5.7 Symp. on Observing and Understanding the Variability of Water in Weather and Climate,
AMS Annual Meeting, Long Beach, CA, Feb. 2003

NCEP REGIONAL REANALYSIS

Fedor Mesinger*, Geoff DiMego+, Eugenia Kalnay# (PIs), Perry Shafran~, Wesley Ebisuzaki^, Yun Fan¬, Robert Grumbine+, Wayne Higgins^, Ying Lin+, Kenneth Mitchell+, David Parrish+, Eric Rogers+, Wei Shi¬, Diane Stokes+, and Jack Woolen~

* NCEP / EMC and UCAR; + NCEP / EMC; # Univ. Maryland; ~ NCEP / EMC and SAIC; ^ NCEP / CPC;
¬ NCEP / CPC and RSIS

Revised, 23 October 2002

1. Introduction

The objective of the NCEP's Regional Reanalysis (RR) is to create a long-term set of consistent climate data on a regional scale, for the North American domain. The RR, on its domain, will be superior to the completed NCEP / NCAR Global Reanalysis (GR), in both resolution and accuracy. This will be achieved using the GR to drive the RR system, and taking advantage of the regional Eta Model, and of the various advances that have been made in regional modeling and data assimilation since the GR system starting time of 1995. These advances include assimilation of precipitation, direct assimilation of radiances, the use of additional data as well as improved data processing efforts, and several Eta Model developments, in particular those arrived at within the NCEP's GCIP-funded land-surface effort.

One of the expectations is that the RR will help answering questions of the variability of water in weather and climate, in particular as it concerns U.S. precipitation patterns. To that end, a special effort will be made to output all native grid time-integrated quantities of water and energy budgets. The RR should have a good handle on extreme events, such as floods and droughts, and should interface with hydrological models as well.

Results of preliminary pilots, produced at 80-km horizontal resolution and 38 layers in the vertical, have been inspected in a variety of ways, as well as reported on at several meetings (e.g., Mesinger et al. 2002a). The assimilation of precipitation during the reanalysis was found to be very successful, obtaining model precipitation quite similar to the analyzed precipitation, especially during the warmer seasons. In the 1998 pilot, temperature and vector wind rms fits to raobs were considerably improved over those of the GR throughout the troposphere, both in January and in July, and both in the analyses and in the first guess fields. Improvements in the 2 m temperatures and 10 m winds were seen as well.

We here report on our first tentative production results, at 32 km / 45 layer resolution. At the time of this writing preliminary production run is in progress, even though some final changes are still being worked on, so that the 32-km reanalysis so far done, the second half of 1987 and the first half of 1988, is being rerun at this time, and will be rerun once again when the system is frozen. It is planned to complete most of the 25 years of RR, 1979-2003, before the end of 2003. Once the 25 years are completed, the RR will continue to be run in real-time, like the "Climate Data Assimilation System" is being run as a real-time continuation of

* *Corresponding author address*: Fedor Mesinger, NCEP Environmental Modeling Center, 5200 Auth Road, Room 207, Camp Springs, MD 20746-4304; e-mail: fedor.mesinger@noaa.gov

the GR. Just as the GR, the final product is planned to contain also free forecasts at regular intervals, useful for predictability studies.

The project is supported by the NOAA Office of Global Programs (OGP), and at the time of this writing has just completed its 4th year of support, thus having ended its Development Stage. An 11 member Scientific Advisory Panel chaired by John Roads provides advice to the project and to OGP.

2. Reanalysis system and data

The RR System is identical to the Eta Model operational 3D-Var Data Assimilation System (EDAS), e.g., Rogers and DiMego (ftp://ftp.ncep.noaa.gov/pub/emc/wd20er/caftimay01/v3_document.htm), except for being augmented to use a variety of additional data sources, and for using an Eta Model at 32 km/45 layer resolution, in which presently some but not all of the model changes following October 2001 are implemented (see http://www.emc.ncep.noaa.gov/mmb/research/eta.log.html). In particular, cloud microphysics is one of October 2001 (Zhao et al. 1997). The system is fully cycled, with a 3-hr forecast from the previous cycle serving as the first guess for the next cycle.

The 32 km/45 layer resolution used for the RR production runs is the same as that of the operational Eta prior to September 2000. The domain however is that of the current operational Eta, including North America and parts of Atlantic and Pacific, and encompassing 106° x 80° of rotated longitude x latitude. The RR domain and topography are shown in Fig. 1.

The data used in most of the lower resolution pilot experiments performed so far, and in the production runs, includes the observations used in the Global Reanalysis, plus several additional sources:

- Precipitation. The assimilation of observed precipitation is the most important addition to the RR. The successful assimilation of these observations (Lin et al. 1999, see also section 3) ensures that the model precipitation during the assimilation is close to that observed, and therefore that the hydrological cycle is more realistic than it would be otherwise. Over the continental United States (ConUS), Mexico, and Canada, precipitation data assimilated are 24 h rain gauge data disaggregated into hourly bins. Over the ConUS area, the disaggregation is performed according to hourly precipitation data (HPD), and using "inverse distance" scheme, and the "mountain mapper" (PRISM). Over Mexico and Canada, disaggregation is based on GR-2 (Kanamitsu et al. 2002) forecasts. Over the remaining areas, with two exceptions, CMAP pentad data (Xie, Arkin, Janowiak) are used, converted to hourly also using the GR-2 precipitation forecasts. The two exceptions are areas north of 49° N if according to the Eta forecast precipitation is likely to be snow, since the CMAP data is then considered unreliable; and over tropical cyclones, since CMAP pentad data then do not have adequate time resolution to be useful. For this purpose, the locations of tropical cyclones are prescribed according to tropical cyclone retrievals of Fiorino (2002);
- TOVS-1b radiances (instead of the NESDIS TOVS retrievals used in GR-1 and GR-2);
- Profilers and Vertical Azimuth Display (VAD) winds;
- Land surface temperature, wind, and moisture;
- Lake surface: Ice cover (Grumbine, personal communication), and lake temperature, to the extent available, as opposed to the global SST used in the GR. For lakes for which temperature is not available, a system was set up to transfer temperatures of lakes nearest to the lakes considered but which do have the temperatures available;

