Integration, Quality Assurance, and Usage of Global Aircraft Observations in CRA

Jie LIAO¹, Huiying WANG¹, Zijiang ZHOU¹*, Zhiquan LIU², Lipeng JIANG¹, and Fang YUAN¹

¹ National Meteorological Information Center, China Meteorological Administration, Beijing 100081, China
² National Center for Atmospheric Research, Boulder, CO 80301, USA

(Received May 30, 2020; in final form November 18, 2020)

ABSTRACT

This paper presents a detailed description of integration, quality assurance procedure, and usage of global aircraft observations for China’s first generation global atmospheric reanalysis (CRA) product (1979–2018). An integration method named “classified integration” is developed. Aircraft observations from nine different sources are integrated into the Integrated Global Meteorological Observation Archive from Aircraft (IGMOAA), a new dataset from the National Meteorological Information Center (NMIC) of the China Meteorological Administration (CMA). IGMOAA consists of global aircraft temperature, wind, and humidity data from the surface to 100 hPa, extending from 1973 to the present. Compared with observations assimilated in the Climate Forecast System Reanalysis (CFSR) of NCEP, the observation number of IGMOAA increased by 12.9% between 2010 and 2014, mainly as a result of adding more Chinese Aircraft Meteorological Data Relay (AMDAR) data. Complex quality control procedures for aircraft observations of NCEP are applied to detect data errors. Observations are compared with ERA-Interim reanalysis from 1979 to 2018 to investigate data quality of different types and aircraft, and subsequently to develop the blacklists for CRA. IGMOAA data have been assimilated in CRA in 2018 and are real-time updated at the CMA Data-as-a-Service (CMADaaS) platform. For CRA, the fits to observations improve over time. From 1994 to 2018, root-mean-square error (RMSE) of observations relative to CRA background decreases from 1.8 to 1.0°C for temperature above 300 hPa, and from 4.5 to 3 m s⁻¹ for zonal wind. The RMSE for humidity appears to exhibit an apparent seasonal variation with larger errors in summer and smaller ones in winter.

Key words: aircraft observation, integration, quality assurance, blacklist, China’s global atmospheric reanalysis (CRA) product

Citation: Liao, J., H. Y. Wang, Z. J. Zho, et al., 2021: Integration, quality assurance, and usage of global aircraft observations in CRA. J. Meteor. Res., 35(1), 1–16, doi: 10.1007/s13351-021-0093-3.

1. Introduction

Data obtained from aircraft have played important roles in meteorological research and operations for over a century (Fleming, 1996; Moninger et al., 2003; Petersen, 2016). Low-cost automated aircraft wind and temperature observations now constitute the third most important dataset for short-range global NWP (Petersen, 2016). Long-term historical aircraft observations have been used to characterize the diurnal variation of the planetary boundary layer (Rahn and Mitchell, 2016; Zhang et al., 2020), and reexamine the wind–pressure relationship in tropical cyclones (Bai et al., 2019).

Automated aircraft wind and temperature observations constitute an important dataset for global NWP and reanalysis. Historical data collection and integration are the two main tasks performed for almost all atmospheric reanalysis projects. For instance, the primary sources of aircraft observations assimilated in ERA-15 were operational GTS data from the ECMWF and the JMA, which were found to contain considerably more aircraft data over the North Pacific than the MARS archive of the ECMWF (Gibson et al., 1999). Aircraft observations assimilated in the NCEP/NCAR reanalysis are available

Supported by the National Innovation Project for Meteorological Science and Technology (CMAGGTD003-5), China Meteorological Administration Special Public Welfare Research Fund (GYHY201506002), and National Key Research and Development Program of China (2017YFC1501801).

*Corresponding author: zzj@cma.gov.cn

© The Chinese Meteorological Society and Springer-Verlag Berlin Heidelberg 2021
from the NCEP GTS source alongside additional data, including data from New Zealand, the GARP Atlantic Tropical Experiment, FGGE, TWERLE, and U.S. Air Force reconnaissance (Kalnay et al., 1996). The bulk of CFSR aircraft observations are taken from the U.S. operational NWS archives. In addition to the U.S. NWS archives, CFSR aircraft observations include a number of archives from U.S. military and national sources, and datasets supplied by NCAR, ECMWF, and JMA (Saha et al., 2010).

The observations received from the GTS and other data sources are usually not quality controlled, and therefore contain gross errors due to instrumental and human errors at measurement recording and during the data transmission. The minimum quality control of aircraft-based observations before their provision to WIS has been written in the Guide to Aircraft-based Observations (WMO, 2017). NCEP operational observation quality control procedures include the ACQC (https://www.emc.ncep.noaa.gov/mmb/data_processing/prepbufr.doc/document.htm). The NCEP ACQC has been used in CFSR and MERRA (Rienecker et al., 2011) before the aircraft observations are assimilated.

In 2014, the CMA started its mission to produce China’s first generation global atmospheric reanalysis (CRA) product for 40 yr (1979–2018; Liu et al., 2017). Before the CRA project implementation, NMIC archived the original global AMDAR reports from GTS and Chinese AMDAR reports from the AMC of the Air Traffic Management Bureau of Civil Aviation Administration of China in the operational database. These data begin from March 2003, and thus do not cover the whole analysis period of CRA. However, there are many additional Chinese AMDAR data sources since 2011, which are restricted to real-time exchange by GTS and available only through NMIC operational systems. To date, no reanalysis has made use of the additional restricted Chinese AMDAR data. However, for various reasons, including differences in decoding practices, some messages transmitted over the GTS are decoded only at some receiving centers but not at others. Thus, to meet the scientific objective of CRA, it is necessary to combine aircraft observations from NMIC’s database and other data sources, especially those datasets that have been used in reanalysis projects (e.g., aircraft observations assimilated in CFSR), into an integration dataset.

