The Real-Time Mesoscale Analysis at NOAA’s National Centers for Environmental Prediction: Current Status and Development

Manuel S. F. V. De Pondeca,* Geoffrey S. Manikin,+ Geoff DiMego,* Stanley G. Benjamin,# David F. Parrish, R. James Purser,* Wan-Shu Wu,† John D. Horel,‡ David T. Myrick,§ Ying Lin,* Robert M. Aune,** Dennis Keyser,† Brad Colman,†† Greg Mann,** and Jamie Vavra@@

* IMSG, Washington, D.C., and NOAA/NWS/NCEP/EMC, Camp Springs, Maryland
† NOAA/NWS/NCEP/EMC, Camp Springs, Maryland
‡ NOAA/NWS/ESRL/GSD, Boulder, Colorado
# University of Utah, Salt Lake City, Utah
§ NOAA/NWS/Western Region Headquarters, Salt Lake City, Utah
** CIMSS, University of Wisconsin—Madison, Madison, Wisconsin
†† NOAA/NWS, Seattle, Washington
@@ NOAA/NWS, Detroit, Michigan
@@ NOAA/NWS/OST, Silver Spring, Maryland

(Manuscript received 21 October 2010, in final form 1 April 2011)

Abstract

In 2006, the National Centers for Environmental Prediction (NCEP) implemented the Real-Time Mesoscale Analysis (RTMA) in collaboration with the Earth System Research Laboratory and the National Environmental, Satellite, and Data Information Service (NESDIS). In this work, a description of the RTMA applied to the 5-km resolution conterminous U.S. grid of the National Digital Forecast Database is given. Its two-dimensional variational data assimilation (2DVAR) component used to analyze near-surface observations is described in detail, and a brief discussion of the remapping of the NCEP stage II quantitative precipitation amount and NESDIS Geostationary Operational Environmental Satellite (GOES) sounder effective cloud amount to the 5-km grid is offered. Terrain-following background error covariances are used with the 2DVAR approach, which produces gridded fields of 2-m temperature, 2-m specific humidity, 2-m dewpoint, 10-m wind components, and surface pressure. The estimate of the analysis uncertainty via the Lanczos method is briefly described. The strength of the 2DVAR is illustrated by (i) its ability to analyze a June 2007 cold temperature pool over the Washington, D.C., area; (ii) its fairly good analysis of a December 2008 mid-Atlantic region high-wind event that started from a very weak first guess; and (iii) its successful recovery of the finescale moisture features in a January 2010 case study over southern California. According to a cross-validation analysis for a 15-day period during November 2009, root-mean-square error improvements over the first guess range from 16% for wind speed to 45% for specific humidity.

1. Introduction

Recent years have seen an increasing demand for high-(1–5 km) resolution meteorological analyses at the National Oceanic and Atmospheric Administration’s (NOAA) National Weather Service (NWS) and in the environmental community at large. In particular, NWS forecasters use high-resolution analyses of sensible weather elements to help create and verify the gridded forecasts of its National Digital Forecast Database (NDFD; Glahn and Ruth 2003). High-resolution analyses are also increasingly necessary for mesoscale modeling and forecasting. They are also sought for climate-related applications, including climate change impact studies on sectors of critical importance to society such as human health, agriculture, energy, water resources, and coastal systems. Furthermore, the finescale details contained in these fields enhance the representativeness of the physical drivers in air quality models, forest fire models, and dispersion models used, for

Corresponding author address: Manuel S. F. V. De Pondeca, NOAA Science Center, Rm. 207, 5200 Auth Rd., Camp Springs, MD 20746.
E-mail: manuel.pondeca@noaa.gov

DOI: 10.1175/WAF-D-10-05037.1

© 2011 American Meteorological Society
example, to simulate the transport of pollutants and other hazardous materials. The creation of high-resolution analyses is also in tune with the current efforts to improve, expand, and strengthen the capabilities of the nation's mesoscale observational networks (National Research Council 2009).

To help meet its demand for high-resolution analyses, the NWS in 2004 established a multiphased program to build the Reanalysis of Record (Horel and Colman 2005). The Reanalysis of Record is an approximate 30-yr record of high spatial and temporal resolution three-dimensional reanalyses paralleling the global National Centers for Environmental Prediction–National Center for Atmospheric Research 40-Year Reanalysis (NCEP–NCAR; Kalnay et al. 1996) and the North American Regional Reanalysis (Mesinger et al. 2006).

The Real-Time Mesoscale Analysis, or RTMA, is the program's initial phase, intended as a proof of concept and baseline for subsequent developments toward the Reanalysis of Record. The RTMA capitalizes on and enhances the existing analysis capabilities at NWS to produce 6–1.25-km-resolution analyses of near-surface conditions on grids that match those of the NDFD. Among other applications, the RTMA is used at various Weather Forecast Offices (WFOs) as a tool to help initialize and adjust their digital forecasts within the framework of the Interactive Forecast Preparation System (IFPS; Ruth 2002) on the Advanced Weather Interactive Processing System (AWIPS; Seguin 2002). It also provides a means by which to assess the quality of the NDFD forecasts, as well as a reference data source to correct biases in model forecasts (Z. Toth 2007, personal communication).

This work reports on the development of the RTMA at NCEP, Earth System Research Laboratory (ESRL), and National Environmental, Satellite, and Data Information Service (NESDIS). The RTMA was implemented at NCEP Central Operations in 2006 for a domain matching the conterminous U.S. NDFD 5-km grid (CONUS-RTMA) and in 2008 for domains matching the 6-km Alaska, 2.5-km Hawaii, and 2.5-km Puerto Rico NDFD grids (Fig. 1). Preparations are under way to run the system for the Guam NDFD grid as well. In all cases, the grids (Fig. 1). Preparations are under way to run the RTMA and in 2008 for domains matching the 6-km CONUS version. The paper is organized as follows. Section 2 introduces the 2DVAR analysis. Section 3 describes the Rapid Update Cycle (RUC)-derived background fields, the observations, and their quality control. The Lanczos-based estimate of the 2DVAR analysis uncertainty is described in section 4. Section 5 presents three case studies that illustrate the performance of the 2DVAR and presents statistics of the fit of the analysis to the observations evaluated with the help of the cross-validation technique. Section 6 describes the mapping of the stage II precipitation amount and NESDIS GOES sounder effective cloud amount onto the RTMA domain. Section 7 describes the system improvements that are scheduled to be implemented during the next RTMA upgrade cycle, as well as the current developmental work. The conclusions follow in section 8.

