Estimation of TAMDAR Observational Error and Assimilation Experiments

FENG GAO
National Center for Atmospheric Research, Boulder, Colorado, and Nanjing University of Information Science and Technology, Nanjing, Jiangsu, China

XIAOYAN ZHANG
National Center for Atmospheric Research, Boulder, Colorado

NEIL A. JACOBS
AirDat LLC, Morrisville, North Carolina

XIANG-YU HUANG AND XIN ZHANG
National Center for Atmospheric Research, Boulder, Colorado

PETER P. CHILDS
AirDat LLC, Morrisville, North Carolina

(Manuscript received 8 October 2011, in final form 16 March 2012)

ABSTRACT

Tropospheric Airborne Meteorological Data Reporting (TAMDAR) observations are becoming a major data source for numerical weather prediction (NWP) because of the advantages of their high spatiotemporal resolution and humidity measurements. In this study, the estimation of TAMDAR observational errors, and the impacts of TAMDAR observations with new error statistics on short-term forecasts are presented. The observational errors are estimated by a three-way collocated statistical comparison. This method employs collocated meteorological reports from three data sources: TAMDAR, radiosondes, and the 6-h forecast from a Weather Research and Forecasting Model (WRF). The performance of TAMDAR observations with the new error statistics was then evaluated based on this model, and the WRF Data Assimilation (WRFDA) three-dimensional variational data assimilation (3DVAR) system. The analysis was conducted for both January and June of 2010. The experiments assimilate TAMDAR, as well as other conventional data with the exception of non-TAMDAR aircraft observations, every 6 h, and a 24-h forecast is produced. The standard deviation of the observational error of TAMDAR, which has relatively stable values regardless of season, is comparable to radiosondes for temperature, and slightly smaller than that of a radiosonde for relative humidity. The observational errors in wind direction significantly depend on wind speeds. In general, at low wind speeds, the error in TAMDAR is greater than that of radiosondes; however, the opposite is true for higher wind speeds. The impact of TAMDAR observations on both the 6- and 24-h WRF forecasts during the studied period is positive when using the default observational aircraft weather report (AIREP) error statistics. The new TAMDAR error statistics presented here bring additional improvement over the default error.

1. Introduction

Aircraft observations, which have been significantly increasing in volume over the past few years due to the expansion of aircraft-based observing systems, as well as the increase in commercial air travel, are becoming an important part in the global observing system (Benjamin et al. 1999, 2010). Operational numerical prediction centers have begun to ingest automated aircraft reports from the Aircraft Communications Addressing and Reporting System (ACARS), which is a digital datalink system for transmission of small messages between aircraft and
ground stations via VHF radio. This is the primary system employed by the Aircraft Meteorological Data Relay (AMDR) program of the World Meteorological Organization (WMO) into regional and global data assimilation systems (Schwartz and Benjamin 1995; Drue et al. 2008).

DiMego et al. (1992) reported forecast improvements from using aircraft data at the National Centers for Environmental Prediction (NCEP). Smith and Benjamin (1994) showed that ACARS reports improved short-range forecasts of upper-level winds and temperatures when added to wind profiler data over the central United States. However, the absence of humidity observations, as well as the high cruise heights, are two shortfalls of the current aircraft observation sets (Moninger et al. 2010). Other than scattered radiosonde soundings (RAOBs), there is a significant lack of routinely collected in situ observations, particularly humidity, from within the region below the tropopause, where the majority of moisture resides and where convective activity originates (Daniels et al. 2006).

To supplement existing technologies, a low-cost sensor called Tropospheric Airborne Meteorological Data Reporting (TAMDAR) was deployed by AirDat, under the sponsorship of a joint National Aeronautics and Space Administration (NASA) and Federal Aviation Administration (FAA) project as part of Aviation Safety and Security Program, according to requirements defined by the FAA, the Global Systems Division (GSD) of the National Oceanic and Atmospheric Administration (NOAA), and WMO. The TAMDAR sensor network has been providing a continuous stream of real-time observations on regional airlines since December 2004. Aircraft equipped with TAMDAR provide coverage over North America, including Alaska, Hawaii, and Mexico, and generate data from locations and times not available from any other observing system. TAMDAR produces thousands of high-frequency daily observations of humidity, icing, and turbulence, as well as conventional temperature, pressure, and winds aloft along with GPS-based coordinates in near–real time. Although TAMDAR will work on any airframe from a transoceanic 777 to a small unmanned aerial vehicle (UAV), commercial regional airlines have been the primary focus because those planes make more daily flights into a greater number of smaller airports, while still serving the major hubs. As a result, a larger number of soundings from a more geographically diverse set of airports are obtained. TAMDAR observations are rapidly becoming a major source of critical data utilized by various assimilation systems for the improvement of mesoscale NWP and the overall safety of aviation in the future (Fischer 2006).

A crucial step in the process of extracting maximal value from this new observation source is to correctly estimate the observational error of TAMDAR measurements. This will provide weighting information among different types of observations and background fields in the data assimilation system in order to obtain a statistically optimal estimated value of the true variables (Lorenc 1986; Benjamin et al. 1999; Barker et al. 2004).

Several previous investigations have addressed various methods for estimating observational error (e.g., Hollingsworth and Lonnberg 1986; Desroziers and Ivanov 2001; Desroziers et al. 2005). Typically, observational error includes instrument error, reporting error (i.e., measurement error), and representativeness error (Daley 1991; Schwartz and Benjamin 1995). Richner and Phillips (1981) used three ascents with two sondes on the same balloon to take simultaneous measurements of the same air mass. The results showed that the deviations between two sondes fell within the accuracies specified by the manufacturer with respect to the instrument error. As previously stated, this is not the only source of error in an observing system; therefore, when a comparison is made between two observations, or a single observation and forecast or model analysis, representativeness error must be taken into account.

Sullivan et al. (1993) updated temperature error statistics for NOAA-10 when the retrieval system in National Meteorological Center (NMC, since renamed the National Centers for Environmental Prediction, or NCEP) changed from a statistical to a physical algorithm. In their study, radiosonde soundings were considered the reference value, and the space–time proximity (i.e., 4 h and 300 km) between those soundings and satellite reports was the constraint. Their research provided an initial methodology regarding the estimation of observational error; however, by the inherent nature of the comparison, it was subject to greater representativeness error because of the large space–time collocation threshold.

In light of an ever-growing wealth of aircraft observations to be included as input for generating initial conditions in NWP, the quality of the data has been subject to several studies. Schwartz and Benjamin (1995) gave statistical characteristics of the difference between radiosonde observations (RAOBs) and ACARS data surrounding the Denver, Colorado, airport as a function of time and distance separation. A standard deviation of 0.97 K in temperature was reduced to 0.59 K through using a more strict collocation match constraint from 150 km and 90 min to 25 km and 15 min, which primarily arose based on the representativeness error decreasing. As a result, this study provided an upper bound on the combined error of ACARS and RAOB data with small representativeness error. Additionally, the authors speculated that the large direction difference was related to mesoscale variability, especially from turbulence in the boundary layer.
In a subsequent study to obtain the independent observation error of ACARS, Benjamin et al. (1999) reported on a collocation study of ACARS reports with different tail numbers to estimate observational error, assuming an equivalent expected error from each aircraft, and the minimization of the representativeness error by a strict match condition of 1.25 km and 2 min and no vertical separation. They reported a temperature root-mean-square error (RMSE) of 0.69–1.09 K and wind vector error of 1.6–2.5 m s\(^{-1}\) in a vertical distribution, which were comparable to RAOB data. The methodology and assumptions employed by Benjamin et al. (1999) are reasonable; however, since TAMDAR-equipped planes frequently fly into airports that do not have operational radiosonde launches, it is difficult to get enough data for a robust statistical comparison, while retaining similar strict collocation constraints.

