SE
93/
4

DOT/FAA/SE-93/4

National Airspace
System, System
Engineering Service
Washington, D.C. 20591

The Next Generation Weather Radar (NEXRAD)/Air Route Surveillance Radar (ARSR) Operational Comparison

Brian Dunbar
Jeff Mittelman

The MITRE Corporation
7525 Colshire Drive
McLean, VA 22101–3481

DEC 9 1993

July 1993

Final Report

This document is available to the public through the National Technical Information Service, Springfield, Virginia 22161.

U.S. Department
of Transportation

**Federal Aviation
Administration**

FAA WJH Technical Center
00019668

NOTICE

This document is disseminated under the sponsorship of the Department of Transportation in the interest of information exchange. The United States Government assumes no liability for its content or use thereof.

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
DOT/FAA/SE-93/4		

4. Title and Subtitle	5. Report Date
The Next Generation Weather Radar (NEXRAD)/Air Route Surveillance Radar (ARSR) Operational Comparison	July 1993
	6. Performing Organization Code
7. Author(s) Brian Dunbar, Jeff Mittelman	8. Performing Organization Report No. 93W0000145
9. Performing Organization Name and Address The MITRE Corporation 7525 Colshire Drive McLean, VA 22101-3481	10. Work Unit No. (TRAIS)
	11. Contract or Grant No. DTFA01-93-C-00001
12. Sponsoring Agency Name and Address Department of Transportation Federal Aviation Administration National Airspace System System Engineering Service Washington, DC 20591	13. Type of Report and Period Covered Final Report
	14. Sponsoring Agency Code ASE-100

15. Supplementary Notes

16. Abstract

The National Weather Service (NWS), Federal Aviation Administration (FAA), and Department of Defense are in the process of fielding the Next Generation Weather Radars (NEXRAD). These doppler weather radars, also known as Weather Surveillance Radar (WSR)-88D, will be replacing the WSR-57 and WSR-74 weather radars in use today. The NEXRAD data will be used by the FAA's Advanced Automation System (AAS) in place of the Air Route Surveillance Radar (ARSR) weather data currently being used by air traffic controllers. Because the NEXRAD's scanning strategy is more time consuming than the ARSR's, there have been some concerns expressed within the FAA about using "untimely" NEXRAD data in an Air Traffic Control (ATC) environment. In response to these concerns, the FAA's Center For Advanced Aviation System Development (CAASD) at MITRE conducted a study, under the sponsorship of the FAA's National Airspace System (NAS) System Engineering Service (ASE), to assess the relative ability of NEXRADs and ARSRs to detect and present significant weather in order to determine the operational impact of using NEXRAD data in lieu of ARSR data. This paper documents the NEXRAD/ARSR operational comparison study.

17. Key Words	18. Distribution Statement
NEXRAD, ARSR, TDWR, ASR-9, WSR-88D, Radar, Weather, Surveillance, Reflectivity, Air Traffic Control	Available to the Public through the National Technical Information Service, Springfield, VA 22161.

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	102	

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

ACKNOWLEDGMENTS

The authors would like to thank the following CAASD/MITRE employees: Steve Zobell for conceiving this study, for his software contributions and consultations, and for his peer review of this document; Graham Glover and Quy Dao for their initial efforts in getting the project underway; Christina Bauhof and Debbie Brown for their software contributions; Karl Barnes and Camille Shiotsuki for their ARSR decoding software and expertise; and Candace Scheytt for her help in preparing this document.

The authors would also like to thank Dave Sharp at the Melbourne Weather Forecast Office for archiving the needed data; Richard Krzepina at the Miami ARTCC for supplying us with the ARSR data and answering our numerous ARSR related questions; Rich Murnan and Bob Coblentz of the NEXRAD Operational Support Facility (OSF) for supplying us with the NEXRAD data; Jim Evans, Joe Cullen, and Ben Stevens of Lincoln Laboratory for supplying us with the ASR-9 and TDWR data, and their expertise; and Craig Goff for his continuing efforts to ensure the completion of the project.

TABLE OF CONTENTS

SECTION		PAGE
1	Introduction	1
1.1	Background	1
1.2	Overview	1
1.3	Related Studies	1
1.4	Scope	2
1.5	Document Organization	2
2	Radar Terminology and Background	3
2.1	Radar Descriptions	3
2.1.1	ARSR Data	3
2.1.2	NEXRAD Data	3
2.1.3	TDWR Data	7
2.1.4	ASR-9 Data	7
2.2	Reflectivity Levels	7
3	Methodology	11
3.1	Study Area of Interest	11
3.2	Data Collection and Assessment Development	11
3.3	Radar Analysis	13
3.4	Accuracy Analysis	13
4	Data Set Selection	15
4.1	Radar Data Types	15
4.1.1	NEXRAD Data Acquisition	15
4.1.2	ARSR Data Acquisition	16
4.1.3	TDWR and ASR-9 Data Acquisition	16
4.2	Scan Strategies	16
4.3	Data Sets	16
4.3.1	23 January 1992	17
4.3.2	02 June 1992	17
4.3.3	04 August 1992	18
5	Qualitative Radar Comparison	19

SECTION		PAGE

5.1 Method 19
5.2 Findings 19
5.2.1 Depiction of Cell Boundaries 19
5.2.2 ARSR Level Thresholds 25
5.2.3 Mutually Exclusive Areas 25
5.2.4 NEXRAD "Lag" 33
5.2.5 Altitude Discrimination 33
5.3 Conclusions 33

6 Accuracy Analysis 47

6.1 Method of Analysis 47
6.1.1 Data Specifications 47
6.1.2 Determining Truth 48
6.1.3 Determining Accuracy 49
6.2 Findings 52
6.2.1 Absolute Analysis 52
6.2.2 Worst Case Analysis 58
6.2.3 Operational Analysis 63
6.3 Conclusions 64

7 Quantitative Radar Comparison 67

7.1 Intensity Depiction Analysis 67
7.1.1 Method of Analysis 67
7.1.2 Findings 67
7.2 Area Analysis 69
7.2.1 Method of Analysis 69
7.2.2 Findings 69
7.3 Penetration Analysis 73
7.3.1 Method of Analysis 73
7.3.2 Findings 77
7.4 Distance Analysis 82
7.4.1 Method of Analysis 82
7.4.2 Findings 82
7.5 Summary 84

8 Conclusions and Recomendations 85

List of References 87

SECTION		PAGE
Appendix A	Display and Analysis Workstation	89
Appendix B	One-on-one Accuracy Analysis	97
Glossary		101

LIST OF FIGURES

SECTION		PAGE
1	Example of ARSR Weather	5
2	Cross Section of NEXRAD Volume Scan	8
3	Study Area of Interest	12
4	Cross Section of Radar Scan Techniques	17
5	Depictions of Reflectivity Values	21
6	Examples of ARSR Depiction of Weather	23
7	Example of Apparent ARSR Insensitivity	27
8	Example of Apparent ARSR Over Sensitivity	29
9	Example of ARSR Depiction of Weather Where NEXRAD Did Not Depict Weather	31
10	Example of 23 January 1992 Data Set	35
11	Example of NEXRAD "Lag"	37
12	Example of NEXRAD Low Layer	39
13	Example of NEXRAD High Layer	41
14	Example of NEXRAD Super-High Layer	43
15	Region Used in Accuracy Analysis	48
16	ARSR Display Technique	49
17	Consensus Approach, Two Radars Equals Consensus	50
18	Calculation of CSI	50
19	Expansion of Search Area	51

SECTION		PAGE
20	Distribution of ARSR Data Relative to NEXRAD Data	68
21	Area of Weather Over Time, 23 January 1992	70
22	Area of Weather Over Time, 2 June 1992	71
23	Area of Weather Over Time, 4 August 1992	72
24	Area of Weather (>40 dBZ) Over Time, 23 January 1992	74
25	Area of Weather (>40 dBZ) Over Time, 2 June 1992	75
26	Area of Weather (>40 dBZ) Over Time, 4 August 1992	76
27	Aircraft Penetrations	77
28	Aircraft Weather Penetrations	78
29	Aircraft Weather Penetrations, Layers	80
30	Average Minimum Distance of Aircraft From Weather	83
A-1	Screen Layout	90
A-2	Data Menu	91
A-3	Delay Menu	93
A-4	Sequence Menu	94
A-5	Map Menu	95

LIST OF TABLES

SECTION		PAGE
1	Reflectivity Values	9
2	CSI Values For Data Greater than or Equal to 41 dBZ	53
3	Raw Truth Statistics For Data Greater Than or Equal to 41 dBZ	54
4	CSI Values For Data Greater Than or Equal to 30 dBZ	56
5	Raw Truth Statistics For Data Greater Than or Equal to 30 dBZ	57
6	NEXRAD Delay Parameters Used In the Accuracy Analysis	58
7	ARSR Delay Parameters Used In the Accuracy Analysis	58
8	Worst Case, CSI Values For Data Greater Than or Equal to 41 dBZ	59
9	Worst Case, Raw Data Statistics For Data Greater Than or Equal to 41 dBZ	60
10	Worst Case, CSI Values For Data Greater Than or Equal To 30 dBZ	61
11	Worst Case, Raw Truth Statistics For Data Greater Than or Equal to 30 dBZ	62
12	CSI Values (1 km Truth, 4 km CSI), Zero Expansion	63
13	Raw Truth Statistics (1 km Truth, 4 km CSI), Zero Expansion	64
14	Distribution of ARSR Levels Relative to NEXRAD Levels	69
15	Aircraft Weather Penetration Data (Greater Than 40 dBZ)	81
16	Aircraft Weather Penetrations, Corrected for Area	82
17	Average Minimum Distance of Aircraft from Weather in Nautical Miles	84
B-1	CSI Values (1 km Truth, 4 km CSI), Zero Expansion, One-on-one	97

SECTION		PAGE

B-2 Raw Truth Statistics (1 km Truth, 4 km CSI), Zero Expansion, One-on-One, Patrick AFB FPS-60 98

B-3 Raw Truth Statistics (1 km Truth, 4 km CSI), Zero Expansion, One-on-one, Fort Lonesome ARSR-3 98

EXECUTIVE SUMMARY

Introduction

The National Weather Service (NWS), Federal Aviation Administration (FAA), and Department of Defense are in the process of fielding the Next Generation Weather Radars (NEXRAD). These doppler weather radars, also known as Weather Surveillance Radar (WSR)-88D, will be replacing the WSR-57 and WSR-74 weather radars in use today. The NEXRAD data will be used by the FAA's Advanced Automation System (AAS) in place of the Air Route Surveillance Radar (ARSR) weather data currently being used by air traffic controllers. Because the NEXRAD's scanning strategy is more time consuming than the ARSR's, there have been some concerns expressed within the FAA about using "untimely" NEXRAD data in an Air Traffic Control (ATC) environment. In response to these concerns, the FAA's Center for Advanced Aviation System Development (CAASD) at MITRE conducted a study under the sponsorship of the FAA's National Airspace System (NAS) System Engineering Service (ASE), to assess the relative ability of NEXRADs and ARSRs to detect and present significant weather in order to determine the operational impacts of using NEXRAD data in lieu of ARSR data.

Background

There are four types of weather radar data used in the study: NEXRAD, ARSR, Terminal Doppler Weather Radar (TDWR), and Airport Surveillance Radar (ASR)-9. Each is described below as well as a description of the reflectivity mapping scheme.

NEXRAD Data

The NEXRAD is a Doppler weather radar being fielded in 113 locations in the conterminous U.S. The FAA is planning to use NEXRAD data in the future enroute air traffic control centers as the primary source of weather data in place of today's ARSR weather information.

The NEXRAD collects data using a pencil beam pattern at various elevation tilts (9 or 14 depending upon the scan strategy) ranging from 0.5 to 19.5 degrees with a range of 248 nm. After all the tilts are collected, they are combined together and many different products are created. For this study, the composite reflectivity products were of interest because they are the products the FAA plans on using for ATC. The composite products are created by vertically combining the data throughout the entire tilt range into four different altitude layers representing the reflectivity at those altitudes. The four composite products are Composite Reflectivity (0-60 k ft), Layer 1 Composite Reflectivity (0-24 k ft), Layer 2 Composite Reflectivity (24-33 k ft), and Layer 3 Composite Reflectivity (33-60 k ft). These three altitude layers nominally

correspond to the same altitude layers used in ATC low, high, and super high sectors. Figure ES-1 shows sample NEXRAD data.

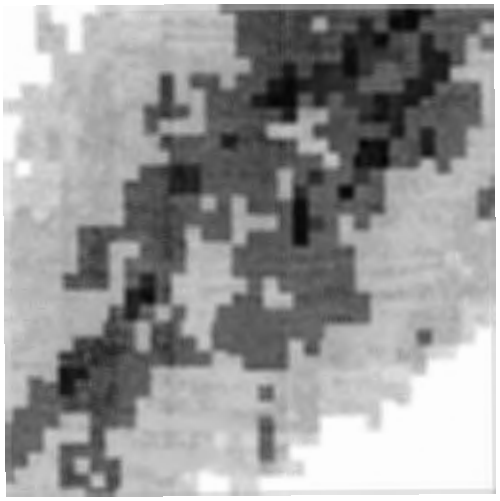


Figure ES-1. NEXRAD Display

The time it takes for the NEXRAD to complete scans of all the tilts, a complete volume scan, could be either 5 or 6 minutes depending upon which scan strategy was in effect at the time. At the end of a volume scan, the data scanned at the first tilt (lowest tilt) is 5 or 6 minutes old. Compound this with processing time within the NEXRAD and the NAS along with waiting for another 5 minutes before the NEXRAD data is updated could result in some of the data in the volume scan being 12 minutes old. The operational impact of this data aging problem, or data "lag" as it is referred to in this paper, is the subject of this study.

ARSR Data

The ARSR is primarily a surveillance radar which also has a weather detecting capability. Today's enroute air traffic control centers use ARSR to track aircraft and observe weather.

ARSR uses a fan beam scanning strategy and has a nominal range of 150 nm. The weather processing portion of the radar offers weather returns in two ranges: moderate, depicted as lines, and heavy, depicted as "H"s. Moderate weather is depicted as the areas under the lines and the heavy weather is depicted as the areas between the pairs of "H"s along the moderate lines. A

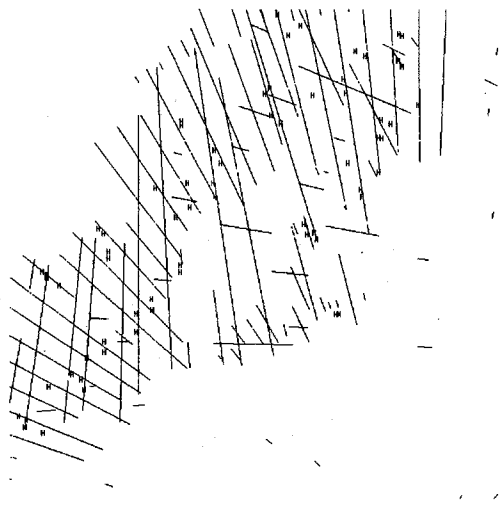


Figure ES-2. ARSR Display

complete scan strategy for most ARSRs includes 12 12-second scans resulting in the weather being updated approximately every 144 seconds. Figure ES-2 shows the ARSR depiction of the same storm depicted in Figure ES-1.

TDWR Data

The TDWR is a Doppler weather radar that will be deployed at about 45 major airports in the US to provide timely and accurate weather information to the terminal area. It employs similar technology to the NEXRAD: pencil beam scanning at multiple elevation tilts. Its range is approximately 60 nm. The TDWR data used in this study was a 7 level reflectivity product produced once per 6 minute volume scan at a 3.7 degree tilt.

ASR-9 Data

The ASR-9 is a surveillance radar for use in terminal areas. Similar to the ARSR, it is a fan beam radar, but has a nominal range of only 60 nm and is equipped with a special weather channel providing a 7 level reflectivity product every minute.

Reflectivity Levels

Table ES-1 shows the mapping between the NWS 6 level reflectivity scheme and the corresponding NEXRAD and ARSR reflectivity levels referred to in this study.

Table ES-1. Reflectivity Values

NWS Levels	NEXRAD Units (dBZ)	ARSR Levels
	< 5	
	5-18	
1	18-30	
2	30-41	Moderate
3	41-46	Heavy
4	46-50	Heavy
5	50-57	Heavy

Method

The study consisted of collecting data from multiple radars over the same geographic area during the same time period, developing an assessment capability, and performing a variety of comparative analyses.

The central Florida region was selected as the study's area of interest for the following reasons:

- One of the few fielded NEXRADs is located in that area (Melbourne)
- Significant weather occurs in that area year round
- TDWR and ASR-9 testbeds are located in that area and could be used to help determine "truth"

There were five ARSRs providing overlapping coverage, with all portions of the study area being covered by at least two ARSRs.

Data was collected for NEXRAD, ARSR, ASR-9, and TDWR in the central Florida area. Three data sets (23 January, 2 June, and 4 August 1992) containing more than 20 hours of severe

weather were obtained, and a display and assessment capability was developed consisting of a workstation capable of displaying all four types of radar data in time sequence.

After the data was collected and an assessment capability was developed, the radar analysis was performed to compare the NEXRAD and ARSR radars in time and space. The first step in this analysis was to perform a visual qualitative comparison. This comparison showed significant differences between NEXRAD and ARSR; therefore, another analysis was needed to determine which radar was more accurate. Following the accuracy analysis, the last portion of the radar analysis was performed with a quantitative comparison of the NEXRAD and ARSR data to one another in the areas of weather intensity and spatial extent to determine their relative strengths and weaknesses. Aircraft track data from the ARSR was also used to observe aircraft weather avoidance as depicted by each radar to help determine the operational impacts of the weather displays.

Qualitative Comparison

The first analysis performed was a qualitative comparison—while the workstation was sequencing through each data set, information was visually analyzed and collected. This comparison concluded the following:

- The ARSR weather display technique gives a coarse depiction of cell boundaries especially when compared to the NEXRAD display.
- The ARSR level thresholds appear to vary considerably. The same cell was often depicted differently by multiple ARSR radars.
- The ARSR appears to miss significant weather and detect weather that is not real. Aircraft frequently avoided areas depicted by NEXRAD and not by ARSR while penetrating areas depicted by ARSR and not by NEXRAD.
- The NEXRAD data aging "lag" is clearly visible with fast moving cells and increases with the speed of the cells.

Because these significant differences were found between NEXRAD and ARSR, an accuracy analysis was conducted in order to determine which radar was closer to the truth.

Accuracy Analysis

A measure of the accuracy of the NEXRAD and ARSR radars was needed to establish which radar was more accurate as well as to explain differences discovered during the qualitative comparison of the radars. In order to determine their accuracy, a definition of truth was

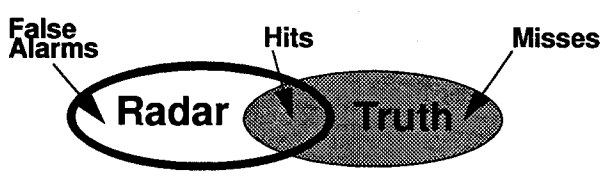
required. Truth was defined using the radar data from NEXRAD, ARSR, TDWR and ASR radars in combination with a consensus approach. Once truth was defined, the NEXRAD and ARSR radar data was compared to this definition of truth to establish a measure of their accuracy.

Determining Truth

In order to determine truth, all of the data was first mapped into a 1.33 km grid. A truth grid was created by scanning through each position in the grid and checking the corresponding location in each of the radar images for data at or above the correct reflectivity threshold. Results were gathered for reflectivity thresholds of 30 dBZ and 41 dBZ. If the appropriate reflectivity level was found at a consensus of the radars (two radars was used to represent consensus), then that element in the truth grid had weather, otherwise it did not. The NEXRAD and ARSR data were then compared to this truth grid to establish a measure of their accuracy.

Assessing Accuracy

The accuracy of a radar was determined by computing its Critical Success Index (CSI), which is computed using three statistics: hits, misses, and false alarms. A hit is defined as the truth grid indicating weather at a grid point and the radar detecting weather at that same grid point. A miss is defined as the truth grid indicating weather at a grid point and the radar not detecting weather at the same grid point. A false alarm is defined as the truth grid indicating that there is no weather at a grid point while the radar is indicating that there is weather at the grid point. The value of the CSI can range from 0 to 100 and it represents the percentage that hits are of the total area of weather plus false alarms (see figure ES-3).



$$CSI = \frac{\text{Hits}}{\text{Hits} + \text{False Alarms} + \text{Misses}} \times 100$$

Figure ES-3. Calculation of CSI

Absolute Analysis

Initially an absolute analysis was conducted where the data was matched in time with all associated sources of delay removed to obtain a measure of the absolute sensing capabilities of the radars as opposed to measuring the operational accuracy that results when the radar data has been delayed by processing (i.e., the NEXRAD lag, or data aging, was artificially removed). This was done in order to determine which radar actually detects weather more accurately.

Worst Case Analysis

The absolute analysis measured the accuracy of the radars at actually sensing weather, but much of the concern over NEXRAD data, as mentioned above, was the problem of data aging. To address this issue, a worst case analysis was conducted which consisted of determining the truth grid with the radars matched in time while the accuracy statistics were computed at the moment the NEXRAD was to be updated, including all processing delays. The NEXRAD image, though, was not updated until after the analysis was complete. In this manner the truth grid had the most up-to-date information while the NEXRAD was at its oldest (worst) age.

Accuracy Results

Table ES-2 presents the results of the accuracy analysis. The first column shows what type of analysis was performed, absolute or worst case, and the second column shows the reflectivity threshold. The last column shows the percent improvement the NEXRAD CSI shows over the ARSR CSI. NEXRAD shows significant improvement over ARSR in both the absolute and worst case. Even in the worst case NEXRAD showed a 92 percent improvement over ARSR.

Table ES-2. Results of Accuracy Analysis

Analysis Type	Reflectivity Threshold in dBZ	% Improvement in NEXRAD CSI over ARSR CSI
Absolute	30	99
Absolute	41	124
Worst Case	30	92
Worst Case	41	111

Quantitative Comparison

A quantitative comparison of the radar data was performed in order to substantiate and measure some of the anomalies that were discovered during the qualitative comparison. The comparison included determining the intensity distribution, in which the data from NEXRAD and ARSR was compared to each other, in addition to analyzing the penetrations of aircraft into the weather data.

Intensity Depiction Analysis

The intensity analysis consisted of comparing ARSR data to the corresponding NEXRAD data to determine the distribution of NEXRAD intensity levels associated with each ARSR data level. This analysis was performed in order to substantiate and quantify the problems that were discovered during the qualitative comparison with ARSRs ability to determine intensity.

The analysis was conducted by recording the NEXRAD data level associated with each point of each line of ARSR data. The NEXRAD composite reflectivity product was compared to the ARSR product while matched in time. Figure ES-4 shows the percentage of ARSR moderate and heavy data that corresponded to the same NEXRAD levels. The most striking observation to be made from the chart is that there is little difference in the distributions of ARSR moderate and heavy data, having an almost random distribution. This data confirms the finding from the qualitative comparison that the ARSR does not accurately determine the intensity of weather.

Penetration Analysis

In addition to the weather data that was acquired from the ARSR radars, track data was also collected for the corresponding times in order to analyze the patterns of the aircraft relative to the weather as depicted by each radar. This analysis was conducted by counting the number of penetrations by aircraft into weather and comparing the results for the two radars. With larger numbers of penetrations there is less confidence in the severity of the weather at the indicated location because of an aircraft's obvious tendency to avoid severe weather (a large portion of the aircraft used in this analysis probably were equipped with on-board weather radar).

Only beacon equipped aircraft transmitting unique beacon codes were used. This eliminated aircraft that were likely flying without controller services. Aircraft below 8000 feet were also eliminated because the study was concerned mainly with en-route control. Only data for 41 dBZ and above is presented here.