2. The RTMA-2DVAR

a. The analysis system and the use of recursive filters

A major component of the RTMA is the NCEP Gridpoint Statistical Interpolation (GSI) Wu et al. (2002) used in the incremental 2DVAR mode. The GSI equations are solved on a Lambert conformal grid with a nominal resolution of 5 km (see Fig. 1). The analysis follows from minimizing the cost function:

$$J(x_k) = \frac{1}{2} x_k^T B^{-1} x_k + \frac{1}{2} (Hx_k - y_o)^T R^{-1} (Hx_k - y_o),$$

(1)

using the iterative preconditioned conjugate gradient algorithm described in Derber and Rosati (1989). The subscript $k$ denotes the iteration number, and $x_k$, the control vector, represents the departure of the estimate of the analysis from the specified background. The elements of $x_k$ are the analysis variables at every grid point, namely, streamfunction, velocity potential, temperature, surface pressure, and pseudo–relative humidity. The latter is defined as the actual specific humidity scaled by the saturated value from the background. The terms $B$ and $R$ are the background and observation error covariance matrices, respectively, and $H$ represents the linearized observation operator and $y_o$ the innovation vector. Observation errors are assumed to be uncorrelated, and $R$ is thus around 50 min past the hour and disseminated to NWS forecasters and external users through AWIPS and the National Digital Guidance Database, a sister to the NDFD. This paper focuses on the CONUS-RTMA. It describes its 2DVAR component in detail and provides a brief description of the derivation of the 5-km resolution precipitation and sky-cover fields. In the remainder of this work, the term “RTMA” implicitly denotes the CONUS version.
a diagonal matrix of observation error variances. Their values are given in section 3. Nonlinear effects in the incremental approach are accounted for by adopting multiple “outer loops” in the minimization and updating the first guess with the partial GSI solution of the previous outer loop at the beginning of the next. The RTMA uses two outer loops with 50 minimization inner loops each. The 2DVAR gridded output fields are for 2-m temperature, 2-m specific humidity, 10-m $U$ and $V$ components of the wind, and surface pressure. A 2-m dewpoint field is derived from the 2-m specific humidity and surface pressure fields.

The background error covariance matrix $\mathbf{B}$ is made up of block-diagonal matrices representing the autocovariances for each analysis variable. Cross covariances across variables can in general be accounted for by the GSI built-in balance equations. However, the RTMA turns off the balance equations, which renders the analysis univariate. Besides being poorly known, storing all the elements of the autocovariance matrix $\mathbf{C}$ is impractical due to its very large dimension. Computational feasibility is attained by (i) adopting a parameterized error model for $\mathbf{C}$ that uses a few easily obtainable measures of the background error statistics as its parameters and (ii) recognizing that the minimization algorithm in fact only requires a “recipe” for evaluating the matrix product $\mathbf{z} = \mathbf{Cf}$, given an arbitrary input vector $\mathbf{f}$.

In the RTMA, the autocovariances are of Gaussian form, with structure functions chosen to follow the underlying terrain field to a controlled degree. The mapping to the terrain is achieved by the explicit use of the local terrain gradient in the autocovariance model. The above matrix operation is accomplished with the help of spatial recursive filters (RFs; see appendix A for details). In the anisotropic RF approach, the effects of the autocovariance on the input vector are determined by three local parameters: (i) the spatial correlation scale, $L_h$, which determines the spatial reach of the covariance; (ii) the background error variance, $\sigma_o^2$, which determines the integrated amplitude of the response function when the
autocovariance is applied to a single, unit observation; and (iii) the function correlation scale, $L_f$, which controls the strength of the anisotropy. As seen from Eqs. (A1) and (A2) of appendix A, smaller (larger) values of $L_f$ correspond to increased (decreased) anisotropy for a given terrain gradient. Alternatively, for a fixed $L_f$ value, the larger the terrain gradient, the stronger the anisotropy. The latter effect translates a shortening of the “effective spatial correlation length” along the direction of the gradient vector. For the sake of argument, one could think in terms of the one-dimensional problem, in which case $S$ in Eq. (A2) becomes a scalar representing an “effective spatial correlation length.” Also, as expected, the special case of a zero terrain gradient yields the isotropic model.

The spatial correlation scales and variances of background errors are latitude dependent and represent a rescaling of those used with the NCEP Nonhydrostatic Mesoscale Model (NMM; Janjić et al. 2001; Janjić 2003) data assimilation system. The rescaling is performed in an extended parameter space that includes the $L_f$ parameters, which for simplicity are assumed constant over the analysis domain. Their initial values are chosen empirically. The rescaling attempts to lower the root-mean-square error of the fit of the analysis to the observations while avoiding overfitting the observations or producing fields of unrealistically small-scale analysis increments. The feedback provided by the WFOs on the RTMA performance plays an important role in the hitherto trial and error rescaling. The domain-averaged values of the spatial correlation scales, background error variances, and function correlation scales are listed in Table 1.

The use of a terrain-following autocovariance would seem intuitively appealing for representing background errors in temperature, surface pressure, and moisture fields, since these fields themselves tend to exhibit a strong dependency on the terrain. The procedure, however, becomes more difficult to justify for errors in the background wind, whose circulations often include flows going up and down mountain slopes. For that reason, the RTMA chooses the $L_f$ parameters such that the anisotropy is very weak for the wind analysis.

Figure 2 shows the autocorrelation functions for temperature and streamfunction for a selected “anchor point” located in a region of strong terrain gradient. The selected $L_f$ parameters lead to a very pronounced

| $\psi$ | 64567 | 3818 | 118.494 m$^2$ s$^{-1}$ |
| $\chi$ | 64764 | 3818 | 115.572 m$^2$ s$^{-1}$ |
| $T$   | 42834 | 636  | 1.7 K            |
| (RH)$_{\text{pseudo}}$ | 40054 | 636  | 0.21            |
| $P_{\text{sfc}}$ | 45003 | 636  | 1.9 hPa          |

**TABLE 1.** The horizontal spatial correlation scale $L_h$, function correlation scale $L_f$, and background error standard deviation $\sigma_b$ used for streamfunction $\psi$, velocity potential $\chi$, temperature $T$, pseudo–relative humidity (RH)$_{\text{pseudo}}$, and surface pressure $P_{\text{sfc}}$.

