NOAA Technical Memorandum ERL WPL-173

A PRELIMINARY COMPARISON OF TEMPERATURE SOUNDINGS OBTAINED FROM SIMULTANEOUS RADIOMETRIC, RADIO-ACOUSTIC, AND RAWINSONDE MEASUREMENTS

Judith A. Schroeder

Wave Propagation Laboratory
Boulder, Colorado
September 1989

noaa NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION / Environmental Research Laboratories

NOAA Technical Memorandum ERL WPL-173

A PRELIMINARY COMPARISON OF TEMPERATURE SOUNDINGS OBTAINED FROM SIMULTANEOUS RADIOMETRIC, RADIO-ACOUSTIC, AND RAWINSONDE MEASUREMENTS

Judith A. Schroeder

Wave Propagation Laboratory
Boulder, Colorado
September 1989

UNITED STATES
DEPARTMENT OF COMMERCE

Robert A. Mosbacher
Secretary

NATIONAL OCEANIC AND
ATMOSPHERIC ADMINISTRATION

John A. Knauss
Under Secretary for Oceans
and Atmosphere/Administrator

Environmental Research
Laboratories

Joseph O. Fletcher
Director

NOTICE

For sale by the National Technical Information Service, 5285 Port Royal Road
Springfield, VA 22161

CONTENTS

A PRELIMINARY COMPARISON OF TEMPERATURE SOUNDINGS OBTAINED FROM SIMULTANEOUS RADIOMETRIC, RADIO-ACOUSTIC, AND RAWINSONDE MEASUREMENTS

Judith A. Schroeder

ABSTRACT. An experiment was conducted to compare the accuracy of atmospheric temperature soundings obtained simultaneously with a ground-based radio-acoustic sounding system (RASS), a ground-based microwave radiometric system, and a combined RASS/radiometric system at Denver, Colorado, during July and August of 1988. Applying a linear statistical retrieval technique to measurements from all three systems permitted three-way comparisons throughout the troposphere. Accuracy was evaluated by comparison with simultaneous in situ rawinsonde temperature measurements. Rms errors for the 22 sets of soundings compared ranged from 0.5 to 2.0°C, depending on altitude. Temperatures retrieved from the radiometer measurements were more accurate than those retrieved from RASS above 5 km and below 0.5 km AGL, where a bias in the low-altitude RASS measurements degraded the retrieved temperatures. Between 0.5 and 5 km the RASS temperature errors were less than 1°C. At the 750- and 700-mb pressure levels, the RASS errors matched the reported precision of rawinsonde measurements. Although the combined system errors were smaller than individual system errors at all altitudes above 0.5 km, the differences among the three systems were too small to conclude that the combined system performance was significantly better than that of either system alone.

1. INTRODUCTION

Knowledge of the three-dimensional temperature structure of the Earth's atmosphere is necessary for accurate weather forecasts, which directly impact aviation, agriculture, and human safety. Traditionally, atmospheric temperature measurements have been obtained by balloon-borne instrument packages called rawinsondes, launched twice daily from selected locations worldwide. Because these measurements are more sparse in space and time than local weather phenomena, the existing ground-based weather observation network limits the accuracy of weather forecasts.

Some improvement in spatial and temporal measurement density has been achieved with remote temperature soundings from satellite platforms. However, the accuracy of these soundings is severely degraded near the Earth's surface, particularly over land (Westwater et al., 1984). Therefore, ground-based measurements of tropospheric temperatures continue to play an important role in the global weather observation network.

Westwater et al. (1984) showed that ground-based remote temperature soundings obtained with a microwave radiometric system complement the satellite-based soundings quite effectively. Since a radiometric system can continuously provide temperature soundings encompassing the entire troposphere, a network of them could interpolate between rawinsonde observations in both space and time. However, the radiometric soundings exhibit very coarse vertical resolution.

Another technique that shows promise for obtaining continuous ground-based remote temperature soundings is the radio-acoustic sounding system (RASS). This system, comprising an acoustic source and a Doppler radar, may be capable of finer vertical resolution than the radiometric system, but altitude coverage is limited by ground clutter, acoustic attenuation, and turbulence effects (May et al., 1988a). In addition, the RASS measures virtual temperature rather than kinetic temperature, so temperature and humidity effects are difficult to separate.

One objective of this research was to investigate whether a combined RASS/radiometric system could produce more accurate temperature soundings (RASS contribution) over a wider altitude range (radiometric contribution) than either system could produce individually. A linear statistical retrieval technique, which uses a priori knowledge of local climatology to estimate temperature soundings from a given set of measurements (Strand and Westwater, 1968), offered a straightforward method for combining the measurements.

Economic considerations motivated a second objective. The costs associated with microwave radiometric systems alone jeopardize their practicality in a network context. Together, the RASS components cost as much as a radiometric system. However, Doppler radars can also measure profiles of wind speed and direction (Strauch et al., 1984). Furthermore, a network of 31 Doppler radars is being installed in the central United States to test the concept of a nationwide wind-profiling network (Chadwick and Hassel, 1987). RASS temperature sounding capability could be added to these radars for less than half the cost of adding a radiometric system. Therefore, the accuracy of temperature soundings obtained by applying the retrieval technique to RASS measurements alone was also evaluated.

These two objectives prompted a preliminary comparison between the accuracy of temperature soundings retrieved from a RASS system, a microwave radiometric system, and a combined RASS/radiometric system, colocated with a rawinsonde launch site at Stapleton International Airport in Denver, Colorado. Although the frequency of the radar involved (915 MHz) is not optimal for RASS operation (May et al., 1988a), the unique opportunity to compare simultaneous RASS, radiometric, and rawinsonde temperature soundings at the same location justified the experiment. During July and August of 1988, simultaneous RASS and radiometric measurements were made during 22 rawinsonde ascents. Despite its imperfections (Hoehne, 1980), the rawinsonde served as the comparison standard.

Four sections follow this introduction. Section 2 describes the measurements, and Section 3 explains the method used to retrieve temperature soundings from the measurements. Comparisons among temperature soundings obtained from each of the four systems are presented and discussed in Section 4. Section 5 summarizes the conclusions suggested by the comparisons.

2. MEASUREMENTS

A National Weather Service (NWS) Forecast Office and rawinsonde launch site are located at Stapleton International Airport in Denver, Colorado. In 1981, a suite of remote sensors, called the Profiler, was installed adjacent to the launch site by the Wave Propagation Laboratory (WPL) of the National Oceanic and Atmospheric Administration (NOAA). Since then, the Profiler has provided continuous soundings of wind speed and direction every hour; pressure, temperature and humidity soundings every 20 min; and column-integrated values of precipitable water vapor and cloud liquid water content every 2 min. The Profiler employs a six-channel microwave radiometric system to measure temperature and moisture, while a 915–MHz ($\lambda \approx 33$ cm) Doppler radar measures winds. Designed for continuous, unattended operation, the systems have no moving parts and are computer controlled. For a complete description of the Profiler system, see Hogg et al. (1983a).

a. Radio-acoustic Sounding System (RASS)

A vertically pointing, 50–watt acoustic source was placed near the base of the Profiler radar antenna to implement the RASS technique. The source transmits a continuous, constant-amplitude wave, whose frequency is modulated by sweeping a range of frequencies centered at about 2 kHz ($\lambda \approx 16.5$ cm) (May et al., 1989). This center frequency was chosen because an acoustic wavelength equal to half the radar wavelength is required to maximize the amount of energy scattered from the acoustic wave back to the radar (Bragg condition). Transmitting a range of frequencies ensures that the Bragg condition is met. Unfortunately, the large acoustic attenuation at this frequency severely limits the altitude range of the RASS measurements (May et al., 1988a).