The analysis was conducted by sampling the positions of each of the aircraft every 30 seconds and determining if they were penetrating any weather. A penetration was defined as the aircraft occupying a location on the display that also contained weather. Figure ES-5 shows the method used in the analysis. The arrows represent the flight path of an aircraft and the crosses represent

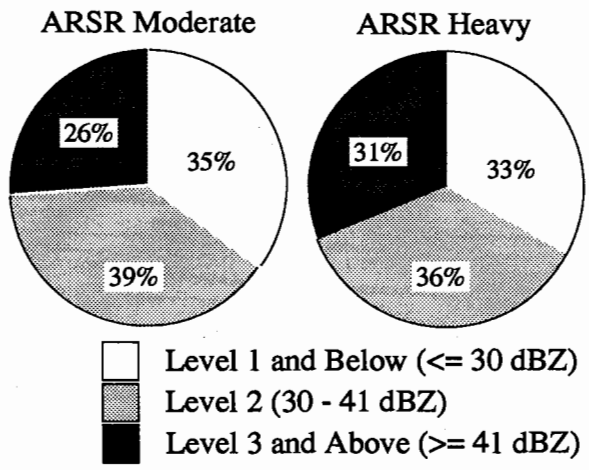


Figure ES-4. Distribution of ARSR Data Relative to NEXRAD

the sample points. Aircraft A has 2 penetrations into the moderate weather while aircraft B has 3 penetrations into the moderate weather and 2 penetrations into the heavy weather. Only aircraft within the coverage altitudes of the corresponding layered product were analyzed (these same aircraft were also compared against the ARSR data).

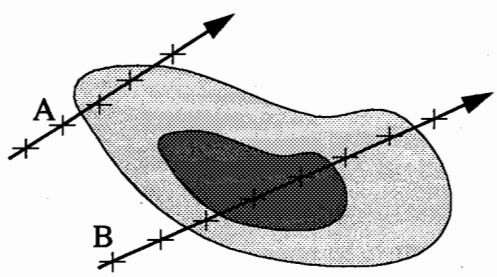


Figure ES-5. Penetration Analysis Sampling

Figure ES-6 presents the results of the penetration analysis for weather data at a level of 41 dBZ and greater. It shows that ARSR has an overwhelmingly larger total number of penetrations than NEXRAD (both with and without lag). There is a slightly larger number of penetrations into NEXRAD with a lag than NEXRAD without, but this is negligible when compared to that of ARSR. These results lead to the conclusion that the NEXRAD data provides a better representation of the pilots view, in addition to supporting the claim that the lag does not significantly impact ATC operational use. The NEXRAD data, though, did have a noticeable number of penetrations in the layer 1 product, some of which can be attributed to the fact that the data covers a significant range of altitudes (0 - 24k feet) and all altitudes in the layer may not contain weather.

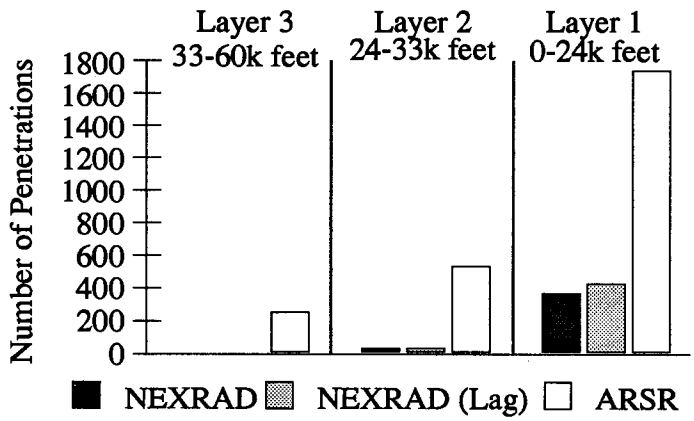


Figure ES-6. Aircraft Weather Penetrations

Conclusions

The qualitative, accuracy, and quantitative portions of the study conclude the following:

- ARSR presents a coarse depiction of cell boundaries; NEXRAD is precise
- ARSR level thresholds appear to vary considerably; NEXRAD levels are accurate and consistent
- ARSR appears to miss significant weather and detect weather that is not real; NEXRAD more accurately depicts the weather

- ARSR updates the position of the weather frequently; NEXRAD data aging "lag" is real but has little operational impact
- ARSR does not discriminate between altitudes; NEXRAD offers altitude discrimination in three layers
- ARSR appears not to match the pilots view of the weather; NEXRAD appears to better depict the weather the pilot is observing at different altitudes
- ARSR weather is not well suited for tactical ATC because of the above problems; NEXRAD is well suited for tactical ATC

Operationally, this means that ARSR can mislead controllers when the ARSRs miss significant weather and produce false weather, possibly reducing capacity. Compared to today's ARSR weather display, NEXRAD will more accurately depict the weather to the controller and better depict the pilot's view of the weather. These features can help tactical ATC operations, possibly improving capacity and reducing the pilot-to-controller communications load. Finally, the lag associated with the NEXRAD data does not significantly affect the operational accuracy of the data, and in comparison to the ARSR data, the NEXRAD data is much more accurate.

The FAA should continue to pursue the use of NEXRAD data, as the primary weather source for the enroute ATC environment, and the programs that deliver the data to the end-users.

SECTION 1

INTRODUCTION

1.1 BACKGROUND

The National Weather Service (NWS), Federal Aviation Administration (FAA), and Department of Defense are in the process of fielding the Next Generation Weather Radars (NEXRAD). These doppler weather radars, also known as Weather Surveillance Radar (WSR)-88D, will be replacing the WSR-57 and WSR-74 weather radars in use today. The NEXRAD data will be used by the FAA's Advanced Automation System (AAS) in place of the Air Route Surveillance Radar (ARSR) weather data currently being used by air traffic controllers. Because the NEXRAD's scanning strategy is more time consuming than the ARSR's, there have been some concerns expressed within the FAA about using "untimely" NEXRAD data in an Air Traffic Control (ATC) environment. In response to these concerns, the FAA's Center For Advanced Aviation System Development (CAASD) at MITRE conducted a study, under the sponsorship of the FAA's National Airspace System (NAS) System Engineering Service (ASE), to assess the relative ability of NEXRADs and ARSRs to detect and present significant weather in order to determine the operational impact of using NEXRAD data in lieu of ARSR data.

1.2 OVERVIEW

The objective of the study was to determine the relative ability of NEXRAD and ARSR to detect and present significant weather, and if differences were found, determine the ATC operational impact of the differences. The study consisted of comparing data, matched in time and space, from one NEXRAD and multiple ARSRs; spatial extent, spatial accuracy, and reflectivity accuracy were compared. To help determine the operational impacts, aircraft track data were also analyzed to determine the weather avoidance patterns of aircraft as depicted by each type of radar. The study was broken down into three parts: qualitative comparison, accuracy analysis, and quantitative comparison. Each part of the study along with the final results will be addressed in this report.

1.3 RELATED STUDIES

The National Center for Atmospheric Research (NCAR) performed a similar study under sponsorship of the FAA's Research and Development Service (ARD) (Dixon, 1992). While the objective of the NCAR study is very similar, it is a complementary study because their analysis

techniques and approach were different from the CAASD/MITRE study and it was performed in a different geographic region of the country.

NCAR's study was performed using the Mile High research radar near Denver. While this is not a true NEXRAD, it is a pre-production NEXRAD that performs very similarly to NEXRAD. Their results were very similar to those found in the CAASD/MITRE study.

NCAR also proposed techniques to further improve the timeliness of the NEXRAD data using alternate scan strategies and processing techniques. These alternatives may be useful if it is found that the NEXRAD data is not timely enough in an operational ATC environment.

The results of both studies were combined together when presented to the FAA, and references to their study are included in this report.

1.4 SCOPE

This report documents the NEXRAD/ARSR Weather Data Operational Comparison. It presents each step of the analysis process leading up to the recommendations for the FAA. This report does not document the software developed to perform the study nor the similar NCAR study.

The study examines the weather data as displayed by each type of weather radar as it might be seen by enroute air traffic controllers. There has been no effort made to try to explain display differences or accuracy deficiencies in terms of the radar hardware, design, or scanning strategy.

It is assumed that the reader has a basic understanding of meteorology, NEXRAD, ARSR, and ATC.

1.5 DOCUMENT ORGANIZATION

The remainder of this document is organized as follows: Section 2 provides additional radar terminology and background information; Section 3 describes the study methodology; Section 4 describes the data acquisition process; Section 5 describes the qualitative comparison; Section 6 describes the accuracy or "truth" analysis; Section 7 describes the quantitative radar comparison; Section 8 presents the conclusions and recommendations.

SECTION 2

RADAR TERMINOLOGY AND BACKGROUND

This section describes additional radar terminology and background information required to understand the study described in the remainder of this report.

2.1 RADAR DESCRIPTIONS

There are four types of weather radar data used in the study: NEXRAD, ARSR, Terminal Doppler Weather Radar (TDWR), and Airport Surveillance Radar (ASR)-9. Each will be described in the following sections. Later in this report will be descriptions on how data from each type of radar are used in the study.

2.1.1 ARSR Data

The ARSR is primarily a surveillance radar which also has a weather detecting capability. Today's enroute air traffic control centers use ARSR to track aircraft and observe weather.

ARSR uses a fan beam scanning strategy and has a nominal range of 150 nm. The weather processing portion of the radar offers weather returns in two ranges: moderate and heavy (see section 2.2). Referring to figure 1, the moderate weather is depicted as the areas under the lines and the heavy weather is depicted as the areas between the pairs of "H"s along the moderate lines. A complete scan strategy for most ARSRs includes 12 12-second scans resulting in the weather being updated approximately every 144 seconds (Federal Aviation Administration, 1989). The aircraft positions are shown as a small red square with an attached data block. The data block shows the aircraft's beacon identifier and altitude (in hundreds of feet).

ARSRs in the study area (see section 3.1) actually consist of three types of radars: ARSR-3s, FPS-60s, and ARSR-1s. No differentiation of these radar types was made in this study. When the ARSR data was displayed for analysis, all the data from all the ARSRs was displayed—no filtering or sorting of the weather data was performed because the ATC operations in the study area do not perform this type of data reduction activities.

2.1.2 NEXRAD Data

The NEXRAD is a Doppler weather radar being fielded in 113 locations in the conterminous U.S. The FAA is planning to use NEXRAD data in the future enroute air traffic control centers as the primary source of weather data in place of today's ARSR weather information (see next section).

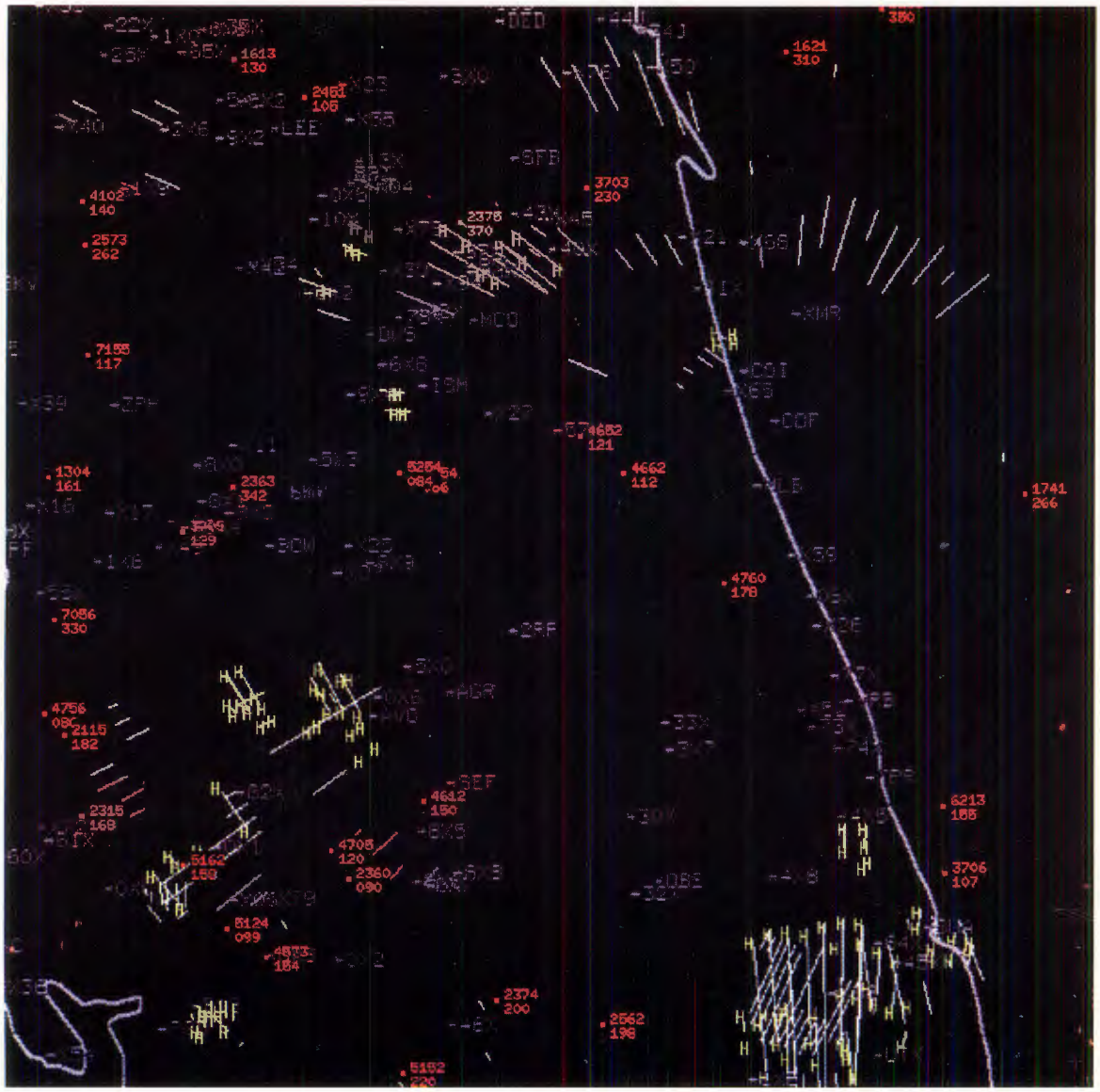


Figure 1. Example of ARSR Weather

The NEXRAD collects data using a pencil beam pattern at various elevation tilts (9 or 12 depending upon the scan strategy) ranging from 0.5 to 19.5 degrees with a range of 248 nm. Figure 2 depicts a cross section of a NEXRAD scan. After all the tilts are collected, they are combined together and many different products are created. For this study, the composite reflectivity products were of interest because they are the products the FAA plans on using for ATC. The composite products are created by vertically combining the data throughout the entire tilt range into four different altitude layers representing the reflectivity at those altitudes. The four composite products are Composite Reflectivity (0-60 k ft msl), Layer 1 Composite Reflectivity (0-24 k ft msl), Layer 2 Composite reflectivity (24-33 k ft msl), and Layer 3 Composite Reflectivity (33-60 k ft msl). (National Oceanic and Atmospheric Administration, 1991) The three altitude layers associated with the layered products nominally correspond to the same altitude layers used in ATC low, high, and super high sectors referred to later in this report. The layered products are created in 8 or 16 level products at 1 or 4 km resolution—8 level, 4 km products were used in this study, unless otherwise stated, because they are the products the FAA will be receiving in the future.

The time it takes for the NEXRAD to complete scans of all the tilts, a complete volume scan, could be either 5 or 6 minutes depending upon which scan strategy was being used at the time. At the end of a volume scan, the data scanned at the first tilt (lowest tilt) is 5 or 6 minutes old. Compound this with processing time within the NEXRAD and the NAS along with waiting for another 5 minutes before the NEXRAD data is updated could result in some of the data in the volume scan being 12 minutes old. The operational impact of this data aging problem, or data "lag" as it is referred to in this report, is examined.

2.1.3 TDWR Data

The TDWR is a Doppler weather radar that will be deployed at about 45 major airports in the U.S. to provide timely and accurate weather information to the terminal area. It employs similar technology to the NEXRAD: pencil beam scanning at multiple elevation tilts. Its range is approximately 60 nm and generates 7 level data at a 1 km resolution.

2.1.4 ASR-9 Data

The ASR-9 is a surveillance radar for use in terminal areas. Similar to the ARSR, it is a fan beam radar, but has a nominal range of only 60 nm and is equipped with a special weather channel providing a 7 level reflectivity product once per minute at a resolution of 1.25 km.

2.2 REFLECTIVITY LEVELS

Table 1 shows how the reflectivity levels from NEXRAD and ARSR map into the 6 level reflectivity scheme currently used by the NWS. Please note that the ARSR "moderate" matches

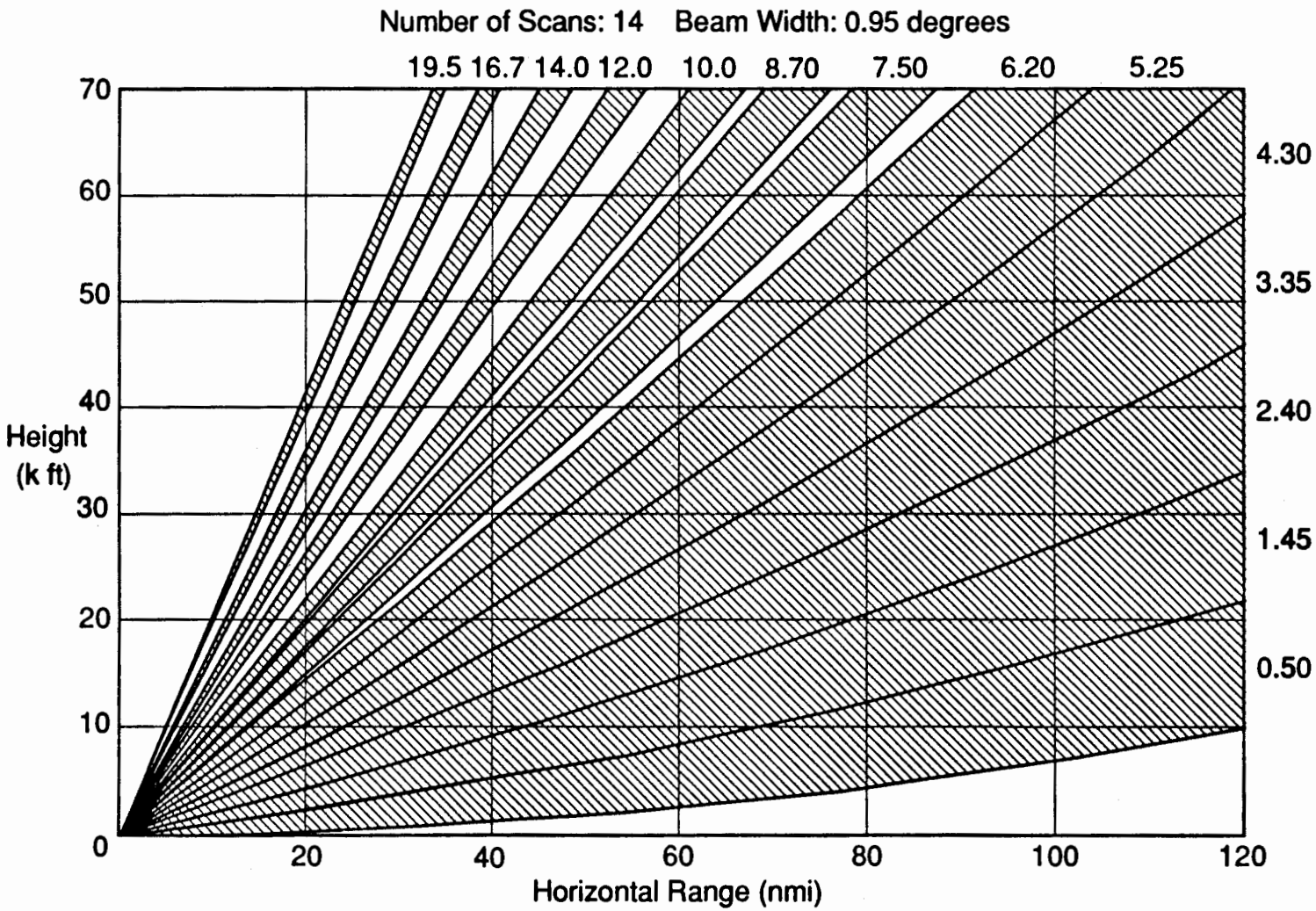


Figure 2. Cross Section of NEXRAD Volume Scan

with the NEXRAD level 4 and the NWS level 2. The ARSR "heavy" level matches with NEXRAD levels 5 and greater and NWS levels 3 and greater. All subsequent comparisons in this report will use these values.

Table 1. Reflectivity Values

NWS Levels	NEXRAD Units (dBZ)	ARSR Levels
	< 5	
	5-18	
1	18-30	
2	30-41	Moderate
3	41-46	Heavy
4	46-50	Heavy
5	50-57	Heavy
6	> 57	Heavy

SECTION 3

METHODOLOGY

The study consisted of collecting data from multiple radars over the same geographic area during the same time period, developing an assessment capability, and performing a variety of comparative analyses. This section provides an overview of the study methodology. Subsequent sections of this report provide additional details on each portion of the study.

3.1 STUDY AREA OF INTEREST

The central Florida region was selected as the study's area of interest for the following reasons:

- One of the few fielded NEXRADs is located in that area (Melbourne)
- Significant weather occurs in that area year round
- The TDWR and ASR-9 testbeds are located in that area and could be used to help determine "truth" (see section 6)

The exact study area was a 248 by 248 nautical miles (nm) area centered on the Melbourne NEXRAD as depicted in the rectangular area in figure 3. Even though the NEXRAD can cover an area approximating 496 by 496 nm, the study area was reduced to 248 nm because the planned adjacent NEXRADs are not yet deployed. Without contributions from these 6 other radars, the ability to detect weather at long range may be somewhat reduced. This smaller study area more accurately reflects the contributions of the Melbourne NEXRAD in the time frame when all the radars are deployed. There are 5 ARSRs in the study area providing overlapping coverage by more than one radar throughout the study area. The reader should be aware that the fact that a single NEXRAD and multiple ARSRs are being used may provide the ARSRs with an advantage in detecting and presenting weather; however, the magnitude of this advantage (if any) is not known at this time.

3.2 DATA COLLECTION AND ASSESSMENT DEVELOPMENT

The first portion of the study was the data collection and assessment development tasks. The data collection consisted of collecting data sets over the same geographic area during the same time period for NEXRAD, ARSR, ASR-9, and TDWR in the study area, central Florida. In order to analyze the data sets, a display and assessment capability was developed consisting of a workstation capable of displaying all four types of radar data in time sequence. This workstation is described in Appendix A.