In 2015, the NMIC of the CMA began the production of IGMOAA to provide aircraft observations to CRA. The goals of the IGMOAA are as follows: 1) to combine as many reliable data sources as possible into one aircraft archive, 2) to reformat these reports into a standard format according to the requirement of CRA, 3) to apply quality control algorithms that remove data with gross errors, and 4) to develop blacklists to exclude observations that are deemed of poor quality.

The IGMOAA aircraft dataset was created over several steps (Fig. 1). First, the data-source priority used during data integration was determined. Second, all data sources were reformatted into a common data format with standard units, variables, and definition of the aircraft POF. Next, a classification integration strategy was used to combine the observations from different data sources into a single data file for every 6 h, in which duplicate observations were eliminated. A comprehensive suite of QA procedures was then applied to the integrated data files. Finally, observations with quality markers were composited into a dataset from January 1973 through December 2018.

This paper describes the details of integration and quality assurance procedure for IGMOAA, and the usage of IGMOAA in the CRA. Data sources and standardization are discussed in Sections 2 and 3. Section 4 provides the detail of the integration method. A summary of the quality assurance procedure is given in Section 5. Section 6 presents the usage of aircraft observations in CRA. A summary is provided in Section 7.

2. Data sources

2.1 History of aircraft-based observations

Manual meteorological observations from commercial aircraft began in the 1930s. Aircraft observations of the operational archives started in 1962 at the NWS of the U.S. (Saha et al., 2010). These early PIREPs were
produced from a pilot talking to a ground observer. The first automated AIREPs began during the FGGE in 1978–1979. Before the 1990s, aircraft observations mainly took place through the traditional AIREP and PIREP measurements.

In general, AMDAR refers to automated meteorological reports from aircraft. The first operational AMDAR program commenced in 1986, with five aircraft producing less than 1000 observations per day. Since then, the program has grown rapidly, with more than 6000 aircraft worldwide contributing approximately 700,000 reports per day. In the U.S., however, these data are usually referred to as ACARS data, sometimes referred to as the MDCRS (Pauley, 2002). ACARS is part of the broader WMO AMDAR program (WMO, 2003). AMDAR data from outside the U.S. are referred to as NUS-AMDAR in this paper. TAMAR sensors were developed in 2004 to improve mesoscale numerical weather prediction (Daniels et al., 2006). These sensors offer high spatiotemporal resolution observations within the troposphere (Gao et al., 2019).

2.2 Data sources of IGMOAA

The primary aim of creating IGMOAA is to supply CRA with aircraft observations with increased spatial and temporal coverage. Nine data sources (Table 1) of IGMOAA are collected from NMIC, NCEI, RDA, and CEDA. These data sources contain PIREP, AIREP, and AMDAR data from 1973 to the present.

The data sources listed in Table 1 have many duplicate observations over the same period. Before integration, PS of nine data sources are calculated to determine their priorities. The PS consists of five components as described in Table 2, and is calculated as follows:

$$PS = P_1 \times P_2 + \sum_{i=3}^{5} P_i.$$  (1)

The priorities of these data sources for IGMOAA are listed in Table 1.

The following three data sources supplied more than 90% of observations and were identified as being of highest priority in the corresponding periods:

1. Aircraft observations assimilated by CFSR for 1979–2014. All conventional CFSR observations within a 6-h time window are archived in a monolithic PREP-BUFR file with the quality markers. CFSR_OBS is the most complete data source, with few data gaps. Moreover, CFSR_OBS provides quality markers that help users to identify observations that are not assimilated by CFSR. Therefore, CFSR_OBS is considered to be the highest priority data source for IGMOAA given its high integrity and high-quality information.

2. Global upper air observational weather data from NCEP GTS (October 1999–present). The dataset includes radiosonde, pibal, and aircraft reports from the GTS, and satellite data from the NESDIS. The reports within a 6-h time window are archived into the NCEP BUFR format. These real-time data are the primary input to the NCEPGDAS, which is used to create the NCEP FNL. NCEP_GTS provides comparatively integrated global aircraft reports with high availability and is thus considered to be the second priority.

3. The Historical Worldwide Aircraft Reports dataset released by the NCEI (DSI-6380) for 1973–1998. Even though DSI-6380 lacks data since 1998, it provides relatively complete global aircraft reports before 1979. Moreover, all reports in DSI-6380 are quality controlled with a syntax check, plausibility check, contradiction check, and diagnostic check (NCDC, 2002). Therefore, DSI-6380 is listed as the third priority.

AMDAR reports collected in real-time by NMIC are an important supplemental data source. Five other supplemental data sources include: 1) AMDAR reports collected by the Met Office, which originate from the CEDA; 2) U.S. NMC Global GTS Aircraft Data released by NCEI (DSI-6106), which are not quality controlled; 3) U.S. Air Force ETAC DATSAV TD57 Global Aircraft Observations; 4) Australian Aircraft Observations; and 5) New Zealand Aircraft Data. The latter three datasets are obtained from NCAR’s RDA.

To ensure the data quality (especially data with high priority) and to avoid data misuse with systematic quality problems, different data sources were evaluated separately. Data from different sources were first each compared with the ERA-Interim reanalysis. We call this process as quality pre-evaluation, which was also applied to the development of the blacklists for CRA. More details are presented in Section 5.2. By pre-evaluation, we confirmed that the biases and RMSEs of CFSR_OBS relative to ERA-Interim are similar to NCEP_GTS, METDB, and other data sources at the same time.

