Data Quality Assessment Methods
for the
Eastern Range 915 MHz Wind Profiler Network

1 June 1998

Prepared by:
Applied Meteorology Unit

ENSCO, Inc.
1980 N. Atlantic Ave., Suite 230
Cocoa Beach, FL 32931
(407)   853-8130
(407)   783-9735

Executive Summary

The Eastern Range installed a network of five 915 MHz Doppler Radar Wind Profilers (DRWP) with Radio Acoustic Sounding Systems (RASS) in the Cape Canaveral Air Station / Kennedy Space Center (CCAS/KSC) area to provide three-dimensional wind speed and direction and virtual temperature estimates in the boundary layer. These profilers were installed to provide high spatial (100 m) and temporal (15 minutes) resolution wind data in the gap between the top of the CCAS/KSC wind tower network (150 m) and the lowest gate (2 km) of the NASA 50 MHz DRWP. The Applied Meteorology Unit (AMU), staffed by ENSCO, Inc., was tasked by the 45th Weather Squadron, the Spaceflight Meteorology Group, and the National Weather Service in Melbourne, Florida to investigate methods which will help forecasters assess profiler network data quality when developing forecasts and warnings for critical ground, launch and landing operations.

Four routines were evaluated in this study: a consensus time period check, a precipitation contamination check, a median filter, and the Weber-Wuertz (WW) algorithm. The consensus time period check flags wind profiles that are created from data collection periods of less than six minutes. The precipitation contamination check uses a relationship between vertical beam radial velocity and signal-to-noise ratio to determine if a wind estimate is affected by rain, a known contaminant of 915 MHz profiler data. The median filter compares the u- and v-components of a wind observation to the median of the components of the surrounding wind observations. If the difference between an observed component and a median component exceeds a certain threshold value, the observation is flagged. This median filter is not related to the Median Filter/First Guess (MFFG) algorithm used on NASA's 50 MHz DRWP. Conclusions about the MFFG performance cannot be made based on the performance of the median filter in this study. WW is a pattern recognition program that is widely used in the profiler community. Wind observations that do not fit into established patterns in the data set are flagged. No routine was able to effectively flag suspect data when used by itself. Therefore, the routines were used in different combinations. An evaluation of all possible combinations revealed two that provided the best results. The precipitation contamination and consensus time routines were used in both combinations. The median filter or WW was used as the final routine in the combinations to flag all other suspect data points.

The routines were tested in both post-analysis and simulated real-time mode. In post-analysis mode, the data are assessed for quality after being collected and archived. This means that WW and the median filter can evaluate continuity with wind estimates collected after the wind estimate being checked. In order to test the routines in a real-time mode, they were modified to use only previously collected data in their continuity checks. Archived data were then input to the routines for the simulated real-time tests. Profiler data for this study were collected during the period 1 May through 31 August 1997. Five days with different weather phenomena were chosen from this period for algorithm development and testing. The use of data collected in diverse weather conditions allowed for a thorough analysis of how the routines would respond in different weather regimes. The results from each routine were evaluated using color displays of wind speed and direction and textual output.

An important conclusion of this study is that more than one quality assessment routine is needed to accurately flag most of the erroneous data. Routines that check for temporal and spatial continuity work well when the ratio of bad to good data is small. However, long-lived contaminants can cause a large amount of erroneous data that will not be properly flagged by these algorithms. Because of the small data set used in the analysis and the fact that the results from the median filter were only slightly better than WW in the simulated real-time mode, neither routine is recommended over the other. It is critical, however, that the consensus time and precipitation contamination checks be used to flag the obviously bad wind estimates before using either of the dependent routines. Both of the combinations are able to identify most of the unreliable data and little of the good data as suspect.

Table of Contents

List of Figures

List of Tables

List of Acronyms

Term	Description
45 WS	45th Weather Squadron
AMU	Applied Meteorology Unit
CCAS	Cape Canaveral Air Station
CPM	Consensus Time/Precipitation Contamination/Median Filter Combination
CPW	Consensus Time/Precipitation Contamination/Weber-Wuertz Combination
DRWP	Doppler Radar Wind Profiler
KSC	Kennedy Space Center
MFFG	Median Filter/First Guess Algorithm
MHz	Mega-Hertz
NOAA	National Oceanic and Atmospheric Administration
NPN	NOAA Profiler Network
NWS MLB	National Weather Service in Melbourne, Florida
QC	Data Quality Assessment
RASS	Radio Acoustic Sounding System
ROCC	Range Operations Control Center
RSA	Range Standardization and Automation
SMG	Spaceflight Meteorology Group
SNR	Signal-to-Noise Ratio
WW	Weber-Wuertz Algorithm

1   Introduction

The purpose of this report is to document the findings of the Applied Meteorology Unit (AMU) task regarding quality assessment of wind data from the 915 MHz Doppler Radar Wind Profiler (DRWP) network on Cape Canaveral Air Station/Kennedy Space Center (CCAS/KSC). The AMU, staffed by ENSCO, Inc., has been tasked by the 45th Weather Squadron (45 WS), the Spaceflight Meteorology Group (SMG), and the National Weather Service in Melbourne, Florida (NWS MLB) to investigate methods which will help forecasters assess profiler data quality when developing forecasts and warnings for critical ground, launch and landing operations.

Section 1 of this report provides an overview of the radars in the network, a description of wind profiling principles and the consensus averaging technique, and a discussion of the need for data quality assessment (QC). A summary of all the QC routines considered and their descriptions are given in Section 2, Section 3 shows the results of the analyses, and the conclusions are given in Section 4.

1.1   Profiler Network Description

The Eastern Range procured and installed a network of five 915 MHz DRWP with Radio Acoustic Sounding Systems (RASS) in the CCAS/KSC area (Heckman et al. 1996). This network can provide three-dimensional wind direction and speed estimates in the boundary layer from 120 m to 4 km (400 to 13100 ft) AGL and virtual temperature (T_v) estimates from 120 m to 1.5 km (400 to 5000 ft) AGL. These profilers were installed to provide high spatial and temporal resolution wind data in the data gap between the top of the KSC/CCAS wind tower network (150 m or 500 ft) and the lowest gate (2 km or 6500 ft) of the NASA 50 MHz DRWP.

The five profilers are arranged in a diamond-like pattern over the Cape area with an average spacing of 10-15 km, as seen in Figure 1.1. The profilers have 4-panel phased-array antennas that are capable of operating in a five-beam configuration: one vertical beam and four oblique beams at a 23.5° angle from the zenith. All profilers currently use a three-beam configuration with one vertical and two oblique beams. The three-beam configuration was chosen to produce the highest possible temporal resolution consensus data without sacrificing accuracy in the three-dimensional wind calculations. The beam directions at each individual site were chosen to minimize potential interference from ground clutter (trees, aircraft, traffic, etc.). Table 1.1 shows the configuration of the oblique beam directions for each profiler. The beams currently used are highlighted in bold type and denoted by an asterisk (*).

1
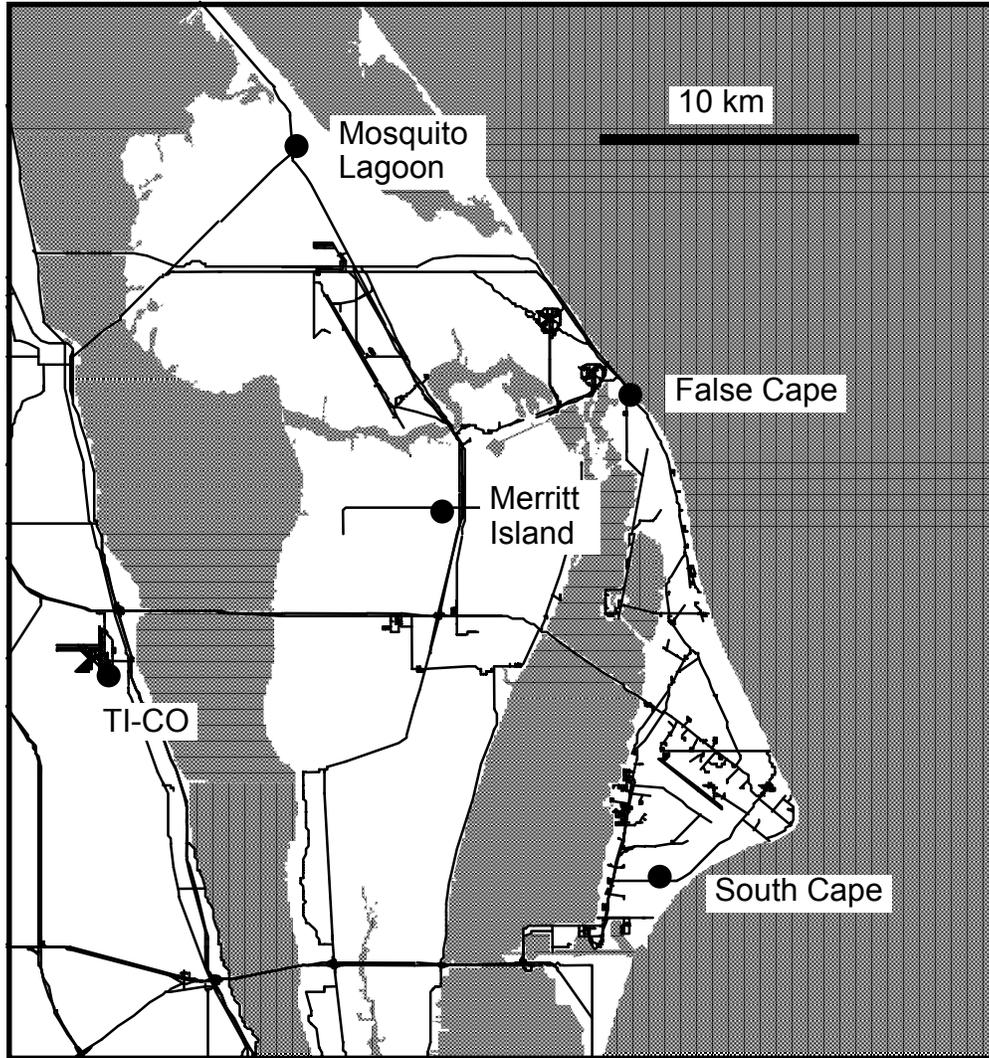
Figure 1.1   Map of CCAS/KSC. The 915 MHz DRWP locations are indicated by a solid circle, and the names of the locations are printed next to the sites. A scale is provided in the upper-right.

Table 1.1   Oblique beam directions for each profiler. Beams in current use are in **bold** type and indicated by a *****.

Radar	North	East	South	West
South Cape	354°	***84°**	***174°**	264°
False Cape	***2°**	92°	182°	***272°**
Merritt Island	***17°**	107°	197°	***287°**
Mosquito Lagoon	40°	***130°**	***220°**	310°
TI-CO Airport	***43°**	***133°**	223°	313°

2

The currently available operational parameter settings for the wind measurements are defined in parameter sets given in Table 1.2 (Radian Corporation 1994). The set currently in use is Parameter Set 2. The parameter set in use can be changed, but the individual settings within each set cannot be modified. Preliminary studies have shown that the settings in Parameter Set 2 produce the greatest height range with acceptable gate and time resolutions.

Table 1.2	The 915 MHz DRWP possible operational parameter settings. The parameter set currently in use is indicated by a **.		
Parameter	**Set 1**	**Set 2****	**Set 3**
Beam width	10°	10°	10°
Pulse width	400 ns	700 ns	1400 ns
Pulse repetition period	23 μs	43 μs	41 μs
Spectral points	64	64	64
# Coherent Avgs	350	188	200
# Spectral Avgs	42	42	42
Averaging time	10 min	10 min	10 min
Number of gates	30	32	22

The operational characteristics of the radars produced by Parameter Set 2 are shown in Table 1.3 (Radian Corporation 1994). For completeness, the characteristics produced by Parameter Sets 1 and 3 are given in parentheses where they differ from those given by Set 2. The first gate and maximum heights define the height range for each wind profile. The gate spacing indicates the height interval between wind estimates in a profile. A 10-minute wind consensus period is followed by a 5-minute RASS consensus period. Thus, wind and RASS consensus data are provided every 15 minutes. The Nyquist velocity is the maximum unambiguous speed that can be detected along a beam toward or away from the radar. The value of the Nyquist velocity is the same for each of the three parameter sets. Speeds in excess of this value will change sign and decrease in magnitude. This is an effect known as velocity folding that will cause an erroneous wind value to be calculated. For the 915 MHz DRWPs in the network, positive and negative values represent wind components toward and away from the radar, respectively. Wind speed and direction profiles are accurate to within 2 knots and 10°.

Table 1.3	The 915 MHz DRWP operational performance characteristics using Parameter Set 2. The characteristics produced by Set 1 and Set 3 are given in parentheses (*Set 1, Set 3*) if they are different.

Parameter	Characteristic
First gate height	383.4 ft (*360.8 ft, 1037.3 ft*)
Maximum height	10170.7 ft (*5592.4 ft, 14296.7 ft*)
Gate spacing	315.7 ft (*180.4 ft, 631.4 ft*)
Time Resolution	15 min
Nyquist velocity	±19.6 kts
Accuracy	Speed: < 2 kts Direction: < 10°

1.2 Principles of Wind Profiling

DRWPs depend on the scattering of electromagnetic energy by variations in the index of refraction of the air to ultimately resolve horizontal and vertical wind speed and direction (Peterson 1988). The index of refraction of the air depends on temperature, pressure, and moisture. Spatial variations in the index are caused by turbulent eddies which are created by horizontal or vertical wind shear. These variations are advected by the wind and are used by DRWPs to determine the mean wind.

If a propagating electromagnetic wave encounters a spatial variation in the refractive index, a small amount of energy is scattered in all directions. Some of the energy is backscattered to the radar where it can be received and analyzed. Backscattering occurs preferentially from variations in the index of refraction that are about half the size of the wavelength of the radar. With a wavelength of ~33 cm, this corresponds to variations on the order of 16.5 cm for 915 MHz profilers. Such small variations occur mostly in the lowest few kilometers of the atmosphere, which is the operating range for these radars.

If the returned energy, or signal, is very weak, the signal processing algorithms may not be able to distinguish it from background noise in the atmosphere and an erroneous wind may be calculated. Certain meteorological conditions have an effect on the returned signal strength. Index of refraction variations of the appropriate size must exist for the energy to be scattered and returned to the radar. If the mean flow is laminar, there will be little or no turbulence to produce the variations and any returned signal will be very weak. The returned signal tends to be stronger when the atmosphere is moist or cloudy. Precipitation also produces a strong signal, but this is a contaminant in 915 MHz profiler data which will be discussed later.

1.3 Wind Calculation

The raw radar waveforms along each individual beam are processed to yield the average Doppler frequency shift at each range gate. The Doppler shifts are then converted into radial velocity estimates along the radar beam. Detailed descriptions of the processes used are beyond the scope of this report, and the interested reader is referred to Peterson (1988). The radial velocities are consensus-averaged over a 10-minute period to remove outliers at each range gate. During this period, the radar produces 6 to 8 velocity estimates for each gate of each beam. The exact number of estimates depends upon the radar beam dwell time. The consensus average is calculated by averaging the values of the largest subset of radial velocity estimates that fall within a defined interval of each other. For the profilers in the network, the velocity estimates must be within 2 ms^{-1} of each other to be used in the calculation of the consensus radial velocity.

The consensus radial velocities are combined into horizontal wind components using the general equations (Peterson 1988)

$$u = (V_1 - w \cos\theta)/ \sin\theta \qquad\qquad 1.1$$

$$v = (V_2 - w \cos\theta)/ \sin\theta \qquad\qquad 1.2$$

$$w = V_3$$

1.3

where V_1 and V_2 are the consensus radial velocities along the oblique beams, V_3 is the measured radial velocity along the vertical beam, and θ is the zenith angle (23.5°). Due to the oblique beams' small zenith angle, a significant portion of the radial velocities measured along the beams could be due to the vertical velocity. The $w\cos\theta$ term on the right-hand side of equations 1.1 and 1.2 is the correction for the vertical velocity in the oblique beam. At least 60% of the individual radial velocity estimates must make up the subset used to calculate the consensus radial velocity. If this condition is not met, a consensus is not reached at the range gate being considered. If a consensus is not reached at a range gate in one or both of the oblique beams, a horizontal wind is not calculated. It is important to note that if a consensus is not reached for the vertical velocity, the horizontal wind will be calculated without applying the vertical velocity correction provided a consensus is reached in both oblique beams. The consensus averaging constraint parameters are listed in Table 1.4.

