6.3 OBSERVATION SENSITIVITY EXPERIMENTS USING THE RAPID UPDATE CYCLE

Barry Schwartz
Stan Benjamin

NOAA Research - Forecast Systems Laboratory, Boulder, Colorado

1. INTRODUCTION

Currently there is much discussion within the meteorological community regarding the design and implementation of our current and future observing systems. Assessing the relative value of observational platforms is useful both for scientific and budgetary interests. Advances in computer speed and economy of storage have made it possible in recent years to run numerical weather prediction (NWP) models retrospectively using archived observational and boundary conditions. This retrospective rerun capability additionally provides an important tool that can be used to assess modifications made to model and data assimilation techniques, code optimizations, and initial conditions with different combinations of observational data.

In this paper we discuss the use of the Rapid Update Cycle (RUC) model (Benjamin et al. 2003a,b), run operationally at the National Centers for Environmental Prediction (NCEP), to assess the relative impact of various observed data sources on forecast accuracy. The RUC was rerun over a 14-day period to assess the relative impact of systematically excluding the following data: rawinsonde, profiler, aircraft, WSR-88D velocity azimuth display (VAD) wind profiles, and surface observations. The RUC is particularly well suited for Observation Sensitivity Experiments (OSEs) since it employs an hourly intermittent 3-dimensional variational analysis (3DVAR) assimilation cycle (Benjamin et al. 2003b) allowing full use of all high frequency asynoptic data.

Corresponding author address: Barry Schwartz, NOAA/OAR/FSL, 325 Broadway, Boulder, CO 80305-6628: email: Barry.E.Schwartz@noaa.gov

Only a general survey of our results is presented here. However, a companion paper that discusses the usefulness and impact of wind profiler data in weather forecasting (Koch et al. 2004, also Benjamin et al. 2004), illustrates in larger detail the impact of denying profiler data from RUC retrospective reruns during significant weather event episodes.

2. EXPERIMENT DESIGN

The operational version of the 20-km RUC model was rerun for a number of 14-day experiments beginning 0000 UTC 4 February 2001. This period corresponds to the same test period used by NCEP to validate the implementation of the most recent versions of both the RUC and Eta models. As in the operational RUC, lateral boundary conditions were specified from the Eta model. Note that the Eta boundary conditions had all data sets available to it so that to a certain degree, the observations cannot be denied completely without rerunning the Eta model and denying them there. Moreover, the same Eta boundaries conditions were used in each experiment, constraining solutions toward those values.

Six retrospective experimental (EXP) runs were performed. The first was a control run containing all data sources (henceforth referred to as CNTL); reruns 2-6 had the following data sources denied individually; rawinsonde, profiler, aircraft, VAD, and surface observations. Verification was performed using conventional 12-hourly rawinsonde observations over two domains. The first domain contains all the rawinsondes located within the RUC domain; the second is a limited area enclosing all of the Midwest profiler stations (see Fig. 3 in Koch et al., 2004). Forecast minus observed (f-o) residuals for 3-, 6-, 9-, and 12 h temperature, relative humidity, and wind forecasts were computed at all rawinsonde sites which are used to compute average error statistics for each experiment. Additionally, quality control was performed on the rawinsonde data by disqualifying rawinsonde data that differed from CNTL RUC analyses by thresholds indicating that the same observations were flagged in the RUC analysis (Benjamin et al. 2003b). EXP minus CNTL averages, henceforth referred to as the data impact are presented at eight mandatory rawinsonde levels (850-150 hPa). The Student-t test is performed on the impact to assess the significance of the differences. Finally, we present verification results using *profiler* data as the verifying observation for the computation of residuals and impacts. The use of the profiler data for verification provides interesting diurnal resolution to the results and poses interesting questions about the verification process itself.

3. RESULTS

The EXP – CNTL impact on the vector wind of denying each data type for the 3-, 6, 9-, and 12 h forecast projections is depicted in Fig 1 for the RUC domain (~91 rawinsonde stations). The impact is greatest on the forecasts of shortest duration and decreases substantially for the 12-h forecasts, with the exception of the rawinsonde impact being the greatest at 12 h. The surface METAR observations are only a small factor at 850 hPa. In general, aircraft data is seen to have the greatest impact on wind forecasts, being greatest at flight levels (300-250 hPa). An unexpected modest VAD impact is seen in the upper troposphere even though these data are generally available only in the lower troposphere in winter.