**FIG. 2.** Autocorrelation functions for (a) temperature and (b) streamfunction for the anchor point marked by the dot. From the outermost to the innermost contours, the autocorrelation values are 0.1, 0.2, 0.5, 0.7, and 0.8. Shaded contours show the underlying terrain used to construct the autocorrelations. Values on the $x$ and $y$ axes are in grid units, whereby ($x = 1, y = 1$) would correspond to the southwestern-most grid point of the RTMA domain. It is noted that the partial domain of (b) is larger than that of (a).
anisotropy for temperature and a nearly isotropic autocorrelation for streamfunction. Results for surface pressure and pseudo–relative humidity are qualitatively similar to those for temperature and are therefore not shown. Similarly, the results for velocity potential closely resemble those for streamfunction and are therefore also not displayed.

b. The handling of land–water contrasts

Analyzing or retaining the correct land–water contrasts can pose a challenge. This is an important aspect especially when analyzing temperature in the vicinity of shorelines. For such cases, an algorithm is desirable to restrict the influence of the warm (cold) land observations on the nearby cold (warm) water. Similarly, marine observations should have their influence mainly restricted to the water body. In the 2DVAR scenarios, this is achieved by sharpening the background error autocovariances along shorelines. The algorithm decreases the effective elevation height on major water bodies arbitrarily by 500 m in the autocovariance construction to sharpen the terrain gradient along the shorelines. Figure 3 illustrates the effects of this artifact on the analysis of a single temperature observation located on the eastern shore of Lake Michigan. Figure 3a shows the analysis increments in the absence of the artificial terrain gradient sharpening, and as expected, the contour lines cross the shoreline into the lake. Figure 3b shows the effects of sharpening the terrain gradient along the shoreline. As desired, the analysis increments are almost exclusively confined to the landmass and display a large gradient near the shoreline as the amplitudes quickly approach zero going into the lake.

3. The RTMA background, the observations, and their quality control

a. The RTMA background

The CONUS-RTMA 2DVAR uses 1-h forecasts from the 13-km Rapid Update Cycle (RUC; Benjamin et al. 2004), downscaled to the 5-km grid as its background. The RUC is a hybrid sigma-isentropic coordinate 3DVAR assimilation system developed at ESRL. Its detailed hourly assimilation of a large suite of both conventional and nonconventional observations, as well as the use of a digital-filter initialization that guarantees a 1-h forecast that is nearly free from model spindown, makes the RUC particularly appropriate for providing the background for further GSI-2DVAR enhancement. The downsampling procedure uses the 3D atmospheric–land surface representation provided by the RUC to generate surface fields that are physically consistent with land surface conditions, land–water contrasts, and terrain elevation of the RTMA grid. The procedure also attempts to preserve other 3D effects that are present in the RUC, such as realistic thermal stability, boundary layer structure, and local circulations. The downscaled RUC fields that are currently used in the RTMA are 2-m temperature, 2-m dewpoint, 2-m specific humidity, 10-m $U$ and $V$ wind components, and surface pressure. The downsampling is detailed in Benjamin et al. (2007).
comprises a bilinear horizontal interpolation of the RUC fields to the higher-resolution RTMA grid, the use of the near-surface vertical gradients from the RUC to adjust the interpolated fields to the RTMA 5-km terrain, and a sharpening of the expected coastline gradients in a manner consistent with the land–water-type indicators from the 5-km RTMA grid. The 5-km terrain of the RTMA, as well as the vegetation type from which the land–water-type indicators are derived, are extracted from the Weather Research and Forecasting Model (WRF) standard initialization program, without any smoothing applied.

For a selected October 2008 day, Fig. 4 illustrates the effects of the downscaling on the temperature field over a portion of the RTMA domain that includes the Rocky Mountains. Not surprisingly, the downscaling introduces local small-scale features reflecting the higher-resolution terrain (not shown). The effects of the terrain adjustment are also apparent in the wind speed displayed in Fig. 5. Compared with the RUC wind speed, the RTMA background displays small-scale patches of enhanced wind speed for most of western Colorado, over the Rocky Mountains.

As local flow features are enhanced, it is also important to ensure that at the minimum the downscaling algorithm does not degrade the domain-averaged statistics of the fit of the fields to the observations. Ideally, the procedure should in fact improve the statistics. Figure 6 compares how the RTMA first guess and the original RUC field fit the observations. The latter are the full pool of quality-controlled observations used in the 2DVAR assimilation. The statistics are for a period of 30 days, from 0000 UTC 2 July 2009 to 2300 UTC 31 July 2009. They represent time and domain-averaged root-mean-square errors (RMSEs) for each hour of the day. For the period chosen, the downscaling improves the

Fig. 4. (a) The original 13-km resolution 1-h RUC temperature forecast and (b) its 5-km downscaled version over a portion of the RTMA domain that includes the Rocky Mountains. Units are °C and the example is for the forecast starting at 1700 UTC 10 Oct 2008.

Fig. 5. As in Fig. 4, but for the 10-m wind speed displayed on a smaller domain that focuses mainly on the state of CO. Units are m s⁻¹.
b. Observations and quality control

The observations are grouped into four categories for their assimilation into the GSI, namely, temperature, moisture, surface pressure, and wind. These observations originate from synoptic, aviation routine weather report (METAR), Mesoscale Network (Mesonet), ship, buoy, tide gauge, and Coastal-Marine Automated Network (C-MAN) stations. The feed for Mesonet observations is the Meteorological Assimilation Data Ingest System (MADIS; Miller et al. 2007) and it is the Telecommunication Operations Center for the other observation types. The assimilation time window is 12 to 12 min centered around the analysis time. Its relatively small extent was chosen so that the observations are representative for the valid analysis time. As discussed in section 7, the forthcoming use of the first guess at the appropriate time (FGAT) technique will justify expanding considerably the assimilation time window for all observation types.