More recently, Moninger et al. (2010) provide error characteristics of TAMDAR by comparing the data to the Rapid Update Cycle (RUC) 1- and 3-h forecasts with a model grid spacing of 20 km. In this study, the RUC 1-h forecast is not considered to be “truth” but is treated as a common denominator by which various aircraft fleets are compared. The results show the RMS difference of 1 K, 8%–20%, and RMS vector difference 4–6 m s\(^{-1}\) in temperature, RH, and wind observations, respectively. Since the Moninger et al. (2010) study, a procedure was implemented to correct for magnetic deviation errors that were degrading the wind observations (Jacobs et al. 2010), and the results reduced the RMS vector difference by 1.3 m s\(^{-1}\) for planes using this type of heading instrumentation. As with previous studies, the RMS difference reported by Moninger et al. (2010) includes the TAMDAR observation error, representativeness error, as well as the RUC 1-h forecast error.

When examining previous studies for the purpose of constructing a methodology for estimating observational error of TAMDAR, two issues become apparent.

a. Collocated observational error sources are typically combined in error statistics

Prior collocation studies consider radiosonde data as truth, and while the uncertainty can be addressed to a limited degree by dual-sensor sonde launches in field experiments (e.g., Schwartz and Benjamin 1995), or by systematic bias adjustments (e.g., Ballish and Kumar 2008; Miloshevich et al. 2009), automated evaluation of TAMDAR observations, as well as TAMDAR-related forecast impacts, for the purposes of deriving error statistics, inherently include radiosonde observational error (Moninger et al. 2010). Ballish and Kumar (2008) highlighted differences between radiosonde temperature and traditional AMDAR data, and suggested correcting temperature biases based on statistics derived against the model background. While this approach may mitigate systematic temperature biases, it is still subject to individual observing system uncertainty and does not address similar issues with wind and moisture observations.

b. The wind error is based on vector notation

There are two areas of interest with respect to the error associated with the assimilation of aircraft wind observations. First, in the Weather Research and Forecasting Model (WRF) data assimilation system (WRFDA; Huang 2009), both the instrumentation error file, as well as the error calculation, use a constant instrumentation error value of 3.6 m s\(^{-1}\) at all levels for aircraft weather reports (AIREPs), which includes TAMDAR. This is not the case for RAOBs, which vary the instrumentation error with height above 800 hPa for vector wind. Second, previous collocation studies that analyze wind error compare the vector wind (e.g., Benjamin et al. 1999; Schwartz et al. 2000; Moninger et al. 2010), which inherently includes both the observed speed error, as well as the directional error embedded in the vector component value. It should be noted that WRFDA and the gridpoint statistical interpolation (GSI) handle wind observations in a slightly different way. In this study, we only address WRFDA, where the observational error of wind speed is assigned as a single error value for both \(u\) and \(v\), and can only be defined by level.

The observed wind vector, \(\mathbf{V}_{W_{TAM}}\), is computed by taking the difference between the ground track vector (i.e., aircraft motion with respect to Earth), \(\mathbf{V}_G\), and the aircraft track vector, \(\mathbf{V}_A\):

\[
\mathbf{V}_{W_{TAM}} = \mathbf{V}_G - \mathbf{V}_A.
\]

The TAMDAR ground track vector is determined by a very accurate Garmin GPS system, and the associated error is at least two orders of magnitude less than the error in \(\mathbf{V}_A\). The aircraft track vector is calculated from the true airspeed (TAS) and the heading angle. The TAS is derived from the difference between the dynamic pressure of the pitot tube and the static pressure. The heading angle is determined by either a laser gyro, or magnetic flux valve, heading system.

This means that there are two primary sources for potential instrumentation error in the wind observation: the TAS and the heading angle. The pitot tube system, which determines TAS, and the heading system, which determines the angle, are essentially two unrelated observing systems. They both provide information to calculate the wind observation, which is then broken down into its \(u\) and \(v\) components. Once this step has occurred, it is not possible to determine if the error in \(u\) or \(v\) originated from
the TAS (i.e., speed) or the heading (i.e., direction). The uncertainty largely depends on the precision of the heading instrumentation, and older magnetic flux gate systems are subject to large deviations that can impact the observed wind accuracy (Moninger et al. 2010). Mulally and Anderson (2011) introduced a magnetic deviation bias filter that corrects this error, and results in wind observations that are comparable to those reported by aircraft with more sophisticated heading instrumentation.

Small errors in the TAS can result in larger observed wind errors due to the sensitivity of the equation to minor fluctuations in dynamic pressure as seen by the pitot tube. Additionally, the assumption is made that the aircraft is in perfect inertial alignment (i.e., no roll, pitch, or yaw), and since this is almost never the case, there is a small angle-based error introduced in addition to the TAS-based speed error (Painting 2003). The latter can be flagged and filtered based on acceptable thresholds as described in Moninger et al. (2003). Schwartz et al. (2000) explained that the RMS differences depend upon the observed wind speeds, and as wind speed increases, so does the RMS vector difference. They state that one of the reasons for this increase is that small directional differences can have a significant impact on vector components when the magnitude of the wind vector is large. Schwartz et al. (2000) showed this by stratifying the mean RMS vector differences by increments in the observed wind speed.

O’Carroll et al. (2008) developed a statistical method for calculating the standard deviation of the observational error for each of three different observation types under the assumption that the observational error of different systems is uncorrelated, which is typically a reasonable assumption in error theory. We revisit this problem with TAMDAR data, but for this study, we characterize the observational error in its original speed and direction notation. We hypothesize that it may be beneficial to define the error in terms of both direction and speed, as opposed to the magnitude of a single vector component, because of the inverse nature of the relationship between the angle and magnitude. In this study, we employ the O’Carroll et al. (2008) method of estimating the TAMDAR observational error in temperature, RH, wind speed, and wind direction. We further analyze the error in both wind speed, as well as direction, as a function of the speed itself. Additionally, we follow strict collocation match conditions as described in Benjamin et al. (1999) for the purpose of minimizing the representativeness error; however, even with this protocol, any time a numerical model representation of the atmosphere is constructed based on observations, an associated space–time representativeness error of that observation will be introduced.

The rest of this paper is laid out as follows. In section 2, the error statistics methodology and the data sources are introduced. Section 3 presents the error statistics results, which include the difference and vertical distribution of observational error. A brief description of the WRFDA system and WRF model configuration is presented in section 4. In section 5, we discuss assimilation sensitivity experiments that compare the default error to the new error statistics. Conclusions and plans for future work are described in section 6.