SST and sea ice: For these data, while were used in the GR, improved processing has been developed for the RR (Stokes, Grumbine, personal communications).

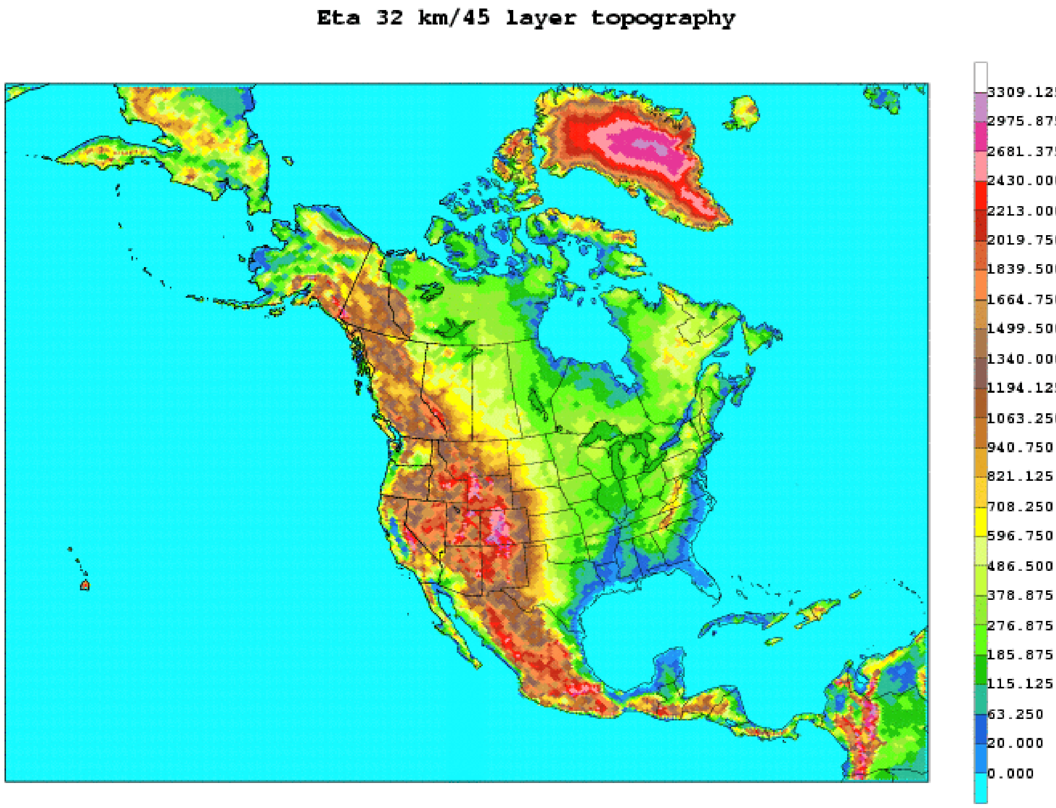

Fig. 1. The NCEP Regional Reanalysis domain and its 32 km / 45 layer topography.

3. Preliminary results

Analyzing a variety of pilot runs that we have performed prior to the current tentative production (e.g., Mesinger et al. 2002a), we have paid attention not only to their realism and sensitivities to RR system changes, but also to their comparison against the results of the NCEP / NCAR Global Reanalysis. Given that the Global Reanalysis data are already available (Kalnay et al. 1996, Kistler et al. 2001), an obvious goal of the RR, in addition to higher resolution, should be to provide a more realistic and accurate data set. Thus, results of various pilots were examined in comparison with the GR data (Mesinger et al. 2002a), testing whether there was increased agreement with observations. We shall do the same now when presenting and commenting upon the results of available production runs. We will show precipitation results, rms fits of analyzed as well as first guess temperature and winds to rawinsonde observations as functions of pressure, and rms fits of 2 m temperatures to surface observations, once again for analyses as well as for the first guess fields.

In Mesinger et al. (2002a) we displayed monthly precipitation totals of our 80 km pilot runs for January and July 1998, "observed" (as available at the time), those of the GR, and those of the RR pilot with precipitation assimilation. The observed precipitation was shown remapped to an 80-km grid, to be a more appropriate verification of the success of the assimilation than the original 1 / 8 degree input precipitation analysis, since the analysis remapped to the Eta 80-km "native" grid was the data assimilated into the RR pilots. Note however that the two 80-km grids were different, the difference being forced by the plotting package we use. We are displaying the same three plots here in Fig 2a ("observed") and 2b (GR, and RR), except for the 32 km preliminary production run for July 1987. The "observed" precipitation is now remapped to a 32-km grid.

