

全球大气再分析常规气象观测资料的 预处理与同化应用

廖捷¹ 胡开喜¹ 江慧¹ 曹丽娟¹ 姜立鹏¹ 李庆雷¹ 周自江¹ 刘志权² 张涛¹ 王蕙莹¹

(1 国家气象信息中心, 北京 100081; 2 美国国家大气研究中心, 博尔德 80301, 美国)

摘要: 用于全球大气再分析的常规气象观测资料由表面(地面站、船舶和浮标等)和高空(无线电探空、测风气球、下投式探空、飞机和风廓线雷达等)观测资料组成。基于已有工作进展,总结了我国全球大气再分析(CRA-40, 1979—2018年)常规观测资料预处理与同化应用研究进展。在CRA-40研发阶段,完成了15个地面、高空、飞机和海洋观测数据源的资料的收集整理、分类整合和质量评估。相对NCEP CFSR同化的常规观测资料,CRA-40同化了更多中国地区的地面站、常规探空观测资料,增加了美国地区的TAMDAR和ACARS飞机报湿度数据。CRA-40采用RAOBCORE 1.4数据集提供的订正值完成了全球探空温度的偏差订正,采用NCEP辐射订正算法减小太阳辐射引起的热敏电阻季节性观测误差。对再分析同化结果带来负面效果的观测(如高海拔地面站观测、模式地形层以下高空观测以及300 hPa以上探空湿度等)在观测资料预处理阶段被剔除。CRA-40采用的黑名单提供了不同时期观测资料存在问题的地面站、高空站、飞机、船舶和浮标等信息。这些信息来自观测资料和ERA-Interim再分析的偏差统计结果,以及NCEP CFSR、ECMWF提供的历史黑名单,后者在使用前进行了严格的诊断评估和确认。经过预处理和质量预评估的全球历史常规资料已应用于CRA-40的10年试验产品研制。

关键词: 再分析, 常规资料, 整合, 质量控制, 黑名单

DOI: 10.3969/j.issn.2095-1973.2018.01.018

Pre-Process and Data Selection for Assimilation of Conventional Observations in the CMA Global Atmospheric Reanalysis

Liao Jie¹, Hu Kaixi¹, Jiang Hui¹, Cao Lijuan¹, Jiang Lipeng¹, Li Qinglei¹,
Zhou Zijiang¹, Liu Zhiquan², Zhang Tao¹, Wang Huiying¹

(1 National Meteorological Information Centre, Beijing 100081 2 National Center for Atmospheric Research, Boulder, CO 80301)

Abstract: The conventional observing system serves as an indispensable constraint to the atmospheric reanalysis. Surface observations were available from ships, drifting buoys, land stations, and airports. In situ measurements of upper-air were available from radiosonde, pilot balloon, aircraft, and wind profile. Data collection and merging of different data source, quality control, and data selection are always important for global atmospheric reanalysis. This article describes the input conventional datasets used to produce the global atmospheric reanalysis project of CMA (CRA-40, 1979-2018), and discusses the processing method of conventional data before being assimilated. During the development stages of CRA-40, conventional data from 15 data sources are integrated, quality controlled and evaluated. Compared with NCEP Climate Forecast System Reanalysis (CFSR), the CRA-40 assimilates more Chinese surface and radiosonde data. The specific humidity of TAMDAR and ACARS reports are assimilated. Observations that are expected to have a negative impact on the reanalysis are rejected. These observations include surface pressure observations over high terrain, radiosonde observations below the model surface, humidity above 300 hPa, and other data of blacklists. Blacklists provide the identifiers of station, aircraft, ship and buoy with systematical quality problem in different periods. Blacklists information of CRA-40 come from the statistic results of the departure between observation and ERA-Interim reanalysis, and the historical blacklists provided by CFSR and ECMWF. Blacklists of CFSR and ECMWF are accessed before being using in CRA-40. Parts of research results are of guiding significance for improving the historical and real-time data processing and application. Quality controlled historical conventional observations are assimilated in CRA -40. Some of conventional observations are serviced by China Integrated Meteorological Information Service System (CIMISS).

Keywords: reanalysis, conventional data, integration, quality control, blacklist

收稿日期: 2017年7月28日; 修回日期: 2017年12月10日

第一作者: 廖捷(1981—), Email: liaoj@cma.gov.cn

资助信息: 国家气象科技创新工程项目(CMAGT003-5);

公益性行业(气象)科研专项(GYHY201506002
和 GYHY201406001)

0 引言

再分析最早起源于第一次全球大气研究实验(FGGE)中观测资料的应用。为了给数值天气预报初始化提供更好的观测,科学家们对观测资料进行了多次分析研究,并很快发现通过再分析可以为大气科学研究提供一个更有价值的数据集^[1]。随后,欧美数值预报机构推出一系列再分析产品。第一代产品包括ERA-15^[2]、NCEP/NCAR再分析数据^[3]和NASA/DAO^[4],第二代产品包括NCEP/DOE^[5]、ERA-40^[6]和JRA-25^[7],第三代产品包括ERA-Interim^[11]、CFSR^[8]、MERRA^[9]和JRA-55^[10]等。各数值预报机构通过使用更好的模式、更好的输入资料 and 同化方法不断提高再分析产品的质量。

大气再分析同化的高空定点观测(in situ)温度、风和湿度来自无线电探空、测风气球、飞机和风廓线雷达;表面气压、2 m温度、2 m相对湿度和近地面风来自船舶、浮标观测地面气象站和机场站。除了飞机报在20世纪90年代后期飞速增长外,其他各类常规观测资料的数量相对稳定。欧洲中期数值天气预报中心(ECMWF)最新再分析系统的评估结果^[11]表明,1979—1990年,常规观测占总观测的比例约35%,对分析的影响权重高达75%;随着卫星观测的快速增加,至2015年,常规观测占总观测的比例约10%,对再分析的影响权重约20%,对大气再分析依然起到不可忽视的约束作用。

各机构研制新版本的全球再分析产品时,所需的常规观测资料通常以国际上已发布的再分析输入观测资料集为主要数据源,同时补充了最新收集的观测资料。如ERA-Interim同化的历史常规观测资料来自ERA-40观测资料集和ECMWF业务同化归档数据。JRA-55在ERA-40同化的观测资料基础上补充了美国国家环境预报中心(NCEP)同化的一些地面站、高空观测,以及俄罗斯水文气象局、美国大学大气研究联合会、美国气象水文局的积雪资料^[10]。ECMWF、美国和日本研制的再分析产品对常规资料的质量控制均采用了本国家、机构的业务质量控制方案,如JRA-25和JRA-55均采用了动力质量控制^[12],NCEP/NCAR、CFSR和MERRA在同化前对探空、飞机报和风廓线资料进行综合质量控制^[13]。

相对第二代再分析产品,第三代再分析产品更加重视探空温度的偏差订正,奥地利维也纳大学于2007年发布了全球首套日定时均一性订正数据集RAOBCORE^[14],为全球1184个探空站点提供的不同高度层温度偏差订正量已被应用到MERRA、JRA-55、

