判断是可信的。

(4) 本文设计的 QC 方案能有效地为 AMDAR 资料在数值模式中的应用提供客观质量依据。不失为一种可以用于数值预报业务的 AMDAR 资料 QC 方案。

需要说明的是,对于数值模式预报结果,一般以观测或分析为参照,通过检验评分和误差均方差作为综合评估指标。而对质量控制系统的,除个例分析外,因为没有参照场或者参照场就是它自身,无法应用数值模式预报结果的综合评估方法。因此质量控制系统的综合评估方法,多采用各要素在各种检查方法的平均偏差、均方差和错误率定性评估它的合理性。另外,任何 QC 方案都不能保证其决策算法是完全正确的,本 QC 方案也不例外。还有需要利用更多的 AMDAR 资料在今后的业务实践中不断完善。

参考文献

- [1] Stickland J J, F Grooters A T. Observations from the Globe AMDAR Programme[OL]. 2004. <http://www.>

- wmo. int/pages/prog/www/IMOP/publications/IOM-82-TECO_2005/Papers/2(01)_Australia_text. pdf
- [2] WMO AMDAR Panel, Aircraft Meteorological Data Relay[OL]. International AMDAR Program, 2004, www. wmo. int/web/aom/amprog/amprog. html
- [3] Be'atrice Pouponeau, Franck Ayrault, Thierry Bergot, And Alain Joly. The impact of aircraft data on an atlantic cyclone analyzed in terms of sensitivities and trajectories[J]. *Weather and Forecasting*, 1999, 14:67-83.
- [4] Richard Mamrosh. Aircraft weather observations improve forecasts[J]. *National Weather Service*, 2003, 2(2):1-7.
- [5] Dean Vickers and Mahrt L. Quality control and flux sampling problems for tower and aircraft data[J]. *Journal of Atmospheric and Oceanic Technology*, 1997, 14:512-526.
- [6] 陶士伟,徐枝芳. 加密自动站资料质量保障体系分析[J]. *气象*, 2007, 33(2):34-41.
- [7] 王海军,杨志. 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.
- [8] 任芝花,熊安元. 地面自动站观测资料三级质量控制业务系统的研制[J]. *气象*, 2007, 33(1):19-24.
- [9] 杨贤为. 气候应用专用数据库气象资料的质量检验[J]. *气象*, 1998, 24(12):33-36.
- [10] 王伯民. 基本气象资料质量控制综合判别法的研究[J]. *应用气象学报*, 2004. 12, 15(增刊):50-59.
- [11] Barnes S L. A technique for maximizing details in numerical weather map analysis [J]. *J Appl Meteor*, 1964, 3, 396-409.