Fig. 2 shows the same wind forecast EXP-CNTL results for observation impact, but here for verification over a smaller geographic subdomain over the central U.S. (~22 rawinsonde stations; 'profiler' subdomain). Here impacts tend to be larger in general than the over the full RUC domain, especially for profiler, which are substantially larger. Over this domain, the profiler shows more impact than the aircraft data below 400 hPa. A more convenient display of these results is presented by forecast projection in Figs. 3 (full domain) and 4 (profiler subdomain). Here it is easier to see that the aircraft data tend to have their greatest impact in the upper troposphere and near the tropopause whereas the profiler data impact wind forecasts in the lower troposphere the most.

The significance of each wind forecast impact as measured using the Student-t test is shown in Table 1 for results over the RUC domain. Values are shown down to the 80th percentile. The t value is dependent on the sample size and the standard deviation of the CNTL – EXP differences used in the calculation. The number of residuals is ~2000 (~ 91 stations X 14 days X 2 observations per day), but varies by altitude because of the availability and quality of verification data. In general, impacts of ~0.20 m s^{-1} approach the 95% confidence limit; however, the 12-h rawinsonde data impact at 700 mb of only 0.15 m s^{-1} is significant at the 95% confidence limit. Note that most of the aircraft data impacts for forecasts out to 6-h are highly significant.

Figure 5 shows the rawinsonde and aircraft impact for 3- and 12-h temperature forecasts verified over the full RUC domain. As with wind, the aircraft (rawinsonde) data appear to have the most impact for 3-h (12-h) forecasts. Figure 5 shows the impact for all data sources for 3-h relative humidity (RH) forecasts verified over the profiler domain. Although only the rawinsonde measures RH above the surface, RH forecasts can be altered by denial of other data types because of advection and other interactions within the model. Predictably, the rawinsonde EXP shows the most impact, but a surprising result is the slight negative impact from aircraft data in the 500-400 hPa layer. In addition, there appears to be a relatively modest impact from profilers and VAD winds.

Verification was also performed using profiler data for verifying observations. This provides verification 8 times a day (00, 03, …21 UTC) instead of just two times per day with rawinsonde verification. The diurnal variation of rawinsonde, profiler, and aircraft 500-hPa 3-h wind forecast impact is shown in Fig. 6. The duplicity and sampling frequency of the different data platforms creates an interesting result that is discussed in the conclusion section.

Finally, as an interesting aside, Table 2 shows the difference between the rawinsonde and profiler CNTL verifications. Here we matched the 0000 and 1200 UTC profiler verification with that from the rawinsonde verification in the profiler domain, both for the CNTL experiment only. The residuals are all far greater for the rawinsonde verification, with differences increasing with altitude. Since the same set of forecasts (CNTL) was verified, the only difference is in the quality of the verifying observations. Thus, Table 2 indicates that the observation wind error for rawinsondes is much larger than for profilers. The increase of observation error for rawinsonde winds is well known and a result of decreasing elevation angle and imprecision with tracking.

4. DISCUSSION

The results presented here raise many interesting questions, not only about the characteristics of each observing system (sampling/reporting frequency, instrument functional precision, and accuracy) but also about the verification in general.

Impacts from denying the various data sources generally increased over the profiler domain, especially for the profiler EXP. This may be attributed to the location of this subdomain relatively far from the fixed boundary conditions, allowing more variation in forecasts from different combinations of observations. The larger impact from profiler data within the profiler domain is predictable because profiler observations are located only in this region. The profiler has a greater impact in the lower troposphere than the rawinsonde or aircraft observations because there are more of them. In the upper troposphere, even within the profiler domain, aircraft data, because of their volume and accuracy, dominate the impact results. Likewise, over the RUC domain, aircraft data dominate the impact everywhere because they are the most numerous in space and time.

Although these results might seem trivial, it has important implications. It appears that aircraft and profiler data are highly complementary in the profiler domain. Although we did not stratify the results by time of day over the entire troposphere, Koch et al. (2004) show that at night, when aircraft observations are not as numerous as during the day, the profiler data also have more of an impact than the ACARS data at flight levels. It seems plausible that an expansion of the profiler network to a national scale network would provide increased accuracy of short-range wind forecasts over the entire troposphere.

