An Evaluation of *GOES-8* Retrievals

P. ANIL RAO AND HENRY E. FUELBERG

Department of Meteorology, The Florida State University, Tallahassee, Florida

(Manuscript received 15 September 1997, in final form 31 March 1998)

ABSTRACT

The Geostationary Operational Environmental Satellite (*GOES-8*) temperature–moisture retrievals were compared with collocated National Weather Service radiosonde observations (RAOBs) to assess retrieval performance. Retrieved values of temperature and dewpoint were evaluated at individual levels. Precipitable water and thickness also were evaluated, and the *GOES-8* retrievals were compared with the first-guess data used in the algorithm. The dataset consisted of 1113 RAOB–retrieval pairs (collocated to within 50 km) over the United States at 1200 UTC during August–November 1995.

GOES-8 temperature retrievals were found to agree better with their RAOB-derived counterparts than did the dewpoints. However, both temperatures and dewpoints were found to be highly dependent on their first-guess data from the Nested Grid Model. Retrievals generally were closer to the RAOBs than was the first guess. However, this was never guaranteed, even for large first-guess discrepancies. In fact, some retrievals did not agree as well with the RAOBs as did the first guess.

GOES-8 and RAOB-derived precipitable water (PW) and thickness showed closer agreement than the level-specific data. Both integrated parameters were dependent on their first guess. However, *GOES-8* and RAOB PW agreed more often in layers above the surface where the guess was less accurate.

Comparison with a previous evaluation of retrievals from the Visible and Infrared Spin Scan Radiometer Atmospheric Sounder (VAS) indicated that *GOES-8* retrievals agreed better with RAOBs than did the VAS versions. This improvement is likely due to *GOES-8*'s increased number of channels and better signal-to-noise values, along with the assumed increase in quality of the first-guess data being used.

1. Introduction

A new series of Geostationary Operational Environmental Satellites (*GOES 8–10*) was launched beginning in April 1994. These satellites replace an older series of geostationary satellites launched in the early 1980s, which contained the Visible and Infrared Spin Scan Radiometer Atmospheric Sounder (VAS) instrument (Menzel et al. 1981). *GOES-8–10* contain separate imager and sounder instruments. The sounder consists of 18 infrared channels that are similar to those of HIRS2 (Susskind et al. 1984) on board the National Oceanic and Atmospheric Administration (NOAA) polar-orbiting satellites. Since these channels sense different layers of the atmosphere, their radiances can be used to obtain vertical profiles of temperature and moisture (retrievals). The horizontal and temporal resolution of GOES retrievals typically is superior to that of National Weather Service (NWS) radiosonde observations (RAOBs) (~50 versus ~250 km and ~1 versus 12 h). Thus, it is hoped that the retrievals will provide information about meteorological features that RAOBs cannot resolve.

Corresponding author address: P. Anil Rao, Department of Meteorology, The Florida State University, Tallahassee, FL 32306-4520.
E-mail: rao@met.fsu.edu

Menzel and Purdom (1994) provide a detailed description of the new GOES instruments and compare them with VAS. Briefly stated, the 18 sounder channels include five longwave infrared bands (VAS contains five), three infrared window bands (VAS includes two), three midtropospheric water vapor bands (VAS contains two), three shortwave infrared bands (VAS includes two), three shortwave window bands (VAS contains one), and an ozone band (VAS has none). It is hoped that this increased number of channels, together with superior horizontal resolution and reduced noise, will make *GOES-8–10* retrievals more accurate than their VAS-derived counterparts.

Both statistical and physical procedures have been used to transform VAS radiance data into retrievals. Lee et al. (1983) described a statistical regression procedure that incorporated surface observations with VAS radiances to determine temperature and moisture profiles. Smith et al. (1985) and Hayden (1988) described a simultaneous physical algorithm for obtaining operational retrievals from VAS radiances and a first-guess temperature–dewpoint profile. VAS retrievals have been used for a variety of purposes, such as locating regions that are conducive to convective storm development (Ellrod 1989; Fuelberg et al. 1995).

Several studies have evaluated VAS retrievals (Velden et al. 1984; Hayden 1988; Franklin et al. 1989).

Fuelberg and Olson (1991) compared VAS retrievals against collocated NWS RAOBs. They found that VAS-derived temperatures compared well with their RAOB counterparts, but moisture showed much less agreement. The retrievals were found to be highly dependent on the first-guess data used in the algorithm.

Since *GOES-8–10* were launched only recently, there have been few papers describing their use. However, descriptions of *GOES-8/9* product imagery and some quantitative *GOES-8/9* products can be found in Hayden et al. (1996) and Gray et al. (1996), respectively. Ellrod (1998) is using GOES data to produce a forecasting product for microbursts. Furthermore, Rao and Fuelberg (1997) used *GOES-8* radiances to determine areas of convective instability. Although Schmit (1996) compared some *GOES-8* moisture retrievals with RAOBs, a comprehensive evaluation of GOES-8 temperature and moisture retrievals has not yet been reported in the formal literature. That is the goal of this paper. Specifically, we assess *GOES-8* retrievals by comparing them with collocated radiosonde observations and with their first-guess input data. Our data and methodologies are described in section 2. Section 3 discusses comparisons of GOES–RAOB temperature, dewpoint, precipitable water, and thickness, along with their relationship to the first guess. Finally, section 4 presents conclusions.

2. Data and methodologies

Our data consisted of NWS RAOB releases and an experimental set of *GOES-8* retrievals during August–November 1995. The RAOBs provide temperature and dewpoint as functions of pressure at both mandatory and significant levels. They were obtained from The Florida State University data archive. Each RAOB was checked for data quality and required to meet several specifications for consistency: 1) temperatures must be reported to at least 100 hPa, 2) dewpoints must be reported to at least 300 hPa, 3) at least 30 data levels must be included, and 4) dewpoint depressions must be less than 30° at all levels. This final requirement was imposed because of the uncertainty associated with radiosonde reports containing very low humidity.

The experimental set of *GOES-8* retrievals was prepared at the Cooperative Institute for Meteorological Satellite Studies (CIMSS) in Madison, Wisconsin. The retrieval algorithm is very similar to the VAS algorithm described by Hayden (1988). Denoted version 1.5, it is a linearized, simultaneous solution for perturbations to first-guess temperature and moisture profiles and an estimated skin temperature. The first-guess data consisted of 12-h forecasts from the Nested Grid Model (NGM) (VAS retrievals were based on the best first guess available at that time, including the NGM and Limited-Area Fine Mesh model). Surface temperatures were incorporated into the NGM-derived first-guess data. This was done by ''blending'' the surface value into the NGM profile up to 850 hPa (Hayden 1988). Thus, our use of ''first guess'' refers to this ''adjusted'' first guess. The first-guess skin temperature was estimated using the GOES infrared window measurements.