The Profiler radar routinely and continuously measures the horizontal and vertical wind components at predetermined altitude levels. While operating in the RASS mode, the radar was used instead to measure the velocity of an acoustic wavefront as it propagated vertically from the acoustic source. Radar receiver adjustments required to detect typical acoustic velocities rather than typical wind velocities interrupted the radar wind measurements. This interruption could have been avoided if a second processor had been allocated to the RASS mode of operation.

Assuming ideal gas behavior, the velocity of sound in moist air can be expressed as (List, 1963),

$$c = \sqrt{\gamma p/\varrho} \ ,$$

where

c = sound velocity
p = barometric pressure
ϱ = moist air density
γ = ratio of specific heat at constant pressure to that at constant volume.

Since, by the ideal gas law (Byers, 1974),

$$p = \varrho R T_v \ ,$$

where

R = gas constant for dry air

T_v = virtual temperature,

the sound velocity becomes

$$c = \sqrt{\gamma R T_v} \ .$$

Using this relationship, virtual temperature (T_v) can be obtained directly from the radar acoustic velocity measurements. Vertical wind velocity, which degrades those measurements, can produce errors in the RASS T_v measurements.

Each set of T_v measurements used in the comparison was obtained from a consensus average (Strauch et al., 1984) of a series of six 1-min samples, beginning just prior to the rawinsonde launch times shown in Table 1. Altitude levels chosen for the T_v measurements ranged from 200 to 2900 m above ground level (AGL), separated by 150 m. However, because acoustic attenuation varies with atmospheric conditions, the maximum altitude of the RASS measurements also varied. The right-hand column in Table 1 shows the level number associated with the maximum altitude of each RASS measurement set. Measurements missing from levels below that altitude were filled by linear interpolation, so the number shown also represents the total number of levels used.

Table 1. Rawinsondes used in the comparison.

Date	Time (UTC)	Type	Highest RASS Level
July 1	20:00	CLASS	10
6	17:00	CLASS	14
6	20:00	CLASS	6
7	17:00	CLASS	15
7	19:59	CLASS	14
11	17:00	CLASS	11
11	20:04	CLASS	10
12	16:59	CLASS	10
14	17:00	CLASS	13
18	20:00	CLASS	12
19	17:00	CLASS	19
19	20:00	CLASS	13
20	11:00	NWS	13
25	11:05	NWS	8
25	17:00	CLASS	11
26	11:00	NWS	9
26	17:10	CLASS	14
28	17:00	CLASS	13
29	17:00	CLASS	15
Aug 3	17:00	CLASS	16
9	17:00	CLASS	13
17	11:00	NWS	16

b. Microwave Radiometric System

The Profiler microwave radiometric system observes radiation emitted by the atmosphere from the zenith direction, reflected by a 45-degree flat-plate reflector through a microwave window into a temperature-controlled building, which houses the radiometer antennas, detectors, and the Profiler data acquisition system. Voltages generated at the detectors are averaged for 2 min and converted to brightness temperatures (*Tb*), a unit directly related to radiance.

The six radiometer channels are located at 20.6, 31.65, 52.85, 53.85, 55.45, and 58.80 GHz. The dots along the total absorption curve in Fig. 1 identify the spectral locations of the channels with respect to the absorption lines of oxygen and water vapor.

The two lowest channels, located near the 22.235 GHz water vapor absorption line and in the neighboring window, repectively, respond to changes in water vapor and cloud liquid. Besides providing estimates of precipitable water vapor, cloud liquid water content, and humidity profiles, these channels correct the measurements between 50 and 60 GHz for the influence of clouds.

The other four channels respond primarily to temperature changes, although some sensitivity to water vapor exists in the 52.85- and 53.85-GHz channels. At 58.80 GHz, where oxygen absorption is highest, the observed emission emanates from the part of the atmosphere nearest the Earth's surface. Since emission depends on temperature, meas-

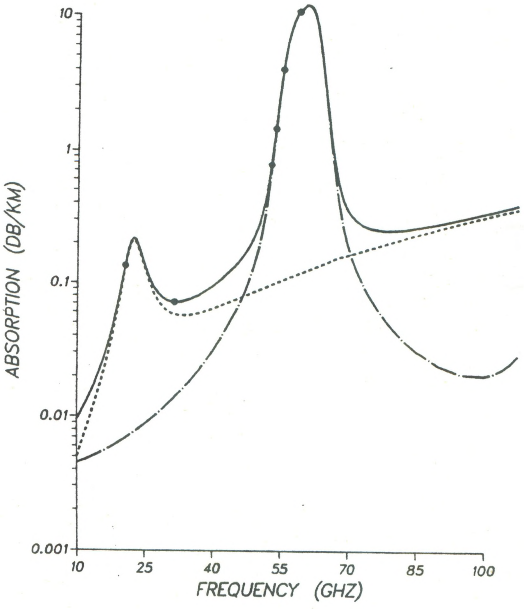


Figure 1. The log of spectral absorption for oxygen (- · -), water vapor (---), and their sum (—). Calculated from the Liebe and Layton (1987) absorption model with interference coefficients from Rosenkranz (1988), assuming average Devner summer surface conditions (pressure = 839 mb, temperature = 21°C, relative humidity = 42%.

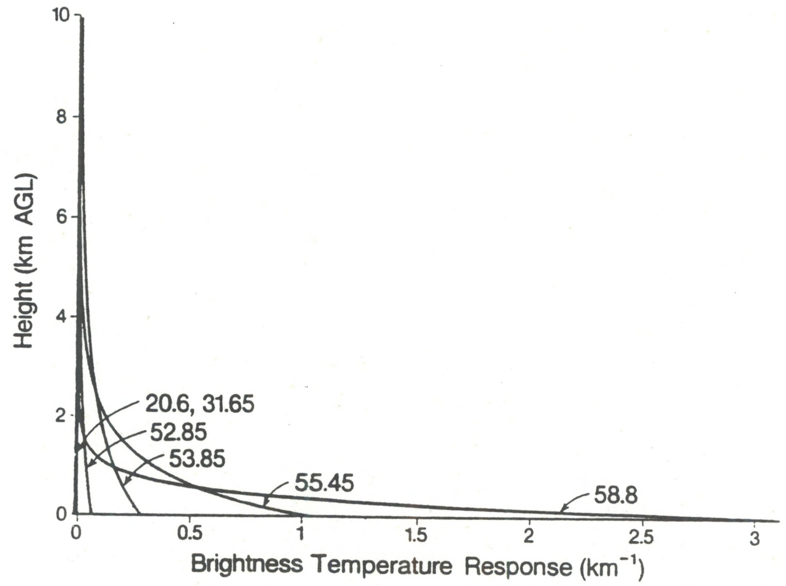


Figure 2. Temperature weighting functions for the six radiometer channels, computed from a mean Denver sounding.

urements at 58.80 GHz can be related to low-altitude temperatures. Emission from higher altitudes may be observed at lower frequencies, where absorption is less. Simultaneous-measurements at the four frequencies, therefore, provide information about temperature as a function of altitude. Figure 2 gives the change in *Tb* resulting from a 1°C change in temperature over a layer 1 km thick at the altitude indicated on the vertical axis. For example, a 1° temperature change near the surface produces a 3° change in the 58.80-GHz channel *Tb*. The figure demonstrates the differences among altitudes observed by the four channels.