ERA-Interim等多套全球大气再分析的探空温度订正中,较好的避免了仪器换型、观测方法改变等因素引起的探空温度时间序列存在的非均一性。ECMWF、NCEP均在同化前区分不同地区、仪器类型订正太阳辐射引起的热敏电阻季节性观测误差,但具体算法和参数存在差异。ECMWF最新研制的第四代再分析产品对历史探空资料的订正量来自RICH^[15],在业务同化系统中采用了探空和飞机报的变分偏差订正^[16-17]。

欧美机构在全球大气再分析产品研发过程中已开展了大量的资料收集整理工作,但国际已发布的再分析产品同化的中国地区常规观测资料均来自全球通信系统(GTS)或双边交换,包含的地面气象站点不及国家级地面气象站的1/3,高空站点约为国内高空站点的2/3,同一探空站包含的垂直层次不到高空月报表包含观测层次的一半,包含的中国飞机报也比国内共享飞机报少。2013年底,中国气象局完成了1951年以来国家级和省级地面、高空纸质报表的数字化,通过多个步骤的质量检测、控制和人工核查订正,建成的中国地面、高空基础数据集相对GTS交换资料的完整性和质量明显提升^[18]。在此基础上,国家气象信息中心自2013年底牵头准备和启动了1979年以来全球大气再分析产品(简称CRA-40)研制。CRA-40是我国自主研发的第一代大气再分析产品。CRA-40的全球常规观测资料准备工作以提高地面、高空、飞机、海洋等基础气象资料的质量及其应用能力为核心,重点解决国内自助研制基础数据集和最新收集的多来源国外观测资料的整合、质量控制等问题,力争提高全球大气再分析常规气象资料的同化应用水平。

1 数据源和方法

1.1 数据源

CRA-40同化的常规资料(简称CRA_OBS)来自CFSR同化输入观测资料集(1979—2014,简称CFSR_OBS)、GDAS同化输入观测资料(2015年以来,简称GDAS_OBS),以及国家气象信息中心(简称NMIC)收集归档的各类常规观测资料。CFSR_OBS的详细介绍参见文献[8]。

依据完整性、和质量状况,可以将NMIC收集归档的各类观测资料分成两类。第一类为经过质量控制的观测数据集。包括:

1) NMIC研制的2400多个国家级地面气象站逐小时基础数据集,其数据源为1979—2014年中国地面气象站纸质月报表数字化资料和NMIC实时数据库(简称NMIC_RDB)资料。

2) 美国NCEI释放的全球综合地面观测资料

(ISD)，包含了气象站和机场站观测资料。

3) NMIC研制的中国高空定时值基础数据集，其数据源为1951—2012年中国高空纸质月报表数字化资料和2013年以来的高空月报数据文件，提供了将近150个无线电探空站的气压、位势高度、温度、露点温度和风场观测数据。数据集所含等压面观测数据相对于GTS交换数据的增量约为50万/年，部分站点提供了世界时06时和18时的加密观测。

4) 美国NCEI释放的综合全球探空归档数据集(IGRA)^[19]，包含了全球范围的无线电探空和测风气球观测资料。

5) 美国NOAA释放的国际综合海洋大气数据集(简称ICOADS 3.0)^[20]，包含了船舶、浮标、海岛站及潮位站观测资料。

6) 美国NCEI释放的全球飞机报数据集(1973—1998年)。

第二类为未经质量控制的观测原始报文。包括：

1) NMIC_RDB保存的全球高空天气报。

2) 国家海洋信息中心、韩国气象局等机构释放的中国、韩国、日本、俄罗斯及澳大利亚周边地区的船舶、浮标、海岛站及潮位站观测报文。

3) NMIC_RDB保存的中国AMDAR报文；美国NCAR释放的NCEP全球飞机观测原始报文(简称NCEP_AD P, 1999年10月至今)，英国CEDA释放的英国气象局实时库全球飞机报(简称MetDB_OBS)。

1.2 方法

CRA-40常规观测资料预处理包含了数据整合和标准化、质量控制、探空温度偏差订正、质量预评估、黑名单等多个处理环节(图1)。

常规观测资料进入CRA-40同化系统前以NCEP业务BUFR格式保存。为此，常规资料整合和标准化处理需要区分不同观测平台将地面、高空、飞机、海洋、风廓线等几大类资料进一步细分成若干数据子类型，以便同化中对不同平台的观测数据采用不同的观测误差。地面资料区分为地面气象站、机场站等；常规高空分为无线电探空、下投式探空、气球测风等；飞机报分为PIREP、AIREP、AMDAR和ACARS等，其中AMDAR资料又将TAMDAR和加拿大AMDAR进行单独分类。风廓线资料按照所在区域进行分类。在观测数据标准化处理前，按照上述资料分类对CRA-40收集的不同来源观测资料进行对比和整合，再进入后续的质量控制、评估、偏差订正和数据选用过程。

为避免质量较差的观测数据对同化造成负面影

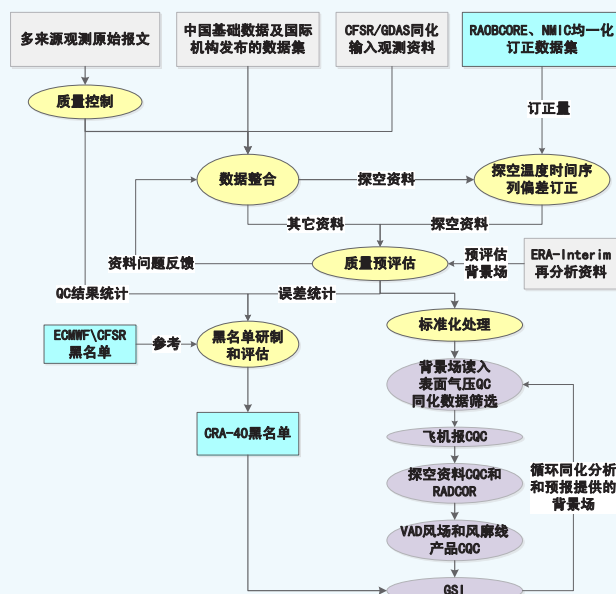


图1 CRA-40常规观测资料处理流程图

Fig. 1 Processing flow of conventional observations in CRA-40

响，全球大气再分析常规资料预处理期间，选用ERA-Interim再分析资料为参考数据，评估了不同数据源在不同时间段、不同区域相对于ERA-Interim的平均偏差(简称BIAS)和均方根误差(简称RMSE)，以及整合前后、质量控制前后、偏差订正前后各类观测数据的质量，观测数据和ERA-Interim再分析偏差过大的站点或数据源需要进行人工诊断分析，这一过程称之为质量预评估。通过质量预评估可避免存在系统性偏差或错误数据进行同化系统，是同化前观测数据质量保障的主要手段。

2 数据整合

地面、高空观测按站点进行数据整合。中国地面气象站观测和高空气象站观测分别来自NMIC研制的2400多个国家级地面气象站逐小时基础数据集和120站高空定时值基础数据集。来自CFSR_OBS和GDAS_OBS的国外地面气象站和全球机场站资料、国外无线电探空站和全球单测风站资料被直接用于CRA-40，NCEI发布的ISD数据集、IGRA数据集提供了少量CFSR_OBS、GDAS_OBS未保存的地面、高空观测数据。此外，从NMIC归档的高空天气报中补充了锦州、丹东、东沙、花莲等高空站点观测数据。由图2可知，2013—2015年，CRA-40每日进入同化的全球地面气象站观测数据约为CFSR的两倍，探空温度观测增加超过1/5。