To reduce the effects of random errors, the *GOES-8* brightness temperatures were averaged over a number of fields of view (FOVs) before being input to the retrieval algorithm. These averaging areas were geographically fixed boxes, that is, not the ''clustering'' approach described by Hillger and Purdom (1990) and Fuelberg et al. (1995). Based on *GOES-8*'s signal-to-noise values, the maximum averaging areas were 25 (5×5) FOVs, approximately 2500 km^2. Conversely, a minimum of nine clear FOVs was required to make a retrieval. In contrast, the VAS instrument provided poorer horizontal resolution (8 or 16 km) and inferior signal-to-noise values. Therefore, its average brightness temperatures were based on a maximum of 121 (11×11) FOVs, an area of approximately 7700 km^2 for 8-km FOVs.

Physical schemes for retrieving temperature and moisture profiles usually apply an empirical bias correction to the observed radiances to account for uncertainties in the transmittances and radiative transfer calculations. The calculation of bias adjustments was updated hourly and followed the method of Fleming et al. (1991) and Uddstrom and McMillin (1994). This hourly bias adjustment was necessary because of an artificial east–west radiance gradient induced by scan mirror deficiencies (Weinreb et al. 1996). This problem has since been corrected.

A nonlinear version of the algorithm (version 1.6) that we evaluate here was placed into operational use in February 1997. Correction of the artificial east–west radiance gradient described above, plus its more advanced physics and numerics, imply that the nonlinear retrievals will show improved results over those described here. Nonetheless, the current retrieval evaluation, especially our in-depth examination of their relation to the first guess, will provide new and useful information. Furthermore, our results will aid in future evaluations of the current operational retrievals.

Our evaluation methodology is patterned after Fuelberg and Olson (1991) so that readers can easily compare current results with those from VAS. Each RAOB was paired with its closest *GOES-8* retrieval. Retrievals at 1200 and 0000 UTC were evaluated separately, and results were generally similar. However, for the sake of brevity, and for consistency with Fuelberg and Olson (1991), only results for 1200 UTC are described here. To minimize retrieval–RAOB discrepancies due to horizontal gradients, the pairings were required to be within 50 km. Since radiosondes typically were released near 1115 UTC, there is a 30–60-min temporal discrepancy, depending on location. There were 1113 pairings at 1200 UTC during the 4-month period of study.

Discrepancies between RAOBs and retrievals are due to two categories of factors: 1) differences due to measurement techniques and 2) differences due to limitations in our evaluation procedures and errors in the sonde data used as the standard. Discrepancies due to measurement techniques are caused by the inherently different ways that satellites and RAOBs sample the atmosphere. Radiosondes are direct sensors, providing point-source measurements. On the other hand, satellites provide volume-averaged data, that is, each channel measures the brightness temperature within a layer and over a horizontal area (FOV) of the atmosphere. Furthermore, since radiances are averaged over an array of FOVs (Menzel and Purdom 1994) to reduce random error, this ''smoothing'' can increase the difference between the retrieval and RAOB at a specific point. The cloud clearing procedure used to remove cloud-contaminated radiance data also is a possible source of error, as are the algorithm's various physical and numerical approximations (Hayden 1988). Finally, errors in the first-guess data can produce errors in the resulting retrievals (Fuelberg and Olson 1991; Hayden 1988).

Although we seek to quantify retrieval–RAOB differences due to measurement techniques, discrepancies also arise due to our methodologies and imperfect radiosonde ''ground truth'' data. These discrepancies cannot be removed completely. Inexact spatial and temporal collocation between the retrievals and soundings is unavoidable, although we have attempted to minimize them by requiring collocation to within 50 km and about 45 min. Even stricter collocation requirements might lead to better retrieval–RAOB comparison statistics than presented here. Inexact collocation also can lead to discrepancies between the first-guess and retrieval comparison. However, since the first guess tends to be a relatively smooth horizontal analysis, the first guess–RAOB comparison will be less sensitive to collocation distance than the retrieval–RAOB comparison. Thus, at any given point, the retrieval might be better than, worse than, or equal to the first guess depending on collocation distance. Bruce et al. (1977) and Garand (1993) give a more detailed discussion of the collocation problem. Finally, the limitations of radiosondes (i.e., data quality, missing data, equipment, and reporting practices) have been well documented (e.g., Pratt 1985; Schwartz and Doswell 1991; Schwartz and Benjamin 1995). Root-mean-square errors as large as 0.8°C and 20% have been noted for sonde-derived temperature and relative humidity, respectively.

In spite of these various limitations to our evaluation procedures, we feel that results will be useful for both research and operational users of *GOES-8* retrievals since we wish to present a simple and objective evaluation of them.