Although the Profiler radiometric system provides *Tb*s every 2 min, the measurements are routinely averaged over 20 min for temperature profiling purposes. Each set of *Tb* measurements used in this study represents a 20-min average, beginning at the time a rawinsonde was launched (Table 1).

c. Rawinsonde

The rawinsonde, a balloon-borne instrument package, measures pressure, temperature, and dewpoint in situ. Two different types of rawinsonde temperature soundings were used in this experiment. The routine NWS soundings provide temperatures at predetermined ("mandatory") pressure levels and at levels where the temperature profile departs from linearity. The National Center for Atmospheric Research (NCAR) also furnished temperature soundings obtained with their Cross-chain Loran Atmospheric Sounding System (CLASS), which happened to be operating at Stapleton during this experiment (Lauritsen et al., 1987). The CLASS soundings provide temperature every 10 s after launch. Table 1 lists the dates, launch times, and types of rawinsonde soundings used in this comparison.

The NWS soundings were used as given, but some editing of CLASS soundings was necessary. Except for surface meteorology measurements made from a tower, the CLASS system is completely self-contained inside a small trailer, equipped with a chute for launching the sondes. Frequently, the first few sounding levels were recorded before the sonde exited the chute. Beginning with the first level, sounding levels were eliminated until the rate of ascent between adjacent levels became consistent with the rates aloft.

The functional precision of NWS rawinsonde temperatures measured at a given height is given by Hoehne (1980) as 0.62°C. Although an NCAR brochure reports 0.5°C for CLASS system performance, a systematic evaluation of CLASS temperature measurement precision comparable to Hoehne's has not been published to my knowledge. Despite their inherent errors, rawinsonde measurements continue to be the standard, or "ground truth," against which remote atmospheric measurements are compared and judged.

The validity of comparisons between rawinsonde and remote measurements are further undermined by the differing natures of the measurements. Rawinsondes measure temperatures at ascending levels sequentially, whereas remote sensors measure all levels simultaneously. Remote measurements characterize the volume of atmosphere directly above them, while the rawinsonde drifts with the wind from one atmospheric volume into another.

d. Surface Meteorology Measurements

The surface temperature, pressure, and dewpoint measurements provided with the rawinsonde soundings constituted the surface meteorology measurements used in this study. Presumably, identical instruments could be deployed anywhere, and differences among surface instruments are not at issue here.

3. RETRIEVAL TECHNIQUE

Since neither RASS nor radiometers measure temperature directly, temperature soundings must be inferred from the measurements using a known physical relationship between the measurements and the temperature at each sounding level. The linear statistical retrieval technique used here represents one approach to this problem. The technique estimates atmospheric temperature at a single height from a linear combination of measurements as (Hogg et al., 1983a),

$$\hat{T}(h) = c_0(h) + \sum_{i=1}^{m} c_i(h) x_i \; ,$$

where

$\hat{T}$ = temperature estimate
h = height coordinate
m = number of measurements

$\vec{x}$ = measurement vector
$\vec{c}$ = retrieval coefficient vector.

An *a priori* data set, independent of the measurements, determines the retrieval coefficient vector. Typically, this data set consists of enough rawinsonde soundings to characterize the local climatology near the instruments. A physical model that relates the measurement vector to these pressure, temperature, and/or dewpoint measurements is used to simulate a measurement vector from each rawinsonde sounding. The retrieval coefficients are then obtained by a least-squares fit of these simulated measurements to the $T(h)$ measured by the same set of rawinsondes. Strand and Westwater (1968) give a mathematical presentation of the technique.

a. Measurement Simulations

Implementing the retrieval technique required simulations of radiometer, surface meteorology, and RASS measurements from an a priori data set. For this study, the data set consisted of 1200 rawinsonde soundings made by NWS during July and August, from 1970 to 1979 at Stapleton Airport in Denver, Colorado.

1) RADIOMETER

The theoretical relationship between the intensity of emitted radiation observed by the radiometers at the ground and the temperature, humidity, and pressure structure of the atmosphere is given by the radiative transfer equation (Askne and Westwater, 1986),

$$Tb = T_{bg} \exp\left(- \int_0^\infty a(h)dh \right) + \int_0^\infty T(h)a(h) \exp\left(- \int_0^h a(h')dh' \right) dh \, , \qquad (1)$$

where

Tb = equivalent brightness temperature
T_{bg} = cosmic background radiation
h = height coordinate
$T(h)$ = temperature at height h
$a(h)$ = attenuation coefficient at height h.

At microwave frequencies, scattering is small compared with absorption, so that attenuation is due primarily to absorption by oxygen, water vapor, and clouds (suspended liquid water droplets).

Tbs at the six radiometer frequencies were calculated from each sounding, using (1). Prior to the calculations, each sounding was extrapolated to 0.1 mb, so that the integration encompassed the entire atmosphere. A model by Liebe and Layton (1987), with interference coefficients updated after Rosenkranz (1988), was used to compute clear-sky $a(h)$ from rawinsonde temperature, pressure, and dewpoint measurements at height h.

$T(h)$ was taken as the rawinsonde temperature at height h. T_{bg} has been documented as 2.75 K (Wilkinson, 1986). These Tbs represented radiometer measurements made under a wide variety of clear summer conditions.

Since rawinsondes do not measure cloud water content, cloudy conditions were modelled after Decker et al. (1978), who assumed a cloud was present at levels where the rawinsonde relative humidity measurement exceeded 95 percent. When this occurred, three additional sets of Tbs were calculated, assuming three different liquid water densities within the cloud layer(s), based on layer thickness. At sub-freezing temperatures, a fraction of the liquid density was treated as ice density, depending on temperature. The microwave attenuation due to clouds, computed after Westwater (1972), was then added to the clear-sky $a(h)$ in (1). The resulting Tbs represented radiometric measurements made during a variety of summer cloud configurations.

The clear and cloudy *Tb*s together represented radiometer observations made during 2225 different summertime conditions. Instrument noise was simulated by adding Gaussian random errors with zero mean and 0.5 K standard deviation to the calculated *Tb*s. Previous simulations using this noise level replicated the temperature-sounding accuracy of the Profiler radiometric system (Schroeder et al., 1989).