图3给出CRA-40同化的地面气象站和机场站的空间分布图。中国2400多个地面气象站基本覆盖我国除

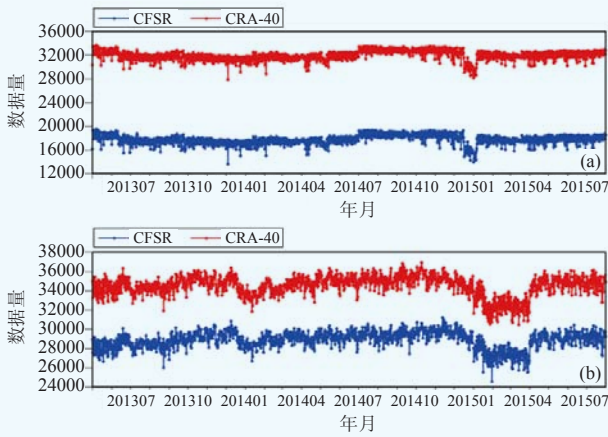


图2 2013年5月—2015年7月每日四个时次（世界时00、06、12、18时，下同）全球地面气象站表面气压（a）和探空温度（b）有效记录数。红线和蓝线分别代表CRA-40和CFSR的输入数据

Fig. 2 Six-hourly observation counts of surface pressure (a) and radiosonde temperature (b) of global SYNOP stations. Red line represents input data of CRA-40. Blue line is the input data of CFSR

青藏高原以外的大部分地区。机场报主要集中在北美和欧洲，亚洲、非洲、南美、澳大利亚等地区的机场报非常少。

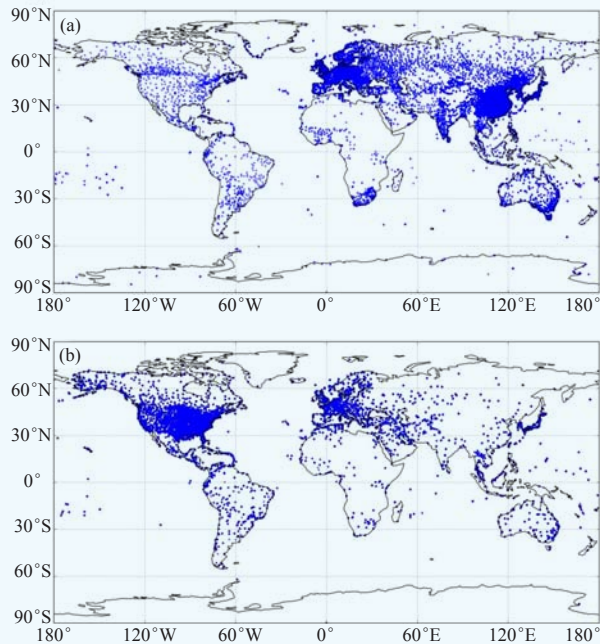


图3 进入CRA-40同化的地面气象站（a）和机场站（b）的站点分布图

Fig. 3 Map of the surface meteorological stations (a) and METAR airports (b) for the CRA-40

CRA-40同化的几类飞机报（AIREP、ACARS、AMDAR）数据源以CFSR_OBS和GDAS_OBS为主，ACARS报文开始于1992年。在此基础上，增加了

几个区域的AMDAR资料，包括NCAR网站提供的TAMDAR资料和加拿大AMDAR资料、MetDB_OBS提供的韩国AMDAR资料，以及NMIC归档的未参加国际交换的中国AMDAR资料等。

图4给出了不同时期各类飞机报观测点的空间分布。AIREP报文集中在美国、北大西洋、北太平洋和澳大利亚周边地区。部分位于北大西洋和北太平洋地区的报文经纬度精度为 5° 。由于观测要素对应的经纬度缺测，这些报文提供的经纬度为估算值。ACARS报文集中在北美和日本地区；AMDAR报文以欧洲到美洲、非洲的航线以及澳大利亚地区航线为主，部分从欧洲到日韩的飞机经过中国大陆地区时提供了少量观测数据。中国AMDAR资料开始于2003年5月，以中东部地区为主。

海洋资料主体来自CFSR_OBS和GDAS_OBS。在此基础上，以站点、船舶呼号为数据组织单元，开展CFSR_OBS、GDAS_OBS和NMIC收集归档的不同来源海洋资料的重复报对比分析和整合；对站点或船舶呼号缺测的报文，通过多源数据的观测时间、位置进行对比和整合。

图5给出整合后不同时期进入CRA-40同化系统的地面、高空、飞机和海洋观测数量。1979—2016年，CRA-40每日同化的常规观测记录量级由 10^5 增加到 10^7 。进入同化的陆地气象站观测（ P_s ）、无线电探空站（ $P_s, T, u/v, q$ ）、测风气球（ u/v ）、下投式探空（ $T, u/v, q$ ）、船舶和浮标观测（ $P_s, 10\text{米}u/v$ ）和飞机报（ $T, u/v$ ）开始于1979年。风廓线开始于1987年，覆盖了北美和日本地区。机场天气报（ P_s ）开始于1997年。

3 质量控制

在观测资料的同化应用之前，需要对整合和标准化处理后的常规观测资料进行质量控制和数据选用。CFSR_OBS、GDAS_OBS，以及中国地面、高空基础数据等经过质量控制的观测数据集可直接进入整合步骤。如数据源为未经质量控制的观测原始报文，整合前需要进行质量控制。整合前，地面气象站、船舶、浮标和海岛站观测数据经过的质量控制步骤包括气候界限值检查、主要变化范围检查、要素间内部一致性检查和时间一致性检查等；探空资料的检查包括气候界限值检查、垂直方向的静力学检查、温度递减率检查、水平风切变检查等；原始飞机报在进入整合前经过了包含航线轨迹检查在内的一系列检查，以及多步检查结果的质量控制结果综合决策。