Second, we analyzed the difference between the Priority 1 data and other data sources. When the observation times and positions of more than two data items from different sources are the same or fall within the similarity thresholds (pressure: 10 hPa; time: 10 minutes; latitude and longitude: 0.1°), the degree of similarity between Priority 1 data and other data sources is diagnosed. In October 1984, only 0.22% of Priority 1 temperatures (CFSR_OBS) are obviously different from all other data sources (absolute difference is more significant than 1°C). In October 2014, the percentage is about 0.07%. In
summary, CFSR_OBS showed no clear difference from other data sources and was thus considered reliable for application as Priority 1 data.

3. Standardization

Besides temporal and spatial coverage, the data format, variables, units, and definition of POF vary among these data sources. To simplify the subsequent integration process, the data sources were first reformatted into a common data format with the standard units, variables, and definition of POF. For each data source, observations of every 24 h (from 2100 UTC of the prior day to 2100 UTC of that day) were grouped into four files in PREPBUFR format (Keyser, 2018), each of which contains observations of a 6-h time window: 0000 ± 0300, 0600 ± 0300, 1200 ± 0300, and 1800 ± 0300 UTC. File names were distinguished by the standard synoptic time (0000, 0600, 1200, and 1800 UTC). Table 3 lists the variables and respective units in the PREPBUFR files.

Pressure ($p$) values of observation points are missing in partial historical data. For example, data decoded from the AMDAR report with TAC format provide only pressure altitude (PALT) to denote the vertical position of aircraft. PALT is calculated from the measured static pressure by using a standard atmosphere. In a few cases,
PALT may depart considerably from the true altitude, or even appear below the sea level when the actual pressure near the surface is greater than the standard atmospheric pressure. For these reports with missing pressure and available PALT, pressure values were recalculated with PALT by the algorithm in WMO (2003)\(^1\).

Water vapor observations are made during all phases of flight and are attached to independently measured temperature and wind data to form a single AMDAR report for transmission to the ground (Petersen et al., 2016). ACARS data in CFSR_OBS contain water vapor observations represented by the variable specific humidity since October 2000. NCEP_GTS contains TAMDAR data from May 2006 through November 2011, which included wind, temperature, and relative humidity. Relative humidity measurements were converted to specific humidity before data integration.

In the data sources, the specification of the POF indicator is different between TAC and BUFR aircraft reports. POF is a key variable of aircraft reports and is helpful to assess the potential system deviation of observations or to implement bias correction of temperature. Ballish and Kumar (2008) and Petersen et al. (2016) showed the dependence on the aircraft’s POF for temperature and humidity. For convenience and ease of use, we unified the specification of POF indicators according to the code table given in WMO (2017).

AMDR and ACARS data use registration number (tail number) or flight number as the identifier of an aircraft. In contrast, PIREP, AIREP, and NUS-AMDAR usually provide only one of these, most of which are registration numbers. We retained both identifiers in the PREPBUFR files. Observations within a 6-h time window were placed into a logical order based on aircraft identifier and observing time.

### 4. Data integration

For every standard synoptic time, there is a set of “mingle groups,” each of which contains one or more PREPBUFR files corresponding to one or more data sources. For example, the mingle group for 0000 UTC 1 February 2014 consists of four PREPBUFR files from four sources (CFSR_OBS, NCEP_GTS, NMIC_RDB, and METDB), and the partial observations of these data overlap.

When the standardization of data source is finished, data in each mingle group are integrated into a single data file for each synoptic time. IGMOAA aims to combine as many observations as possible into the archive and eliminate duplicate data. For this purpose, an integrat-

---

\(^1\) For PALT equal to or less than 36,089 ft, static pressure (SP) is related to PALT (ft) by the following expression:

\[
SP (hPa) = 1013.25(1 - 10^{-6} \times 6.8756 \times PALT)^{5.2559}.
\]

If PALT is greater than 36,089 ft, static pressure is given by:

\[
SP (hPa) = 226.32 \times \exp[-(PALT - 36089)/20805].
\]

---

### Table 3. Parameters archived in IGMOAA

| Parameter                        | Unit       | Type   | Meaning of flag values                                                                 |
|----------------------------------|------------|--------|----------------------------------------------------------------------------------------|
| Aircraft registration number     | –          | Character | –                                                                                   |
| Aircraft flight number           | –          | Character | –                                                                                   |
| Report type                      | –          | Integer | 130/230: AIREP/PIREP<br>131/231: NUS-AMDAR<br>133/233: ACARS<br>134/234: TAMDAR<br>135/235: Canadian AMDAR |
| Year                             | –          | Integer | –                                                                                   |
| Month                            | –          | Integer | –                                                                                   |
| Day                              | –          | Integer | –                                                                                   |
| Hour                             | –          | Integer | 00UTC/06UTC/12UTC/18UTC                                                            |
| Relative time                    | h          | Float   | Observation time minus cycle time                                                    |
| Latitude                         | °          | Float   | –                                                                                   |
| Longitude                        | °          | Float   | –                                                                                   |
| Pressure                         | hPa        | Float   | –                                                                                   |
| Geopotential height              | gpm        | Float   | –                                                                                   |
| Phase of aircraft flight         | –          | Integer | See WMO (2017), page 89                                                             |
| Temperature                      | °C         | Float   | –                                                                                   |
| \(U\) wind                       | m s\(^{-1}\) | Float   | –                                                                                   |
| \(V\) wind                       | m s\(^{-1}\) | Float   | –                                                                                   |
| Specific humidity                | g kg\(^{-1}\) | Float   | –                                                                                   |
| Quality control marker           | –          | Integer | –                                                                                   |
tion method named “classified integration” was developed. Figure 2 shows the data integration process. The classified integration method comprises two different integration strategies: route integration and point-to-point integration. At the beginning of the integration, the data source of the highest priority is first selected as the initial core data. Then the integration procedure is run one or more times according to the number of data sources at this synoptic time. After each integration, the core data are appended to new observations and updated for each integration. The details of data integration are described as follows.