3. RAOB comparison and relation to first guess

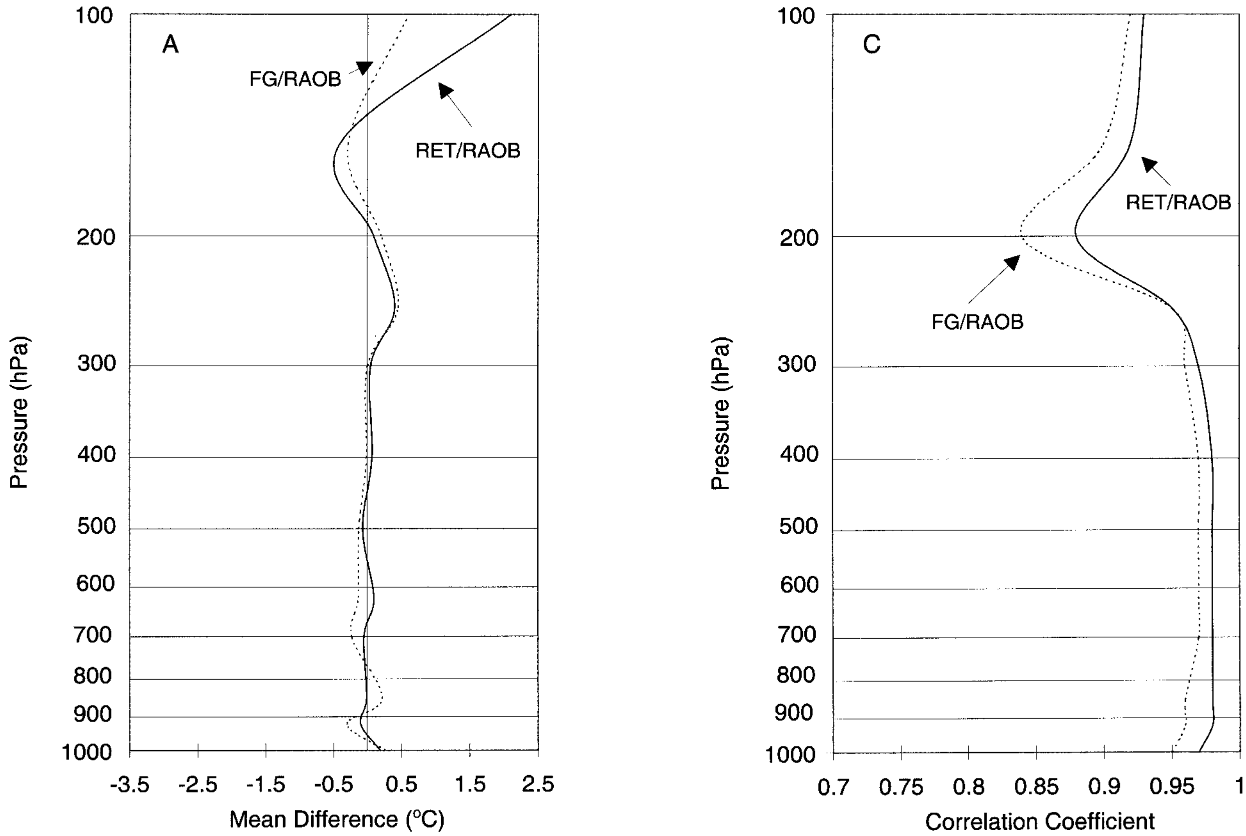
a. Temperature

Solid lines in Fig. 1 show mean temperature differences (retrieval minus RAOB, °C), standard deviations of the temperature differences (°C), and correlation coefficients of temperature for the 1113 retrieval–RAOB pairs at 1200 UTC. Mean differences between the retrievals and RAOBs (Fig. 1a) are near zero from the surface to near 300 hPa. Above this level, values are between $-0.5°$ and 0.5°C below 125 hPa. These differences are attributable to an inability of *GOES-8* retrievals to properly resolve the altitude of the tropopause and its associated change in lapse rate. Standard deviations of retrieval–RAOB temperature differences (Fig. 1b) are approximately 1.2°C from the surface to 250 hPa, exhibiting little change with height. Above this level the values increase slightly due to the limited resolution of the tropopause, as discussed earlier. The profile of correlation coefficients (Fig. 1c) is consistent with that of the previous two statistics, that is, large values (>0.97) below 250 hPa, with decreased agreement above.

Figure 2 contains scatterplots of *GOES-8* versus RAOB-retrieved temperatures. Levels of relatively good (700 hPa, Fig. 2a) and relatively poor (200 hPa, Fig. 2b) agreement are shown. These figures also include the best-fit linear regression line obtained from the least squares method. At 700 hPa, the two versions of temperature exhibit a strong linear relationship with very little scatter, and this is evident in the correlation coefficient of 0.98. The best-fit linear regression line has a slope of 1.0, indicating no differential bias. The *GOES-8* temperatures at 200 hPa (Fig. 2b) exhibit somewhat less agreement with the RAOBs. There is a good linear fit but more scatter, as reflected in the lower correlation coefficient (0.88). The slope of the regression line is 1.1, implying only a small differential bias.

Current results for *GOES-8* temperatures (Figs. 1 and 2) are slightly better than previous results from VAS (Fuelberg and Olson 1991) at nearly all levels. In the lower troposphere, correlation coefficients for *GOES-8* temperatures usually are several percent larger than corresponding values from VAS. For example, at 700 hPa, the correlation coefficient between *GOES-8* and RAOB temperature is 0.98, while its VAS–RAOB counterpart is 0.96. Furthermore, *GOES-8* temperatures exhibit smaller mean differences than determined previously for VAS, for example, $-0.06°$ and $-0.25°$C for *GOES-8*–RAOB and VAS–RAOB, respectively. Note that the VAS evaluation was conducted during March–July, while the current study occurs in August–November. The effect of this somewhat different season is unknown.

VAS retrievals were found to depend greatly on the first-guess data that were input to the algorithm (Hayden 1988; Fuelberg and Olson 1991; Knabb and Fuelberg 1997). We next examine this issue using *GOES-8* retrievals. The dashed lines in Fig. 1 denote mean temperature differences (first guess minus RAOB, °C), standard deviations of these differences (°C), and correlation coefficients for the first-guess–RAOB pairs. Mean differences between the NGM-derived first guess and RAOBs (Fig. 1a) vary from $-0.4°$ to $+0.3°$C in the



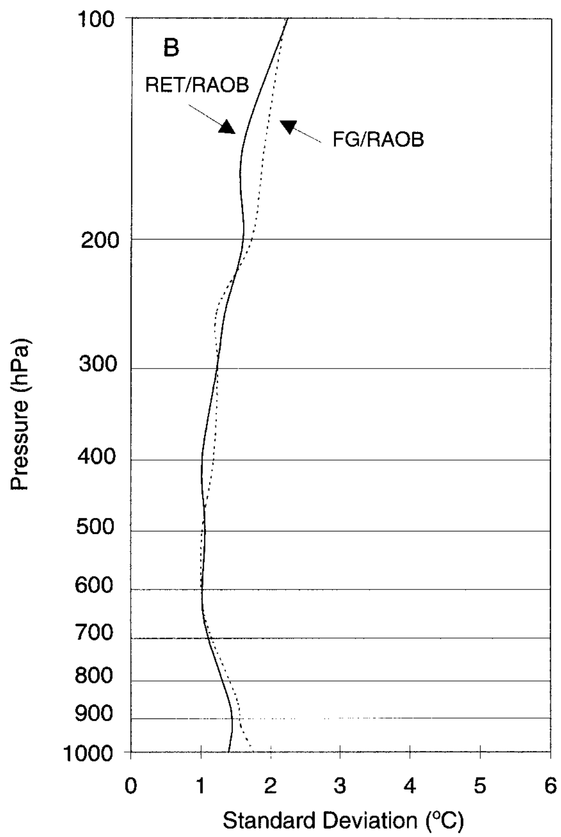
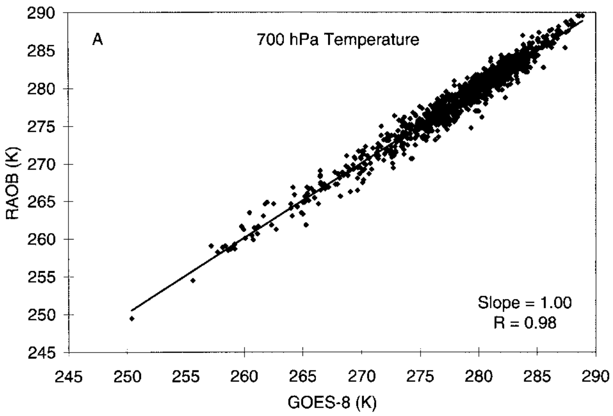


FIG. 1. (a) Mean differences (°C), (b) standard deviations of differences (°C), and (c) correlation coefficients between retrievals and RAOBs (solid) and first guesses and RAOBs (dashed) temperature. Differences are calculated as retrievals–first guess minus RAOB values.



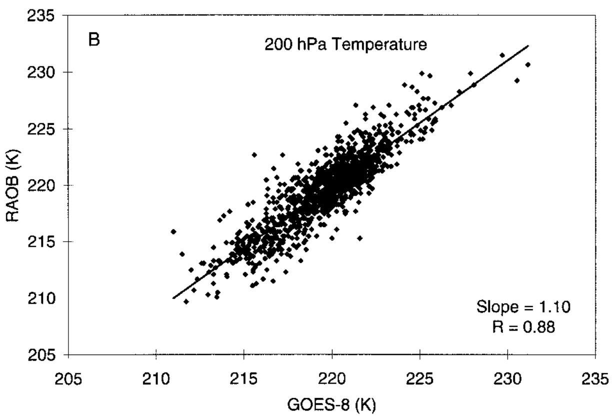


FIG. 2. Scatter diagrams of RAOB vs *GOES-8* retrievals (a) 700-hPa temperature (K) and (b) 200-hPa temperature (K).