Although the GR captures some of the characteristics of the large scale precipitation for July 1987, it fails to depict regional features, in particular over the ConUS area. It shows, for example, increased precipitation over the southeastern United States, as opposed to the Midwest. The RR, on the other hand, reproduces extremely accurately regional features over the ConUS area, even those of a very small scale; with few exceptions. Agreement is quite good over the oceans also, but not to the same extent.

Unfortunately, following July 1987 production run precipitation assimilation was inadvertently switched off in our system, and we have only relatively recently become aware of it. Thus, at this time, this period is being rerun, and we do not have the January and July 1988 32 km results that we otherwise would have. For these and other more recent results, please see http://www.emc.ncep.noaa.gov/mmb/rreanl where an updated version of this paper will be found. In the meantime, we recall results of the 80 km pilot shown in Mesinger et al. (2002a). In January 1998, a generally realistic result was produced by the GR as far as the larger scale features were concerned, with some exceptions. An even more realistic result was obtained with the RR; but not quite as realistic as the one for July 1998, and the 32 km one shown here for July 1987.

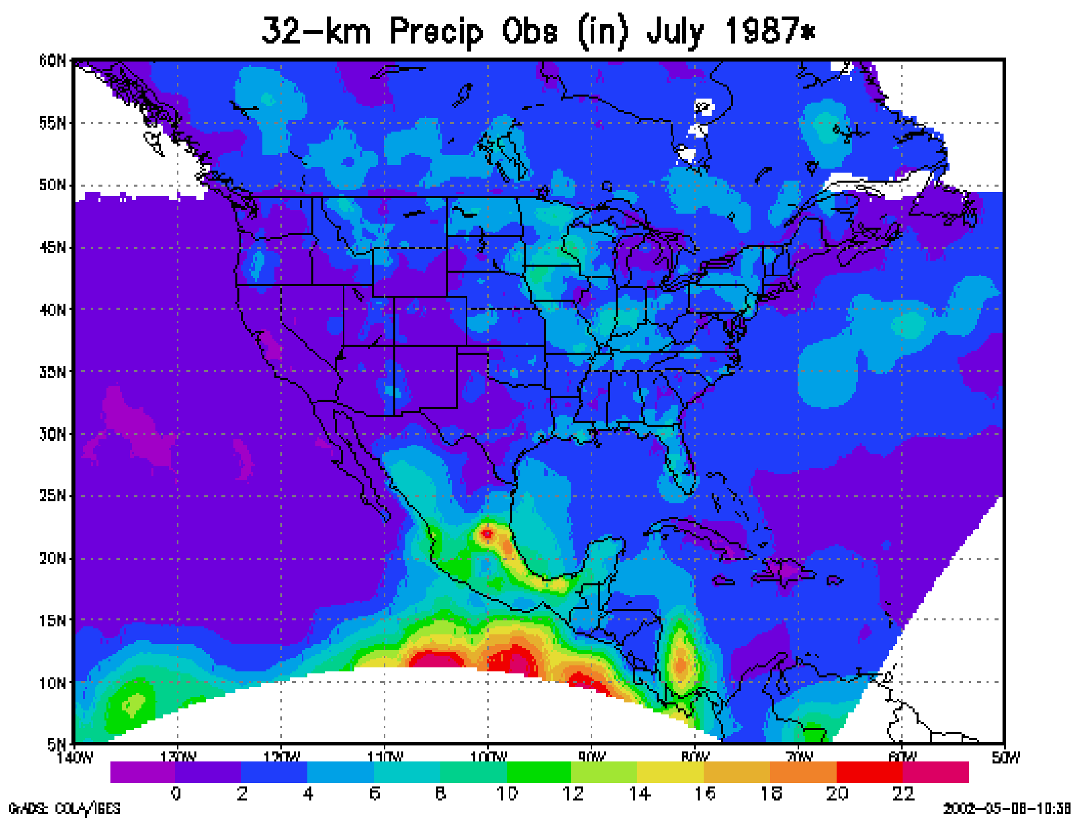

Fig 2a. “Observed” (see text) precipitation total for July 1987 (inches).