Brightness temperatures at 20.6 and 31.65 GHz were converted to absorption, because previous studies showed absorption to be a better estimator than *Tb* at these frequencies, where the range of *Tb*s is large. The Rayleigh-Jeans approximation to the Planck function allows the conversion (Hogg et al., 1983b),

$$\tau = \ell n[(T_{mr} - T_{bg})/(T_{mr} - Tb)] ,$$

where

τ = total absorption (nepers)
T_{bg} = cosmic background Tb = 2.75 K
T_{mr} = mean radiating temperature (K)
nepers = ℓn (power in/power out).

2) SURFACE METEOROLOGY

Surface meteorology measurements were simulated from the surface values of temperature (T), pressure (P), and relative humidity (RH) provided with the rawinsonde soundings. Random noise added to these values had zero mean and standard deviations of 0.5 K, 0.3 mb, and 4.4 percent, respectively, based on previous comparisons between Profiler and NWS surface measurements (Falls, 1987).

3) RASS

Simulating RASS measurements required interpolating the rawinsonde temperature, pressure, and dewpoint measurements onto the heights of the RASS measurement levels (Section 2). Then virtual temperature at each height was calculated from them as (Byers, 1974),

$$T_v = pT \,/\, [\epsilon e + (p - e)] \ ,$$

where

T_v = virtual temperature (K)
T = temperature (K)
p = total pressure (mb)
ϵ = ratio of molecular weights, water vapor:dry air
= .62197
e = vapor pressure (mb) = $e_s(T_d)$
T_d = dewpoint temperature (K)
e_s = saturation vapor pressure (mb)

The Goff-Gratch formula was used to compute e from T_d (List, 1963). Random noise with zero mean and 0.9°C standard deviation was added to each simulated RASS measurement, based on previous experience with RASS at 915 MHz (May, 1988b).

b. Retrieval Coefficients

Table 2 summarizes the measurement vectors associated with each remote sensing system involved in the comparison. A retrieval coefficient vector corresponding to each measurement vector was calculated for each height level requiring a temperature estimate. A set of 85 retrieved temperature levels, chosen to preserve the vertical resolution of the RASS and the radiometer, provided temperature estimates for the entire troposphere. The pressure-dependence of the radiometer measurements suggested a vertical coordinate of pressure, rather than height. Difference from surface pressure was used, rather than absolute pressure, to circumvent problems associated with varying surface pressures among the rawinsonde soundings. Closely spaced levels (5–mb increments) within the first 280 mb above the surface (about 4 km AGL) preserved the 150–m RASS height resolution. A 10–mb spacing was used between 280 and 500 mb difference pressure (about 4 to 7 km AGL), where the radiometer is still somewhat sensitive. Above 500 mb difference pressure, a coarse 25–mb grid sufficed. Rawinsonde temperature measurements were interpolated to fit this difference pressure grid.

Table 2. Measurement vectors associated with each system.

System	Measurement vector
Radiometric	Surface T, P, RH, 2τ, 4 Tb
RASS	Surface T, P, RH, 6–19 T_v
Combined	Surface T, P, RH, 2τ, 4 Tb, 6–19 T_v

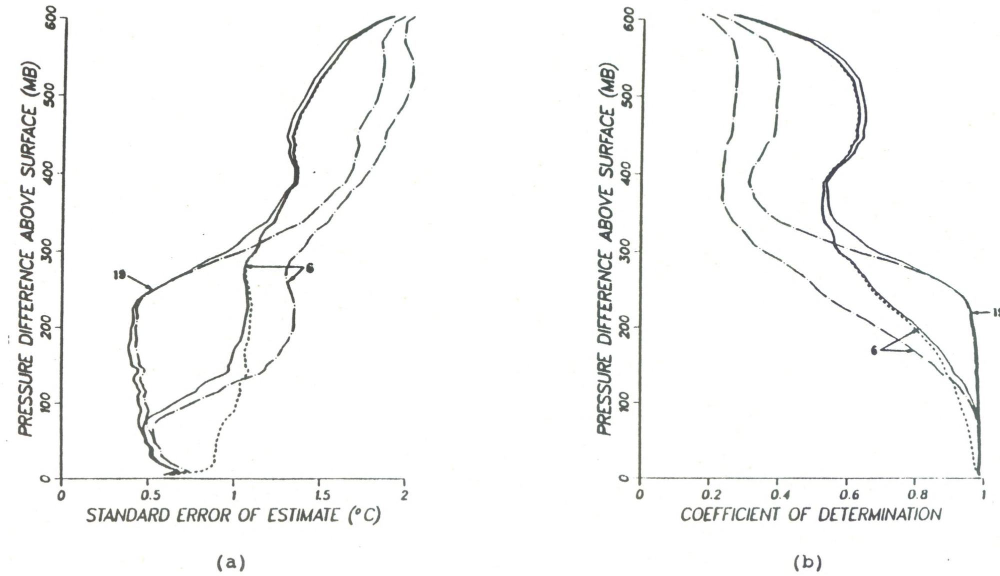


Figure 3. (a) Standard error of temperature estimates and (b) fraction of temperature variance explained by the regression model for radiometric (---), RASS (- · -), and combined RASS/radiometer (—) systems. The two curves for RASS and combined systems represent the fewest (6) and the largest (19) number of RASS virtual temperature levels measured during this experiment.

Commercial statistical software simplified retrieval coefficient calculations considerably. First, a covariance matrix involving all simulated measurements and the 85-level rawinsonde temperatures was calculated. Then, using appropriate rows and columns of that matrix, a multiple linear regression model for each of the 85 levels was fitted to the simulated measurement vectors. The regression coefficients obtained served as retrieval coefficients for the respective systems.

Figure 3 shows how well the regression models estimated the temperatures from which they were derived. The dashed curves represent the radiometric system, the dash-dot curves represent the RASS system, and the solid curves represent the combined system. Two curves appear in each plot for the systems involving the RASS measurements because of the variable number of measurement levels. The results shown for 6 levels and 19 levels illustrate the two extremes encountered in this experiment. Curves corresponding to other numbers of levels lie between the two shown.

At each sounding level, Fig. 3a shows the standard error of the estimate, or root mean square (rms) difference, between rawinsonde and retrieved temperatures. Errors differ greatly, depending on the number of RASS measurement levels involved. Within the altitude range of the RASS measurements, the RASS and combined system errors are within the rawinsonde precision reported by Hoehne (1980). Above that range, the RASS system temperature estimates are poorer than the radiometer estimates.

The coefficient of determination, plotted in Fig. 3b, puts the size of the errors in perspective by showing the fraction of the temperature variance explained by the regression model for each system. The model for a RASS system with 19 levels accounted for nearly all of the temperature variance in the first 250 mb above the surface. The effectiveness of all models deteriorates with increasing altitude.

4. DATA ANALYSIS AND INTERPRETATION

Temperature soundings were retrieved from radiometer, RASS, and combined RASS/radiometer measurements obtained during the ascent of each rawinsonde listed in Table 1, resulting in 22 soundings from each measurement system. Three different types of comparisons among the four systems are presented in this section. First, temperatures from all levels are compared simultaneously without segregating levels. Then I present comparisons at selected NWS mandatory pressure levels. Finally, a multivariate comparison examines all levels simultaneously, while preserving the level-dependence of the temperatures.