2. Methodology

a. Error analysis

Following the simultaneous equations for three-way collocation statistics given by O’Carroll et al. (2008), where the observation \( x \) is expressed as the sum of the true value of the variable, \( x_T \), the bias, or mean error, \( b \), and the random error \( e \), which has a mean of zero by definition, we begin with

\[
x = b + e + x_T.
\]

For a set of three collocated observation types \( i, j, \) and \( k \), the following set of equations are obtained:

\[
x_i = b_i + e_i + x_T
\]

\[
x_j = b_j + e_j + x_T
\]

\[
x_k = b_k + e_k + x_T.
\]

The variance of the difference, \( V \), between two observation types can be expressed as

\[
V_{ij} = \sigma_i^2 + \sigma_j^2 - 2r_{ij}\sigma_i\sigma_j
\]

\[
V_{jk} = \sigma_j^2 + \sigma_k^2 - 2r_{jk}\sigma_j\sigma_k
\]

\[
V_{ki} = \sigma_k^2 + \sigma_i^2 - 2r_{ki}\sigma_k\sigma_i
\]

(4)

where \( \sigma \) is the standard deviation, \( \sigma^2 \) is the variance, and \( r_{ij} \) is the correlation coefficient. Solving this set of equations allows the error variance of each observation type to be estimated from the statistics of the differences between the three types, which can be expressed as

\[
\sigma_i^2 = \frac{1}{2}(V_{ij} + V_{jk} - V_{ki})
\]

\[
\sigma_j^2 = \frac{1}{2}(V_{jk} + V_{ki} - V_{ij})
\]

\[
\sigma_k^2 = \frac{1}{2}(V_{ki} + V_{ij} - V_{jk}).
\]

(5)

A complete derivation of Eq. (5) is presented in the appendix.

If the representativeness error is taken into consideration, \( \sigma^2 \) from Eq. (5) can be expressed as
\[ \sigma_i^2 = \frac{1}{2}(V_{ij} - V_{jk} + V_{ki} - \sigma_{\text{rep}_{ij}}^2 + \sigma_{\text{rep}_{jk}}^2 - \sigma_{\text{rep}_{ki}}^2), \] (6)

where \( \sigma_{\text{rep}}^2 \) is the variance of the representativeness error, which is inherent when comparing any two observation types in different space–time locations. If the collocation match conditions between any type (e.g., \( k = fg \); WRF 6-h forecast) and two other types (e.g., \( i = \text{TAMDAR} \) and \( j = \text{RAOB} \)) are met, we can assume \( \sigma_{\text{rep}_{fg}}^2 = \sigma_{\text{rep}_{fg}}^2 \), where \( fg \) has a constant resolution based on the model grid spacing. Thus, in terms of the representativeness error of TAMDAR and RAOB data, Eq. (5) will give an upper bound of error calculated in Eq. (6). The assumption discussed above eliminates two terms in Eq. (6); however, the contribution from the fourth term, in this case \( \sigma_{\text{rep}_{TAMDAR-RAOB}}^2 \), may still be apparent, and is discussed below. Additionally, for this study, it is assumed that the short-term forecast error of a nonlinear numerical model is uncorrelated to the error of subsequent observations of any observing system.

b. Data

Based on the methodology discussed in the previous section, TAMDAR, RAOB, and WRF 6-h forecasts are selected as the three sources of data employed to estimate the TAMDAR observational error.

1) TAMDAR

The high-frequency TAMDAR observations, which number in the tens of thousands daily, are collected with multifunction in situ atmospheric sensors on aircraft (Daniels et al. 2004; Moninger et al. 2010). In addition to the conventional measurements of temperature and winds aloft, the observations contain measurements of humidity, pressure, icing, and turbulence, along with GPS-derived coordinates.

For humidity, the fundamental physical parameter that the TAMDAR capacitive sensor technology responds to is the density of \( \text{H}_2\text{O} \) molecules. The sensor uses a polymer material that either absorbs or desorbs water molecules based on the RH with respect to water. This in turn affects the capacitance; the relationship is monotonic, so in principle, a given capacitance, which can be measured and turned into a voltage, represents a certain RH.

The observations are relayed via satellite in real time to a ground-based network operations center where they are received, processed, quality controlled, and available for distribution or model assimilation in less than a minute from the sampling time. These observations are reported at 10-hPa pressure intervals up to 200 hPa, with the largest time-based interval during cruise being no more than 3 min.

From 2005 through 2011, NOAA/GSD played a central role in the distribution, evaluation, and initial quality control (i.e., reports formatting or units error) of TAMDAR data. In this study, TAMDAR observations are collected via the Meteorological Assimilation Data Ingest System (MADIS) dataset from NOAA/GSD. The TAMDAR observations used in this study came from fleets that covered most of the airports in the east-central and northwest continental United States (CONUS; Fig. 1). The dataset in this study uses a winter month (January) and a summer month (1–25 June) in 2010. The time series of TAMDAR observation counts are displayed in Fig. 2. The wind observations are fewer than those for RH, which is less than for the temperature. This is because the wind observation requires an accurate aircraft heading reading, so whenever the plane is banking or rolling in a turn over a frame-specific threshold, the wind data are flagged. On occasion, RH data will also be flagged, which typically happens during brief icing events; these data were withheld based on the current quality control flagging system.

2) RADIOSONDE

The radiosonde observations are transmitted to a receiving station where the height of the package is sequentially computed in incremental layers at each reporting level using the hypsometric equation. The drift speed and direction at various levels are determined from the ground-based radio antenna that tracks the instrument package as it is carried by the wind during the ascent. These observations are processed, tabulated, and encoded for transmission over various communication networks. The National Weather Service launches radiosondes from 92
stations in North America and the Pacific Islands twice daily. Nearly all routine launches occur approximately 45 min before the official observation times of 0000 and 1200 UTC to allow time for a relaunch should there be a mechanical failure. Therefore, the only observations that are truly sampled at these synoptic hours are in the middle of the profile. In this study, both the TAMDAR and RAOB data were selected over the same time windows from the National Center for Atmospheric Research (NCAR) database, which is routinely transmitted over the Global Telecommunication System (GTS).

The comparison described below was tested with and without a drift applied to the radiosonde using ascent rate statistics from Seidel et al. (2011). For the highest level one would expect to see in TAMDAR observations, the average drift was 16 km in June and 26 km in January with a standard deviation of roughly 12 km. Below 450 hPa, which would include the majority of the TAMDAR observations, the mean drift results were approximately 7 and 10 km for June and January, respectively. In general, the number of matched pairs remained unchanged when considering drift for these cases. Presently, WRFDA does not account for radiosonde drift during the assimilation. One would expect accounting for drift to still improve the accuracy slightly at lower levels; however, when considering the mean statistics over a 1-month period, it was not significant enough to alter the error values to the present decimal place used by WRFDA.

3) WRF FORECAST

When using tight constraints for the collocation, there are no other real observations besides RAOB and TAMDAR that share this small window of three-dimensional space over the same time. Thus, the third data source employed for this three-way collocation methodology during January and June of 2010 is the 6-h forecast from the Advanced Research core of the WRF (WRF-ARW; Skamarock et al. 2008). The configuration of the WRF 6-h forecast is discussed in section 4.

4) DATA COLLOCATION

According to Eq. (5), the standard deviation of the TAMDAR observational error ($\sigma_T$) can be expressed as

$$\sigma_T = \left[ \frac{1}{2} (V_{T-R} + V_{R-F} + V_{F-T}) \right]^{1/2},$$  \hspace{1cm} (7)

where $T$, $R$, and $F$ stand for TAMDAR, RAOB, and WRF 6-h forecast, respectively, and $V$ is the variance of the difference between two observation types. Therefore, data pairs between any two observation types must be calculated first to determine the value of $V$ in Eq. (7).

To compare with both TAMDAR and WRF, the RAOB data are interpolated to every 5 hPa. The data pairs between TAMDAR and RAOB were created for every TAMDAR observation that occurred within

![Fig. 2. Time series of the observation count of the experiment period for temperature, RH, and wind from TAMDAR during (top) January and (bottom) June 2010.](image-url)
a certain space–time interval of a RAOB report. This means that both data types are assumed to represent the mean value of a small volume of air within a certain spatial and temporal range. To assure confidence in the quality of data pairs, as well as to decrease the effect of representativeness error, different collocation match conditions are applied on different levels for the purpose of maintaining statistically significant matched-pair volume counts. A maximum temporal (vertical spatial) separation difference of 1 h (25 m) was applied in conjunction with three horizontal spatial separation limits of 10 km below 775 hPa, 20 km between 775 and 450 hPa, and 30 km above 450 hPa. The main reason why the collocation matches become fewer with height is that the number of TAMDAR measurements decreases with height because not all flights achieve maximum cruise altitude.