CRA-40采用NCEP业务观测数据质量控制软件

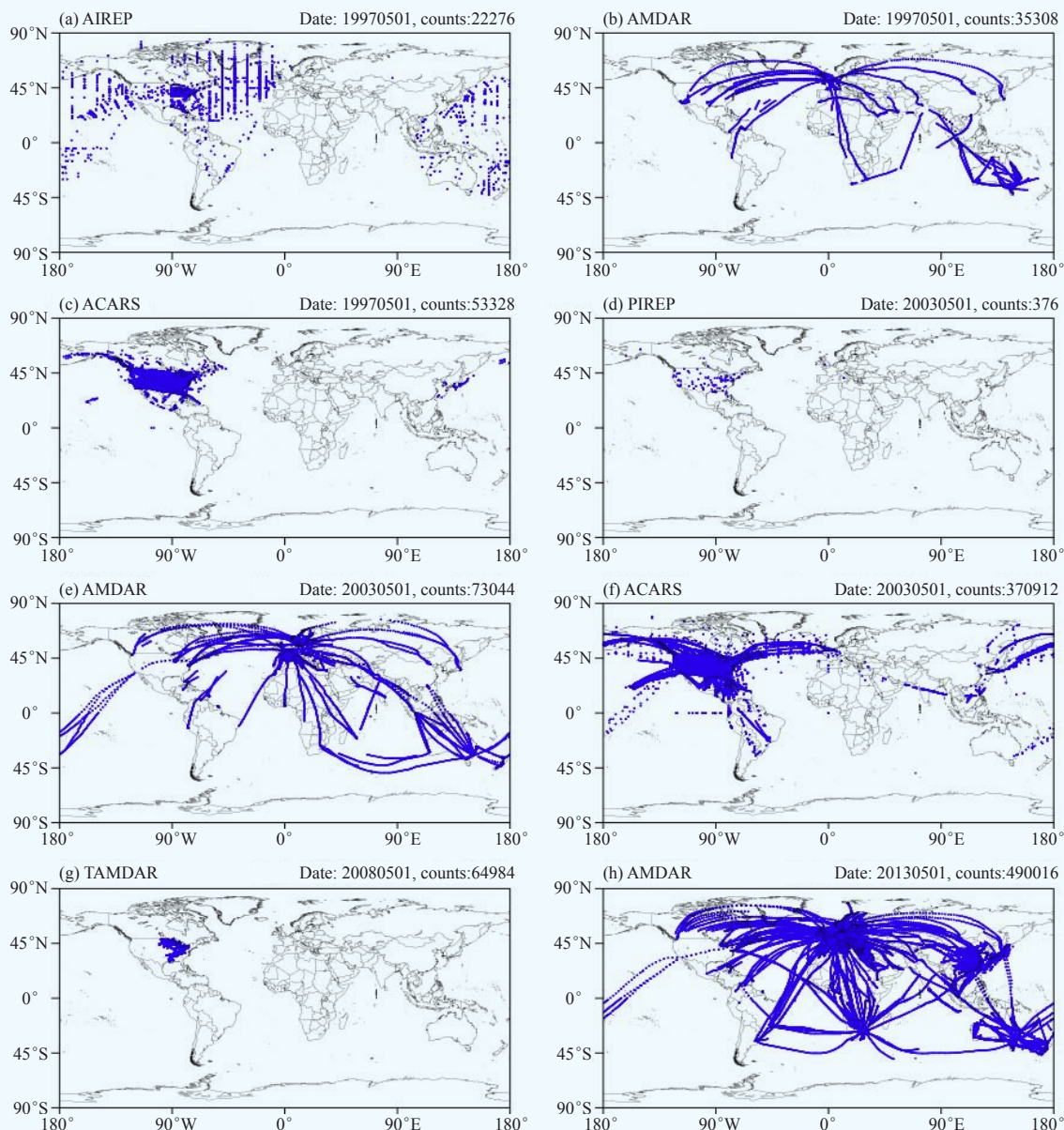


图4 不同时期的飞机报空间分布图
Fig. 4 Map of the aircraft reports in different period

实施循环同化前观测数据的质量控制，具体包括背景场数据读入、表面气压的背景场质量控制、探空资料的综合质量控制和辐射误差订正、飞机报的重复报检查和综合质量控制等。质量控制算法详见 www.emc.ncep.noaa.gov/mmb/data_processing/prepbuftr.doc/document.htm。

表1给出了2006年7月9日00时几类常规观测数据的观测数量、质量控制剔除比例以及进入同化系统的观测比例。300 hPa以上的湿度观测在同化数据筛查模块被标注为不同化。飞机报剔除的数据比例相对较高。这是由于采用BUFR格式编码的飞机报时间精度已逐步提高到秒，NCEP业务质量控制软件包含的飞

机报重复报检查对部分观测密度过大的时段或地区起到较好的稀疏化效果。

4 偏差订正

4.1 探空温度偏差订正

观测仪器和观测方式的变化会导致探空温度时间序列产生漂移。CRA-40使用RAOBCORE 1.4提供的订正量^[14]来订正2013年以前全球探空温度观测的系统偏差。RAOBCORE 1.4以ERA-40和ERA-Interim的模式背景场作为参考序列进行探空温度断点监测和均一性订正。对RAOBCORE 1.4未提供订正量的部分中国站点，参照RAOBCORE 1.4的思路计算订正量。此

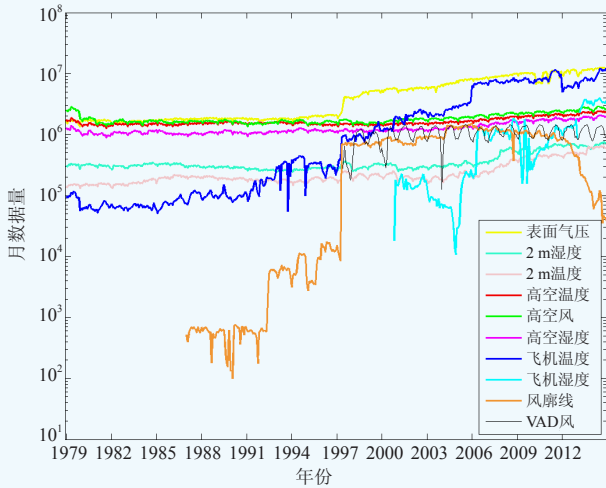


图5 不同平台观测的温度、湿度和风场的逐月数据量
Fig. 5 Monthly counts, on a logarithmic scale, of observations assimilated in the atmospheric analysis component of CRA-40

外, 太阳辐射能导致白天观测探空温度存在系统偏差, 该偏差大小与太阳高度角、仪器型号有关。经过

RAOBCORE 1.4温度订正后, 再采用NCEP的辐射偏差订正模块(简称NCEP RADCOR)处理问题, 具体参见文献[21-22]。

图6给出了2006年7月2—12日采用RAOBCORE 1.4订正值和NCEP RADCOR前后中国探空站和国外探空站温度相对于ERA-Interim再分析BIAS和RMSE的垂直分布。中国站点从2003年开始不再进行辐射偏差订正。RAOBCORE订正后中国探空温度和国外探空温度相对于ERA-Interim的RMSE都明显减小, 温度平均偏差绝对值在70 hPa以上显著下降, 和ERA-Interim更接近。在采用RAOBCORE基础上增加RADCOR订正值后, 国外探空温度的RMSE进一步减小。但两次温度调整后, 国外探空站在200 hPa附近相对ERA-Interim的温度负偏差有所增大。

从订正前后误差的空间分布图看(图7), 订正前中国探空温度在50 hPa为明显负偏差。印度、俄罗斯地区的温度为明显正偏差, 个别站点的正偏差超过5 °C。利用RAOBCORE1.4数据集实施订正后, 中

表1 2006年7月9日00时CRA-40常规观测数据量和质量控制统计表

Table 1 Data counts and quality control statistics for conventional data in CRA-40 for 0000 UTC 9 July 2006