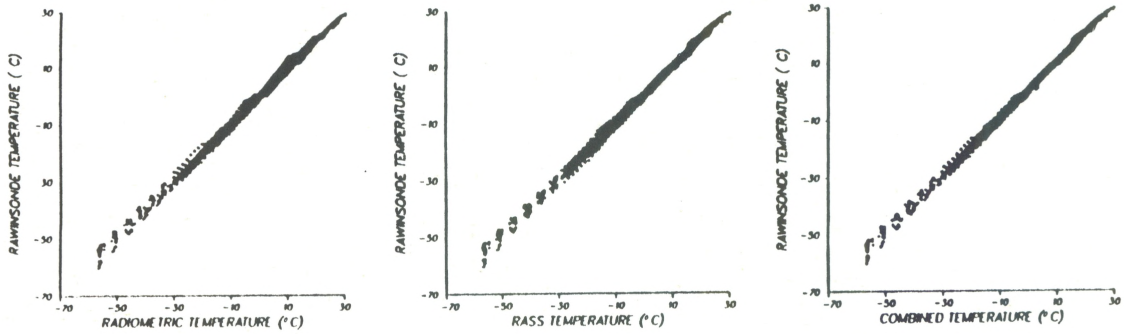
Since the rawinsonde temperature measurements served as the standard of comparison, I defined error as retrieved temperature minus rawinsonde temperature. A vertical coordinate common to all four systems was required to compute these errors. Because rawinsonde measurement levels vary from sounding to sounding, rawinsonde temperatures were interpolated onto the 85 retrieved sounding levels for all comparisons except those at mandatory levels, which required the interpolation of all soundings.

a. Overall Univariate Comparison

Hoehne (1980) presented a comparison of temperature soundings obtained by two NWS rawinsonde packages flown on the same balloon, separated vertically by 5 m, launched weekly for 50 weeks. His statistical analysis of the data did not distinguish among sounding levels, so that his reported mean, standard deviation, and rms differences characterized the overall precision of the NWS rawinsonde instruments. Since I know of no comparable study for the CLASS rawinsondes, I extend Hoehne's results to characterize all of the rawinsondes used in this experiment.

In order to compare directly with Hoehne's results, I compared retrieved temperatures with rawinsonde temperatures without segregating them by sounding level. The scatter plots in Fig. 4 show the temperatures measured by rawinsonde versus those retrieved from each of the three systems. Notice the lack of outliers and the similarity among the three plots. The axis endpoints illustrate the range of temperatures involved in tropospheric temperature profiling. There is slightly more scatter at very low temperatures, which occur at high altitudes.

The overall mean, standard deviation, and rms errors associated with the plots in Fig. 4 appear with Hoehne's rawinsonde precision results (line 4) in Table 3. The rms errors shown compare favorably with the NWS rawinsonde rms precision. Although Hoehne attributes the reported rawinsonde bias to heating caused by the upper sonde's battery, the retrieved soundings produced a similar bias. In fact, only the RASS system's overall mean error differed significantly from the rawinsonde bias, when tested statistically ($\alpha = .05$).



(a) (b) (c)

Figure 4. Temperatures measured by rawinsonde versus temperatures retrieved from (a) radiometric, (b) RASS, and (c) combined RASS/radiometer measurements.

Although this type of analysis provides no information about accuracy as a function of height, it is one way to summarize and compare system performance. The overall rms errors for the radiometric and RASS systems differ by only 0.06°C. The combined system improves on the RASS and radiometric systems by 0.15 and 0.21°C, respectively. Ninety-five percent confidence intervals for the error standard deviations of each system did not overlap, indicating that the observed differences are statistically significant. However, since the data used in this study were not randomly sampled, and the temperatures at adjacent levels are not independent, the assumptions underlying this statistical test were probably violated.

Table 3. Statistics for overall univariate comparison.

System	rms (C)	Mean (C)	Std. Dev. (C)
Radiometric	1.15	−0.14	1.14
RASS	1.09	−0.21	1.07
Combined	0.94	−0.12	0.93
Rawinsonde	0.62	−0.13	0.61

b. Comparison at Mandatory Levels

Since numerical models and other analyses used by weather forecasters require temperature measurements at so- called mandatory pressure levels, the accuracy of temperatures obtained from the three remote sensing systems at selected mandatory levels were compared. In order to avoid comparisons involving extrapolated rawinsonde soundings, comparisons were restricted to levels below 150 mb.

Figure 5 shows scatter plots of the errors associated with each system at three mandatory levels, 700, 500, and 200 mb. Plots for 750, 400, 300, and 250 mb were similar to those at neighboring levels. At 700 mb, which corresponds approximately to the height of