For winds, the Student-t significance tests with rawinsonde verification (every 12 h) over the RUC domain (Table 1) indicated that aircraft, profiler, and rawinsonde impact was significant at or over the 90th percent confidence limit. For this verification against rawinsonde data, aircraft data appeared to have the greatest impact at 3 h and 6h , and rawinsonde data had the largest impact at 12 h.

Table 2 shows that the CNTL wind forecast residuals were smaller for the verification using profiler data, with differences between the rawinsonde and profiler verifications increasing with height. This is likely a result of rawinsonde drift (i.e., the rawinsonde is displaced downstream by the wind) and inaccuracies aloft in tracking the balloon as it disappears below the horizon.

The rawinsonde data were shown to have the most impact at 12-h. The 3-h impact shown in Figs. 1 and 2 is misleadingly small because it is based on twice a day rawinsonde data as the verifying observation. As a result, 3-h forecasts that verify at 0000 and 1200 UTC come from RUC forecasts initialied at 2100 and 0900 UTC that are 9 h removed from the last rawinsonde observation. For the other platforms, observations are continuously input to the RUC hourly assimilation cycle. However, when the EXPs were verified against 3-h profiler data as in Fig. 5, the 3-h rawinsonde impact valid at 03 and 15 UTC was competitive with the other platforms.

Observation sensitivity experiment verification results need to be examined carefully keeping in mind the spatial and temporal frequency of the observations, the measurement characteristics of the various platforms, and the verification data.

5. REFERENCES

Benjamin, S.G., G. Grell, J. Brown, T. Smirnova, and R. Bleck, 2003a: Mesoscale weather prediction with the RUC hybrid isentropic/terrain following coordinate model. *Mon. Wea. Rev*.,**131**, (in press).

Benjamin, S.G., D. Devenyi, S. Weygandt, K. Brundage, J. Brown, G. Grell, D. Kim, B. Schwartz, T. Smirnova, T.L. Smith, and G.S. Manikin, 2003b: An hourly assimilation/forecast cycle: The RUC. *Mon Wea Rev.,* **131**, (in press).

Benjamin, S.G. B.E. Schwartz, E.J. Szoke, and S.E. Koch, 2004: The value of wind profiler data in U.S. weather forecasting. *Bull. Amer. Meteor. Soc.,* **85,** submitted.

Koch, S.E., S.G. Benjamin, B.E. Schwartz, and E.J. Szoke: The value of wind profiler data in U.S. weather forecasting. This preprint volume, paper 6.1

Table 1. Wind forecast mean rawinsonde station residual difference (EXP-CNTL) and the Student- t significance value (in parentheses) for each denial experiment. Results for 3- 6-, and 12-h forecasts during the period 4 -16 Feb 2001 test period verified over the RUC domain. Number of residuals is ~2000.

	Rawinsonde			Profiler			Aircraft			VAD		
P(hPa)												
proj	3	6	12	3	6	12	3	6	12	3	6	12
850	.04	.01	.14 (90)	.04	.01	.04	,06	.07	.02	.05	.00	.01
700	.10 (85)	.12 (85)	.15 (95)	.14 (95)	.07	.04	.20 (99)	.17 (95)	.03	.07	.04	.03
500	.09 (80)	.15 (90)	.17 (90)	.11 (80)	.09	.06	.26 (99)	.24 (99)	.05	.04	.07	.04
400	.17 (90)	.17 (90)	.30 (99)	.19 (95)	.10	.13 (80)	.41 (99)	.32 (99)	.10	.12 (80)	.09	.11
300	.22 (90)	,23 (90)	.16 (80)	.17 (85)	.16 (85)	.07	.51 (99)	.44 (99)	.16 (80)	.15 (80)	.17 (85)	.06
250	.26 (95)	.18 (85)	.16 (80)	.21 (85)	.14 (80)	.11	.59 (99)	.50 (99)	.16 (80)	.20 (85)	.12	.12
200	.21 (85)	.16 (80)	.13	.17 (80)	.10	.08	.38 (99)	.42 (99)	.13	.16 (80)	.13	.10