When applying these constraints over the study window, 23,551 matched data pairs were obtained. Despite having less strict horizontal collocation criteria for upper levels (i.e., 30 km), most of the data pairs still fell into the <10-km bin, similar to the lower levels. Figure 3 depicts the distribution of these pairs by distance and time separation. Approximately 70%, 71%, and 61% of TAMDAR–RAOB matched pairs for temperature, RH, and wind, respectively, have a spatial separation of less than 10 km.

As discussed for Fig. 2, more wind data (RH data) were withheld because of aircraft maneuvering (icing). This is also evident in Fig. 3a, where for a collocation threshold below 10 km, despite more net matched pairs, there are fewer matched wind observations. This is because most RAOB sites are located at airports, and during the final approach, aircraft tend to make more banking maneuvers, which result in flagged wind data on descents near airports (i.e., near RAOB launch sites). There is also a notable increase in matched pairs in Fig. 3b in the 60 min prior to 0000 and 1200 UTC, and fewer after, which is a function of RAOB launch time and ascent rate, and the flight schedules, which are routed to avoid passing too close to the ascending balloon during the cruise phase of flight. The data pairs between the WRF 6-h forecast and both RAOB and TAMDAR are easily obtained by gridpoint interpolation to the observation location.

Figure 4 presents the RMSE difference of the data pairs based on the spatial and temporal separation. Generally, the data-pair difference for temperature (Fig. 4a) and RH (Fig. 4b) steadily increases, as the space–time separation grows, which is expected according to the representativeness error. The RH values above 400 hPa do not appear to follow this trend. This is likely a result of the numerically small mixing ratio values; however, based on the low correlation, any meaningful trend would be hard to identify. The order of the lines based on RMSE difference is consistent with temperature becoming more homogeneous with height, and larger RH RMSE resulting from very dry air at higher levels.

Conversely, the wind vectors can be significantly affected by synoptic flow patterns. The wind speed (Fig. 4c) follows the same trend as the previously mentioned scalars with the exception of the lowest level, which still increases, but not at the same rate. This may be a result of the more chaotic flow observed closer to the surface combined with smaller speed magnitudes.

In Fig. 4d, there is greater directional error closer to the surface, and the magnitude of the directional error decreases with height. As we will discuss below, higher speeds typically have less directional error associated with them despite having larger speed error. Thus, it is somewhat expected to see the direction error from the three layers stratified in opposite order of the speed error in Fig. 4c.

The small decrease in error for the middle layer may be related to the amount of RAOB drift. The top of this layer is high enough to be slightly affected by drift, but still low enough that it lacks the homogeneity of the upper-most layer, which can offset this variability. Since we are using the launch position of the radiosonde site for the spatial coordinate, rapidly drifting sondes should have a lower wind direction error. Another possible cause may occur when large spatial collocation match conditions are employed. In
this case, similar wind directions may be obtained; however, the actual observations are a full wavelength apart. Either situation can greatly increase the variance of the representativeness error term \( \sigma^2_{\text{rep-TAMDAR-RAOB}} \) discussed above, and as a result, we employed strict collocation match conditions, described in the beginning of this section, to minimize the impact of this term.

3. Error estimation analysis

a. Difference distribution

Figures 5 and 6 present the collocated matched-pair Gaussian-like distribution patterns between TAMDAR and RAOBs, and TAMDAR and WRF 6-h forecasts, respectively. The dotted line is the Gaussian distribution according to the mean value and standard deviation of the difference. Observation pairs with a difference more than 7 K, 40%, 10 m s\(^{-1}\), and 60° in the temperature, RH, speed, and direction variables, respectively, are rejected, since abnormal differences are often a result of representativeness error or mesoscale perturbations. In both Figs. 5 and 6, the matched-pair counts approach zero well before the error threshold limits are encountered; therefore, only clear outliers would fall beyond this limit and, as such, would be caught by the upstream in-line quality control procedure. The statistics of the differences are shown in Table 1, which includes bias \((B)\), standard deviation \((\sigma)\), and confidence intervals of 68%, 90%, and 99% (i.e., 1, 1.64, and 2.58\(\sigma\)) of the normal Gaussian distribution.

In general, TAMDAR data quality is comparable to RAOB data based on the biases \((-0.1s)\) with the exception of direction (Table 1). One standard deviation away from the mean, the differences between all the metrics are more than 68.3%, which means that the collocated matched pairs exceeded the normal distribution threshold while maintaining comparable quality. Both difference distributions also follow a similar trend at the 1.64\(\sigma\) confidence interval, and have a percent area of around 90%, which shows that the three observation types have a reasonable error range.

It is expected that mesoscale perturbations will lead to some observation pairs with abnormal differences, which lie outside 2.58\(\sigma\); however, the collocated data pairs with large differences resulting from either coarse match conditions, or the occasional bad observation that slipped
through the quality control filter, are rare enough so as not to significantly affect the error estimation of the observations. In this respect, variables such as RH, wind speed, and direction should be closely monitored by quality control and gross error check procedures during operational preprocessing. The total $\sigma$ of observed RH is typically smaller in the critical lower levels because the error in the upper levels increases remarkably based on the very low water vapor values. An inherent characteristic of the wind observations is that the direction error feeds back into the wind vector observations. As a result, mesoscale variability and fluctuations in accurate heading information can cause larger directional error at lower wind speeds, which can increase the frequency of bad observations.
b. Vertical distribution of observational error

1) TEMPERATURE

The standard deviation of the temperature errors of TAMDAR and RAOB data in January and June are shown in Fig. 7. The temperature error for TAMDAR is 0.6–0.9 K depending on the level. It is comparable to RAOB data below 700 hPa for the June period and slightly smaller than RAOB data for the January period. The maximum difference of about 0.15 K occurs around 500 hPa in June, and around 700 hPa in January. Based on the vertical distribution described in the following section, higher impacts between 850 and 700 hPa are expected. The difference seen at 500 hPa in June may be related to the increased frequency of small-scale perturbations from the convective mixed layer during this time of year.

The instrument error, as specified by the manufacturers of both TAMDAR and the radiosondes, is 0.5 K, and anything above this is likely representativeness error, which must always be taken into account when comparing two data sources within an assimilation system.

2) WIND

Wind error can be affected by multiple factors such as mesoscale perturbations, aircraft heading information, and even as a function of the wind speed itself. According to the wind speed observation statistics in Fig. 8, the RAOB error is about 0.5–1.0 m s\(^{-1}\) less than for TAMDAR; however, the average error of 2.0–2.5 m s\(^{-1}\) is still quite small. It is also interesting to note that the results shown in Fig. 8 for RAOBs appear to validate the default RAOB error in WRFDA. The wind direction error can be seen in Fig. 9. Both RAOB and TAMDAR have a significant decrease in error with height, and the magnitude of the error in the lower levels is quite large. Three possible reasons for these large error statistics are presented:

- The TAMDAR wind observations are dependent on the accuracy of the heading information supplied by the aircraft instrumentation, and even the most accurate avionics will still introduce an additional source of error. This error source would have more impact on lower wind speeds.
- The error of the wind speed and direction changing as a function of wind speed is a typical characteristic of wind observations.
- The frequency of TAMDAR is weighted toward the 1800 and 0000 UTC cycles, whereas the RAOB data have equal numbers for both 0000 and 1200 UTC. An increase in diurnal instability tends to be more prevalent in the 0000 UTC cycle, compared to 1200 UTC, especially during the summer months because of the heating and length of day.