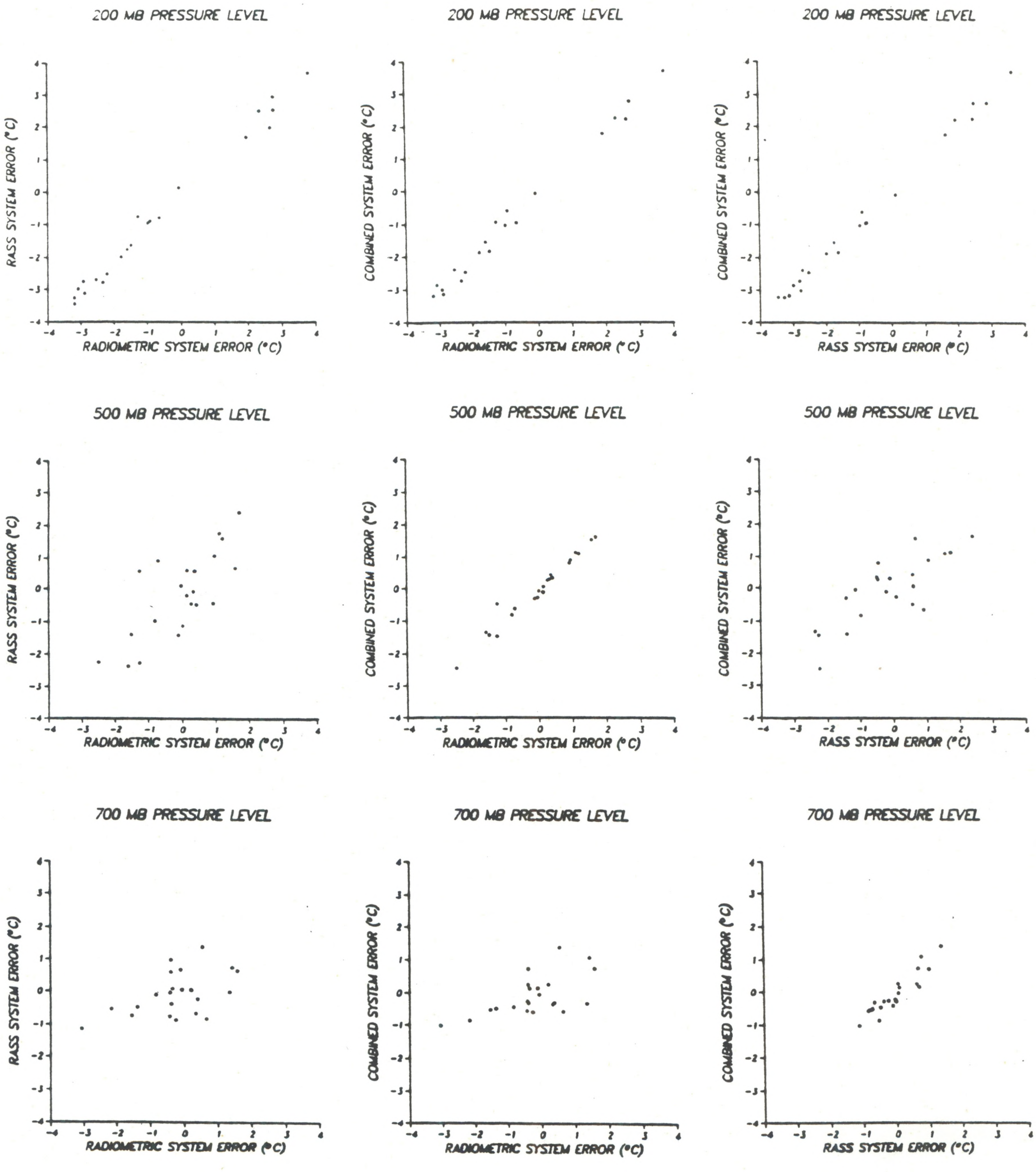


Figure 5. Comparison of errors (retrieved temperature minus rawinsonde temperature) at 700–, 500–, and 200–mb mandatory levels.

the tenth RASS measurement level, the RASS system errors are much smaller than the radiometric system errors, and they dominate the combined system performance (r = 0.93). Above the RASS measurements, at 500 mb, the radiometric and combined systems are highly correlated (r = 0.98), even though the range of errors in the RASS and radiometric soundings is about the same. At 200 mb the errors are larger, and all three systems are highly correlated with one another (r > 0.99), indicating that the climatology

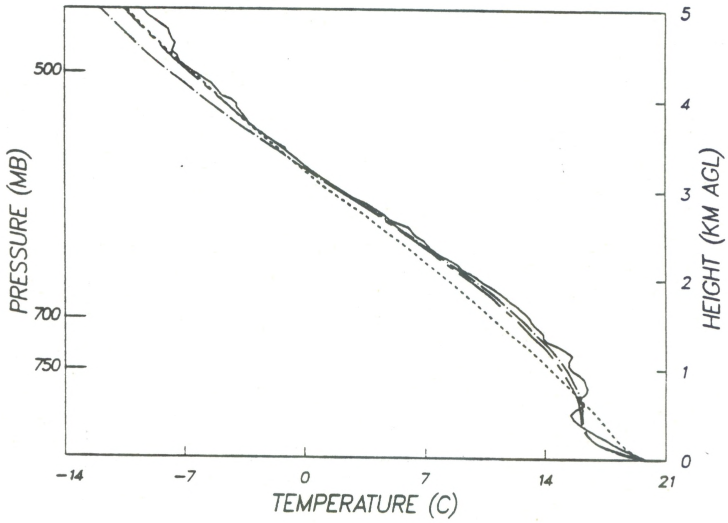


Figure 6. 1700 UTC 3 August 1988; 16 RASS measurement levels; CLASS temperature sounding (—) with temperature soundings retrieved from radiometric (---), RASS (- · · -), and combined (—-—) systems.

used in the retrieval process, rather than the measurements, may be producing the temperature estimates at this high-altitude level.

Figure 6 exemplifies the results shown in Fig. 5. A CLASS sounding is plotted with temperature soundings retrieved from the radiometric, RASS, and combined RASS/radiometric measurements. The length of the tick marks on the pressure axis represents the uncertainty in the rawinsonde temperature measurement (twice the rms precision reported by Hoehne). Most of the small structure in the CLASS sounding is within the measurement uncertainty, but there is evidence of a temperature inversion just below 1 km. Although all three remote systems miss the inversion, the RASS and combined soundings show some indication of structure at that height. The inversion might have been recovered from the RASS and combined systems if a lower noise level had been assigned to the RASS measurements in the retrieval scheme. At 700 mb, within the RASS measurement region, the RASS and combined soundings lie closer to the CLASS sounding than the radiometric sounding does. Above the RASS measurements, at 500 mb, the radiometric and combined soundings show better agreement than RASS does.

Table 4 contains means ($\bar{x}$), standard deviations (s), and respective 95% confidence intervals (C.I.) for the errors associated with each remote sensing system at each of the seven mandatory levels below 150 mb. All confidence intervals for the mean error contain both zero and the rawinsonde bias reported by Hoehne (1980), indicating that the mean errors for the remote sensing systems are not significantly different from each other or the rawinsonde. At 750 and 700 mb, the confidence intervals for the standard deviation of the RASS and combined system errors contain Hoehne's standard deviation for the rawinsonde. This implies that, at 750 and 700 mb, the accuracy of temperatures retrieved from the RASS and combined systems is within the functional precision of the NWS rawinsonde temperature measurements. At every mandatory level, the standard deviation confi

Table 4. Error statistics at selected mandatory levels.

Level (mb)	System	$\bar{x}$ (C)	C.I. ($\bar{x}$) (C)	s (C)	C.I. (s) (C)
750	Radiom.	0.21	(−0.16, 0.58)	0.84	(0.65, 1.20)
750	RASS	−0.24	(−0.54, 0.06)	0.68	(0.52, 0.97)
750	Comb.	−0.20	(−0.45, 0.04)	0.56	(0.43, 0.79)
700	Radiom.	−0.25	(−0.74, 0.25)	1.12	(0.86, 1.60)
700	RASS	−0.07	(−0.37, 0.23)	0.67	(0.52, 0.96)
700	Comb.	−0.07	(−0.34, 0.21)	0.62	(0.48, 0.88)
500	Radiom.	−0.03	(−0.52, 0.47)	1.11	(0.85, 1.59)
500	RASS	−0.13	(−0.72, 0.46)	1.34	(1.03, 1.91)
500	Comb.	−0.03	(−0.50, 0.43)	1.05	(0.81, 1.50)
400	Radiom.	−0.16	(−0.73, 0.40)	1.28	(0.98, 1.82)
400	RASS	−0.23	(−0.92, 0.47)	1.57	(1.20, 2.24)
400	Comb.	−0.04	(−0.57, 0.48)	1.18	(0.91, 1.69)
300	Radiom.	−0.21	(−0.84, 0.41)	1.41	(1.08, 2.01)
300	RASS	−0.24	(−0.79, 0.31)	1.24	(0.96, 1.77)
300	Comb.	−0.10	(−0.67, 0.47)	1.28	(0.98, 1.82)
250	Radiom.	−0.40	(−1.25, 0.46)	1.94	(1.49, 2.76)
250	RASS	−0.48	(−1.21, 0.25)	1.66	(1.27, 2.37)
250	Comb.	−0.38	(−1.20, 0.44)	1.85	(1.42, 2.64)
200	Radiom.	−0.67	(−1.70, 0.35)	2.31	(1.78, 3.30)
200	RASS	−0.77	(−1.80, 0.25)	2.30	(1.77, 3.29)
200	Comb.	−0.74	(−1.75, 0.27)	2.29	(1.76, 3.27)

dence intervals for the three remote systems overlap, indicating that the error standard deviations are not significantly different. Since the soundings considered here do not represent a random sample, these significance tests apply only to this experiment, not to the three remote sensing systems in general.

c. Overall Multivariate Comparison

The most effective way to compare the performance of the three remote sensing systems is to examine the errors associated with each sounding level simultaneously, while preserving the level-dependence of the temperatures.