Table 1.4	Consensus Constraints.
Parameter	**Constraint**
Vertical Velocity Correction	Yes*
Consensus Avg Window	2 ms^{-1} (4 kts)
% Required for Consensus	60%
Total # Samples in Period	6 - 8
* *Applied when consensus vertical velocity is available.*	

The flow diagram in Figure 1.2 illustrates the process by which a horizontal wind is calculated from the raw radar data at one gate along one beam, labeled Beam 1 in the diagram. In one beam sample, the raw radar data are received, the Doppler shift is determined, and the radial velocity at the gate is calculated. In this example, six sample beams are taken and a radial velocity is determined for each. Five of the six samples are within 2 ms^{-1} of each other. This meets the criterion that at least 60% of the samples be within 2 ms^{-1} of each other in order to calculate a horizontal wind. Estimate 1 in the figure is an outlier and consequently not used in the consensus average. The remaining five values are averaged to produce a consensus radial velocity of 6.2 ms^{-1} for Beam 1 at that gate. It is important to note that an average radial velocity for the vertical and oblique beams will still be calculated if less than 60% of the estimates are within 2 ms^{-1} of each other. If the 60% criterion is not met in one or both of the oblique beams, the horizontal wind will not be calculated.

This process takes place concurrently in both oblique beams and the vertical beam during the 10-minute consensus period. When the consensus radial velocities have been calculated at each gate for all three beams, they are used in equations 1.1 - 1.3 to determine the horizontal winds.
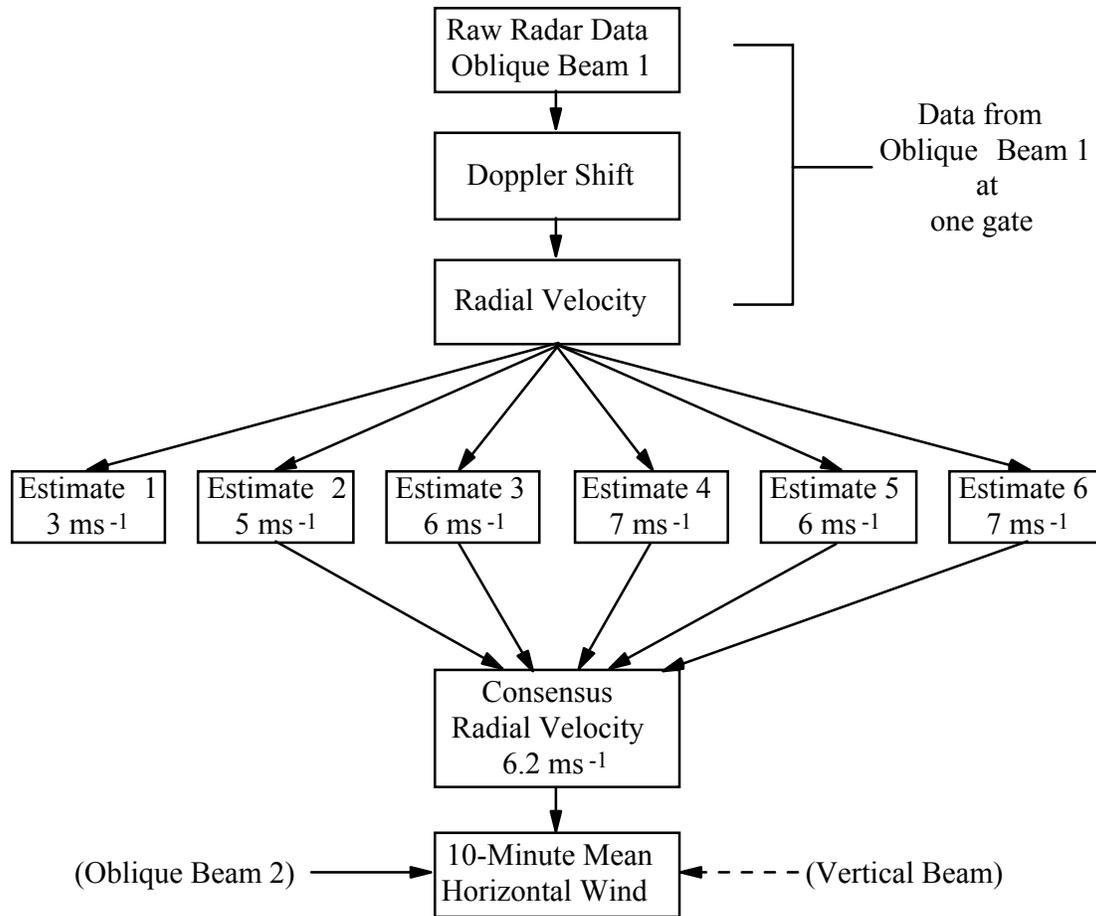
Figure 1.2   Flow diagram illustrating how a horizontal wind is derived from raw radar data. This represents the data flow for one gate along Oblique Beam 1. (Oblique Beam 2) and (Vertical Beam) represent the same process to produce consensus radial velocities in the other two beams. The dashed line from (Vertical Beam) indicates that input from this beam is not needed to calculate a horizontal wind.

1.4   Need for Data Quality Assessment

There are several factors that have a negative effect on 915 MHz profiler data. Wind data can be contaminated by reflections from nearby traffic, trees, ocean waves or sea spray, aircraft, birds, or precipitation, and also by velocity folding. The consensus averaging technique is generally successful at eliminating any short-lived contaminant that affects a gate or gates less than 40% of the time during the consensus period. However, if a contaminant produces consistent values at least 60% of the time it will likely be used in the wind calculation thereby producing an erroneous wind observation. There is currently no quality assessment of the network data beyond the consensus constraints. Thus, the need exists to develop quality assessment routines.

Each of the profilers is located in an area which is prone to certain long-lived data contaminants (refer to Figure 1.1). The TI-CO profiler is located next to a small airport. Depending on the landing configuration, heavy air traffic could contaminate the data. The False Cape profiler is located next to a road that can have heavy automobile traffic in the mornings and afternoons. This traffic could create a return signal from antenna sidelobes thereby contaminating the data. This profiler is also located on the coast. However, there is a rise in the topography between it and the ocean so that sidelobe contamination

from ocean waves or spray is a rare occurrence. The other three profilers are located near areas of dense vegetation and trees which can also cause sidelobe contamination.

All five profilers are prone to velocity folding, and bird and precipitation contamination. The nearby wildlife refuge experiences large bird migrations in the spring and fall. Although data from the network has not been examined with bird contamination in mind, this is a potential data contaminant. Precipitation is a known contaminant of 915 MHz profiler data (Ralph 1995). The CCAS/KSC area receives close to 50" of rainfall annually. There is a great deal of convective rainfall in the warm season and periods of extended rainfall in the cool season. Because of this, precipitation is a very common contaminant of data in the network.

The data from the network will be used by operational forecasters to support launch and landing operations and general day-to-day ground operations. The acceptable wind speed and direction thresholds for each of these operations are very specific. These data will also be used to develop methods for forecasting thunderstorm initiation and other weather phenomena over CCAS/KSC. In both of these cases, it is important that the users know whether or not the data are reliable. The algorithms developed in this task must be able to identify most, if not all, of the unreliable data and little, if any, of the good data as suspect.

2 Quality Assessment Routines

Five QC algorithms were evaluated in this study and are described in this section. Each of the routines were tested extensively using network data sets that were known to contain both good and contaminated data. The routines were first tested individually to determine the appropriate threshold and parameter settings and to rule out any routines that could not properly flag suspect data. The routines which produced favorable individual results were then tested in several different combinations with each other. This involved determining the best combinations, the appropriate order of the routines, and further modification of their threshold and parameter settings.

These routines use the 10-minute consensus data to determine wind data reliability. Specifically, they use the calculated horizontal winds, the consensus radial velocities, and the consensus signal-to-noise ratios (SNR). It is a common practice to use individual beam data to QC profiler data. These data were not used here because the individual beam data will no longer be transmitted from the profiler sites by the end of 1998. Any routine that uses these data would have to be modified extensively to use the consensus data or not used at all. However, good results are obtained from using the consensus data as will be shown in Section 3.

2.1 Routine Descriptions

This section provides descriptions of the five routines considered and analyzed in the task. Some were found in the literature and others were developed based on characteristics seen in suspect data. The routines were evaluated and revised using color graphical displays and textual output of the network data. Wind speed and direction were displayed separately as color-fill time-height profiles over a 24-hour period. These displays facilitated the detection of erroneous wind speed and direction changes. Once the times and heights of suspect winds were determined, the data values in the textual output at those times and heights were examined. The data values were used to help determine the appropriate threshold and parameter settings in the QC routines.

The data were then QC'd with the new routines. The output from the routines was examined again using the graphical and textual data to determine if the routines were able to flag the suspect data without flagging the good data. If the results were not satisfactory, the thresholds and parameters were modified and the process repeated.

2.1.1 Consensus Time

This algorithm is necessary because of a system check that will reset a profiler if its computer time is more than 5 seconds off the time on the central computer located in the Range Operations Control Center (ROCC). When a radar is reset during a consensus period all data collected up to the time of the reset is erased. However, the profiler will continue to collect data through the end of the allotted period and calculate a consensus wind. If the reset occurs toward the end of the period, a horizontal wind estimate is calculated from data collected over a very short time as long as the consensus criteria are met. Analysis of the data showed that wind profiles calculated from these shortened time periods were highly inconsistent in time and space. This reset procedure rarely occurs, but because it does occur the time periods must be checked. This is easily done as the number of minutes in the period are transmitted with the consensus data.

The consensus time algorithm is a very simple check that will flag a data point as suspect if its consensus period is less than the threshold value of six minutes. This value was derived from the standard consensus time period of ten minutes and the consensus constraint that 60% of the samples must be used to calculate a horizontal wind, i.e. 60% of ten minutes is six minutes. This ensures that a sufficient number of samples have been collected to produce a reliable wind estimate.

2.1.2   Precipitation Contamination

Obviously erroneous horizontal winds were frequently calculated during periods of observed rainfall over CCAS/KSC. Large downward vertical velocities and high SNRs were regularly associated with these winds. A review of the literature revealed several sources which confirm that rain does contaminate 915 MHz profiler data (Ralph 1995, Williams et al. 1995, and Ralph et al. 1996). Ralph (1995) concluded that precipitation produces a relatively unambiguous signature in the radial vertical velocities. In addition, a stronger return signal can be used as an additional signature to identify precipitation in the profiles. A stronger return signal can be inferred from high SNR values. As rain occurs frequently in the CCAS/KSC area, this is likely a significant cause of erroneous horizontal wind estimates.

In order to determine appropriate data value thresholds that would indicate rain in profiles, vertical beam consensus radial velocities were plotted against vertical beam consensus SNRs for clear air data and data known to be contaminated by rain. Approximately 15,000 points were used in the analysis. These plots revealed two distinct populations separating clear-air and rain contaminated data points, as seen in Figure 2.1. This allowed development of a discriminant function as found in Panofsky and Brier (1968). This is a linear regression technique in which two or more predictors are used to determine a linear function that will most successfully discriminate between two groups of events. The following equation is used to distinguish between rain contaminated data points and clear-air data points:

$$L = -1.731 + 0.298(VV) + 0.014(SNR) \hspace{3cm} 2.1$$

where VV is the vertical velocity in knots and SNR is in decibels (dB). If the result, L, is positive, the data point is considered to be contaminated by rain and is flagged as suspect.

It is important when using this algorithm to ensure that at least 60% of the individual estimates are used to calculate the consensus radial velocity and SNR in the vertical beam. Otherwise, the values of the estimates may be unrepresentative of the current environment. If less than 60% of the estimates are used, the vertical velocity will not be used in the calculation of the horizontal wind. Therefore, the vertical beam data are checked for the number of samples used in their calculation before being used in the discriminant function. If the number is less than 60% of the total number of estimates, it is not checked in the discriminant function and the horizontal wind estimate is not flagged.

In Figure 2.1 there are a few rain contaminated points located to the left of the discriminant function line in the clear-air region. This is a result of the analysis technique in which all estimates in a profile were assumed to be contaminated by rain if more than 75% of the vertical velocities were above 6 knots and more than 75% of the SNRs were above 15 dB. This assumption caused estimates in those profiles to be included in the rain contaminated data sets that did not meet the consensus constraints as well as those that had values less than 10 knots and 20 dB. The inclusion of these points in the analysis may have created inaccurate constant values in the function, but the results in Section 3 indicate that the equation works well in its current form.
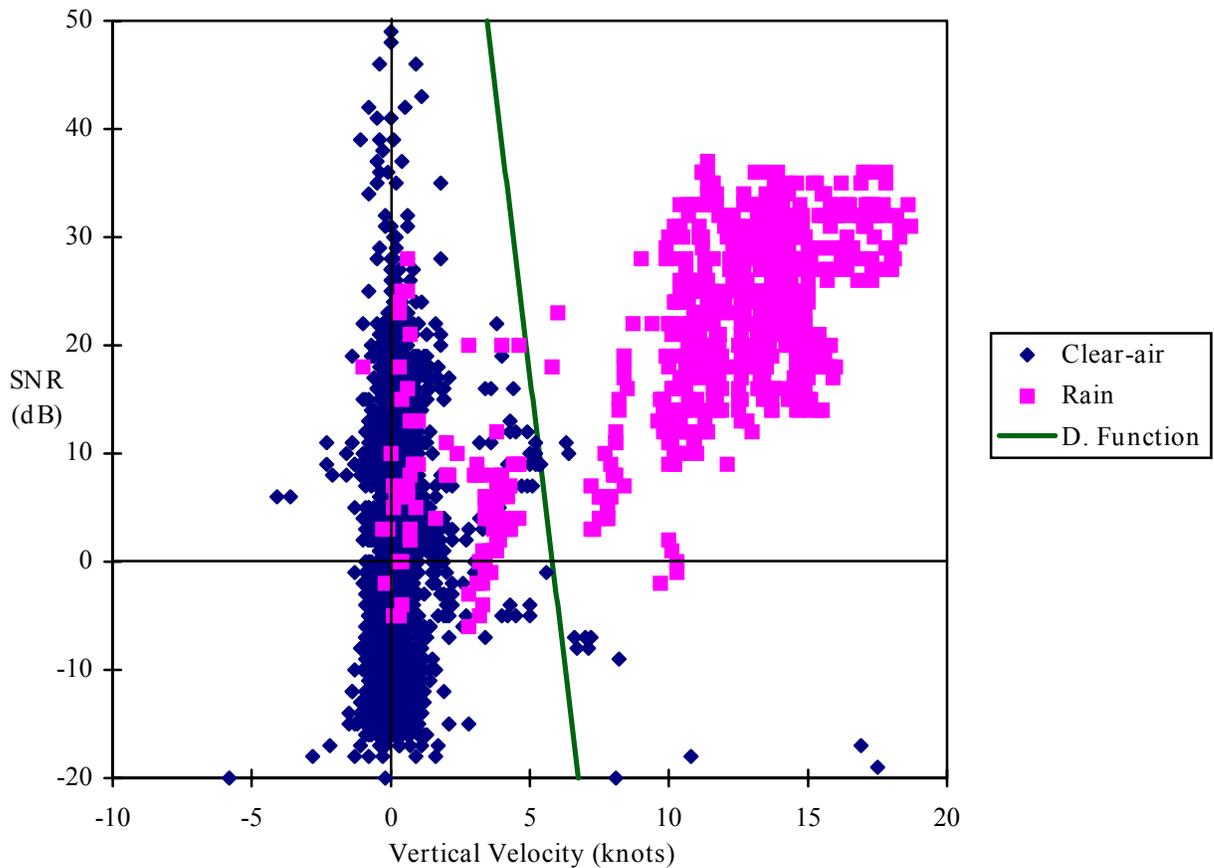
Figure 2.1   Plot of consensus SNR (dB) versus consensus vertical velocity (knots) for known rain contaminated and clear-air data on 17 June 1997 at the False Cape profiler. This data set was used in the development of the discriminant function, which is represented by the thick line. The lines at SNR=0 and VV=0 are drawn for reference.

2.1.3   Median Filter

The median filter described herein is not related in any form to the Median Filter/First Guess (MFFG) algorithm (Wilfong 1993) which is used only on the NASA 50 MHz DRWP data. The MFFG is applied to the individual beam spectral data as a temporal filter to remove outliers. The median filter in this study uses equations that are adaptations of the routines developed in Carr et al (1995). It tests the temporal and spatial consistency of the u- and v-components of the calculated horizontal wind in an attempt to identify suspect data. The reader should not attempt to relate the performance of the MFFG to the performance of the Weber-Wuertz (WW) algorithm based on the comparisons of the median filter algorithm and WW algorithm contained within this report.

The algorithm compares the u- and v-components of a horizontal wind observation to the medians of the u- and v-components of the surrounding (space and time) observations. If the difference between the observed wind component, u_i and/or v_i, and the median of the neighboring observations, u_m and/or v_m, exceeds a critical threshold, T_u and/or T_v, then the wind observation, u_i and v_i, is flagged as suspect.

The critical threshold values, T_u and T_v, are computed as follows:

$$T_u = \max(T_{u1}, T_2) \text{ and}$$

$$T_v = \max(T_{v1}, T_2)$$

where

$$T_{u1} = 0.2 \cdot |u_m + u_i|,$$

$$T_{v1} = 0.2 \cdot |v_m + v_i|, \text{ and}$$

$$T_2 = a \cdot (Ah^2 + Bh + C).$$

In T_2,

a   =   1.3,

h   =   height of the wind observation (feet),

A   = -5.695×10^{-9},

B   = 3.66×10^{-4}, and

C   =   7.3834.