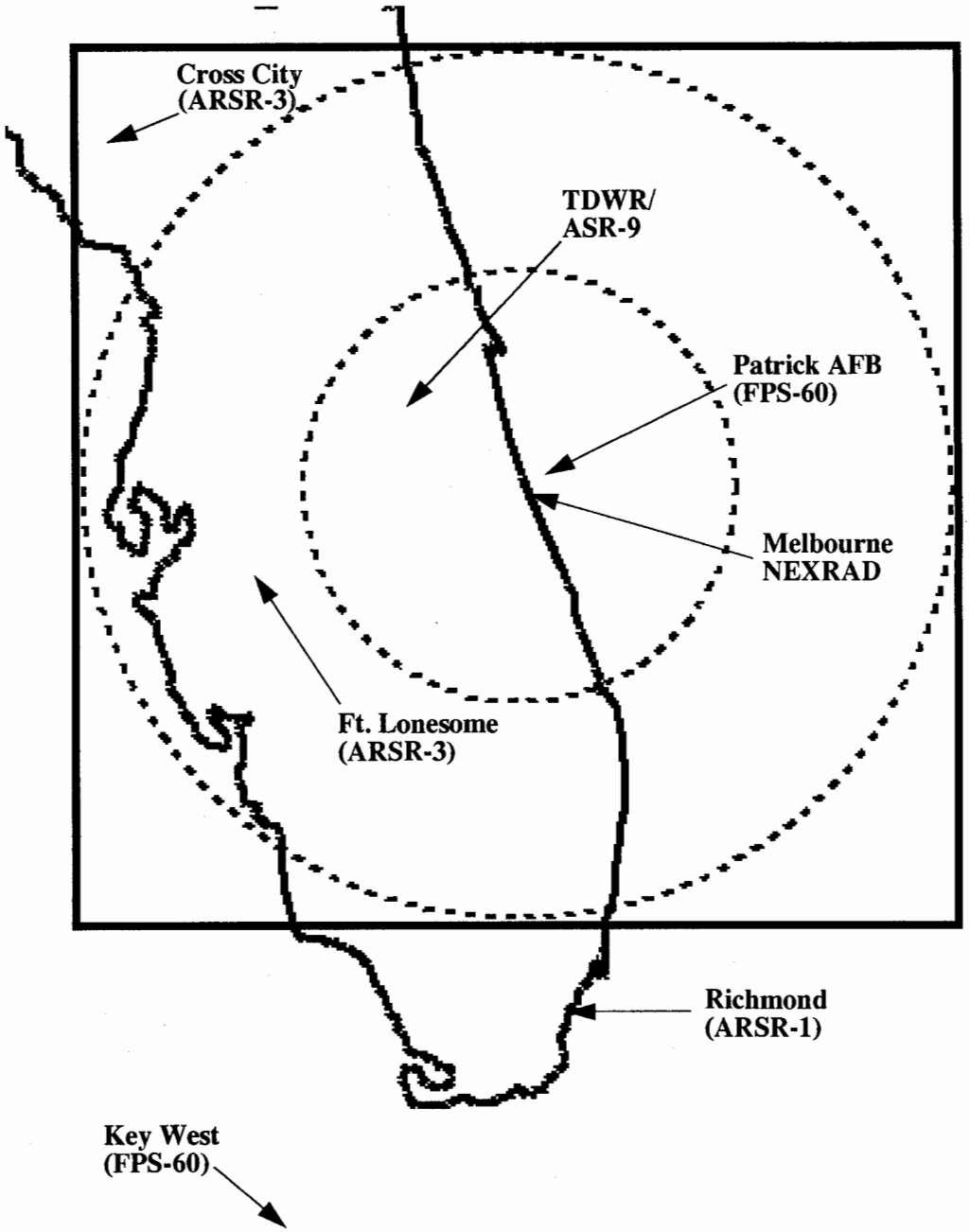


Figure 3. Study Area of Interest

3.3 RADAR ANALYSIS

After the data was collected and an assessment capability was developed, the radar analysis was performed to compare the two radar types in time and space. The first step in this analysis was to perform a visual qualitative comparison. This comparison showed significant differences between NEXRAD and ARSR; therefore, an accuracy analysis was needed to determine which radar was correct before the radar analysis could be completed (see next section). Following the accuracy analysis, the last portion of the radar analysis was performed with a quantitative comparison of the NEXRAD and ARSR data to one another in the areas of weather intensity and spatial extent to determine their relative strengths and weaknesses. NEXRAD data with and without the lag was compared. Aircraft track data from the ARSR was also used to observe aircraft weather avoidance as depicted by each radar. Sections 5 and 7 further describe the qualitative and quantitative comparisons, respectively.

3.4 ACCURACY ANALYSIS

The purpose of the accuracy analysis was to determine which radar was correct when the radars did not agree in their depiction of the weather situation. The technique selected for this study was to use additional weather radars in the study area: the TDWR and ASR-9 Testbeds near the Orlando International Airport (see figure 3). These radars are pre-production units still in a research and development phase. A "consensus" approach, using data from all four radar types, was used to determine "truth." The NEXRAD (with and without the lag) and ARSR data were then compared to this definition of truth to determine the accuracy of the radars. The results of this analysis were fed back into the radar comparison so it could be completed. Section 6 further describes the accuracy analysis.

SECTION 4

DATA SET SELECTION

Key to the study's successful completion was the acquisition of the necessary data. Data, coordinated in time and space, was required from four different types of radars, only one of which was a truly operational radar (i.e., ARSR)—acquiring the desired data from the other radars was not always possible. Even though all the desired data was not ultimately acquired, enough data was collected to complete the study.

4.1 RADAR DATA TYPES

The primary focus of this study was the comparison of NEXRAD and ARSR weather data. However, additional data was also needed to support the study. The following is a list of all the data types needed for the central Florida area to complete the study:

- Melbourne NEXRAD data
  - Composite Reflectivity, 1 and 4 km resolutions
  - Layer Composite Reflectivity: 0-24 k ft, 24-33 k ft, and 33-60 k ft
- Miami Air Route Traffic Control Center (ARTCC) ARSR data
  - Weather data for all 5 ARSRs covering central Florida
  - Aircraft beacon data for the same ARSRs
- TDWR Testbed data
  - Reflectivity data similar to the NEXRAD's composite reflectivity product
- ASR-9 Testbed data
  - Reflectivity data similar to the NEXRAD's composite reflectivity product

4.1.1 NEXRAD Data Acquisition

The Melbourne NEXRAD data was acquired in coordination with the NEXRAD Joint System Program Office (JSPO), the NEXRAD Operational Support Facility (OSF), and the Melbourne Weather Service Office (WSO). To accommodate our needs, the WSO changed their archiving

scheme to record the required data every volume scan rather than every third volume scan. When a date and time for a data set was determined, CAASD contacted the OSF, and they prepared a tape of the data from the WSO's archive when it was received from the WSO.

4.1.2 ARSR Data Acquisition

To collect the ARSR data, coordination was made with the FAA's Air Traffic Southern Region and the Miami ARTCC. When a date and time for a data set was determined, CAASD contacted the ARTCC within 15 days of the event, and they prepared a tape. For this study, the weather and aircraft beacon data was required from the five ARSR radars covering central Florida (see section 3.1).

4.1.3 TDWR and ASR-9 Data Acquisition

To collect the TDWR and ASR-9 data, coordination was made with the Massachusetts Institute of Technology's (MIT) Lincoln Laboratory (LL), the custodians of these radars while they are undergoing research tests for the FAA. When a date and time for a data set was determined, CAASD contacted LL, and they prepared a tape. Because these radars are being used in research activities, data for many of the requested dates was not available.

The TDWR data needed for the study was a reflectivity product that was as close as possible to the NEXRAD composite reflectivity product and the ARSR data. A 7 level product produced once per 6 minute volume scan at a 3.7 degree tilt was the best match—equivalent composite type products were not available.

4.2 SCAN STRATEGIES

As was noted in the section 2, the various radars use different scan strategies. Figure 4 shows a simplification of the strategies. In order to compare similar weather products, the NEXRAD composite reflectivity product was used when compared to the ARSR unless otherwise noted. The TDWR product is dissimilar to all the other products because of its one tilt at an elevation angle of 3.7 degrees—a lower angle would have provided a more similar product. This difference in the TDWR should be noted when reviewing the remainder of the study.

4.3 DATA SETS

In order to have a representative sample of data, at least three different data sets were required, preferably with differing weather conditions. Weather conditions of interest included newly initiating and dissipating thunderstorms, heavy thunderstorms, fast moving storms, and any other severe weather that could affect aviation.

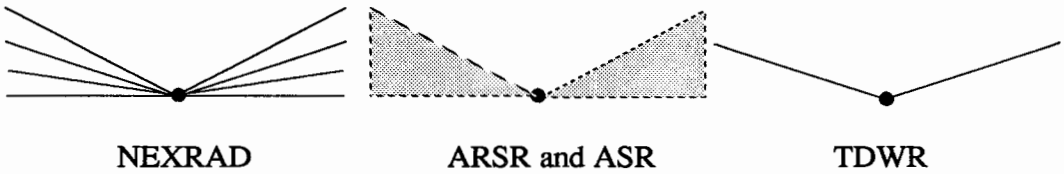


Figure 4. Cross Section of Radar Scan Techniques

The data sets were selected by monitoring the weather in central Florida using the national radar summary product. When conditions included level 5 or above, the date and inclusive times were recorded. Numerous dates were selected, but because of the 15 day time constraint for requesting the ARSR data, the "best" data sets were periodically selected and requested. Overall, 7 data sets were requested over a six month period; however, because of archiving problems at the NEXRAD site, data for only three dates was acquired. Each data set is described in the following sections.

4.3.1 23 January 1992

Depicted in this data set is a large frontal passage moving over central Florida from 16Z to 00Z on 23-24 January 1992. The precipitation area associated with the front is about 300 nautical miles long and 100 nautical miles wide (as depicted by the NEXRAD) and is moving southeast at about 20 knots. Along the front edge of this large area are 6 small cells moving at 35-40 knots in a northeast direction. The storm was already well developed at the beginning of the data set and dissipates during the data set time period.

Volume scans are 5 minutes apart, but because of archiving problems at the NEXRAD site, only every third volume scan was archived resulting in 10 minute gaps throughout the data set. Furthermore, there are two 45-60 minute gaps in the data resulting from the radar going out-of-service twice during the period.

There was no TDWR or ASR-9 data available for this data set.

4.3.2 02 June 1992

Depicted in this data set is the initiation of an afternoon and evening thunderstorm situation over central and southern Florida. The data set time period is 17Z to 06Z on 2-3 June 1992. Once fully developed, the precipitation from scattered thunderstorms covered about 25,000 square nautical miles (as depicted by the NEXRAD) and moved very slowly to the east.

Thunderstorm tops in some of the cells exceeded 50k feet and reflectivity values exceeded 57 dBZ. Volume scans are 6 minutes apart.

Data from all the radars except the ASR-9 were available for this date.

4.3.3 04 August 1992

Depicted in this data set is another afternoon and evening thunderstorm situation except this set of storms is more organized and faster moving (20-25 knots). The data set time period is 17Z to 20Z on 4 August 1992. Similar to the June data set, precipitation from these storms also covered about 25,000 square nautical miles at their peak, some storm tops exceeded 50k feet, and reflectivity values exceeded 57 dBZ. Volume scans are 6 minutes apart.

All data was available for this date; therefore, this data set was used in the accuracy analysis (see section 6).

SECTION 5

QUALITATIVE RADAR COMPARISON

The first portion of the radar analysis performed was a qualitative comparison in order to better understand the problem to be solved. Because significant differences were found in this comparison, an accuracy analysis needed to be performed to try to isolate the differences (see section 6). Once this accuracy analysis was complete, the quantitative comparison was begun (see section 7). This section describes the qualitative comparison.

5.1 METHOD

The qualitative comparison was performed visually while the workstation was sequencing through the data. It was performed on each data set for the duration of the data set. Visual samples of the results were collected, and are presented here for clarity.

5.2 FINDINGS

Because the qualitative comparison was a visual examination, it is discussed in the same way by example. In the following discussion of the findings, refer to table 1 and figure 5 for descriptions of the reflectivitiy values and their depiction in the following figures.

Unless otherwise stated, the following figures depict the NEXRAD and ARSR data perfectly matched in time, eliminating the NEXRAD "lag". Examples of the lag are also presented.

5.2.1 Depiction of Cell Boundaries

The first observation made was that the ARSR depiction of the weather with slashes and "H"s results in a display with poorly defined cell boundaries (see figure 6). From a large area view (e.g., ARTCC area), the depiction appears adequate; however, in a more realistic smaller area depiction (e.g., a sector area) the depiction may be very coarse. If multiple ARSRs are covering the area, the overall cell boundaries are reasonably depicted, but with only one or two ARSRs covering the area, the depiction is so coarse that it is difficult to determine the cell boundaries. The heavy area ("H"s) boundaries are especially hard to distinguish, and even more so if multiple radars are painting the same area because it is not clear which "H"s are paired together to show the bounds of the area.

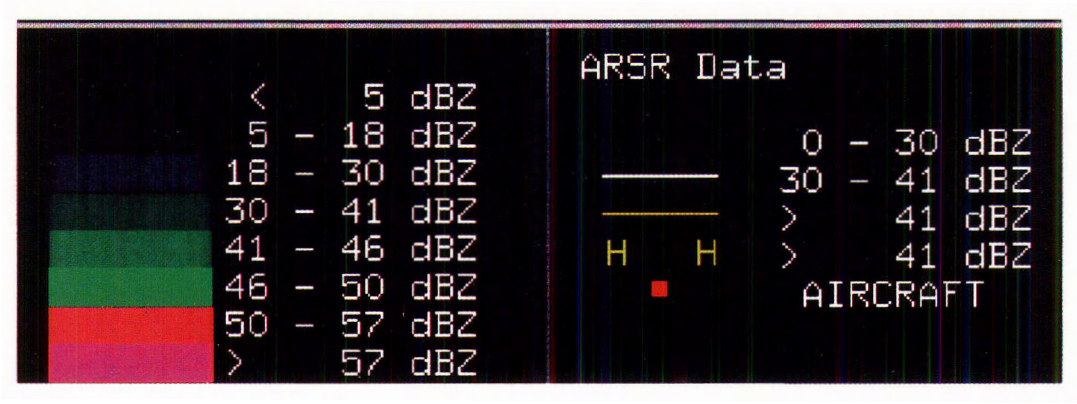


Figure 5. Depictions of Reflectivity Values

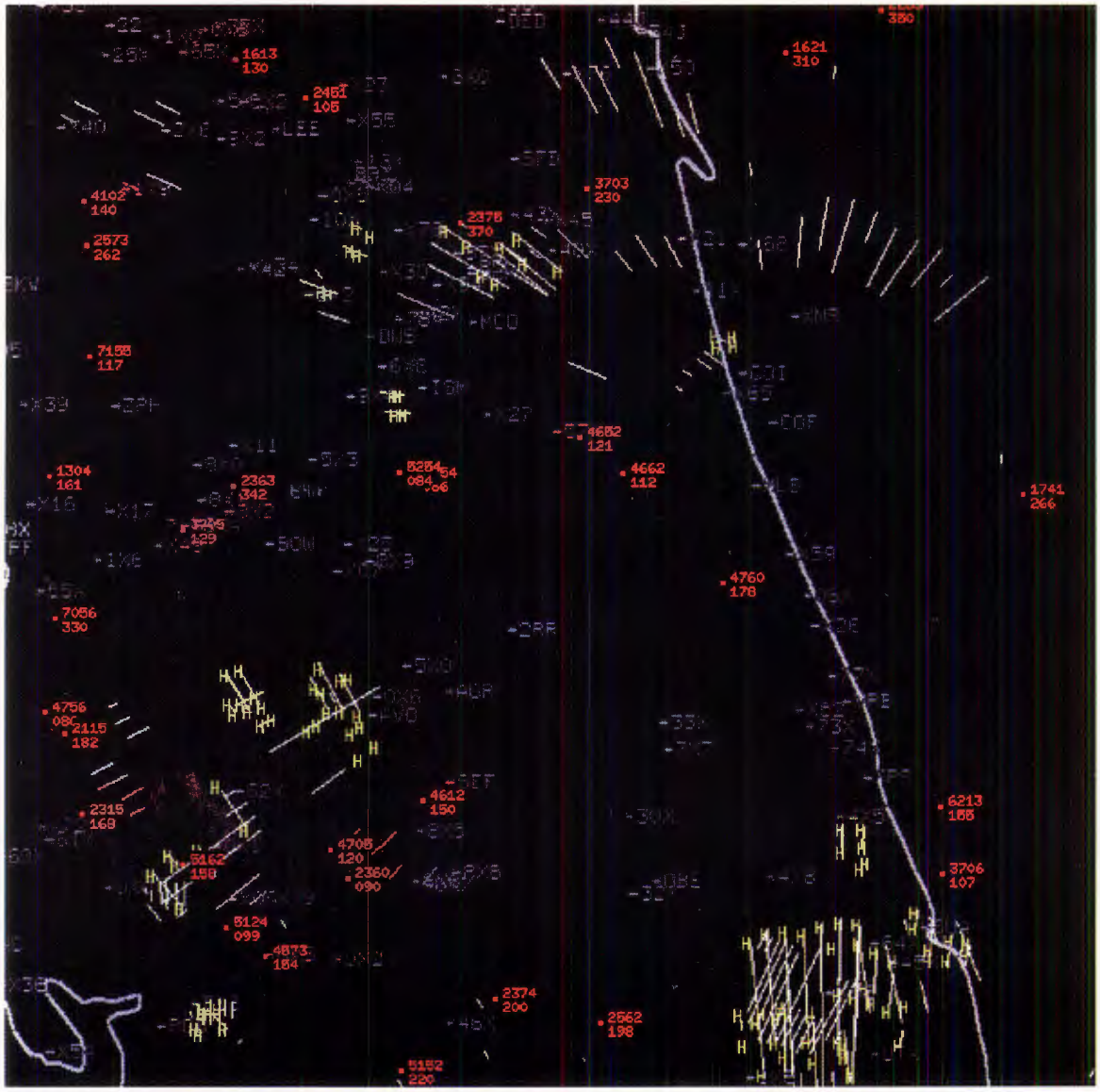


Figure 6. Examples of ARSR Depiction of Weather

5.2.2 ARSR Level Thresholds

The most striking finding was the variation amongst ARSRs in depicting a weather cell. The thresholds for defining moderate and heavy levels appeared to vary considerably from radar to radar and from data set to data set for a given radar. In areas of overlap among ARSRs, some ARSRs would depict an area as heavy and others would not detect it at all. In areas of overlap with NEXRAD, there appeared to be little correlation of moderate and heavy ARSR levels to the equivalent NEXRAD levels.

For example, in figure 7, the ARSR moderate lines should match the darker green areas of the NEXRAD; however, there is clearly more area covered by the NEXRAD than by the ARSR indicating the possibility that the NEXRAD is more sensitive. In figure 8, the large blue area depicted by the NEXRAD as very light is depicted by 2 ARSRs as heavy and by 1 ARSR as moderate. This example leads to the possibility that the ARSRs are more sensitive and different ARSRs are not using the same threshold values. This apparent variation in threshold values seems to indicate that the ARSRs are not calibrated similarly and that they may not be able to effectively distinguish between various weather levels.

5.2.3 Mutually Exclusive Areas

There were also areas where ARSR depicted weather and the NEXRAD did not, and the NEXRAD depicted weather and ARSR did not.

In the first case, the weather depicted by the ARSR appeared to be spurious because low altitude aircraft routinely penetrated what was sometimes depicted as heavy weather. This apparent problem may be related to the level threshold problem discussed above. For example, in figure 9, ARSR depicts a large area of heavy weather towards the top of the figure but NEXRAD depicts virtually nothing.

The second case was a little more unusual. The instances of NEXRAD detecting weather and ARSR missing it were typically seen when severe weather was nearby the ARSR. If another ARSR overlapped that area sufficiently, there was no apparent loss of information. However, there appeared to be many instances where there was not sufficient overlap and severe weather (as depicted by the NEXRAD) went undetected. For example, in figure 9, in the center of the figure is an area of very intense weather as depicted by the NEXRAD in red, but the nearby ARSR (and any other ARSR) did not detect this weather nor much of the less intense weather nearby the radar.

This problem with severe weather near the ARSR may be explainable. Apparently, if the radar front end saturates due to the high reflectivity so close to the radar, an area around the radar is automatically blanked out. This area is in the 20-25 nm range. No official documentation