4.1 Route integration

To date, the number of daily aircraft observations has exceeded 0.7 million, most of which are AMDAR data. AMDAR data usually contain the identifier of the aircraft, which is important to eliminate duplicate data of the same route by comparing the identifiers of two data sources. Thus, a route integration strategy was selected for the AMDAR data. The CFSR_OBS and NCEP_GTS data were selected as the initial core data for 1979–2014 and 2015–2018, respectively.

Each aircraft identifier of AMDAR in the supplemental source first undergoes a basic plausibility check, named the “invalid identifier check.” An aircraft identifier is reported as a string of fewer than eight characters. If an identifier is missing or contains less than five characters, it will be considered as invalid. Point-to-point integration is applied for those AMDAR data with an invalid identifier. Then the valid identifiers in the supplemental source are compared with the identifiers in the core data. Observations of the same aircraft from the two data sources are sorted according to the observation time. When there are two observations at the same time and position, the core data item is retained if it has passed through QC procedures. If it does not pass through QC procedures, but the supplemental data item does pass, the supplemental data item is the preference. Otherwise, the core data item is retained. METDB provides more dense observations of Japanese airlines than CFSR_OBS, NCEP_GTS, and NMIC_RDB. Based on the route integration, these additional Japanese AMDAR data are integrated into the core data.

If an identifier from a supplemental data source does not appear in the core data, the identifier will be con-

Fig. 2. Flowchart illustrating the data integration process for developing IGMOAA.
sidered to be from a new route. All the observations from a new route are added to the core data. For AMDAR data, the area of origin of the aircraft is easily known according to the prefix characters of the identifier. We found that CFSR_OBS did not contain TAMAR, Canadian AMDAR, and Korean AMDAR data. As a result, TAMAR and Canadian AMDAR data of IGMOAA before 2015 were taken from NCEP_GTS. Through this approach, IGMOAA filled two gaps in CFSR_OBS (23 August 2010 and 12 March 2014), and thus provided more AMDAR data than CFSR_OBS. These additional data included Korean AMDAR data from March 2011 through September 2014, Hong Kong AMDAR data since June 2012, and restricted Chinese mainland AMDAR data since 2011. Figure 3 presents an example. NMIC_RDB provides 2.4% of IGMOAA temperature at 0000 UTC 5 October 2014. The percentage of data supplied by NCEP_GTS and METDB is about 0.8% and 0.3%, respectively. Figure 4 shows monthly temperature observation numbers of global aircraft reports in IGMOAA and CFSR_OBS. Compared with CFSR_OBS, the observation number of IGMOAA increased by 19.2% in 2006–2009, and 12.9% in 2010–2014.

4.2 Point-to-point integration

Most of the global aircraft observations before the mid-1990s are AIREP and PIREP, and observation numbers are usually less than 10,000 per day. To date, numbers of AIREP and PIREP observations have only a small increase. Aircraft identifiers of many PIREP and AIREP observations (especially before the mid-1990s) are missing or invalid. Thus, the point-to-point integration strategy was applied to process AIREP and PIREP data.

In this point-to-point integration, every AIREP or PIREP observation from the supplemental source is compared with all the AIREP and PIREP observations in the core data. Taking into account the differences in processing procedures among the various data sources, two observational records are considered to be closely matching each other if the differences between pairs of values fall within the similarity thresholds. If the difference of observation time is less than 10 minutes, the similarity thresholds are 10 hPa and 0.1° for pressure and latitude/longitude, respectively. If the difference of observation time exceeds 10 minutes but is less than 30 minutes, the similarity thresholds are 1 hPa and 0.01° for pressure and latitude/longitude, respectively. If such a match is unavailable from the core data, the AIREP or PIREP data will be added into the core data. Based on this point-to-point integration, more historical observations over Australia, New Zealand, North Atlantic, North Pacific Ocean, Mediterranean, and West Asia have been added into DSI-6380, the initial core data before 1979.

4.3 Distribution of observations in different periods

Almost all observations are below 100 hPa. Data above 100 hPa represent less than 0.008% of the total observations, most of which are AIREPs. Figure 5 shows monthly numbers of aircraft observations on a logarithmic scale. Before 1979, there are less than 0.2 million
The number of aircraft observations assimilated in CRA has increased from approximately $10^5$ per month on average at the beginning of the 1980s, to nearly $10^6$ per month in 1998, and has exceeded $10^7$ in recent years. The number of wind observations is very close to that of temperature observations. IGMOAA contains humidity observations from TAMDAR (May 2006 to November 2011) and ACARS (October 2000 to the present). Before 2009, there are less than 0.1 million monthly number of humidity observations from ACARS, a quarter of that from TAMDAR. In subsequent years, humidity observations from ACARS have increased rapidly. To date, the number of humidity observations per month has reached about 0.8 million.

As shown in Fig. 6, the geographic distribution of IGMOAA is very uneven. To examine this in more detail, the data distribution of temperature observations from aircraft was plotted for the same month in 1988, 1998, 2008, and 2018. The first thing that can be seen is the difference in the total number of observations (the scale...
is the same in the four pictures). In 1988, daily average data counts on most $2^\circ \times 2^\circ$ grid boxes are less than 20, and the observations are mainly distributed in North Atlantic, Mediterranean, and the east coast of Australia. Daily average data counts for most $2^\circ \times 2^\circ$ grid boxes exceed 200 in North America in 1998, mainly as a result of the wide application of MDCRS in the U.S. (Martin et al., 1993; Petersen, 2016). With the implementation of the WMO AMDAR program, Europe and China became new regions with high observation density in 2008. Until 2018, there are almost 0.8 million AMDAR reports every day. The number of observations has exceeded 10 thousand per day on $2^\circ \times 2^\circ$ grids in some areas of North America, Europe, China, and Australia.