The following two sections describe how the median filter was applied in the post-analysis and simulated real-time modes.

2.1.3.1 Post-Analysis Mode

The number (N) of good wind observations (i.e., not missing or not previously flagged) in the eight surrounding points in a 3X3 space and time grid box (representing 30 minutes and 200 m) is determined (Figure 2.2) for the wind observation under evaluation. If N is greater than 3, then the medians of the u- and v-components of the good wind observations are calculated and the median filter test (described above) is performed.
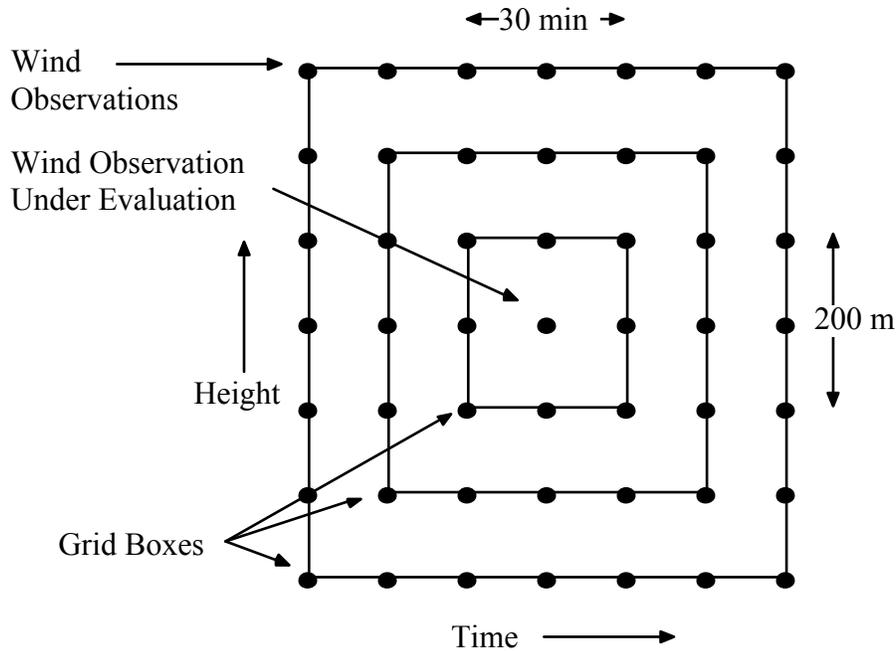

Figure 2.2   Illustration of the method in which the median filter is operated in post-analysis mode. The dots indicate the time-height wind observations and the center dot is the observation being QC'd.

If N is less than 3 for the inner 3X3 box, then the test grid box is enlarged to a 5X5 space and time grid box (representing 1 hour and 400 m) and the same test procedure is invoked. If N is greater than 3 for the 24 surrounding points in the 5X5 grid box, then the medians of the u- and v-components of the good wind observations are calculated and the median filter test is performed.

If N is still less than 3 for the 5X5 grid box, then the test grid box is enlarged for the final time to a 7X7 space and time grid box (representing 1.5 hours and 600 m) and the same test procedure is invoked. If N is greater than 3 for the 48 surrounding points in the 7X7 grid box, then the medians of the u- and v-components of the good wind observations are calculated and the median filter test is performed.

If N is less than 3 for the 7X7 grid box, then the median filter test is not applied and the wind observation under evaluation is not flagged. This process is repeated for all good wind observations in the data set.

2.1.3.2 Real-Time Mode

In real-time mode the grid box used is not symmetrical because the wind observations used in the algorithm must be from the current or previous time periods. Future data are not available as they are in post-analysis mode. The number of good wind observations in the five surrounding points in a 3X2 space and time grid box (representing 15 minutes and 200 m) is determined (Figure 2.3) for the wind observation under evaluation. If N is greater than 3, then the medians of the u- and v-components of the good wind observations are calculated and the median filter test (described above) is performed.

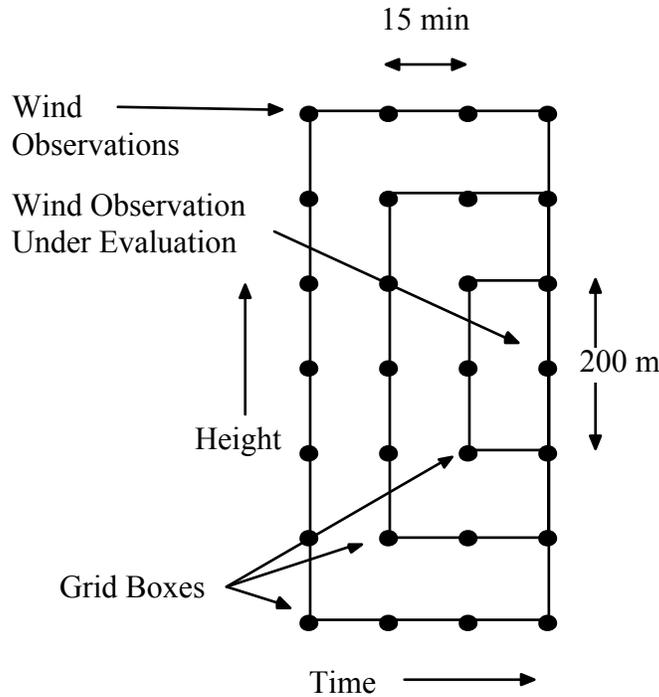Figure 2.3   Illustration of the method in which the median filter is operated in real-time mode. The dots indicate the time-height wind observations and the dot at right-center is the observation being QC'd.

If N is less than 3 for the inner 3X2 box, then the test grid box is enlarged to a 5X3 space and time grid box (representing 30 minutes and 400 m) and the same test procedure is invoked. If N is greater than

3 for the 14 surrounding points in the 5X3 grid box, then the medians of the u- and v-components of the good wind observations are calculated and the median filter test is performed.

If N is still less than 3 for the 5X3 grid box, then test grid box is enlarged for the final time to a 7X4 space and time grid box (representing 45 minutes and 600 m) and the same test procedure is invoked. If N is greater than 3 for the 27 surrounding grid points in the 7X4 grid box, then the medians of the u- and v-components of the good wind observations are calculated and the median filter test is performed.

If N is less than 3 for the 7X4 grid box, then the median filter test is not applied and the wind observation under evaluation is not flagged. This process is repeated for all good wind observations in the data set.

2.1.4   Weber-Wuertz

The Weber-Wuertz QC algorithm (WW) is widely used and well known within the profiler community. It is currently being used to QC real-time data from the NOAA Profiler Network (NPN) (Barth et al. 1997). It therefore has potential to be an appropriate QC algorithm for the 915 MHz DRWP network data. This algorithm uses pattern recognition techniques that identify winds which do not satisfy mathematical definitions of temporal and spatial continuity (Weber and Wuertz 1991).

WW is a mathematically complex routine that recognizes patterns in one- or two-dimensional arrays of any desired data type. In this study, it is used to recognize data patterns in time and space of the consensus radial velocities in each of the three beams. All data in an individual beam data set are input at the beginning of the routine. A set of six subroutines are then used to establish neighbors, branches, and patterns of connected branches in the data set. Neighboring points with similar values are placed in the same branch of a pattern. Neighboring points with large differences in their values are placed in a different branch in the same pattern or in a branch in a completely different pattern. After the patterns are established, four quality control routines are used to flag branches and patterns that are clearly inconsistent with the majority of the data and patterns that contain a very small number of data points.

Patterns are established and QC'd separately in each of the three beams' consensus radial velocity data. If a radial velocity estimate in any of the beams is flagged then the corresponding calculated horizontal wind is flagged, with one exception: if a flagged vertical velocity estimate is not used in the calculation of the horizontal wind, then that wind is not flagged provided that neither of the corresponding radial velocity estimates in the two oblique beams are flagged.

Certain parameters must be set that dictate how WW will establish and QC the patterns. Several iterations of tests using several data sets were performed to determine the appropriate settings. The parameters and their settings are shown in Table 2.1.

Table 2.1 Weber-Wuertz Algorithm Parameter Settings.	
Parameter	*Value*
dx: time space (along radial)	30 minutes (2 time periods) 631.4 feet (194 m or 2 gates)
dy: oblique beams vertical beam	4 knots 2 knots
nmin	32 data points

WW defines dx as the neighborhood size in both time and space and dy as the acceptable change in wind speed over the neighborhood size. In other words, dy/dx can be viewed as the maximum allowable derivative for continuous data within a pattern. Because large changes in wind speed and direction in time and space can actually occur in the boundary layer, relatively small intervals were chosen for the time and

space neighborhood sizes. Smaller values are less restrictive and allow larger changes over larger intervals. There are separate values of change in wind speed for the vertical and oblique beams. Vertical wind speeds tend to be very small (< 1 knot) and the temporal and spatial changes tend to be very small as well. The oblique radial velocities contain a component of the horizontal wind and tend to be much stronger than the vertical velocities with larger changes in magnitude. Thus, dy is smaller for the vertical beam than for the oblique beams. The minimum number of data points any pattern can have is defined by nmin. Patterns with less than 32 data points are considered unreliable and are flagged as suspect by the algorithm.

The following two sections describe how WW was applied in the post-analysis and simulated real-time modes.

2.1.4.1 Post-Analysis Mode

Although any time period can be used, in this study WW checks 24 hours of data at a time in post-analysis mode. This was convenient because each file contains 24 hours of data, which equates to 3072 points (96 profiles X 32 gates). All the data are input to the routine, then the patterns are determined for the whole period. Thus, data collected before and after the wind estimate in question are used to determine whether or not the estimate fits into a pattern.

2.1.4.2 Real-Time Mode

When used in real-time, a QC routine will only check the data in the current profile for reliability. Any QC algorithm that flags data based on continuity, such as WW, only has present and past data available to it. Since it may be considered for use in real-time, the WW routine was modified to QC archived data in a real-time mode. Simulations using archived data sets will give an indication of the performance that can be expected from using WW in real-time.

The technique used in this study is similar to that used in Barth et al. (1997). It uses a sliding time window to establish patterns, then checks all data points in the time window but only retains the results for the current profile. WW must establish new patterns when a new profile is introduced and uses a sliding window in order to use a consistent time period over which patterns are established. The QC information from the previous profiles is retained when establishing new patterns, i.e. wind estimates previously flagged by WW and any other routine within the current time window are not used to quality control the current profile. The time period chosen must be large enough for the algorithm to establish legitimate patterns, yet be small enough to allow the algorithm enough time to process the data before the next profile arrives and so that the current profile is displayed in a timely manner. After several tests a time period of 6 hours, or 25 profiles, was chosen.
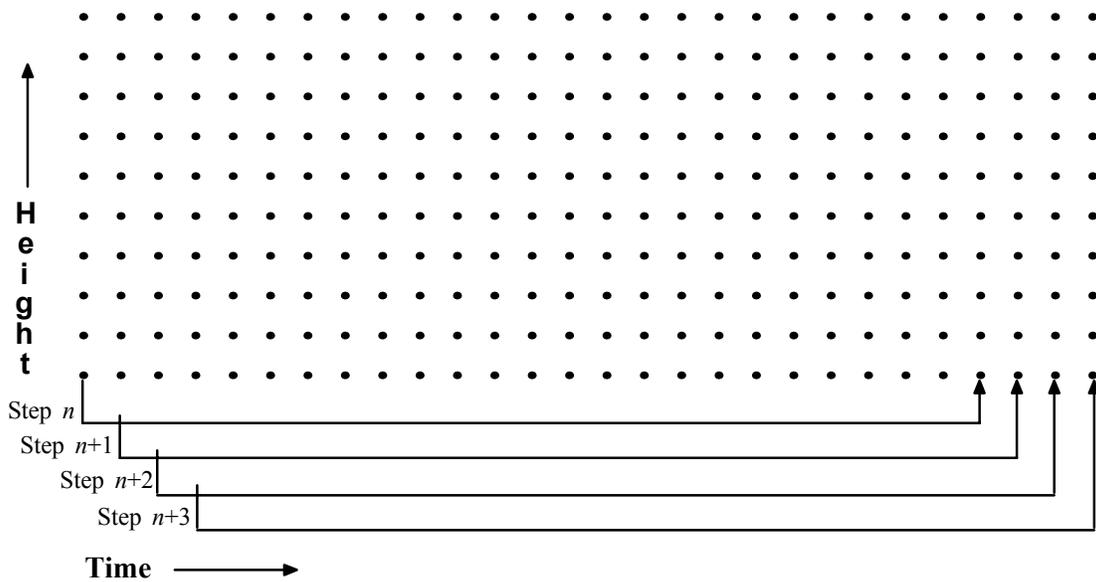
Figure 2.4   Illustration of the method in which WW is operated in real-time mode. The dots indicate the time-height data points, the four brackets below the grid labeled Step *n*, Step *n*+1, and so on represent the sliding time window, and the arrows on the right hand side of the brackets point to the current profile being QC'd.

In Figure 2.4 a grid of dots representing a portion of the data in each of the profiles over 28 time periods is shown. The first bracket, labeled Step *n*, encompasses the 25 profiles used in WW to establish data patterns. These patterns are used to QC the current profile, which is indicated by the arrow on the right-hand side of the bracket. Suspect data are flagged in this profile alone during this step. When a new profile is produced, the window slides to the right, as shown by the bracket labeled Step *n*+1, and new patterns are established in WW as the process is repeated.

2.1.5   SNR Threshold

This routine was considered because low SNR values can indicate a weak returned signal, strong background noise, or combination of both such that the calculated horizontal wind may be regarded as suspect. SNR threshold algorithms have been used in other studies (Merceret 1997) to effectively QC profiler data. The routine developed for this study flagged data that had consensus SNR values below a threshold value of -15 dB. A wind value was flagged as suspect if the consensus SNR in one or both of the oblique beams was below the threshold. The data point was also flagged if the vertical velocity was used in the calculation of the horizontal wind and its SNR was below the threshold.

This routine was eliminated from the analysis in the early stages of the study. It was ineffective at flagging suspect data and tended to flag winds that were consistent in time and space with surrounding wind values. Tests were conducted in which the threshold value of -15 dB was lowered to -20 dB in increments of 1 dB with the same results. It appears that a low consensus SNR value is not a reliable indicator of a suspect consensus wind in this network.

2.2   Effective Routine Combinations

No routine was able to effectively flag suspect data when used on its own. Combinations of the four routines with favorable individual evaluations were examined, and two of those combinations were found most effective:

Combination 1 (CPW)	Combination 2 (CPM)
Consensus Time	Consensus Time
Precipitation Contamination	Precipitation Contamination
Weber-Wuertz	Median Filter

The precipitation contamination and consensus time routines were used in both combinations. Both of these routines were very effective at flagging data that met the specific criteria of each routine. They were not able, however, to flag suspect data that had other sources of contamination. The median filter and WW routines were used in the combinations to flag all other suspect wind estimates.

The algorithms must be used in a specific order. The consensus time period and rain contamination checks do not depend on time or space continuity of the data, but only on the data values associated with the wind estimate being checked. They are used to QC the data first to remove the obviously bad wind estimates. This ensures that WW will not see the areas of bad data as legitimate patterns and that the median filter will not use the bad values in the calculation of the median values. When run in the order shown in the above lists, the algorithms do very well in flagging most of the bad data points while flagging few good data points.

3   Examination of the Results

As stated in the previous section, the routines were developed and evaluated by examining the output in graphical and textual form. In this section the graphical output from the CPW and CPM combinations will be examined. Due to extensive processing time and computer memory limits, and in the interest of brevity, data from the False Cape profiler only are shown. There were many hardware problems with several of the profilers in the 1997 warm season. The False Cape profiler is used because it operated relatively reliably during this period and was the only profiler that was operating on all five days chosen for the evaluation.

3.1   Cases Used in the Analysis

Profiler data for this study were collected during the period 1 May through 31 August 1997. Five days with different weather phenomena were chosen from this period for algorithm development and testing. They are summarized in Table 3.1. The use of data collected in diverse weather conditions allows for a thorough analysis of how the routines respond in different weather regimes.

Table 3.1	Summary of case days and their associated weather phenomena.
Date	***Meteorological Conditions***
2 May 1997	Dry conditions, afternoon sea breeze
12 May 1997	Rain over network all day
15 May 1997	Nocturnal jet, afternoon sea breeze
18 May 1997	Convection near network, afternoon sea breeze
17 June 1997	Nocturnal jet, convection over network, afternoon sea breeze

The following paragraphs describe each of the case days in more detail. The color graphical images used in the analysis (described in Section 2.1) are given with each description. The images contain the raw wind speed and direction data before the QC routines were applied.