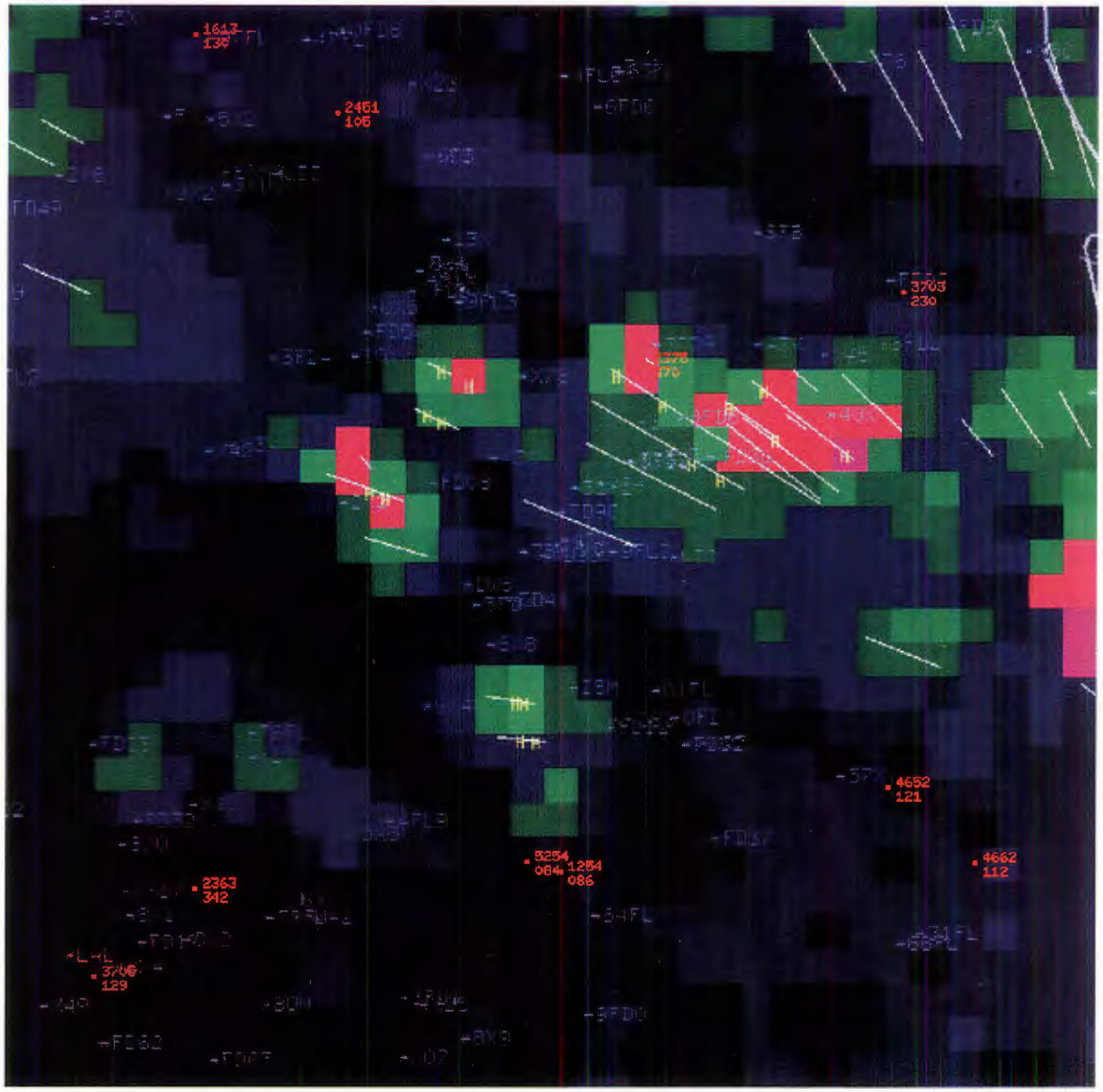


Figure 7. Example of Apparent ARSR Insensitivity

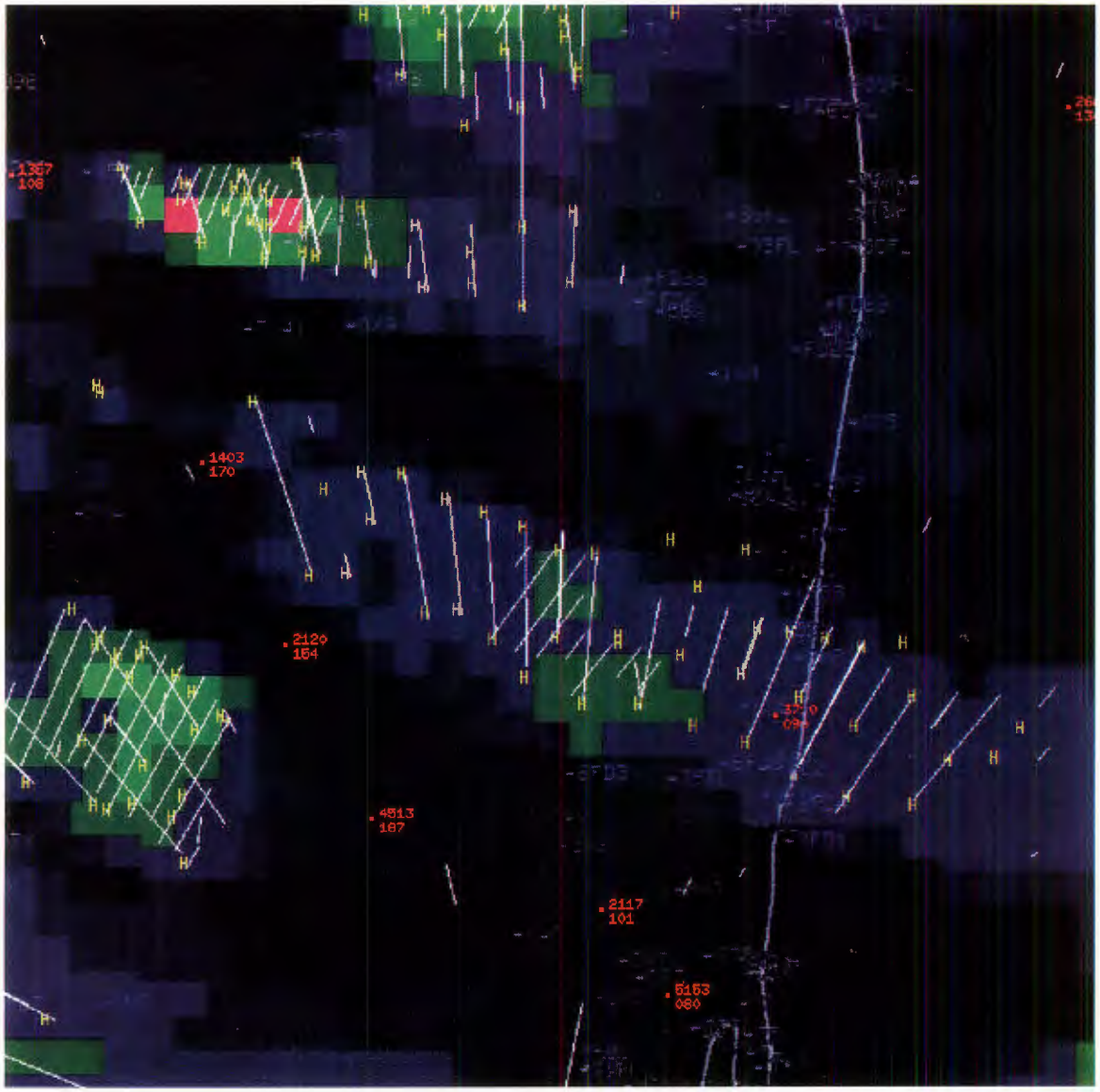


Figure 8. Example of Apparent ARSR Over Sensitivity

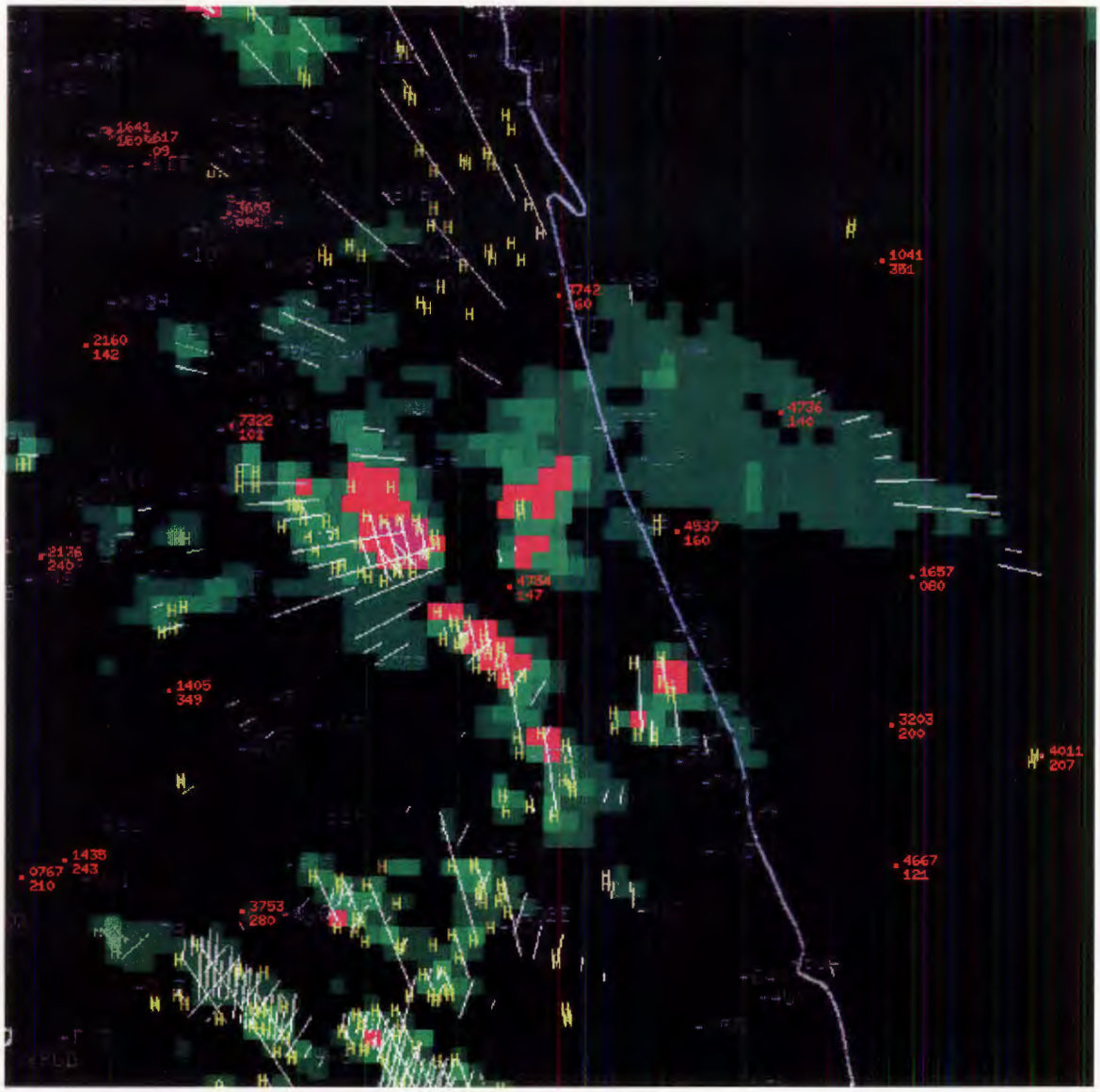


Figure 9. Example of ARSR Depiction of Weather
Where NEXRAD Did Not Depict Weather

stating this feature has been found. If this is an ARSR "feature," then controllers may be routinely missing significant weather on their displays, today.

The point to keep in mind is that from an air traffic controller's view, the ARSR appears to miss significant weather and show weather that does not exist—both of which could negatively impact ATC operations.

5.2.4 NEXRAD "Lag"

In the January data set, there is a front moving through central Florida in a southeast direction at about 20 knots, and at the leading edge of the weather mass associated with the front are small cells moving relatively fast (35-40 knots) to the northeast along the edge of the mass (see figure 10). Figure 11 shows a closer view of the small cells with no correction to the time, as a controller would see the situation. In this figure, there is clearly a horizontal offset of 3-6 nm between the NEXRAD and ARSR depictions with what appears to be the NEXRAD depiction lagging behind the ARSR depiction. This is an example of the NEXRAD data aging, or lag, problem. The operational impact of the lag was determined in the quantitative comparison and is presented in section 7.

5.2.5 Altitude Discrimination

Because the ARSR is a fan beam radar, it is not expected to be able to discriminate weather at different altitudes. However, the NEXRAD can discriminate weather in three altitude layers matching the nominal heights for the ATC low, high, and super high sectors. As part of the qualitative comparison, the area covered by weather at the different altitude sectors was examined. As can be seen in figures 12-14, the higher the altitude, the less weather is detected by the NEXRAD. This is typical of a thunderstorm weather situation. The significance of this is that for the high and super high sectors, using these NEXRAD products would show a more accurate depiction of the weather at altitude and "free up" a great deal of airspace.

Another altitude related observation that was made was that very few aircraft penetrate significant weather at their altitude as depicted by NEXRAD using the NEXRAD layered products. Only aircraft flying within the NEXRAD altitude layers were observed. However, the aircraft do appear to penetrate the weather as depicted by the ARSR. The importance of this is that the NEXRAD appears to better represent what the pilot is observing.

5.3 CONCLUSIONS

The qualitative comparison concludes the following:

- The ARSR gives a coarse depiction of cell boundaries

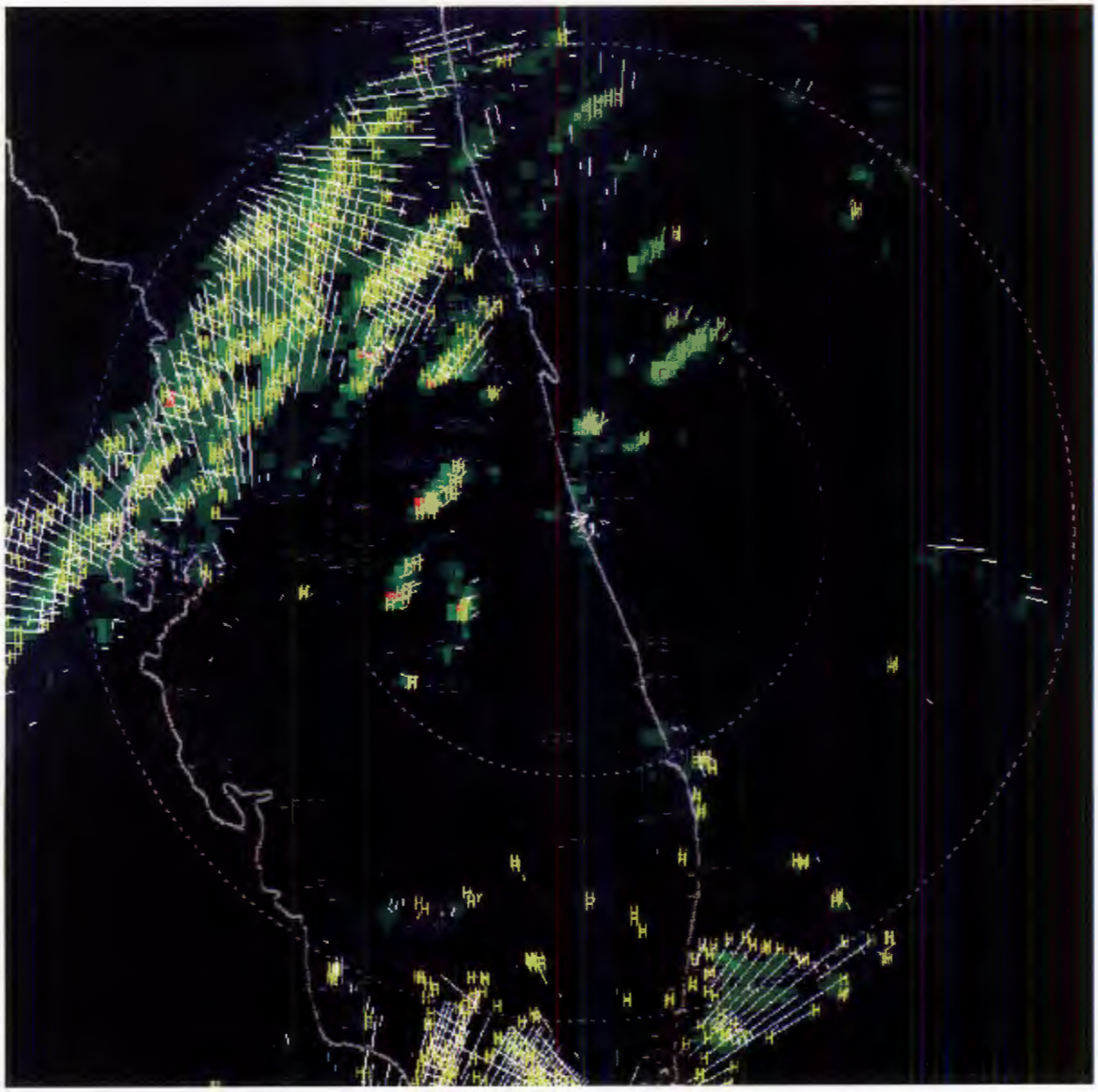


Figure 10. Example of 23 January 1992 Data Set

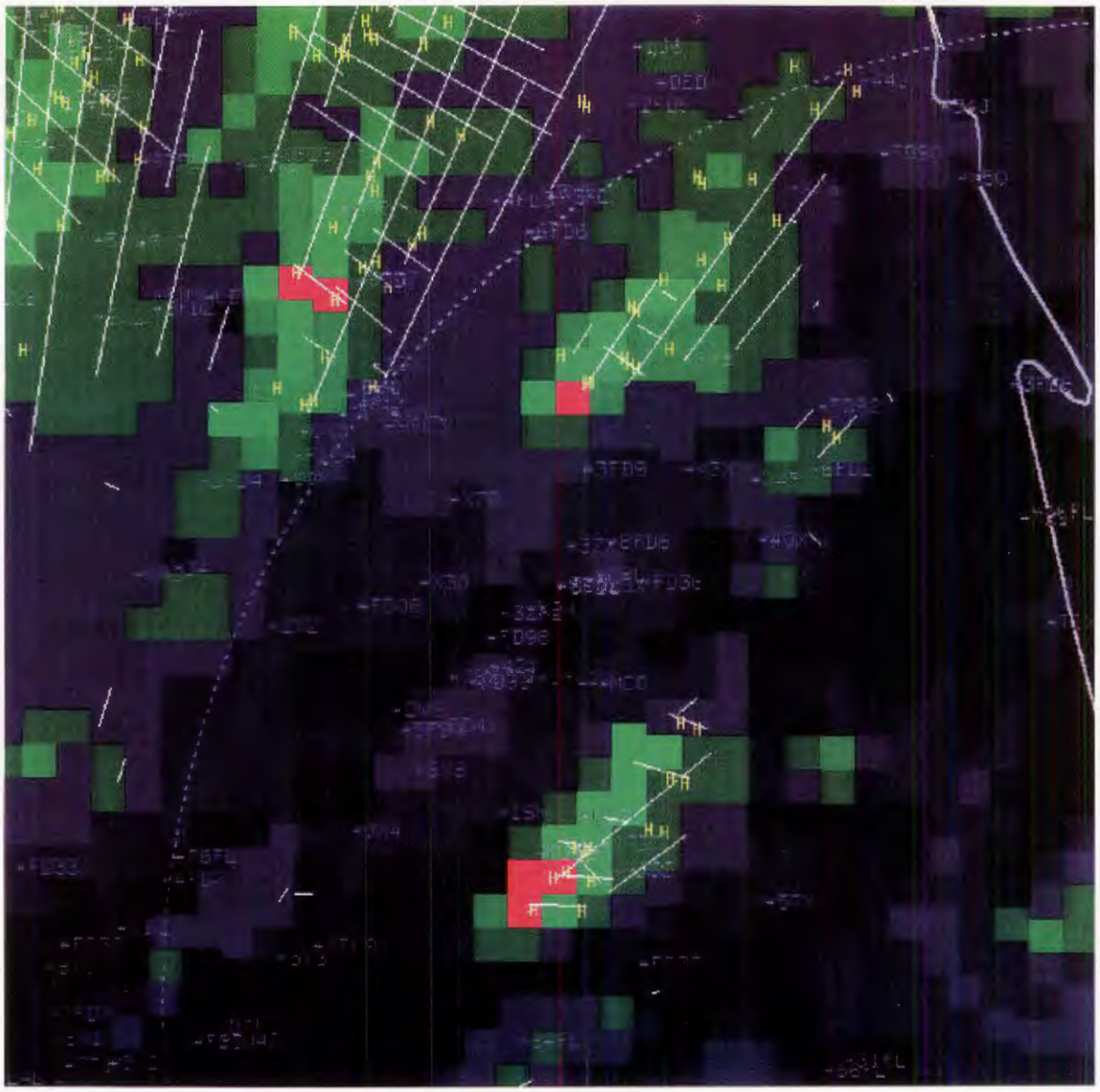


Figure 11. Example of NEXRAD "Lag"