5. Quality assurance

Quality assurance of IGMOAA includes three aspects (Fig. 7). First, the systematic temperature errors of Chinese AMDAR from March to May in 2007 are corrected. Second, all observations are subject to a suite of quality control procedures, which include invalid identifier check and ACQC. Last, data selection rules specified in the so-called blacklists are applied to exclude observations deemed to be of poor quality.

5.1 Correction of historical Chinese aircraft temperature observations

The Chinese AMDAR aircraft reports from AMC contained errors from March to May in 2007. In these cases,
positive (> 0°C) temperatures were reported as negative Celsius temperatures (Tavolato and Isaksen, 2015). These errors occurred in ascent or descent over airports. To identify these erroneous data points, the height of the freezing level of each flight route over Chinese airports was estimated by ERA-Interim reanalysis (Dee et al., 2011) data from March to May in 2007. These negative Celsius temperatures were correspondingly corrected below the height of the freezing level. The use of estimation height of the freezing level was convenient to identify these incorrect temperatures, even though some differences occurred between the estimated height and the true height. For example, Fig. 8 represents the vertical profiles from radiosonde and aircraft for 0000 UTC 8 May 2007. After the correction, temperature profiles from aircraft had a similar vertical variability to the radiosonde temperatures. The large deviation found for a few observations may be associated with the differences in observing time and position between the aircraft and radiosonde.

5.2 Quality control

Before entering ACQC, the validity of aircraft identifiers is checked. Observations are grouped according to the identifiers and sorted by the observation time for the track check, an important step of ACQC. The invalid identifier is likely to influence the effect of ACQC. Invalid identifiers are very common in the 1980s and 1990s. For example, there are 508 identifiers archived in the AIREP data at 1800 UTC 31 March 1985. Of these, 66 contain less than five characters, and 23 contain only three characters. These abnormally short identifiers are considered to be invalid identifiers and may be the result of coding, decoding, and transmission problems. Observations with the same short identifier may come from different aircraft. The proportion of abnormally short identifiers of AIREPs exceeds 8% for observations before 1987. For observations after January 1987, the proportion falls to 2%. Observations with invalid identifiers do not enter ACQC but enter the invalid observing value check.

Observations with valid identifiers enter the ACQC procedures. ACQC consists of a sequence of targeted algorithms, each of which is designed to look for a typical type of gross error. Gross errors most commonly occur as a result of equipment failure, data transmission problems, and data processing mistakes. Typical characteristics of such errors include unrealistic repetitions of values, invalid reports, anomalously temporal variations of values, internal or position inconsistencies of observations, and unrealistic flying speed derived from observations.

5.3 Blacklisting

Blacklisting of information is vitally important for improving analysis quality in assimilation system. It may be based on prior knowledge about instrument performance; for example, to remove observations observed by a malfunctioning temperature probe. A good approach to investigating the quality of observations from aircraft-specific data and determining which aircraft should be blacklisted is to compare the aircraft observation with a high-quality reference observation, such as radiosonde. However, this approach may yield too few samples because of the restricted locations and observing times of radiosonde data.

In general, quality feedback information generated through the assimilation system is used to estimate biases in observations. Here, the differences between aircraft observations and the ERA-Interim reanalysis (observation minus reanalysis, OMERA, \( \xi \)) are considered as “quality feedback information.” Forty years of quality feedback information can be easily obtained with this method. An assessment of the accuracy of aircraft wind, temperature, and humidity observations has been made by comparison with ERA-Interim from 1979 through 2018. To interpolate the reanalysis from the model to observation locations, the reanalysis was first interpolated horizontally to observation locations on model levels, and then vertical interpolation was conducted based on an assumption that model fields are linear in the logarithm of pressure between model levels.
5.3.1 Selection of report type

The bias and RMSE of temperatures and winds relative to ERA-Interim were calculated first for the separate report types: PIREP, AIREP, NUS-AMDAR, and ACARS (Fig. 9). Of these, PIREP has the poorest quality. The RMSE of temperatures of PIREP is larger than 4°C in the middle troposphere, and the RMSE of wind is larger than 7 m s⁻¹. Based on these statistical results, PIREP data were blacklisted and not assimilated in CRA.

5.3.2 Selection of humidity observation

In addition to analysis of the bias and RMSE of aircraft humidity, we also analyzed the bias and RMSE of radiosonde humidity relative to ERA-Interim across North America. A clear negative bias of ACARS humidity was observed before 2010, and the RMSE of ACARS humidity was greater than 1.5 g kg⁻¹, which was much higher than the RMSE of radiosonde humidity. Relative humidity observations from TAMDAR were closer to those from radiosonde before 2010. The systematic differences in ACARS and radiosonde humidity observations have decreased gradually since 2011. Thus, the specific humidity data from ACARS before 2011 were blacklisted.

5.3.3 Selection of temperature and wind observations

Based on the OMERA results, observations from those aircraft with the following three situations were blacklisted: 1) unacceptable magnitude of gross error; 2) bias greater than predetermined levels; and 3) RMSE greater than predetermined levels.

The following criteria were used as the limits for excess gross difference from the ERA-Interim reanalysis: 5°C for temperature and 10 m s⁻¹ for wind. For an aircraft, a variable with 25% gross errors is considered unacceptable in a month, and the variables observed by this aircraft in that month are blacklisted.