观测类型	要素	所有层次				<100 hPa				100 - 1000 hPa				>1000 hPa			
		总数据量	同化比例	FGrej	CQCrej	总数据量	同化比例	FGrej	CQCrej	总数据量	同化比例	FGrej	CQCrej	总数据量	同化比例	FGrej	CQCrej
地面气象站	ps	372230	94.8%	3.7%	1.6%					216651	93.6%	4.0%	2.4%	155579	96.4%	3.2%	0.4%
机场站	ps	751175	96.6%	2.4%	1.0%					436330	95.9%	2.7%	1.4%	314845	97.6%	2.1%	0.4%
船舶/浮标	ps	59349	96.4%	2.3%	1.3%					4006	80.7%	8.5%	10.8%	55343	97.7%	1.8%	0.5%
船舶/浮标	q	14343	88.0%	11.2%	0.8%					507	92.9%	0.0%	7.1%	13836	87.8%	11.6%	0.5%
船舶/浮标	t	29405	88.3%	10.2%	1.5%					821	69.3%	0.0%	30.7%	28584	89.0%	10.6%	0.4%
船舶/浮标	u/v	28498	97.7%	0.8%	1.5%					490	60.7%	1.6%	37.7%	28008	98.8%	0.7%	0.5%
探空	q	214001	60.3%	0.7%	39.0%	21426	0.0%	100.0%	0.0%	183994	71.7%	0.7%	27.6%	8581	95.5%	4.3%	0.2%
探空	t	327150	98.0%	0.7%	1.4%	82514	99.2%	0.2%	0.6%	236049	97.6%	0.7%	1.7%	8587	95.2%	4.3%	0.5%
探空	u/v	370517	98.8%	1.1%	0.1%	125262	97.6%	2.4%	0.0%	237745	99.4%	0.4%	0.2%	7510	99.5%	0.2%	0.3%
飞机报(AIREP)	t	340908	97.2%	0.7%	2.1%	82514	99.2%	0.2%	0.6%	249807	96.7%	0.7%	2.6%	8587	95.2%	4.3%	0.5%
飞机报(AIREP)	u/v	383887	98.1%	1.1%	0.8%	125262	97.6%	2.4%	0.0%	251115	98.3%	0.5%	1.2%	7510	99.5%	0.2%	0.3%
飞机报(WMO-AMDA)	t	435166	87.4%	0.8%	11.8%	82514	99.2%	0.2%	0.6%	342629	85.4%	0.8%	13.8%	10023	70.7%	3.9%	25.4%
飞机报(WMO-AMDA)	u/v	522496	98.2%	0.8%	1.0%	125262	97.6%	2.4%	0.0%	385458	98.4%	0.3%	1.3%	11776	96.1%	0.1%	3.8%
飞机报(ACARS)	q	214001	60.3%	0.7%	39.0%	21426	0.0%	100.0%	0.0%	183994	71.7%	0.7%	27.6%	8581	95.5%	4.3%	0.2%
飞机报(ACARS)	t	829126	90.5%	0.7%	8.9%	82518	99.2%	0.2%	0.6%	732305	90.0%	0.6%	9.5%	14303	71.9%	5.8%	22.4%
飞机报(ACARS)	u/v	914740	96.6%	0.5%	2.9%	125266	97.6%	2.4%	0.0%	774044	96.5%	0.2%	3.3%	15430	91.9%	0.1%	8.0%
飞机报(Can-AMDA)	t	992490	91.1%	0.7%	8.2%	82518	99.2%	0.2%	0.6%	893945	90.8%	0.7%	8.6%	16027	72.6%	6.6%	20.8%
飞机报(Can-AMDA)	u/v	1054728	96.4%	0.5%	3.2%	125266	97.6%	2.4%	0.0%	912240	96.3%	0.2%	3.6%	17222	92.2%	0.1%	7.7%
飞机报(TAMDA)	t	896184	91.0%	0.6%	8.3%	82518	99.2%	0.2%	0.6%	799363	90.6%	0.6%	8.8%	14303	71.9%	5.8%	22.4%
飞机报(TAMDA)	u/v	958182	96.7%	0.5%	2.8%	125266	97.6%	2.4%	0.0%	817486	96.6%	0.2%	3.2%	15430	91.9%	0.1%	8.0%
飞机报(TAMDA)	q	276491	67.1%	0.7%	32.3%	21426	0.0%	100.0%	0.0%	246484	77.4%	0.6%	21.9%	8581	95.5%	4.3%	0.2%

注:FG rej指背景场质量控制剔除数据比例。CQC rej指综合质量控制剔除数据比例。

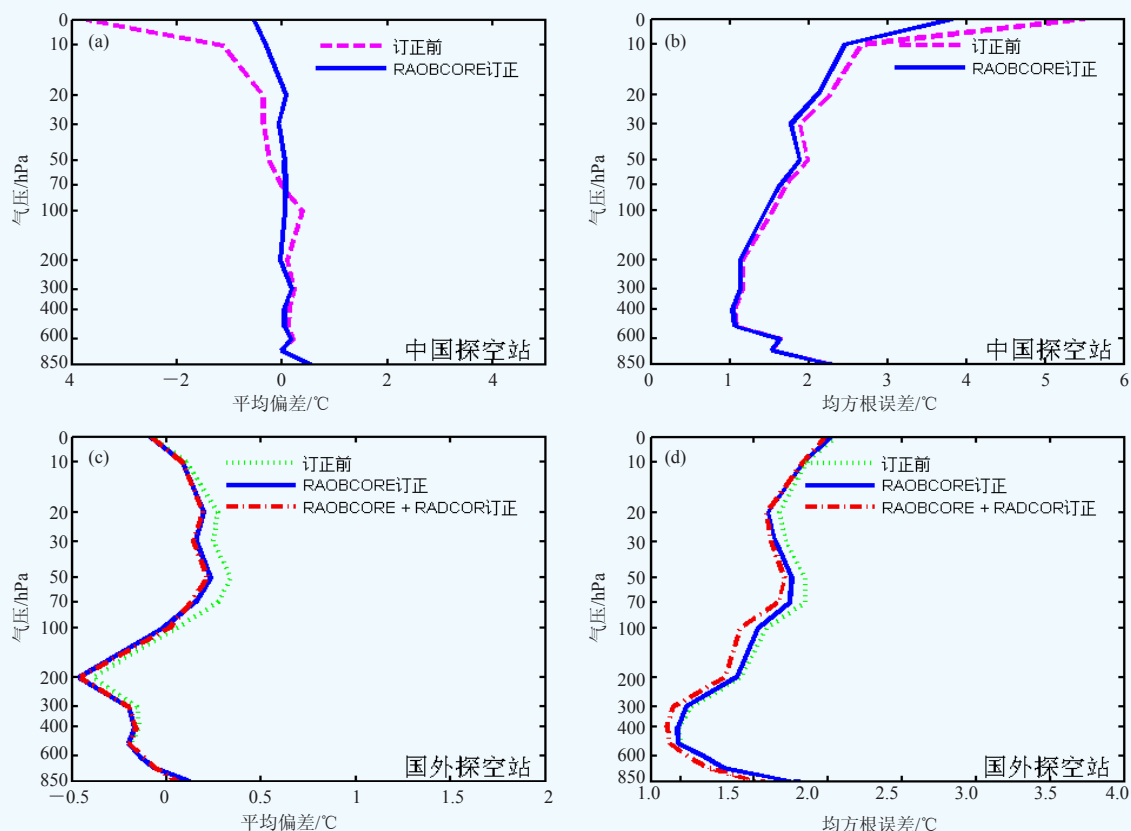


图6 探空温度相对于ERA-Interim的BIAS (a, c) 和RMSE (b, d) 垂直分布 (单位: °C)