2 May 1997

2 May was characterized by very dry air above 6000 ft. As can be seen in the False Cape profiler data in Figures 3.1 and 3.2, this resulted in a limited number of consensus winds above 6000 ft and also a small number of consensus winds below 6000 ft until 1600 UTC. After that time the boundary layer became more moist and likely more turbulent with the afternoon heating, thus creating conditions more conducive to producing consensus winds. A sea breeze occurred although the wind speed (Figure 3.1) and direction (Figure 3.2) changes associated with its passage were not abrupt. A gradual change in direction from northerly to easterly in the lowest 2000 ft occurred during the period from 1200 UTC to 1700 UTC. Subsequently, the flow gradually became southeasterly and remained that way through the end of the day. Wind speeds in the sea breeze layer were less than 10 knots from 1200 to 1700 UTC, but eventually increased to 15 knots by 2200 UTC. This day was chosen to evaluate the performance of WW and the median filter in the presence of large amounts of missing data.
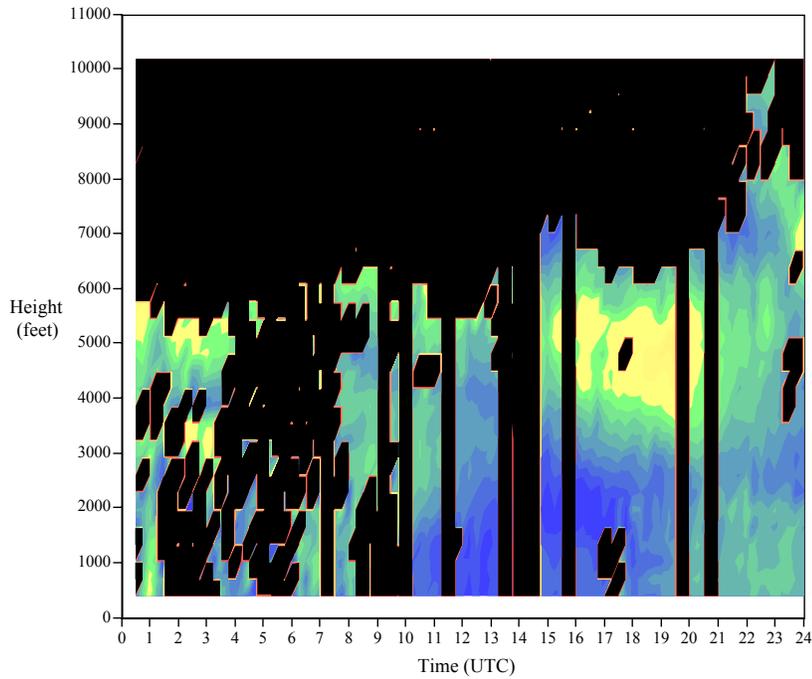
Figure 3.1   Wind speeds from the False Cape profiler on 2 May 1997. Speeds are in knots (legend at right). The warm colors represent winds greater than 20 knots and the cool colors represent winds less than 20 knots. Areas of black indicate gates where a consensus wind was not calculated.
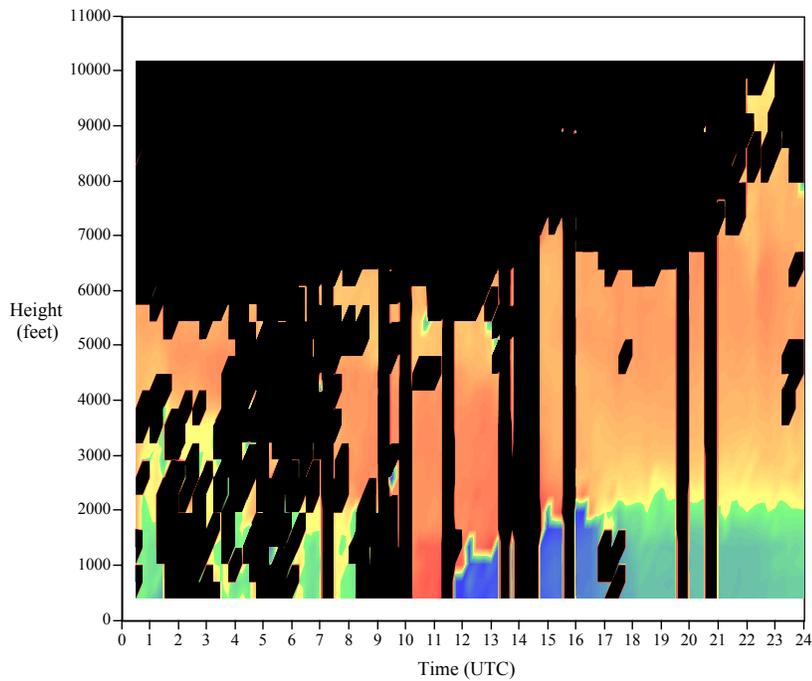

Figure 3.2   Wind directions from the False Cape profiler on 2 May 1997. Directions are in degrees (legend at right). The warm colors represent winds with a westerly component and the cool colors represent winds with an easterly component. Areas of black indicate gates where a consensus wind was not calculated.

12 May 1997

Rain was observed over CCAS/KSC from 0830 UTC to approximately 2300 UTC on 12 May. Every profiler in the network was affected by this rain over the entire period. The actual wind speeds (Figure 3.3) and directions (Figure 3.4) were difficult to resolve due to the rain contamination of the data. This contamination can be seen in the False Cape profiler data mainly as temporally inconsistent wind speeds and directions from profile to profile. This day was chosen primarily to evaluate the performance of the rain contamination check. This check, which is used first, would likely flag a large amount of data as bad. The flagged data would be seen as missing by WW and the median filter, which are used last. Therefore, this also provides an evaluation of WW and the median filter in the presence of large amounts of missing data.
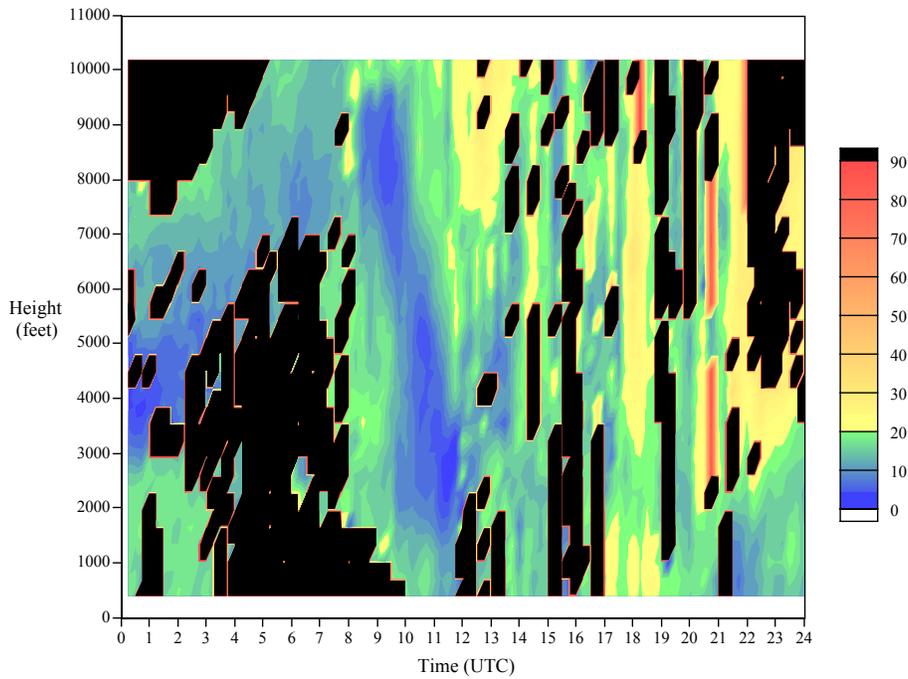

Figure 3.3   Wind speeds from the False Cape profiler on 12 May 1997. Speeds are in knots (legend at right). Areas of black indicate gates where a consensus wind was not calculated.

Figure 3.4   Wind directions from the False Cape profiler on 12 May 1997. Directions are in degrees (legend at right). Areas of black indicate gates where a consensus wind was not calculated.

15 May 1997

15 May is also characterized by somewhat dry conditions with few consensus winds above 7000 ft and limited consensus winds up to 1400 UTC at all levels in the profiles. The other meteorological phenomena that occurred on this day were an early morning southerly low-level jet and a sea breeze passage, which can be seen in the False Cape profiler wind speed and direction data in Figures 3.5 and 3.6, respectively. The low-level jet was contained below 3000 ft and reached its peak speed of 25 knots between 0300 and 0530 UTC (Figure 3.5). The sea breeze passage was indicated by a veering of the winds below 2000 ft from northerly to easterly by 1530 UTC and continuing to southeasterly by the end of the period (Figure 3.6). Speeds in the sea breeze at 1530 UTC were less than 10 knots, but gradually increased to between 15 and 20 knots by 2200 UTC. This day was chosen to evaluate the performance of WW and the median filter when there are changes in wind speed and direction, particularly at the sea breeze front. If either of these algorithms flagged the sea breeze passage as suspect, it would be considered unacceptable and the algorithm would be modified or eliminated from use.

Figure 3.5   Wind speeds from the False Cape profiler on 15 May 1997. Speeds are in knots (legend at right). Areas of black indicate gates where a consensus wind was not calculated.
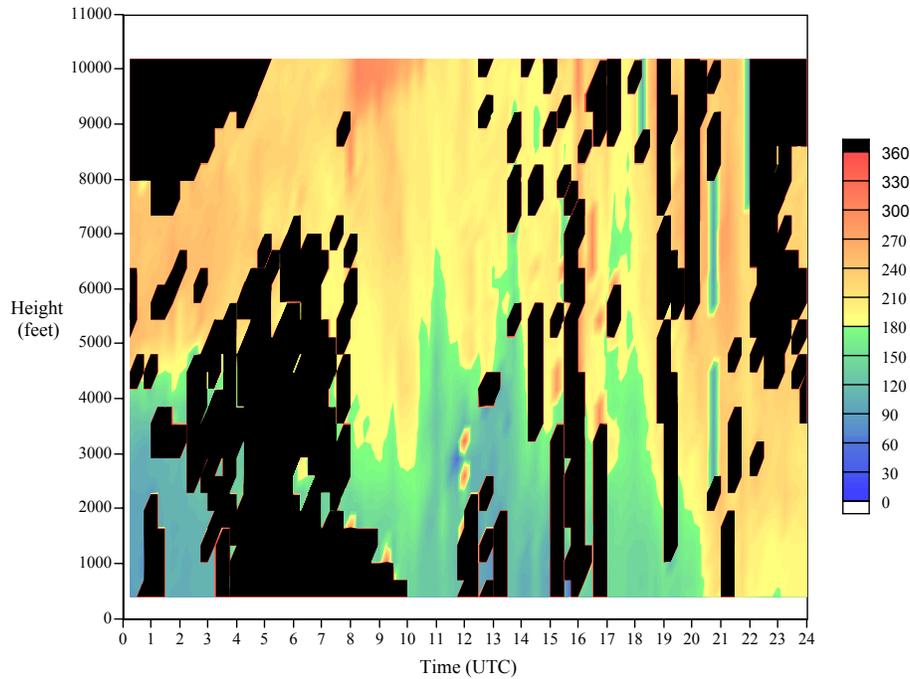

Figure 3.6   Wind directions from the False Cape profiler on 15 May 1997. Directions are in degrees (legend at right). Areas of black indicate gates where a consensus wind was not calculated.

18 May 1997

18 May was more moist than the previously discussed days. Convection occurred during the afternoon west of CCAS/KSC, but none occurred over any of the profilers. As can be seen in Figures 3.7 and 3.8 the False Cape profiler was able to produce consensus winds most of the time at all levels up to 0600 UTC, but only limited consensus winds above approximately 8000 ft after 0600 UTC. The winds were light and variable, but had an easterly and/or southerly component at most levels over the entire period. This made it difficult to detect the sea breeze in the direction data (Figure 3.8), but it could be seen in the speeds at the False Cape (Figure 3.7). The easterly winds began to increase below 3000 ft at 1430 UTC from less than 5 knots to 15 knots by 1700 UTC. The easterly winds continued increasing below 4000 ft to a maximum of 20 knots by 2300 UTC. This day was chosen to evaluate the performance of WW and the median filter in light and variable wind conditions and during the passage of a sea breeze.
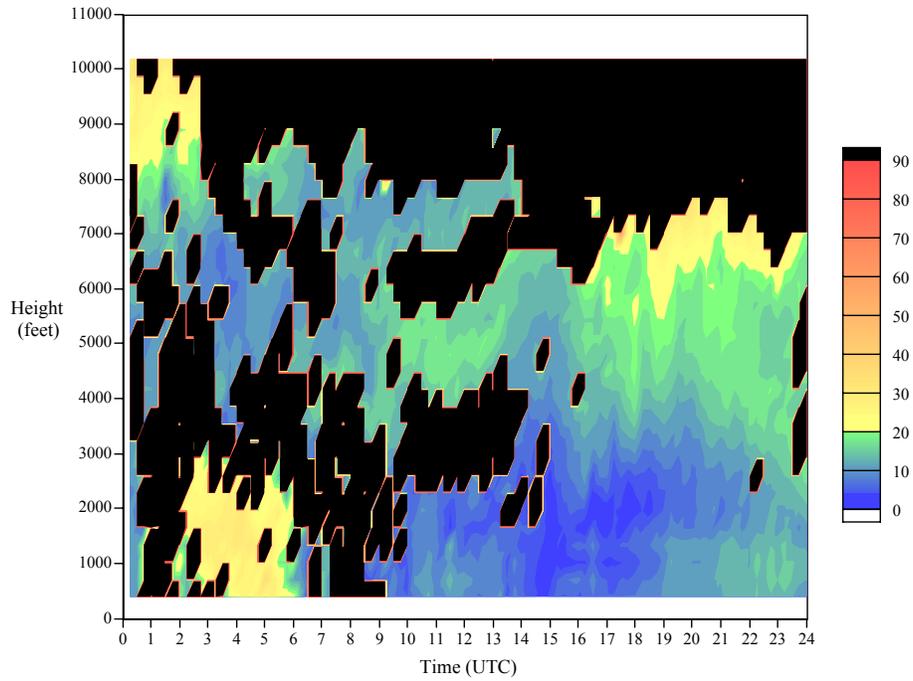

Figure 3.7   Wind speeds from the False Cape profiler on 18 May 1997. Speeds are in knots (legend at right). Areas of black indicate gates where a consensus wind was not calculated.
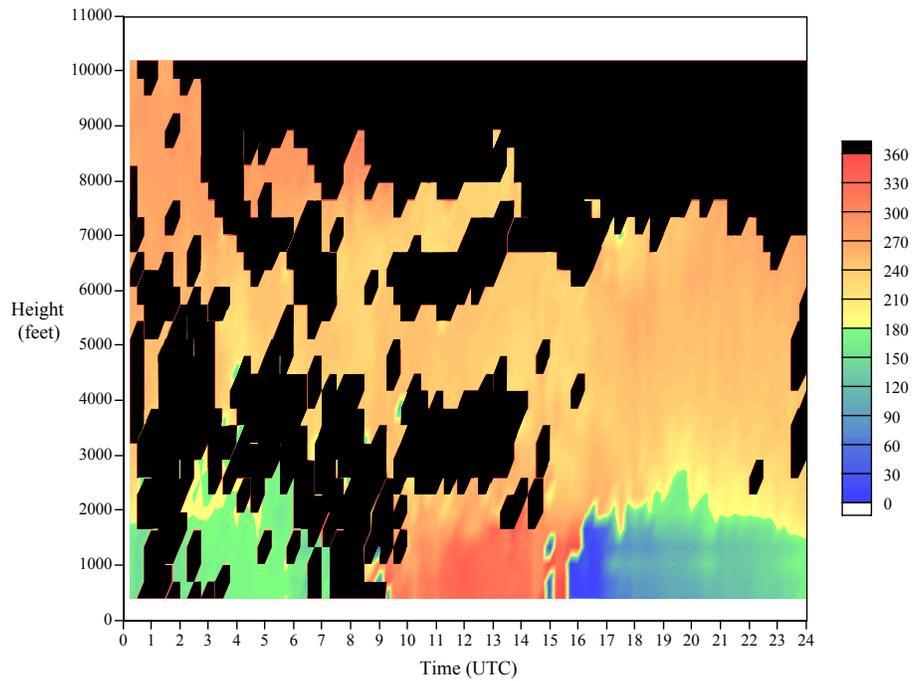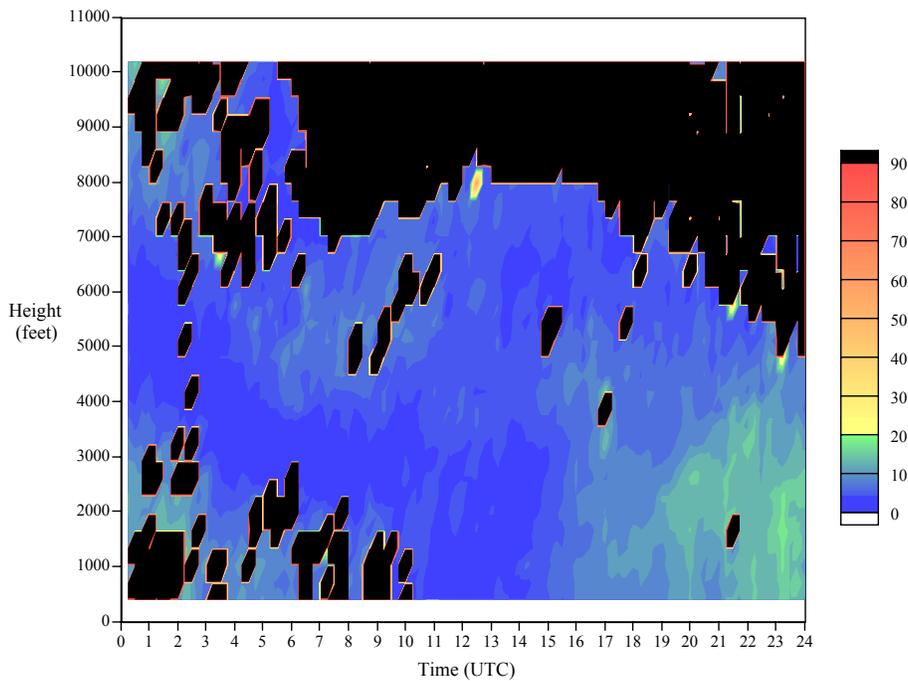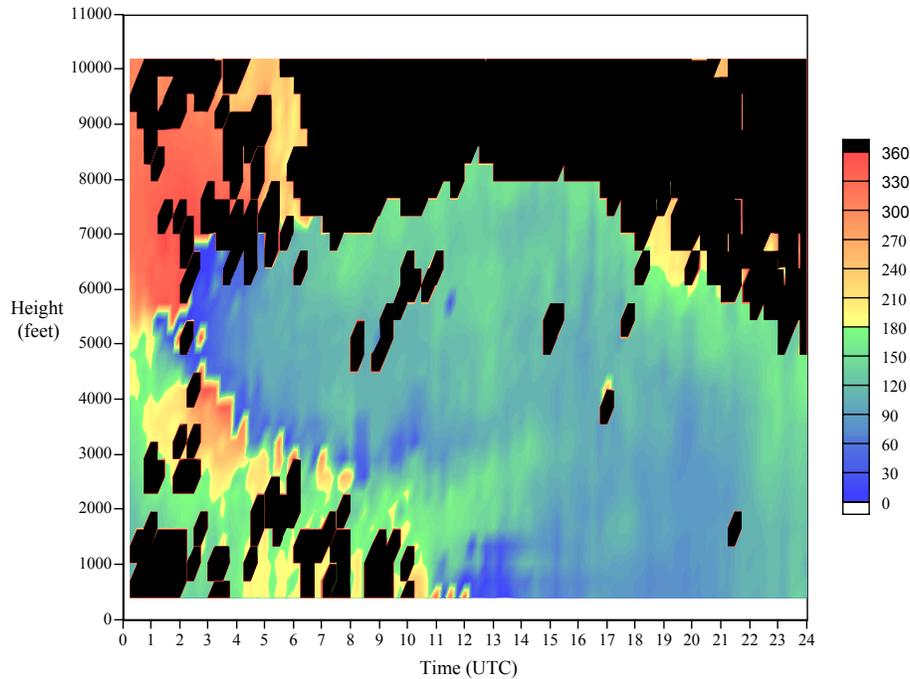
Figure 3.8   Wind directions from the False Cape profiler on 18 May 1997. Directions are in degrees
(legend at right). Areas of black indicate gates where a consensus wind was not calculated.