Figure 7 displays a representative map of the temperature stations found in the RTMA observation file for a given analysis time. The example shown is from 1500 UTC 20 November 2009. It contains 14 299 stations, of which 12 081 are land Mesonet stations, 2049 are land synoptic and METAR stations, and 169 are surface marine stations (ship, buoy, C-MAN, and tide gauge). Mesonet stations alone account for 84.5% of the stations, while land synoptic and METARs account for 14.3%, and marine stations account for 1.2%. These percentage numbers also hold approximately true for moisture, surface pressure, and wind observations. Likewise, Fig. 7 bears good resemblance with similar maps for these three other observation categories.

There are four layers of quality control in the RTMA-2DVAR:

1) Observations preflagged by the MADIS quality control process are also flagged by the RTMA.
2) The so-called gross-error check is performed at the beginning of each outer loop of the minimization: the absolute value of the ratio of the observation increment to the observation error is evaluated, and the observation is flagged when that ratio exceeds a preestablished threshold. For each observation type, Table 2 summarizes the observation errors and the thresholds used for the gross-error check.
3) Observation blacklists are used for each observation type. They combine four distinct lists: (i) a static list updated periodically by the WFOs, (ii) a static list composed from the MADIS monthly quality control statistics, (iii) a static list containing stations flagged by the RTMA developers during routine evaluation of the RTMA or following feedback from NWS field forecasters, and (iv) a dynamic list updated hourly by the RTMA. The latter is constructed using the gross-error statistics from the previous three to six analyses, and is applied to the next analysis.

4) In addition, for Mesonet winds the RTMA uses a list of “trusted providers,” as well as a list of “trusted stations,” both borrowed from the RUC system. Both lists are constructed from long-term statistics of the differences between the observed and the RUC model wind speeds. However, the “provider list” groups the stations according to the names of their providers and subproviders (information online at http://madis.noaa.gov/mesonet_providers.html), while the “station list” treats each station individually (Benjamin et al. 2010). Mesonet winds are only used in the assimilation if they belong to at least one of these lists.

The percentage of stations flagged by the various quality control mechanisms is around 10% for temperature and moisture, 25% for surface pressure, and a staggering 55% for wind. The latter figure stems mainly from the fact that about 60% of the Mesonet wind stations fail the provider- and station-list criteria for acceptability.

The quality control procedure for Mesonet winds is

<table>
<thead>
<tr>
<th>Observation type</th>
<th>Temperature</th>
<th>Pseudo-RH</th>
<th>Surface pressure</th>
<th>Wind</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\sigma_0$ (K)</td>
<td>$R$</td>
<td>$\sigma_0$ (%)</td>
<td>$R$</td>
</tr>
<tr>
<td>Land METAR</td>
<td>1.0</td>
<td>7.5</td>
<td>5.9</td>
<td>5.0</td>
</tr>
<tr>
<td>Land synoptic</td>
<td>1.0</td>
<td>5.0</td>
<td>5.9</td>
<td>5.0</td>
</tr>
<tr>
<td>Land Mesonet</td>
<td>1.2</td>
<td>7.0</td>
<td>7.1</td>
<td>3.0</td>
</tr>
<tr>
<td>Marine</td>
<td>1.0</td>
<td>5.0</td>
<td>5.9</td>
<td>5.0</td>
</tr>
</tbody>
</table>

Table 2. The square root of the observation error covariance ($\sigma_0$) for each observation type found in the input error table for the RTMA. Note that the actual observation error used may be inflated in the RTMA to account for the occurrence of duplicate observations or in response to the input quality control flag from MADIS. Also shown in the table is the ratio $R$ between the maximum acceptable error and $\sigma_0$ used in the gross-error check for each observation type. It is also noted that, although specific humidity is the moisture observation, the gross-error check is performed using pseudo–relative humidity (RH). The actual observation error for specific humidity is that of pseudo-RH multiplied by the local value of the background specific humidity.
indeed a challenging matter. Compared with the background, these data tend to exhibit a low speed bias arguably due primarily to questionable anemometer siting strategies (Benjamin et al. 2010). However, some of the perceived low speed biases may also be reflective of the lack of enough resolution and adequate boundary layer physics in the background. The example given later in section 5b illustrates the critical importance that quality control plays for the wind analysis.

4. Estimate of the analysis uncertainty

Associated with each analysis is the analysis error covariance matrix, \( A = (B^{-1} + H^T R^{-1} H)^{-1} \), which reflects the spatial error correlations for the observations and the background fields. The RTMA computes an estimate of \( A \) and takes the square root of its diagonal elements (i.e., of the analysis error variances) to represent the analysis uncertainty. The shortcoming of this definition rests in recognizing that the 2DVAR-associated \( A \) matrix assumes zero systematic errors in both the observations and background fields. Therefore, a more representative measure of the analysis uncertainty should in the future naturally incorporate these systematic errors. For now, they are accounted for in a simplistic manner by arbitrarily inflating the original estimates of the analysis error. Full details on how \( A \) is estimated will be presented in a forthcoming paper. The procedure is an adaptation of the Lanczos method to the GSI described by Fisher and Courtier (1995). It uses “by-products” of the minimization algorithm, namely, gradient vectors and step sizes, to estimate the \( A \) matrix. Appendix B describes the three steps of the estimate.

Figure 8 displays an example of the analysis error for temperature over a selected region in the states of Utah, Idaho, Wyoming, and Nevada. As expected, the analysis error is equal to the (inflated) background error in data-void regions. At the observation sites and in their vicinities, the analysis error is reduced compared with the background errors. This is the expected effect from the information contained in the observations. It is also apparent that the patterns of the analysis error are locally consistent with the shapes of the underlying, terrain-following background error covariances.

As discussed in appendix B, although the pattern of the analysis error is estimated quite accurately, the magnitude of the error is unfortunately underestimated. Calibration procedures for the error magnitude are still being pursued within the broader context of the GSI. At present, the RTMA analysis uncertainty is intended mainly for identifying regions of “small” and “large” errors in a relative sense.