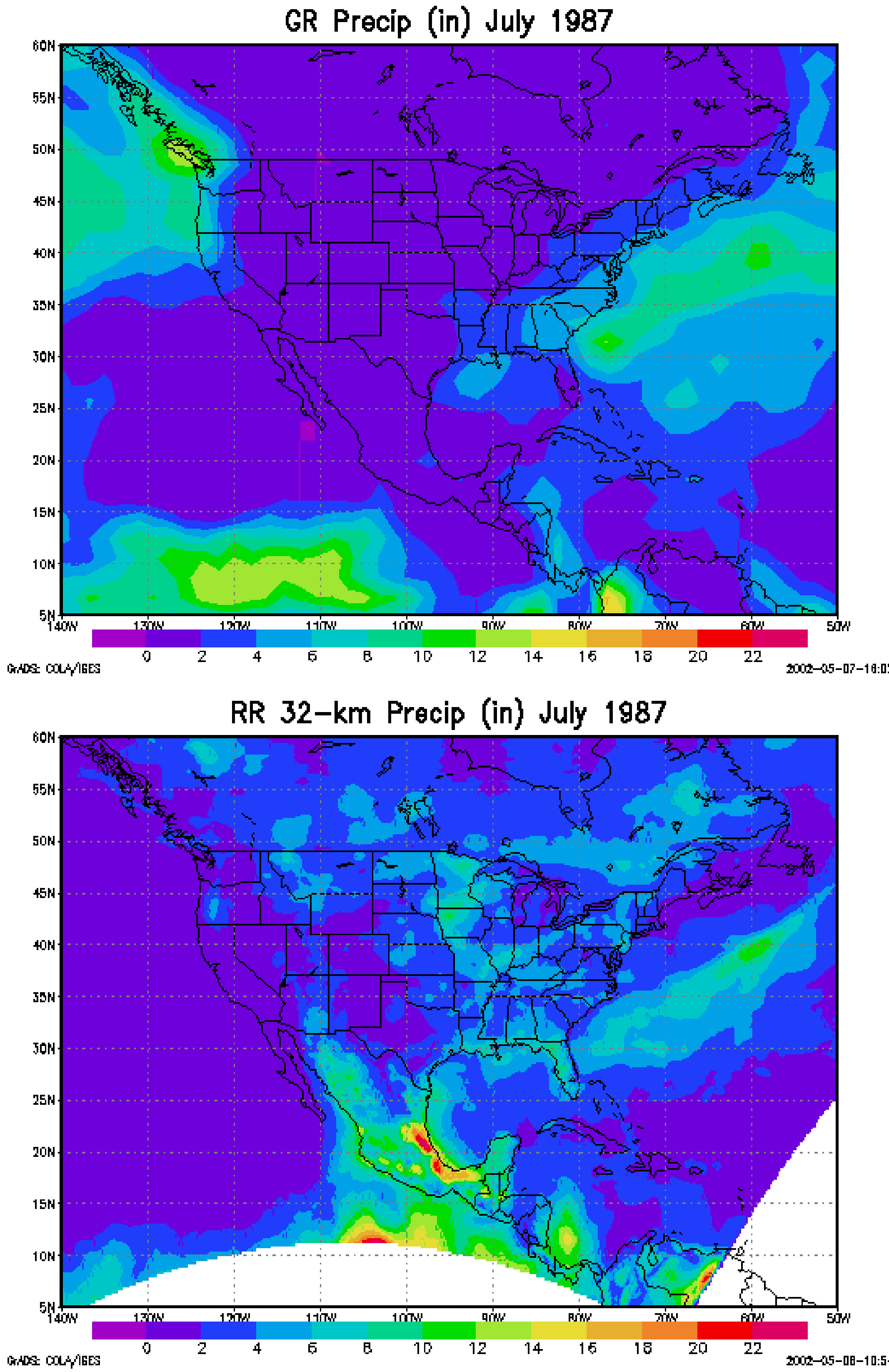

Fig 2b. GR (upper panel) and RR (lower panel) precipitation totals for July 1987 (inches).

It should be stressed that the RR precipitation is model produced; it is the latent heat, derived from observations, that is assimilated (e.g., Lin et al. 1999). A better agreement of the RR in summer, when precipitation forecasts are more difficult, thus may be found surprising. It likely indicates that our RR assimilation method works better with the convective precipitation, dominant in summer, than it does with the large-scale precipitation.

In Mesinger et al. (2002a) we have shown 80 km pilot temperature and vector wind RR rms fits to raobs, for January and July 1998, as functions of pressure. RR had shown a considerably better fit to raobs than the GR, in particular for winds. A puzzling feature was no improvement over the GR at the surface (1000 mb). This was later identified to have been due to an error in the NCEP Forecast Verification System (FVS): it was assuming the observed temperatures to be virtual temperatures and thus was including an erroneous conversion. When this was corrected the RR improvement over the GR at the surface was similar to that at the other levels.

We now present in Fig. 3 a similar set of four plots but for 1988 and the 32 km run. The advantage of the RR over that of the GR is seen to be considerable, in particular for winds. As in the pilot results, the advantage in winds over the GR is greatest in the upper troposphere, and in particular in winter, January. This is a feature which we believe would generally not have been expected. It is however consistent with results of the operational Eta that have been reviewed by Mesinger et al. (2002b).

As to the comparison with the corresponding plots of the 80 km pilot for 1998, the advantage of the RR over the GR is similar, or slightly better. In this respect, two points might be mentioned. First, it is probably generally accepted that the impact of resolution when comparing rms fits to data is not necessarily beneficial, as more detail even if mostly correct may not reduce the rms difference because of the negative impact of placement errors. Second, the 1988 results we have at the time of this writing and are showing in Fig. 3 are obtained with no precipitation assimilation. Improved results are expected when these months are rerun with the precipitation assimilation included.

Since observations are used in the analysis, the fit of the analysis to the observations shown in Fig. 3 is influenced both by the choice of background and observation error covariances, and by the degree of balance imposed on the analysis. The fit will be worse the stronger the balance constraint imposed in the analysis scheme. For this reason, the fit of the first guess to the observations is generally considered a better independent validation of the quality of the analysis system. For example, the changes implemented in the operational Eta 3D-Var in May 2001 (web site given in section 2) resulted in improved RR fits to rawinsondes in the first guess (3-h forecasts) but made them somewhat worse in the analysis. We therefore compare the RR and GR first guess fits to data, similar to the rms differences shown in Fig. 3 but prior to entering the 3D-Var analysis. From a practical point of view, most users of the RR will want to use the analyses (which are the best estimates) for the variables analyzed, but will use the first guess for non-analyzed fields such as surface fluxes.