Then, the bias \( \bar{\xi}_k(z, m) \) of \( \bar{\xi} \) of the measurements (e.g., temperature \( T \) or zonal wind speed \( u \)) made by the \( k \) aircraft at level \( z \) at month \( m \) was calculated as follows:

\[
\bar{\xi}_k(z, m) = O_{k,i}(z, m) - ERA_{k,i}(z, m),
\]

\[
\bar{\xi}_k(z, m) = \frac{1}{n_k(z, m)} \sum_{i=1}^{n_k(z, m)} \bar{\xi}_{k,i}(z, m),
\]

Fig. 9. (a) Total observation numbers of PIREP (cyan), AIREP (black), NUS-AMDAR (blue), and ACARS (red). (b) Bias and RMSE of temperatures against ERA-Interim data of each report type. (c) Same as (b) but for U wind. (d) Same as (b) but for V wind. The statistical period is from 1981 to 2010.
where \( n_k(z, m) \) is the number of measurements taken by a specific aircraft \( k \) at level \( z \) at month \( m \); \( z = 1, 2, 3 \), representing low-level \((p > 700 \text{ hPa})\), middle-level \((700 \leq p \leq 300 \text{ hPa})\), and high-level \((p < 300 \text{ hPa})\) observations, respectively; \( O_{k,i}(z, m) \) is the \( i \)th measurement taken by a specific aircraft \( k \) at level \( z \) at month \( m \); and \( \text{ERA}_{k,i}(z, m) \) is the corresponding analysis value of ERA-Interim.

The RMSE of \( \xi \) of the measurements made by the \( k \) aircraft at level \( z \) at month \( m \) was calculated as follows:

\[
\Delta \xi_k(z, m) = \sqrt{\frac{\sum_{i=1}^{n_k(z,m)} \xi_{k,i}^2(z,m)}{n_k(z,m)}}. \tag{4}
\]

A total of 4763 aircraft were selected, which included 165 Chinese aircraft. These aircraft observed for at least three months in 2014 and conducted more than 60 observations per month. Each \( \bar{\xi}_k \) and \( \Delta \xi_k \) of these aircraft were calculated per month in 2014; 99.9th quantiles of the \( \bar{\xi}_k \) and \( \Delta \xi_k \) were then calculated and used as the predetermined criteria. Aircraft with \( \bar{\xi}_k \) or \( \Delta \xi_k \) greater than predetermined criteria were blacklisted. The monthly proportion of rejected temperature observations was about 1%–2% from 2001 to 2014. No more than 0.3% of wind observations were rejected in this way.

Based on the above algorithm, four Chinese airplanes were blacklisted in consecutive months since January 2011. To further confirm the credibility of the blacklists, we analyzed the systematic difference in aircraft and radiosonde temperatures, separating data from different airplanes. Here, the high-vertical-resolution radiosonde temperatures (sampling of 1 second or 1 hPa below 850 hPa) were used. To count as a collocation, the aircraft and radiosonde temperatures had to be within 150 km of each other, with a time separation of 1.5 h or less. Radiosonde temperatures were vertically interpolated to the observing level of the aircraft. Figure 10 shows the bias \( \bar{\xi}_k(z,m) \) and RMSE \( \Delta \xi_k(z,m) \) of temperatures from 165 Chinese aircraft during 2014. The temperatures observed by the blacklisted aircraft have warmer biases than other aircraft, and the RMSE exceeds 2°C.

### 6. Data usage in CRA

Various studies have noted that aircraft temperatures have a generally warm bias relative to radiosonde data at around 200 hPa (Cardinali et al., 2003; Ballish and Kumar, 2008; Zhu et al., 2015). For this reason, temperature measurements from aircraft were assimilated with VarBC in CRA. VarBC of aircraft temperatures is incorporated in the NCEP GSI analysis system. Details of the implementation of the VarBC scheme in CRA are presented in Zhu et al. (2015).

The confidence of CRA can be partly verified by a comparison with observations. To assess the performance of CRA, the bias and RMSE of observations relative to CRA’s analysis and background are used as metrics. Figure 11 shows the bias and RMSE for all assimilated aircraft temperature for the period of January 1994–December 2018. The RMSE reduces with time as the observing system improves. This is especially evident after 1997. This is likely because of the rapid increase of AMDAR and ACARS data availability, for which the quality is much better than AIREP. The bias of temperature relative to CRA analysis is spatially inhomogeneous in January 1988 (Fig. 12), indicating that the quality of AIREP observations is less stable. Aircraft temperatures in Africa have higher positive deviation than those in other areas in 1998 and 2008, which is probably related to the small number of observations. By
Fig. 11. (a) Bias and (b) RMSE relative to CRA analysis (blue) and background (red) above 300 hPa for global aircraft temperatures assimilated in CRA. Solid curves are 3-month running averages, and lightly shaded dots show hourly values at 0000, 0600, 1200, and 1800 UTC.

Fig. 12. Distributions of bias of aircraft temperature relative to CRA analysis on a 2° × 2° grid at pressure levels between 200 and 300 hPa during January (a) 1988, (b) 1998, (c) 2008, and (d) 2018. The minimum number of observations on a 2° × 2° grid is 20.
2018, the biases of temperatures at high altitude ($p < 300$ hPa) do not exceed $0.3^\circ$C in most areas.

Figure 13 shows the RMSE of wind data from aircraft. Beginning in 1998, the RMSE gradually decreases to around $2 \text{ m s}^{-1}$ at the middle level near the end of 2018, and $2.5 \text{ m s}^{-1}$ at the high level. This result indicates an improvement over time in the quality of CRA winds. Figure 14 shows the bias and RMSE of aircraft-observed specific humidity in the low and middle troposphere. The statistical result of humidity does not vary notably. However, a strong seasonal cycle of RMSE of humidity is evident (larger in summer and smaller in winter).