Table 2. Differences (ms^{-1}) between verifications using rawinsonde observations vs. profiler observations. For vector wind forecasts for the 4-16 Feb 2001 test period. Results for 00 UTC and 12 UTC combined.

pressure	**3h forecasts**	**6h forecasts**	**9h forecasts**	**12h forecasts**
850	**0.42**	**0.34**	**0.17**	**0.23**
700	**0.50**	**0.25**	**0.21**	**0.18**
500	**0.36**	**0.43**	**0.52**	**0.33**
400	**0.62**	**0.57**	**0.53**	**0.34**
300	**1.17**	**1.09**	**1.02**	**0.82**
250	**1.53**	**1.47**	**1.49**	**1.35**
200	**2.30**	**1.92**	**1.65**	**1.59**
150	**2.22**	**1.66**	**1.88**	**2.02**

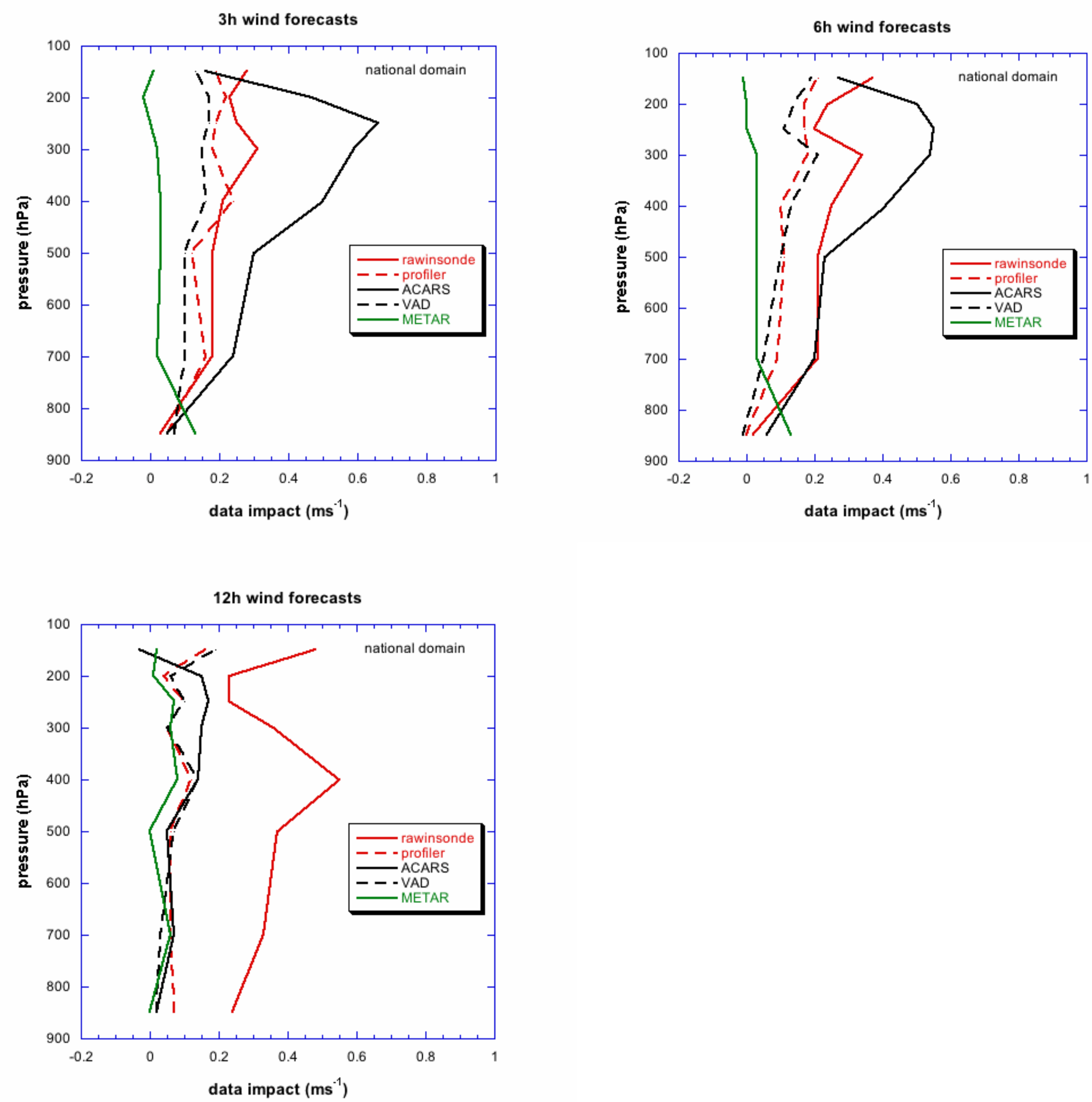


Figure 1: Data impact (EXP-CNTL) for 3-, 6-, and 12-h forecasts verified over the full RUC domain.