The RR fits to raobs are still considerably better than those of the GR for the first guess fields, even though the improvement is not as large as for the analysis fields. In comparing these plots to those of Fig. 3, a point may come to mind additional to that of the balance imposed in the 3D-Var scheme. It is the difference in time intervals of using the 3D-Var analyses in the two assimilation systems: the GR is using 6 h assimilation intervals, as opposed to the 3 h intervals of the EDAS and the RR. This results in a considerable fraction of the data being used at more correct times, and also makes the first guess closer to the initialization time so that there is less time for the model error grows to take place, both favoring the RR. However, being closer to the initialization time also allows less time for the gravity waves created by

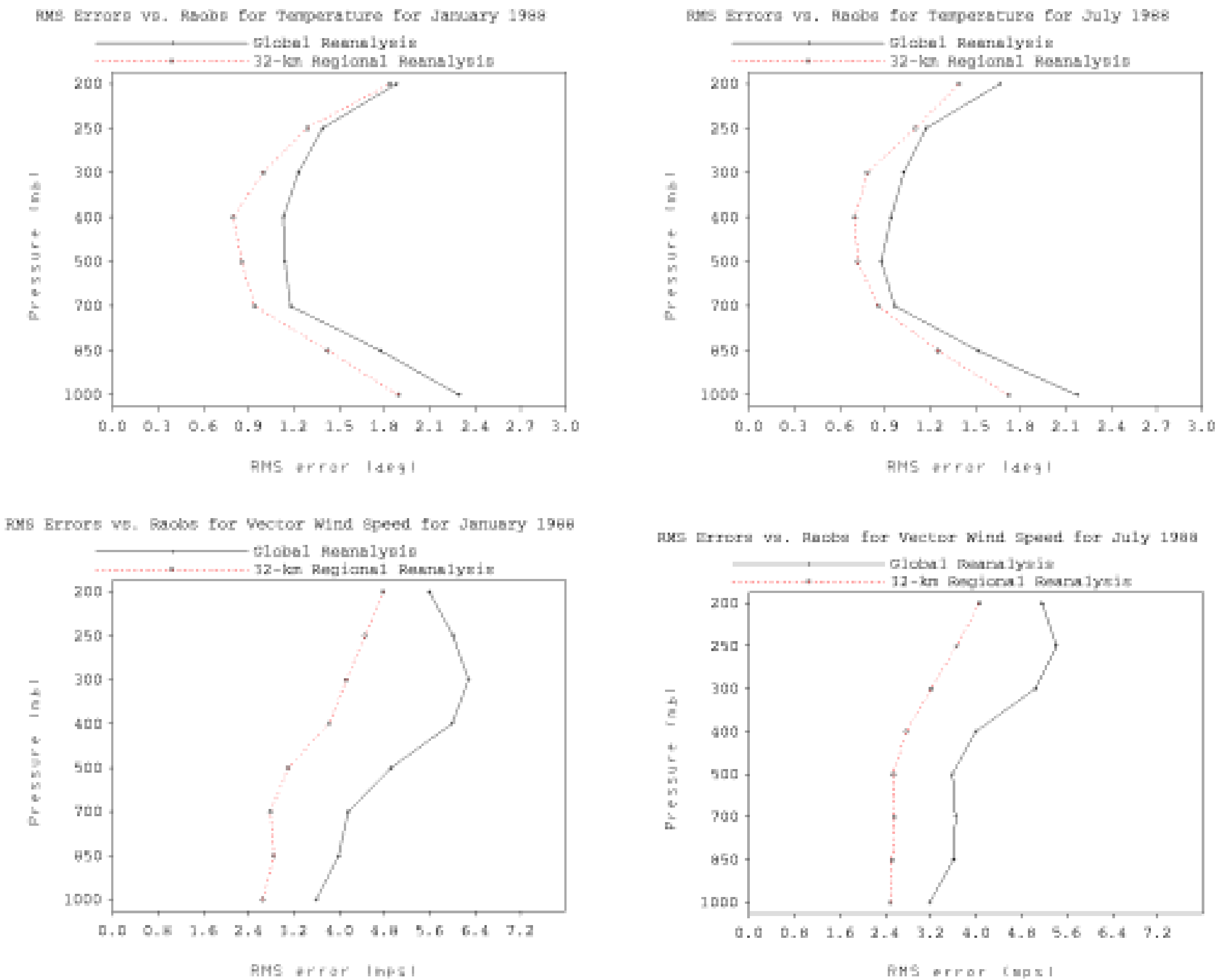

Fig. 3. RR rms fits to raobs as a function of pressure, dashed lines, for temperature (upper panels), and for vector wind (lower panels), for January (left panels) and July 1988 (right panels). Same, but for the GR, solid lines.

the initial imbalance to settle down, putting the RR first guess at a disadvantage in terms of fitting the observations.

We now compare in Figs. 5 and 6 the bias and the rms fits to surface observations of the first guess 2 m temperature, again for January and July 1988 for both the RR (dashed lines) and the GR (solid lines), as a function of time. Recall that, just as in Figs. 3 and 4, these results are also obtained without precipitation assimilation, so that somewhat improved RR fits to observations are hoped for once these months are rerun. Stations inside the so-called grid 212 are chosen for the verifications shown, encompassing most of Mexico to the south and up to a considerable fraction of Canada to the north, resulting in about 450 stations. A large majority of these stations however are within the ConUS area.