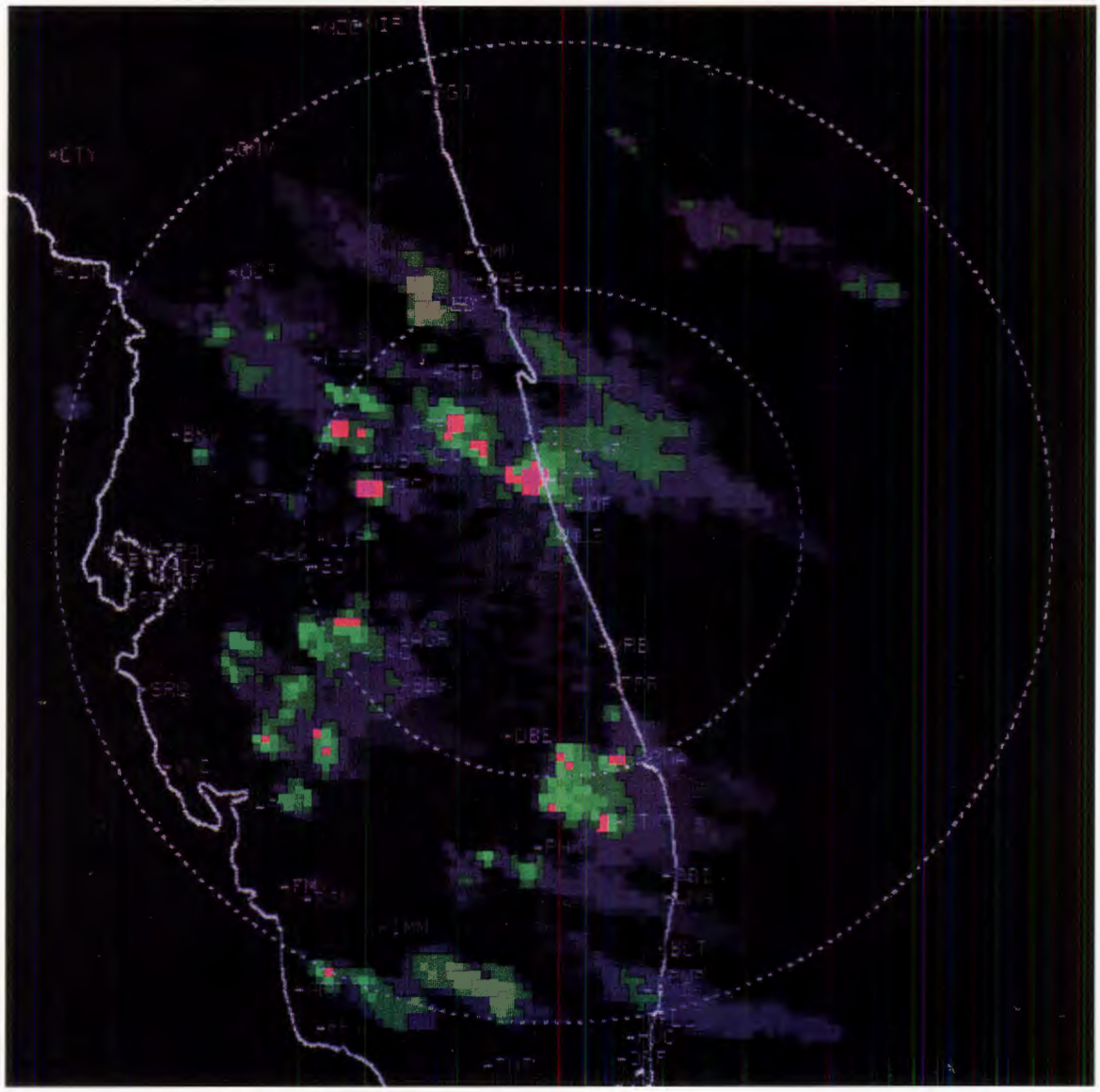


Figure 12. Example of NEXRAD Low Layer

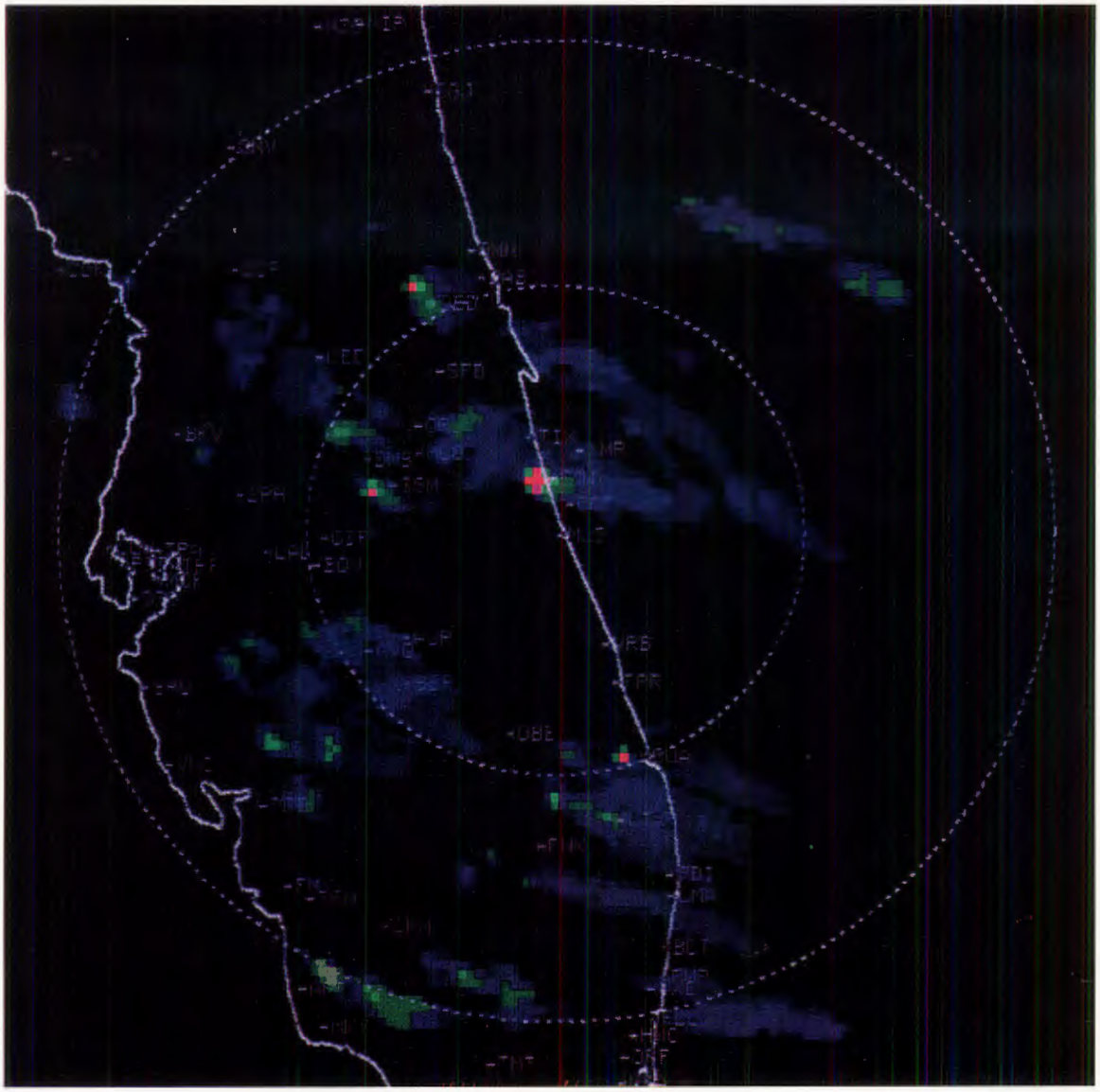


Figure 13. Example of NEXRAD High Layer

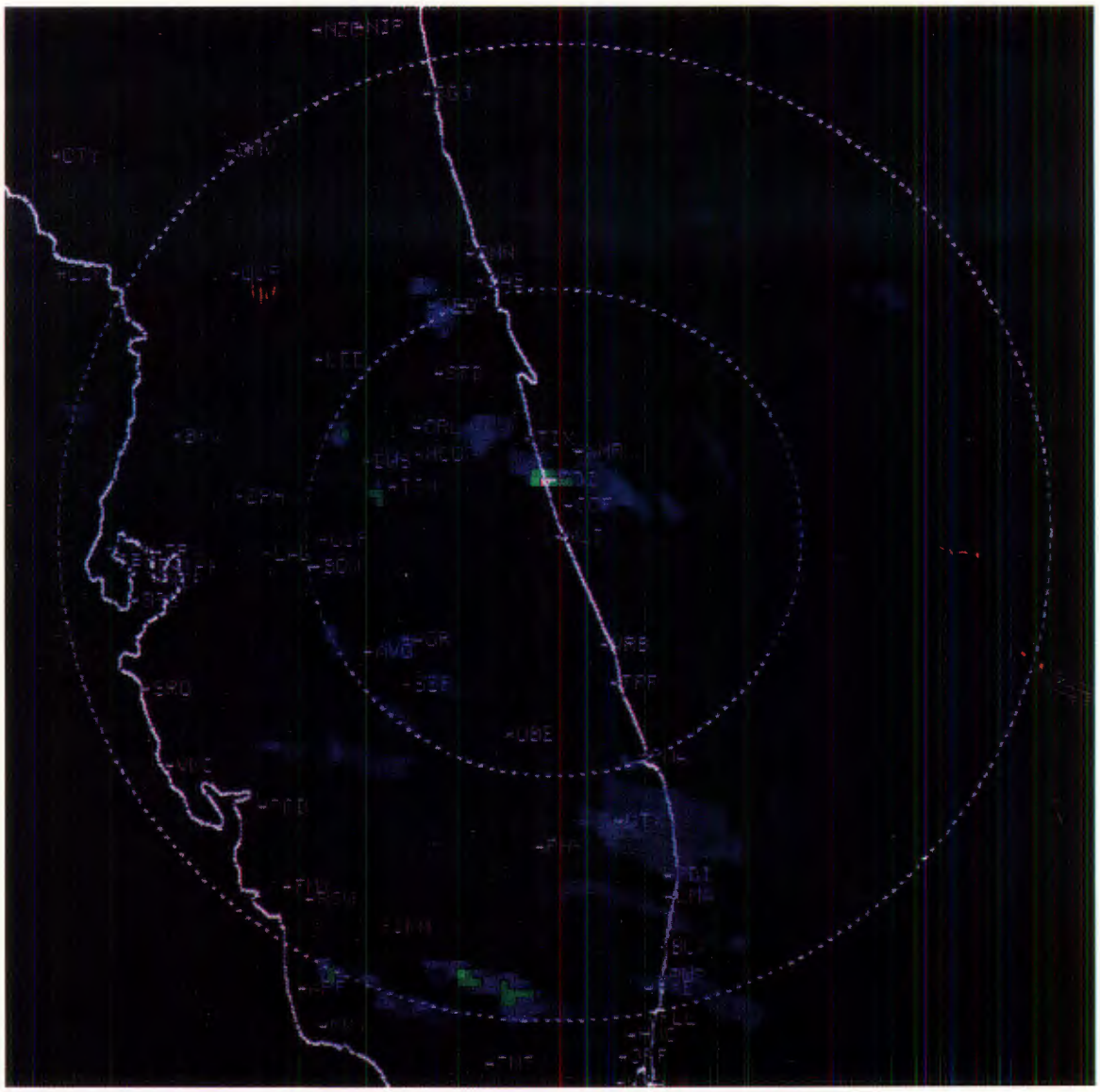


Figure 14. Example of NEXRAD Super-High Layer

- The ARSR level thresholds appear to vary considerably
- The ARSR appears to miss significant weather and detect spurious weather
- The NEXRAD data aging "lag" is clearly visible with fast moving cells
- NEXRAD offers altitude discrimination
- NEXRAD appears to better depict the weather the pilot is observing at different altitudes

In order to verify these qualitative findings, a determination of which radar (if any) is more accurate needed to be performed—that is the accuracy analysis described in the next section.

SECTION 6

ACCURACY ANALYSIS

The study required a measure of the accuracy of the NEXRAD and ARSR radars in order to establish which radar was more accurate as well as to explain differences discovered during the qualitative comparison of the radars. Determining the accuracy the radar data required that a definition of truth be constructed. The accuracy analysis defined truth using the radar data from NEXRAD, ARSR, TDWR and ASR radars in combination with a consensus approach. Once truth was defined, the NEXRAD and ARSR radar data was compared to this definition of truth to establish a measure of their accuracy. The accuracy analysis included both an absolute analysis (with the NEXRAD lag removed) and a worst case analysis (where the NEXRAD data is evaluated when it is at its oldest time in the cycle). This section describes the method used to determine truth as well as the results of the truth comparison.

6.1 METHOD OF ANALYSIS

This subsection describes the method used to assess the accuracy of the NEXRAD and ARSR radars, including the specifics of the data used, the procedure used to determine truth as well as the method of comparing the data and assigning a value of accuracy to each radar.

6.1.1 Data Specifications

The accuracy analysis used a square region measuring 133 km on a side and centered at the location of the Orlando ASR radar. The box in figure 15 delimits the region used in the accuracy analysis (the rings delimit distances of 62 nm and 124 nm from the Melbourne NEXRAD). This was the largest square area that the radar types adequately covered. The analysis used snapshots of the data between the times of 17:45 and 20:00 from the August 4 data set. The snapshots were taken at different times relative to the NEXRAD's scan strategy, and are described in the following sections. All of the data presented is drawn from an aggregate of all the snapshots. The accuracy analysis was not performed on the January and June data sets because the ASR and TDWR were unavailable for those data sets. Section 2 describes the specifics of each of the 4 types of radar data. The analysis did not use the standard method of displaying ARSR data as lines and H's for moderate and heavy data, respectively, since this is a simplification done for display reasons. Instead the data was drawn as trapezoids or "pie wedges" centered on the appropriate azimuth and extending to the midpoint between the current azimuth and the adjacent azimuth on both sides. Figure 16 gives an example of this technique. The dashed lines represent the possible reported azimuths of the radar data. The solid lines represent where the ARSR data specified that the radar detected weather. The shaded

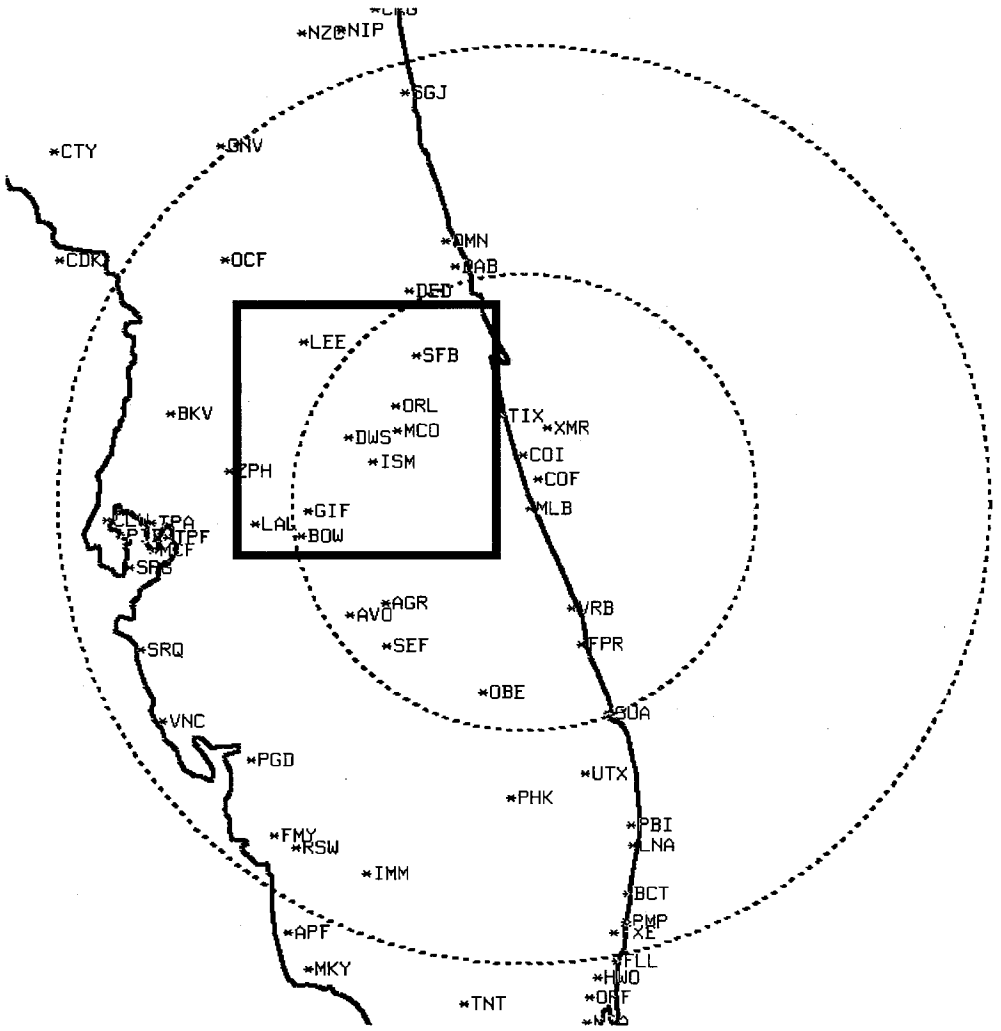


Figure 15. Region Used in Accuracy Analysis

"pie wedges" represent the area that was used in our analysis. This method gives the ARSR data continuous coverage which is a better representation for the actual radar coverage.

6.1.2 Determining Truth

The analysis used a 100 element by 100 element grid with each element having a resolution of 1.33 km to represent a truth grid. Each element of the truth grid was scanned and then the

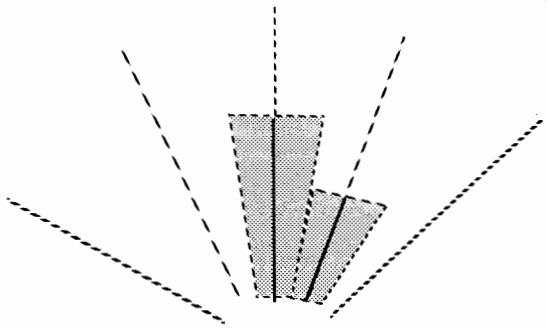


Figure 16. ARSR Display Technique

corresponding location in each of the radar images was checked for data at or above the correct reflectivity threshold in order to determine consensus. The analysis included gathering data for reflectivity thresholds of 30 dBZ and 41 dBZ. If the correct reflectivity level was found at a consensus of the radars, then that element in the truth grid had weather, otherwise it did not. Since the TDWR product that was received was not a multi-elevation angle product, the analysis used both two and three radars for the consensus value (the number of radars used for consensus refers to the number of radar types used, where all the ARSR radars constitute one radar type). Using two radars for consensus helps keep the TDWR data from reducing the truth area which was desired because the TDWR's beam rises quickly as range increases (because of the 3.7 degree elevation angle) which could cause it to only see the top of a storm and miss much of the data. Truth grids were created using both 1km and 4 km resolution NEXRAD product. The NEXRAD and ARSR data were then compared to this truth grid to establish a measure of their accuracy. Figure 17 gives an example of the consensus approach where two was used as the definition of consensus.

6.1.3 Determining Accuracy

The Critical Success Index determined the accuracy of a radar (CSI) (Dixon, 1992) and is computed using three statistics: hits, misses and false alarms (see figure 18). A hit meant that the truth grid indicated weather at a grid point and the radar detected weather at that same grid point. A miss meant that the truth grid indicated weather at a grid point and the radar did not detect weather at the same grid point. A false alarm meant that the truth grid indicated that there was no weather at a grid point while the radar indicated that there was weather at the same grid point. The analysis defined the CSI as the number of hits multiplied by 100 divided by the total of the number of hits plus the number of misses plus the number of false alarms. The value of

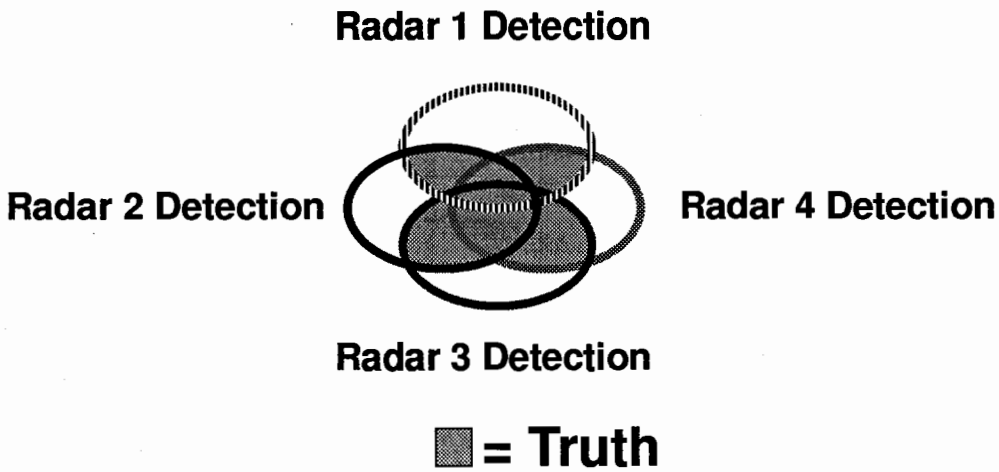
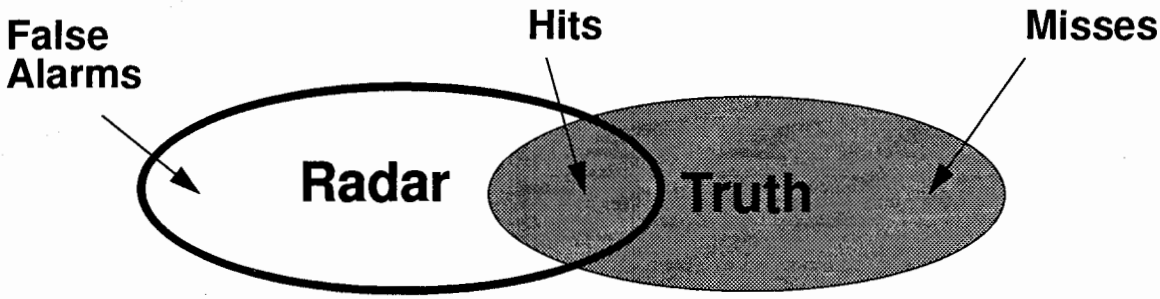


Figure 17. Consensus Approach, Two Radars Equals Consensus



$$\text{CSI} = \frac{\text{Hits}}{\text{Hits} + \text{False Alarms} + \text{Misses}} \times 100$$

Figure 18. Calculation of CSI

the CSI can range from 0 to 100 and it represents the percentage that hits are of the total area of weather plus false alarms.

In addition to the exact match criteria described above, the definition of the accuracy parameters was relaxed for another analysis which is referred as the tactical analysis. The relaxation of the parameters is referred to as the expansion of the search area. This expansion allowed a radar to score better if it came 'close' to determining the position of the weather rather than requiring an exact match. For a search area expansion value of X grid points, the new definition of a hit meant that the radar detected weather at the current grid point and the truth indicated weather at that exact grid point or any of the grid points X away from it. For an expansion value of one, the nine grid points adjacent to and including the current grid point were searched. Figure 19 shows how the expansion was performed. The nine numbered grid points represent the grid points that would be searched for an expansion value of 1 when the current grid point was at the grid point labeled 5. A miss is where the radar did not detect weather at the current grid point and truth indicated that there was weather at the current grid point as well as all grid points X away from it. A false alarm is where the radar detected weather at the current grid point and truth indicated that there was no weather at the current grid point as well as no weather at all grid points X away from it.

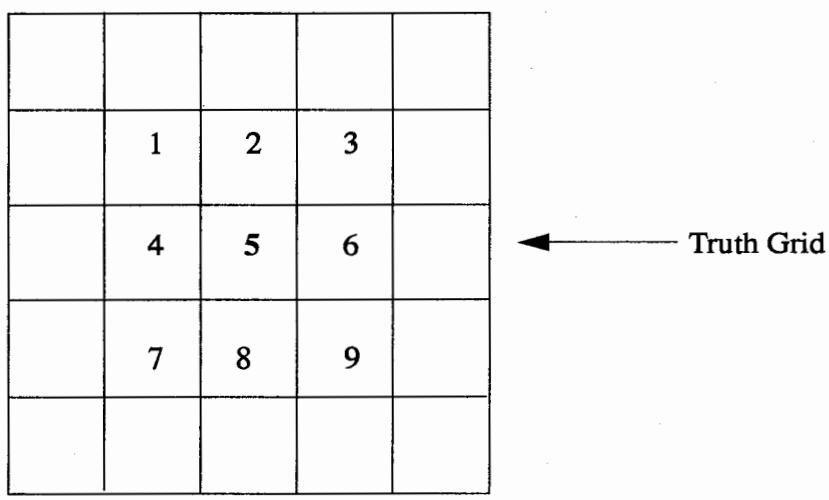


Figure 19. Expansion of Search Area

6.2 FINDINGS

The truth analysis, as described above, used two approaches, absolute and worst case. In the absolute analysis, the data was matched in time (i.e. with no time lags for NEXRAD) in order to determine the absolute accuracy of the radars at detecting weather. The worst case analysis analyzed the data at the moment that the NEXRAD data was at its oldest in order to get a worst case measure of the operational accuracy of the NEXRAD data. The following subsections describe the findings of these analyses. It should be kept in mind that the ARSR CSI values presented below were determined using input from all ARSRs that detected weather within the truth area.

6.2.1 Absolute Analysis

The absolute analysis used data that was matched in time (with all associated sources of delay removed) with the net result being that the absolute sensing capabilities of the radars was being measured as opposed to the operational accuracy (that results when the radar data has been delayed by processing) being measured. This determined which radar actually detects weather more accurately. Tables 2 through 5 show the results of the absolute analysis.

Table 2 shows the results of comparing the NEXRAD and ARSR data corresponding to reflectivity levels greater than or equal to 41 dBZ (NWS level three and greater) to a truth grid constructed for the same levels. The first column shows the resolution of the NEXRAD data that was used to determine truth as well as that used to determine the NEXRAD CSI. The second column shows the number of radars that determined consensus. The third column shows the number of grid units that the search area was expanded. The fourth and fifth columns contain the NEXRAD and ARSR CSI values respectively. The last column shows the percent improvement that NEXRAD showed over ARSR. Examination of the data in the table shows

that the NEXRAD showed significant improvement over the ARSR, showing at least a 130 percent improvement.

Table 2. CSI Values For Data Greater than or Equal to 41 dBZ

NEXRAD Resolution (km)	# of Radars = Consensus	Expansion of Search Area (1.3 km grid)	NEXRAD CSI	ARSR CSI	% Improvement
4	2	0	70.53	30.72	130
		1	85.11	34.89	144
		2	91.46	38.52	137
	3	0	28.35	11.75	141
		1	41.18	13.93	196
		2	56.43	16.12	250
1	2	0	73.85	20.82	255
		1	93.55	24.82	277
		2	96.55	28.20	242
	3	0	46.51	9.86	372
		1	69.12	11.54	499
		2	81.78	13.39	511

The best measure of the NEXRAD and ARSR CSIs for the absolute analysis are 73.85 and 20.82 respectively, and can be found in the row corresponding to one km resolution NEXRAD data, two radars equalling consensus, and an exact match (zero expansion). The one km resolution data was chosen because that is what the NEXRAD's actual resolution is (for this product) and that is also the most accurate representation to use as the NEXRAD's input into the truth process. Two radars equalling consensus was chosen because of the scanning characteristics associated with the TDWR. Using two as the consensus value enabled the inclusion of the data that was detected by the TDWR while keeping the areas that were not detected by the TDWR (possibly because they were not in the beams path) from artificially

reducing the area of the weather in the truth. Although the percent improvement for NEXRAD was greater when three radars were used for consensus, the CSI values were significantly lower for both NEXRAD and ARSR which is mainly due to the limiting affect the TDWR had on the size of truth.

Table 3 shows the raw data that was used to compute the CSI values that are presented in table 2. The first three columns correspond to the first three columns in table 2. The next three

Table 3. Raw Truth Statistics For Data Greater Than or Equal to 41 dBZ

Resolution	Consensus	Expansion	NEXRAD			ARSR		
			Hits	Misses	False Alarms	Hits	Misses	False Alarms
4	2	0	9101	384	3418	9365	120	20998
		1	8794	18	1521	9059	6	16902
		2	8509	0	795	8769	0	13998
	3	0	3555	19	8964	3570	4	26793
		1	3427	0	4896	3443	0	21266
		2	3287	0	2538	3303	0	17192
1	2	0	5333	1102	786	6345	90	24040
		1	5135	48	306	6102	2	18483
		2	4928	1	175	5853	0	14906
	3	0	2895	105	3224	2997	3	27388
		1	2787	2	1243	2885	0	22124
		2	2675	0	596	2770	0	17924

columns present the three statistics that were used to compute the CSI value for the NEXRAD and the last three columns present the statistics used to compute the CSI value for ARSR. Examination of the data shows that the number of hits and misses for both radars was similar

while there was a significant difference in the number of false alarms with the ARSR showing the larger number of false alarms. Close examination of the number of hits and misses for each radar shows that the ARSR had a slightly larger number of hits and a slightly smaller number of misses which is to be expected with the large number of false alarms that are present in the ARSR data. This is due to the fact that if a radar painted its entire coverage area, it would hit everything and miss nothing, but its false alarm rate would be extremely high. This is the reason that the CSI value is a good measure of a radars overall accuracy because it takes into account all three of the parameters.

A peculiarity that may be noticed in table 3 is that the number of hits did not increase as the search area was expanded. This is due to the fact that when the search area is expanded by x grid cells, the overall area that can be tested is reduced by x around the perimeter of the truth area. If the expansion was two, for example, the width and height of the truth area would be reduced by four (two on each side). For this reason, the CSI values in table 1 should be used if comparisons between rows are to be performed instead of comparing data rows in table 3 to one another.

Table 4 shows the results of the same analysis as in table 2, but with a reflectivity threshold of 30 dBZ (NWS level two and above). Comparison with the data contained in table 2 will show that the trends are the same even though the 30 dBZ threshold data shows slightly improved absolute CSI values with a slight reduction in the relative improvement of the NEXRAD over the ARSR. The NEXRAD, though, again showed significant improvement in accuracy over that of the ARSR in all cases.

Table 4. CSI Values For Data Greater Than or Equal to 30 dBZ

NEXRAD Resolution (km)	# of Radars = Consensus	Expansion of Search Area (1.3 km grid)	NEXRAD CSI	ARSR CSI	% Improvement
4	2	0	77.10	38.51	100
		1	89.95	43.80	105
		2	94.81	49.00	93
	3	0	40.36	18.88	114
		1	55.20	22.02	151
		2	69.56	25.05	178
1	2	0	79.37	28.79	176
		1	95.69	34.19	180
		2	98.37	39.30	150
	3	0	60.84	16.75	263
		1	81.54	19.39	321
		2	90.30	22.18	307

Table 5 shows the raw data that was used to compute the CSI values contained in table 4.

Table 5. Raw Truth Statistics For Data Greater Than or Equal to 30 dBZ

Resolution	Consensus	Expansion	NEXRAD			ARSR		
			Hits	Misses	False Alarms	Hits	Misses	False Alarms
4	2	0	18267	853	4573	18980	140	30171
		1	17490	45	1910	18205	2	23353
		2	16734	4	912	17432	0	18146
	3	0	9237	44	13603	9281	0	39870
		1	8840	0	7175	8882	0	31463
		2	8449	0	3697	8491	0	25405
1	2	0	12099	2196	948	14215	80	35074
		1	11560	196	325	13609	1	26196
		2	11015	23	159	12984	0	20050
	3	0	8058	197	4989	8254	1	41035
		1	7688	3	1738	7871	0	32722
		2	7318	0	786	7492	0	26286

Comparison with the data in table 3 shows that the trends were the same (similar number of hits for the NEXRAD and ARSR, much larger number of false alarms for ARSR). There are no significant differences in the data and it is provided for the purposes of completeness.