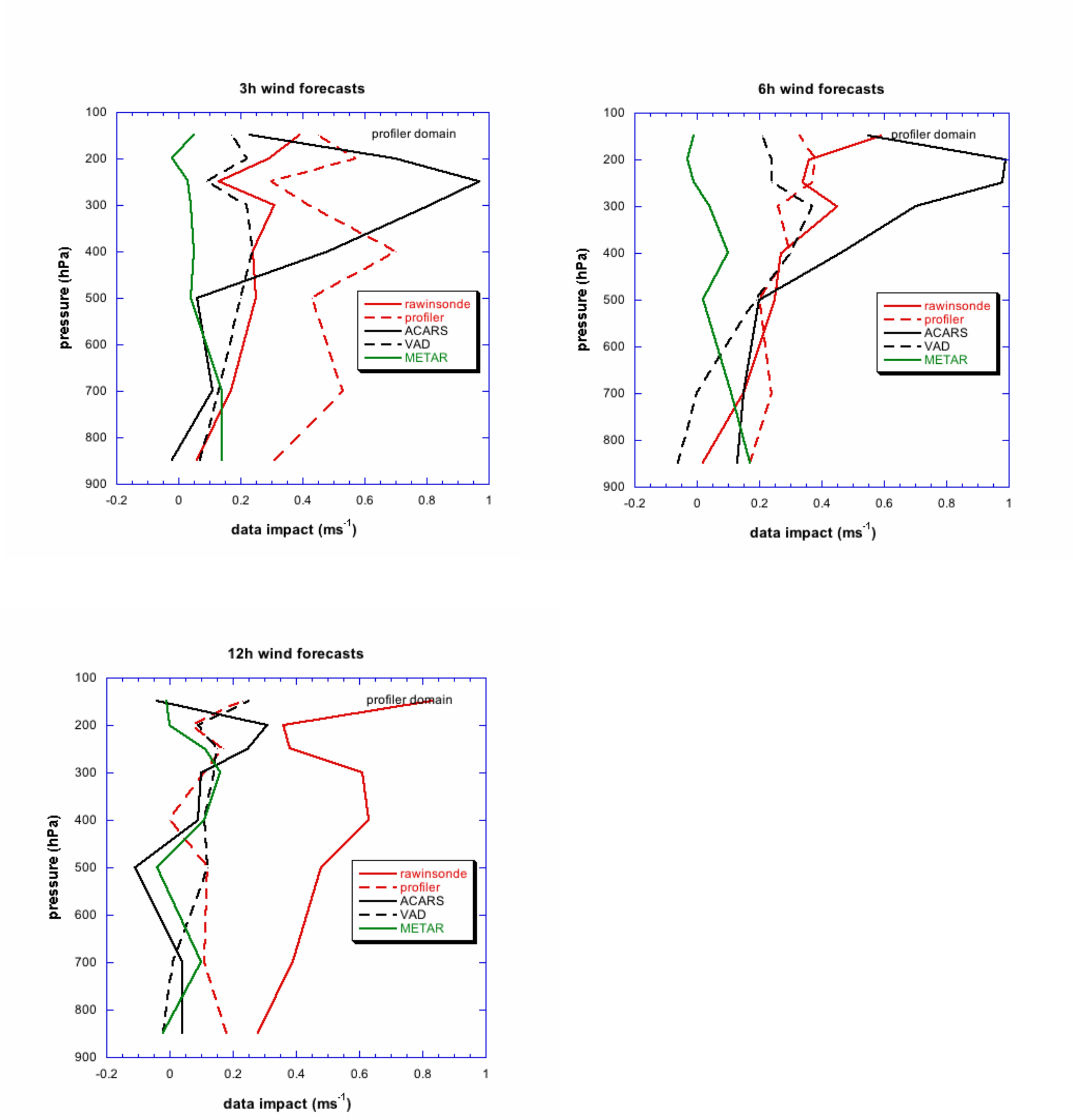
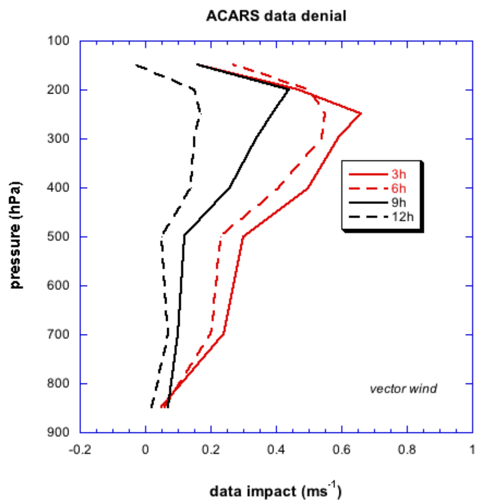
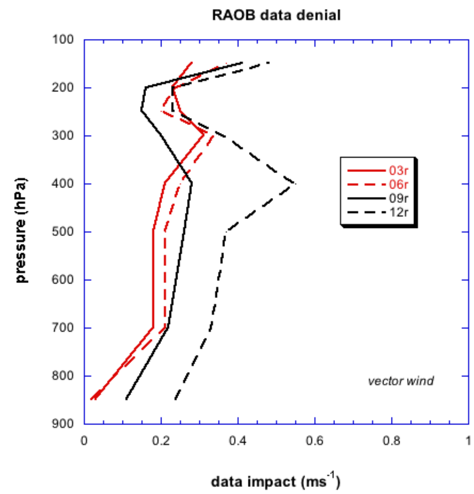
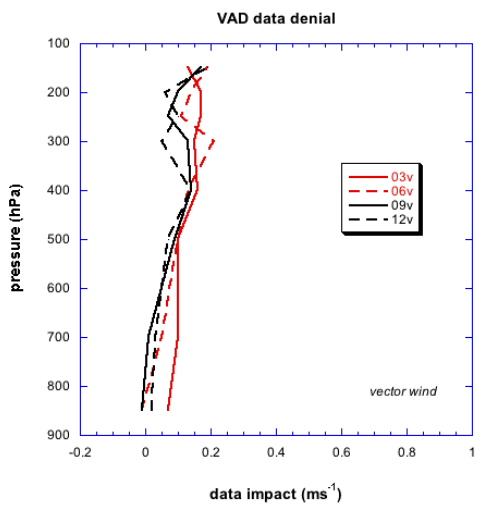
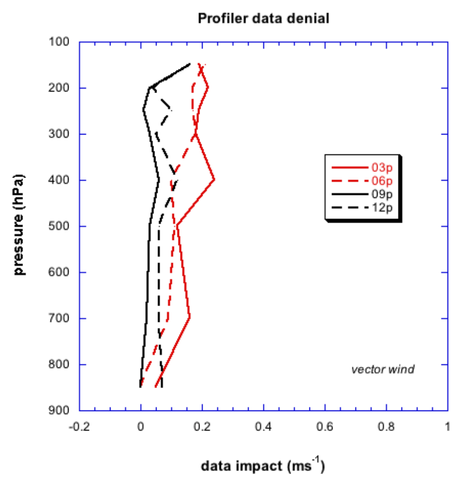


Figure 2. Same as Fig. 1, but for verification over the smaller profiler domain.









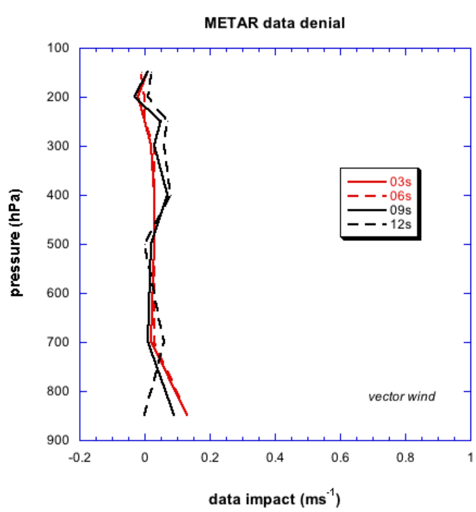


Figure 3. Wind forecast data impact for each EXP denial experiment over full RUC domain.