5. Case studies and validation statistics

a. Analysis of a cold pool in the Washington, D.C., area

The RTMA has been shown to effectively capture convectively induced cold surface temperature pools, which are often absent from mesoscale model analyses and forecasts. Figures 9a and 9b show the RTMA first-guess temperature and analysis, respectively, for 2100 UTC 13 June 2007. The region of interest in this case description is the greater Washington, D.C., area, with emphasis on northeastern Virginia. That region is roughly centered at 38.8°N, 77°W. Figures 9c and 9d display the observations used during the first and second outer loops of the minimization, respectively. The colors of the small circles used to mark the location of the observations follow the color code shown below each panel. In the region of focus, both panels show a cold temperature pool, with many of the observations displaying values between 20° and 24° C in northeastern Virginia. The observations also show a tight temperature gradient, particularly across the District of Columbia and Maryland, with temperatures increasing rapidly away from the cold pool. Comparing Fig. 9a with Figs. 9c and 9d, it
Fig. 9. (a) The background temperature and (b) analysis in °C at 2100 UTC 13 Jun 2007. The small circles show the observations used during the (c) first and (d) second outer loops of the minimization. Their colors follow the color code shown at the bottom of the figure. (e),(f) Observations rejected during the first and second outer loops of the minimization, respectively.
becomes apparent that the first guess is too warm in the region of interest, with temperatures ranging from 27° to 29°C. Figure 9b, on its turn, shows the analysis displaying cooler temperatures that correspond very well with the observed values. The cold pool is captured well.

This example also illustrates the benefit of the two-outter-loop configuration described in section 1 for handling the observation quality control. It is noted that the cold pool is better defined in Fig. 9d than it is in Fig. 9c. The explanation goes as follows. The gross-error check at the start of the first outer loop rejected many of the observations, since they were too cold compared with the first guess. However, at the beginning of the second outer loop, the temperature of the updated first guess had been lowered sufficiently that nearly all of the observations previously rejected were now able to pass the gross-error check (see Figs. 9e and 9f).

b. Wind analysis during a high-impact event in the mid-Atlantic region

Mesonet winds provide an extraordinarily valuable source of observations. However, their usage in an assimilation must be undertaken with caution due to numerous quality control challenges. It is noted that it was not until the year 2007 that the RUC system and the NCEP North America Data Assimilation System (NDAS) began to assimilate this data type on account of the quality control challenges (Benjamin et al. 2010). Furthermore, the Global Forecast System (GFS; Sela 1980, 1982), which is one of NOAA’s most widely used atmospheric data assimilation–forecast systems, does not yet make use of these data.

The challenges associated with the assimilation of Mesonet winds during a high-impact event are illustrated by the 31 December 2008 mid-Atlantic region windstorm. Power outages and scattered damage were reported during the storm, which was associated with a strong pressure gradient. At 1500 UTC on that day, Fig. 10a displays the “non-Mesonet” winds that were available to the assimilation. The observations in Fig. 10a originate from 91 stations, of which three are surface marine stations, and the remaining 88 are METAR stations. It is noted that the METAR network is the most trusted of all networks available to the RTMA thanks to its accuracy and consistency. The reader is referred to NOAA (2005, chapter 3) for the quality control standards for METARs in the United States. Where available, the METAR winds can therefore be considered to be a good representation of the true winds. The eastern half of Virginia, as well as Maryland and the District of Columbia, are dominated by wind speeds stronger than 10 m s⁻¹. There are only a few exceptions, with reports between 5 and 8 m s⁻¹. The RTMA first guess, which is displayed in Fig. 10b, clearly underestimates the wind speed in the above region of focus. With the exception of small patches in central Virginia, and northeastern Virginia along the border with Maryland, where the winds are between 8 and 10 m s⁻¹, the whole area is dominated by lighter winds. Figure 10c displays all Mesonet wind speeds faster than 8 m s⁻¹. The picture agrees well with, and complements, that of Fig. 10a, which is composed of the most reliable stations. However, a highly contradicting picture is seen in Fig. 10d, which depicts all Mesonet winds with values lower than 8 m s⁻¹. Figure 10c and 10d, combined, reveal that the Mesonets winds are dominated by slow-biased reports. Most of the stations report values below 8 m s⁻¹, and in fact many of them report between 0 and 4 m s⁻¹ in the region of focus. It is therefore clear that quality control will play a crucial role in determining the quality of the analysis for this case. Figure 10e shows the RTMA analysis when the only active layers of quality control are the MADIS flags and the GSI internal gross-error check. Just as with the first guess, Fig. 10e reveals an analysis that is dominated by winds lighter than those reported by the trusted stations in Fig. 10a. Some improvements are however present in southern, east-central, and southeastern Virginia. Finally, Fig. 10f displays the analysis when all quality control layers are activated, namely, the MADIS flags, the gross-error check, the lists of trusted providers and trusted stations, the static blacklist, and the dynamic blacklist. Although not perfect, the analysis shown in Figure 10f represents a large improvement upon the first guess, with winds above 9 m s⁻¹ for most of the region of focus. Notice the improvement in the Washington, D.C., area and in most areas of Maryland and Virginia. However, the need to continue to improve the quality control procedure also becomes apparent. For example, although less severe than its counterpart in Fig. 10e, the “bull’s-eye” of light wind speed on the eastern shore of the Chesapeake Bay stems from a shortcoming in the quality control process. Since a tighter gross-error check could have the adverse effect of rejecting some of the good observations, further improvements to the quality control must be achieved mainly by improving the list of “trusted stations” and the dynamic blacklist.

c. Dewpoint “analysis” example for southern California

As mentioned in section 2, the RTMA uses the 2DVAR gridded fields of 2-m specific humidity and surface pressure to provide a diagnostic 2-m dewpoint field. Field forecasters in general have a preference for dewpoint rather than specific humidity for their routine operations.
Figure 11 illustrates how the RTMA is able to capture finescale details evident in the observations in a case when the first guess is already fairly good. Figure 11 refers to the 1800 UTC 28 January 2010 analysis, upon converting the original first guess, analysis, analysis increments, and observations expressed in terms of specific humidity into the equivalent dewpoint quantities. Figure 11a shows the first-guess dewpoint, and Fig. 11b...
displays the RTMA analysis. Figure 11c and 11d show the analysis increments and the observations, respectively. A comparison of Figs. 11a and 11d reveals the first guess in general reflects the observations well. A closer inspection, however, shows finescale details in the observations that are not present in the first guess. Some of these details are highlighted by the three black circles in Fig. 11a. The dewpoint values inside the northernmost and middle focus circles show a moister first guess than is revealed by the observations. Most observations report between 0° and 4° C in those two areas, while the first guess shows dewpoint values of 6° C and higher. In contrast, the values inside the southernmost circle show a first guess that is slightly drier than the observations. Some of the observations report over 10° C, while the first guess shows values below 10° C. Comparing the analysis in Fig. 11b with the first guess and the observations, it becomes apparent that the RTMA is able to correct the first guess in the above circled areas to bring it closer to the observations. Negative dewpoint increments are applied inside the northernmost and middle focus circles, while a positive increment is applied inside the southernmost circle, as revealed by Fig. 11c. It should also be noted that, just as desired, the derived dewpoint analysis increments also follow the contour lines of the underlying topography to some extent.