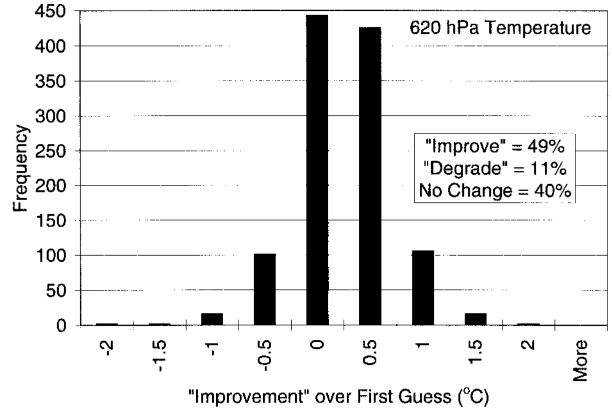


FIG. 3. Histogram of *GOES-8* improvement over first guess for 620-hPa temperature (°C). Positive values denote improvement over the first guess; negative values indicate a degradation of the guess. The term improvement (degradation) is used to indicate that a retieval is closer to (farther from) the RAOB than its first guess.

lowest 500 hPa but then are nearly zero from 500 to 300 hPa. Above this level, the values show larger positive and negative differences. Thus, like the retrievals, the first-guess data also have difficulty placing the tropopause. Standard deviations of first-guess–RAOB differences (Fig. 1b) are approximately 1.7°C near the surface. Values then decrease to near 1.0°C through the remainder of the troposphere, increasing slightly above the tropopause. Correlation coefficients (Fig. 1c) are near 0.96 from the surface to 250 hPa, with decreasing values above.

The statistics in Fig. 1 reveal close similarity between the *GOES-8*–RAOB comparisons (solid lines) and first-guess–RAOB comparisons (dashed lines). Correlation coefficients are slightly larger for the retrieval–RAOB comparisons than the first-guess–RAOB comparisons (e.g., 0.98 versus 0.96 from the surface to 250 hPa) (Fig. 1c). However, mean differences and standard deviations of differences show no consistent contrasts between the two datasets. Relations between retrievals and their first guess now are examined in greater detail.

Figure 3 is a histogram showing the number of cases in which the first-guess NGM temperatures at 620 hPa are ''improved'' or ''degraded'' by the *GOES-8* retrieval scheme. Results at 620 hPa are similar to those of other levels (not shown). We use the term ''improvement'' (''degradation'') to mean that a retrieval is closer to (farther from) the RAOB than is its first guess. This assumes that the collocated RAOB can be used as the standard for evaluation. Of course, as noted earlier, that methodology has limitations. However, based on this assumption, positive values indicate an improvement over the NGM first guess, while negative numbers mean that the retrieval is ''worse'' than its first guess. For the 620-hPa temperature (Fig. 3), the first guess is improved approximately 38% more often than it is degraded (i.e., 49% versus 11%) with virtually no change in about 40% of the cases (within ±0.25°C). Although the retrievals exhibit more improvements than degradations of the first guess, most improvements (84%) are less than 0.5°C.

We next plotted retrieval improvement–degradation of the first guess as a function of the absolute value of the first-guess discrepancy. This is patterned after Fuelberg and Olson (1991). The plot for the 620-hPa temperature is shown in Fig. 4. If the retrieval algorithm had performed flawlessly and if there were no collocation or radiosonde data errors, all points would lie on a straight line beginning at zero improvement and have a slope of 1. Inspection of Fig. 4 shows that this is not the case. The slope for the regression line is 0.25. Although small, this positive slope does indicate that larger differences in the first guess tend to be associated with greater improvements by the retrievals. On the other hand, the significant scatter is reflected in the correlation coefficient of 0.36. There are many cases in which the retrievals are worse than the first guess (negative improvement), even at relatively large values of first-guess discrepancy. In addition, there are numerous examples in which a nearly perfect first guess is degraded considerably.

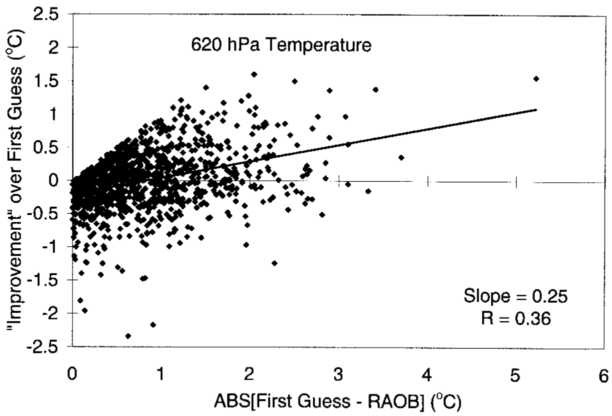


FIG. 4. Scatter diagram of *GOES-8* improvement over first guess vs the absolute value of the differences between the first guess and RAOB for 620-hPa temperature (°C).

b. Dewpoint

The evaluation of *GOES-8* dewpoint retrievals is presented similar to that of temperature in the previous section. Solid lines in Fig. 5 denote mean dewpoint differences (retrieval minus RAOB), standard deviations of these differences (°C), and correlation coefficients for the retrieval–RAOB pairs. The shapes of these lines differ greatly from those of temperature (Fig. 1, solid lines) with the satellite-derived dewpoints exhibiting considerably more disagreement with the RAOBs. *GOES-8* dewpoints are more moist (warmer) at all levels (Fig. 5a). These mean differences are less than 0.5°C from the surface to 620 hPa but increase steadily at higher levels. This may be partly due to the bias correction used to prepare this set of retrievals. Standard deviations of dewpoint differences (Fig. 5b) range from 2.5° to 5.5°C within the troposphere with maximum values between 850 and 500 hPa. Correlation coefficients (Fig. 5c) generally exceed 0.75 with values exceeding 0.90 near the surface.

Figure 6 contains scatter diagrams of *GOES-8* versus RAOB dewpoints. The dewpoints exhibit considerably more scatter than the temperatures (Fig. 2). There is less scatter at 700 hPa (Fig. 6a) than at 500 hPa (Fig. 6b) with correlation coefficients of 0.87 and 0.74, respectively. A given sonde-derived dewpoint corresponds to a large range of retrieved values. For example, at 700 (500) hPa, a RAOB dewpoint of 260 (245) K is associated with retrieved values between 245 and 270 (235 and 260) K. Both plots also show significant differential bias; slopes for the best-fit regression lines are 0.83 and 0.63 at 700 and 500 hPa, respectively. Specifically, *GOES-8* dewpoints are colder (drier) at low RAOB values and warmer (more humid) at high RAOB values. Since moisture is more variable than temperature, space–time collocation errors between the RAOBs and retrievals are more important for dewpoint than for temperature. Therefore, a greater portion of the dewpoint difference probably is due to inexact collocation.