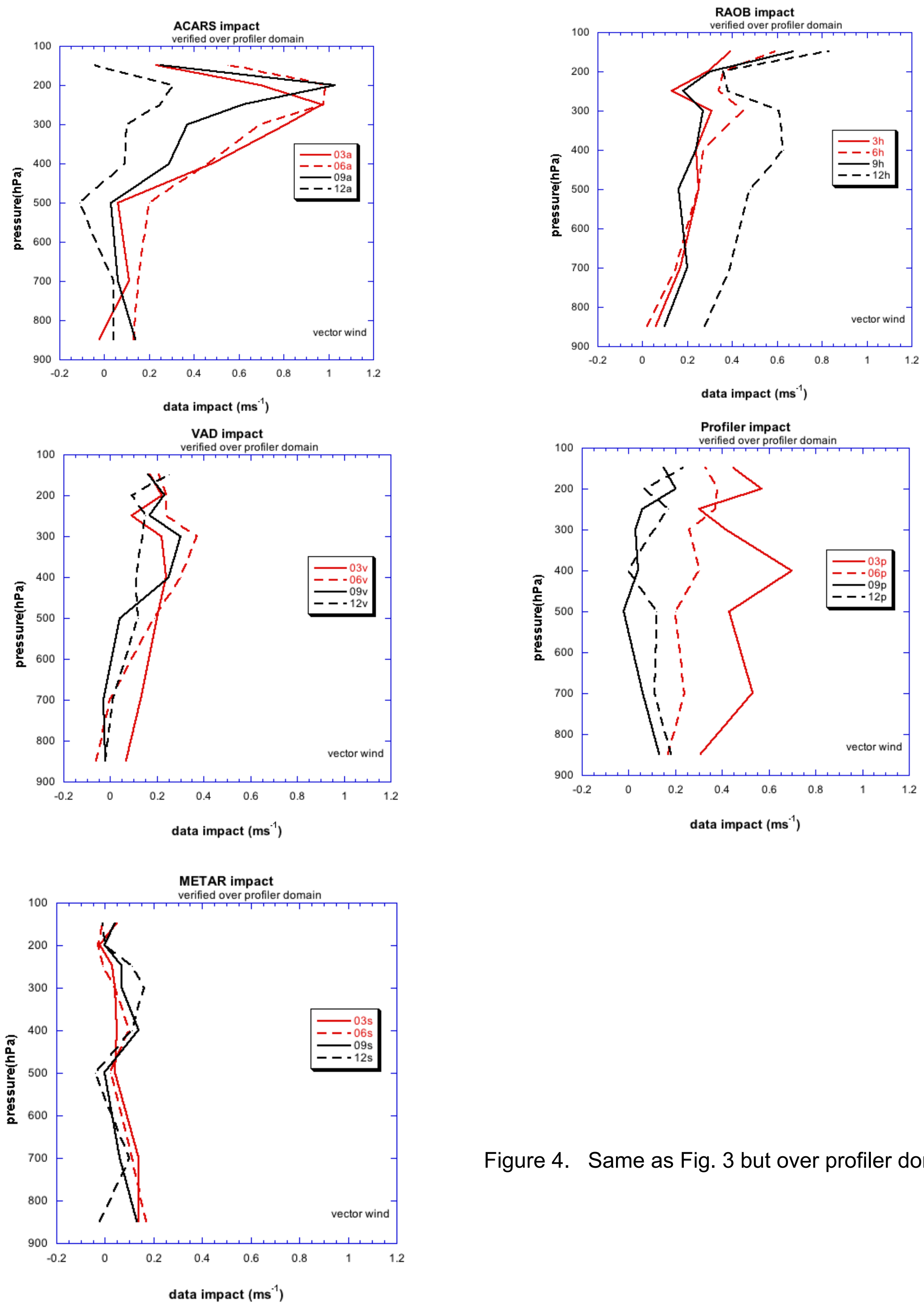


Figure 4. Same as Fig. 3 but over profiler domain.