Figure 10 further illustrates this error dependence on wind speed. Both the wind direction error and the wind speed error for the TAMDAR and RAOB data are plotted based on bins of wind speed magnitude. In general, wind speed (direction) error increases (decreases) with wind speed. RAOBs have less error than TAMDAR for speeds below 15 m s\(^{-1}\), while the opposite is true for speeds above 15 m s\(^{-1}\), but the differences typically remain less than 0.5 m s\(^{-1}\).

Recent improvements in correcting the heading bias seen on some aircraft are discussed in Jacobs et al. (2010) and Mulally and Anderson (2011). The error is reduced by approximately 2 kt, or 1 m s\(^{-1}\), which is roughly a 45% decrease in error for this subset of magnetic-heading-equipped planes, and brings the quality of the data in line with the remainder of the fleet. It should also be mentioned that this subset of planes makes up a small fraction (i.e., 15 out of more than 250) of the expanding TAMDAR fleet, which typically rely on more sophisticated avionics.

3) RELATIVE HUMIDITY

The abundant RH observations of TAMDAR should provide a substantial supplement to present observation types. The quality of TAMDAR RH observations can be seen in Fig. 11, where during the month of January the error of 7%–9% was similar to the RAOB result. In June, the TAMDAR error ranged from 6% near the

| TABLE 1. The statistics of the differences between TAMDAR and RAOB, and TAMDAR and WRF 6-h forecasts. |
|---------------------------------|---------------------------------|---------------------------------|---------------------------------|
| RH   | T       | Speed | Direction | RH       | T       | Speed | Direction |
| N    | 9994    | 12.549 | 5384      | 4810     | 284586  | 374198 | 149044    | 143572   |
| B    | -1.57 (%) | 0.04 (K) | -0.004 (m s\(^{-1}\)) | -1.22 (%) | -0.37 (%) | -0.16 (K) | 0.15 (m s\(^{-1}\)) | -1.39 (%) |
| \(\sigma\) | 12.18 (%) | 1.06 (K) | 2.83 (m s\(^{-1}\)) | 20.13 (%) | 13.81 (%) | 1.16 (K) | 2.88 (m s\(^{-1}\)) | 16.83 (%) |
| 1.6\(\sigma\) | 72.10% | 74.90% | 71.50% | 73.60% | 69.80% | 70.10% | 70.10% | 74.30% |
| 2.58\(\sigma\) | 97.80% | 97.50% | 97.50% | 97.80% | 98.30% | 98.30% | 98.50% | 97.10% |
FIG. 7. The standard deviation of temperature error for TAMDAR and RAOB in (a) January and (b) June.

FIG. 8. As in Fig. 7, but for wind speed error.
surface to 8% above 400 hPa, and had a maximum reduction in error of 3% RH compared to RAOB at 500 hPa. The estimated error range is consistent with the findings of 5%–10% from Daniels et al. (2006).

4. Model configuration and experiment design

To evaluate the performance of TAMDAR with the observational error estimated above, three parallel experiments are conducted. These experiments were performed during June 2010, and are based on WRFDA three-dimensional variational data assimilation (3DVAR) and the WRF ARW (version 3.2).

In the present WRFDA system, the AIREP error is the default table for any type of observation collected by an aircraft regardless of airframe, phase of flight, or instrumentation type. Since TAMDAR is based on different operating principles compared to traditional AIREPs, it is not realistic to assume it possesses similar error characteristics. The TAMDAR error statistics derived above will be used in the following assimilation experiment as a substitute for the default AIREP error. Both the original default AIREP error, and the new
TAMDAR-specific error, are presented in Table 2. The WRFDA error file contains values that correspond to levels up to 10 hPa. Since TAMDAR data would never appear above 200 hPa, any value entered in the table above this height would merely serve as a placeholder.

In section 3, both the wind speed error and the wind direction error were characterized. However, in this part, only the wind speed error (Fig. 8) was employed to modify the WRFDA observational error statistics in Table 2. The results from Figs. 9 and 10 help shed light on some of the error-related trends, but those results were not applied to the error table for the NEWerr_T run, as this would require a modification to the observation operator discussed in the following section.

a. WRF ARW and WRFDA 3DVAR

The WRF-ARW is a fully compressible and Euler nonhydrostatic model with a vertical coordinate of terrain-following hydrostatic pressure, and features time-split integration using a third-order Runge–Kutta scheme with a smaller time step for acoustic and gravity wave modes and multiple dynamical cores with high-order numerics to improve accuracy. A detailed description of the model can found in Skamarock et al. (2008).

The WRFDA 3DVAR system originates from the 3DVAR system in the fifth-generation Pennsylvania State University–NCAR Mesoscale Model (MM5) developed by Barker et al. (2004), and based on an incremental formulation (Courtier et al. 1994). Following Lorenc et al. (2000), the control variables are the streamfunction, unbalanced velocity potential, unbalanced temperature, unbalanced surface pressure, and pseudo RH, which are used in the minimization process of the first term of Eq. (8) below.

The basic goal of 3DVAR is to obtain statistically optimal estimated values of the true atmospheric state at a desired analysis time through an iterative minimization of the prescribed cost function (Ide et al. 1997):

\[
J(x) = \frac{1}{2}(x - x_b)^T B^{-1}(x - x_b)
+ \frac{1}{2}[y_0 - H(x)]^T R^{-1}[y_0 - H(x)],
\]

(8)

where \(B\) is the background error covariance matrix, \(x_b\) is the background state, \(H\) is the nonlinear observation operator, \(y\) is the data vector, \(R\) is the observation error covariance matrix, and the state vector is defined as

\[
x = [u^T, v^T, T^T, q^T, p_s^T]^T.
\]

(9)

If an iterative solution can be found from Eq. (9) that minimizes Eq. (8), the result represents a minimum
variance estimate of the true atmospheric state given the background $x_b$ and observations $y_0$, as well as $B$ and $R$ (Lorenc 1986). The conjugate gradient method is used to minimize the incremental cost function. A detailed description of this system can be found in Barker et al. (2004), as well as at the WRFDA web site (http://www.mmm.ucar.edu/wrf/users/wrfda/pub-doc.html). Additional background on TAMDAR assimilation by WRFDA can also be found in Wang et al. (2009).

b. Model configuration

The data assimilation experiments and WRF forecasts were performed on a single 400 × 250 grid with 20-km spacing that covered the United States and surrounding oceanic regions. There were 35 vertical levels with a top of 50 hPa. The model produced a 24-h forecast. While this configuration is not necessarily optimal for assimilation of high-resolution asynoptic data like TAMDAR, it was sufficient to reach conclusions and illustrate the necessity to consider other methods of determining wind observation error statistics.