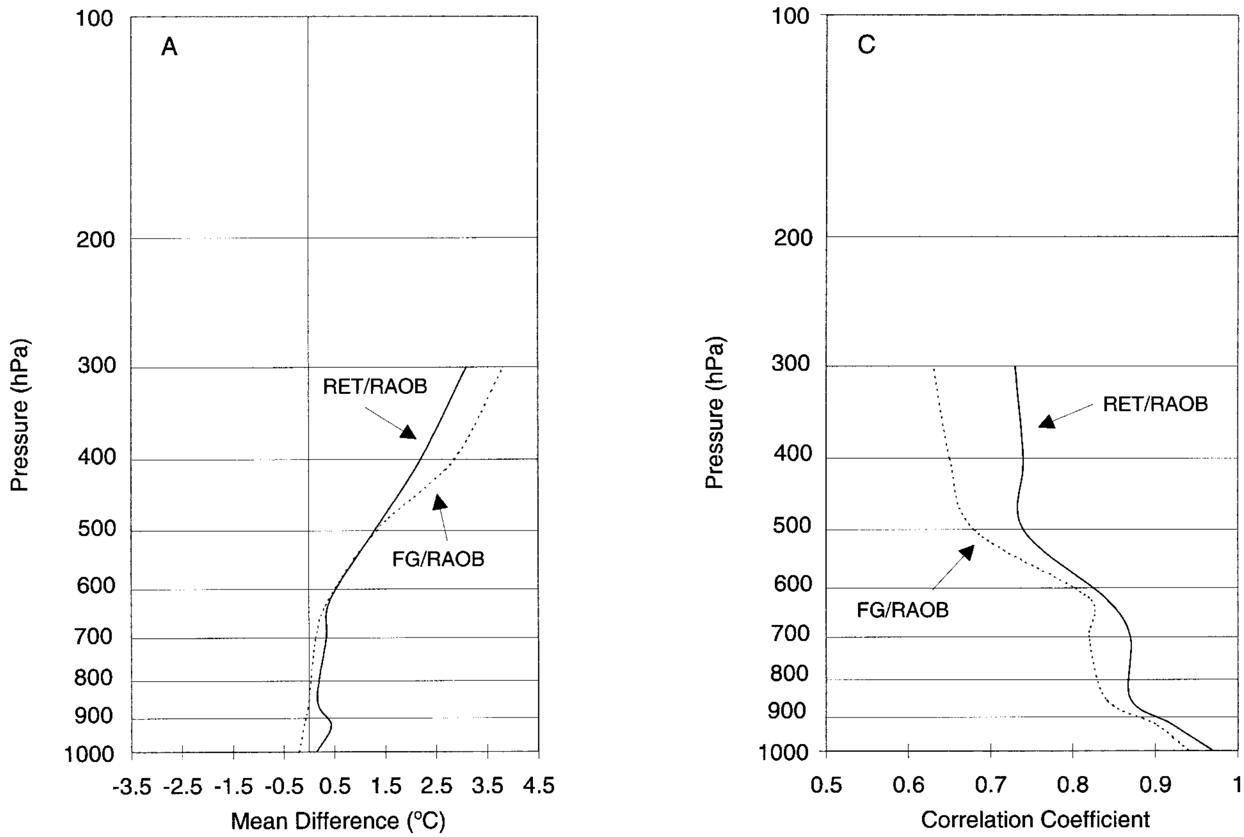
The *GOES-8* dewpoint retrievals (Figs. 5 and 6) generally are closer to the RAOBs than were their VAS-derived counterparts (Fuelberg and Olson 1991). In the lower troposphere, correlation coefficients for *GOES-8* dewpoint are usually several percent larger than corresponding values from VAS (e.g., 0.89 versus 0.81 at 700 hPa). As noted for the *GOES-8* temperature retrievals, the *GOES-8* dewpoints also exhibit smaller mean differences than observed for VAS.

First-guess dewpoints now are compared to RAOB-derived values using a method similar to the one used previously for temperature. The dashed lines in Fig. 5 denote mean first guess–RAOB dewpoint differences, standard deviations of these differences (°C), and correlation coefficients between the first-guess–RAOB pairs. Mean differences (Fig. 5a) are near zero from the surface to approximately 700 hPa but then increase steadily with height. Standard deviations of differences are largest near 500 hPa (~5.5°C). Correlation coefficients range from 0.94 near the surface to 0.61 at 300 hPa.

The first-guess–RAOB dewpoint statistics (Fig. 5, dashed lines) are not greatly different from those of the retrieval–RAOB pairs (Fig. 5, solid lines). The retrieval–RAOB pairs have larger mean differences and standard deviations of differences than their first-guess–RAOB counterparts below approximately 700 hPa. However, this relationship is reversed at higher levels. The correlation coefficient for the retrieval–RAOB comparisons exceeds that of the first-guess–RAOB pairs throughout the atmosphere with the greatest differences above 500 hPa. In summary, the three statistics in Fig. 5 indicate that retrieved dewpoints are slightly closer to the RAOBs than are the first-guess values in the upper troposphere. However, statistics for the lower levels indicate that the retrievals are not as close to the RAOBs as are the first guess. This is investigated further below.

Figure 7 shows the distribution of cases in which the first-guess NGM dewpoints at 620 hPa are improved or degraded by the *GOES-8* retrieval scheme. The terms improved and degraded are used in the same manner as described previously for temperature (section 3b), and Fig. 7 corresponds to Fig. 3 for temperature. Results show that the first guess is improved only 17% more often than it is degraded (51% versus 34%). Magnitudes of dewpoint improvement–degradation are larger than for temperature, for example, all changes to first-guess temperature are within ±2°C (Fig. 3), whereas only 62% of the changes to first-guess dewpoint fall within this interval. Some of the improvements–degradations in dewpoint exceed 10°C.

We again plotted dewpoint retrieval improvement–degradation over the first guess as a function of the absolute value of the first-guess discrepancy. The plot for 620 hPa, a representative level, is shown in Fig. 8. The relatively large amount of scatter is reflected in the correlation coefficient of 0.52. The slope for the regression line is 0.49. This positive slope indicates that