Fig. 6 Vertical distribution of bias (a, c) and RMSE (b, d) relative to ERA-Interim of radiosonde temperature for 2-12 July 2006. Dash lines are unadjusted. Solid lines are adjusted with RAOBCORE 1.4. Dot-dash lines are adjusted with RAOBCORE 1.4 and RADCOR

国、印度、俄罗斯地区的误差明显下降，中国个别探空站的负偏差接近-2 °C。采用NCEP RADCOR订正后，部分站点的偏差有所下降，但印度、俄罗斯地区探空温度的偏差过大现象依然存在。

4.2 飞机观测温度偏差订正

研究表明，飞机观测温度相对探空温度在200 hPa附近存在暖偏差。飞机观测温度的系统偏差和飞机型号、飞行状态等因素有关。CRA-40基于NCEP GSI同化系统对飞机观测温度进行变分偏差订正，订正算法考虑了飞机的上升和下降速度。鉴于飞机观测温度偏差订正不在同化前的观测预处理环节，本文不过多阐述，具体细节参见文献[23]。

5 质量评估与同化数据筛选

5.1 黑名单应用

大气再分析在进入同化环节前关闭造成负面同化效果影响的观测资料开关，这一过程称之为数据选用 (data selection) 或黑名单 (blacklist)。除了借助质量控制剔除错误数据，CRA-40通过质量控制前的观测

资料预处理程序避免了观测站点高度与模式高度差异悬殊的资料、模式地形层以下的探空资料、300 hPa以上探空湿度等观测资料进入同化系统。

此外，CRA-40同化系统会读取外置黑名单文件。黑名单通常来自和观测资料质量相关的一些先验信息，如已知质量较差的观测站点，或者观测仪器误差较大的飞机、船舶等，或者通过质量控制、比较观测和背景场的偏差等方式来确定。对CFSR_OBS的质量控制码进行分析可以获取CFSR同化所用的部分观测数据黑名单。ECMWF早期提供的同化前观测数据选用信息包含了无线电探空、气球测风、飞机、地面气象站、船舶和浮标等几类观测数据被同化关闭的站号、船舶呼号、浮标号或者飞机标识符，但未提供具体的时间范围。由于不同时期观测数据质量存在差异，CFSR、ECMWF的黑名单并不完全适用于CRA-40。一些经验性的做法可能导致过多观测资料被关闭。如CFSR_OBS将2005—2007年的中国飞机报均标注为黑名单，但分析发现，2005—2007年的中国飞机报仅2007年4月7日—5月25日中国飞机观测温度存

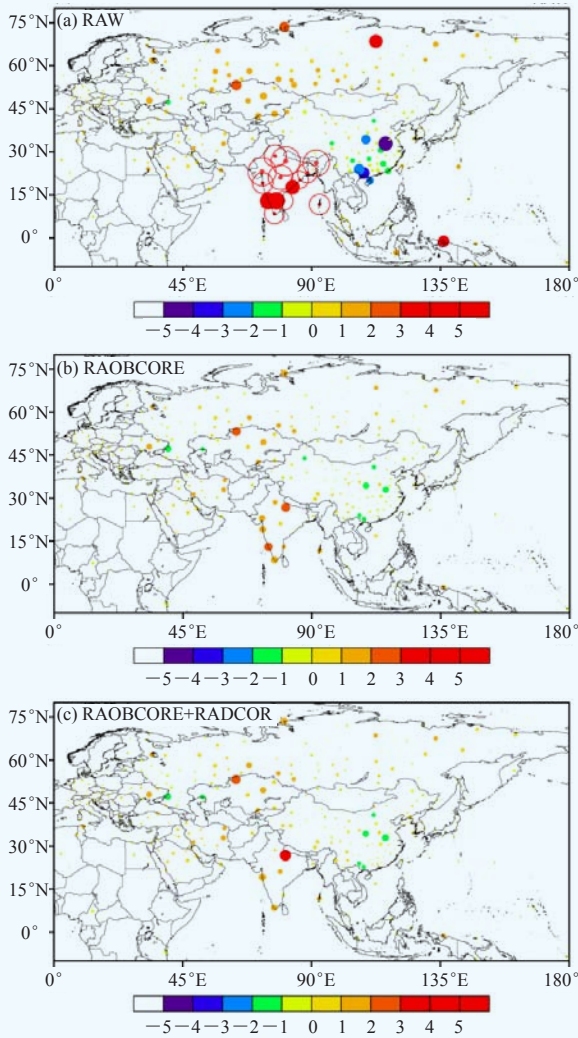


图7 2006年7月2—12日50 hPa探空温度相对于ERA-Interim的平均偏差

(a) 订正前; (b) 采用RAOBCORE 1.4订正量; (c) 采用RAOBCORE 1.4 +RADCOR订正量。空心圈表示平均偏差超过5℃

Fig. 7 Observation bias relative to ERA-Interim of radiosonde temperature on 50 hPa for 2-12 July 2006. (a) unadjusted, (b) adjusted with RAOBCORE 1.4, (c) adjusted with RAOBCORE 1.4 and RADCOR. Hollow circles represent the value is greater than 5℃

在符号错误, 其他月份观测数据未存在系统偏差或错误。CRA-40常规资料预处理阶段完成了所有常规观测资料和ERA-Interim再分析的对比, 以及CFSR_OBS和ECMWF黑名单信息的检验评估, 基于评估结果确定了观测数据的黑名单。

如图8所示, 进入CRA-40黑名单的全球探空温度和湿度数据在2010年以前比例超过了2%, 进入风场黑名单的比例相对较低, 约为0.6%。2010年开始进入探空黑名单的观测数据比例明显下降, 至2014年, 仅0.5%的探空资料被纳入黑名单。剔除黑名单站点的观测数据

后, 50~70 hPa全球探空温度观测相对于ERA-Interim的RMS有所减小, 个别月份减小约0.3℃(图9)。

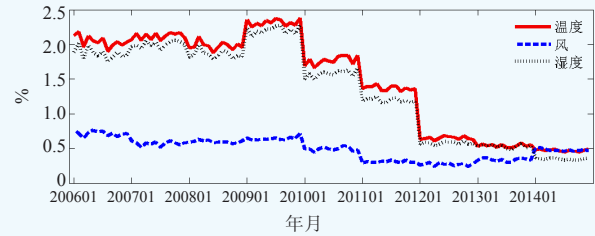


图8 2006—2014年进入探空黑名单的逐月观测数据比例
Fig. 8 Rejecting percentage by blacklisting check of radiosonde data for 2006-2014. Solid line is temperature. Dash line is wind. Dot line is humidity