In this study, the Kain–Fritsch cumulus parameterization was employed (Kain 2004), along with the Goddard cloud microphysics scheme, and the Yonsei University (YSU) planetary boundary layer parameterization (Hong et al. 2006).

c. Experiment design

Three parallel WRF runs were performed during 1–20 June 2010:

- ‘DEFerr_noT’ is used as a control run, conventionally assimilating GTS data with default error statistics in WRFDA, including surface synoptic observations (SYNOP), aviation routine weather reports (METARs), PROFILER, RAOB, ground-based GPS precipitable water, SHIP, and BUOY, but excluding all non-TAMDAR automated aircraft data (e.g., the Meteorological Data Collection and Reporting System, MDCRS);
- ‘DEFerr_T’ is identical to ‘DEFerr_noT’ in every way except that it also assimilates TAMDAR wind, temperature, and RH observations; in this run, the default AIREP error is applied to the TAMDAR observations; and
- ‘NEWerr_T’ is identical to ‘DEFerr_T’ in every way except it uses the TAMDAR error statistics introduced in section 3 and Table 2 instead of the default AIREP error.

The cycling process employed here begins with a 6-h WRF forecast based on the GFS at the four analysis times (i.e., 0000, 0600, 1200, and 1800 UTC) for initial background and lateral boundary conditions (LBCs), after which, the 6-h WRF forecasts are used as the background or first guess for all three runs in the experiment. The LBCs are updated on every run by the latest GFS. All three assimilation versions produce eighty 24-h forecasts over the 20-day period during June 2010. The window of time for the 3D-Var data assimilation process is 2 h on either side of the analysis time; thus, TAMDAR data from 0300, 0900, 1500, and 2100 UTC were not included. Ideally, a more rapid cycling 3DVAR (or 4DVAR) assimilation process would be able to extract greater value from the asynoptic observations, but the objective of this study was only to derive and test the error statistics.

Before assimilation, a basic quality control procedure, which includes a vertical consistency check (superadiabatic

<table>
<thead>
<tr>
<th>Level (RH,T)</th>
<th>New RH (%)</th>
<th>New T (K)</th>
<th>Default RH (%)</th>
<th>Default T (K)</th>
<th>Level (wind)</th>
<th>New wind (m s⁻¹)</th>
<th>Default wind (m s⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td>200</td>
<td>8.00</td>
<td>0.8</td>
<td>10.00</td>
<td>1.0</td>
<td>200</td>
<td>2.7</td>
<td>3.6</td>
</tr>
<tr>
<td>250</td>
<td>8.00</td>
<td>0.8</td>
<td>10.00</td>
<td>1.0</td>
<td>250</td>
<td>2.7</td>
<td>3.6</td>
</tr>
<tr>
<td>300</td>
<td>8.00</td>
<td>0.8</td>
<td>10.00</td>
<td>1.0</td>
<td>300</td>
<td>2.7</td>
<td>3.6</td>
</tr>
<tr>
<td>400</td>
<td>7.50</td>
<td>0.8</td>
<td>10.00</td>
<td>1.0</td>
<td>400</td>
<td>2.5</td>
<td>3.6</td>
</tr>
<tr>
<td>500</td>
<td>7.00</td>
<td>0.9</td>
<td>10.00</td>
<td>1.0</td>
<td>500</td>
<td>2.5</td>
<td>3.6</td>
</tr>
<tr>
<td>700</td>
<td>7.00</td>
<td>0.9</td>
<td>10.00</td>
<td>1.0</td>
<td>700</td>
<td>2.5</td>
<td>3.6</td>
</tr>
<tr>
<td>850</td>
<td>6.50</td>
<td>0.7</td>
<td>10.00</td>
<td>1.0</td>
<td>850</td>
<td>2.5</td>
<td>3.6</td>
</tr>
<tr>
<td>1000</td>
<td>5.50</td>
<td>0.7</td>
<td>15.00</td>
<td>1.0</td>
<td>1000</td>
<td>2.0</td>
<td>3.6</td>
</tr>
</tbody>
</table>

AUGUST 2012 GAO ET AL.

869
check and wind shear check), and dry convective adjustment, is performed on all observations including TAMDAR. This is following the initial quality assurance protocol performed by AirDat on the TAMDAR data when it is sampled (Anderson 2006). Additionally, observations that differ from the background by more than 5 times the observational error are also rejected.

The NMC method (Parrish and Derber 1992) was applied over a period of 1 month prior to each study window to generate the background error covariances using monthly statistics of differences between WRF 24- and 12-h daily forecasts. The verifications presented here are from both the 6- and 24-h forecasts, and are based on the average results of the 80 forecast cycles.

5. New TAMDAR error statistics in WRFDA

a. Previous studies

Several studies have been conducted on the TAMDAR dataset, which underwent a lengthy operational evaluation during the Great Lakes Fleet Experiment (GLFE; e.g., Mamrosh et al. 2006; Jacobs et al. 2008; Moninger et al. 2010). Mamrosh et al. (2006) found that TAMDAR, with high spatial and temporal resolution, was valuable when used in marine (lake breeze) forecasting, convective forecasting, and aviation forecasting. Moninger et al. (2010) conducted a detailed analysis of the additional forecast skill provided by TAMDAR to the Rapid Update Cycle (RUC) model over a period of more than three years. The estimated temperature, wind, and RH 3-h forecast errors in RUC were reduced by up to 0.4 K, 0.25 m s\(^{-1}\), and 3\%, respectively, by assimilating TAMDAR, which corresponds to an estimated maximum potential improvement in RUC of 35\%, 15\%, and 50\% for temperature, winds, and RH, respectively. These experiments were conducted during the initial phase of installation of the sensors, so the improvements seen as time progresses are a function of additional sensors being deployed in the field.

b. Observations assimilated

Figure 12 presents the time series and vertical profile of RAOB and TAMDAR temperature observation counts used in assimilation experiments. RAOBs are routinely launched twice daily at 0000 and 1200 UTC in
the CONUS, which leaves a void of upper-air observational data to initialize the 0600 and 1800 UTC forecast cycles. Supplementing both the spatial gaps, as well as the temporal gaps, is critical for improving forecast skill. It is evident in Fig. 12a that TAMDAR not only is coincident with the 0000 and 1200 UTC cycles, but also peaks in data volume during the 1800 UTC cycle. The daily mean observation count for TAMDAR over the study period for the 0600 and 1800 UTC cycles, respectively, is 798 and 7402 for temperature, 756 and 6409 for RH, and 230 and 2145 for wind. Figure 12b presents the observation count per mandatory level for the duration of the study period. It is hypothesized that these extra observations will make a statistically significant contribution to forecast skill.

c. Results analysis

The TAMDAR error in the DEFerr_T run is based on the default AIREP error statistics (i.e., generic aircraft data) in the obsproc program within WRFDA. The default errors, especially the wind observational error, which is a constant 3.6 m s\(^{-1}\), are larger than the new error statistics. In this situation, TAMDAR will receive unreasonably small weight among the other observation types and background. The new error statistics derived here will correct this issue.

Because we wanted to highlight the difference in TAMDAR impact using the default (DEFerr_T) versus new errors (NEWerr_T), we withheld MDCRS data, which would be assimilated as a default AIREP in WRFDA. If the MDCRS data were included as a default AIREP, we would expect the relative impact of TAMDAR, especially for wind and temperature, to be noticeably reduced.

To evaluate the impact of the new error statistics, the root-mean-square errors (RMSEs) of the 6- and 24-h forecast are calculated using mainly radiosondes, as well as various near-surface observation types (e.g., SYNOP and METAR) available between the surface and 200 hPa as verification over the entire CONUS domain. Since the verification package referenced all of the conventional observations distributed by NCEP over the GTS, a small fraction of the TAMDAR feed (\(\sim 2.7\%\)) was also present in this database. The percentage of improvement is calculated by \(\%\text{IMP} = (\alpha - \beta)/\beta \times 100\), where simulation \(\alpha\) is compared to simulation \(\beta\), and appears contextually below as an improvement of \(\alpha\) over \(\beta\) (Brooks and Doswell 1996).