The data presented thus far supports the conclusion that NEXRAD is indeed the more accurate radar at detecting weather. Because the data was matched in time (with the lag removed) this analysis represents a measurement of the radars' absolute ability to detect weather, and therefore did not address the major concern about NEXRAD's suitability for tactical ATC use

which is data aging. The worst case analysis addressed these concerns and is described in the following section.

6.2.2 Worst Case Analysis

For the worst case analysis, the truth grid and the ARSR CSI values were determined as in the absolute analysis. The only difference in the worst case analysis was that the NEXRAD CSI value was computed using the previous NEXRAD image. In this manner the truth grid had the most up-to-date information while the NEXRAD was at its oldest (worst) age. The processing time for the ARSR (see table 7) was subtracted from the NEXRAD delay with the net result being a comparison of the operational worst case accuracy of the NEXRAD compared to ARSR. Tables 8 through 11 show the results of the worst case analysis.

Table 6. NEXRAD Delay Parameters Used In the Accuracy Analysis

Delay Parameter	Seconds
Volume scan time	360
WSR88D-RWP transmission	45
Real-time Weather Processor (RWP) processing	20
RWP-Area Control Computer Complex (ACCC) transmission	1
ACCC processing	30

Table 7. ARSR Delay Parameters Used In the Accuracy Analysis

Delay Parameter	Seconds
Weather Fixed Map Unit (WFMU) processing and transmission	10
Host processing	1

Table 8 shows the results of doing a worst case (for the NEXRAD) accuracy analysis for reflectivity levels greater than 40 dBZ. The NEXRAD does show a reduction in accuracy when

the worst case CSI values are compared to the corresponding absolute accuracy CSI values in table 2. The important detail, though, is that NEXRAD is still the more accurate radar. Even in the worst case, NEXRAD showed a significant improvement over ARSR, showing an improvement of at least 48 percent.

Table 8. Worst Case, CSI Values For Data Greater Than or Equal to 41 dBZ

NEXRAD Resolution (km)	# of Radars = Consensus	Expansion of Search Area (1.3 km grid)	NEXRAD CSI	ARSR CSI	% Improvement
4	2	0	47.53	32.07	48
		1	65.32	36.94	77
		2	79.25	41.16	93
	3	0	20.57	11.47	79
		1	28.63	13.51	112
		2	37.12	15.70	136
1	2	0	35.81	22.60	58
		1	66.04	27.21	143
		2	85.63	31.08	176
	3	0	20.25	8.67	134
		1	32.92	10.04	228
		2	45.74	11.72	290

Table 9 presents the raw data that was used to compute the CSI values contained in table 8. Examination of the data will reveal that the disparity in the number of hits and misses between the NEXRAD and ARSR increased. This, again, is to be expected since this is the worst case

for the NEXRAD data. ARSR still showed a large number of false alarms in comparison to NEXRAD.

Table 9. Worst Case, Raw Data Statistics For Data Greater Than or Equal to 41 dBZ

Resolution	Consensus	Expansion	NEXRAD			ARSR		
			Hits	Misses	False Alarms	Hits	Misses	False Alarms
4	2	0	7350	3332	4781	10537	145	22179
		1	7067	1003	2749	10154	5	17327
		2	6803	221	1560	9793	0	14000
	3	0	2710	1043	9421	3572	1	28964
		1	2611	91	6418	3616	0	23147
		2	2521	0	4270	3488	0	18722
1	2	0	3503	3924	2355	7369	58	25173
		1	3359	762	965	7077	1	18935
		2	3206	79	459	6758	0	14987
	3	0	1462	1362	4396	2822	2	29720
		1	1403	107	2752	2713	0	24316
		2	1348	1	1598	2615	0	19704

Table 10 presents the results of the worst case analysis done for reflectivity levels greater than 30 dBZ. The trends are nearly identical to those in table 8 in terms of NEXRAD's improvement

over ARSR. Both radars showed an improvement in the absolute CSI value with the NEXRAD showing a slight reduction in its improvement over ARSR.

Table 10. Worst Case, CSI Values For Data Greater Than or Equal To 30 dBZ

NEXRAD Resolution (km)	# of Radars = Consensus	Expansion of Search Area (1.3 km grid)	NEXRAD CSI	ARSR CSI	% Improvement
4	2	0	54.05	38.73	40
		1	70.58	44.45	59
		2	83.04	49.95	66
	3	0	31.63	18.54	71
		1	42.14	21.50	96
		2	52.17	24.40	114
1	2	0	45.86	29.66	55
		1	71.64	35.24	103
		2	88.44	40.39	119
	3	0	34.00	15.47	135
		1	51.37	17.81	188
		2	65.93	20.31	225

Table 11 presents the raw data that was used to compute the CSI values in table 10. There are no significant differences in the trends of tables 9 and 11; the data is provided in the interests of completeness.

Table 11. Worst Case, Raw Truth Statistics For Data Greater Than or Equal to 30 dBZ

Resolution	Consensus	Expansion	NEXRAD			ARSR		
			Hits	Misses	False Alarms	Hits	Misses	False Alarms
4	2	0	15221	5451	7489	20481	191	32208
		1	14524	1837	4217	19593	1	24482
		2	13858	531	2300	18724	0	18763
	3	0	7805	1968	14905	9771	2	42918
		1	7446	362	9861	9332	0	34068
		2	7098	35	6473	8917	0	27633
1	2	0	8990	6781	3832	15686	85	37118
		1	8552	1848	1537	14988	1	27547
		2	8101	334	725	14262	0	21051
	3	0	5325	2842	7497	8167	0	44637
		1	5056	465	4322	7792	0	35960
		2	4781	62	2409	7420	0	29111

The worst case analysis shows that NEXRAD is the more accurate radar even when the NEXRAD data is used at its oldest moment. Using the one km data for truth and for determination of the CSI, though, is valid only for determining the absolute accuracy of the NEXRAD, but using the four km data for truth and for determining the CSI is not an accurate assessment of the operational accuracy since there was a better representation of truth in the one km data. The operational analysis eliminated these problems, and this analysis is described in the following section.

6.2.3 Operational Analysis

The operational analysis used the one km NEXRAD data in the determination of truth, while the four km data was used in determining the NEXRAD's CSI value in order to obtain the best possible determination of the operational accuracy of the radars. The four km NEXRAD data was used to determine the NEXRAD CSI because the four km data is the resolution that will be displayed for ATC use. A value of two radars was used for the consensus for the same reasons discussed in section 6.2.1. This analysis used only a zero expansion (exact match)—both reflectivity thresholds and both absolute and worst case analyses were done. The results of this analysis represent the best measure of the operational accuracy of NEXRAD and ARSR and are presented in tables 12 and 13.

The first column of table 12 shows the type of analysis was performed, absolute or worst case, and the second column shows the reflectivity threshold. The third and fourth columns show the NEXRAD CSI and ARSR CSI values respectively and the last column shows the percent improvement NEXRAD showed over ARSR. NEXRAD again showed significant improvement over ARSR in both the absolute and worst case.

Table 12. CSI Values (1 km Truth, 4 km CSI), Zero Expansion

Analysis Type	Reflectivity Threshold	NEXRAD CSI	ARSR CSI	% Improvement
Absolute	30	55.74	28.01	99
Absolute	41	45.92	20.45	124
Worst Case	30	52.01	27.08	92
Worst Case	41	42.96	20.32	111

Table 13 presents the raw data that was used in the computation of the CSI values in table 12. Similar trends appear in table 13 as have been noted in the previous tables with the number of

hits and misses for both radars being similar (the disparity growing in the worst case) and the number of false alarms being produced by ARSR being the most notable fact.

Table 13. Raw Truth Statistics (1 km Truth, 4 km CSI), Zero Expansion

Analysis Type	Reflectivity Threshold	NEXRAD			ARSR		
		Hits	Misses	False Alarms	Hits	Misses	False Alarms
Absolute	30	11938	888	8590	12776	50	32793
Absolute	41	5419	408	5973	5775	52	22411
Worst Case	30	10178	1425	7965	11553	50	31058
Worst Case	41	4611	795	5327	5361	45	20972

The results of the accuracy analysis presented above provide the best measure of the operational accuracy of the radars. The data was collected using the inputs from all five of the ARSR radars and one NEXRAD, though and could lead to questions regarding the results of using one ARSR and one NEXRAD. For this reason, a one-on-one analysis was performed using two different ARSRs in separate comparisons. The results of this analysis are presented in appendix B.

The operational analysis shows that the four km NEXRAD data that will be used in ATC operations is more accurate than ARSR data in both the absolute and worst case. The NEXRAD provided a 92 percent improvement over ARSR in the worst case for data greater than or equal to 30 dBZ and 111 percent improvement for data greater than or equal to 41 dBZ.

6.3 CONCLUSIONS

The absolute analysis showed that NEXRAD is more accurate at detecting weather than ARSR with the operational analysis demonstrating that, even with the associated lag, the NEXRAD data is far more accurate than the ARSR data. The analyses were conducted in several different ways with many different parameters and they all pointed toward the conclusion that NEXRAD is the superior weather radar.

As mentioned above, these results contained analysis from the August data set only. The qualitative comparison of the other two data sets, though, showed that the results would have been similar. The ARSR did not appear to have as high of a false alarm rate in the June and January data sets, but appeared to have a higher number of misses.

SECTION 7

QUANTITATIVE RADAR COMPARISON

The quantitative comparison of the radar data measured the anomalies that were discovered during the qualitative comparison.The comparison included determining the intensity distribution, comparing the area of weather data coverage, analyzing the distances of aircraft from weather data, and analyzing the penetrations of aircraft into the weather data. The quantitative comparison was completely automated to ensure completeness and accuracy. This section describes the methods used to perform the comparison and discusses the findings.

7.1 INTENSITY DEPICTION ANALYSIS

The intensity analysis compared ARSR data to the corresponding NEXRAD data to determine the distribution of NEXRAD intensity levels associated with each ARSR data level. This analysis substantiated and quantified some of the problems that were discovered during the qualitative comparison with ARSRs ability to determine intensity. The following subsections provide a description of the method used to perform the analysis as well as a detailed discussion of the data obtained.

7.1.1 Method of Analysis

The analysis consisted of recording the NEXRAD data level associated with each point of each line of ARSR data and comparing the NEXRAD 4 km composite product to the ARSR product while matched in time (without the associated lag). The resolution of the display was 1.33 km. The analysis used all three data sets for the entire study area (see figure 3). The distribution of ARSR moderate and heavy data was plotted to show the trends.

7.1.2 Findings

Figure 20 shows the percentage of ARSR moderate and heavy data that corresponded to the NEXRAD data ranges (combined to match the specified ARSR ranges). The most striking observation to be made from the chart is that there was almost no difference in the distributions of ARSR moderate and heavy data. This supports the evidence found during the qualitative comparison that the ARSR is not accurate at determining the intensity of the weather.

Table 14 shows the computed percentages that are presented in figure 20. Optimally, 100 percent of the ARSR moderate data should be in the Level 2 range (30 to 41 dBZ) and 100 percent of the ARSR heavy data should be in the Level 3 and above range (greater than 40 dBZ). It should be observed from the data in the table that only 39.29 percent of the ARSR

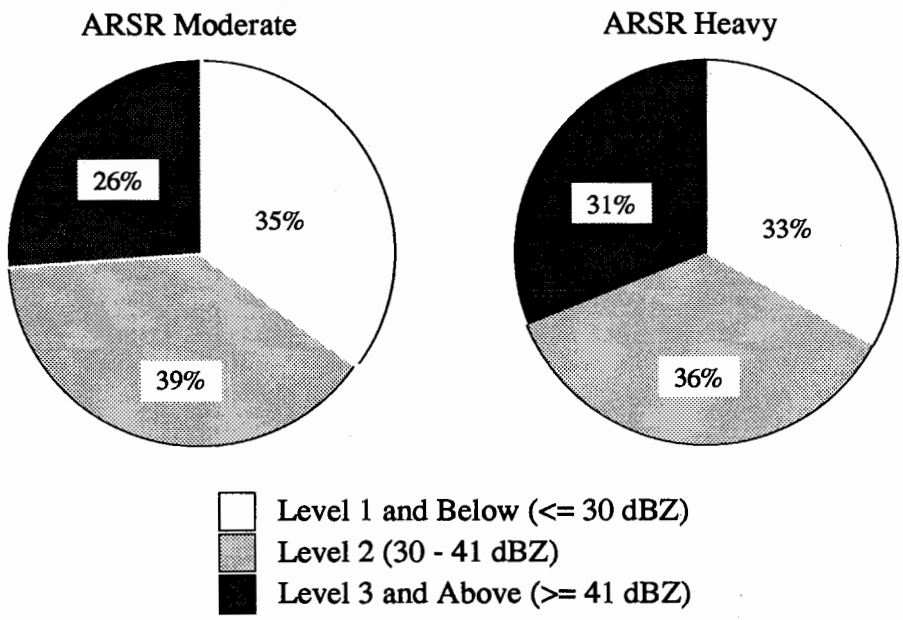


Figure 20. Distribution of ARSR Data Relative to NEXRAD Data

moderate data was given a correct intensity, 30 - 41 dBZ (NWS Level 2), 26.11 percent of the data should have been assigned a heavy intensity, and 34.60 percent should not have been presented at all. In addition, only 31.28 percent of the ARSR heavy data was given the correct

intensity, greater than or equal to 41 dBZ (NWS Level 3 and above), 36.21 percent should have instead been rated at moderate intensity, and 32.52 should not have been presented.

Table 14. Distribution of ARSR Levels Relative to NEXRAD Levels

ARSR Level	NEXRAD Level		
	< 30 dBZ	30-41 dBZ	>40 dBZ
Moderate	34.60	39.29	26.11
Heavy	32.52	36.21	31.28

The distribution analysis shows that ARSR has difficulty determining the intensity of the weather, with there being little difference in the distribution of moderate and heavy data. This confirms the same finding from the qualitative comparison.

7.2 AREA ANALYSIS

The area analysis tracked over time the area of weather that each radar depicted. This helped determine if the radars showed trends over time which is difficult to determine using the other analyses which reduced the results down to single values. The following subsections present the method of analysis as well as the findings that resulted from the analysis.

7.2.1 Method of Analysis

The area analysis consisted of calculating the area of weather covered by NEXRAD and ARSR after each NEXRAD update and recording the values over time. The analysis used a 466 km by 466 km region centered on the NEXRAD with a display resolution of 1.33 km. The analysis used the NEXRAD composite product with a resolution of 4 km (with the associated lag removed) and the ARSR data was displayed using the trapezoid drawing method as described in section 6.1.1. The following subsection presents the findings of the area analysis.

7.2.2 Findings

Figures 21 through 23 show the results of the area analysis conducted on all three data sets for reflectivity intensities greater than or equal to 30 dBZ. The data from 4 January and 2 June data

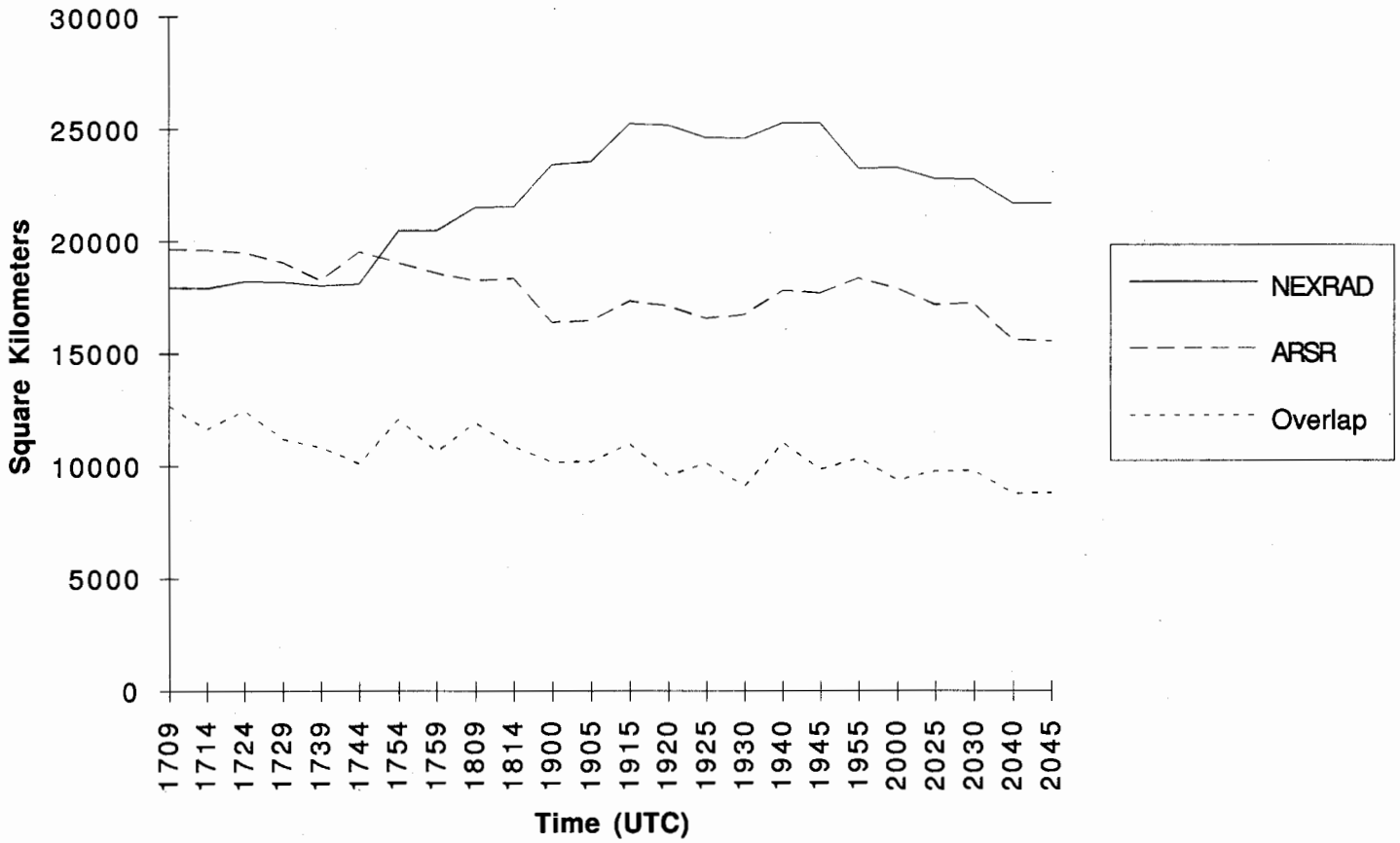


Figure 21. Area of Weather Over Time, 23 January 1992

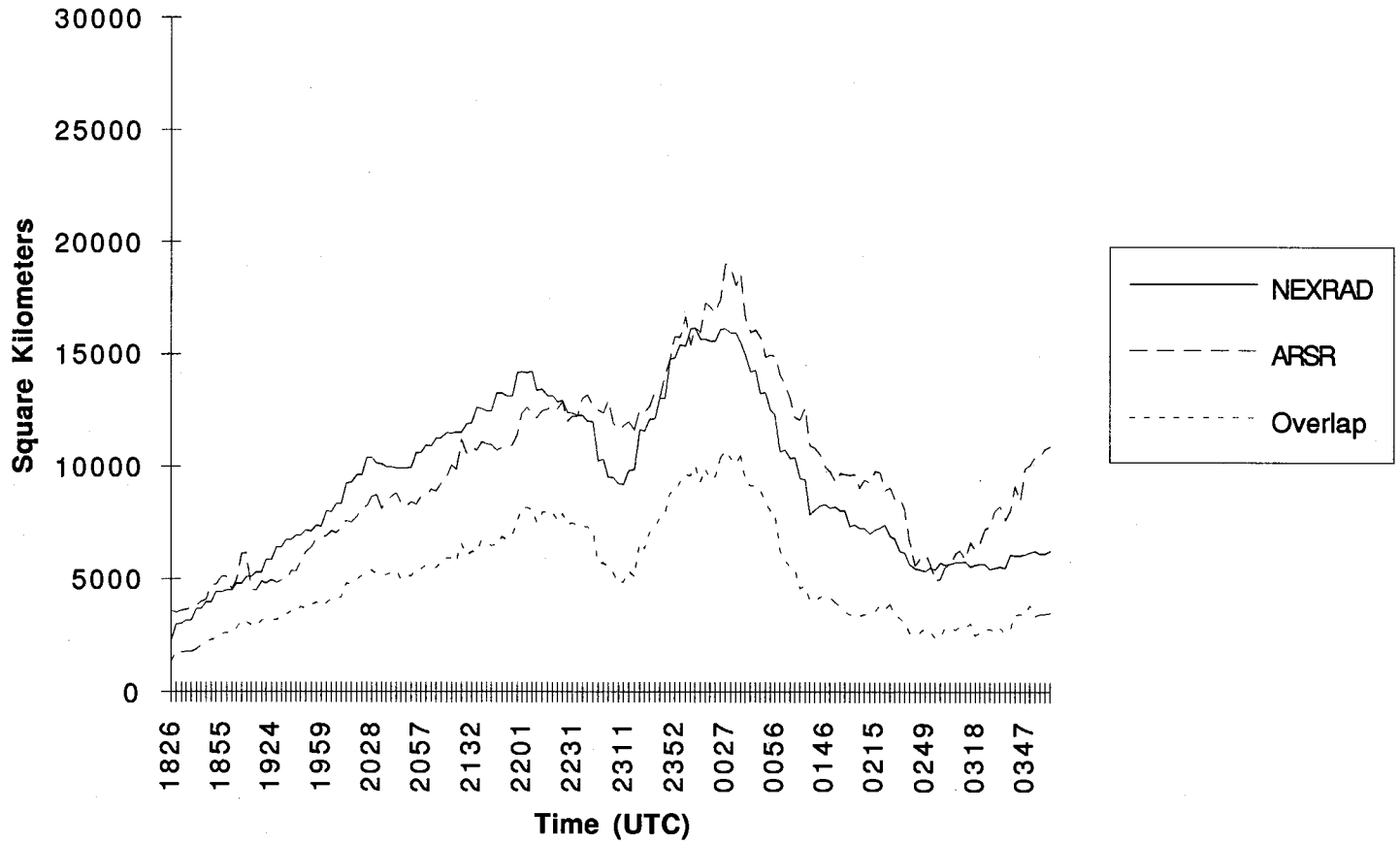


Figure 22. Area of Weather Over Time, 2 June 1992

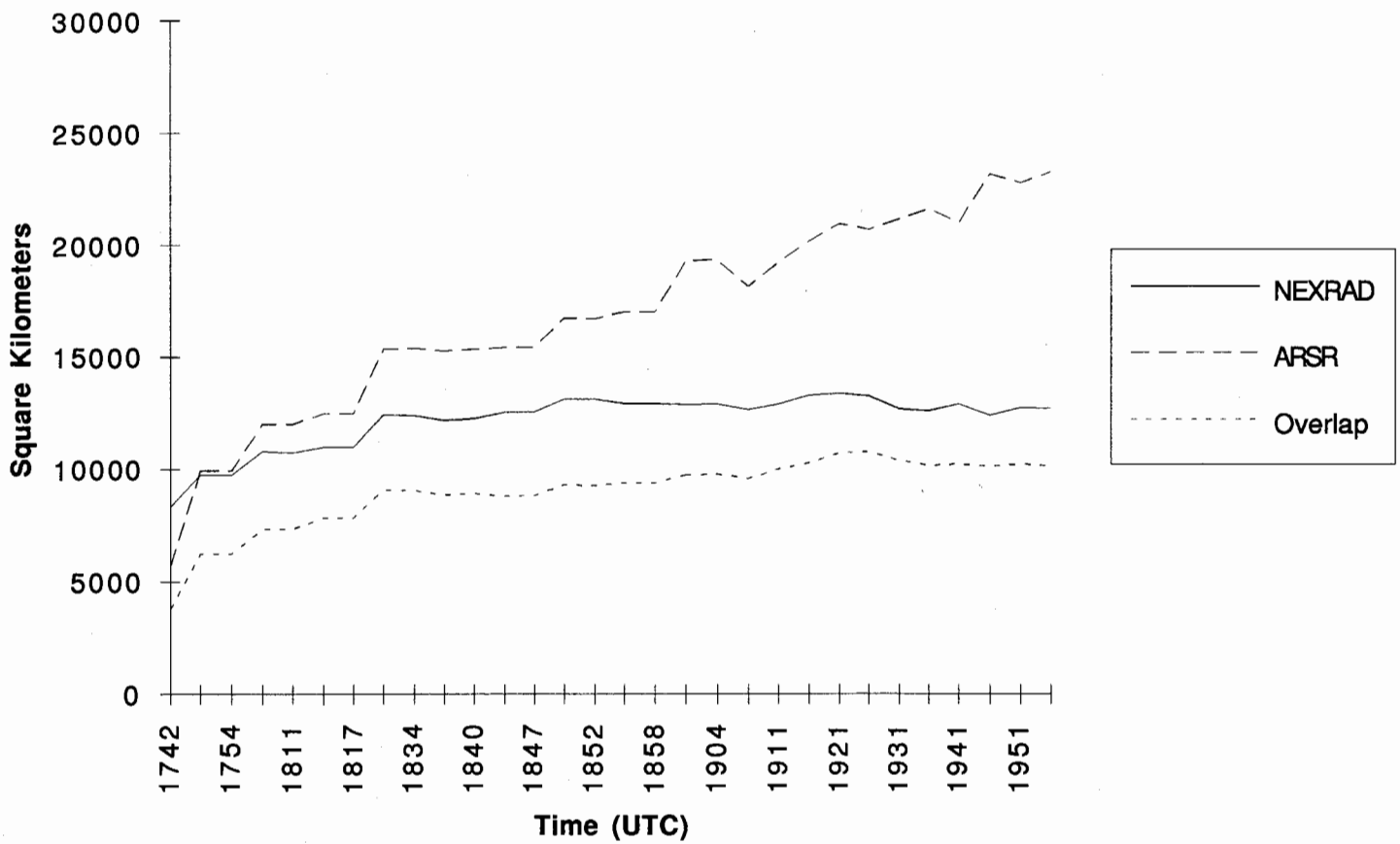


Figure 23. Area of Weather Over Time, 4 August 1992

sets show very similar amounts for NEXRAD and ARSR with about a 50 percent overlap. The data from 4 August data set shows that ARSR paints a larger area of weather than the NEXRAD with about 75 percent of the NEXRAD data overlapping with ARSR and about 50 percent of ARSR data overlapping with NEXRAD data.