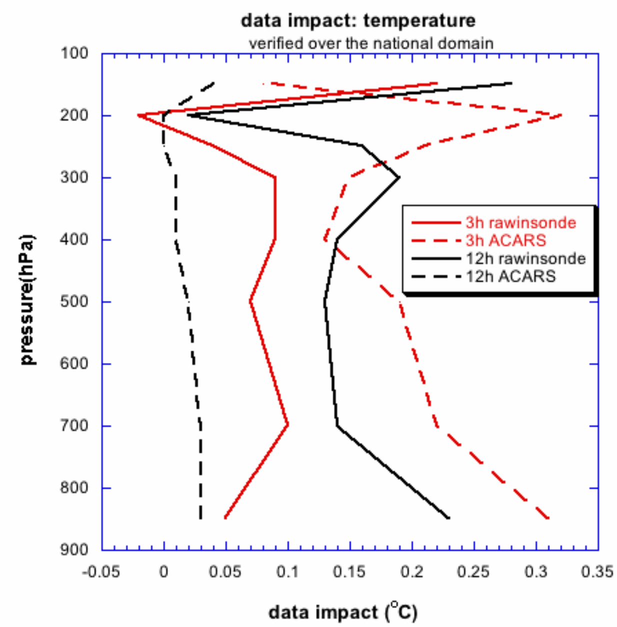


Fig. 5. Impact on 3 and 12-h temperature forecasts from aircraft and rawinsonde observations verified over the RUC domain.

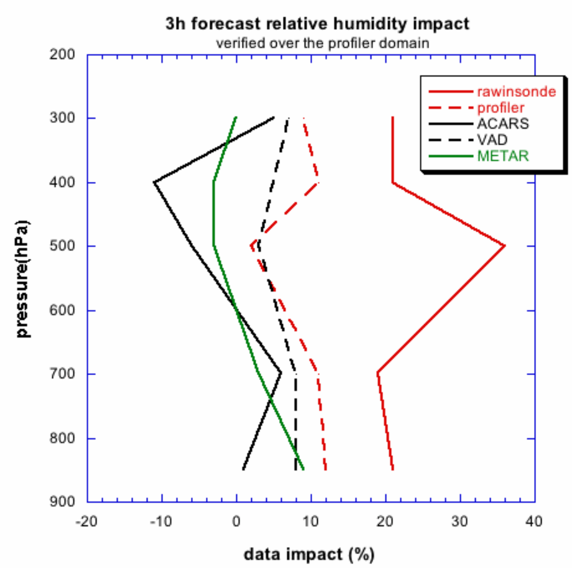


Fig. 6. Impact on 3-h relative humidity forecasts of all EXP experiments verified over the profiler domain.

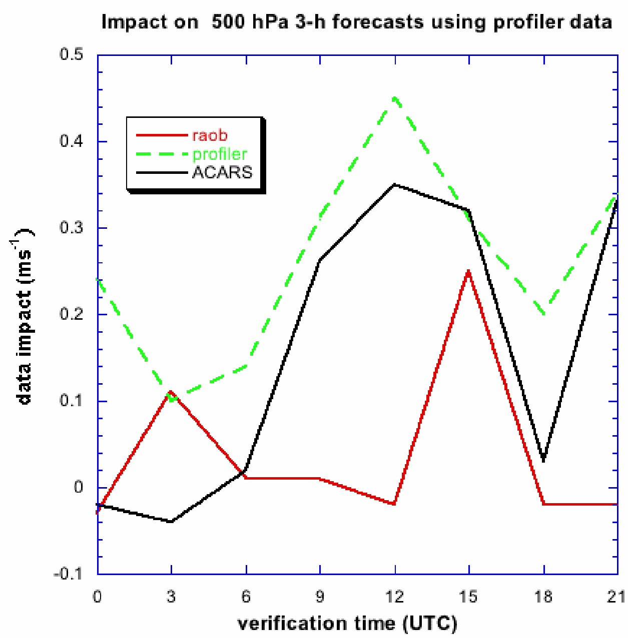


Fig. 7. Diurnal variation of impact of rawinsonde, profiler, and aircraft data denial verified against profiler data at 500 hPa.