d. Average cross-validation statistics

In applications of data assimilation for the sole purpose of generating gridded fields of weather elements, cross-validation techniques represent perhaps the only robust methods for verifying the analysis. Cross validation denotes the use of a subset of the observed data to verify the analysis performed with these selected data withheld. The reliability and robustness of the cross-validation method can be improved by using multiple disjoint validation subsets to produce an aggregate validation statistic. It is desirable that each validation subset contains representative data from all of the
geographical regions observed but without the redundancy of close pairs or tight clusters, even if these are present in the original dataset. For RTMA applications, the desired properties of representativeness and nonredundancy are achieved fairly well and efficiently with the help of a simple application of a continuous space-filling Hilbert curve (De Pondeca et al. 2006; Purser et al. 2009; Tyndall et al. 2010). All the data are mapped and sorted in the order of the curve’s single real parameter to simultaneously construct multiple disjoint subsets. It is worth noting that a commonly used alternative approach was adopted by Horel and Dong (2010). They performed similar cross-validation experiments in which all observations used in a control analysis were used except for one. Then, this leave-one-out cross-validation procedure was repeated, removing the next, etc.

The aggregate statistics in this work represent averages over five validation subsets, each containing approximately 10% of the observations. Separate subsets are constructed for each observation category (i.e., temperature, specific humidity, wind, and surface pressure). For temperature, Fig. 12 illustrates the fifth subset constructed from the pool of observations shown in Fig. 7. It displays good geographical representativeness (better than the original data) and keeps the occurrence of close pairs of observations and clustering to an acceptable minimum.

Cross-validation root-mean-square errors (cv-RMSEs) were computed for a period of 15 days from 0000 UTC 15 November 2009 through 2300 UTC 29 November 2009 at 3-h intervals. Figure 13 displays the results for the analyzed 2-m temperature, 2-m specific humidity, surface pressure, and 10-m wind speed, and the respective backgrounds. They represent average values for the RTMA domain, computed as a function of time of day. Compared with the background, Fig. 13 reveals a consistent decrease of the cv-RMSEs in the analysis for all four fields. The 15-day time-averaged statistics are summarized in Table 3. In particular, the percent improvements are 34% for temperature, 45% for specific humidity, 29% for surface pressure and 16% for wind.

In interpreting the cross-validation error statistics, it is important to bear in mind that they include not only the errors in the analysis, but also the measurement and representativeness errors from the cross-validating observations themselves. One implication is that the cv-RMSE tends to overestimate the true error in the analysis. Conversely, the cross-validation percentage improvement tends to underestimate the true percentage improvement in the analysis errors. Finally, it should also be pointed out that the statistics remain weighted to the data-dense regions of the central and eastern United States, simply because there are more stations in those regions. Scores are likely lower in regions of complex terrain.

6. RTMA precipitation and sky cover grids

The RTMA quantitative precipitation amount is obtained by bilinearly interpolating the so-called early version of the NCEP stage II multisensor hourly precipitation analysis from the 4-km Hydrologic Rainfall Analysis Project grid (Lin and Mitchell 2005) to the 5-km NDFD grid. A snapshot of the analysis is shown in
The stage II analysis is made from radar hourly precipitation estimates and hourly rain gauge data received at NCEP. The real-time radar data are from the real-time digital precipitation arrays produced at the approximately 150 CONUS Weather Surveillance Radar-1988 Doppler (WSR-88D) sites. The early version of the stage II is run at 35 min after each hour, using the approximately 1500 Automated Surface Observing System (ASOS) sites obtained from the METAR files, which are usually available within 30 min after the hour. For the benefit of clarity, the existence of the so-called mid and late versions of the stage II data is also noted. They are run at NCEP 6 and 18 h later, using the more complete set of approximately 5000 gauge reports from the Hydrometeorological Automated Data System, most of which are not received in time for the RTMA.

The RTMA sky cover represents a bilinear interpolation of the 10-km resolution NESDIS effective cloud amount (ECA) from GOES sounders (e.g., Schreiner et al. 1993; Bayler et al. 2000) to the 5-km NDFD grid. It should be noted that the ECA is a proxy for the actual sky cover as defined in NOAA (2005, chapter 9). The ECA measures the opacity of the cloud. A 100% ECA value refers to a cloud that is totally opaque to radiation from below such that none reaches the satellite sensor. In contrast, an ECA value of 10% can result from thick clouds covering 10% of the satellite field of view (FOV), or a thin cloud covering the entire FOV.

Fig. 14. The stage II analysis is made from radar hourly precipitation estimates and hourly rain gauge data received at NCEP. The real-time radar data are from the real-time digital precipitation arrays produced at the approximately 150 CONUS Weather Surveillance Radar-1988 Doppler (WSR-88D) sites. The early version of the stage II is run at 35 min after each hour, using the approximately 1500 Automated Surface Observing System (ASOS) sites obtained from the METAR files, which are usually available within 30 min after the hour. For the benefit of clarity, the existence of the so-called mid and late versions of the stage II data is also noted. They are run at NCEP 6 and 18 h later, using the more complete set of approximately 5000 gauge reports from the Hydrometeorological Automated Data System, most of which are not received in time for the RTMA.

The RTMA sky cover represents a bilinear interpolation of the 10-km resolution NESDIS effective cloud amount (ECA) from GOES sounders (e.g., Schreiner et al. 1993; Bayler et al. 2000) to the 5-km NDFD grid. It should be noted that the ECA is a proxy for the actual sky cover as defined in NOAA (2005, chapter 9). The ECA measures the opacity of the cloud. A 100% ECA value refers to a cloud that is totally opaque to radiation from below such that none reaches the satellite sensor. In contrast, an ECA value of 10% can result from thick clouds covering 10% of the satellite field of view (FOV), or a thin cloud covering the entire FOV. The cloud

Fig. 13. Average cross-validation RMSE statistics for the 15-day period from 0000 UTC 15 Nov through 2300 UTC 29 Nov 2009 computed at 3-h intervals for (a) 2-m temperature, (b) 2-m specific humidity, (c) surface pressure, and (d) 10-m wind speed. Lines with open (closed) circles show the errors in the first guess (analysis). The statistics represent averages over five disjoint validation sets, each containing approximately 10% of the observations.