Figures 24 through 28 show the results of the area analysis conducted on all three data sets for reflectivity intensities greater than or equal to 41 dBZ. All three data sets show that ARSR shows a much larger area of data than NEXRAD. The overlap consists of about 60 percent of the NEXRAD data and between 25 and 50 percent of the ARSR data.

The data shows that ARSR depicted much more heavy (greater than 40 dBZ) area than NEXRAD and that the overlap between the radars is limited. These results add supporting evidence that ARSR appears to have trouble determining the intensity of the weather (by depicting too large an area of weather) as well as showing the difference in coverages that is hinted to in the accuracy analysis.

7.3 PENETRATION ANALYSIS

The penetration analysis utilized aircraft track data in order to analyze the patterns of the aircraft relative to the weather as depicted by each radar. This analysis was conducted by counting the number of penetrations by aircraft into weather and comparing the results for the two radars. With larger numbers of penetrations there is less confidence in the severity of the weather at the indicated location because of an aircraft's tendency to avoid severe weather (a large portion of the aircraft used in this analysis could have been equipped with on board radar). The following subsections present the method used to obtain this data, as well as the detailed findings.

7.3.1 Method of Analysis

The penetration analysis used a 466 km by 466 km square centered on the location of the NEXRAD radar as its analysis region. The analysis covered both the composite and the layered NEXRAD products (both with a 4 km resolution) with and without the associated lag. The analysis utilized all three of the data sets, although, only the NEXRAD's composite data was analyzed for the January data set because the layered products were unavailable. Section 6.1.1 describes the trapezoidal drawing method used for the ARSR data. The analysis used a 1.33 km resolution. The analysis included only beacon equipped aircraft transmitting unique beacon codes in order to eliminate aircraft that are likely flying without controller services. The analysis also eliminated aircraft below 8000 feet because the study was concerned mainly with en-route control.

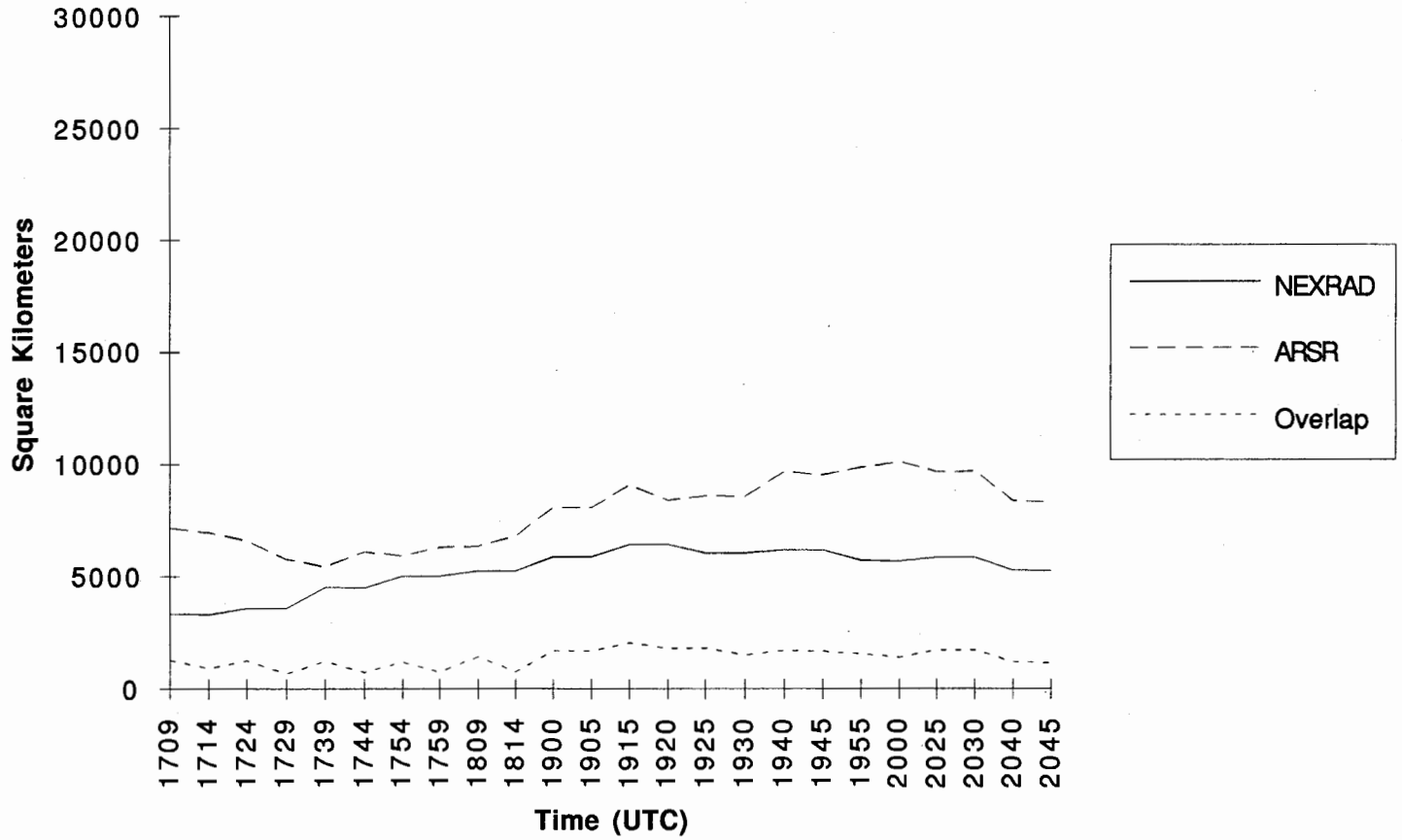


Figure 24. Area of Weather (>40 dBZ) Over Time, 23 January 1992

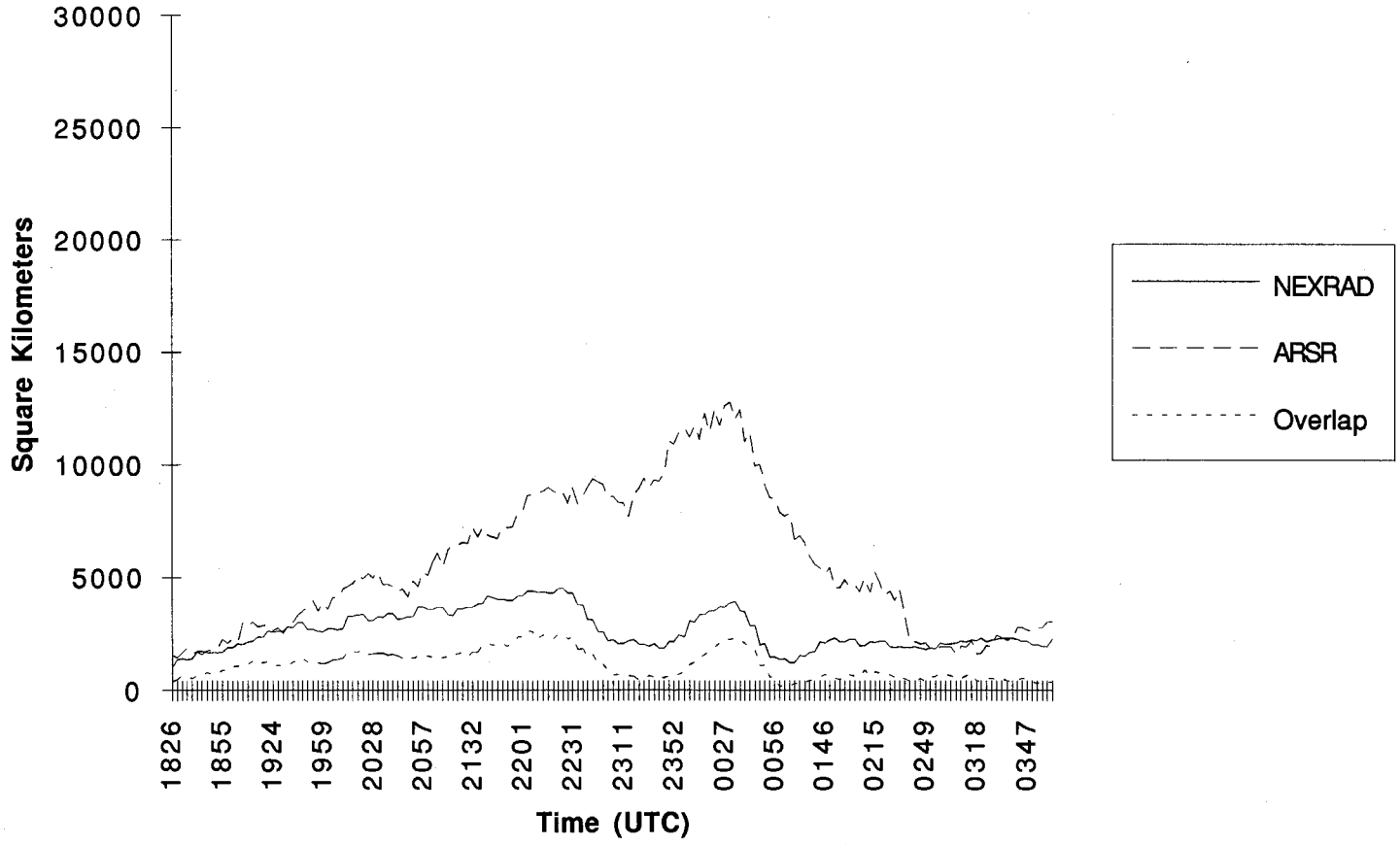


Figure 25. Area of Weather (>40 dBZ) Over Time, 2 June 1992

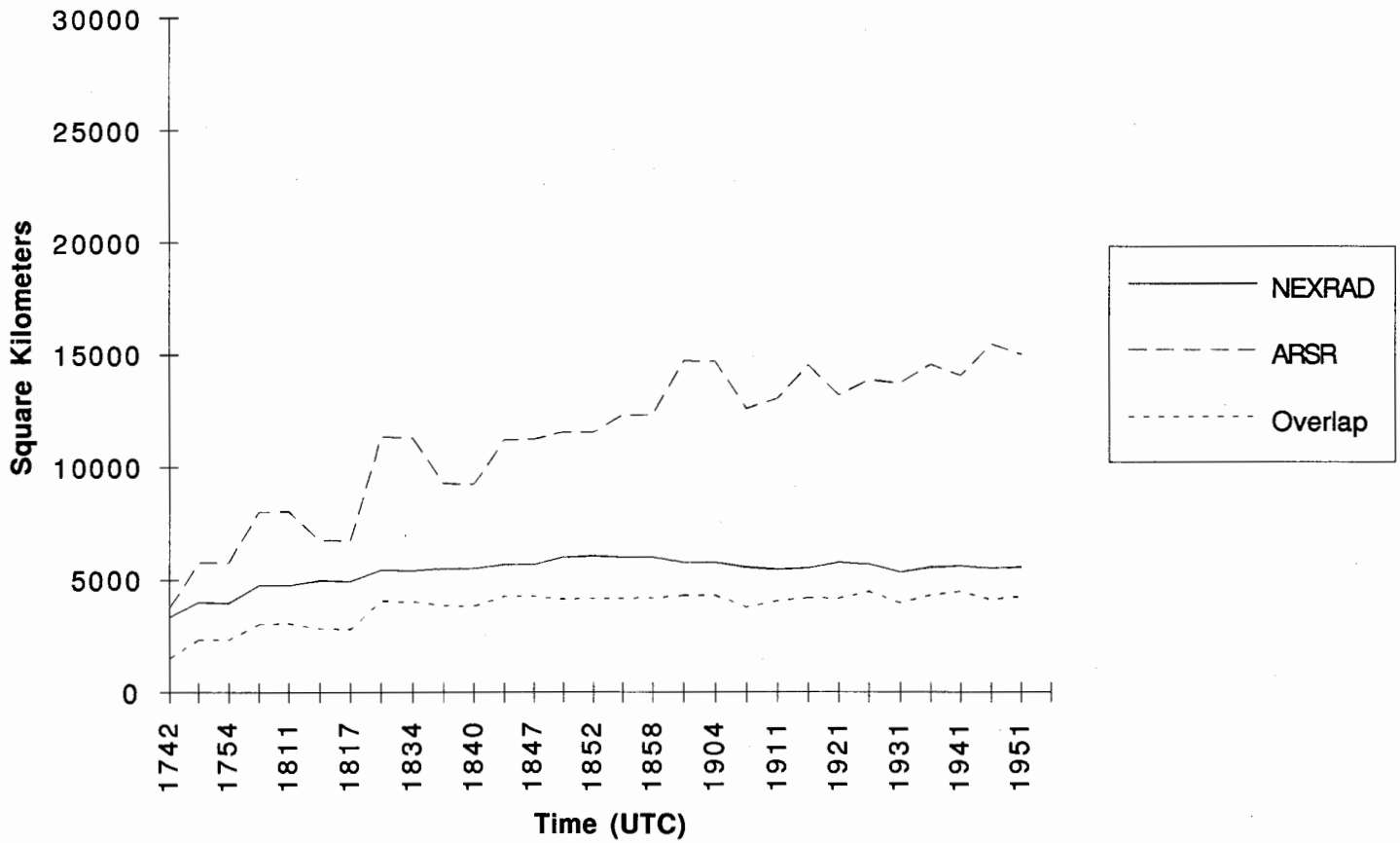


Figure 26. Area of Weather (>40 dBZ) Over Time, 4 August 1992

The positions of each of the aircraft were sampled every 30 seconds to test for penetrations. The analysis defined a penetration as the aircraft occupying a location on the display that also contained weather. Figure 27 shows an example of two aircraft and the resulting penetrations. The arrows represent the flight path of a single aircraft, and the shaded regions represent weather. The cross marks represent penetrations of the weather when sampled at a 30 second interval. Aircraft A has three penetrations into the lightly shaded weather and two penetrations into the darkly shaded weather. For the layered products, the analysis included only aircraft within the coverage altitudes (these same aircraft were also compared against the ARSR data).

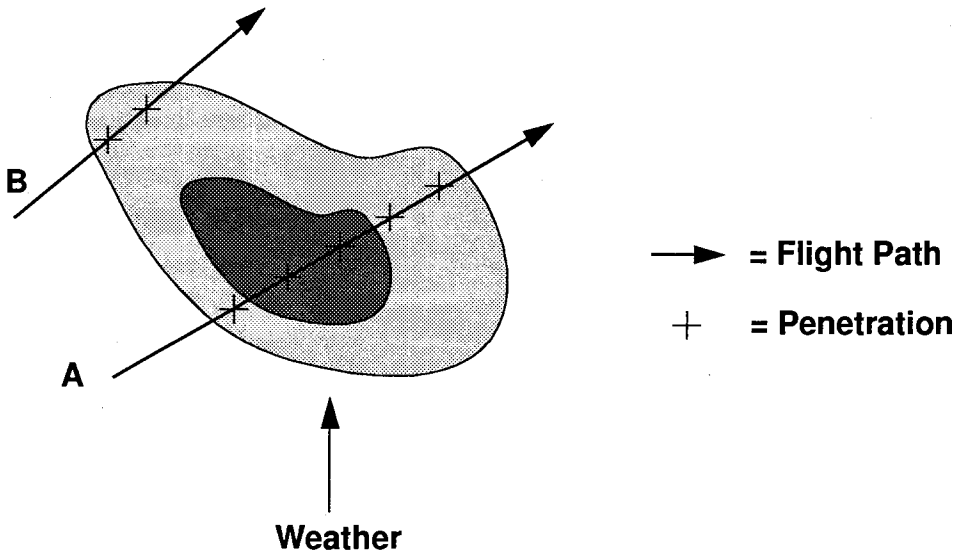


Figure 27. Aircraft Penetrations

7.3.2 Findings

Figure 28 presents the results of the penetration analysis performed on the composite NEXRAD product and the ARSR data for weather data at a level of 41 dBZ and greater. It shows that ARSR had an overwhelmingly larger total number of penetrations than NEXRAD (both with and without lag) as well as the fact that there was little difference in the number of penetrations into ARSR heavy and moderate. Penetrations into NEXRAD, though, decreased sharply as the reflectivity level increases. A final observation about the data in figure 28 is that there was only a slightly larger number of penetrations into NEXRAD with a lag than NEXRAD without, showing that the delay does not significantly affect the operational accuracy of the NEXRAD data. The NEXRAD data, though, did have a noticeable number of

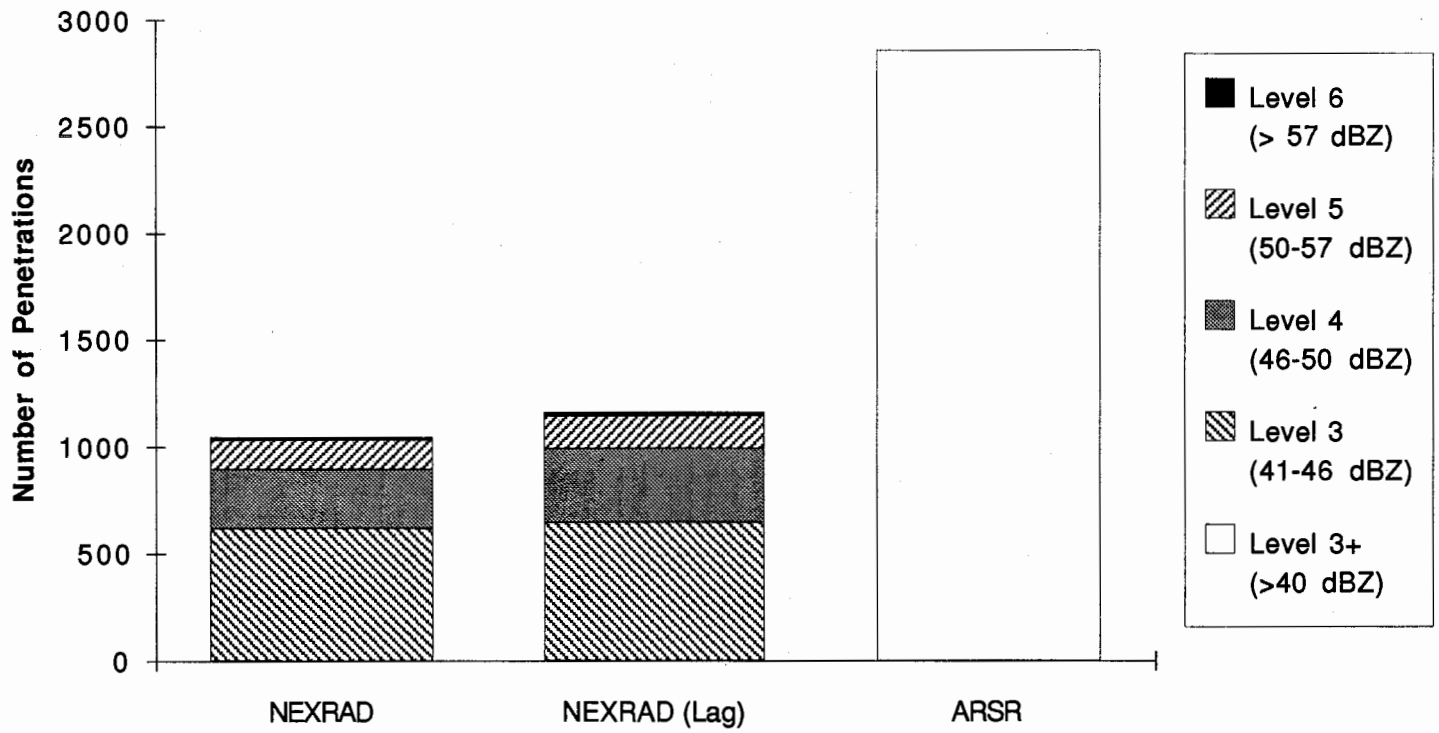


Figure 28. Aircraft Weather Penetrations

penetrations, some of which can be attributed to the fact that the data covers a significant range of altitudes and all altitudes in the layer may not contain weather.

Figure 28 shows the penetrations by aircraft at all altitudes (greater than 8000 ft.) into weather at all altitudes. The NEXRAD, though, provides layered products which closely match the sector altitude levels found in ATC. Figure 29 shows the penetration data for the NEXRAD layered products by aircraft that are in the associated altitude range compared to the penetrations into ARSR by the same aircraft. The penetration data is for weather intensities of 41 dBZ and greater. The NEXRAD layered products reduce the amount of weather to only that which is in the relevant altitude layers, and thereby reduce the number of penetrations into NEXRAD depicted data.

Table 15 presents penetration data that was displayed in the previous charts.As mentioned above, the composite data is from all three data sets but the penetration data for the layered products is only for the 2 June and 4 August data sets.

The argument could be made that the radar that displays the smallest area of weather, regardless of truth (NEXRAD in this case), is most likely to do the best in the penetration analysis. In order to analyze its implications, the ARSR penetration data was corrected to account for its display of a larger area of weather. Calculating the amount of area presented by ARSR and dividing it by the amount of weather presented by NEXRAD produces the correction factor. The number of penetrations produced by ARSR was then divided by the correction factor. Table 16 presents the results of the penetration analysis with the corrections. NEXRAD still showed significantly less penetrations than ARSR. The corrected results, though, are not considered to be the best representation of aircraft weather penetrations because the correction process can penalize the more accurate radar if it shows less area than the competing radar. This correction was computed to demonstrate that even if ARSR had displayed the same amount of weather as NEXRAD, NEXRAD (with the lag) would still reduce the number of penetrations significantly.