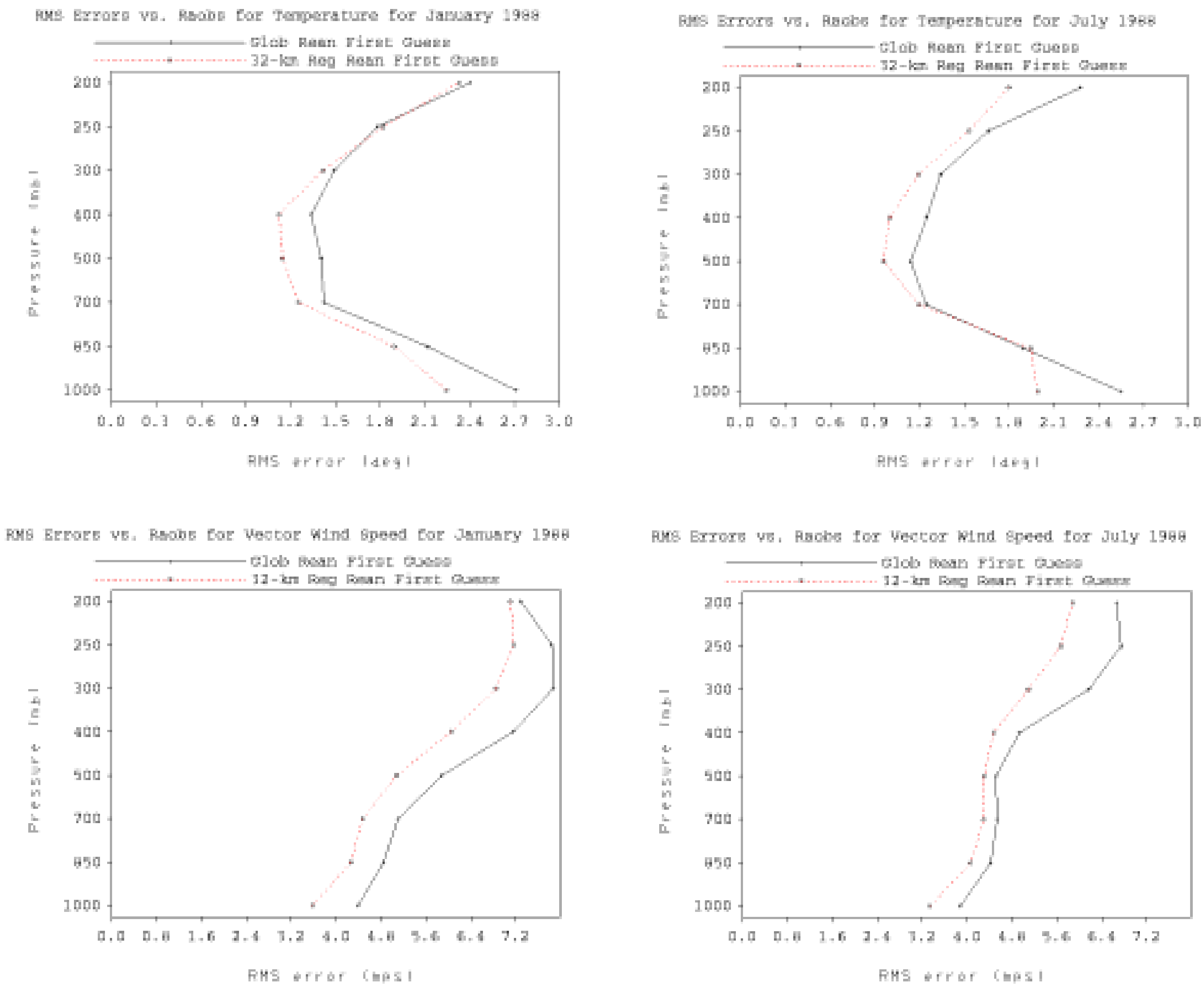

Fig. 4. RR first guess rms fits to raobs as a function of pressure, dashed lines, for temperature (upper panels), and for vector wind (lower panels), for January (left panels) and July 1988 (right panels). Same, but for the GR, solid lines.

We first show bias plots given that the bias is perhaps of particular interest when it comes to surface observations. In January, the RR has a very small overall bias, as opposed to a negative bias of the GR. In July, the RR displays a negative bias, but smaller and with much less of a diurnal oscillation than the GR which has a strong positive bias at 0000 UTC.

The rms plots (Fig. 6) show that in both winter and summer the RR rms fits are consistently less than those of the GR. However, in the winter there are occasional “unpredictable” periods that affect both the RR and the GR short-range forecasts. In the summer the diurnal cycle in the rms errors shows strong maxima at 0000 UTC, especially for the GR, and there is less of a relationship between exceptionally large errors in the two systems than in the winter. It is worth noting that the 1200 UTC July RR rms differences