### Table 3. Time- and domain-averaged cross-validation RMSEs for the 15-day period from 0000 UTC 15 Nov through 2300 UTC 29 Nov 2009.

<table>
<thead>
<tr>
<th>Field</th>
<th>FG</th>
<th>ANL</th>
<th>IMPROV (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T$</td>
<td>1.85 K</td>
<td>1.38 K</td>
<td>34</td>
</tr>
<tr>
<td>SH</td>
<td>0.68 g kg$^{-1}$</td>
<td>0.47 g kg$^{-1}$</td>
<td>45</td>
</tr>
<tr>
<td>$P_{slc}$</td>
<td>1.60 hPa</td>
<td>1.24 hPa</td>
<td>29</td>
</tr>
<tr>
<td>WSP</td>
<td>2.15 m s$^{-1}$</td>
<td>1.86 m s$^{-1}$</td>
<td>16</td>
</tr>
</tbody>
</table>
Algorithm accurately detects the presence of a cloud better than 95% of the time. Clear FOVs (0% ECA) are falsely detected (cloud missed) less than 5% of the time. The accuracy is affected by the season and time of the day. It tends to be closer to 95% in winter, and closer to 98% in summer. Also, the detection accuracy is better at local noon than at local sunrise. The algorithm is designed to minimize the occurrence of false detection, although on rare occasions, warm low clouds and very thin high clouds are incorrectly detected.

Figure 15 shows an example of the RTMA ECA for a selected analysis time.
7. Planned system improvements and ongoing work

a. Planned system improvements

The following new features have been tested and are planned for implementation during the next operational RTMA upgrade.

1) The nominal resolution of the CONUS-RTMA is to be refined from 5 to 2.5 km.
2) The so-called first guess at the appropriate time (FGAT) is to be added. In this scheme, the observation increment is computed at the exact report time for the observation by interpolating between several input first-guess fields valid at different times within the assimilation time window. The interpolated observation increment is then assumed valid at the analysis time. With this procedure in place, the assimilation time window is to be expanded from ±12 to ±30 min for all observation types, so that more data can be used in the assimilation.
3) A bias correction scheme for the first-guess temperature adapted from the unbiased sequential analysis algorithm of Dee and da Silva (1998) and Dee and Todling (2000) is to be added.
4) The ability to blend the first-guess winds from the RUC-based first guess with the winds from the Hurricane Weather Research and Forecasting Model (HWRF; Tallapragada et al. 2008) is to be added to provide a much-needed, improved first guess for tropical systems (Manikin and Pondeca 2009).

b. Ongoing work

Current work includes the following.

1) Testing the assimilation of low-orbit satellite winds, and of ocean surface winds derived from the satellite-borne Advanced Scatterometer (ASCAT; Gelsthorpe et al. 2000) and the multifrequency polarimetric microwave radiometer (WindSat; Gaiser et al. 2004).
2) Exploring the use of time-dependent autocovariances built from background field diagnostics, such as background potential temperature.
3) Adding the capability to analyze more meteorological parameters in the RTMA, including wind gusts (Zhu et al. 2009), planetary boundary layer height, maximum and minimum temperatures, and mean sea level pressure.
4) Improving the estimate of the analysis uncertainty to remove the ambiguity in the amplitude and at the same time incorporate information about the systematic errors from the background and the observations. The approach consists of using the statistics gained from routine cross-validation experiments to tune the numerical values of the error reduction computed from the Lanczos method.
5) Continuing to improve the quality control for Mesonet winds. The recently proposed building of the “National Mesonet of Mesonets” is expected to greatly aid in the quality control process thanks to the detailed metadata information that is to be collected along with the observations themselves. In particular, one can envisage an adaptive quality control for Mesonet winds that takes into account maps of the obstacles to the flow. It must be stressed that the advent of the National Mesonet of Mesonets is a response to the recommendations from the National Academy of Science (National Research Council 2009) for the need to improve coordination among the government agencies, commercial firms, and educational institutions that provide access to near-surface observational data across the United States.
6) Testing a nonlinear quality control for the observations.
7) Using aggregates of RUC model forecasts to derive more appropriate, location-dependent RTMA background error statistics, and applying the cross-validation technique to aid with their fine-tuning. For example, Tyndall et al. (2010) incorporate cross validation in their estimates of the background error variance of the RTMA temperature background fields as well as the observation error variance as a function of observation type (their Table 1). Their results suggest that using an observation to background error variance ratio greater than one may be beneficial. Cross validation is also a useful tool to help minimize the occurrence of analysis overfitting, which has been seen occasionally in the RTMA.
8) Producing a “quality mark” array accompanying each precipitation RTMA analysis, with information such as whether a grid point is outside of an effective radar coverage area, radar beam height, gauge density, etc., to give users some measure of the reliability of the analysis at each grid point.

8. Conclusions

A description of the NCEP Real-Time Mesoscale Analysis (RTMA), applied to the CONUS NDFD grid, was presented. The system applies the 2DVAR technique to analyze near-surface observations and produce 5-km-resolution gridded fields of 2-m temperature, 2-m specific humidity, 10-m U and V components of the wind, and surface pressure. In addition, it computes a diagnostic
field of dewpoint temperature. The background error autocovariances of the 2DVAR are prescribed to follow the contour lines of the terrain field to a controlled degree. The 2DVAR uses the 13-km 1-h forecast from the RUC downscaled to the 5-km-resolution RTMA domain as its first guess. Along with the analysis, the 2DVAR computes gridded fields of the analysis uncertainty, which are equal to the Lanczos-based estimate of the square root of the analysis error variances. The analysis is performed hourly using observations from METARs, and synoptic, Mesonet, and marine stations. Besides honoring the MADIS quality control flags, the RTMA quality control for the observations consists of the GSI-internal gross-error check, and the use of station blacklists. For Mesonet winds, it also uses a list of trusted providers, as well as a list of trusted stations. The RTMA also creates grids of quantitative precipitation amount and sky cover by remapping the NCEP stage II precipitation data and the NESDIS GOES sounder effective cloud amount to the 5-km grid.