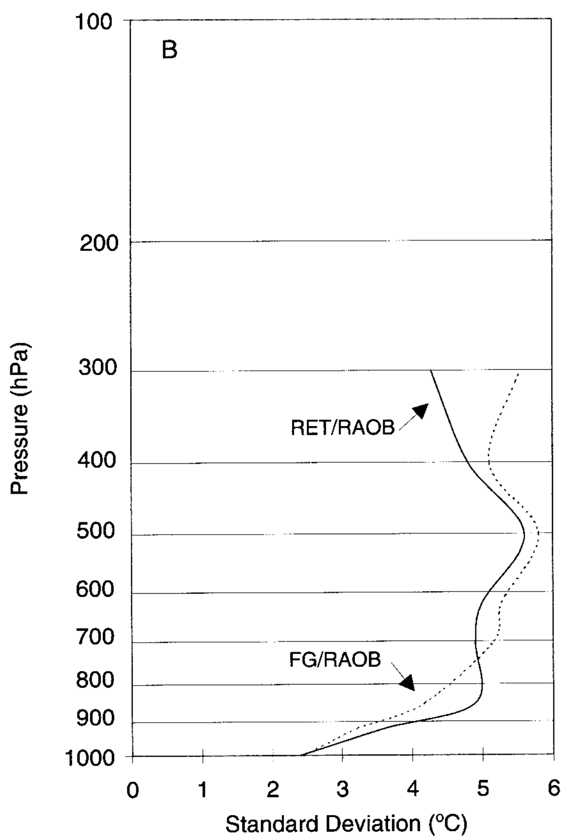


FIG. 5. (a) Mean differences (°C), (b) standard deviations of differences (°C), and (c) correlation coefficients between retrievals and RAOBs (solid) and first guesses and RAOBs (dashed) dewpoint. Differences are calculated as retrievals–first guess minus RAOB values.

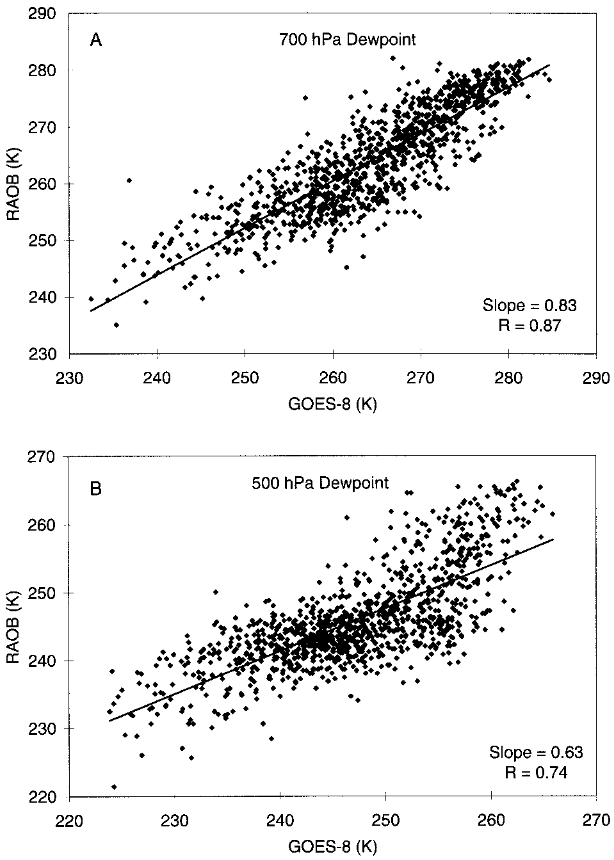


FIG. 6. Scatter diagrams of RAOB vs *GOES-8* retrievals (a) 700-hPa dewpoint (K) and (b) 500-hPa dewpoint (K).

larger differences in the first guess tend to be associated with greater improvements by the retrievals. However, there are many cases in which good first-guess values were degraded by the retrieval process. Furthermore, there are numerous cases where relatively large first-guess differences were improved only slightly or not at all.

This relationship between *GOES-8* temperature and dewpoint retrievals and their first-guess profiles is similar to that found by Fuelberg and Olson (1991) for VAS retrievals. They found that VAS retrievals were degraded about as often as they were improved; there were greater improvements for dewpoint than temperature. Their magnitudes of retrieval improvement–degradation for VAS are similar to ours for *GOES-8.* For example, their (our) plots of first-guess improvements versus first-guess discrepancies for 700 hPa (620 hPa) have slopes of 0.28 (0.25) for temperature (Fig. 4) and 0.45 (0.49) for dewpoint (Fig. 8).

c. Layer parameter comparison

Since satellites measure upwelling radiation from atmospheric layers, not specific levels, one might assume that vertically integrated *GOES-8* parameters will compare more favorably with their radionsonde-derived counterparts than will retrieval data at individual levels (Fuelberg and Olson 1991). To test this assumption, we evaluated precipitable water (PW) and thickness for several layers. The layers were chosen to include various depths and altitudes and were consistent with those selected by Fuelberg and Olson (1991) in their evaluation of VAS retrievals. These calculations utilized data for all levels within a given layer and were performed for the RAOBs, retrievals, and first-guess profiles.

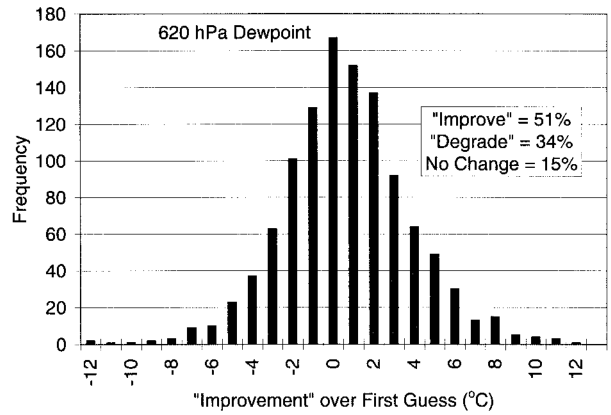


FIG. 7. Histogram of *GOES-8* improvement over first guess for 620-hPa dewpoint (°C). Positive values denote improvement over the first guess; negative values indicate a degradation of the guess.

The layers chosen for PW were the surface–300, 850–300, and 500–400 hPa. Correlation coefficients between retrievals and RAOBs in these layers are 0.96, 0.94, and 0.83, respectively. Each of these values is greater than any single-level correlation coefficient within the respective layers (Fig. 5c, solid line). This finding supports the hypothesis that vertically integrated satellite parameters compare better with their sonde-derived counterparts than do satellite data at individual levels. Furthermore, the amount of correlation is directly proportional to the thickness of the layer, that is, the deep

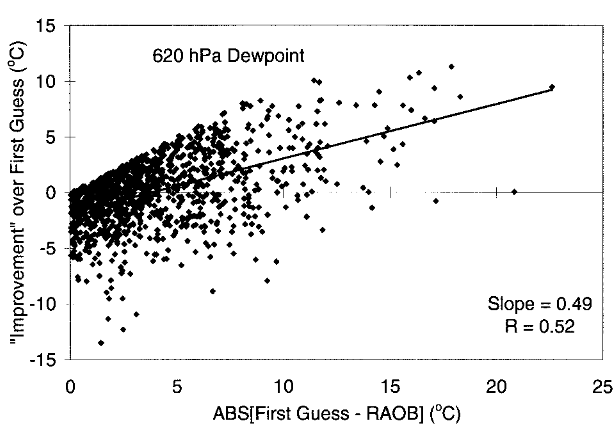


FIG. 8. Scatter diagram of *GOES-8* improvement over first guess vs the absolute value of the differences between the first guess and RAOB for 620-hPa dewpoint (°C).