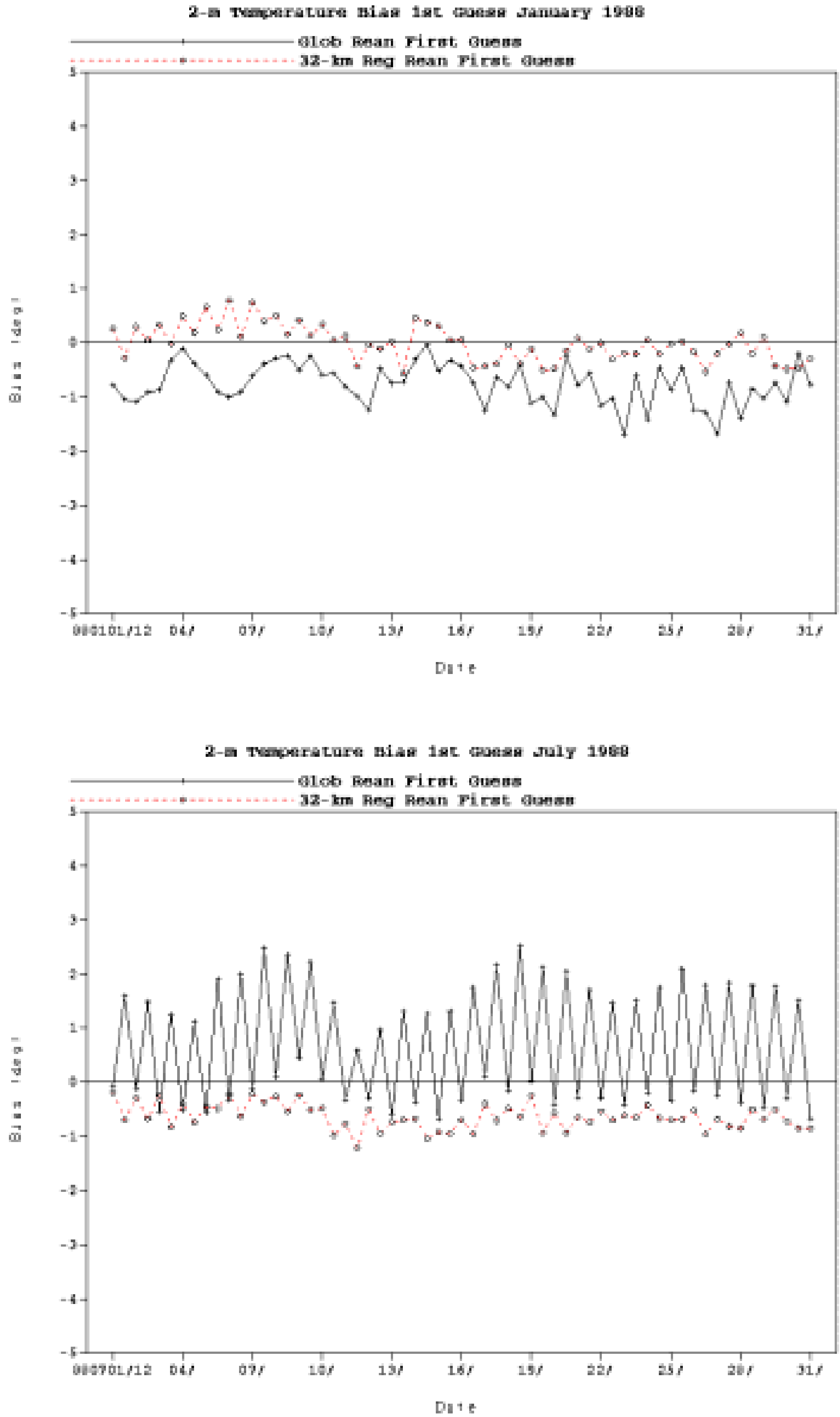

Fig 5. Bias of the first guess 2 m temperatures of the RR (dashed lines) and the GR (solid lines), for January (upper panel) and July 1988 (lower panel), as functions of time.