The rms difference, mean difference, and the standard deviation of the differences between rawinsonde temperature measurements and those retrieved from each of the three systems appear in Fig. 7 as a function of pressure difference above the surface (AGL). The mean surface pressure for this set of soundings was 842 mb, with a standard deviation of 4 mb. Subtracting the vertical coordinate given in the figure from 842 gives an approximate value for the absolute pressure. The dots along the vertical axis, showing

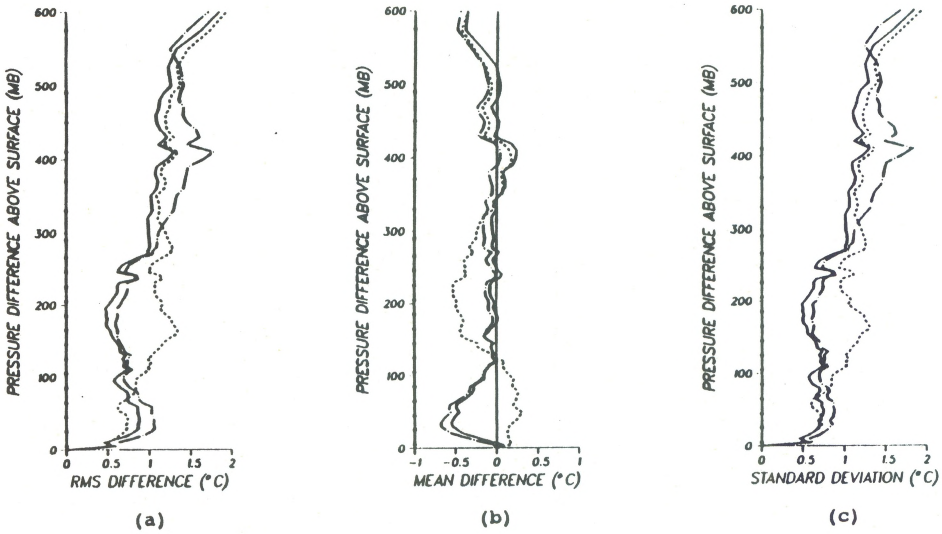


Figure 7. Comparison of differences between rawinsonde-measured temperatures and those retrieved from radiometric (---), RASS (- · -), and combined RASS/radiometric (—) systems. Dots on y-axis represent every fourth level compared.

every fourth level retrieved, illustrate the decreasing resolution with height chosen for the retrieved temperature levels. The heights of the RASS measurements obtained during this experiment fell between approximately 20 and 250 mb AGL, although most of them fell below about 180 mb AGL. In all three graphs, the dashed curve represents the radiometric system, the dot-dash curve represents the RASS system, and the solid curve represents the combined system.

The mean and rms differences shown in Figs. 7a and 7b for the radiometric system are similar in shape and magnitude to the those obtained by Westwater et al. (1984) with the same system during July 1982 (47 profiles). The radiometric system errors are smaller than the RASS errors at all levels except those in the range of the RASS virtual temperature measurements. In that region, the RASS errors are about half as large as the radiometric system errors.

The combined system appears to outperform both systems everywhere above the first 60 mb AGL. Figure 7b shows a large bias in the RASS soundings near the surface that persists in the combined soundings. The bias was probably due to ground clutter affecting the radar measurements at those heights (Strauch, 1988). When that bias was removed, the accuracy of the RASS and combined systems at those levels improved noticeably (Fig. 7c).

The statistical significance of the observed differences in performance among the three systems was assessed by multivariate hypothesis testing. Assuming that the temperature errors associated with all three systems were distributed normally, differences in accuracy could be identified by testing differences between mean vectors and covariance matrices of errors associated with each system.

The first test concerned whether or not the error covariance matrices of the three systems differed significantly, using a simultaneous test for equality of several covariance matrices described in detail by Box (1949). Initially, the test was conducted using 85 by 85 covariance matrices, which included all retrieved temperature levels. However, high correlations between errors at adjacent levels caused numerical problems, so every fourth level, indicated by the 21 dots along the vertical axes in Fig. 7, was used instead. Test results showed no statistically significant differences ($\alpha = 0.05$) among the error covariance matrices associated with the three systems.

Assuming equal covariance matrices, Hotelling's T^2 statistic was used to test the differences among the mean error vectors of the three systems (Anderson, 1958). When the test was applied to the same 21 levels described above, differences among mean vectors were not significant ($\alpha = 0.05$).

These test results suggest that the capabilities of the radiometric, RASS, and combined systems to mimic the rawinsonde temperature measurements are not significantly different. However, since the small sample of soundings considered here may not satisfy the underlying randomness and normality assumptions, the hypothesis test results presented may not apply to temperature sounding in general.

5. CONCLUSIONS

Twenty-two sets of temperature soundings obtained simultaneously from four different measurement systems at Stapleton Airport in Denver, Colorado, during July and August of 1988 have been compared. A linear statistical retrieval technique (Strand and Westwater, 1968), which requires a priori knowledge of local climatology, was used to estimate temperatures throughout the troposphere from measurements made by a six-channel microwave radiometer, a 915-MHz/2-kHz radio-acoustic sounding system (RASS), and a combination of the two systems. Rawinsonde temperature soundings served as the comparison standard.

The rms difference between rawinsonde-measured temperatures and those retrieved from each system seldom exceeded 1.5°C below the tropopause. The combined system rms error was smaller than the errors of the other two systems at all levels above about 0.5 km. Below this level, the radiometric system outperformed the RASS and combined systems, whose retrieved temperatures were degraded by a bias at the first few RASS measurement levels. Above 5 km, the radiometric system performance approached that of the combined system. In between, where unbiased RASS virtual temperature measurements were made, the radiometric system could not compete. In fact, the accuracy of temperatures retrieved from the RASS and combined systems at 750 and 700 mb was within the functional precision of the NWS rawinsonde (Hoehne, 1980).