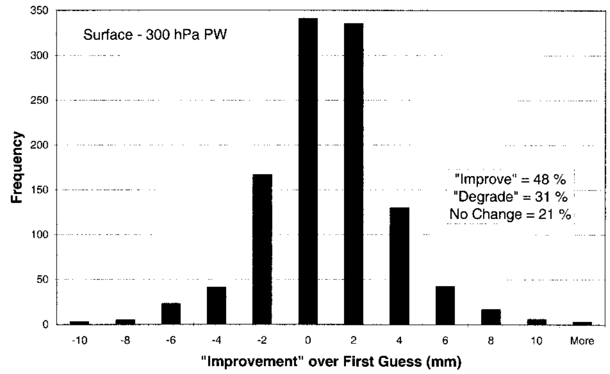


FIG. 9. Histogram of *GOES-8* improvement over first guess for surface–300-hPa PW (mm). Positive values denote improvement over the first guess; negative values indicate a degradation of the guess.

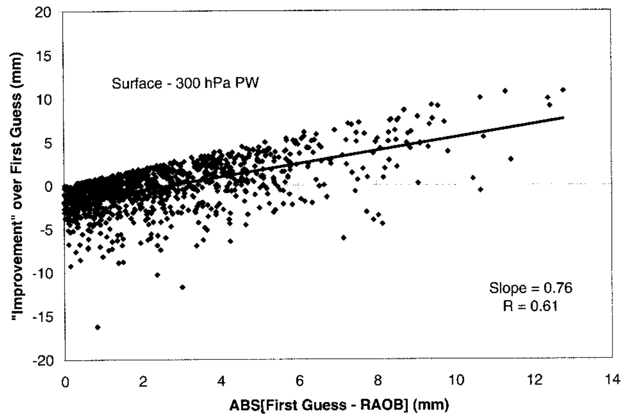


FIG. 10. Scatter diagram of *GOES-8* improvement over first guess vs the absolute value of the differences between the first guess and RAOB for surface–300-hPa PW (mm).

surface–300-hPa layer with 0.96 correlation versus the thinner 500–400-hPa layer with 0.83 correlation.

The first-guess dependence noted with the single-level data was investigated for PW. Figure 9 shows the number of degradation–improvements of the first guess for the surface–300-hPa layer. It corresponds to Figs. 3 and 7 discussed previously. The first guess is improved 48% of the time and degraded in only 31% of the cases. However, as with the single-level calculations, many of the changes are relatively small (e.g., 75% are within 2 mm of the first guess). Figure 10 is a plot of improvement over the first guess as a function of the quality of the guess, that is, the absolute value of the first guess minus RAOB. Here, the correlation coefficient and slope are 0.61 and 0.76, respectively. This indicates that the first guess generally is being improved by the retrieval. These statistics also are greater than those for any single level (e.g., Fig. 8), which is further proof of the superiority of satellite integrated parameters. Nonetheless, there still are many cases in which a good first guess is degraded considerably.

Correlation coefficients between first guess and RAOB PW for the surface–300-hPa, 850–300-hPa, and 500–400-hPa layers are 0.94, 0.93, and 0.71, respectively. Thus, differences between the RAOB and first-guess PW decrease as the layer moves away from the surface. This is expected since surface data are incorporated into the first guess. In addition, the RAOB-derived humidity data also become less reliable at the higher, colder levels. These results are consistent with the vertical profiles of first-guess–RAOB correlation for dewpoint (Fig. 5, dashed line). Finally, the difference between correlation coefficients of the two datasets (i.e., RAOBs versus retrievals and RAOBs versus first guess) also increases as the layer moves upward (e.g., 0.02 for the surface–300-hPa layer versus 0.12 for the 500–400-hPa layer). This suggests that the retrieval process improves the first-guess moisture data more in the upper layer than the lower one.

Thickness values also were calculated for three layers: 1000–500, 700–500, and 1000–850 hPa. The RAOB and retrieved versions of thickness have a near-perfect correlation (0.99) for each layer. Correlation coefficients between the first guess and RAOB data are only 1% smaller (i.e., 98%) for each layer. These results are consistent with those of the single-level data (Fig. 1c), indicating that vertical integration provides little improvement over the high-quality first-guess temperatures.

Current results for *GOES-8* are slightly better than those found previously using VAS data (Fuelberg and Olson 1991). For example, correlations between VAS- and RAOB-derived PW (thickness) ranged from 0.82 (0.96) to 0.94 (0.99) for the layers investigated. The agreement of VAS- and RAOB-derived PW also was found to be proportional to the thickness of the layer being investigated.

4. Summary and conclusions

A set of experimental *GOES-8* temperature–dewpoint retrievals from the simultaneous physical algorithm has been compared with collocated NWS RAOB releases during the period August–November 1995 over the continental United States. Relations between the retrievals and their first-guess input soundings also were investigated.

Results showed that *GOES-8* temperatures agreed closely with the RAOBs. Correlation coefficients, mean differences, and standard deviations of differences were quite good, with values for levels below 300 hPa being near 0.97, 0.0°C, and 1.2°C, respectively. The retrievals exhibited almost no differential bias at any level. The inability of the retrieval scheme to precisely locate the change in lapse rate associated with the tropopause is believed to cause the decreased agreements at higher levels.

GOES-8-derived dewpoints did not compare as well with the sonde data as did the temperatures. *GOES-8* dewpoints were colder (drier) at low RAOB values and warmer (more humid) at high RAOB values. Correlation coefficients generally were larger than 0.75, but *GOES-8* consistently indicated greater moisture above 620 hPa. This difference may be due in part to retrieval dependence on first guess and in part to the radiosonde's lack of skill in detecting upper-tropospheric humidity.

GOES-8 temperatures and dewpoints were found to be highly dependent on their first-guess data from the NGM model. The retrieval algorithm improved the first-guess data more often than not. For example, at 620 hPa, temperature (dewpoint) values were improved approximately 38% (17%) more often than degraded. Relatively large differences in the first guess generally were associated with greater improvements by the retrievals. However, improvement to the first guess was never guaranteed, even when the guess exhibited considerable difference with the RAOB. In addition, some good NGM first-guess data were degraded considerably by the retrieval algorithm.

Results for PW and thickness were superior to those of dewpoint and temperature at the individual levels. Correlation coefficients between *GOES-8*- and RAOB-derived PW were larger than all single-level correlations within the layer. *GOES-8*- and RAOB-derived thicknesses exhibited an almost perfect correlation. Nonetheless, the retrieved PW was found to be dependent on the first guess. However, this dependence decreased in layers above the surface where the first guess differed more with the RAOBs. These results support the hypothesis that satellite-derived integrated parameters compare better with their sonde-derived counterparts than do level-specific data.

Current results indicate that *GOES-8* retrievals are somewhat better than VAS retrievals (Fuelberg and Olson 1991). On the other hand, *GOES-8* retrievals continue to exhibit the strong dependence on the first-guess data that was observed with VAS.