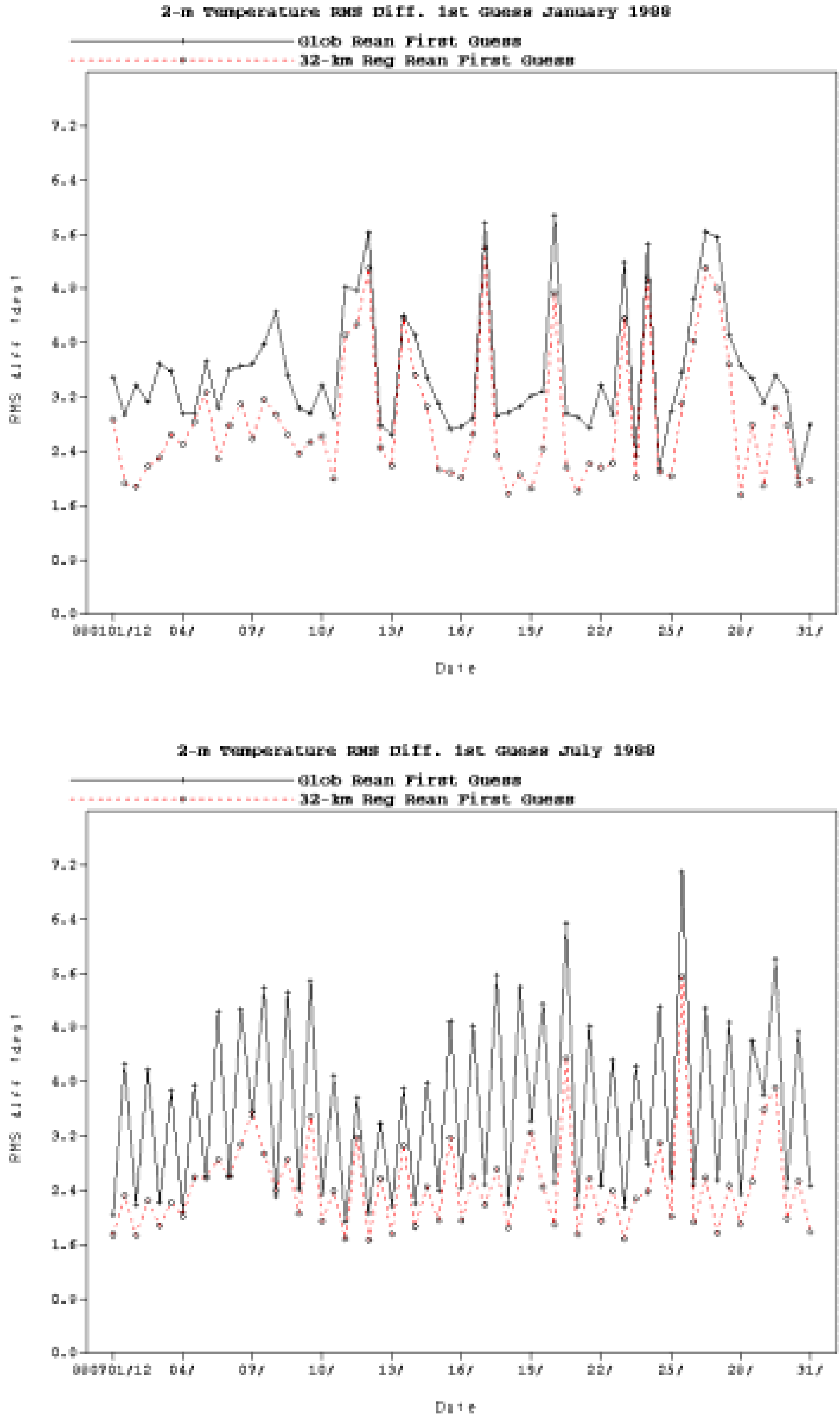

Fig 6. RMS fits to observations of the first guess 2 m temperatures of the RR (dashed lines) and the GR (solid lines), for January (upper panel) and July 1988 (lower panel), as functions of time.

are also consistently smaller than the their GR counterparts, even though their negative biases tend to be greater than those of the GR.

The bias and the rms plots of the first guess 10 m vector wind biases and rms fits (not shown) indicate that both the RR and the GR have a systematic negative bias, but the RR bias, on the order of up to about 1 m/s, is less than that of the GR, which tended to be about 1-2 m/s. However, this bias advantage of the RR did not result in a clear rms advantage. The rms values of the RR and of the GR in fact were remarkably similar. A tentative explanation for this puzzling result is that the gravity wave noise in the RR, at its 3 h first guess time, may be more intense that that of the 6 h noise of the GR.

4. Work in progress and plans

NCEP management has recently announced that half of the current NCEP mainframe computer IBM SP system will be devoted to the production of the RR, starting in early 2003. This will obviously increase manifold the production speed. In the meantime, we are proceeding with our tentative production, and at the same time working intensely on the preparation of various data sets ahead of time, so as to be ready for the accelerated production. The production will be carried out in two streams, one being the present stream that has started in July 1987 to be continued into the future in real time as in CDAS, and the other to start in 1978, eventually to overlap with the former one. A monitoring system to quality control the RR products and keep up with the production speed that we expect is under development.

Results are now being posted for evaluation by the expected user community as they become available, at http://www.emc.ncep.noaa.gov/mmb/rreanl. Comments on the results posted are most welcome and are hereby solicited.

References

Fiorino, M., 2002: Analyses and forecasts of tropical cyclones in the ECMWF 40-year Reanalysis (ERA-40). *25th Conf. on Hurricanes and Tropical Meteorology*. San Diego, CA, Amer. Meteor. Soc., 261-265.

Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.*, **77**, 437-471.

Kanamitsu, M., and Coauthors, 2002: NCEP/DOE AMIP-II Reanalysis (R-2). *Bull. Amer. Meteor. Soc.*, **83**, (in press).

Kistler, R., and Coauthors, 2001: The NCEP-NCAR 50-Year Reanalysis: Monthly means CD-ROM and documentation. *Bull. Amer. Meteor. Soc.*, **82**, 247-267.

Lin, Y., K. E. Mitchell, E. Rogers, M. E. Baldwin, and G. J. DiMego, 1999: Test assimilations of the real-time, multi-sensor hourly precipitation analysis into the NCEP Eta model. *8th Conf. on Mesoscale Meteorology*, Boulder, CO, Amer. Meteor. Soc., 341-344.

Mesinger, F., G. DiMego, E. Kalnay, P. Shafran, E. Berbery, W. Collins, W. Ebisuzaki, W. Higgins, J. Huang, Y. Lin, K. Mitchell, D. Parrish, E. Rogers, and coauthors, 2002a: NCEP Regional Reanalysis. *Symp. on Observations, Data Assimilation, and Probabilistic Prediction*, Orlando, FL, Amer. Meteor. Soc., J59-J63.

Mesinger, F., T. Black, K. Brill, H.-Y. Chuang, G. DiMego, and E. Rogers, 2002b: A decade + of the Eta performance, including that beyond two days: Any lessons for the road ahead? *19th Conf. on Weather Analysis and Forecasting/15th Conf. on Numerical Weather Prediction*, San Antonio, TX, 387-390.

Zhao, Q., T. L. Black, and M. E. Baldwin, 1997: Implementation of the cloud prediction scheme in the Eta Model at NCEP. *Wea. Forecasting*, **12**, 697-712.