17 June 1997

Several interesting meteorological phenomena occurred on 17 June. The formation and decay of a
southerly low-level jet was observed during the early hours of the day, a sea breeze passed in the late
morning hours, and convection developed over the network through most of the afternoon and early
evening. Both the wind speeds (Figure 3.9) and directions (Figure 3.10) from the False Cape profiler
clearly illustrate these features.

The winds were generally westerly throughout the 24-hour period. The sea breeze from the previous
afternoon can be seen between 0000 and 0200 UTC as easterly winds (Figure 3.10) below 2000 ft that
were weakening (Figure 3.9) with time. The winds began to veer and become more southerly and the
speeds increased from 0200 to 0500 UTC. The jet had a maximum speed of over 20 knots from the south
between 0500 and 0700 UTC (0100-0300 EDT) and below 3000 ft. The winds veered and weakened until
they were westerly at 10 knots by 0900 UTC. This was also observed in the Merritt Island profiler data,
the only other profiler operating that day. The sea breeze passed abruptly at 1530 UTC as noted by a
change in wind direction below 2000 ft from westerly to easterly. The wind speeds were quite weak, but a
slight increase occurred just after the wind direction change. Convection over the False Cape profiler
began just after 1700 UTC and continued through 2200 UTC. The rain contamination can be seen as
temporally inconsistent wind speeds and directions from profile to profile during this period.

This case was chosen for the evaluation to test the precipitation contamination check and to test WW
and the median filter in the presence of the wind shifts associated with the low-level jet and the sea breeze.
A large amount of data will be flagged by the precipitation contamination check which will be seen as
missing data by WW and the median filter. Thus, the effect of missing data on the performance of WW
and the median filter will also be tested.
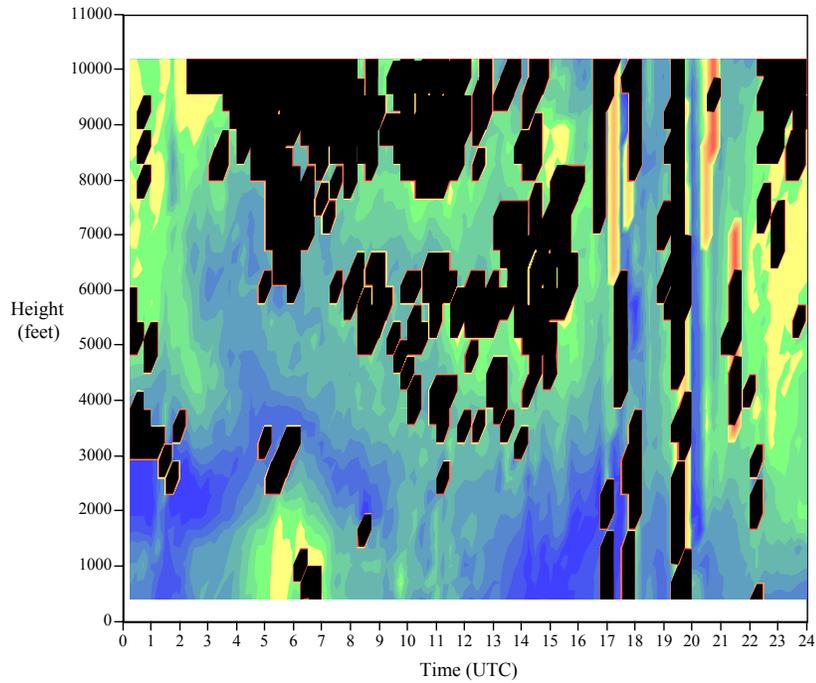
23

Figure 3.9   Wind speeds from the False Cape profiler on 17 June 1997. Speeds are in knots (legend at right). Areas of black indicate gates where a consensus wind was not calculated.
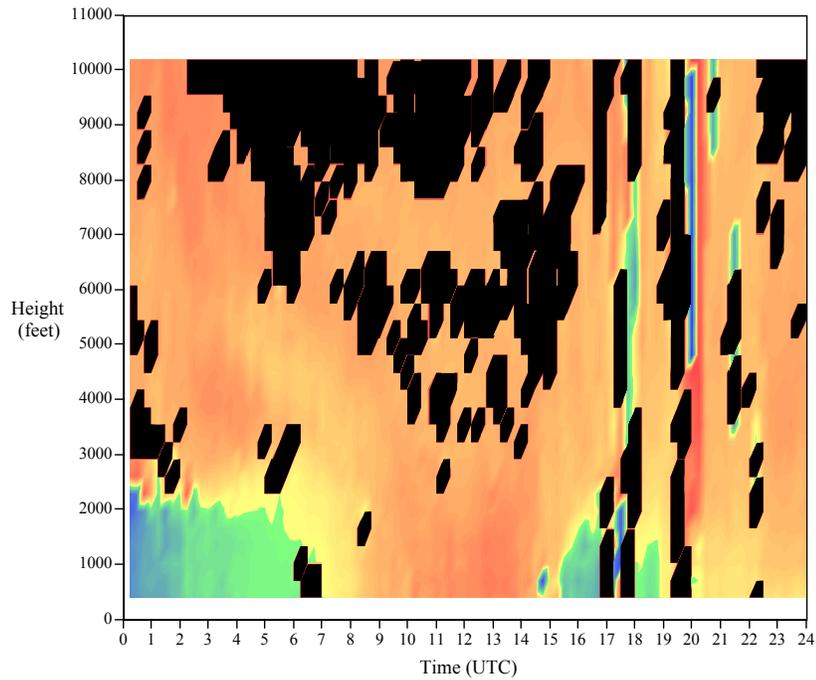

Figure 3.10   Wind directions from the False Cape profiler on 17 June 1997. Directions are in degrees (legend at right). Areas of black indicate gates where a consensus wind was not calculated.

3.2 Post-Analysis Mode

In post-analysis mode, the data are quality controlled after they are collected and archived. This means that WW and the median filter can evaluate continuity with wind estimates collected after the wind estimate being checked.

Table 3.2 summarizes the statistics for the quality control of the False Cape profiler data on each of the case days. The first two rows give a summary of the consensus data before the QC. The first row shows the number of consensus wind estimates and the second row shows the number of times the profiler was unable to produce a consensus wind estimate, or missing data. The next four rows show the number of wind estimates flagged (i.e., selected as suspect) by each routine when used in the CPM and CPW combinations. The fifth row shows the number of common wind estimates flagged by CPM and CPW. These are the wind estimates that were flagged as suspect by both WW and the median filter. Finally, the last two rows show the number of consensus wind estimates left after the QC combination runs.

The consensus time and precipitation contamination routines flag the same number of points in both combinations. This is because they are independent checks: they do not depend on time and space continuity of the data but only on the value of a wind estimate's consensus time period or vertical beam radial velocity and SNR. The consensus time routine flagged no wind estimates on any of the days in False Cape profiler data. The rain contamination algorithm performed very well flagging very few wind estimates on days with no rain (2 May, 15 May, 18 May) and flagging a large number of wind estimates on days when rain was observed (12 May and 17 June). An examination of the data values (not shown) reveals skillful performance of this algorithm to flag the rain contaminated wind estimates.

Table 3.2	Post-analysis QC summary for the False Cape Profiler. The total possible number of consensus winds for a 24-hour period is 3072.				
Point Summary	***2 May***	***12 May***	***15 May***	***18 May***	***17 June***
Data Prior to QC:					
Consensus Wind Estimates	1438	2356	1928	2255	2509
Missing Data	1634	716	1144	817	563
Number of Consensus Winds Flagged by Each Routine:					
Consensus Time	0	0	0	0	0
Precipitation Contamination	1	1271	3	4	326
Median Filter (CPM)	37	17	22	6	49
Weber-Wuertz (CPW)	35	157	16	26	55
Estimates Flagged by CPM and CPW	14	14	8	4	44
Number of "Good" Consensus Winds After QC:					
Number of "Good" Estimates - CPM	1400	1068	1903	2245	2134
Number of "Good" Estimates - CPW	1402	928	1909	2225	2128

The median filter and WW performed similarly in number of wind estimates flagged on most days. Note, however, the large difference on 12 May in which WW flagged 140 more wind estimates than the median filter. The bulk of this difference can be explained in the way both algorithms treat areas of isolated data. In the case of the median filter, if there are not enough wind estimates to determine a median value, the wind estimate is not flagged. However, if a pattern contains less than 32 wind estimates, WW will flag all wind estimates in that pattern. This can occur in areas where there are small groups of wind estimates surrounded by many missing (or previously flagged) winds. The WW algorithm can not connect these isolated groups with other patterns in the data. After the precipitation contamination algorithm was run in the 12 May data, such small pockets of isolated data existed.

This is best illustrated by comparing Figures 3.11 and 3.12 (see Figure 3.3 for the raw wind speed data). In Figure 3.11 there are some data left by CPM between 0600 and 0800 UTC at 2000 ft. CPW identified these isolated consensus winds as patterns with less than 32 wind estimates and flagged them as suspect. There are other areas of data seen as vertical lines at 1200, 1645, 2015, and 2130 UTC in Figure 3.11 that are not seen in Figure 3.12 for this reason.
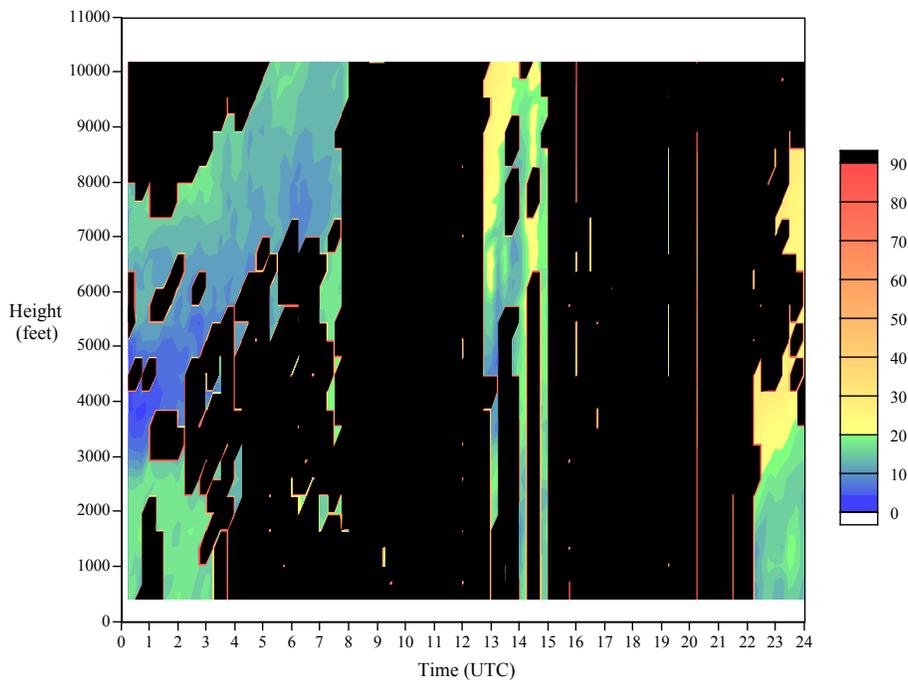

Figure 3.11   Wind speeds from the False Cape profiler on 12 May 1997 after CPM. Speeds are in knots (legend at right). Areas of black indicate gates where a consensus wind was not calculated and where the data were flagged by the algorithms.
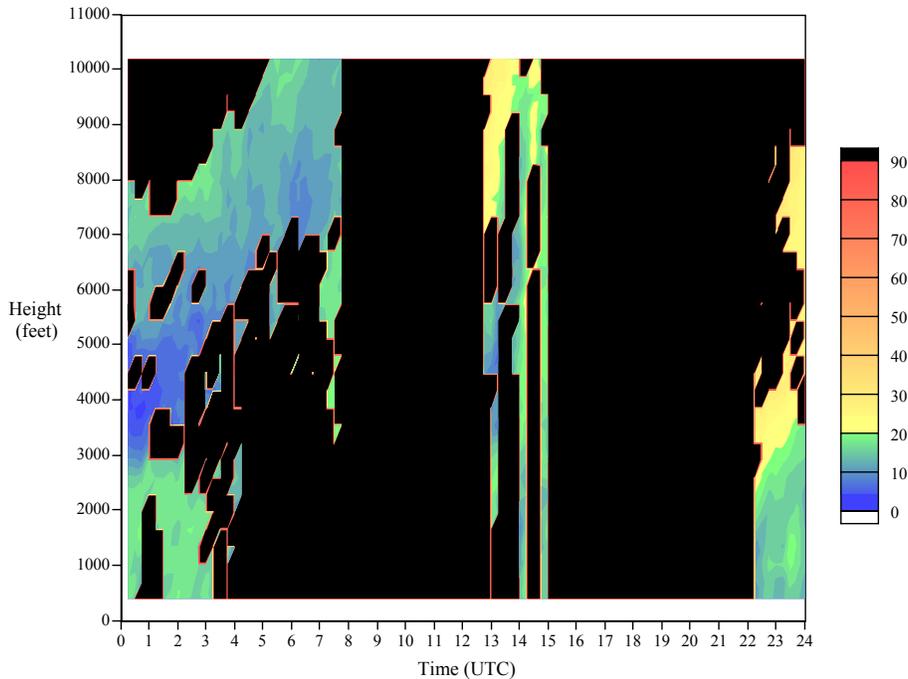
Figure 3.12   Wind speeds from the False Cape profiler on 12 May 1997 after CPW. Speeds are in knots
(legend at right). Areas of black indicate gates where a consensus wind was not calculated
and where the data were flagged by the algorithms.

Another important item to note from Table 3.2 is the number of wind estimates commonly flagged by CPM and CPW. In an ideal situation, erroneous radial velocities would create an erroneous horizontal wind to the extent that the wind estimate in question would be flagged by both algorithms. The commonly flagged wind estimates in Table 3.2 likely represent such data, but the other flagged wind estimates present a dilemma. The critical question to be answered is why would a wind estimate be considered suspect by one algorithm and not the other. The answer lies in the differences in the algorithms and data types used in the routines (refer to Section 2). In summary, the median filter compares the median of the u- and v-components of the winds using the surrounding points in a grid box with the u- and v-components of the center point. WW uses a pattern recognition technique which evaluates each of the three radial beam velocities to determine if a wind estimate is erroneous.