1) TEMPERATURE

The average impact of the TAMDAR temperature observations on the 6- and 24-h WRF forecasts are presented in Fig. 13. In general, temperature error decreases with elevation through the troposphere, and the inaccuracy of forecasted temperatures associated with errors within the planetary boundary layer are greatest between 1000 and 900 hPa for a 24-h forecast. In Fig. 13b, the vertical profile of 24-h forecast error has a maximum around 925 hPa, which is consistent with the results from Moninger et al. (2010) discussed in section 1.

Both the DEFerr_T and NEWerr_T results have noticeably less error when compared to DEFerr_noT, particularly below 500 hPa, which is not surprising given the vertical distribution of the TAMDAR observations (cf. Figs. 12b and 13). The maximum difference in the 6-h forecast RMSE between DEFerr_noT and DEFerr_T is 0.15 K at 925 hPa, and 0.09 K at the same level for the 24-h forecast. The NEWerr_T simulation further reduces the error, and this reduction is attributed to the new error statistics. The profile pattern is similar, but for the 6-h forecast RMSE the maximum difference between NEWerr_T and DEFerr_noT is 0.27 K, which is a 44%
improvement over the DEFerr_T run. For the 24-h forecast, the maximum difference between the NEWerr_T and DEFerr_noT is 0.2 K, which is roughly a 55% improvement over the DEFerr_T run at the same level.

2) WIND

Despite having a wind observation error larger than a radiosonde, the TAMDAR wind data still produce a notable improvement in forecast skill. In Fig. 14, there are two interesting comparisons that are observed. First, both the DEFerr_T and NEWerr_T simulations produce a similar decrease in RMSE around 400 hPa. This is more easily seen in Fig. 14a for the 6-h forecast, and is a combination of a larger volume of data compared to 300 and 200 hPa and similar error statistics.

However, below 900 hPa, the $U$ and $V$ wind RMSEs of the DEFerr_T run do not improve much, if any, over the DEFerr_noT run. This is largely a result of the error statistics in the default error file, which employ a constant value of 3.6 m s$^{-1}$ for every level. While this is acceptable for levels around 400 hPa and higher, it is likely too large for levels below 800 hPa, especially when dealing with observations of lower wind speed or aircraft maneuvering on the final approach of the descent. As a result, fewer observations are rejected at lower levels. Adjustments made to the error statistics in the NEWerr_T run mitigate this problem, and as a result, the $U$ and $V$ wind forecast RMSEs are reduced. This serves as an alternative to withholding all of the descent wind observations. A maximum reduction in error of roughly 0.25 m s$^{-1}$ is observed around 850 hPa.

3) HUMIDITY

The positive impacts of the TAMDAR RH observations can be seen in Fig. 15. For the 6-h forecast, the magnitude of the impact below 500 hPa is approximately 0.1 g kg$^{-1}$; however, above 500 hPa, there is the improvement is much less clear, which converges to undetectable above 200 hPa. This is partly because the volume of TAMDAR decreases with height, but it is primarily a function of the water vapor magnitudes approaching very small numerical values, as height increases above 500 hPa. A similar trend is seen for the 24-h
The impacts of the new error statistics are positive but small with the exception of those at the 850-hPa level. Although the TAMDAR RH error in the NEWerr_T run was changed, the observations impacted by the change based on RAOB data and the model background remained similar, so that in the window examined here, the positive impact from the new error statistics is relatively small. This is likely a function of atmospheric variability, as well as dynamic events and seasonal fluctuations, which may produce larger differences.

6. Summary and outlook

The TAMDAR sensor network provides abundant meteorological data with high spatial and temporal resolution over most of North America. The large diversity of regional airport coverage, real-time reporting, adjustable vertical resolution, as well as humidity measurements are some of the advantages that this dataset provides above traditional aircraft observations (e.g., the Aircraft Communication, Addressing, and Reporting System, ACARS), which have ascents and descents only at larger airport hubs. We have estimated the TAMDAR observational error and evaluated the subsequent impacts when employing the new error statistics in WRF through three assimilation experiments. The error estimation results are summarized as follows:

- The observational error of the TAMDAR RH is approximately 5.5%–9.0%, which is comparable to RAOB in winter months, and smaller than RAOB during summer months.
- The TAMDAR temperature error of 0.6–1.0 K is comparable to RAOB. The largest difference between RAOB and TAMDAR error at any time or level was 0.15 K.
- With respect to wind observations, the RAOB data have less error for speeds below 15 m s\(^{-1}\), while the opposite is true for speeds above this threshold. The differences typically remained less than 0.5 m s\(^{-1}\), with June producing slightly larger variance. The average magnitude of the TAMDAR wind error was approximately 2 m s\(^{-1}\).

In general, for both RAOB and TAMDAR, a slight increase in speed error was seen as a function of increasing wind speed; however, with this same increase in wind speed came a notable decrease in direction error from more than 40° for winds <3 m s\(^{-1}\) to roughly 10° for winds >15 m s\(^{-1}\).

Improvements in forecast skill are seen in both the 6- and 24-h forecasts throughout the altitude range where the TAMDAR data are collected. Additionally, greater gains, sometimes exceeding 50%, are achieved when the new error statistics are applied. As mentioned previously, because we wanted to highlight the difference in TAMDAR impact using the new error statistics, we withheld MDCRS data, which would be assimilated as a default AIREP in WRFDA. If the MDCRS data were included as a default AIREP, we would expect the relative impact of TAMDAR and the improvements discussed below, especially for wind and temperature, to be noticeably reduced.

- The 6-h (24-h) temperature forecast errors are reduced by up to 0.27 K (0.2 K), and the new error statistics are responsible for as much as 0.1 K of that difference (~37%).
- The impact of the TAMDAR wind observations was largest at 850 hPa, where it reduced the RMSE by 0.25 m s\(^{-1}\). Nearly all of this reduction was a result of the new error statistics at this level. Across the entire profile, average reductions between 0.1 and 0.2 m s\(^{-1}\) were noted, and approximately 50%–75% of that was attributable to the revised error.
Based on the evidence that the new error statistics play a leading role in the improvement in wind forecast skill below 500 hPa, it is concluded that the default AIREP wind error is large enough to inhibit the contribution of the TAMDAR wind observations. Thus, unique level-specific error statistics are warranted.

- The impact of TAMDAR RH is generally close to 0.1 g kg\(^{-1}\) from 925 to 700 hPa for the 6-h forecast. The maximum difference of either with-TAMDAR run over DEFer_nO\(\text{T}\) was 0.1 g kg\(^{-1}\) at 850 hPa. In the case of water vapor, the revised error statistics produced only a slight positive impact, which is not unexpected. In terms of RH, this change was approximately 2\% for the 6-h forecast, which is consistent with the findings presented in Moninger et al. (2010), where the RH error reduction in the 3-h RUC forecast that was attributed to TAMDAR varied between 1\% and 3\% RH. With an analysis RMSE of roughly 5\%-6\% RH, the improvement of 2\% RH, according to the estimated maximum potential improvement (EMPI; Moninger et al. 2010), would be approximately 35\%-40\%.

Due to the high temporal and spatial resolution of the TAMDAR data, real-time changes in the boundary layer and midtroposphere with respect to temperature and humidity, several stability parameters can be monitored (Szoke et al. 2006; Fischer 2006), and positive impacts on quantitative precipitation forecast (QPF) are achieved (Liu et al. 2010).