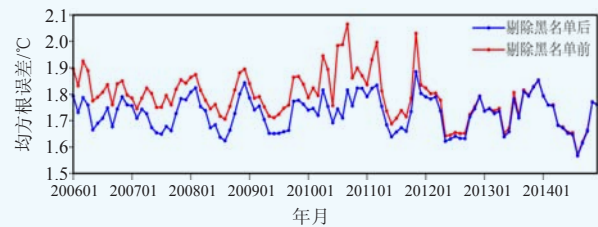


图9 剔除黑名单前、后的70~100 hPa全球探空温度相对于ERA-Interim再分析的逐月均方根误差
Fig. 9 Monthly RMSE of global radiosonde temperatures relative to ERA-Interim on 70~100 hPa for 2006 - 2014

CRA-40的飞机报黑名单来自各架飞机相对ERA-Interim再分析的BIAS和RMSE统计分析。不同时期偏差过大的飞机标识符被纳入CRA-40的黑名单清单。评估发现全球飞机报包含的PIREP报文误差过大, 温度相对AIREP、AMDAR、ACARS三类飞机报偏暖幅度接近2℃, 因此CRA-40关闭了PIREP报文的同化。加入黑名单检查后, 全球飞机报的整体误差水平有所降低, 在对流层高层对观测数据的质量改善效果相对更明显, 温度RMSE降低了约0.1℃(图10)。

5.2 湿度的评估与同化应用

国际已有评估^[24]表明, 飞机观测湿度的质量和探空湿度相当, 有助于提高局地短期的数值天气预报效果。目前ECMWF、NCEP的全球和区域业务数值天气预报模式均同化了飞机观测湿度。CRA-40常规资料准备过程中, 完成了2000年10月至今的ACARS报文和2006年5月—2011年11月的TAMDAR报文两类飞机报的湿度资料整理, 这些湿度资料主要位于美国地区。为了决策是否在全球大气再分析中同化飞机观测湿度, 对北美地区(50°—130°W, 15°—55°N)探空、ACAR和TAMDAR的比湿相对于ERA-Interim的BIAS和RMSE进行了对比分析。结果表明, 三者相对ERA-Interim的RMSE均有明显的季节变化, 夏季误差高于冬季。TAMDAR湿度的误差接近探空湿度。但2011年

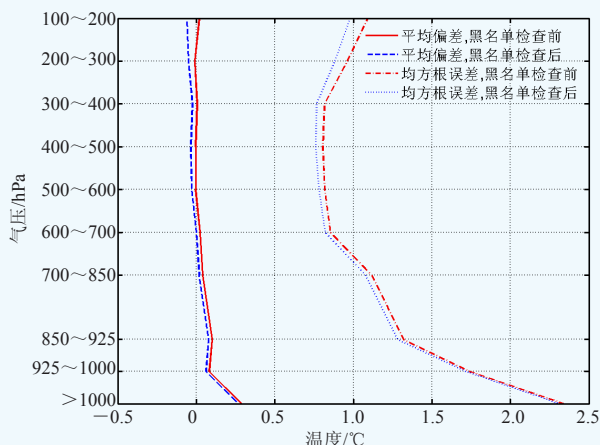


图10 黑名单检查前后2014年7月全球飞机观测温度相对于ERA-Interim再分析的平均偏差和均方根误差垂直分布
Fig. 10 Vertical distributions of bias and RMSE of global aircraft observations relative to ERA-Interim before and after blacklisting for July 2014

以前，ACARS的湿度偏大。基于评估结果，1000 hPa以上的TAMDAR湿度和2011年以来的ACARS湿度进入了CRA-40同化应用。

分别选取2008年和2014年7—8月北美地区TAMDAR、ACARS、探空观测湿度，计算观测与分析场的差异（observation-minus-analysis, OMA），以及观测与背景场的差异（observations-minus-background, OMB）。结果（图11）表明，飞机湿度的OMB、OMA和探空湿度的OMB、OMA垂直分布接近，飞机湿度OMA的平均偏差和均方根误差总体小于探空湿度的OMA。在近地面层，探空和CRA-40比湿的平均偏差超过0.5 g/kg，ACARS比湿和CRA-40分析场的平均偏差不超过0.2 g/kg。TAMDAR、ACARS比湿相对CRA-40分析场的RMSE在对流层低层小于1 g/kg，比探空湿度对应的RMSE要低0.5 g/kg。和ERA-Interim分析场相比，几类湿度观测更接近CRA-40分析场。

6 小结

CRA-40常规资料预处理的主要目的是为CRA-40及未来中国新一代的全球大气再分析提供高质量的观测资料。针对40年全球大气再分析（CRA-40）需求，国家气象信息中心从2015年开始启动全球常规资料预处理工作。以CFSR同化输入常规观测资料为主要数据源，收集和整理多个国家和机构最新释放的全球观测数据集或归档的原始观测报文，完成了多源观测资料的整合。相对国际已有再分析产品，CRA-40同化应用的全球常规资料在中国及周边地区的地面、高空、海洋和飞机观测数据量明显提升。偏差订正后中

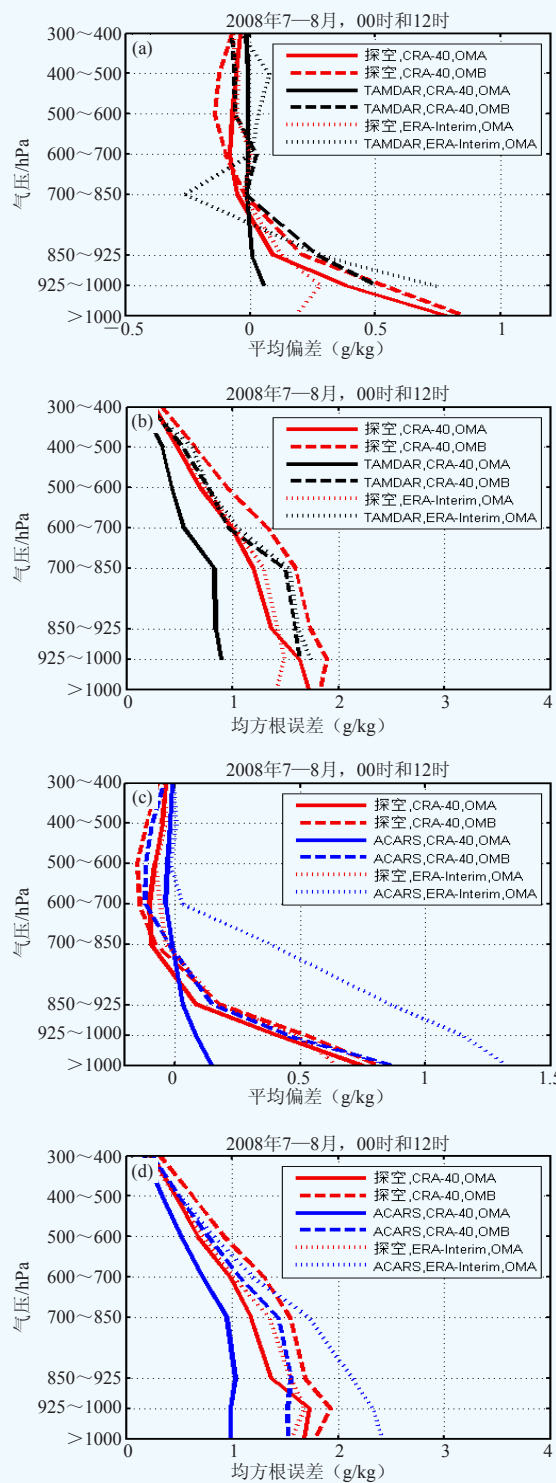


图11 2008年和2014年7—8月北美地区（50°—130°W，15°—55°N）比湿观测相对于CRA-40分析场、背景场和ERA-Interim分析场的平均偏差（a, c）和均方根误差（b, d）
Fig. 11 Bias (a, c) and RMSE (b, d) of specific humidity observations relative to CRA-40 analysis (solid lines), background (dash lines) and ERA-Interim reanalysis (dot lines) in North America for July – August in 2008 and 2014. Red lines represent radiosonde. Blue lines indicate TAMDAR. Black lines denote ACARS