The penetration data shows that aircraft tend to penetrate weather as depicted by ARSR more frequently than as depicted by NEXRAD. Aircraft, especially those with on-board radar, are normally going to avoid known severe weather, so this leads to a higher confidence in the weather data from the NEXRAD than from the ARSR. The excess penetrations into the ARSR data supports the finding from the area analysis that ARSR paints too much weather as well as the finding from the intensity analysis that ARSR has trouble determining the intensity of the weather. Based upon the data that is provided by this analysis, the NEXRAD data seems to be the more accurate data and appears to provide a closer representation of what the pilot is seeing

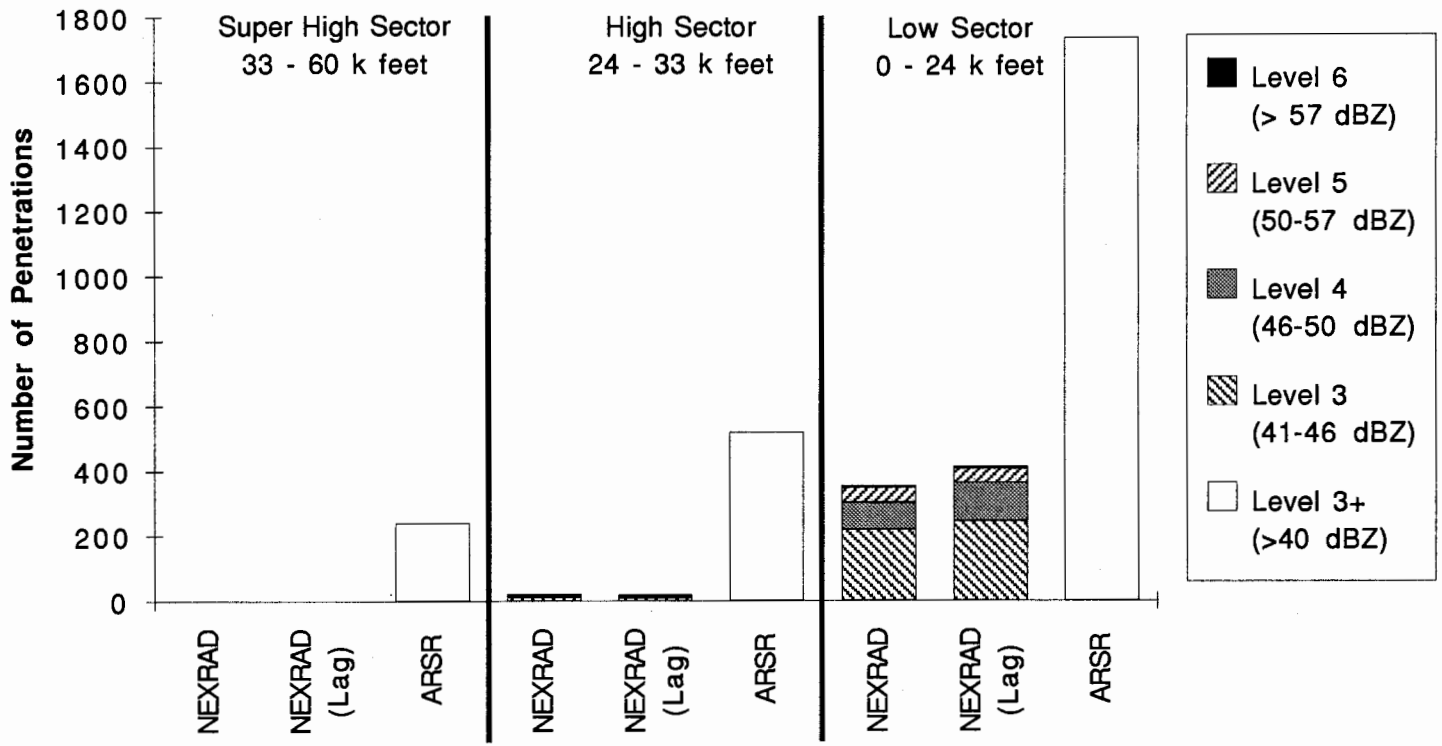


Figure 29. Aircraft Weather Penetrations, Layers

Table 15. Aircraft Weather Penetration Data (Greater Than 40 dBZ)

Product	Total	41-46 dBZ	46-50 dBZ	50-57 dBZ	> 57 dBZ
NEXRAD Composite	1045	619	275	137	14
NEXRAD Composite (lag)	1161	647	344	151	19
ARSR	2858				
NEXRAD Layer 1 (Aircraft at 0 - 24 k feet)	353	219	82	48	4
NEXRAD Layer 1 (lag) (Aircraft at 0 - 24 k feet)	413	244	118	44	7
ARSR (Aircraft at 0 - 24 k feet)	1734				
NEXRAD Layer 2 (Aircraft at 24 - 33 k feet)	21	13	4	4	0
NEXRAD Layer 2 (lag) (Aircraft at 24 - 33 k feet)	20	11	3	6	0
ARSR (Aircraft at 24 - 33 k feet)	520				
NEXRAD Layer 3 (Aircraft at 33 - 60 k feet)	0	0	0	0	0
NEXRAD Layer 3 (lag) (Aircraft at 33 - 60 k feet)	0	0	0	0	0
ARSR (Aircraft at 33 - 60 k feet)	240				

Table 16. Aircraft Weather Penetrations, Corrected for Area

Product	NEXRAD Lag	Corrected ARSR	Uncorrected ARSR
Composite	1161	1398	2858
Layer 1	413	763	1734
Layer 2	20	34	520
Layer 3	0	2	240

7.4 DISTANCE ANALYSIS

The distance analysis consisted of determining the average minimum distance that each aircraft came to each level of weather. The purpose of the distance analysis was to determine whether or not aircraft tended to keep a greater distance from the weather as presented by ARSR or by NEXRAD. The following subsections describe the method of analysis as well as the findings.

7.4.1 Method of Analysis

The distance analysis consisted of sampling the positions of each aircraft every thirty seconds to determine its distance from each level of weather as depicted by ARSR and NEXRAD. The analysis area included the entire study area (466 x 466 km), though, only the 43 km by 43 km square centered on each aircraft was searched for weather. The distance analysis consisted of determining the minimum distance to each weather level for every aircraft (zero being the lowest value, meaning it penetrated the weather, 22.6 km being the largest value) and computing the average of these minimum distances (1 per aircraft). This analysis used the NEXRAD composite product (with and without the associated lag), and the trapezoidal drawing technique as described in section 6.1.1 for the ARSR data, both of which were displayed at a resolution of 1.44 km. The analysis utilized all three data sets, the results of which are presented in the following section.

7.4.2 Findings

Figure 30 shows the results of the distance analysis for all three data sets which were combined by using a weighted average based on the number of aircraft in the data set. The data shows that aircraft keep a greater distance from the weather as depicted by NEXRAD and that the lag does not reduce this amount significantly. Table 17 shows the actual average distances presented in figure 30.

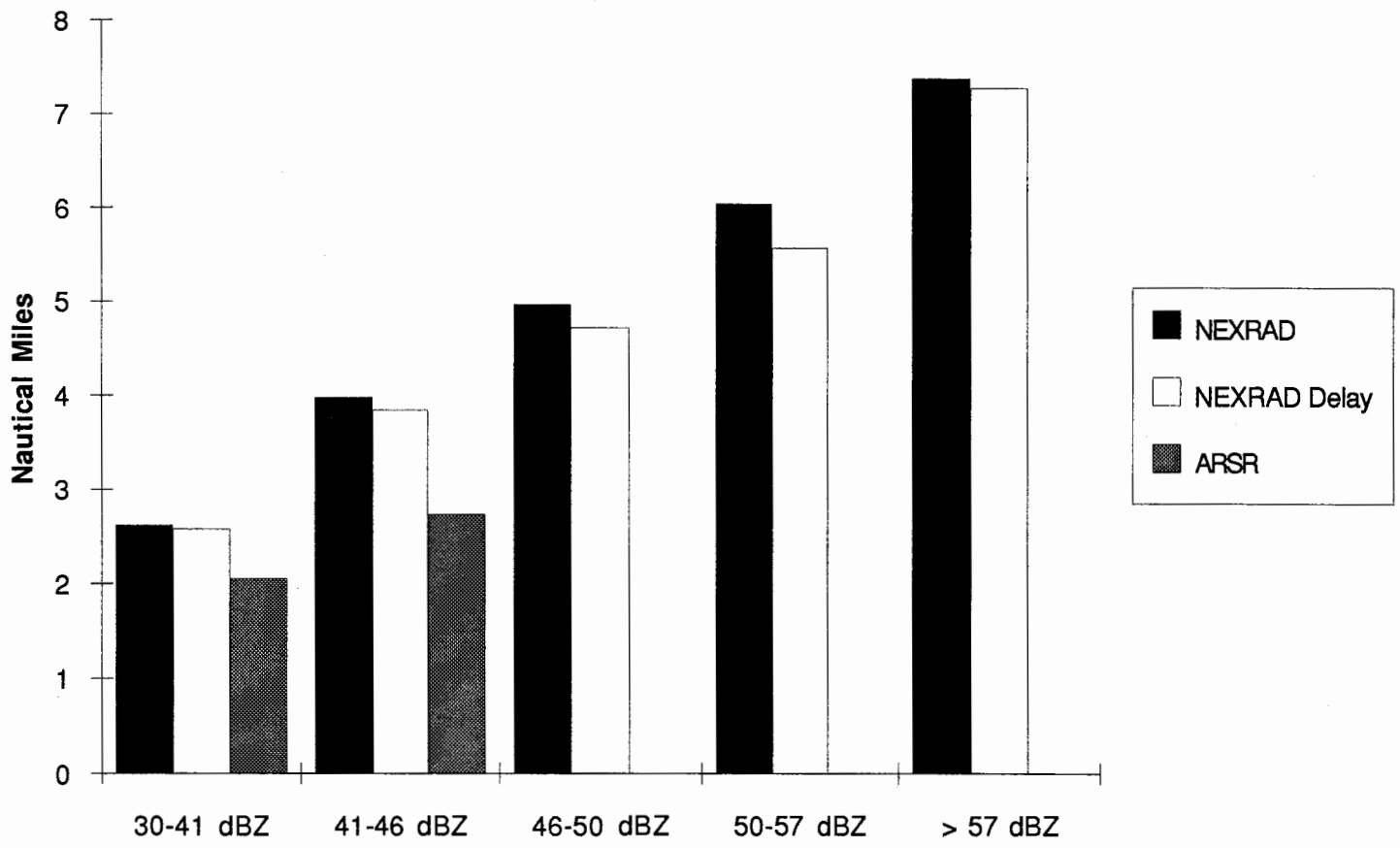


Figure 30. Average Minimum Distance of Aircraft From Weather

The data from the distance analysis can be interpreted as meaning that the NEXRAD data is more accurate because pilots tend to keep more distance away from the weather as depicted by NEXRAD. Another interpretation could be that the ARSR data was more accurate because the aircraft were coming closer to the more severe weather as depicted by ARSR meaning that the ARSR was more credible because the pilots were flying on the "edge" of the weather. Because of these problems, the distance analysis was deemed to be inconclusive.

Table 17. Average Minimum Distance of Aircraft from Weather in Nautical Miles

Product	30-41 dBZ	41-46 dBZ	46-50 dBZ	50-57 dBZ	> 57 dBZ
NEXRAD Composite	2.62	3.98	4.96	6.03	7.37
NEXRAD Composite (lag)	2.58	3.84	4.72	5.57	7.27
ARSR	2.06	2.74			

7.5 SUMMARY

The data collected during the quantitative comparison indicates that NEXRAD is the more accurate weather radar for operational ATC use. ARSR has difficulty in determining the intensity of the weather which was shown by the weather data analysis. Operationally, this means that ARSR can miss (or not sufficiently portray the hazards of) potentially dangerous airspace while depicting as hazardous areas of usable airspace. This is supported by the aircraft analysis which indicates that the weather as depicted by NEXRAD is a more accurate depiction of the pilots view, especially when using the layered products. The quantitative comparison supports the findings of the truth analysis and the qualitative comparison, that NEXRAD is the superior weather radar.

SECTION 8

CONCLUSIONS AND RECOMENDATIONS

The qualitative, accuracy, and quantitative portions of the study conclude the following:

- ARSR presents a coarse depiction of cell boundaries; NEXRAD is concise
- ARSR level thresholds appear to vary considerably; NEXRAD levels are accurate and consistent
- ARSR appears to miss significant weather and detect weather that is not real; NEXRAD more accurately depicts the weather
- ARSR updates the position of the weather frequently; NEXRAD data aging "lag" is real but has little operational impact
- ARSR does not discriminate between altitudes; NEXRAD offers altitude discrimination in three layers
- ARSR appears not to match the pilot's view of the weather; NEXRAD appears to better depict the weather the pilot is observing at different altitudes
- ARSR is not well suited for tactical ATC because of the above problems; NEXRAD is well suited for tactical ATC

Operationally, this means that ARSR can mislead controllers when the ARSRs miss significant weather and produce false weather, possibly reducing capacity. Compared to today's ARSR weather display, NEXRAD will more accurately depict the weather to the controller and better depict the pilot's view of the weather. These features can help tactical ATC operations, possibly improving capacity and reducing the pilot-to-controller communications load. Finally, the lag associated with the NEXRAD data does not significantly affect the operational accuracy of the data, and in comparison to the ARSR data, the NEXRAD data is much more accurate.

The FAA should continue to pursue the use of NEXRAD data, as the primary weather source for the enroute ATC environment, and the programs that deliver the data to the end-users.

LIST OF REFERENCES

Dixon, M., 15 December 1992, *ARSR/NEXRAD Comparison Study Phase One (Third Draft)*, National Center For Atmospheric Research, Boulder, CO.

Federal Aviation Administration, 4 August 1989, "Multiple Radar Processing," *National Airspace System Configuration Management Document*, NAS-MD-320, Federal Aviation Administration Technical Center, Atlantic City International Airport, NJ.

National Oceanic and Atmospheric Administration, February 1991, "Doppler Meteorological Observations", *Federal Meteorological Handbook Number 11*, Part C Federal Coordinator for Meteorological Services and Supporting Research, Washington, D.C.

APPENDIX A

DISPLAY AND ANALYSIS WORKSTATION

In order to analyze the acquired data, a display and analysis workstation was developed. This workstation allows the weather and aircraft data to be visualized and manipulated in time sequence allowing qualitative analyses to be performed. Additional modifications of the workstation including image processing techniques allowed the quantitative data to be collected (see section 7). This section describes the major functions of this workstation.

A.1 SCREEN LAYOUT

The workstation screen is divided into 4 sections (see figure A-2): WSR-88D, ARSR, combined, and legend. The WSR-88D section displays only the WSR-88D data; the ARSR section displays only the ARSR data; and the combined section displays the WSR-88D and ARSR data overlaid on one another. The legend section displays the product legends for all the products as well as the product times.

A.2 DATA MENU

The data menu allows the selection and display of the desired data sets and data types (see figure A-2). After each item is selected, it is displayed in the appropriate portion of the screen as described above. The following sections describe this menu.

A.2.1 Data Sets

This portion of the menu allows a data set to be selected. When selected, the data set is loaded and the start time of the data set is based on the time set in the Times menu (see section A.3). Only one data set may be selected at any given time.

A.2.2 NEXRAD Products

This portion of the menu allows the display of NEXRAD products. The five choices are Layer 1, Layer 2, Layer 3, and Composite Reflectivity products, as well as Echo Tops. Only one product may be selected at any given time.

The NEXRAD products are displayed at a 4 km resolution in 8 data levels as described in the legend. The value of the data levels are different for reflectivity products and echo tops products-the reflectivity legend is shown.

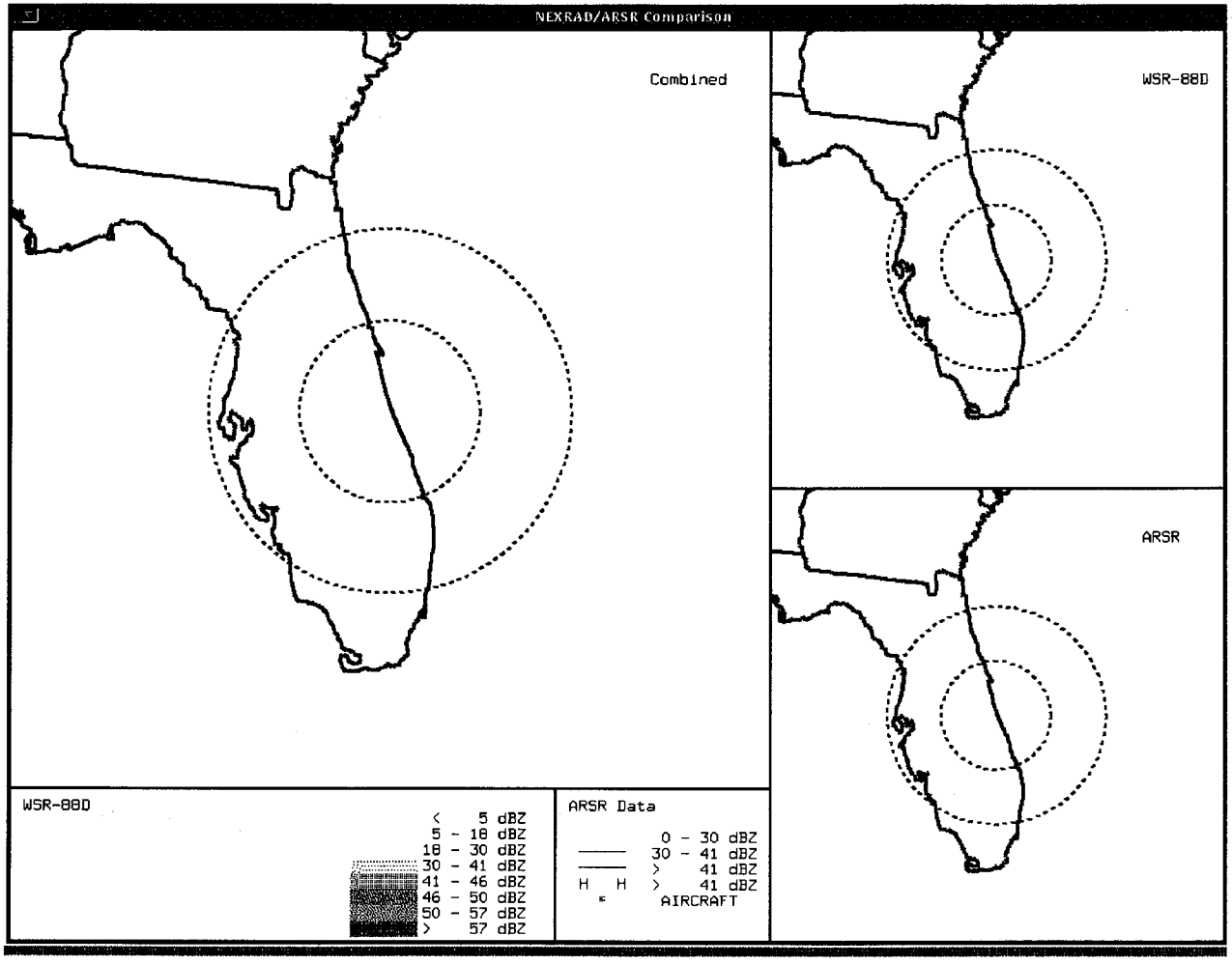


Figure A-1. Screen Layout

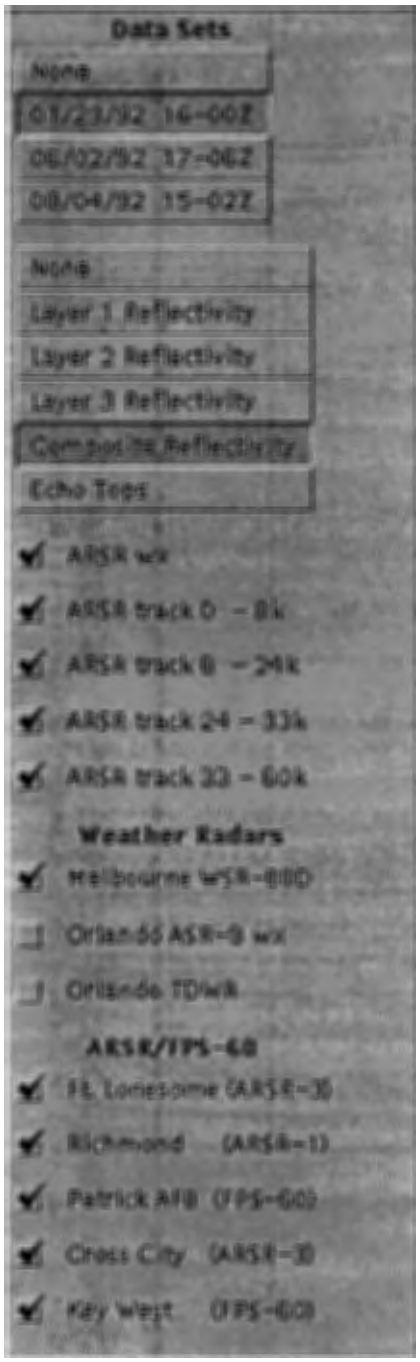


Figure A-2. Data Menu

Table B-2. Raw Truth Statistics (1 km Truth, 4 km CSI), Zero Expansion, One-on-One, Patrick AFB FPS-60

Analysis Type	Reflectivity Threshold	NEXRAD			ARSR		
		Hits	Misses	False Alarms	Hits	Misses	False Alarms
Absolute	30	9683	687	8811	9953	417	20343
Absolute	41	4735	340	5637	4939	136	18615
Worst Case	30	9312	1058	8831	9953	417	20343
Worst Case	41	4379	696	5559	4939	136	18615

data), causing a higher false alarm rate for NEXRAD, reducing its CSI value. The ARSR-3 (Fort Lonesome), though showed a marked improvement in CSI values for the 40 dBZ threshold over those presented in the aggregate analysis. The CSI went from a value of 20 to a value of 40, showing nearly a 100 percent improvement. The changes that resulted in the truth slightly reduced NEXRAD's accuracy. This significant improvement in the ARSR value is most likely due to the fact that the ARSR-3 is the best of the ARSR radars and it shows up as a significant reduction in the false alarm rate (for the data greater than 40 dBZ) when the data in table B-3 and table 13 are compared.

Table B-3. Raw Truth Statistics (1 km Truth, 4 km CSI), Zero Expansion, One-on-one, Fort Lonesome ARSR-3

Analysis Type	Reflectivity Threshold	NEXRAD			ARSR		
		Hits	Misses	False Alarms	Hits	Misses	False Alarms
Absolute	30	9796	440	8698	9484	752	15877
Absolute	41	5577	787	5752	3797	2567	3126
Worst Case	30	9235	1001	8908	9484	752	15877
Worst Case	41	4925	965	5013	3608	2282	3054

Even though the ARSR data showed improvements in the CSI value when they were singled out, the NEXRAD data was still more accurate in all cases. This data, though, does not represent the ATC operational accuracy of the radars, and is presented for the purposes of providing a complete analysis. The accuracy data from the mixed resolution analysis (see section 6) represents the best measure of the ATC operational accuracy of the radars.

GLOSSARY

ACRONYMS

AAS	Advanced Automation System
ARD	Research and Development Service
ARSR	Air Route Surveillance Radar
ARTCC	Air Route Traffic Control Center
ASE	Systems engineering Service
ASR	Airport Surveillance Radar
ATC	Air Traffic Control
CSI	Critical Success Index
FAA	Federal Aviation Administration
FPS	Flight Planning System
JSPO	Joint System Program Office
LL	Lincoln Laboratory
MIT	Massachusetts Institute of Technology
MSL	Mean Sea Level
NAS	National Airspace System
NCAR	National Center for Atmospheric Research
NEXRAD	Next Generation Weather Radar
NWS	National Weather Service
OSF	Operational Support Facility
TDWR	Terminal Doppler Weather Radar
WSO	Weather Service Office
WSR	Weather Surveillance Radar