A thorough quantitative analysis was not done to determine exactly why a particular wind was flagged by one algorithm and not the other. But some of the reasons can be deduced based on the algorithms' procedures. It is important to note that if two adjacent radial velocity estimates do not meet the continuity requirements to be in the same pattern in WW, they may still be part of different acceptable patterns and not be flagged. However, a large difference between adjacent radial velocities could create a large difference between adjacent horizontal wind estimates. The median filter would, therefore, flag one of those wind estimates provided the median of the 3X3 grid box is largely influenced by points that are part of one of the WW patterns. Conversely, WW can flag data not flagged by the median filter when there are isolated small patterns of data, as discussed above.

WW flags wind estimates based on discontinuities in the vertical radial velocity, which the median filter does not consider. This can also cause wind estimates to be flagged by WW but not the median filter. Figures 3.11 and 3.12 show an example of this. There are wind speed estimates at 6000 ft to 7000 ft between 1300 and 1500 UTC that change rather quickly over time. These were flagged by WW but not the median filter. The data show that there were changes in the vertical velocity at these times and heights that were significant enough to be flagged by WW. However, the resulting horizontal wind change was not great enough to be flagged by the median filter.

These are very likely just a few of the reasons why certain wind estimates are not commonly flagged by both algorithms. There may be other reasons, but a quantitative analysis of a very large data set from all the profilers would need to be done to determine the range of possible causes.

3.3 Real-time Mode

In order to test the routines in real-time mode, they were modified to use only previously collected data in their continuity checks. Archived data were then input to the routines for the simulated real-time QC tests. The differences between the real-time CPW and CPM are presented in this section.

Table 3.3 summarizes the statistics for the quality control of the False Cape profiler data on each of the case days. As with Table 3.2, the first two rows give a summary of the consensus data before the QC, the next four rows show the number of wind estimates flagged by each routine, the fifth row shows the number of common wind estimates flagged by CPM and CPW, and the last two rows show the number of "good" consensus wind estimates left after the QC combination runs. The consensus time and precipitation contamination routines flag the same wind estimates in both combinations and in both modes (see Table 3.2). Because they are independent checks they will flag the same winds in the same data set whether in real-time or post-analysis mode.

Table 3.3	Real-time analysis QC summary for the False Cape Profiler. The total possible number of consensed winds for a 24-hour period is 3072.				
Point Summary	**2 May**	**12 May**	**15 May**	**18 May**	**17 June**
Data Prior to QC:					
# of Consensed Points	1438	2356	1928	2255	2509
# of Non-consensed Points	1634	716	1144	817	563
Number of Consensus Winds Flagged by Each Routine:					
Consensus Time	0	0	0	0	0
Precipitation Contamination	1	1271	3	4	326
Median Filter (CPM)	54	21	26	5	61
Weber-Wuertz (CPW)	32	139	19	8	110
# Common Points CPM/CPW	8	11	6	3	30
Number of "Good" Consensus Winds After QC:					
Number of "Good" Estimates - CPM	1383	1064	1899	2246	2122
Number of "Good" Estimates - CPW	1405	946	1906	2243	2073

A comparison between Tables 3.2 and 3.3 reveals many similarities in the number of wind estimates flagged by CPW and CPM, as well as in the number commonly flagged by both combinations. The cause of the large difference in the number of wind estimates flagged between CPW and CPM on 12 May is the same as that described in Section 3.2, as is the disparity in the number of commonly flagged winds. A new large difference of 49 wind estimates between CPW and CPM appeared on 17 June. Figures 3.13 and 3.14 show the wind speeds after CPM and CPW, respectively. For comparison, the raw wind speeds for 17 June can be found in Figure 3.9. The large difference in the number of wind estimates flagged can be found mainly in the two profiles at 1815 and 2000 UTC.

CPW flagged the two complete profiles at 1815 and 2000 UTC. Both of these profiles came immediately after a group of missing profiles that had been flagged by the precipitation contamination routine. Most of the winds were flagged because they did not fit the patterns that existed before the missing wind profiles. CPM did not flag any wind estimates in the 1815 UTC profile, and flagged 11 wind estimates in the 2000 UTC profile between 4000 - 5500 ft, 7500 - 8500 ft, and the top-most wind estimate (see white arrows in Figure 3.13). At 1815 UTC the missing data affected CPM such that it would only have the data above and below a wind estimate in the same profile to calculate the median (see Figure 2.3). The winds in this profile were spatially consistent at all levels and none were flagged. This was not the case for the profile at 2000 UTC. Substantially stronger values from the 1915 UTC profile were used in the calculation of the medians for the 2000 UTC wind estimates causing several of the wind estimates at this time to be flagged. CPW flagged two complete profiles (64 wind estimates) in which CPM only flagged 11 wind estimates. This accounts for most of the difference in the number of wind estimates flagged between the two combinations on 17 June.

Note that CPW did not flag the 6 wind estimates at 2030 UTC between 6500 - 8500 ft (see white arrow in Figure 3.14). These data are obviously erroneous and should have been flagged. The entire profile at 2030 UTC was affected by rain and all other wind estimates in the profile were flagged by the precipitation contamination algorithm. The data in question were not flagged by the precipitation contamination algorithm because less than 60% of the individual beam estimates were used in the calculation of the consensus vertical velocity (see Section 2.1.2). WW will also not flag a wind estimate based on an inconsistent vertical velocity if less than 60% of the estimates were used in its calculation. The oblique beam consensus velocities fit into previous patterns in WW and, therefore, the wind estimates were not flagged by this algorithm. The horizontal wind estimates, however, are highly inconsistent with the estimates in the previous profiles. Therefore, CPM successfully flagged these erroneous estimates because the difference between their u- and v-components and the median u- and v-components exceeded the threshold value in the median filter routine.

Figure 3.13   Wind speeds from the False Cape profiler on 17 June 1997 after real-time CPM. Speeds are in knots (legend at right). Areas of black indicate gates where a consensus wind was not calculated and where the data were flagged by the algorithms.
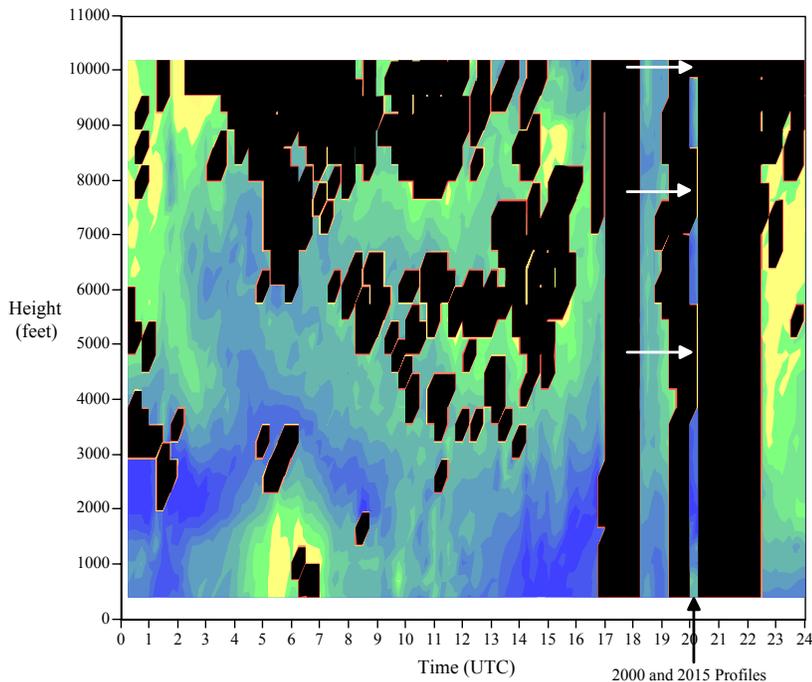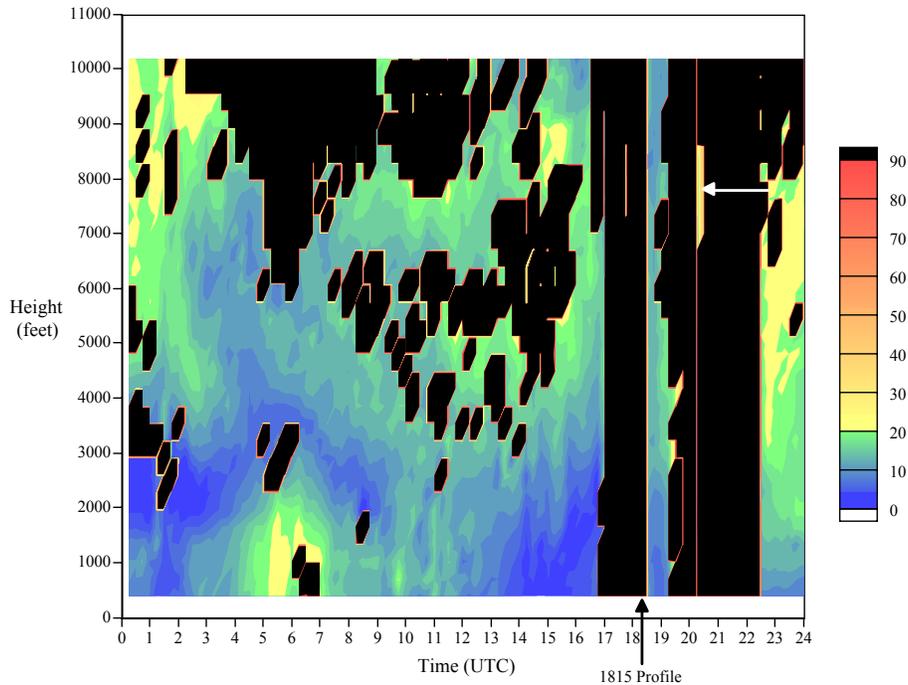
Figure 3.14   Wind speeds from the False Cape profiler on 17 June 1997 after real-time CPW. Speeds are in knots (legend at right). Areas of black indicate gates where a consensus wind was not calculated and where the data were flagged by the algorithms.

3.4   Mode Comparison

The differences in the way CPW and CPM input data in real-time and post-analysis mode have caused some differences in which wind estimates are flagged. One difference is that there are fewer wind estimates used in real-time mode to calculate the medians in the median filter or determine patterns in WW. In the median filter, up to approximately 40% less wind estimates are available to use in the calculation of the medians in real-time mode than in post-analysis mode. The post-analysis WW inputs 24 hours of data before calculating patterns whereas the real-time WW inputs 6 hours of data. Another difference is that neither real-time routine can use wind estimates collected after the estimate being evaluated to check for continuity. Thus, no future trend information is available to influence the determination of patterns in WW or the value of the medians in the median filter.

A comparison of four of the five case days shows very few differences in performance between the two combinations in the real-time and post-analysis modes. However, there is a notable difference in performance between the real-time and post-analysis WW on 17 June. Figure 3.15 shows the 17 June wind speed results from the post-analysis WW (the raw data and real-time WW output are found in Figures 3.9 and 3.14, respectively). The good profiles at 1815 and 2000 UTC that were flagged in the real-time mode were not flagged in the post-analysis mode, and the bad wind estimates between 6500 ft – 8500 ft at 2030 UTC that were not flagged in the real-time mode were flagged in the post-analysis mode. These differences are a result of the fact that future trend information was available to WW in the post-analysis mode and it was, therefore, able to correctly determine which wind estimates were erroneous and which were not.
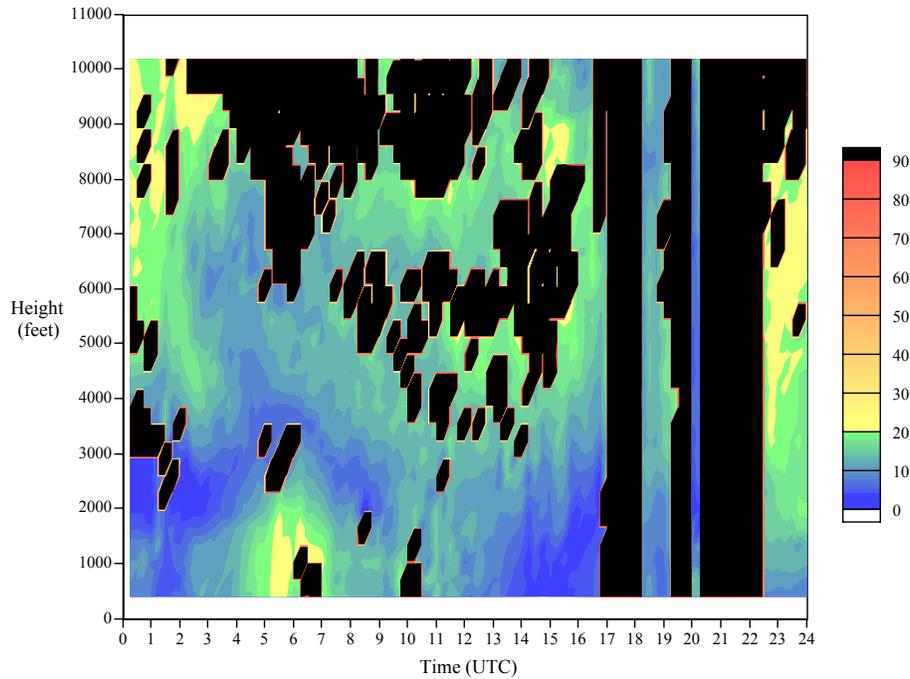
30

Figure 3.15   Wind speeds from the False Cape profiler on 17 June 1997 after post-analysis CPW. Speeds are in knots (legend at right). Areas of black indicate gates where a consensus wind was not calculated and where the data were flagged by the algorithms.

It is important to note that even though the differences in algorithm performance between the two modes for the other four days were small, there was usually a noticeable degradation in the output from CPW. The differences between the modes from CPM were very small and not noticeably worse for the real-time mode. A possible reason for this is that the median filter uses fewer wind estimates than WW to evaluate continuity in the post-analysis mode. The reduction in the number of wind estimates in the real-time mode might then be relatively small compared to that of WW and may have only a small effect on the median values. This would result in very few differences between the output of the real-time versus the post-analysis mode in CPM.

More wind speed and direction images to facilitate a comparison between the real-time and post-analysis modes are contained in the Appendix.

31

4   Summary and Conclusions

There is currently no quality assessment of the data from the 915 MHz DRWP network at CCAS/KSC. Because of the existence of many possible contaminants of these data, quality assessment routines are needed. In this study, the AMU has successfully developed and/or configured routines that are able to identify most of the unreliable data and little of the good data as suspect.

Several routines were tested both individually and in combination with other routines.   Two independent routines that check the value of certain data types and two dependent checks that check the space and time continuity of the wind data were analyzed. The independent checks include

- A consensus time algorithm that checks the number of minutes in a profile's consensus period (Section 2.1.1) and

- A precipitation contamination algorithm that determines whether a wind estimate is erroneous due to the presence of rain (Section 2.1.2).

The dependent checks include

- The Weber-Wuertz algorithm, a well known and widely used pattern recognition program (Section 2.1.4) and

- A median filter that compares the value of the u-and v-components of a wind estimate to the median value of the components of the surrounding wind estimates (Section 2.1.3).

It is important to understand that the median filter used in this study is not related in any form to the MFFG algorithm used on NASA's 50 MHz DRWP (see Section 2.1.3 for a brief description).   Therefore, a comparison between the performances of the MFFG and WW algorithms cannot be made based on the results of the median filter and WW algorithm comparisons contained within this report.

Perhaps the most important result of this study is that no single routine performed satisfactorily on its own. The independent routines each check for one specific type of contaminant and cannot flag erroneous wind estimates caused by other contaminants. The dependent routines can flag an erroneous wind caused by any contaminant as long as it is inconsistent with the surrounding winds. However, an erroneous wind estimate caused by a long-lasting contaminant, such as rain, may be seen as consistent with the other surrounding winds and not flagged. The best results are obtained when the two independent routines are run first followed by one of the dependent routines. This ensures that the obviously bad wind estimates are flagged first so that WW will not see them as areas of legitimate patterns and the median filter will not use the bad wind values in the calculation of the median values.

4.1   Advantages and Disadvantages of Each Routine

Although the routines perform well when used in the CPM and CPW combinations, each routine has certain disadvantages that can adversely affect the quality assessment results.

The advantage of the consensus time period check is that it is effective at flagging erroneous profiles caused by too few samples from which to estimate correct winds. This algorithm has one disadvantage. Profiles calculated from shortened consensus time periods usually appear inconsistent with previous and subsequent profiles. However, there may be times when the profile will be consistent with the other profiles and will be correct. It is possible that the correct winds will be measured at some, if not all, of the gates in a profile with only one or two samples. Even though these winds would be consistent and correct, they would be flagged. Since this algorithm is rarely invoked, good wind estimates will rarely be flagged.