Three examples were presented to illustrate the performance of the RTMA-2DVAR. In the first example, the system was shown to capture a temperature cold pool in the Washington, D.C., area, which was present in the observations but absent from the first guess. The second example demonstrated the critical importance quality control plays for Mesonet winds, a data type known to exhibit a slow bias compared with the first guess. Finally, focusing on southern California, the third example demonstrated the RTMA capability to capture finescale details in the dewpoint that were present in the observations but absent from an otherwise quite reasonable first guess. Domain-averaged, cross-validation statistics were also presented for a 15-day period in November 2009. Measured in terms of the RMSE, the analysis was shown to be a significant improvement upon the background. A list of system improvements that are scheduled for implementation during the next RTMA upgrade cycle was presented. Of note is the planned doubling of the resolution to 2.5 km. The ongoing work, including testing the use of wind observations from the ASCAT and WindSat observing platforms and improving the estimate of the analysis uncertainty with the help of cross-validation statistics was described.

A main goal of this project is to see RTMA become a basic resource tool for NWS forecasters to use in their routine operations, and an important data resource serving the academic community and other federal users. Forecasters are thus expected to continue to send their feedback as the various changes are tested, and the broader community of users is urged to follow suit. Indeed, the continued system evaluation and broadening of the pool of evaluators is crucial to guaranteeing the quality of the RTMA and its multiphased upgrading toward the Reanalysis of Record.

Acknowledgments. We thank Drs. Joshua Watson, Dave Radell, and Steven Lazarus for their many contributions to this project. This research is, in part, in response to requirements and funding by the Federal Aviation Administration (FAA). The views expressed are those of the authors and do not necessarily represent the official policy or position of the FAA.

APPENDIX A

The 2DVAR Autocovariance Model of Background Errors

The autocovariances of background error are synthesized with the help of spatial recursive filters (RFs) in conjunction with the line-filtering “triad” algorithm (Purser et al. 2003a,b; Purser 2005). This quasi-diffusive combination of filters allows for the implementation of any arbitrary anisotropy, as quantified by a second-moment “aspect tensor,” S. The direct role of this aspect tensor is to act as an effective diffusivity guiding the application of the filters, and, when its spatial variation is effectively negligible, the implied autocovariance between two points separated by ∆x is of the form

$$C(\Delta x) = \sigma_o^2 \exp \left( -\frac{1}{2} \Delta x^T S^{-1} \Delta x \right),$$  \hspace{1cm} \text{(A1)}

where \( \sigma_o \) is the background error standard deviation. Equation (A1) is only approximate, however, when \( S \) is varying, as the covariance is synthesized as the result of a diffusion-like process.

The mapping of the autocovariance to the terrain follows a variant of the method of Riishøjgaard (1998). Specifically, the following form is assigned to the matrix of the inverse aspect tensor at each grid point:

$$S^{-1} = \frac{1}{L_h^2} + \frac{1}{L_f^2} (\nabla h) (\nabla h)^T.$$

\hspace{1cm} \text{(A2)}

The first term on the right-hand side of Eq. (A2) represents the inverse isotropic aspect tensor, where \( L_h \) is the horizontal spatial correlation scale and \( I \) the \( 2 \times 2 \) identity matrix. The second term accounts for the anisotropy, where \( \nabla h \) is the local terrain gradient and \( L_f \) the function correlation scale.
Estimation of the Analysis Error via the Lanczos Algorithm

It can be shown that the adaptation of the Lanczos method described by Fisher and Courtier (1995) to take into account the B preconditioning of the GSI leads to the following steps for estimating the analysis error.

1) Compose the tridiagonal matrix $T$ of the following elements:

$$ T_{kl} = \begin{cases} \delta_k & : l = k \\ \beta_l & : l = k + 1 \\ 0 & : |l - k| > 1 \end{cases}, $$

(B1)

where $k$ is the minimization inner-loop counter, $\delta_{k+1} = (1/\alpha_k + \gamma_k/\alpha_k)$, and $\beta_{k+1} = -c^k(\alpha_{k+1})$. Here, $\alpha_k$ is the minimization step size, $c_k = 1/\sqrt{\mathbf{g}_k^T \mathbf{Q} \mathbf{g}_k}$, and $\gamma_k = (\mathbf{g}_k^T \mathbf{Q} \mathbf{g}_{k+1})/(\mathbf{g}_k^T \mathbf{Q} \mathbf{g}_k)$, whereby $\mathbf{g}_k$ is the gradient of the cost function with respect to the control vector, and $\mathbf{Q} = \mathbf{B} \mathbf{g}_k$. These vectors, along with the step size, are readily available as by-products of the 2DVAR minimization. The dimension $n$ of the gradient vectors is that of the control vector. In the RTMA, $n = 3966485$. Denoting by $m$ the total number of inner loops, the maximum dimension of $T$ is $m \times m$.

2) Find the eigenvectors and eigenvalues $\{\mathbf{f}_k, \lambda_k\}$ of $T$.

3) Compute the estimate of the analysis error covariance matrix as

$$ \mathbf{A} \approx \mathbf{B} - \sum_k^m (1 - \lambda_k^{-1})(\mathbf{Q} \mathbf{f}_k)(\mathbf{Q} \mathbf{f}_k)^T, \tag{B2} $$

where $\mathbf{Q}$ is the $n \times m$ matrix whose columns are the $\{\mathbf{q}_k\}$ vectors. The analysis error is taken to be the square root of the diagonal elements of $\mathbf{A}$.

Since the number of leading vectors $\{\mathbf{Q} \mathbf{f}_k\}$ used in practice is very small compared with the rank of the full analysis error covariance matrix, Eq. (B2) implies an overestimation of the analysis error variances. However, sensitivity tests with respect to the number of eigenpairs retained to estimate $\mathbf{A}$ suggest that the method is very good at estimating the patterns of the analysis error variances (see also Fisher and Courtier 1995). Section 7 discusses how the Lanczos method might be complemented with cross-validation information to yield more accurate magnitudes for the errors.


Purser, R. J., 2005: A geometrical approach to the synthesis of smooth anisotropic covariance operators for data assimilation. NOAA/NCEP Office Note 447, 60 pp.