Hypothesis tests were used to assess the statistical significance ($\alpha = 0.05$) of the accuracy differences observed. When all temperatures were compared together, as if they were all at the same level, differences among the three systems were statistically significant. However, the degrees of freedom involved were inflated, because temperatures at adjacent levels are not independent. The accuracies of temperatures retrieved from the three systems at each of seven mandatory levels were not significantly different. Multivariate tests, which evaluate all the temperature sounding levels simultaneously while preserving the level-dependence of the temperatures, also indicated that differences among the RASS, radiometric, and combined system errors were not statistically significant. Due to probable violations of the assumptions underlying these tests, the results only describe this particular experiment, not the performance of the three remote sensing systems in general.

Statistically significant or not, the accuracy differences observed are probably too small to justify the expense of a combined system, even though it out-performed either system individually. Given the similarity between RASS and radiometric temperature soundings shown in this study, cost considerations certainly favor the RASS system in situations where a wind-profiling radar is already in place. A RASS system involving a 404.37–MHz radar, such as those being built for the U.S. wind profiler network, would suffer less acoustic attenuation than the RASS system used in this study, permitting measurements over a wider altitude range. The additional measurements should produce more accurate temperature estimates at higher altitudes.

This experiment was conducted during a season characterized by relatively unchallenging temperature profiles. A similar experiment conducted during a season when elevated temperature inversions are common might identify more significant differences among the three systems considered here. The research presented explores just one way of combining the RASS and radiometer measurements. Other retrieval techniques might produce more synergistic results, justifying the expense of a combined system.

ACKNOWLEDGMENTS

The research presented here could not have been accomplished without the cooperation of several groups of people. The WPL Thermodynamic Profiling Group, headed by E. R. Westwater, funded the research and provided the radiometer data. In particular, M. T. Decker guided data interpretation and calibration; R. H. Beeler provided computer access to radiometer and rawinsonde data; J. R. Jordan maintained radiometer data integrity; M. J. Falls compiled the a priori rawinsonde data set.

R. G. Strauch, P. T. May, K. P. Moran, and D. A. Merritt of the WPL Profiler Research Program furnished the RASS data and guided its interpretation. R. McBeth and J. Murphy of NCAR provided the CLASS data, which N. Gonzalez of the WPL Radar Meteorology Program transferred by computer. The NWS furnished their routine rawinsonde data. J. Flueck, head of the NOAA Program for Regional Observing and Forcasting Services Evaluation Group, provided insight for data analysis and display.

Criticisms contributed by E. A. Howard III, J. C. Emerick, and W. W. Whitman of Colorado School of Mines resulted in a clearer presentation. R. Bonino and M. Birchfield of the WPL Message Center assisted with manuscript preparation.

This work was supported in part by the U.S. Customs Service under interagency agreement 1X890084, and we are grateful for the assistance of Rajiv Singh and Thomas Henneberger.

REFERENCES

Anderson, T.W., 1958. *An Introduction to Multivariate Statistical Analysis*. John Wiley and Sons, New York, 374 pp.

Askne, J.I.H., and E.R. Westwater, 1986. A review of ground-based remote sensing of temperature and moisture by passive microwave radiometers. *IEEE Trans. Geosci. Remote Sensing*, **GE-24:**340–52.

Box, G.E.P., 1949. A general distribution theory for a class of likelihood criteria. *Biometrika*, **36**:317–47.

Byers, H.R., 1974. *General Meteorology*. McGraw-Hill, Inc., 461 pp.

Chadwick, R.B., and N. Hassel, 1987. Profiler: the next generation surface-based atmospheric sounding system. Third International Conference on Interactive Information and Processing Systems for Meteorology, Oceanography and Hydrology, Jan. 12–16, 1987, New Orleans, La. American Meteorological Society, Boston, Mass., 15–21.

Decker, M.T., E.R. Westwater, and F.O. Guiraud, 1978. Experimental evaluation of ground-based microwave radiometric sensing of atmospheric temperature and water vapor profiles. *J. Appl. Meteor.*, **17**:1788–95.

Falls, M.J., 1987. Personal communication.

Hoehne, W.E., 1980. Precision of National Weather Service upper-air measurements. NOAA Tech. Memo. NWS T&ED–16.

Hogg, D.C., M.T. Decker, F.O. Guiraud, K.B. Earnshaw, D.A. Merritt, K.P. Moran, W.B. Sweezy, R.G. Strauch, E.R. Westwater, and C.G. Little, 1983a. An automatic Profiler of the temperature, wind and humidity in the troposphere. *J. Appl. Meteorol.*, **22**:807–31.

Hogg, D.C., F.O. Guiraud, J.B. Snider, M.T. Decker, and E.R. Westwater, 1983b. A steerable dual-channel microwave radiometer for measurement of water vapor and liquid in the troposphere. *J. Appl. Meteorol.*, **22**:789–806.

Lauritsen, D., Z. Malekmadani, C. Morel, and R. McBeth, 1987. The Cross-chain Loran Atmospheric Sounding System (CLASS). Sixth Symposium Meteorological Observations and Instrumentation, Jan. 12–16, 1987, New Orleans, La. American Meteorological Society, Boston, Mass., 340–43.

Liebe, H.J., and D.H. Layton, 1987. Millimeter-wave properties of the atmosphere: Laboratory studies and propagation modeling. NTIA Report 87–224, 74 pp.

List, R.J., 1963. *Smithsonian Meteorological Tables.* Smithsonian Institution, Washington, D. C., 527 pp.

May, P.T., R.G. Strauch, and K.P. Moran, 1988a. The altitude coverage of temperature measurements using RASS with wind profiler radars. *Geophys. Res. Lett.*, **15**:1381–84.

May, P.T., 1988b. Personal communication.

May, P.T., R.G. Strauch, K.P. Moran, and W.L. Ecklund, 1989. Temperature sounding by RASS with wind profiler radars: A preliminary study. *IEEE Trans. Geosci. Remote Sensing* (in press).

Rosenkranz, P.W., 1988 Interference coefficients for overlapping oxygen lines in air. *J. Quant. Spectrosc. Radiat. Transfer*, **39**:287–97.

Schroeder, J.A., J.R. Jordan, and M.T. Decker, 1989. Design considerations for a network of thermodynamic profilers. *J. Atmos. Oceanic Technol.* October issue (in press).

Strand, O.N., and E.R. Westwater, 1968. Minimum rms estimation of the numerical solution of a Fredholm integral equation of the first kind. *SIAM J. Numer.Anal.*, **5**:287–95.

Strauch, R.G., 1988. Personal communication.

Strauch, R.G., D.A. Merritt, K.P. Moran, K.B. Earnshaw, and D. van de Kamp, 1984. The Colorado wind-profiling network. *J. Atmos. Oceanic Technol.*, **1**:37–49.

Westwater, E.R., 1972. Microwave emission from clouds. NOAA Tech. Report ERL 219-WPL 18, 43 pp.

Westwater, E.R., W. Sweezy, L. McMillin, and C. Dean, 1984. Determination of atmospheric temperature profiles from a statistical combination of ground-based Profiler and operational NOAA 6/7 satellite retrievals. *J. Clim. Appl. Meteor.*, **23**:689-703.

Wilkinson, D.T., 1986. Anisotropy of the cosmic blackbody radiation. *Science*, **232**:1517–22.