As rain occurs quite frequently over the CCAS/KSC area, the precipitation contamination algorithm is necessary and has proven to be very effective at flagging individual erroneous wind estimates contaminated by rain. This algorithm will not check the vertical beam radial velocity and SNR if less than 60% of the individual beam estimates are used in the calculation of the values. This is because these vertical velocities will have no effect on the horizontal wind estimate as they will not be used in the

horizontal wind calculation (Equations 1.1 – 1.2, Section 1.3). This can be both an advantage and disadvantage. The advantage is that many times these vertical velocities and SNRs may be consistent with those found in rain, but will actually be erroneous and inconsistent with surrounding values. The horizontal wind estimates will be unaffected by these vertical velocities and will not be flagged. The disadvantage occurs in rain-contaminated profiles when accurate vertical velocities and SNRs are calculated from less than 60% of the individual beam estimates. This occurred in the 17 June data as can be seen in Figure 3.10. The six bad wind estimates at 2030 UTC between 6500 - 8500 ft would have been flagged by this algorithm if the minimum individual beam limit had not been considered. Instead, the bad wind estimates were assumed to be good.

The median filter and WW both check the temporal and spatial consistency of the wind estimates. If a wind estimate does not meet the consistency requirements of either algorithm, it is flagged as suspect. These algorithms rely on the assumption that the ratio of bad data to good is small. These algorithms are at a disadvantage when there is a large amount of bad data present. Erroneous wind estimates can influence the consistency checks to the point that good wind estimates are flagged as bad and bad wind estimates are not flagged. It is for this reason that the independent checks are run first to remove any large blocks of obviously bad data. Nonetheless, a large amount of erroneous wind estimates caused by contaminants other than rain or a short consensus time period may exist in the data set which could degrade the performance of the median filter or WW.

The median filter has another disadvantage not found in WW. If there are not enough wind estimates to calculate component median values to compare with the wind observation being checked, then the wind observation is not flagged. A wind observation surrounded by mostly missing or flagged wind estimates may very well be erroneous but, because of the aforementioned limitation, will not be flagged.

4.2 Recommendations for Operational Use

The two routine combinations designated as CPM and CPW in this report (Section 2.2) are compared in post-analysis and real-time modes in Section 3 and in the Appendix. Both routines perform similarly in both modes with some notable differences that have already been discussed. These differences indicate that CPW is the better routine combination when used in the post-analysis mode, while CPM performs better in the real-time mode.

Care should be taken when choosing one of the routines over the other, however, as they were closely scrutinized on only five 24-hour data sets. An important point to consider is that the median filter (Carr et al. 1995) was newly modified in the AMU and has not been tested on many data sets, while WW was developed several years ago (Weber and Wuertz, 1991) and has been widely tested and used in the profiler community. In light of the small data set used in the analysis and the fact that the results from the median filter were only slightly better than WW in the simulated real-time mode, it is difficult to recommend one routine over the other. Either routine could be used in real-time. Whichever routine is chosen, it is critical that the consensus time and precipitation contamination checks be used to flag the obviously bad wind estimates before using either of the dependent routines. Neither the median filter or WW performed acceptably when used alone.

Complete descriptions of the routines in CPW and CPM are given in Section 2.1. A summary of the important parameters of each routine and their currently recommended settings are given in Table 4.1.

Table 4.1	Brief descriptions of and currently recommended parameter settings for the routines in CPW and CPM.
Routine	***Parameters and Settings***
Consensus Time Period Check Flags profiles with short consensus time periods	Short consensus time period: <6 minutes
Precipitation Contamination Check Flags wind estimates contaminated by rain using a discriminant function relating vertical beam radial velocity (VV) and SNR	$L = -1.731 + 0.298(VV) + 0.014(SNR)$ For $L > 0$, wind estimate contaminated by rain

Weber-Wuertz Algorithm
 	
Pattern recognition program. Uses ratio dy/dx as acceptable change in wind speed to create patterns. Flags estimates that do not fit established patterns and patterns with less than nmin estimates.
	dx: time 30 min
	       space (along radial) 632 ft
	
	dy: oblique beams 4 kts
	       vertical beam 2 kts
	
	nmin: 32 estimates

Median Filter (Carr et al. 1995) Compares the u- and v- component values of a wind observation to the median of the u- and v- component values of the surrounding wind estimates. If the difference between observation and the median exceeds T_u or T_v, the observation is flagged as suspect.	$T_{u,v} = \max(T_{(u,v)1}, T_2)$ Where: $T_{(u,v)1} = 0.2 \cdot \| (u,v)_m + (u,v)_i \|$ $T_2 = a \cdot (Ah^2 + Bh + C)$ $a = 1.3,$ $h = $ height (feet) $A = -5.695 \times 10^{-9}$ $B = 3.66 \times 10^{-4}$ $C = 7.3834$

4.3 Future Work

Most accepted quality assessment routines, such as WW, have been developed and tested over long periods of time using large wind data sets collected from every season in varying weather conditions. The routines developed and tested in this study have only been used with a limited warm season data set. Cool season precipitation signatures in the data may be different from those of the warm season. Thus, the discriminant function likely needs modification. The real-time WW routine does not perform as well as the post-analysis version. Modifications to the parameters are likely needed to improve its performance. As more data are analyzed in future tasks, the quality assessment routines will also be analyzed and modified as necessary to improve their performance.

The next objective to be completed is quality assessment of the RASS data. Steps similar to that of developing wind data routines will be followed. Hardware and software upgrades to the profiler network through the Range Standardization and Automation (RSA) contract are scheduled for late July and early August 1998. The subsequent changes to the wind data may require changes to the algorithms developed in this study. These two issues are discussed in the following two subsections.

4.3.1 RASS Data Quality Assessment

When in RASS mode, the radar emits a vertically propagating sound wave, then measures the speed of that wave using the radar vertical beam. An equation that relates the speed of sound to T_v is used to derive the individual T_v estimates. A 5-minute consensus period is used to estimate the final T_v profile in the same manner as a wind profile. The vertical T_v profile ranges from 112 m to 1.4 km with a gate spacing of 97 m.

The T_v distribution over the CCAS/KSC area will be analyzed for its usefulness in forecasting thunderstorm formation. It is important that the quality of the data be determined before this analysis is done. The preliminary work to archive and process this data type is completed. The next step is to develop and test quality assessment routines. The WW algorithm will be modified and tested and a lapse-rate algorithm will be developed. Analysis of the data may result in the development of other routines, such as the precipitation contamination check, that can flag erroneous data created by a specific contaminant.

The majority of this work is expected to be completed by May 1998. The product will be a final report describing how RASS data are acquired and the results of the quality assessment routine evaluation.

4.3.2 RSA Data Processing Upgrade

As stated earlier, hardware and software upgrades to the radars in the network by RSA are scheduled for late July and early August 1998. One scheduled improvement will modify system software such that when a radar is reset in the middle of a consensus period, the data collected previous to the reset will not be deleted and will be used in the consensus calculation. This will allow for the elimination of the consensus time period check routine. The basic data processing techniques, however, will remain the same as those described in Section 1.3. Changes may be made to the maximum height of the profile, the number of gates in the profile, the gate spacing, and the consensus time period. The format of the data transmitted to the ROCC will also be different. Exact details of all the changes affecting the data have not yet been provided. But, any changes in these parameters may require changes to the quality assessment routines and the software that ingests the data.

RSA personnel have indicated that WW will be used to QC the data before it is available to operational personnel. It is not known at this time how WW will be implemented, i.e. which data types will be used or how the parameters will be set. In any case, it is clear from this study that some form of a precipitation contamination algorithm must be used before WW, otherwise the performance of WW will be degraded when rain is affecting the data.

4.4 Conclusions

An important conclusion of this study is that more than one quality assessment routine is needed to accurately flag most of the erroneous data. Routines that check for temporal and spatial continuity work well when the ratio of bad to good data is small. However, long-lived contaminants can cause a large amount of erroneous data that will not be properly flagged by these algorithms. Because of the small data set used in the analysis and the fact that the results from the median filter were only slightly better than WW in the simulated real-time mode, neither routine is recommended over the other. It is critical, however, that the consensus time and precipitation contamination checks be used to flag the obviously bad wind estimates before using either of the dependent routines. Both of the combinations are able to identify most of the unreliable data and little of the good data as suspect.

Appendix

This Appendix contains images that allow for a closer comparison of CPW and CPM in post-analysis and real-time modes. Wind speed data from 12 May and wind direction data from 17 June are shown.

In the 12 May case, the performance of CPW was superior to that of CPM in post-analysis mode. However, CPW performance was degraded in the real-time mode such that the results were similar to CPM in both modes. The performance of CPM did not differ significantly between post-analysis and real-time modes. In the 17 June case, CPW and CPM performed similarly in post-analysis mode. Again, a degradation occurred in the real-time performance of CPW, but not in CPM.

12 May 1997

Figure A.1 shows the raw wind speed data, Figures A.2 and A.3 show the post-analysis results for CPW and CPM, respectively, and Figures A.4 and A.5 show the real-time results for CPW and CPM, respectively. A comparison of Figures A.2 and A.3 shows that CPW is more effective than CPM at flagging isolated wind estimates that are surrounding by a large amount of flagged or missing data. A degradation in that ability is seen by comparing Figures A.2 and A.4. The real-time results in Figures A.4 and A.5 are very similar in that isolated areas of wind estimates were not flagged by either CPW or CPM.
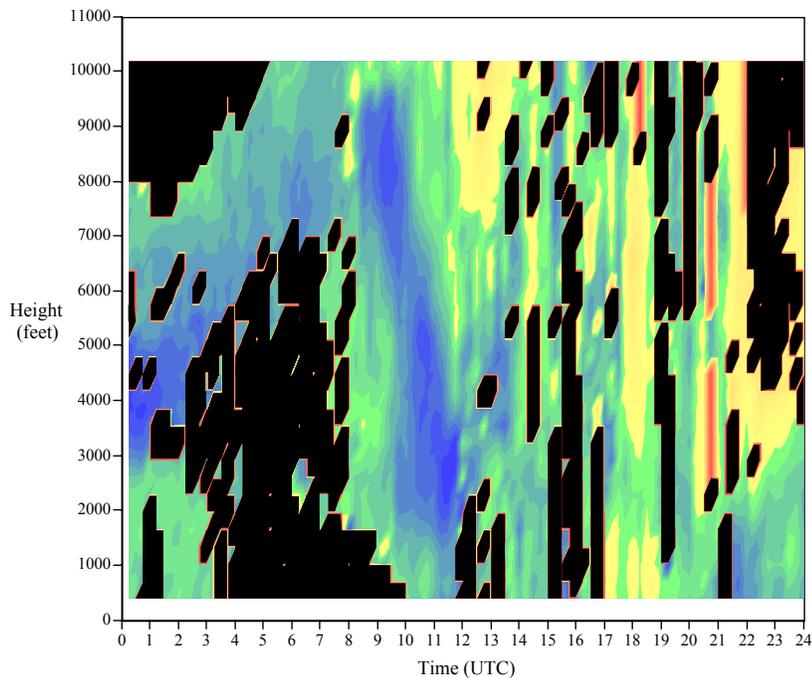

Figure A.1   Wind speeds from the False Cape profiler on 12 May 1997. Speeds are in knots. Areas of black indicate gates where a consensus wind was not calculated.
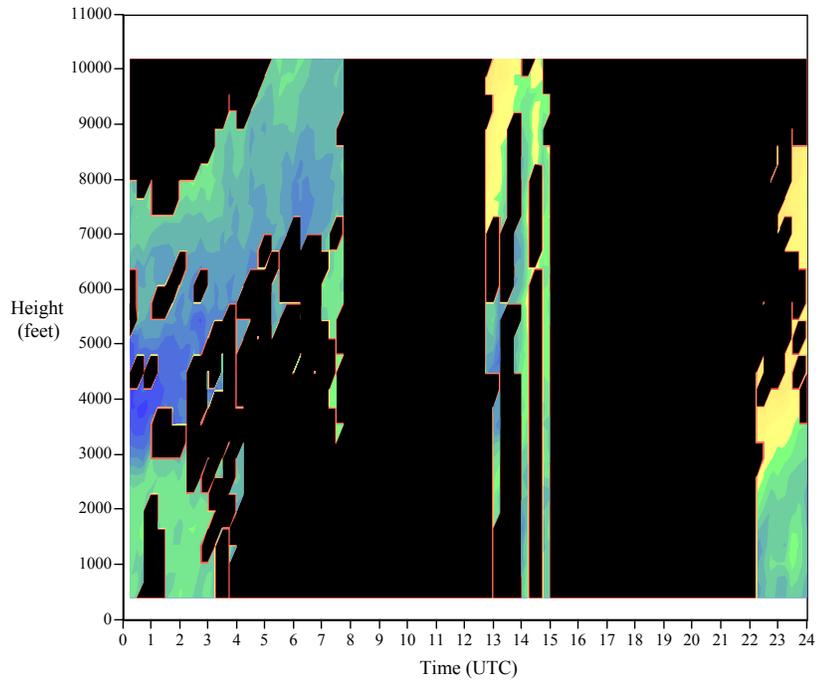
Figure A.2   Wind speeds from the False Cape profiler on 12 May 1997 after post-analysis CPW. Speeds are in knots. Areas of black indicate gates where a consensus wind was not calculated and where the data were flagged by the algorithms.
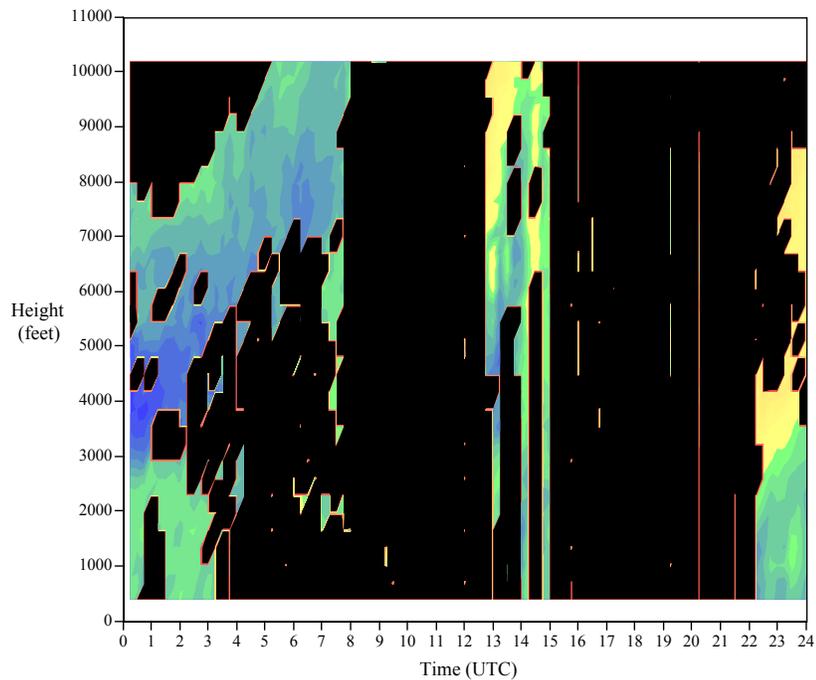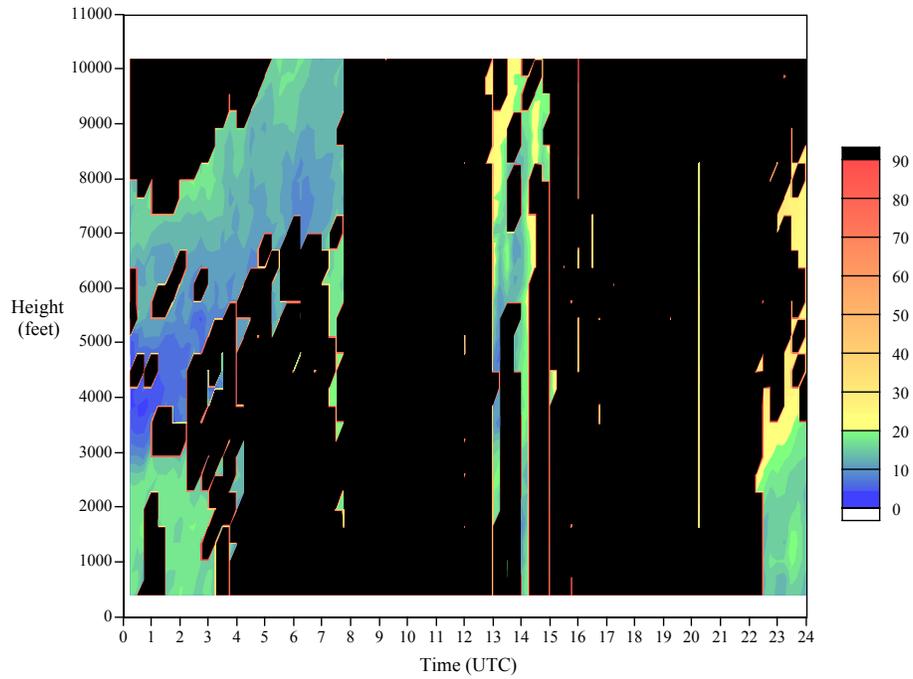

Figure A.3   Wind speeds from the False Cape profiler on 12 May 1997 after post-analysis CPM. Speeds are in knots. Areas of black indicate gates where a consensus wind was not calculated and where the data were flagged by the algorithms.
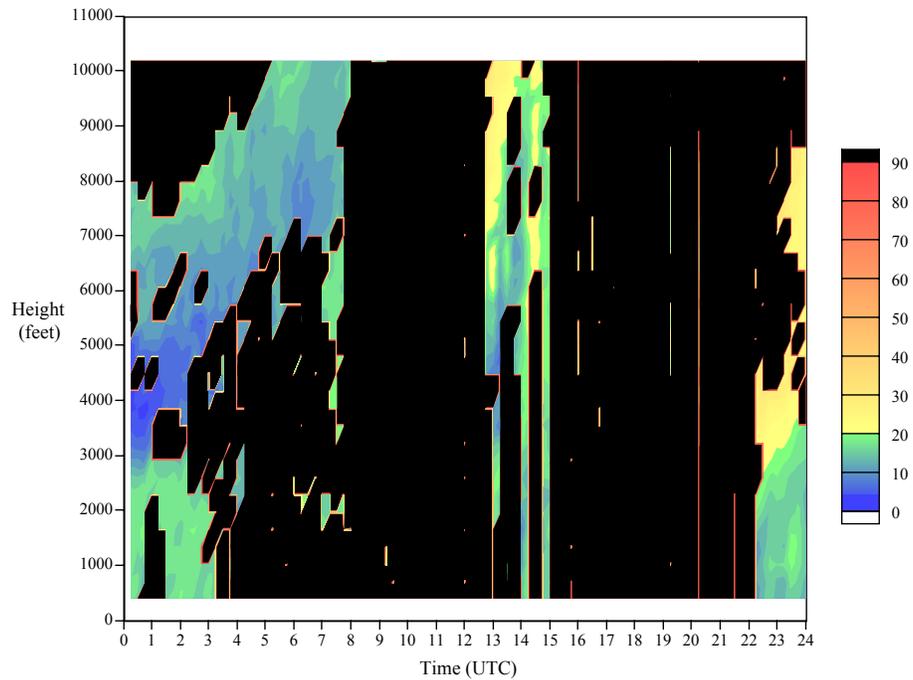
Figure A.4   Wind speeds from the False Cape profiler on 12 May 1997 after real-time CPW. Speeds are in knots. Areas of black indicate gates where a consensus wind was not calculated and where the data were flagged by the algorithms.