**Fig. 13.** As in Fig. 11b, but for zonal wind observations (a) between 300 and 700 hPa, and (b) above 300 hPa.

**Fig. 14.** As in Fig. 11, but for specific humidity observations at the middle levels (300–700 hPa).
7. Summary

This paper has highlighted the integration algorithm of observations from different sources, procedures for quality control, and blacklists of global aircraft observations for CRA. Each of these aspects has affected the quality of the reanalysis. The integration algorithm named “classified integration” combines aircraft observations from nine data sources into a new archive and eliminates duplicate data. The integration algorithm was applied to the development of IGMOAA, a new global aircraft dataset for the period 1973–2018. IGMOAA exhibits a higher level of completeness and integrity than existing data sources, especially in East Asia. The blacklist algorithm considers the differences between aircraft observations and the ERA-Interim reanalysis to provide quality feedback information. The credibility of blacklists is confirmed by careful comparison between aircraft observations and vertically high-resolution radiosonde observations. By using 40 years of quality feedback information, monthly blacklists from January 1979 to December 2018 were established and applied to exclude observations that were expected to have a negative impact on CRA.

This paper highlights the assimilation features of global aircraft observations in CRA from time series and spatial distribution. Overall changes in data usage for aircraft observations can be associated with improved quality of the observing system. These assimilation features inspire a certain level of confidence in the accuracy and stability of the observational record from aircraft.

In addition to the application to global atmospheric reanalysis, IGMOAA is suitable for a wide range of applications, such as comparisons with measurements from radiosonde and other observing systems, verification of output from model simulations, and improving understanding of the diurnal climatology of the lower atmosphere. IGMOAA is real-time updated by using the global AMDAR data from GTS at the CMADaaS platform since 2019.

Acknowledgments. The authors thank Tao Zhang and Hui Jiang for their assistance in collection of the historical aircraft observations, and the agencies for providing the source data for IGMOAA.

Appendix: Acronyms

| Acronym   | Description                                                   |
|-----------|---------------------------------------------------------------|
| ACARS     | Aircraft Communications Addressing and Reporting System       |
| ACQC      | Complex Quality Control for Aircraft observations             |
| AIREP     | aircraft report                                               |
| AMC       | Aviation Meteorological Center                                |
| AMDAR     | Aircraft Meteorological Data Relay                            |
| CEDA      | Centre for Environmental Data Analysis                        |
| CFSR      | Climate Forecast System Reanalysis                            |
| CMA       | China Meteorological Administration                           |
| CMADaaS   | CMA Data-as-a-Service                                        |
| CRA       | China’s first generation of a 40-yr global atmospheric reanalysis product |
| ECMWF     | European Centre for Medium-Range Weather Forecasts            |
| ERA-15    | The 15-yr ECMWF Re-Analysis                                   |
| ERA-Interim | ECMWF’s interim reanalysis                                      |
| ETAC      | Environmental Technical Applications Center                   |
| FGGE      | First GARP Global Experiment                                  |
| FNL       | Final Operational Global Analysis                             |
| GARP      | Global Atmospheric Research Program                            |
| GDAS      | Global Data Assimilation System                               |
| GSI       | Gridpoint Statistical Interpolation                           |
| GTS       | Global Telecommunication System                               |
| IGMOAA    | Integrated Global Meteorological Observation Archive from Aircraft |
| JMA       | Japan Meteorological Agency                                   |
| MARS      | Meteorological Archive and Retrieval System                  |
| MDCRS     | Meteorological Data Communications and Reporting System        |
| MERRA     | Modern-Era Retrospective Analysis for Research and Applications |
| NCAR      | National Center for Atmospheric Research                      |
| NCDC      | National Climatic Data Center                                 |
| NCEI      | National Centers for Environmental Information               |
| NCEP      | National Centers for Environmental Prediction                |
| NESDIS    | National Environmental Satellite, Data, and Information Service |
| NMC       | National Meteorological Center                                |
| NMIC      | National Meteorological Information Center                     |
| NUS-AMDAR | AMDAR data from outside the U.S.                              |
| NWP       | numerical weather prediction                                  |
| NWS       | National Weather Service                                      |
| OMA       | observation minus analysis of CRA                              |
| OMERA     | observation minus ERA-Interim reanalysis                      |
| PALT      | pressure altitude                                             |
| PIREP     | pilot report                                                  |
| POF       | phase of flight                                               |
| PS        | priority score                                                |
| QA        | quality assurance                                             |
| QC        | quality control                                               |
| RDA       | Research Data Archive                                         |
| RMSE      | root-mean-square error                                        |
| TAC       | Traditional Alphanumeric Codes                                |
| TAMDAR    | Tropospheric Airborne Meteorological Data Reporting            |
| TS        | threshold                                                     |
| TWERLE    | Tropical Wind, Energy Conversion, and Reference Level Experiment |
| U.S.      | United States                                                 |
| VarBC     | variational bias correction                                   |
| WIS       | WMO Information System                                        |
| WMO       | World Meteorological Organization                             |
Ballish, B. A., and V. K. Kumar, 2008: Systematic differences in aircraft and radiosonde temperatures: Implications for NWP and climate studies. Bull. Amer. Meteor. Soc., 89, 1689–1708, doi: 10.1175/2008BAMS2332.1.

Cardinali, C., L. Isaksen, and E. Andersson, 2003: Use and impact of automated aircraft data in a global 4DVAR data assimilation system. Mon. Wea. Rev., 131, 1865–1877, doi: 10.1175/2569.1.