附录：英文缩写定义

ACARS	Aircraft Communication Addressing and Reporting System
AIREP	Air report
AMDAR	Aircraft Meteorological Data Relay
C-AMDAR	Canadian AMDAR
CEBD	Centre for Environmental Data Analysis
CFSR	Climate Forecast System Reanalysis
CQC	Complex quality control
CRA-40	CMA Global Atmospheric Reanalysis for 40 years
DOE	Department of Energy
ERA-15	ECMWF reanalysis for 15 years
ERA-40	ECMWF reanalysis for 40 years
FGGE	First GARP Global Experiment
GARP	Global Atmospheric Research Program
IGRA	Integrated Global Radiosonde Archive
ISD	Integrated Surface Database
JRA-25	Japanese 25-year Reanalysis Project
JRA-55	Japanese 55-year Reanalysis Project
MERRA	Modern Era Retrospective-analysis for Research and Applications
NCEI	National Centers for Environmental Information
PIREP	pilot reports
RAOBCORE	Radiosonde Observation Correction using Reanalysis
RICH	Radiosonde Innovation Composite Homogenization
TAMDAR	Tropospheric Airborne Meteorological Data Reporting

国探空温度在平流层的负偏差减弱，全球探空温度相对ERA-Interim的RMSE在垂直各层总体减小。剔除黑名单对应观测数据可提高全球常规观测资料的整体质量。此外，CRA-40再分析产品研制新增了2006年以来部分飞机观测湿度的同化。

2017年6月，经过预处理的2006—2016年全球常规资料已进入CRA-40的10年中间产品研制。下一步，CRA-40常规资料预处理工作将继续完成全球40年常规资料的预处理工作，重点解决40年资料黑名单的建立和评估问题。前期评估已发现，由于采用的元数据信息准确度不高，RAOBCORE 1.4对1980—1990年代中国探空温度的订正量存在一些问题。CRA-40计划在40年中国探空温度订正中采用自主研发的均一性订正结果，并对RADCOR采用的中国历史探空站辐射偏差订正系数进行优化。

参考文献

[1] Dee D P, Uppala S M, Simmons A J, et al. The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological*, 2011, 137(656): 553–597.

[2] Gibson J K, Kallberg P, Uppala S M, et al. ERA description. ERA-15 Report Series, No. 1, ECMWF, 1997.

[3] Kalnay E, Kanamitsu M, Kistler R, et al. The NCEP /NCAR 40-year reanalysis project. *Bulletin of the American Meteorological Society*, 1996, 77: 437–471.

[4] Schubert S D, Park C K, Wu C Y, et al. A multi-year assimilation with the GEOS -1 systems: Overview and results. *NASA Technical Memorandum*, 1995.

[5] Kanamitsu M, Ebisuzaki W, Woollen J, et al. NCEP-DOE AMIP-II Reanalysis (R-2). *Bull Amer Meteor Soc*, 2002, 83: 1631–1643.

[6] Uppala S M, Kállberg P W, Simmons A J, et al. The ERA-40 reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 2005, 131: 2961–3012.

[7] Onogi K, Tsutsui J, Koide H, et al. The JRA-25 reanalysis. *Journal of the Meteorological Society of Japan*, 2007, 85(3): 369–432.

[8] Saha S, Moorthi S, Wu X, et al. The NCEP climate forecast system version 2. *J Climate*, 2014, 27: 2185–2208.

[9] Rienecker M, Suarez M J, Ronald G, et al. MERRA: NASA's modern-era retrospective analysis for research and applications. *J Climate*, 2013, 24: 3624–3648.

[10] Kobayashi S, Yukinari O, Harada Y, et al. The JRA-55 reanalysis: General specifications and basic characteristics. *Journal of the Meteorological Society of Japan*, 2015, 93: 5–48.

[11] András H, Berrisford P, Biavati G, et al. The impact of observations in the latest ECMWF reanalysis system. 6th Workshop on the Impact of Various Observing Systems on NWP. Shanghai, China. May 2016.

[12] Onogi K. A data quality control method using forecasted horizontal gradient and tendency in a NWP system: dynamic QC. *J Meteor Soc Japan*, 2009, 76: 497–516.

[13] Collins W G. The operational complex quality control of radiosonde heights and temperatures at the National Centers for Environmental Prediction. Part I: Description of the method. *J Appl Meteor*, 2010, 40: 137–151.

[14] Haimberger L. Homogenization of radiosonde temperature time series using innovation statistics. *Journal of Climate*, 2007, 20(7): 1377–1403.

[15] Haimberger L, Tavolato C, Sperka S. Homogenization of the global radiosonde temperature dataset through combined comparison with reanalysis background series and neighboring stations. *Journal of Climate*, 2011, 25: 8108–8131.

[16] Milan M, Haimberger L. Predictors and grouping for variational bias correction of radiosonde. *ECMWF ERA Report Series*, 2015, No. 2.

[17] Isaksen L, Vasiljevic D, Dee D, et al. Bias correction of aircraft data implemented in November 2011. *ECMWF Newsletter*, 2012, 131: 6.

[18] 熊安元. 中国地面和高空气候变化数据产品研发技术. 北京: 气象出版社, 2015.

[19] Durre I, Vose R S, Wuertz D B. Overview of the integrated global radiosonde archive. *J Climate*, 2006, 19(1): 53–68.

[20] Woodruff S, Worley S, Lubker S, et al. ICOADS release 2.5: extensions and enhancements to the surface marine meteorological archive. *International Journal of Climatology*, 2010, doi:10.1002/joc.2103.

[21] Collins W. Determination of new adjustment tables in order to bring radiosonde temperature and height measurements from different sonde types into relative agreement. *EMC/NCEP/NOAA*, 1999.

[22] Ballish B A, Kumar V K. Systematic difference in aircraft and radiosonde temperatures: Implications for NWP and climate studies. *Bull Amer Meteor Soc*, 2008, 89(11):1689–1708.

[23] Zhu Y, Derber J, Purser R J, et al. Variational correction of aircraft temperature bias in the NCEP's GSA analysis system. *Mon Wea Rev*, 2015, 143: 3774–3803.

[24] Petersen R A, Cnonce L, Mamrosh R, et al. On the impact and future benefits of AMDAR observations in operational forecasting. Part II: Water Vapor Observations. *Bull Amer Meteor Soc*, 2016, 11: 2117–2134.