Future work will focus on refining the ability of the data assimilation methodology to extract the maximum benefit of the RH observations for the purpose of improving QPF skill. Additionally, significant work will be performed on characterizing the wind errors based on the phase of flight, space–time position in the atmosphere, and the observed wind direction and magnitude.

Acknowledgments. The authors thank Yong-Run Guo (NCAR), Daniel Mulally (AirDat), and three anonymous reviewers for their valuable comments and suggestions. We are grateful for the code support of TAMDAR assimilation provided by Hongli Wang (NCAR). The National Center for Atmospheric Research is sponsored by the National Science Foundation. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

APPENDIX

Error Analysis for Three-Way Collocation Statistics

Following the methodology presented by O’Carroll et al. (2008), below is the complete derivation of the set of equations used for obtaining the three-way collocation statistics. For a set of three collocated observation types \(i, j,\) and \(k\), the following set of equations is obtained:

\[
\begin{align*}
    x_i &= b_j + e_i + x_T \\
    x_j &= b_j + e_j + x_T \\
    x_k &= b_k + e_k + x_T,
\end{align*}
\]

where the observation \(x\) is expressed as the sum of the true value of the variable, \(x_T\), the bias, or mean error, \(b\); and the random error \(e\), which, over a reasonable sample size, has a mean of zero by definition. The mean difference between two observation types (e.g., \(i\) and \(j\)) is

\[
\bar{x}_i - \bar{x}_j = b_i - b_j + \bar{e}_i - \bar{e}_j + \bar{e}_{rep} + \bar{e}_{rep}, \tag{A2}
\]

where the representativeness error, which is always present when comparing two observation types, is broken into its mean (\(\bar{e}_{rep}\)) and random (\(e_{rep}\)) components. Because it is assumed that \(e\) has a mean of zero, and the mean of \(b\) is just \(b\), Eq. (A2) can be written as

\[
\bar{x}_i - \bar{x}_j = b_i - b_j + \bar{e}_{rep}. \tag{A3}
\]

Using Eqs. (A2) and (A3), we can express the difference between these two observation types as

\[
x_i - x_j = \bar{x}_i - \bar{x}_j + e_i - e_j + e_{rep}. \tag{A4}
\]

After rearranging (A4), squaring both sides, and expanding the terms, we obtain

\[
(x_i - \bar{x}_i)^2 + (x_j - \bar{x}_j)^2 - 2(x_i - \bar{x}_i)(x_j - \bar{x}_j) = (e_i - e_j + e_{rep})^2. \tag{A5}
\]

If we apply Eq. (A5) to a sample size of \(N\) collocated observations, we can express the equation as

\[
\frac{1}{N} \sum_{k=1}^{N} (x_k - \bar{x}_k)^2 + \frac{1}{N} \sum_{k=1}^{N} (x_k - \bar{x}_j)^2 - \frac{2}{N} \sum_{k=1}^{N} (x_k - \bar{x}_i)(x_k - \bar{x}_j) = \frac{1}{N} \sum_{k=1}^{N} (e_i - e_j + e_{rep})^2, \tag{A6}
\]

where the first two terms in (A6) are the variances of \(i\) and \(j\).
\[
\sigma_i^2 = \frac{1}{N} \sum_{k=1}^{N} (x_{ik} - \bar{x}_i)^2, \quad \sigma_j^2 = \frac{1}{N} \sum_{k=1}^{N} (x_{jk} - \bar{x}_j)^2, \quad \sigma_k^2 = \frac{1}{N} \sum_{k=1}^{N} (x_{ki} - \bar{x}_k)^2,
\]

and term three in (A6) is the covariance of \(i\) and \(j\), which can be expressed in terms of the standard deviation of \(i\) and \(j\):

\[
\frac{1}{N} \sum_{k=1}^{N} (e_{ik} - e_{jk} + e_{i\text{rep}}) \sigma_i \sigma_j = \sigma_i^2 + \sigma_j^2 + \sigma_{i\text{rep}}^2 - 2\sigma_i \sigma_j \sigma_{i\text{rep}}.
\]

Therefore, the variance of the difference between observation types \(i\) and \(j\), which is defined as

\[
V_{ij} = \text{var}(i - j) = \text{var}(i) + \text{var}(j) - 2\text{cov}(i, j),
\]

(A10)

can be written in terms of (A9), and applied to a three-way comparison between observation types \(i, j,\) and \(k\) to obtain the following set of equations:

\[
V_{ij} = \sigma_i^2 + \sigma_j^2 + \sigma_{i\text{rep}}^2 + 2r_{i\text{rep}} \sigma_i \sigma_{i\text{rep}} - 2r_{ij} \sigma_i \sigma_j
\]

\[
V_{jk} = \sigma_j^2 + \sigma_k^2 + \sigma_{j\text{rep}}^2 + 2r_{j\text{rep}} \sigma_j \sigma_{j\text{rep}} - 2r_{jk} \sigma_j \sigma_k
\]

\[
V_{ki} = \sigma_k^2 + \sigma_i^2 + \sigma_{k\text{rep}}^2 + 2r_{k\text{rep}} \sigma_k \sigma_{k\text{rep}} - 2r_{ki} \sigma_k \sigma_i
\]

(A11)

The reasonable assumption is made that the instrumentation error produced by completely independent techniques will be independent, and the instrumentation error is also independent of the representativeness error; thus, \(r = 0\), and we obtain

\[
V_{ij} = \sigma_i^2 + \sigma_j^2 + \sigma_{i\text{rep}}^2
\]

\[
V_{jk} = \sigma_j^2 + \sigma_k^2 + \sigma_{j\text{rep}}^2
\]

\[
V_{ki} = \sigma_k^2 + \sigma_i^2 + \sigma_{k\text{rep}}^2
\]

(A12)

Solving this system of equations for the variance of error for each of the three observation types yields

\[
\sigma_i^2 = \frac{1}{2} \left( V_{ij} - V_{jk} + V_{ki} - \sigma_{i\text{rep}}^2 - \sigma_{j\text{rep}}^2 - \sigma_{k\text{rep}}^2 \right)
\]

\[
\sigma_j^2 = \frac{1}{2} \left( V_{jk} - V_{ki} + V_{ij} - \sigma_{i\text{rep}}^2 - \sigma_{j\text{rep}}^2 - \sigma_{k\text{rep}}^2 \right)
\]

\[
\sigma_k^2 = \frac{1}{2} \left( V_{ki} - V_{ij} + V_{jk} - \sigma_{i\text{rep}}^2 - \sigma_{j\text{rep}}^2 - \sigma_{k\text{rep}}^2 \right).
\]

(A13)

where \(r_{ij}\) is the correlation coefficient. Applying Eqs. (A7) and (A8) to Eq. (A6) yields

\[
\sigma_{ij} = \frac{1}{N} \sum_{k=1}^{N} \left( x_{ik} - \bar{x}_i \right) \left( x_{jk} - \bar{x}_j \right) = r_{ij} \sigma_i \sigma_j.
\]

If strict collocation match conditions are met, as discussed in section 2, the variance of the representativeness error terms will be very small, and (A13) can be reduced to

\[
\sigma_i = \left[ \frac{1}{2} \left( V_{ij} - V_{jk} + V_{ki} \right) \right]^{1/2}
\]

\[
\sigma_j = \left[ \frac{1}{2} \left( V_{jk} - V_{ki} + V_{ij} \right) \right]^{1/2}
\]

\[
\sigma_k = \left[ \frac{1}{2} \left( V_{ki} - V_{ij} + V_{jk} \right) \right]^{1/2}
\]

(A14)

REFERENCES


AUGUST 2012 GAO ET AL.