The retrievals evaluated in this research (version 1.5) are from one of several schemes presently being examined by CIMSS. Since February 1997, operational retrievals have been based on a nonlinear algorithm (version 1.6, T. Schmit 1997, personal communication). In addition, the east–west gradient bias has been corrected. Compared to the current results, these retrievals may show improved agreement with the RAOB-derived ''ground truth'' and greater improvement over the first guess. A superior transmittance model in even later versions also may improve results. The results presented in this paper should aid in evaluating these new operational retrievals.

It should be noted that the numerical forecast models that provide the first-guess data to the retrieval algorithms are continually being improved. For retrievals to be useful over data-rich areas such as the United States, they should provide improvement over this numerical guidance. Since October 1997, the National Centers for Environmental Prediction Eta model has been assimilating three layers of *GOES-8/9* moisture information from the nonlinear (version 1.6) algorithm (W. P. Menzel 1998, personal communication). Initial results have been promising. Threat scores for 24-h forecasts of precipitation have improved. However, this early result must be validated over a much longer period of time. Compared to data-rich areas, one can assume that retrievals will provide greater improvement to numerical guidance over data-sparse regions of the globe. Nonetheless, accurate retrievals over data-rich areas (e.g., some land areas) should continue to be a high priority for achieving the correct balance in the observing system of the twenty-first century.

Acknowledgments. We extend our deep gratitude to Tim Schmit of NOAA/NESDIS for providing the *GOES-8* retrievals and for answering many questions about them. This research could not have been completed without his input. We also thank Dr. Christopher Hayden of NOAA/NESDIS for his assistance. This research was supported by NASA Grant NAG8-911 under the auspices of the Marshall Space Flight Center.

REFERENCES

Bruce, R. E., L. D. Duncan, and J. H. Pierluissi, 1977: Experimental study of the relationships between radiosonde temperatures and satellite-derived temperatures. *Mon. Wea. Rev.,* **105,** 493–496.

Ellrod, G., 1989: Environmental conditions associated with the Dallas microburst storm determined from satellite soundings. *Wea. Forecasting,* **4,** 469–484.

——, cited 1998: Experimental GOES microburst products. [Available online at http://orbit-net.nesdis.noaa.gov/ora/fpdt1/mb.html.]

Fleming, H. E., N. C. Grody, and E. J. Katz, 1991: The forward problem and corrections for the SSM/T satellite microwave temperature sounder. *IEEE Trans. Geosci. Remote Sens.,* **29,** 571–583.

Franklin, J. L., C. S. Velden, C. M. Hayden, and J. Kaplan, 1989: A comparison of VAS and ODW data around a subtropical cold low. Preprints, *Fourth Conf. on Satellite Meteorology and Oceanography,* San Diego, CA, Amer. Meteor. Soc., 141–144.

Fuelberg, H. E., and S. R. Olson, 1991: An assessment of VAS-derived retrievals and parameters used in thunderstorm forecasting. *Mon. Wea. Rev.,* **119,** 795–814.

——, P. A. Rao, and D. W. Hillger, 1995: Clustering of satellite sounding radiances to investigate intense low-level humidity gradients. *J. Appl. Meteor.,* **34,** 1525–1535.

Garand, L., 1993: A pattern recognition technique for retrieving humidity profiles from METEOSAT or GOES imagery. *J. Appl. Meteor.,* **32,** 1592–1607.

Gray, D. G., C. M. Hayden, and W. P. Menzel, 1996: Review of quantitative satellite products derived from *GOES-8/9* imager and sounder instrument data. Preprints, *Eighth Conf. on Satellite Meteorology and Oceanography,* Atlanta, GA, Amer. Meteor. Soc., 159–163.

Hayden, C. M., 1988: GOES–VAS simultaneous temperature–moisture retrieval algorithm. *J. Appl. Meteor.,* **27,** 705–733.

——, G. S. Wade, and T. J. Schmit, 1996: Derived product imagery from *GOES-8. J. Appl. Meteor.,* **35,** 153–162.

Hillger, D. W., and J. F. W. Purdom, 1990: Clustering of satellite sounding radiances to enhance mesoscale meteorological retrievals. *J. Appl. Meteor.,* **29,** 1344–1351.

Knabb, R. D., and H. E. Fuelberg, 1997: A comparison of the first-guess dependence of precipitable water estimates from three techniques using GOES data. *J. Appl. Meteor.,* **36,** 417–427.

Lee, T. H., D. Chesters, and A. Mostek, 1983: The impact of con-

ventional surface data upon VAS regression retrievals in the lower troposphere. *J. Climate Appl. Meteor.*, **22,** 1853–1874.
Menzel, W. P., and J. F. W. Purdom, 1994: Introducing GOES-I: The first of a new generation of geostationary operational environmental satellites. *Bull. Amer. Meteor. Soc.,* **75,** 757–781.
——, W. L. Smith, and L. D. Herman, 1981: Visible infrared spin-scan radiometer atmospheric sounder radiometric calibration: An inflight evaluation from intercomparisons with HIRS and radiosonde measurements. *Appl. Opt.*, **20,** 3641–3644.
Pratt, R. W., 1985: Review of radiosonde humidity and temperature errors. *J. Atmos. Oceanic Technol.*, **2,** 404–407.
Rao, P. A., and H. E. Fuelberg, 1997: Diagnosing convective instability from *GOES-8* radiances. *J. Appl. Meteor.*, **36,** 350–364.
Schmit, T. J., 1996: Sounder bias correction of the east–west radiance gradient. Preprints, *GOES-8 and Beyond,* Denver, CO, SPIE, 630–637.
Schwartz, B., and C. A. Doswell III, 1991: North American rawinsonde observations: Problems, concerns, and a call to action. *Bull. Amer. Meteor. Soc.*, **72,** 1885–1896.
——, and S. G. Benjamin, 1995: A comparison of temperature and wind measurements from ACARS-equipped aircraft and rawinsondes. *Wea. Forecasting,* **10,** 528–544.
Smith, W. L., H. M. Woolf, and A. J. Schreiner, 1985: Simultaneous retrieval of surface atmospheric parameters: A physical and analytically direct approach. *Advances in Remote Sensing,* A. Deepak, H. E. Fleming, and M. T. Chahine, Eds., A. Deepak Publishing, 221–232.
Susskind, J., J. Rosenfield, D. Reuter, and M. T. Chahine, 1984: Remote sensing of weather and climate parameters from HIRS2/MSU on TIROS-N. *J. Geophys. Res.*, **89,** 4677–4697.
Uddstrom, M. J., and L. M. McMillin, 1994: System noise in the NESDIS TOVS forward model. Part I: Specification. *J. Appl. Meteor.*, **33,** 919–938.
Velden, C. S., W. L. Smith, and M. Mayfield, 1984: Application of VAS and TOVS to tropical cyclones. *Bull. Amer. Meteor. Soc.*, **65,** 1059–1067.
Weinreb, M. P., J. X. Johnson, J. C. Bremer, E. C. Wack, and O. Chen, 1996: Algorithm to compensate for variation of reflectance of *GOES-8* and *-9* scan mirrors with scan angle. Preprints, *Eighth Conf. on Satellite Meteorology and Oceanography*, Atlanta, GA, Amer. Meteor. Soc., 110–114.