Figure A.5   Wind speeds from the False Cape profiler on 12 May 1997 after real-time CPM. Speeds are in knots. Areas of black indicate gates where a consensus wind was not calculated and where the data were flagged by the algorithms.

17 June 1997

Figure A.6 shows the raw wind direction data, Figures A.7 and A.8 show the post-analysis results for CPW and CPM, respectively, and Figures A.9 and A.10 show the real-time results for CPW and CPM, respectively. A comparison of Figures A.7 and A.8 shows very little difference in the results of the two routine combinations. A degradation can seen in a comparison of Figures A.7 and A.9. The obviously erroneous winds at 2030 UTC between 6500 - 8500 ft in Figure A.9 were not flagged by the real-time CPW, but were flagged by the post-analysis CPW. The reason for this is explained in Section 3.3. The real-time CPM flagged more wind estimates between 2000 and 2015 UTC as seen in Figure A.10, but the results were otherwise very similar to the post-analysis CPM results in Figure A.8.
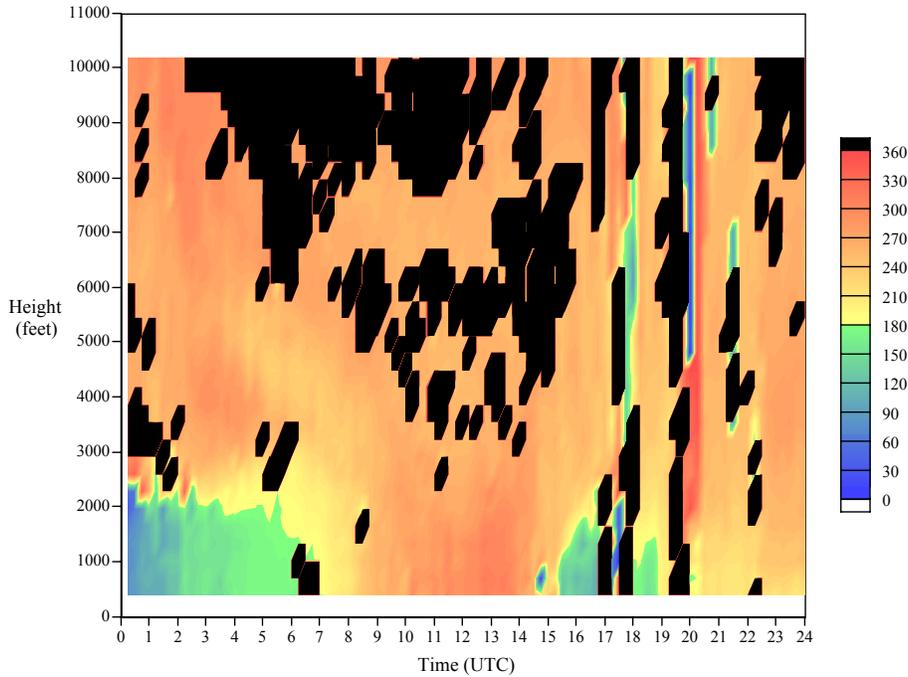

Figure A.6   Wind directions from the False Cape profiler on 17 June 1997. Directions are in degrees. Areas of black indicate gates where a consensus wind was not calculated.
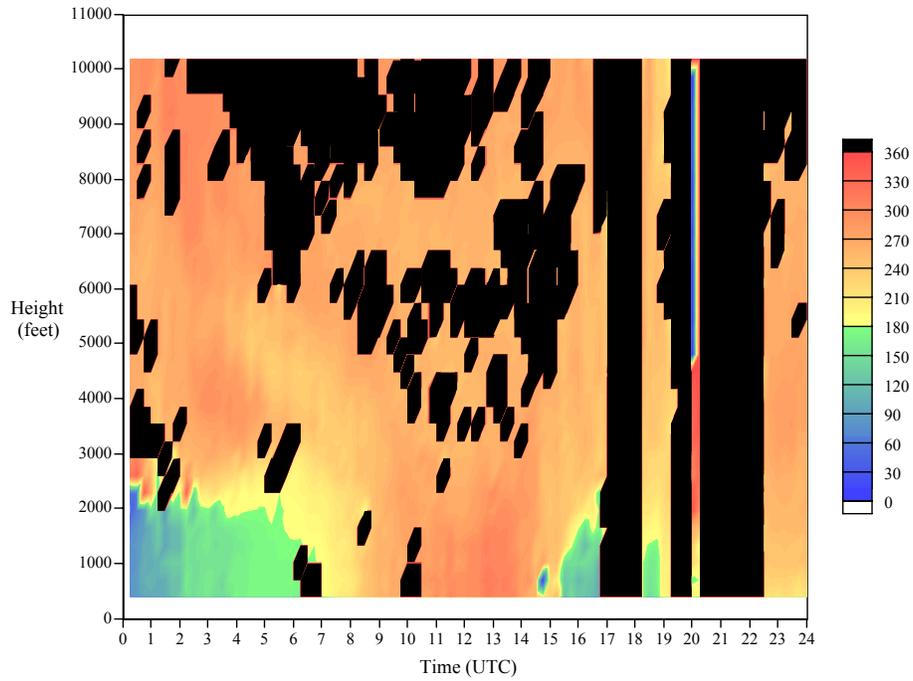
Figure A.7   Wind directions from the False Cape profiler on 17 June 1997 after post-analysis CPW. Directions are in degrees. Areas of black indicate gates where a consensus wind was not calculated and where the data were flagged by the algorithms.
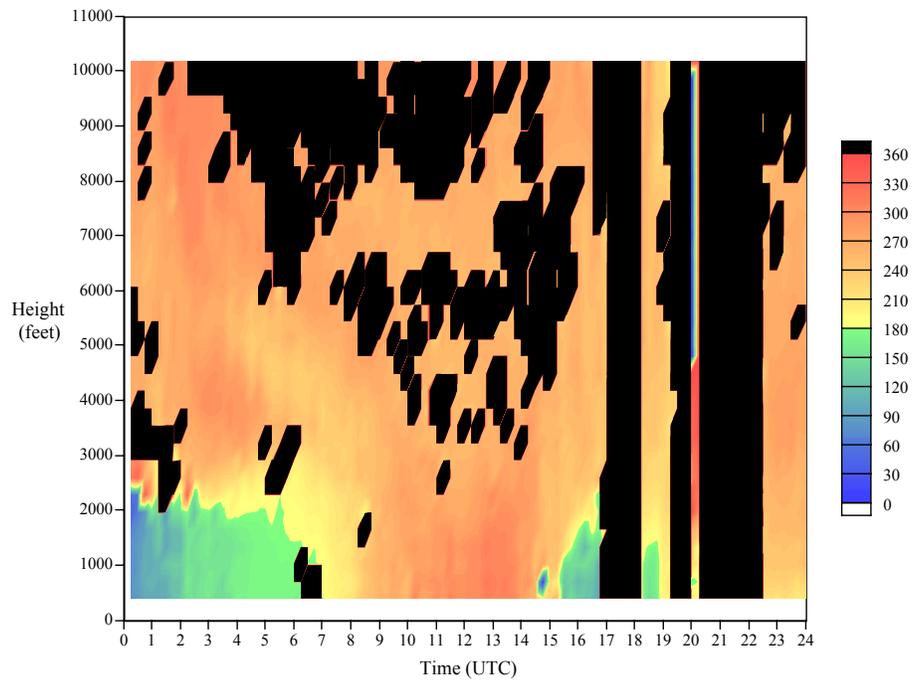

Figure A.8   Wind directions from the False Cape profiler on 17 June 1997 after post-analysis CPM. Directions are in degrees. Areas of black indicate gates where a consensus wind was not calculated and where the data were flagged by the algorithms.
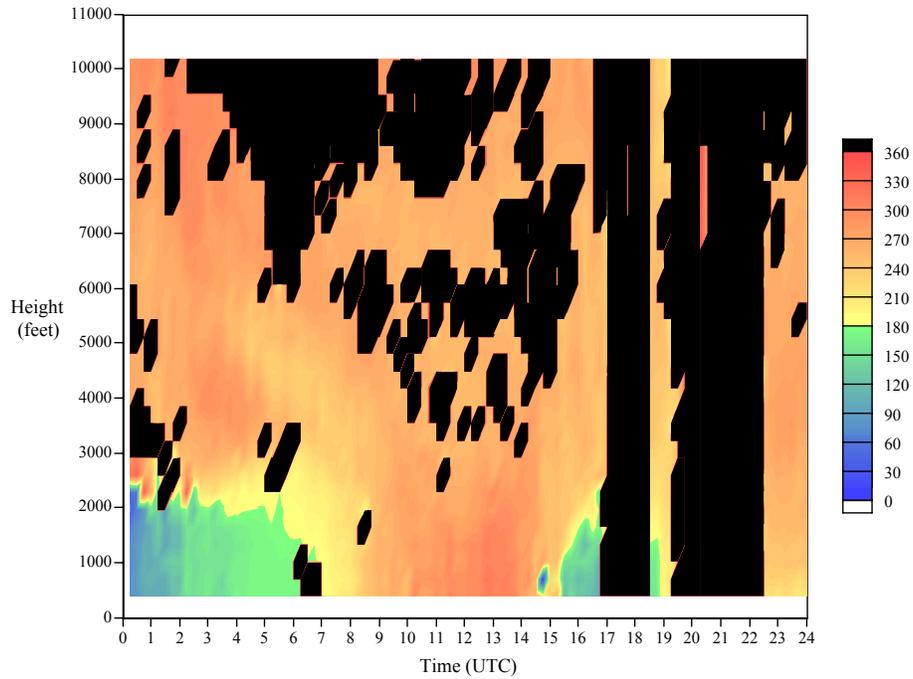
Figure A.9   Wind directions from the False Cape profiler on 17 June 1997 after real-time CPW. Directions are in degrees. Areas of black indicate gates where a consensus wind was not calculated and where the data were flagged by the algorithms.
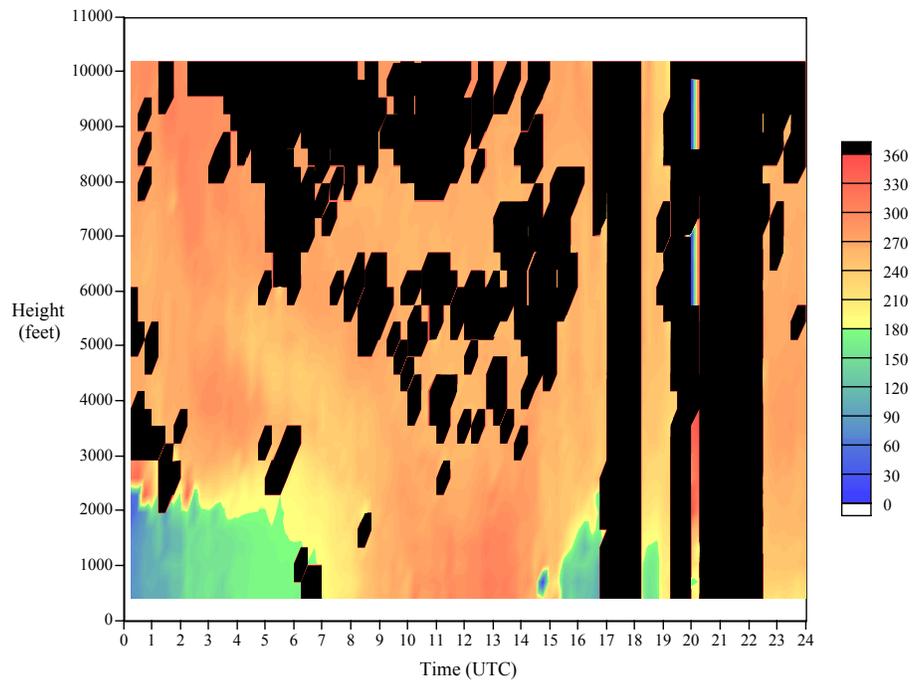

Figure A.10   Wind directions from the False Cape profiler on 17 June 1997 after real-time CPM. Directions are in degrees. Areas of black indicate gates where a consensus wind was not calculated and where the data were flagged by the algorithms.

References

Barth, M. F., R. B. Chadwick, W. M. Faas, 1997: The Forecast Systems Laboratory Boundary Layer Profiler Data Acquisition Project. *First Symposium on Integrated Observing Systems*, Long Beach, CA. Amer. Meteor. Soc., 130 - 137.

Carr, F. H., P. L. Spencer, C. A. Doswell III, and J. D. Powell, 1995: A comparison of two objective analysis techniques for profiler time-height data. *Mon. Wea. Rev.*, **123**, 2165-2180.

Heckman, S. T., M. W. Maier, W. P. Roeder, J. B. Lorens, and B. F. Boyd, 1996: The operational use of a boundary layer profiler network at the Eastern Range and Kennedy Space Center, *27th Conference on Radar Meteorology*, Vail, CO, 346-348.

Merceret, F. J., 1997: Rapid temporal changes of midtropospheric winds. *J. Appl. Meteor.*, **36**, 1567-1575.

Peterson, V. L., 1988: *Wind Profiling, The history, principles, and applications of clear-air Doppler radar.* Tycho Technology, Inc., Boulder, CO, 64 pp.

Panofsky, H. A., G. W. Brier, 1968: *Some Applications of Statistics to Meteorology.* The Pennsylvania State University, University Park, PA, 224 pp.

Radian Corporation, 1994: *The LAP™-3000 Operation and Maintenance Manual*, 2990 Center Green Court, South, Boulder, CO 80301, Document Control No. 80018004, 310 pp.

Ralph, F. M., 1995: Using radar-measured radial vertical velocities to distinguish precipitation scattering from clear-air scattering. *J. Atmos. Oceanic Tech.*, **12**, 257-267.

Ralph, F. M., P. J. Neiman, D. Ruffieux, 1996: Precipitation identification from radar wind profiler spectral moment data: Vertical velocity histograms, velocity variance, and signal power-vertical velocity correlations. *J. Atmos. Oceanic Tech.*, **13**, 545-559.

Weber, B. L. and D. B. Wuertz, 1991: Quality control algorithm for profiler measurements of winds and temperatures. NOAA Tech. Memo. ERL WPL-212, 32 pp.

Wilfong, T. L., S. A. Smith, and R. L. Creasy, 1993: High temporal resolution velocity estimates from a wind profiler. *J. Spacecraft and Rockets*, **30**, 348-354.

Williams, C. R., W. L. Ecklund, and K. S. Gage, 1995: Classification of precipitating clouds in the tropics using 915-MHz wind profilers. *J. Atmos. Oceanic Tech.*, **12**, 996-1012.