Daniels, T. S., W. R. Moninger, and R. D. Mamrosh, 2006: Tropospheric Airborne Meteorological Data Reporting (TAMDAR) overview. 10th Symposium on Integrated Observing and Assimilation Systems for the Atmosphere, Oceans, and Land Surface, Amer. Meteor. Soc., Atlanta, GA, 9.1.

Dee, D. P., S. M. Uppala, A. J. Simmons, et al., 2011: The ERA-Interim reanalysis: configuration and performance of the data assimilation system. Quart. J. Roy. Meteor. Soc., 137, 553–597, doi: 10.1002/qj.828.

Fleming, R. J., 1996: The use of commercial aircraft as platforms for environmental measurements. Bull. Amer. Meteor. Soc., 77, 2229–2242, doi: 10.1175/1520-0477(1996)077<2229:TNFSAE>2.0.CO;2.

Gao, F., Z. Q. Liu, J. H. Ma, et al., 2019: Variational bias correction of TAMDAR temperature observations in the WRF data assimilation system. Mon. Wea. Rev., 147, 1927–1945, doi: 10.1175/MWR-D-18-0025.1.

Gibson, J. K., P. Källberg, S. Uppala, et al., 1999: ECMWF Re-Analysis Project Report Series: 1. ERA-15 Description (Version 2 - January 1999). Available online at https://www.ecmwf.int/sites/default/files/elibrary/1999/5984-era-description.pdf. Accessed on 25 December 2020.

Kalnay, E., M. Kanamitsu, R. Kistler, et al., 1996: The NCEP/NCAR 40-year reanalysis project. Bull. Amer. Meteor. Soc., 77, 437–472, doi: 10.1175/1520-0477(1996)077<0437:TPRARA>2.0.CO;2.

Keyser, D., 2018: PREPBUFR processing at NCEP. Available online at http://www.emc.ncep.noaa.gov/mmb/data_processing/prepbfur/doc/document.htm. Accessed on 25 December 2020.

Liu, Z. Q., C. X. Shi, Z. J. Zhou, et al., 2017: CMA global reanalysis (CRA-40): Status and plans. Proc. 5th International Conference on Reanalysis, 13–17 November 2017, Copernicus Climate Change Service, Rome, Italy, 16 pp.

Martin, R. C., M. M. Wolfson, and R. G. Hallowell, 1993: MD-CRS: Aircraft observations collection and uses. Fifth Int. Conf. on Aviation Weather Systems, Amer. Meteor. Soc., Vienna, VA, 317–321.

Moninger, W. R., R. D. Mamrosh, and P. M. Pauley, 2003: Automated meteorological reports from commercial aircraft. Bull. Amer. Meteor. Soc., 84, 203–216, doi: 10.1175/BAMS-84-2-203.

NCDC, 2002: National Climatic Data Center Data Document for Data Set 6380 (DSI-6380): Worldwide Aircraft Reports. NCDC, Asheville, USA, 14 pp.

Pauley, P. M., 2002: Telling north from south: An example of an error in automated aircraft data. Wea. Forecasting, 17, 334–336, doi: 10.1175/1520-0434(2002)017<0334:TNFSAE>2.0.CO;2.

Petersen, R. A., 2016: On the impact and benefits of AMDAR observations in operational forecasting—Part I: A review of the impact of automated aircraft wind and temperature reports. Bull. Amer. Meteor. Soc., 97, 585–602, doi: 10.1175/BAMS-D-14-00055.1.

Petersen, R. A., L. Cronce, R. Mamrosh, et al., 2016: On the impact and future benefits of AMDAR observations in operational forecasting: Part II: Water vapor observations. Bull. Amer. Meteor. Soc., 97, 2117–2133, doi: 10.1175/BAMS-D-14-00211.1.

Rahn, D. A., and C. J. Mitchell, 2016: Diurnal climatology of the boundary layer in southern California using AMDAR temperature and wind profiles. J. Appl. Meteor. Climatol., 55, 1123–1137, doi: 10.1175/JAMC-D-15-0234.1.

Rienecker, M. M., M. J. Suarez, R. Gelaro, et al., 2011: MERRA: NASA’s Modern-Era Retrospective Analysis for Research and Applications. J. Climate, 24, 3624–3648, doi: 10.1175/JCLI-D-11-00015.1.

Saha, S., S. Moorthi, H.-L. Pan, et al., 2010: The NCEP Climate Forecast System Reanalysis. Bull. Amer. Meteor. Soc., 91, 1015–1058, doi: 10.1175/2010BAMS3001.1.

Tavolato, C., and L. Isaksen, 2015: On the use of a Huber norm for observation quality control in the ECMWF 4D-Var. Quart. J. Roy. Meteor. Soc., 141, 1514–1527, doi: 10.1002/qj.2440.

WMO, 2003: Aircraft Meteorological Data Relay (AMDAR) Reference Manual. WMO, WMO-No. 958, Geneva, 80 pp. Available online at https://library.wmo.int/doc_num.php?explnum_id=4120. Accessed on 12 January 2021.

WMO, 2017: Guide to Aircraft-Based Observations. WMO, WMO-No. 1200, Geneva, 132 pp. Available online at https://library.wmo.int/doc_num.php?explnum_id=4120. Accessed on 12 January 2021.

Zhang, Y. J., L. Wang, J. A. Santanello, et al., 2020: Aircraft observed diurnal variations of the planetary boundary layer under heat waves. Atmos. Res., 235, 104801, doi: 10.1016/j.atmosres.2019.104801.

Zhu, Y. Q., J. C. Derber, R. J. Purser, et al., 2015: Variational correction of aircraft temperature bias in the NCEP’s GSI analysis system. Mon. Wea. Rev., 143, 3774–3803, doi: 10.1175/MWR-D-14-00235.1.